UNIVERSITY OF LEEDS

Using big data and statistics to understand melanoma skin cancer

Adam Jacob Trower

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

The University of Leeds

Faculty of Medicine and Health

School of Medicine

January 2021

Intellectual Property

The candidate confirms that the work submitted is his own and that appropriate credit has been given where reference has been made to the work of others.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

Analysis conducted in this thesis has previously been published or is in preparation for submission. Analysis previously published is the non-red hair colour analysis pipeline conducted in Chapter 7, which has been published in a melanoma meta-analysis paper (Landi et al. 2020). Analysis currently in preparation for submission is the work conducted to understand the genetic architecture between red hair and non-red hair in Chapter 4, the pigmentation-based genome-wide association studies conducted in Chapter 5, and the polygenic risk scores generated in Chapter 6.

Signed 

Acknowledgements

I would like to acknowledge the Alan and Joan Webster scholarship for funding my PhD study which has made this thesis possible.

Professor Tim Bishop, my main supervisor and the Head of Section of Epidemiology and Biostatistics, for welcoming me into the world of statistical genetics and melanoma research, for providing a wonderful learning atmosphere where I was able to create lasting relationships that better my professional skills constantly; and for all the help and support your have provided throughout the whole experience.

Dr Mark Iles, my co-supervisor, for sharing with me your melanoma GWAS data, your constant help whenever I found myself down a difficult path, and also for your valuable comments and suggestions to improve my thesis.

Professor Julia Newton-Bishop, my co-supervisor, for opening my eyes to the underlying biology of all my work, and for always giving a different prospective when I was presenting my findings.

Associate Professor David Duffy for providing naevus GWAS summary statistics utilised in my analysis.

Prof. Jenny Barrett for always providing useful comments on any findings presented.

May Chan for your continuous help with the data and helping me to understand the Leeds

melanoma cohort.

Dr John Davies for always supporting me when I hit a wall, whether that was analysis-related or emotionally.

Dr Juliette Randerson-Moor for helping me with any questions, for welcoming me into the wider-research group, and for the great times we spent at Melgen conferences.

John Taylor, Faye Elliott, Dr Jeremie Nsengimana, Dr Mark Harland, Dr Hara Hatzigeorgiou, Dr Vinny Davies, Theofanis Tsismentzoglou, Natalie Chaddock, and Michal Zulcinski for providing me with support in the office whenever it was needed.

My friends, Joe Mount, Joe Yates, Leon Lok, Ash Chapman, Jay-Jay Luckman, for your constant support when I was going through difficult periods.

My family, Kevin Trower, Geraldine Trower, Marc Trower, for their never-ending support.

Finally, My partner Amy Rebbeck, for your love, incredible understanding and support for these past few years of my PhD journey.

Abstract

Human pigmentation, defined as the coloration of external tissues on the body, can be measured in many different ways and is highly heritable. The extent to which the genetic factors underlying one pigmentary measure influence another pigmentary measure is largely unknown. Pigmentation is one of the strongest risk factors melanoma and many of the genetic loci associated with risk of melanoma are known to be related to pigmentation. Thus it is hoped that insights into the genetics of pigmentation may improve our understanding of melanoma. It is reported in this thesis the results of genome-wide association analyses of the 350,000 UK Biobank participants for five pigmentary measures and three combined pigmentation scores. 500 signals from 322 loci associated are identified; 109 of these loci are novel. By generating polygenic risk scores (PRS) for four pigmentary measure and a combined score, it was demonstrated the utility these PRS have when modelling melanoma risk as they performed similar to melanoma-based PRS. Constructing joint-analysis pipelines that supplement current melanoma GWAS summary statistics with large-scale pigmentation-based GWAS summary statistics uncovered 88 potential novel melanoma loci. The discovery of many new loci improves our understanding of how the genetic architecture of melanoma relates to pigmentation. The results in this thesis increase our understanding of the genetic architectures of distinct pigmentary measurements by identifying 109 novel loci; how the genetics of pigmentary characteristics overlap; provide evidence of the equivalency of using pigmentation-derived PRS (AUC=0.64) to a melanoma-specific PRS (AUC=0.65) when modelling melanoma risk; and the viability of jointly-analysing melanoma with known risk pathway phenotypes to identify novel melanoma loci, highlighting that large gains in our understanding of the genetic architecture of melanoma may still be achieved with larger melanoma data sets and by supplementing with known risk pathway phenotypes.

Contents

List of Tables

List of Figures

List of abbreviations

AP	Analysis Pipeline
AUC	Area under the receiver operating characteristic curve
CFA	Confirmatory factor analysis
Chr	Chromosome
EFA	Exploratory factor analysis
eQTL	expression Quantitative Trait Loci
FA	Factor analysis
GCWJA	Genetic correlation weighted joint-analysis
GWAS	Genome-wide association study
LD	Linkage disequilibrium
MAF	Minor allele frequency
n	Number of samples
ODF	Overlap dataframe
OR	Odds ratio
PC	Principal component
PCA	Principal components analysis
Pigscore 1	Overall pigmentation score 1
Pigscore 2	Overall pigmentation score 2
Pigscore 3	Overall pigmentation score 3
PRS	Polygenic risk score
QC	Quality control
SD	Standard deviation
SE	Standard error
SNP	Single nucleotide polymorphism
SSWJA	Sample size weighted joint-analysis
TWAS	Transciptome-wide association study
UV	Ultraviolet radiation

Chapter 1

Introduction

1.1 Pigmentation

Human pigmentation can be defined as the different colorations of external tissues, such as skin, hair and iris. Determination of these different colorations is primarily caused by the quantity, distribution and ratio of the two main types of melanin: eumelanin which determines brown to black tones; and pheomelanin which determines reddish-brown to blond tones (Lin and David E. Fisher 2007). These two types of melanin are synthesised in the melanocytes which are located in the basal layer of the epidermis, hair bulb, and iris.

Changes in pigmentation are believed to have been driven by early human populations migrating away from Africa. When these humans found themselves in cooler regions with lower levels of UV, selective pressure on pigmentation phenotypes is postulated to have changed. Around the equator, a darker pigmentation complex was beneficial to early humans as it offered protection from high UV exposure (Caldwell et al. 1998; Madronich et al. 1998; Agar and Young 2005), which has been found to be associated with birth defects (N. Jablonski and George Chaplin 2010). However, with migration to climates with lower UV, it's suggested the selective advantage of darker pigmentation reduced. Lower levels of UV meant that synthesising a sufficient amount of vitamin D became more difficult for darker pigmentations, and it has been argued that this was a factor for the increase in lighter pigmentation amongst those living in these regions with lower UV (Loomis 1967; Murray 1934).

Hair colour is one of the most prominent phenotypic traits of humans; ranging from blond, red, brown, and black it plays a central role in cultural and social identity for many peoples. Although a diverse phenotype, it is believed around 70-90% of the variation in hair colour is determined by genetic factors (Morgan et al. 2018). Due to this high heritability, hair colour is highly correlated with ethnicity and therefore differs between ethnic populations. These differences can be seen in the prevalence of hair colour across populations, with a wider range of pigmentation being more, if not exclusively, prevalent in Europeans when compared to African and Asian populations (Loussouarn et al. 2016). As well as ethnicity, hair colour correlates with other human pigmentation traits, such as skin colour and eye colour.

Similarly, skin colour variations is also one of the most prominent phenotypic traits of humans; ranging from very fair skin to black, it has played an similar role to hair colour in social and cultural identity, and correlates strongly with geographical location likely due to UV intensity. This geographical correlation tends to have darker skin pigmentation at lower latitudes where UV is most intense, and becoming lighter as the latitude increases (Relethford 1997). Although this association with latitude is present for skin colour, the variation is believed to be predominately genetic (around 82%) (Paik et al. 2012), with with strong similarities in the distribution of the gene *SLC24A5* with skin colour distribution patterns across the globe. (Figure 1.1).

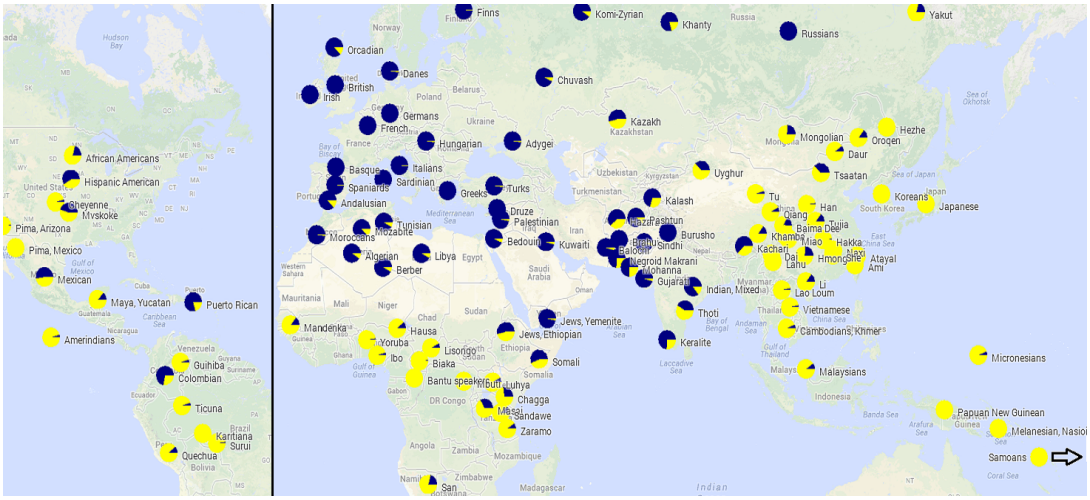


Figure 1.1: Global frequency distribution of the SLC24A5 gene's ancestral Ala111 allele (yellow) and its derived Ala111Thr allele (blue).
Image by FonsScientiae - Own workbased on the Yale School of Medicine's Allele Frequency Database, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=27455242

1.1.1 Biosynthetic pathway for the production of eumelanin and pheomelanin

Melanin pigment is synthesised in follicular melanocytes founds in melanosomes organelles, which are transported to nearby keratinocytes to induce pigmentation. Within the melanosome, *L*-tyrosine is converted to *L*-dopaquinone using the enzyme tyrosinase as a catalyst, which is the initial and key step of melanogenesis (Shosuke Ito, Nakanishi, et al. 2011). One of the key characteristics of Dopaguinone is that it is an ortho-quinone which reacts quickly with thiol compounds and slowly with amines. When thiol compounds, such as *L*-cysteine, are absent in melanosomes, dopaguinone undergoes a reaction with amino groups to create dopachrome (Simon et al. 2009). Over time dopachrome is gradually converted to 5,6-dihydroxyindole (DHI) or is rapidly converted by dopachrome tautomerase to form 5,6-dihydroxyindole-2-carboxylic acid (DHICA) (Tsukamoto et al. 1992). Through the oxidative polymerisation of DHI and DHICA, in various ratios, eumelanin is formed. Conversely, when *L*-cysteine is present in melanosomes, it interacts with dopaquinone to form 5-S- and 2-S-cysteinyldopa (Boissy et al. 1998). The oxidation of these cysteinyldopa isomers with dopaquinone leads to the production of pheomelanin (Kazumasa Wakamatsu et al. 2009) (Figure 1.2).

This biosynthetic pathway for eumelanin and pheomelanin production is the same between all external tissues on the body, such as skin, hair and eye colour. These distinct phenotypic characteristics are considered to be the result a copolymerisation of eumelanin and pheomelanin (mixed melanogenesis) (Shosuke Ito and Kazumasa Wakamatsu 2003), where the biochemical environment effects the quantities, ratio and distribution of these two types of melanin and thus cause the difference in a person's pigment (S. Ito and K. Wakamatsu 2011).

1.1.2 Melanin and pigmentation

Skin pigmentation is observed to follow an additive model of eumelanin from light to dark skin colour with pheomelanin quantity and ratio varying between skin colours. Although this eumelanin/skin pigmentation model appears to following a simple additive trend, the biochemical environment in which the emulation is produced differs between skin colours (S. Ito and K. Wakamatsu 2011). Notably, melanosomes from melanocytes of fair-skinned individuals (typically people of European descent) are more acidic, whereas dark-skinned individuals have more

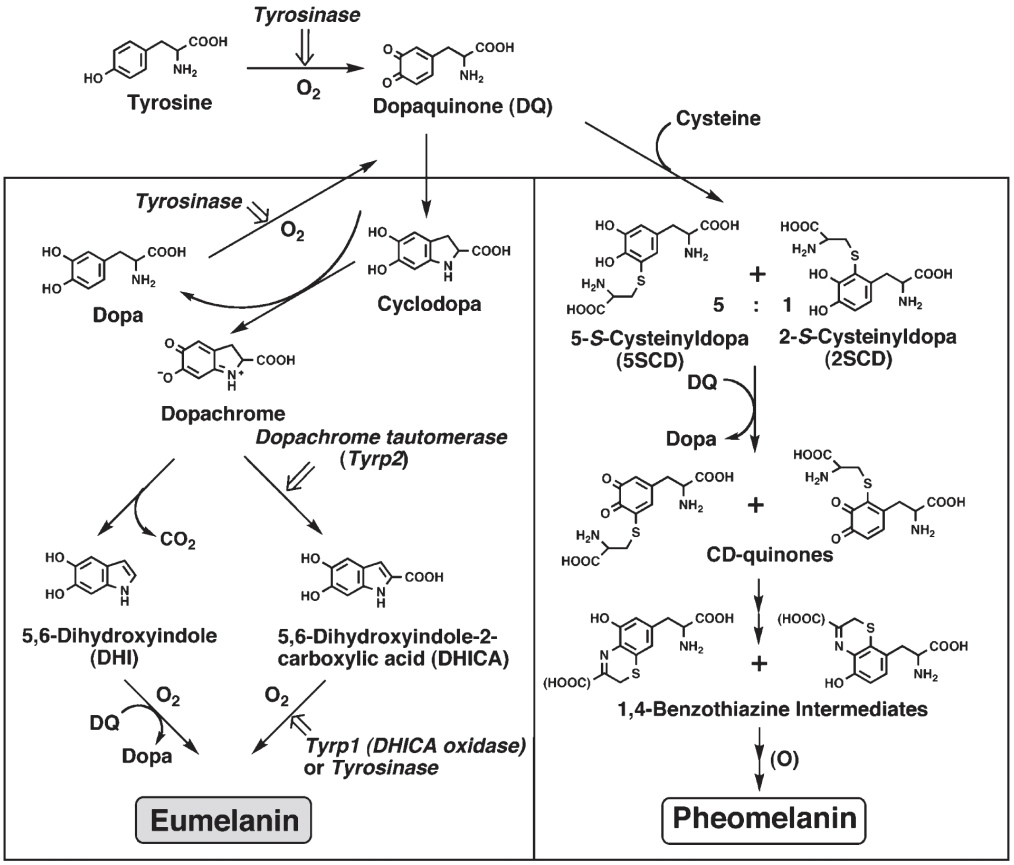


Figure 1.2: Biosynthetic pathways leading to eumelanin and pheomelanin production. Figure taken from Shosuke Ito and Kazumasa Wakamatsu (2011).

neutral melanosomes (Cassady and Sturm 1994; Bellono and Oancea 2014). This distinction effects one of the intermediary stages of eumelanin production, where dopachrome is gradually converted to 5,6-dihydroxyindole (DHI) or is rapidly converted by dopachrome tautomerase to form 5,6-dihydroxyindole-2-carboxylic acid (DHICA) in the presence of a lower pH, which results in two distinct eumelanin componds being produced: DHI-eumeumelanin and DHICA-eumeumelanin. When comparing between skin pigmentation and hair pigmentation, studies investigating melanin components of melanocytes have suggested skin pigmentation has a greater variation in eumelanin/pheomelanin ratios compared to hair pigmentation (Simon et al. 2009; S. Ito and K. Wakamatsu 2011).

For hair pigmentation, natural hair phenotypes are also considered to be copolymers of eumelanin and pheomelanin (mixed melanogenesis) (Shosuke Ito and Kazumasa Wakamatsu 2003), where the differing quantities, ratios, and distribution of these melanin determine the hair colour phenotype. Studies (Shosuke Ito, Nakanishi, et al. 2011; S. Ito and K. Wakamatsu 2011), in-

vestigating the levels of eumelanin and pheomelanin, in individuals of European descent, have identified similar ratios of eumelanin/pheomelanin between blond, light brown, medium brown, dark brown, and black hair phenotypes (between 82% to 96% in favour of eumelanin); and that total eumelanin levels are the defining aspect between these non-red hair colour phenotypes (4.7, 8.7, 10.4, 14.6, 22.2 μg/mg, respectively) (Shosuke Ito, Nakanishi, et al. 2011). Conversely, a distinct characteristic of the red hair colour phenotype is that there is a much higher ratio of pheomelanin (55%) and a lower overall eumelanin concentration (3.7 μg/mg) (Shosuke Ito, Nakanishi, et al. 2011; S. Ito and K. Wakamatsu 2011).

Notably, the blond hair phenotype is the lightest in colour with around 21% of the eumelanin of black hair (S. Ito and K. Wakamatsu 2011). This dilution of eumelanin and hair shape is believed to be the cause of a blond hair phenotype rather than from a contribution of pheomelanin (Morgan et al. 2018). Although this is believed to be the main attribute to blond hair, an over expression of *ASIP* has been linked to synthesis of only yellow pheomelanin in mice, and has been found to have an established role in the red hair phenotype in humans (Ollmann et al. 1998).

Although natural hair colours are considered copolymers of eumelanin and pheomelanin, hair greying is not. This is a natural age-associated process uniquely determined by distinct characteristics compared to other natural hair phenotypes (Jo et al. 2018). The primary cause of greying hair is the the loss of pigment in the hair shaft. Recent studies identify this loss of pigment to be a consequence of an overall and specific depletion of bulb and outer root sheath melanocytes of hair which may be caused by incomplete melanocyte stem cell maintenance (Jo et al. 2018). As hair greying is largely related to a lack of melanocytes (and thus a lack of overall melanin), and not due to a eumelanin and pheomelanin combination, it can be considered as a separate condition rather than an additional hair colour phenotypes. Thus, it is not appropriate to consider grey hair when considering an ordinal scale of hair colour when investigating the underlying genetics of natural hair colour.

1.1.3 Known pigmentary genes

Until recently, previous genome-wide association studies (GWAS) have identified only a limited number of loci associated with hair colour with a total of eleven genes associated (*MC1R, ASIP,*

OCA2, SLC45A2, KITLG, TYR, TYRP1, TPCN2, IRF4) (Beaumont et al. 2011; A. L. Cook et al. 2009; Bin et al. 2015; Kasamatsu et al. 2008; Jagirdar et al. 2014; C. Y. Li et al. 2004; Ambrosio et al. 2016; Chhabra et al. 2018). A recent study by Hysi et al (Hysi et al. 2018) performed a GWAS meta-analysis of hair colour on 300,000+ individuals of European descent from UK Biobank and 23andMe cohorts. In line with the polygenic nature of hair colour, their analyses identified over 120 loci genome-wide significantly ($P < 5 \times 10^{-8}$) associated with hair colour of which 107 were novel. Of these new loci, strong associations were found in genes whose mutation reportedly cause impairment of pigmentation (*EDNRB, MITF, HPS5, FGF5, TWIST2*); two solute carrier 45 A family genes (*SLC45A1, SLC45A2*); the forkhead box family genes (*FOX06, FOXE1*); and the sex-determining regions Y box genes (*SOX5, SOX6*). Of these newly identified genes, many have been reported to be involved in melanocyte development (*EDNRB, MITF, HPS5, FGF5, SEX5, TWIST2*) and highlight overlap in susceptibility loci of pigmentation and the originating cells of cutaneous melanoma. These analyses also explained heritable variation in hair colour with 34.6% for the red hair phenotype, 24.8% for the blond hair phenotype, and 26.1% for the black hair phenotype.

A further study by Morgan et al (Morgan et al. 2018) conducted GWAS of hair colour on 343,000 UK Biobank individuals with a focus on individual hair phenotypes and *MC1R*. The analyses identified 200 genetic variants associated with hair colour over 163 genes, of which 96 were previously reported by Hysi et al. Of the newly reported genes, several (*ERRF11, FRAS1, HOXC13, PADI3, KRTAP, PEX14, LGR4*) were found to be associated with blond hair and have been previously reported for hair growth and texture. Highlighting the the importance of hair shape on colour perception.

Although GWAS of hair colour have generally had larger sample size, and thus power to detect genetic associations, these studies have considered hair colour to be a single ordinal scale with an assumption that each hair pigment has the same underlying genetic architecture. Since the increase in pigmentary data, it is becoming increasingly viable to separate red and non-red hair into individual GWAS and retain power to detect loci of association for both.

In comparison to hair colour, current knowledge on the genetic architecture of skin colour and

tanning ability is limited. This is partly due to the population specific influence of genetic variants observed between distinct human populations (G. Chaplin and N. G. Jablonski 2013; Nina G. Jablonski and George Chaplin 2000). Typically, GWAS of skin colour are performed within homogeneous population groups, such as people of European descent (F. Liu et al. 2015), African descent (Crawford et al. 2017; Martin et al. 2017), South Asian (Stokowski et al. 2007), or Latin American descent (Adhikari et al. 2019). Identified associations of skin colour/tanning ability derived from GWAS conducted on these population groups highlight the extent of the population specific genetic influence, as the genetic architectures of skin colour and tanning ability are largely incongruous between these populations with only the six major loci identified across different population groups. As a result of this, it is possible to derive a population specific list of genetic variants that determine skin colour, but a general-human wide one would primarily contain a lists of genetic variants important to specific populations. Recent GWAS of skin colour conducted on people of European descent have identified these six major associated loci (*TYR, SLC45A2, IRF4, HERC2/OCA2, MC1R* and *ASIP*) (F. Liu et al. 2015); whilst a recent GWAS of tanning ability, using 175,000 UK Biobank participants of European descent, identified 20 associated loci (Visconti et al. 2018).

1.1.4 Pigmentation genes in a wider context

In a wider context pigmentation genes have been found to be associated with many other phenotypes. Commonly, a large number of cancer types have been associated with previously identified pigmentation genes. The most strongly linked of these is melanoma skin cancer with all six major pigmentation genes having a prior identified association. The link between melanoma and pigmentation can easily be identified by considering UV exposure leading to cutaneous damage and then to melanoma, where the lighter a person's pigmentation the more easily UV damage can occur on the skin. In a similar manner, non-melanoma skin cancer has also been widely been found to associate with the six major pigmentation genes and is believed to act through a similar exposure pathway. As well as melanoma and non-melanoma skin cancer, current literature has identified polymorphisms in *MC1R* are associated with a lower risk of prostate cancer (Weinstein et al. 2013). It's believed *MC1R* may influence prostate cancer development either directly, through genetic effects or regulatory mechanisms. Other cancers and diseases have also been linked to these pigmentation genes, such as lymphoid leukemia (*IRF4*, Berndt et al. 2016), lung carcinoma (*IRF4*, McKay et al. 2017), crohn's disease (*IRF4*,J. Z. Liu et al. 2015),

autoimmune disease (*IRF4*, F. Liu et al. 2015), Alzheimer's disease (*ASIP*, Herold et al. 2016) and parkinson's disease (*OCA2*, Latourelle et al. 2009).

In addition to diseases-based phenotypes, these pigmentation genes have been found to associate with other visual phenotypic traits. Unsurprisingly, overlapping with additional pigmentation-based phenotypes, such as eye colour (*OCA2*, *SLC45A2*, *IRF4*, and *TYR*) (Simcoe et al. 2021) and the presence of freckles (*ASIP*, *SLC45A2*, *MC1R*, *IRF4*, and *TYR*) (Jacobs et al. 2015). In addition to these pigmentation-based phenotypes, a common shared trait across these genes is balding, with *ASIP*, *MC1R*, and *IRF4* being identified as having a contributing affect (Pirastu et al. 2017). Recent GWAS have also identified an association between the appearance of skin aging and pigmentation genes (*SLC45A2*, *MC1R*, and *IRF4*) (Law, Medland, et al. 2017). Skin aging is believed to be acting through a UV damage pathway, where lighter skin is more likely to be damaged from UV exposure, causing epidermal thinning and a reduction in DNA repair.

1.2 Cutaneous Melanoma

At the basal layer of the epidermis sit the melanocytes, which are responsible for the production of pigment melanin through the absorption of ultraviolet radiation (UV). Although they have a large influence on a person's pigmentation, they make up a minority cell population within the epidermis with only 1500 melanocytes per square millimetre, and divide infrequently (typically less than twice a year). With UV stimulus, keratinocytes produce α-melanocyte stimulating hormone (α-MSH), which binds to the melanocortin 1 receptor (*MC1R*) on melanocytes signalling to induce eumelanin/pheomelanin synthesis (Williams et al. 2011) (See 1.1.1). The resulting melanin is transported from the melanocytes to nearby keratinocytes by the use of finger-like projections (Seiberg 2001). These keratinocytes use the melanin to protect it's nuclei from the mutagenic effects of UV. Once matured, keratinization occurs within these keratinocytes which causes the keratinocytes to anucleate and die. Thus, from these two mechanisms, the outer layer of skin is protected by both melanin pigmentation, and a layer of dead karatinocytes that act as a barrier to protect cells beneath from UV exposure.

Cutaneous melanoma (melanoma) begins in the melanocytes that reside in the basal layer of the epidermis. It is the deadliest form of skin cancer, and where the cancerous growth oc-

curs when genetically mutated skin cells multiply rapidly and form tumorous masses. If not treated appropriately, these tumorous masses can metastasise. As with common complex traits, melanoma can arise through a number of often interacting pathways such as high penetrance genetic variants (Aitken et al. 1994) (familial cases: around 10% of all melanomas), multiple additive low penetrance genetic variants (Law, Bishop, et al. 2015; Ransohoff et al. 2017; Landi et al. 2020), or environmental exposure to UV resulting in cutaneous damage (Autier et al. 1998).

1.2.1 Diagnosis, prognosis, and treatment of melanoma

Diagnosis and staging

Early classification methods for melanoma were based on where the tumour arose from, which can typically be from an existing naevus, or acquired melanocytic lesion. This region based classification system remained in place until the 1960s, where a classification system based on histological features was developed (R. A. Scolyer et al. 2011). This new system significantly improved the way melanoma was diagnosed, by being able to group melanoma cases into three categories: superficial spreading melanoma, lentigo maligna melanoma, and nodular melanoma, which are still broadly used today. This sub-grouping of melanoma provided the insight that it is actually a heterogeneous disease, where variants have different characteristics, prognosis, and that they cannot be treated in a uniform manner (Rebecca et al. 2012). Later classification systems have utilised the measured thickness of the tumour (the Breslow thickness) to assess the aggressiveness of the melanoma (Breslow 1970). Typically this measured thickness is then compared to a five stage Breslow classification system, where stage I (less than or equal to 0.75 mm thickness) is the least aggressive and stage V (greater than 3.0 mm) is the most aggressive form of melanoma. As the measured thickness is associated to aggressiveness of the tumour, it allows a simple prognosis of survival, and provides an insight into the likelihood of region or distant metastasis. More recently, newer staging systems have been created that make use of the wide availability of data on patient treatment and survival. The American Joint Committee on Cancer (AJCC) has been instrumental in the creation of the TNM (tumor, node, metastasis) staging system (R. A. Scolyer et al. 2011; Dickson and Gershenwald 2011) (Table 1.1). The AJCC staging system provides pathologists and clinicians a guideline for staging patients diagnosed with melanoma. By combining histologic attributes of the primary tumor (T), the presence and extent of regional lymph node disease (N), and the presence and extent of distant

metastasis (M), clinicians are able to assign patients a stage grouping that is strongly linked to survival and prognosis (R. A. Scolyer et al. 2011; Dickson and Gershenwald 2011; Bartlett and Karakousis 2015).

Tumour Staging	**Breslow Thickness**
T1	Less than or equal to 1.0 mm
T2	1.01mm to 2.0mm
T3	2.01mm to 4.0mm
T4	Greater than 4.0mm
Node Staging	**No. of metastatic nodes**
N0	There are no melanoma cells in the nearby lymph nodes.
N1	There are melanoma cells in one lymph node or there are in-transit, satellite or microsatellite metastases.
N2	There are melanoma cells in 2 or 3 lymph nodes or there are melanoma cells in one lymph node and there are also in-transit, satellite or microsatellite metastases.
N3	There are melanoma cells in 4 or more lymph nodes or there are melanoma cells in 2 or 3 lymph nodes and there are in-transit, satellite or microsatellite metastases or there are melanoma cells in any number of lymph nodes and they have stuck to each other (matted lymph nodes)
Metastasis Staging	**Site**
M0	The cancer has not spread to another part of the body.
M1	The cancer has spread to another part of the body.

Table 1.1: The American Joint Committee on Cancer TNM staging system.

Prognosis and prevalence

The prognosis for melanoma overall is good with 80-100% survival for five years after diagnosis at stages I or II. As with most cancers, survival decreases sharply at stages III and IV with 50% and 10% survival respectively for 5 years (Cancer Research UK 2016). Although the majority of melanoma can be cured with surgical removal, a large proportion relapse and develop an advanced disease with poorer prognosis. Thus, there is a pressing need to identify factors that could prevent melanoma occurrence as relapse rates are high. Melanoma has seen a substantial increase in incidence over the last four decades. Incidence of melanoma is almost exclusively in fair skinned populations, with only rare acral lentiginous melanoma affecting darker skinned people. Compared to other cancers, melanoma incidence is increasing faster than any other form worldwide (Lens and Dawes 2004). Australia has the highest incidence at 51.6 per 100 000 p.a. for males and 40.7 per 100 000 p.a. for females, while China has the lowest incidence at 0.21 per 100 000 p.a. for males and 0.17 per 100 000 p.a. for females. The UK is ranked 14th with an incidence rate of 8.3 per 100 000 p.a. for males and 11.3 per 100 000 p.a. for females (Bray et al. 2013). Comparing melanoma to non-melanoma skin cancers, melanoma is responsible for only 2.3% of all skin cancer cases, but is responsible for over 75% of skin cancer deaths in the UK (Corrie et al. 2014). A sharp rise in incidence has been observed, with incidence rates doubling for at least 4 decades, resulting in melanoma becoming the fifth most common cancer in the UK (Cancer Research UK 2014).

Treatment

As melanoma is a heterogeneous disease a number of treatment options are available for patients which vary on invasiveness and on melanoma characteristics. The most common treatment is surgical resection. This involves the surgical removal of the tumour and surrounding healthy tissue, and is typically followed by a lymph node biopsy for patients with a melanoma Breslow thickness greater than 0.8mm, or if the tumour has ulcerated (C. Lee et al. 2013). If melanoma cells are detected in the biopsy, the remaining lymph nodes are removed. Generally, this treatment route is selected for lower stages of disease progression, as metastasized tumours would require additional treatments.

Until recently, for metastatic cancer the only treatment option available to patients was chemother-

apy. The most common drug therapy to use being dacarbazine which is considered the standard care option for metastatic melanoma and offers a median survival from 5-11 months and a one year survival rate of 27% (C. Lee et al. 2013; Rebecca et al. 2012; Koller et al. 2016). Advances in our understanding of melanoma have provided other options for metastatic melanoma, which are targeted therapies, and immunotherapies. Both have become more prominent in recent years, and have had success in increasing the survival rates of melanoma. Multiple targeted therapies have been developed to take advantage of molecular defects present in melanoma. The most promising ones are a *BRAF* inhibitors, which shrink or slow the growth of metastatic melanoma in participants whose tumours have a *BRAF* mutation (Rebecca et al. 2012; Chapman et al. 2011). Although highly effective in around half of these patients, many of these patients go on to develop secondary resistance within a relatively short amount of time (R. A. Scolyer et al. 2011; Rebecca et al. 2012; Chapman et al. 2011).

The most recent treatment development for late stage melanoma are immunotherapies. Generally, these treatment options are designed to engage a patients immune system into detecting and fighting the disease. There are a variety of biological mechanisms that allow this, such as adoptive cell therapy, cancer vaccines, immunomodulators, monoclonal antibodies, or oncolytic virus. The most effect treatment option of metastatic melanoma are immune checkpoint inhibitors (N. Lee et al. 2016; Koller et al. 2016; Prieto et al. 2012). Typically this involves the introduction of antibodies against PD1, PD-1/2, and CTLA-4, which effectively block binding to the respective ligands, and the corresponding signal that causes immune tolerance of the tumour cells (N. Lee et al. 2016; Koller et al. 2016; Topalian et al. 2015; Prieto et al. 2012; Wolchok et al. 2010). By blocking these signals, it allows stimulation of an immune response, and thus increases patient survival rates. Although these established treatments are promising, there are still a significant subset of patients that do not respond to the treatment, and many patients who do respond but develop a secondary resistance to the treatment. Significant amounts of research is dedicated to detecting biomarkers that can predict how a patient will respond to treatment (N. Lee et al. 2016; Byrne and D. E. Fisher 2017; George et al. 2017).

1.2.2 Risk factors for melanoma

It has been shown that a person's genetic susceptibility and distinct patterns of sun (UV) exposure contributes to the increase in melanoma incidence (Crombie 1979; Holman and B.

Armstrong 1984). Previous studies investigating the role of sun exposure on melanoma risk have found that intermittent intense sunlight exposure is a risk factor for melanoma, and that the age at which this occurs is important, with those at adolescence or early adulthood considered to be the most sensitive (Veierod et al. 2003). Although intense sunlight exposure has been found to be a risk factor of melanoma, long-term continuous exposure has been found to have little or no effect on risk (Rosso et al. 2008). The major genetically mediated risk factors that have been highlighted are: skin pigmentation, naevi count and telomere length (Meyle and Guldberg 2009; Ribero et al. 2015; Nan et al. 2009). Further research into these risk factors is therefore vital to fully understand the relationship between them and melanoma, with the aims to identify persons at higher risk and to promote preventative measures for those most vulnerable.

High-penetrance melanoma susceptibility genes

Although most melanomas occur with no family history, multi-case families do occur with a dominant pattern of susceptibility. These have been researched extensively. Through these family studies, high-penetrance melanoma susceptibility genes have been identified. Germline mutations in the *CDKN2A* locus and *CDK4* were found in melanoma-prone families (Kamb et al. 1994). A study consisting of 466 high-risk families with three or more melanoma cases found 66 different *CDKN2A* mutations for the recruited families (Goldstein et al. 2006). Median age at melanoma occurrence was lower for the families carrying the *CDKN2A* mutations compared to patients with wildtype *CDKN2A*. More recent familial studies have identified further high-penetrance melanoma susceptibility genes: *BRCA2* (The Breast Cancer Linkage 1999), *TERT* (Harland et al. 2016), and *POT1* (Robles-Espinoza et al. 2014). In addition to these, germline *BAP1* (Harbour et al. 2010) mutations have been found in families with uveal melanomas while somatic mutations have been found in metastasising uveal melanomas.

Low-penetrance melanoma susceptibility genes

Many epidemiological studies have identified pigmentation traits as a risk factor for melanoma, such as: poor tanning response, red hair, freckling, and fair skin (Bliss et al. 1995). The genetic determinants of these traits have been studied extensively through systematic analysis of candidate genes and GWAS. A large number of significant SNPs have been discovered, most notably SNPs found in the melanin synthesis pathway (Meyle and Guldberg 2009). The genes associated with these SNPs have been highlighted as low-penetrance melanoma susceptibility

genes. These genes are: *MC1R* (Rees and Healy 1997), *ASIP* (Gudbjartsson et al. 2008; Kevin M Brown et al. 2008), *MITF* (Yokoyama et al. 2011), *TYR* and *TYRP1* (Gudbjartsson et al. 2008). More recent studies, using a GWAS approach, validated these findings (Bishop et al. 2009; Falchi et al. 2009; Gerstenblith et al. 2010) excluding *TYRP1* which was not found to be genome-wide significant. They also highlighted additional regions of interest (Barrett et al. 2011): *ATM*, *MX2*, around *CASP8*, and close to *CCND1*. A meta-analysis approach, that used five previous GWAS studies, replicated the associations between significant GWAS SNPs and melanoma risk; SNPs in the following gene regions were found to be significant: *CDK10*, *AFG3LIP*, *SLC45A2*, *TYR*, *MTAP*, *PIGU*, *PLA2G6* (Kocarnik et al. 2014; Landi et al. 2020). In total, population-based GWAS of melanoma have identified 68 SNPs from 54 loci strongly associated with melanoma (Landi et al. 2020). The majority (76%) of these risk SNPs are jointly associated with at least one of three major heritable pathways to melanoma risk: pigmentation, naevus count, and telomere length.

Sample power issues and alternative approaches

Due to the inherent study power issues present in melanoma GWAS, other approaches have been adopted to supplement power. Most commonly, transcriptome-wide association studies (TWAS) have recently been employed as an approach that can link GWAS and gene expression studies to further identify genes associated with complex traits such as melanoma. Typically, this is achieved by imputing the *cis*-genetic components of gene expression (eQTLs) from an independent dataset into a larger GWAS study (Nicolae et al. 2010; Schadt 2006; Gusev et al. 2016; Z. Zhu et al. 2016). The primary benefits of conducting this analysis, is that only summary level data is required from the GWAS, and that it provides an improvement on statistical power, as external information is included into the model in the form of eQTLs. Recently, a study utilising this methodology conducted a TWAS on melanocyte eQTLs with a melanoma GWAS variant set and uncovered four novel susceptibility loci, where imputed expression levels of five genes (*ZFP90*, *HEBP1*, *MSC*, *CBWD1*, and *RP11-383H13.1*) were associated with melanoma at genome-wide significant levels (p-value $< 5 \times 10^{-8}$) (Zhang et al. 2018).

1.3 Genetics

1.3.1 Deoxyribonucleic Acid

DNA is the hereditary material in humans and almost all other organisms. Found inside the nucleus of cells, it consisted of chained nucleotides which form a stand of DNA. These nucleotides are made up from a phospate group, a sugar group, and one of four types of nitrogen base, which are: adenine (A), thymine (T), guanine (G) and cytosine (C). The order, or sequence, of these bases determines what biological instructions are contained in a strand of DNA. The DNA is then tightly coiled many times around proteins called histones that support its structure to form a chromosome. Each chromosome has a constriction point called the centromere, which divides the chromosome into two sections or arms: the short section (p arm) and the long section (q arm). The human genome consists of three billion nucleotide pairs across 23 pairs of chromosomes. A gene is defined as a sequence of nucleotides that encode the synthesis of a gene product, such as a protein.

1.3.2 Single nucleotide polymorphisms

SNPs are polymorphisms that are caused by point mutations that give rise to different alleles containing alternative bases at a given position of nucleotide within a locus. The abundance of SNPs in the genome is high with around one occurring every 1000 nucleotides on average, equating to around four to five million SNPs sites in the human genome. Due to the high density, SNPs serve as the predominant marker type in genetic studies, and can occur in noncoding regions of the genome as well as in genes (both exons and introns). SNPs vary across human populations resulting in the allele frequency differing between these population groups, the MAF represents the least frequent allele present in a population group.

1.3.3 Population Stratification

Population stratification is defined as the differences in allele frequency between population groups. It's believed humans left Africa around 150,000 to 190,000 years ago, and groups expanded into far reaching part of the globe (McDougall et al. 2005; White et al. 2003). This large migration (and sequential geographical isolation of sub-populations), as well as mating between population groups and other hominins have resulted in allele frequency diversifying across these population groups (R. E. Green et al. 2010; Meyer et al. 2012). These varying allele frequen-

cies have been detected when observing genotypes of different population groups (Novembre et al. 2008), and can introduce confounding factors into association studies. These confounding factors can cause the detection of false-positive association. Commonly, principal components of the genetic variation across a study group are included into any genetic association tests to compensate for this population structure confounding.

1.3.4 Linkage disequilibrium

One of the predominant characteristics of SNPs is the non-random association between two or more alleles at differing points across the genome. This non-random link between two point on the genome is known as linkage disequilibrium (LD). When alleles are in LD, haplotypes do not occur at the expected frequencies, and can appear more regularly together (in positive LD), or less regularly together (negative LD). Multiple factors affect the LD between alleles, such as the rate of genetic recombination, genetic distance, mutation rate, genetic drift, the system of mating, population structure, and genetic linkage. As so many factors affect LD, it is able to provide an insight into the past and it constrains the potential response to both natural and artificial selection. More specifically to genome-wide based studies, LD patterns can create difficulties in discovery rates of causal variants due to surrounding alleles also appearing significantly associated to the studied trait.

1.4 Data Sources

1.4.1 UK Biobank

UK Biobank is a UK-wide multi-site cohort study established by the Wellcome Trust medical charity, the Medical Research Council, Department of Health, Scottish Government, and North-west Regional Development Agency (Sudlow et al. 2015). It investigates genetic predisposition and environmental exposures with the aims to improve: prevention, diagnosis and treatment for a wide range of illnesses. More specifically; UK Biobank holds health outcome, phenotypic, exposure, and genotypic data on its 500,000+ participants. These 500,000+ participants were recruited at baseline, aged 40-69 years old, between the years 2006-2010 from across the UK. Following this, a second stage assessment of 20,000 previously recruited participants was conducted between August 2012 and June 2013 to assess reproducibility. Both baseline and secondary assessments were used in part self-reported questionnaire methods for phenotypic data collection,

with cancer incidence confirmation being acquired from the cancer registry, and DNA extraction for genetic data collection.

The baseline information in UK Biobank continues to be extended. For example, repeat assessments are planned to be conducted in subsets of the cohort every few years, to enable calibration of measurements, adjustment for regression dilution, and estimation of longitudinal change. Additional information such as physical activity, naevus count, carotid ultrasound and a whole body dual-energy X-ray absorptiometry of the bones and joints are planned to be implemented into the dataset. In addition to expanding the current baseline information, all participants of the UK Biobank have provided consent for follow-up through linkage to their health-related records. As of May 2018, there were over 14,000 deaths, 79,000 participants with cancer diagnoses, and 400,000 participants with at least one hospital admission (Sudlow et al. 2015). These have been recorded through national datasets including primary care, screening programs, and disease-specific registries, as well as asking participants directly about health-related outcomes.

UK Biobank genotyped data

The UK Biobank genetic data contains genotypes for 488,377 participants. These were assayed using two very similar genotyping arrays. A subset of 49,950 participants involved in the UK Biobank Lung Exome Variant Evaluation (UK BiLEVE) study were genotyped at 807,411 markers using the Applied Biosystems UK BiLEVE Axiom Array by Affymetrix. Following this, 438,427 participants were genotyped using the closely related array (812,428 markers) that shares 95% of marker content with the UK BiLEVE Axiom Array. UK Biobank provides a lists of participants whose genetic results should be excluded on the basis of poor performance or close relatedness. Through conducted analyses in this thesis, non-European outliers were identified using the same approach that UK Biobank apply to their "Caucasian" definition: the 'aberrant' routine in R 29 was applied to PCs 1&2, 3&4 and 5&6, but anyone declaring themselves to be White was included in this analysis (where UK Biobank automatically exclude "Irish" and "any other white background"); the lambda parameter used was 100.

UK Biobank variables

Upon recruitment, UK Biobank participants are interviewed by trained interviewers, and provide detailed information on a wide array of phenotypes. More specifically to the thesis, self-reported pigmentation information and cancer history is provided. For pigmentation four questions were completed via touch screen and considered multiple aspects of a person pigmentation: their hair colour, skin colour, tanning ability, and the number of childhood sunburns (Table 1.2). Prior cancer diagnoses were also recorded by touch screen, and were cross-validated by linking participant information with the UK Cancer Registry.

Pigmentary question	**UK Biobank variable ID**	**Sample size**	**Response options**
What best describes the colour of your skin without tanning?	1717	370,260	Very fair 1; fair 2; light olive 3; dark olive 4; brown 5; black 6
What would happen to your skin if it was repeatedly exposed to bright sunlight without any protection?	1727	367,229	Get very tanned 1; Get moderately tanned 2; Get mildly or occasionally tanned 3; Never tan, only burn 4
Before the age of 15, how many times did you suffer sunburn that was painful for at least 2 days or caused blistering?	1737	280,306	Number input keypad
What best describes your natural hair colour? (If your hair colour is grey, the colour before you went grey)	1747	352,662	Blond 1; red 2; light brown 3; dark brown 4; black 5

Table 1.2: The identity of the UK Biobank pigmentation variables and their possible responses.

1.4.2 Leeds melanoma cohort

The Leeds Melanoma Cohort (LMC) is a population-based melanoma study. Participants were recruited on incidents of melanoma that were diagnosed between September 2000 and December 2012 from the Yorkshire county situated in the northern region of England, UK. Cases were confirmed by clinicians, pathology registers, and the Northern and Yorkshire Cancer Registry. All controls and the first 960 cases recruited were examined by trained interviewers who conducted an examination of the skin and recorded naevi count information. Throughout the study, except for 18 months, recruitment was restricted to melanoma diagnoses with a Breslow thickness of at least 0.75mm. Controls were selected as specifically melanoma-free and were matched to cases by age and GP surgery. Both cases and controls were genotyped on the Illumina Infinium HumanOmniExpressExome (San Diego, CA, USA) array. Participants consented to studies approved by the Northern and Yorkshire Research Ethics Committee.

1.4.3 Use of Data Sources in the Thesis

Throughout this thesis the UK Biobank and LMC have been the primary datasets used to explore pigmentation and melanoma genetics. To reduce overfitting and maximise the utilisation of the UK Biobank, the dataset was was split into two independent sub-samples post standard QC (2.3.4). Firstly, a melanoma-based sample which consisted of the 5,301 UK Bioabnk participants that had a self-report or pathological confirmation of melanoma and 21,204 control samples. The second, a pigmentation-based sample which consisted of all remaining UK Biobank participants (460,049) that passed standard QC. Thus leaving three independent datasets: the pigmentation-based UK Biobank sub-sample, the melanoma-based UK Biobank sub-sample, and the LMC dataset. These datasets were used in the following analysis for each chapter:

Chapter 3: The full UK Biobank dataset was used for pigmentation variable QC, the reliability of cancer matching and factor analysis.

Chapter 4: The pigmentation-based UK Biobank sub-sample was used for each pairwise hair colour GWAS and complementary analyses, the non-red vs red hair colour GWAS and complementary analyses, both linear GWAS (including and excluding red hair colour), the hair colour-based polygenic risk scores (PRS), and the pairwise skin colour and tanning ability GWAS.

Chapter 5: The pigmentation-based UK Biobank sub-sample was used for the principal components

analysis used top derive overall pigmentation score and their subsequent GWAS, and the five pigmentation measure GWAS.

Chapter 6: The pigmentation-based UK Biobank sub-sample was used to create PRS from the four pigmentation measure GWAS and overall pigmentation score GWAS. The melanoma PRS was derived from a melanoma meta-analysis GWAS which contained the melanoma-based UK Biobank sub-sample. The full UK Biobank dataset was used to select the optimum significance level for each pigmentation PRS when modelling melanoma risk. The LMC was used to validate the performance of each derived PRS.

Chapter 7: For both joint-analysis methodologies, the melanoma GWAS summary statistics were derived from a melanoma meta-analysis that contained the melanoma-based UK Biobank sub-sample, and the pigmentation GWAS summary statistics were derived from the pigmentation-based UK Biobank sub-sample.

Chapter 8: All penetrance and epistasis analyses were conducted on GWAS summary statistics derived from the pigmentation-based UK Biobank sub-sample.

1.5 Aims of the thesis

1.5.1 Aims & Objectives

Aims

I. Quantify the reliability of melanoma cancer self-reports

II. Better understand the relationship (phenotypic self-report and genetic architecture) between distinct pigmentary characteristics

III. Better understand the genetic architecture of a person's pigmentation (and how each pigmentation trait's genetic architecture overlaps) and how it related to the genetic architecture of melanoma (via the risk pathway)

IV. Improve risk prediction models for melanoma by utilising risk pathway information

V. Increase power to detect melanoma risk loci by utilising pigmentation information

VI. Detect epistasis between *MC1R* and other genes across pigmentation phenotypes.

Objectives

1. Conduct a factor-analysis to assess the underlying correlation across UK Biobank pigmentation variables.
2. Create a methodology-pipeline that matches UK Biobank melanoma self-reports with the UK Cancer Registry records.
3. Create and overall pigmentation score by conducting a principal components analysis on UK Biobank pigmentation variables
4. Determine the genetic architecture of distinct pigmentation characteristics by conducting GWAS for all UK Biobank pigmentation variables, and overall pigmentation score variables.
5. Assess the overlap in genetic architectures across each pigmentation measures and melanoma
6. Create PRS from pigmentation GWAS results to predict melanoma, and compare performance with melanoma derived PRS
7. Create an analysis-pipeline to meta-analysis melanoma and pigmentation GWAS summary statistics to increase melanoma detection power.
8. Quantify the penetrance of *MC1R* for distinct pigmentation characteristics.
9. Model epistasis effects *MC1R* may be having on other genes for distinct pigmentation characteristics.

Chapter 2

Methods

2.1 Positive predictive value and Sensitivity

When evaluating cancer self-reports and confirmations in Chapter 3, Positive predictive value (PPV) (a measure of specificity) and sensitivity statistics are used. PPV is the proportion of positive results that are true positives. This is calculated using the following equation:

$$PPV = \frac{\text{Number of true-positives}}{\text{Number of true-positives} + \text{Number of false-positives}} \tag{2.1}$$

Similarly, sensitivity is defined as the proportion of positives that are correctly identified. This is calculated using the following equation:

$$\text{Sensitivity} = \frac{\text{Number of true-positives successfully identified}}{\text{Number of true-positives}} \tag{2.2}$$

2.2 Principal Components Analysis and Factor Analysis

2.2.1 Principal components analysis

Principal components analysis (PCA) reduces the complexity of multi-dimensional data. This is achieved by producing summary components that maximise the overall variation in the data captured, derived from underlying trends and patterns in the analysis data. PCA is unsupervised and requires no prior knowledge. It utilises the correlation structure between variables and minimises the distances between the data and their projection onto principal components (PC)

to produce linear combinations of the original variables. This is different from linear regression which minimises the distance between the response variable and predicted value. The first PC captures the maximum amount of variation in the data sample, with the second PC being orthogonal to the first and capturing the maximum of the remaining variation.

In genetic studies PCA captures genetic variation due to ancestry across a population sample when this effect is sufficiently big. Individuals with similar ancestry are expected to have similar values for the first few PC from the analysis (Zou et al. 2010; Price et al. 2006). This aspect of PCA on genetic data has been demonstrated in European samples where the first two PCs were able to accurately map individuals to their European country of origin for whom all four grandparents originated in the same country (Novembre et al. 2008).

2.2.2 Factor Analysis

Factor Analysis (FA) is a multivariate statistical method. It is widely used in the social sciences where variables of interest are not directly measurable but can be recorded via latent variables (related variables that can act as indicators for the variable of interest). The fundamental understanding of FA is that, given a set of observed variables, it is assumed that they are linearly related to a smaller number of unobservable factors (latent variables). Much like PCA, FA uses correlations between variables and minimises the distances between the data and their projection onto the latent variables. FA falls into two distinct analysis plans: an exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). Although the underlying mechanics of these methods are the same, the objectives of the approaches differ. EFA is considered unsupervised and requires no prior knowledge; the aim of the analysis is to determine the optimum number of latent variables and their relationship, to maximise model fit. Conversely, CFA is conducted on a pre-determined model, with the number of factor loadings already decided. This approach is generally favoured as it allows the inclusion of supplementary information that may not be captured in a factor model derived from EFA. The underlying model followed for both FA

approaches can be written in matrix form as:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda_{11} & \lambda_{21} & \cdots & \lambda_{m1} \\ \lambda_{12} & \lambda_{22} & \cdots & \lambda_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{1n} & \lambda_{2n} & \cdots & \lambda_{mn} \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_m \end{pmatrix} + \begin{pmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_n \end{pmatrix} \tag{2.3}$$

where x are observed variables; δ are the error terms of observed variables; λ are factor loadings; ψ are latent variables; n is the number of observed variables; and m is the number of latent variables.

As FA minimises the distance between the data and their projection onto the latent variables, rotations of the data are paramount to achieve this. Many rotation methods are available to be used depending on model criteria, such as type of data used or model interpretation. The most commonly used rotation is the Varimax method (Kaiser 1958). This method is an orthogonal rotation that minimises the number of variables that have high loadings on each latent variable, which benefits the interpretation of the latent variables. Similar to this is the Quartimax method (Jackson 2005) which minimises the number of factors needed to explain each variable. This benefits the interpretation of the given variables. A combination of these methods leads to a Equamax method (Kaiser 1974), which minimises the number of variables that have high loadings and the number of latent variables needed to explain the variable. In addition to orthogonal rotations, non-orthogonal methods are available, such as the Oblimin method (Clarkson and Jennrich 1988), which allows for latent variables to be dependent and correlated to each other.

To determine model fit in FA three model fit statistics are utilised. Firstly the Tucker-Lewis index (Tucker and Lewis 1973) which summarises the discrepancy between the χ^2 value of the hypothesised model and the χ^2 of the null model. Typically this statistic ranges between 0-1 and a threshold of > 0.95 is an indicator of good model fit (Hu and Bentler 1999). The second is the root mean square error of approximation (RMSEA) which summarises the discrepancy between the hypothesised model and null model with optimally chosen parameter estimated. The RMSEA ranges between 0-1 and a value < 0.06 is considered to indicate good model fit

(Hu and Bentler 1999). Finally the third model fit statistics used is the χ^2 test which indicates the difference between the observed and expected covariance matrix of the observed variables.

2.2.3 Polychloric correlations

One of the fundamental characteristics across both FA and PCA, is the use of correlation to understand the relationship between observed variables. As the UK Biobank pigmentation data is ordinal in nature, and the mean and covariance for ordinal data is meaningless, Pearson's correlation is not an appropriate method to calculate correlation in the FA and PCA (Jaehwa Choi et al. 2010). Many alternative correlation calculations for ordinal data exist, such as Kendall's correlations, Spearman's and polychoric correlation, that are able to determine correlation between variables whilst not relying on mean and covariate information. Previous studies (Vittadini 1999) evaluating these methods have consistently found polychoric correlations provide the closest estimation of correlations between ordinal variants. The method works by assuming each ordinal variable has an underlying continuous Normally distributed latent variable, which is linked by a monotonic relationship.

2.2.4 Principal Components Analysis and Factor Analysis in the Thesis

CFA and EFA are conducted in Chapter 3 of the thesis,with the aim of understanding the underlying relationship between pigmentation traits. The analysis is conducted under the assumption that pigmentary measures collected by the UK Biobank are distinct questions measuring the same thing: a person's overall pigmentation. Thus it was optimal to consider this an FA model, where overall pigmentation is the latent variable, rather than use other methods such as a PCA. To investigate this assumption, an EFA was conducted to determine how many latent variables may underlie the UK Biobank pigmentary traits, and a CFA was conducted to provide factor loadings for the latent pigmentation model determined by the EFA. To produce an overall pigmentation score which included each pigmentation measure, PCA was conducted in preference to FA. This method creates PCs that are linear combinations of observed variables, in this case pigmentation measures.

2.3 Genome-wide Association study

A combination of the completion of the sequence of the human genome (Lander et al. 2001), continued improvement of SNP genotyping technology in terms of accuracy and costs (Christensen et al. 2015), and the increasing size of public databases of genetic variation (Sachidanandam et al. 2001), have led to GWAS becoming commonplace as a means of identifying the genetic factors influencing susceptibility to common and complex diseases. This approach relies on conducting individual tests for each genotyped SNP across a study sample group, which allows for large amounts of the genome to be surveyed for associated genetic variants. As no underlying assumptions are made about where in the genome associated variants may lie, an unbiased and comprehensive understanding of the genetic architecture of a trait is achieved, requiring no prior knowledge regarding the function or location of causal genes.

Although one of the primary strengths of GWAS is its agnostic approach to testing the genome, this leads to concern about the number of SNPs to be tested. As genotyping arrays have increased the number of variants included in studies, the number of independent tests rises which also increases the rate of false-positive associations. Therefore, a balance between power and type 1 error is required. Typically, to balance these two aspects a p-value threshold of 5×10^{-8} (Jannot et al. 2015) is selected, based on performing a naive Bonferroni correction for the number of independent common SNPs across the human genome (Xu et al. 2014). This p-value threshold is widely considered to be overly conservative due to the assumption that all genetic variants tested are independent of the one another. Over time, use of this p-value threshold has become standard practice, although recent research (Fadista et al. 2016) has found this significance threshold to not cope with the rarer variants now found in genotyping arrays. Thus it has been proposed that a more stringent threshold is required for lower frequency variants: $p < 1 \times 10^{-6}$, 7×10^{-7}, 5×10^{-7}, and 3×10^{-7}, for MAF>5%, MAF>1%, MAF>0.5% and MAF>0.1%, respectively.

2.3.1 Methods

TTwo risk association models are commonly implemented for GWAS depending on the measured trait being considered. Logisitic regression is implemented to test for potential association between genetic variants and a dichotomous dependent variable, such as the case/control status

of a participant. These dichotomous data follow the logistic curve pattern found in equation 2.4.

$$P = \frac{e^{a+bX}}{1 + e^{a+bX}} \tag{2.4}$$

where P is the probability of an individual being a case (1), e is the base of the natural logarithm, and a and b are model parameters.

Equation 2.4 can be rearranged to use the logistic link function $log(\frac{P}{1-P}) = Y$ to create the following risk model:

$$Y = log(\frac{P}{1-P}) = a + bX \tag{2.5}$$

where where P is the probability of a case (1), Y logit function, X is an independent variable value, a and b are model parameters, and ϵ is an error term in the model.

Equation 2.5 can also be expanded to include multiple independent variables (covariates) to form $Y = a + b_1X_1 + b_2X_2 + ... + b_nX_n$: where n is the number of included variables. In the case of GWAS, for equation 2.5 logit function is Y, a is a model constant, X modelled SNP status or covariate variables, and b is the effect size the modelled SNP is having on the study trait.

The second risk association model commonly used in GWAS is linear regression. Linear regression is used to model the relationship between independent variables and a continuous dependent variable. This model assumes that the relationship between independent and dependent variables is linear relationship as in equation 2.6:

$$Y = a + b_1X_1 + b_2X_2 + ... + b_nX_n \tag{2.6}$$

where Y is dependent variable outcome value, Xi are the independent variables, a and b are model parameters, and n is the number of independent variables included in the model.

2.3.2 Genetic inflation factor and QQ plots

Population structure, relatedness in the sample, trait polygenicity, and low MAFs of tested SNPs can cause spurious associations in GWAS. Relatedness, low and MAF are easy to detect

but excess significance due to population stratification or polygenicity is harder to detect and correct for where necessary. Such inflation of test statistics can be identified through QQ plots and the genetic inflation factor. These measures assume that the vast majority of SNPs are not associated with the trait being studied and so follow the null distribution. However, differences in allele frequencies due to population stratification will affect all SNPs and so the test statistics will be inflated throughout the genome. The genetic inflation factor is defined as the ratio of the median of the empirically observed distribution of the test statistic (χ^2) to the expected median of the null hypothesis χ^2 distribution. If the genetic factor is > 1 this indicates inflation in association for the GWAS results. A QQ plot is a graphical representation of this inflation across the whole genome: They are produced by plotting the observed p-values largest to smallest and plotted against expected values of the null hypothesis χ^2 distribution. These two methods have been conducted throughout the thesis as standard post-GWAS analyses. However with very large sample sizes very small genuine effects across the entire genome may be detected and QQ plots and the genetic inflation factor do not distinguish between inflation due to population stratification, which we want to correct for, and polygenicity, which we do not.

2.3.3 LD Score regression

As stated previously (2.3.2), trait polygenicity, and confounding biases, such as population structure or relatedness in the sample, can both cause inflation in the distribution of test statistics for GWAS results. Although these factors are known, and as a whole quantifiable, their individual contributions are not easy to disentangle. Inflation caused by true polygenic signals are an expected and welcome genetic association, however, the inflation caused by confounding biases can contribute to an increase in false-positive associations. LD score regression allows the estimation of the true polygenic signals and stratification effect by considering the LD between each SNP's test statistics in GWAS results and using this to partition the inflation of test statistics (Brendan K Bulik-Sullivan, Loh, et al. 2015). This is achieved by regressing the χ^2 test statistics on an LD score derived for each SNP using the sum of r^2 values between near by SNPs (defined by a genomic distance window). As the LD score is derived from r^2 between SNPs, inflation from population stratification or relatedness will not correlate with the LD score. Therefore when regressing the χ^2 statistics on LD score, the intercept from the regression will estimate the mean contribution of the confound bias to the inflation in the test statistics, and the slope estimates the polygenic signals. Genetic correlation between two traits are calculated using a cross-trait

LD score regression (Brendan K Bulik-Sullivan, Finucane, et al. 2015), where combined test statistics are regressed on LD score for both traits, and a genetic correlation value is derived from this regression model.

2.3.4 QC of UK Biobank dataset

Within UK Biobank, biological samples were available for genetic analysis from 488,000 participants. The majority of participants were genotyped using a purpose designed UK Biobank Applied Biosystems Axiom array assessing 826,000 SNPs and indels. The quality control and imputation approaches applied are described elsewhere (Sudlow et al. 2015). UK Biobank provides a list of participants whose genetic results should be excluded on the basis of poor performance or close relatedness; these persons were excluded from analysis conducted in this thesis. Non-European outliers were identified using the same approach that UK Biobank apply to their "Caucasian" definition: the 'aberrant' R-package (Bellenguez et al. 2012) was applied to PCs 1&2, 3&4 and 5&6, but anyone declaring themselves to be White was included in this analysis (where UK Biobank automatically exclude "Irish" and "any other white background"); the lambda parameter used was 100. Imputed genotypes were used from UK Biobank and the same QC was applied to select 'Europeans' as UK Biobank did for their 'Caucasian' definition (Field 22006), but included those who said they were 'White - Irish' or 'White -Any other white background' as well as 'White - British'.

2.3.5 GWAS in the Thesis

GWAS have been utilised extensively throughout the analysis conducted in the thesis as the standard procedure for understanding the genetic architecture of a trait (pigmentation and melanoma). For the main GWAS, both logistic and linear regression have been conducted: melanoma, red hair status, and childhood sunburn number utilising logistic regression, and skin colour, non-red hair colour, and tanning ability using linear regression despite these traits having been recorded as ordinal variables. It is understood that the greater the number of categories in an ordinal scale, the more it will resemble a continuous characteristic when a linear-regression will be a suitable assumption for the data (Borgatta 1968; Bollen and Barb 1981). In the case of the thesis: non-red hair colour had four categories, skin colour had six categories, and tanning ability had four categories. Standard QC was conducted before each GWAS (see 2.3.4). GWAS were conducted in the dedicated software PLINK 2.0 using the -glm command for logistic-based

regression and -linear command for linear-based regression. GWAS summary statistics were collated in R for post-GWAS analysis.

One of the limitations of GWAS is the computational resource required to successfully conduct an analysis. This has been a common theme throughout the history of GWAS, and remains to be as the SNP arrays used for genotyping study participants have steadily increased in size. The cause of this computational burden is the number of models required. Typically, GWAS make use of simple relationship models (linear models or generalised linear models), but have to conduct these simple models for each SNP included in the study (commonly in the millions). To counteract this computational burden, dedicated software is available that can efficiently and rapidly fit large numbers of models. In addition to this, as these models are independent, it is possible to break up genomes into smaller clusters, that can be run independently of one another. To minimise this burden, these two solutions were adopted for all GWAS conducted in this thesis.

2.4 Joint-analyses of GWAS summary statistics

Joint-analysis of GWAS summary statistics is an increasingly popular methodology to increase study power in complex genetic diseases and traits (Bakker et al. 2008). It provides an efficient strategy for detecting low-penetrance variants with subtle effect sizes by combining summary level statistics of multiple GWAS to increase study power (Skol et al. 2007). The method of how these GWAS summary statistics are combined varies across approaches, but largely follows a design where study contribution is calculated through a weighted averaged across included studies.

2.4.1 METAL

METAL is a computationally efficient tool for meta-analysis of GWAS summary statistics. It's primary approach is to utilises a sample size weighted joint-analysis method that combines multiple GWAS summary statistics and weights their contribution in the analysis on their relative sample size compared to the other included studies. METAL allows for the combination of different traits and effect scales by utilising the Stouffer method which only requires p-value, sample size and estimated direction of effect to meta-analyse each study. METAL converts the

direction of effect and p-value observed in each study into a signed Z-score, so large negative Z-scores indicate a small p-value and an allele associated with lower disease risk or quantitative trait levels, whereas large positive Z-scores indicate a small p-value and an allele associated with higher disease risk or quantitative trait levels. Z-scores for each allele are combined across traits in a weighted sum, with weights proportional to the square root of the sample size for each study. This method is particularly useful for combining GWAS summary statistics that have an unequal number of cases and controls as results are weighted on effective sample size.

2.4.2 MTAG

The primary idea of MTAG is that when GWAS estimates from different traits are correlated, the effect estimates from each trait can be improved by appropriately incorporating information contained in the GWAS estimates from the other traits. Generally, correlation between GWAS arises from two scenarios. Firstly, the traits are genetically correlated, and thus the underlying effect sizes of the SNPs are also correlated across the traits. Secondly, the estimation error of SNPs' effects may be correlated, which could be the result of nonzero phenotypic correlations due to either sample overlap, biases in the estimates from population stratification or cryptic relatedness. MTAG is able to boost statistical power by incorporating information about these two sources of correlation by estimation through cross-trait LD score regression (2.3.3), which allows for a genetic correlation-based weighted estimate of effect across tested traits.

2.4.3 GWAS-PW

To assess similarities in genetic architecture of distinct ohenotypic traits, GWAS-PW uses a Bayesian approach to assess pairwise GWAS summary statistics and determine the probability of pre-defined genomic windows's association with one, both, or neither of the pairwise traits. This is achieved by calculating Z-values (using $Z = \frac{log(OR)}{SE}$) and variance (using $V = SE^2$) for each SNP that is present in both GWAS. GWAS-PW then estimates the posterior probability of association (PPA) of each of four possible models: model 1 (region is associated solely with trait 1), model 2 (trait 2 alone), model 3 (Pleiotropic), model 4 (co-located but independent) and model 5 (no association found with either trait). GWAS-PW performs this analysis both at the level of the individual SNP, and also by region after assigning SNPs to blocks of the genome using their recommended blocks based on LD (https://bitbucket.org/nygcresearch/ldetect-data) or splitting the genome into consistent sized blocks.

2.4.4 Conducting Joint-Analysis in the Thesis

These three methods were adopted in the analysis pipelines and genetic weighted joint-analyses conducted in Chapter 7 of the thesis. By considering a sample size weighted joint-analysis through METAL and GWAS-PW, it allowed for an approach which targeted separate genomics regions that had prior evidence of joint association and provided a conservative list of novel signals which were likely to contribute to a person's pigmentation and melanoma risk. By considering a genetic correlation weighted joint-analysis through MTAG, the entire genome could be considered during analysis, and thus it was possible to highlight a greater number of regions potentially associated with melanoma.

2.5 Measures of linkage disequilibrium

There are several measures available to assess the LD between genomic variants. The LD measure most commonly implemented is the r^2 statistics. This measure represents the statistical coefficient of determination, which is equivalent to the correlation between a pair of variants. It is defined as:

$$r^2 = \frac{D^2}{f(A)f(a)f(B)f(b)}$$

where A and a are the alleles at the first SNP, B and b are the alleles at the second SNP, $f(X)$ represents the frequency of the allele X, and D represents the coefficient of LD given by $D = f(AB) - f(A)f(B)$.

When considering which measure of LD to use throughout the thesis, the r^2 measure was adopted. As the r^2 scales by MAF, LD can be more readily compared between pairs of variants with different allele frequencies. Furthermore, as many of the variants detected throughout the thesis have low MAF, the r^2 statistic gives a robust LD score. Similarly, as the r^2 statistic is exclusively used for most of the software and programs implemented in the thesis, the r^2 statistic was adopted to retain a universal measure throughout.

2.6 Identifying Loci of Association in GWAS Summary Statistics

2.6.1 Peak sorting algorithm

As multiple signals can be present in associated locus identified through GWAS, selecting the single most significant SNP residing in that locus is not an optimal approach as this can overlook other legitimate signals that may be present. To produce a better understanding of the signals presented in an associated locus, a genomic-distance approach between associated SNPs may result in the identification of addition signals per locus, and thus, additional association information. One of the simplest methods for representing multiple associated signals in genomic loci for post-GWAS analysis, is to selected the most significant (p-value) SNP for any given genomic window. This was implemented by selecting the most significant SNP from GWAS summary statistics and creating a pre-determined genomic distance window around that SNP. All SNPs within that window were removed from the algorithm. The next most significant SNP was then selected, and again, SNPs that were genomically close were removed. This repeated until all SNPs above a pre-determined p-value cut-off were selected. This was conducted in Chapter 4, for the generation of polygenic risk scores. The peak sorting algorithm code can be found in Appendix A.1. This method is computationally the least intensive, and provides a single SNP that represents a genomic distance-based locus.

2.6.2 Conditional analysis: GCTA

Conditional analysis has been used as a tool to identify multiple signals of association within the same locus. This is achieved by conditioning on the primary associated SNP to determine whether any SNPs at the same locus are also significantly associated. The approach can utilise GWAS summary results. Generally these methods follow a stepwise strategy beginning with the most associated SNP (determined by p-value) and adding additional SNPs to the model if they significantly improve the model. To ensure these methods are robust to the LD structure between SNPs, an r^2threshold is typically defined to remove any SNPs exceeding the tolerance level from being included in the association model. The methodology for mapping the LD structure between SNPs varies by approach.

One of the conditional analysis approaches used through the thesis is implemented through the GCTA software package (J. Yang et al. 2011). Uniquely to GCTA, it calculates the LD structure

between SNPs, as this is done by a subset of GWAS participants' genotype data. Once the LD structure is understood, the conditional analysis is conducted on the GWAS summary statistics using the following steps:

1. Select the most significant SNP in the GWAS summary statistics across the whole genome with p-values below a pre-defined cut-off
2. Calculate the p-values for all remaining SNPs when the SNP already selected is in the model. To avoid issues caused by colinearity, if the squared multiple correlation between a SNP to be tested and the selected SNP(s) is larger than a pre-determined cut-off value (r^2 value), the conditional p-value for that SNP will be set to one.
3. Select the SNP with the most significant conditional p-value that is lower than the defined p-value cut-off. If adding the new SNP causes additional colinearity issues between any already select SNPs and the others, the new SNP is dropped and the SNP that has the next most significant conditioned p-value is select and the process is repeated.
4. Conduct an association model with all selected SNPs included and drop the SNP with the largest p-value that is greater than the pre-defined cut-off p-value.
5. Repeat step (2), (3), and (4) until no SNPs can be added or removed from the association model.

This methodology was adopted in Chapter 5, as it provided the most robust way to identify signals of association, so they could be compared across pigmentation traits.

2.6.3 FUMA

An additional approach used in the thesis to identify multiple signals within a single genomic locus, is through the FUMA software (Watanabe et al. 2017). This software is similar in approach to GCTA, but is less robust in mapping the LD structure across SNPs, and in the identification of independent signals within a locus. To reduce computational burden, FUMA implements a pre-calculated LD structure derived from the 1000G of relevant reference populations (in this thesis, people of European-descent). The method is implemented by firstly identifying all independent (determined by a r^2 threshold) significant (determined by a p-value threshold) SNPs. To allow for annotation of these signals, any SNP that has an $r^2 > 0.6$ are retained for GWAS catalog (Buniello et al. 2019) searches. SNPs not present in the GWAS summary statistics, but present

in the 1000G reference panel are also included for annotation purposes. Based on the identified independent significant SNPs, independent lead SNPs are defined if they are independent from each other at a pre-defined level (default $r^2 = 0.1$). This creates LD blocks of SNPs that are in high LD with each other. These LD blocks of signals are then grouped together into genomic loci by considering the genomic distance between blocks. If any SNP in a block is within a pre-determined genomic distance threshold, the LD blocks are grouped into the same genomic locus. Each genomic locus can thus contain multiple independent significant SNPs and lead SNPs. These genomic loci are then annotated by performing ANNOVAR (a variant annotation tool used to obtain functional consequences of SNPs on gene functions) (K. Wang et al. 2010), and matching to supplementary databases (such as GWAS catalog (Buniello et al. 2019)). By having a pre-calculated LD structure, FUMA allows independent signals to be determined quickly with little computational burden. This method was adopted in Chapter 7, as it was quick and not computationally intense, but offered a more robust signal definition for the post-GWAS analyses in the chapter compared to the peak sorting algorithm.

2.6.4 Identifying Loci of Association in the Thesis

Throughout the thesis, the peak sorting algorithm, conditional analysis, and FUMA were used to define signals and loci of association post-GWAS. Each methodology was chosen due to their characteristics and utilisation of the results. Firstly, in Chapter 4, a peak sorting algorithm was implemented. This was due to the computationally efficient nature of the method. As these results were used to investigate the viability of including red-hair in a hair colour ordinal scale, selecting one SNP per genomic window using the peak sorting algorithm provided the necessary evidence without the need for a computationally intensive approach.

When conducting joint-analyses between melanoma and pigmentation GWAS, the FUMA approach was adopted. As it is a non-computationally intensive approach that provides multiple signals per locus, and supplementary information, such as prior associations. It also allows for an efficient analysis pipeline to be produced. Furthermore, the fundamental difference between GCTA and FUMA is the use of a reference panel to estimate the LD structure between participants included in analysis; and as the two GWAS that were jointly-analysed used independent samples, it was optimal to use a reference panel to estimate the LD structure, as a combining of GWAS samples would have been necessary to get an accurate LD structure estimation

Finally, to identify independent signals in the pigmentation GWAS conducted in the thesis, conditional analysis was adopted. This approach was selected as it provides the most accurate estimation of independent associated signals across GWAS results (due to calculate the LD structure for GWAS participants), and the LD structure was able to be calculated for the GWAS sample because all participants' genotyped information was available and didn't need to be combined.

2.7 Polygenic Risk Scores

A polygenic risk score (PRS) is a combination of the estimated effect on the trait of interest across multiple genetic variants into a single score. This variable reflects each individual's estimated genetic risk or value (in the case of a continuous trait) for the given phenotype. Typically the PRS is constructed by summing the estimated effect at each genetic variant (Equation 2.7).

$$S = \sum_{i=1}^{n} \beta_i x_i \tag{2.7}$$

where S is the polygenic risk score value, n is the number of genetic variants used to calculate the score, β is the effect size of the score, and x is the allele status of the genetic variant (0,1, or 2).

Whilst the equation for determining scores has been widely adopted as a standard, the methodology for determining which genetic variants are included in the score varies. One of the commonest of these is a pruning and thresholding approach. The difficulty of the approach is in the selection of which genetic variants to include in the PRS. Typically there are two metrics that are considered when determining whether genetic variants should be included, the LD measure (r^2 value), and the p-value inclusion threshold. The effect of each genetic variant in a PRS is based on its estimated effect on the trait in a univariate analysis. Effect estimates from genetic variants in LD will not be independent and may all represent the exact same effect (either from one of these genotyped variants or from another unobserved variant). Thus including in the PRS multiple variants in LD may lead to the same effect being included in the score multiple times, rendering the PRS suboptimal. Thus by not accounting for the non-random association strong LD presents between variants the predictive accuracy of a PRS may suffer as a result. A typical

inclusion criterion to account for LD between variants is to not include any two variants with $r^2 > 0.05$. The second inclusion criteria is the p-value threshold for variants. For GWAS 5×10^{-8} as widely been adopted as a suitable p-value threshold that balances type II and multiple-testing type I errors. Although PRS are similar, one distinct difference is present: it is possible to optimise the balance between detectability and type I errors by maximising the predictive ability of the score. This therefore allows the multiple-testing of different sets of variants of various p-value thresholds. If numerous false-positives are included into the score due to a less stringent p-value threshold additional noise can be included which can reduce predictability. However, the option to include less significant variants can also benefit the performance of highly polygenic traits, where multiple low-penetrance variants may represent a substantial proportion of overall risk, and would be missed using a more conservative p-value threshold (such as 5×10^{-8}) (Ware et al. 2017).

Another commonly adopted approach to PRS is to use a Bayesian framework. This approach attempts to model the genetic architecture of a given phenotypic trait, and thus the distribution of effect sizes for the given trait. The most widely used methods uses LD prediction to set a variant's PRS weighting to equal the average of its posterior distribution after LD has been accounted for. The last commonly adopt approach is penalised regression methods, such as LASSO or ridge regression. This is achieved by placing informative prior probabilities on how many genetic variants are expected to affect a trait, and the distribution of their effect sizes.

2.7.1 Polygenic Risk Scores in the Thesis

Throughout the thesis (Chapter 4 and 6), PRS have been adopted as the standard methodology to determine the overall effect of many genetic variants have on a trait, using a pruning and thresholding methodology for SNP inclusion. For Chapter 4, the most significant SNP for each locus was selected for use in the PRS generation. In Chapter 6 variants are chosen for the PRS with the aim of maximising the variation explained. For each PRS, an extensive conditional analysis methdology was used to minimise LD effects by the use of a threshold $r^2 > 0.05$. Multiple significance inclusion levels were also tested to maximise PRS performance.

Studies have suggested that pruning and thresholding approaches underperform in comparison

to Bayesian and penalised regression methods. However, this approach was considered for PRS generation due to it being less computationally intensive, especially as multiple pigmentation traits PRS are produced throughout the thesis. Additionally, pruning and thresholding allows for the 'manual' inclusion of variants that may not have been identified through post-GWAS analysis such as *MC1R* variants, and for the combing of variants identified across GWAS of different traits, such as with pigmentation and melanoma, to produce a hybrid PRS. Furthermore, pruning and thresholding is naturally complimentary to post GWAS results as variants weightings are the GWAS' estimated effect size for the given trait, which allows for an easier comparison of the PRS across traits.

2.8 Penetrance and Epistasis

2.8.1 Penetrance

Penetrance describes the likelihood of an associated phenotype given a particular genotype. Variants are commonly classified into high penetrance variants, where the presence of a particular allele is highly likely to result in the occurrence of a phenotype, and low penetrance variants, where the presence of an allele has a weak effect on determining a phenotypic trait. Although commonly used to category associated variants, the boundary between high and low penetrance is somewhat arbitrary, with no real definition to classify between high and low penetrance groups.

As well as being hard to categorise, penetrance can be difficult to estimate as numerous external factors can have a bearing on determining a phenotype, such as environmental factors, genetic epistasis, and age-related cumulative risk and ascertainment of samples is rarely clear cut. Historically, family-based methods were adopted to estimate penetrance, where a study would consider everyone with an observed genotype and determine how many have the phenotypic trait, or how many have the phenotypic trait by a pre-determined age. Due to the nature of family studies, these penetrance estimates suffered ascertainment bias (Minikel et al. 2014).

More recently, as genotyping has become cheaper and more widely available, population-based methods have become possible. These population-based studies overcome the biases present in family-based penetrance methods, and offer a good estimation of penetrance as long as good

estimates of allele frequency are used. A common way to calculate penetrance across a population group is to consider the following equation:

$$P = \frac{af_c}{af_{(c+con)}} \tag{2.8}$$

where P is the estimated penetrance, af_c is the summation of cases with a specific allele paring, and $af_{(c+con)}$ is the summation of cases and controls with the same specific allele paring.

In Chapter 8 of the thesis, this penetrance model is used to estimate the penetrance of different combinations of *MC1R* alleles on four pigmentation phenotypes for UK Biobank participants.

2.8.2 Epistasis

Epistasis occurs when the effect of a genetic variant on the expression is modulated by the genotype at another variant. Sometimes referred to as gene-gene interaction, epistasis is fundamentally important in understanding the structure and function of genetic pathways, as it allows for the effect of a broader genomic system rather than independent genetic variants on phenotypes. There a numerous ways to detect and estimate, the most common being regression-based analysis. The regression-based approach compares a model with the two genetic variants fitted as independent variables (the main effects) with another model that includes an interaction term between the two variables.. The two risk models are compared to assess whether the interaction term significantly improves the modelling of the phenotypic trait, usually with a χ^2 test.

Further to regression-based approaches, recent developments in machine learning have been implemented to detect non-additive epistasis effects between genes. These non-additive effects may be missed from traditional GLM regression-based approaches. One of these approaches is multi-factor dimensionality reduction (Motsinger and Ritchie 2006), which has been designed specifically for detecting non-additive gene-gene interactions in the absence of statistically detectable independent effects. It creates new variables by utilising genotypes from multiple SNPs, and combining these variables into a multi-locus genotype combination that are a function of contributing variables. A multi-locus genotype combination is considered high risk if the ratio of cases to controls in that group exceeds the overall ratio of cases to controls in the dataset, which would be considered to be caused by a gene-gene interaction

When considering the regression-based models through the thesis, the epistasis models were further extensions of the regression-based models defined for GWAS (see 2.3.1). If the phenotypic trait modelled was binary in nature, logistic-based regression models were selected for analysis, one with an interaction term (Equation 2.9) and one without (Equation 2.10).

$$logit(Y) = \beta_0 + \beta_a X_a + \beta_b X_2 b + \beta ab X_{ab} + \alpha_1 Y_1 + ... + \alpha_n Y_n \tag{2.9}$$

$$logit(Y) = \beta_0 + \beta_a X_a + \beta_b X_2 b + \alpha_1 Y_1 + ... + \alpha_n Y_n \tag{2.10}$$

where Y is the dependent phenotypic variables, β_0 is the model intercept, $\beta_a X_a$ is the regression coefficient for signal a, $\beta_b X_2 b$ is the regression coefficient for signal b,$\beta ab X_{ab}$ is the interaction terms between signal a and b, $\alpha_1 Y_1$ to $\alpha_n Y_n$ are model covariates, and ϵ is an error term. Covariates included in the model were the first ten genetic principal components to account for population structure. Analysis was conducted on the dedicated Cassi software (https://www.staff.ncl.ac.uk/richard.howey/cassi/index.html) to allow for efficient computation and inclusion of covariates, as other dedicated software (PLINK) restrict models.

If a phenotypic variable was recorded on an ordinal scales, linear-based regression models were selected for analysis, one with an interaction term (Equation 2.11) and one without (Equation 2.12).

$$Y = \beta_0 + \beta_a X_a + \beta_b X_2 b + \beta ab X_{ab} + \alpha_1 Y_1 + ... + \alpha_n Y_n \tag{2.11}$$

$$Y = \beta_0 + \beta_a X_a + \beta_b X_2 b + \alpha_1 Y_1 + ... + \alpha_n Y_n \tag{2.12}$$

where Y is the dependent phenotypic variables, β_0 is the model intercept, $\beta_a X_a$ is the regression coefficient for signal a, $\beta_b X_2 b$ is the regression coefficient for signal b,$\beta ab X_{ab}$ is the interaction terms between signal a and b, $\alpha_1 Y_1$ to $\alpha_n Y_n$ are model covariates, and ϵ is an error term. Covariates included in the model were the first ten genetic principal components to account for population structure. Analysis was conducted on the dedicated Cassi software (https://www.staff.ncl.ac.uk/richard.howey/cassi/index.html) to allow for efficient computation and inclusion of covariates, as other dedicated software (such as PLINK) restrict models due to calculation burden.

2.8.3 Penetrance and Epistasis in the Thesis

Penetrance and epistasis effects are investigated in Chapter 8. For penetrance, a population-based estimation was adopted to model the effects different combinations of *MC1R* variants have on pigmentation measures for UK Biobank participants. This was largely due to the UK Biobank data being population-based in nature, and due to population-based approaches being able to overcome many of the biases (particularly ascertainment bias) that are present in the older family-based methods.

For epistasis, a GLM-regression model was adopted to detect epistatic effects between *MC1R* variants and other significantly associated genomic regions for four UK Biobank pigmentation measures. If a pigmentation trait had a binary outcome a logistic regression-based model was adopted; if a pigmentation variable was represented by an ordinal scale a linear-regression based model was used. The primary reason to select a regression-based approach was due to limited resources available for implementing other methods. Although multi-factor dimensional approaches have some ability to detect non-linear epistatic effects in multiple loci, they can be hard to interpret (Dai et al. 2012), can include higher levels of type I error (Hsieh et al. 2011), are resource-intensive and require a large amounts of storage, memory, and computational power to successfully implement. This is particularly true if a large locus window is defined. Due to the nature of resources to conduct these analyses, they were impractical to conduct for this thesis.

Chapter 3

Reliability of cancer reporting and factor analysis of pigmentation in the UK Biobank

3.1 Introduction

Distinctions can be drawn between hair and skin melanin composition, where hair-based pigmentation is largely influenced by the presence of pheomelanin, and skin-based pigmentation largely determined by the ratio and quantity of DHICA-eumelanin and DHI-eumelanin (Cassady and Sturm 1994; Bellono and Oancea 2014). These distinctions in melanin composition show subtle differences in pigmentation across tissue types, and highlight that these pigmentation variables could be driven by closely related, but distinctly different biological processes. If this is the case, then assuming that these different types of pigmentation are due to the same 'overall' pigmentation processes may introduce bias when further investigating pigmentation, such as with GWAS.

When considering melanoma GWAS, current population-based studies combine the results from multiple melanoma datasets to maximise study detection power; where consistency across these datasets can be lacking with many biases being introduced throughout recruitment and interviewing. In particular, the criteria for recording melanoma status can vary considerably, with some studies (UK Biobank) using trained interviewer and external registry confirmation, and

others using online un-moderated forms (23andMe). The extent of incorrect melanoma status is largely unknown across these datasets with many accepting self-reported melanoma status without further validation.

3.1.1 Originality of research

Many studies aiming to compare the pigmentation of external tissues types have considered genetic factors. These studies utilise GWAS to identify the genomic regions that are associated with pigmentation. The number of loci known to be associated with each pigmentary trait varies, most likely due to differences in power between the various studies. The most recent is a large scale GWAS of hair colour that utilised one third of UK Biobank (Morgan et al. 2018; Hysi et al. 2018) which identified 200 loci. Other pigmentary traits such as skin colour and tanning ability have not been as successful with only six loci of association found for skin colour and 20 for tanning ability (F. Liu et al. 2015; Visconti et al. 2018). This lack of power has limited understanding of how these distinct pigmentation characteristics relate to one another: only the six loci mentioned previously have been consistent across tested pigmentary traits. By considering the self-reported phenotypic values for each pigmentation trait in UK Biobank and assessing the underlying correlation between them, it should be possible to further our understanding of how these pigmentary traits relate to one another, especially as they are measured on the same (large) number of people.

Recall bias in relation to melanoma has been investigated in only a few studies and for a limited number of risk factors (Parr et al. 2008; Walter, Marrett, from, et al. 1990). Several studies have assessed the reproducibility of melanoma risk factors separately for cases and controls (Walter, Marrett, from, et al. 1990; Walter, Marrett, Shannon, et al. 1992; Weinstock et al. 1991; Cockburn et al. 2001). However, no studies have assessed the recall bias in melanoma diagnosis, and the only large studies conducted to assess discrepancies in patient-reported and doctor-reported traits have been for non-cancer traits, looking at patient intervals and date of first symptom presentation (Parr et al. 2008; Han et al. 2006; Weinstock et al. 1991). Therefore linking data between cancer self-report and the UK Cancer Registry in UK Biobank presents a novel way to further our understanding of the accuracy of cancer type reporting.

3.1.2 Chapter Aims & Objectives

Aims

I. Determine how different pigmentation types (skin-based or hair-based) relate to one another

II. Determine the reliability of melanoma self-reports in the UK Biobank

Objectives

1. Conduct a factor analysis on UK Biobank pigmentation variables to assess the underlying correlation between them
2. Create a lexicon of cancer self-reports to cancer pathological confirmations
3. Create and implement a rule-set to assess the agreement across melanoma self-reports and pathological confirmations
4. Compare agreement of melanoma self-reports and pathological confirmations with other cancer types

3.1.3 Overview to achieve aims & objectives

Here we consider the self-report pigmentation responses for five measures (hair colour, red hair colour, skin colour, tanning ability, and number of childhood skin burns) for 350,000 UK Biobank participants. By conducting an exploratory factor analysis on a subsetted UK Biobank dataset, it was possible to demonstrate that a single overall pigmentation latent variable is optimal to summarise most of the variation in pigmentation (red hair colour, hair colour, skin colour and tanning ability). Number of childhood sunburns didn't correlate well with other pigmentary variables, likely due to high levels of recall bias. Furthermore, by conducting a confirmatory factor analysis, it was possible to demonstrate that a person's 'overall' pigmentation explains much of the variation in each distinct pigmentation attributes

Here we also report on the analysis of the 350,000 UK Biobank participants' self-reported cancer status for melanoma, non-melanoma skin cancer, and breast cancer compared to the UK Cancer Registry. By creating a rule-set that assesses the level of agreement between cancer self-reports and UK Cancer Registry entries, it was possible to investigate the reliability of self-reported

cancer types and how likely a participant with a cancer diagnosis was to report it. It was shown that, of the cancers studied, breast cancer has the highest level of reliability in self-reports and had the highest successful reporting if a UK Cancer Registry entry was present. A distinct difference was observed between melanoma and non-melanoma skin cancers, where UK Biobank participants were more likely to report melanoma if they had a prior diagnosis compared to non-melanoma skin cancer.

3.2 Methods

All analysis was conducted in R (R Core Team 2017) using UK Biobank data (Sudlow et al. 2015).

3.2.1 QC of pigmentation data

The UK Biobank offers four pigmentation measures of it's 488,377 genotyped participants. For each of these pigmentation measures descriptive statistics were produced to determine the differences across sex, geographical location using recruitment centre and melanoma status. Additionally, concurrence across each UK Biobank participant's pigmentation measures were investigated by highlighting the number of discordant responses for each pigmentation measure. For hair colour, a discordant response was determined for participant's reporting 'red hair' with 'brown' or 'black' skin colour and 'red hair' with 'never burn only tan' tanning ability; for skin colour, reporting 'very fair' skin with 'black' hair colour and 'very fair' skin with 'never burn only tan' tanning ability; for tanning ability, reporting 'never tan, only burn' with 'black' hair and 'never tan, only burn' with 'brown' or 'black' skin. An overall estimation of how many participant's in the UK Biobank provide discordant pigmentation measure responses was calculated by summing the total number participant's with discordant responses across the four pigmentation measures.

3.2.2 Cancer matchings

UK Biobank participants self-report a wide range of diseases. There are 82 unique self-report codes related to cancer used within UK Biobank. In addition to these self-report codes, truncated ICD9 and ICD10 codes are gathered from the UK Cancer Registry to identify confirmed cancers for participants. These truncated codes retain only category information (the first three characters) and include limited information on aetiology, anatomical site, severity, and other vital details. In order to match self-reports to a UK Cancer Registry entry, a rule-set was produced that first linked ICD9 and ICD10 codes (as there are discrepancies between the category information), and then linked the UK Biobank self-report IDs to these.

To link IC9 and ICD10 codes, a two-step clustering approach was adopted. Firstly within both the ICD9 and ICD10 code lists, any cancer types sharing the same category information code

(the first three characters) were grouped together, as these were inseparable without the whole ICD9/ICD10 code for each cancer type. Once cancer types were grouped together within each ICD9 and ICD10 list, these groups were then clustered together based on overlap. If a cancer type grouped with a different cancer type in the ICD9 list, and also grouped with another different cancer type in the ICD10 list, all three cancer types would be considered as the same overall cancer group in the linked list. An example of this can be found in Table 3.1. The overall linked groups were then matched with the UK Biobank self-report IDs using the same method to give the final groupings. Each group of cancer types was assigned a code and each UK Biobank participant's self-reports and confirmed cancers were converted into these linked codes, enabling them to be matched.

ICD9	ICD9 grouping	ICD10	ICD10 Grouping	Linked group	Cancer Type
153 1540		C19			Large bowl cancer/colorectal
153 1540		C18			Large bowl cancer/colorectal
153		D126 C18			Large bowl cancer/colorectal
154		C19 C20			Large bowl cancer/colorectal
154		C21			Large bowl cancer/colorectal

Table 3.1: Example matching process for ICD9 and ICD10 matchings for UK Biobank pathological confirmations using colour matchings.

UK Biobank Cancer matchings

Post-grouping, positive predictive value (PPV) and sensitivity were calculated for melanoma skin cancer, non-melanoma skin cancer and breast cancer using the UK Cancer Registry as the gold standard. Participants who self-reported 'skin cancer' were excluded from analysis as there was no specification of the type of skin cancer. False-positive reports occurred when a cancer was self-reported by the participants in the absence of a UK Cancer Registry record. Conversely, false-negative reports occurred when a cancer was not self-reported in the presence of a UK Cancer Registry entry.

PPV was further investigated for other cancers by matching each participant's self-reports to their entire UK Cancer Registry log, and sensitivity was further investigated by matching participant's entire UK Cancer Registry log with their self-reports. By matching self-reports and cancer registry logs together, it is possible to identify: (1) mix-ups that occur frequently; (2) the level of accuracy in self-reports; (3) the proportion of self-reports with no UK Cancer Registry matching; (4) the proportion of self-reports apparently matching to a different cancer types.

Within UK Biobank, participants had up to 7 individual cancer self-reports and up to 14 cancer confirmations from the UK Cancer Registry. To overcome the problem of having multiple data records for a single participant, a sorting algorithm was created that matched UK Cancer Registry entries to the most appropriate self-report and that matched self-reports to the most appropriate UK Cancer Registry entry for each of the 502,000 participants in UK Biobank. This was achieved by using a 4 step matching method for each self-report (1-7) or cancer registry log (1-14) to produce two matrices. Matrix 1 (M1): participant number $\times$ self-report number 1-7 (or cancer registry number 1-14), with elements consisting of the matched confirmation cancer code (or self-report). Matrix 2 (M2): participant number $\times$ self-report number 1-7 (or cancer registry number 1-14), with elements consisting of a, b, c, or d where: a is a match on cancer type and the self-report date and confirmation date are a year or less than a year apart; b is a match on cancer type with self-report date and confirmation date being greater than a year apart; c is a participant did not have a record of the self-reported (or a cancer registry entry) cancer, but did have a record of a different type of cancer within a year of the self-report (or cancer registry) date; and d is no match was found for the self-report (or cancer registry entry).

A generalised code for the sorting algorithm follows:

I. IF self-report (or UK Cancer Registry entry) matches cancer code with any confirmation (1-14) (or self-report (1-7)) and date difference is within one year.
   Output: M1 = matched cancer code; M2 = a

II. IF self-report (or UK Cancer Registry entry) matches cancer code with any confirmation (1-14) (or self-report (1-7)) and date difference is over one year.
   Output: M1 = matched cancer code; M2 = b

III. IF self-report (or UK Cancer Registry entry) does not have a cancer code match but has a UK Cancer Registry entries (or self-report) for a different type of cancer within a date difference of one year.
   Output: M1 = cancer code; M2 = c

IV. No match on cancer code or date.
   Output: M1 = -10; M2 = d

3.2.3 Factor analyses

FA can identify underlying factors that are not easily measured directly. Two distinct FA approaches were implemented in this chapter: an EFA and CFA. By conducting these factor analyses it allowed for the characterising of latent variables that underlie those that can be measured/observed.

Preparing the data

The EFA and CFA were conducted on equal-sized independent subsets of the UK Biobank to allow for model fitting in the EFA and model testing in the CFA. Prior to analysis, data preparation was conducted. Each UK Biobank participant's pigmentation measures were extracted from their interview responses (ease of tanning, skin colour, number of childhood sunburns and hair colour). Hair colour was split into a pseudo-eumelanin (labelled hair colour) and pseudo-pheomelanin (labelled red hair) scale (1.1.1). Red hair, hair colour, skin colour, and tanning ability were considered on an ordinal scale: For hair colour: 'red' = 1 'blond' = 2, 'light brown' = 3, 'dark brown' = 4, 'black' = 5; for red hair: 'non-red hair' = 1, 'red hair' = 2; for skin colour: 'very fair' = 1, 'fair' = 2, 'light olive' = 3; 'dark olive' = 4, 'brown' = 5, 'black' = 6; For tanning ability: 'get very tanned' = 1, 'get moderately tanned' = 2, 'get mildly or occasionally tanned' = 3, 'never tan only burn' = 4. Number of childhood sunburn incidents was converted into an ordinal scale ('None to two incidents' = 1, 'one to three incidents' = 2, 'over three incidents' = 3) since over a third of entries were between 0-2, with the rest ranging between three and 999. All five of these pigmentation responses were combined in a single dataset and dichotomised into two equal-sized independent samples (one for each FA approach): an EFA dataset and a CFA dataset.

Scree plot & EFA

To determine the approximate number of latent variables in a latent model of UK Biobank pigmentation measures (tanning ability, skin colour, hair colour, red-hair colour, and number of childhood sunburns), a scree plot approach was adopted to test the amount of variation explained for up to five latent variables included in the model. Analysis was conducted on the EFA dataset. Polychoric correlations (due to the measures being ordinal) were calculated between the five pigmentation measures and combined into a single polychoric correlation matrix. Five

eigenvalues were then calculated from this correlation matrix. The first eigenvalue represents the variation explained by the first latent variable of the model, and each following eigenvalue corresponds to the additional variation explained by each subsequent latent variable. The five eigenvalues were labelled one to five and plotted in a scree plot. The approximate number of latent variables for the latent model of pigmentation was determined by declaring at which point the rate of increase in variation explained with increasing latent variables falls (known as the 'elbow'). The two latent models with the number of latent variants situated either side of the 'elbow' were selected for EFA.

For each of the EFA conducted post-scree plot, a maximum likelihood approach with polychoric correlations (as the pigmentation variables are ordinal) was used. To maximise model fit, an Oblimin rotation was used as this is an oblique rotation that doesn't assume independence between factors (as red hair and hair colour are not independent). A factor loading threshold was declared at 0.3, where variables below this threshold were not considered to be effected by the latent variable. If no difference in pigmentary variables effected by the latent variables between EFA models was detected for the relationship of pigmentation and latent variables, the lowest number of latent variables was selected for the optimal number.

Once the optimal number of latent variables had been identified, sensitivity analysis was conducted on the optimal model, by assessing the influence individual variables had on the analysis by excluding each variable in turn and re-conducting the EFA. A large influence was declared if a difference of 0.3 in the factor loadings of retained variables was detected.

CFA

The optimal set of latent variables identified in the EFA was selected for CFA and conducted on the CFA dataset. All five pigmentation variables were included in the analysis, with red hair and hair colour having an interaction term due to their high correlation. A robust maximum likelihood approach was adopted for model fitting as this allowed polychoric correlations between pigmentation variables to be used. Model fit was assessed by χ^2 test statistics, root mean square error of approximation (RMSEA), and Tucker-Lewis index. Model fit was considered good if the χ^2 p-value > 0.05; the RMSEA < 0.05; and the Tucker-Lewis index was > 0.95. If model fit

was good, standardised factor loadings for the effect of the latent variables on the pigmentation variables were tabulated. Sensitivity analysis was conducted by individually omitting variables in turn and re-running the CFA to determine their influence on results by detecting a difference of 0.3 in the factor loadings of retained variables.

3.3 Results

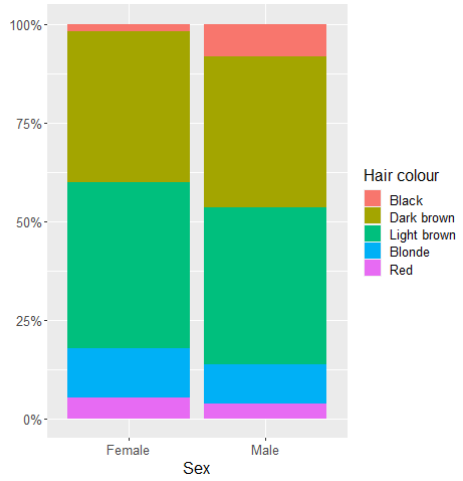
3.3.1 QC of pigmentation data

Assessing the distribution of categories for the four UK Biobank pigmentation variables across sex, geographical location, and melanoma status highlighted biases present across these assessed groups. For hair colour, 465,505 UK Biobank participants provided a pre-greying hair colour at interview; with the median response being 'light brown'. When distributing hair colour responses across sex a greater percentage of male (7%) gave 'black hair' as a response compared to females (1%). Similarly, females provided a higher percentage of 'red hair' (4%) and 'blond hair' (13%) responses compared to males (3% and 10%, respectively) (Figure 3.1). This could be due reporting bias where females may prefer to reporter lighter hair colours compared to men who may prefer to have a darker hair colour. Comparing across geographical location found difference in distribution of hair colour across and UK Biobank participants (Figure 3.1). Unsurprisingly, the distribution of hair colour across melanoma case and controls found lighter hair colours (blond and red hair) to be a higher percentage in melanoma cases (27% combined) compared to controls (14% combined), as lighter pigmentation is a risk factor for melanoma (Figure 3.1).

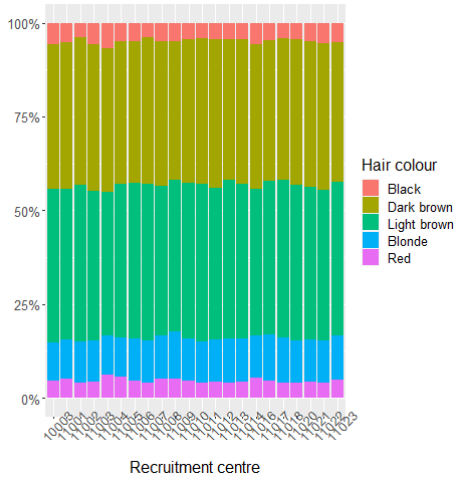
For skin colour, 466,443 UK Biobank participants provided a skin colour response at interview; with the median response being 'fair skin'. When distributing skin colour responses across sex little difference identified between males and females (Figure 3.2). Similarly, there was little variation in the distribution of skin colour responses across recruitment centres (Figure 3.2). When comparing skin colour responses between melanoma cases and controls, a higher percentage of 'very fair' (14%) and 'fair' (74%) responses were observed for cases compared to controls (8% and 71%, respectively) (Figure 3.2). This observation is unsurprising due to lighter skin being associated with higher melanoma risk.

For tanning ability, 459,601 UK Biobank participants provided a tanning ability response at interview; with the median response being 'get moderately tanned'. Distributing tanning ability responses by sex found a higher percentage of females (47%) responding 'never tan, only burn' or 'get mild, or occasionally tanned' compared to males (31%), who responded with 'get very tanned' or 'get moderately tanned' more frequently (69%) compared to females (53%) (Figure 3.3). Across recruitment centres, centre 1002 had a slightly different response distribution compared to other centres with a lower percentage of 'Never tan, only burn' responses: 5% compared to around 12% for other centres (Figure 3.3). As with other pigmentation variables, comparing the distribution of tanning ability response across melanoma cases and controls found 50% of melanoma cases responding 'never tan, only burn' or 'get mild, or occasionally tanned' compared to melanoma controls (38%), which again is expected due to poor tanning ability being associated with higher melanoma risk (Figure 3.3).

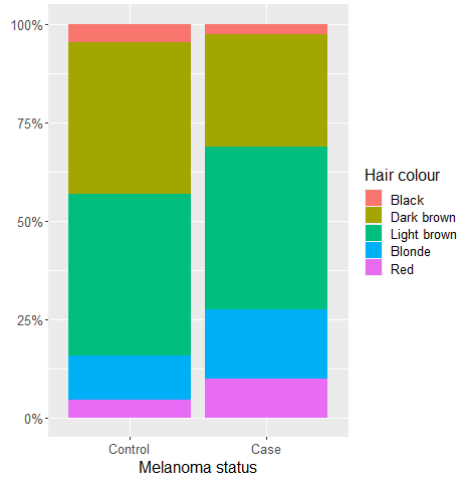
For number of childhood sunburns, 283,286 UK Biobank participants provided a number for childhood sunburns when interviewed; with the median response value of 0. When categorising number of childhood sunburns into no incidents of sunburn and at least one incident of sunburn, the distribution of responses by sex and location found little difference between males and females or geographical location with around 25% responding at least one incident across males and females and across each recruitment centre (Figure 3.4). A slight difference was observed between melanoma cases and controls with melanoma cases having a higher percentage (33%) of 'at least one incident' responses compared to melanoma cases (25%) (Figure 3.4).



(a)

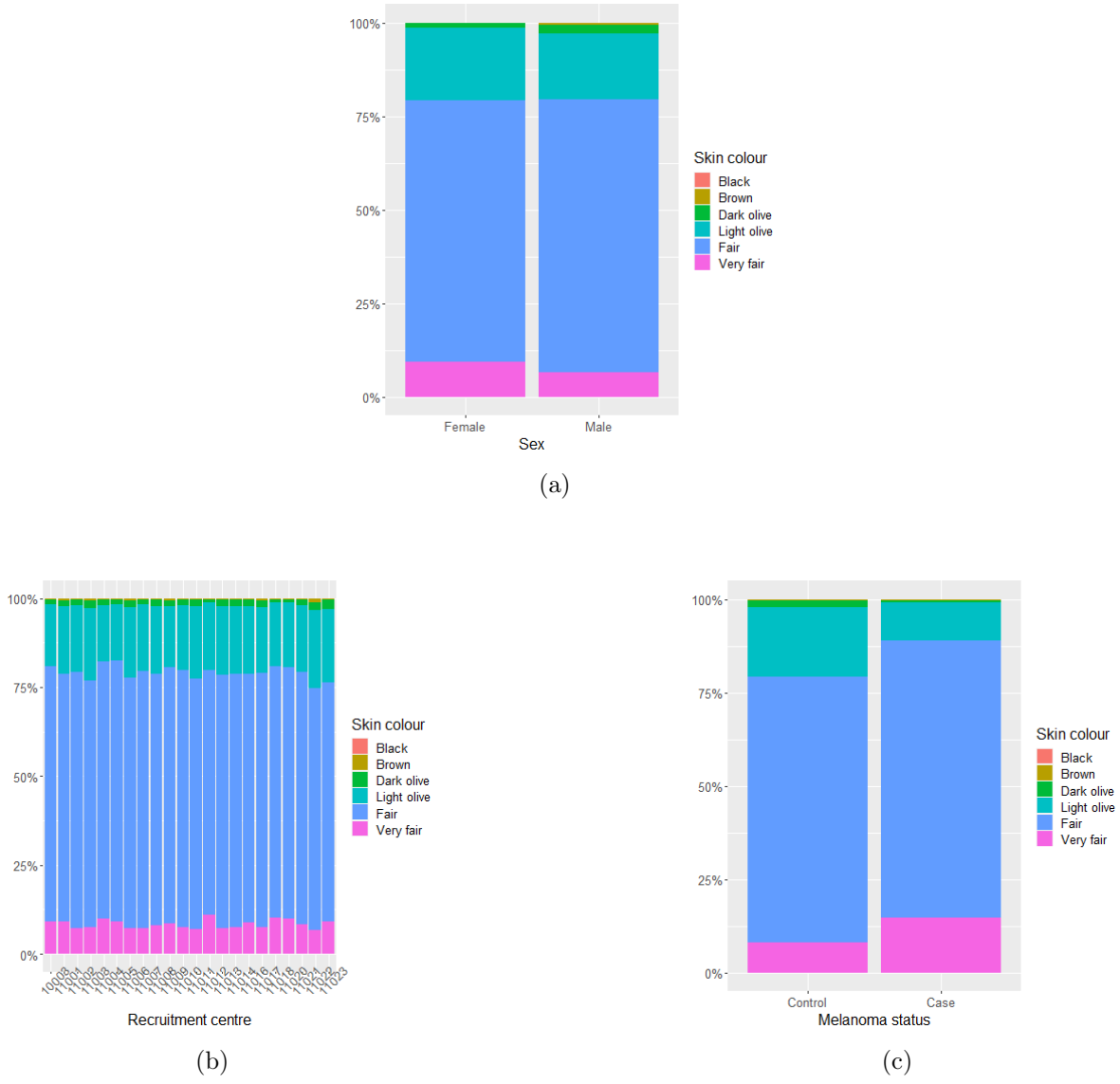


(b)



(c)

Figure 3.1: Stacked bar plots for the responses of hair colour in UK Biobank: (a) Hair colour and sex; (b) Hair colour and location; (c) Hair colour and melanoma status.

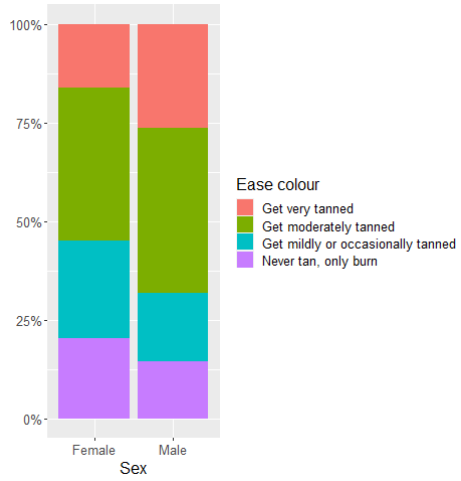


(a)

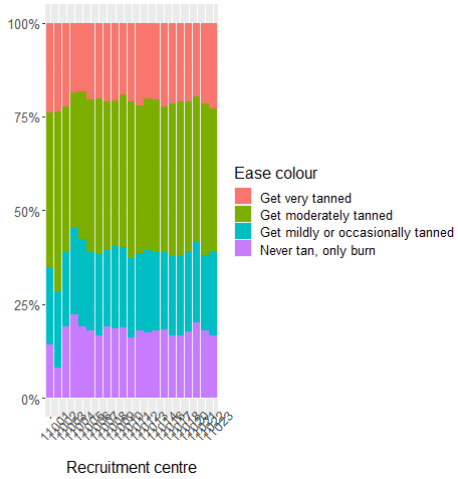
(b)

(c)

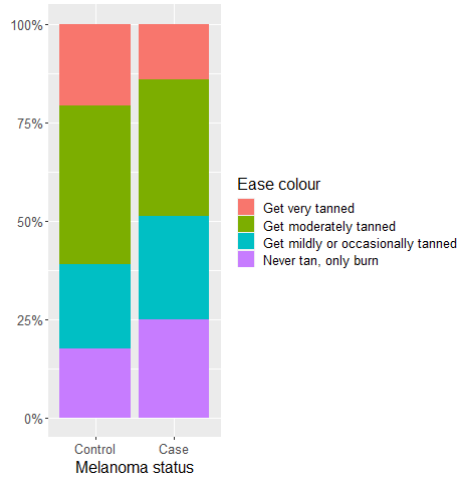
Figure 3.2: Stacked bar plots for the responses of skin colour in UK Biobank: (a) Skin colour and sex; (b) Skin colour and location; (c) Skin colour and melanoma status.



(a)

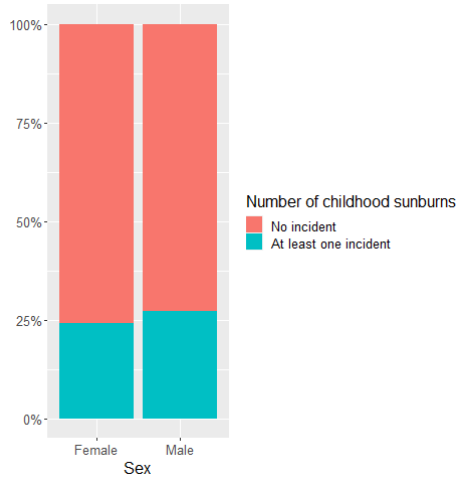


(b)

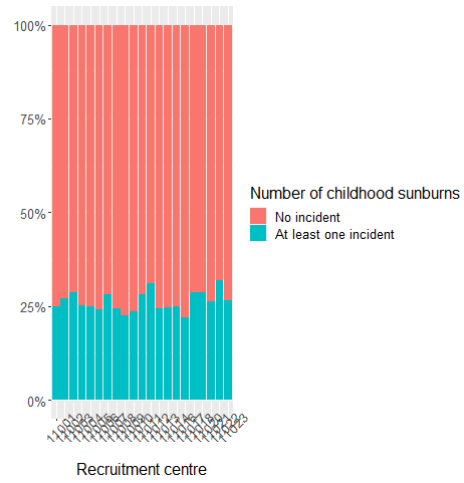


(c)

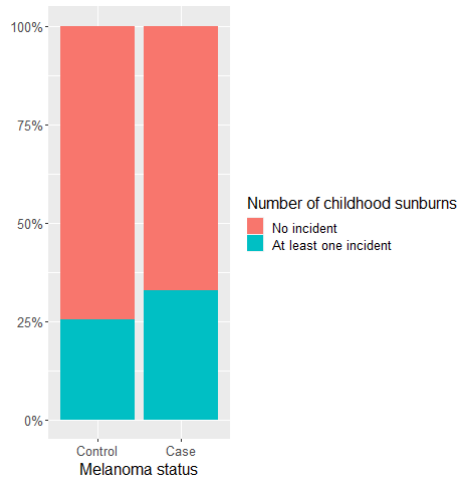
Figure 3.3: Stacked bar plots for the responses of tanning ability in UK Biobank: (a) Tanning ability and sex; (b) Tanning ability and location; (c) Tanning ability and melanoma status.



(a)



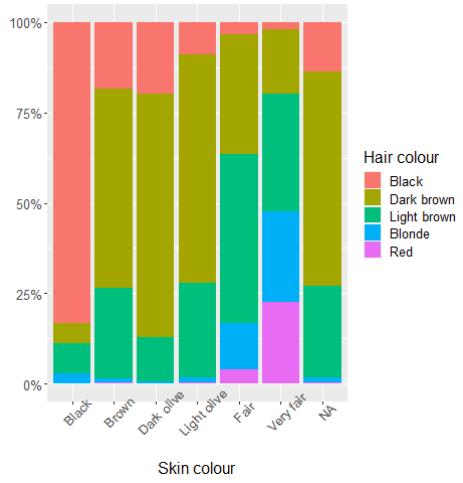
(b)



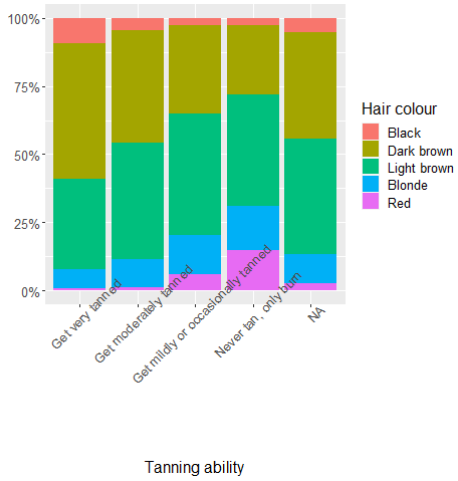
(c)

Figure 3.4: Stacked bar plots for the responses of number of childhood sunburns in UK Biobank: (a) Number of childhood sunburns and sex; (b) Number of childhood sunburns and location; (c) Number of childhood sunburns and melanoma status.

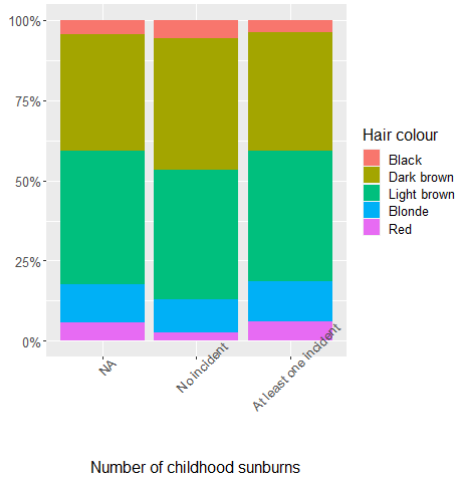
To determine the extent of discordant responses between the four UK Biobank pigmentation measures for each participant, the following matched responses were grouped together: if a participant reported their hair colour as 'red' but also their skin colour as 'black' or 'brown' (one participant), or their tanning ability as 'get very tanned' (504 participants); if a participant reported their skin colour as 'very fair' but also their hair colour as 'black' (285 participants), or their tanning ability as 'get very tanned' (1,921 participants); or if a participant reported their tanning ability to be 'never tan, only burn' but their hair colour as 'black' (67 participants) or their skin colour as 'brown' or 'black (2,060 participants). Overall there are 4,838 (1.02%) participants in the UK Biobank that provided discordant responses across their pigmentation measures. When considering the distribution of each UK Biobank pigmentation measures across the others, good correlation between the measures are observed (Table 3.9), with largely concordant responses provided by participants (Figure 3.5, 3.6, 3.7, 3.8).



(a)

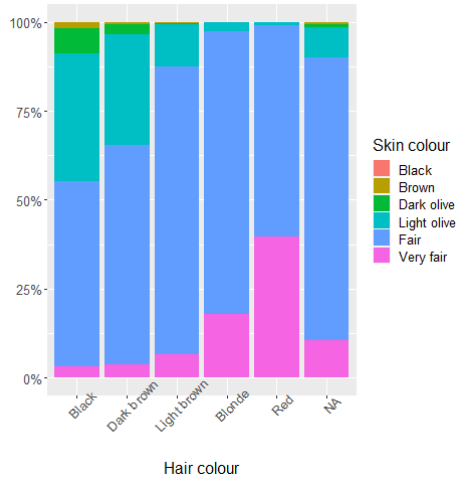


(b)

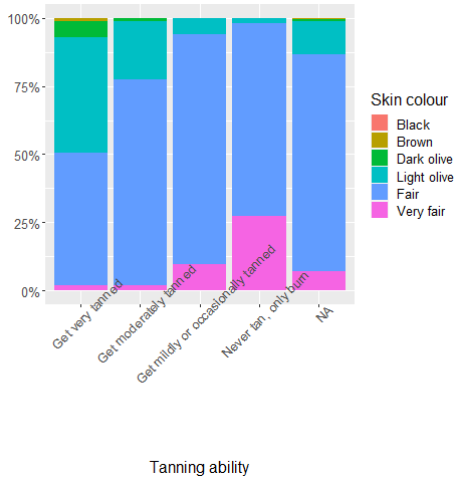


(c)

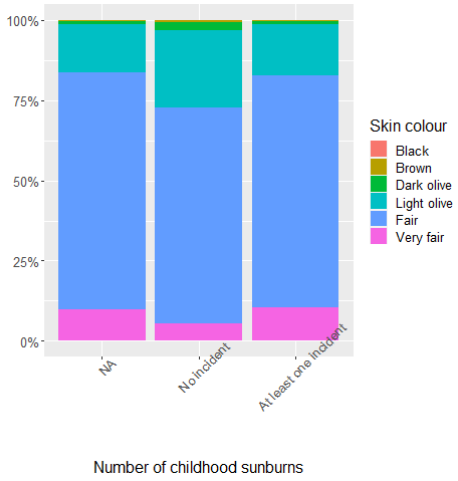
Figure 3.5: Stacked bar plots for the responses of hair colour in UK Biobank: (a) Hair colour and skin colour; (b) Hair colour and tanning ability; (c) Hair colour and number of childhood sunburns.



(a)

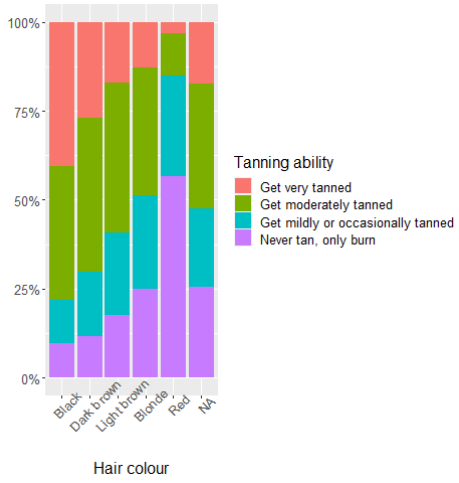


(b)

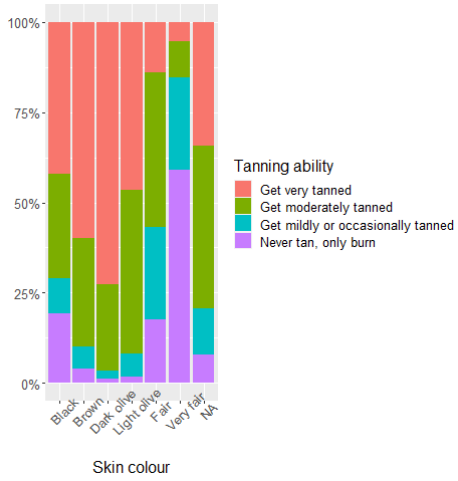


(c)

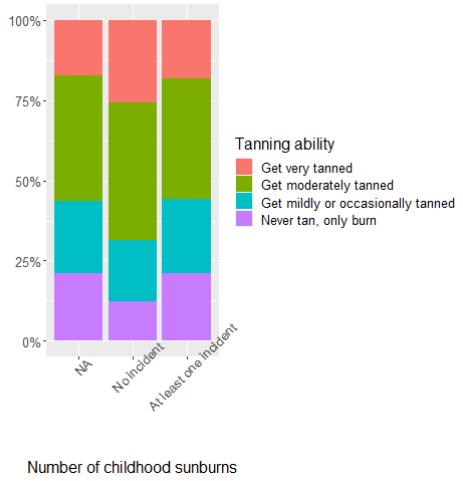
Figure 3.6: Stacked bar plots for the responses of skin colour in UK Biobank: (a) Skin colour and hair colour; (b) Skin colour and tanning ability; (c) Skin colour and number of childhood sunburns.



(a)

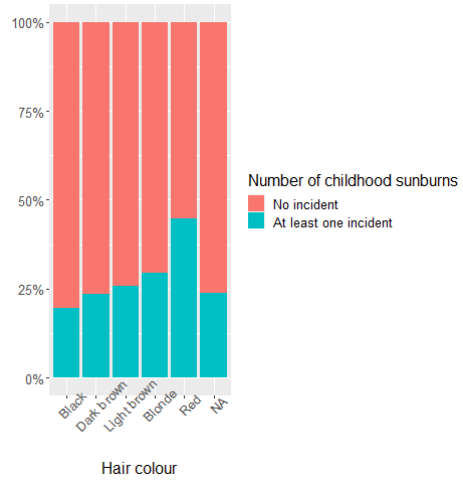


(b)

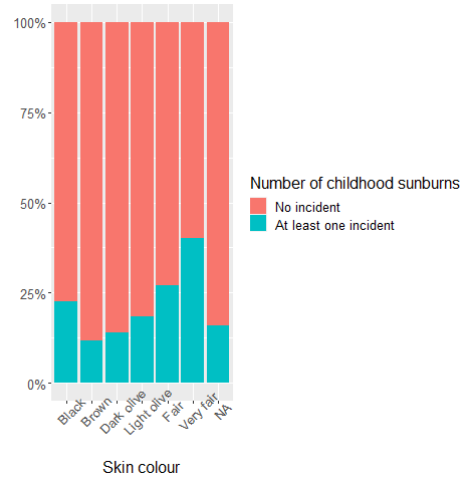


(c)

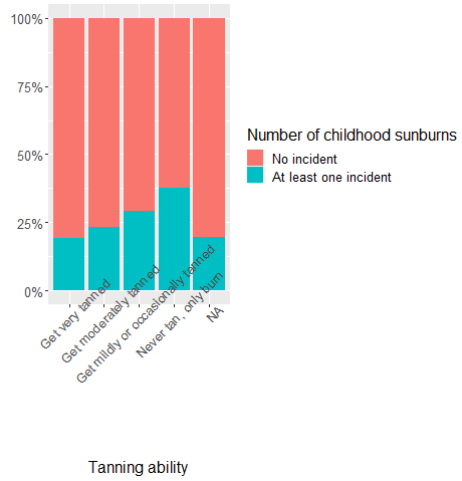
Figure 3.7: Stacked bar plots for the responses of tanning ability in UK Biobank: (a) Tanning ability and hair colour; (b) Tanning ability and skin colour; (c) Tanning ability and number of childhood sunburns.



(a)



(b)



(c)

Figure 3.8: Stacked bar plots for the responses of number of childhood sunburns in UK Biobank: (a) Number of childhood sunburns and hair colour; (b) Number of childhood sunburns and skin colour; (c) Number of childhood sunburns and tanning ability.

Childhood Sunburn, melanoma, non-melanoma skin cancer and confounding

When investigating the association of number of childhood sunburns with melanoma and non-melanoma skin cancer, the variable was significant ($p < 0.001$) in both the melanoma risk model and non-melanoma skin cancer risk model. For predicting melanoma risk, the effect estimate was 0.36 between no incidents and at least one sunburn incident before 16 years of age, where having at least one incident of sunburn increased melanoma risk (Table 3.2). Similarly, for the prediction of non-melanoma skin cancer, the effect estimate was 0.26 where having at least one incident of sunburn before 16 years of age increased non-melanoma skin cancer risk (Table 3.2). Evaluating confounding between the number of childhood sunburns with age, deprivation (Townsend score), geographical location, and education level when modelling melanoma and non-melanoma skin cancer risk, identified confounding between number of childhood sunburns and age for both melanoma and non-melanoma skin cancer risk. This confounding resulted in an overestimation (positive confounding) of the effect for number of childhood sunburn in melanoma risk with an effect estimate increase of 0.08 (22% increase), and non-melanoma skin cancer risk with an effect estimation increase of 0.17 (65% increase) (Table 3.2 and 3.3). This identified confounding is likely due to the constant trend of increased ease and reduction of the associated costs for foreign travel, meaning older participants in the UK Biobank may have been less likely to travel to sunnier climates compared to younger participants, and thus having less UV exposure before the age of 16.

Model:	**Variable**	**Effect estimate**	**Std. error**	**p-value**
CS, age, Townsend score, location, education level	CS	0.43	0.056	< 0.001
	age	-0.04	0.004	< 0.001
	Townsend score	-0.07	0.01	< 0.001
	location	-0.02	0.005	< 0.001
	education level	0.005	0.007	0.45
CS	CS	0.36	0.056	< 0.001
CS, age,	CS	0.44	0.056	< 0.001
	age	-0.04	0.004	< 0.001
CS, Townsend score,	CS	0.36	0.056	< 0.001
	Townsend score	-0.07	0.01	< 0.001
CS, location,	CS	0.36	0.056	< 0.001
	location	-0.02	0.005	< 0.001
CS, education level,	CS	0.36	0.056	< 0.001
	education level	0.005	0.007	0.45

Table 3.2: Included variable effect sizes, SEs, and p-value for melanoma risk models when considering age, Townsend score, location, and education level as confounders for CS. CS = Number of Childhood sunburns

Model:	**Variable**	**Effect estimate**	**Std. error**	**p-value**
CS, age, Townsend score, location, education level	CS	0.43	0.02	< 0.001
	age	-0.09	0.004	< 0.001
	Townsend score	-0.04	0.004	< 0.001
	location	-0.008	0.002	< 0.001
	education level	0.003	0.037	0.189
CS	CS	0.26	0.023	< 0.001
CS, age,	CS	0.43	0.023	< 0.001
	age	-0.09	0.002	< 0.001
CS, Townsend score,	CS	0.25	0.023	< 0.001
	Townsend score	-0.05	0.004	< 0.001
CS, location,	CS	0.26	0.023	< 0.001
	location	-0.02	0.002	< 0.001
CS, education level,	CS	0.27	0.023	< 0.001
	education level	0.025	0.003	< 0.001

Table 3.3: Included variable effect sizes, SEs, and p-value for non-melanoma risk models when considering age, Townsend score, location, and education level as confounders for CS. CS = Number of Childhood sunburns

3.3.2 UK Biobank matching table

Grouping cancer types based on the combination of ICD9, ICD10 and UK Biobank cancer codes created 50 unique cancer groupings, an 'other' group and a melanoma *in situ* group. Of the 50 unique cancer groups, three were related to cutaneous cancer: Group 0: Skin cancer, Group 1: Melanoma, Group 3: Non-melanoma skin cancer (Table 3.4).

Self-report code	ICD9	ICD10	Group Number	Group Description
1003			0	Skin cancer
1059	172	C43	1	Malignant melanoma
1060	173	C44	2	Non-melanoma skin cancer
1061	173	C44	2	Non-melanoma skin cancer
1062	173	C44	2	Non-melanoma skin cancer
1073	173	C44	2	Non-melanoma skin cancer
1001	162	C34	3	Lung cancer/trachea cancer
1027	162	C34	3	Lung cancer/trachea cancer
1028	162	C34	3	Lung cancer/trachea cancer
1080	162	C33	3	Lung cancer/trachea cancer
1005	142	C08 C07	4	Salivary gland cancer/parotid gland cancer
1015	142	C07	4	Salivary gland cancer/parotid gland cancer
1016	142	C08	4	Salivary gland cancer/parotid gland cancer
1020	153 1540	C19	5	Large bowl cancer/colorectal cancer/anal cancer/colon cancer/appendix cancer
1021	154	C21	5	Large bowl cancer/colorectal cancer/anal cancer/colon cancer/appendix cancer
1022	153	D126 C18	5	Large bowl cancer/colorectal cancer/anal cancer/colon cancer/appendix cancer
1023	154	C19 C20	5	Large bowl cancer/colorectal cancer/anal cancer/colon cancer/appendix cancer
1086	153 1540	C18	5	Large bowl cancer/colorectal cancer/anal cancer/colon cancer/appendix cancer
1026	157	C25	6	Pancreatic cancer/malignant insulinoma
1088	157	C25	6	Pancreatic cancer/malignant insulinoma
1030	190	C69	7	Eye cancer/adnexal cancer/reinoblastoma
1075	190	C69	7	Eye cancer/adnexal cancer/reinoblastoma
1031	192	C72 C70	8	Meningeal cancer/malignant meningioma/ spinal cord or cranial nerve cancer
1033	192	C72 C70	8	Meningeal cancer/malignant meningioma/ spinal cord or cranial nerve cancer
1037	184	C51 C57	9	Female genital tract cancer/vaginal cancer/vulval cancer
1042	184	C52	9	Female genital tract cancer/vaginal cancer/vulval cancer
1043	184	C51	9	Female genital tract cancer/vaginal cancer/vulval cancer
1038	187	C60 C63	10	Male genital tract cancer/Penis cancer
1046	187	C60 C63	10	Male genital tract cancer/Penis cancer
1039	183	C56 C57	11	Ovarian cancer/fallopian tube cancer
1087	183	C57	11	Ovarian cancer/fallopian tube cancer
1047	200 201 202	C96 C91 C85 C83 C85 C81 C82 C84	12	Lymphoma/Leukaemia
1052	201	C96 C91 C85 C83 C85 C81 C82 C84	12	Lymphoma/Leukaemia
1053	202	C96 C91 C85 C83 C85 C81 C82 C84	12	Lymphoma/Leukaemia
1048	204 205 206 207 208	C91 C92 C93 C94 C95	12	Lymphoma/Leukaemia
1055	204	C91 C92 C93 C94 C95	12	Lymphoma/Leukaemia
1056	205	C91 C92 C93 C94 C95	12	Lymphoma/Leukaemia
1058	204 205 206 207 208	C91 C92 C93 C94 C95	12	Lymphoma/Leukaemia
1074	205	C91 C92 C93 C94 C95	12	Lymphoma/Leukaemia
1063	170	C40 C41	14	Primary bone cancer/Sarcoma/fibrosarcoma
1068	170 171	C40 C41 C49	14	Primary bone cancer/Sarcoma/fibrosarcoma
1064	163	C38 C45	15	Mesothelioma/thymus cancer/heart cancer
1081	164	C37	15	Mesothelioma/thymus cancer/heart cancer
1082	164	C38	15	Mesothelioma/thymus cancer/heart cancer
1066	194	C74 C75	16	Parathyroid cancer/adrenal cancer
1067	194	C74 C75	16	Parathyroid cancer/adrenal cancer
1078	146	C09	17	Tonsil cancer/oropharynx/oropharyngeal cancer
1079	146	C10	17	Tonsil cancer/oropharynx/oropharyngeal cancer
1071	199	C80	18	Metastatic unknown primary/unclassifiable
99999	199	C80	18	Metastatic unknown primary/unclassifiable
1002	174	C50	19	Breast cancer
1004	-	-	20	Cancer of lip/mouth/pharynx/oral/cavity
1006	161 149	C32 C14	21	Larynx/throat cancer
1007	160	C31 C30	22	Nasal cavity cancer
1008	-	-	23	Primary ear cancer
1009	-	-	24	Sinus cancer
1010	140	C00	25	Lip cancer
1011	141	C01 C02	26	Tongue cancer
1012	143	C03	27	Gum cancer
1017	150	C15	28	Oesophageal cancer
1018	151	C16	29	Stomach cancer
1019	152	C17	30	Small intestine/small bowl cancer
1024	155	C22	31	Liver/hepatocellular cancer
1025	156	C23 C24	32	Gallbladder/bile duct cancer
1029	-	-	33	Peripheral nerve/autonomic nerve cancer
1032	191	C71	34	Brain cancer/primary malignant brain tumour
1034	189	C64 C65	35	Kidney/renal cell cancer/other cancer of urinary tract
1036	189	C66 C68	35	Kidney/renal cell cancer/other cancer of urinary tract
1035	188	C67	36	Bladder cancer
1040	179 1820	C55 C54	38	Uterine/endometrial cancer
1041	180	C53	39	Cervical cancer
1044	185	C61	40	Prostate cancer
1045	186	C62	41	Testicular cancer
1050	203	C90	42	Multiple myeloma
1051	-	-	43	myofibroblast/myelodysplasia
1065	193	C73	44	Thyroid cancer
1070	196	-	45	Malignant lymph node (unspecified)
1072	2331	D069	46	Cin/precancer cells cervix
1076	-	-	47	Kaposi's sarcoma
1077	145	C06 C05	48	Mouth cancer
1084	165	C39	49	Respiratory/intrathoracic cancer
1085	198	C79	50	Bone metastases/bony secondaries
other	other	other	51	Other
•		D03	52	Melanoma in situ

Table 3.4: Group cancer types based on the combination of ICD9, ICD10, and UK Biobank cancer codes.

Melanoma had a high sensitivity (85.1%) and low PPV (49.5%) compared to non-melanoma skin cancer which had a low sensitivity (39.4%) and a higher PPV (61.5%). Breast cancer had the highest sensitivity (98.7%) and PPV (81.9%) compared to the other two cancers investigated, indicating it has the most reliable self-reports. The high sensitivity and low PPV for melanoma indicates that if a UK Biobank participant had a melanoma UK Cancer Registry entry, there is a high chance they will report it. Conversely, non-melanoma skin cancer had a low sensitivity and a higher PPV (Tables 3.5).

	Sensitivity (%)	Positive predictive value (%)
Melanoma	85.05	49.5
Breast cancer	98.75	81.86
Non-melanoma skin cancer	39.43	61.55

Table 3.5: Sensitivity and positive predictive values (PPV) of melanoma, breast cancer and non-melanoma skin cancer from the UK Biobank.

Only 9% of breast cancer self-reports were matched with a different type of cancer when compared to the UK Cancer Registry entries for UK Biobank participants. 3% of non-melanoma skin cancer self-reports were matched to a different type of cancer when compared to the UK Cancer Registry entries for UK Biobank participants. 28% of non-melanoma skin cancer self-reports had no cancer match, while this was true for only 7% of breast cancer self-reports. 32% of melanoma self-reports did not match with melanoma or any other type of cancer when comparing to the UK Cancer Registry entries for UK Biobank participants. Additionally, 11% of melanoma self-reports matched with a different type of cancer (including non-melanoma skin cancer), and 4% of melanoma self-reports matched with melanoma *in situ*. 10% of UK Cancer Registry confirmations of melanoma had no corresponding self-report (melanoma or other cancer). 14% of UK Cancer Registry confirmations of melanoma matched to a self-report of a different cancer (Tables 3.6, 3.7, Appendix B: Tables B.1, B.2, B.3, B.4, B.5, B.6, B.7).

	UK Biobank Self-reports							
Matched Cancer registry confirmations	Melanoma		Non-melanoma skin cancer		Breast cancer		Skin cancer	
No Match	1286	32.48%	2010	27.85%	802	6.79%	945	61.05%
Melanoma in situ	171	4.32%	26	0.36%	2	0.02%	19	1.23%
Bladder cancer			1	0.01%				
Bone metastases/bony secondaries	2	0.05%			3	0.03%	4	0.26%
Brain cancer/primary malignant brain tumour	1	0.03%						
Breast cancer	2	0.05%	3	0.04%	9995	84.62%		
Cervical cancer			1	0.01%				
Cin/precancer cells cervix	1	0.03%	1	0.01%	83	0.70%	1	0.06%
Eye cancer/adnexal cancer/reinoblastoma	38	0.96%						
Female genital tract cancer/vaginal cancer/vulval cancer	2	0.05%	6	0.08%			1	0.06%
Gum cancer			2	0.03%				
Kidney/renal cell cancer/other cancer of urinary tract	1	0.03%	1	0.01%			1	0.06%
Large bowl cancer/colorectal cancer/anal cancer/colon cancer/appendix cancer	1	0.03%	12	0.17%	1	0.01%	1	0.06%
Larynx/throat cancer			1	0.01%				
Lip cancer	2	0.05%	2	0.03%			1	0.06%
Lymphoma/Leukaemia	3	0.08%	3	0.04%			5	0.32%
Male genital tract cancer/Penis cancer	2	0.05%	4	0.06%			2	0.13%
Malignant lymph node (unspecified)	1	0.03%		0.00%	3	0.03%		
Malignant melanoma	2062	52.08%	38	0.53%	1	0.01%	76	4.91%
Metastatic unknown primary/unclassifiable	6	0.15%			4	0.03%		
Mouth cancer			2	0.03%				
Multiple myeloma	2	0.05%						
Nasal cavity cancer	1	0.03%					1	0.06%
Non-melanoma skin cancer	319	8.06%	5033	69.73%	6	0.05%	442	28.55%
Ovarian cancer/fallopian tube cancer					1	0.01%		
Primary bone cancer/Sarcoma/fibrosarcoma	5	0.13%	4	0.06%			2	0.13%
Prostate cancer	1	0.03%	3	0.04%			2	0.13%
Salivary gland cancer/parotid gland cancer	1	0.03%					1	0.06%
Thyroid cancer	1	0.03%	1	0.01%				
Tongue cancer			6	0.08%				
Tonsil cancer/oropharynx/oropharyngeal cancer			8	0.11%				
Uterine/endometrial cancer			1	0.01%				
Other	48	1.21%	49	0.68%	911	7.71%	44	2.84%
Total	3959		7218		11812		1548	

Table 3.6: Melanoma, non-melanoma, breast cancer, and 'skin cancer' self-report and closest pathological confirmation within one year to the self-report

	UK Biobank Cancer Registry Confirmations					
Matched Cancer Self-reports	Melanoma		Non-melanoma skin cancer		Breast cancer	
No Match	240	9.64%	5955	52.22%	121	1.11%
Bladder cancer			7	0.06%		
Breast cancer	3	0.12%	35	0.31%	10730	98.80%
Cancer of lip/mouth/pharynx/oral/cavity			3	0.03%		
Cervical cancer			1	0.01%		
Cin/precancer cells cervix			2	0.02%		
Eye/adnexal cancer/reinoblastoma			14	0.12%		
Female genital tract cancer/vaginal cancer/vulval cancer	1	0.04%	2	0.02%		
Kidney/renal cell cancer/other cancer of urinary tract	1	0.04%	2	0.02%		
Large bowl cancer/colorectal cancer/anal cancer/colon cancer/appendix cancer			18	0.16%		
Larynx/throat cancer			2	0.02%		
Lip cancer			4	0.04%		
Liver/hepatocellular cancer			1	0.01%		
Lung/trachea cancer	1	0.04%				
Lymphoma/Leukaemia	3	0.12%	9	0.08%		
Malignant lymph node (unspecified)			2	0.02%	2	0.02%
Malignant melanoma	2131	85.62%	341	2.99%		
Metastatic unknown primary/unclassifiable	1	0.04%	13	0.11%	4	0.04%
Mouth cancer			2	0.02%		
Multiple myeloma			1	0.01%		
Nasal cavity cancer	1	0.04%	7	0.06%		
Non-melanoma skin cancer	34	1.37%	4521	39.64%		
Oesophageal cancer			2	0.02%		
Ovarian/fallopian tube cancer			2	0.02%		
Pancreatic cancer/malignant insulinoma	2	0.08%	2	0.02%		
Primary bone cancer/Sarcoma/fibrosarcoma			14	0.12%	3	0.03%
Primary ear cancer			5	0.04%		
Prostate cancer			26	0.23%		
Skin Cancer	72	2.89%	405	3.55%		
Small intestine/small bowl cancer			1	0.01%		
Stomach cancer			1	0.01%		
Testicular cancer			2	0.02%		
Thyroid cancer			1	0.01%		
Uterine/endometrial cancer			1	0.01%		
Total	2490		11404		10860	

Table 3.7: Melanoma, non-melanoma, and breast cancer UK Cancer Registry confirmation and closest cancer self-report within one year to the confirmation

Assessing the association between the accuracy of melanoma and non-melanoma skin cancer self-reporting with sex, age, deprivation (Townsend score), education level, and years from diagnosis identified the importance of underlying factors in self-reporting accuracy. For the melanoma accuracy model, sex (effect: -0.21, $p < 0.05$), education level (effect: 0.02, $p < 0.001$), and year from diagnosis (effect: -0.003, $p < 0.001$) were all significant ($p < 0.05$) in predicting an accurate melanoma self-report (Table 3.8): where females, less time between diagnosis and self-report, and higher education level were associated with increased self-report accuracy. For the non-melanoma skin cancer accuracy model, only years from diagnosis (effect: -0.04, $p < 0.05$) was found to associate with accurately self-reporting non-melanoma skin cancer (Table 3.8): where less time between diagnosis and self-report was associated with increased self-report accuracy. For both melanoma and non-melanoma skin cancer self-reporting accuracy, these identified associated variables have been previously highlighted in recent research which determined distinctions across self-report accuracy between males and females for all cancer types, with higher education increasing self-report accuracy and with more time between diagnosis and self-report decrease accuracy (Cowdery et al. 2020).

Model:	**Variable**	**Effect estimate**	**Std. error**	**p-value**
Accuracy of melanoma	sex (male compared to female)	-0.21	0.81	< 0.05
self-reports. Using:	age	0.006	0.005	0.23
sex, age, location	Townsend score	-0.02	0.015	0.23
education level,	location	0.004	0.007	0.31
Townsend score,	education level	0.02	0.01	< 0.05
years from diagnosis	years from diagnosis	-0.003	0.037	< 0.001
Accuracy of non-melanoma	sex (male compared to female)	-0.13	0.33	0.69
self-reports. Using:	age	-0.01	0.026	0.61
sex, age, location	Townsend score	-0.03	0.06	0.61
education level,	location	-0.002	0.032	0.93
Townsend score,	education level	0.005	0.045	0.91
years from diagnosis	years from diagnosis	-0.04	0.019	< 0.05

Table 3.8: Included variable effect sizes, SEs, and p-value for accuracy of melanoma and non-melanoma self-reports.

3.3.3 Factor analysis

Exploratory factor analysis

When plotting the eigenvalues from the polychoric correlation matrix of five UK Biobank pigmentation measures, the elbow was determined to be between eigenvalue one and two, as there was a sharp drop between them. This indicates that a one or two latent variable model would be optimal for modelling the five UK Biobank pigmentation measures. (Figure 3.9).

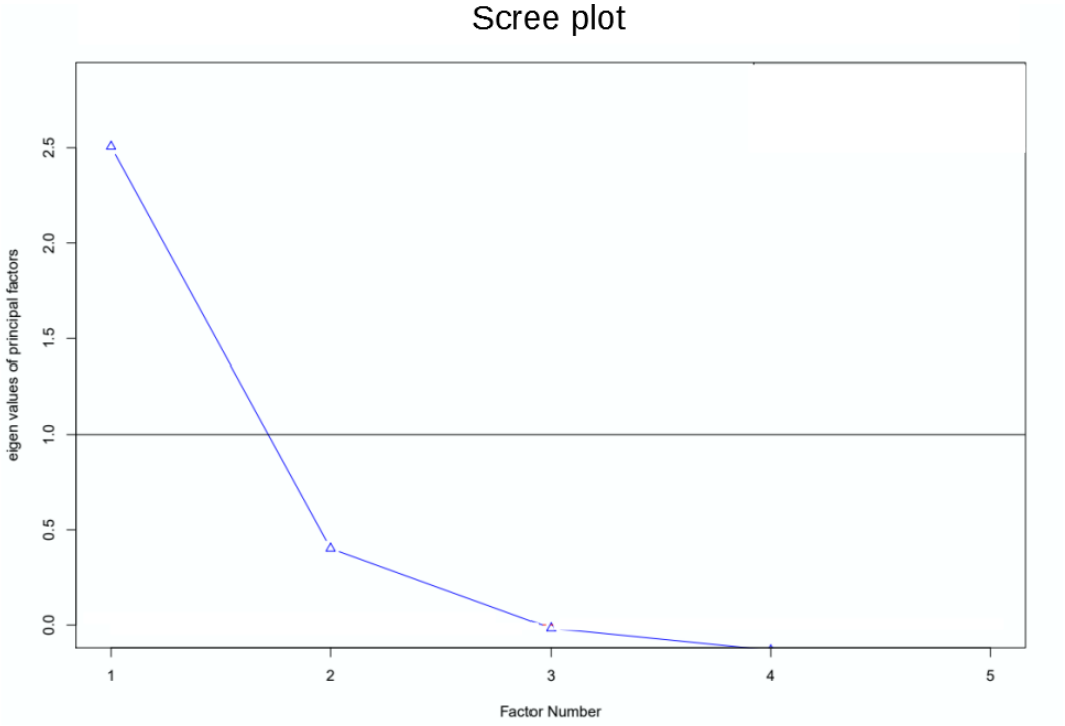


Figure 3.9: Scree plot to determine the optimum number of factors in a latent model including the following variables: hair colour, red hair colour, tanning ability, skin colour, and number of childhood sunburns.

As the one and two latent variable models were selected from the scree plot, two EFA were conducted, one with a single latent variable, and the other with two latent variables. Factor loadings in both models were similar, and showed only one latent variable was observed to effect the pigmentation variables, as no pigmentation variable passed the 0.3 effect threshold for latent variable 2 in the two latent variant model. The latent variable that effects the pigmentation traits in both models had the largest effect on red hair (factor loading of 1 for both models), second largest effect on hair colour (factor loadings of -0.8 and -0.9 for the one and two latent variables models respectively, and a modest effect on skin colour and tanning ability (factor loadings between -0.6 and -0.7 on skin colour for a one and two latent factor models, and 0.6

on tanning ability for both one and two latent factor models). Number of childhood sunburns did not pass the factor loading threshold (0.3) for any latent variable in either model. This is likely due to the inherent bias and external contributing factors to the measure (Figures 3.10, 3.10). For both models, the direction of effect for all the loadings suggest that a higher score represents a pale/red hair phenotype, as the pigmentation variables with positive loadings are recorded in a dark to light pigmentation scale (Red hair: 'non-red hair' = 0; 'red hair' = 1, Tanning ability: 'get very tanned' = 1 to 'never tan only burn' = 4), and the variables with negative loadings are recorded in a light to dark pigmentation scale (Hair colour: 'red' = 1 to 'black' = 5; Skin colour: 'very fair' = 1 to 'black' = 6). These observations suggest a single latent variable factor model is optimal for the confirmatory factor analysis. Sensitivity analysis found no single variable dominant in model determination, and systematically removing each pigmentation variable did not have an effect on the model findings.

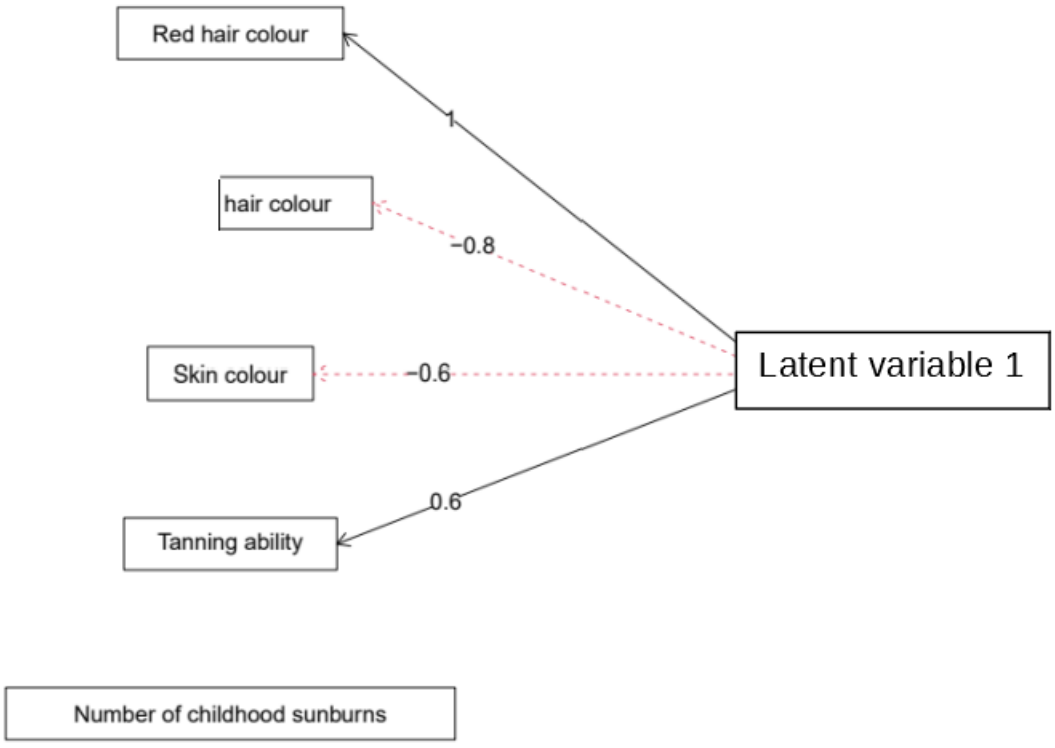


Figure 3.10: A one underlying factor latent model for a exploratory factor analysis consisting of the ordinal UK Biobank variables: ease of tanning, skin colour, hair colour, red hair, and number of childhood sunburn incidents. A maximum likelihood approach was adopted and with Oblimin rotation.

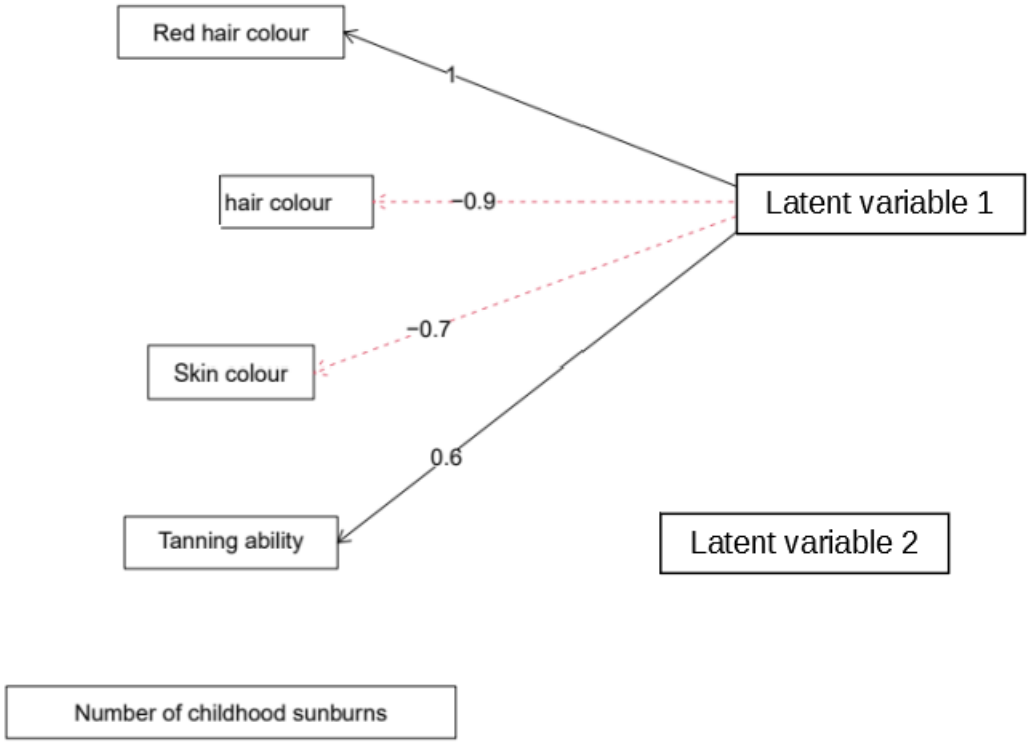


Figure 3.11: A two underlying factor latent model for an exploratory factor analysis consisting of the ordinal UK Biobank variables: ease of tanning, skin colour, hair colour, red hair, and number of childhood sunburn incidents. A maximum likelihood approach was adopted and with Oblimin rotation.

Confirmatory factor analysis

For the five pigmentation variables, the CFA model produced factor loadings, error terms, and an interaction factor between the red hair and hair colour variables. The factor model for the CFA can be found in Figure 3.12. Strong correlations were observed across skin colour, tanning ability, hair colour, and red hair in the polychoric correlation matrix, with skin colour and tanning ability being the most strongly correlated (-0.46), and red hair and skin colour having the weakest (-0.21). Number of childhood sunburns had weak polychoric correlations with the other pigmentation measures, with the strongest between skin colour and ease of tanning (-0.17 and 0.17, respectively), and the weakest between red hair and hair colour (0.11 and -0.11, respectively) (Table 3.11). As with the EFA models, the direction of effect for all the loadings in the CFA model suggest that a higher score represents a pale/red hair phenotype, as the pigmentation variables with positive loadings are recorded in a dark to light pigmentation scale (Red hair: 'non-red hair' = 0; 'red hair' = 1, Tanning ability: 'get very tanned' = 1 to 'never tan

only burn' = 4; Childhood sunburns: 'none to two incidence' = 1 to 'over three incidence' = 3), and the variables with negative loadings are recorded in a light to dark pigmentation scale (Hair colour: 'red' = 1 to 'black' = 5; Skin colour: 'very fair' = 1 to 'black' = 6). Model fit was determined to be good as the RMSEA value was 0.037, p-value <0.05, and a Tucker-Lewis index of 1.00. This suggests that a single latent variable model was a suitable fit. The standardised loadings represent the correlations between the variables and the latent variable (Table 3.9). Skin colour had the strongest factor loading of 0.89 and the lowest error term 0.21. Red hair and ease of tanning had similar factor loadings, 0.71 and 0.67, and similar error terms, 0.49 and 0.56 respectively. The hair colour variable had the fourth strongest factor loading of 0.47 and fourth highest error term of 0.78. Number of childhood sunburns had the weakest factor loading of 0.29 and the largest error term of 0.91. The strong correlation factor loading between the red hair and hair colour variables was observed at 1.07.

	Skin	Ease	Childhood sunburns	Red	hair
Skin	1	-0.46	-0.17	-0.21	0.35
Ease	-	1	0.17	0.23	-0.26
Childhood sunburns	-	-	1	0.11	-0.11
Red	-	-	-	1	-0.54
Hair	-	-	-	-	1

Table 3.9: Observed polychloric correlation between the five UK Biobank pigmentation variables: skin colour, ease of tanning, childhood sunburn incidents, red hair, and hair colour.

	Standardised factor loadings	SE of standardised factor loadings
Skin	0.89	0.21
Ease	-0.67	0.56
Childhood sunburns	-0.29	0.91
Red	-0.71	0.49
Hair	0.47	0.78

Table 3.10: Standardised factor loadings for a confirmatory factor analysis latent model which includes five UK Biobank pigmentation variables (tanning ability, skin colour, hair colour, red hair colour, and number of childhood sunburns). Standardised factor loadings represent the polychoric correlation between the observed variable and latent underlying factor (considered to be overall pigmentation).

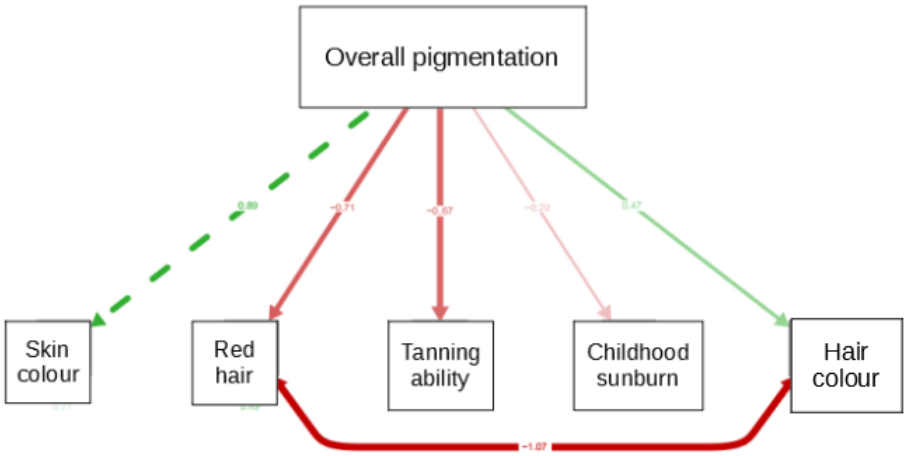


Figure 3.12: A single factor confirmatory factor analysis latent model for a confirmatory factor analysis on UK Biobank variables: ease of tanning, skin colour, hair colour, red hair, and number of childhood sunburn incidents.

When conducting sensitivity analysis, each variable was omitted in turn which had little effect on factor loadings and so these can be considered robust. For both models, the variables factor loadings continued to go from strongest to weakest in the following order: skin colour, red hair, ease of tanning, hair colour, and number of childhood sunburn incidents.

3.4 Discussion

Here we have studied the reliability of self-reporting of melanoma, non-melanoma skin cancer, and breast cancer. We have also investigated the possibility of one or more pigmentation variables underlying the various pigmentation traits recorded in UK Biobank using factor analysis.

For the EFA, one of the equal-sized subsets of UK Biobank pigmentation variables was used to provide a large scale investigation into the number of optimal latent variables in a latent model of self-reported pigmentary traits (hair colour, red hair colour, skin colour, tanning ability, and number of childhood sunburns). An initial scree plot indicated a one or two latent variable model was optimal when modelling these pigmentation traits. To investigate this further, a one and two latent variable model EFA was conducted. Both EFA identified a single latent variable to be affecting four of the five pigmentation variables (skin colour, tanning ability, hair colour, and red hair colour), with the remaining variable (number of childhood sunburns) not passing the factor loading threshold in either model for any latent variable. As all variables were affected by the

same single latent variable, it was considered to represent a person's overall pigmentation. CFA was conducted on the second sample of dichotomised UK Biobank pigmentation variables. The CFA provided reliable factor loadings (influence) for the latent variable on the five pigmentary variables. These loadings further provided evidence that overall pigmentation equally influences a person's pigmentation measures (excluding number of childhood sunburns).

For the factor analyses conducted, one of the biggest issues was considering how to incorporate red hair as a variable into the analysis. As red hair and non-red have been observed to be distinctly different (in terms of the biosynthetic processes involved in melanin production), including red hair with non-red hair as a single variable could have introduced greater variation in factor loading. Two pseudo-melanin variables were therefore considered for hair colour. When considering just eumelanin, previous studies into the distribution, quantity, and ratio of eumelanin within hair colour have found that the quantity consistently decreases across all hair colours, from dark to light, with red hair being at the lowest point in the scale. The hair colour variable used in analyses, was therefore considered a pseudo-eumelanin scale, and red hair was moved to the lightest end of the scale. Similarly, for the red hair variable used in analyses, it was considered a pseudo-pheomelanin scale, as the amount of pheomelanin is constant across non-red hair colours, and considerably higher for red hair.

Investigating melanoma self-reports and the accuracy of reporting post diagnosis used the entire UK Biobank phenotypic dataset, and provided an insight into the cancer trait report characteristics. Breast cancer was the most reliable of the self-reports investigated. This is likely due to breast cancer having a higher public profile, as well as presenting a greater burden on participants both in terms of risk and treatment compared to skin cancer. When comparing melanoma and non-melanoma skin cancer, it was observed that melanoma was more likely to be successfully self-reported if a participant had a previous diagnosis. However, melanoma self-reports were less reliable than non-melanoma skin cancer self-reports, as there was a large misreporting of other skin cancer types present, likely due to participants with non-melanoma skin cancer confusing melanoma and other skin cancers.

Although the accuracy of reporting of melanoma, non-melanoma skin cancer, and breast cancer

was identified through a matching algorithm, the source of the incorrect reporting cannot be simply quantified. The nature of these incorrect reports are likely the result of a combination of biases that may be present within the data, especially as the UK Biobank phenotypic data is retrospective. Primarily, mismatched reporting is likely the result of recall bias. Particularly for melanoma, results highlighted a higher proportion of mismatching self-reports in comparison to non-melanoma skin cancer and breast cancer, and a lower proportion of unsuccessful diagnosis reporting. This observation could be due to various factors such as confusion about terminology, with participants with differing cancers incorrectly recalling their diagnosis as melanoma or participants with multiple cancer types incorrectly recalling specific cancer diagnosis. Additionally, advances in detecting, diagnosing, recording and treating cancer has increased accuracy of cancer cases in recent years, but, as the UK Biobank participants were all age 40-69 at recruitment in 2006-2010 and may well have been diagnosed when registries were less complete, the reliability of the cancer registry will potentially be diminished. Indeed, the UK Cancer Registry wasn't national until 1970. This inconsistency may mean there are missing records of historic cases, and thus appear as incorrect self-reports in the analysis. Furthermore, selection bias may also be present in these historic cases as the further back the initial case the longer a participant has had to remain cancer free until UK Biobank recruitment.

Chapter 4

Investigating the distinct genetic architecture of red hair compared to non-red hair colour

4.1 Introduction

As it is believed the variation in hair colour is largely determined by genetic factors (1.1), and that the eumelanin/pheomelanin profile of red hair appears distinct compared to non-red hair colours, investigating the underlying genetic architectures of hair colour will provide an insight into the validity of a single hair colour ordinal scale for GWAS.

4.1.1 Investigating the genetic architecture of individual hair colour phenotypes - originality of research

Previous GWAS investigating the genetic architecture of hair colour have considered it as a single ordinal variable based on phenotype pigment. The biggest contention between these studies has been the placement of red hair within an ordered hair colour scale (blond-light brown-dark brown-black), with most studies electing to place red hair between blond and light brown hair (based on the darkness of the colour) and other placing red hair prior to blond (based on eumelanin quantity). Although there is no established position for red hair in an ordered scale of hair colour, these previous studies have nonetheless included it as an ordinal value. This inclusion of red hair, however, may not even be appropriate due to the unique ratio and quantity of eume-

lanin/pheomelanin in red hair compared to non-red hair phenotypes previously outlined (1.1.2) and the distinct environments in the biosynthetic pathways involved in the production of eumelanin and pheomelanin (1.1.1). This underlying biology suggests that there may be differences between the genetic architectures of red hair and non-red hair phenotypes.

Although the eumelanin/pheomelanin profile of red hair appears distinct compared to non-red hair colours, the similarities in the melanin characteristics between non-red hair phenotypes do not equate to similar genetic architectures. Other factors may contribute to the genetic architectures of non-red hair. Indeed, previous studies into hair colour have highlighted growth and texture to share an aetiology with blond hair only; which provides evidence that the genetic architectures across non-red hair colour may not be the same (Hysi et al. 2018).

One primary assumption made when combining hair pigmentation phenotypes into a single ordered scale is that the underlying genetic architecture is uniform across the phenotypes. However, this assumption has not been thoroughly investigated. It is therefore key, when studying the genetic architecture of hair colour, to understand how the genetic architectures of different hair colour phenotypes are related and their ability to predict overall hair colour; differing genetic architectures could result in false-positive associations and leave true-positives undetected (Type I and Type II errors) if a single underlying genetic mechanism for all hair colours is falsely assumed. Furthermore, it is also important to consider the extent of misreporting when considering hair colour as a single ordered scale, as people might misreport their hair colour — e.g. light brown instead of blond – causing genetic architecture overlap. The distinct characteristics between hair colour phenotypes have been previously stated, and it is expected that there is a large overlap in the underlying genetics across these phenotypes. By conducting this research, it will highlight the extent of this overlap, and quantify the scale of misreported findings.

4.1.2 Chapter Aims & Objectives

Aims

I. Determine whether the genetic architecture of red hair differs from that of non-red hair colours in terms of the genetic variants involved and their relative effect sizes.

II. Investigate the potential for false-positives and false-negatives regions of association for

hair colour if red hair is included in an ordinal scale of hair colour.

III. Determine whether hair colour prediction models benefit from utilising separate polygenic risk scores for red and non-red hair.

Objectives

1. Investigate pairwise GWAS of hair colour and a red vs non-red GWAS with an ordinal hair colour GWAS excluding red hair to detect differences in genetic architectures across hair colours.
2. Identify regions that are only associated with hair colour if red hair is not in the ordinal scale and those that are only associated if it isn't in the scale.
3. Create a PRS for having red/non-red hair, another PRS for ordinal hair colour (excluding red) hair and a third PRS for ordinal hair colour including red hair. Assess the extent to which these scores are associated with red/non-red hair as a binary trait.

4.1.3 Plan to achieve aims & objectives

To achieve the objects laid out in 4.1.2 the following was conducted:

For objectives (1) and (2) four GWAS were conducted. These four GWAS were based on a logistic model and were for the following comparisons: red hair vs all other hair colours; blond hair vs light brown hair; light brown hair vs dark brown hair; dark brown hair vs black hair. For objective (1), the beta coefficients (log odds ratios) for these four GWAS were compared against one another to determine whether there are any SNPs that have an effect on just a single hair colour. To achieve objective (2), the genome was split into independent LD blocks; a Bayesian approach (GWAS-PW) was then used to identify any differences in significance between these LD blocks for the four GWAS. For objective (3), two linear model GWAS were conducted: one that included red hair on a linear scale, and one that excluded red hair. This meant results from a GWAS similar to that reported by previous studies could be directly compared to a linear model GWAS that excluded red hair. The GWAS beta coefficients for the traits were compared to determine differences between the two approaches.

4.2 Methods

Analyses were conducted in PLINK 2.0 (Purcell et al. 2007), GWAS-PW (Pickrell et al. 2016), and R (R Core Team 2017) using the UK Biobank (Sudlow et al. 2015).

4.2.1 Conducting the GWAS

Initial hair colour GWAS:

Four genome-wide association studies of binary hair colour (red vs non-red hair, blond vs light brown, light brown vs dark brown, dark brown vs black) were conducted on non-overlapping subgroups of 391,679 UK Biobank cohort participants. To minimise population stratification, the analyses excluded individuals of non-European descent and adjusted for the first ten genetic principal components, sex, as well as for the genotyping array used.

As the hair colour variable in UK Biobank is self-reported, individuals who reported: 'don't know', 'prefer not to say', or did not answer were removed from the UK Biobank hair colour data. After removing these participants and conducting standard QC (see 2.3.4), individuals were distributed into the following non-overlapping hair colour sub-groups: 'red hair': 18,373; 'non-red hair': 107,420; 'blond hair': 52,991; 'light brown hair' group one: 74,438; 'light brown hair' group two: 74,438; 'dark brown hair' group one: 71,948; 'dark brown hair' group two: 71,948; and 'black hair': 41,138. Sub-groups were then paired to create four GWAS samples: (1) red/non-red hair colour of 125,793 individuals (coded 1 non-red, 2 red), (2) blond/light brown hair colour of 127,429 individuals (coded 1 light brown, 2 blond), (3) light brown/dark brown hair colour of 146,386 individuals (coded 1 dark brown, 2 light brown), and (4) dark brown/black hair colour of 76,086 individuals (coded 1 black, 2 dark brown). Phenotype files (containing participant ID and hair colour codings) and covariates files (including participant ID, the first ten genetic principal components (to reduce population stratification), sex, and genotyping array used) were created for the four sub-groups. The four binary hair colour phenotype files were used in logistic regression-based GWAS analyses whilst using the corresponding covariates file to adjust the model. PLINK 2.0 was used to conduct analyses, with summary statistic downloaded and further analysed in R.

Linear non-red hair model and logistic red hair model GWAS

Given the strong correlation seen between the estimated genetic effects from the GWAS of pairwise non-red hair colours, these hair colours were combined on a single ordinal scale. The hope was to provide further insight into the distinct genetic architecture of red hair compared to non-red hair colours. Firstly, to combine non-red hair into an ordinal scale an appropriate ordering had to be adopted. This ordering was achieved by considering the overall eumelanin quantities (as pheomelanin doesn't vary across non-red hair colours), giving a low to high eumelanin scale of 1='blond', 2='light brown', 3='dark brown', and 4='black'.

Once a single non-red hair colour ordinal scale was created, UK Biobank participant's who self-reported their hair as 'don't know', 'prefer not to say' or did not answer were, again, removed pre-analyses. Standard QC was conducted on remaining individuals (see 2.3.4), with those who passed being dichotomised into two independent samples. The first containing red haired participants (18,373) and a non-red haired sample five times the size (107,420). Separate phenotype files containing participant study ID and hair colour information were compiled for each independent sample, as well as a corresponding covariate file containing participant ID, the first ten genetic principal components (to reduce population stratification), sex, and genotyping array used. The red/non-red phenotype file was used in logistic regressed-based GWAS analysis and the non-red hair phenotype file was used in a linear regression-based GWAS analysis with the model adjusted using the variables in the covariates file. PLINK 2.0 was used to conduct analyses, with summary statistic downloaded and further analysed in R

The replication GWAS model

Having observed clear differences between the genetic architecture of non-red and red hair, I investigated the effect of conducting a GWAS of hair colour where red hair was incorrectly included into the ordinal scale. To begin, the previously defined hair colour ordinal scale (1='blond', 2='light brown', 3='dark brown', and 4='black') was expanded to incorporate red hair into the scale as the second lightest value to coincide with previously conducted hair colour GWAS (Morgan et al. 2018; Hysi et al. 2018) giving a scale of 1='blond', 2='red', 3='light brown', 4='dark brown' 5='black'.

As two hair colour ordinal scales were defined (one including red hair and one excluding red hair), UK Biobank cohort participants who self-reported their hair colour (and did not respond 'don't know', 'prefer not to say' or did not answer) and passed QC (see 2.3.4) were dichotomised into two equal sized independent samples. The first contained red haired participants and non-red haired sample (244,138 consisting of 'blond': 25,414; 'red': 18,373; 'light brown': 91,785, 'dark brown': 88,856; 'black': 19,710), and the second consisted only of non-red haired participants (247,148 consisting of 'blond': 27,820; 'light brown': 89,529; 'dark brown': 97,336; 'black': 32,463). Separate phenotype files containing participant study ID and hair colour information were compiled for each independent sample. Both phenotype files were used in linear regression-based GWAS analyses adjusted for the covariates stated previously. PLINK 2.0 was used to conduct analyses, with summary statistic downloaded and further analysed in R. To assess the potential of false-positives and false-negatives, signals (lead SNPs) uniquely present in only one of the GWAS were identified by the peak sorting algorithm, and were tabulated with rsID, position, chromosome and p-value for both conducted GWAS.

4.2.2 Investigating GWAS results

Genomic inflation was investigated by producing QQ plots and calculating the genomic inflation factor (λ) for all GWAS conducted.

Peak sorting

A peak sorting algorithm was developed to define distance based loci of association ($p < 5 \times 10^{-8}$) and identify a lead SNP for each genomic-based signal. Peaks were defined by ordering all SNPs below a p-value threshold of $p < 5 \times 10^{-8}$ and selecting the most significant SNP and all other SNPs within a $5MB$ window either side of this lead SNP. This process was then repeated for all remaining SNPs until every SNP below the p-value threshold had been allocated to a peak. A detailed R script of this peak sorting algorithm can be found in Appendix A.1.

4.2.3 Beta coefficient plotting

To compare the estimated effect size of the associated loci for each GWAS, the following summary statistics for each SNP were recorded: (1) the beta coefficient or odds ratio; (2) chromosome; (3) chromosomal position; (4) p-value for the GWAS; (5) rs-number. The most significant variant at each locus was investigated by considering significance across all GWAS results.

If a logistic model GWAS was conducted for a trait an OR was calculated. To convert these OR into beta coefficients the following equation was used:

$$\beta = ln(OR) \tag{4.1}$$

where β is the beta coefficient and OR is the odds ratio.

To assess the reliability of the genotyping/imputation of the peak SNPs, found in the peak dataset, The MAF was compared against an external dataset. The Haplotype Reference Consortium dataset was used (McCarthy et al. 2016) and merged with the peak dataset.

To plot the beta coefficients of two traits, any peak that was genome-wide significant ($p < 5 \times 10^{-8}$) for either of the two traits was selected. The lead SNP at each selected locus was chosen: if a locus was significant for only one trait then the most significant SNP for that trait was chosen as the 'lead' SNP; if traits shared a locus of association, the most significant SNP may differ between the two traits. This was addressed by using a rank thinning code (Appendix A.3). For any shared locus of association, all SNPs within a 1MB region of each trait's top SNP were assessed. The SNPs in this window were then ranked according to their significance for both trait one and trait two. A mean rank for each SNP was calculated by averaging the ranks from both traits. The SNP with the highest mean rank was used as the lead SNP for plotting. Selecting the highest mean ranked SNP minimised bias towards either trait in selection of the lead SNP to 'represent' the effect at the locus. Using these lead SNPs, beta coefficients for both traits were then plotted in R.

4.2.4 GWAS-PW

GWAS-PW was implemented to estimate the posterior probability of association (PPA) of each of four possible models: model 1 (region is associated solely with melanoma), model 2 (pigmentation alone), model 3 (Pleiotropic), model 4 (co-located but independent and model 5 (no association found with either trait).

GWAS-PW setup files

Using the summary statistics from each trait's GWAS, a setup file containing the following information on each SNP was created: (1) Chromosome; (2) Chromosomal position (BP); (3) SNP ID; (4) odds ratio or beta coefficient for trait one; (5) SE of the beta for trait one; (6) odds ratio or beta coefficient for trait two; (7) SE of the beta for trait two. Z-scores and variances representing each trait were produced for all SNPs using the following equations:

$$Z = \frac{ln(OR)}{SE} \text{ or } Z = \frac{\beta}{SE} \tag{4.2}$$

where Z is the Z-score, OR is the odds ratio, β is the beta coefficient, and SE is the SE.

$$Var = SE^2 \tag{4.3}$$

where Var is the variance and SE is the SE.

The SNP ID, chromosome, chromosomal position, Z-scores and variances for all individual SNPs were submitted to GWAS-PW for analysis.

GWAS-PW analysis results

GWAS-PW performs this analysis both at the level of the individual SNP, and also after assigning SNPs to blocks of the genome using their recommended blocks based on LD (https://bitbucket.org/nygcresearch/ldetect-data). When assigning SNPs and loci to one of the five models we used the model for the entire LD block (rather than the per SNP PPA scores), with a PPA > 0.7 selected as the most probable model. If an LD block had no PPA of > 0.7 for any model, there was uncertainty in the association and was dropped from further analysis. For each LD block, model 4 was removed and the remaining models were scaled up to simulate the distribution of model 4's PPA to these models. Results from models I,II and III were plotted using the "qqman" package in R with the following colourings: yellow - PPA for model 1 is greater than 0.7; blue - PPA for model 1 is greater than 0.7; green - PPA for model 3 is greater than 0.7.

4.2.5 Non-red/red hair colour prediction using polygenic risk scores

In addition to demonstrating that the genetic architecture of red hair is different to non-red hair colours, and that including red hair in an ordinal scale of hair colour may produce misleading associations, I sought to determine whether considering red hair and non-red hair colours as separate factors when modelling hair colour could increase hair colour prediction compared to a combined red/non-red hair factor. To achieve this, three polygenic risk scores (PRS) were produced from the using the summary statistics from the previously-conducted GWAS in UK Biobank of red vs non-red, non-red hair colour scale and red/non-red combined hair colour scale (4.2.1). For the three GWAS, the lead SNPs representing genome-wide significant ($p < 5 \times 10^{-8}$) regions of association were selected using a peak sorting algorithm (4.2.2); the beta coefficients (log odds ratio) for these lead SNPs were used as weights in the PRS. These three PRS models represented a red hair (v non-red hair) (PRS_r), non-red hair (linear scale) PRS (PRS_{nr}), and a full hair colour (linear scale) (PRS_f) that combined red and non-red hair colours. To assess red and non-red hair colour prediction using these PRS, participants were scored from the independent Leeds melanoma cohort dataset, and regressed both red hair and non-red hair on the first ten genetic principal components with each of the three PRS in turn, and then with the red (PRS_r) and non-red (PRS_{nr}) scores included in the same model. Prediction performance was quantified with AUC analyses and by estimating the proportion of variation in each trait occurrence explained for each model.

4.2.6 Skin colour and tanning ability

In addition to investigating red and non-red hair, the same QC, sample distribution, and GWAS covariates (see 4.2.1) were used to determine the whether the UK Biobank ordinal skin colour scale and ordinal tanning ability scale shared an underlying genetic architecture. For skin colour, individuals were distributed into the following non-overlapping skin colour sub-groups: 'very fair' 29,429, 'fair' group one 128,258, 'fair' group two 128,259, 'light olive' group one 34,988, 'light olive' group two 34,988, 'dark olive' group one 5,665, 'dark olive' group two 5,666, and 'brown' 3,007. Sub-groups were then paired to create four non-overlapping: (1) 'very fair/fair' for 157,687 individuals (coded 'very fair' 1, 'fair' 2); (2) 'fair/light olive' for 163,247 individuals (coded 'fair' 1, 'light olive' 2); (3) 'light olive/dark olive' for 40,653 individuals (coded 'light olive' 1, 'dark olive' 2); (4) 'dark olive/brown' of 8,673 individuals (coded 'dark olive' 1, 'brown' 2). Similarly

for tanning ability, individuals were divided into the following non-overlapping tanning ability sub-groups: 'Never tan, only burn' 63,608, 'occasionally tan, mostly burn' group one 39,022, 'occasionally tan, mostly burn group' two 39,022, 'moderately tan, occasionally burn' group one 73,099, 'moderately tan, occasionally burn' group two 73,099, and 'only tan never burn' 79,381. Sub-groups were then paired to create three non-overlapping GWAS samples: (1) 'Never tan, only burn/occasionally tan, mostly burn' for 102,630 individuals (coded 'never tan, only burn': 1, and 'occasionally tan, mostly burn' 2); (2) 'occasionally tan, mostly burn/moderately tan, occasionally burn' for 112,121 individuals (Coded 'occasionally tan, most burn': 1, 'moderately tan, occasionally burn':2); (3) 'moderately tan, occasionally burn/only tan, never burn' for 152,480 individuals (coded 'moderately tan, occasionally burn': 1, 'only tan, never burn': 2).

For both traits, phenotype files (containing participant ID and the corresponding skin colour/-tanning ability codings) and covariate files (including participant ID, the first ten genetic principal components to reduce population stratification, sex, and genotyping array used) were created for the sub-groups. Each phenotype file created was used in logistic regressed-based GWAS analyses whilst using the corresponding covariate file to adjust the models. PLINK 2.0 was used to conduct analyses, with summary statistics further analysed in R.

Peak sorting was conducted on all seven GWAS results (see 4.2.2) and beta coefficients (log odds) were plotted to determine the similarity between the underlying genetic architectures across skin colour and tanning ability (see 4.2.3).

4.3 Results

4.3.1 Binary GWAS

Genomic inflation factors from the four binary hair colour GWAS were red vs non-red: 0.90; blond vs light brown: 1.00; light brown vs dark brown: 1.00; dark brown vs black: 0.92 (Figure 4.2). Genome-wide significant ($p < 5 \times 10^{-8}$) association, through the peak sorting algorithm, was detected for red hair vs non-red at 14 loci, for blond vs light brown at 49 loci, for light brown vs dark brown at 43 loci, and for dark brown vs black at 17 loci (Figure 4.1).

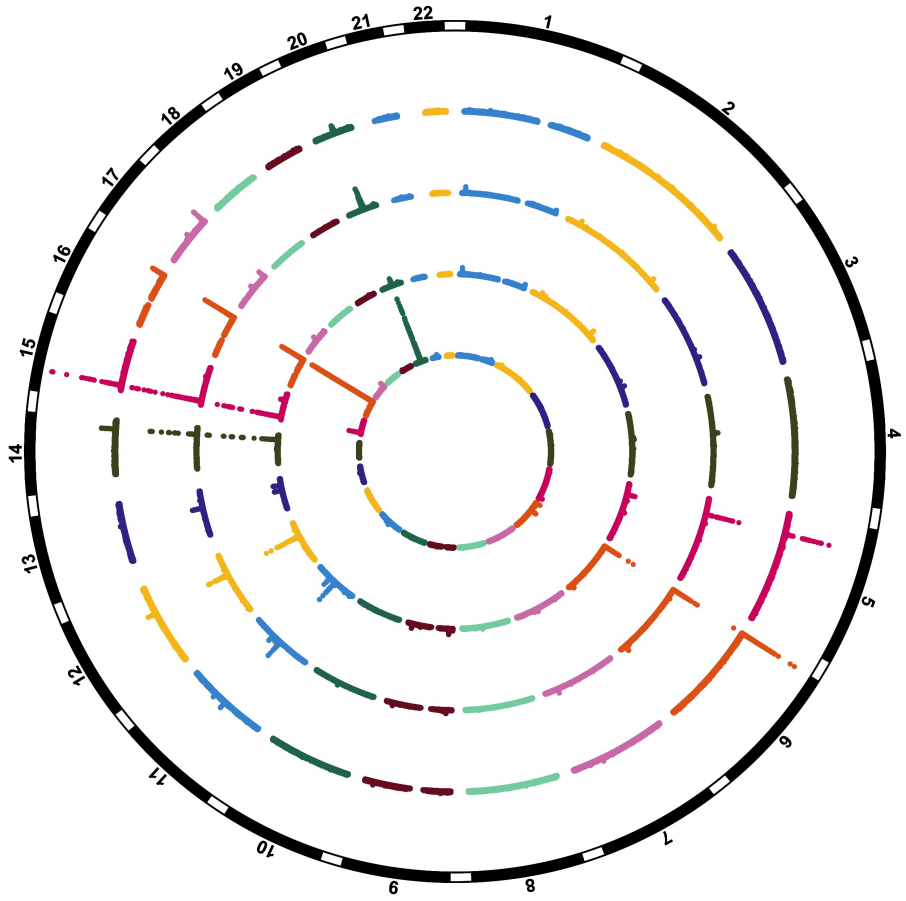


Figure 4.1: Circular Manhattan plot for four binary hair colour GWAS from UK Biobank. From the centre out: (1) red/non-red hair colour of 125,793 individuals (coded 1 non-red, 2 red), (2) blond/light brown hair colour of 127,429 individuals (coded 1 light brown, 2 blond), (3) light brown/dark brown hair colour of 146,386 individuals (coded 1 dark brown, 2 light brown), and (4) dark brown/black hair colour of 76,086 individuals (coded 1 black, 2 dark brown)

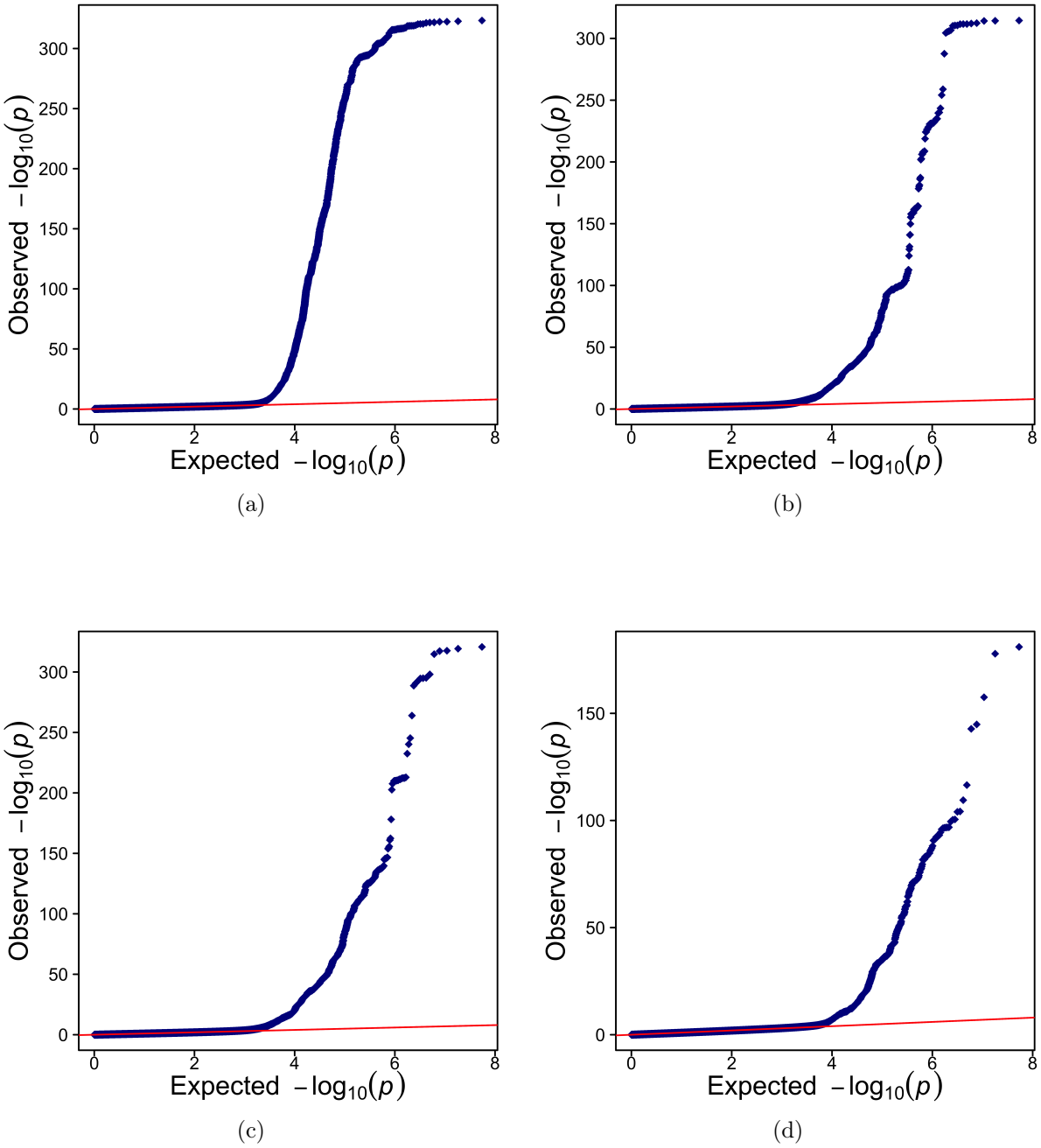
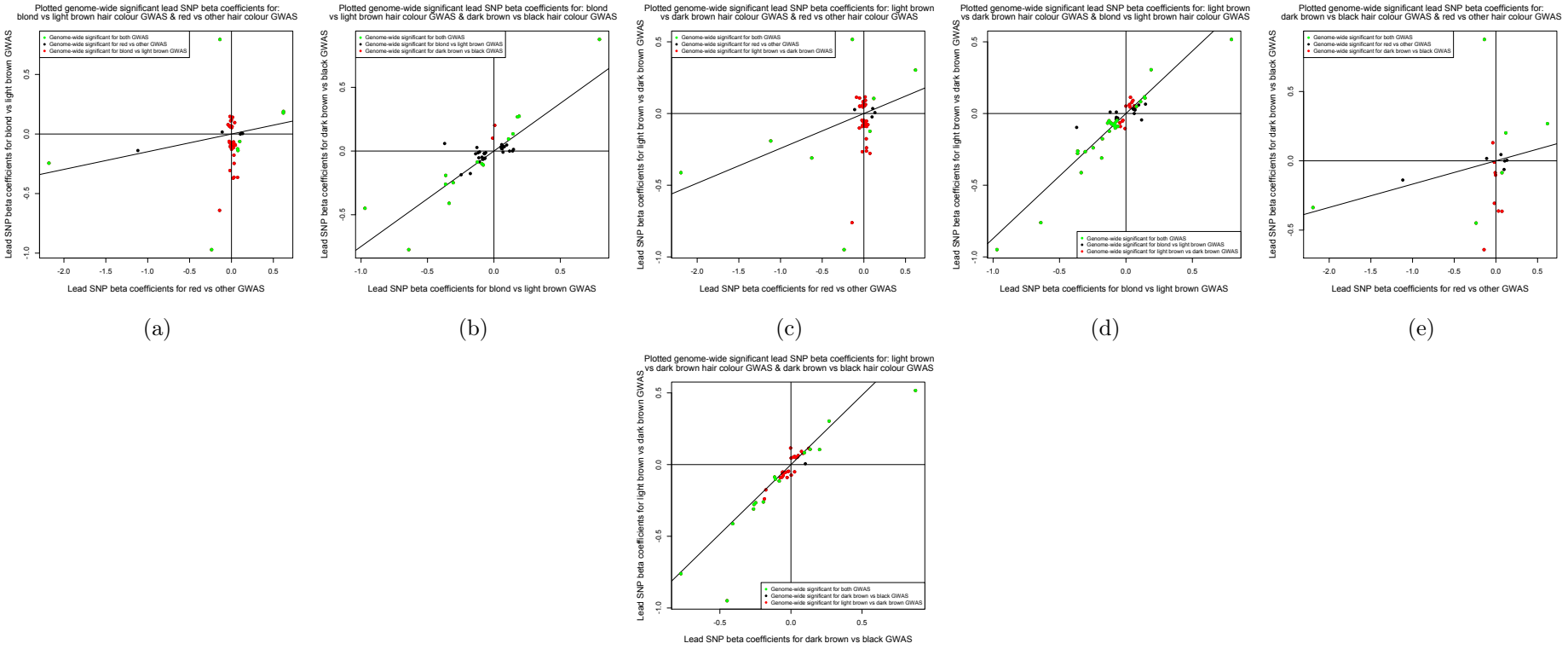


Figure 4.2: Q-Q plots for four binary hair colour GWAS conducted: (a) red/non-red hair colour of 125,793 individuals (coded 1 non-red, 2 red), (b) blond/light brown hair colour of 127,429 individuals (coded 1 light brown, 2 blond), (c) light brown/dark brown hair colour of 146,386 individuals (coded 1 dark brown, 2 light brown), and (d) dark brown/black hair colour of 76,086 individuals (coded 1 black, 2 dark brown)

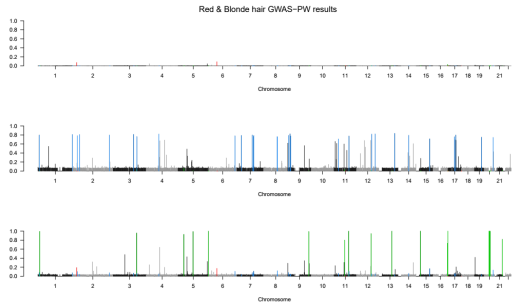
The betas (log odds ratios) for the lead SNPs at each locus of association were highly correlated for the non-red hair colours, ranging from $\rho = 0.94$ (light brown vs dark brown against blond vs light brown) to $\rho = 0.87$ (blond vs light brown against dark brown vs black). In comparison the red hair log odds ratios were only weakly correlated with those from the non-red hair GWAS, ranging from $\rho = 0.36$ (red vs non-red against light brown vs dark brown) to $\rho = 0.21$ (red vs non-red against blond vs light brown) (Figure 4.3).

A joint analysis of non-red hair trait GWAS using GWAS-PW to find shared regions of association found that any locus genome-wide significant for a non-red GWAS had a posterior probability > 0.7 for being associated with any other non-red GWAS. However, when comparing the red hair trait against non-red hair traits, GWAS-PW reported 54 loci with a high posterior probability (> 0.7) of being associated only with non-red hair colours. No regions were found to have a posterior probability (> 0.7) of being associated only with red hair (Figure 4.4). When considering a more stringent posterior probability threshold (> 0.9) a similar result was observed for non-red hair colours, with GWAS-PW only identifying shared loci across all non-red hair phenotypes and no unique associations across loci (Blond vs light brown compared with light brown vs dark brown: 47 shared loci; Blond vs light brown compared with dark brown vs black hair: 38 shared loci; and light brown vs dark brown compared with dark brown vs black: 72 shared loci). Comparing red and non-red GWAS using a more stringent posterior probability (> 0.9), GWAS-PW identified no unique loci between red and non-red hair colour GWAS (compared to the less stringent posterior probability (> 0.7) which identified 54 non-red hair colour only loci) and 19 shared genomic loci between red hair and non-red joint analyses.

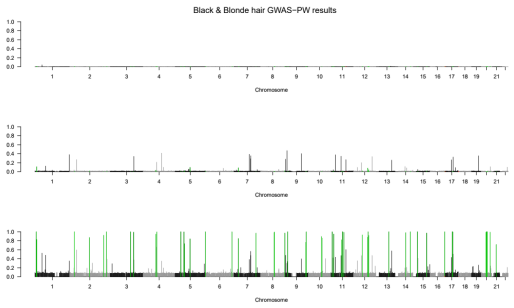


(a) (b) (c) (d) (e) (f)

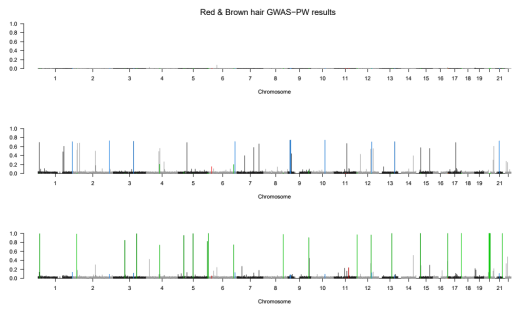
Figure 4.3: Comparison of estimated effect sizes for pairwise GWAS of adjacent categories of hair colour. Beta coefficients (log odds) of Lead SNPs for each GWAS are plotted for: (a) Blond vs light brown plotted against dark brown vs black GWAS; (b) Blond vs light brown plotted against light brown vs dark brown GWAS; (c) Dark brown vs black plotted against light brown vs dark brown; (d) red vs non-red plotted against dark brown vs black; (e) red vs non-red plotted against blond vs light brown; (f) red vs non-red plotted against light brown vs dark brown. Green points represent Lead SNPs genome-wide significant ($p < 5 \times 10^{-8}$) for both traits; red points represent Lead SNPs genome-wide significant ($p < 5 \times 10^{-8}$) for the trait on the x-axis; and black points represent Lead SNPs genome-wide significant ($p < 5 \times 10^{-8}$) for the trait on the y-axis.



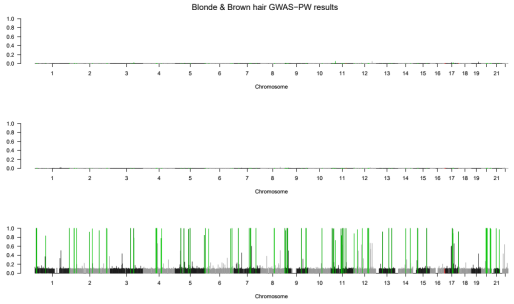
(a)



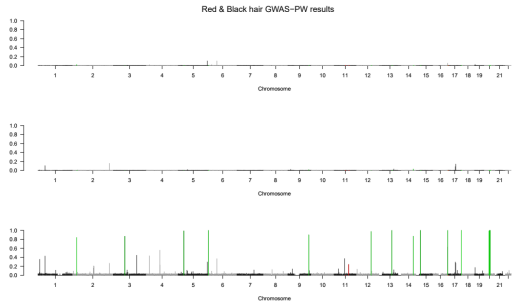
(b)



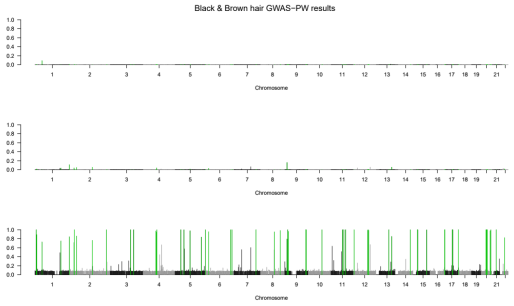
(c)



(d)



(e)



(f)

Figure 4.4: GWAS-PW analysis between binary hair colour GWAS: (a) Blond vs light brown plotted against dark brown vs black GWAS; (b) Blond vs light brown plotted against light brown vs dark brown GWAS; (c) Dark brown vs black plotted against light brown vs dark brown; (d) red vs non-red plotted against dark brown vs black; (e) red vs non-red plotted against blond vs light brown; (f) red vs non-red plotted against light brown vs dark brown. The top plot indicates the probability of each window being uniquely associated with first hair colour (model I); the middle plot indicates the probability of each window being associated with the second hair colour (model II); and the bottom plot indicates the probability of each window being associated uniquely associated with both hair colours (model III). windows are coloured yellow if the probability is > 0.7 for model I, blue for probability is > 0.7 for model II, and green for probability is > 0.7 for model III.

4.3.2 Non-red hair vs red hair

As the genetic correlation between non-red hair colours was strong, an ordinal non-red hair colour GWAS and binary red v non-red hair colour GWAS were conducted. Genomic inflation factors from the non-red and red hair colour GWAS were 0.91 and 1.11, respectively (Figure 4.6). This is expected due to the large sample cohort used for analyses and the power to detect polygenic effects (Jian Yang et al. 2011). Genome-wide significant ($p < 5 \times 10^{-8}$) loci of association, through the peak sorting algorithm, were detected for red hair at 14 loci, and for non-red hair at 107 loci. Of all associated loci, three were uniquely associated with red hair, 96 uniquely with non-red hair, and 11 associated with both red and non-red hair (Figure 4.5).

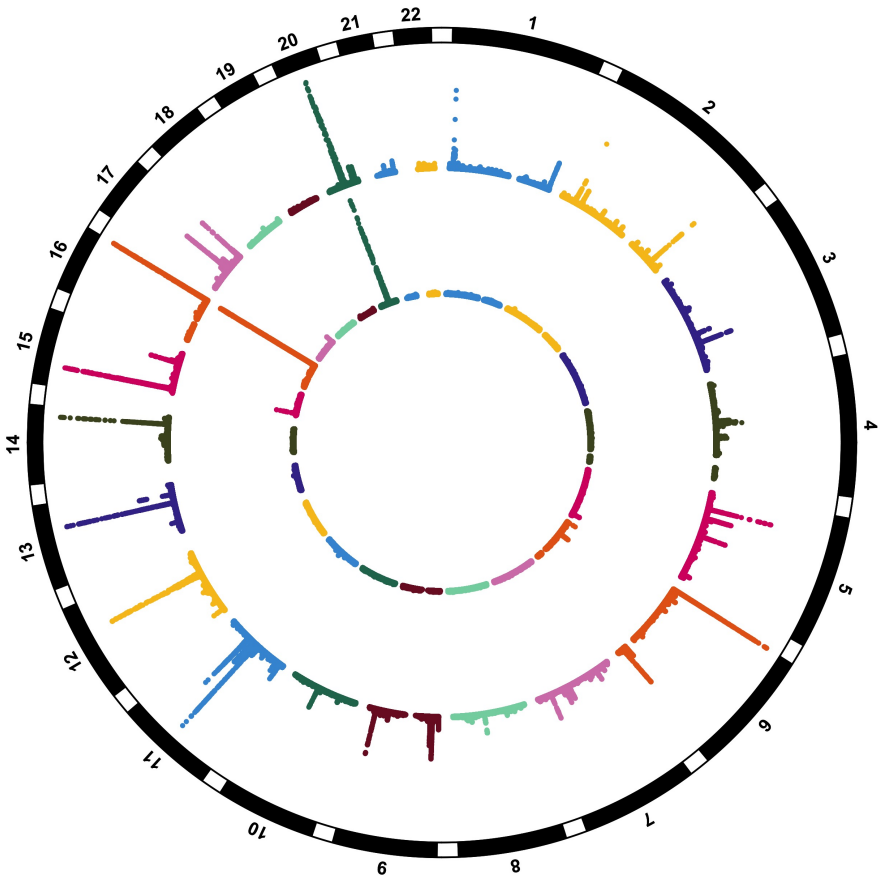
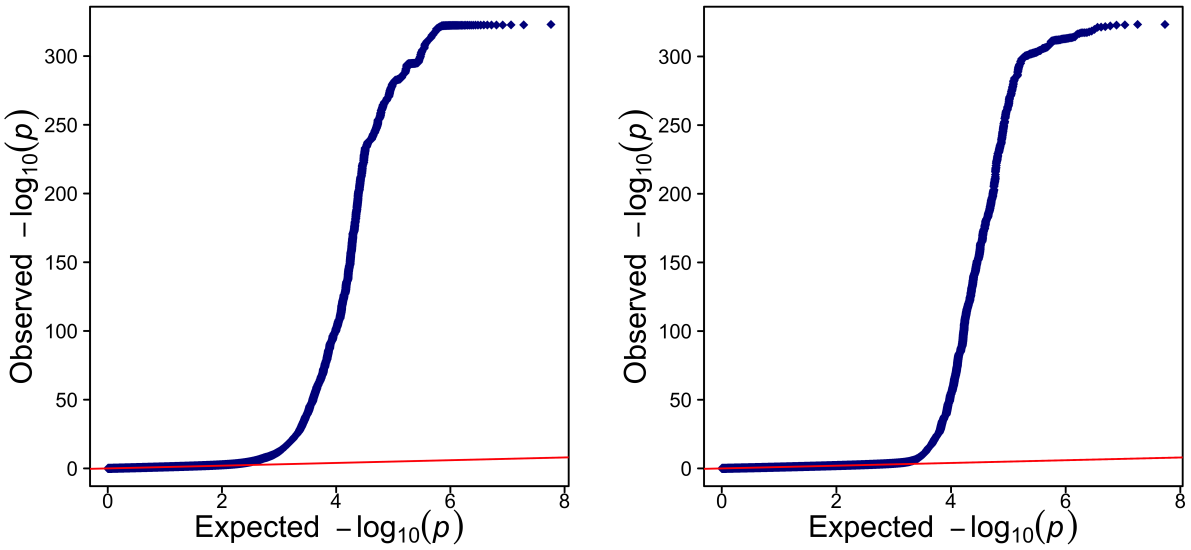


Figure 4.5: Circular Manhattan plot for a red hair and non-red hair GWAS from UK Biobank. From the centre out: (1) Red hair vs non-red hair (binary GWAS of 18,373 red hair and 107,420 non-red hair individuals) and (2) a non-red hair colour GWAS (Ordinal GWAS of 243,623 individuals)



(a) (b)

Figure 4.6: Q-Q plots for the two hair colour GWAS conducted: Red hair vs non-red hair (binary GWAS of 18,373 red hair and 107,420 non-red hair individuals) and a non-red hair colour GWAS (Ordinal GWAS of 243,623 individuals)

The lead SNP log odds ratios (betas) from the loci of association were only weakly correlated between red hair and non-red hair (0.29). Notably, three SNPs (rs2025753 in the *PKHD1* gene, rs60780889 in the *LOC105377739* gene, rs76645364 in the *POMC* gene) reached genome-wide significance in the red/non-red hair GWAS but not in the linear non-red hair GWAS. When exploring these three SNPs, no previous association with red hair has been reported (Figure 4.7).

When identifying joint regions of association through GWAS-PW, results were found to be consistent with previous findings on the relationship of red hair when compared to non-red hair colours. Comparing the ordinal non-red hair GWAS to the binary red hair GWAS identified 36 jointly associated regions (Model III probability > 0.7), of which, all were found to be genome-wide significantly associated with red hair (Figure 4.8). This is likely due to the effective sample size for the red/non-red GWAS being much smaller than the linear non-red hair GWAS. When considering a more stringent posterior probability threshold (> 0.9) a similar result was observed with 26 jointly-associated regions being identified between the binary red hair GWAS and ordinal non-red hair GWAS; all of which were genome-wide significantly associated with red hair and non-red hair in their corresponding GWAS.

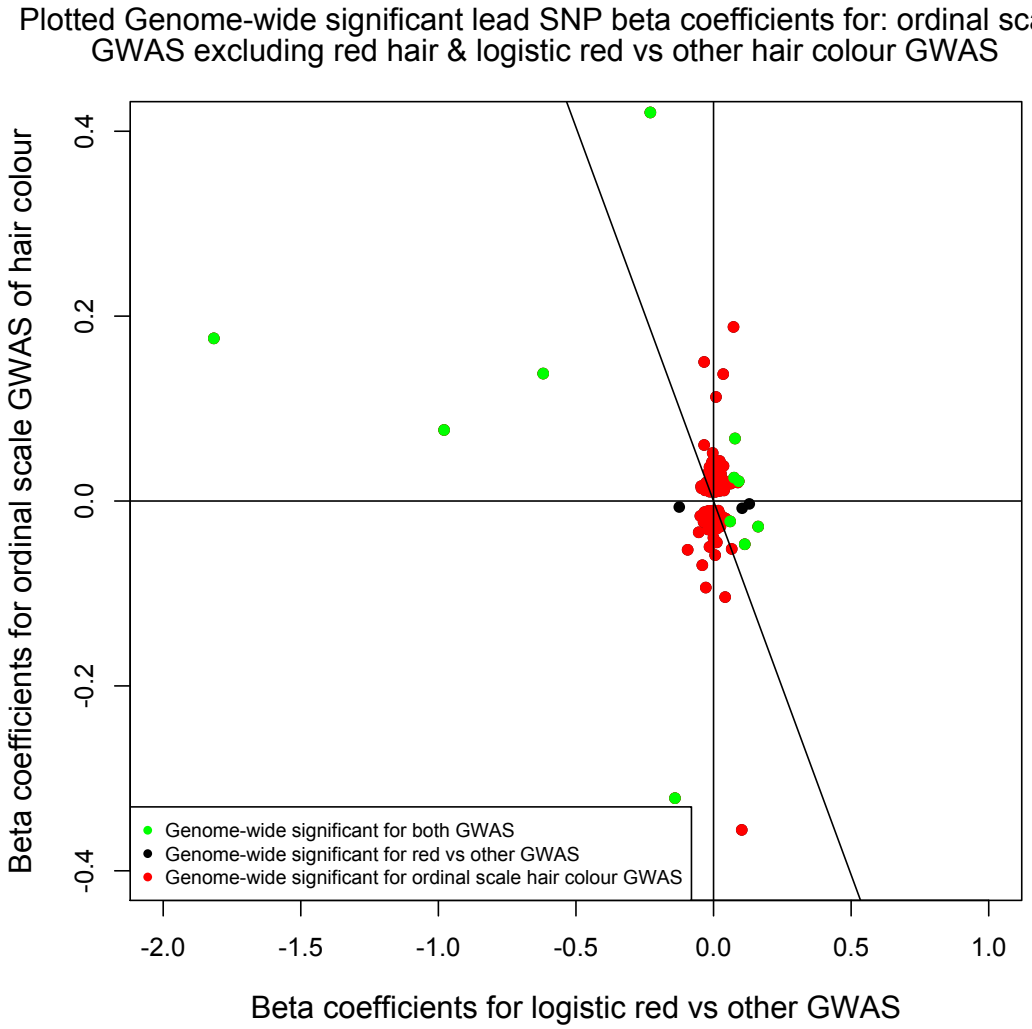


Figure 4.7: Comparison of estimated effect sizes of two hair colour GWAS: Red hair vs non-red hair (binary GWAS of 18,373 red hair and 107,420 non-red hair individuals) against a non-red hair colour GWAS (Ordinal GWAS of 243,623 individuals. Green points represent Lead SNPs genome-wide significant ($p < 5 \times 10^{-8}$) for both traits; red points represent Lead SNPs genome-wide significant ($p < 5 \times 10^{-8}$) for the trait on the x-axis; and black points represent Lead SNPs genome-wide significant ($p < 5 \times 10^{-8}$) for the trait on the y-axis.

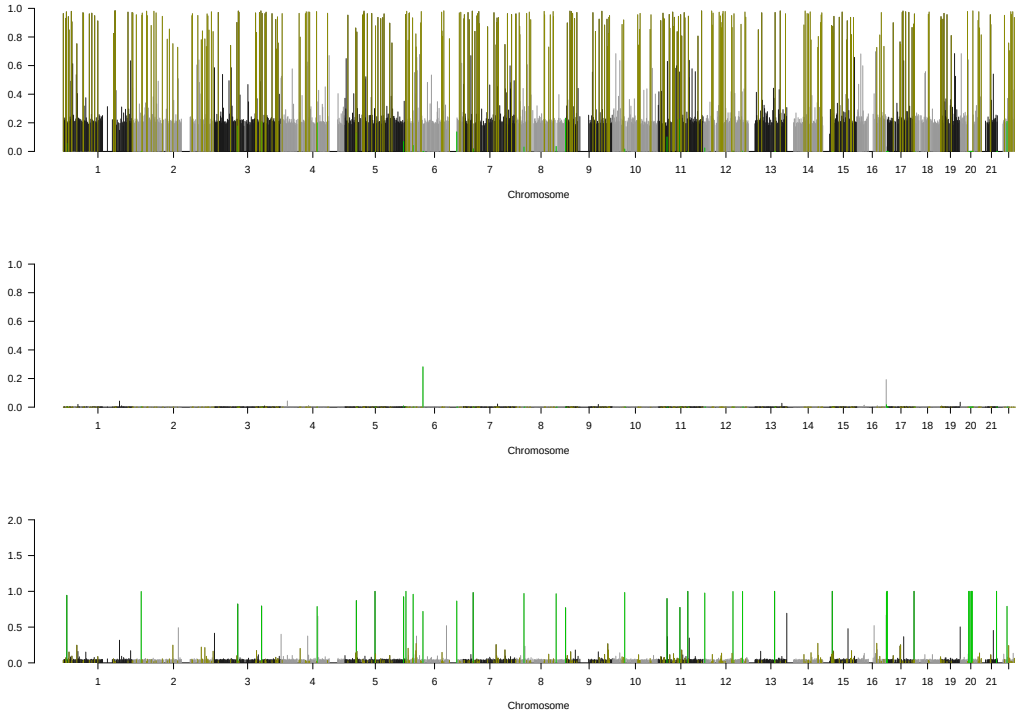


Figure 4.8: GWAS-PW analysis between non-red hair and red hair. The top plot indicates the probability of each window being uniquely associated with non-red hair (model I); the middle plot indicates the probability of each window being associated with red hair (model II); and the bottom plot indicates the probability of each window being associated uniquely associated with non-red hair and red hair (model III). windows are coloured yellow if the probability is > 0.7 for model I, blue for probability is > 0.7 for model II, and green for probability is > 0.7 for model III.

4.3.3 Non-red vs replication

The genomic inflation factors from the ordinal hair colour GWAS including red hair and the ordinal non-red hair colour GWAS were 1.06 and 1.07, respectively (Figure 4.10). Using the peak sorting algorithm, genome-wide significant loci of association were detected for ordinal hair colour including red hair at 84 loci, and for ordinal non-red hair at 107 (Figure 4.9).

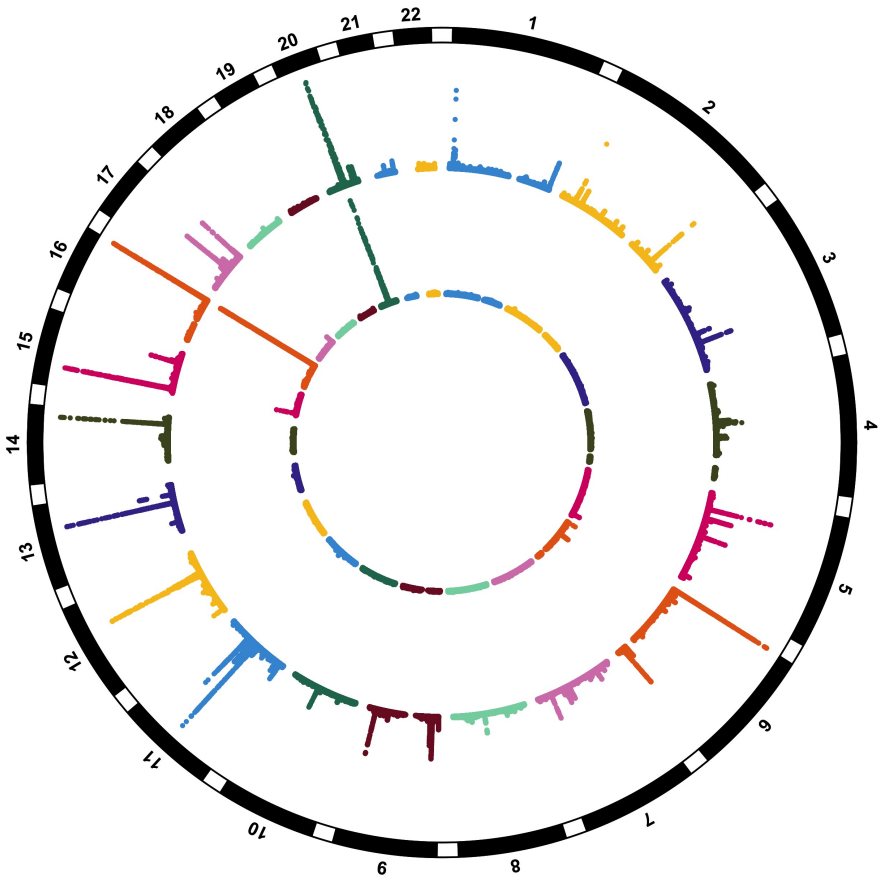
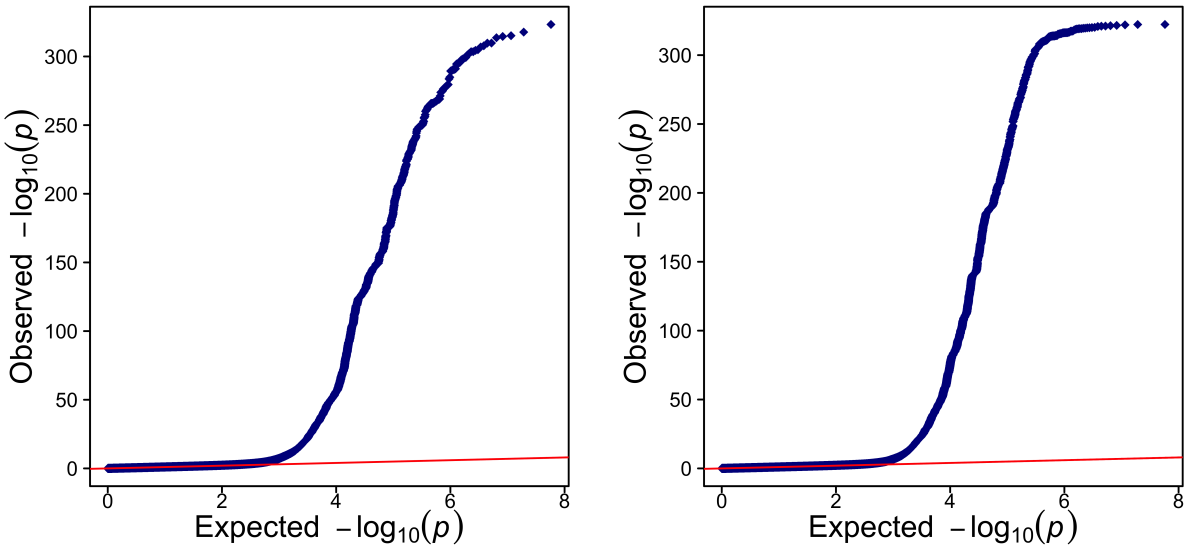


Figure 4.9: Circular Manhattan plot for two hair colour GWAS conducted: ordinal scale hair colour GWAS including red hair and an ordinal scale hair colour GWAS excluding red hair



(a) (b)

Figure 4.10: Q-Q plots for two hair colour GWAS conducted: ordinal scale hair colour GWAS including red hair and an ordinal scale hair colour GWAS excluding red hair

Lead SNPs from each locus of association were identified and, unsurprisingly, a strong Pearson's correlation between the beta coefficients for the hair colour GWAS including and excluding red hair (0.99, 95%CI $(0.991, 0.995)$, p-value < 0.05) was observed. However, the linear trend suggests that including red hair may marginally dilute the effect size of the beta coefficients within a linear regression model (Regression coefficient $= 0.79$) when regressing the beta coefficients from the GWAS including red hair on the beta coefficients from the GWAS excluding red hair (Figure 4.11).

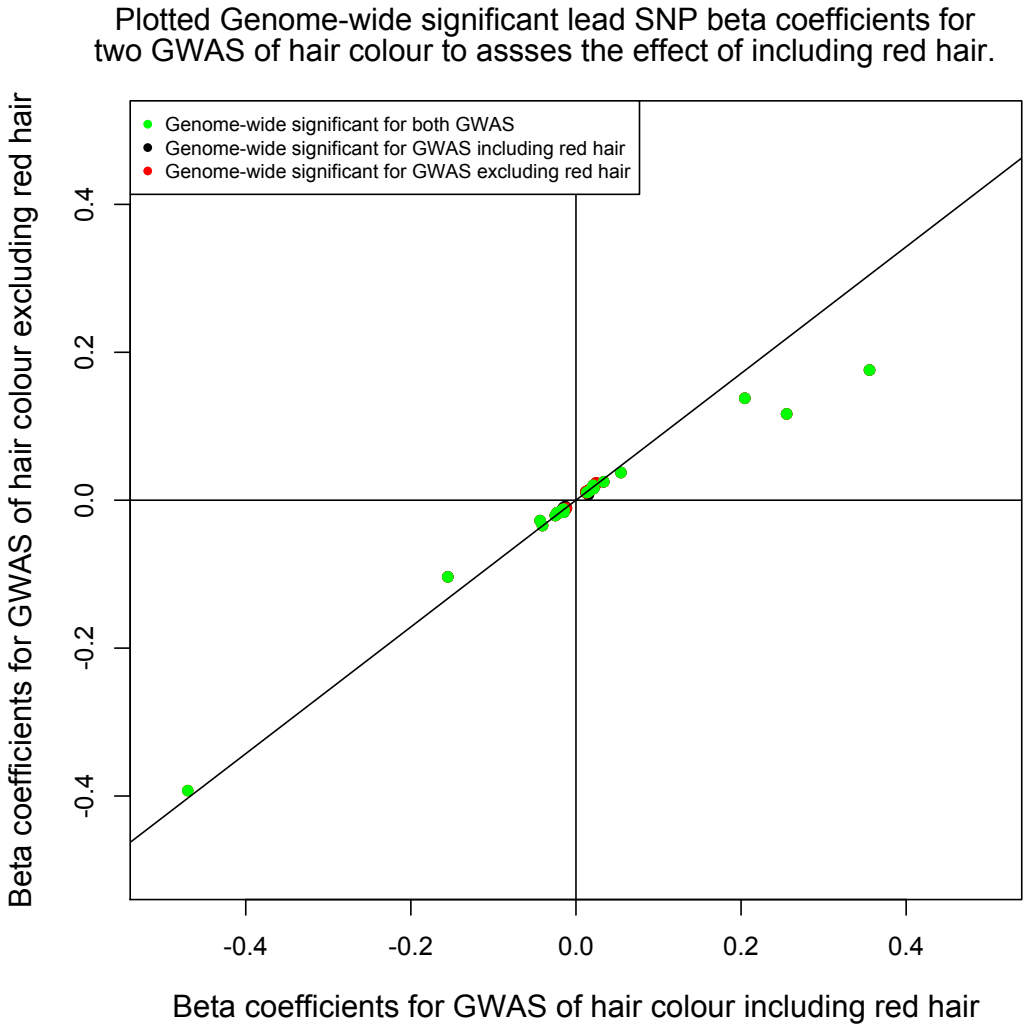


Figure 4.11: Comparison of estimated effect sizes of two hair colour GWAS: ordinal scale hair colour GWAS including red hair and an ordinal scale hair colour GWAS excluding red hair. Green points represent Lead SNPs genome-wide significant ($p < 5 \times 10^{-8}$) for both traits; red points represent Lead SNPs genome-wide significant ($p < 5 \times 10^{-8}$) for the trait on the x-axis; and black points represent Lead SNPs genome-wide significant ($p < 5 \times 10^{-8}$) for the trait on the y-axis.

Furthermore, the potential for misleading results was assessed, such as false-positives and false-negatives, by assessing the lead SNPs that were genome-wide significant for only one of the conducted GWAS. Seven lead SNPs were found to be uniquely significant when including red hair into the ordinal scale, and sixteen were found to be uniquely significant when excluding red hair (Table 4.1).

rsID	Position	CHR	p-value inc red hair in GWAS	p-value exc red hair in GWAS
rs2474890	51807089	6	2.95171E-10	4.44604E-05
rs1265564	111708458	12	1.31851E-09	2.91695E-07
rs4660116	236028841	1	1.99895E-09	5.47575E-08
rs9924516	16975959	16	6.82105E-09	1.64745E-07
rs7784327	35824415	7	1.67622E-08	9.47967E-08
rs62117462	41900813	19	3.4011E-08	1.04442E-06
rs1624162	33697765	18	4.89072E-08	4.83721E-07
rs10402247	18522805	19	5.11809E-08	2.2419E-08
rs7999377	114825471	13	6.35967E-08	6.18835E-12
rs10802861	240485258	1	7.65554E-08	1.86448E-09
rs698083	186997742	3	7.8482E-08	5.71612E-09
rs2060296	139323257	5	1.00585E-07	2.30555E-08
rs11701104	36238517	21	1.08892E-07	8.07866E-10
rs4282273	66789202	5	1.1757E-07	5.62139E-10
rs11029693	26769140	11	1.80297E-07	3.97213E-11
rs77717551	125368342	12	2.20236E-07	1.70343E-09
rs188216331	33950201	5	2.26533E-07	1.70589E-08
rs2617515	98123347	5	5.81895E-07	1.34187E-08
rs3749110	109605767	2	6.33689E-07	2.02815E-08
rs45510602	118917587	11	6.66587E-07	1.56687E-08
rs4785217	50589243	16	2.69466E-06	4.10642E-08
rs7583880	207999713	2	3.42901E-06	1.58625E-08
rs11698187	62960292	20	1.29834E-05	3.29946E-08

Table 4.1: Lead SNPs that were uniquely genome-wide significant ($p < 5 \times 10^{-8}$) in a GWAS of hair colour which included red hair in the analysis or were uniquely genome-wide significant in a GWAS of hair colour which excluded red hair in the analysis

4.3.4 Polygenic risk scores

Three polygenic risk scores were generated using (i) the red/non-red hair colour GWAS defined in 4.2.1 (PRS_r) with 14 variants using the associated beta coefficients as weightings; (ii) the ordinal non-red hair colour GWAS defined in 4.2.1 (PRS_{nr}) using 107 variants with the associated beta coefficients as weightings; and (iii) the ordinal hair colour GWAS (including red hair) defined in 4.2.1 (PRS_f) using 84 variants with the associated beta coefficient as weightings. All SNPs included in PRS passed the significance threshold ($p < 5 \times 10^{-8}$).

When predicting non-red hair, the PRs for the non-red hair score (PRS_{nr}) explained the most variation ($R^2 = 0.122$) compared to the red score (PRS_r, $R^2 = 0.013$) and full hair colour score (PRS_f, $R^2 = 0.110$) models. When combining our PRS_{nr} and PRS_r into the same model a marginal increase on variation explained was observed ($R^2 = 0.123$) compared to the non-red (PRS_{nr}) PRS alone. When considering red hair prediction, we found red hair to be predicted significantly better by our red hair score (PRS_r, $R^2 = 0.254$) compared to our non-red and full hair colour scores ($PRS_{nr} : R^2 = 0.018$ & $PRS_f : R^2 = 0.046$). Additionally, including our red (PRS_r) and non-red scores (PRS_{nr}) into the same model saw no increase in variation explained ($R^2 = 0.254$). All scores contained a single 'top' variants for *MC1R*, selected through the peak sorting algorithm.

4.3.5 Skin colour and tanning ability

All four skin colour GWAS and all three tanning ability GWAS conducted had high genetic inflation factors ('very fair' vs 'fair': 0.95; 'fair' vs 'light olive': 0.97; 'light olive' vs 'dark olive': 0.76; 'dark olive' vs 'brown': 0.63; and' never tan only burn' vs 'occasionally tan mostly burn': 1.00; 'occasionally tan mostly burn' vs 'moderately tan occasionally burn': 0.98; 'moderately tan occasionally burn' vs 'never burn only tan': 0.95; (Figures 4.14, 4.15). For the skin colour GWAS, genome-wide significant loci of association, through the peak sorting algorithm, were detected for 'very fair' vs 'fair' at 15 loci, for 'fair' vs 'light olive' at 37 loci, for 'light olive' vs 'dark olive' at 1 loci, and 'dark olive' vs 'brown' at 2 loci (Figure 4.12). For tanning ability, genome-wide significant loci of association, through the peak sorting algorithm, were detected for 'never tan only burn' vs 'occasionally tan mostly burn' at 13 loci, for 'occasionally tan mostly burn' vs 'moderately tan occasionally burn' at 8 loci, and 'moderately tan occasionally burn' vs 'never burn only tan' at 18 loci (Figure 4.13).

For skin colour, the betas (log odds ratios) for the lead SNPs at each locus of association were highly correlated across each skin colour comparison, these correlations ranged from $\rho = 0.93$ ('very fair' vs 'fair' with 'dark olive' vs 'brown') to $\rho = -0.57$ ('very fair' vs 'fair' with 'light olive' vs 'dark olive') (Figure 4.16). For tanning ability, the betas (log odds ratios) for the lead SNPs at each locus of association were highly correlated across each skin colour comparison, these

correlations ranged from $\rho = 0.82$ ('occasionally tan' vs 'never tan only burn' with 'moderately tan occasionally burn' vs 'occasionally tan') to $\rho = 0.74$ ('moderately tan' vs 'occasionally tan' with 'only tan, never burn' vs 'moderately tan') (Figure 4.17).

Figure 4.12: Circular Manhattan plot for four binary skin colour GWAS from UK Biobank. From the centre out: (1) very fair/fair of 157,687 individuals (coded very fair 1, fair 2); (2) fair/light olive of 163,247individuals (coded fair 1, light olive 2); (3) light olive/dark olive of 40,653 individuals (coded light olive 1, dark olive 2); (4) dark olive/brown of 8,673 individuals (coded dark olive 1, brown2)

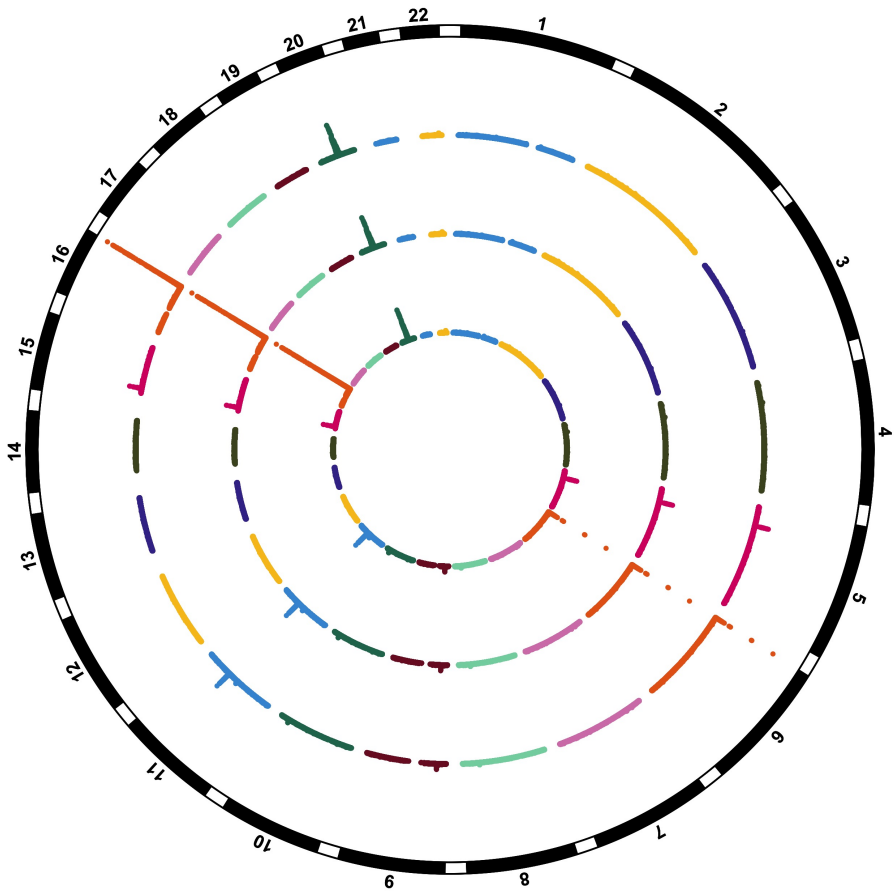
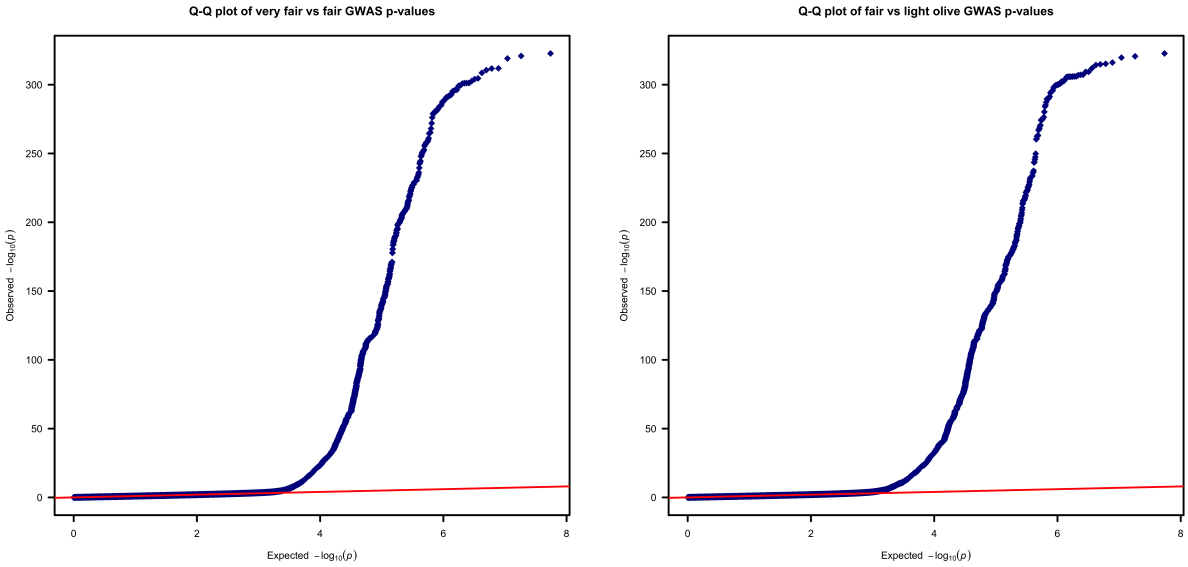
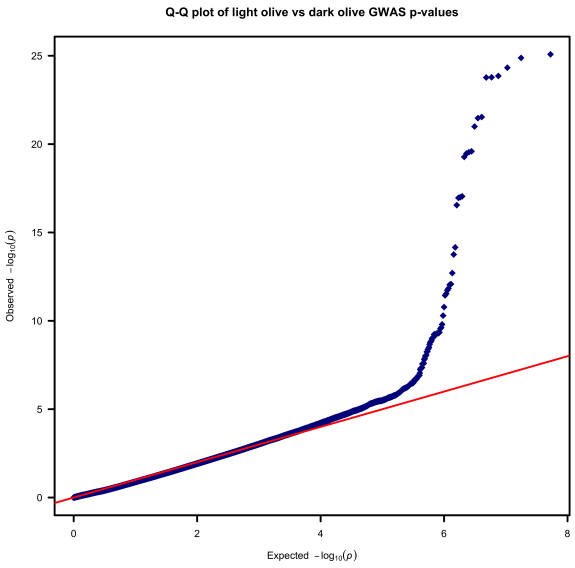


Figure 4.13: Circular Manhattan plot for three binary tanning ability GWAS from UK Biobank. From the centre out: (1) Never tan, only-burn/occasionally tan, mostly burn of 102,630 individuals (coded never tan, only burn: 1, and occasionally tan, mostly burn 2); (2) occasionally tan, mostly burn/moderately tan, occasionally burn of 112,121 individuals (coded occasionally tan, most burn: 1, moderately tan, occasion-ally burn:2); (3) moderately tan, occasionally burn/only tan, never burn of 152,480 individuals (coded moderately tan, occasionally burn: 1, only tan, never burn:2)

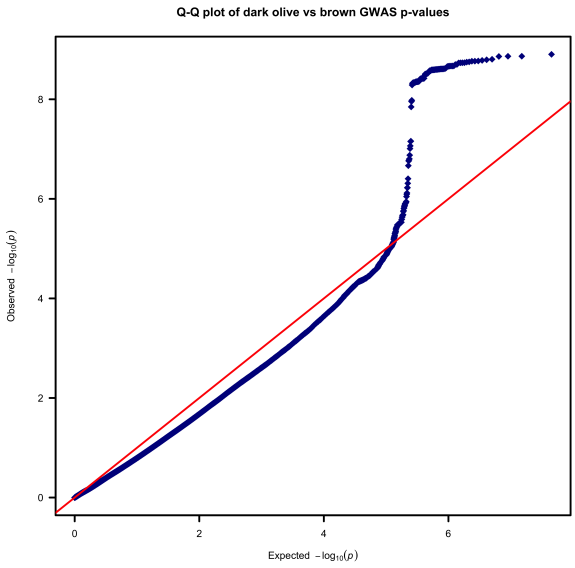


(a)

(b)

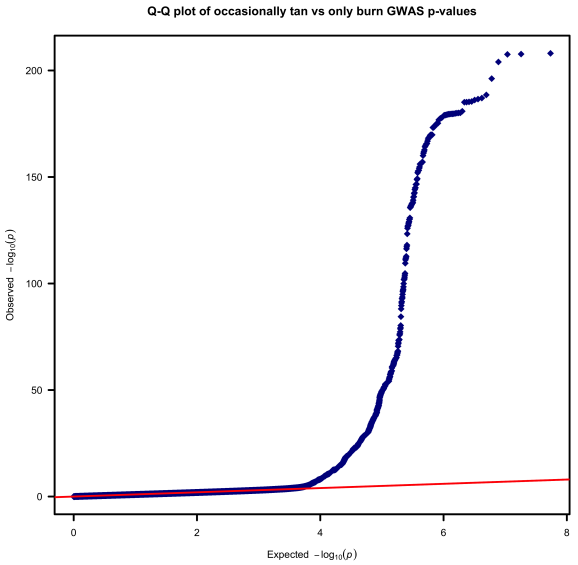


(c)

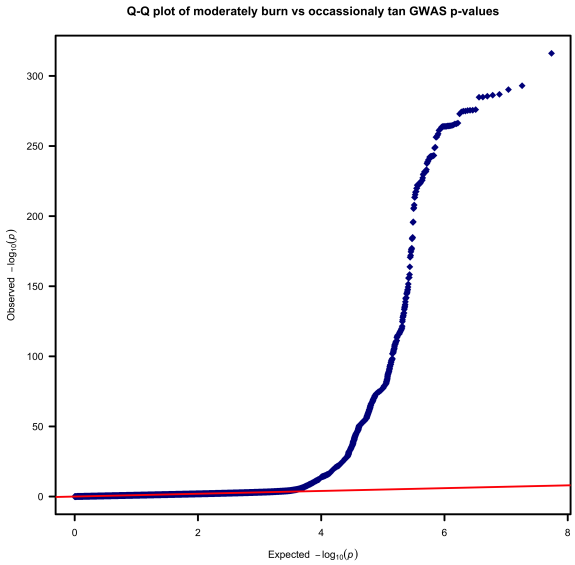


(d)

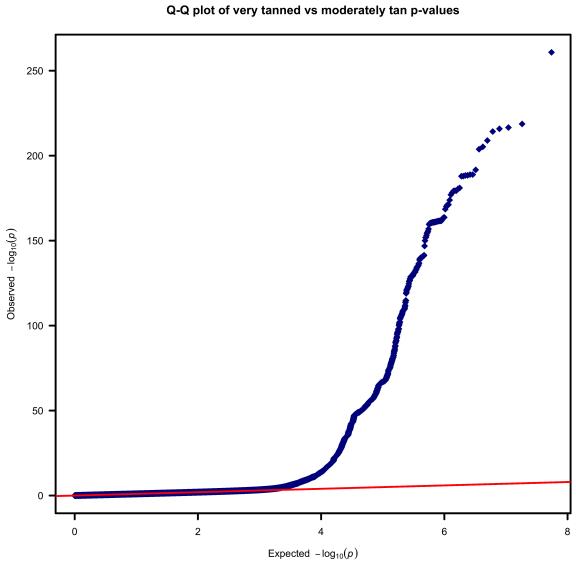
Figure 4.14: Q-Q plots for four binary skin colour GWAS conducted: (a) very fair/fair of 157,687 individuals (coded very fair 1, fair 2); (b) fair/light olive of 163,247individuals (coded fair 1, light olive 2); (c) light olive/dark olive of 40,653 individuals (coded light olive 1, dark olive 2); (d) dark olive/brown of 8,673 individuals (coded dark olive 1, brown2)



(a)



(b)



(c)

Figure 4.15: Q-Q plots for three binary tanning ability GWAS conducted: (a) Never tan, only-burn/occasionally tan, mostly burn of 102,630 individuals (coded never tan, only burn: 1, and occasionally tan, mostly burn 2); (b) occasionally tan, mostly burn/moderately tan, occasionally burn of 112,121 individuals (coded occasionally tan, most burn: 1, moderately tan, occasionally burn:2); (c) moderately tan, occasionally burn/only tan, never burn of 152,480 individuals (coded moderately tan, occasionally burn: 1, only tan, never burn:2

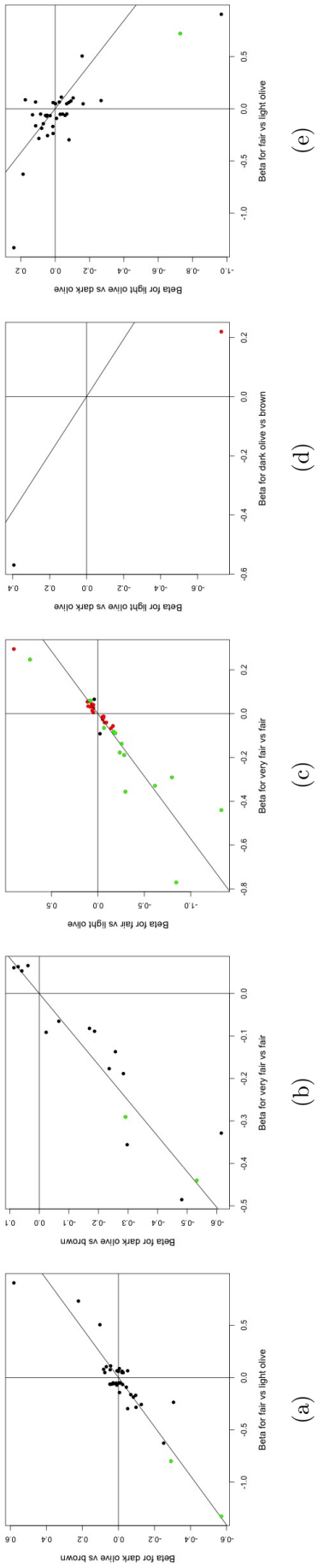
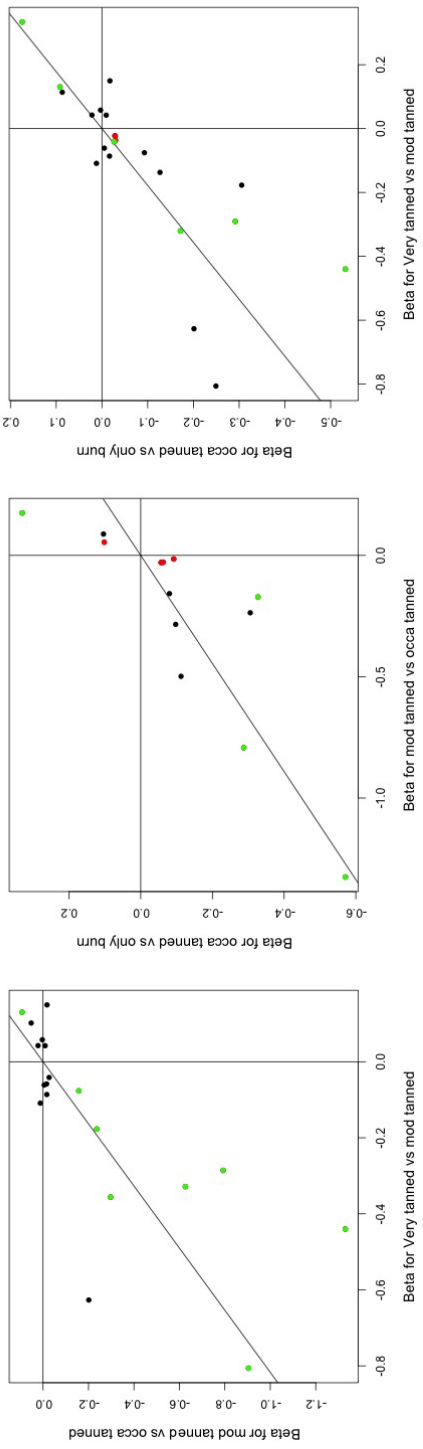


Figure 4.16: Comparison of estimated effect sizes for pairwise GWAS of adjacent categories of skin colour. Beta coefficients (log odds) of Lead SNPs for each GWAS are plotted for: (a) fair vs light olive plotted against dark olive vs brown; (b) very fair vs fair plotted against dark olive vs brown; (c) very fair vs fair plotted against fair vs light olive; (d) dark olive vs brown plotted against light olive vs dark olive; (e) fair vs light olive plotted against light olive vs dark olive; (f) very fair vs fair plotted against light olive vs dark olive. Green points represent Lead SNPs genome-wide significant ($p < 5 \times 10^{-8}$) for both traits; red points represent Lead SNPs genome-wide significant ($p < 5 \times 10^{-8}$) for the trait on the x-axis; and black points represent Lead SNPs genome-wide significant ($p < 5 \times 10^{-8}$) for the trait on the y-axis.



(a) (b) (c)

Figure 4.17: Comparison of estimated effect sizes for pairwise GWAS of adjacent categories of ease of tanning. Beta coefficients (log odds) of Lead SNPs for each GWAS are plotted for: (a) very tanned vs moderately tanned plotted against moderately tanned vs occasionally tanned; (b) moderately tanned vs occasionally tanned plotted against occasionally tanned vs only burn; (c) very tanned vs moderately tanned plotted against occasionally tanned vs only burn. Green points represent Lead SNPs genome-wide significant ($p < 5 \times 10^{-8}$) for both traits; red points represent Lead SNPs genome-wide significant ($p < 5 \times 10^{-8}$) for the trait on the x-axis; and black points represent Lead SNPs genome-wide significant ($p < 5 \times 10^{-8}$) for the trait on the y-axis.

4.4 Discussion

This work has identified distinct differences between the genetic architectures of red and non-red hair phenotypes. It has been shown that combining red and non-red hair in the same ordinal scale, under the assumption they share identical underlying genetics, may lead to the reporting of misleading associations and missing some evidence of association. Analyses concluded 11 loci were jointly associated with red and non-red hair, 96 were uniquely associated with non-red hair, and 3 were uniquely associated with red hair. Bayesian analysis supported this finding by identifying uniquely associated LD-based regions in non-red hair compared to red hair. Comparing a hair colour GWAS that included red hair to a non-red hair GWAS highlighted discrepancies between associated regions supporting the hypothesis that type I and type II errors may be present when including red hair into an ordinal scale of hair colour.

Additionally, this work has highlighted the genetic similarities between non-red hair phenotypes, providing supportive evidence for combining these phenotypes into a single ordinal scale, under the assumption that the underlying genetics are uniform. Furthermore, this work has highlighted the uniformity of the genetics underlying the ordinally coded variables for tanning ability and skin colour, supporting combining skin colour and tanning ability into single ordinal scales. Analysis concluded substantial overlap in regions of association between non-red hair phenotypes, tanning ability phenotypes and skin colour phenotypes as lead SNPs representing each region identify as associated across all phenotypes whilst having the same direction of effect.

Throughout this work there has been the potential for population structure to bias results, as hair colour follows a unique geographical distribution. Where possible, effort was taken to reduce this bias in analyses by including the first ten principal components into GWAS models, and PRS trait prediction models.

Throughout analyses loci of association were delimited by physical distance (using the peak sorting algorithm). This combination and annotation of associated genetic signals based on physical distance may underestimate the number of genetic regions involved in defining hair colour. For instance, within *MC1R* there are multiple functional coding variants strongly associated with

the red hair phenotype, known to be in negative LD and so independently associated with the hair colour phenotype. However, across my analyses, because a purely distance based approach has been adopted, only a single SNP has been selected to represent the *MC1R* region.

In addition to underestimating associated regions by combining signals into distance based loci, epistasis may have underlying effects across associated regions. Previous studies (Duffy, Box, et al. 2004; Hysi et al. 2018) have established that *MC1R* causal coding variants have differing penetrance with regards to red hair ("R" variants having high penetrance, and "r" having low penetrance). However, the epistatic effects these variants may be having across other associated regions are not well defined. In general, epistatic effects are believed to be much smaller than main effects; typically in the case of polygenic traits (such as hair colour) these main effects are already small. However subtle, it would be advantageous to further investigate these effects as it would lead to a clearer understanding of the associated regions and their relationship to one another.

Previous studies into hair colour have highlighted the importance of hair shape and texture in determining hair colour; and that it is particularly relevant in the case of blond hair compared to other hair colours (Morgan et al. 2018). By conducting a similar methodology used here for red hair on blond hair, further genes uniquely associated with blond hair that are involved in hair shape and texture may be identified. These new genes could provide a much needed insight into the underlying biology as to why shape and texture seems specifically important for blond hair determination compared to non-blond hair colour. Current literature suggests shape and texture may change the refractive and reflective properties of individual hairs which could influence the overall perceived colour (Bustard and R. W. Smith 1991). Further to this, an inverse correlation is observed between light to dark hair colour and hair thickness with lighter hair being thinner than darker hair (Vaughn et al. 2009). As blond hair is thinner and has less melanin compared to darker hair colours it may be more susceptible to changes to the overall perceived colour if shape and texture differs.

In conclusion, these analyses have improved our understanding of the underlying genetics between hair colour phenotypes. In particular, it has highlighted the distinct genetic architecture

of red hair compared to non-red hair phenotypes and provided evidence against a single ordinal hair colour scale that includes red hair. Additionally, it has furthered the case of a single ordinal non-red hair colour scale based upon colour darkness by demonstrating a large overlap between non-red hair genetics. Finally, analyses have offered and insight into the problem of misreporting of genetic associations for overall hair colour due to the inclusion of red hair into a single scale by detecting Type I and Type II errors in overall hair colour GWAS where red hair was included into the phenotypic scale.

Chapter 5

The genetic architecture of pigmentation measures in UK Biobank

5.1 Introduction

Current pigmentation GWAS have highlighted over 200 associated loci for hair colour, six loci for skin colour, and 20 associated loci for tanning ability. Given the large sample size and multiple pigmentation traits measured on the same cohort, utilising UK Biobank could further our knowledge in understanding the genetic architecture of these pigmentation traits and how the genetic architectures of UK Biobank pigmentation traits overlap between one another.

5.1.1 Originality of research

Although there is an understanding of the individual underlying genetics of pigmentary measurements, research into the overlap in genetic architecture of these pigmentary traits is limited. The identical biosynthetic pathway for the production of eumelanin and pheomelanin for each of the external tissues on the body (epidermis, hair bulb, and iris) is well-established (S. Ito and K. Wakamatsu 2011; Cassady and Sturm 1994; Bellono and Oancea 2014; Hysi et al. 2018), and is supported by the observation of strong correlation between pigmentary measures (self-reported or otherwise) (Cust, Pickles, et al. 2015); but the full extent of overlap between the underlying genetics of these pigmentary measurements is still relatively unknown. More specifically, multiple studies into the genetics of pigmentation have highlighted the importance of six major loci of association (*TYR, SLC45A2, IRF4, HERC2/OCA2, MC1R*, and *ASIP*), with these six con-

sistently appearing as genome-wide significant in GWAS of hair colour, skin colour, and tanning ability (Crawford et al. 2017; F. Liu et al. 2015; Morgan et al. 2018). Although these loci are genome-wide significant across pigmentary traits, their similarity of effect sizes is still unclear. Furthermore, recent hair colour GWAS have identified large numbers of novel hair colour loci that appear to have subtle effects on determining hair colour, it is still unknown whether these are associated with other pigmentary traits, and if their effect sizes are consistent across them. With the availability of large databanks contain pigmentary information, conducting GWAS of each pigmentary trait would yield further novel associations for pigmentary traits, and identify the role the subtle hair colour loci have on skin colour and ease of tanning.

5.1.2 Chapter Aims & Objectives

Aims

I. Determine the genetic architectures of the UK Biobank pigmentation traits.

II. Determine the genetic architectures of the overall pigmentation scores creating from a PCA of the UK Biobank pigmentation traits.

III. Investigate the distinct characteristics between the pigmentary GWAS results.

IV. Investigate the distinct characteristics between the overall pigmentation GWAS results.

V. Determine how well these pigmentary GWAS correlate with each other and melanoma

Objectives

1. Conduct GWAS on each pigmentary trait in the UK Biobank.
2. Conduct GWAS on each overall phenotypic pigmentation score.
3. Investigate the differences in regions of association for each pigmentation GWAS
4. Investigate the differences in regions of association for each overall pigmentation score GWAS.
5. LD-score regression to estimate correlation between pigmentation traits and melanoma.

5.1.3 Rationale to achieve aims & objectives

Here we report the analysis of the 350,000 UK Biobank participants. By performing eight genome-wide association studies of pigmentary measures we were able to identify 109 novel pigmentary loci and assess their effects across each pigmentary measure. I was able to further provide evidence of the importance of the pigmentation pathway on melanoma by identifying a large overlap between know melanoma loci and pigmentation loci identified through the GWAS conducted here.

5.2 Methods

All analysis was conducted in PLINK 2.0 (Purcell et al. 2007), GCTA (J. Yang et al. 2011), ldsc (Brendan K. Bulik-Sullivan et al. 2015), and R (R Core Team 2017) using the UK Biobank (Sudlow et al. 2015) dataset.

Each pigmentation variable in the UK Biobank was considered as an ordinal scale. For non-red hair colour: 'blonde' = 1, 'light brown' = 2, 'dark brown' = 3, 'black' = 4; for red hair: 'non-red hair' = 1, 'red hair' = 2; for skin colour: 'very fair' = 1, 'fair' = 2, 'light olive' = 3; 'dark olive' = 4, 'brown' = 5; 'black' = 6; For tanning ability: 'get very tanned' = 1, 'get moderately tanned' = 2. 'get mildly or occasionally tanned' = 3, 'never tan only burn' = 4.

5.2.1 Creating overall pigmentation scores derived PCA analysis

PCA was conducted to produce three overall pigmentation scores for the 316,518 individuals in the UK Biobank that passed standard QC for the genetic information and had pigmentary information for hair colour, ease of tanning and skin colour. The analysis was conducted on the UK Biobank pigmentary variables excluding number of childhood sunburn incidents to retain sample power and minimise reporting bias. Correlation-based PCA was selected (over calculating a covariance matrix) for analysis as the pigmentary variables were on different scales; and due to the ordinal nature of the pigmentary variables, polychloric correlations were calculated between variables in preference to linear correlations methods (Pearson and Spearman). Component number was determined by setting a variation capture threshold of greater than 0.9 across tanning ability, skin colour, red hair and non-red hair colour. An individual's score for each component was then derived using the pigmentary traits coefficients (as weightings) from each

component model. These scores were then used as phenotypic variables in linear regression-based GWAS analyses.

5.2.2 Conducting Pigmentary GWAS

Five pigmentary GWAS (non-red hair colour, red hair colour, ease of tanning, skin colour, and number of childhood incidence of sunburn) were conducted on UK Biobank cohort individuals. In addition to these five pigmentary GWAS, three overall pigmentation score (pigscore 1, pigscore 2, and pigscore 3) GWAS were conducted which were derived from principal components analysis (PCA) of four UK Biobank pigmentary traits (5.2.1). To minimise population stratification, the analyses were conducted on individuals of European descent and adjusted for the first ten genetic principal components, as well as for the array individuals were genotyped on.

Due to the pigmentary variables in the UK Biobank being self-reported, individuals who responded: 'don't know', 'prefer not to say', or did not answer for the respective pigmentary GWAS were dropped from analysis. After removing these participants and conducting standard QC (see 2.3.4) for genetic data, $370,260$ individuals were retained: for the skin colour GWAS, $367,229$ for the ease of tanning GWAS, $352,662$ for the non-red hair colour GWAS, $280,811$ for the number of childhood sunburn incidents GWAS, $103,111$ for the red hair colour GWAS with $18,373$ red haired cases, and $316,518$ for three overall pigmentation GWAS. Pigmentary traits were considered to either follow an ordinal scale or were binary, and thus ordered by light/dark pigmentation with assigned increasing numerical values (Table 5.1) (hair: 'blond' to 'black'; skin: 'very fair' to 'black'; ease of tanning: 'get very tanned' to 'never tan only burn'; red: 'non-red' or 'red'; childhood sunburn incidents: 'no incidents' or at 'least one incident'). These pigmentary ordinal scales were then used as phenotypic variables in linear regression-based GWAS analyses, and binary in logistic regressed-based GWAS analyses.

5.2.3 Genomic inflation

Genomic inflation was investigated by producing a QQ plot and calculating the genomic inflation factor for each of the eight GWAS results by calculating the ratio (λ) of the median of the observed GWAS test statistic (where p-values are used to calculate a Z test statistic and then converted into χ^2 values) to the expected median (χ^2 distribution with one degree of freedom). Comparing this ratio (λ) quantifies the extent of inflation and false-positives within results

Pigmentary Trait	**Sample Size**	**Ordinal scale/Score coefficients**
Skin colour	370,260	Very fair 1; fair 2; light olive 3; dark olive 4; brown 5; black 6
Ease of tanning	367,229	Get very tanned 1; Get moderately tanned 2; Get mildly or occasionally tanned 3; Never tan, only burn 4
Non-red hair colour	352,662	Blond 1; light brown 2; dark brown 3; black 4
Red hair	Red: 18,373 Non-red: 84,772	Non-red 1; red 2
Number of childhood sunburn incidents	at least one incident: 55,202 no incident: 225,104	No childhood sunburn incidents 1; At least one incident 2
Overall pigmentation score 1	316,518	**Score coefficients:** 0.78 skin; −0.69 ease; 0.87 non-red hair; −0.95 red hair
Overall pigmentation score 2	316,518	**Score coefficients:** −0.40 skin; 0.61 ease; 0.50 non-red hair; −0.31 red hair
Overall pigmentation score 3	316,518	**Score coefficients:** 0.49 skin; 0.39 ease; −0.02 non-red hair; 0.10 red hair

Table 5.1: Sample sizes and ordinal scales/score coefficients used in the eight pigmentary GWAS on UK Biobank participants that passed standard QC and are of European descent. pigscore 1-3 were created using principal components analysis on skin colour, ease of tanning, non-red hair colour, and red hair pigmentary variables with the aim to create overall pigmentation scores.

which can be caused by population stratification, sample duplications, large population samples, polygenicity, unknown familial relationships, or inappropriate GWAS modelling.

5.2.4 Genetic correlation

Additionally, to investigate the polygenicity and genetic correlation between the UK Biobank pigmentary traits, and between the UK Biobank pigmentary traits and melanoma, LD score regression was conducted on the summary statistics of the eight GWAS. As polygenicity and confounding bias (such as population stratification, or unknown familial relationships) both produce genomic inflation in GWAS, it is important to investigate the nature of this inflation to understand the validity of results. LD score regression estimates the individual contribution of polygenicity and confounding bias (such as population structure) in GWAS summary statistics by examining the relationship between test statistics and LD. As variants in LD with a causal variant will show an increase in their test statistic proportional to the the LD with the causal variant; and as inflation from confounding biases (between populations or studies) will not correlate with LD, it is possible to quantify the likely cause of inflation. Similarly, the genetic correlation between two traits can be estimated using a simple extension of LD score regression where two traits are incorporated rather than a single trait.

To estimate a trait's polygenicity and the genetic correlation between pigmentary traits and between pigmentary traits and melanoma, GWAS summary statistics (from the UK Biobank pigmentation GWAS and melanoma meta-analysis GWAS) were prepped for analysis by removing rare variants (MAF < 0.01), and omitting SNPs with missing information. The following information were extracted for each SNP from all pigmentary GWAS summary statistics: (1) SNP ID; (2) chromosome; (3) base position; (4) effect allele; (5) reference allele; (6) GWAS n; (7) beta coefficient/odds ratio; (8) SE; (9) p-value; (10) effect allele frequency. To ensure the SNP's effect direction was consistent between GWAS, the effect and reference allele was compared across all results; if a mis-matched allele was detected, the effect and reference allele was switched and beta coefficient sign or odds ratio was inverted. LD score regression analysis (using the ldsc package) was conducted on each prepped GWAS summary statistics to determine the polygenicity/confounding bias inflation and on pairwise GWAS summary statistics to determine the genetic correlation. Strong correlation was defined by an rg^2 threshold of greater than 0.6 and weak correlation by an rg^2 threshold of less than 0.2.

5.2.5 Identifying signals and signal overlap

Independent associated signals

The program GCTA was used to conduct conditional analysis to identify independent associated signals from each pigmentary GWAS with a p-value significance threshold of $P < 5 \times 10^{-8}$ and collinearity threshold of $r^2 > 0.05$. Functional loci were defined by grouping genome-wide significant ($p < 5 \times 10^{-8}$) independent associated signals that were within a 500kb window of each other.

Recurring signals across pigmentary traits

To identify recurring signals across the eight pigmentary GWAS results, All previously defined independent associated signals were combined from each GWAS analysis into an overlap dataframe (ODF), and omitted any duplicated SNPs. Using PLINK 2.0's '–r2' command and the UK Biobank genetic dataset, an r^2 value was calculated between each pair of SNPs within the ODF. A collinearity sorting algorithm (implemented in R) was then used to consider each SNP within the ODF and cluster signals that had a collinearity threshold r^2 value below 0.05. If any two SNPs did not have an r^2 below the collinearity threshold, but did pair with a common SNP, they were classified as the same signal. Once recurring signals had been identified across the pigmentary GWAS, functional loci were defined by grouping signals that were within a 500kb window of one another. The collinearity sorting algorithm R script can be found in Appendix A.2.

5.2.6 Comparing between overall pigmentation scores: beta plots

As the components in a PCA are orthogonal to one other (2.2.1), the overall pigmentation GWAS that used scores derived from PCA weightings (5.2.1) would also be expected to be orthogonal and highlight distinct underlying pigmentary effects. This was investigated for each pigmentation score (one to three) GWAS summary statistics by selecting the genome-wide significant independent signals (5.2.5) and matching them to the other two pigmentation score GWAS results. Once matched, the beta coefficients (log odds) were plotted and correlations between them calculated. This process was repeated for each overall pigmentation score GWAS.

5.2.7 Comparing between pigmentary traits: beta plots of signals and of loci

To investigate the genetic architecture across UK Biobank hair colour, skin colour, and tanning ability traits and to highlight discrepancies of signals effect across pigmentary trait, independent associated signals beta coefficients (log odds), from the GWAS of non-red hair colour, of red hair, of ease of tanning, and of skin colour, were plotted against each trait.

To compare the effect sizes of the independently associated signals derived from the pigmentary GWAS results, the following information was extracted for every SNP in the ODF: (1) SNP ID; (2) base position; (3) MAF; (4) Overlap locus number; (5) Overlap signal number; (6) GWAS of hair colour p-value; (7) GWAS of hair colour beta coefficient; (8) GWAS of hair colour SE; (9) GWAS of skin colour p-value; (10) GWAS of skin colour beta coefficient; (11) GWAS of skin colour SE; (12) GWAS of ease of tanning p-value; (13) GWAS of ease of tanning beta coefficient; (14) GWAS of ease of tanning SE.

Beta coefficients (log odds) for each pairwise combination of pigmentary traits were plotted by compiling lead SNPs from independent signal associated with either considered traits into a data frame. If a signal was jointly associated with both traits a signal window was defined (using all the SNPs within the signal as a size guide whilst ensuring there was no overlap with any other signal) and then ordered by each trait's GWAS p-value and ranked. A mean rank for each trait's p-value was calculated. SNP's were ordered by mean rank and the lowest mean rank was used as the 'fairest most representative' SNP for plotting. For each SNP, the SE found in the corresponding pigmentary GWAS was used as error bars on the plot.

Prior associations and identifying signals of interest

Prior pigmentation and melanoma associations were identified for pigmentation signals by merging the ODF with GWAS-catalog (Buniello et al. 2019) data and known pigmentation and melanoma GWAS not present in GWAS-catelog databases at time of analysis.

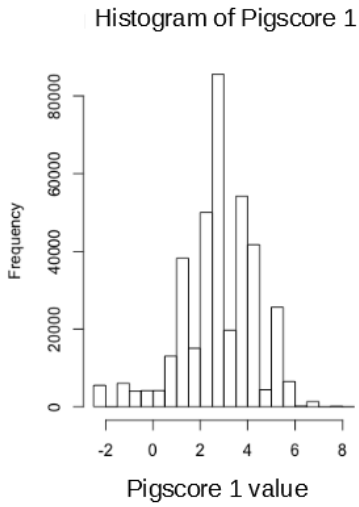
Signals of interest were identified through beta plots (5.2.7) and the ODF table by assessing signal significance and effect size across all traits with the aim to identifying SNPs that are uniquely genome-wide significant with a large effect for a single pigmentary trait. Additionally,

when highlighting signals of interest, the overall relationship between pigmentary trait's signals were taken into account. By considering these relationships of effect sizes between traits, it was possible to identify discrepancies between the overall effects these signals may have on pigmentary risk and highlight signals with distinct effects across the traits.

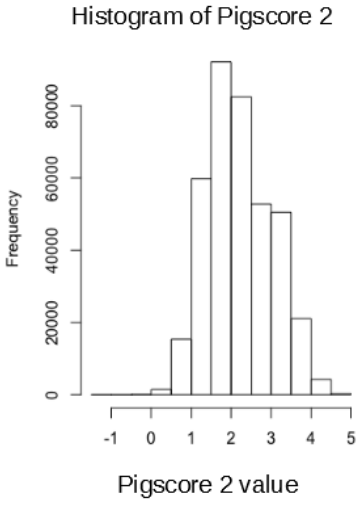
5.3 Results

5.3.1 Principal components analysis

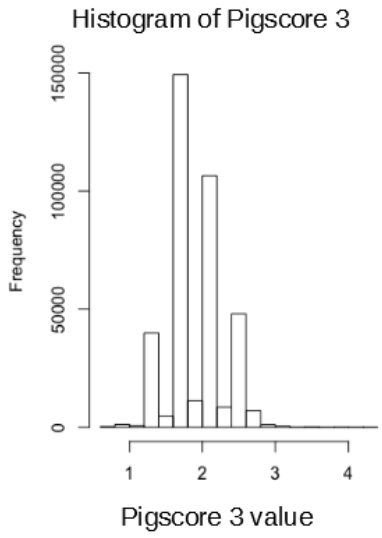
For the 316,518 UK Biobank participants that had all the required pigmentary information, principal component analysis found 99% of pigmentation variance was accounted for by three PC. The first PC had high weightings for all pigmentation variables which followed the same pigmentation effect direction (light to dark pigmentation) and accounted for 60% of variation between pigmentation measures. The second PC had more modest weighted coefficients for the pigmentation variables with skin colour and tanning ability pigmentation effect directions switched compared to PC1 (dark to light). The third PC observed effects weighted towards skin-based variables (tanning ability and skin colour) with tanning ability (light to dark) and skin colour (dark to light) pigmentation effects direction opposite to one another (Table 5.1). When plotting histograms of each participant's three pigmentation scores, none of the three score distributions were skewed positively or negatively. The first pigmentation had the widest score range (-2 to 8) and the third score had the narrowest (1 to 4). No outliers were identified through the histograms (Figure 5.1).



(a)



(b)



(c)

Figure 5.1: Histogram plots for the frequency of score values for overall phenotypic pigmentation scores: (a) Pigscore 1; (b) Pigscore 2; (c) Pigscore 3

5.3.2 Pigmentation GWAS

Genomic inflation factors from the five pigmentary GWAS were 1.086 for ease of tanning, 1.047 for skin colour, 1.015 for childhood sunburn, 1.104 for hair colour (excluding red hair), and 0.911 for red hair (Figure 5.2). The three combined pigmentation scores had genomic inflation factors of 1.065 for pigscore 1, 1.093 for pigscore 2, and 1.040 for pigscore 3 (Figure 5.3). This level of inflation is expected in high power large sample cohorts that have the ability to detect large polygenic effects (Jian Yang et al. 2011). Unsurprisingly, the genetic correlation between the pigmentary traits, estimated by LD score regression was uniformly high; all had a genetic correlation > 0.37 with skin colour, red hair/childhood sunburn having the highest genetic correlation (0.98) and non-red hair colour being the least correlated with other pigmentary traits (< 0.68) (Table 5.2).

	Red	Ease	Skin	Childhood sunburns
Hair	0.37	0.50	0.68	0.34
Red	-	0.51	0.56	0.98
Ease	-	-	0.88	0.52
Skin	-	-	-	0.57

Table 5.2: Genetic correlation for five pigmentary GWAS: Tanning ability (ease); Skin colour (Skin); Non-red hair colour (Hair); Red hair colour (Red); Number of childhood sunburn incidents (Childhood sunburns)

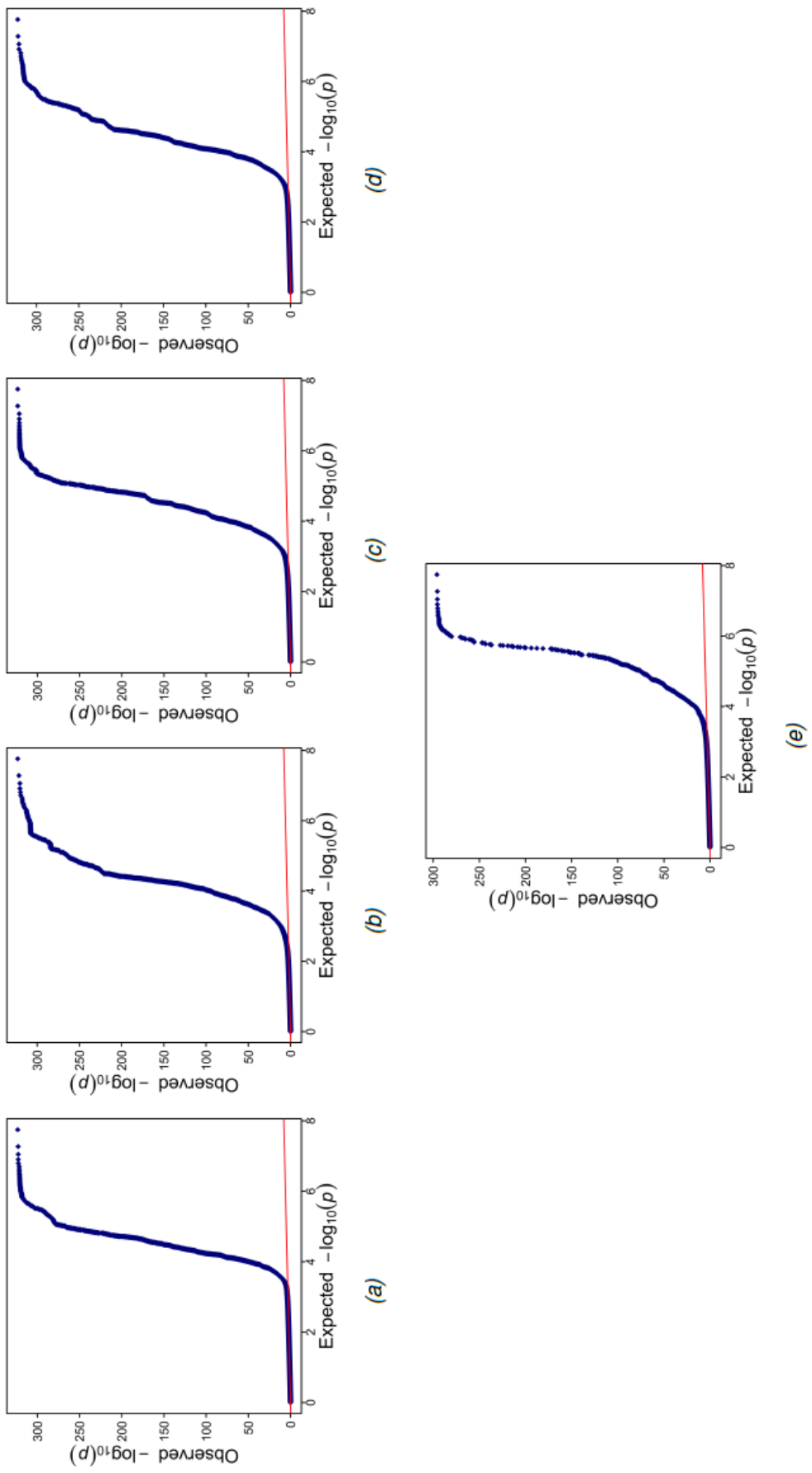


Figure 5.2: QQ plots for the five pigmentary GWAS conducted: (a) Red vs non-red hair colour; (b) Hair colour (excluding red hair); (c) Skin colour; (d) Ease of tanning; and (e) Number of childhood sunburns

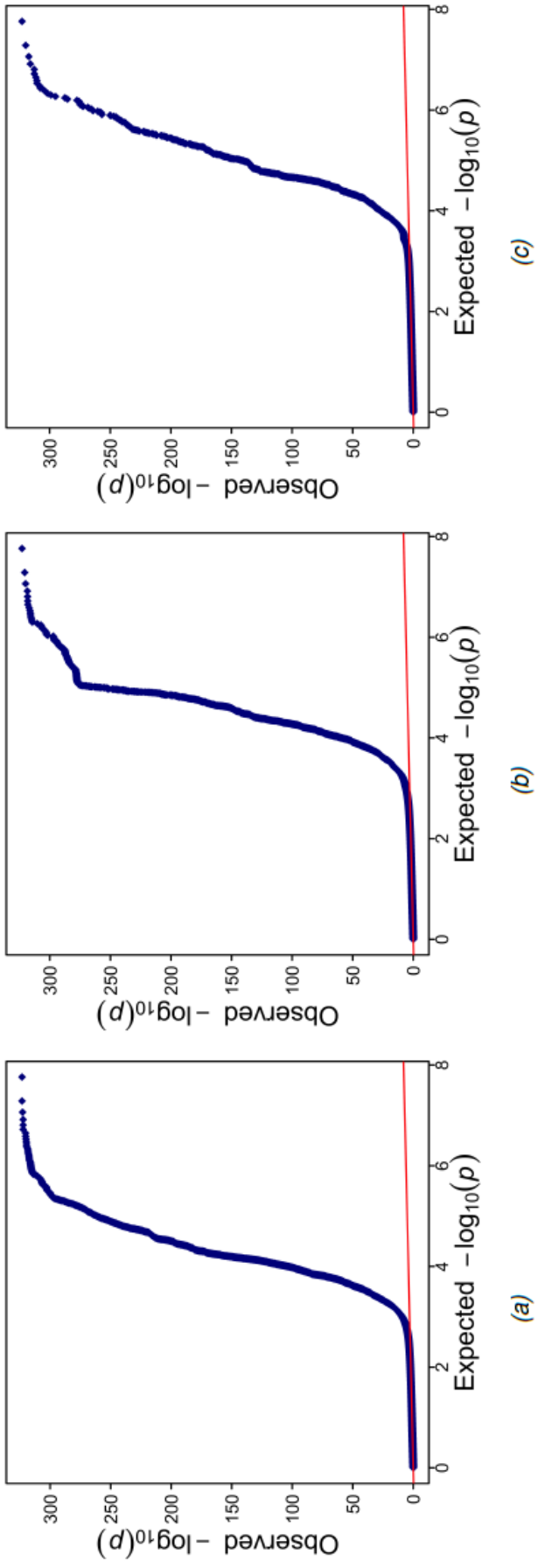


Figure 5.3: QQ plots for the three combined pigmentary trait GWAS conducted: (a) Pigscore 1; (b) Pigscore 2; and (c) Pigscore 3

Genome-wide significant association with pigmentation ($p < 5 \times 10^{-8}$) was detected for tanning ability at 119 independent genetic variants across 80 loci (Appendix C.3), for skin colour at 123 independent variants across 96 loci (Appendix C.4), for number of incidents of childhood sunburn at 50 independent genetic variants across 29 loci (Appendix C.5), for non-red hair colour at 323 independent variants across 241 loci (Appendix C.1), and for red/non-red hair at 25 independent genetic variants across 14 loci (Appendix C.2). Similarly, genome-wide significant association with PC-derived combined pigmentation scores was detected for pigscore 1 at 181 independent genetic variants from 135 loci (Appendix D.1), pigscore 2 at 146 independent genetic variants from 119 loci (Appendix D.2), and pigscore 3 at 29 independent genetic variants from 20 loci (Appendix D.3) (Figure 5.4).

5.3.3 Comparing pigmentation score GWAS

When comparing the effect sizes of the identified independent signals from the pigscore 1 GWAS with their corresponding effect sizes from the pigscore 2 and pigscore 3 GWAS, mixed correlation was observed ranging between -0.61 (with pigscore 2) and -0.34 (with pigscore 3). When comparing the effect sizes of identified independent signals from the pigscore 2 GWAS with the corresponding effect sizes from the pigscore 1 and pigscore 3 GWAS, correlation was strong with pigscore 2 and pigscore 1 having a correlation of -0.66 and pigscore 2 and pigscore 3 having a correlation of 0.65. Comparing the effect sizes of the identified independent signals from the pigscore 3 GWAS with their corresponding effect sizes from the pigscore 2 and pigscore 1 GWAS, mixed correlation was observed with strong correlation identified between the effect sizes for pigscore 3 and pigscore 2 (0.89), and weak correlation identified between the effect sizes for pigscore 3 and pigscore 1 (-0.37).

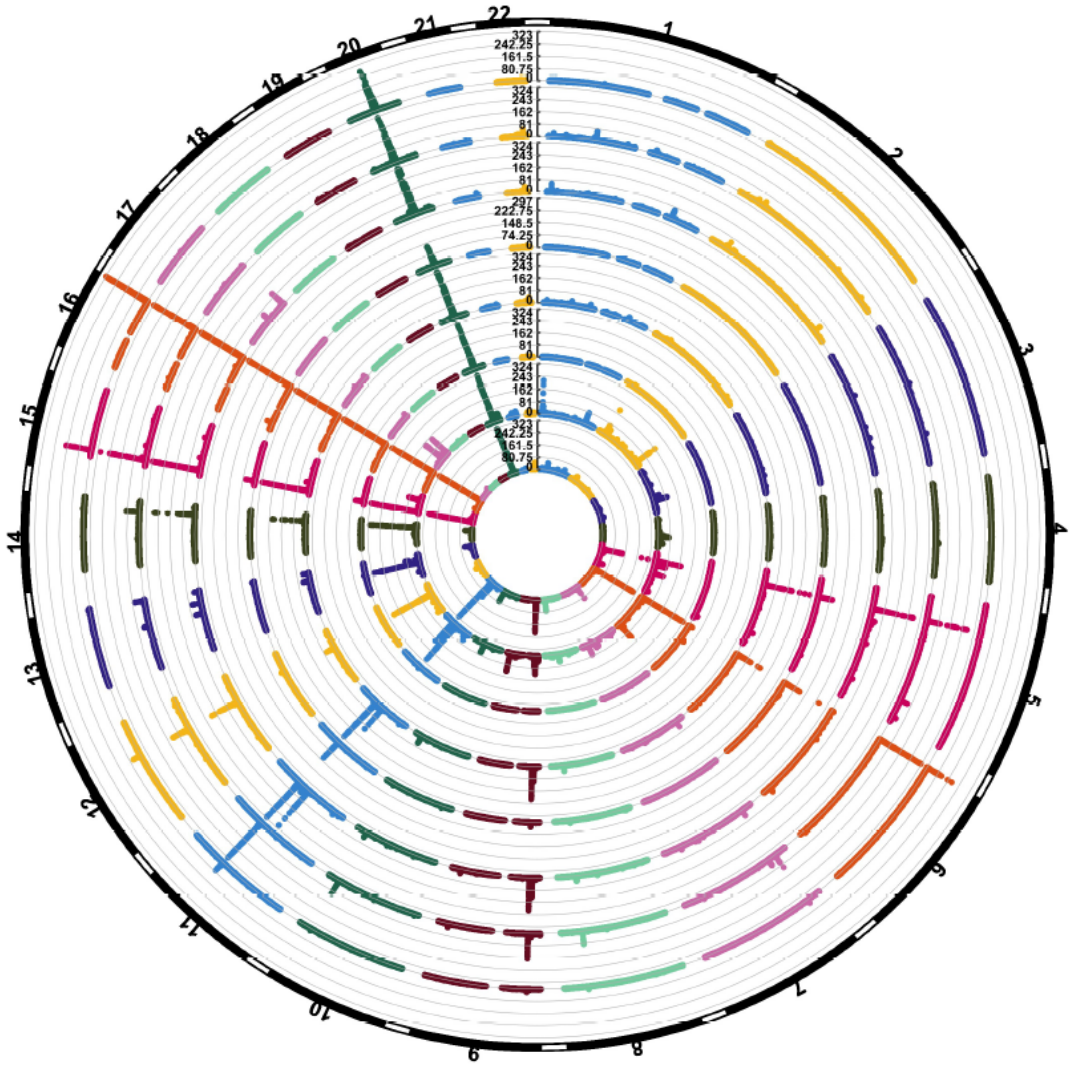


Figure 5.4: Circular Manhattan plot for eight pigmentary GWAS from UK Biobank. From the centre out: pigscore 1 (316,518 samples, quantitative), pigscore 2 (316,518 samples, quantitative), pigscore 3 (316,518 samples, quantitative), ease of tanning (367,229 samples, ordinal), skin colour (370,260 samples, ordinal), number of childhood sun-burns (280,811 samples, binary), non-red hair colour (352,662 samples, ordinal), Red hair colour (18,373 red hair, 84,738 non-red hair, binary).

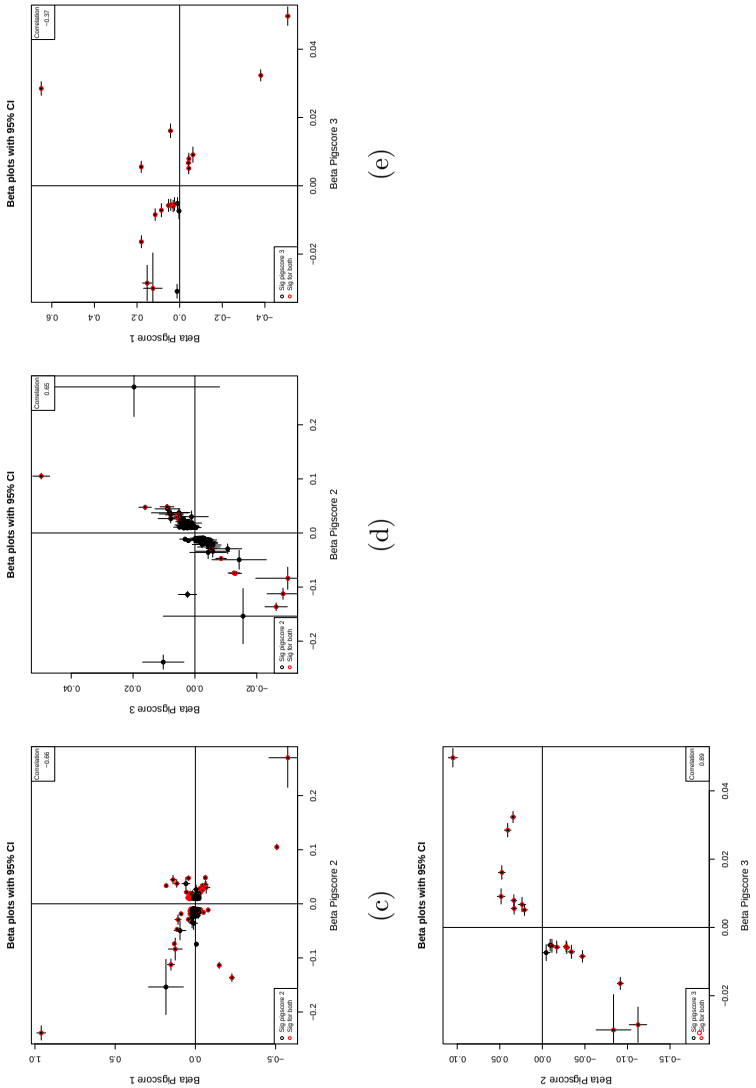


Figure 5.5: Lead SNP beta coefficients (log odds) for independent associated loci of pigmentary trait GWAS. (a) Beta coefficients of pigscore 1 vs pigscore 2; (b) Beta coefficients of pigscore 1 vs pigscore 3; (c) Beta coefficients of pigscore 2 vs pigscore 1; (d) Beta coefficients of pigscore 2 vs pigscore 3 (excluding red); (e) Beta coefficients of pigscore 3 vs pigscore 1; (f) Beta coefficients of pigscore 3 vs pigscore 2. Independent signals from each GWAS were defined through conditional analysis with a p-value significance threshold of $P < 5 \times 10^{-8}$ and a collinearity threshold of $r^2 > 0.05$.

5.3.4 Identifying signals of interest

When plotting the beta coefficients (log odds ratios) of the lead signal SNPs of identified loci through conditional analysis (see 5.2.5), high correlation across the pigmentation traits were observed indicating high overlap in the genetic architecture and effects between the measured pigmentation traits (skin colour, tanning ability, non-red hair, and red hair). The highest correlated beta coefficients were between skin colour and tanning ability ($\rho = -0.93$) and the least correlated beta coefficients between red hair and non-red hair ($\rho = 0.3$) (Table 5.3).

	Red	Ease	Skin
Hair	0.30	-0.61	0.81
Red	-	0.69	-0.57
Ease	-	-	-0.94

Table 5.3: Beta coefficient correlation for the four pigmentary GWAS: Tanning ability (ease); Skin colour (Skin); Non-red hair colour (Hair); Red hair colour (Red)

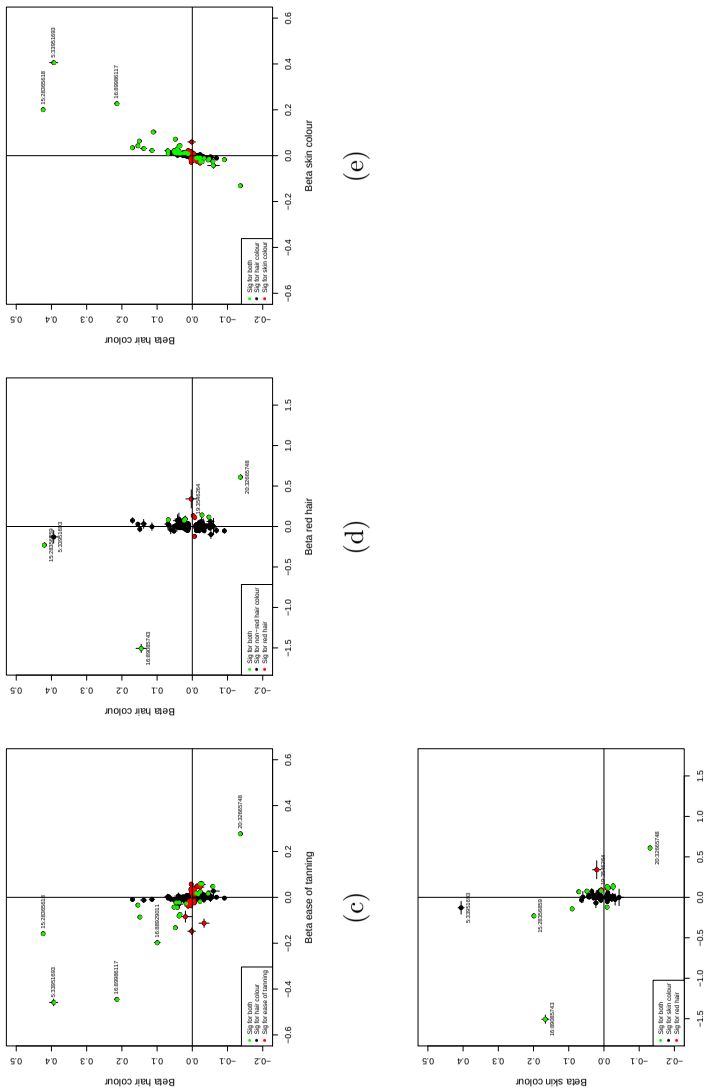


Figure 5.6: Comparison of estimated effect sizes for different pigmentary traits. Beta coefficients (log odds) for Lead SNPs at independently associated loci are plotted for those variants reaching genome-wide significance for either of each pair of traits. Green points represent Lead SNPs genome-wide significant ($P < 5 \times 10^{-8}$) for both traits; red points represent Lead SNPs genome-wide significant ($P < 5 \times 10^{-8}$) for the trait on the x-axis; and black points represent Lead SNPs genome-wide significant ($P < 5 \times 10^{-8}$) for the trait on the y-axis. (a) Beta coefficients of tanning ability vs skin colour; (b) Beta coefficients of tanning ability vs red hair colour; (c) Beta coefficients of tanning ability vs non-red hair colour; (d) Beta coefficients of red hair vs non-red hair colour; (e) Beta coefficients of skin colour vs non-red hair colour; (f) Beta coefficients of red hair vs skin colour (excluding red). Lead SNPs from each GWAS were defined through conditional analysis with a p-value significance threshold of $P < 5 \times 10^{-8}$ and collinearity threshold of $r^2 = 0.05$.

Many polymorphisms significantly associated ($P < 5 \times 10^{-8}$) here with at least one of the pigmentation traits have previously been reported as associated with pigmentation in the literature and are recorded in recent GWAS studies (Morgan et al. 2018; Hysi et al. 2018; F. Liu et al. 2015; Visconti et al. 2018) and in the GWAS-catalog (Buniello et al. 2019). 500 independent variants ($r^2 < 0.05$) across 322 loci were found here to be genome-wide significantly associated with at least one pigmentation measurement or combined pigmentation score; of these loci, 109 had not previously been reported as associated with any pigmentary traits (Appendix E.1). Across all loci found here to be associated with any pigmentary trait, 8 were found to be associated only with number of childhood sunburn incidents with 6 of these novel, 12 only with ease of tanning with 4 novel, 179 only with non-red hair colour with 44 of these novel, 1 only with red hair and also novel, and 11 only with skin colour with 6 novel (Figure 5.7) (Appendix E.1). The PC-derived combined pigmentation traits identified novel loci: pigscore 1 identified 5 uniquely associated loci, of which 3 were novel, pigscore 2 identified 17 uniquely associated loci, of which 11 were novel, and pigscore 3 identified 1 uniquely associated loci which was novel. Additionally, of the 322 loci found to be associated with any pigmentation trait, 39 of them have previously been reported as associated with melanoma risk (Landi et al. 2020).

Previous studies have consistently reported six loci to be strongly associated with these pigmentary traits, these strong associations ($p< 5 \times 10^{-100}$) were also found in these pigmentation GWAS results: *SLC45A2* (rs16891982), *SLC24A4* (rs941799), *MC1R* (rs1805007), *HERC2* (rs12913832), *TYR* (rs-), and *IRF4* (rs12203592), which were associated with all eight pigmentary GWAS (Appendix C.1, C.2, C.3, C.4, C.5, D.1, D.2, D.3). Many identified signals here have also had prior-reported associations with other pathways not relating to the four pigmentation measures used or melanoma. Most commonly, further pigmentation pathways were repeatedly highlighted by identified loci across all eight pigmentation GWAS, with the freckling pathway appearing in all GWAS results: tanning ability: rs251468, and rs12203592; skin colour: rs12203592; red hair: rs12203592; non-red hair: rs17833789, and pigscore 1,2, and 3: rs1805007. As well as the eye colour pathway also being highlighted in each GWAS result: tanning ability: rs12203592, and rs12913832; skin colour: rs12913832, and rs16891982; red hair: rs12203592; hair colour: rs12913832, and rs12913832; and pigscore 1,2, and 3: rs12913832, and rs16891982. In addition to further pigmentation traits, evidence of a shared pathway between pigmentation and non-melanoma skin cancer was present as many identified loci had previously been reported for

non-melanoma skin cancer. In particular loci identified in the region of *MC1R* also had prior-associations between Basal cell carcinoma (rs1805007) and tanning ability, non-red hair colour, skin colour, and pigscore 1); and loci identified in the regions of *TYR* highlighted a shared risk pathway between Squamous cell carcinoma (rs1126809) and tanning ability, skin colour, pigscore 1, and non-red hair colour).

For the 109 novel loci identified through the eight pigmentation GWAS, with the highest amount (13) situated on chromosome 1, and lowest (1) amount situated on chromosome 14 (Appendix C.1, C.2, C.3, C.4, C.5, D.1, D.2, D.3). Highlighting coincident associations between these novel pigmentation loci and non-pigmentation phenotypes highlighted shared pathways between tanning ability (rs12203592) and baldness acting through the *IRF4* gene, on chromosome 2, which is important in the regulation of interferons in response to infection by virus; and between non-red hair colour (rs68088846) and baldness acting through the *RUNX1* gene, found on chromosome 21, which is thought to be involved in the development of normal hematopoiesis and to be associated with blood platelet disorders. Additionally, evidence of a shared pathway between skin-based pigmentation and age spots was also observed with skin colour (rs251468), on chromosome 5, in the *PPARGC1B* gene which may be involved in fat oxidation, non-oxidative glucose metabolism, and tanning ability (rs35563099) in the *RP11-355F22.1* gene on chromosome 10.

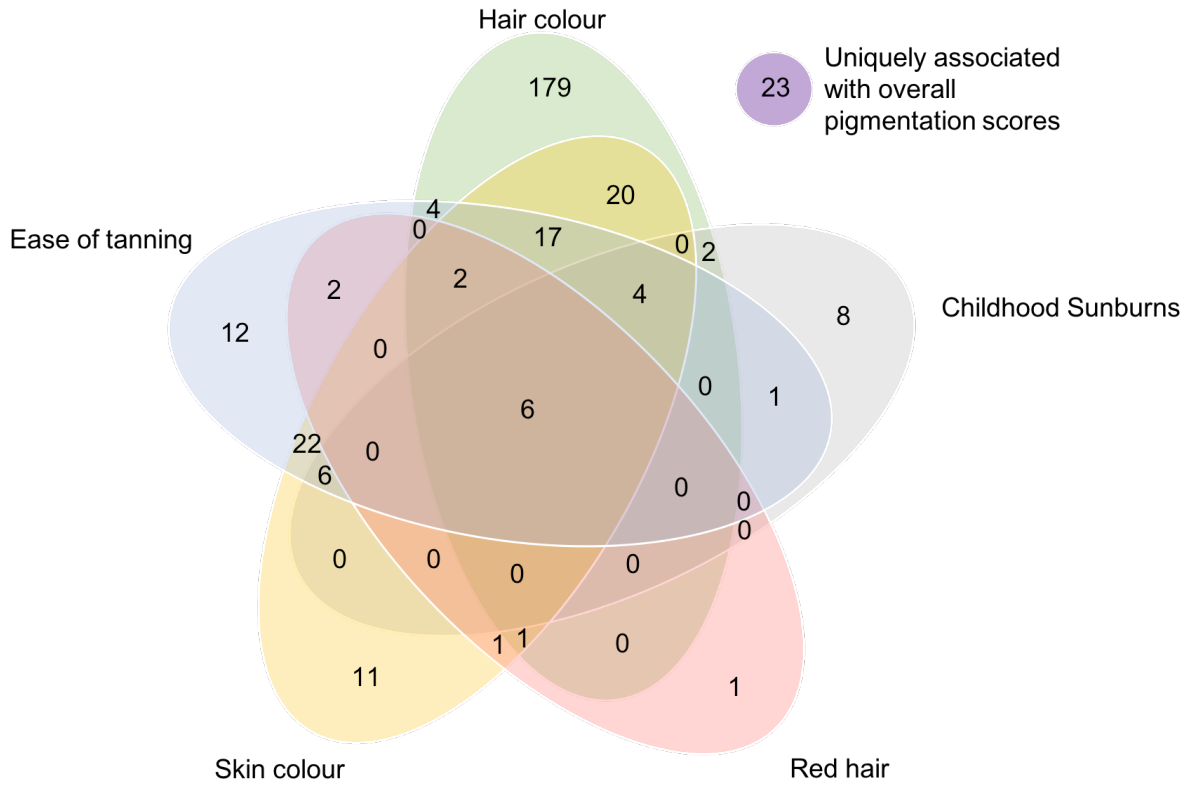


Figure 5.7: Overlap in loci of association for five pigmentary traits (hair colour (excluding red hair), red hair, tanning ability, skin colour and number of childhood sunburns). Independent signals from each GWAS were defined through conditional analysis with a p-value significance threshold of $P < 5 \times 10^{-8}$ and a collinearity threshold of $r^2 > 0.05$. Independent signals were classified as the same across pigmentary traits if they were within a collinearity threshold of $r^2 > 0.05$ and were then grouped into functional loci based on 500kb windows around the grouped signals.

5.4 Discussion

Here it is reported the results of multiple GWAS of pigmentation, including the largest GWAS ever conducted of either skin colour or tanning ability, two hair colour GWAS with hair colour separated into non-red and red hair, a childhood sunburn GWAS; and the results of a GWAS of combined pigmentation. Overall we identified 322 loci across all the traits, 109 of which have never previously been associated with a pigmentation trait. The discovery of many new loci improves our understanding of how the genetic architectures of distinct pigmentary measurements relate to one another. In particular, LD score regression estimated there to be strong genetic correlation between skin colour, tanning ability, and red hair; a moderate correlation between non-red hair with red hair, tanning ability, and skin colour; and strong correlation between the three combined pigmentation scores and the five pigmentary measures.

PCA analysis showed that an overall pigmentation was able to be modelled using the pigmentation variables with 99% of the variation between the variables being captured. The first pigmentation score captured over 60% of the variation between the pigmentation variables and captured an overall score for pigmentation as the coefficients for the pigmentation variables were all weighted evenly and followed the same pigmentary direction (light to dark). The second score appeared to capture a similar score profile, however the skin colour and tanning ability direction was switched; indicating a darker skinned, less likely to tan, but with light hair risk profile. The final score captured a primarily skin-based risk profile where hair colour had little weighting. The skin and tanning ability in this score indicate dark skinned but burn easily risk profile. When plotting the coefficients of the independent signals identified from the GWAS of these three scores, correlations between these scores were mixed. Low correlation was expected as these score would be orthogonal to one another, and thus, capture distinct pigmentary profiles. However, correlations ranged between -0.34 to 0.89, these mixed correlation are likely caused by the nature of the pigmentation profiles captured from the PCA – due to the similarity in coefficient effect size and direction in the pigmentation score models.

Plotting the beta coefficients (log odds) from the GWAS of pairs of categorical values from each pigmentary trait against one another revealed strong correlations, confirming the ordinal nature of these pigmentary traits. The only exception to this was hair colour: the genetic architecture of red hair is quite different to that of non-red hair colours. This is unsurprising given the different biosynthetic processes underlying eumelanin and pheomelanin production, and how these relate to the pigmentation of red hair (equal quantities of eumelanin and pheomelanin) and non-red hair (a small constant quantity of pheomelanin and gradually increasing quantities of eumelanin). This observation provides further evidence that red hair should be considered a disctinct trait in future genetic studies of hair colour. Plotting the beta coefficients of associated loci from the various pigmentary measures (now as ordinal measures) against one another, and estimating their correlation, demonstrated the strong underlying relationship between these traits. In particular, skin colour and ease of tanning had a high correlation of 0.88 as well as the high degree of associated variant overlap that also have similar effect sizes. Red hair was highly correlated with skin colour and ease of tanning, but with fewer shared variants than seen between

the skin-based traits; this could be because of the limited effective sample size for red hair. Non-red hair colour had the lowest correlation and the fewest shared genome-wide significant variants when compared to the other pigmentary traits; which highlights the uniqueness of the genetic architecture of non-red hair colour.

Chapter 6

Modelling pigmentation and melanoma risk with polygenic risk scores

6.1 Introduction

Current population-based GWAS of melanoma have identified 68 SNPs from 54 loci strongly associated with melanoma (Landi et al. 2020). The majority (76%) of these risk SNPs are jointly associated with at least one of three major heritable pathways to melanoma risk: pigmentation, naevus count, and telomere length. Given the well-established relationship between these three traits and melanoma, utilising pigmentation GWAS summary statistics to predict melanoma risk could further our current risk prediction models.

6.1.1 Originality of research

Studies aiming to identify high risk individuals have focused on factors such as family history (C. M. Olsen et al. 2010), phenotypic characteristics (skin colour, hair colour, eye colour, and naevus counts) (Sara Gandini et al. 2005), and UV exposure (solar or artificial) (Cust, B. K. Armstrong, et al. 2011). Studies that included genetic information commonly limited their models to include only a small number of well-established SNPs that are strongly associated with melanoma such as *MC1R* (Cust, Drummond, et al. 2018; Fang et al. 2014; Kypreou et al. 2016), resulting in marginal gains in risk prediction. By utilising large databases containing genetic and phenotypic information on the three functional pathway traits it should be possible

to create PRS for each trait to improve our current melanoma risk prediction models. This is largely due to these large databases allowing for a greater detection power and thus are able to provide a greater number of associated SNPs that are acting through these three functional pathways to be included into the risk model.

6.1.2 Chapter Aims & Objectives

Aims

I. Determine polygenic risk scores to predict pigmentary traits.

II. Determine pigmentary traits which can be utilised to predict melanoma.

Objectives

1. Create polygenic risk scores for each UK Biobank pigmentary trait and for an overall pigmentation score to be used to model pigmentation.

2. Create polygenic risk scores for each pigmentary trait and for an overall pigmentation score to be used in a logistic regression melanoma risk prediction model.

6.2 Methods

All analysis was conducted in PLINK 2.0 (Purcell et al. 2007), GCTA (J. Yang et al. 2011), and R (R Core Team 2017) using the UK Biobank (Sudlow et al. 2015) and LMC.

6.2.1 Predicting pigmentary traits with PRS'

How well skin colour, ease of tanning, non-red hair colour and red hair colour can be modelled using genetic variants was investigated. To achieve this five polygenic risk score (PRS) models for each pigmentation GWAS (skin colour, ease of tanning, non-red hair colour, red hair colour, and pigscore 1) were produced using independent associated variant's betas (log odds ratios), as associated weights, derived from GCTA using different thresholds for inclusion of a SNP based on significance level ($p < 5 \times 10^{-4}$, $p < 5 \times 10^{-5}$, $p < 5 \times 10^{-6}$, $p < 5 \times 10^{-7}$, and $p < 5 \times 10^{-8}$) and a collinearity threshold of $r^2 < 0.05$. For each pigmentary trait that was modelled (skin colour, ease of tanning, non-red hair colour, and red hair colour), the optimal significance level for each GWAS based PRS was selected by scoring European only UK Biobank participants and

regressing the modelled trait on each significance level score and the first ten genetic principal components. As the first overall pigmentation score (pigscore 1) explained the majority of variation between UK Biobank pigmentation traits (60%) and showed an overall pigmentation profile compared to the other two overall pigmentation scores (see 5.3.1), pigscore 2 and pigscore 3 were not considered for modelling melanoma risk through PRS. Similarly, as childhood sunburn number didn't group well with other pigmentation traits in FA (see 3.3.3) and likely has inherent reporting bias, it was not considered for modelling melanoma risk through PRS.

Once optimal significance levels for our GWAS based PRS were determined when predicting skin colour, ease of tanning, non-red hair and red hair colour, AUC analysis and estimation of the proportion of variance explained through them was conducted by scoring participants in the independent Leeds melanoma cohort study (LMC), and regressing the modelled traits on the scored participants and the first ten genetic principal components.

6.2.2 Predicting melanoma risk with PRS

To investigate how pigmentation predicted melanoma risk, our previously defined pigmentation PRS were applied to melanoma data. The aim was to identify distinct pigmentary characteristic that best predict melanoma risk. I began by identifying which of the five varying significance levels, for each of our give pigmentation GWAS, optimally predicted melanoma in the UK Biobank. For each of our five pigmentation GWAS result, European descent participants in the UK Biobank were, again, scored using the five different significance level thresholds of SNP inclusion PRS. Melanoma case/control status was then regressed on each of the five scores separately with the first ten genetic principal components. Once the optimum significance level PRS for each of our five pigmentary traits were identified, additionally including the eight known functional *MC1R* SNPs (rs1805005, rs1805006 rs2228479, rs11547464, rs1805007, rs1805008 rs885479, rs1805009) in every model (as they are known to influence both pigmentation and melanoma risk and are not always well-identified by regression), scored participants in the independent LMC, and regressed melanoma case/control status on our scores with the first ten genetic principal components. To compare performance between our scores, AUC analysis and the estimated proportion of variance explained (R^2) through each score were used with the genetic principal components as performance measurements.

The ability of a naevus-based PRS and a melanoma-specific PRS to predict melanoma risk were also estimated for comparison. Naevus PRS were produced using the effect estimates for independently associated genetic variants from a naevus dataset (Duffy, Gu Zhu, et al. 2018) of 52,506 samples at five different significance thresholds ($p < 5 \times 10^{-4}$, $p < 5 \times 10^{-5}$, $p < 5 \times 10^{-6}$, $p < 5 \times 10^{-7}$, and $p < 5 \times 10^{-8}$) and a collinearity threshold of $r^2 > 0.05$. The melanoma PRS were produced using the same methodology from a large-scale meta-analysis dataset (Landi et al. 2020) of 36,760 cases. The optimal significance level for melanoma prediction was determined (as in section 6.2.1) using the UK Biobank data. The PRS for melanoma and naevus count were calculated for each of the 10127 individuals in the independent LMC and used 1383 and 14 SNPs, respectively. Melanoma status was then regressed on both the naevus and pigmentation PRS in the same model (as distinct pathways to melanoma risk), and on the melanoma PRS alone (which should include all pathways concurrently). The relative performance of these melanoma risk prediction models was compared using their AUC and estimated pseudo-R^2. While a melanoma PRS should be, in theory, the ideal predictor individual pathways based on commonly recorded traits have the advantage of being based on larger samples sizes so effect estimates will be more accurate and more variants will reach the significance threshold.

Additionally, a melanoma risk prediction model was investigated that contained our optimum melanoma and pigmentation PRS. To minimise double weighting joint regions of association, variants associated across both traits ($p < 5 \times 10^{-8}$) were taken into account by removing them from the melanoma PRS and retaining in the pigmentation PRS. LMC participants were then scored using these two PRS'. Self-reported melanoma case/control status was regressed on our melanoma (with no pigmentation variants) PRS (MelPR PRS), our pigmentation PRS (pigscore 1 PRS), and the first ten genetic principal components. Again, AUC analysis and the estimated proportion of variance was used to compare model performance.

In addition to determining whether an distinct pigmentary characteristic is superior at melanoma prediction, I investigated whether modelling melanoma risk prediction using genetic information in the form of a PRS is better than using the phenotypic self-reported data our GWAS are based on. Firstly, the pigmentary trait scales from the UK Biobank and LMC were altered to compliment each other by: reordering the LMC non-red hair colour scale by collapsing it into 'blond'

= 1, 'light brown' = 2, 'medium brown' = 2.5 and 'dark brown' = 3, 'black' = 4; and collapsing the UK Biobank skin colour variables to 'very fair' = 1, 'fair' = 2, 'olive brown' = 3, or 'darker than olive brown' = 4. Self-reported melanoma case/controls status was then regressed on the phenotypic skin colour, ease of tanning, non-red and red hair colour variables in the UK Biobank European only dataset. To minimise over fitting, LMC participants, who had no missing self-reported pigmentary data were scored using the pigmentation variable coefficients derived from the UK Biobank model. Melanoma case/control status was regressed on this phenotypic defined score with the first ten principal components and on our highest performing pigmentation PRS with the first ten principal components. Performance between the melanoma risk prediction models was assessed using AUC analysis and the estimated the proportion of variance explained (R^2) through the model variables

Finally, a 'best fit' melanoma prediction model was produced by regressing melanoma case/control status on the best-performing pigmentary PRS, self-reported sun behaviour at the age of 10 (were you ever sun-burnt and number of hours spend outdoors 9-5 for holidays), age, sex, and self-reported naevus count. To minimise over-fitting, the melanoma risk model was fitted using 90% of the LMC participants with all relevant data ($n = 978$) and tested on the remaining 10% ($n = 97$).

6.2.3 PRS prediction for Sub-group analysis

Age and sex

To investigate the efficacy of the 'best fit' melanoma risk model across different population phenotype groups, two sub-group analyses were conducted. Firstly, a sex-based analysis, where the LMC was dichotomised into males and females; and secondly, an age-based analysis, where the LMC was dichomised by the median LMC participant's age (58 years), to create 58 and under and over 58 groups. As with the previous 'best fit' melanoma risk models, each participant in the LMC had an overall pigmentation (pigscore 1) PRS, using 213 variants derived from the overall pigmentation GWAS conducted in 5.2.2. Each participant's naevi count, exposure data, age, and sex was matched with their pigmentation PRS and melanoma case control status in the LMC. For the sex-based analysis, melanoma risk was modelled using logistic regression with naevi count, age, exposure data, and their pigmentation PRS as model variables; and for the

age-based analysis, melanoma risk was modelled using logistic regression with naevi count, sex, exposure data, and their pigmentation PRS as model variables. Each pairwise sub-grouped melanoma risk model's goodness of fit was tested using the DeLong's test (DeLong et al. 1988) for two correlated ROC curves where a p-value $< 0.05/5$ was deemed to be statistically significant to account for multiple sub-group testing.

Melanoma type

The accuracy of 'best fit' melanoma risk model was compared across multiple tumour sites. LMC partipants were split into four set-up groups: (1) melanoma controls; (2) melanoma cases on limbs; (3) melanoma cases on trunks; and (4) melanoma cases on heads. Each melanoma site group was merged with the melanoma control group, to create three sub-groups for logistic regression analysis: trunkal melanomas vs controls; limb-sited melanomas vs controls; and head-sited melanomas vs controls. Each of these sub-groups were modelled using logistic regressions with the overall pigmentation PRS, exposure data, naevus count, age, and sex as model variables. AUC and R^2 were used to assess model fit, and the goodness of fit across analyses was compared using DeLong's test (DeLong et al. 1988) for two correlated ROC curves with a significant threshold of $p < 0.05/5$ to account for multiple sub-group testing.

6.2.4 Collinearity r^2 threshold

To investigate the affect LD collinearity had on SNP selection for pigmentation PRS, and the effectiveness of these pigmentation PRS when modelling melanoma risk, conditional analysis was conducted using the methodology previously defined (see 6.2.1) with additional r^2 thresholds ($r^2 = 0.9$, $r^2 = 0.5$) for pigscore 1. For each additional r^2 value, the same p-value thresholds were used ($p < 5 \times 10^{-4}$, $p < 5 \times 10^{-5}$, $p < 5 \times 10^{-6}$, $p < 5 \times 10^{-7}$, and $p < 5 \times 10^{-8}$), with the optimum being selected using the UK Biobank dataset. Melanoma risk was then modelled using the LMC dataset, and each optimal p-value model's goodness of fit was compared across the other r^2 values optimal p-value models using DeLong's test (DeLong et al. 1988) for two correlated ROC curves with a statistical difference threshold of $0.05/3$ to account for multiple testing.

6.3 Results

The optimal p-value thresholds for PRS predicting the pigmentation traits were 5×10^{-7} for tanning ability (consisting of 172 SNPs; Appendix F.4), and 5×10^{-8} for the remaining traits: non-red hair colour consisting of 374 SNPs (Appendix F.1); red hair colour consisting of 33 SNPs (Appendix F.2); skin colour consisting of 145 SNPs (Appendix F.3); and pigscore 1 consisting of 213 SNPs (Appendix F.5). PRS for each pigmentary trait explained at least 11% of the variation in the same trait in the independent LMC dataset. No PRS explained less than 2% of the variation or had an AUC value of less than 0.6 in any other pigmentary trait in the LMC (Table 6.1).

		Baseline	Pigscore 1	Ease	Skin	Non-red hair	Red
Ease	AUC	0.58	0.70	0.73	0.71	0.60	0.65
	R^2	0.01	0.11	0.15	0.12	0.02	0.07
		Baseline	Pigscore 1	Ease	Skin	Non-red hair	Red
Skin	AUC	0.58	0.75	0.74	0.75	0.68	0.65
	R^2	0.01	0.11	0.10	0.11	0.05	0.05
		Baseline	Pigscore 1	Ease	Skin	Non-red hair	Red
Non-red hair	AUC	0.58	0.72	0.62	0.69	0.81	0.61
	R^2	0.02	0.18	0.05	0.13	0.34	0.04
		Baseline	Pigscore 1	Ease	Skin	Non-red hair	Red
Red hair	AUC	0.56	0.85	0.82	0.82	0.65	0.89
	R^2	0.01	0.27	0.21	0.22	0.04	0.35

Table 6.1: AUC and R^2 of pigmentation-based PRS when predicting pigmentation traits. All PRS were optimised on the UK Biobank and tested on the LMC dataset

When modelling melanoma risk with the pigmentation PRS, all individual pigmentation traits had similar AUCs, with hair colour as the least predictive (AUC=0.59) and skin colour as the most predictive (AUC=0.61); however the PRS from the combined pigmentation scores (pigscore 1) was the most predictive of melanoma risk (AUC=0.64) (Figure 6.1). This suggests either that the combined pigmentation PRS benefits by capturing multiple aspects of pigmentation or that, given the similar performance of the individual-trait PRS in predicting melanoma risk, the combined pigmentation PRS is improved by amalgamating multiple measures of a single factor underlying melanoma risk, into a single more accurate measure. The inclusion of the naevus

PRS with the combined pigmentation (pigscore 1) PRS made little difference to the prediction of melanoma risk (Figure 6.2). Testing pigmentation PRS calculated using previously identified novel signals from the corresponding pigmentation GWAS (5.3.4), found minimal contribution for melanoma risk prediction with novel signals from the tanning ability, pigscore 1, and skin colour GWAS being the most predicted (AUC: 0.5923, 0.5902, 0.5895, respectively) and red and non-red hair least predictive (AUC: 0.5171 and 0.5458, respectively) (Figure 6.3). This result is unsurprising as recent hair colour-based GWAS (Hysi et al. 2018; Morgan et al. 2018) have utilised large data sets to detect rare and low-penetrance signals compared to skin-based pigmentation GWAS (F. Liu et al. 2015; Visconti et al. 2018) resulting in hair-based novel signals identified in this thesis (5.3.4) numbering less and having less contribution to melanoma risk prediction compared to novel skin-based signals.

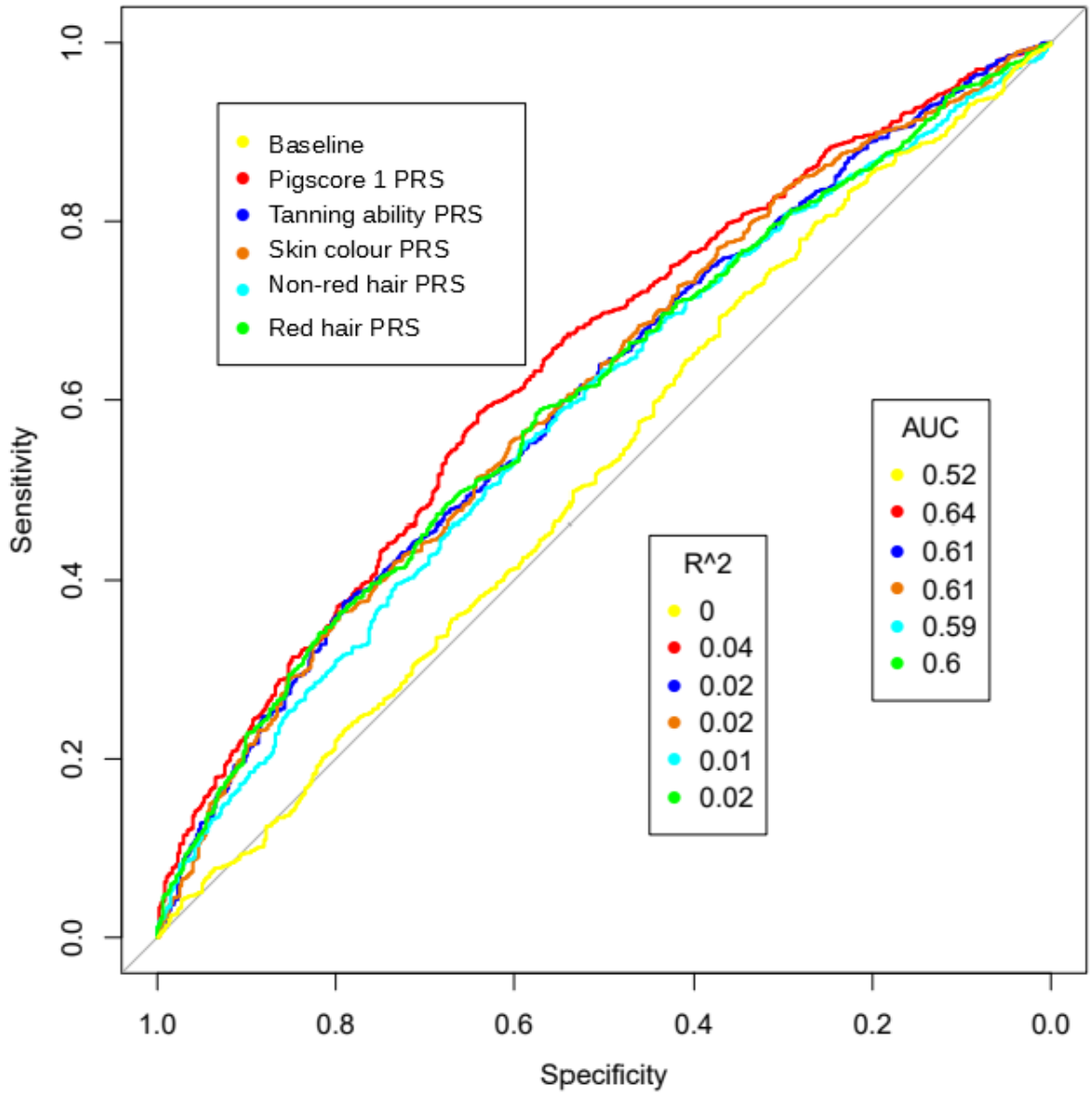


Figure 6.1: ROC curves for the prediction of melanoma using PRS derived from: pigscore 1, tanning ability, skin colour, non-red hair colour, red hair, and a baseline prediction model with no PRS included. All PRS were optimized on the UK Biobank and tested on the LMC

When comparing the combined pigmentation PRS (pigscore 1 PRS) with the melanoma-specific PRS, the two performed similarly in predicting melanoma risk (AUC of 0.65 for melanoma PRS, AUC of 0.64 for pigscore 1 PRS). There was only a slight gain by jointly modelling melanoma risk using the MelPR PRS and combined pigmentation PRS concurrently (AUC=0.66) (Figure 6.2).

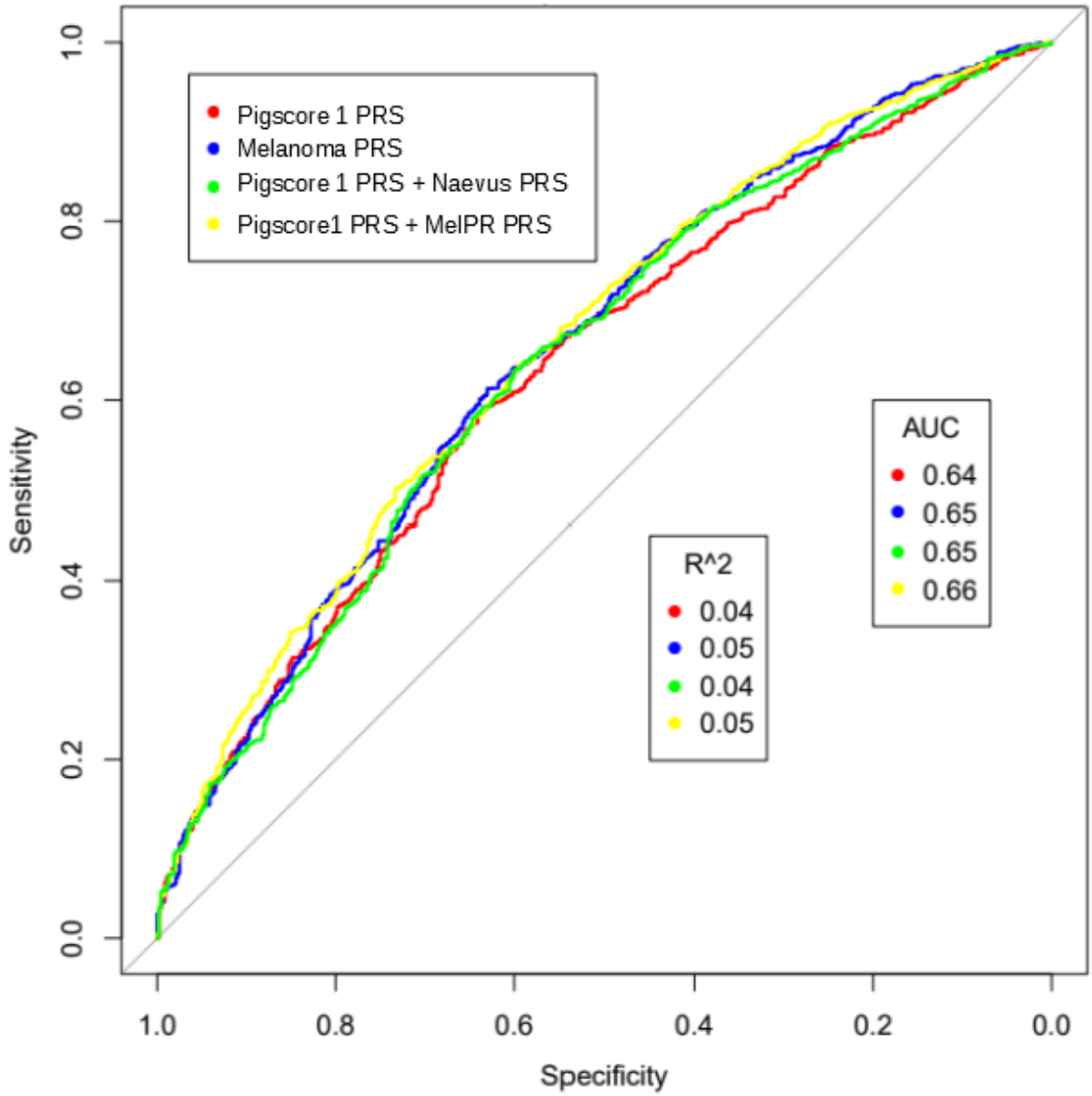


Figure 6.2: ROC curves plot of melanoma prediction using PRS derived from: pigscore 1, an independent melanoma PRS, pigscore 1 and naevus count, and a hybrid combination of melanoma and pigscore 1 by removing pigmentary loci from the melanoma PRS and replacing them with pigscore 1 loci. All PRS were optimized on the UK Biobank and tested on the LMC

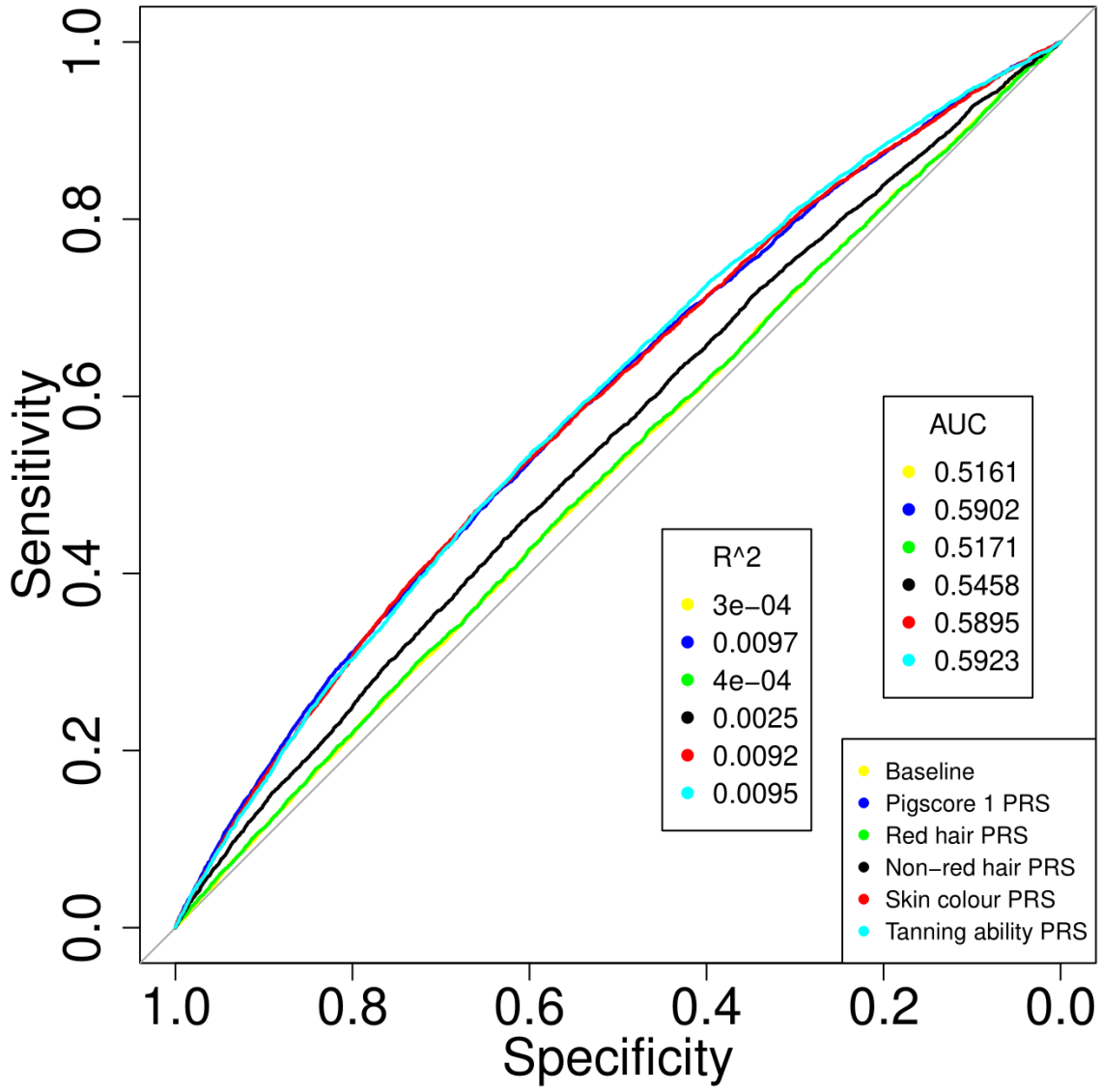


Figure 6.3: ROC curves for the prediction of melanoma using PRS derived using novel signals from: pigscore 1, tanning ability, skin colour, non-red hair colour, red hair, and a baseline prediction model with no PRS included. All PRS were tested on the UK Biobank

A comparison of the performance of self-reported pigmentation variables to the pigscore 1 genetic pigmentation score in predicting melanoma risk found that the genetic score slightly outperformed the self-reported pigmentation in predicting melanoma risk (AUC of 0.65 for pigscore 1; AUC of 0.62 for self-reported pigmentation) (Figure 6.4).

The inclusion of traditional factors in the model (naevus count, sex, age, UV exposure) in addition to the combined pigmentation genetic score (pigscore 1) resulted in markedly improved

melanoma prediction, with an AUC of 0.76 and proportion of variation explained of 0.18. This outperformed a similar model, that differed only by using self-reported pigmentation rather than the combined pigmentation genetic score (pigscore 1) (AUC=0.74, R^2=0.17) (Figure 6.4). For the melanoma risk model using self-reported pigmentation, only skin colour had a significant effect ($p < 0.05$) on model performance with non-red hair colour, red hair colour, and tanning ability not reaching the significance threshold in the model (Table 6.2). Similar performance was observed when comparing between the overall pigmentation (pigscore 1) PRS and a melanoma specific PRS, with the melanoma PRS having a marginally higher AUC (0.77) and lower R^2 value (0.18).

Model:	**Variable**	**Effect estimate**	**Std. error**	**p-value**
Melanoma PRS, sex, age, exposure data, number of naevi	Melanoma PRS	0.96	0.38	< 0.01
	Sex	0.11	0.16	0.46
	Age	0.004	0.008	0.63
	Number of naevi	0.07	0.003	< 0.005
	Sun-burnt age 10	0.15	0.21	0.48
	Hours outside age 10	-0.001	0.01	0.88
Pigmentation PRS, sex, age, exposure data, number of naevi	Pigmentation PRS	-0.80	0.20	< 0.005
	Sex	0.13	0.16	0.43
	Age	0.004	0.008	0.63
	Number of naevi	0.03	0.003	< 0.005
	Sun-burnt age 10	0.15	0.21	0.46
	Hours outside age 10	-0.002	0.003	0.80
hair, skin and tanning ability self-report, sex, age, exposure data, number of naevi	non-red hair colour self-report	-0.23	0.11	0.04
	red hair colour self-report	0.70	0.39	0.86
	skin colour self-report	-0.59	0.17	< 0.005
	tanning ability self-report	0.03	0.11	0.82
	Sex	0.06	0.16	0.69
	Age	0.004	0.008	0.65
	Number of naevi	0.03	0.003	< 0.005
	Sun-burnt age 10	0.12	0.20	0.57
	Hours outside age 10	-0.0004	0.009	0.97

Table 6.2: Included variable effect sizes, SEs, and p-value for melanoma risk models.

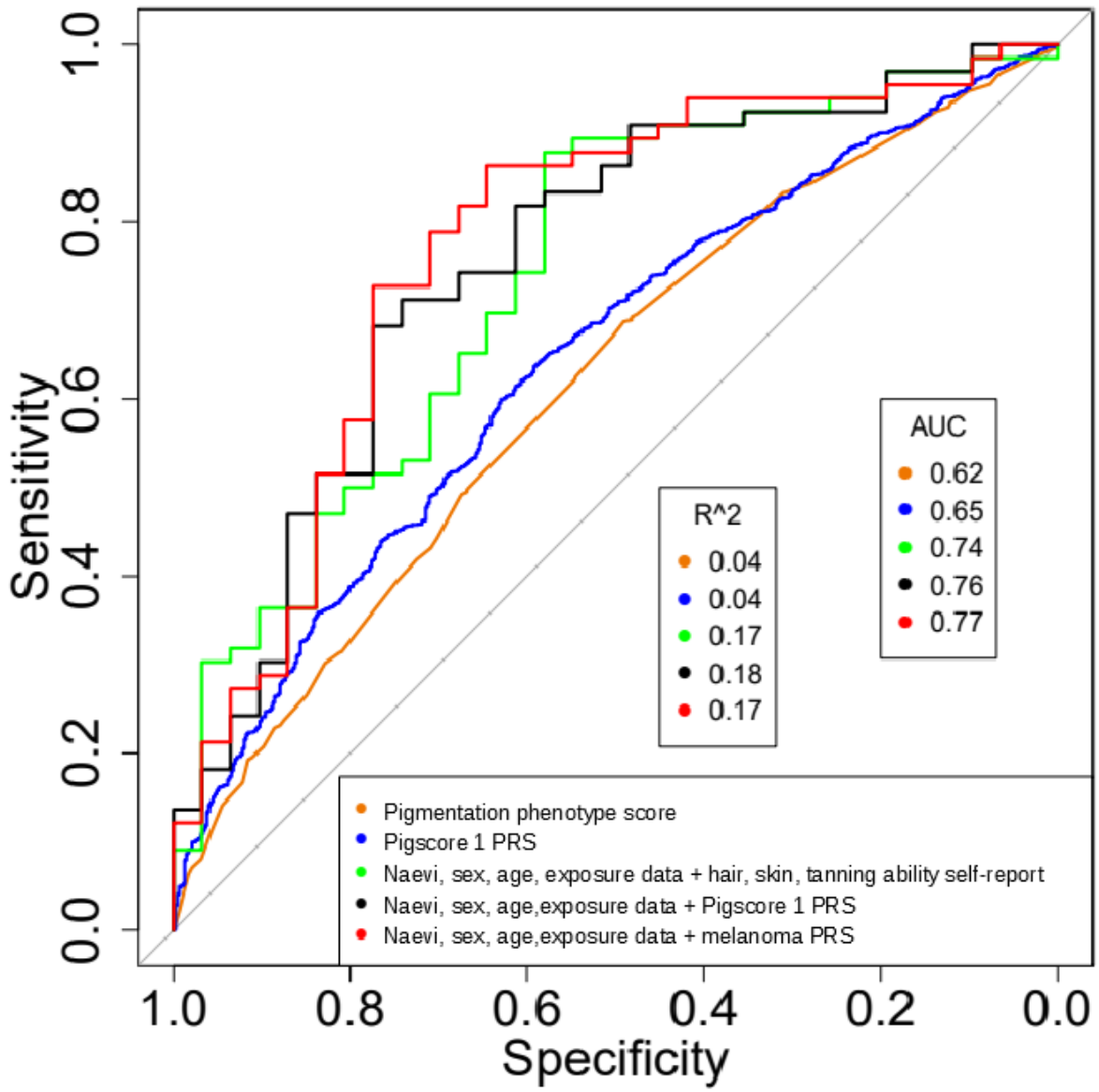


Figure 6.4: ROC curves plot of the prediction of melanoma using: an overall pigmentation score derived from the phenotypic variables, pigscore 1, and two best melanoma prediction models that utilise age, sex, UV exposure data, self-reported naevus count, and either, a melanoma based PRS, the pigscore 1 PRS or overall pigmentation phenotype score. All PRS were optimized on the UK Biobank and tested on the LMC

When sub-grouping the LMC dataset on particpant's sex, an increase of AUC (0.81) and R^2 (0.23) was observed for the male group when modelling melanoma using the pigmentation PRS and traditional factors (naevi, age, and exposure data), and a decrease in AUC (0.65) and R^2 (0.19) for the female group when compared to the un-grouped LMC dataset models. However, when comparing between the male and female melanoma risk models using DeLong's test for two correlated ROC curves found no statistical significance ($p = 0.18$) (Figure 6.5). Sub-grouping the LMC dataset by participant's age produced minimal differences between AUC (58 and un-

der: 0.75; Over 58: 0.77) and R^2 (58 and under: 0.26; Over 58: 0.15) between the two age groups when modelling melanoma risk using the pigmentation PRS and traditional factors (Figure 6.6); and when compared to the un-grouped LMC dataset, there was minimal difference in AUC and R^2. Testing goodness of fit using DeLong's test for two correlated ROC curves found no statistically significant difference between the models ($p = 0.86$).

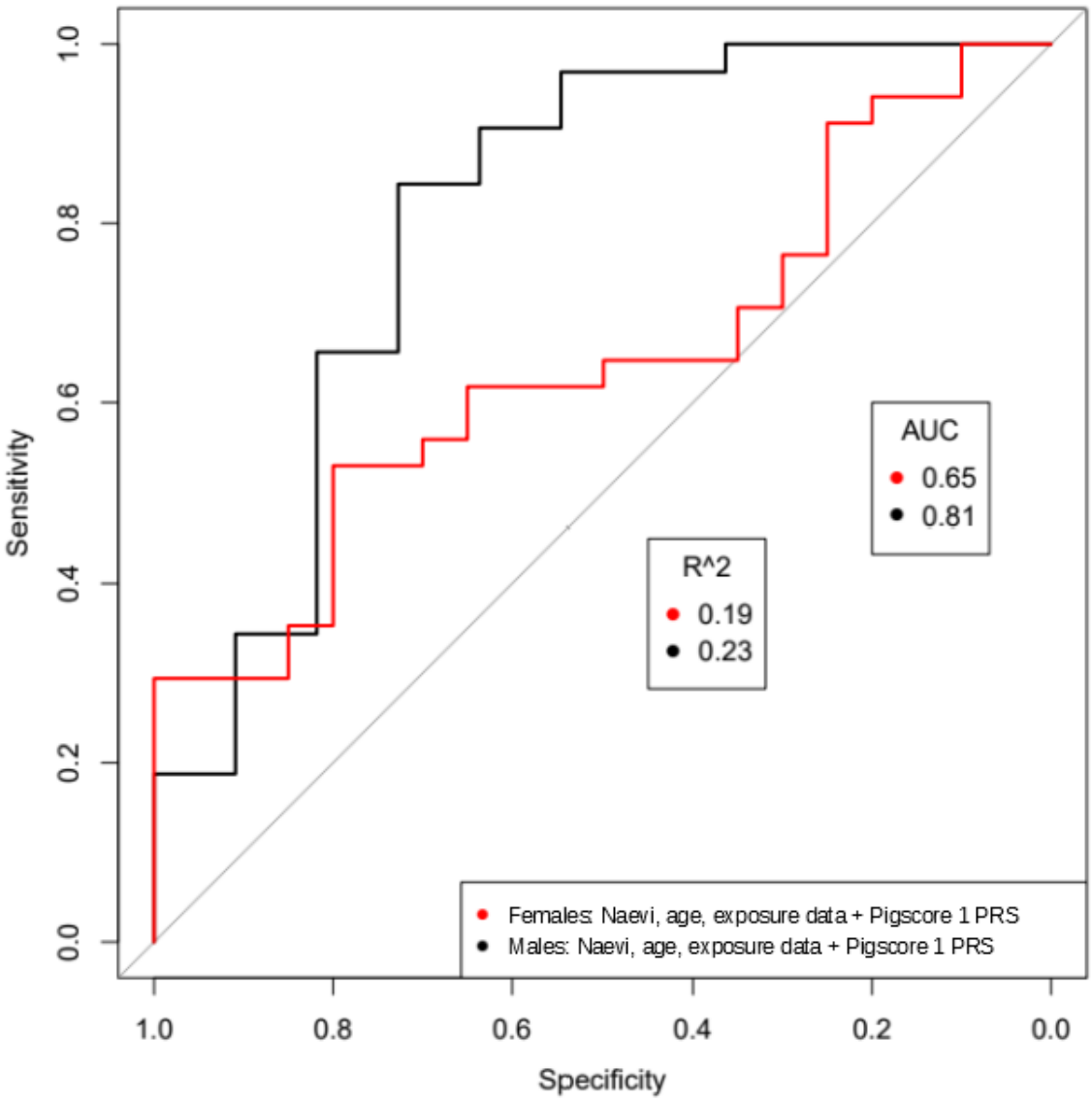


Figure 6.5: ROC curves plot of the prediction of melanoma sub-grouped by sex (male or female). Melanoma prediction models utilises age, UV exposure data, self-reported naevus count information and the overall pigmentation PRS derived from pigscore 1.

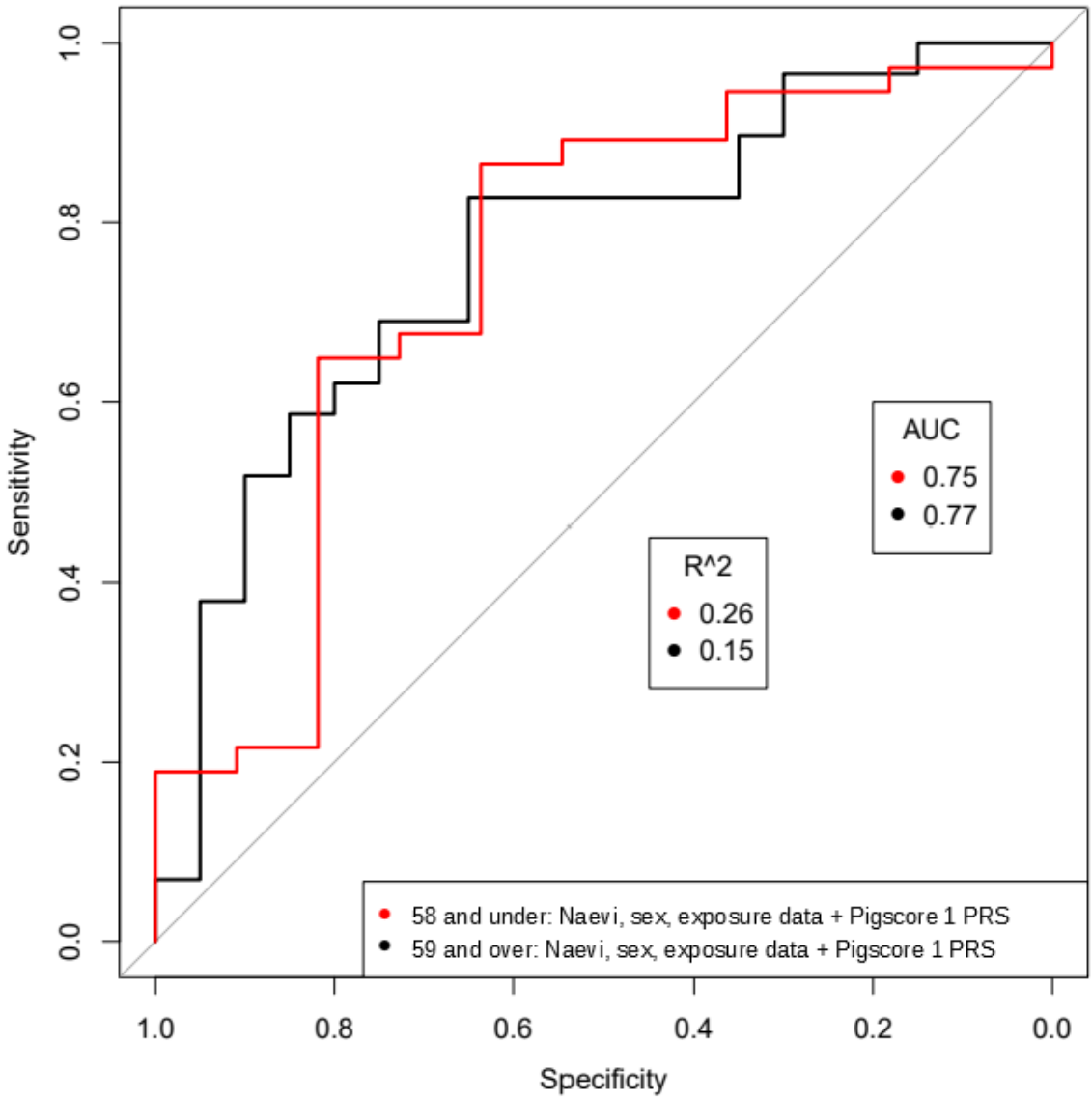


Figure 6.6: ROC curves plot of the prediction of melanoma sub-grouped by age (58 and under, and 59 and over). Melanoma prediction models utilises sex, UV exposure data, self-reported naevus count information and the overall pigmentation PRS derived from pigscore 1.

Sub-grouping the LMC dataset into tumour site (limb, head and trunk) and modelling these site based melanoma cases using the traditional factors and pigmentation PRS, found melanoma cases on limbs had the highest AUC (0.8) and second highest R^2 (0.24), compared to melanoma cases on heads (AUC 0.78, R^2 0.15) and melanoma cases trunks (AUC 0.62 R^2 0.27) (Figure 6.7). Although a difference between AUC and R^2 between the three site was identified, when comparing the goodness of fit between the three site models using DeLong's test for two correlated ROC curves, no statistical significant difference was detected across any of the models (Limb vs head: $p = 0.84$; limb vs trunk: $p = 0.16$; and head vs trunk $p = 0.025$).

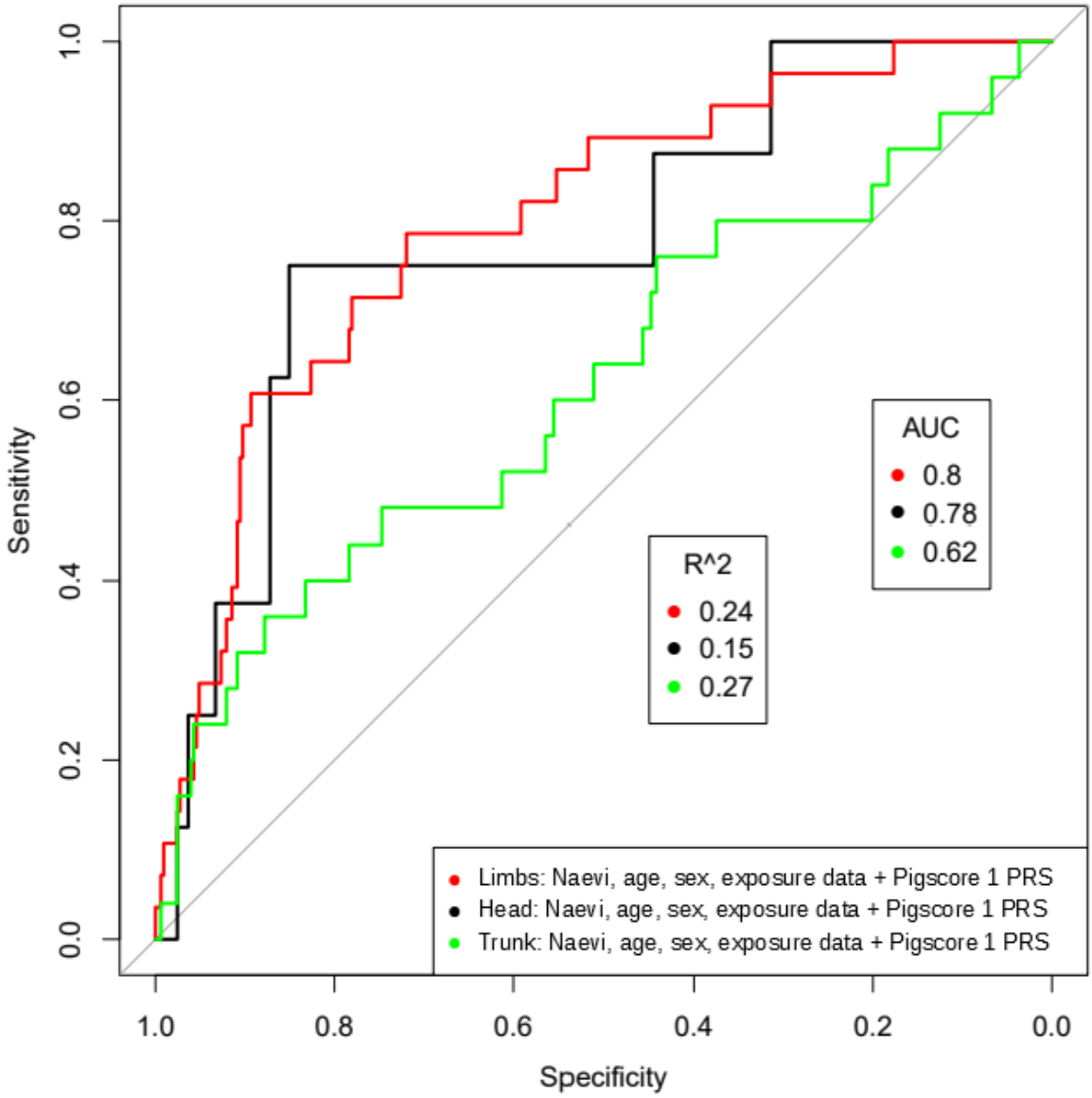


Figure 6.7: ROC curves plot of the prediction of melanoma sub-grouped by tumour site (limbs, head, or trunk). Melanoma prediction models utilises age, UV exposure data, self-reported naevus count information and the overall pigmentation PRS derived from pigscore 1.

Finally, investigating the effect that differing the collinearity threshold had on pigmentation PRS prediction of melanoma risk found little difference between the three thresholds used (r^2 = 0.9, 0.5, 0.05). SNP numbers where similar across thresholds with r^2=0.9 470 snps; r^2=0.5 470 snps r^2=0.05 214 snps at the $p < 5 \times 8^{-8}$ level; to r^2=0.9 2159 snps; r^2=0.5 2026 snps r^2=0.05 1786 snps at the p04 level. After selecting optimal models for each threshold (r^2 = 0.9: $p < 5 \times 10^{-4}$; r^2 = 0.5: $p < 5 \times 10^{-8}$; r^2 = 0.05: $p < 5 \times 10^{-8}$) the pigmentation PRS derived from an r^2 threshold of 0.05 had the highest AUC (0.638) and R^2 (0.039), and the lowest being the pigmentation PRS derived from a collinearity threshold of 0.9 (AUC: 0.631,

R^2: 0.035). When testing for model fit difference between the three thresholds using DeLong's test for two correlated ROC curves found no pairwise statistical difference between any of the models (r^2=0.05 vs r^2=0.5 $p = 0.36$; r^2=0.05 vs r^2=0.9 $p = 0.39$; r^2=0.5 vs r^2=0.9 $p = 0.67$) (Figure 6.8).

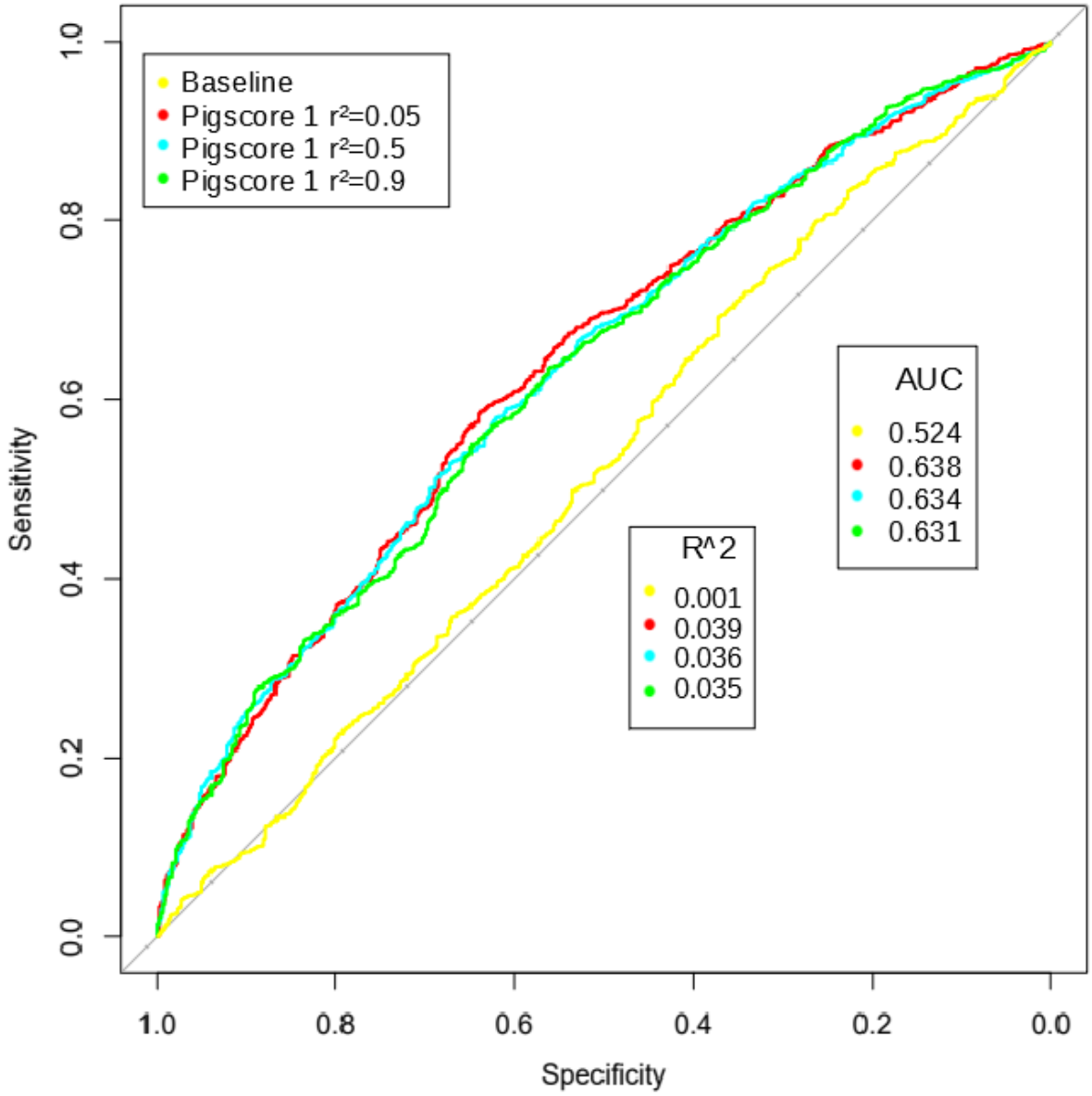


Figure 6.8: ROC curves plot of the prediction of melanoma using three collinearity thresholds ($R^2 = 0.05, 0.5, 0.9$) for signal selection for the pigscore 1 PRS.

6.4 Discussion

Here is reported the effectiveness of using pigmentation-based PRS in predicting melanoma risk. There was no particular pigmentary trait whose PRS was notably more strongly or weakly associated with melanoma risk, suggesting that there is no specific aspect of pigmentation this is

more predictive of melanoma risk than any other. A PRS based on a PCA-derived combined pigmentation score was the most predictive of melanoma risk. This may be due to the combined pigmentation score bringing together multiple aspects of pigmentation; more likely the combined pigmentation score benefits from increased accuracy by combining information from multiple measures all reflecting the same underlying genetic factors (given the strong correlation between them).

Several factors in this chapter have improved the reliability of our results. Firstly I have avoided over-fitting and demonstrated validity across datasets by defining our predictive models in one dataset (UK Biobank) and testing them on another (the Leeds Melanoma Cohort). We reduced the effects of population stratification by including the first ten genetic principal components in all GWAS, and by conducting analyses on a homogeneous population group (people of European descent resident in the UK).

Importantly it was observed that a combined pigmentation PRS is slightly better at predicting melanoma risk than self-reported pigmentation; which suggests melanoma risk is better predicted by the genetic drivers of phenotypic risk factors than the self-reported phenotypes themselves. This is likely due to the genetic-based scores avoidance of self-reporting error or bias (e.g. a lack of clarity in the scale used, or bias towards reporting particular coloration, cultural bias, effect of UV exposure or age on pigmentation), and that the genetic-based scores also utilise results of very large sampled sized to give an overall more precise measurement of pigmentation. In addition to the benefits of result accuracy, replacing self-reported pigmentation with genetic scores allow comparability across studies (of the same ethnicity), avoids bias, and will only improve over time as studies of pigmentation get bigger.

Also, very similar performance between a melanoma-specific PRS (AUC=0.65) and the combined pigmentation (pigscore 1) PRS (AUC=0.64) was observed. This appears counter-intuitive as the melanoma-specific PRS is designed explicitly to predict melanoma risk while the pigmentation PRS only reflects one pathway in the genetic architecture of melanoma risk. However, the pigmentation PRS is based on a far larger dataset than is available for melanoma risk, resulting in more accurate effect estimates for associated genetic variants and hence a PRS based

on more variants; this increase in informativeness is the likely reason that the pigmentation PRS performs so well.

The inclusion of naevus count, UV exposure data, age, and sex information in these models sharply increased melanoma risk prediction, although this is likely due to naevus count as it had a significant contribution to the melanoma risk model. This increase in melanoma risk prediction demonstrates the importance of these phenotypic data to melanoma risk. When comparing to previous melanoma risk models (Cust, Drummond, et al. 2018) that have utilised exposure, phenotypic and limited genetic data, an improvement was observed due to the large amount of genetic information in our models (AUC=0.76 compared to AUC=0.68). This improvement is likely due to polygenic nature of melanoma risk, and by increasing the number of subtle genetic variant, the PRS is able to explain more variation of melanoma risk, and thus the melanoma risk model has higher prediction power.

When sub-grouping the LMC cohort into sex, age, and tumour site, no statistically significant difference was observed between groups. Although previous research has identified differences in PRS performance across melanoma tumour sites (Mark's PRS work), the non-statistical significant observations was likely due a lack of power to detected differences between models caused by small sample sizes within the melanoma risk models. Conducting the PRS with differing the colinearity threshold (r^2: 0.9, 0.5, 0.05) produced similar results, which was likely caused by the similarities in the number of SNPs selected in the conditional analysis for each r^2 threshold.

As the effective sample size of melanoma GWAS increases, evidence mounts of the key role of the three heritable pathways (pigmentation, naevus count and telomere length) in the genetic risk of melanoma and their potential for melanoma risk prediction. Current GWAS of melanoma have identified 54 associated loci with 34 also associated with pigmentation, 14 also associated with naevus count, and 5 also associated with telomere length (Landi et al. 2020) (11 of these associated with more than one pathway). This is further supported by the strong genetic correlation identified between melanoma and the pigmentary traits (see 5.3.2). The pigmentation PRS achieved similar melanoma risk prediction to a melanoma-specific PRS. However, when considering the other two genetic risk pathways, current naevus (Duffy, Gu Zhu, et al. 2018)

and telomere length (C. Li et al. 2020) GWAS have much smaller sample sizes (naevus: 52,506 samples, telomere: 78,592 samples) and have identified five and 42 loci of association respectively. This limitation restricts the potential value of PRS derived from these results to model melanoma risk prediction; however, given the importance of these pathways to melanoma risk (S. Gandini et al. 2005; Iles et al. 2014) and their high heritability (G. Zhu et al. 1999; Wachsmuth et al. 2001; Slagboom et al. 1994; Vasa-Nicotera et al. 2005; Njajou et al. 2007; Bakaysa et al. 2007) increased sample sizes for GWAS of naevus count and telomere length will undoubtedly uncover many more associated loci and increase the utility of PRS based on these to predict melanoma risk. The potential impact of just a naevus PRS can be seen from the increase in the AUC for melanoma prediction achieved by including self-reported naevus count (0.68 to 0.76). Risk prediction of melanoma based on joint modelling by PRS of pigmentation, naevus count and telomere length has the potential to be very powerful.

Recent GWAS of breast cancer have uncovered multiple susceptibility variants, each carrying a small individual risk, but in combination accounting for a substantial increase in risk. Current breast cancer PRS contain over 300 SNPs and achieve an AUC of 0.63 whilst only adjusting for country and no other prediction factor in regression models (Mavaddat et al. 2019). This high predictability has led to clinical studies that aim to use the breast cancer PRS to inform targeted breast screening programs (Shieh et al. 2017; D. G. Evans et al. 2016). In comparison, for melanoma, the primary limiting factor in PRS risk prediction has been relative lack of power to detect susceptibility variants. This limitation could be overcome by utilising pigmentation, naevus, and telomere length PRS derived from well-powered GWAS. With increased of power for GWAS of these traits, a substantial increase in the ability to predict melanoma risk can be expected, with the future potential for targeted advice or even screening for those at highest risk.

In conclusion, these results highlight the power of pigmentary genetics in predicting melanoma risk by deriving PRS of combined pigmentary measures that outperform self-reported pigmentation and markedly improve on current melanoma prediction models. Thus we demonstrate that in predicting disease risk there are substantial gains to be made by utilising PRS of heritable disease-associated common or continuous traits, which benefit from large datasets, rather than just PRS that model the disease itself.

Chapter 7

Joint-analysis of pigmentation and melanoma

7.1 Introduction

As it's well-established the majority of melanoma risk SNPs are jointly associated with at least one of three major heritable pathways to melanoma risk: pigmentation, naevus count, and telomere length. Using pigmentation GWAS summary statistics to supplement melanoma meta-analysis GWAS power may further our understanding of the genetic architecture of melanoma.

Recently, many more loci have been identified as being associated with melanoma, largely due to the availability of large databases of genetic information on participants, the overarching limitation to such studies being sample size and power. This sample size issue is primarily caused by the rarity of melanoma in the population. As it is the rarest form of skin cancer – only accounting for 2.3% of all skin cancer cases (Corrie et al. 2014) – acquiring sufficient genotyped cases can be problematic for studies. Previous solutions to this have been to meta-analyse multiple melanoma studies to increase the overall power (Landi et al. 2020). Although this has been successful and identified further loci of association, the polygenic nature of melanoma means there are likely numerous subtle associations yet to be identified. This has been further backed up by studies investigating the overall genetic contribution of melanoma and estimating the heritability (h^2) to be 0.085 (Landi et al. 2020).

7.1.1 Originality of research

Although increasing sample power by meta-analysing multiple melanoma studies has increased our melanoma risk understanding, recent methods to jointly analyse multiple phenotypic traits using GWAS summary statistics have become popular. Broadly, these methods take GWAS summary statistics from phenotypic trait A and combine the results with the GWAS summary statistics of phenotypic trait B to increase the effective sample size to detect genetic associations for both trait A and trait B. The primary constraint to utilising these methods is understanding the relationship between the two traits being jointly analysed. This is due to the fact that there must be genetic overlap between the two traits; otherwise there may be numerous false-positive genetic associations in the results. More specifically to melanoma, there is a well-established relationship between the genetic architectures of naevus count, telomere length and pigmentation with melanoma, which indicates that these methods could be used to exploit common characteristics relating to these three traits with respect to melanoma risk, such as the abundance of genetic and phenotypic pigmentation data available in comparison to melanoma.

7.1.2 Chapter Aims & Objectives

Aims

I. Determine whether the abundant pigmentary trait data can be used to increase power to detect melanoma risk loci.

Objectives

1. Create an analysis pipeline (AP) that can identify regions of joint association between melanoma and a pigmentary trait and jointly analyse regions of shared association.
2. Conduct joint-analysis of melanoma and pigmentation using weighted meta-analysis methods

7.1.3 Rationale to achieve aims & objectives

This chapter explores two distinct approaches that are able to jointly analyse melanoma risk data with UK Biobank pigmentation traits and an overall pigmentation score derived in previous chapters. The first approach is an AP which adopts the most basic joint-analysis methodology (a sample size weighted joint-analysis (SSWJA)) to identify regions that reach genome-wide sig-

nificance when traits are combined (and might not with melanoma alone) and combines it with a Bayesian methodology which is implemented to ensure identified regions from the SSWJA associated with pigmentation and melanoma are not reaching genome-wide significance because one trait (pigmentation) is having a large effect on estimates and the other (melanoma) is not. Combining these two steps in a pipeline results in the identification of multiple small regions associated with melanoma risk that benefit from the large amount pigmentation data.

The second approach: a genetic correlation weighted joint-analysis (GCWJA) considers the entire genome when combining pigmentation and melanoma GWAS summary statistics. But, rather than targeting specific regions thought to be associated with both traits, a weighting derived from the genetic correlation between both traits is applied when jointly analysing melanoma risk and pigmentation. The two methodologies are then compared to each other and also with the original melanoma meta-analysis GWAS results to highlight the potential new melanoma risk associations identified through these joint-analyses.

7.2 Methods

All analysis was conducted in PLINK 2.0 (Purcell et al. 2007), FUMA (Watanabe et al. 2017), GWAS-PW (Pickrell et al. 2016), ldsc (Brendan K. Bulik-Sullivan et al. 2015), MTAG (Turley et al. 2018), METAL (Willer et al. 2010), and R (R Core Team 2017) using UK Biobank (Sudlow et al. 2015).

7.2.1 Identifying signals for melanoma skin cancer

In order to distinguish between genome-wide significant regions associated with pigmentation, with melanoma or with both, the program GCTA was used. Conditional analysis were conducted on melanoma meta-analysis summary results to identify independently associated signals with a p-value significance threshold of $p < 5 \times 10^{-8}$ and collinearity threshold of $r^2 < 0.05$. Any such independent signals less than 500kb apart were declared to be part of the same functional locus. These defined functional loci would form the reference loci for joint-analyses of melanoma and pigmentation.

7.2.2 Approach one: Analysis Pipeline

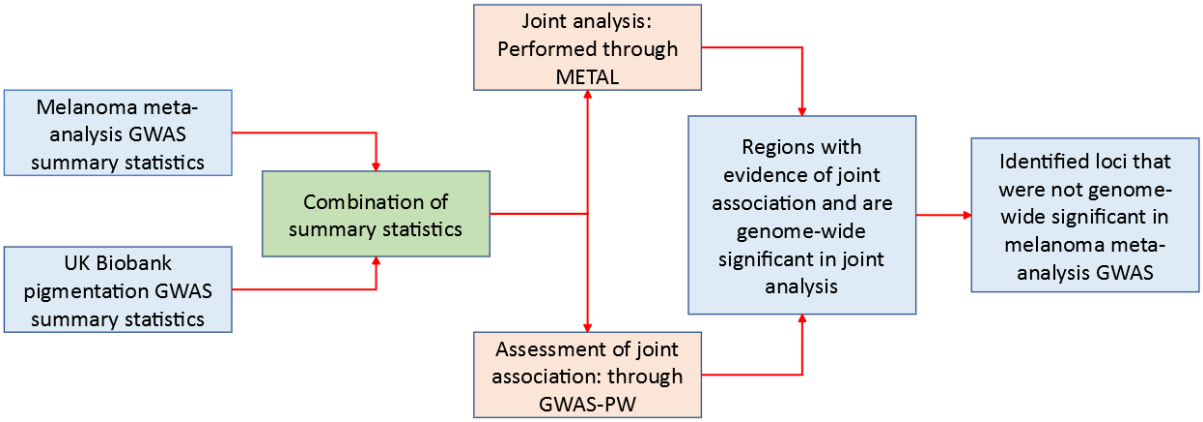


Figure 7.1: Overview approach one: an AP for joint-analysis of UK Biobank pigmentation GWAS summary statistics and melanoma meta-analysis GWAS summary statistics.

Highlighting regions of joint association between melanoma and pigmentation

To highlight regions of joint association between melanoma and pigmentation, the similarity in genetic architecture was investigated between melanoma and each pigmentary trait using a Bayesian approach. GWAS-PW was adopted to assess pairwise GWAS summary statistics for

melanoma and each pigmentary trait in turn to determine the probability of sequential 50 SNP windows being associated with one, both or neither trait. Each pigmentary GWAS and the melanoma GWAS summary statistics were prepared for analysis by retaining only the SNPs they shared in common, removing rare variants (MAF < 0.01), omitting SNP with missing information, calculating the Z-value (using $Z = \frac{log(OR)}{SE}$) and variance (using $V = SE^2$) for each SNP. The rs number for each SNP, its chromosome, base pair position, Z-value for trait 1, variance for trait 1, Z-value for trait 2, and variance for trait 2 were combined into a single data frame and analysed using GWAS-PW. Regions were classified as jointly associated with the pigmentary trait used and melanoma if the probability of them being associated with both traits was greater than 0.7.

METAL software

To identify regions that reach genome-wide significance when traits are combined in the analysis pipeline, a SSWJA that jointly analyses melanoma with each individual pigmentation trait and score was implemented in the software METAL. For pigmentation and melanoma GWAS summary statistics to be valid for the SSWJA, the data had to firstly be formatted to ensure that each SNP's effect size, for each trait, followed the same directional scale. This was achieved by converting pigmentation traits to follow the typical melanoma risk direction (1 = non-melanoma case and 2 = melanoma case): red hair became 1 (non-red hair) and 2 (red hair); non-red hair colour became 'black' = 1, 'dark brown' = 2, 'light brown' = 3, and 'blond' = 4; skin colour became 'black' = 1, 'brown' = 2, 'dark olive' = 3, 'light olive' = 4, 'fair' = 5, and 'very fair' = 6; and tanning ability became 'don't burn only tan' = 1, 'mainly tan partially burn' = 2, 'partially tan mainly burn' = 3, and 'don't tan only burn' = 4. Converted scales were then combined with the melanoma meta-analysis data and SNP QC was conducted. The QC consisted of retaining SNPs that were present in both the selected pigmentation dataset and melanoma dataset; removing any SNP with a MAF < 0.01; removing any SNP where the number of samples was below 90% of the total sample size for the dataset. To ensure the effect size correctly followed the melanoma risk direction, SNP alleles were compared to ensure the effect and reference allele were the same between the two datasets. If a discrepancy between the alleles was detected, the pigmentation reference and effect allele were flipped to match that of the melanoma database and the Z-score sign inverted.

Post QC datasets containing one of the pigmentation datasets matched with the melanoma dataset were then uploaded into the METAL software and analysed. To identify independent associated signals, completed jointly analysed results were downloaded and then input into FUMA to identify independent loci using the methodology described in 2.6.3.

Regions identified as appropriate for joint-analysis

As the sample size joint-analyses gives results on the single-variant level while GWAS-PW gives results for a region. Post joint-analyses and joint association validation, cross-validation was implemented to map each 50 SNP window association with both melanoma and each pigmentation trait with the detected genome-wide associated ($p < 5 \times 10^{-8}$) loci from the corresponding joint-analyses. This was completed by compiling each 50 SNP window associated with both the pigmentary traits and melanoma. Any independent signal identified through conditional analysis of the SSWJA was represented by a lead SNP, and searched for in the compiled list of 50 SNP windows that were associated with both pigmentation and melanoma. If multiple lead SNPs were found to reside in any given 50 SNP window, further investigation took place to ensure each loci were truly independent by re-evaluating the LD between the two lead SNPs and verifying it was below the cutoff 0.05.

Combining regions of joint association with melanoma meta-analysed GWAS results

For each individual pigmentation trait and melanoma AP, lead SNPs were linked to the original pigmentation trait's GWAS summary statistics and original melanoma meta-analysis summary statistics. The beta coefficients (log odds) for these lead SNPs in each the original GWAS were plotted and compared to ensure the direction and value in the original melanoma meta-analysis matched with the pigmentation trait. A 95% confidence interval for each lead SNP's beta coefficient was calculated for each pigmentation trait and for the melanoma meta-analysis by adding/taking away the respective standard error times 1.96 (normal tail value for bell curve for 95% confidence); these were then included in the beta coefficient plots. SNP characteristics (MAF, total number of participants with the SNP, melanoma meta-analysis p-value, previous associations) for newly identified associations were investigated to draw commonalities between them and to review the validity of the newly found association.

Identifying cross-pigmentary trait associations in the joint-analysis.

To identify recurring pigmentation-associated signals across the five joint-analyses (melanoma analysed with non-red hair colour, red hair colour, skin colour, tanning ability, and pigscore 1), any pair of genome-wide significant variants which were within a region identified as appropriate to jointly analyse (see 7.2.2) and which were in LD with $r^2 > 0.05$ were declared as equivalent and classified as the same signal. Once recurring signals had been identified, they were again grouped into loci if they were less than 500kb apart. Each locus were then compiled into a list of regions found to associated with both melanoma and at least one pigmentary trait.

The compiled list of associations were dichotomised into previously known melanoma risk signals and newly identified melanoma risk signals. The newly identified signals were input into FUMA software and GWAS catalog (Buniello et al. 2019) to investigate previously-identified trait associations. If any signal was flagged as associated with a trait, a further literature review was conducted focusing on the biological process and phenotypic/genetic risk factors to determine whether there could be a meaningful link with melanoma. Further to this, any 500kb window around a signal was also investigated for previous associations with any known melanoma risk pathways, as signals associated with two or more pathway traits have been found to affect the magnitude and direction with respect to melanoma risk.

7.2.3 Approach two: Genetic correlation weighted meta-analysis

As well as the AP to highlight joint association between melanoma and pigmentation, a second methodology was adopted that utilizes LD score regression between different phenotypic traits to create a GCWJA of GWAS summary statistics.

LD score regression between melanoma and pigmentary traits

In order to be able to weight the joint-analyses using genetic correlation, firstly LD score regression was conducted on melanoma and each pigmentary trait and score. The method relies on the fact that GWAS effect size estimates (beta coefficients/log odds) for any given SNP incorporate the effects of all SNPs in LD with that SNP; thus, SNPs with a higher LD will more likely have a higher χ^2 value compared to SNPs with a lower LD, and that for nonzero genetic correlations this holds true for Z scores. By utilising these traits, it is possible for the method to derive a

genetic correlation between two different traits.

To derive the genetic correlation between the pigmentation and melanoma GWAS summary statistics, datasets had to be formatted and combined into a set-up file. Firstly, SNP ID, chromosome, base pair position, SNP sample size, p-value, beta coefficient/odds ratio (linear or logistic model GWAS output), and SE information were extracted for all SNPs from the pigmentary GWAS summary statistics and melanoma meta-analysis summary statistics. For each dataset, Z-scores were derived for each SNP by dividing the beta coefficients/log odds ratios by their estimated SE. For the pigmentary traits, the effect size directions were converted to follow same direction as for melanoma risk in the melanoma meta-analysis (1 = non-melanoma case and 2 = melanoma case): red hair became 1 (non-red hair) and 2 (red hair hair); non-red hair colour became 'black' = 1, 'dark brown' = 2, 'light brown' = 3, and 'blond' = 4; skin colour became 'black' = 1, 'brown' = 2, 'dark olive' = 3, 'light olive' = 4, 'fair' = 5, and 'very fair' = 6; and tanning ability became 'don't burn only tan' = 1, 'mainly tan partially burn' = 2, 'partially tan mainly burn' = 3, and 'don't tan only burn' = 4. Converted scales were then combined with the melanoma meta-analysis data and SNP QC was conducted. The QC consisted of retaining SNPs that were present in both the selected pigmentation dataset and melanoma dataset; removing any SNP with a MAF < 0.01; removing any SNP where the number of samples was below 90% of the total sample size for the dataset. To ensure the effect size correctly followed the melanoma risk direction, SNP alleles were compared to ensure the effect and reference allele were the same in the two datasets. If a discrepancy between the alleles was detected, the pigmentation reference and effect allele were flipped to match that of the melanoma database and the beta coefficient (log odds) multiplied by -1.

Post QC, the formatted files were input into the ldsc software where LD score regression was conducted. The genetic correlation outputs for each combination of pigmentary trait or score with melanoma was extracted and retained to be used as the basis for the weighting in the joint-analysis. A strong correlation was defined as $rg^2 > 0.6$ and weak correlation as $rg^2 < 0.2$.

MTAG software

To conducted the genetic correlated weight joint-analysis between the pigmentation and melanoma GWAS summary statistics, datasets were formatted and QCed using the same methodology for the ldsc software (See 7.2.3). Post formatting and QC, each pigmentation trait and melanoma file and the corresponding genetic correlation (calculated in 7.2.3) was input into the MTAG software where GCWJA was conducted. The output files for melanoma risk, each weighted using the genetic correlation between melanoma and one of the eight pigmentary traits, were collated; the following information was extracted: SNP ID, chromosome, base pair position, pre-analysis Z-score, pre-analysis p-value, pre-analysis SE, post-analysis Z-score, post-analysis p-value, and post-analysis SE. FUMA was used on each of the five files using the post-analysis GWAS summary statistics to identify independent signals of association for melanoma (see 2.6.3 for methods).

MTAG results

For each individual pigmentation trait and melanoma joint-analysis, lead SNPs were linked to the original pigmentation trait's GWAS summary statistics and original melanoma meta-analysis summary statistics. The beta coefficients (log odds) for these lead SNPs in each the original GWAS were plotted and compared to ensure the direction and value in the original melanoma meta-analysis matched with the pigmentation trait. A 95% confidence interval for each lead SNP's beta coefficient was calculated for each pigmentation trait and for the melanoma meta-analysis by calculating $+/-1.96\times$ SE; these were then included in the beta coefficient plots.

Identifying overlap in associations

To identify recurring pigmentation-associated signals across the five GCWJA (melanoma weighted using the genetic correlation between melanoma and non-red hair colour, red hair colour, skin colour, tanning ability, and pigscore 1), any pair of genome-wide significant variants within 500kb distance and with LD of $r^2 > 0.05$ were declared as equivalent and classified as the same signal. All loci were then compiled into a list of regions that are found to associated with melanoma and at least one pigmentary trait.

The compiled list of associations was dichotomised into either previously known melanoma risk

signals or newly identified melanoma risk signals. These newly identified signals were input into FUMA software and GWAS catalog (Buniello et al. 2019) to investigate any previously-identified trait associations. If any of these loci were identified as being previously associated with a trait, a further literature review was conducted focusing on the biological process and phenotypic/genetic risk factors to determine whether there could be a meaningful link with melanoma. Further to this, any 500kb window around a signal was also investigated for previously know associations with any known melanoma risk pathways, since signals associated with two or more pathway traits have been found to affect the effect size and direction with melanoma risk.

7.2.4 Overlap of signals in the pipeline to the GCWJA

To identify recurring pigmentation-associated signals across the five APs and five GCWJA (melanoma analysed with non-red hair colour, red hair colour, skin colour, tanning ability, and pigscore 1), any pair of genome-wide significant variants from either approach which were in LD with $r^2 > 0.05$ were declared as equivalent and classified as the same signal. Once recurring signals had been identified through the analyses, they were grouped into loci if they were less than 500kb apart. Each locus were then compiled into a list of regions found to be associated with both melanoma and at least one pigmentary trait.

To assess the number of previously known melanoma-associated loci that were not identified through either the APs or GCWJA, prior known melanoma signals identified through a previous melanoma GWAS study (Landi et al. 2020) or GWAS catalog (Buniello et al. 2019) were combined with the pigmentation-associated signals identified through the APs or GCWJA. The same analysis outlined above (7.2.4) was conducted with the included melanoma-only signals that were not identified in the either joint-analysis approaches. Once analysis was conducted, each locus were then compiled into a list of regions found to be associated with both melanoma and at least one pigmentary trait, or found to be associated with melanoma only and wasn't identified in either joint-analysis approaches. The compiled list of melanoma-associated locus was input into FUMA and GWAS-catelog (Buniello et al. 2019) to identify any previously known associations with naevus count or telomere length.

7.3 Results

7.3.1 Analysis Pipeline

Joint-analysis of melanoma and pigmentary traits using GWAS-PW identified required a posterior probability > 0.7 for association with both traits. This identified 4416 50-SNP windows for melanoma and skin colour (Figure 7.5); 4508 50-SNP windows for melanoma and ease of tanning (Figure 7.2); 2540 50-SNP windows for melanoma and non-red hair colour (Figure 7.3); 1698 50-SNP windows for melanoma and red hair (Figure 7.4); and 3931 50-SNP windows for melanoma and the pigscore 1 (Figure 7.6).

To determine which of the 50 SNP windows demonstrating evidence for association (posterior probability of Model III > 0.7) with both melanoma and a pigmentation trait showed genome-wide significant levels of association, these were combined with the SSWJA summary statistics. Combining these analyses and comparing to the original GWAS for each trait identified 70 regions for tanning ability and melanoma, where 44 regions had prior genome-wide significance for tanning ability in the original GWAS, 3 regions had prior significance for melanoma in the original GWAS, and 23 regions had prior significance in both original GWAS; 64 regions for skin colour and melanoma, where 36 regions had prior genome-wide significance for the original skin colour GWAS, 4 regions had prior significance for the original melanoma GWAS, and 24 regions had prior significance in both original GWAS; 44 regions for non-red hair colour and melanoma, where 21 regions had prior genome-wide significance in the original non-red hair colour GWAS, 2 regions had prior significance in the original melanoma GWAS, and 21 regions had prior significance in both original GWAS; 22 regions for red hair and melanoma, where 10 regions had prior genome-wide significance in the original red hair GWAS, 3 regions had prior significance in the original melanoma GWAS, and 9 regions had prior significance in both original GWAS; and 69 regions for pigscore 1 and melanoma, where 36 regions had prior genome-wide significance in the pigscore 1 GWAS, 3 regions had prior significance in the melanoma GWAS, and 30 regions had prior significance in both original GWAS.

When considering a more stringent Model III posterior probability (> 0.9) for GWAS-PW, which equates to a 90% false discovery rate, 1789 50 SNPs windows were identified for non-red

hair colour and melanoma; 2962 for tanning ability and melanoma; 2827 for skin colour and melanoma; 1158 for red hair and melanoma; and 2667 pigscore 1 and melanoma. Combining the more stringent Model III posterior probability thresholds with the SSWJA summary statistics identified 33 regions for non-red hair colour and melanoma, 46 regions for tanning ability and melanoma; 39 for skin colour and melanoma; 14 for red hair and melanoma; and 49 regions for pigscore 1 and melanoma.

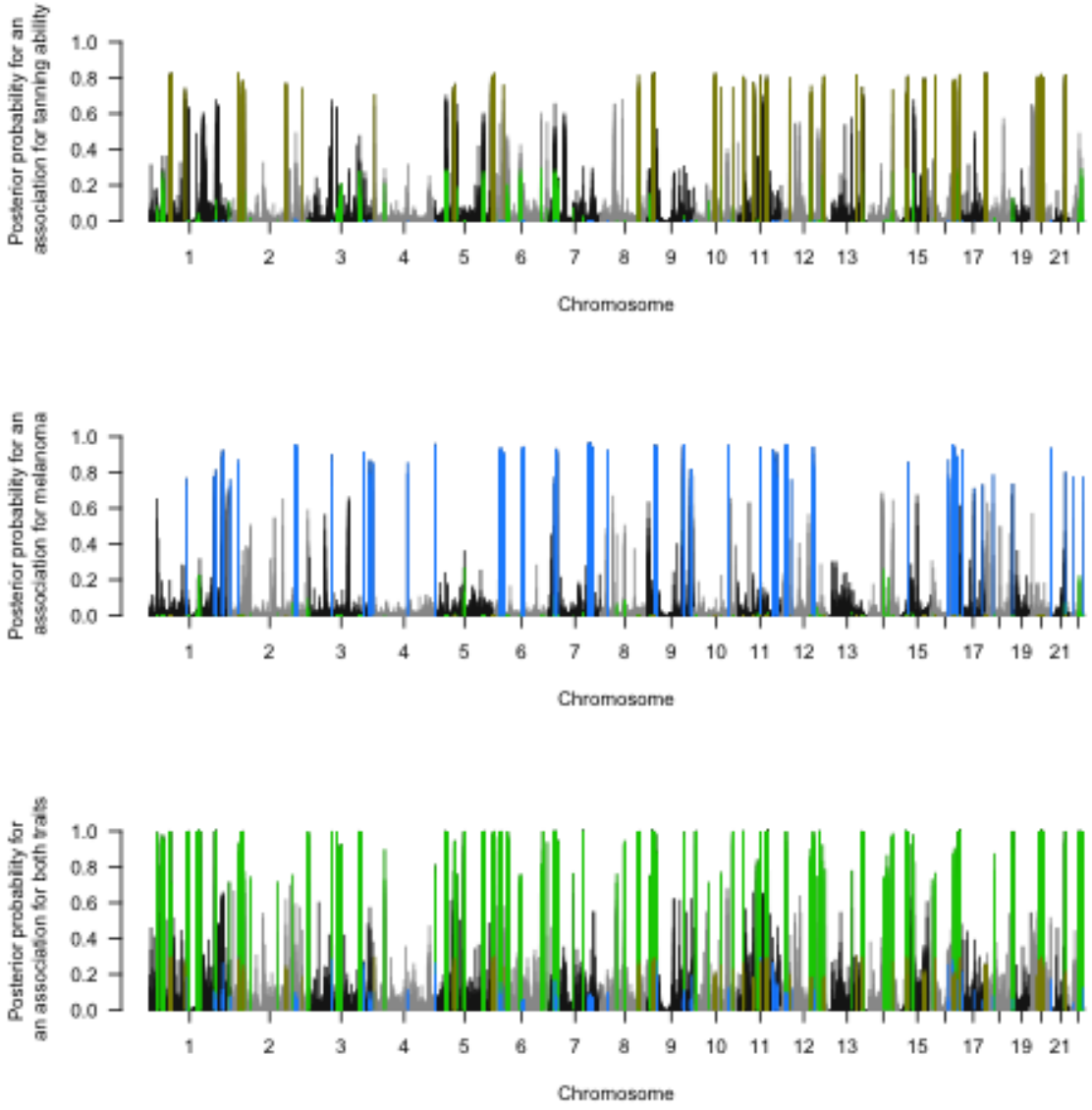


Figure 7.2: 50 SNP window GWAS-PW analysis between melanoma and tanning ability. The top plot indicates the probability of each 50 SNP window being uniquely associated with tanning ability (model I); the middle plot indicates the probability of each 50 SNP window being associated uniquely associated with melanoma (model II); and the bottom plot indicates the probability of each 50 SNP window being associated with tanning ability and melanoma (model III). 50 SNP windows are coloured yellow if the probability is > 0.7 for model I, blue for probability is > 0.7 for model II, and green for probability is > 0.7 for model III.

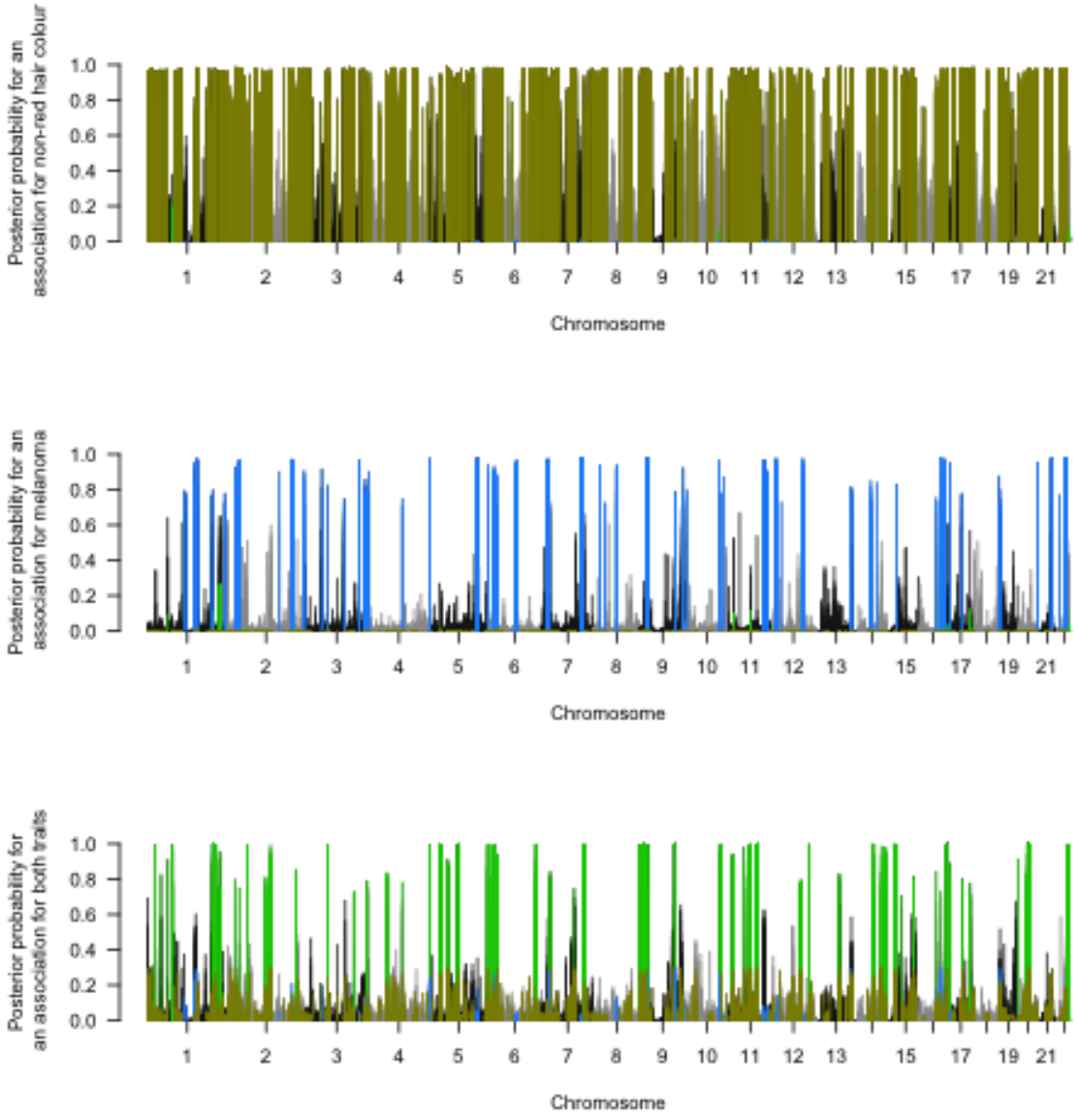


Figure 7.3: 50 SNP window GWAS-PW analysis between melanoma and non-red hair colour. The top plot indicates the probability of each 50 SNP window being uniquely associated with non-red hair colour (model I); the middle plot indicates the probability of each 50 SNP window being associated uniquely associated with melanoma (model II); and the bottom plot indicates the probability of each 50 SNP window being associated with non-red hair colour and melanoma (model III). 50 SNP windows are coloured yellow if the probability is > 0.7 for model I, blue for probability is > 0.7 for model II, and green for probability is > 0.7 for model III.

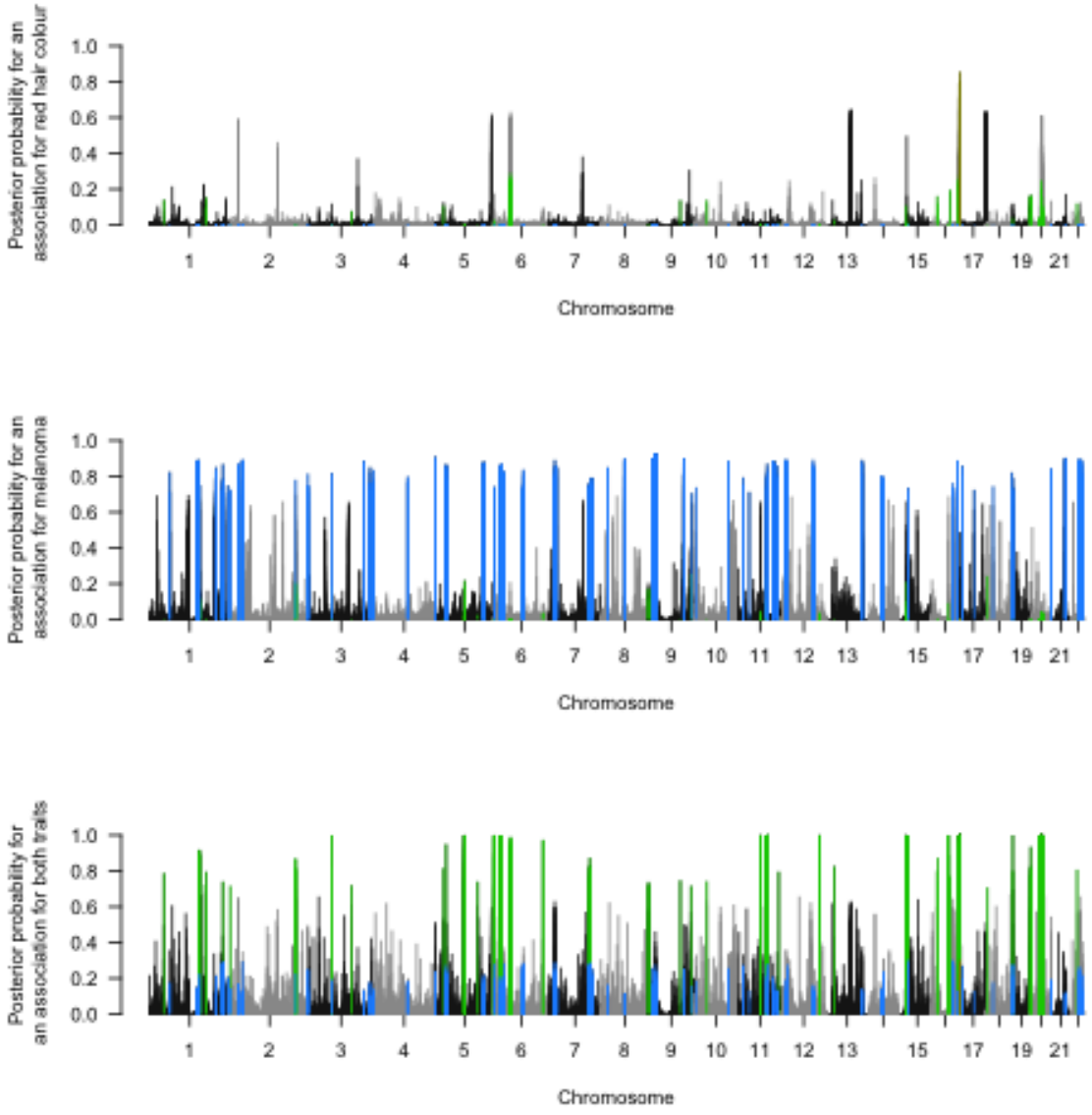


Figure 7.4: 50 SNP window GWAS-PW analysis between melanoma and red hair colour. The top plot indicates the probability of each 50 SNP window being uniquely associated with red hair colour (model I); the middle plot indicates the probability of each 50 SNP window being associated uniquely associated with melanoma (model II); and the bottom plot indicates the probability of each 50 SNP window being associated with red hair colour and melanoma (model III). 50 SNP windows are coloured yellow if the probability is > 0.7 for model I, blue for probability is > 0.7 for model II, and green for probability is > 0.7 for model III.

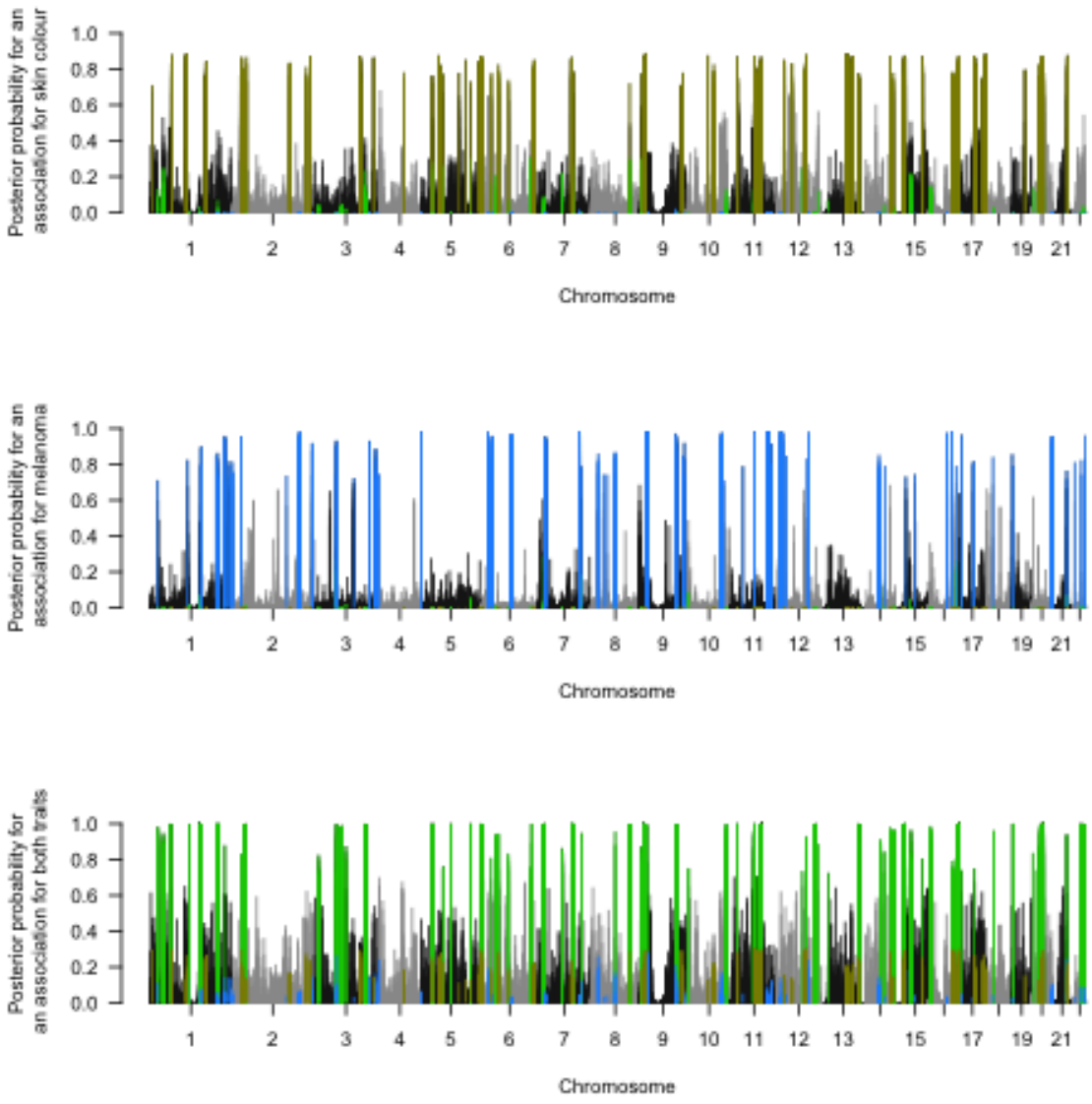


Figure 7.5: 50 SNP window GWAS-PW analysis between melanoma and skin colour. The top plot indicates the probability of each 50 SNP window being uniquely associated with skin colour (model I); the middle plot indicates the probability of each 50 SNP window being associated uniquely associated with melanoma (model II); and the bottom plot indicates the probability of each 50 SNP window being associated with skin colour and melanoma (model III). 50 SNP windows are coloured yellow if the probability is > 0.7 for model I, blue for probability is > 0.7 for model II, and green for probability is > 0.7 for model III.

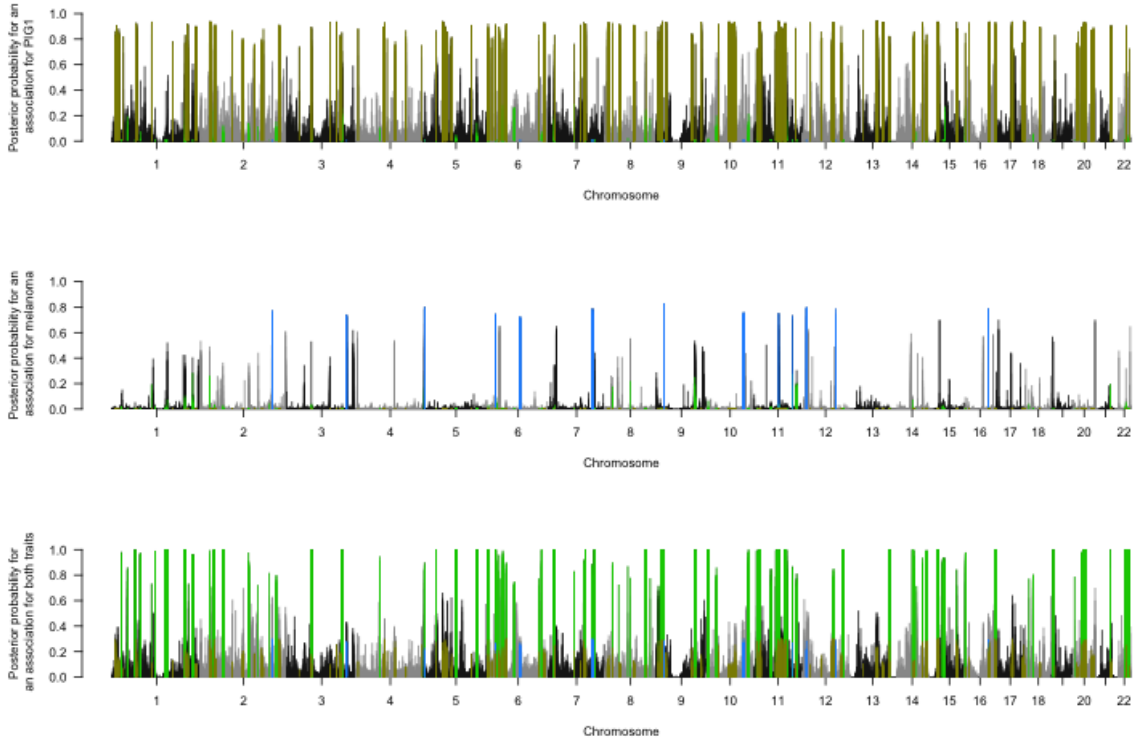


Figure 7.6: 50 SNP window GWAS-PW analysis between melanoma and pigscore 1. The top plot indicates the probability of each 50 SNP window being uniquely associated with pigscore 1 (model I); the middle plot indicates the probability of each 50 SNP window being associated uniquely associated with melanoma (model II); and the bottom plot indicates the probability of each 50 SNP window being associated with pigscore 1 and melanoma (model III). 50 SNP windows are coloured yellow if the probability is > 0.7 for model I, blue for probability is > 0.7 for model II, and green for probability is > 0.7 for model III.

The beta coefficients (log odds ratios) of the lead SNPs at those loci identified through the AP were compared across pigmentation traits (Figures 7.7, 7.8, 7.9, 7.10, 7.11). Correlations (Spearman's) were high between the various pigmentation traits and melanoma, indicating high overlap in the genetic architecture between the measured pigmentation traits (skin colour, tanning ability, non-red hair, red hair, and pigscore 1 with melanoma). The most strongly correlated beta coefficients were between pigscore 1 and melanoma ($\rho = 0.77$) and the most weakly correlated beta coefficients between red hair and melanoma ($\rho = 0.05$) (Table 7.1).

	Ease	Skin	Non-red	Red	Pigscore 1
Melanoma	0.71	0.63	0.65	0.05	0.77

Table 7.1: Identified lead SNP beta coefficient correlation between five SSWJA: tanning ability and melanoma, skin colour and melanoma, non-red hair colour and melanoma, red hair colour and melanoma, and pigscore 1 and melanoma.

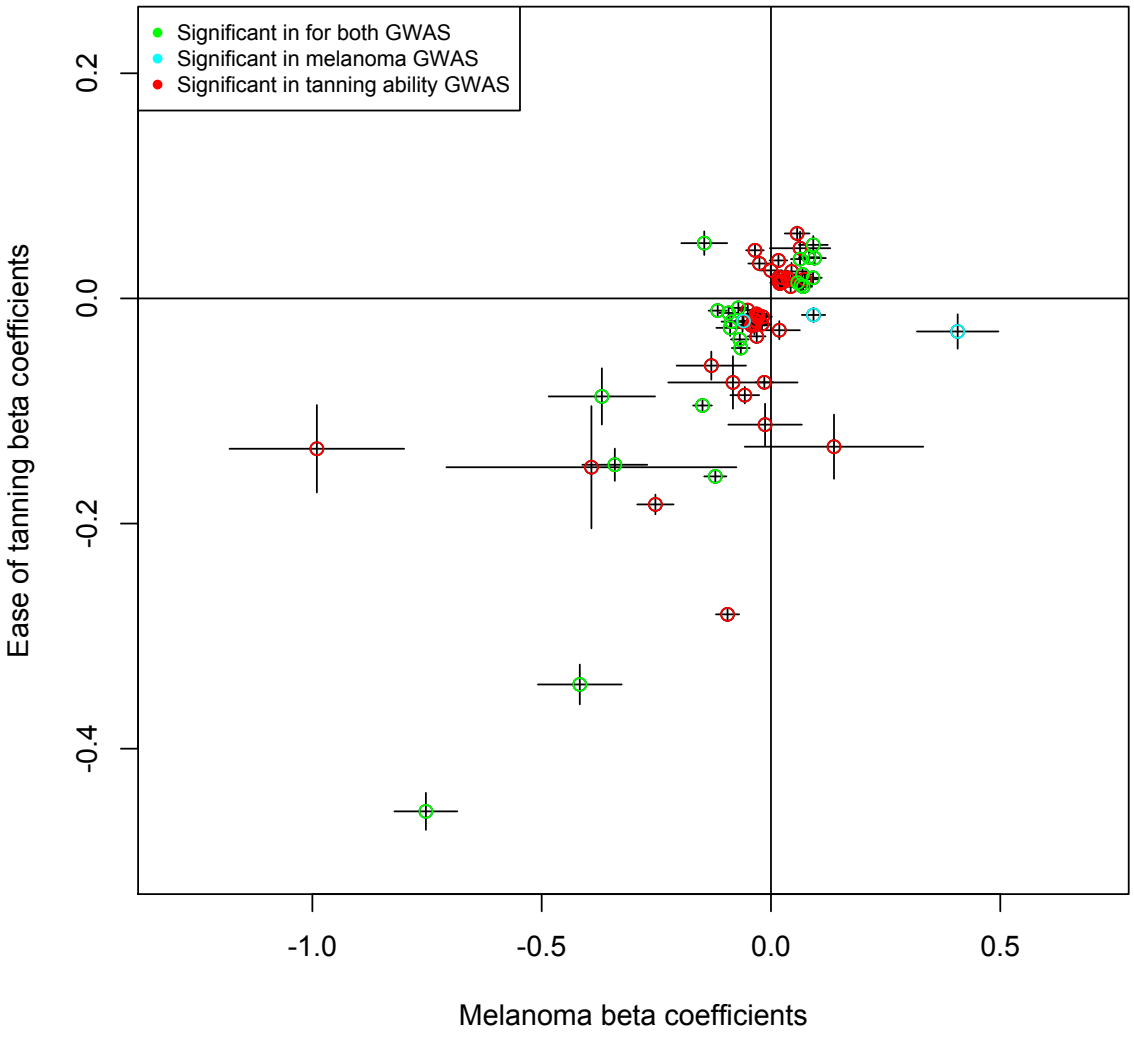


Figure 7.7: Lead SNP beta coefficients (log odds) for genome-wide significant loci of associated for SSWJA of a GWAS of tanning ability and a meta-analysis GWAS of melanoma risk. Lead SNPs are coloured to indicate whether they were significant in the contributing GWAS.

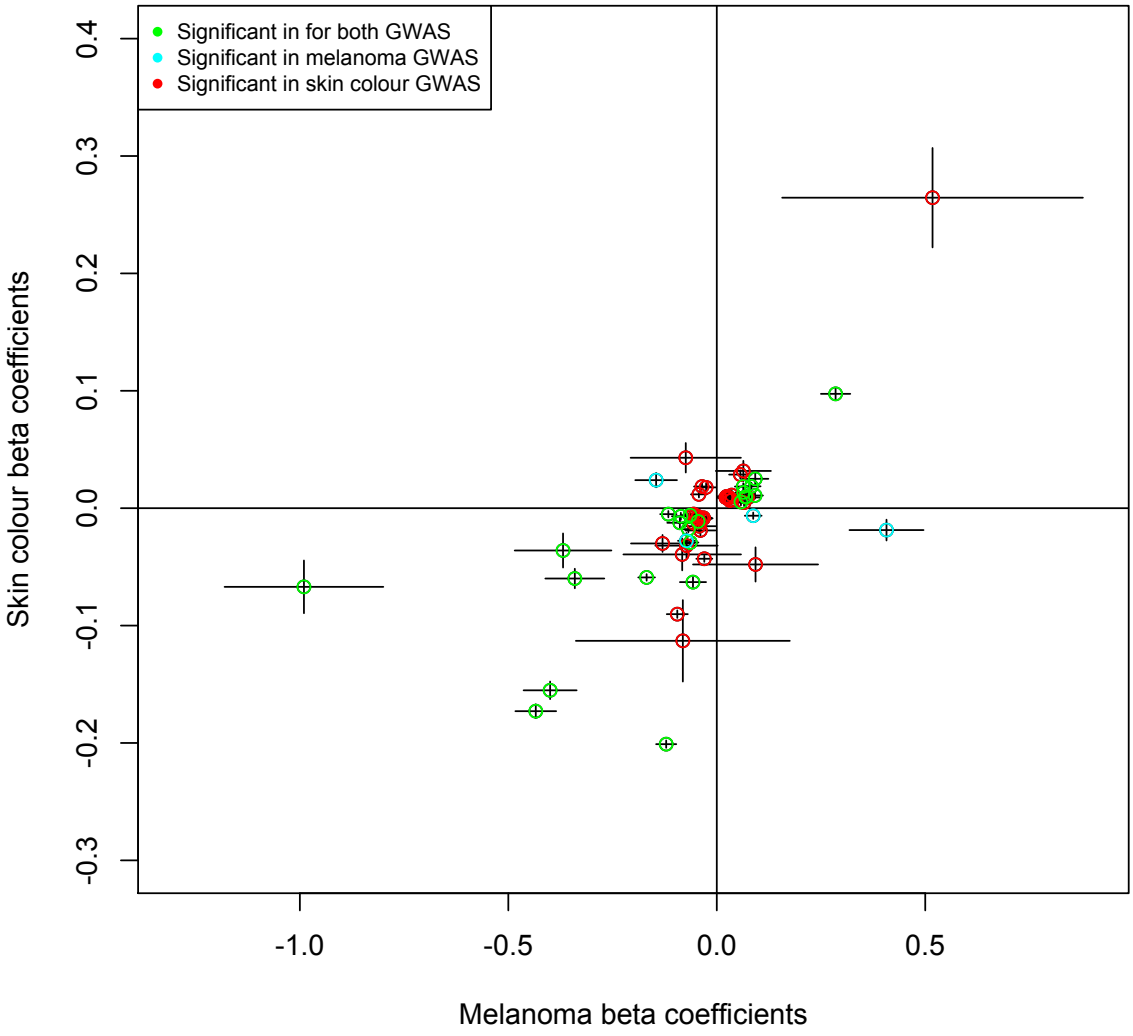


Figure 7.8: Lead SNP beta coefficients (log odds) for genome-wide significant loci of associated for SSWJA of a GWAS of skin colour and a meta-analysis GWAS of melanoma risk. Lead SNPs are coloured to indicate whether they were significant in the contributing GWAS.

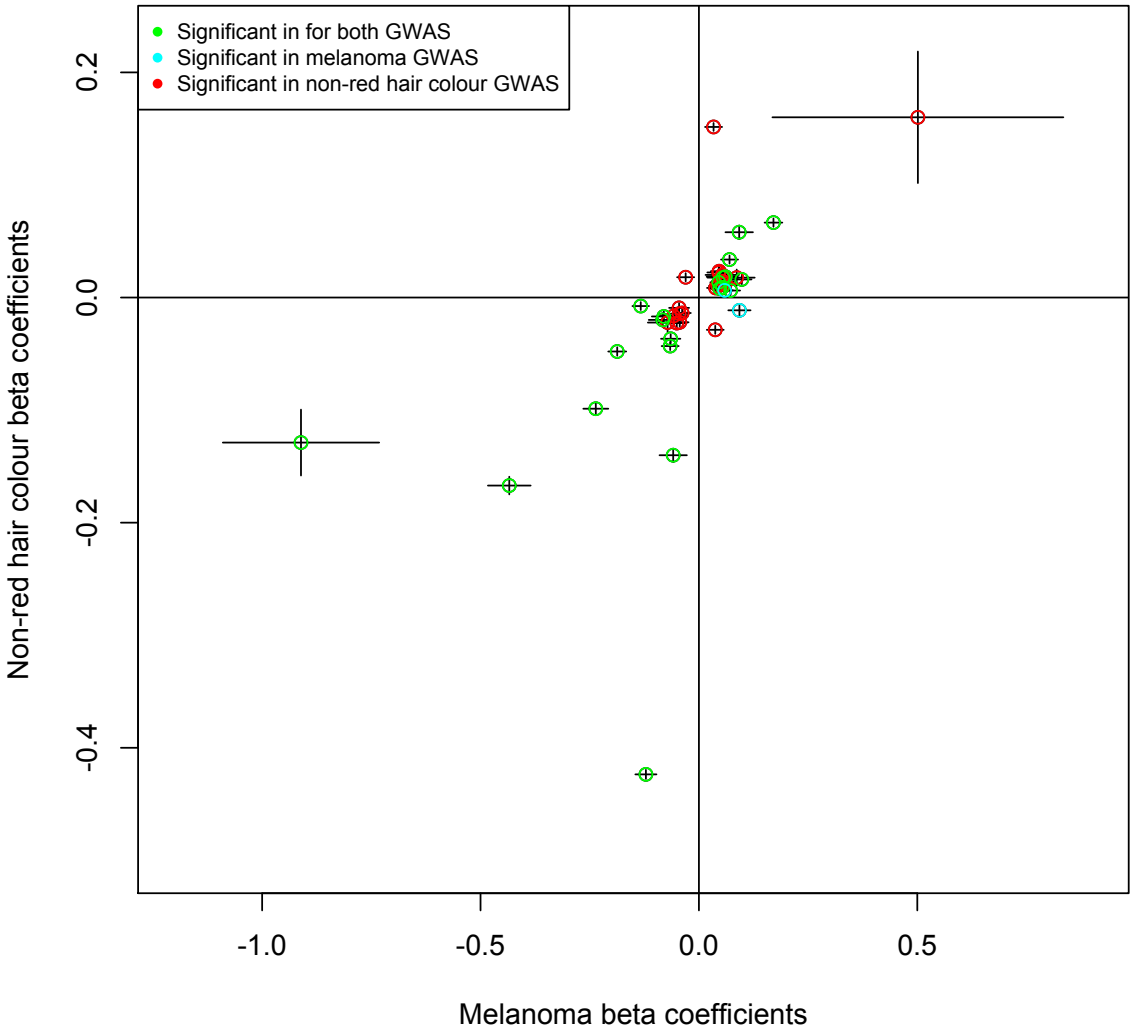


Figure 7.9: Lead SNP beta coefficients (log odds) for genome-wide significant loci of associated for SSWJA of a GWAS of non-red hair colour and a meta-analysis GWAS of melanoma risk. Lead SNPs are coloured to indicate whether they were significant in the contributing GWAS.

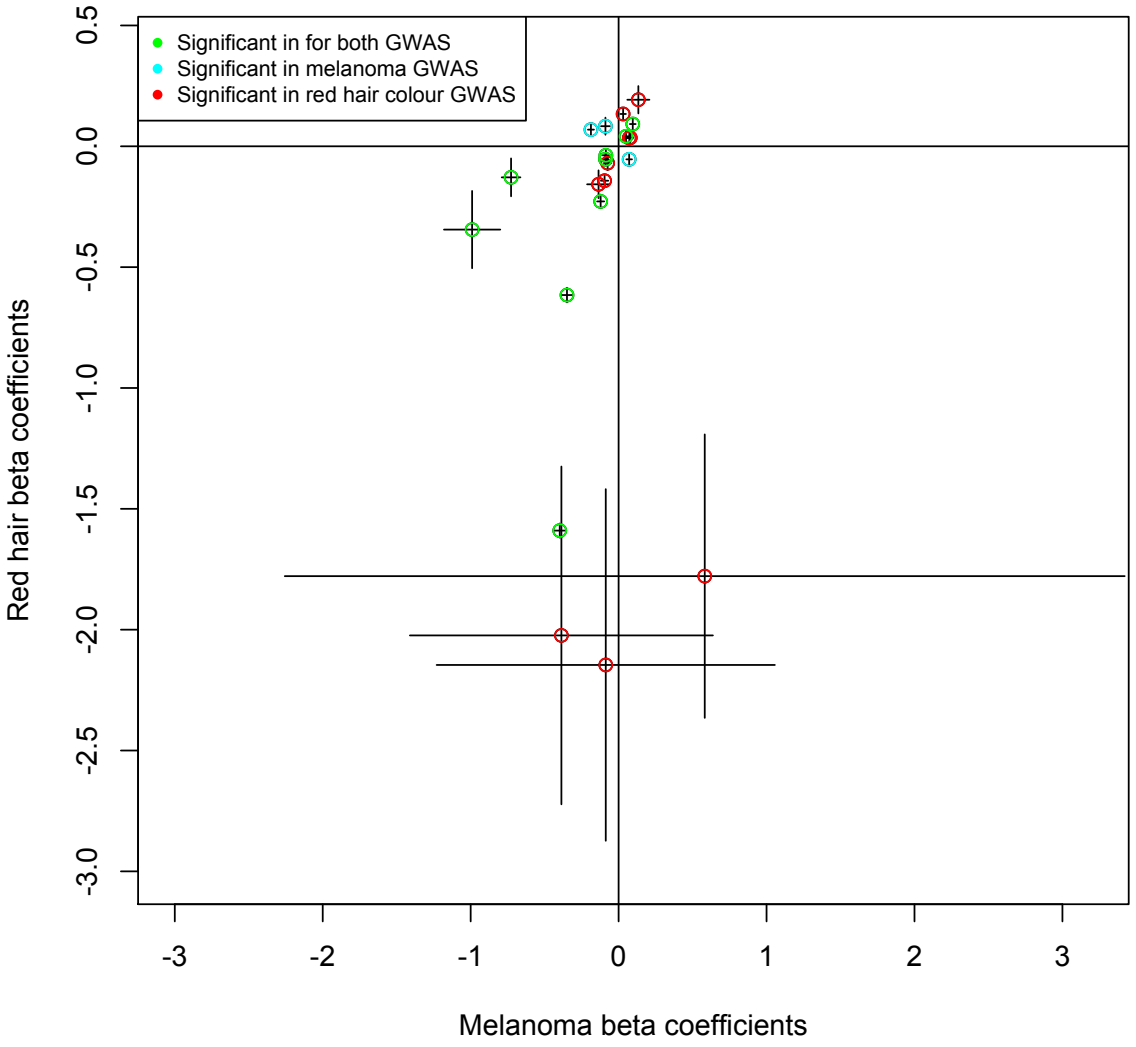


Figure 7.10: Lead SNP beta coefficients (log odds) for genome-wide significant loci of associated for SSWJA of a GWAS of red hair and a meta-analysis GWAS of melanoma risk. Lead SNPs are coloured to indicate whether they were significant in the contributing GWAS.

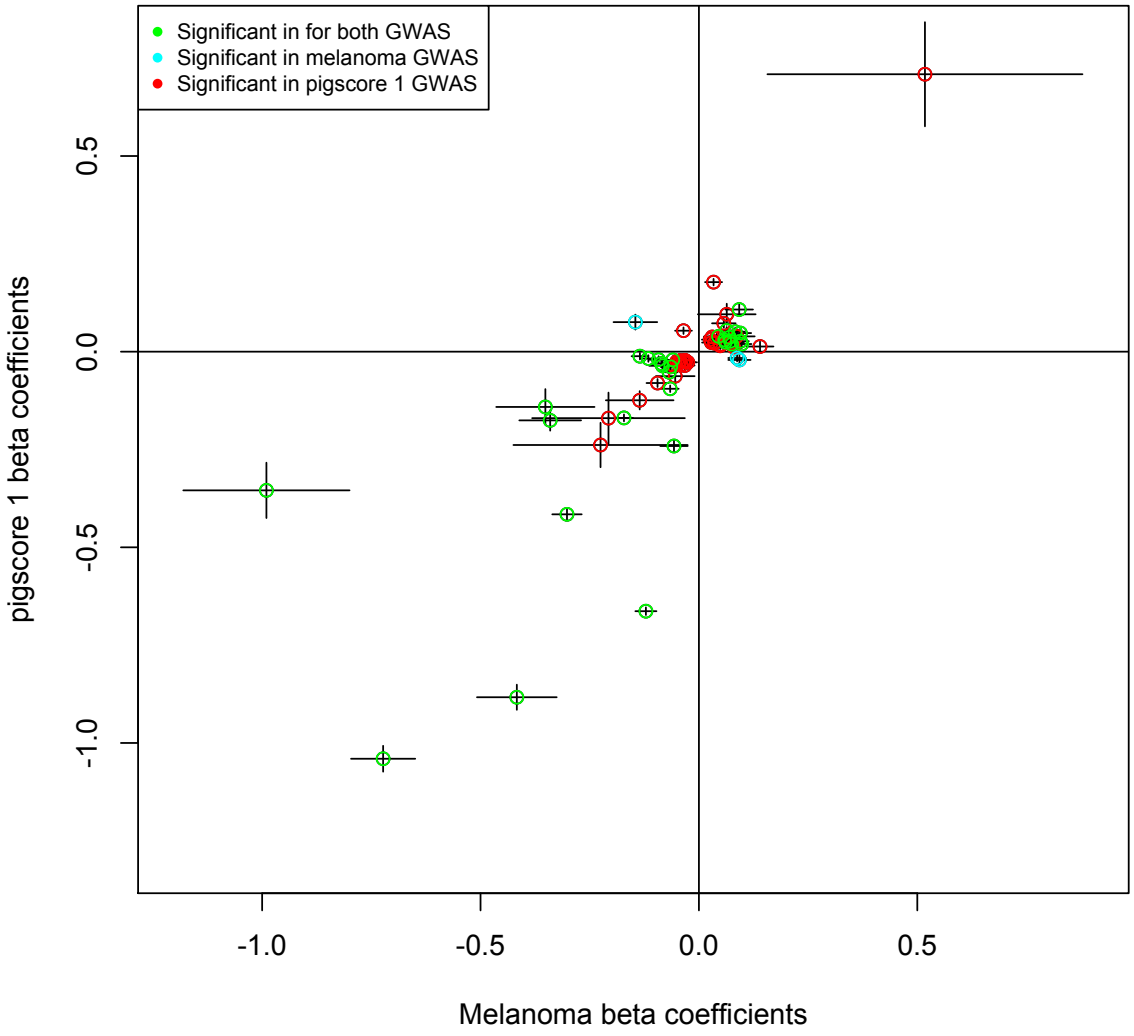


Figure 7.11: Lead SNP beta coefficients (log odds) for genome-wide significant loci of associated for SSWJA of a GWAS of overall pigmentation and a meta-analysis GWAS of melanoma risk. Lead SNPs are coloured to indicate whether they were significant in the contributing GWAS.

The overlap in significantly associated loci from each AP of melanoma and the five pigmentations trait was determined. Of the 37 loci found here to be associated with melanoma, 23 were found to be associated with pigscore 1 with none of these unique to the phenotype, 30 with ease of tanning with 3 unique to the phenotype, 22 with non-red hair colour with none of these unique to the phenotype, 11 with red hair with none of these unique to the phenetype, and 28 with skin colour with none unique to the phenotype (Figure 7.12)

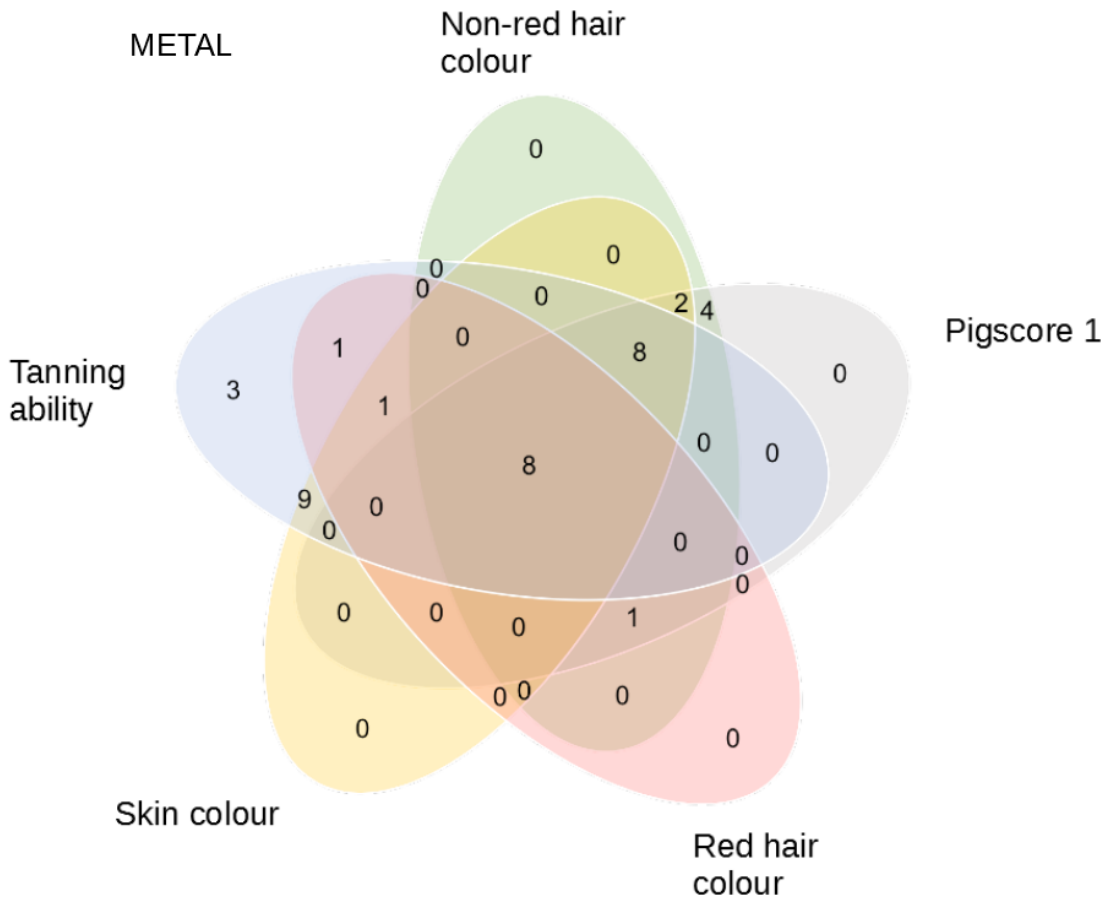


Figure 7.12: Overlap in number of loci for sample size weighted based joint-association of five pigmentary traits (non-red, red hair, tanning ability, skin colour and pigscore 1) and melanoma.

7.3.2 MTAG

When plotting the beta coefficients (log odds ratios) of the lead signal SNPs of identified loci identified through GCWJA, high correlation (Spearman's) across the pigmentation traits when compared with melanoma were observed indicating high overlap in the genetic architecture between the measured pigmentation traits (skin colour, tanning ability, non-red hair, red hair, and pigscore 1 with melanoma). The most strongly correlated beta coefficients were those of tanning ability and melanoma ($\rho = 0.78$) and the most weakly correlated between red hair and melanoma ($\rho = 0.4$) (Table 7.2).

	Ease	Skin	Non-red	Red	PIG1
Melanoma	0.78	-0.67	-0.55	0.40	-0.78

Table 7.2: Identified lead SNP beta coefficient correlation between five GCWJA: tanning ability and melanoma, skin colour and melanoma, non-red hair colour and melanoma, red hair colour and melanoma, and pigscore 1 and melanoma.

To determine which trait is the source of the evidence for those loci identified through the GCWJA, each locus' lead SNP was examined for the corresponding pigmentation and melanoma GWAS. For tanning ability and melanoma, 57 loci were identified, of which 22 had prior genome-wide significance for tanning ability, 16 had prior significance for melanoma, and 19 had prior significance in both conducted GWAS (Figure 7.13); for skin colour and melanoma, 58 loci were identified, of which 21 had prior genome-wide significance for skin colour, 18 had prior significance for melanoma, and 19 had prior significance in both conducted GWAS (Figure 7.14); for non-red hair colour and melanoma, 88 loci were identified, of which 48 had prior genome-wide significance for non-red hair colour, 27 had prior significance for melanoma, and 13 had prior significance in both conducted GWAS (Figure 7.15); for red hair and melanoma, 56 loci were identified, of which 8 loci had prior genome-wide significance for red hair, 42 had prior significance for melanoma, and 6 had prior significance in both conducted GWAS (Figure 7.16); and for pigscore 1 and melanoma, 70 loci were identified, of which 34 had prior genome-wide significance for pigscore 1, 15 had prior significance for melanoma, and 21 had prior significance in both conducted GWAS (Figure 7.17).

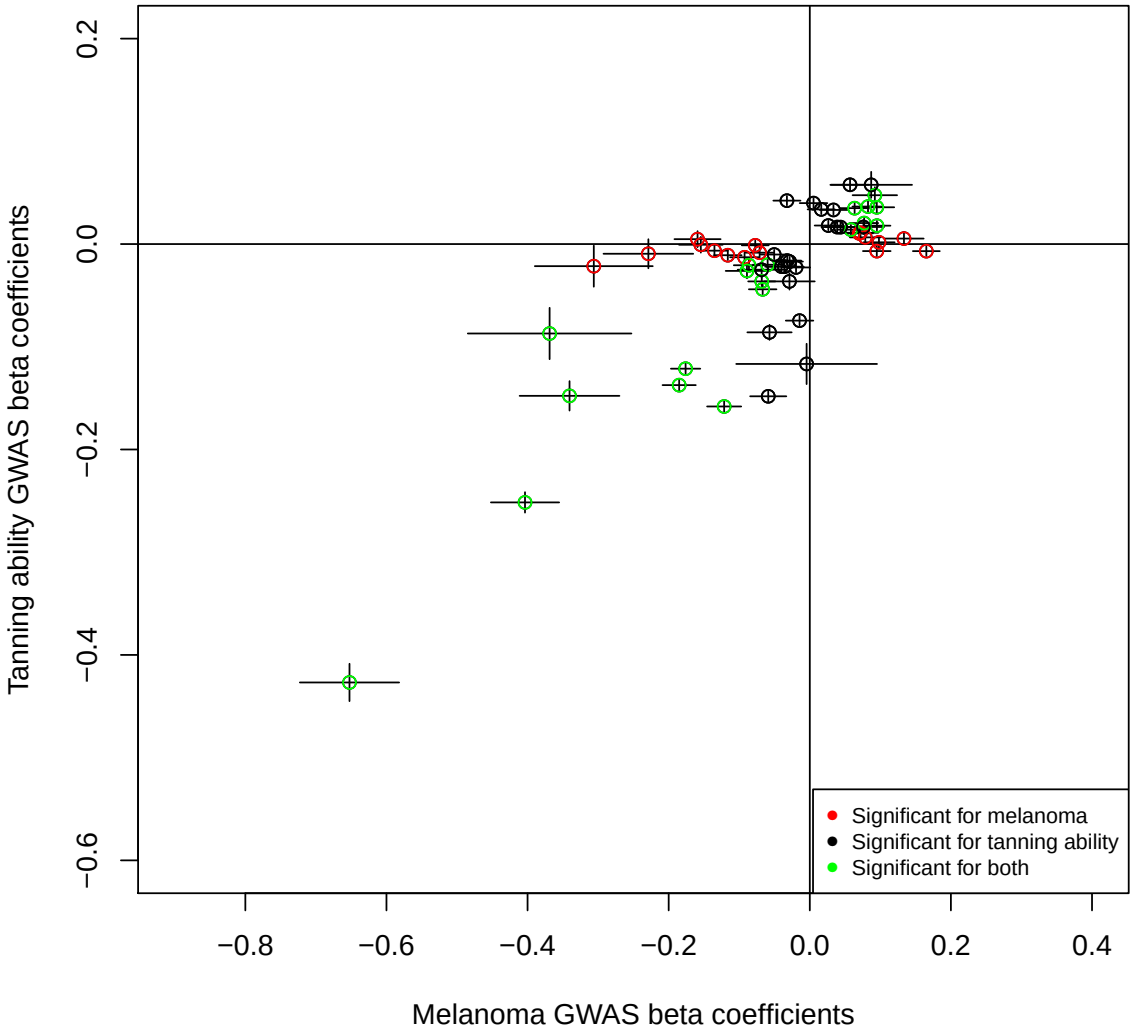


Figure 7.13: Lead SNP beta coefficients (log odds) for genome-wide significant loci of associated for GCWJA of a GWAS of tanning ability and a meta-analysis GWAS of melanoma risk. Lead SNPs are coloured to indicate whether they were significant in the contributing GWAS.

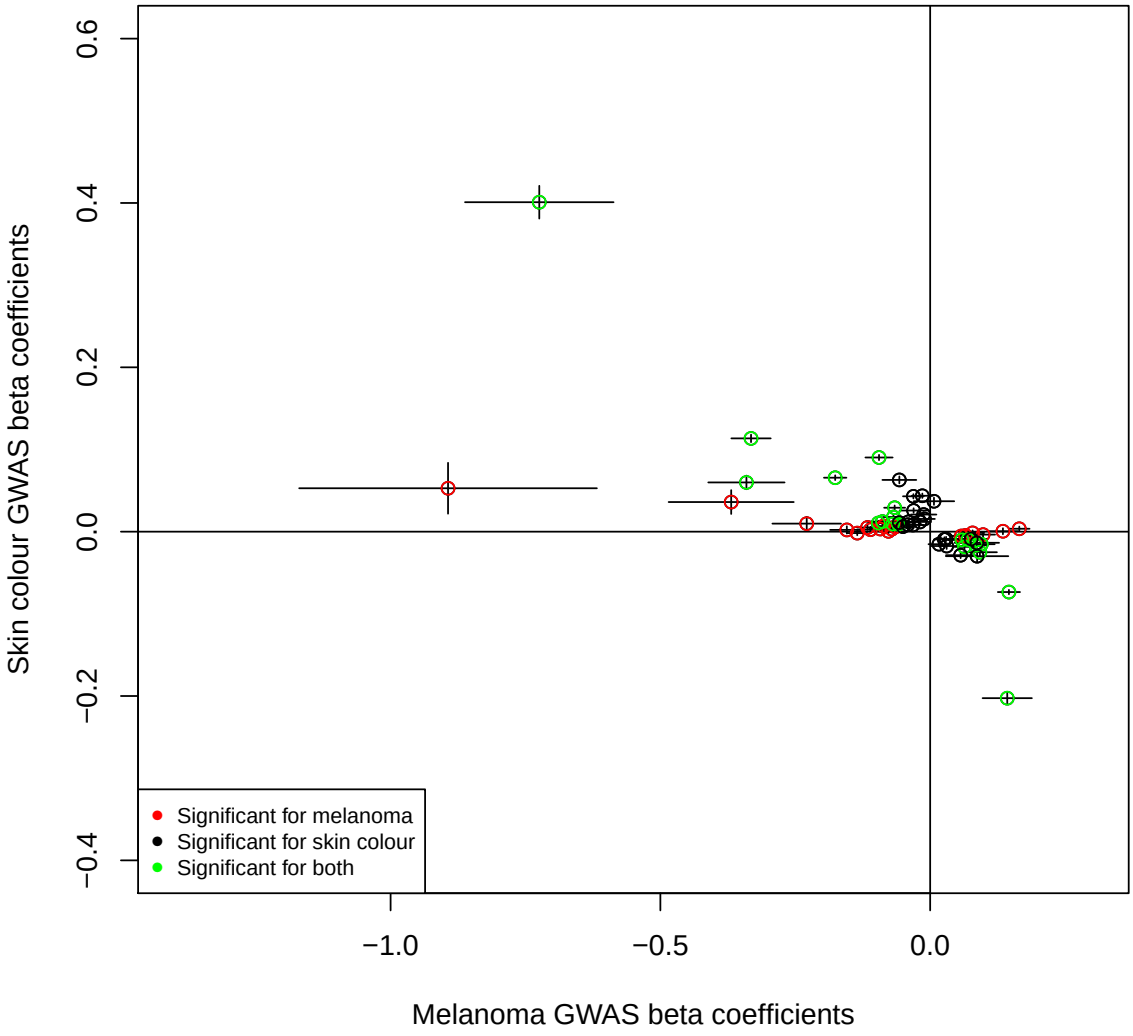


Figure 7.14: Lead SNP beta coefficients (log odds) for genome-wide significant loci of associated for GCWJA of a GWAS of skin colour and a meta-analysis GWAS of melanoma risk. Lead SNPs are coloured to indicate whether they were significant in the contributing GWAS.

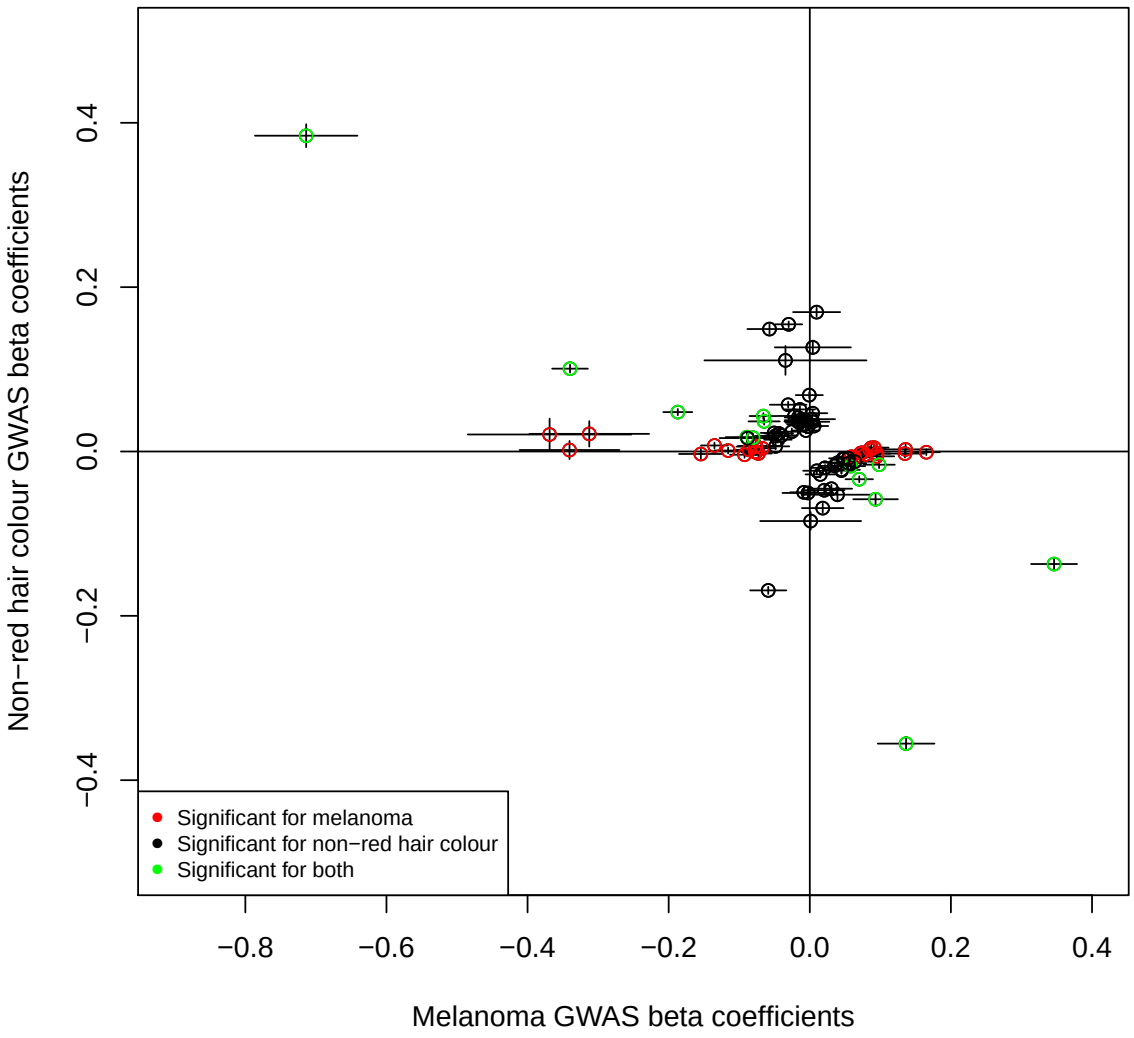


Figure 7.15: Lead SNP beta coefficients (log odds) for genome-wide significant loci of associated for GCWJA of a GWAS of non-red hair colour and a meta-analysis GWAS of melanoma risk. Lead SNPs are coloured to indicate whether they were significant in the contributing GWAS.

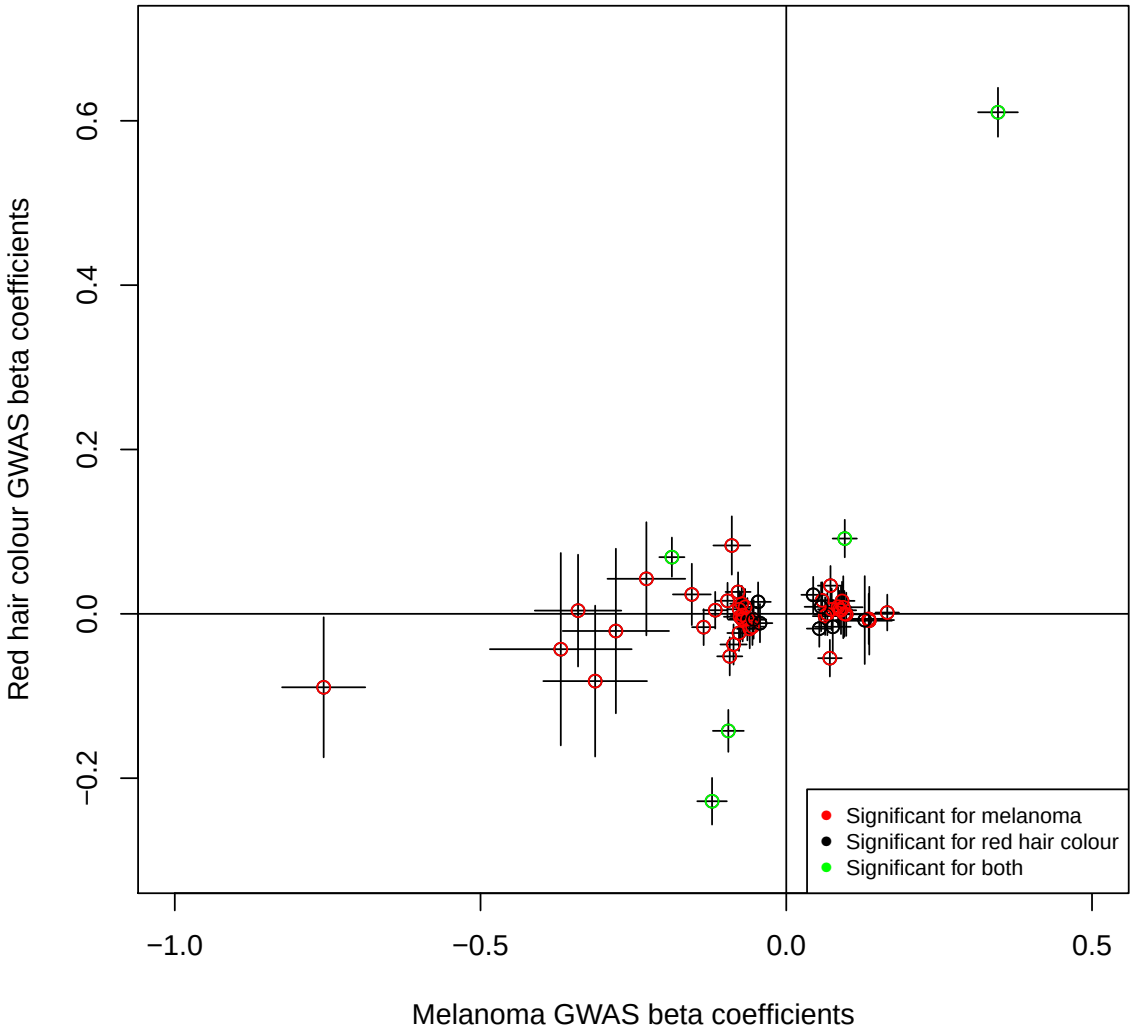


Figure 7.16: Lead SNP beta coefficients (log odds) for genome-wide significant loci of associated for GCWJA of a GWAS of red hair and a meta-analysis GWAS of melanoma risk. Lead SNPs are coloured to indicate whether they were significant in the contributing GWAS.

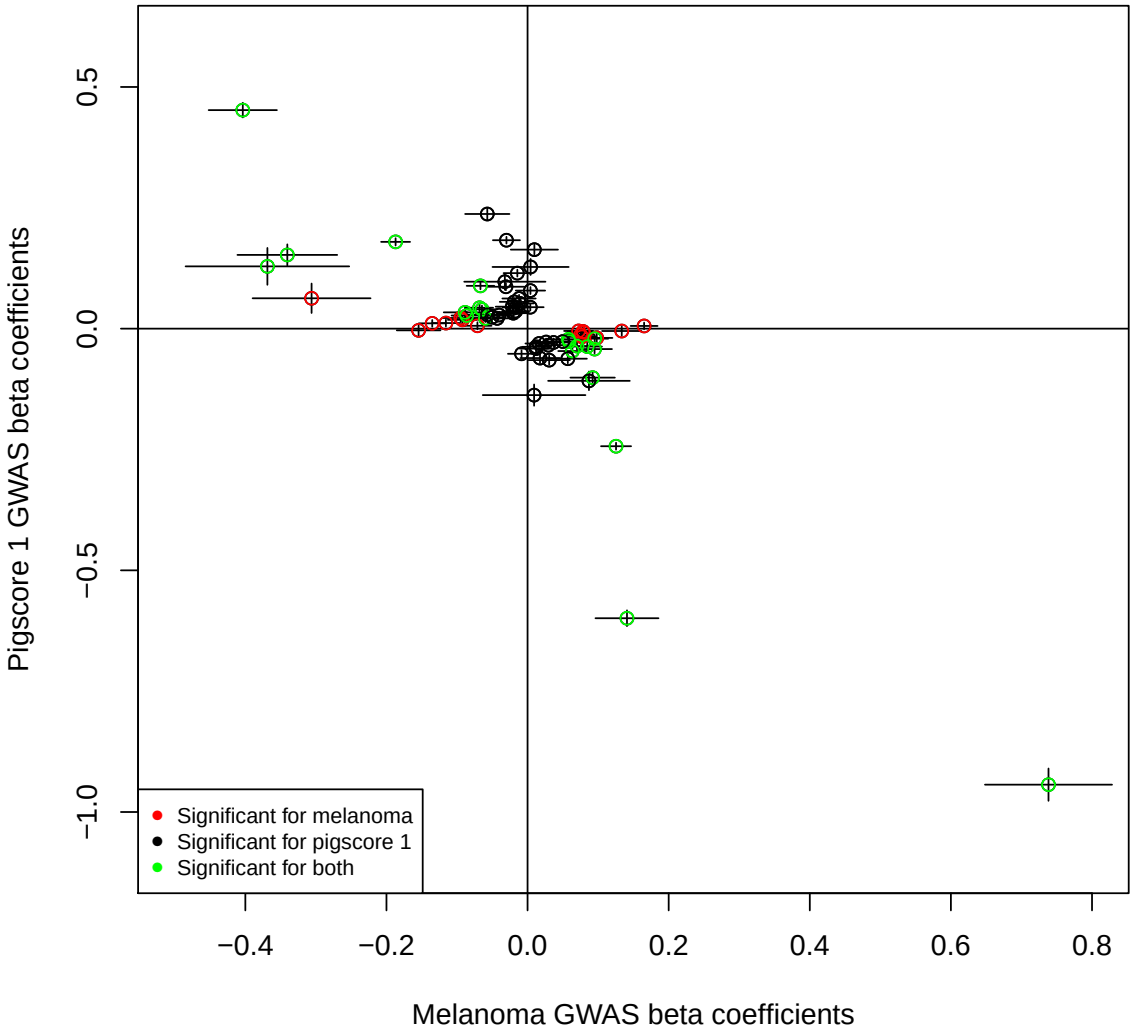


Figure 7.17: Lead SNP beta coefficients (log odds) for genome-wide significant loci of associated for GCWJA of a GWAS of overall pigmentation and a meta-analysis GWAS of melanoma risk. Lead SNPs are coloured to indicate whether they were significant in the contributing GWAS.

Determining the overlap in detected associations across the five pigmentary and melanoma joint-analysis, across all 133 loci found here to be associated with melanoma through genetic correlated weighted joint-analysis with pigmentation, 69 were found to be associated with pigscore 1 with 4 of these unique to the phenotype, 57 with ease of tanning with 8 unique to the phenotype, 88 with non-red hair colour with 28 of these unique to the phenotype, 56 with red hair with 11 of these unique to the phenetype, and 58 with skin colour with 4 unique to the phenotype (Figure 7.18)

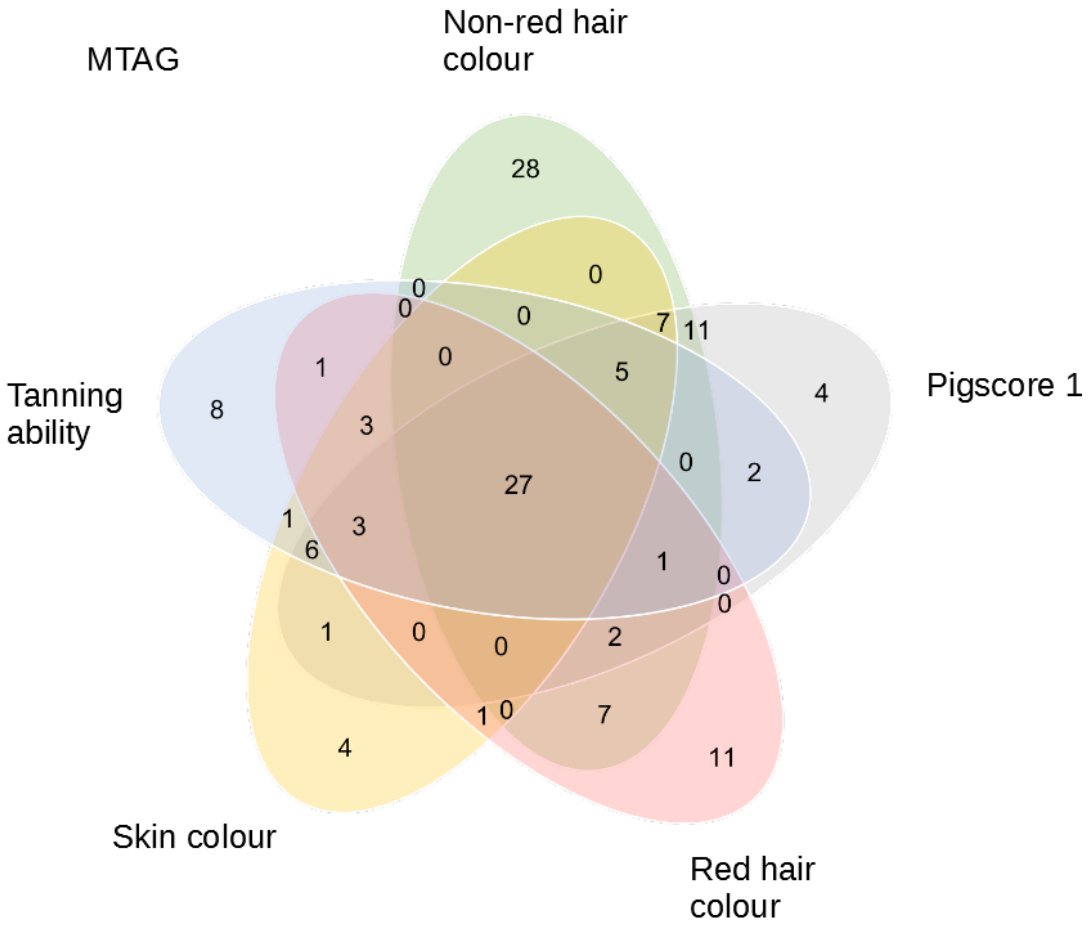


Figure 7.18: Overlap in loci for genetic correlation based joint-association of five pigmentary traits (non-red hair, red hair, tanning ability, skin colour and pigscore 1) and melanoma.

7.3.3 Combining identified loci from the APs and genetic correlated weight joint-analyses

Many polymorphisms significantly associated ($p < 5 \times 10^{-8}$) here with at least one of the pigmentation traits and melanoma from the APs or GCWJA have previously been reported as associated with melanoma in the literature and are recorded in a recent GWAS study (Landi et al. 2020) and in the GWAS catalog (Buniello et al. 2019). 215 independent variants across 139 loci were found here to be genome-wide significantly associated with at least one pigmentation measurement or combined pigmentation score and melanoma when conducting genetically correlated weighted joint-analyses and the APs (Figure 7.19, Appendix G.1); of these loci, 88 had not been previously identified as associated with melanoma (Appendix G.2). For these newly identified loci, across the two methods conducted, 80 loci were only identified through GCWJA, 5 loci were only identified through the APs, and 3 were identified in both the GCWJA and APs

(Tables 7.3, 7.4). Additionally, when considering previously reported melanoma association loci, six regions did not appear in the GCWJA or APs. Of these six melanoma-only loci, only one of these had previously been reported to be associated with a known risk pathway (naevus count), and the other five melanoma-only loci had not previously been associated with any melanoma risk pathway phenotype and also had a relatively low p-values ($p < 5 \times 10^{-9}$) in the melanoma meta-analysis GWAS.

Comparing the identified loci from the GCWJA and APs to other melanoma risk pathway phenotypes (naevus count and telomere length) found many previously known association with naevus and telomere length to be present in these results. 64 of the identified loci had prior reported associations with naevus count or telomere length, with 53 loci having had only a prior reported association with naevus count, three having had only a prior reported association with telomere length, and eight having had a prior reported association with naevus score and telomere length (Appendix G.2). Further comparing the novel identified loci from the GCWJA and APs to pathways not involved with melanoma risk found multiple loci that had prior reported associations with non-melanoma skin cancer: for both the GCWJ and APs rs6059655 was identified for tanning ability, skin colour, and non-red hair and has a prior known association with basal cell carcinoma and squamous cell carcinoma. This loci is on chromosome 20 in the *RALY* gene which is involved in pre-mRNA splicing and in embryonic development. Additionally evidence was observed between a shared pathways between these identified loci and vitiligo on chromosome 11 (rs11021172) through the *GRM5* gene.

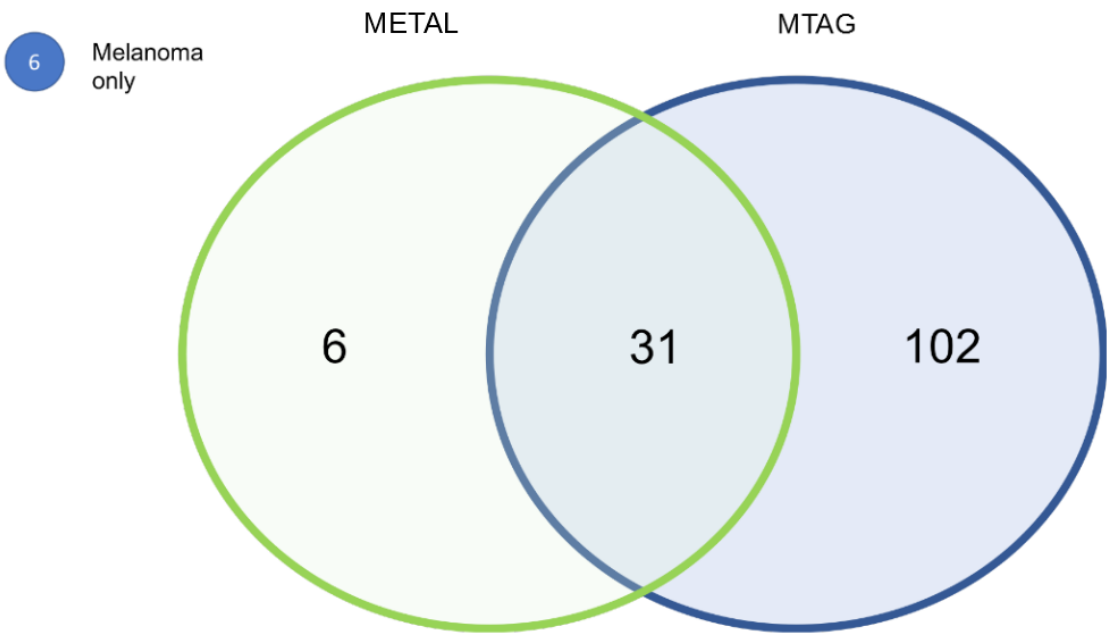


Figure 7.19: Venn diagram of the number of melanoma associated locus and the analysis they were identified in.

Fulltraits	count
GCWJA hair	27
GCWJA hair, GCWJA pigscore 1	11
GCWJA red	11
GCWJA hair, GCWJA red	8
AP ease, AP skin, AP hair, AP pigscore 1, GCWJA ease, GCWJA skin, GCWJA hair, GCWJA red, GCWJA pigscore 1	7
GCWJA ease	7
GCWJA ease, GCWJA skin, GCWJA pigscore 1	6
AP ease, AP skin, GCWJA ease, GCWJA skin, GCWJA hair, GCWJA red, GCWJA pigscore 1	6
GCWJA ease, GCWJA skin, GCWJA hair, GCWJA red, GCWJA pigscore 1	6
GCWJA skin, GCWJA hair, GCWJA pigscore 1	5
AP ease, AP skin, AP hair, AP red, AP pigscore 1, GCWJA ease, GCWJA skin, GCWJA hair, GCWJA red, GCWJA pigscore 1	4
GCWJA ease, GCWJA skin, GCWJA hair, GCWJA pigscore 1	4
GCWJA pigscore 1	4
GCWJA skin	3
AP ease, AP skin, AP hair, AP red, AP pigscore 1	3
GCWJA ease, GCWJA skin, GCWJA red	2
GCWJA hair, GCWJA red, GCWJA pigscore 1	2
AP hair, AP pigscore 1, GCWJA ease, GCWJA skin, GCWJA hair, GCWJA red, GCWJA pigscore 1	2
GCWJA ease, GCWJA pigscore 1	2
AP skin, AP hair, AP pigscore 1, GCWJA skin, GCWJA hair, GCWJA pigscore 1	2
GCWJA ease, GCWJA skin	1
AP ease	1
AP ease, AP skin, AP red, GCWJA ease, GCWJA skin, GCWJA hair, GCWJA red, GCWJA pigscore 1	1

AP ease, GCWJA ease	1
AP ease, AP skin, AP hair, AP red, AP pigscore 1, GCWJA ease, GCWJA skin, GCWJA pigscore 1	1
AP ease, AP skin, GCWJA ease, GCWJA skin, GCWJA red, GCWJA pigscore 1	1
AP ease, AP red, GCWJA ease, GCWJA skin, GCWJA hair, GCWJA red, GCWJA pigscore 1	1
GCWJA skin, GCWJA pigscore 1	1
AP ease, GCWJA ease, GCWJA skin, GCWJA red, GCWJA pigscore 1	1
GCWJA ease, GCWJA skin, GCWJA red, GCWJA pigscore 1	1
AP hair, AP red, AP pigscore 1, GCWJA ease, GCWJA red	1
AP ease, AP skin, GCWJA ease, GCWJA skin, GCWJA hair, GCWJA pigscore 1	1
AP ease, AP skin, AP hair, AP pigscore 1	1
GCWJA skin, GCWJA red	1
AP hair, AP pigscore 1	1
AP ease, AP skin, GCWJA ease, GCWJA skin, GCWJA red	1
AP hair, AP pigscore 1, GCWJA hair	1

Table 7.4: Table of the origin and frequency of genome-wide associated signals (5×10^{-8}) identified in an AP of melanoma and pigmentation, and a GCWJA of melanoma and pigmentation.

Fulltraits	count
GChair	26
GChair, GCps1	11
GCease, GCskin, GCps1	7
APease, APskin, APhair, APps1, GCease, GCskin, GChair, GCred, GCps1, mel	7
GCease	7
APease, APskin, GCease, GCskin, GChair, GCred, GCps1, mel	6
GChair, GCred, mel	6
GCred, mel	6
GCease, GCskin, GChair, GCred, GCps1, mel	6
mel	6
GCred	5
GCskin, GChair, GCps1	5
APease, APskin, APhair, APred, APps1, GCease, GCskin, GChair, GCred, GCps1, mel	4
GCease, GCskin, GChair, GCps1	4
GCps1	4
GCskin	3
APease, APskin, APhair, APred, APps1	3
APhair, APps1, GCease, GCskin, GChair, GCred, GCps1, mel	2
GCease, GCps1	2
GCease, GCskin	1
APease, mel	1
GCease, GCskin, GCred	1
APease, APskin, APred, GCease, GCskin, GChair, GCred, GCps1, mel	1
GChair, GCred	1
GChair, GCred, GCps1, mel	1
APease, GCease	1
APease, APskin, APhair, APred, APps1, GChair, GCred, mel	1
APease, APskin, GCease, GCskin, GCred, GCps1, mel	1
APease, APred, GCease, GCskin, GChair, GCred, GCps1, mel	1
GCskin, GCps1	1
APskin, APhair, APps1, GCskin, GChair, GCps1, mel	1
APease, GCease, GCskin, GCred, GCps1, mel	1
GCease, GCskin, GCred, GCps1	1
APhair, APred, APps1, GCease, GCred, mel	1
APease, APskin, GCease, GCskin, GChair, GCps1	1
APskin, APhair, APps1, GCskin, GChair, GCps1	1
GChair, GCred, GCps1	1
APease, APskin, APhair, APps1	1
GCease, GCskin, GCred, mel	1
GChair, mel	1
GCskin, GCred, mel	1
APhair, APps1	1
APease, APskin, GCease, GCskin, GCred, mel	1
APhair, APps1, GChair, mel	1

Table 7.3: Table of the origin and frequency of genome-wide associated signals (5×10^{-8}) identified in a melanoma GWAS, AP of melanoma and pigmentation, and a GCWJA of melanoma and pigmentation. GC is GCWJA. ps1 is pigscore 1

7.4 Discussion

Here is reported the results of two joint-analysis methods that utilise the power of pigmentation GWAS summary statistics to supplement melanoma GWAS results and detect additional associated loci. The first method uses a targeted AP that separately identifies regions of joint-association through GWAS-PW and conducts a sample-size weighted joint-analysis for each pigmentation GWAS summary statistics with a melanoma meta-analysis GWAS summary statistics to ensure that any locus identified both reaches genome-wide significance in the combined melanoma/pigmentation analysis and is associated with both traits. The second method conducts a GCWJA that inflates/deflates regions of association for each of the pigmentation GWAS summary statistics with the melanoma meta-analysis GWAS summary statistics. Overall, these two methods identify 139 independent associated loci for melanoma; of these 139 independently associated loci, 88 were novel. The discovery of many new loci improves our understanding of how the genetic architecture of melanoma relates to pigmentation, and highlights the possibility that large gains in our understanding of the genetic architecture of melanoma may still be achieved with larger melanoma datasets and by supplementing GWAS with information from known risk pathway phenotypes.

For the genetic correlation joint-analyses, the entire genome could be considered during analysis, and thus it was possible to highlight a greater number of regions potentially associated with melanoma. Typically a higher number of candidate loci would ordinarily provide a greater insight into a trait's genetic architecture. However in the case of melanoma and pigmentation, it is well-established that naevus count and telomere length are known genetic risk pathways concurrent with pigmentation. As these three risk pathway phenotypes act simultaneously on melanoma risk, the trait-specific effect of each identified locus and its effect on melanoma can vary. This is partly due to of pleiotropy and partly because the three phenotypes may not be entirely independent (e.g. telomere length may influence naevus count). This may result in the trait-specific effect 'over influencing' the melanoma effect at these locus influenced by more than one risk pathway trait as the other risk other trait-specific effects might not be accounted for. Conversely, this three risk pathway characteristic will also affect the polygenic/more powered phenotype (in this case pigmentation), by creating a scenario where loci that are not associated with pigmentation but are with melanoma could be inflated for in the pigmentation

results, due to the higher correlation present between pigmentation and melanoma, leading to more uncertain positive associations with little to no evidence/literature to back up the findings.

Compared to the GCWJA method, the AP uses a more conservative approach by only selecting regions that show joint association between pigmentation and melanoma through GWAS-PW. Although more conservative, a similar issues with 'complicated' regions that are associated with melanoma, pigmentation and other melanoma risk pathway phenotypes is present, and may result in these region being missed or bias the effect size on melanoma. For example, if a locus only influences pigmentation and melanoma, the allele that leads to paler skin will also lead to an increase in the risk of melanoma and the ratio of the betas are fairly constant. However, if the 'pale skin' allele is also associated with naevus count (having a different effect direction and is more strongly associated than pigmentation), the 'pale skin' allele may appear protective for melanoma.

In addition to potential biases that are present in the GCWJA and APs, one of the primary concerns when combining pigmentary and melanoma GWAS summary statistics was with pleiotropic signals between pigmentation and the other two genetic risk pathways (naevus count and telomere length). In particular loci influencing both pigmentation and naevus count, as polymorphisms that shared association between these two traits did not necessarily share a common direction of effect. More specifically of the loci identified through the GCWJA and the APs of melanoma and pigmentation, 61 shared associations with naevus count. For these 61 loci, many influenced the two pathway phenotypes in opposite directions in terms of impact on melanoma risk. So the allele associated with an increased number of naevi (and so increased melanoma risk) was also associated with darker skin (and so decreased melanoma risk). Likely the extent of bias is limited due to the polygenic nature of pigmentation and naevus count, and by the fact this overlap does not seem to bare any effect on the big six pigmentation hits – likely due to the magnitude of their effect on pigmentation – which carry the vast majority of risk when considering melanoma. The loci identified to be effected by this bias have a far more subtle effect size, and accounting for this in further research, such as predicting melanoma risk with pigmentation and PRS, would have minimal effects on findings.

Previous research that has utilised a targeted approach for melanoma and naevus data has also succeeded in uncovering potential new melanoma variants (Landi et al. 2020). Using a similar methodology to that outlined in this chapter, 66 loci were identified through analysis, with 11 being novel. 2 of these 11 novel loci were also identified through the pigmentation joint-analyses in this chapter. Combining these naevus count and melanoma results with the pigmentation and melanoma risk from this chapter, 97 novel candidate variants have been highlighted as association with melanoma risk. This increase in variant detection from utilising complimentary data indicates the potential power in joint-analysis. However, although new potential variants have been identified through this methodology, ultimately there is currently no independent GWAS literature to validate these findings. This is because meta-analysis incorporate as many GWAS as possible so replication only happens if new data have been genotyped since the study began or individual SNPs can be genotyped just for replication. Also, studies are becoming so big that a single study does not have the power to replicate the small effects they identify. Furthermore, evidence for association is provided through a wider-literature search of biological processes, which has the potential to introduce confirmation bias into the process, as previous research into this type of literature search has shown that spurious links are in-distinguishable from genuine links when explaining identified associations from GWAS results. As this is the case, validation of these potential variants can also be supported indirectly by more functional studies, such as TWAS results, which integrates GWAS and eQTL data, ideally from melanocytes or skin tissue samples and assessing overlap; or to simply compare identified potential variants with future melanoma GWAS results.

One of the most important aspects of pigmentation data in comparison to melanoma, naevus count, and telomere length data is its abundance, as it is routinely measured in many large studies. For this reason utilizing pigmentation GWAS data is an ideal way to increase power for melanoma GWAS. However, when considering the other two genetic risk pathways, current naevus (Duffy, Gu Zhu, et al. 2018) and telomere length (C. Li et al. 2020) GWAS have much smaller sample sizes (naevus: 52,506 samples, telomere: 78,592 samples) and have identified five and 42 associated loci respectively. This limitation limits the potential for joint-analysis with melanoma. However, given the importance of these pathways for melanoma risk and their high heritability (G. Zhu et al. 1999; Wachsmuth et al. 2001; Slagboom et al. 1994; Vasa-Nicotera et al. 2005; Njajou et al. 2007; Bakaysa et al. 2007) increased sample sizes for GWAS of naevus

count and telomere length will undoubtedly uncover many more associated loci and improve their utility when jointly analysed with melanoma GWAS. Jointly analysing melanoma with pigmentation, naevus count and telomere length has the potential to be a very powerful way of identifying melanoma-associated loci.

Chapter 8

Penetrance and Epistasis estimates for *MC1R* variants on pigmentation

8.1 Introduction

Different pigmentation genes have been associated with colour diversity in external tissues on the body, such as skin, hair, and eye colour. On the basis of the literature, the six top candidates for contributors to the variation in colour of these external human tissues are: *MC1R*, *ASIP*, *HERC2/OCA2*, *IRF4*, *SLC45A2*, and *TYR*. Although it is well-established the influence these genes have on pigmentation, previous research into these interactions have primarily been focused on animal furs/wools, where strong indicators of epistasis effects for *MC1R* on other genes have been found (Hepp et al. 2016). When considering human pigmentation, research has been limited, with only a handful of studies investigating this (Zorina-Lichtenwalter et al. 2019; Morgan et al. 2018), due to the need for extensive sample size, in previous genetic studies of pigmentation.

In particular, *MC1R* is a seven-transmembrane G-protein coupled receptor that activates adenylate cyclase to elevate cAMP levels upon hormone stimulation. These increased cAMP levels play a role in the melanogenic pathways, by helping to increase the *L*-cysteine present in the melanosome, which interacts with dopaquinone to form 5-S- and 2-S-cysteinyldopa. The oxidation of these cysteinyldopa isomers with dopaquinone leads to the production of pheomelanin – the yellow/red coloured melanin. Gene expression studies focusing on eumelanogenesis have

shown increased expression of tyrosinase (*TYR*), tyrosinase-related protein 1 (*TYRP1*) and tyrosinase-related protein 2 (*DCT*) and the P-protein genes that are directed through *MC1R* stimulation (García-Borrón et al. 2014). However, in the absence of a strong *MC1R* signal lower levels of *TYR* and P-protein together with the absence of *TYRP1* and *DCT* may be enough to switch synthesis to pheomelanin. This observation provides biological evidence that *MC1R* could be influencing the impact of these other pigmentary genes.

8.1.1 Originality of research

Epistasis in this context is the interaction between multiple genes which influence pigmentation. Detecting epistasis in complex traits is challenging. Generally and by their nature, epistatic effects require considerably larger sample sizes to detect than direct main effects. However, pigmentation, and more specifically the red hair phenotype, has the potential to uncover epistatic effects due to the substantial effects of specific genetic variants have on the phenotype, and thus a more tractable model may be obtainable to detect epistasis. In addition to this, the increase in sample-size power provided by the UK Biobank also increases the detection power for epistasis. As previous pigmentary datasets have not had this increased sample-size power, utilising UK Biobank offers a unique scenario to detect novel epistatic interactions.

8.1.2 Chapter Aims & Objectives

Aims

I. Investigate penetrance of *MC1R* coding variants in red hair, non-red hair, tanning ability, and skin colour. By penetrance, I mean the frequency that persons with specific genotypes self-report red hair.

II. Investigate epistasis between known *MC1R* coding variants and other red hair-associated variants.

III. Investigate epistasis between known *MC1R* coding variants and identified associated variants in tanning ability, skin colour, and non-red hair.

Objectives

1. Create a penetrance matrix for the *MC1R* coding variant for red hair case/control status

2. Investigate whether *MC1R* coding variants influence other known red hair associated regions by including an interaction term in a logistic regression model

3. Create a combined *MC1R* coding variant factor and determine whether this overall *MC1R* factor influences other known red hair associated regions by including an interaction term in a logistic regression model

8.1.3 Rationale to achieve aims & objectives

This chapter explores the penetrance and epistatic effects of the *MC1R* coding variants on red hair, non-red hair, skin colour, and tanning ability. Firstly, by calculating the penetrance of *MC1R* coding variants for red hair by simply calculating the proportion of red hair/non-red hair participants with the number of *MC1R* coding variants (0,1,2). Calculating this penetrance on UK Biobank participants will be the largest sampled-sized study of its kind, and aims to provide the most accurate penetrance mapping of *MC1R* coding variants on the red hair phenotype. Secondly, I investigate epistasis for *MC1R* coding variants and other loci using two methodologies for red hair. The first method, considers all *MC1R* coding variants separately and detects individual epistasic effects based on each coding variants; and the second considering a combined approach where participants were scored by the number of *MC1R* coding variants they possess (0, 1, or 2). This combined methodology is then extended for other pigmentation traits (non-red hair, skin colour, and tanning ability). By modelling epistasis on UK Biobank participant phenotypes, it will provide the largest sample-sized study that investigates the epistasis effects *MC1R* has on other genes when relating to human pigmentation.

8.2 Methods

All analysis was conducted in PLINK 2.0 (Purcell et al. 2007), Cassi software (https://www.staff.ncl.ac.uk/richard.howey/cassi/index.html), GCTA (J. Yang et al. 2011), and R (R Core Team 2017) using the UK Biobank (Sudlow et al. 2015).

8.2.1 Penetrance of *MC1R* in pigmentation

To calculate estimates for the penetrance of *MC1R* coding variants across the UK Biobank pigmentation variables, four penetrance models were conducted for each of the pigmentary traits (red hair, non-red hair, tanning ability, and skin colour) carried out on the UK Biobank cohort individuals. To minimise population stratification, the analyses were conducted on individuals of European descent (QC methods 2.3.4). Due to the pigmentary variables in the UK Biobank being self-reported, individuals who responded: 'don't know', 'prefer not to say', or did not answer for the respective pigmentary penetrance model were dropped from analysis. After removing these participants, 370,260 individuals were retained for the skin colour penetrance model, 367,229 for the ease of tanning penetrance model, 352,662 for the non-red hair penetrance model, and 103,111 for the red hair penetrance model. Pigmentary traits that were not already binary (tanning ability, skin colour and non-red hair colour) were dichtomised based on median values for each trait's sample size and then ordered from dark to light pigmentation: hair group one: Black, dark brown, group two: blond and light brown hair; skin colour: group one black, brown, dark olive, light olive, group two: very fair and fair; tanning ability: get very tanned. moderately tanned, group two: occasionally tanned, never tan only burn.

In this analysis, I define penetrance to be the percentage of persons with a specific genotype who have the phenotype under discussion. To facilitate analysis, the result of Zorina-Lichtenwalter et al. 2019 (doi: 10.1093/hmg/ddx018) shows that the *MC1R* coding variants occur on distinct haplotypes. To obtain penetrance estimates for each combination of *MC1R* coding variant SNPs (i.e. genotype), the thirteen identified *MC1R* coding variant SNPs were extracted from each participant in UK Biobank, and scored based on the number of effect alleles present for each SNP (0,1,2). I should note that for this analysis, only data on SNPs that were available were selected so that the few, very rare frameshift mutations were excluded e.g. rs555179612 whose MAF is 0.002. Participants ID, and the thirteen *MC1R* coding variant scores were combined with the

phenotype data (see previous paragraph), to create a penetrance dataset for each pigmentary trait that consisted of: (1) Participants ID; (2) thirteen individual scores for each *MC1R* coding variant; (3) the relative pigmentation trait group status (coded 0 or 1). For each combination of *MC1R* coding variants, a penetrance estimate was calculated by dividing the total number of participants with that variant coding combination by the number of participants with that variant coding combination and were phenotype positive (red hair), these penetrance estimates were then compiled into a variant matrix for comparison. Penetrance estimates were then compared with previous estimates calculating from prior-studies.

Variant rs-number	**Effect allele**	**Wild-type allele**	**Classification**
rs3212379	T	C	Strong-R
rs1805005	T	G	Weak-r
rs34474212	C	T	Strong-R
rs1805006	A	C	Strong-R
rs2228479	A	C	Weak-r
rs34158934	T	C	Strong-R
rs11547464	A	G	Strong-R
rs1805007	T	C	Strong-R
rs201326893	A	C	Strong-R
rs1110400	C	T	Weak-r
rs1805008	T	C	Strong-R
rs885479	A	G	Weak-r
rs555179612	TC		Strong-R
rs200000734	T	C	Strong-R
rs1805009	C	G	Strong-R
rs368507952	A	G	Strong-R

Table 8.1: Table of *MC1R* coding variants used in analysis with effect allele and wild-type allele used as the baseline in penetrance and epistasis models, and classification of Strong-R or weak-r

8.2.2 Red hair GWAS sub-grouped on number of coding alleles zero or two

To highlight the impact *MC1R* coding variants have on other regions of the genome for red hair, two allele-based sub-grouped GWAS were conducted. Firstly, UK Biobank individuals who responded: 'don't know', 'prefer not to say', or did not answer for their hair colour response were dropped from both GWAS analyses. After removing these participants and conducting standard QC (QC methods 2.3.4), individuals were grouped into two groups: those who had no *MC1R* coding variants, and those who had two *MC1R* coding variants, leaving 51,244 individuals for the red hair GWAS based on individuals with no *MC1R* coding variants, and 15,222 individuals

for the red hair GWAS based on individuals with two *MC1R* coding variants. To minimise population stratification, the analyses were conducted on individuals of European descent and adjusted for the first ten genetic principal components, as well as for the chip individuals were genotyped on; additionally for the two *MC1R* allele group, the thirteen known *MC1R* variants were adjusted for in analysis. Red hair was considered as a binary phenotype and used as a phenotypic variable in logistic regressed-based GWAS analyses.

For both analyses, to identify regions of association all SNPs with a p-value threshold of $p < 5 \times 10^{-8}$ were extract, and compared against all other extracted SNPs. Independent associated regions were defined through condition analysis using GCTA software, and declared SNPs with a collinearity threshold of $r^2 > 0.05$ and within a maximum distance window of 500kb as a locus.

8.2.3 Epistasis of *MC1R* coding alleles in red hair

To identify epistasis effects in the thirteen *MC1R* coding variants on other genomic regions associated with red hair, logistic regression and likelihood ratio tests were conducted on every combination of *MC1R* variant and previously identified red hair signal. To minimise population stratification, the analyses were conducted on individuals of European descent and adjusted for the first ten genetic principal components, sex, and the chip individuals were genotyped on. Due to the pigmentary variables in the UK Biobank being self-reported, individuals who responded: 'don't know', 'prefer not to say', or did not answer for their hair colour response were dropped from analysis. After removing these participants, 371,035 individuals were retained for analyses.

To test for epistasis each signal identified to be associated with red hair through conditional analysis (5.2.5) was extracted with a 1.5MB window from UK Biobank genetyped data, and merged with individuals phenotype (case control red hair) information. For each combination of the thirteen *MC1R* coding variant and previously identified red hair signal, two logistic regression models consisting of red hair case/control status as the independent variable, and the *MC1R* coding variant and signal as dependent variables with one model having an interaction term between the two dependent variables, and one without. Additionally, SNP alleles that co-occur on the same haplotypes, but are in imperfect LD, may generate the false impression of interactions (Wood et al. 2014). To account for this, *MC1R* variants were included as covariates

in SNPs tested in chromosome 16. To determine whether the two models differed by more than a sampling error, a likelihood ratio test was conducted to assess the goodness of fit of the two logistic regression models. This analysis was conducted through Cassi software. Evidence of epistasis between the *MC1R* coding variant and previously identified signal was declared if the χ^2 p-value was less than 0.05/135,264 (adjusted for multiple testing) in the likelihood ratio test. Results for each combination were tabulated.

8.2.4 Epistasis of combined *MC1R* alleles in red hair, tanning ability, skin colour and non-red hair colour

To assess overall epistasis effects within the *MC1R* coding regions, and to increase model sample size, coding variants within the *MC1R* regions were combined together into a single coded score based upon the number of effect alleles an individual had in the *MC1R* region. To achieve this, the thirteen identified *MC1R* coding variant SNPs were extracted from each UK Biobank individual, and the number of effect alleles were summed across all thirteen SNPs, giving each individual a score between zero and two based on the number of *MC1R* effect alleles present. For each pigmentary trait (red hair, non-red hair colour, skin colour, tanning ability), individual's ID, combined *MC1R* effect allele score, and phenotype were combined with the identified signals for their respective GWAS. Epistasis effects were then identified for each of the pigmentary traits using the methods describes previously (see 8.2.3) where the individual *MC1R* coding variants were replaced with the combined *MC1R* effect allele score.

8.3 Results

8.3.1 Penetrance

Overall *MC1R* coding variant penetrance estimations varied across pigmentary traits, with hair-based phenotypes observing the highest overall penetrance estimates compared to skin-based phenotypes. For red hair, it was possible to precisely quantifying the penetrance of effect alleles whether homozygotes or combination. As with other penetrance studies, homozygotes variants ranged from less than 1% (rs201326893) to over 90% (rs1805009) penetrance of red hair. Combination penetrance ranged from less than 1% (many found to be 0% see Table 8.1) to over 90% (rs368507952 with rs1805009). This large range in detected penetrance coincides with a previous study that categorised coding variants 'R' and 'r' for high and low penetrance (Hysi

et al. 2018). Lower penetrance estimates results were observed for the other four pigmentation phenotypes. For Non-red hair, the penetrance estimates for homozygotes variants ranged between 0% (rs368507952) to over 37% (rs1805007). Combined effect alleles had a wider range of penetrance with many combinations estimating penetrance at 0% and the highest penetrance estimate being 75% (rs1805007 with rs201326893). Penetrance estimates of *MC1R* coding variants on skin colour ranged between 0% and 75% (rs1805007 and rs34158934). The penetrance estimates of *MC1R* coding variants on tanning ability had the lowest overall values with many variant combinations estimated penetrance to be 0%, and the highest penetrance estimates being 50% (rs1805006 with rs34474212, and rs34474212 with rs1805009). Across pigmentation traits similar penetrance patterns were observed for homozygotes and combination *MC1R* coding variant pairings with the paring of rs1805007 and rs201326893 having a high penetrance across red hair, non-red hair, and skin colour, with tanning ability not following this pattern. Comparing the penetrance estimates for strong-R and weak-r variants across pigmentary traits found weak-r variants having a lower penetrance effect for red hair compared to strong-R variants whereas the other pigmentary traits didn't follow this pattern, having a much more similar penetrance effect estimation between strong-R and weak-r variants.

	rs3212379	rs1805005	rs34474212	rs1805006	rs2228479	rs34158934	rs11547464	rs1805007	rs201326893	rs1110400	rs1805008	rs885479	rs555179612	rs200000734	rs1805009	rs368507952
rs3212379	59.09	7.77	23.08	58.33	5.40	NA	58.70	67.26	0.00	31.17	57.95	5.42	17.13	50.00	65.32	33.33
rs1805005	7.77	0.44	2.16	3.52	0.20	0.00	3.67	7.78	8.93	1.36	4.49	0.30	2.05	2.74	4.26	17.39
rs34474212	23.08	2.16	50.00	9.09	1.31	NA	23.53	13.24	33.33	10.00	10.87	1.50	4.32	0.00	12.35	NA
rs1805006	58.33	3.52	9.09	19.18	1.47	0.00	38.27	52.33	80.00	11.73	28.32	1.20	11.83	11.11	53.50	57.14
rs2228479	5.40	0.20	1.31	1.47	0.17	0.00	1.63	2.59	3.33	0.88	1.35	0.13	0.62	1.33	2.04	3.45
rs34158934	NA	0.00	NA	0.00	0.00	NA	0.00	50.00	NA	50.00	50.00	0.00	16.67	NA	33.33	NA
rs11547464	58.70	3.67	23.53	38.27	1.63	0.00	78.79	65.35	100.00	14.29	29.84	1.63	15.06	20.00	83.67	80.00
rs1805007	67.26	7.78	13.24	52.33	2.59	50.00	65.35	73.69	89.47	23.50	54.33	2.33	12.83	35.29	78.60	86.36
rs201326893	0.00	8.93	33.33	80.00	3.33	NA	100.00	89.47	0.00	62.50	73.53	4.55	23.91	NA	66.67	NA
rs1110400	31.17	1.36	10.00	11.73	0.88	50.00	14.29	23.50	62.50	2.22	12.72	3.77	5.26	11.11	18.36	33.33
rs1805008	57.95	4.49	10.87	28.32	1.35	50.00	29.84	54.33	73.53	12.72	35.77	1.81	10.85	22.95	52.27	65.85
rs885479	5.42	0.30	1.50	1.20	0.13	0.00	1.63	2.33	4.55	3.77	1.81	0.26	0.89	0.00	1.48	0.00
rs555179612	17.13	2.05	4.32	11.83	0.62	16.67	15.06	12.83	23.91	5.26	10.85	0.89	4.18	4.23	16.55	16.07
rs200000734	50.00	2.74	0.00	11.11	1.33	NA	20.00	35.29	NA	11.11	22.95	0.00	4.23	NA	33.33	0.00
rs1805009	65.32	4.26	12.35	53.50	2.04	33.33	83.67	78.60	66.67	18.36	52.27	1.48	16.55	33.33	92.36	90.91
rs368507952	33.33	17.39	NA	57.14	3.45	NA	80.00	86.36	NA	33.33	65.85	0.00	16.07	0.00	90.91	66.67

Figure 8.1: Penetrance (defined as the percentage of persons with genotype with red hair) estimates matrix for *MC1R* variants on the red hair phenotype

	rs3212379	rs1805005	rs34474212	rs1805006	rs2228479	rs34158934	rs11547464	rs1805007	rs201326893	rs1110400	rs1805008	rs885479	rs555179612	rs200000734	rs1805009	rs368507952
rs3212379	12.5	18.37	NA	18.18	14.79	NA	15.79	17.24	NA	21.57	18.62	21.18	0	0	17.07	0
rs1805005	18.37	12.81	13.89	15.13	11.01	20	15.95	20.33	20.83	15.17	18.79	13.21	21.58	19.72	19.04	10.81
rs34474212	NA	13.89	NA	NA	0	NA	NA	0	NA	NA	0	22.22	NA	NA	0	NA
rs1805006	18.18	15.13	NA	29.09	13.91	0	34.04	31.26	0	22.3	24.46	13.99	0	14.29	23.2	0
rs2228479	14.79	11.01	0	13.91	8.69	25	17	15.82	14.29	11.18	14.29	8.73	8.76	13.7	12.99	15.38
rs34158934	NA	20	NA	0	25	NA	NA	0	NA	0	50	50	NA	NA	0	NA
rs11547464	15.79	15.95	NA	34.04	17	NA	33.33	32.55	NA	23.94	28.86	15	0	0	43.48	0
rs1805007	17.24	20.33	0	31.26	15.82	0	32.55	37.77	75	22.51	33.06	17.91	37.5	27.91	34.72	33.33
rs201326893	NA	20.83	NA	0	14.29	NA	NA	75	NA	0	14.29	9.52	NA	NA	33.33	NA
rs1110400	21.57	15.17	NA	22.3	11.18	0	23.94	22.51	0	17.44	24.07	12.23	16.67	37.5	22.11	50
rs1805008	18.62	18.79	0	24.46	14.29	50	28.86	33.06	14.29	24.07	30.36	17.16	43.48	31.11	29.69	50
rs885479	21.18	13.21	22.22	13.99	8.73	50	15	17.91	9.52	12.23	17.16	6.07	22.67	7.69	15.01	16.67
rs555179612	0	21.58	NA	0	8.76	NA	0	37.5	NA	16.67	43.48	22.67	11.23	0	NA	NA
rs200000734	0	19.72	NA	14.29	13.7	NA	0	27.91	NA	37.5	31.11	7.69	0	NA	50	NA
rs1805009	17.07	19.04	0	23.2	12.99	0	43.48	34.72	33.33	22.11	29.69	15.01	NA	50	22.22	0
rs368507952	0	10.81	NA	0	15.38	NA	0	33.33	NA	50	50	16.67	NA	NA	0	0

Figure 8.2: Penetrance estimates matrix for *MC1R* variants on the non-red hair phenotype

	rs3212379	rs1805005	rs34474212	rs1805006	rs2228479	rs34158934	rs11547464	rs1805007	rs201326893	rs1110400	rs1805008	rs885479	rs555179612	rs200000734	rs1805009	rs368507952
rs3212379	36.36	12.35	NA	38.33	12.77	NA	41.3	36.48	0	24.68	32.01	14.44	27.27	25	41.94	33.33
rs1805005	12.35	7.79	12.5	12.92	7.34	0	10.3	13.91	12.5	10.97	12.03	6.77	20.44	8.22	14.27	4.35
rs34474212	NA	12.5	NA	0	13.33	NA	50	57.89	NA	0	25	10	NA	NA	50	NA
rs1805006	38.33	12.92	0	31.51	13.24	0	43.21	33.75	40	22.84	26.37	12.33	45.83	44.44	33.57	28.57
rs2228479	12.77	7.34	13.33	13.24	7.2	12.5	9.77	13.15	26.67	9.95	10.38	5.5	17.76	6.67	12.71	17.24
rs34158934	NA	0	NA	0	12.5	NA	0	75	NA	0	25	0	NA	NA	33.33	NA
rs11547464	41.3	10.3	50	43.21	9.77	0	57.58	34.79	50	15.48	22.79	6.84	60	40	53.74	60
rs1805007	36.48	13.91	57.89	33.75	13.15	75	34.79	38.49	44.74	25.05	29.2	11.78	46.63	32.35	43.9	43.18
rs201326893	0	12.5	NA	40	26.67	NA	50	44.74	0	25	29.41	13.64	0	NA	33.33	NA
rs1110400	24.68	10.97	0	22.84	9.95	0	15.48	25.05	25	14.44	17.56	12.84	29.17	22.22	25	33.33
rs1805008	32.01	12.03	25	26.37	10.38	25	22.79	29.2	29.41	17.56	22.1	9.87	43.54	29.51	31.17	43.9
rs885479	14.44	6.77	10	12.33	5.5	0	6.84	11.78	13.64	12.84	9.87	4.39	16.09	18.52	14.05	12.5
rs555179612	27.27	20.44	NA	45.83	17.76	NA	60	46.63	0	29.17	43.54	16.09	7.6	0	52.38	NA
rs200000734	25	8.22	NA	44.44	6.67	NA	40	32.35	NA	22.22	29.51	18.52	0	NA	33.33	0
rs1805009	41.94	14.27	50	33.57	12.71	33.33	53.74	43.9	33.33	25	31.17	14.05	52.38	33.33	56.25	45.45
rs368507952	33.33	4.35	NA	28.57	17.24	NA	60	43.18	NA	33.33	43.9	12.5	NA	0	45.45	33.33

Figure 8.3: Penetrance estimates matrix for *MC1R* variants on the skin colour phenotype

	rs3212379	rs1805005	rs34474212	rs1805006	rs2228479	rs34158934	rs11547464	rs1805007	rs201326893	rs1110400	rs1805008	rs885479	rs555179612	rs200000734	rs1805009	rs368507952
rs3212379	4.55	15.97	NA	3.33	12.95	NA	13.04	5.69	0	7.79	6.69	9.75	0	0	4.03	0
rs1805005	15.97	17.09	10	9.33	16.7	0	11.24	8.42	8.93	12.87	10.89	18.36	9.94	9.59	8.96	13.04
rs34474212	NA	10	NA	50	0	NA	0	5.26	NA	0	8.33	20	NA	NA	50	NA
rs1805006	3.33	9.33	50	5.48	9.8	0	1.23	3.73	0	3.7	4.64	10.1	0	11.11	3.5	0
rs2228479	12.95	16.7	0	9.8	18.18	0	11.73	8.96	10	13.56	10.81	21.56	5.92	6.67	8.81	0
rs34158934	NA	0	NA	0	0	NA	0	0	NA	0	12.5	0	NA	NA	0	NA
rs11547464	13.04	11.24	0	1.23	11.73	0	3.03	3.54	0	2.38	5.74	14.33	10	20	2.04	0
rs1805007	5.69	8.42	5.26	3.73	8.96	0	3.54	3.86	5.26	6.44	4.47	10.32	3.68	4.41	3.08	4.55
rs201326893	0	8.93	NA	0	10	NA	0	5.26	0	12.5	5.88	22.73	0	NA	0	NA
rs1110400	7.79	12.87	0	3.7	13.56	0	2.38	6.44	12.5	14.44	7.26	10.45	4.17	11.11	3.52	33.33
rs1805008	6.69	10.89	8.33	4.64	10.81	12.5	5.74	4.47	5.88	7.26	5.45	12.32	3.4	4.92	3.99	4.88
rs885479	9.75	18.36	20	10.1	21.56	0	14.33	10.32	22.73	10.45	12.32	27.24	13.79	3.7	9.89	20.83
rs555179612	0	9.94	NA	0	5.92	NA	10	3.68	0	4.17	3.4	13.79	21.02	0	2.38	NA
rs200000734	0	9.59	NA	11.11	6.67	NA	20	4.41	NA	11.11	4.92	3.7	0	NA	0	0
rs1805009	4.03	8.96	50	3.5	8.81	0	2.04	3.08	0	3.52	3.99	9.89	2.38	0	1.39	0
rs368507952	0	13.04	NA	0	0	NA	0	4.55	NA	33.33	4.88	20.83	NA	0	0	0

Figure 8.4: Penetrance estimates matrix for *MC1R* variants on the tanning ability phenotype

8.3.2 GWAS

Genome-wide significant association with red hair ($p < 5 \times 10^{-8}$) was detected for individuals with no *MC1R* effect alleles at 119 independent genetic variants across 80 loci, and for individuals with two *MC1R* effect alleles at 123 independent variants across 96 loci. This discrepancy in number of identified regions between the two GWAS provide evidence of epistasis in *MC1R* coding variants across other associated regions with red hair colour. Further to the detection of loci identfied in both red hair GWAS, statistical power between the two GWAS also differed with the sample size for the GWAS of individuals with no *MC1R* coding variants was around three times the sample size of the GWAS of individuals with two *MC1R* coding variants, and thus with the additional statistical power it would be expected to detect these additional associated loci. To further provide evidence of epistasis between *MC1R* coding variants and other genomic regions using the red hair GWAS, the peaks on chromosome 16 and chromosome 20 across the two red hair GWAS were investigated. For chromosome 16, the effect size (log OR) for the top signal in the red hair GWAS on individuals with no *MC1R* coding variants had a similar effect estimate (rs144013020: -1.83) compared to the red hair GWAS on individuals with two *MC1R* coding variants (rs144013020: -1.58). For the chromosome 20 peaks, the red hair GWAS on individuals with no *MC1R* coding variants had a weaker effect estimate (log OR) (rs62211989: -1.37) compared to red hair GWAS on individuals with two *MC1R* coding variants (rs62211989: -2.07), which may be due to epistatic effects between the *MC1R* coding variants and the peak region on chromosome 20. When further investigating the peaks on chromosomes 16 and 20 for both red hair GWAS, the red hair GWAS conducted on individuals with no *MC1R* coding variants found association at chromosome 16 with the most significant signal rs144013020 found in the *TCF25* gene, and at chromosome 20 with the most significant signals rs62211989 found near *ASIP*; the red hair GWAS conducted individuals with two *MC1R* coding variants found association at chromosome 16 with the most significant signal rs11641201 found upstream of *MC1R* and at chromosome 20 with the most significant signal rs6059655 found in *ASIP*.

Figure 8.5: Manhattan plot of red hair colour GWAS from UK Biobank conducted on 51,244 participants that do not have any *MC1R* coding variants.

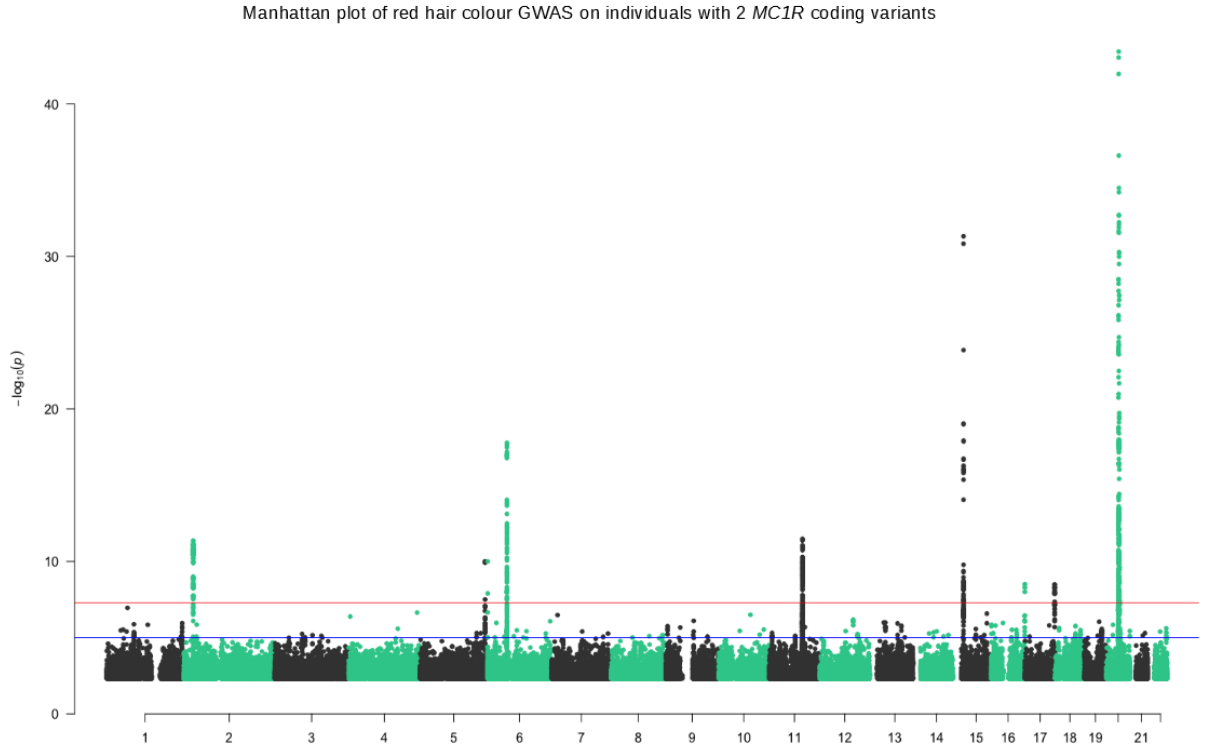


Figure 8.6: Manhattan plot of red hair colour GWAS from UK Biobank conducted on 15,222 participants that have two *MC1R* coding variants.

8.3.3 Epistasis

Testing the thirteen associated genetic variants of *MC1R* identified through a UK Biobank-based red hair GWAS using logistic models with and without interaction terms, found six regions where *MC1R* coding variant had epistasis effects on genomic regions. The strongest evidence of epistasis was found in chromosome 20 in the *ASIP* region (rs6059655) where nine *MC1R* coding variants (rs9921065, rs34357723, rs1805005, rs2228479, rs1805007, rs1110400, rs885479, rs1805009, and rs11859970) were found to affect this region. The effect sizes (log odds) of the interaction terms ranged between -0.28 to 1.16, indicating a complex relation between phenotype, present coding variants, and epistasis effects. Of these six detected regions, two were novel rs60780889 on chromosome 5, and rs9544609 on chromosome 13 (Table 8.2). When further investigating the interaction between SNPs with the strongest p-value (rs18005005 and rs6059655), the OR for the interaction term between rs1805005 and rs6059655 was 1.06, and suggested that a combination of these SNP's effect alleles (rs1805005: T and rs6059655: A) would further increase penetrance compared to individual effects. This was observed when calculating penetrance estimations for allele pairings where two effect alleles for rs1805005 and two wild-types for rs6059655 estimated penetrance of red hair at 0.46% and two effect alleles for rs6059655 and two wild-types for rs1805005 estimated penetrance of red hair at 0.44%. Combining effect alleles for both traits (rs1805005: TT and rs6059655: AA) found higher penetrance estimate (1.14%) if compared to a simple additive penetrance effect model of red hair using the two individual pairing penetrance estimates (rs1805005: TT and rs6059655: GG: 0.46% + rs1805005: GG and rs6059655: AA 0.44%) (Table 8.3).

			CHR 5			CHR 6			CHR 13			CHR 15			CHR 20					
RsID	CHR	BP	RsID	LR OR	LR P	RsID	LR OR	LR P	RsID	LR OR	LR P	RsID	LR OR	LR P	RsID	LR OR	LR P	RsID	LR OR	LR P
rs1110400	16	89986130													rs6059655	0.51	4E-11			
rs117030214	16	89749165										rs1129038	0.26	5.5E-08						
rs117614496	16	88895529										rs1129038	0.22	4.4E-07						
rs11859970	16	90018389													rs6059655	0.23	1E-09			
rs164737	16	89666511	rs60780889	-0.09	1E-08							rs1129038	-0.10	6.2E-09						
rs1805005	16	89985844	rs60780889	0.21	5.4E-09	rs2025753	-0.20	2.5E-09	rs9544609	0.17	6.6E-07	rs1129038	-0.28	1.5E-09	rs6059655	1.06	1E-162	rs4519594	0.30	7.1E-08
rs1805007	16	89986117										rs1129038	0.18	6.2E-16	rs6059655	-0.28	2.3E-27			
rs1805008	16	89986144	rs60780889	0.12	3.4E-10															
rs1805009	16	89986546										rs1129038	0.21	3E-09	rs6059655	-0.29	1.9E-12			
rs2228479	16	89985940													rs6059655	1.16	1.6E-65			
rs34357723	16	89886519													rs6059655	-0.29	1.9E-34			
rs885479	16	89986154													rs6059655	0.83	7.3E-21			
rs9921065	16	89066873										rs1129038	-0.60	1.8E-13	rs6059655	0.56	1.8E-07			

Table 8.2: Epistasis effect estimates for *MC1R* variants across other associated genomic regions for red hair

		rs6059655		
		A A	G A	G G
rs1805005	G G	0.44%	0.09%	0.01%
	G T	0.76%	0.55%	0.17%
	T T	1.14%	1.39%	0.46%

Table 8.3: Red hair penetrance matrix for rs1805005 and rs6059655 alleles.

Combining the thirteen associated genetic variants of *MC1R* into a single *MC1R* score (0,1,2), identified fewer epistatic effects in red hair, with the only interaction identified residing in the *ASIP* region on chromosome 20 (rs6059655) and gave an overall interaction term effect (log odds) of -0.57 (Table 8.4). For tanning ability, no epistatic effects were identified for a combined *MC1R* coding variant score (Table 8.5); for skin colour, no epistasis effects were identified for a combined *MC1R* coding variant score (Table 8.7); and for non-red hair, no epistatic effects were identified for a combined *MC1R* coding variant score (Table 8.6).

CHR1	ID1	Region 1	CHR2	ID2	BP2	OR	SE	p-value
16	rscombine	*MC1R* region	17	rs9747347	79606820	-0.115659	0.0361385	0.00137224
16	rscombine	*MC1R* region	2	rs76645364	25393388	0.0861509	0.045594	0.0588218
16	rscombine	*MC1R* region	20	rs6059655	32665748	-0.565474	0.0522653	2.79E-27
16	rscombine	*MC1R* region	20	rs4519594	34960201	-0.110991	0.0608902	0.0683324
16	rscombine	*MC1R* region	6	rs12529074	51713402	-0.116437	0.0611602	0.0569356

Table 8.4: Epistasis effect estimates for combined *MC1R* variants across other associated genomic regions for red hair

CHR1	ID1	Region 1	CHR2	ID2	BP2	Beta	SE	p-value
16	rscombine	*MC1R* region	1	rs670318	63727542	-0.0242434	0.00888241	0.00634574
16	rscombine	*MC1R* region	2	rs780094	27741237	0.00842987	0.00398839	0.0345502
16	rscombine	*MC1R* region	3	rs9821965	98698414	0.00764703	0.00395333	0.0530736
16	rscombine	*MC1R* region	3	rs9818780	156492758	-0.00761175	0.00389027	0.0503939
16	rscombine	*MC1R* region	6	rs2517674	29935806	-0.0131407	0.00469392	0.00511827
16	rscombine	*MC1R* region	7	rs117132860	17134708	-0.0302399	0.0121669	0.0129403
16	rscombine	*MC1R* region	12	rs61935849	116637641	0.012462	0.00613224	0.0421333
16	rscombine	*MC1R* region	12	rs4760521	129306308	0.00894797	0.00403984	0.0267655
16	rscombine	*MC1R* region	14	rs1885194	92777462	0.00684041	0.0038832	0.0781478
16	rscombine	*MC1R* region	15	rs12913832	28365618	-0.00985243	0.0046769	0.0351516
16	rscombine	*MC1R* region	20	rs151525	3522015	0.0131124	0.00457239	0.00413456

Table 8.5: Epistasis effect estimates for combined *MC1R* variants across other associated genomic regions for tanning ability

CHR1	ID1	Region 1	CHR2	ID2	BP2	BETA	SE	p-value
16	rscombine	*MC1R* region	1	rs78555129	1178482	-0.00834188	0.00492694	0.0904347
16	rscombine	*MC1R* region	1	rs17377295	61710179	0.0120755	0.00594014	0.0420668
16	rscombine	*MC1R* region	2	rs4952542	42148519	-0.00537325	0.00307542	0.0806108
16	rscombine	*MC1R* region	2	rs3749110	109605767	0.00778077	0.0043023	0.0705277
16	rscombine	*MC1R* region	2	rs9287636	239680992	0.00936396	0.00324109	0.00386325
16	rscombine	*MC1R* region	3	rs2293252	138123854	-0.0054527	0.00317757	0.0861637
16	rscombine	*MC1R* region	4	rs9998015	105816898	-0.00617097	0.00315308	0.0503336
16	rscombine	*MC1R* region	6	rs34451818	31583841	0.010756	0.00623184	0.0843528
16	rscombine	*MC1R* region	8	rs330926	9000965	0.00595747	0.00338289	0.0782291
16	rscombine	*MC1R* region	9	rs10811647	22065002	-0.00579698	0.00304345	0.0568149
16	rscombine	*MC1R* region	11	rs2049129	18305333	0.00790903	0.00389738	0.0424265
16	rscombine	*MC1R* region	11	rs9645690	62206288	0.00572921	0.00323773	0.0768089
16	rscombine	*MC1R* region	11	rs1042602	88911696	-0.00514605	0.00310337	0.0972753
16	rscombine	*MC1R* region	12	rs12368980	91116069	-0.00634662	0.00341056	0.0627627
16	rscombine	*MC1R* region	12	rs1716169	123716930	-0.00737626	0.00375813	0.0496766
16	rscombine	*MC1R* region	12	rs77717551	125368342	-0.0124076	0.00612401	0.0427591
16	rscombine	*MC1R* region	20	rs12481673	57845936	0.00640838	0.00329404	0.0517219

Table 8.6: Epistasis effect estimates for combined *MC1R* variants across other associated genomic regions for non-red hair colour

CHR1	ID1	Region 1	CHR2	ID2	BP2	BETA	SE	p-value
16	rscombine	*MC1R* region	1	rs535930	110724488	-0.00499416	0.00229364	0.0294516
16	rscombine	*MC1R* region	1	rs3894771	150789961	-0.00528373	0.00229718	0.0214437
16	rscombine	*MC1R* region	1	rs3851294	205130413	-0.0107663	0.00393332	0.00619664
16	rscombine	*MC1R* region	2	rs7591631	175292326	-0.00409735	0.00228291	0.072688
16	rscombine	*MC1R* region	3	rs9867857	156491160	-0.00470388	0.00227984	0.03909
16	rscombine	*MC1R* region	5	rs285894	113996889	-0.00410548	0.00231397	0.076028
16	rscombine	*MC1R* region	6	rs13201830	51696355	0.00531974	0.00228305	0.0198014
16	rscombine	*MC1R* region	7	rs56134000	100499789	-0.00758542	0.00263487	0.00399138
16	rscombine	*MC1R* region	11	rs7478729	18319915	0.00397546	0.00235515	0.0914143
16	rscombine	*MC1R* region	15	rs28456199	31395538	-0.00997373	0.00424873	0.0189025
16	rscombine	*MC1R* region	15	rs28371837	83383187	-0.00660083	0.00285079	0.0205893
16	rscombine	*MC1R* region	17	rs6420484	79612397	0.00493514	0.0023614	0.0366258

Table 8.7: Epistasis effect estimates for combined *MC1R* variants across other associated genomic regions for skin colour

8.4 Discussion

I report the results of penetrance estimation for red hair colour for *MC1R* coding variants. I also report the results of estimates of the epistatic effect for *MC1R* coding variants for tanning ability, skin colour, non-red hair colour, and red hair colour. For the *MC1R* penetrance estimates, the large UK Biobank sample size is utilised to create the current most accurate effect estimates that homozygotes and combinations of *MC1R* coding variants have on an individual's pigmentation. Results showed a wide variation in penetrance, with strong-R variants having high penetrance estimates (over 90% penetrance) for homozygotes and combinations, and weak-r variants having low penetrance estimates (less than 1% penetrance). This results concur with previous studies investigating *MC1R* coding penetrance on human pigmentation. Similarly, the identification of epistatic effects in *MC1R* coding variants on human pigmentation utilised the large UK Biobank sample size to provide the current largest power to detect the presence of epistasis in associated pigmentation regions with *MC1R*. These results provided evidence of epistasis in six previously associated regions where *MC1R* effects were present for red hair. Of these six, two were novel and not previously reported. No significant epistasis was identified in tanning ability, skin colour, or non-red hair colour for *MC1R* coding variants.

When conducting two red hair GWAS, one for individuals with two *MC1R* coding variants and one for individuals with no *MC1R* coding variants, a difference in the number of associated loci was observed between GWAS results. Surprisingly, a high significance was detected around the *MC1R* gene in the red hair GWAS for individuals with no *MC1R* coding variants. By removing individuals with any *MC1R* coding variant, it was expected that *MC1R* would have a minimal effect on the determination of red hair. Investigating this peak identified rs144013020 in the nearby *TCF25* gene, which has no prior known associations through GWAS catalog (Buniello et al. 2019). The likely cause of the association may be due to LD with a non-coding variant (eg prompter variant) situated in or up/downstream of *MC1R*. Furthermore, for the red hair GWAS conducted on individuals with two *MC1R* coding variants, a peak was observed around the *MC1R* gene also. Inspecting this peak identified rs11641201, upstream of the *MC1R* gene, which also provides evidence of further variants yet to be discovered for *MC1R*.

While the results presented here highlight the strong relationship between *MC1R* coding vari-

ants and their effect on hair colour, particularly red hair, their effect on skin colour are far more subtle when investigating their penetrance. This quantitative difference in penetrance across pigmentary traits for *MC1R* coding variants is evidence of a fundamental difference in response of melanocytes found in follicles compared to the epidermis. Within the skin, each melanocyte is surrounded by keratinocytes, which form an epidermal melanin unit (EMU). These EMU acquire, transport and metabolizes the melanin produced from one cell. Due to this self-contained property, epidermal pigmentation can be consider as a mosaic of individual melanocytes contributing millions of EMUs to the surface of the skin. Conversely, hair follicles have a density of melanocytes placed as a cone surrounding the dermal papilla and adjoining cells from differing hair shafts. Follicular melanocytes, keratinocytes and dermal papilla cells have mutual interactions; the dermal papilla signals to melanocytes with *ASIP*, the melanocytes transfer melanin granules into the keratinocytes. Perturbations of these interactions could affect the amount and type of melanin delivered to the hair. Given these inherent differences in cutaneous environment between these two pigmentation, it is possible that the level or accessibility ligands vital to pheomelanin production is different across the two, and thus, explains how the *MC1R* coding variants are having a greater effect on follicle-based pigmentation compared to epidemis-based pigmentation.

Furthermore to the biological evidence provided as to why the hair-based pigmentation traits were observed to have higher *MC1R* penetrance estimates compare to skin-based traits, the underlying reliability and bias of the pigmentation traits may also contribute. This is likely caused by incorrect self-reporting of individuals pigmentation measures. In particular, tanning ability may be more susceptible to this form of reporting bias compared to other traits, as tanning ability could be a more subjective measure compared to hair colour and skin colour, i.e. individuals are more likely to accurately report a more observed physical pigmentation trait (such as hair colour or skin colour) compared to tanning ability. Additionally, the lower *MC1R* penetrance estimations observed for skin colour may be caused by the distribution of responses in UK Biobank for pigmentary measures. As analysis was conducted on participants of European descent the dominant responses were: 'very fair', 'fair' and 'light olive' limiting the skin colour scale to consist of closely related subjective responses compared to hair colour which had a larger spread of responses. This limitation in the skin colour scale may cause higher reporting bias, and thus, lower penetrance estimates for *MC1R*.

In addition to the highlighted difference of *MC1R* coding variants across pigmentary traits, epistasis was only detected between *MC1R* and other genes for red hair. The mechanism that links *ASIP* and *MC1R* is the dermal papilla, which are used as intermediary between the two. This uniqueness found in hair follicles compared to epidermal melanocytes goes some ways in explaining the strong evidence of epistasis in red hair compared to other pigmentary traits.

Chapter 9

Conclusions

The primary outcome of the thesis has been to further our understanding of the genetic architecture of pigmentation and melanoma by utilising large datasets to supplement sample size issues commonly found in GWAS studies. To achieve this, six aims were developed that were derived to understand different aspects of melanoma and how it relates to pigmentation. Firstly for Aim I, the reliability of self-reported melanoma was assessed. This was conducted to determine whether the assumption that including self-reported melanoma cases within the UK Biobank was appropriate. As previous studies had assessed recall bias in relation to a limited number of risk factors, and the reproducibility of melanoma risk factors separately for cases and controls, but no previous studies had directly assessed the recall bias in melanoma self-reported diagnosis. To quantify this bias, UK Biobank participant's cancer self-reports were linked with UK Cancer Registry entries with the use of a matching algorithm. From the analysis, it was observed that 32% of melanoma self-reports did not match with melanoma or any other type of cancer when comparing to the UK Cancer Registry. Melanoma was more likely to be successfully self-reported if a participant had a previous diagnosis, and that melanoma self-reports were less reliable compared to breast cancer and non-melanoma skin cancer, likely due to participants with non-melanoma skin cancer believing melanoma as a general term of skin cancer.

Secondly for Aim II and Aim III, the relationship between pigmentation traits was investigated. By understanding the inter-relationship across pigmentary traits allowed for optimal efficacy when implementing the data for supplementary melanoma/pigmentation analyses. This was achieved by conducting factor analysis. Firstly by conducting an exploratory factor analysis,

it identified that the optimal number of latent variants acting on these pigmentation measures (based on correlation between measures) was one. This single latent variant was considered to be a person's overall pigmentation. The confirmatory factor analysis allowed for an accurate estimate of the effect this overall pigmentation latent variable has on pigmentation measures in the form of factor loadings. These factor loadings indicated that not any one pigmentation trait is dominated by a person's overall pigmentation, and that a person's overall pigmentation equally contributes to a person's individual pigmentation characteristics. From the analysis, childhood sunburns had a significantly lower loading which was likely due to recall bias and external factors influencing the reporting of this trait. These findings led to the creation of an overall pigmentation variables, which was created using PCA and gave the following weightings to the pigmentation measures: 0.78 for skin; -0.69 for ease; 0.87 for non-red hair; and -0.95 and red hair. This findings further supports the hypothesis no one pigmentation measure dominates a person's overall pigmentation phenotype.

Additional to understanding the latent model between pigmentation measures, the genetic architecture of pigmentation was modelled to allow for the comparison between the genetic architecture of melanoma with the genetic architecture of pigmentation, prediction of melanoma risk, and combination with melanoma in further analyses. Firstly, the suitability of measuring each pigmentation trait using an ordinal scale was assessed and found distinct characteristic between the genetic architectures of red and non-red hair phenotypes; and provided evidence that including these into the same ordinal scale, under the assumption that there is a uniform genetic architecture across hair colour, may lead to the reporting of misleading associations. Red hair was therefore separated from non-red hair for further analyses. Tanning ability and skin colour did not find any structural genetic architecture differences across their respective ordinal scales, and therefore it was deemed suitable to retain the UK Biobank determined ordinal scales. After confirming the validity of the ordinal scales for the three UK Biobank pigmentation variables, eight pigmentation-based GWAS were conducted: two skin-based GWAS with skin colour and tanning ability, two hair colour GWAS with hair colour separated into non-red and red hair, a childhood sunburn GWAS; and three results of three GWAS of combined pigmentation scores. Overall 322 loci across all the traits were identified, 109 of which have never previously been associated with a pigmentation trait. The discovery of many new loci improves our understanding of how the genetic architectures of distinct pigmentary measurements relate to one another.

There was large overlap between pigmentation-derived loci and melanoma-derived loci.

From the investigation of melanoma self-reports, the inter-pigmentation relationship, and genetic architecture of pigmentation and how it overlaps with melanoma, it was determined to be appropriate to supplement pigmentation data with melanoma data. For Aim IV, the pigmentation GWAS summary statistics were used to predict melanoma risk. A PRS of hair colour (red and non-red separately), skin colour, tanning ability, and three overall pigmentation scores were produced and optimised using the UK Biobank dataset. The PRSs were tested in the LMC, and high effectiveness was observed across all scores (AUC: 0.59 to 0.64), suggesting that there is no one specific aspect of pigmentation that is any more genetically-related to melanoma risk than others. The effectiveness of combining genetic information for multiple melanoma risk pathways was observed by creating a melanoma risk model that included a pigmentation PRS and naevus PRS that out performed a single pigmentation PRS-based melanoma risk model. When considering a 'best prediction' model, the utility of large pigmentary datasets found as a 'best fit' model that included a pigmentation-based PRS over a melanoma-based PRS performed equally well.

For Aim V, two methodologies were produced to provide evidence of the feasibility of detecting novel melanoma risk signals. The first method uses a targeted analysis pipeline that separately identifies regions of joint-association and conducts sample-size weighted joint-analyses for each pigmentation GWAS summary statistics with a melanoma meta-analysis GWAS summary statistics. The second method conducts separate genetic correlation weighted joint-analysis that inflates/deflates regions of association for each pigmentation GWAS summary statistics with a melanoma meta-analysis GWAS summary statistics. Overall, these two methods identified 139 independent associated loci for melanoma; of these 139 independently associated loci, 88 were novel. These novel loci provide further insight into the genetic architecture of melanoma, and in time will be verified with ever growing sample-size melanoma meta-analysis GWAS.

Finally, for Aim VI, as the UK Biobank provides a uniquely large sample size, to detect subtle genetic effects, the penetrance of *MC1R* coding variants on red hair, non-red hair, tanning ability, and skin colour were estimated, and the epistasis effects of *MC1R* coding variants on

red hair, non-red hair, skin colour, and tanning ability for other significantly associated genomic regions. For these pigmentation traits, pentrance analyses concluded strong-R variants had high penetrance estimates (over 90% penetrance) for homozygotes and combinations, and weak-r variants had low penetrance estimates (less than 1% penetrance). These results are in-line with previous studies of *MC1R* coding variants penetrance on pigmentation. Detection of epistasis was found for *MC1R* coding variants on six genomic regions for red hair with the largest in the *ASIP* gene region. No other epistatic effects were found for other pigmentary traits (non-red hair, skin colour, or tanning ability). These analyses furthered our understanding of the relationship of *MC1R* with pigmentation and how MC1R affects other known pigmentation genes.

9.1 Strengths

- One of the biggest strength of the work presented in this thesis is the availability of detailed phenotypic and genomic data in a large cohort of over 500,000 participants, allowing high-powered GWAS capable of detecting subtle signals associated with pigmentation, and for the utilisation of these subtle signals in further analysis to supplement melanoma study power.
- Several precautions were taken to ensure bias was not introduce in analysis. Pre-GWAS QC ensured a homogeneous population sample was select for analysis with population structure adjusted for in models as covariate variables. Validation datasets were used when training and testing could be separated, and dichotomising of data into independent sample when a validation dataset was not available, ensured confirmation bias was not present in findings. Confounding variables were adjusted for in analysis where appropriate; these included genotype array used, age and sex. To ensure misleading results caused by low allele frequency were not present in findings, robust MAF cut-offs were selected for analysis, and robust r^2 values selected to ensure independent signals were genuine and not caused by strong LD factors.
- The availability of multiple pigmentation measures allowed for a more complete understanding of the genetic architecture of a person's overall pigmentation. UK Biobank participant's were asked four pigmentation questions at interview with each highlighting a distinct aspect of pigmentation. The recording of four unique pigmentary aspects for each

participant allowed a robust overall pigmentation score to be produced that correlated well with all pigmentation aspects measured in the UK Biobank. Furthermore, individually testing each different aspects of a person's pigmentation allowed for the detection of overlap between the traits to ensure pigmentation wasn't sub-grouped by characteristics.

9.2 Limitations

- One of the fundamental limitations of the UK Biobank dataset is the retrospective nature of variable generation. Although the prospective aspect of UK Biobank is well underway, and will periodically update cancer self-reports/diagnoses, the pigmentation variables will always be self-reported. Measures were implemented in UK Biobank to assess the misreporting of these initial phenotypic variables gathered with a replication interview on 10,000 particpants. Although this would identify changes in response, the entire accuracy of phenotypic responses isn't fully quantified.

- The identification of melanoma signals through the utilisation of pigmentation data offered two novel methodologies to address the power issues found in melanoma genetic studies. Although the methodologies were carefully considered, the signals found are not able to be validated through a widely-adopted approach. To be able to confirm these signals, sample-size issues often found for melanoma GWAS will have to be resolved. This is a likely outcome as melanoma GWAS meta-analyses are becoming ever larger through collaboration and compiling of melanoma genetic data.

- Throughout analysis, overlap between genetic loci in pigmentation and naevus were identified. In particular, many of these jointly associated loci were found to have substantially different effect sizes and directions of effects. These differences explained some of effect size discrepancies when supplementing pigmentation and melanoma. Commonly, if a locus had opposite effect direction between naevus and pigmentation, the effect size and direction would be an amalgamation of both traits in melanoma. For the known naevus signals, this was easily observed. However, as naevus count GWAS are typically much lower in power compared to pigmentation, and as naevus count is known to be highly heritable, there may be many more of these associations found across melanoma. This could also be the case of telomere length which also suffers power issues for detecting associated genomic regions.

9.3 Future perspectives

Throughout this thesis, the three fundamental aspects of the pigmentation/melanoma relationship were: increase power to detect melanoma risk loci, better understand the role of these pathways in melanoma risk, and improve risk prediction models for melanoma. All three of these aspects have been addressed (Chapter 5, 6, 7), and offer a valid and novel insight into the way large abundant genetic phenotype data can be used to supplement power-limited phenotypic and genotypic trait-based data. With the ever increasing data-banks that provide large-scale genotype and phenotype data on participants, these methodologies outlined here have the potential to further increase our current knowledge when adapted to similar scenarios. More specifically to melanoma, naevus and telomere datasets are becoming larger, and could be utilised in a similar way to pigmentation. Many of these methodologies derived in this thesis (Chapter 6, 7, specifically) can be further adapted to allow for more than one risk pathway trait to be utilised at the same time to further our melanoma understanding. This was briefly explored with a limited-power naevus dataset, that offered marginal gains in melanoma risk prediction.

Furthermore, with the ever decreasing costs associated with computation resources, many of the analyses conducted through the thesis could benefit from further exploration. For instances, many of the pigmentation GWAS were conducted using a linear-based regression model on ordinal scales when some of the effects may be better modelled as recessive. This mismatch of data to model is common place for current GWAS studies, but as computation resources become more readily available, the utilisation of PLINK 2.0 with other software (such as R) to produce an ordinal-regression based GWAS is achievable. This would provide more accurate effect sizes for signals, and further increase the utilisation of pigmentation GWAS findings to supplement melanoma-based genetic studies.

References

Adhikari, Kaustubh, Mendoza-Revilla, Javier, Sohail, Anood, Fuentes-Guajardo, Macarena, Lampert, Jodie, Chacón-Duque, Juan Camilo, Hurtado, Malena, Villegas, Valeria, Granja, Vanessa, Acuña-Alonzo, Victor, Jaramillo, Claudia, Arias, William, Lozano, Rodrigo Barquera, Everardo, Paola, Gómez-Valdés, Jorge, Villamil-Ramírez, Hugo, Silva de Cerqueira, Caio C., Hunemeier, Tábita, Ramallo, Virginia, Schuler-Faccini, Lavinia, Salzano, Francisco M., Gonzalez-José, Rolando, Bortolini, Maria-Cátira, Canizales-Quinteros, Samuel, Gallo, Carla, Poletti, Giovanni, Bedoya, Gabriel, Rothhammer, Francisco, Tobin, Desmond J., Fumagalli, Matteo, Balding, David, and Ruiz-Linares, Andrés (2019). "A GWAS in Latin Americans highlights the convergent evolution of lighter skin pigmentation in Eurasia". In: *Nature Communications* 10.1, p. 358. ISSN: 2041-1723. DOI: 10.1038/s41467-018-08147-0. URL: https://doi.org/10.1038/s41467-018-08147-0.

Agar, Nita and Young, Antony R. (2005). "Melanogenesis: a photoprotective response to DNA damage?" In: *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* 571.1, pp. 121–132. ISSN: 0027-5107. DOI: https://doi.org/10.1016/j.mrfmmm.2004.11.016. URL: http://www.sciencedirect.com/science/article/pii/S0027510704004907.

Aitken, Joanne F., Duffy, David L., Green, Adèle, Youl, Philippa, MacLennan, Robert, and Martin, Nicholas G. (1994). "Heterogeneity of Melanoma Risk in Families of Melanoma Patients". In: *American Journal of Epidemiology* 140.11, pp. 961–973. ISSN: 0002-9262. DOI: 10.1093/oxfordjournals.aje.a117203. URL: https://doi.org/10.1093/oxfordjournals.aje.a117203.

Ambrosio, Andrea, Boyle, Judith, Aradi, Al, Christian, Keith, and Pietro, Santiago (2016). "TPC2 controls pigmentation by regulating melanosome pH and size". In: *Proceedings of the National Academy of Sciences* 113, p. 201600108. DOI: 10.1073/pnas.1600108113.

Autier, Philippe, Doré for, Epimel, and Eortc Melanoma Cooperative Group, Jean-François (1998). "Influence of sun exposures during childhood and during adulthood on melanoma risk". In: *International Journal of Cancer* 77.4, pp. 533–537. ISSN: 0020-7136. DOI: 10.1002/(SICI)1097-0215(19980812)77:4<533::AID-IJC10>3.0.CO;2-7. URL: https://doi.org/10.1002/(SICI)1097-0215(19980812)77:4%3C533::AID-IJC10%3E3.0.CO;2-7.

Bakaysa, S. L., Mucci, L. A., Slagboom, P. E., Boomsma, D. I., McClearn, G. E., Johansson, B., and Pedersen, N. L. (2007). "Telomere length predicts survival independent of genetic influences". In: *Aging Cell* 6.6, pp. 769–74. ISSN: 1474-9726 (Electronic) 1474-9718 (Linking). DOI: 10.1111/j.1474-9726.2007.00340.x. URL: https://www.ncbi.nlm.nih.gov/pubmed/17925004.

Bakker, Paul I.W. de, Ferreira, Manuel A.R., Jia, Xiaoming, Neale, Benjamin M., Raychaudhuri, Soumya, and Voight, Benjamin F. (Oct. 2008). "Practical aspects of imputation-driven meta-analysis of genome-wide association studies". In: *Human Molecular Genetics* 17.R2, R122–R128. ISSN: 0964-6906. DOI: 10.1093/hmg/ddn288. eprint: https://academic.oup.com/hmg/article-pdf/17/R2/R122/8223466/ddn288.pdf. URL: https://doi.org/10.1093/hmg/ddn288.

Barrett, Jennifer H., Iles, Mark M., Harland, Mark, Taylor, John C., Aitken, Joanne F., Andresen, Per Arne, Akslen, Lars A., Armstrong, Bruce K., Avril, Marie-Francoise, Azizi, Esther, Bakker, Bert, Bergman, Wilma, Bianchi-Scarrà, Giovanna, Bressac-de Paillerets, Brigitte, Calista, Donato, Cannon-Albright, Lisa A., Corda, Eve, Cust, Anne E., Dębniak, Tadeusz, Duffy, David, Dunning, Alison M., Easton, Douglas F., Friedman, Eitan, Galan, Pilar, Ghiorzo, Paola, Giles, Graham G., Hansson, Johan, Hocevar, Marko, Höiom, Veronica, Hopper, John L., Ingvar, Christian, Janssen, Bart, Jenkins, Mark A., Jönsson, Göran, Kefford, Richard F., Landi, Giorgio, Landi, Maria Teresa, Lang, Julie, Lubiński, Jan, Mackie, Rona, Malvehy, Josep, Martin, Nicholas G., Molven, Anders, Montgomery, Grant W., Nieuwpoort, Frans A. van, Novakovic, Srdjan, Olsson, Håkan, Pastorino, Lorenza, Puig, Susana, Puig-Butille, Joan Anton, Randerson-Moor, Juliette, Snowden, Helen, Tuominen, Rainer, Van Belle, Patricia,

Stoep, Nienke van der, Whiteman, David C., Zelenika, Diana, Han, Jiali, Fang, Shenying, Lee, Jeffrey E., Wei, Qingyi, Lathrop, G. Mark, Gillanders, Elizabeth M., Brown, Kevin M., Goldstein, Alisa M., Kanetsky, Peter A., Mann, Graham J., Macgregor, Stuart, Elder, David E., Amos, Christopher I., Hayward, Nicholas K., Gruis, Nelleke A., Demenais, Florence, Bishop, Julia A. Newton, Bishop, D. Timothy, and Geno, M. E. L. Consortium (2011). "Genome-wide association study identifies three new melanoma susceptibility loci". In: *Nature genetics* 43.11, pp. 1108–1113. ISSN: 1546-1718 1061-4036. DOI: 10.1038/ng.959. URL: https://pubmed.ncbi.nlm.nih.gov/21983787%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3251256/.

Bartlett, E. K. and Karakousis, G. C. (2015). "Current staging and prognostic factors in melanoma". In: *Surg Oncol Clin N Am* 24.2, pp. 215–27. ISSN: 1558-5042 (Electronic) 1055-3207 (Linking). DOI: 10.1016/j.soc.2014.12.001. URL: https://www.ncbi.nlm.nih.gov/pubmed/25769707.

Beaumont, Kimberley A., Wong, Shu S., Ainger, Stephen A., Liu, Yan Yan, Patel, Mira P., Millhauser, Glenn L., Smith, Jennifer J., Alewood, Paul F., Leonard, J. Helen, and Sturm, Richard A. (2011). "Melanocortin MC1 receptor in human genetics and model systems". In: *European journal of pharmacology* 660.1, pp. 103–110. ISSN: 1879-0712 0014-2999. DOI: 10.1016/j.ejphar.2010.11.040. URL: https://pubmed.ncbi.nlm.nih.gov/21199646%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3095693/.

Bellenguez, Céline, Strange, Amy, Freeman, Colin, Consortium, Wellcome Trust Case Control, Donnelly, Peter, and Spencer, Chris C. A. (Jan. 2012). "A robust clustering algorithm for identifying problematic samples in genome-wide association studies". eng. In: *Bioinformatics (Oxford, England)* 28.1. btr599[PII], pp. 134–135. ISSN: 1367-4811. DOI: 10.1093/bioinformatics/btr599. URL: https://doi.org/10.1093/bioinformatics/btr599.

Bellono, Nicholas W. and Oancea, Elena V. (2014). "Ion transport in pigmentation". In: *Archives of biochemistry and biophysics* 563, pp. 35–41. ISSN: 1096-0384 0003-9861. DOI: 10.1016/j.abb.2014.06.020. URL: https://pubmed.ncbi.nlm.nih.gov/25034214%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4497562/.

Berndt, S.I., Camp, N.J., Skibola, Christine, Vijai, Joseph, Wang, Zhaoming, Gu, J., Nieters, A., Kelly, Rachel, Smedby, K.E., Monnereau, Alain, Cozen, W., Cox, Angela, Wang, Sophia, Lan, Qing, Teras, Lauren, Machado, Moara, Yeager, Meredith, Brooks-Wilson, A.R., Hartge, Patricia, and Slager, S.L. (Mar. 2016). "Meta-analysis of genome-wide association studies discovers multiple loci for chronic lymphocytic leukemia." In: *Nature Communications* 7.

Bin, Bum-Ho, Bhin, Jinhyuk, Yang, Seung Ha, Shin, Misun, Nam, Yeon-Ju, Choi, Dong-Hwa, Shin, Dong Wook, Lee, Ai-Young, Hwang, Daehee, Cho, Eun-Gyung, and Lee, Tae Ryong (2015). "Membrane-Associated Transporter Protein (MATP) Regulates Melanosomal pH and Influences Tyrosinase Activity". In: *PLOS ONE* 10.6, e0129273. DOI: 10.1371/journal.pone.0129273. URL: https://doi.org/10.1371/journal.pone.0129273.

Bishop, D. Timothy, Demenais, Florence, Iles, Mark M., Harland, Mark, Taylor, John C., Corda, Eve, Randerson-Moor, Juliette, Aitken, Joanne F., Avril, Marie-Francoise, Azizi, Esther, Bakker, Bert, Bianchi-Scarrà, Giovanna, Bressac-de Paillerets, Brigitte, Calista, Donato, Cannon-Albright, Lisa A., Chin-A-Woeng, Thomas, Debniak, Tadeusz, Galore-Haskel, Gilli, Ghiorzo, Paola, Gut, Ivo, Hansson, Johan, Hocevar, Marko, Höiom, Veronica, Hopper, John L., Ingvar, Christian, Kanetsky, Peter A., Kefford, Richard F., Landi, Maria Teresa, Lang, Julie, Lubiński, Jan, Mackie, Rona, Malvehy, Josep, Mann, Graham J., Martin, Nicholas G., Montgomery, Grant W., Nieuwpoort, Frans A. van, Novakovic, Srdjan, Olsson, Håkan, Puig, Susana, Weiss, Marjan, Workum, Wilbert van, Zelenika, Diana, Brown, Kevin M., Goldstein, Alisa M., Gillanders, Elizabeth M., Boland, Anne, Galan, Pilar, Elder, David E., Gruis, Nelleke A., Hayward, Nicholas K., Lathrop, G. Mark, Barrett, Jennifer H., and Bishop, Julia A. Newton (2009). "Genome-wide association study identifies three loci associated with melanoma risk". In: *Nature genetics* 41.8, pp. 920–925. ISSN: 1546-1718 1061-4036. DOI: 10.1038/ng.411. URL: https://pubmed.ncbi.nlm.nih.gov/19578364%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2741419/.

Bliss, Judith M., Ford, Deborah, Swerdlow, Anthony J., Armstrong, Bruce K., Cristofolini, Mario, Elwood, J. Mark, Green, Adèle, Holly, Elizabeth A., Mack, Thomas, Mackie, Rona M., ØSterlind, Anne, Walter, Stephen D., Peto, Julian, and Easton, Douglas F. (1995). "Risk of cutaneous melanoma associated with pigmentation characteristics and freckling: Systematic overview of 10 case-control studies". In: *International Journal of Cancer* 62.4, pp. 367–376.

ISSN: 0020-7136. DOI: https://doi.org/10.1002/ijc.2910620402. URL: https://doi.org/10.1002/ijc.2910620402.

Boissy, Raymond E., Sakai, Chie, Zhao, Huiquan, Kobayashi, Takeshi, and Hearing, Vincent J. (1998). "Human tyrosinase related protein-1 (TRP-1) does not function as a DHICA oxidase activity in contrast to murine TRP-1". In: *Experimental Dermatology* 7.4, pp. 198–204. ISSN: 0906-6705. DOI: https://doi.org/10.1111/j.1600-0625.1998.tb00324.x. URL: https://doi.org/10.1111/j.1600-0625.1998.tb00324.x.

Bollen, Kenneth A. and Barb, Kenney H. (1981). "Pearson's R and Coarsely Categorized Measures". In: *American Sociological Review* 46.2, pp. 232–239. ISSN: 00031224. URL: http://www.jstor.org/stable/2094981.

Borgatta, Edgar F. (1968). "My Student, the Purist: A Lament1". In: *The Sociological Quarterly* 9.1, pp. 29–34. DOI: https://doi.org/10.1111/j.1533-8525.1968.tb02284.x. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1533-8525.1968.tb02284.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1533-8525.1968.tb02284.x.

Bray, Freddie, Ren, Jian-Song, Masuyer, Eric, and Ferlay, Jacques (2013). "Global estimates of cancer prevalence for 27 sites in the adult population in 2008". In: *International Journal of Cancer* 132.5, pp. 1133–1145. ISSN: 0020-7136. DOI: https://doi.org/10.1002/ijc.27711. URL: https://doi.org/10.1002/ijc.27711.

Breslow, A. (1970). "Thickness, cross-sectional areas and depth of invasion in the prognosis of cutaneous melanoma". In: *Ann Surg* 172.5, pp. 902–8. ISSN: 0003-4932 (Print) 0003-4932 (Linking). DOI: 10.1097/00000658-197011000-00017. URL: https://www.ncbi.nlm.nih.gov/pubmed/5477666.

Brown, Kevin M, Macgregor, Stuart, Montgomery, Grant W., Craig, David W., Zhao, Zhen Zhen, Iyadurai, Kelly, Henders, Anjali K., Homer, Nils, Campbell, Megan J., Stark, Mitchell, Thomas, Shane, Schmid, Helen, Holland, Elizabeth A., Gillanders, Elizabeth M., Duffy, David L., Maskiell, Judith A., Jetann, Jodie, Ferguson, Megan, Stephan, Dietrich A., Cust, Anne E., Whiteman, David, Green, Adele, Olsson, Håkan, Puig, Susana, Ghiorzo, Paola, Hansson, Johan, Demenais, Florence, Goldstein, Alisa M., Gruis, Nelleke A., Elder, David E., Bishop,

Julia Newton, Kefford, Richard F., Giles, Graham G., Armstrong, Bruce K., Aitken, Joanne F., Hopper, John L., Martin, Nicholas G., Trent, Jeffrey M., Mann, Graham J., and Hayward, Nicholas K. (2008). "Common sequence variants on 20q11.22 confer melanoma susceptibility". In: *Nature genetics* 40.7, pp. 838–840. ISSN: 1546-1718 1061-4036. DOI: 10.1038/ng.163. URL: https://pubmed.ncbi.nlm.nih.gov/18488026%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2755512/.

Bulik-Sullivan, Brendan K, Finucane, Hilary K, Anttila, Verneri, Gusev, Alexander, Day, Felix R, Loh, Po-Ru, ReproGen, Consortium, Psychiatric Genomics, Consortium, Genetic Consortium for Anorexia Nervosa of the Wellcome Trust Case Control, Consortium, Duncan, Laramie, Perry, John R B, Patterson, Nick, Robinson, Elise B, Daly, Mark J, Price, Alkes L, and Neale, Benjamin M (2015). "An atlas of genetic correlations across human diseases and traits". In: *Nature Genetics* 47, p. 1236. DOI: 10.1038/ng.3406https://www.nature.com/articles/ng.3406#supplementary-information. URL: https://doi.org/10.1038/ng.3406.

Bulik-Sullivan, Brendan K, Loh, Po-Ru, Finucane, Hilary K, Ripke, Stephan, Yang, Jian, Schizophrenia Working Group of the Psychiatric Genomics, Consortium, Patterson, Nick, Daly, Mark J, Price, Alkes L, and Neale, Benjamin M (2015). "LD Score regression distinguishes confounding from polygenicity in genome-wide association studies". In: *Nature Genetics* 47, p. 291. DOI: 10.1038/ng.3211https://www.nature.com/articles/ng.3211#supplementary-information. URL: https://doi.org/10.1038/ng.3211.

Bulik-Sullivan, Brendan K., Loh, Po-Ru, Finucane, Hilary K., Ripke, Stephan, Yang, Jian, Patterson, Nick, Daly, Mark J., Price, Alkes L., Neale, Benjamin M., and Schizophrenia Working Group of the Psychiatric Genomics, Consortium (2015). "LD Score regression distinguishes confounding from polygenicity in genome-wide association studies". In: *Nature Genetics* 47.3, pp. 291–295. ISSN: 1546-1718. DOI: 10.1038/ng.3211. URL: https://doi.org/10.1038/ng.3211.

Buniello, A., MacArthur, J. A. L., Cerezo, M., Harris, L. W., Hayhurst, J., Malangone, C., McMahon, A., Morales, J., Mountjoy, E., Sollis, E., Suveges, D., Vrousgou, O., Whetzel, P. L., Amode, R., Guillen, J. A., Riat, H. S., Trevanion, S. J., Hall, P., Junkins, H., Flicek, P., Burdett, T., Hindorff, L. A., Cunningham, F., and Parkinson, H. (2019). "The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary

statistics 2019". In: *Nucleic Acids Res* 47.D1, pp. D1005–D1012. ISSN: 1362-4962 (Electronic) 0305-1048 (Linking). DOI: 10.1093/nar/gky1120. URL: https://www.ncbi.nlm.nih.gov/pubmed/30445434.

Bustard, Helen K. and Smith, Robin W. (Aug. 1991). "Investigation into the scattering of light by human hair". In: *Appl. Opt.* 30.24, pp. 3485–3491. DOI: 10.1364/AO.30.003485. URL: http://ao.osa.org/abstract.cfm?URI=ao-30-24-3485.

Byrne, E. H. and Fisher, D. E. (2017). "Immune and molecular correlates in melanoma treated with immune checkpoint blockade". In: *Cancer* 123.S11, pp. 2143–2153. ISSN: 1097-0142 (Electronic) 0008-543X (Linking). DOI: 10.1002/cncr.30444. URL: https://www.ncbi.nlm.nih.gov/pubmed/28543699.

Caldwell, M. M., Björn, L. O., Bornman, J. F., Flint, S. D., Kulandaivelu, G., Teramura, A. H., and Tevini, M. (1998). "Effects of increased solar ultraviolet radiation on terrestrial ecosystems". In: *Journal of Photochemistry and Photobiology B: Biology* 46.1, pp. 40–52. ISSN: 1011-1344. DOI: https://doi.org/10.1016/S1011-1344(98)00184-5. URL: http://www.sciencedirect.com/science/article/pii/S1011134498001845.

Cancer Research UK (2014). URL: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/skincancer#heading-zero.

– (2016). URL: http://www.cancerresearchuk.org/about-cancer/melanoma/survival.

Cassady, Jennifer L. and Sturm, Richard A. (1994). "Sequence of the human dopachrome tautomerase-encoding TRP-2 cDNA". In: *Gene* 143.2, pp. 295–298. ISSN: 0378-1119. DOI: https://doi.org/10.1016/0378-1119(94)90114-7. URL: http://www.sciencedirect.com/science/article/pii/0378111994901147.

Chaplin, G. and Jablonski, N. G. (2013). "The Human Environment and the Vitamin D Compromise: Scotland as a Case Study in Human Biocultural Adaptation and Disease Susceptibility". In: *Human Biology* 85.4, pp. 529–552. ISSN: 0018-7143. DOI: Doi10.3378/027.085.0402. URL: %3CGo%20to%20ISI%3E://WOS:000330230200002.

Chapman, P. B., Hauschild, A., Robert, C., Haanen, J. B., Ascierto, P., Larkin, J., Dummer, R., Garbe, C., Testori, A., Maio, M., Hogg, D., Lorigan, P., Lebbe, C., Jouary, T., Schadendorf,

D., Ribas, A., O'Day, S. J., Sosman, J. A., Kirkwood, J. M., Eggermont, A. M., Dreno, B., Nolop, K., Li, J., Nelson, B., Hou, J., Lee, R. J., Flaherty, K. T., McArthur, G. A., and Group, Brim- Study (2011). "Improved survival with vemurafenib in melanoma with BRAF V600E mutation". In: *N Engl J Med* 364.26, pp. 2507–16. ISSN: 1533-4406 (Electronic) 0028-4793 (Linking). DOI: 10.1056/NEJMoa1103782. URL: https://www.ncbi.nlm.nih.gov/pubmed/21639808.

Chhabra, Yash, Yong, Hilary X. L., Fane, Mitchell E., Soogrim, Arish, Lim, Wen, Mahiuddin, Dayana Nur, Kim, Reuben S. Q., Ashcroft, Melinda, Beatson, Scott A., Ainger, Stephen A., Smit, Darren J., Jagirdar, Kasturee, Walker, Graeme J., Sturm, Richard A., and Smith, Aaron G. (2018). "Genetic variation in IRF4 expression modulates growth characteristics, tyrosinase expression and interferon-gamma response in melanocytic cells". In: *Pigment Cell and Melanoma Research* 31.1, pp. 51–63. ISSN: 1755-1471. DOI: https://doi.org/10.1111/pcmr.12620. URL: https://doi.org/10.1111/pcmr.12620.

Choi, Jaehwa, Peters, Michelle, and Mueller, Ralph (2010). "Correlational analysis of ordinal data: From Pearson's r to Bayesian polychoric correlation". In: *Asia Pacific Education Review* 11, pp. 459–466. DOI: 10.1007/s12564-010-9096-y.

Christensen, Kurt D., Dukhovny, Dmitry, Siebert, Uwe, and Green, Robert C. (2015). "Assessing the Costs and Cost-Effectiveness of Genomic Sequencing". In: *Journal of personalized medicine* 5.4, pp. 470–486. ISSN: 2075-4426. DOI: 10.3390/jpm5040470. URL: https://pubmed.ncbi.nlm.nih.gov/26690481%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4695866/.

Clarkson, Douglas B. and Jennrich, Robert I. (May 1988). "Quartic rotation criteria and algorithms". In: *Psychometrika* 53.2, pp. 251–259. ISSN: 1860-0980. DOI: 10.1007/BF02294136. URL: https://doi.org/10.1007/BF02294136.

Cockburn, Myles, Hamilton, Ann, and Mack, Thomas (May 2001). "Recall Bias in Self-reported Melanoma Risk Factors". In: *American Journal of Epidemiology* 153.10, pp. 1021–1026. ISSN: 0002-9262. DOI: 10.1093/aje/153.10.1021. eprint: https://academic.oup.com/aje/article-pdf/153/10/1021/279734/1021.pdf. URL: https://doi.org/10.1093/aje/153.10.1021.

Cook, Anthony L., Chen, Wei, Thurber, Amy E., Smit, Darren J., Smith, Aaron G., Bladen, Timothy G., Brown, Darren L., Duffy, David L., Pastorino, Lorenza, Bianchi-Scarra, Giovanna, Helen Leonard, J., Stow, Jennifer L., and Sturm, Richard A. (2009). "Analysis of Cultured Human Melanocytes Based on Polymorphisms within the SLC45A2/MATP, SLC24A5/NCKX5, and OCA2/P Loci". In: *Journal of Investigative Dermatology* 129.2, pp. 392–405. ISSN: 0022-202X. DOI: 10.1038/jid.2008.211. URL: https://doi.org/10.1038/jid.2008.211.

Corrie, Pippa, Hategan, Mirela, Fife, Kate, and Parkinson, Christine (2014). "Management of melanoma". In: *British Medical Bulletin* 111.1, pp. 149–162. ISSN: 0007-1420. DOI: 10.1093/bmb/ldu019. URL: https://doi.org/10.1093/bmb/ldu019.

Cowdery, Stephanie P., Stuart, Amanda L., Pasco, Julie A., Berk, Michael, Campbell, David, and Williams, Lana J. (2020). "Validity of self-reported cancer: Comparison between self-report versus cancer registry records in the Geelong Osteoporosis Study". In: *Cancer Epidemiology* 68, p. 101790. ISSN: 1877-7821. DOI: https://doi.org/10.1016/j.canep.2020.101790. URL: https://www.sciencedirect.com/science/article/pii/S1877782120301247.

Crawford, Nicholas G., Kelly, Derek E., Hansen, Matthew E. B., Beltrame, Marcia H., Fan, Shaohua, Bowman, Shanna L., Jewett, Ethan, Ranciaro, Alessia, Thompson, Simon, Lo, Yancy, Pfeifer, Susanne P., Jensen, Jeffrey D., Campbell, Michael C., Beggs, William, Hormozdiari, Farhad, Mpoloka, Sununguko Wata, Mokone, Gaonyadiwe George, Nyambo, Thomas, Meskel, Dawit Wolde, Belay, Gurja, Haut, Jake, Rothschild, Harriet, Zon, Leonard, Zhou, Yi, Kovacs, Michael A., Xu, Mai, Zhang, Tongwu, Bishop, Kevin, Sinclair, Jason, Rivas, Cecilia, Elliot, Eugene, Choi, Jiyeon, Li, Shengchao A., Hicks, Belynda, Burgess, Shawn, Abnet, Christian, Watkins-Chow, Dawn E., Oceana, Elena, Song, Yun S., Eskin, Eleazar, Brown, Kevin M., Marks, Michael S., Loftus, Stacie K., Pavan, William J., Yeager, Meredith, Chanock, Stephen, and Tishkoff, Sarah A. (2017). "Loci associated with skin pigmentation identified in African populations". In: *Science* 358.6365, eaan8433. DOI: 10.1126/science.aan8433. URL: http://science.sciencemag.org/content/358/6365/eaan8433.abstract.

Crombie, I. K. (1979). "Variation of melanoma incidence with latitude in North America and Europe". In: *British journal of cancer* 40.5, pp. 774–781. ISSN: 0007-0920 1532-1827. DOI:

10.1038/bjc.1979.260. URL: https://pubmed.ncbi.nlm.nih.gov/508580%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2010099/.

Cust, Anne E., Armstrong, Bruce K., Goumas, Chris, Jenkins, Mark A., Schmid, Helen, Hopper, John L., Kefford, Richard F., Giles, Graham G., Aitken, Joanne F., and Mann, Graham J. (2011). "Sunbed use during adolescence and early adulthood is associated with increased risk of early-onset melanoma". In: *International Journal of Cancer* 128.10, pp. 2425–2435. ISSN: 0020-7136. DOI: 10.1002/ijc.25576. URL: https://doi.org/10.1002/ijc.25576.

Cust, Anne E., Drummond, Martin, Kanetsky, Peter A., Mann, Graham J., Cust, Anne E., Schmid, Helen, Hopper, John L., Aitken, Joanne F., Armstrong, Bruce K., Giles, Graham G., Holland, Elizabeth, Kefford, Richard F., Jenkins, Mark A., Newton Bishop, Julia A., Affleck, Paul, Barrett, Jennifer H., Bishop, D. Timothy, Harrison, Jane, Iles, Mark M., Randerson-Moor, Juliette, Harland, Mark, Taylor, John C., Whittaker, Linda, Kukalizch, Kairen, Leake, Susan, Karpavicius, Birute, Haynes, Sue, Mack, Tricia, Chan, May, Taylor, Yvonne, Davies, John, King, Paul, Goldstein, Alisa M., Barrett, Jennifer H., MacGregor, Stuart, Law, Matthew H., Iles, Mark M., Bui, Minh, Hopper, John L., Brossard, Myriam, Demenais, Florence, Taylor, John C., Hoggart, Clive, Brown, Kevin M., Landi, Maria Teresa, Newton-Bishop, Julia A., Mann, Graham J., and Bishop, D. Timothy (2018). "Assessing the Incremental Contribution of Common Genomic Variants to Melanoma Risk Prediction in Two Population-Based Studies". In: *Journal of Investigative Dermatology* 138.12, pp. 2617–2624. ISSN: 0022-202X. DOI: https://doi.org/10.1016/j.jid.2018.05.023. URL: http://www.sciencedirect.com/science/article/pii/S0022202X18320463%20https://spiral.imperial.ac.uk:8443/bitstream/10044/1/66082/2/Assessing%20the%20Incremental%20Contribution%20of%20Common%20Genomic%20Variants%20to%20Melanoma%20Risk%20Prediction%20in%20Two%20Population-Based%20Studies.pdf.

Cust, Anne E., Pickles, Kristen M., Goumas, Chris, Vu, Thao, Schmid, Helen, Nagore, Eduardo, Kelly, John, Aitken, Joanne F., Giles, Graham G., Hopper, John L., Jenkins, Mark A., and Mann, Graham J. (2015). "Accuracy of self-reported nevus and pigmentation phenotype compared with clinical assessment in a population-based study of young Australian adults". In: *Cancer epidemiology, biomarkers and prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*

24.4, pp. 736–743. ISSN: 1538-7755 1055-9965. DOI: 10.1158/1055-9965.EPI-14-1203. URL: https://pubmed.ncbi.nlm.nih.gov/25628333%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4698346/.

Dai, Hongying, Bhandary, Madhusudan, Becker, Mara, Leeder, J. Steven, Gaedigk, Roger, and Motsinger-Reif, Alison A. (May 2012). "Global tests of P-values for multifactor dimensionality reduction models in selection of optimal number of target genes". In: *BioData Mining* 5.1, p. 3. ISSN: 1756-0381. DOI: 10.1186/1756-0381-5-3. URL: https://doi.org/10.1186/1756-0381-5-3.

DeLong, Elizabeth R., DeLong, David M., and Clarke-Pearson, Daniel L. (1988). "Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach". In: *Biometrics* 44.3, pp. 837–845. ISSN: 0006341X, 15410420. URL: http://www.jstor.org/stable/2531595.

Dickson, P. V. and Gershenwald, J. E. (2011). "Staging and prognosis of cutaneous melanoma". In: *Surg Oncol Clin N Am* 20.1, pp. 1–17. ISSN: 1558-5042 (Electronic) 1055-3207 (Linking). DOI: 10.1016/j.soc.2010.09.007. URL: https://www.ncbi.nlm.nih.gov/pubmed/21111956.

Duffy, David L., Box, Neil F., Chen, Wei, Palmer, James S., Montgomery, Grant W., James, Michael R., Hayward, Nicholas K., Martin, Nicholas G., and Sturm, Richard A. (Jan. 2004). "Interactive effects of MC1R and OCA2 on melanoma risk phenotypes". In: *Human Molecular Genetics* 13.4, pp. 447–461. ISSN: 0964-6906. DOI: 10.1093/hmg/ddh043. eprint: https://academic.oup.com/hmg/article-pdf/13/4/447/1984041/ddh043.pdf. URL: https://doi.org/10.1093/hmg/ddh043.

Duffy, David L., Zhu, Gu, et al. (2018). "Novel pleiotropic risk loci for melanoma and nevus density implicate multiple biological pathways". In: *Nature Communications* 9.1, p. 4774. ISSN: 2041-1723. DOI: 10.1038/s41467-018-06649-5. URL: https://doi.org/10.1038/s41467-018-06649-5%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6235897/pdf/41467_2018_Article_6649.pdf.

Evans, D. Gareth, Astley, Susan, Stavrinos, Paula, Harkness, Elaine, Donnelly, Louise S., Dawe, Sarah, Jacob, Ian, Harvie, Michelle, Cuzick, Jack, Brentnall, Adam, Wilson, Mary, Harrison,

Fiona, Payne, Katherine, and Howell, Anthony (2016). "Improvement in risk prediction, early detection and prevention of breast cancer in the NHS Breast Screening Programme and family history clinics: a dual cohort study". In: URL: http://europepmc.org/abstract/MED/27559559%20http://europepmc.org/books/NBK379488%20https://www.ncbi.nlm.nih.gov/books/NBK379488.

Fadista, João, Manning, Alisa K., Florez, Jose C., and Groop, Leif (2016). "The (in)famous GWAS P-value threshold revisited and updated for low-frequency variants". In: *European journal of human genetics : EJHG* 24.8, pp. 1202–1205. ISSN: 1476-5438 1018-4813. DOI: 10.1038/ejhg.2015.269. URL: https://pubmed.ncbi.nlm.nih.gov/26733288%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4970684/.

Falchi, Mario, Bataille, Veronique, Hayward, Nicholas K., Duffy, David L., Bishop, Julia A. Newton, Pastinen, Tomi, Cervino, Alessandra, Zhao, Zhen Z., Deloukas, Panos, Soranzo, Nicole, Elder, David E., Barrett, Jennifer H., Martin, Nicholas G., Bishop, D. Timothy, Montgomery, Grant W., and Spector, Timothy D. (2009). "Genome-wide association study identifies variants at 9p21 and 22q13 associated with development of cutaneous nevi". In: *Nature genetics* 41.8, pp. 915–919. ISSN: 1546-1718 1061-4036. DOI: 10.1038/ng.410. URL: https://pubmed.ncbi.nlm.nih.gov/19578365%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3080738/.

Fang, Shenying, Han, Jiali, Zhang, Mingfeng, Wang, Li-e, Wei, Qingyi, Amos, Christopher I., and Lee, Jeffrey E. (2014). "Joint Effect of Multiple Common SNPs Predicts Melanoma Susceptibility". In: *PLOS ONE* 8.12, e85642. DOI: 10.1371/journal.pone.0085642. URL: https://doi.org/10.1371/journal.pone.0085642.

Gandini, S., Sera, F., Cattaruzza, M. S., Pasquini, P., Abeni, D., Boyle, P., and Melchi, C. F. (2005). "Meta-analysis of risk factors for cutaneous melanoma: I. Common and atypical naevi". In: *Eur J Cancer* 41.1, pp. 28–44. ISSN: 0959-8049 (Print) 0959-8049 (Linking). DOI: 10.1016/j.ejca.2004.10.015. URL: https://www.ncbi.nlm.nih.gov/pubmed/15617989.

Gandini, Sara, Sera, Francesco, Cattaruzza, Maria Sofia, Pasquini, Paolo, Abeni, Damiano, Boyle, Peter, and Melchi, Carmelo Francesco (2005). "Meta-analysis of risk factors for cutaneous melanoma: I. Common and atypical naevi". In: *European Journal of Cancer* 41.1,

pp. 28–44. ISSN: 0959-8049. DOI: https://doi.org/10.1016/j.ejca.2004.10.015. URL: http://www.sciencedirect.com/science/article/pii/S0959804904008329.

García-Borrón, Jose C., Abdel-Malek, Zalfa, and Jiménez-Cervantes, Celia (Sept. 2014). "MC1R, the cAMP pathway, and the response to solar UV: extending the horizon beyond pigmentation". eng. In: *Pigment cell & melanoma research* 27.5. PMC4150834[pmcid], pp. 699–720. ISSN: 1755-148X. DOI: 10.1111/pcmr.12257. URL: https://doi.org/10.1111/pcmr.12257.

George, D. D., Armenio, V. A., and Katz, S. C. (2017). "Combinatorial immunotherapy for melanoma". In: *Cancer Gene Ther* 24.3, pp. 141–147. ISSN: 1476-5500 (Electronic) 0929-1903 (Linking). DOI: 10.1038/cgt.2016.56. URL: https://www.ncbi.nlm.nih.gov/pubmed/27834353.

Gerstenblith, Meg R., Shi, Jianxin, and Landi, Maria Teresa (2010). "Genome-wide association studies of pigmentation and skin cancer: a review and meta-analysis". In: *Pigment cell and melanoma research* 23.5, pp. 587–606. ISSN: 1755-148X 1755-1471. DOI: 10.1111/j.1755-148X.2010.00730.x. URL: https://pubmed.ncbi.nlm.nih.gov/20546537%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3179913/.

Goldstein, Alisa M., Chan, May, Harland, Mark, Gillanders, Elizabeth M., Hayward, Nicholas K., Avril, Marie-Francoise, Azizi, Esther, Bianchi-Scarra, Giovanna, Bishop, D. Timothy, Bressac-de Paillerets, Brigitte, Bruno, William, Calista, Donato, Cannon Albright, Lisa A., Demenais, Florence, Elder, David E., Ghiorzo, Paola, Gruis, Nelleke A., Hansson, Johan, Hogg, David, Holland, Elizabeth A., Kanetsky, Peter A., Kefford, Richard F., Landi, Maria Teresa, Lang, Julie, Leachman, Sancy A., MacKie, Rona M., Magnusson, Veronica, Mann, Graham J., Niendorf, Kristin, Newton Bishop, Julia, Palmer, Jane M., Puig, Susana, Puig-Butille, Joan A., Snoo, Femke A. de, Stark, Mitchell, Tsao, Hensin, Tucker, Margaret A., Whitaker, Linda, and Yakobson, Emanuel (2006). "High-risk Melanoma Susceptibility Genes and Pancreatic Cancer, Neural System Tumors, and Uveal Melanoma across GenoMEL". In: *Cancer Research* 66.20, p. 9818. DOI: 10.1158/0008-5472.CAN-06-0494. URL: http://cancerres.aacrjournals.org/content/66/20/9818.abstract.

Green, Richard E., Krause, Johannes, Briggs, Adrian W., Maricic, Tomislav, Stenzel, Udo, Kircher, Martin, Patterson, Nick, Li, Heng, Zhai, Weiwei, Fritz, Markus Hsi-Yang, Hansen,

Nancy F., Durand, Eric Y., Malaspinas, Anna-Sapfo, Jensen, Jeffrey D., Marques-Bonet, Tomas, Alkan, Can, Prüfer, Kay, Meyer, Matthias, Burbano, Hernán A., Good, Jeffrey M., Schultz, Rigo, Aximu-Petri, Ayinuer, Butthof, Anne, Höber, Barbara, Höffner, Barbara, Siegemund, Madlen, Weihmann, Antje, Nusbaum, Chad, Lander, Eric S., Russ, Carsten, Novod, Nathaniel, Affourtit, Jason, Egholm, Michael, Verna, Christine, Rudan, Pavao, Brajkovic, Dejana, Kucan, Željko, Gušic, Ivan, Doronichev, Vladimir B., Golovanova, Liubov V., Lalueza-Fox, Carles, Rasilla, Marco de la, Fortea, Javier, Rosas, Antonio, Schmitz, Ralf W., Johnson, Philip L. F., Eichler, Evan E., Falush, Daniel, Birney, Ewan, Mullikin, James C., Slatkin, Montgomery, Nielsen, Rasmus, Kelso, Janet, Lachmann, Michael, Reich, David, and Pääbo, Svante (2010). "A draft sequence of the Neandertal genome". In: *Science (New York, N.Y.)* 328.5979, pp. 710–722. ISSN: 1095-9203 0036-8075. DOI: 10.1126/science.1188021. URL: https://pubmed.ncbi.nlm.nih.gov/20448178%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5100745/.

Gudbjartsson, Daniel F., Sulem, Patrick, Stacey, Simon N., Goldstein, Alisa M., Rafnar, Thorunn, Sigurgeirsson, Bardur, Benediktsdottir, Kristrun R., Thorisdottir, Kristin, Ragnarsson, Rafn, Sveinsdottir, Steinunn G., Magnusson, Veronica, Lindblom, Annika, Kostulas, Konstantinos, Botella-Estrada, Rafael, Soriano, Virtudes, Juberías, Pablo, Grasa, Matilde, Saez, Berta, Andres, Raquel, Scherer, Dominique, Rudnai, Peter, Gurzau, Eugene, Koppova, Kvetoslava, Kiemeney, Lambertus A., Jakobsdottir, Margret, Steinberg, Stacy, Helgason, Agnar, Gretarsdottir, Solveig, Tucker, Margaret A., Mayordomo, José I., Nagore, Eduardo, Kumar, Rajiv, Hansson, Johan, Olafsson, Jon H., Gulcher, Jeffrey, Kong, Augustine, Thorsteinsdottir, Unnur, and Stefansson, Kari (2008). "ASIP and TYR pigmentation variants associate with cutaneous melanoma and basal cell carcinoma". In: *Nature Genetics* 40.7, pp. 886–891. ISSN: 1546-1718. DOI: 10.1038/ng.161. URL: https://doi.org/10.1038/ng.161.

Gusev, Alexander, Ko, Arthur, Shi, Huwenbo, Bhatia, Gaurav, Chung, Wonil, Penninx, Brenda W. J. H., Jansen, Rick, Geus, Eco J. C. de, Boomsma, Dorret I., Wright, Fred A., Sullivan, Patrick F., Nikkola, Elina, Alvarez, Marcus, Civelek, Mete, Lusis, Aldons J., Lehtimäki, Terho, Raitoharju, Emma, Kähönen, Mika, Seppälä, Ilkka, Raitakari, Olli T., Kuusisto, Johanna, Laakso, Markku, Price, Alkes L., Pajukanta, Päivi, and Pasaniuc, Bogdan (2016). "Integrative approaches for large-scale transcriptome-wide association studies". In: *Nature Genetics* 48.3,

pp. 245–252. ISSN: 1546-1718. DOI: 10.1038/ng.3506. URL: https://doi.org/10.1038/ng.3506.

Han, Jiali, Colditz, Graham A, and Hunter, David J (Aug. 2006). "Risk factors for skin cancers: a nested case–control study within the Nurses' Health Study". In: *International Journal of Epidemiology* 35.6, pp. 1514–1521. ISSN: 0300-5771. DOI: 10.1093/ije/dyl197. eprint: https://academic.oup.com/ije/article-pdf/35/6/1514/1878391/dyl197.pdf. URL: https://doi.org/10.1093/ije/dyl197.

Harbour, J. William, Onken, Michael D., Roberson, Elisha D. O., Duan, Shenghui, Cao, Li, Worley, Lori A., Council, M. Laurin, Matatall, Katie A., Helms, Cynthia, and Bowcock, Anne M. (2010). "Frequent mutation of BAP1 in metastasizing uveal melanomas". In: *Science (New York, N.Y.)* 330.6009, pp. 1410–1413. ISSN: 1095-9203 0036-8075. DOI: 10.1126/science.1194472. URL: https://pubmed.ncbi.nlm.nih.gov/21051595%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3087380/.

Harland, Mark, Petljak, Mia, Robles-Espinoza, Carla Daniela, Ding, Zhihao, Gruis, Nelleke A., Doorn, Remco van, Pooley, Karen A., Dunning, Alison M., Aoude, Lauren G., Wadt, Karin A. W., Gerdes, Anne-Marie, Brown, Kevin M., Hayward, Nicholas K., Newton-Bishop, Julia A., Adams, David J., and Bishop, D. Timothy (2016). "Germline TERT promoter mutations are rare in familial melanoma". In: *Familial cancer* 15.1, pp. 139–144. ISSN: 1573-7292 1389-9600. DOI: 10.1007/s10689-015-9841-9. URL: https://pubmed.ncbi.nlm.nih.gov/26433962%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4698275/.

Hepp, Diego, Gonçalves, Gislene Lopes, Moreira, Gilson Rudinei Pires, and Freitas, Thales Renato Ochotorena de (June 2016). "Epistatic Interaction of the Melanocortin 1 Receptor and Agouti Signaling Protein Genes Modulates Wool Color in the Brazilian Creole Sheep". In: *Journal of Heredity* 107.6, pp. 544–552. ISSN: 0022-1503. DOI: 10.1093/jhered/esw037. eprint: https://academic.oup.com/jhered/article-pdf/107/6/544/19576989/esw037.pdf. URL: https://doi.org/10.1093/jhered/esw037.

Herold, C, Hooli, BV, Mullin, K, Liu, T, Roehr, JT, Mattheisen, M, Parrado, AR, Bertram, L, Lange, C, and Tanzi, RE (Nov. 2016). "Family-based association analyses of imputed genotypes reveal genome-wide significant association of Alzheimer's disease with OSBPL6,

PTPRG, and PDCL3". In: *Molecular psychiatry* 21.11, pp. 1608–1612. ISSN: 1359-4184. DOI: 10.1038/mp.2015.218. URL: https://europepmc.org/articles/PMC4970971.

Holman, C. and Armstrong, B. (1984). "Pigmentary traits, ethnic origin, benign nevi, and family history as risk factors for cutaneous malignant melanoma." In: *Journal of the National Cancer Institute* 72 2, pp. 257–66.

Hsieh, Ai-Ru, Hsiao, Ching-Lin, Chang, Su-Wei, Wang, Hui-Min, and Fann, Cathy S.J. (2011). "On the use of multifactor dimensionality reduction (MDR) and classification and regression tree (CART) to identify haplotype–haplotype interactions in genetic studies". In: *Genomics* 97.2, pp. 77–85. ISSN: 0888-7543. DOI: https://doi.org/10.1016/j.ygeno.2010.11.003. URL: http://www.sciencedirect.com/science/article/pii/S0888754310002405.

Hu, Li-tze and Bentler, Peter M. (1999). "Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives". In: *Structural Equation Modeling: A Multidisciplinary Journal* 6.1, pp. 1–55. DOI: 10.1080/10705519909540118. eprint: https://doi.org/10.1080/10705519909540118. URL: https://doi.org/10.1080/10705519909540118.

Hysi, Pirro G., Valdes, Ana M., Liu, Fan, Furlotte, Nicholas A., Evans, David M., Bataille, Veronique, Visconti, Alessia, Hemani, Gibran, McMahon, George, Ring, Susan M., Smith, George Davey, Duffy, David L., Zhu, Gu, Gordon, Scott D., Medland, Sarah E., Lin, Bochao D., Willemsen, Gonneke, Jan Hottenga, Jouke, Vuckovic, Dragana, Girotto, Giorgia, Gandin, Ilaria, Sala, Cinzia, Concas, Maria Pina, Brumat, Marco, Gasparini, Paolo, Toniolo, Daniela, Cocca, Massimiliano, Robino, Antonietta, Yazar, Seyhan, Hewitt, Alex W., Chen, Yan, Zeng, Changqing, Uitterlinden, Andre G., Ikram, M. Arfan, Hamer, Merel A., Duijn, Cornelia M. van, Nijsten, Tamar, Mackey, David A., Falchi, Mario, Boomsma, Dorret I., Martin, Nicholas G., Hinds, David A., Kayser, Manfred, Spector, Timothy D., and The International Visible Trait Genetics, Consortium (2018). "Genome-wide association meta-analysis of individuals of European ancestry identifies new loci explaining a substantial fraction of hair color variation and heritability". In: *Nature Genetics* 50.5, pp. 652–656. ISSN: 1546-1718. DOI: 10.1038/s41588-018-0100-5. URL: https://doi.org/10.1038/s41588-018-0100-5.

Iles, Mark M., Bishop, D. Timothy, Taylor, John C., Hayward, Nicholas K., Brossard, Myriam, Cust, Anne E., Dunning, Alison M., Lee, Jeffrey E., Moses, Eric K., Akslen, Lars A., Investigators, Amfs, Andresen, Per A., Avril, Marie-Françoise, Azizi, Esther, Scarrà, Giovanna Bianchi, Brown, Kevin M., Dębniak, Tadeusz, Elder, David E., Friedman, Eitan, Ghiorzo, Paola, Gillanders, Elizabeth M., Goldstein, Alisa M., Gruis, Nelleke A., Hansson, Johan, Harland, Mark, Helsing, Per, Hočevar, Marko, Höiom, Veronica, investigators, I. B. D., Ingvar, Christian, Kanetsky, Peter A., Landi, Maria Teresa, Lang, Julie, Lathrop, G. Mark, Lubiński, Jan, Mackie, Rona M., Martin, Nicholas G., Molven, Anders, Montgomery, Grant W., Novaković, Srdjan, Olsson, Håkan, Puig, Susana, Puig-Butille, Joan Anton, Qmega, Investigators, Qtwin, Radford-Smith, Graham L., Randerson-Moor, Juliette, Group, S. D. H. Study, Stoep, Nienke van der, Doorn, Remco van, Whiteman, David C., MacGregor, Stuart, Pooley, Karen A., Ward, Sarah V., Mann, Graham J., Amos, Christopher I., Pharoah, Paul D. P., Demenais, Florence, Law, Matthew H., Newton Bishop, Julia A., Barrett, Jennifer H., and Geno, M. E. L. Consortium (2014). "The effect on melanoma risk of genes previously associated with telomere length". In: *Journal of the National Cancer Institute* 106.10, dju267. ISSN: 1460-2105 0027-8874. DOI: 10.1093/jnci/dju267. URL: https://pubmed.ncbi.nlm.nih.gov/25231748%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4196080/.

Ito, S. and Wakamatsu, K. (2011). "Diversity of human hair pigmentation as studied by chemical analysis of eumelanin and pheomelanin". In: *Journal of the European Academy of Dermatology and Venereology* 25.12, pp. 1369–1380. ISSN: 0926-9959. DOI: 10.1111/j.1468-3083.2011.04278.x. URL: https://doi.org/10.1111/j.1468-3083.2011.04278.x.

Ito, Shosuke, Nakanishi, Yukiko, Valenzuela, Robert K., Brilliant, Murray H., Kolbe, Ludger, and Wakamatsu, Kazumasa (2011). "Usefulness of alkaline hydrogen peroxide oxidation to analyze eumelanin and pheomelanin in various tissue samples: application to chemical analysis of human hair melanins". In: *Pigment Cell and Melanoma Research* 24.4, pp. 605–613. ISSN: 1755-1471. DOI: https://doi.org/10.1111/j.1755-148X.2011.00864.x. URL: https://doi.org/10.1111/j.1755-148X.2011.00864.x.

Ito, Shosuke and Wakamatsu, Kazumasa (2003). "Quantitative Analysis of Eumelanin and Pheomelanin in Humans, Mice, and Other Animals: a Comparative Review". In: *Pigment*

Cell Research 16.5, pp. 523–531. ISSN: 0893-5785. DOI: https://doi.org/10.1034/j.1600-0749.2003.00072.x. URL: https://doi.org/10.1034/j.1600-0749.2003.00072.x.

– (2011). "Human hair melanins: what we have learned and have not learned from mouse coat color pigmentation". In: *Pigment Cell and Melanoma Research* 24.1, pp. 63–74. ISSN: 1755-1471. DOI: https://doi.org/10.1111/j.1755-148X.2010.00755.x. URL: https://doi.org/10.1111/j.1755-148X.2010.00755.x.

Jablonski, Nina and Chaplin, George (2010). "Human Skin Pigmentation as an Adaptation to UV Radiation". In: *Proceedings of the National Academy of Sciences of the United States of America* 107 Suppl 2, pp. 8962–8. DOI: 10.1073/pnas.0914628107.

Jablonski, Nina G. and Chaplin, George (2000). "The evolution of human skin coloration". In: *Journal of Human Evolution* 39.1, pp. 57–106. ISSN: 0047-2484. DOI: https://doi.org/10.1006/jhev.2000.0403. URL: http://www.sciencedirect.com/science/article/pii/S0047248400904032.

Jackson, J. Edward (2005). "Quartimax Rotation". In: *Encyclopedia of Biostatistics*. American Cancer Society. ISBN: 9780470011812. DOI: https://doi.org/10.1002/0470011815.b2a10076. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/0470011815.b2a10076. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/0470011815.b2a10076.

Jacobs, Leonie C., Hamer, Merel A., Gunn, David A., Deelen, Joris, Lall, Jaspal S., van Heemst, Diana, Uh, Hae-Won, Hofman, Albert, Uitterlinden, André G., Griffiths, Christopher E.M., Beekman, Marian, Slagboom, P. Eline, Kayser, Manfred, Liu, Fan, and Nijsten, Tamar (2015). "A Genome-Wide Association Study Identifies the Skin Color Genes IRF4, MC1R, ASIP, and BNC2 Influencing Facial Pigmented Spots". In: *Journal of Investigative Dermatology* 135.7, pp. 1735–1742. ISSN: 0022-202X. DOI: https://doi.org/10.1038/jid.2015.62. URL: https://www.sciencedirect.com/science/article/pii/S0022202X15373280.

Jagirdar, Kasturee, Smit, Darren J., Ainger, Stephen A., Lee, Katie J., Brown, Darren L., Chapman, Brett, Zhen Zhao, Zhen, Montgomery, Grant W., Martin, Nicholas G., Stow, Jennifer L., Duffy, David L., and Sturm, Richard A. (2014). "Molecular analysis of common polymorphisms within the human Tyrosinase locus and genetic association with pigmentation traits".

In: *Pigment cell and melanoma research* 27.4, pp. 552–564. ISSN: 1755-148X 1755-1471. DOI: 10.1111/pcmr.12253. URL: https://pubmed.ncbi.nlm.nih.gov/24739399%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4119297/.

Jannot, Anne-Sophie, Ehret, Georg, and Perneger, Thomas (2015). "P 5 times 10 to the power -8 has emerged as a standard of statistical significance for genome-wide association studies". In: *Journal of Clinical Epidemiology* 68.4, pp. 460–465. ISSN: 0895-4356. DOI: 10.1016/j.jclinepi.2015.01.001. URL: https://doi.org/10.1016/j.jclinepi.2015.01.001.

Jo, Seong Kyeong, Lee, Ji Yeon, Lee, Young, Kim, Chang Deok, Lee, Jeung-Hoon, and Lee, Young Ho (2018). "Three Streams for the Mechanism of Hair Graying". In: *Annals of dermatology* 30.4, pp. 397–401. ISSN: 1013-9087 2005-3894. DOI: 10.5021/ad.2018.30.4.397. URL: https://pubmed.ncbi.nlm.nih.gov/30065578%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6029974/.

Kaiser, Henry F. (Sept. 1958). "The varimax criterion for analytic rotation in factor analysis". In: *Psychometrika* 23.3, pp. 187–200. DOI: 10.1007/BF02289233.

– (1974). "A NOTE ON THE EQUAMAX CRITERION". In: *Multivariate Behavioral Research* 9.4. PMID: 26754605, pp. 501–503. DOI: 10.1207/s15327906mbr0904_9. eprint: https://doi.org/10.1207/s15327906mbr0904_9. URL: https://doi.org/10.1207/s15327906mbr0904_9.

Kamb, A., Shattuck-Eidens, D., Eeles, R., Liu, Q., Gruis, N. A., Ding, W., Hussey, C., Tran, T., Miki, Y., Weaver-Feldhaus, J., McClure, M., Aitken, J. F., Anderson, D. E., Bergman, W., Frants, R., Goldgar, D. E., Green, A., MacLennan, R., Martin, N. G., Meyer, L. J., Youl, P., Zone, J. J., Skolnick, M. H., and Cannon-Albright, L. A. (1994). "Analysis of the p16 gene (CDKN2) as a candidate for the chromosome 9p melanoma susceptibility locus". In: *Nature Genetics* 8.1, pp. 22–26. ISSN: 1546-1718. DOI: 10.1038/ng0994-22. URL: https://doi.org/10.1038/ng0994-22.

Kasamatsu, Shinya, Hachiya, Akira, Higuchi, Kazuhiko, Ohuchi, Atsushi, Kitahara, Takashi, and Boissy, Raymond E. (2008). "Production of the Soluble Form of KIT, s-KIT, Abolishes Stem Cell Factor-Induced Melanogenesis in Human Melanocytes". In: *Journal of Investigative*

Dermatology 128.7, pp. 1763–1772. ISSN: 0022-202X. DOI: 10.1038/jid.2008.9. URL: https://doi.org/10.1038/jid.2008.9.

Kocarnik, Jonathan M., Park, Sungshim Lani, Han, Jiali, Dumitrescu, Logan, Cheng, Iona, Wilkens, Lynne R., Schumacher, Fredrick R., Kolonel, Laurence, Carlson, Chris S., Crawford, Dana C., Goodloe, Robert J., Dilks, Holli, Baker, Paxton, Richardson, Danielle, Ambite, José Luis, Song, Fengju, Quresh, Abrar A., Zhang, Mingfeng, Duggan, David, Hutter, Carolyn, Hindorff, Lucia A., Bush, William S., Kooperberg, Charles, Le Marchand, Loic, and Peters, Ulrike (2014). "Replication of associations between GWAS SNPs and melanoma risk in the Population Architecture Using Genomics and Epidemiology (PAGE) Study". In: *The Journal of investigative dermatology* 134.7, pp. 2049–2052. ISSN: 1523-1747 0022-202X. DOI: 10.1038/jid.2014.53. URL: https://pubmed.ncbi.nlm.nih.gov/24480881%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4057959/.

Koller, K. M., Wang, W., Schell, T. D., Cozza, E. M., Kokolus, K. M., Neves, R. I., Mackley, H. B., Pameijer, C., Leung, A., Anderson, B., Mallon, C. A., Robertson, G., and Drabick, J. J. (2016). "Malignant melanoma-The cradle of anti-neoplastic immunotherapy". In: *Crit Rev Oncol Hematol* 106, pp. 25–54. ISSN: 1879-0461 (Electronic) 1040-8428 (Linking). DOI: 10.1016/j.critrevonc.2016.04.010. URL: https://www.ncbi.nlm.nih.gov/pubmed/27637351.

Kypreou, Katerina P., Stefanaki, Irene, Antonopoulou, Kyriaki, Karagianni, Fani, Ntritsos, Georgios, Zaras, Alexios, Nikolaou, Vasiliki, Kalfa, Iro, Chasapi, Vasiliki, Polydorou, Dorothea, Gogas, Helen, Spyrou, George M., Bertram, Lars, Lill, Christina M., Ioannidis, John P. A., Antoniou, Christina, Evangelou, Evangelos, and Stratigos, Alexander I. (2016). "Prediction of Melanoma Risk in a Southern European Population Based on a Weighted Genetic Risk Score". In: *Journal of Investigative Dermatology* 136.3, pp. 690–695. ISSN: 0022-202X. DOI: https://doi.org/10.1016/j.jid.2015.12.007. URL: http://www.sciencedirect.com/science/article/pii/S0022202X15002134.

Lander, Eric S. et al. (2001). "Initial sequencing and analysis of the human genome". In: *Nature* 409.6822, pp. 860–921. ISSN: 1476-4687. DOI: 10.1038/35057062. URL: https://doi.org/10.1038/35057062.

Landi, Maria Teresa et al. (2020). "Genome-wide association meta-analyses combining multiple risk phenotypes provide insights into the genetic architecture of cutaneous melanoma susceptibility". In: *Nature Genetics* 52.5, pp. 494–504. ISSN: 1546-1718. DOI: 10.1038/s41588-020-0611-8. URL: https://doi.org/10.1038/s41588-020-0611-8%20https://www.nature.com/articles/s41588-020-0611-8.

Latourelle, Jeanne C, Pankratz, Nathan, Dumitriu, Alexandra, Wilk, Jemma B, Goldwurm, Stefano, Pezzoli, Gianni, Mariani, Claudio B, DeStefano, Anita L, Halter, Cheryl, Gusella, James F, Nichols, William C, Myers, Richard H, Foroud, Tatiana, PROGENI Investigators, Coordinators and Molecular Genetic Laboratories, and GenePD Investigators, Coordinators and Molecular Genetic Laboratories (Sept. 2009). "Genomewide association study for onset age in Parkinson disease". In: *BMC medical genetics* 10, p. 98. ISSN: 1471-2350. DOI: 10.1186/1471-2350-10-98. URL: https://europepmc.org/articles/PMC2758866.

Law, Matthew H., Bishop, D. Timothy, Lee, Jeffrey E., Brossard, Myriam, Martin, Nicholas G., Moses, Eric K., Song, Fengju, Barrett, Jennifer H., Kumar, Rajiv, Easton, Douglas F., Pharoah, Paul D. P., Swerdlow, Anthony J., Kypreou, Katerina P., Taylor, John C., Harland, Mark, Randerson-Moor, Juliette, Akslen, Lars A., Andresen, Per A., Avril, Marie-Françoise, Azizi, Esther, Scarrà, Giovanna Bianchi, Brown, Kevin M., Dębniak, Tadeusz, Duffy, David L., Elder, David E., Fang, Shenying, Friedman, Eitan, Galan, Pilar, Ghiorzo, Paola, Gillanders, Elizabeth M., Goldstein, Alisa M., Gruis, Nelleke A., Hansson, Johan, Helsing, Per, Hočevar, Marko, Höiom, Veronica, Ingvar, Christian, Kanetsky, Peter A., Chen, Wei V., Geno, M. E. L. Consortium, Essen-Heidelberg, Investigators, Group, S. D. H. Study, Q, Mega, Investigators, Qtwin, Investigators, Amfs, Group, Athens Melanoma Study, Landi, Maria Teresa, Lang, Julie, Lathrop, G. Mark, Lubiński, Jan, Mackie, Rona M., Mann, Graham J., Molven, Anders, Montgomery, Grant W., Novaković, Srdjan, Olsson, Håkan, Puig, Susana, Puig-Butille, Joan Anton, Qureshi, Abrar A., Radford-Smith, Graham L., Stoep, Nienke van der, Doorn, Remco van, Whiteman, David C., Craig, Jamie E., Schadendorf, Dirk, Simms, Lisa A., Burdon, Kathryn P., Nyholt, Dale R., Pooley, Karen A., Orr, Nick, Stratigos, Alexander J., Cust, Anne E., Ward, Sarah V., Hayward, Nicholas K., Han, Jiali, Schulze, Hans-Joachim, Dunning, Alison M., Bishop, Julia A. Newton, Demenais, Florence, Amos, Christopher I., MacGregor, Stuart, and Iles, Mark M. (2015). "Genome-wide meta-analysis identifies five new susceptibility loci for cutaneous malignant melanoma". In: *Nature genetics* 47.9, pp. 987–995.

ISSN: 1546-1718 1061-4036. DOI: 10.1038/ng.3373. URL: https://www.ncbi.nlm.nih.gov/pubmed/26237428%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4557485/.

Law, Matthew H., Medland, Sarah E., Zhu, Gu, Yazar, Seyhan, Viñuela, Ana, Wallace, Leanne, Shekar, Sri Niranjan, Duffy, David L., Bataille, Veronique, Glass, Dan, Spector, Tim D., Wood, Diane, Gordon, Scott D., Barbour, Julie M., Henders, Anjali K., Hewitt, Alex W., Montgomery, Grant W., Sturm, Richard A., Mackey, David A., Green, Adèle C., Martin, Nicholas G., and MacGregor, Stuart (2017). "Genome-Wide Association Shows that Pigmentation Genes Play a Role in Skin Aging". In: *Journal of Investigative Dermatology* 137.9, pp. 1887–1894. ISSN: 0022-202X. DOI: https://doi.org/10.1016/j.jid.2017.04.026. URL: https://www.sciencedirect.com/science/article/pii/S0022202X17314926.

Lee, C., Collichio, F., Ollila, D., and Moschos, S. (2013). "Historical review of melanoma treatment and outcomes". In: *Clin Dermatol* 31.2, pp. 141–7. ISSN: 1879-1131 (Electronic) 0738-081X (Linking). DOI: 10.1016/j.clindermatol.2012.08.015. URL: https://www.ncbi.nlm.nih.gov/pubmed/23438377.

Lee, N., Zakka, L. R., Mihm M. C., Jr., and Schatton, T. (2016). "Tumour-infiltrating lymphocytes in melanoma prognosis and cancer immunotherapy". In: *Pathology* 48.2, pp. 177–87. ISSN: 1465-3931 (Electronic) 0031-3025 (Linking). DOI: 10.1016/j.pathol.2015.12.006. URL: https://www.ncbi.nlm.nih.gov/pubmed/27020390.

Lens, M. B. and Dawes, M. (2004). "Global perspectives of contemporary epidemiological trends of cutaneous malignant melanoma". In: *British Journal of Dermatology* 150.2, pp. 179–185. ISSN: 0007-0963. DOI: https://doi.org/10.1111/j.1365-2133.2004.05708.x. URL: https://doi.org/10.1111/j.1365-2133.2004.05708.x.

Li, C., Stoma, S., Lotta, L. A., Warner, S., Albrecht, E., Allione, A., Arp, P. P., Broer, L., Buxton, J. L., Da Silva Couto Alves, A., Deelen, J., Fedko, I. O., Gordon, S. D., Jiang, T., Karlsson, R., Kerrison, N., Loe, T. K., Mangino, M., Milaneschi, Y., Miraglio, B., Pervjakova, N., Russo, A., Surakka, I., Spek, A. van der, Verhoeven, J. E., Amin, N., Beekman, M., Blakemore, A. I., Canzian, F., Hamby, S. E., Hottenga, J. J., Jones, P. D., Jousilahti, P., Magi, R., Medland, S. E., Montgomery, G. W., Nyholt, D. R., Perola, M., Pietilainen, K. H., Salomaa, V., Sillanpaa, E., Suchiman, H. E., Heemst, D. van, Willemsen, G., Agudo, A., Boeing, H.,

Boomsma, D. I., Chirlaque, M. D., Fagherazzi, G., Ferrari, P., Franks, P., Gieger, C., Eriksson, J. G., Gunter, M., Hagg, S., Hovatta, I., Imaz, L., Kaprio, J., Kaaks, R., Key, T., Krogh, V., Martin, N. G., Melander, O., Metspalu, A., Moreno, C., Onland-Moret, N. C., Nilsson, P., Ong, K. K., Overvad, K., Palli, D., Panico, S., Pedersen, N. L., Penninx, Bwjh, Quiros, J. R., Jarvelin, M. R., Rodriguez-Barranco, M., Scott, R. A., Severi, G., Slagboom, P. E., Spector, T. D., Tjonneland, A., Trichopoulou, A., Tumino, R., Uitterlinden, A. G., Schouw, Y. T. van der, Duijn, C. M. van, Weiderpass, E., Denchi, E. L., Matullo, G., Butterworth, A. S., Danesh, J., Samani, N. J., Wareham, N. J., Nelson, C. P., Langenberg, C., and Codd, V. (2020). "Genome-wide Association Analysis in Humans Links Nucleotide Metabolism to Leukocyte Telomere Length". In: *Am J Hum Genet* 106.3, pp. 389–404. ISSN: 1537-6605 (Electronic) 0002-9297 (Linking). DOI: 10.1016/j.ajhg.2020.02.006. URL: https://www.ncbi.nlm.nih.gov/pubmed/32109421.

Li, C. Y., Gao, T. W., Wang, G., Han, Z. Y., Shen, Z., Li, T. H., and Liu, Y. F. (2004). "The effect of antisense tyrosinase-related protein 1 on melanocytes and malignant melanoma cells". In: *British Journal of Dermatology* 150.6, pp. 1081–1090. ISSN: 0007-0963. DOI: https://doi.org/10.1111/j.1365-2133.2004.05929.x. URL: https://doi.org/10.1111/j.1365-2133.2004.05929.x.

Lin, Jennifer Y. and Fisher, David E. (2007). "Melanocyte biology and skin pigmentation". In: *Nature* 445.7130, pp. 843–850. ISSN: 1476-4687. DOI: 10.1038/nature05660. URL: https://doi.org/10.1038/nature05660%20https://www.nature.com/articles/nature05660.pdf.

Liu, Fan, Visser, Mijke, Duffy, David L., Hysi, Pirro G., Jacobs, Leonie C., Lao, Oscar, Zhong, Kaiyin, Walsh, Susan, Chaitanya, Lakshmi, Wollstein, Andreas, Zhu, Gu, Montgomery, Grant W., Henders, Anjali K., Mangino, Massimo, Glass, Daniel, Bataille, Veronique, Sturm, Richard A., Rivadeneira, Fernando, Hofman, Albert, Ijcken, Wilfred F. J. van, Uitterlinden, André G., Palstra, Robert-Jan T. S., Spector, Timothy D., Martin, Nicholas G., Nijsten, Tamar E. C., and Kayser, Manfred (2015). "Genetics of skin color variation in Europeans: genome-wide association studies with functional follow-up". In: *Human Genetics* 134.8, pp. 823–835. ISSN: 1432-1203. DOI: 10.1007/s00439-015-1559-0. URL: https://doi.org/10.1007/s00439-015-1559-0.

Liu, Jimmy Z, Sommeren, Suzanne van, Huang, Hailiang, Ng, Siew C, Alberts, Rudi, Takahashi, Atsushi, Ripke, Stephan, Lee, James C, Jostins, Luke, Shah, Tejas, Abedian, Shifteh, Cheon, Jae Hee, Cho, Judy, Dayani, Naser E, Franke, Lude, Fuyuno, Yuta, Hart, Ailsa, Juyal, Ramesh C, Juyal, Garima, Kim, Won Ho, Morris, Andrew P, Poustchi, Hossein, Newman, William G, Midha, Vandana, Orchard, Timothy R, Vahedi, Homayon, Sood, Ajit, Sung, Joseph Y, Malekzadeh, Reza, Westra, Harm-Jan, Yamazaki, Keiko, Yang, Suk-Kyun, International Multiple Sclerosis Genetics Consortium, International IBD Genetics Consortium, Barrett, Jeffrey C, Alizadeh, Behrooz Z, Parkes, Miles, Bk, Thelma, Daly, Mark J, Kubo, Michiaki, Anderson, Carl A, and Weersma, Rinse K (Sept. 2015). "Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations". In: *Nature genetics* 47.9, pp. 979–986. ISSN: 1061-4036. DOI: 10.1038/ng.3359. URL: https://europepmc.org/articles/PMC4881818.

Loomis, W. Farnsworth (1967). "Skin-Pigment Regulation of Vitamin-D Biosynthesis in Man". In: *Science* 157.3788, p. 501. DOI: 10.1126/science.157.3788.501. URL: http://science.sciencemag.org/content/157/3788/501.abstract.

Loussouarn, Geneviève, Lozano, Isabelle, Panhard, Ségolène, Collaudin, Catherine, El Rawadi, Charles, and Genain, Gilles (2016). "Diversity in human hair growth, diameter, colour and shape. An in vivo study on young adults from 24 different ethnic groups observed in the five continents". In: *European Journal of Dermatology* 26.2, pp. 144–154. ISSN: 1952-4013. DOI: 10.1684/ejd.2015.2726. URL: https://doi.org/10.1684/ejd.2015.2726.

Madronich, S., McKenzie, R. L., Björn, L. O., and Caldwell, M. M. (1998). "Changes in biologically active ultraviolet radiation reaching the Earth's surface". In: *Journal of Photochemistry and Photobiology B: Biology* 46.1, pp. 5–19. ISSN: 1011-1344. DOI: https://doi.org/10.1016/S1011-1344(98)00182-1. URL: http://www.sciencedirect.com/science/article/pii/S1011134498001821.

Martin, Alicia R., Lin, Meng, Granka, Julie M., Myrick, Justin W., Liu, Xiaomin, Sockell, Alexandra, Atkinson, Elizabeth G., Werely, Cedric J., Möller, Marlo, Sandhu, Manjinder S., Kingsley, David M., Hoal, Eileen G., Liu, Xiao, Daly, Mark J., Feldman, Marcus W., Gignoux, Christopher R., Bustamante, Carlos D., and Henn, Brenna M. (2017). "An Unexpectedly Complex Architecture for Skin Pigmentation in Africans". In: *Cell* 171.6, 1340–1353.e14. ISSN:

1097-4172 0092-8674. DOI: 10.1016/j.cell.2017.11.015. URL: https://www.ncbi.nlm.nih.gov/pubmed/29195075%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5884124/.

Mavaddat, Nasim et al. (2019). "Polygenic Risk Scores for Prediction of Breast Cancer and Breast Cancer Subtypes". In: *The American Journal of Human Genetics* 104.1, pp. 21–34. ISSN: 0002-9297. DOI: https://doi.org/10.1016/j.ajhg.2018.11.002. URL: http://www.sciencedirect.com/science/article/pii/S0002929718304051.

McCarthy, Shane et al. (Oct. 2016). "A reference panel of 64,976 haplotypes for genotype imputation". In: *Nature Genetics* 48.10, pp. 1279–1283. ISSN: 1546-1718. DOI: 10.1038/ng.3643. URL: https://doi.org/10.1038/ng.3643.

McDougall, Ian, Brown, Francis H., and Fleagle, John G. (2005). "Stratigraphic placement and age of modern humans from Kibish, Ethiopia". In: *Nature* 433.7027, pp. 733–736. ISSN: 1476-4687. DOI: 10.1038/nature03258. URL: https://doi.org/10.1038/nature03258.

McKay, James D et al. (July 2017). "Large-scale association analysis identifies new lung cancer susceptibility loci and heterogeneity in genetic susceptibility across histological subtypes". In: *Nature genetics* 49.7, pp. 1126–1132. ISSN: 1061-4036. DOI: 10.1038/ng.3892. URL: https://europepmc.org/articles/PMC5510465.

Meyer, Matthias, Kircher, Martin, Gansauge, Marie-Theres, Li, Heng, Racimo, Fernando, Mallick, Swapan, Schraiber, Joshua G., Jay, Flora, Prüfer, Kay, Filippo, Cesare de, Sudmant, Peter H., Alkan, Can, Fu, Qiaomei, Do, Ron, Rohland, Nadin, Tandon, Arti, Siebauer, Michael, Green, Richard E., Bryc, Katarzyna, Briggs, Adrian W., Stenzel, Udo, Dabney, Jesse, Shendure, Jay, Kitzman, Jacob, Hammer, Michael F., Shunkov, Michael V., Derevianko, Anatoli P., Patterson, Nick, Andrés, Aida M., Eichler, Evan E., Slatkin, Montgomery, Reich, David, Kelso, Janet, and Pääbo, Svante (2012). "A high-coverage genome sequence from an archaic Denisovan individual". In: *Science (New York, N.Y.)* 338.6104, pp. 222–226. ISSN: 1095-9203 0036-8075. DOI: 10.1126/science.1224344. URL: https://pubmed.ncbi.nlm.nih.gov/22936568%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3617501/.

Meyle, Kathrine Damm and Guldberg, Per (2009). "Genetic risk factors for melanoma". In: *Human Genetics* 126.4, pp. 499–510. ISSN: 1432-1203. DOI: 10.1007/s00439-009-0715-9. URL: https://doi.org/10.1007/s00439-009-0715-9.

Minikel, Eric Vallabh, Zerr, Inga, Collins, Steven J., Ponto, Claudia, Boyd, Alison, Klug, Genevieve, Karch, André, Kenny, Joanna, Collinge, John, Takada, Leonel T., Forner, Sven, Fong, Jamie C., Mead, Simon, and Geschwind, Michael D. (2014). "Ascertainment bias causes false signal of anticipation in genetic prion disease". In: *American journal of human genetics* 95.4, pp. 371–382. ISSN: 1537-6605 0002-9297. DOI: 10.1016/j.ajhg.2014.09.003. URL: https://pubmed.ncbi.nlm.nih.gov/25279981%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4185115/.

Morgan, Michael D., Pairo-Castineira, Erola, Rawlik, Konrad, Canela-Xandri, Oriol, Rees, Jonathan, Sims, David, Tenesa, Albert, and Jackson, Ian J. (2018). "Genome-wide study of hair colour in UK Biobank explains most of the SNP heritability". In: *Nature Communications* 9.1, p. 5271. ISSN: 2041-1723. DOI: 10.1038/s41467-018-07691-z. URL: https://doi.org/10.1038/s41467-018-07691-z%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6288091/pdf/41467_2018_Article_7691.pdf.

Motsinger, Alison A. and Ritchie, Marylyn D. (Mar. 2006). "Multifactor dimensionality reduction: an analysis strategy for modelling and detecting gene-gene interactions in human genetics and pharmacogenomics studies". eng. In: *Human genomics* 2.5. PMC3500181[pmcid], pp. 318–328. ISSN: 1479-7364. DOI: 10.1186/1479-7364-2-5-318. URL: https://doi.org/10.1186/1479-7364-2-5-318.

Murray, Frederick G. (1934). "PIGMENTATION, SUNLIGHT, AND NUTRITIONAL DISEASE". In: *American Anthropologist* 36.3, pp. 438–445. ISSN: 0002-7294. DOI: 10.1525/aa.1934.36.3.02a00100. URL: https://doi.org/10.1525/aa.1934.36.3.02a00100.

Nan, Hongmei, Kraft, Peter, Qureshi, Abrar A., Guo, Qun, Chen, Constance, Hankinson, Susan E., Hu, Frank B., Thomas, Gilles, Hoover, Robert N., Chanock, Stephen, Hunter, David J., and Han, Jiali (2009). "Genome-wide association study of tanning phenotype in a population of European ancestry". In: *The Journal of investigative dermatology* 129.9, pp. 2250–2257.

ISSN: 1523-1747 0022-202X. DOI: 10.1038/jid.2009.62. URL: https://pubmed.ncbi.nlm.nih.gov/19340012%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2758927/.

Nicolae, Dan L., Gamazon, Eric, Zhang, Wei, Duan, Shiwei, Dolan, M. Eileen, and Cox, Nancy J. (2010). "Trait-Associated SNPs Are More Likely to Be eQTLs: Annotation to Enhance Discovery from GWAS". In: *PLOS Genetics* 6.4, e1000888. DOI: 10.1371/journal.pgen.1000888. URL: https://doi.org/10.1371/journal.pgen.1000888.

Njajou, Omer T., Cawthon, Richard M., Damcott, Coleen M., Wu, Shih-Hsuan, Ott, Sandy, Garant, Michael J., Blackburn, Elizabeth H., Mitchell, Braxton D., Shuldiner, Alan R., and Hsueh, Wen-Chi (2007). "Telomere length is paternally inherited and is associated with parental lifespan". In: *Proceedings of the National Academy of Sciences* 104.29, p. 12135. DOI: 10.1073/pnas.0702703104. URL: http://www.pnas.org/content/104/29/12135.abstract.

Novembre, John, Johnson, Toby, Bryc, Katarzyna, Kutalik, Zoltán, Boyko, Adam R., Auton, Adam, Indap, Amit, King, Karen S., Bergmann, Sven, Nelson, Matthew R., Stephens, Matthew, and Bustamante, Carlos D. (2008). "Genes mirror geography within Europe". In: *Nature* 456.7218, pp. 98–101. ISSN: 1476-4687 0028-0836. DOI: 10.1038/nature07331. URL: https://pubmed.ncbi.nlm.nih.gov/18758442%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2735096/.

Ollmann, M. M., Lamoreux, M. L., Wilson, B. D., and Barsh, G. S. (1998). "Interaction of Agouti protein with the melanocortin 1 receptor in vitro and in vivo". In: *Genes and development* 12.3, pp. 316–330. ISSN: 0890-9369. DOI: 10.1101/gad.12.3.316. URL: https://pubmed.ncbi.nlm.nih.gov/9450927%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC316484/.

Olsen, Catherine M., Carroll, Heidi J., and Whiteman, David C. (2010). "Familial Melanoma: A Meta-analysis and Estimates of Attributable Fraction". In: *Cancer Epidemiology Biomarkers and Prevention* 19.1, p. 65. DOI: 10.1158/1055-9965.EPI-09-0928. URL: http://cebp.aacrjournals.org/content/19/1/65.abstract.

Paik, Seung Hwan, Kim, Hyun-Jin, Son, Ho-Young, Lee, Seungbok, Im, Sun-Wha, Ju, Young Seok, Yeon, Je Ho, Jo, Seong Jin, Eun, Hee Chul, Seo, Jeong-Sun, Kwon, Oh Sang, and Kim, Jong-Il (2012). "Gene mapping study for constitutive skin color in an isolated Mongolian

population". In: *Experimental and Molecular Medicine* 44.3, pp. 241–249. ISSN: 2092-6413. DOI: 10.3858/emm.2012.44.3.020. URL: https://doi.org/10.3858/emm.2012.44.3.020.

Parr, Christine L., Hjartåker, Anette, Laake, Petter, Lund, Eiliv, and Veierød, Marit B. (Nov. 2008). "Recall Bias in Melanoma Risk Factors and Measurement Error Effects: A Nested Case-Control Study Within the Norwegian Women and Cancer Study". In: *American Journal of Epidemiology* 169.3, pp. 257–266. ISSN: 0002-9262. DOI: 10.1093/aje/kwn363. eprint: https://academic.oup.com/aje/article-pdf/169/3/257/199269/kwn363.pdf. URL: https://doi.org/10.1093/aje/kwn363.

Pickrell, Joseph K., Berisa, Tomaz, Liu, Jimmy Z., Ségurel, Laure, Tung, Joyce Y., and Hinds, David A. (2016). "Detection and interpretation of shared genetic influences on 42 human traits". In: *Nature genetics* 48.7, pp. 709–717. ISSN: 1546-1718 1061-4036. DOI: 10.1038/ng.3570. URL: https://pubmed.ncbi.nlm.nih.gov/27182965%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5207801/.

Pirastu, Nicola, Joshi, Peter K, Vries, Paul S de, Cornelis, Marilyn C, McKeigue, Paul M, Keum, NaNa, Franceschini, Nora, Colombo, Marco, Giovannucci, Edward L, Spiliopoulou, Athina, Franke, Lude, North, Kari E, Kraft, Peter, Morrison, Alanna C, Esko, Tõnu, and Wilson, James F (Nov. 2017). "GWAS for male-pattern baldness identifies 71 susceptibility loci explaining 38% of the risk". In: *Nature communications* 8.1, p. 1584. ISSN: 2041-1723. DOI: 10.1038/s41467-017-01490-8. URL: https://europepmc.org/articles/PMC5691155.

Price, Alkes L., Patterson, Nick J., Plenge, Robert M., Weinblatt, Michael E., Shadick, Nancy A., and Reich, David (2006). "Principal components analysis corrects for stratification in genome-wide association studies". In: *Nature Genetics* 38.8, pp. 904–909. ISSN: 1546-1718. DOI: 10.1038/ng1847. URL: https://doi.org/10.1038/ng1847.

Prieto, P. A., Yang, J. C., Sherry, R. M., Hughes, M. S., Kammula, U. S., White, D. E., Levy, C. L., Rosenberg, S. A., and Phan, G. Q. (2012). "CTLA-4 blockade with ipilimumab: long-term follow-up of 177 patients with metastatic melanoma". In: *Clin Cancer Res* 18.7, pp. 2039–47. ISSN: 1078-0432 (Print) 1078-0432 (Linking). DOI: 10.1158/1078-0432.CCR-11-1823. URL: https://www.ncbi.nlm.nih.gov/pubmed/22271879.

Purcell, Shaun, Neale, Benjamin, Todd-Brown, Kathe, Thomas, Lori, Ferreira, Manuel A. R., Bender, David, Maller, Julian, Sklar, Pamela, Bakker, Paul I. W. de, Daly, Mark J., and Sham, Pak C. (2007). "PLINK: a tool set for whole-genome association and population-based linkage analyses". In: *American journal of human genetics* 81.3, pp. 559–575. ISSN: 0002-9297 1537-6605. DOI: 10.1086/519795. URL: https://pubmed.ncbi.nlm.nih.gov/17701901%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1950838/.

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: https://www.R-project.org/.

Ransohoff, Katherine J., Wu, Wenting, Cho, Hyunje G., Chahal, Harvind C., Lin, Yuan, Dai, Hong-Ji, Amos, Christopher I., Lee, Jeffrey E., Tang, Jean Y., Hinds, David A., Han, Jiali, Wei, Qingyi, and Sarin, Kavita Y. (2017). "Two-stage genome-wide association study identifies a novel susceptibility locus associated with melanoma". In: *Oncotarget* 8.11, pp. 17586–17592. ISSN: 1949-2553. DOI: 10.18632/oncotarget.15230. URL: https://www.ncbi.nlm.nih.gov/pubmed/28212542%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5392271/.

Rebecca, Vito W., Sondak, Vernon K., and Smalley, Keiran S. M. (2012). "A brief history of melanoma: from mummies to mutations". In: *Melanoma research* 22.2, pp. 114–122. ISSN: 1473-5636 0960-8931. DOI: 10.1097/CMR.0b013e328351fa4d. URL: https://pubmed.ncbi.nlm.nih.gov/22395415%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3303163/.

Rees, J. L. and Healy, E. (1997). "Melanocortin receptors, red hair, and skin cancer". In: *J Investig Dermatol Symp Proc* 2.1, pp. 94–8. ISSN: 1087-0024 (Print) 1087-0024 (Linking). DOI: 10.1038/jidsymp.1997.18. URL: https://www.ncbi.nlm.nih.gov/pubmed/9487023.

Relethford, John H. (1997). "Hemispheric difference in human skin color". In: *American Journal of Physical Anthropology* 104.4, pp. 449–457. ISSN: 0002-9483. DOI: https://doi.org/10.1002/(SICI)1096-8644(199712)104:4<449::AID-AJPA2>3.0.CO;2-N. URL: https://doi.org/10.1002/(SICI)1096-8644(199712)104:4%3C449::AID-AJPA2%3E3.0.CO;2-N.

Ribero, Simone, Davies, John R., Requena, Celia, Carrera, Cristina, Glass, Daniel, Rull, Ramon, Vidal-Sicart, Sergi, Vilalta, Antonio, Alos, Lucia, Soriano, Virtudes, Quaglino, Pietro, Traves, Victor, Newton-Bishop, Julia A., Nagore, Eduardo, Malvehy, Josep, Puig, Susana, and Bataille, Veronique (2015). "High nevus counts confer a favorable prognosis in melanoma

patients". In: *International Journal of Cancer* 137.7, pp. 1691–1698. ISSN: 0020-7136. DOI: https://doi.org/10.1002/ijc.29525. URL: https://doi.org/10.1002/ijc.29525.

Robles-Espinoza, Carla Daniela, Harland, Mark, Ramsay, Andrew J., Aoude, Lauren G., Quesada, Víctor, Ding, Zhihao, Pooley, Karen A., Pritchard, Antonia L., Tiffen, Jessamy C., Petljak, Mia, Palmer, Jane M., Symmons, Judith, Johansson, Peter, Stark, Mitchell S., Gartside, Michael G., Snowden, Helen, Montgomery, Grant W., Martin, Nicholas G., Liu, Jimmy Z., Choi, Jiyeon, Makowski, Matthew, Brown, Kevin M., Dunning, Alison M., Keane, Thomas M., López-Otín, Carlos, Gruis, Nelleke A., Hayward, Nicholas K., Bishop, D. Timothy, Newton-Bishop, Julia A., and Adams, David J. (2014). "POT1 loss-of-function variants predispose to familial melanoma". In: *Nature genetics* 46.5, pp. 478–481. ISSN: 1546-1718 1061-4036. DOI: 10.1038/ng.2947. URL: https://pubmed.ncbi.nlm.nih.gov/24686849%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4266105/.

Rosso, Stefano, Sera, Francesco, Segnan, Nereo, and Zanetti, Roberto (2008). "Sun exposure prior to diagnosis is associated with improved survival in melanoma patients: Results from a long-term follow-up study of Italian patients". In: *European Journal of Cancer* 44.9, pp. 1275–1281. ISSN: 0959-8049. DOI: 10.1016/j.ejca.2008.03.009. URL: https://doi.org/10.1016/j.ejca.2008.03.009.

Sachidanandam, Ravi, Weissman, David, Schmidt, Steven C., Kakol, Jerzy M., Stein, Lincoln D., Marth, Gabor, Sherry, Steve, Mullikin, James C., Mortimore, Beverley J., Willey, David L., Hunt, Sarah E., Cole, Charlotte G., Coggill, Penny C., Rice, Catherine M., Ning, Zemin, Rogers, Jane, Bentley, David R., Kwok, Pui-Yan, Mardis, Elaine R., Yeh, Raymond T., Schultz, Brian, Cook, Lisa, Davenport, Ruth, Dante, Michael, Fulton, Lucinda, Hillier, LaDeana, Waterston, Robert H., McPherson, John D., Gilman, Brian, Schaffner, Stephen, Van Etten, William J., Reich, David, Higgins, John, Daly, Mark J., Blumenstiel, Brendan, Baldwin, Jennifer, Stange-Thomann, Nicole, Zody, Michael C., Linton, Lauren, Lander, Eric S., Altshuler, David, The International, S. N. P. Map Working Group, Cold Spring Harbor, Laboratories, National Center for Biotechnology, Information, The Sanger, Centre, Washington University in St, Louis, and Whitehead, M. I. T. Center for Genome Research (2001). "A map of human genome sequence variation containing 1.42 million single nucleotide polymor-

phisms". In: *Nature* 409.6822, pp. 928–933. ISSN: 1476-4687. DOI: 10.1038/35057149. URL: https://doi.org/10.1038/35057149.

Schadt, E. E. (2006). "Novel integrative genomics strategies to identify genes for complex traits". In: *Animal genetics* 37 Suppl 1.s1, pp. 18–23. ISSN: 0268-9146 1365-2052. DOI: 10.1111/j.1365-2052.2006.01473.x. URL: https://pubmed.ncbi.nlm.nih.gov/16886998%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2367618/.

Scolyer, Richard A., Long, Georgina V., and Thompson, John F. (2011). "Evolving concepts in melanoma classification and their relevance to multidisciplinary melanoma patient care". In: *Molecular oncology* 5.2, pp. 124–136. ISSN: 1878-0261 1574-7891. DOI: 10.1016/j.molonc.2011.03.002. URL: https://pubmed.ncbi.nlm.nih.gov/21482206%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5528281/.

Seiberg, Miri (2001). "Keratinocyte-Melanocyte Interactions During Melanosome Transfer". In: *Pigment cell research / sponsored by the European Society for Pigment Cell Research and the International Pigment Cell Society* 14, pp. 236–42. DOI: 10.1034/j.1600-0749.2001.140402.x.

Shieh, Y., Eklund, M., Madlensky, L., Sawyer, S. D., Thompson, C. K., Stover Fiscalini, A., Ziv, E., Van't Veer, L. J., Esserman, L. J., Tice, J. A., and Athena Breast Health Network, Investigators (2017). "Breast Cancer Screening in the Precision Medicine Era: Risk-Based Screening in a Population-Based Trial". In: *J Natl Cancer Inst* 109.5. ISSN: 1460-2105 (Electronic) 0027-8874 (Linking). DOI: 10.1093/jnci/djw290. URL: https://www.ncbi.nlm.nih.gov/pubmed/28130475.

Simcoe, Mark, Valdes, Ana, Liu, Fan, Furlotte, Nicholas A., Evans, David M., Hemani, Gibran, Ring, Susan M., Smith, George Davey, Duffy, David L., Zhu, Gu, Gordon, Scott D., Medland, Sarah E., Vuckovic, Dragana, Girotto, Giorgia, Sala, Cinzia, Catamo, Eulalia, Concas, Maria Pina, Brumat, Marco, Gasparini, Paolo, Toniolo, Daniela, Cocca, Massimiliano, Robino, Antonietta, Yazar, Seyhan, Hewitt, Alex, Wu, Wenting, Kraft, Peter, Hammond, Christopher J., Shi, Yuan, Chen, Yan, Zeng, Changqing, Klaver, Caroline C. W., Uitterlinden, Andre G., Ikram, M. Arfan, Hamer, Merel A., Duijn, Cornelia M. van, Nijsten, Tamar, Han, Jiali, Mackey, David A., Martin, Nicholas G., Cheng, Ching-Yu, Hinds, David A., Spector, Timothy

D., Kayser, Manfred, and Hysi, Pirro G. (2021). "Genome-wide association study in almost 195,000 individuals identifies 50 previously unidentified genetic loci for eye color". In: *Science Advances* 7.11. DOI: 10.1126/sciadv.abd1239.

Simon, John D., Peles, Dana, Wakamatsu, Kazumasa, and Ito, Shosuke (2009). "Current challenges in understanding melanogenesis: bridging chemistry, biological control, morphology, and function". In: *Pigment Cell and Melanoma Research* 22.5, pp. 563–579. ISSN: 1755-1471. DOI: 10.1111/j.1755-148X.2009.00610.x. URL: https://doi.org/10.1111/j.1755-148X.2009.00610.x.

Skol, Andrew D., Scott, Laura J., Abecasis, Gonçalo R., and Boehnke, Michael (2007). "Optimal designs for two-stage genome-wide association studies". In: *Genetic Epidemiology* 31.7, pp. 776–788. DOI: https://doi.org/10.1002/gepi.20240. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/gepi.20240. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/gepi.20240.

Slagboom, P. E., Droog, S., and Boomsma, D. I. (1994). "Genetic determination of telomere size in humans: a twin study of three age groups". In: *Am J Hum Genet* 55.5, pp. 876–82. ISSN: 0002-9297 (Print) 0002-9297 (Linking). URL: https://www.ncbi.nlm.nih.gov/pubmed/7977349.

Stokowski, R. P., Pant, P. V., Dadd, T., Fereday, A., Hinds, D. A., Jarman, C., Filsell, W., Ginger, R. S., Green, M. R., Ouderaa, F. J. van der, and Cox, D. R. (2007). "A genomewide association study of skin pigmentation in a South Asian population". In: *Am J Hum Genet* 81.6, pp. 1119–32. ISSN: 1537-6605 (Electronic) 0002-9297 (Linking). DOI: 10.1086/522235. URL: https://www.ncbi.nlm.nih.gov/pubmed/17999355.

Sudlow, Cathie, Gallacher, John, Allen, Naomi, Beral, Valerie, Burton, Paul, Danesh, John, Downey, Paul, Elliott, Paul, Green, Jane, Landray, Martin, Liu, Bette, Matthews, Paul, Ong, Giok, Pell, Jill, Silman, Alan, Young, Alan, Sprosen, Tim, Peakman, Tim, and Collins, Rory (2015). "UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age". In: *PLOS Medicine* 12.3, e1001779. DOI: 10.1371/journal.pmed.1001779. URL: https://doi.org/10.1371/journal.pmed.1001779.

The Breast Cancer Linkage, Consortium (1999). "Cancer Risks in BRCA2 Mutation Carriers". In: *JNCI: Journal of the National Cancer Institute* 91.15, pp. 1310–1316. ISSN: 0027-8874. DOI: 10.1093/jnci/91.15.1310. URL: https://doi.org/10.1093/jnci/91.15.1310.

Topalian, S. L., Drake, C. G., and Pardoll, D. M. (2015). "Immune checkpoint blockade: a common denominator approach to cancer therapy". In: *Cancer Cell* 27.4, pp. 450–61. ISSN: 1878-3686 (Electronic) 1535-6108 (Linking). DOI: 10.1016/j.ccell.2015.03.001. URL: https://www.ncbi.nlm.nih.gov/pubmed/25858804.

Tsukamoto, K., Jackson, I. J., Urabe, K., Montague, P. M., and Hearing, V. J. (1992). "A second tyrosinase-related protein, TRP-2, is a melanogenic enzyme termed DOPAchrome tautomerase". In: *The EMBO journal* 11.2, pp. 519–526. ISSN: 0261-4189 1460-2075. URL: https://pubmed.ncbi.nlm.nih.gov/1537333%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC556482/.

Tucker, Ledyard R. and Lewis, Charles (Mar. 1973). "A reliability coefficient for maximum likelihood factor analysis". In: *Psychometrika* 38.1, pp. 1–10. ISSN: 1860-0980. DOI: 10.1007/BF02291170. URL: https://doi.org/10.1007/BF02291170.

Turley, Patrick, Walters, Raymond K., Maghzian, Omeed, Okbay, Aysu, Lee, James J., Fontana, Mark Alan, Nguyen-Viet, Tuan Anh, Wedow, Robbee, Zacher, Meghan, Furlotte, Nicholas A., Agee, Michelle, Alipanahi, Babak, Auton, Adam, Bell, Robert K., Bryc, Katarzyna, Elson, Sarah L., Fontanillas, Pierre, Furlotte, Nicholas A., Hinds, David A., Hromatka, Bethann S., Huber, Karen E., Kleinman, Aaron, Litterman, Nadia K., McIntyre, Matthew H., Mountain, Joanna L., Northover, Carrie A. M., Sathirapongsasuti, J. Fah, Sazonova, Olga V., Shelton, Janie F., Shringarpure, Suyash, Tian, Chao, Tung, Joyce Y., Vacic, Vladimir, Wilson, Catherine H., Pitts, Steven J., Magnusson, Patrik, Oskarsson, Sven, Johannesson, Magnus, Visscher, Peter M., Laibson, David, Cesarini, David, Neale, Benjamin M., Benjamin, Daniel J., Research, Team andMe, and Social Science Genetic Association, Consortium (2018). "Multi-trait analysis of genome-wide association summary statistics using MTAG". In: *Nature Genetics* 50.2, pp. 229–237. ISSN: 1546-1718. DOI: 10.1038/s41588-017-0009-4. URL: https://doi.org/10.1038/s41588-017-0009-4.

Vasa-Nicotera, Mariuca, Brouilette, Scott, Mangino, Massimo, Thompson, John R., Braund, Peter, Clemitson, Jenny-Rebecca, Mason, Andrea, Bodycote, Clare L., Raleigh, Stuart M., Louis, Edward, and Samani, Nilesh J. (2005). "Mapping of a Major Locus that Determines Telomere Length in Humans". In: *The American Journal of Human Genetics* 76.1, pp. 147–151. ISSN: 0002-9297. DOI: https://doi.org/10.1086/426734. URL: http://www.sciencedirect.com/science/article/pii/S0002929707625518.

Vaughn, Michelle R., Brooks, Elizabeth, Oorschot, Roland A.H. van, and Baindur-Hudson, Swati (2009). "A Comparison of Macroscopic and Microscopic Hair Color Measurements and a Quantification of the Relationship between Hair Color and Thickness". In: *Microscopy and Microanalysis* 15.3, pp. 189–193. DOI: 10.1017/S1431927609090357.

Veierod, Marit Bragelien, Weiderpass, Elisabete, Thorn, Magnus, Hansson, Johan, Lund, Eiliv, Armstrong, Bruce, and Adami, Hans-Olov (2003). "A Prospective Study of Pigmentation, Sun Exposure, and Risk of Cutaneous Malignant Melanoma in Women". In: *JNCI: Journal of the National Cancer Institute* 95.20, pp. 1530–1538. ISSN: 0027-8874. DOI: 10.1093/jnci/djg075. URL: https://doi.org/10.1093/jnci/djg075.

Visconti, Alessia, Duffy, David L., Liu, Fan, Zhu, Gu, Wu, Wenting, Chen, Yan, Hysi, Pirro G., Zeng, Changqing, Sanna, Marianna, Iles, Mark M., Kanetsky, Peter A., Demenais, Florence, Hamer, Merel A., Uitterlinden, Andre G., Ikram, M. Arfan, Nijsten, Tamar, Martin, Nicholas G., Kayser, Manfred, Spector, Tim D., Han, Jiali, Bataille, Veronique, and Falchi, Mario (2018). "Genome-wide association study in 176,678 Europeans reveals genetic loci for tanning response to sun exposure". In: *Nature Communications* 9.1, p. 1684. ISSN: 2041-1723. DOI: 10.1038/s41467-018-04086-y. URL: https://doi.org/10.1038/s41467-018-04086-y.

Vittadini, Giorgio (1999). "Analysis of Qualitative Variables in Structural Models with Unique Solutions". In: *Classification and Data Analysis*. Ed. by Maurizio Vichi and Otto Opitz. Springer Berlin Heidelberg, pp. 203–210. ISBN: 978-3-642-60126-2.

Wachsmuth, R. C., Gaut, R. M., Barrett, J. H., Saunders, C. L., Randerson-Moor, J. A., Eldridge, A., Martin, N. G., Bishop, T. D., and Newton Bishop, J. A. (2001). "Heritability and gene-environment interactions for melanocytic nevus density examined in a U.K. adolescent twin study". In: *J Invest Dermatol* 117.2, pp. 348–52. ISSN: 0022-202X (Print) 0022-202X

(Linking). DOI: 10.1046/j.0022-202x.2001.01415.x. URL: https://www.ncbi.nlm.nih.gov/pubmed/11511314.

Wakamatsu, Kazumasa, Ohtara, Kazumi, and Ito, Shosuke (2009). "Chemical analysis of late stages of pheomelanogenesis: conversion of dihydrobenzothiazine to a benzothiazole structure". In: *Pigment Cell and Melanoma Research* 22.4, pp. 474–486. ISSN: 1755-1471. DOI: https://doi.org/10.1111/j.1755-148X.2009.00580.x. URL: https://doi.org/10.1111/j.1755-148X.2009.00580.x.

Walter, Stephen D., Marrett, Loraine D., from, Lynn, Hertzman, Clyde, Shannon, Harry S., and Roy, Peter (Feb. 1990). "THE ASSOCIATION OF CUTANEOUS MALIGNANT MELANOMA WITH THE USE OF SUNBEDS AND SUNLAMPS". In: *American Journal of Epidemiology* 131.2, pp. 232–243. ISSN: 0002-9262. DOI: 10.1093/oxfordjournals.aje.a115493. eprint: https://academic.oup.com/aje/article-pdf/131/2/232/320154/131-2-232.pdf. URL: https://doi.org/10.1093/oxfordjournals.aje.a115493.

Walter, Stephen D., Marrett, Loraine D., Shannon, Harry S., From, Lynn, and Hertzman, Clyde (Apr. 1992). "The Association of Cutaneous Malignant Melanoma and Fluorescent Light Exposure". In: *American Journal of Epidemiology* 135.7, pp. 749–762. ISSN: 0002-9262. DOI: 10.1093/oxfordjournals.aje.a116361. eprint: https://academic.oup.com/aje/article-pdf/135/7/749/156119/135-7-749.pdf. URL: https://doi.org/10.1093/oxfordjournals.aje.a116361.

Wang, Kai, Li, Mingyao, and Hakonarson, Hakon (Sept. 2010). "ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data". eng. In: *Nucleic acids research* 38.16. gkq603[PII], e164–e164. ISSN: 1362-4962. DOI: 10.1093/nar/gkq603. URL: https://doi.org/10.1093/nar/gkq603.

Ware, Erin B., Schmitz, Lauren L., Faul, Jessica, Gard, Arianna, Mitchell, Colter, Smith, Jennifer A., Zhao, Wei, Weir, David, and Kardia, Sharon L. R. (2017). "Heterogeneity in polygenic scores for common human traits". In: *bioRxiv*, p. 106062. DOI: 10.1101/106062. URL: http://biorxiv.org/content/early/2017/02/05/106062.abstract.

Watanabe, Kyoko, Taskesen, Erdogan, Bochoven, Arjen van, and Posthuma, Danielle (2017). "Functional mapping and annotation of genetic associations with FUMA". In: *Nature com-*

munications 8.1, pp. 1826–1826. ISSN: 2041-1723. DOI: 10.1038/s41467-017-01261-5. URL: https://pubmed.ncbi.nlm.nih.gov/29184056%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5705698/.

Weinstein, S. J., Virtamo, J., and Albanes, D. (Aug. 2013). "Pigmentation-related phenotypes and risk of prostate cancer". In: *British Journal of Cancer* 109.3, pp. 747–750. ISSN: 1532-1827. DOI: 10.1038/bjc.2013.385. URL: https://doi.org/10.1038/bjc.2013.385.

Weinstock, Martin A., Colditz, Graham A., Willet, Walter C., Stampfer, Meir J., Rosner, Bernard, and Speizer, Frank E. (Feb. 1991). "Recall (Report) Bias and Reliability in the Retrospective Assessment of Melanoma Risk". In: *American Journal of Epidemiology* 133.3, pp. 240–245. ISSN: 0002-9262. DOI: 10.1093/oxfordjournals.aje.a115868. eprint: https://academic.oup.com/aje/article-pdf/133/3/240/236502/133-3-240.pdf. URL: https://doi.org/10.1093/oxfordjournals.aje.a115868.

White, Tim D., Asfaw, Berhane, DeGusta, David, Gilbert, Henry, Richards, Gary D., Suwa, Gen, and Clark Howell, F. (2003). "Pleistocene Homo sapiens from Middle Awash, Ethiopia". In: *Nature* 423.6941, pp. 742–747. ISSN: 1476-4687. DOI: 10.1038/nature01669. URL: https://doi.org/10.1038/nature01669.

Willer, Cristen J., Li, Yun, and Abecasis, Gonçalo R. (2010). "METAL: fast and efficient meta-analysis of genomewide association scans". In: *Bioinformatics (Oxford, England)* 26.17, pp. 2190–2191. ISSN: 1367-4811 1367-4803. DOI: 10.1093/bioinformatics/btq340. URL: https://pubmed.ncbi.nlm.nih.gov/20616382%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2922887/.

Williams, Patricia F., Olsen, Catherine M., Hayward, Nicholas K., and Whiteman, David C. (2011). "Melanocortin 1 receptor and risk of cutaneous melanoma: A meta-analysis and estimates of population burden". In: *International Journal of Cancer* 129.7, pp. 1730–1740. ISSN: 0020-7136. DOI: https://doi.org/10.1002/ijc.25804. URL: https://doi.org/10.1002/ijc.25804.

Wolchok, J. D., Neyns, B., Linette, G., Negrier, S., Lutzky, J., Thomas, L., Waterfield, W., Schadendorf, D., Smylie, M., Guthrie T., Jr., Grob, J. J., Chesney, J., Chin, K., Chen, K., Hoos, A., O'Day, S. J., and Lebbe, C. (2010). "Ipilimumab monotherapy in patients with pre-

treated advanced melanoma: a randomised, double-blind, multicentre, phase 2, dose-ranging study". In: *Lancet Oncol* 11.2, pp. 155–64. ISSN: 1474-5488 (Electronic) 1470-2045 (Linking). DOI: 10.1016/S1470-2045(09)70334-1. URL: https://www.ncbi.nlm.nih.gov/pubmed/20004617.

Wood, Andrew R., Tuke, Marcus A., Nalls, Mike A., Hernandez, Dena G., Bandinelli, Stefania, Singleton, Andrew B., Melzer, David, Ferrucci, Luigi, Frayling, Timothy M., and Weedon, Michael N. (2014). "Another explanation for apparent epistasis". In: *Nature* 514.7520, E3–E5. ISSN: 1476-4687 0028-0836. DOI: 10.1038/nature13691. URL: https://pubmed.ncbi.nlm.nih.gov/25279928%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6478385/.

Xu, ChangJiang, Tachmazidou, Ioanna, Walter, Klaudia, Ciampi, Antonio, Zeggini, Eleftheria, Greenwood, Celia M. T., and Consortium, Uk K. (2014). "Estimating genome-wide significance for whole-genome sequencing studies". In: *Genetic epidemiology* 38.4, pp. 281–290. ISSN: 1098-2272 0741-0395. DOI: 10.1002/gepi.21797. URL: https://pubmed.ncbi.nlm.nih.gov/24676807%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4489336/.

Yang, J., Lee, S. H., Goddard, M. E., and Visscher, P. M. (2011). "GCTA: a tool for genome-wide complex trait analysis". In: *Am J Hum Genet* 88.1, pp. 76–82. ISSN: 1537-6605 (Electronic) 0002-9297 (Linking). DOI: 10.1016/j.ajhg.2010.11.011. URL: https://www.ncbi.nlm.nih.gov/pubmed/21167468.

Yang, Jian, Weedon, Michael N., Purcell, Shaun, Lettre, Guillaume, Estrada, Karol, Willer, Cristen J., Smith, Albert V., Ingelsson, Erik, O'Connell, Jeffrey R., Mangino, Massimo, Mägi, Reedik, Madden, Pamela A., Heath, Andrew C., Nyholt, Dale R., Martin, Nicholas G., Montgomery, Grant W., Frayling, Timothy M., Hirschhorn, Joel N., McCarthy, Mark I., Goddard, Michael E., Visscher, Peter M., and Consortium, Giant (2011). "Genomic inflation factors under polygenic inheritance". In: *European journal of human genetics : EJHG* 19.7, pp. 807–812. ISSN: 1476-5438 1018-4813. DOI: 10.1038/ejhg.2011.39. URL: https://pubmed.ncbi.nlm.nih.gov/21407268%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3137506/.

Yokoyama, Satoru, Woods, Susan L., Boyle, Glen M., Aoude, Lauren G., MacGregor, Stuart, Zismann, Victoria, Gartside, Michael, Cust, Anne E., Haq, Rizwan, Harland, Mark, Taylor, John C., Duffy, David L., Holohan, Kelly, Dutton-Regester, Ken, Palmer, Jane M., Bonazzi,

Vanessa, Stark, Mitchell S., Symmons, Judith, Law, Matthew H., Schmidt, Christopher, Lanagan, Cathy, O'Connor, Linda, Holland, Elizabeth A., Schmid, Helen, Maskiell, Judith A., Jetann, Jodie, Ferguson, Megan, Jenkins, Mark A., Kefford, Richard F., Giles, Graham G., Armstrong, Bruce K., Aitken, Joanne F., Hopper, John L., Whiteman, David C., Pharoah, Paul D., Easton, Douglas F., Dunning, Alison M., Newton-Bishop, Julia A., Montgomery, Grant W., Martin, Nicholas G., Mann, Graham J., Bishop, D. Timothy, Tsao, Hensin, Trent, Jeffrey M., Fisher, David E., Hayward, Nicholas K., and Brown, Kevin M. (2011). "A novel recurrent mutation in MITF predisposes to familial and sporadic melanoma". In: *Nature* 480.7375, pp. 99–103. ISSN: 1476-4687 0028-0836. DOI: 10.1038/nature10630. URL: https://pubmed.ncbi.nlm.nih.gov/22080950%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3266855/.

Zhang, Tongwu, Choi, Jiyeon, Kovacs, Michael A., Shi, Jianxin, Xu, Mai, Program, Nisc Comparative Sequencing, Melanoma Meta-Analysis, Consortium, Goldstein, Alisa M., Trower, Adam J., Bishop, D. Timothy, Iles, Mark M., Duffy, David L., MacGregor, Stuart, Amundadottir, Laufey T., Law, Matthew H., Loftus, Stacie K., Pavan, William J., and Brown, Kevin M. (2018). "Cell-type-specific eQTL of primary melanocytes facilitates identification of melanoma susceptibility genes". In: *Genome research* 28.11, pp. 1621–1635. ISSN: 1549-5469 1088-9051. DOI: 10.1101/gr.233304.117. URL: https://pubmed.ncbi.nlm.nih.gov/30333196%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6211648/.

Zhu, G., Duffy, D. L., Eldridge, A., Grace, M., Mayne, C., O'Gorman, L., Aitken, J. F., Neale, M. C., Hayward, N. K., Green, A. C., and Martin, N. G. (1999). "A major quantitative-trait locus for mole density is linked to the familial melanoma gene CDKN2A: a maximum-likelihood combined linkage and association analysis in twins and their sibs". In: *American journal of human genetics* 65.2, pp. 483–492. ISSN: 0002-9297 1537-6605. DOI: 10.1086/302494. URL: https://pubmed.ncbi.nlm.nih.gov/10417291%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1377947/.

Zhu, Zhihong, Zhang, Futao, Hu, Han, Bakshi, Andrew, Robinson, Matthew R., Powell, Joseph E., Montgomery, Grant W., Goddard, Michael E., Wray, Naomi R., Visscher, Peter M., and Yang, Jian (2016). "Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets". In: *Nature Genetics* 48.5, pp. 481–487. ISSN: 1546-1718. DOI: 10.1038/ng.3538. URL: https://doi.org/10.1038/ng.3538.

Zorina-Lichtenwalter, Katerina, Lichtenwalter, Ryan N, Zaykin, Dima V, Parisien, Marc, Gravel, Simon, Bortsov, Andrey, and Diatchenko, Luda (Jan. 2019). "A study in scarlet: MC1R as the main predictor of red hair and exemplar of the flip-flop effect". In: *Human Molecular Genetics* 28.12, pp. 2093–2106. ISSN: 0964-6906. DOI: 10.1093/hmg/ddz018. eprint: https://academic.oup.com/hmg/article-pdf/28/12/2093/28765644/ddz018.pdf. URL: https://doi.org/10.1093/hmg/ddz018.

Zou, Fei, Lee, Seunggeun, Knowles, Michael R., and Wright, Fred A. (2010). "Quantification of population structure using correlated SNPs by shrinkage principal components". In: *Human heredity* 70.1, pp. 9–22. ISSN: 1423-0062 0001-5652. DOI: 10.1159/000288706. URL: https://pubmed.ncbi.nlm.nih.gov/20413978%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2912642/.

Appendix A

Codes used throughout Thesis

A.1 Peak sorting algorithm R script

```
setwd("~/Documents/PhD/Projects/MTAG")
MAF<-fread("HRC.r1-1.GRCh37.wgs.mac5.sites.txt", select=c(3,8))
setwd("~/Documents/PhD/Projects/GN Compare/2nd GWAS/linfull2")

data2<-subset(datablonde, P<=(0.5*10^(-7)))
data3 <- data2[order(data2$P),]
data3a<-data3[which(data3$P>0),]
data3b<-merge(data3a, MAF, by="ID")
data3c<-data3b[which(data3b$AF>0.01 & data3b$AF<0.99)]
data3d <- data3c[order(data3c$P),]
data4<-data.frame(data3d,NA)

while(sum(is.na(data4[,8]))!=0){

  n1<-which(is.na(data4[,8]))[1]

  data4[n1,8]<-1


  pos1<-data4[n1,4]
  chr1<-data4[n1,3]
  close<-which(is.na(data4[,8]) & data4[,3]==chr1 & abs(data4[,4]-pos1) <
      5000000)

```

```
  data4[close,8]<-0

  print(mean(is.na(data4[,8])))
}

sum(data4$NA.)

Final<-subset(data4, data4$NA.==1)
write.table(Final, file="Peaks SNPs blonde2.txt", row.names = F, quote=F)
rm(data2,data3,data4, Final, chr1, close, n1, pos1, data3a, data3b, data3c,
    data3d)

data2<-subset(datared, P<=(0.5*10^(-7)))
data3 <- data2[order(data2$P),]
data3a<-data3[which(data3$P>0),]
data3b<-merge(data3a, MAF, by="ID")
data3c<-data3b[which(data3b$AF>0.01 & data3b$AF<0.99)]
data3d <- data3c[order(data3c$P),]
data4<-data.frame(data3d,NA)


while(sum(is.na(data4[,8]))!=0){

  n1<-which(is.na(data4[,8]))[1]

  data4[n1,8]<-1


  pos1<-data4[n1,4]
  chr1<-data4[n1,3]
  close<-which(is.na(data4[,8]) & data4[,3]==chr1 & abs(data4[,4]-pos1) <
      5000000)

  data4[close,8]<-0

  print(mean(is.na(data4[,8])))
}

sum(data4$NA.)
```

```
61

Final<-subset(data4, data4$NA.==1)
63 write.table(Final, file="Peaks SNPs red2.txt", row.names = F, quote=F)
```

A.2 Collinearity sorting algorithm

```
#### putting all hits from the traits into one file and then getting it ready for
     the sorting code
full2<-full[which(full$p<=5*10^(-8)),]
full2.5<-full2[,c(1,3,2)]
full3<-full2.5[!duplicated(full2.5), ]
full4<-full3[order(full3$Chr, full3$bp),]
rmatrix2<-fread("R2for_paired_SNPS2.txt")
starter<-cbind(full4, "a")

rm(ease, hair, skin, red, psnocs1, psnocs2, psnocs3, full2, full2.5, full3, full4
    , full, cs)


g<-1
rmatrix2$test<-0
while(g<=nrow(starter)){
  listo<-as.character(starter[g,3])
  rmatrix2$test<-ifelse(rmatrix2$SNP2==listo, 1, rmatrix2$test)
  g<-g+1
}
rmatrix<-rmatrix2[which(rmatrix2$test!=0),c(1,2,3)]
rm(rmatrix2)



#### This is the sorting code that groups hits together based of an r^2 value
    which has to be specified in the code.
n<-1
c<-data.frame()
while(n<=22){
  fullchr<-starter[which(starter$Chr==n)]
  v<-1
  while(v<=nrow(fullchr)){
    if (fullchr[v,4]=="a"){
      rs<-as.character(fullchr[v,3])
      fullchr[v,4]<-rs
      rmat<-rmatrix[which(rmatrix$SNP1==rs),]
      greater_than_0.5_snps<-rmat[which(rmat$r>0.05), c(3,2)]
      colnames(greater_than_0.5_snps)<-c("r2", "locus")
      if(is.na(greater_than_0.5_snps[1,1])){
```

```
37        } else {
            x<-1
39          while(x<=nrow(greater_than_0.5_snps)){
              rschecker<-as.list(greater_than_0.5_snps[x,2])
41            if(fullchr[which(fullchr$SNP==rschecker),4]=="a"){
                fullchr[which(fullchr$SNP==rschecker),4]<-rs
43            } else if (fullchr[which(fullchr$SNP==rschecker),4]!="a"){
                rsreplace<-fullchr[which(fullchr$SNP==rschecker),4]
45              fullchr[which(fullchr$SNP==rsreplace),4]<-rs
              }
47            x<-x+1
            }
49        }


51


53      }
        v<-v+1
55

    }
57  c<-rbind(c,fullchr)
    n<-n+1
59  }
#### Once the code is run, it uses rs SNP names as group names - each hit group
    needs a number instead of one of their rsIDs
61  c2<-c[order(c$Chr, c$bp),]
    c2$locus<-0
63  c4<-c[which(!duplicated(c$V2)), 4]
    j<-1
65  while(j<=nrow(c4)){
      rs<-as.character(c4[j])
67    c2[which(c2$V2==rs),5]<-j
      j<-j+1
69  }
    colnames(c2)<-c("CHR", "BP", "RS", "RSlead", "locus")
71  c3<-c2[,c(5,1,2,3)]
    write.table(c3, "Combined_pig_locus2.txt", quote=F, row.names = F)
73  write.csv(c3, "Combined_pig_locus2.csv")
```

A.3 R code for beta coefficient plotting

```
rm(list = ls(all.names = TRUE))
setwd("~/OneDrive - University of Leeds/Projects/MPGWAS/Lead SNPS from GCTA p08
    r0.05")
overlap<-fread("overlap_with_locus2.csv")
t<-0
c<-data.frame()
while(t<=max(overlap$signal)){
  sub<-overlap[which(overlap$signal==t),]
  sub$checker2<-0
  sub$checker3<-0
  sub[,15]<-ifelse(!is.na(sub[,6]), 1, 0)
  sub[,16]<-ifelse(!is.na(sub[,7]), 1, 0)
  sub$fullcheck<-0
  if(nrow(sub)>1){

    if(colSums(sub[,15])>=1 & colSums(sub[,16])>=1){sub$fullcheck<-3}
    if(colSums(sub[,15])>=1 & colSums(sub[,16])==0){sub$fullcheck<-1}
    if(colSums(sub[,15])==0 & colSums(sub[,16])>=1){sub$fullcheck<-2}



  } else if(nrow(sub)==1) {

    if(sub[,15]>=1 & sub[,16]>=1){ sub$fullcheck<-3}
    if(sub[,15]>=1 & sub[,16]==0){sub$fullcheck<-1}
    if(sub[,15]==0 & sub[,16]>=1){sub$fullcheck<-2}
    else {}

  } else {}

    c<-rbind(c,sub)
  t<-t+1
}

c2<-c[which(c$fullcheck!=0),]
final<-c2[which(!duplicated(c2$signal)),]
setwd("~/OneDrive - University of Leeds/Projects/MPGWAS/Other pheno/skincolour/
    Results")
```

```
skin<-fread("pdtotal.txt")
setwd("~/OneDrive - University of Leeds/Projects/MPGWAS/R")
ease<-fread("pdtotal.txt")

Finalease<-merge(final, ease[,c(3,8,9,11)], by.x="RS", by.y="ID")
finalfinal<-merge(Finalease, skin[,c(3,8,9,11)], by.x="RS", by.y="SNP")


names(finalfinal)[names(finalfinal) == "BETA.x"] <- "BETAease"
names(finalfinal)[names(finalfinal) == "BETA.y"] <- "BETAskin"
names(finalfinal)[names(finalfinal) == "SE.x"] <- "SEease"
names(finalfinal)[names(finalfinal) == "SE.y"] <- "SEskin"
names(finalfinal)[names(finalfinal) == "P.x"] <- "Pease"
names(finalfinal)[names(finalfinal) == "P.y"] <- "Pskin"

plot(finalfinal$BETAease, finalfinal$BETAskin)

dev.new(width=30,hight=10, units="in")
plot(finalfinal$BETAease, finalfinal$BETAskin, pch = 16, xlim=c(-0.6,0.6),ylim=c
    (-0.2,0.5), ylab = "Beta skin colour", xlab = "Beta ease of tanning", cex.lab
    =1.2,col="red")
points(finalfinal$BETAease[which(finalfinal$Pease<=(5*10^(-8)))],finalfinal$
    BETAskin[which(finalfinal$Pease<=(5*10^(-8)))], col = "black", pch = 16);
    abline(h=0); abline(v=0)
points(finalfinal$BETAease[which(finalfinal$Pease<=(5*10^(-8)) & finalfinal$Pskin
    <=(5*10^(-8)))],finalfinal$BETAskin[which(finalfinal$Pease<=(5*10^(-8)) &
    finalfinal$Pskin<=(5*10^(-8)))], col = "green", pch = 16)
mtext("Plotted independent signals lead SNP beta coefficients for:", cex=1.2,
    side=3, outer=TRUE, line=-2) ##Change the name to the phenotype
mtext("ease of tanning GWAS and skin colour GWAS", cex=1.2, side=3, outer=TRUE,
    line=-3) ##Change the name to the phenotype

#legend('bottomleft', legend=c("Genome-wide significant for both GWAS", "Genome-
    wide significant for red vs other GWAS", "Genome-wide significant for ordinal
    scale hair colour GWAS"), pch=c(16,16,16), col=c("green", "black", "red"), cex
    =0.755,bg="white")
legend('bottomleft', legend=c("Sig for both", "Sig for skin colour", "Sig for
    ease of tanning"), pch=c(16,16,16), col=c("green", "red", "black"), cex=0.755,
    bg="white")
```

Appendix B

Individual full cancer self-report and confirmation matching tables

UK Biobank melanoma self-reports						
Cancer registry confirmation	Self-report 1	Self-report 2	Self-report 3	Self-report 4	Self-report 5	Total
No Match	1167	102	16	1		1286
Malignant melanoma	1870	168	19	4	1	2062
Non-melanoma skin cancer	279	37	1		2	319
Salivary gland cancer/parotid gland cancer	1					1
Large bowl cancer/colorectal cancer/anal cancer/colon cancer/appendix cancer	1					1
Eye cancer/adnexal cancer/reinoblastoma	35	2	1			38
Female genital tract cancer/vaginal cancer/vulval cancer	1	1				2
Male genital tract cancer/Penis cancer	2					2
Lymphoma/Leukaemia	1	2				3
Primary bone cancer/Sarcoma/fibrosarcoma	4	1				5
Metastatic unknown primary/unclassifiable	5		1			6
Breast cancer	1			1		2
Nasal cavity cancer	1					1
Lip cancer	2					2
Brain cancer/primary malignant brain tumour	1					1
Prostate cancer	1					1
Multiple myeloma	2					2
Malignant lymph node (unspecified)	1					1
Cin/precancer cells cervix	1					1
Bone metastases/bony secondaries	2					2
Other	42	6				48
Melanoma in situ	156	14	1			171
Brain cancer/primary malignant brain tumour		1				1
Kidney/renal cell cancer/other cancer of urinary tract		1				1
Thyroid cancer		1				1
Total	3576	336	39	6	3	3960

UK Biobank melanoma self-reports (sifint)						
Selfint	Self-report 1	Self-report 2	Self-report 3	Self-report 4	Self-report 5	Total
a	1398	109	11	4	0	1522
b	472	59	8	0	1	540
c	539	66	4	1	2	612
d	1167	102	16	1	0	1286
Total	3576	336	39	6	3	3960

*a cancer match and dates are within 1 year
*b cancer match dates are over 1 year
*c no cancer match but another cancer within 1 year of reported date
*d no cancer match and no other cancer within 1 year

Table B.1: Melanoma self-report and closest pathological confirmation to the self-report. Closeness determined by differences in data between report and confirmation, and cancer type.

UK Biobank melanoma Confirmations															
Cancer self-report	Confirmation 1	Confirmation 2	Confirmation 3	Confirmation 4	Confirmation 5	Confirmation 6	Confirmation 7	Confirmation 8	Confirmation 9	Confirmation 10	Confirmation 11	Confirmation 12	Confirmation 13	Confirmation 14	Total
No Match	194	41	4	1											240
Skin Cancer	58	14													72
Malignant melanoma	1745	323	52	7	2					1	1				2131
Non-melanoma skin cancer	28	5	1												34
Pancreatic cancer/malignant insulinoma	2														2
Lung/trachea cancer		1													1
Female genital tract cancer/vaginal cancer/vulval cancer	1														1
Lymphoma/Leukaemia	1	1	1												3
Metastatic unknown primary/unclassifiable	1														1
Breast cancer	3														3
Nasal cavity cancer	1														1
Kidney/renal cell cancer/other cancer of urinary tract	1														1
Total	2035	385	58	8	2	0	0	0	0	1	1	0	0	0	2490

UK Biobank melanoma confirmations								
	Confirmation 1	Confirmation 2	Confirmation 3	Confirmation 4	Confirmation 5	Confirmation 10	Confirmation 11	Total
a	1322	212	30	2				1566
b	423	111	22	5	2	1	1	565
c	96	21	2					119
d	194	41	4	1				240
Total	2035	385	58	8	2	1	1	2490

Table B.2: Melanoma UK Cancer Registry confirmation and closest cancer self-report to the confirmation. Closeness determined by differences in data between report and confirmation, and cancer type.

UK Biobank non-melanoma skin cancer self-reports							
Cancer registry confirmation	Self-report 1	Self-report 2	Self-report 3	Self-report 4	Self-report 5	Self-report 6	Total
No Match	1822	157	23	5	2	1	2010
Non-melanoma skin cancer	4347	579	87	17	2	1	5033
Malignant melanoma	35	3					38
Large bowl cancer/colorectal cancer/anal cancer/colon cancer/appendix cancer	11	1					12
Female genital tract cancer/vaginal cancer/vulval cancer	6						6
Male genital tract cancer/Penis cancer	4						4
Lymphoma/Leukaemia	3						3
Primary bone cancer/Sarcoma/fibrosarcoma	4						4
Tonsil cancer/oropharynx/oropharyngeal cancer	8						8
Larynx/throat cancer	1						1
Lip cancer	2						2
Tongue cancer	6						6
Gum cancer	2						2
Kidney/renal cell cancer/other cancer of urinary tract	1						1
Bladder cancer	1						1
Uterine/endometrial cancer	1						1
Cervical cancer	1						1
Thyroid cancer	1						1
Cin/precancer cells cervix	1						1
Mouth cancer	1	1					2
Other	46	3					49
Melanoma in situ	25	1					26
Breast cancer		2	1				3
Prostate cancer		3					3
Total	6329	750	111	22	4	2	7218

UK Biobank non-melanoma skin cancer self-reports (slfint)							
Selfint	Self-report 1	Self-report 2	Self-report 3	Self-report 4	Self-report 5	Self-report 6	Total
a	2458	301	34	7			2800
b	1889	278	53	10	2	1	2233
c	160	14	1				175
d	1822	157	23	5	2	1	2010
Total	6329	750	111	22	4	2	7218

*a cancer match and dates are within 1 year
*b cancer match dates are over 1 year
*c no cancer match but another cancer within 1 year of reported date
*d no cancer match and no other cancer within 1 year

Table B.3: Non-melanoma skin cancer self-report and closest pathological confirmation to the self-report. Closeness determined by differences in data between report and confirmation, and cancer type.

UK Biobank "skin cancer" self-reports					
Cancer registry confirmation	Self-report 1	Self-report 2	Self-report 3	Self-report 4	Total
No Match	872	63	9	1	945
Malignant melanoma	59	15	2		76
Non-melanoma skin cancer	398	40	4		442
Salivary gland cancer/parotid gland cancer	1				1
Large bowl cancer/colorectal cancer/anal cancer/colon cancer/appendix cancer	1				1
Female genital tract cancer/vaginal cancer/vulval cancer	1				1
Male genital tract cancer/Penis cancer	2				2
Lymphoma/Leukaemia	5				5
Primary bone cancer/Sarcoma/fibrosarcoma	2				2
Nasal cavity cancer			1		1
Lip cancer	1				1
Prostate cancer	1		1		2
Cin/precancer cells cervix	1				1
Bone metastases/bony secondaries	2	2			4
Other	42	2			44
Melanoma in situ	15	4			19
Kidney/renal cell cancer/other cancer of urinary tract	1				1
Total	1404	126	17	1	1548

Table B.4: Skin cancer self-report and closest pathological confirmation to the self-report. Closeness determined by differences in data between report and confirmation, and cancer type.

UK Biobank non-melanoma skin cancer confirmations															
Cancer self-report	Confirmation 1	Confirmation 2	Confirmation 3	Confirmation 4	Confirmation 5	Confirmation 6	Confirmation 7	Confirmation 8	Confirmation 9	Confirmation 10	Confirmation 11	Confirmation 12	Confirmation 13	Confirmation 14	Total
No Match	5030	768	122	22	8	2	1		1	1					5955
Skin Cancer	353	42	8	2											405
Non-melanoma skin cancer	3913	491	93	16	4	2	1	1							4521
Malignant melanoma	279	51	8	3											341
Large bowl cancer/colorectal cancer/anal cancer/colon cancer/appendix cancer	15	2			1										18
Pancreatic cancer/malignant insulinoma	2														2
Eye/adnexal cancer/reinoblastoma	12	2													14
Ovarian/fallopian tube cancer	2														2
Female genital tract cancer/vaginal cancer/vulval cancer		1	1												2
Lymphoma/Leukaemia	6	1	2												9
Primary bone cancer/Sarcoma/fibrosarcoma	11	3													14
Metastatic unknown primary/unclassifiable	12	1													13
Cancer of lip/mouth/pharynx/oral/cavity	2	1													3
Larynx/throat cancer	2														2
Lip cancer	4														4
Nasal cavity cancer	5	1	1												7
Primary ear cancer	4		1												5
Oesophageal cancer	1	1													2
Small intestine/small bowl cancer			1												1
Stomach cancer		1													1
Liver/heptatocellular cancer	1														1
Kidney/renal cell cancer/other cancer of urinary tract	2														2
Bladder cancer	3	3	1												7
Uterine/endometrial cancer	1														1
Cervical cancer		1													1
Thyroid cancer		1													1
Cin/precancer cells cervix			1	1											2
Testicular cancer	2														2
Mouth cancer	1	1													2
Malignant lymph node (unspecified)	1	1													2
Breast cancer	17	15	3												35
Multipl myeloma		1													1
Thyroid cancer				1											1
Prostate cancer	9	16	1												26
Total	9690	1405	243	45	13	4	2	1	1	1	0	0	0	0	11405

Table B.5: Skin cancer UK Cancer Registry confirmation and closest cancer self-report to the confirmation. Closeness determined by differences in data between report and confirmation, and cancer type.

UK Biobank Breast cancer self-reports								
Cancer registry confirmation	Self-report 1	Self-report 2	Self-report 3	Self-report 4	Self-report 5	Self-report 6	Self-report 7	Total
No Match	762	39	1					802
Breast cancer	9452	492	36	7	5	1	2	9995
Malignant melanoma	1							1
Non-melanoma skin cancer	6							6
Large bowl cancer/colorectal cancer/anal cancer/colon cancer/appendix cancer	1							1
Ovarian cancer/fallopian tube cancer	1							1
Metastatic unknown primary/unclassifiable	3	1						4
Malignant lymph node (unspecified)	3							3
Cin/precancer cells cervix	79	4						83
Bone metastases/bony secondaries	3							3
Other	844	59	6	2				911
Melanoma in situ	1	1						2
Total	11156	596	43	9	5	1	2	11812

UK Biobank Breast cancer self-reports (slfint)								
Selfint	Self-report 1	Self-report 2	Self-report 3	Self-report 4	Self-report 5	Self-report 6	Self-report 7	Total
a	8638	405	27	5	2	1	2	9080
b	814	87	9	2	3	0	0	915
c	942	65	6	2	0	0	0	1015
d	762	39	1		0		0	802
Total	11156	596	43	9	5	1	2	11812

*a cancer match and dates are within 1 year
*b cancer match dates are over 1 year
*c no cancer match but another cancer within 1 year of reported date
*d no cancer match and no other cancer within 1 year

Table B.6: Breast cancer self-report and closest pathological confirmation to the self-report. Closeness determined by differences in data between report and confirmation, and cancer type.

UK Biobank Breast Cancer Confirmations															
Cancer self-report	Confirmation 1	Confirmation 2	Confirmation 3	Confirmation 4	Confirmation 5	Confirmation 6	Confirmation 7	Confirmation 8	Confirmation 9	Confirmation 10	Confirmation 11	Confirmation 12	Confirmation 13	Confirmation 14	Total
No Match	98	17	4	2											121
Primary bone cancer/Sarcoma/fibrosarcoma	3														3
Metastatic unknown primary/unclassifiable	3	1													4
Breast cancer	8909	1564	212	41	2	1			1						10730
Malignant lymph node (unspecified)	2														2
Total	9015	1582	216	43	2	1	0	0	1	0	0	0	0	0	10860

UK Biobank Breast cancer confirmations								
	Confirmation 1	Confirmation 2	Confirmation 3	Confirmation 4	Confirmation 5	Confirmation 6	Confirmation 9	Total
a	8248	1271	137	25			1	9682
b	661	293	75	16	2	1		1048
c	8	1						9
d	98	17	4	2				121
Total	9015	1582	216	43	2	1	1	10860

Table B.7: Breast cancer UK Cancer Registry confirmation and closest cancer self-report to the confirmation. Closeness determined by differences in data between report and confirmation, and cancer type.

Appendix C

Pigmentation GWAS summary statistics

C.1 Non-red hair

Locus	Signal	CHR	BP	SNP	EA	NEA	EAF	Nearest Gene	BETA	SE	P
103	1	1	1178482	rs78555129	A	G	0.89	FAM132A	0.02	0.00	1.81E-15
11	2	1	8207579	rs80293268	G	C	0.95	RP11-431K24.1	0.14	0.00	2.08E-217
11	3	1	8208573	rs74865018	G	A	0.92	RP11-431K24.1	0.05	0.00	3.46E-42
35	4	1	10596341	rs12375	C	T	0.68	PEX14	0.01	0.00	1.16E-09
35	5	1	11037447	rs56168930	A	G	0.77	C1orf127	0.03	0.00	7.02E-53
159	6	1	16293592	rs848200	C	T	0.90	ZBTB17	0.02	0.00	5.55E-11
95	7	1	17601165	rs11585357	C	T	0.84	PADI3	0.02	0.00	1.23E-16
218	8	1	24810069	rs195700	A	G	0.51	NIPAL3	-0.01	0.00	4.28E-09
232	9	1	25463142	rs12077712	T	A	0.66	RP4-781L3.1	0.01	0.00	1.58E-08
164	10	1	27284913	rs79598313	C	T	0.98	C1orf172	0.04	0.01	8.62E-11
198	11	1	41898581	rs1937999	G	T	0.63	RNA5SP45	0.01	0.00	1.73E-09
231	12	1	59053745	rs12568356	C	T	0.80	TACSTD2	0.01	0.00	1.27E-08
141	13	1	61710179	rs17377295	G	A	0.94	NFIA	-0.02	0.00	5.18E-12
201	14	1	78450517	rs34517439	C	A	0.88	DNAJB4:GIPC2	0.02	0.00	2.06E-09
177	15	1	85528006	rs12034421	C	A	0.62	WDR63	0.01	0.00	2.63E-10
166	16	1	94132788	rs236285	G	A	0.80	BCAR3	-0.01	0.00	1.16E-10
171	17	1	103345744	rs12722976	C	G	0.51	COL11A1	0.01	0.00	1.56E-10
144	18	1	153189978	rs1410860	A	G	0.50	PRR9	0.01	0.00	6.85E-12

106	19	1	155106227	rs4745	A	T	0.49	EFNA1	-0.01	0.00	2.84E-15
97	20	1	192530548	rs2760524	A	G	0.17	RP5-1011O1.2	-0.02	0.00	2.24E-16
26	21	1	205130413	rs3851294	A	G	0.09	DSTYK	-0.06	0.00	1.85E-78
102	22	1	211344388	rs28826269	G	A	0.84	RP11-543B16.1	-0.02	0.00	1.18E-15
183	23	1	214597586	rs113614987	C	T	0.52	PTPN14	0.01	0.00	4.43E-10
183	24	1	214673271	rs7533482	T	C	0.81	PTPN14	-0.01	0.00	1.31E-09
186	25	1	227366626	rs59116340	A	T	0.86	CDC42BPA:RP11-1B20.1	-0.02	0.00	4.97E-10
207	26	1	232618974	rs16857370	T	C	0.97	SIPA1L2	-0.03	0.01	2.51E-09
233	27	1	236075564	rs12129097	C	T	0.97	RNU6-968P	-0.03	0.01	1.88E-08
235	28	1	240395615	rs1947536	C	T	0.35	FMN2	0.01	0.00	2.58E-08
223	29	2	1642070	rs12714332	C	T	0.18	PXDN	0.01	0.00	5.70E-09
173	30	2	11309644	rs72785456	A	G	0.95	PQLC3	0.03	0.00	1.75E-10
146	31	2	27730940	rs1260326	T	C	0.39	GCKR	-0.01	0.00	8.69E-12
12	32	2	28613302	rs71443018	G	C	0.94	AC104695.3	0.11	0.00	4.22E-190
12	33	2	28647650	rs55870117	G	A	0.70	RP11-373D23.3	-0.02	0.00	1.02E-25
39	34	2	42148519	rs4952542	T	C	0.39	AC104654.2	0.03	0.00	2.01E-46
143	35	2	43626521	rs57696714	T	C	0.90	THADA	0.02	0.00	6.35E-12
165	36	2	70545164	rs2706764	C	T	0.70	FAM136A	0.01	0.00	9.38E-11
61	37	2	88581126	rs7608166	A	G	0.88	AC012671.3	0.03	0.00	7.35E-25
208	38	2	109605767	rs3749110	G	A	0.85	EDAR	-0.02	0.00	2.52E-09

67	39	2	119538982	rs6716872	G	A	0.75	RP11-19E11.1	-0.02	0.00	2.50E-23
152	40	2	121081603	rs114153232	G	T	0.93	AC012363.13	0.02	0.00	2.38E-11
70	41	2	135407285	rs6724774	C	T	0.42	TMEM163	-0.02	0.00	1.21E-22
118	42	2	176991857	rs71421546	C	A	0.96	HOXD-AS2	0.03	0.00	9.53E-14
86	43	2	177582940	rs13431878	T	C	0.58	AC092162.1	0.02	0.00	2.68E-18
161	44	2	192117238	rs13030978	C	T	0.72	MYO1B	-0.01	0.00	6.71E-11
74	45	2	202839971	rs12693954	T	C	0.43	AC007358.1	-0.02	0.00	1.58E-21
228	46	2	207999713	rs7583880	G	A	0.31	KLF7	-0.01	0.00	1.06E-08
239	47	2	214082380	rs56020963	T	G	0.89	RP11-105N14.2	-0.02	0.00	3.90E-08
14	48	2	222051419	rs10169459	T	C	0.42	AC011233.2	-0.05	0.00	7.35E-165
14	49	2	222089797	rs17349283	A	G	0.56	AC011233.2	0.04	0.00	1.60E-119
14	50	2	222220804	rs7570475	G	C	0.67	EPHA4	0.02	0.00	1.63E-19
14	51	2	222348434	rs13029372	A	G	0.58	EPHA4	-0.02	0.00	2.56E-29
48	52	2	223024442	rs744174	A	G	0.35	PAX3	-0.02	0.00	1.98E-15
48	53	2	223025055	rs116254882	G	T	0.96	PAX3	-0.05	0.00	7.68E-33
48	54	2	223110512	rs12618431	A	G	0.87	PAX3	-0.02	0.00	2.44E-13
48	55	2	223159119	rs2303948	C	A	0.78	PAX3	-0.02	0.00	4.10E-16
189	56	2	233737132	rs812383	C	A	0.44	C2orf82	0.01	0.00	6.80E-10
84	57	2	239680992	rs9287636	G	A	0.31	AC144525.1	-0.02	0.00	9.28E-19
84	58	2	239949681	rs4075018	G	A	0.79	RP11-648F7.1	0.02	0.00	7.86E-14

176	59	3	250758	rs9809528	A	G	0.57	CHL1	0.01	0.00	2.58E-10
104	60	3	11660766	rs2437689	C	T	0.30	VGLL4	0.02	0.00	2.24E-15
153	61	3	66789001	rs754821	A	G	0.41	RPL21P41	-0.01	0.00	2.70E-11
77	62	3	69797413	rs2014520	G	A	0.59	MITF	0.02	0.00	6.05E-21
109	63	3	72397279	rs9809116	A	G	0.58	RYBP	-0.01	0.00	9.97E-15
96	64	3	73862616	rs586936	G	A	0.60	RP11-20B7.1	-0.02	0.00	2.19E-16
38	65	3	122526816	rs9847240	G	A	0.34	DIRC2	0.03	0.00	6.47E-47
135	66	3	125993274	rs115182912	G	A	0.98	RP11-71E19.1	0.04	0.01	2.38E-12
210	67	3	133607321	rs2715610	A	T	0.47	RAB6B	-0.01	0.00	2.81E-09
160	68	3	138123854	rs2293252	C	T	0.35	MRAS	-0.01	0.00	6.03E-11
20	69	3	141094209	rs6440003	G	A	0.55	ZBTB38	0.04	0.00	8.98E-94
20	70	3	141179873	rs73869619	T	C	0.98	KRT18P35	-0.05	0.01	3.78E-10
89	71	3	141634056	rs3804772	G	A	0.88	ATP1B3	0.02	0.00	8.35E-18
187	72	3	151904190	rs323610	T	G	0.36	RP11-246A10.1	-0.01	0.00	5.05E-10
202	73	3	153786394	rs1713843	C	G	0.85	ARHGEF26-AS1	0.01	0.00	2.17E-09
185	74	3	171037521	rs34128525	G	A	0.72	TNIK	-0.01	0.00	4.54E-10
157	75	3	181517982	rs833268	G	A	0.38	SOX2-OT	-0.01	0.00	5.31E-11
236	76	3	189214006	rs62289589	C	T	0.58	TP63	-0.01	0.00	2.71E-08
120	77	4	4388874	rs10015223	C	G	0.05	NSG1	-0.03	0.00	1.29E-13
149	78	4	54509016	rs6554121	C	A	0.70	FIP1L1:LNX1	-0.01	0.00	1.74E-11

127	79	4	57838690	rs1107674	T	G	0.64	NOA1	-0.01	0.00	8.17E-13
204	80	4	74442349	rs17804499	G	C	0.94	RASSF6	0.02	0.00	2.47E-09
42	81	4	75328479	rs1874202	C	G	0.59	AREG	-0.02	0.00	3.48E-39
32	82	4	79281682	rs371273	A	G	0.34	FRAS1	0.03	0.00	5.85E-61
29	83	4	81199966	rs1458046	G	A	0.58	FGF5	0.02	0.00	1.01E-39
29	84	4	81661426	rs72661730	G	A	0.78	C4orf22	0.04	0.00	1.36E-72
73	85	4	86602918	rs17392334	C	T	0.52	ARHGAP24	-0.02	0.00	1.27E-21
73	86	4	86929983	rs28483422	T	C	0.85	RP13-514E23.1	0.02	0.00	2.72E-11
56	87	4	105816898	rs9998015	T	C	0.63	RP11-556I14.2	0.02	0.00	6.38E-27
56	88	4	106028823	rs17429682	G	A	0.68	RP11-556I14.1	-0.01	0.00	1.74E-10
52	89	4	109012183	rs7673917	T	C	0.99	LEF1	-0.07	0.01	5.93E-14
52	90	4	109076822	rs17038688	C	A	0.73	LEF1	-0.01	0.00	3.58E-10
52	91	4	109252929	rs220625	C	T	0.09	LEF1-AS1	0.03	0.00	5.19E-16
52	92	4	109350880	rs219493	T	C	0.17	EXOC7P1	0.02	0.00	9.04E-24
52	93	4	109478108	rs11731416	C	G	0.49	RPL34-AS1	-0.02	0.00	1.08E-30
151	94	4	139446206	rs1584590	T	A	0.19	RP11-173E2.1	0.02	0.00	1.85E-11
200	95	4	145511194	rs1996020	A	C	0.84	KRT18P51	0.01	0.00	1.85E-09
75	96	4	149805677	rs72719803	C	T	0.87	CTB-191D16.1	0.03	0.00	3.36E-21
237	97	4	154000968	rs4696396	C	T	0.70	RP11-285C1.2	-0.01	0.00	3.63E-08
230	98	5	9547021	rs1651282	C	T	0.40	SNHG18	0.01	0.00	1.10E-08

4	99	5	33951693	rs16891982	C	G	0.07	SLC45A2	0.39	0.01	0
4	100	5	33986409	rs13289	C	G	0.39	AMACR	-0.01	0.00	9.88E-10
209	101	5	34700000	rs6881866	A	T	0.78	RAI14	-0.01	0.00	2.57E-09
27	102	5	53018404	rs77325285	G	A	0.90	NDUFS4	-0.04	0.00	5.85E-40
27	103	5	53112624	rs6875907	C	T	0.48	CTD-2081C10.1	0.03	0.00	1.35E-74
27	104	5	53484470	rs3846492	G	A	0.80	ARL15	-0.01	0.00	9.31E-09
80	105	5	56011357	rs7714232	A	T	0.83	AC008940.1	-0.02	0.00	4.76E-20
162	106	5	57135050	rs1835873	C	A	0.43	AC116606.1	0.01	0.00	6.83E-11
215	107	5	66789202	rs4282273	C	T	0.71	RP11-434D9.1	-0.01	0.00	3.48E-09
129	108	5	79695370	rs259035	T	G	0.90	ZFYVE16	-0.02	0.00	1.53E-12
28	109	5	90263347	rs11957689	G	T	0.62	GPR98	-0.03	0.00	1.73E-74
241	110	5	106889936	rs163888	C	T	0.67	EFNA5	-0.01	0.00	4.05E-08
195	111	5	112541296	rs9326880	G	C	0.20	MCC	-0.01	0.00	1.23E-09
136	112	5	116231256	rs10519488	A	G	0.95	CTC-472C24.1	-0.03	0.00	2.56E-12
133	113	5	133852468	rs10479082	G	A	0.76	RN7SL541P	0.01	0.00	2.10E-12
66	114	5	173831110	rs2936938	C	T	0.62	RP11-267A15.1	-0.02	0.00	2.50E-17
66	115	5	174156168	rs4242182	T	C	0.11	MSX2	-0.03	0.00	4.03E-24
158	116	5	180661980	rs17714046	T	C	0.95	TRIM41	-0.03	0.00	5.42E-11
2	117	6	192288	rs61376093	T	C	0.95	NA	0.06	0.00	9.40E-31
2	118	6	376962	rs71550004	C	T	0.97	IRF4	0.08	0.01	2.21E-33

2	119	6	396321	rs12203592	C	T	0.79	IRF4	-0.32	0.00	0
2	120	6	466518	rs60729976	G	A	0.67	EXOC2	0.04	0.00	1.68E-61
2	121	6	521147	rs9378896	T	C	0.36	EXOC2	-0.09	0.00	0
110	122	6	10580294	rs594329	C	T	0.83	GCNT2	0.02	0.00	1.51E-14
59	123	6	20627777	rs2223620	C	T	0.67	CDKAL1	-0.02	0.00	2.58E-25
184	124	6	21920773	rs9368374	A	G	0.70	CASC15	-0.01	0.00	4.47E-10
128	125	6	31351541	rs9266756	A	G	0.93	NA	-0.03	0.00	6.45E-12
128	126	6	31583841	rs34451818	C	G	0.94	NA	-0.03	0.00	1.20E-12
128	127	6	31726850	rs17207524	G	T	0.94	NA	0.02	0.00	1.01E-09
85	128	6	45739785	rs62400428	C	T	0.74	RUNX2	-0.01	0.00	4.68E-12
85	129	6	45901916	rs9349337	G	A	0.70	CLIC5	-0.02	0.00	2.19E-18
191	130	6	107426530	rs10457158	G	C	0.76	BEND3	0.01	0.00	7.08E-10
99	131	6	113304903	rs2127965	A	G	0.40	RNU6-1163P	0.02	0.00	2.88E-16
212	132	6	130374461	rs7740107	T	A	0.26	L3MBTL3	0.01	0.00	3.14E-09
148	133	6	131388483	rs6901548	T	C	0.77	EPB41L2	-0.01	0.00	1.36E-11
137	134	6	134609291	rs4896038	A	C	0.78	SGK1	0.01	0.00	3.49E-12
138	135	6	148570402	rs34286635	A	G	0.77	RP11-631F7.1	0.01	0.00	4.00E-12
138	136	6	148737041	rs4897010	C	A	0.82	SASH1	0.01	0.00	1.37E-09
44	137	6	151577739	rs10434895	A	T	0.55	AKAP12	0.02	0.00	2.01E-36
19	138	6	159248431	rs9457478	C	G	0.50	EZR-AS1	0.04	0.00	3.06E-120

225	139	6	163391444	rs9365518	C	T	0.43	PACRG	0.01	0.00	6.74E-09
182	140	7	2577781	rs1043291	T	C	0.66	BRAT1	-0.01	0.00	1.70E-08
182	141	7	2822933	rs2266921	C	T	0.81	GNA12	0.01	0.00	4.17E-10
92	142	7	14015318	rs2282867	T	A	0.28	ETV1	-0.02	0.00	4.13E-17
46	143	7	28180556	rs864745	T	C	0.51	JAZF1	-0.02	0.00	2.50E-34
76	144	7	28818245	rs10224680	T	C	0.63	CREB5	-0.02	0.00	4.70E-21
125	145	7	41327697	rs117441897	A	G	0.95	AC005022.1	-0.03	0.00	2.94E-09
125	146	7	41408001	rs77969347	A	G	0.97	AC005022.1	0.04	0.01	6.14E-13
91	147	7	46801428	rs1464841	T	C	0.75	AC011294.3	-0.02	0.00	2.86E-17
91	148	7	46905525	rs11764140	C	G	0.61	AC004901.1	0.01	0.00	7.07E-14
41	149	7	90867230	rs80030895	C	T	0.90	FZD1	-0.04	0.00	1.15E-39
41	150	7	90913120	rs73217177	T	C	0.94	FZD1	0.03	0.00	1.60E-13
41	151	7	91209001	rs2157725	C	T	0.63	RP11-142A5.1	0.01	0.00	9.49E-10
58	152	7	100401825	rs314349	T	G	0.63	EPHB4	-0.02	0.00	1.45E-18
58	153	7	100451732	rs12535629	C	T	0.72	SLC12A9	0.02	0.00	6.74E-26
83	154	7	105416560	rs2529369	C	A	0.71	ATXN7L1	-0.02	0.00	2.08E-19
30	155	7	130742066	rs7803075	A	G	0.28	LINC-PINT	0.04	0.00	9.31E-72
238	156	7	144076764	rs1612590	A	G	0.39	ARHGEF5	-0.01	0.00	3.88E-08
179	157	7	155091406	rs9767875	G	T	0.86	INSIG1	-0.02	0.00	3.40E-10
226	158	7	156151175	rs56326046	A	T	0.55	AC073133.2	0.01	0.00	6.87E-09

213	159	8	9000965	rs330926	C	A	0.69	PPP1R3B:RP11-10A14.3	-0.01	0.00	3.36E-09
139	160	8	22600953	rs2048651	T	C	0.36	RP11-459E5.1:PEBP4	0.01	0.00	4.05E-12
178	161	8	29852712	rs2595041	T	C	0.65	MAP2K1P1	0.01	0.00	2.74E-10
82	162	8	38568215	rs57128498	C	A	0.84	RP11-495O10.1	0.02	0.00	1.89E-19
36	163	8	82704598	rs6473306	T	C	0.40	RP13-923O23.6	0.02	0.00	4.90E-17
36	164	8	82720760	rs2600605	G	A	0.74	RP13-923O23.6:SNX16:HNRNPA1P36	0.03	0.00	1.41E-50
216	165	8	101076864	rs2446928	A	T	0.48	RGS22	0.01	0.00	3.65E-09
81	166	8	108994382	rs1389985	G	A	0.96	RSPO2	0.02	0.00	3.21E-08
81	167	8	109092699	rs446454	C	A	0.75	RSPO2	-0.02	0.00	1.17E-19
98	168	8	116823453	rs1526460	G	T	0.27	TRPS1	-0.02	0.00	2.45E-16
134	169	8	119133623	rs2445922	C	T	0.85	EXT1	-0.02	0.00	2.12E-12
43	170	9	211762	rs520015	C	G	0.51	C9orf66	-0.02	0.00	4.14E-39
122	171	9	5896161	rs7037587	T	C	0.72	KIAA2026:MLANA	0.01	0.00	3.06E-13
18	172	9	12716762	rs1326797	T	G	0.39	RP11-3L8.3	0.04	0.00	6.17E-123
196	173	9	15866367	rs770197	C	G	0.13	CCDC171	0.02	0.00	1.32E-09
23	174	9	16560854	rs7858025	A	T	0.91	BNC2	0.02	0.00	3.12E-10
23	175	9	16802973	rs2254330	A	G	0.75	BNC2	-0.03	0.00	1.88E-35
23	176	9	16885017	rs12350739	G	A	0.42	BNC2	0.03	0.00	4.77E-80
117	177	9	22065002	rs10811647	C	G	0.58	CDKN2B-AS1	-0.01	0.00	8.60E-14
219	178	9	27510360	rs10812605	C	T	0.36	MOB3B	0.01	0.00	4.44E-09

155	179	9	34107505	rs11557154	C	T	0.87	DCAF12	0.02	0.00	3.18E-11
180	180	9	82289001	rs1475675	T	C	0.09	TLE4	-0.02	0.00	3.78E-10
63	181	9	100616583	rs3021523	T	C	0.26	FOXE1	0.02	0.00	1.13E-24
114	182	9	109054417	rs10739220	C	T	0.23	RP11-308N19.1	-0.02	0.00	6.79E-14
94	183	9	116387254	rs540599	C	T	0.41	RP11-168K11.3	-0.02	0.00	8.19E-17
17	184	9	126705122	rs57425397	G	A	0.95	DENND1A	-0.03	0.00	1.26E-10
17	185	9	126768582	rs969585	T	C	0.42	LHX2:RP11-85O21.2	0.03	0.00	7.33E-52
17	186	9	126808006	rs58979150	C	T	0.89	LHX2	-0.07	0.00	2.26E-125
17	187	9	126808021	rs16936765	C	G	0.82	LHX2	0.05	0.00	3.80E-101
17	188	9	126984382	rs953470	A	G	0.76	NEK6	0.02	0.00	7.11E-13
17	189	9	126991185	rs10818930	T	G	0.20	NEK6	0.04	0.00	7.30E-55
172	190	9	129822536	rs7032500	G	A	0.51	RALGPS1	0.01	0.00	1.64E-10
174	191	10	32679078	rs17230340	C	G	0.83	RNU6-1244P	-0.02	0.00	1.89E-10
93	192	10	35464172	rs12264698	G	A	0.69	CREM	0.02	0.00	8.17E-17
147	193	10	74068998	rs4433500	G	A	0.64	DNAJB12	0.01	0.00	8.85E-12
31	194	10	80944147	rs703978	C	G	0.41	ZMIZ1	0.03	0.00	4.95E-65
224	195	10	112704740	rs4918614	G	A	0.36	SHOC2	-0.01	0.00	6.22E-09
34	196	11	7543519	rs11041426	G	A	0.38	PPFIBP2	0.03	0.00	1.15E-54
62	197	11	15677913	rs75319234	A	G	0.99	RP11-396O20.1	0.05	0.01	5.86E-09
62	198	11	15710084	rs7108738	T	G	0.82	RP11-396O20.1	-0.02	0.00	8.48E-25

62	199	11	15858293	rs1920672	T	G	0.34	RP11-222N13.1	-0.01	0.00	2.22E-09
62	200	11	16189834	rs77535014	G	A	0.98	SOX6	-0.04	0.01	5.23E-10
65	201	11	16354653	rs297343	T	G	0.37	SOX6	-0.02	0.00	2.52E-18
65	202	11	16500554	rs151060912	C	T	0.98	SOX6	0.04	0.01	3.73E-11
65	203	11	16586767	rs11023988	G	T	0.79	SOX6	-0.02	0.00	3.79E-24
206	204	11	18305333	rs2049129	T	C	0.82	HPS5	-0.01	0.00	2.49E-09
170	205	11	26769140	rs11029693	G	A	0.53	SLC5A12	0.01	0.00	1.27E-10
217	206	11	27495259	rs4923447	C	T	0.06	RP11-159H22.2	-0.02	0.00	3.73E-09
131	207	11	32927778	rs145678014	G	T	0.96	QSER1	-0.03	0.00	1.77E-12
90	208	11	43922776	rs11037723	G	A	0.77	ALKBH3	-0.01	0.00	3.26E-10
90	209	11	44330610	rs66716358	C	T	0.50	ALX4	0.02	0.00	8.69E-18
90	210	11	44451792	rs1902961	G	A	0.65	RP11-58K22.1	-0.01	0.00	1.22E-12
113	211	11	47245803	rs7395496	A	G	0.17	DDB2	-0.02	0.00	6.13E-14
72	212	11	61618169	rs61897795	A	G	0.84	FADS2	0.02	0.00	4.97E-22
53	213	11	62206288	rs9645690	C	T	0.68	AHNAK	0.02	0.00	2.47E-30
64	214	11	65422853	rs11227247	A	C	0.86	RELA	0.03	0.00	2.49E-24
7	215	11	68831364	rs72928978	G	A	0.90	TPCN2	0.15	0.00	0
7	216	11	68855363	rs3829241	G	A	0.61	TPCN2	0.05	0.00	5.63E-143
24	217	11	78130325	rs11237489	G	A	0.83	RP11-452H21.1	-0.05	0.00	1.81E-79
13	218	11	88879915	rs2187128	T	C	0.53	TYR	0.05	0.00	6.96E-77

13	219	11	88911696	rs1042602	C	A	0.65	TYR	0.05	0.00	1.33E-174
13	220	11	89204911	rs77368047	T	C	0.84	NOX4	0.04	0.00	3.19E-62
175	221	11	129601467	rs492335	G	A	0.61	RP11-507F16.1	-0.01	0.00	2.55E-10
45	222	12	4317563	rs3764032	T	C	0.95	CCND2-AS1	-0.05	0.00	4.53E-35
79	223	12	23979791	rs9971729	A	C	0.43	SOX5	-0.02	0.00	3.15E-20
220	224	12	24846522	rs6487406	T	C	0.05	RP11-625L16.1	0.03	0.00	5.15E-09
199	225	12	28221404	rs11049319	T	C	0.79	CCDC91	-0.01	0.00	1.78E-09
142	226	12	46774771	rs7979418	A	G	0.38	RP11-96H19.1	-0.01	0.00	5.94E-12
112	227	12	49402393	rs10875910	G	C	0.65	RP11-386G11.5:PRKAG1	-0.01	0.00	2.86E-14
234	228	12	54301044	rs1470321	G	A	0.84	HOXC13-AS	0.01	0.00	2.37E-08
119	229	12	57594955	rs11172124	G	A	0.75	LRP1	0.02	0.00	1.04E-13
123	230	12	65147371	rs1147094	T	C	0.41	GNS	0.01	0.00	3.28E-13
33	231	12	85391419	rs1818930	G	T	0.76	TSPAN19	0.02	0.00	2.26E-12
33	232	12	85693252	rs12425342	G	A	0.95	ALX1	0.07	0.00	2.38E-58
25	233	12	88143958	rs73200863	G	C	0.91	CYCSP30	0.07	0.00	4.86E-79
10	234	12	88707968	rs10858711	C	T	0.85	TMTC3	-0.09	0.00	1.67E-285
6	235	12	88956625	rs1798011	T	C	0.08	KITLG	0.04	0.00	1.07E-35
6	236	12	89328335	rs12821256	T	C	0.89	RP11-13A1.1	0.17	0.00	0
6	237	12	89355709	rs9651934	C	A	0.93	RP11-13A1.1	-0.04	0.00	1.83E-29
6	238	12	89675376	rs2638470	G	A	0.72	RP11-13A1.3	0.03	0.00	9.34E-38

6	239	12	89684892	rs77495873	C	T	0.94	RP11-13A1.3	-0.04	0.00	6.49E-15
88	240	12	89921277	rs10777186	G	A	0.77	RP11-734K2.4	-0.02	0.00	5.26E-18
88	241	12	90078824	rs17782847	T	C	0.96	ATP2B1	-0.03	0.00	3.26E-11
71	242	12	91116069	rs12368980	C	T	0.73	RP11-20L19.1	-0.02	0.00	2.08E-22
227	243	12	111905245	rs73201772	C	T	0.89	ATXN2	0.02	0.00	8.81E-09
154	244	12	116683588	rs61937364	C	T	0.90	MED13L	0.02	0.00	3.12E-11
154	245	12	116948892	rs6490074	C	T	0.63	RP11-148B3.1	0.01	0.00	8.76E-11
229	246	12	123716930	rs1716169	A	T	0.20	MPHOSPH9	0.01	0.00	1.07E-08
205	247	12	125368342	rs77717551	G	A	0.93	SCARB1	0.02	0.00	2.48E-09
68	248	13	39343822	rs9603422	C	T	0.88	FREM2	-0.03	0.00	4.44E-23
9	249	13	78391757	rs1279403	C	T	0.40	EDNRB-AS1	0.07	0.00	0.00E+00
21	250	13	95196559	rs6492711	C	T	0.64	TGDS	0.04	0.00	2.61E-93
203	251	13	111187839	rs9515244	T	A	0.67	RAB20	0.01	0.00	2.21E-09
150	252	13	114825471	rs7999377	G	A	0.58	RASA3	-0.01	0.00	1.83E-11
130	253	14	50655357	rs72681869	G	C	0.99	SOS2	-0.06	0.01	1.53E-12
111	254	14	54107791	rs210381	G	A	0.43	AL163953.3	-0.01	0.00	2.12E-14
111	255	14	54336842	rs56412931	C	T	0.83	AL162759.1	0.01	0.00	2.51E-09
101	256	14	60782189	rs1951116	G	A	0.66	CTD-2568P8.1	0.02	0.00	1.13E-15
78	257	14	64390030	rs10873172	G	C	0.28	SYNE2	-0.02	0.00	1.53E-20
115	258	14	68423215	rs28649231	G	A	0.82	RAD51B	-0.02	0.00	7.09E-14

87	259	14	69214219	rs12148044	G	A	0.83	RNU6-921P	-0.02	0.00	3.12E-18
3	260	14	92652216	rs113005382	C	T	0.94	CPSF2	-0.02	0.00	1.11E-08
3	261	14	92777462	rs1885194	T	C	0.56	SLC24A4	0.15	0.00	0
3	262	14	92815008	rs76519749	T	C	0.89	SLC24A4	0.11	0.00	4.17E-306
121	263	14	104010198	rs879552	G	A	0.63	RNU7-160P	-0.01	0.00	2.57E-13
1	264	15	28130503	rs35591793	C	T	0.97	OCA2	-0.06	0.01	8.19E-17
1	265	15	28150846	rs12909167	G	A	0.61	OCA2	-0.03	0.00	1.43E-61
1	266	15	28365618	rs12913832	A	G	0.26	HERC2	0.42	0.00	0
140	267	15	49189487	rs143692029	G	C	0.98	SHC4	0.04	0.01	4.12E-12
156	268	15	57542893	rs28653088	C	G	0.83	TCF12	-0.02	0.00	3.96E-11
22	269	15	81526302	rs12324786	C	T	0.79	IL16	0.04	0.00	3.23E-85
60	270	15	83964925	rs1877024	G	A	0.73	RP11-382A20.4	0.02	0.00	5.29E-25
221	271	15	86050204	rs17554929	T	C	0.67	AKAP13	0.01	0.00	5.41E-09
197	272	15	90734426	rs34560261	C	T	0.84	SEMA4B	-0.01	0.00	1.53E-09
240	273	16	54019653	rs76536391	C	T	0.96	FTO	-0.02	0.00	3.91E-08
132	274	16	57689385	rs1801257	C	G	0.51	GPR56	0.01	0.00	2.03E-12
190	275	16	72122706	rs12708925	A	T	0.59	TXNL4B	0.01	0.00	7.01E-10
5	276	16	89587080	rs116918202	C	T	0.94	SPG7	0.14	0.00	1.80E-286
5	277	16	89655877	rs144988021	C	T	0.98	CPNE7	-0.06	0.01	6.36E-21
5	278	16	89799124	rs371629999	G	A	0.90	NA	0.05	0.00	1.13E-46

5	279	16	89959387	rs73276549	A	G	0.98	TCF25	0.10	0.01	9.87E-37
5	280	16	89985844	rs1805005	G	T	0.88	MC1R:AC092143.1:TUBB3	0.06	0.00	3.78E-103
5	281	16	89985918	rs1805006	C	A	0.99	MC1R:AC092143.1:TUBB3	0.15	0.01	6.06E-80
5	282	16	89986117	rs1805007	C	T	0.90	MC1R:AC092143.1:TUBB3	0.21	0.00	0
5	283	16	89986546	rs1805009	G	C	0.98	MC1R:AC092143.1:TUBB3:RP11-566K11.4	0.14	0.01	3.28E-105
116	284	17	18134354	rs2746026	C	T	0.60	LLGL1	-0.01	0.00	7.14E-14
169	285	17	38279036	rs117827273	C	G	0.85	MSL1	0.02	0.00	1.25E-10
69	286	17	39551099	rs117612447	C	T	0.97	KRT31	-0.05	0.01	5.67E-23
69	287	17	39641932	rs897418	G	C	0.08	KRT36	-0.02	0.00	8.32E-12
107	288	17	45309693	rs11079764	A	G	0.44	MYL4	-0.02	0.00	5.36E-15
15	289	17	45950721	rs72833470	A	G	0.73	RP11-6N17.2	0.05	0.00	1.22E-146
15	290	17	45968591	rs7219107	G	A	0.87	RP11-6N17.4	0.02	0.00	2.47E-14
15	291	17	46027257	rs16949720	G	A	0.65	RP11-6N17.9	-0.03	0.00	5.41E-42
47	292	17	48008683	rs9303554	C	T	0.55	RP11-304F15.6	-0.02	0.00	1.57E-19
47	293	17	48022018	rs55788912	G	A	0.73	RP11-304F15.5	-0.02	0.00	4.31E-34
49	294	17	55230628	rs17833789	C	A	0.56	AKAP1	0.02	0.00	2.74E-32
163	295	17	63533768	rs1133683	G	A	0.34	AXIN2	0.01	0.00	8.04E-11
16	296	17	79385044	rs56931525	C	T	0.74	RP11-1055B8.7	-0.02	0.00	3.21E-23
16	297	17	79564930	rs71373084	C	G	0.65	NPLOC4	0.05	0.00	8.72E-141
16	298	17	79909537	rs34683731	G	A	0.41	NOTUM	0.02	0.00	4.72E-27

55	299	17	80752132	rs75596709	G	T	0.90	TBCD	0.03	0.00	7.85E-29
188	300	18	42601019	rs7236870	A	G	0.65	SETBP1	0.01	0.00	5.89E-10
108	301	18	43536260	rs9964220	C	T	0.84	EPG5	-0.02	0.00	6.21E-15
167	302	18	77305852	rs12954491	A	G	0.34	NFATC1	-0.01	0.00	1.19E-10
194	303	19	1252827	rs12609746	T	C	0.76	MIDN	0.01	0.00	1.02E-09
192	304	19	4086807	rs350818	A	G	0.23	MAP2K2	0.01	0.00	9.08E-10
211	305	19	10808770	rs4804519	C	T	0.33	QTRT1	0.01	0.00	3.02E-09
222	306	19	31200501	rs16964556	A	C	0.59	ZNF536	0.01	0.00	5.61E-09
168	307	20	22075642	rs6113507	G	C	0.59	RP11-125P18.1	-0.01	0.00	1.23E-10
8	308	20	32665748	rs6059655	A	G	0.10	RALY	-0.14	0.00	0
51	309	20	34960201	rs4519594	T	C	0.24	DLGAP4	-0.03	0.00	3.40E-31
193	310	20	38101545	rs6028446	T	A	0.47	RN7SL194P	0.01	0.00	9.56E-10
37	311	20	52642793	rs55901013	C	T	0.95	BCAS1	0.06	0.00	1.68E-47
54	312	20	55409093	rs6127868	G	A	0.85	RNU6-929P	0.03	0.00	6.86E-30
40	313	20	57845936	rs12481673	T	C	0.70	ZNF831	-0.03	0.00	5.18E-46
57	314	21	26868270	rs2829780	C	T	0.89	AP000221.1	-0.03	0.00	3.51E-26
181	315	21	28227149	rs229103	G	A	0.65	ADAMTS1	-0.01	0.00	3.86E-10
214	316	21	36208167	rs68088846	G	A	0.79	RUNX1	-0.01	0.00	3.46E-09
50	317	21	44752768	rs73220980	G	A	0.97	LINC00322	0.05	0.01	8.99E-19
50	318	21	44793448	rs672948	A	T	0.40	AP001046.6	-0.02	0.00	3.26E-32

100	319	22	19953059	rs165895	T	C	0.66	COMT	-0.02	0.00	4.05E-16
105	320	22	29525694	rs13056416	C	T	0.83	KREMEN1	0.02	0.00	2.27E-15
124	321	22	33259104	rs137487	A	G	0.48	SYN3	0.01	0.00	4.22E-13
145	322	22	39739187	rs9611155	C	T	0.69	SYNGR1	-0.01	0.00	8.61E-12
126	323	22	50722408	rs79966207	T	C	0.83	PLXNB2	0.02	0.00	6.74E-13

Table C.1: Independent signals for non-red hair colour GWAS

C.2 Red hair

Locus	Signal	CHR	BP	SNP	EA	NEA	EAF	Nearest Gene	BETA	SE	P
11	1	2	25393388	rs76645364	A	G	0.82	POMC	0.11	0.01	4.84E-14
9	2	5	173976464	rs60780889	C	T	0.73	SUMO2P6	-0.12	0.01	1.46E-22
6	3	6	396321	rs12203592	C	T	0.79	IRF4	-0.14	0.01	7.94E-28
5	4	6	51697567	rs2025753	T	C	0.53	PKHD1	0.13	0.01	7.60E-33
5	5	6	51713402	rs12529074	G	A	0.88	PKHD1	0.14	0.02	6.82E-14
10	6	11	69388143	rs652190	G	A	0.34	CCND1	0.09	0.01	2.93E-15
13	7	11	88966584	rs4512823	A	G	0.66	TYR	0.08	0.01	4.26E-10
12	8	13	78392656	rs9544609	C	A	0.41	EDNRB-AS1	0.08	0.01	1.14E-13
4	9	15	28356859	rs1129038	C	T	0.26	HERC2	-0.23	0.01	2.48E-56
2	10	16	88895529	rs117614496	C	T	0.98	GALNS	-1.32	0.02	0.00E+00
2	11	16	89029902	rs117791738	C	T	0.98	CBFA2T3	-0.26	0.04	2.44E-13
2	12	16	89066873	rs9921065	A	G	0.02	CTD-2555A7.1	-0.68	0.07	9.59E-24
2	13	16	89085743	rs112584165	A	G	0.98	CTD-2555A7.1	-1.51	0.03	0.00E+00
2	14	16	89133302	rs12925392	A	C	0.92	CTD-2555A7.2	0.37	0.02	1.88E-50
1	15	16	89454991	rs117984432	T	C	0.97	ANKRD11	-0.46	0.02	9.19E-77
1	16	16	89568288	rs191414071	C	T	0.98	SPG7:RP11-104N10.1	-1.40	0.03	0
1	17	16	89666511	rs164737	A	G	0.72	CPNE7	-0.61	0.01	0
1	18	16	89715564	rs61482648	A	C	0.78	CHMP1A	-0.91	0.01	0

1	19	16	89749165	rs117030214	T	C	0.97	CDK10:RP11-368I7.4	-0.87	0.03	6.83E-243
1	20	16	89886519	rs34357723	C	T	0.74	SPIRE2	-2.17	0.01	0
1	21	16	90018389	rs11859970	T	C	0.87	DEF8	0.71	0.02	1.59E-221
7	22	17	79606820	rs9747347	T	C	0.35	NPLOC4:TSPAN10	0.12	0.01	2.26E-24
14	23	19	3546264	rs34878396	C	T	0.99	MFSD12:AC005786.7	0.34	0.06	2.30E-09
3	24	20	32665748	rs6059655	A	G	0.10	RALY	0.61	0.02	0
8	25	20	34960201	rs4519594	T	C	0.24	DLGAP4	0.14	0.01	3.59E-23

Table C.2: Independent signals for red hair colour GWAS

C.3 Tanning ability

Locus	Signal	CHR	BP	SNP	EA	NEA	EAF	Nearest Gene	BETA	SE	P
48	1	1	25251326	rs1507101	G	A	0.43	RUNX3	-0.02	0.00	8.22E-12
54	2	1	42118766	rs4393146	A	G	0.34	HIVEP3	-0.02	0.00	1.05E-10
29	3	1	63727542	rs670318	T	C	0.05	LINC00466	0.05	0.01	4.31E-20
19	4	1	66895085	rs1613999	T	G	0.57	PDE4B	0.03	0.00	2.09E-46
72	5	1	110720400	rs6689641	A	G	0.47	SLC6A17:RP5-1028L10.2	-0.01	0.00	5.67E-09
30	6	1	150789961	rs3894771	A	T	0.54	ARNT	0.02	0.00	7.43E-20
35	7	1	154994978	rs76798800	G	T	0.74	DCST2	-0.02	0.00	3.79E-15
24	8	1	205130413	rs3851294	A	G	0.09	DSTYK	0.05	0.00	2.10E-32
68	9	1	212417130	rs12727947	T	C	0.79	RP11-15I11.3	0.02	0.00	4.52E-09
47	10	2	25373298	rs13428823	G	A	0.35	EFR3B	-0.02	0.00	1.54E-12
40	11	2	27741237	rs780094	T	C	0.38	GCKR	0.02	0.00	3.83E-14
22	12	2	38298139	rs1800440	T	C	0.82	CYP1B1	0.04	0.00	2.09E-33
22	13	2	38298877	rs162561	T	G	0.18	CYP1B1	0.03	0.00	2.42E-32
67	14	2	43105088	rs12612692	T	A	0.64	AC098824.6	-0.01	0.00	4.11E-09
36	15	2	172381948	rs35380972	A	C	0.63	CYBRD1	0.02	0.00	6.85E-15
79	16	2	221964207	rs1541966	T	C	0.30	AC011233.2	-0.01	0.00	3.12E-08
37	17	3	85625852	rs7617323	G	T	0.38	CADM2	-0.02	0.00	1.32E-14
64	18	3	98698414	rs9821965	A	G	0.61	CTD-2021J15.1	0.01	0.00	2.53E-09

25	19	3	156492758	rs9818780	T	C	0.52	LINC00886	-0.02	0.00	5.03E-21
61	20	4	4346065	rs4689314	T	C	0.06	NSG1	-0.03	0.00	1.48E-09
4	21	5	33951693	rs16891982	C	G	0.07	SLC45A2	-0.46	0.01	0
27	22	5	59028853	rs11954036	T	C	0.68	PDE4D	-0.02	0.00	2.74E-20
62	23	5	60696108	rs74852363	C	T	0.88	ZSWIM6	0.02	0.00	2.05E-09
76	24	5	66347990	rs469394	T	G	0.60	MAST4	0.01	0.00	9.41E-09
21	25	5	149194485	rs251468	C	T	0.75	PPARGC1B	0.04	0.00	8.87E-42
38	26	5	173973627	rs6883391	C	T	0.73	RP11-267A15.1	0.02	0.00	2.81E-14
57	27	5	176731452	rs28362590	G	T	0.25	MXD3:PRELID1	0.02	0.00	6.62E-10
2	28	6	192288	rs61376093	T	C	0.95	NA	0.06	0.01	1.84E-18
2	29	6	375759	rs7753579	T	C	0.28	IRF4	-0.08	0.00	7.79E-139
2	30	6	396321	rs12203592	C	T	0.79	IRF4	-0.28	0.00	0
32	31	6	1142731	rs10214796	G	C	0.70	AL033381.1	-0.02	0.00	2.57E-16
66	32	6	3355595	rs9405194	G	A	0.68	SLC22A23	-0.01	0.00	3.87E-09
75	33	6	19538436	rs115949579	C	T	0.97	RP1-167F1.2	-0.04	0.01	8.61E-09
58	34	6	29935806	rs2517674	G	T	0.77	NA	-0.02	0.00	9.04E-10
42	35	6	41760990	rs6938966	T	G	0.25	USP49	-0.02	0.00	1.85E-13
77	36	6	82473953	rs1034241	T	A	0.91	RP5-991C6.2	-0.02	0.00	1.15E-08
65	37	6	146864384	rs73783709	T	G	0.97	RAB32	0.04	0.01	3.55E-09
53	38	6	164516218	rs62435865	T	C	0.89	RP1-155D22.2	-0.02	0.00	6.61E-11

11	39	7	16990804	rs1721040	T	A	0.59	AC098592.7	-0.02	0.00	3.39E-22
11	40	7	17041434	rs34585474	C	T	0.89	AC098592.8	-0.03	0.00	6.51E-18
11	41	7	17134708	rs117132860	G	A	0.97	AC098592.8	-0.15	0.01	5.88E-91
41	42	7	28177338	rs849138	G	A	0.51	JAZF1	-0.02	0.00	1.11E-13
63	43	7	104846634	rs56269269	G	A	0.48	SRPK2	0.01	0.00	2.41E-09
14	44	8	116613258	rs2721929	G	A	0.43	TRPS1	0.04	0.00	1.05E-74
13	45	9	12671566	rs10960749	G	A	0.40	TYRP1	-0.04	0.00	8.45E-76
9	46	9	16799109	rs10810636	A	G	0.27	BNC2	-0.06	0.00	3.63E-93
9	47	9	16885017	rs12350739	G	A	0.42	BNC2	-0.07	0.00	2.19E-213
50	48	10	5808086	rs1129614	G	A	0.79	GDI2	0.02	0.00	2.06E-11
45	49	10	64564892	rs2236295	G	T	0.61	RP11-436D10.3:ADO	0.02	0.00	3.86E-13
69	50	10	82199415	rs10788623	G	A	0.54	RP11-137H2.6	-0.01	0.00	4.63E-09
56	51	10	102102612	rs7086846	G	A	0.63	RP11-34D15.2	0.01	0.00	6.48E-10
15	52	10	119573178	rs7098111	C	T	0.84	RP11-355F22.1	0.06	0.00	2.45E-74
33	53	11	16217413	rs10766301	C	T	0.58	SOX6	-0.02	0.00	3.70E-11
33	54	11	16615883	rs72632979	A	G	0.83	SOX6	0.02	0.00	7.33E-16
80	55	11	47439444	rs2053979	A	G	0.67	PSMC3	-0.01	0.00	3.89E-08
10	56	11	68831364	rs72928978	G	A	0.90	TPCN2	-0.09	0.00	2.59E-116
10	57	11	68855363	rs3829241	G	A	0.61	TPCN2	-0.02	0.00	5.71E-18
74	58	11	78131408	rs10899501	C	T	0.83	RP11-452H21.1	0.02	0.00	8.29E-09

8	59	11	88303869	rs117796400	G	A	0.95	GRM5	0.04	0.01	1.10E-12
8	60	11	88484644	rs11021172	C	G	0.30	GRM5	-0.08	0.00	3.22E-218
6	61	11	89005267	rs183139540	G	A	0.99	TYR	0.17	0.01	7.98E-58
6	62	11	89017961	rs1126809	G	A	0.71	TYR	-0.13	0.00	0
6	63	11	89180697	rs78043739	G	A	0.97	NOX4	0.06	0.01	1.69E-16
6	64	11	89323846	rs200522762	T	G	0.86	NOX4:RP11-643G5.6	-0.04	0.00	2.28E-23
43	65	11	89822066	rs140100229	C	T	0.97	UBTFL1	0.06	0.01	2.50E-13
59	66	12	23979199	rs10771034	T	A	0.44	SOX5	0.01	0.00	9.70E-10
70	67	12	41123651	rs10879095	G	A	0.36	CNTN1:RP11-367O10.1	0.01	0.00	4.69E-09
23	68	12	88829294	rs11104870	C	T	0.31	Y_RNA	0.02	0.00	4.84E-15
23	69	12	88942980	rs11104947	G	A	0.98	KITLG	-0.11	0.01	1.67E-32
23	70	12	88962365	rs77081360	C	T	0.98	KITLG	-0.07	0.01	2.34E-15
46	71	12	116637641	rs61935849	A	C	0.89	MED13L	-0.03	0.00	1.05E-12
52	72	12	125387877	rs68099344	A	T	0.83	RNU6-927P	-0.02	0.00	5.67E-11
73	73	12	129306308	rs4760521	G	A	0.64	SLC15A4	-0.01	0.00	6.51E-09
31	74	13	95171058	rs9590030	G	C	0.68	DCT	-0.02	0.00	4.11E-19
16	75	13	113532990	rs1278761	T	C	0.45	ATP11A	-0.04	0.00	2.84E-55
18	76	14	92777462	rs1885194	T	C	0.56	SLC24A4	-0.03	0.00	1.18E-47
5	77	15	28062125	rs78114576	G	A	0.94	OCA2	-0.05	0.00	8.23E-29
5	78	15	28103388	rs142306097	T	C	0.97	OCA2	0.06	0.01	2.55E-16

5	79	15	28114816	rs75137181	G	A	0.98	OCA2	0.05	0.01	1.80E-10
5	80	15	28143305	rs139615925	T	C	0.96	OCA2	-0.05	0.01	4.89E-16
5	81	15	28157441	rs72710600	T	A	0.98	OCA2	-0.07	0.01	6.81E-14
5	82	15	28206623	rs141514981	C	T	0.96	OCA2	-0.09	0.01	2.47E-41
5	83	15	28230378	rs117886461	G	A	0.98	OCA2	-0.26	0.01	2.09E-152
5	84	15	28264167	rs77516282	G	T	0.96	OCA2	0.07	0.01	2.09E-30
5	85	15	28365618	rs12913832	A	G	0.26	HERC2	-0.16	0.00	0
26	86	15	31388787	rs17228129	G	A	0.80	TRPM1	0.03	0.00	1.39E-20
78	87	15	82427605	rs2455717	A	T	0.56	EFTUD1	0.01	0.00	1.54E-08
39	88	16	14010994	rs6498485	G	A	0.64	ERCC4	0.02	0.00	3.70E-14
51	89	16	71634811	rs4788815	A	T	0.34	RP11-432I5.1	-0.02	0.00	2.18E-11
55	90	16	73143479	rs4238968	C	A	0.39	HCCAT5	-0.01	0.00	3.00E-10
7	91	16	88929011	rs75898184	A	G	0.95	TRAPPC2L:PABPN1L	-0.20	0.01	2.29E-302
7	92	16	89121692	rs72815590	C	T	0.92	CTD-2555A7.2	0.09	0.00	3.66E-106
7	93	16	89368895	rs149610032	G	A	0.97	ANKRD11	0.12	0.01	6.03E-72
7	94	16	89370630	rs150020387	G	A	0.98	ANKRD11	0.09	0.01	7.27E-24
1	95	16	89592248	rs143743429	G	A	0.98	SPG7	-0.29	0.01	3.33E-240
1	96	16	89631186	rs59013041	A	G	0.94	RPL13	0.13	0.01	1.06E-139
1	97	16	89675362	rs139031001	C	T	0.98	DPEP1	0.12	0.01	1.61E-35
1	98	16	89749244	rs11641963	C	G	0.96	CDK10:RP11-368I7.4	0.17	0.01	2.60E-181

1	99	16	89799124	rs371629999	G	A	0.90	NA	-0.16	0.00	1.05E-228
1	100	16	89959387	rs73276549	A	G	0.98	TCF25	-0.20	0.01	6.73E-81
1	101	16	89985918	rs1805006	C	A	0.99	MC1R:AC092143.1:TUBB3	-0.36	0.01	1.95E-276
1	102	16	89986117	rs1805007	C	T	0.90	MC1R:AC092143.1:TUBB3	-0.44	0.00	0
1	103	16	90004693	rs150909008	G	A	0.85	TUBB3	-0.03	0.00	1.51E-18
1	104	16	90107861	rs117711828	G	C	0.91	GAS8:URAHP	-0.13	0.00	6.61E-227
28	105	17	79549250	rs34635363	G	A	0.65	NPLOC4	-0.02	0.00	2.75E-20
49	106	19	3537184	rs77733715	A	G	0.99	FZR1	-0.08	0.01	1.88E-11
49	107	19	3565253	rs6510760	G	A	0.92	MFSD12	-0.03	0.00	3.34E-09
60	108	20	3522015	rs151525	C	T	0.77	ATRN	-0.02	0.00	1.45E-09
71	109	20	21465380	rs6137358	A	T	0.66	RP5-984P4.1	-0.01	0.00	4.98E-09
3	110	20	32665748	rs6059655	A	G	0.10	RALY	0.28	0.00	0
3	111	20	32752091	rs138420028	C	T	0.93	ASIP	0.04	0.00	8.45E-21
3	112	20	32793500	rs6120568	T	C	0.91	ASIP	0.07	0.00	2.26E-54
3	113	20	32944549	rs13042880	C	T	0.32	NA	0.04	0.00	1.33E-33
17	114	20	33865588	rs6060343	G	C	0.87	NA	0.06	0.00	4.39E-48
12	115	20	34960201	rs4519594	T	C	0.24	DLGAP4	0.06	0.00	1.06E-84
44	116	21	43429646	rs399907	G	A	0.33	ZBTB21:ZNF295-AS1	0.02	0.00	3.17E-13
34	117	22	38244216	rs4821721	C	T	0.10	ANKRD54	0.02	0.00	7.90E-10
34	118	22	38572440	rs4384	G	C	0.54	PLA2G6	0.02	0.00	3.65E-15

20	119	22	45622014	rs6007506	C	T	0.66	KIAA0930	0.04	0.00	2.19E-46

Table C.3: Independent signals for tanning ability GWAS

C.4 Skin colour

Locus	Signal	CHR	BP	SNP	EA	NEA	EAF	Nearest Gene	BETA	SE	P
29	1	1	7731730	rs4908473	T	C	0.92	CAMTA1	0.01	0.00	1.76E-09
29	2	1	8207579	rs80293268	G	C	0.95	RP11-431K24.1	0.03	0.00	5.67E-21
73	3	1	25259884	rs3845302	C	G	0.42	RUNX3	0.01	0.00	1.57E-09
60	4	1	42110888	rs6696511	T	C	0.37	HIVEP3	0.01	0.00	5.12E-11
44	5	1	63727542	rs670318	T	C	0.05	LINC00466	-0.02	0.00	1.49E-14
23	6	1	66895085	rs1613999	T	G	0.57	PDE4B	-0.02	0.00	9.29E-29
22	7	1	110724488	rs535930	G	A	0.47	SLC6A17	0.02	0.00	8.63E-31
49	8	1	150789961	rs3894771	A	T	0.54	ARNT	-0.01	0.00	1.71E-12
61	9	1	170510306	rs4656781	C	T	0.38	GORAB	-0.01	0.00	6.87E-11
25	10	1	205130413	rs3851294	A	G	0.09	DSTYK	-0.03	0.00	6.87E-27
52	11	2	27742603	rs780093	T	C	0.38	GCKR	-0.01	0.00	6.91E-12
41	12	2	28613302	rs71443018	G	C	0.94	AC104695.3	0.02	0.00	7.06E-15
77	13	2	33001100	rs7584123	T	C	0.61	TTC27	-0.01	0.00	6.05E-09
27	14	2	38291967	rs10169939	T	A	0.59	RMDN2:RMDN2-AS1	-0.01	0.00	9.92E-23
70	15	2	43302571	rs10185673	G	A	0.69	RNU6-242P	0.01	0.00	5.24E-10
66	16	2	172378893	rs3731976	G	A	0.61	CYBRD1	-0.01	0.00	2.39E-10
93	17	2	175292326	rs7591631	C	A	0.55	SCRN3	-0.01	0.00	2.31E-08
30	18	2	222034085	rs62186153	A	G	0.79	AC011233.2	0.02	0.00	1.84E-19

30	19	2	222072535	rs10192020	G	A	0.59	AC011233.2	0.01	0.00	1.85E-14
42	20	2	234668570	rs887829	C	T	0.68	UGT1A8:UGT1A10: UGT1A9:UGT1A7: UGT1A6:UGT1A5: UGT1A4:UGT1A3	-0.01	0.00	8.48E-15
96	21	3	85434260	rs11921010	T	C	0.41	CADM2	0.01	0.00	4.80E-08
43	22	3	141154542	rs6440008	T	C	0.62	ZBTB38	0.01	0.00	1.06E-14
43	23	3	141576863	rs4683632	A	G	0.65	HMGN2P25	0.01	0.00	1.03E-09
32	24	3	156491160	rs9867857	C	T	0.52	LINC00886	0.01	0.00	9.20E-19
46	25	3	181511951	rs1345417	C	G	0.40	SOX2-OT	-0.01	0.00	1.54E-13
87	26	4	75341885	rs6819541	A	G	0.59	AREG	-0.01	0.00	1.70E-08
3	27	5	33887419	rs4866399	G	A	0.71	ADAMTS12	-0.01	0.00	2.31E-12
3	28	5	33936721	rs79014396	C	T	0.99	RXFP3	-0.08	0.01	6.97E-32
3	29	5	33951693	rs16891982	C	G	0.07	SLC45A2	0.41	0.00	0
47	30	5	53112624	rs6875907	C	T	0.48	CTD-2081C10.1	0.01	0.00	2.10E-13
39	31	5	59018067	rs7720119	C	G	0.67	PDE4D	0.01	0.00	3.57E-15
53	32	5	66321379	rs2662225	G	C	0.58	NA	-0.01	0.00	1.18E-11
50	33	5	90277037	rs6887203	A	G	0.62	GPR98	-0.01	0.00	3.49E-12
94	34	5	113996889	rs285894	G	C	0.52	NA	-0.01	0.00	2.62E-08
89	35	5	133861756	rs329120	C	T	0.58	JADE2	0.01	0.00	2.00E-08

21	36	5	149195603	rs251466	C	G	0.75	PPARGC1B	-0.02	0.00	1.92E-32
78	37	5	173975145	rs923520	T	C	0.73	SUMO2P6	-0.01	0.00	6.45E-09
5	38	6	396321	rs12203592	C	T	0.79	IRF4	0.09	0.00	0
71	39	6	19538436	rs115949579	C	T	0.97	RP1-167F1.2	0.02	0.00	1.13E-09
79	40	6	29719561	rs909728	T	C	0.42	NA	0.01	0.00	6.58E-09
88	41	6	41892809	rs10947996	G	T	0.75	BYSL	-0.01	0.00	1.92E-08
59	42	6	51696355	rs13201830	A	T	0.53	PKHD1	-0.01	0.00	4.71E-11
92	43	6	82225675	rs9443928	G	A	0.75	FAM46A	0.01	0.00	2.29E-08
48	44	6	146847519	rs1407373	G	C	0.97	RAB32	-0.03	0.00	5.29E-13
45	45	6	159242809	rs3123139	A	G	0.20	EZR-AS1	0.01	0.00	1.19E-13
17	46	7	17004301	rs1721019	C	A	0.60	AC098592.7	0.01	0.00	2.22E-14
17	47	7	17041434	rs34585474	C	T	0.89	AC098592.8	0.01	0.00	4.00E-08
17	48	7	17134708	rs117132860	G	A	0.97	AC098592.8	0.06	0.00	3.55E-45
95	49	7	73012042	rs35332062	G	A	0.87	MLXIPL	-0.01	0.00	2.90E-08
40	50	7	100499789	rs56134000	T	C	0.76	RN7SL549P	0.01	0.00	4.43E-15
33	51	7	104849737	rs7794285	A	G	0.46	SRPK2	-0.01	0.00	3.17E-18
18	52	8	116601902	rs6469606	C	T	0.42	TRPS1	-0.02	0.00	6.56E-42
18	53	8	116823046	rs6981915	T	C	0.97	TRPS1	0.03	0.01	4.86E-09
86	54	9	5874777	rs4601374	G	A	0.76	KIAA2026	0.01	0.00	1.60E-08
12	55	9	12675342	rs75630385	A	G	0.98	TYRP1	-0.05	0.00	1.24E-27

12	56	9	12712157	rs1137134	G	A	0.42	RP11-3L8.3	0.03	0.00	5.77E-100
9	57	9	16802973	rs2254330	A	G	0.75	BNC2	-0.03	0.00	1.76E-74
9	58	9	16885017	rs12350739	G	A	0.42	BNC2	0.04	0.00	2.58E-217
75	59	9	125884052	rs10985890	C	T	0.92	STRBP	-0.02	0.00	1.65E-09
84	60	9	129801021	rs76798990	T	C	0.89	RALGPS1	0.01	0.00	1.30E-08
69	61	10	64564892	rs2236295	G	T	0.61	RP11-436D10.3:ADO	-0.01	0.00	5.19E-10
76	62	10	82089493	rs55901776	T	G	0.52	RP11-36D19.8	0.01	0.00	3.18E-09
62	63	10	102102612	rs7086846	G	A	0.63	RP11-34D15.2	-0.01	0.00	8.76E-11
14	64	10	119572403	rs35563099	C	T	0.84	RP11-355F22.1	-0.03	0.00	7.68E-55
34	65	11	16217413	rs10766301	C	T	0.58	SOX6	0.01	0.00	1.50E-16
34	66	11	16615883	rs72632979	A	G	0.83	SOX6	-0.02	0.00	3.30E-18
63	67	11	18319915	rs7478729	G	A	0.38	HPS5	0.01	0.00	1.02E-10
10	68	11	68817897	rs150527451	G	A	0.90	TPCN2	0.06	0.00	1.29E-184
10	69	11	68855363	rs3829241	G	A	0.61	TPCN2	0.02	0.00	2.94E-38
26	70	11	78128047	rs881361	C	G	0.83	GAB2:RP11-452H21.1	-0.02	0.00	3.79E-24
11	71	11	88484644	rs11021172	C	G	0.30	GRM5	0.04	0.00	1.16E-171
6	72	11	89005318	rs144650507	C	T	0.97	TYR	-0.07	0.00	8.09E-62
6	73	11	89017961	rs1126809	G	A	0.71	TYR	0.07	0.00	0
81	74	11	89966202	rs10830452	A	G	0.67	CHORDC1	-0.01	0.00	1.03E-08
65	75	12	23979199	rs10771034	T	A	0.44	SOX5	-0.01	0.00	1.47E-10

54	76	12	41097386	rs2405295	G	T	0.34	CNTN1	-0.01	0.00	1.30E-11
83	77	12	46774771	rs7979418	A	G	0.38	RP11-96H19.1	-0.01	0.00	1.19E-08
57	78	12	85693252	rs12425342	G	A	0.95	ALX1	0.02	0.00	2.63E-11
13	79	12	88847783	rs11104887	C	T	0.89	Y_RNA	-0.02	0.00	2.94E-22
13	80	12	88942980	rs11104947	G	A	0.98	KITLG	0.05	0.01	1.58E-16
13	81	12	89328335	rs12821256	T	C	0.89	RP11-13A1.1	0.03	0.00	2.07E-62
51	82	12	90485617	rs61924662	G	A	0.66	RP11-567C2.1	0.01	0.00	5.98E-12
72	83	12	116745298	rs113159861	G	A	0.90	MED13L	0.01	0.00	1.14E-09
67	84	12	125389272	rs7304293	T	C	0.24	RNU6-927P	-0.01	0.00	3.31E-10
64	85	13	78382705	rs1823554	A	C	0.40	EDNRB-AS1	0.01	0.00	1.02E-10
16	86	13	95171058	rs9590030	G	C	0.68	DCT	0.02	0.00	3.48E-46
19	87	13	113532990	rs1278761	T	C	0.45	ATP11A	0.02	0.00	4.80E-40
56	88	14	50655357	rs72681869	G	C	0.99	SOS2	-0.04	0.01	2.10E-11
8	89	14	92776825	rs941799	C	T	0.56	SLC24A4	0.04	0.00	8.08E-223
8	90	14	92815008	rs76519749	T	C	0.89	SLC24A4	0.03	0.00	3.23E-35
1	91	15	28130503	rs35591793	C	T	0.97	OCA2	-0.03	0.00	3.99E-09
1	92	15	28230378	rs117886461	G	A	0.98	OCA2	0.13	0.01	3.51E-108
1	93	15	28266235	rs746861	T	C	0.52	OCA2	-0.02	0.00	1.70E-50
1	94	15	28365618	rs12913832	A	G	0.26	HERC2	0.20	0.00	0
35	95	15	31395538	rs28456199	G	A	0.92	TRPM1	-0.02	0.00	4.31E-17

74	96	15	50276245	rs57249121	T	C	0.78	ATP8B4	-0.01	0.00	1.59E-09
58	97	15	81526279	rs72744156	C	T	0.80	IL16	0.01	0.00	4.29E-11
68	98	15	83383187	rs28371837	T	G	0.81	AC105339.1	0.01	0.00	5.10E-10
80	99	15	91116386	rs12901150	T	G	0.25	CRTC3	-0.01	0.00	8.84E-09
85	100	16	4434198	rs758045	T	C	0.23	CORO7-PAM16:CORO7	0.01	0.00	1.60E-08
37	101	16	71990424	rs8054175	G	A	0.26	PKD1L3	0.01	0.00	1.66E-15
7	102	16	89067671	rs113894462	G	A	0.94	CTD-2555A7.1	0.10	0.00	1.07E-288
2	103	16	89713938	rs11648089	T	C	0.89	CHMP1A	-0.09	0.00	0
2	104	16	89777078	rs56288641	G	A	0.98	VPS9D1	0.18	0.00	0
2	105	16	89799124	rs371629999	G	A	0.90	NA	0.08	0.00	2.38E-189
2	106	16	89808241	rs75092208	T	C	0.99	FANCA	-0.10	0.01	2.44E-60
2	107	16	89985918	rs1805006	C	A	0.99	MC1R:AC092143.1:TUBB3	0.17	0.01	3.24E-180
2	108	16	89986117	rs1805007	C	T	0.90	MC1R:AC092143.1:TUBB3	0.23	0.00	0
2	109	16	90107861	rs117711828	G	C	0.91	GAS8:URAHP	0.07	0.00	1.75E-170
31	110	17	45921906	rs2325752	T	C	0.46	SP6	0.01	0.00	4.45E-19
55	111	17	48430291	rs739992	C	T	0.54	XYLT2	0.01	0.00	1.90E-11
91	112	17	73426048	rs149789627	G	A	0.92	Y_RNA	-0.01	0.00	2.25E-08
28	113	17	79612397	rs6420484	A	G	0.35	NPLOC4:TSPAN10	-0.01	0.00	5.28E-21
82	114	19	39437559	rs1865048	C	T	0.66	SARS2:CTC-360G5.8:FBXO17	-0.01	0.00	1.05E-08
90	115	20	21457672	rs2249048	C	A	0.67	RP5-984P4.1	0.01	0.00	2.20E-08

4	116	20	32665748	rs6059655	A	G	0.10	RALY	-0.13	0.00	0
4	117	20	32793500	rs6120568	T	C	0.91	ASIP	-0.04	0.00	1.41E-46
4	118	20	32944549	rs13042880	C	T	0.32	NA	-0.02	0.00	3.43E-23
24	119	20	33865588	rs6060343	G	C	0.87	NA	-0.03	0.00	1.03E-27
15	120	20	34960201	rs4519594	T	C	0.24	DLGAP4	-0.03	0.00	4.82E-49
38	121	21	44793448	rs672948	A	T	0.40	AP001046.6	-0.01	0.00	2.53E-15
36	122	22	38612604	rs34066050	G	A	0.48	MAFF	0.01	0.00	7.67E-16
20	123	22	45622684	rs5766565	A	G	0.66	KIAA0930	-0.02	0.00	1.63E-38

Table C.4: Independent signals for skin colour GWAS

C.5 Number of childhood sunburns

Locus	Signal	CHR	BP	SNP	EA	NEA	EAF	Nearest Gene	BETA	SE	P
11	1	2	25393388	rs76645364	A	G	0.82	POMC	0.11	0.01	4.84E-14
9	2	5	173976464	rs60780889	C	T	0.73	SUMO2P6	-0.12	0.01	1.46E-22
6	3	6	396321	rs12203592	C	T	0.79	IRF4	-0.14	0.01	7.94E-28
5	4	6	51697567	rs2025753	T	C	0.53	PKHD1	0.13	0.01	7.60E-33
5	5	6	51713402	rs12529074	G	A	0.88	PKHD1	0.14	0.02	6.82E-14
10	6	11	69388143	rs652190	G	A	0.34	CCND1	0.09	0.01	2.93E-15
13	7	11	88966584	rs4512823	A	G	0.66	TYR	0.08	0.01	4.26E-10
12	8	13	78392656	rs9544609	C	A	0.41	EDNRB-AS1	0.08	0.01	1.14E-13
4	9	15	28356859	rs1129038	C	T	0.26	HERC2	-0.23	0.01	2.48E-56
2	10	16	88895529	rs117614496	C	T	0.98	GALNS	-1.32	0.02	0.00E+00
2	11	16	89029902	rs117791738	C	T	0.98	CBFA2T3	-0.26	0.04	2.44E-13
2	12	16	89066873	rs9921065	A	G	0.02	CTD-2555A7.1	-0.68	0.07	9.59E-24
2	13	16	89085743	rs112584165	A	G	0.98	CTD-2555A7.1	-1.51	0.03	0.00E+00
2	14	16	89133302	rs12925392	A	C	0.92	CTD-2555A7.2	0.37	0.02	1.88E-50
1	15	16	89454991	rs117984432	T	C	0.97	ANKRD11	-0.46	0.02	9.19E-77
1	16	16	89568288	rs191414071	C	T	0.98	SPG7:RP11-104N10.1	-1.40	0.03	0
1	17	16	89666511	rs164737	A	G	0.72	CPNE7	-0.61	0.01	0
1	18	16	89715564	rs61482648	A	C	0.78	CHMP1A	-0.91	0.01	0

1	19	16	89749165	rs117030214	T	C	0.97	CDK10:RP11-368I7.4	-0.87	0.03	6.83E-243
1	20	16	89886519	rs34357723	C	T	0.74	SPIRE2	-2.17	0.01	0
1	21	16	90018389	rs11859970	T	C	0.87	DEF8	0.71	0.02	1.59E-221
7	22	17	79606820	rs9747347	T	C	0.35	NPLOC4:TSPAN10	0.12	0.01	2.26E-24
14	23	19	3546264	rs34878396	C	T	0.99	MFSD12:AC005786.7	0.34	0.06	2.30E-09
3	24	20	32665748	rs6059655	A	G	0.10	RALY	0.61	0.02	0
8	25	20	34960201	rs4519594	T	C	0.24	DLGAP4	0.14	0.01	3.59E-23

Table C.5: Independent signals for number of childhood sunburns GWAS

Appendix D

Overall pigmentation GWAS summary statistics

D.1 Pigscore 1

Locus	Signal	CHR	BP	SNP	EA	NEA	EAF	Nearest Gene	BETA	SE	P
16	1	1	8206827	rs112147823	T	C	0.91	RP11-431K24.1	0.04	0.01	4.00E-11
16	2	1	8207579	rs80293268	G	C	0.95	RP11-431K24.1	0.14	0.01	1.21E-57
72	3	1	11034165	rs12081181	T	G	0.77	C1orf127	0.03	0.00	1.63E-12
82	4	1	17607501	rs11584287	G	C	0.85	PADI3	0.03	0.00	4.08E-11
108	5	1	42112236	rs6702659	T	C	0.34	HIVEP3	0.02	0.00	1.54E-09
62	6	1	63727542	rs670318	T	C	0.05	LINC00466	-0.06	0.01	2.71E-14
47	7	1	66895085	rs1613999	T	G	0.57	PDE4B	-0.03	0.00	1.45E-17
73	8	1	110720400	rs6689641	A	G	0.47	SLC6A17:RP5-1028L10.2	0.02	0.00	3.18E-12
125	9	1	150618961	rs9733	T	G	0.38	GOLPH3L	0.02	0.00	1.49E-08
75	10	1	154994978	rs76798800	G	T	0.74	DCST2	0.03	0.00	1.21E-11
15	11	1	205130413	rs3851294	A	G	0.09	DSTYK	-0.10	0.01	1.44E-61
90	12	1	212426172	rs351376	C	G	0.41	RP11-15I11.3	0.02	0.00	1.42E-10
111	13	1	236045229	rs4660121	G	A	0.12	LYST	-0.03	0.01	2.47E-09
40	14	2	25393388	rs76645364	A	G	0.82	POMC	-0.04	0.00	1.36E-20
40	15	2	25532969	rs13036246	C	T	0.49	DNMT3A	-0.02	0.00	6.52E-09
135	16	2	26337154	rs112272576	T	A	0.90	RAB10	-0.03	0.01	3.55E-08
63	17	2	27741237	rs780094	T	C	0.38	GCKR	-0.03	0.00	2.89E-14
21	18	2	28613302	rs71443018	G	C	0.94	AC104695.3	0.12	0.01	2.30E-51

41	19	2	38298139	rs1800440	T	C	0.82	CYP1B1	-0.04	0.00	1.52E-20
41	20	2	38298877	rs162561	T	G	0.18	CYP1B1	-0.04	0.00	9.21E-20
51	21	2	42140464	rs1962386	A	T	0.39	AC104654.2	0.03	0.00	2.90E-16
107	22	2	88556133	rs6730279	T	C	0.86	RNY4P15	0.03	0.01	1.51E-09
106	23	2	119538982	rs6716872	G	A	0.75	RP11-19E11.1	-0.02	0.00	1.50E-09
120	24	2	135353183	rs13410301	C	T	0.41	TMEM163	-0.02	0.00	5.84E-09
100	25	2	172381948	rs35380972	A	C	0.63	CYBRD1	-0.02	0.00	6.56E-10
22	26	2	222051419	rs10169459	T	C	0.42	AC011233.2	-0.05	0.00	2.13E-47
22	27	2	222089797	rs17349283	A	G	0.56	AC011233.2	0.05	0.00	1.49E-41
78	28	2	223121674	rs1430663	T	C	0.87	PAX3	-0.04	0.01	2.52E-11
102	29	2	234668570	rs887829	C	T	0.68	UGT1A8:UGT1A10: UGT1A9:UGT1A7:UGT1A6:UGT1A5: UGT1A4:UGT1A3	-0.02	0.00	6.98E-10
93	30	3	69934529	rs17006579	A	G	0.63	MITF	-0.02	0.00	2.87E-10
57	31	3	122526816	rs9847240	G	A	0.34	DIRC2	0.03	0.00	4.18E-15
35	32	3	141154542	rs6440008	T	C	0.62	ZBTB38	0.04	0.00	5.43E-25
35	33	3	141634056	rs3804772	G	A	0.88	ATP1B3	0.04	0.01	7.54E-12
54	34	3	156491160	rs9867857	C	T	0.52	LINC00886	0.03	0.00	1.76E-15
81	35	4	4346065	rs4689314	T	C	0.06	NSG1	0.05	0.01	3.37E-11
134	36	4	38820986	rs7696175	T	C	0.46	TLR1	0.02	0.00	3.29E-08

64	37	4	75328479	rs1874202	C	G	0.59	AREG	-0.03	0.00	1.14E-13
74	38	4	79256036	rs436034	T	C	0.26	FRAS1	0.03	0.00	4.05E-12
66	39	4	81661426	rs72661730	G	A	0.78	C4orf22	0.03	0.00	1.45E-13
95	40	4	86602918	rs17392334	C	T	0.52	ARHGAP24	-0.02	0.00	3.04E-10
130	41	4	109246930	rs999870	A	G	0.11	LEF1-AS1	0.03	0.01	3.12E-08
4	42	5	33887419	rs4866399	G	A	0.71	ADAMTS12	-0.03	0.00	2.37E-13
4	43	5	33930012	rs116125333	T	G	0.98	RXFP3	-0.07	0.01	1.65E-09
4	44	5	33936721	rs79014396	C	T	0.99	RXFP3	-0.18	0.02	5.24E-27
4	45	5	33951693	rs16891982	C	G	0.07	SLC45A2	0.99	0.01	0
27	46	5	52902307	rs427527	A	G	0.61	NDUFS4	-0.02	0.00	8.05E-11
27	47	5	53112624	rs6875907	C	T	0.48	CTD-2081C10.1	0.04	0.00	7.61E-36
42	48	5	59018067	rs7720119	C	G	0.67	PDE4D	0.03	0.00	4.16E-20
42	49	5	59133111	rs62370543	C	T	0.96	PDE4D	-0.06	0.01	1.47E-10
70	50	5	66321379	rs2662225	G	C	0.58	NA	-0.03	0.00	1.29E-12
38	51	5	90277797	rs60325490	T	C	0.62	GPR98	-0.04	0.00	7.04E-23
98	52	5	133852468	rs10479082	G	A	0.76	RN7SL541P	0.03	0.00	6.16E-10
36	53	5	149195603	rs251466	C	G	0.75	PPARGC1B	-0.04	0.00	1.18E-23
109	54	5	174156168	rs4242182	T	C	0.11	MSX2	-0.04	0.01	1.87E-09
48	55	6	515393	rs9392618	C	T	0.41	EXOC2	-0.03	0.00	1.46E-16
131	56	6	1144979	rs57354745	C	A	0.71	AL033381.1	0.02	0.00	3.20E-08

110	57	6	10545282	rs77571968	A	G	0.89	GCNT2	0.03	0.01	2.44E-09
87	58	6	19538436	rs115949579	C	T	0.97	RP1-167F1.2	0.06	0.01	1.12E-10
103	59	6	20658012	rs4710940	A	C	0.64	CDKAL1	0.02	0.00	8.85E-10
85	60	6	30982544	rs9394026	G	A	0.77	NA	-0.03	0.00	7.98E-11
71	61	6	41892809	rs10947996	G	T	0.75	BYSL	-0.03	0.00	1.29E-12
39	62	6	51696355	rs13201830	A	T	0.53	PKHD1	-0.03	0.00	3.35E-22
128	63	6	113377048	rs2185710	A	G	0.39	FCF1P10	0.02	0.00	2.20E-08
114	64	6	146847519	rs1407373	G	C	0.97	RAB32	-0.07	0.01	4.87E-09
76	65	6	151577739	rs10434895	A	T	0.55	AKAP12	0.02	0.00	1.65E-11
26	66	6	159234178	rs3123135	A	G	0.88	EZR	-0.03	0.01	1.96E-08
26	67	6	159247859	rs1471836	T	C	0.20	EZR-AS1	0.06	0.00	5.39E-36
24	68	7	16990804	rs1721040	T	A	0.59	AC098592.7	0.03	0.00	6.54E-14
24	69	7	17134708	rs117132860	G	A	0.97	AC098592.8	0.15	0.01	1.01E-42
91	70	7	73012042	rs35332062	G	A	0.87	MLXIPL	-0.03	0.01	1.75E-10
69	71	7	90905531	rs59201311	C	T	0.85	FZD1	-0.04	0.00	1.10E-12
53	72	7	100451732	rs12535629	C	T	0.72	SLC12A9	0.03	0.00	9.36E-16
83	73	7	104849737	rs7794285	A	G	0.46	SRPK2	-0.02	0.00	5.92E-11
49	74	7	105416560	rs2529369	C	A	0.71	ATXN7L1	-0.03	0.00	1.84E-16
37	75	7	130761235	rs10954300	G	A	0.29	LINC-PINT	0.04	0.00	3.51E-23
113	76	8	9229139	rs11779019	T	G	0.83	RP11-115J16.1	-0.03	0.00	4.57E-09

77	77	8	22600953	rs2048651	T	C	0.36	RP11-459E5.1:PEBP4	0.02	0.00	2.29E-11
97	78	8	42045933	rs2070711	G	A	0.88	PLAT	0.03	0.01	3.34E-10
46	79	8	82774442	rs12545508	C	T	0.74	SNX16	0.03	0.00	1.42E-17
28	80	8	116555360	rs12543799	C	A	0.44	TRPS1	-0.04	0.00	1.59E-35
28	81	8	116837897	rs7842594	G	A	0.97	TRPS1	0.08	0.01	1.53E-08
61	82	9	5874777	rs4601374	G	A	0.76	KIAA2026	0.03	0.00	2.58E-14
11	83	9	12596242	rs75123768	T	A	0.98	TYRP1	-0.13	0.01	3.64E-22
11	84	9	12712157	rs1137134	G	A	0.42	RP11-3L8.3	0.09	0.00	7.93E-133
9	85	9	16799623	rs10962605	A	G	0.27	BNC2	0.09	0.00	2.48E-94
9	86	9	16885017	rs12350739	G	A	0.42	BNC2	0.11	0.00	1.41E-217
105	87	9	27544943	rs10967976	G	A	0.48	C9orf72	0.02	0.00	1.42E-09
86	88	9	100660136	rs35324451	A	T	0.38	C9orf156	0.02	0.00	8.28E-11
104	89	9	102733298	rs1055187	T	C	0.68	STX17	0.02	0.00	1.33E-09
32	90	9	126808021	rs16936765	C	G	0.82	LHX2	0.05	0.00	7.97E-27
32	91	9	126814579	rs4838097	T	C	0.87	LHX2	-0.06	0.01	4.79E-29
32	92	9	126991185	rs10818930	T	G	0.20	NEK6	0.03	0.00	2.83E-12
65	93	10	35496626	rs1148246	C	T	0.35	CREM	-0.03	0.00	1.45E-13
84	94	10	64579460	rs1888967	T	A	0.56	EGR2	-0.02	0.00	7.36E-11
79	95	10	74068998	rs4433500	G	A	0.64	DNAJB12	0.02	0.00	2.90E-11
45	96	10	80944147	rs703978	C	G	0.41	ZMIZ1	0.03	0.00	1.08E-17

92	97	10	102103508	rs56014906	A	G	0.63	RP11-34D15.2	-0.02	0.00	1.81E-10
25	98	10	119572403	rs35563099	C	T	0.84	RP11-355F22.1	-0.06	0.00	4.04E-38
67	99	11	7541511	rs4078279	C	T	0.42	PPFIBP2	0.03	0.00	6.51E-13
59	100	11	15710084	rs7108738	T	G	0.82	RP11-396O20.1	-0.03	0.00	7.45E-13
59	101	11	15918766	rs4756838	C	T	0.84	RP11-222N13.1	0.04	0.00	1.28E-14
34	102	11	16305351	rs12277740	G	A	0.94	SOX6	-0.06	0.01	8.24E-18
34	103	11	16381813	rs7929530	C	A	0.85	SOX6	0.03	0.00	5.50E-10
34	104	11	16615883	rs72632979	A	G	0.83	SOX6	-0.05	0.00	2.36E-25
60	105	11	18319915	rs7478729	G	A	0.38	HPS5	0.03	0.00	1.40E-14
116	106	11	44330610	rs66716358	C	T	0.50	ALX4	0.02	0.00	5.33E-09
123	107	11	62203865	rs10897275	G	A	0.68	AHNAK	0.02	0.00	1.40E-08
8	108	11	68817441	rs72917317	T	G	0.90	TPCN2	0.24	0.01	0
8	109	11	68855363	rs3829241	G	A	0.61	TPCN2	0.07	0.00	5.87E-76
23	110	11	78131408	rs10899501	C	T	0.83	RP11-452H21.1	-0.07	0.00	4.57E-45
58	111	11	88303869	rs117796400	G	A	0.95	GRM5	-0.06	0.01	4.21E-15
6	112	11	88920675	rs1942493	C	T	0.89	TYR	-0.18	0.01	3.21E-219
6	113	11	89005267	rs183139540	G	A	0.99	TYR	-0.24	0.02	6.83E-50
6	114	11	89017961	rs1126809	G	A	0.71	TYR	0.18	0.00	0
6	115	11	89323846	rs200522762	T	G	0.86	NOX4:RP11-643G5.6	0.07	0.01	2.28E-22
6	116	11	89421092	rs2634309	C	T	0.79	FOLH1B	-0.04	0.00	4.65E-20

68	117	11	90432150	rs17201363	C	T	0.97	DISC1FP1	-0.08	0.01	7.59E-13
56	118	12	23979791	rs9971729	A	C	0.43	SOX5	-0.03	0.00	3.84E-15
126	119	12	46774771	rs7979418	A	G	0.38	RP11-96H19.1	-0.02	0.00	1.67E-08
55	120	12	85693252	rs12425342	G	A	0.95	ALX1	0.07	0.01	2.00E-15
43	121	12	88143958	rs73200863	G	C	0.91	CYCSP30	0.06	0.01	2.13E-18
19	122	12	88707968	rs10858711	C	T	0.85	TMTC3	-0.08	0.01	8.99E-53
10	123	12	88957352	rs10858760	A	C	0.97	KITLG	0.09	0.01	9.92E-10
10	124	12	89328335	rs12821256	T	C	0.89	RP11-13A1.1	0.16	0.01	3.28E-196
94	125	12	111708458	rs1265564	A	C	0.58	CUX2	-0.02	0.00	2.92E-10
99	126	12	116535976	rs61939692	G	A	0.89	MED13L	0.03	0.01	6.19E-10
88	127	13	39343822	rs9603422	C	T	0.88	FREM2	-0.04	0.01	1.23E-10
18	128	13	78385068	rs1766344	A	T	0.40	EDNRB-AS1	0.06	0.00	3.62E-53
14	129	13	95171411	rs7337275	A	G	0.65	DCT	0.06	0.00	1.15E-61
30	130	13	113532990	rs1278761	T	C	0.45	ATP11A	0.04	0.00	1.30E-34
118	131	14	50655357	rs72681869	G	C	0.99	SOS2	-0.10	0.02	5.57E-09
127	132	14	64390030	rs10873172	G	C	0.28	SYNE2	-0.02	0.00	1.88E-08
5	133	14	92776825	rs941799	C	T	0.56	SLC24A4	0.18	0.00	0
5	134	14	92815008	rs76519749	T	C	0.89	SLC24A4	0.13	0.01	5.00E-98
89	135	14	104184322	rs7359090	A	G	0.64	ZFYVE21	-0.02	0.00	1.23E-10
2	136	15	28014947	rs146704333	C	A	0.98	OCA2	-0.13	0.01	6.98E-20

2	137	15	28038935	rs35690998	C	G	0.97	OCA2	-0.11	0.01	5.52E-25
2	138	15	28201539	rs17567007	C	T	0.95	NA	-0.05	0.01	1.39E-08
2	139	15	28230378	rs117886461	G	A	0.98	OCA2	0.46	0.02	6.87E-209
2	140	15	28266235	rs746861	T	C	0.52	OCA2	-0.05	0.00	1.96E-44
2	141	15	28271247	rs56108911	G	A	0.93	OCA2	-0.08	0.01	7.69E-34
2	142	15	28365618	rs12913832	A	G	0.26	HERC2	0.65	0.00	0
129	143	15	31410084	rs4779826	A	T	0.37	TRPM1	0.02	0.00	2.56E-08
33	144	15	81526302	rs12324786	C	T	0.79	IL16	0.05	0.00	1.51E-25
112	145	15	83260436	rs4779035	T	C	0.81	CPEB1	0.03	0.00	3.06E-09
115	146	16	4527109	rs11639998	T	C	0.27	NMRAL1:HMOX2	0.02	0.00	4.95E-09
52	147	16	71987452	rs12935795	G	C	0.26	PKD1L3	0.03	0.00	3.71E-16
7	148	16	88711819	rs9940427	G	T	0.97	CYBA	-0.13	0.01	6.54E-39
7	149	16	88739665	rs112787727	G	A	0.95	SNAI3-AS1	0.35	0.01	0
7	150	16	88896100	rs116004099	T	C	0.92	GALNS	-0.15	0.01	6.21E-115
1	151	16	89494760	rs9923267	C	G	0.66	ANKRD11	0.06	0.00	1.38E-44
1	152	16	89618530	rs4785577	G	A	0.73	SPG7	-0.30	0.00	0
1	153	16	89754335	rs72807518	G	A	0.98	CDK10	0.43	0.01	6.00E-228
1	154	16	89866177	rs34427386	G	A	0.99	FANCA	0.29	0.02	2.55E-72
1	155	16	89896481	rs7185897	C	A	0.54	SPIRE2	0.12	0.00	1.35E-223
1	156	16	89986117	rs1805007	C	T	0.90	MC1R:AC092143.1:TUBB3	1.13	0.01	0

1	157	16	90135457	rs137938320	C	T	0.98	PRDM7	0.95	0.01	0
20	158	17	45950721	rs72833470	A	G	0.73	RP11-6N17.2	0.06	0.00	5.39E-52
132	159	17	48431032	rs739990	G	A	0.76	XYLT2	0.02	0.00	3.25E-08
117	160	17	55231168	rs62060349	T	C	0.56	AKAP1	0.02	0.00	5.44E-09
13	161	17	79606820	rs9747347	T	C	0.35	NPLOC4:TSPAN10	-0.08	0.00	1.71E-103
96	162	18	43536260	rs9964220	C	T	0.84	EPG5	-0.03	0.01	3.15E-10
80	163	19	3537184	rs77733715	A	G	0.99	FZR1	0.13	0.02	3.31E-11
119	164	19	7574918	rs2161568	G	A	0.91	CTD-2207O23.12	-0.04	0.01	5.61E-09
133	165	20	22305785	rs6113632	A	T	0.69	RP5-1004I9.2	-0.02	0.00	3.25E-08
3	166	20	32665748	rs6059655	A	G	0.10	RALY	-0.51	0.01	0
3	167	20	32752091	rs138420028	C	T	0.93	ASIP	-0.08	0.01	3.78E-29
3	168	20	32783827	rs6087561	T	C	0.91	ASIP	-0.11	0.01	3.47E-72
3	169	20	32944549	rs13042880	C	T	0.32	NA	-0.07	0.00	1.69E-45
17	170	20	33836110	rs6088778	T	C	0.89	NA	-0.07	0.01	9.56E-26
17	171	20	33865588	rs6060343	G	C	0.87	NA	-0.10	0.01	2.36E-54
12	172	20	34960201	rs4519594	T	C	0.24	DLGAP4	-0.10	0.00	1.35E-111
12	173	20	35298668	rs221314	G	A	0.60	NDRG3	-0.04	0.00	1.70E-28
50	174	20	52659595	rs73132906	G	C	0.95	BCAS1	0.07	0.01	2.38E-16
124	175	20	55409093	rs6127868	G	A	0.85	RNU6-929P	0.03	0.00	1.42E-08
44	176	20	57843358	rs6123874	G	A	0.70	ZNF831	-0.03	0.00	9.04E-18

122	177	21	43429646	rs399907	G	A	0.33	ZBTB21:ZNF295-AS1	-0.02	0.00	8.77E-09
31	178	21	44793448	rs672948	A	T	0.40	AP001046.6	-0.04	0.00	2.13E-31
121	179	22	38545942	rs132941	T	C	0.55	PLA2G6	-0.02	0.00	6.68E-09
101	180	22	39662703	rs11089938	C	T	0.41	AL031590.1	0.02	0.00	6.97E-10
29	181	22	45622684	rs5766565	A	G	0.66	KIAA0930	-0.05	0.00	2.57E-35

Table D.1: Independent signals for pigscore 1 GWAS

D.2 Pigscore 2

Locus	Signal	CHR	BP	SNP	EA	NEA	EAF	Nearest Gene	BETA	SE	P
24	1	1	8207579	rs80293268	G	C	0.95	RP11-431K24.1	0.04	0.00	1.18E-27
109	2	1	23425939	rs489644	A	C	0.17	LUZP1	-0.01	0.00	1.72E-08
38	3	1	25250830	rs10794666	C	T	0.44	RUNX3	-0.01	0.00	1.32E-17
86	4	1	42110888	rs6696511	T	C	0.37	HIVEP3	-0.01	0.00	2.42E-09
81	5	1	46897560	rs78378261	C	T	0.98	RP5-1109J22.1	-0.03	0.01	1.23E-09
33	6	1	63727542	rs670318	T	C	0.05	LINC00466	0.04	0.00	5.21E-20
14	7	1	66888542	rs1308048	T	C	0.57	PDE4B	0.03	0.00	1.47E-66
40	8	1	110720400	rs6689641	A	G	0.47	SLC6A17:RP5-1028L10.2	-0.01	0.00	3.29E-17
61	9	1	120506786	rs2493411	T	C	0.87	NOTCH2	0.02	0.00	5.09E-11
22	10	1	150449701	rs78778914	T	C	0.95	RPRD2	-0.03	0.00	3.48E-12
22	11	1	150798577	rs10305720	C	G	0.63	ARNT	0.02	0.00	1.73E-28
71	12	1	154994978	rs76798800	G	T	0.74	DCST2	-0.01	0.00	6.15E-10
71	13	1	155205331	rs1800473	T	C	0.70	GBA	0.01	0.00	1.56E-10
116	14	1	166159622	rs1432119	G	A	0.26	FAM78B	0.01	0.00	2.72E-08
72	15	1	192513661	rs1853433	T	C	0.21	RP5-1011O1.2	-0.01	0.00	2.08E-10
54	16	1	205138321	rs35845538	G	C	0.95	DSTYK	-0.03	0.00	2.34E-12
89	17	1	219931448	rs4846586	A	G	0.22	SLC30A10	-0.01	0.00	3.01E-09
112	18	1	226654252	rs2099380	G	A	0.88	RN7SKP165	0.01	0.00	2.17E-08

58	19	2	25393254	rs3754860	C	T	0.69	POMC	-0.01	0.00	9.43E-12
27	20	2	28613302	rs71443018	G	C	0.94	AC104695.3	0.04	0.00	3.23E-25
99	21	2	33004062	rs7578659	C	A	0.59	TTC27	0.01	0.00	7.82E-09
20	22	2	38311111	rs11684710	C	T	0.45	CYP1B1:CYP1B1-AS1	0.02	0.00	1.12E-35
20	23	2	38404386	rs400681	G	A	0.52	CYP1B1-AS1	0.01	0.00	2.12E-09
87	24	2	42156497	rs11691937	A	G	0.39	AC104654.2	0.01	0.00	2.61E-09
36	25	2	172380873	rs6730957	C	A	0.63	CYBRD1	0.01	0.00	7.90E-18
31	26	2	222041066	rs7603664	T	C	0.42	AC011233.2	-0.02	0.00	1.49E-20
31	27	2	222078384	rs10498102	C	A	0.56	AC011233.2	0.01	0.00	1.36E-08
88	28	2	223102757	rs1367409	G	T	0.55	PAX3	-0.01	0.00	2.61E-09
77	29	2	239654296	rs7604616	T	C	0.58	AC113618.1	0.01	0.00	7.87E-10
32	30	3	85646525	rs1947221	T	A	0.38	CADM2	-0.02	0.00	4.96E-20
85	31	3	98613773	rs2439222	C	T	0.63	DCBLD2	0.01	0.00	2.32E-09
102	32	3	122521576	rs6438766	G	A	0.34	DIRC2	0.01	0.00	9.35E-09
117	33	3	141094769	rs4683605	C	A	0.55	ZBTB38	0.01	0.00	3.45E-08
29	34	3	156491160	rs9867857	C	T	0.52	LINC00886	-0.02	0.00	4.68E-24
73	35	4	79282759	rs1385131	G	A	0.34	FRAS1	0.01	0.00	2.10E-10
30	36	4	81661426	rs72661730	G	A	0.78	C4orf22	0.02	0.00	6.12E-22
4	37	5	33951693	rs16891982	C	G	0.07	SLC45A2	-0.25	0.01	0
97	38	5	53116123	rs17248377	G	A	0.77	CTD-2081C10.1	0.01	0.00	5.67E-09

69	39	5	57046967	rs10054371	T	C	0.22	RP11-772C9.1	0.01	0.00	1.30E-10
43	40	5	59021562	rs7702956	C	T	0.67	PDE4D	-0.01	0.00	1.36E-14
16	41	5	149195603	rs251466	C	G	0.75	PPARGC1B	0.03	0.00	1.75E-54
23	42	5	173973627	rs6883391	C	T	0.73	RP11-267A15.1	0.02	0.00	3.10E-28
98	43	5	176731452	rs28362590	G	T	0.25	MXD3:PRELID1	0.01	0.00	7.82E-09
1	44	6	192288	rs61376093	T	C	0.95	NA	0.07	0.00	3.94E-50
1	45	6	375759	rs7753579	T	C	0.28	IRF4	-0.10	0.00	0
1	46	6	396321	rs12203592	C	T	0.79	IRF4	-0.35	0.00	0
1	47	6	531401	rs9378914	G	T	0.92	NA	0.07	0.00	1.34E-86
110	48	6	1137160	rs240685	A	C	0.53	AL033381.1	-0.01	0.00	1.77E-08
96	49	6	29935806	rs2517674	G	T	0.77	NA	-0.01	0.00	5.54E-09
104	50	6	41915351	rs6933697	A	G	0.73	CCND3	0.01	0.00	1.40E-08
114	51	6	146386018	rs1546966	G	T	0.48	GRM1	0.01	0.00	2.40E-08
114	52	6	146866242	rs77568325	A	G	0.97	RAB32	0.03	0.01	3.01E-08
113	53	6	148570402	rs34286635	A	G	0.77	RP11-631F7.1	0.01	0.00	2.24E-08
48	54	6	159231098	rs7754951	T	G	0.49	EZR	0.01	0.00	1.44E-13
39	55	6	164516218	rs62435865	T	C	0.89	RP1-155D22.2	-0.02	0.00	3.05E-17
10	56	7	16990804	rs1721040	T	A	0.59	AC098592.7	-0.02	0.00	8.82E-27
10	57	7	17041434	rs34585474	C	T	0.89	AC098592.8	-0.03	0.00	4.92E-23
10	58	7	17134708	rs117132860	G	A	0.97	AC098592.8	-0.11	0.01	2.09E-102

18	59	7	28189411	rs1635852	T	C	0.50	JAZF1	-0.02	0.00	2.91E-44
82	60	7	46385815	rs6966398	T	G	0.91	AC092657.2	-0.02	0.00	1.82E-09
118	61	7	105998889	rs17152960	T	C	0.66	CTB-111H14.1	-0.01	0.00	4.29E-08
46	62	7	130758576	rs12669378	C	T	0.29	LINC-PINT	0.01	0.00	4.13E-14
52	63	8	82748928	rs12674770	T	C	0.74	SNX16	0.01	0.00	7.92E-13
93	64	8	109115122	rs429884	T	C	0.54	RSPO2	-0.01	0.00	4.79E-09
13	65	8	116614543	rs2737207	C	A	0.43	TRPS1	0.03	0.00	2.96E-89
47	66	9	205964	rs478882	A	G	0.50	C9orf66	-0.01	0.00	9.63E-14
45	67	9	12655186	rs1925238	C	G	0.27	TYRP1	-0.01	0.00	6.73E-14
45	68	9	12656787	rs118087215	A	G	0.98	TYRP1	0.04	0.01	2.96E-14
7	69	9	16799109	rs10810636	A	G	0.27	BNC2	-0.03	0.00	9.80E-68
7	70	9	16885017	rs12350739	G	A	0.42	BNC2	-0.05	0.00	5.50E-166
26	71	9	126770250	rs10818907	C	G	0.55	LHX2:RP11-85O21.2	0.02	0.00	1.21E-26
26	72	9	126807459	rs12555707	C	T	0.82	LHX2	0.01	0.00	1.07E-10
50	73	10	5808086	rs1129614	G	A	0.79	GDI2	0.01	0.00	3.95E-13
53	74	10	64564892	rs2236295	G	T	0.61	RP11-436D10.3:ADO	0.01	0.00	2.10E-12
115	75	10	82203069	rs10788627	T	C	0.52	RP11-137H2.6	-0.01	0.00	2.54E-08
107	76	10	102102612	rs7086846	G	A	0.63	RP11-34D15.2	0.01	0.00	1.69E-08
11	77	10	119572403	rs35563099	C	T	0.84	RP11-355F22.1	0.05	0.00	4.50E-101
83	78	11	16163426	rs10832552	T	C	0.56	SOX6	-0.01	0.00	1.91E-09

84	79	11	58983401	rs1938594	G	A	0.71	RN7SL42P	0.01	0.00	2.26E-09
94	80	11	61120076	rs7940665	C	T	0.02	DAK:CYB561A3	-0.13	0.02	5.30E-09
78	81	11	87597790	rs592128	T	C	0.46	RP11-665E10.3	0.01	0.00	9.88E-10
8	82	11	88484644	rs11021172	C	G	0.30	GRM5	-0.05	0.00	3.69E-149
2	83	11	89005267	rs183139540	G	A	0.99	TYR	0.09	0.01	1.17E-35
2	84	11	89017961	rs1126809	G	A	0.71	TYR	-0.09	0.00	0
2	85	11	89323846	rs200522762	T	G	0.86	NOX4:RP11-643G5.6	-0.03	0.00	2.23E-16
59	86	12	41123651	rs10879095	G	A	0.36	CNTN1:RP11-367O10.1	0.01	0.00	9.76E-12
55	87	12	65102065	rs1147100	A	G	0.39	GNS	0.01	0.00	2.56E-12
44	88	12	88143958	rs73200863	G	C	0.91	CYCSP30	0.02	0.00	2.51E-14
9	89	12	88953659	rs4590952	A	G	0.22	KITLG	0.05	0.00	4.49E-122
9	90	12	89328335	rs12821256	T	C	0.89	RP11-13A1.1	0.06	0.00	3.55E-102
28	91	12	89560578	rs189785589	A	G	0.99	RP11-13A1.3	-0.08	0.01	1.77E-24
75	92	12	90245378	rs2553101	T	C	0.51	RP11-654D12.2	0.01	0.00	5.86E-10
51	93	12	116512277	rs61939687	T	G	0.89	MED13L	-0.02	0.00	4.88E-13
37	94	12	125387877	rs68099344	A	T	0.83	RNU6-927P	-0.02	0.00	1.07E-17
101	95	12	129306308	rs4760521	G	A	0.64	SLC15A4	-0.01	0.00	9.02E-09
19	96	13	78377834	rs765377	T	C	0.41	EDNRB-AS1	0.02	0.00	4.99E-37
15	97	13	113533653	rs1765762	G	T	0.45	ATP11A	-0.03	0.00	3.36E-65
67	98	14	55996992	rs10134019	T	G	0.43	KTN1-AS1	0.01	0.00	1.03E-10

12	99	14	92775967	rs746586	C	T	0.56	SLC24A4	0.03	0.00	3.13E-90
12	100	14	92815008	rs76519749	T	C	0.89	SLC24A4	0.03	0.00	2.44E-25
3	101	15	28062125	rs78114576	G	A	0.94	OCA2	-0.03	0.00	1.51E-20
3	102	15	28230318	rs1800407	C	T	0.92	OCA2	-0.13	0.00	0
3	103	15	28230378	rs117886461	G	A	0.98	OCA2	-0.11	0.01	1.72E-54
3	104	15	28250156	rs4778136	T	C	0.90	OCA2	0.05	0.00	1.96E-76
3	105	15	28267027	rs2122005	G	A	0.83	OCA2	-0.03	0.00	3.53E-26
21	106	15	31395346	rs28441327	C	T	0.81	TRPM1	0.03	0.00	1.98E-31
106	107	15	49749735	rs11639111	C	T	0.60	FAM227B:FGF7	-0.01	0.00	1.48E-08
90	108	15	50320678	rs11854435	T	C	0.58	ATP8B4	0.01	0.00	4.18E-09
66	109	15	82427605	rs2455717	A	T	0.56	EFTUD1	0.01	0.00	8.02E-11
57	110	15	85535071	rs8028550	T	C	0.40	PDE8A	-0.01	0.00	4.20E-12
92	111	15	91116386	rs12901150	T	G	0.25	CRTC3	0.01	0.00	4.57E-09
70	112	16	1498197	rs2235578	A	G	0.60	CLCN7	-0.01	0.00	1.41E-10
62	113	16	14008797	rs7186212	T	C	0.64	ERCC4	0.01	0.00	5.25E-11
111	114	16	71634811	rs4788815	A	T	0.34	RP11-432I5.1	-0.01	0.00	2.02E-08
68	115	16	73144902	rs6499612	T	C	0.39	C16orf47	-0.01	0.00	1.18E-10
105	116	16	81603771	rs8044524	G	A	0.40	CMIP	-0.01	0.00	1.48E-08
5	117	16	89645437	rs369230	G	T	0.31	CPNE7	0.03	0.00	1.06E-66
5	118	16	89983276	rs11641201	T	C	0.88	RP11-566K11.7:MC1R	-0.10	0.00	0

5	119	16	89985844	rs1805005	G	T	0.88	MC1R:AC092143.1:TUBB3	-0.05	0.00	3.47E-100
5	120	16	89985918	rs1805006	C	A	0.99	MC1R:AC092143.1:TUBB3	-0.11	0.01	1.15E-50
5	121	16	89986546	rs1805009	G	C	0.98	MC1R:AC092143.1:TUBB3:RP11-566K11.4	-0.05	0.01	3.40E-22
119	122	17	26279598	rs9895999	A	T	0.19	RP11-218F4.1	-0.01	0.00	4.46E-08
80	123	17	38344485	rs2314338	T	C	0.73	RAPGEFL1	0.01	0.00	1.17E-09
60	124	17	39467500	rs72828120	G	A	0.93	KRTAP16-1	-0.02	0.00	2.26E-11
41	125	17	46025052	rs1962201	A	G	0.32	PNPO:RP11-6N17.9:RP11-6N17.6	-0.01	0.00	1.16E-15
91	126	17	48008264	rs79571620	G	C	0.92	RP11-304F15.6	0.02	0.00	4.28E-09
64	127	17	79612397	rs6420484	A	G	0.35	NPLOC4:TSPAN10	-0.01	0.00	6.28E-11
108	128	18	46465665	rs4939832	A	G	0.75	SMAD7	0.01	0.00	1.71E-08
42	129	19	3537184	rs77733715	A	G	0.99	FZR1	-0.05	0.01	3.18E-09
42	130	19	3565092	rs8103610	T	G	0.80	MFSD12	-0.02	0.00	7.10E-15
103	131	19	4394310	rs73234	G	C	0.45	SH3GL1	-0.01	0.00	1.06E-08
79	132	19	10745764	rs2163832	T	C	0.33	SLC44A2	0.01	0.00	1.08E-09
63	133	19	18404936	rs4808779	A	G	0.37	RPL39P38	-0.01	0.00	5.54E-11
65	134	20	3332062	rs6051776	T	C	0.79	C20orf194	-0.01	0.00	6.62E-11
76	135	20	21467208	rs11697152	A	G	0.66	RP5-984P4.1	-0.01	0.00	5.89E-10
6	136	20	32664917	rs200948404	G	A	0.98	RALY	0.05	0.01	2.55E-16
6	137	20	32665748	rs6059655	A	G	0.10	RALY	0.11	0.00	0
34	138	20	33865588	rs6060343	G	C	0.87	NA	0.03	0.00	1.77E-19

25	139	20	34960201	rs4519594	T	C	0.24	DLGAP4	0.02	0.00	5.59E-27
74	140	20	52628933	rs73133429	C	T	0.98	BCAS1	0.04	0.01	4.57E-10
100	141	20	57857023	rs3026622	G	A	0.70	EDN3	-0.01	0.00	8.42E-09
95	142	21	26859304	rs2186299	T	C	0.89	snoU13	-0.02	0.00	5.35E-09
49	143	21	43429646	rs399907	G	A	0.33	ZBTB21:ZNF295-AS1	0.01	0.00	2.73E-13
56	144	22	35708790	rs4465	C	T	0.36	TOM1	-0.01	0.00	2.81E-12
35	145	22	38572440	rs4384	G	C	0.54	PLA2G6	0.01	0.00	5.60E-19
17	146	22	45622014	rs6007506	C	T	0.66	KIAA0930	0.03	0.00	1.45E-49

Table D.2: Independent signals for pigscore 2 GWAS

D.3 Pigscore 3

Locus	Signal	CHR	BP	SNP	EA	NEA	EAF	Nearest Gene	BETA	SE	P
13	1	1	66938005	rs783326	G	A	0.43	SGIP1	-0.01	0.00	5.54E-12
20	2	1	154986091	rs3753639	T	C	0.76	ZBTB7B	-0.01	0.00	4.40E-08
17	3	2	38298877	rs162561	T	G	0.18	CYP1B1	0.01	0.00	3.66E-10
2	4	6	375759	rs7753579	T	C	0.28	IRF4	-0.02	0.00	1.63E-50
2	5	6	396321	rs12203592	C	T	0.79	IRF4	-0.06	0.00	0
7	6	7	17134708	rs117132860	G	A	0.97	AC098592.8	-0.03	0.00	5.05E-27
10	7	8	116621214	rs2737212	C	T	0.44	TRPS1	0.01	0.00	1.37E-20
9	8	9	16799109	rs10810636	A	G	0.27	BNC2	-0.01	0.00	5.89E-13
9	9	9	16885017	rs12350739	G	A	0.42	BNC2	-0.01	0.00	2.00E-22
11	10	10	119573178	rs7098111	C	T	0.84	RP11-355F22.1	0.01	0.00	2.41E-15
19	11	11	47439444	rs2053979	A	G	0.67	PSMC3	-0.01	0.00	7.38E-09
5	12	11	88941950	rs35300336	G	A	0.89	TYR	0.02	0.00	3.76E-37
5	13	11	89017961	rs1126809	G	A	0.71	TYR	-0.02	0.00	2.90E-72
6	14	12	88953659	rs4590952	A	G	0.22	KITLG	0.02	0.00	1.27E-55
6	15	12	89184348	rs2100883	C	T	0.95	RNU1-117P	0.02	0.00	8.22E-17
14	16	13	113543499	rs1278772	C	T	0.46	ATP11A	-0.01	0.00	4.63E-11
15	17	14	92766340	rs61977801	C	A	0.56	SLC24A4	0.01	0.00	4.96E-11
4	18	15	28230378	rs117886461	G	A	0.98	OCA2	-0.04	0.00	5.69E-34

4	19	15	28250156	rs4778136	T	C	0.90	OCA2	0.01	0.00	7.55E-22
4	20	15	28365618	rs12913832	A	G	0.26	HERC2	0.03	0.00	1.47E-169
1	21	16	89708267	rs77651727	C	T	0.92	CHMP1A	-0.05	0.00	1.85E-258
1	22	16	89985918	rs1805006	C	A	0.99	MC1R:AC092143.1:TUBB3	-0.07	0.00	2.94E-81
1	23	16	89986117	rs1805007	C	T	0.90	MC1R:AC092143.1:TUBB3	-0.09	0.00	0
1	24	16	89986546	rs1805009	G	C	0.98	MC1R:AC092143.1:TUBB3:RP11-566K11.4	-0.08	0.00	1.63E-204
16	25	17	44365098	rs77471755	G	A	0.90	NA	-0.01	0.00	1.25E-10
3	26	20	32665748	rs6059655	A	G	0.10	RALY	0.05	0.00	1.25E-278
12	27	20	33865588	rs6060343	G	C	0.87	NA	0.01	0.00	7.39E-15
8	28	20	34960201	rs4519594	T	C	0.24	DLGAP4	0.01	0.00	1.86E-26
18	29	22	45631518	rs34980158	C	T	0.53	KIAA0930	0.01	0.00	1.02E-09

Table D.3: Independent signals for pigscore 3 GWAS

Appendix E

Overlap in pigmentation GWAS

E.1 Overlap table

Locus	Signal	SNP	GWAS Association	Pigmentary GWAS Catalog Association	Pigmentary GWAS Catalog Association SNP	Melanoma GWAS Catalog Association	Melanoma GWAS Catalog Association SNP
1	1	rs78555129	Non-red	Hair colour, Blond vs. brown/black hair colour	rs79361800, rs75972122, rs78555129		
2	3	rs4908473	Skin, Non-red	Blond vs. brown/black hair colour, Hair colour, Brown vs. black hair colour	rs11582820, rs76648881, rs80293268, rs147458259, rs77905678, rs72635790		
2	2	rs112147823	Skin, Non-red	Blond vs. brown/black hair colour, Hair colour, Brown vs. black hair colour	rs11582820, rs76648881, rs80293268, rs147458259, rs77905678, rs72635790		
2	3	rs80293268	Skin, Non-red	Blond vs. brown/black hair colour, Hair colour, Brown vs. black hair colour	rs11582820, rs76648881, rs80293268, rs147458259, rs77905678, rs72635790		

2	4	rs74865018	Skin, Non-red	Blond vs. brown/black hair colour, Hair colour, Brown vs. black hair colour	rs11582820, rs76648881, rs80293268, rs147458259, rs77905678, rs72635790
3	6	rs12375	Non-red	Blond vs. brown/black hair colour, Hair colour	rs6687430, rs12090714, rs6685766, rs112115136, rs6689838
3	5	rs12081181	Non-red	Blond vs. brown/black hair colour, Hair colour	rs6687430, rs12090714, rs6685766, rs112115136, rs6689838
3	6	rs56168930	Non-red	Blond vs. brown/black hair colour, Hair colour	rs6687430, rs12090714, rs6685766, rs112115136, rs6689838
4	7	rs848200	Non-red	Blond vs. brown/black hair colour, Hair colour	rs12738340, rs848188
5	8	rs11585357	Non-red	Blond vs. brown/black hair colour, Hair colour	rs144080386, rs11203346, rs2977303, rs72646785

5	8	rs11584287	Non-red	Blond vs. brown/black hair colour, Hair colour	rs144080386, rs11203346, rs2977303, rs72646785
6	9	rs489644	Pigscore 2		
7	11	rs195700	Ease, Skin, Non-red	Hair colour	rs195720
7	12	rs10794666	Ease, Skin, Non-red	Hair colour	rs195720
7	11	rs1507101	Ease, Skin, Non-red	Hair colour	rs195720
7	10	rs3845302	Ease, Skin, Non-red	Hair colour	rs195720
7	11	rs12077712	Ease, Skin, Non-red	Hair colour	rs195720
8	13	rs79598313	Non-red	Blond vs. brown/black hair colour	rs112535818
9	14	rs1937999	Ease, Skin, Non-red	Hair colour	rs1937999, rs3856254
9	15	rs6696511	Ease, Skin, Non-red	Hair colour	rs1937999, rs3856254

9	15	rs6702659	Ease, Skin, Non-red	Hair colour	rs1937999, rs3856254		
9	15	rs4393146	Ease, Skin, Non-red	Hair colour	rs1937999, rs3856254		
10	16	rs78378261	Pigscore 2				
11	17	rs12568356	Non-red				
12	18	rs17377295	Non-red	Hair colour, Blond vs. brown/black hair colour	rs76732304, rs17377218, rs17377232		
13	19	rs670318	Ease, Skin			Melanoma paper	rs670318
14	20	rs1308048	Ease, Skin	Low tan response	rs1308048		
14	20	rs1613999	Ease, Skin	Low tan response	rs1308048		
14	20	rs783326	Ease, Skin	Low tan response	rs1308048		
15	21	rs34517439	Non-red	Hair colour	rs34517439		
16	22	rs12034421	Non-red	Hair colour, Blond vs. brown/black hair colour	rs12034421		
17	23	rs236285	Non-red	Hair colour	rs1011903		
18	24	rs12722976	Non-red				
19	25	rs6689641	Ease, Skin				

19	25	rs535930	Ease, Skin		
20	26	rs2493411	Pigscore 2		
21	28	rs78778914	Ease, Skin	Nevus count or cutaneous melanoma, melanoma, Melanoma paper	rs72704658, rs7412746, rs8444, rs1722784
21	28	rs9733	Ease, Skin	Nevus count or cutaneous melanoma, melanoma, Melanoma paper	rs72704658, rs7412746, rs8444, rs1722784

21	27	rs3894771	Ease, Skin			Nevus count or cutaneous melanoma, melanoma, Melanoma paper	rs72704658, rs7412746, rs8444, rs1722784
21	28	rs10305720	Ease, Skin			Nevus count or cutaneous melanoma, melanoma, Melanoma paper	rs72704658, rs7412746, rs8444, rs1722784
22	29	rs61816761	CS				
23	30	rs1410860	Non-red	Hair colour	rs1329101		
24	32	rs3753639	Ease, Non-red	Hair colour	rs4390169	Melanoma paper	rs76798800

24	31	rs76798800	Ease, Non-red	Hair colour	rs4390169	Melanoma paper	rs76798800
24	32	rs4745	Ease, Non-red	Hair colour	rs4390169	Melanoma paper	rs76798800
24	31	rs1800473	Ease, Non-red	Hair colour	rs4390169	Melanoma paper	rs76798800
25	33	rs12043212	CS	Hair colour	rs12134456		
26	34	rs1432119	Pigscore 2				
27	35	rs4656781	Skin				
28	36	rs1853433	Non-red	Hair colour	rs1323292, rs3011685		
28	36	rs2760524	Non-red	Hair colour	rs1323292, rs3011685		
29	38	rs3851294	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour	rs3851294, rs2369633	Melanoma paper	rs2369633
29	37	rs35845538	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour	rs3851294, rs2369633	Melanoma paper	rs2369633
30	39	rs28826269	Non-red	Blond vs. brown/black hair colour, Hair colour	rs1338356, rs3125847, rs6698338, rs1338349		
31	40	rs12727947	Ease	Hair colour	rs351376		
31	40	rs351376	Ease	Hair colour	rs351376		

32	41	rs113614987	Non-red	Hair colour	rs113614987		
32	42	rs7533482	Non-red	Hair colour	rs113614987		
33	43	rs4846586	Pigscore 2				
34	44	rs2099380	Pigscore 2			Melanoma, Nevus count or cutaneous melanoma, Melanoma paper	rs3219090, rs2695237
35	45	rs59116340	Non-red	Hair colour, Blond vs. brown/black hair colour	rs112779011, rs11806180		
36	46	rs16857370	Non-red	Hair colour	rs4649269, rs12749578		
37	48	rs4660121	Non-red	Hair colour	rs10803124		
37	47	rs12129097	Non-red	Hair colour	rs10803124		
38	49	rs1947536	Non-red	Hair colour, Blond vs. brown/black hair colour	rs9661140, rs7550088		
39	50	rs12714332	Non-red	Hair colour	rs12714332		
40	51	rs72785456	Non-red				

41	53	rs13428823	Ease, Red	Hair colour	rs4665773, rs28562485	Melanoma paper	rs12473635
41	52	rs3754860	Ease, Red	Hair colour	rs4665773, rs28562485	Melanoma paper	rs12473635
41	52	rs76645364	Ease, Red	Hair colour	rs4665773, rs28562485	Melanoma paper	rs12473635
41	52	rs13036246	Ease, Red	Hair colour	rs4665773, rs28562485	Melanoma paper	rs12473635
42	54	rs112272576	Pigscore 1				
43	55	rs1260326	Ease, Skin, Non-red	Hair colour	rs4665978		
43	55	rs780094	Ease, Skin, Non-red	Hair colour	rs4665978		
43	55	rs780093	Ease, Skin, Non-red	Hair colour	rs4665978		
44	57	rs71443018	Skin, Non-red	Blond vs. brown/black hair colour, Hair colour	rs4665412, rs71443018, rs11680860, rs62139588		
44	56	rs55870117	Skin, Non-red	Blond vs. brown/black hair colour, Hair colour	rs4665412, rs71443018, rs11680860, rs62139588		

45	58	rs7584123	Skin				
45	58	rs7578659	Skin				
46	59	rs10169939	Ease, Skin	Low tan response	rs2855655	Cutaneous malignt melanoma, melanoma, Nevus count or cutaneous melanoma, Melanoma paper	rs6750047, rs4670813, rs1800440

46	59	rs1800440	Ease, Skin	Low tan response	rs2855655	Cutaneous malignt melanoma, melanoma, Nevus count or cutaneous melanoma, Melanoma paper	rs6750047, rs4670813, rs1800440
46	59	rs162561	Ease, Skin	Low tan response	rs2855655	Cutaneous malignt melanoma, melanoma, Nevus count or cutaneous melanoma, Melanoma paper	rs6750047, rs4670813, rs1800440

46	59	rs11684710	Ease, Skin	Low tan response	rs2855655	Cutaneous malignt melanoma, melanoma, Nevus count or cutaneous melanoma, Melanoma paper	rs6750047, rs4670813, rs1800440
46	60	rs400681	Ease, Skin	Low tan response	rs2855655	Cutaneous malignt melanoma, melanoma, Nevus count or cutaneous melanoma, Melanoma paper	rs6750047, rs4670813, rs1800440

47	61	rs1962386	Non-red	Hair colour, Blond vs. brown/black hair colour	rs2888814, rs13419021, rs4952542
47	61	rs4952542	Non-red	Hair colour, Blond vs. brown/black hair colour	rs2888814, rs13419021, rs4952542
47	61	rs11691937	Non-red	Hair colour, Blond vs. brown/black hair colour	rs2888814, rs13419021, rs4952542
48	63	rs12612692	Ease, Skin, Non-red	Hair colour	rs6544650
48	62	rs10185673	Ease, Skin, Non-red	Hair colour	rs6544650
48	64	rs57696714	Ease, Skin, Non-red	Hair colour	rs6544650
49	65	rs2706764	Non-red	Hair colour	rs2169838, rs13028186
50	66	rs6730279	Non-red	Blond vs. brown/black hair colour, Hair colour	rs6707137, rs56165474, rs73952210
50	66	rs7608166	Non-red	Blond vs. brown/black hair colour, Hair colour	rs6707137, rs56165474, rs73952210
51	67	rs3749110	Non-red		
52	68	rs6716872	Non-red		

53	69	rs114153232	Non-red		
54	70	rs13410301	Non-red	Hair colour	rs6728095
54	70	rs6724774	Non-red	Hair colour	rs6728095
55	71	rs3731976	Ease, Skin		
55	71	rs6730957	Ease, Skin		
55	71	rs35380972	Ease, Skin		
56	72	rs7591631	Skin	Red vs. brown/black hair colour	rs1205151
57	73	rs71421546	Non-red	Hair colour, Blond vs. brown/black hair colour	rs12693099
58	74	rs13431878	Non-red		
59	75	rs13030978	Non-red		
60	76	rs12693954	Non-red	Hair colour	rs726357
61	77	rs7583880	Non-red		
62	78	rs56020963	Non-red	Hair colour	rs78992409
63	80	rs1541966	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour	rs11900971, rs10169459, rs116254882, rs13017777, rs12618431, rs12618491, rs2303948

63	81	rs62186153	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour	rs11900971, rs10169459, rs116254882, rs13017777, rs12618431, rs12618491, rs2303948
63	81	rs7603664	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour	rs11900971, rs10169459, rs116254882, rs13017777, rs12618431, rs12618491, rs2303948
63	83	rs10169459	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour	rs11900971, rs10169459, rs116254882, rs13017777, rs12618431, rs12618491, rs2303948
63	79	rs10192020	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour	rs11900971, rs10169459, rs116254882, rs13017777, rs12618431, rs12618491, rs2303948

63	81	rs10498102	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour	rs11900971, rs10169459, rs116254882, rs13017777, rs12618431, rs12618491, rs2303948
63	80	rs17349283	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour	rs11900971, rs10169459, rs116254882, rs13017777, rs12618431, rs12618491, rs2303948
63	82	rs7570475	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour	rs11900971, rs10169459, rs116254882, rs13017777, rs12618431, rs12618491, rs2303948
63	80	rs13029372	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour	rs11900971, rs10169459, rs116254882, rs13017777, rs12618431, rs12618491, rs2303948
64	85	rs744174	Non-red		
64	86	rs116254882	Non-red		
64	84	rs1367409	Non-red		

64	86	rs12618431	Non-red					
64	87	rs1430663	Non-red					
64	84	rs2303948	Non-red					
65	88	rs812383	Non-red	Hair greying	rs213546			
66	89	rs887829	Skin					
67	91	rs7604616	Non-red	Hair colour	rs11684254		Nevus count or cutaneous melanoma	rs55875066
67	90	rs9287636	Non-red	Hair colour	rs11684254		Nevus count or cutaneous melanoma	rs55875066
67	90	rs4075018	Non-red	Hair colour	rs11684254		Nevus count or cutaneous melanoma	rs55875066
68	92	rs9809528	Non-red	Blond vs. brown/black hair colour, Hair colour	rs9809528, rs13082190	rs4685448,		

69	93	rs2437689	Non-red	Hair colour, Blond vs. brown/black hair colour	rs2437689, rs2443723, rs67001479, rs2574715		
70	94	rs754821	Non-red	Hair colour	rs754821		
71	95	rs2014520	Non-red	Blond vs. brown/black hair colour, Hair colour	rs9825958, rs11716109, rs9823839	Melanoma paper	rs149617956
71	95	rs17006579	Non-red	Blond vs. brown/black hair colour, Hair colour	rs9825958, rs11716109, rs9823839	Melanoma paper	rs149617956
72	96	rs9809116	Non-red	Hair colour	rs9809116, rs9821691		
73	97	rs586936	Non-red	Hair colour	rs586936		
74	98	rs11921010	Ease, Skin				
74	98	rs7617323	Ease, Skin				
74	98	rs1947221	Ease, Skin				
75	99	rs2439222	Ease				
75	99	rs9821965	Ease				
76	100	rs6438766	Non-red	Hair colour, Blond vs. brown/black hair colour	rs9847240		
76	100	rs9847240	Non-red	Hair colour, Blond vs. brown/black hair colour	rs9847240		
77	101	rs115182912	Non-red	Hair colour	rs115182912		

78	102	rs2715610	Non-red		
79	103	rs2293252	Non-red	Hair colour, Blond vs. brown/black hair colour	rs2054468, rs6782181
80	106	rs6440003	Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour	rs6440008, rs4683605, rs1344672, rs35225290, rs79228062, rs2288634, rs3804772
80	104	rs4683605	Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour	rs6440008, rs4683605, rs1344672, rs35225290, rs79228062, rs2288634, rs3804772
80	106	rs6440008	Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour	rs6440008, rs4683605, rs1344672, rs35225290, rs79228062, rs2288634, rs3804772
80	104	rs73869619	Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour	rs6440008, rs4683605, rs1344672, rs35225290, rs79228062, rs2288634, rs3804772

80	104	rs4683632	Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour	rs6440008, rs4683605, rs1344672, rs35225290, rs79228062, rs2288634, rs3804772
80	105	rs3804772	Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour	rs6440008, rs4683605, rs1344672, rs35225290, rs79228062, rs2288634, rs3804772
81	107	rs323610	Non-red	Blond vs. brown/black hair colour, Hair colour	rs325712, rs323610
82	108	rs1713843	Non-red	Hair colour	rs453699
83	109	rs9867857	Ease, Skin		
83	109	rs9818780	Ease, Skin		
84	110	rs34128525	Non-red	Hair colour	rs9829436
85	111	rs1345417	Skin, Non-red	Hair colour	rs60733335
85	111	rs833268	Skin, Non-red	Hair colour	rs60733335
86	112	rs62289589	Non-red	Hair colour	rs9820421

87	113	rs111391498	CS		
88	115	rs4689314	Ease, Non-red	Hair colour	rs10015223
88	114	rs10015223	Ease, Non-red	Hair colour	rs10015223
89	116	rs7696175	Pigscore 1		
90	117	rs6554121	Non-red	Hair colour	rs2117599
91	118	rs1107674	Non-red		
92	119	rs17804499	Non-red	Hair colour, Blond vs. brown/black hair colour	rs1874202
93	120	rs1874202	Skin, Non-red	Hair colour	rs1874202
93	121	rs6819541	Skin, Non-red	Hair colour	rs1874202
94	122	rs436034	Non-red		
94	122	rs371273	Non-red		
94	122	rs1385131	Non-red		
95	123	rs1458046	Non-red	Hair colour, Blond vs. brown/black hair colour	rs72661730, rs62302224, rs7681907

95	124	rs72661730	Non-red	Hair colour, Blond vs. brown/black hair colour	rs72661730, rs62302224, rs7681907
96	125	rs17392334	Non-red	Hair colour	rs1026873, rs72656294
96	126	rs28483422	Non-red	Hair colour	rs1026873, rs72656294
97	128	rs9998015	Non-red	Hair colour	rs142402773, rs9998015
97	127	rs17429682	Non-red	Hair colour	rs142402773, rs9998015
98	133	rs7673917	Non-red	Blond vs. brown/black hair colour, Hair colour	rs11731416, rs1436501, rs1436502, rs7672648
98	130	rs17038688	Non-red	Blond vs. brown/black hair colour, Hair colour	rs11731416, rs1436501, rs1436502, rs7672648
98	132	rs999870	Non-red	Blond vs. brown/black hair colour, Hair colour	rs11731416, rs1436501, rs1436502, rs7672648
98	131	rs220625	Non-red	Blond vs. brown/black hair colour, Hair colour	rs11731416, rs1436501, rs1436502, rs7672648
98	129	rs219493	Non-red	Blond vs. brown/black hair colour, Hair colour	rs11731416, rs1436501, rs1436502, rs7672648
98	131	rs11731416	Non-red	Blond vs. brown/black hair colour, Hair colour	rs11731416, rs1436501, rs1436502, rs7672648
99	134	rs1584590	Non-red		

100	135	rs1996020	Non-red					
101	136	rs72719803	Non-red	Hair colour	rs72737816			
102	137	rs4696396	Non-red					
103	138	rs1651282	Non-red					
104	140	rs4866399	Ease, Skin, Non-red, CS	Hair colour, Hair morphology traits, Black vs. non-black hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Skin pigmentation traits, Skin pigmentation, Low tan response, Tanning, Skin, hair and eye pigmentation (multivariate alysis), Black vs. blond hair colour, Black vs. red hair colour, Skin colour saturation, Perceived skin darkness	rs111217601, rs16891982, rs35391, rs116887602, rs201259497, rs35412, rs13289	rs993459, rs185146, rs28777, rs183671,	Melanoma, Melanoma paper	rs35407, rs16891982, rs35390

104	143	rs116125333	Ease, Skin, Non-red, CS	Hair colour, Hair morphology traits, Black vs. non-black hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Skin pigmentation traits, Skin pigmentation, Low tan response, Tanning, Skin, hair and eye pigmentation (multivariate alysis), Black vs. blond hair colour, Black vs. red hair colour, Skin colour saturation, Perceived skin darkness	rs111217601, rs16891982, rs35391, rs116887602, rs201259497, rs35412, rs13289	rs993459, rs185146, rs28777, rs183671,	Melanoma, Melanoma paper	rs35407, rs16891982, rs35390

104	142	rs79014396	Ease, Skin, Non-red, CS	Hair colour, Hair morphology traits, Black vs. non-black hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Skin pigmentation traits, Skin pigmentation, Low tan response, Tanning, Skin, hair and eye pigmentation (multivariate alysis), Black vs. blond hair colour, Black vs. red hair colour, Skin colour saturation, Perceived skin darkness	rs111217601, rs16891982, rs35391, rs116887602, rs201259497, rs35412, rs13289	rs993459, rs185146, rs28777, rs183671,	Melanoma, Melanoma paper	rs35407, rs16891982, rs35390

104	139	rs16891982	Ease, Skin, Non-red, CS	Hair colour, Hair morphology traits, Black vs. non-black hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Skin pigmentation traits, Skin pigmentation, Low tan response, Tanning, Skin, hair and eye pigmentation (multivariate alysis), Black vs. blond hair colour, Black vs. red hair colour, Skin colour saturation, Perceived skin darkness	rs111217601, rs16891982, rs35391, rs116887602, rs201259497, rs35412, rs13289	rs993459, rs185146, rs28777, rs183671,	Melanoma, Melanoma paper	rs35407, rs16891982, rs35390

104	141	rs13289	Ease, Skin, Non-red, CS	Hair colour, Hair morphology traits, Black vs. non-black hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Skin pigmentation traits, Skin pigmentation, Low tan response, Tanning, Skin, hair and eye pigmentation (multivariate alysis), Black vs. blond hair colour, Black vs. red hair colour, Skin colour saturation, Perceived skin darkness	rs111217601, rs993459, rs16891982, rs185146, rs35391, rs28777, rs116887602, rs201259497, rs183671, rs35412, rs13289	Melanoma, Melanoma paper	rs35407, rs16891982, rs35390
105	144	rs6881866	Non-red	Hair colour	rs6893700		
106	145	rs427527	Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour	rs1504212, rs271205, rs73754267, rs62370277, rs148906333, rs6875907, rs17248377, rs138015091		

106	147	rs77325285	Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour	rs1504212, rs271205, rs73754267, rs62370277, rs148906333, rs6875907, rs17248377, rs138015091
106	145	rs6875907	Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour	rs1504212, rs271205, rs73754267, rs62370277, rs148906333, rs6875907, rs17248377, rs138015091
106	146	rs17248377	Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour	rs1504212, rs271205, rs73754267, rs62370277, rs148906333, rs6875907, rs17248377, rs138015091
106	145	rs3846492	Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour	rs1504212, rs271205, rs73754267, rs62370277, rs148906333, rs6875907, rs17248377, rs138015091

107	148	rs7714232	Non-red		
108	149	rs10054371	Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour	rs34196215, rs7714232, rs61055995, rs1835873
108	149	rs1835873	Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour	rs34196215, rs7714232, rs61055995, rs1835873
109	150	rs7720119	Ease, Skin	Hair colour	rs7700279
109	151	rs7702956	Ease, Skin	Hair colour	rs7700279
109	150	rs11954036	Ease, Skin	Hair colour	rs7700279
109	150	rs62370543	Ease, Skin	Hair colour	rs7700279
110	152	rs74852363	Ease		
111	153	rs2662225	Ease, Skin, Non-red		
111	154	rs469394	Ease, Skin, Non-red		
111	153	rs4282273	Ease, Skin, Non-red		
112	155	rs259035	Non-red	Hair colour	rs259035, rs11953891

113	156	rs11957689	Skin, Non-red	Hair colour	rs1995774	Melanoma paper	rs12523094
113	156	rs6887203	Skin, Non-red	Hair colour	rs1995774	Melanoma paper	rs12523094
113	156	rs60325490	Skin, Non-red	Hair colour	rs1995774	Melanoma paper	rs12523094
114	157	rs163888	Non-red				
115	158	rs9326880	Non-red				
116	159	rs285894	Skin				
117	160	rs10519488	Non-red	Hair colour	rs10519488		
118	161	rs10479082	Skin, Non-red	Hair colour	rs10051152, rs12054866		
118	161	rs329120	Skin, Non-red	Hair colour	rs10051152, rs12054866		
119	162	rs251468	Ease, Skin, CS			Melanoma paper	rs32578
119	162	rs251466	Ease, Skin, CS			Melanoma paper	rs32578

119	162	rs109077	Ease, Skin, CS			Melanoma paper	rs32578
120	163	rs2936938	Ease, Skin, Non-red, Red	Hair colour, Blond vs. brown/black hair colour, Red vs. brown/black hair colour	rs4242182, rs2936938,	rs2964049, rs6876712	
120	166	rs6883391	Ease, Skin, Non-red, Red	Hair colour, Blond vs. brown/black hair colour, Red vs. brown/black hair colour	rs4242182, rs2936938,	rs2964049, rs6876712	
120	165	rs923520	Ease, Skin, Non-red, Red	Hair colour, Blond vs. brown/black hair colour, Red vs. brown/black hair colour	rs4242182, rs2936938,	rs2964049, rs6876712	
120	164	rs60780889	Ease, Skin, Non-red, Red	Hair colour, Blond vs. brown/black hair colour, Red vs. brown/black hair colour	rs4242182, rs2936938,	rs2964049, rs6876712	
120	164	rs4242182	Ease, Skin, Non-red, Red	Hair colour, Blond vs. brown/black hair colour, Red vs. brown/black hair colour	rs4242182, rs2936938,	rs2964049, rs6876712	
121	167	rs28362590	Ease				
122	168	rs17714046	Non-red				

123	170	rs61376093	Ease, Skin, Non-red, Red, CS	Low tan response, Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Hair morphology traits, Hair greying, Skin, hair and eye pigmentation (multivariate alysis), Black vs. blond hair colour, Black vs. red hair colour, Skin pigmentation traits, Freckling, Tanning, Non-melanoma skin cancer, Skin colour saturation, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Freckles	rs2235874, rs7767355, rs9405192, rs2671427, rs74758148, rs12203592, rs3778607, rs12211228, rs9392504, rs4246064, rs62389423, rs62389424, rs143615986, rs1540771, rs12202284, rs12210050, rs974455, rs950039, rs13192740, rs6918152, rs71550013, rs9405909, rs7765892, rs139066356, rs115074082	Melanoma	rs62389423

123	172	rs7753579	Ease, Skin, Non-red, Red, CS	Low tan response, Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Hair morphology traits, Hair greying, Skin, hair and eye pigmentation (multivariate alysis), Black vs. blond hair colour, Black vs. red hair colour, Skin pigmentation traits, Freckling, Tanning, Non-melanoma skin cancer, Skin colour saturation, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Freckles	rs2235874, rs7767355, rs9405192, rs2671427, rs74758148, rs12203592, rs3778607, rs12211228, rs9392504, rs4246064, rs62389423, rs62389424, rs143615986, rs1540771, rs12202284, rs12210050, rs974455, rs950039, rs13192740, rs6918152, rs71550013, rs9405909, rs7765892, rs139066356, rs115074082	Melanoma	rs62389423

123	169	rs71550004	Ease, Skin, Non-red, Red, CS	Low tan response, Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Hair morphology traits, Hair greying, Skin, hair and eye pigmentation (multivariate alysis), Black vs. blond hair colour, Black vs. red hair colour, Skin pigmentation traits, Freckling, Tanning, Non-melanoma skin cancer, Skin colour saturation, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Freckles	rs2235874, rs7767355, rs9405192, rs2671427, rs74758148, rs12203592, rs3778607, rs12211228, rs9392504, rs4246064, rs62389423, rs62389424, rs143615986, rs1540771, rs12202284, rs12210050, rs974455, rs950039, rs13192740, rs6918152, rs71550013, rs9405909, rs7765892, rs139066356, rs115074082	Melanoma	rs62389423

123	171	rs12203592	Ease, Skin, Non-red, Red, CS	Low tan response, Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Hair morphology traits, Hair greying, Skin, hair and eye pigmentation (multivariate alysis), Black vs. blond hair colour, Black vs. red hair colour, Skin pigmentation traits, Freckling, Tanning, Non-melanoma skin cancer, Skin colour saturation, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Freckles	rs2235874, rs7767355, rs9405192, rs2671427, rs74758148, rs12203592, rs3778607, rs12211228, rs9392504, rs4246064, rs62389423, rs62389424, rs143615986, rs1540771, rs12202284, rs12210050, rs974455, rs950039, rs13192740, rs6918152, rs71550013, rs9405909, rs7765892, rs139066356, rs115074082	Melanoma	rs62389423

123	170	rs60729976	Ease, Skin, Non-red, Red, CS	Low tan response, Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Hair morphology traits, Hair greying, Skin, hair and eye pigmentation (multivariate alysis), Black vs. blond hair colour, Black vs. red hair colour, Skin pigmentation traits, Freckling, Tanning, Non-melanoma skin cancer, Skin colour saturation, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Freckles	rs2235874, rs7767355, rs9405192, rs2671427, rs74758148, rs12203592, rs3778607, rs12211228, rs9392504, rs4246064, rs62389423, rs62389424, rs143615986, rs1540771, rs12202284, rs12210050, rs974455, rs950039, rs13192740, rs6918152, rs71550013, rs9405909, rs7765892, rs139066356, rs115074082	Melanoma	rs62389423

123	173	rs9392618	Ease, Skin, Non-red, Red, CS	Low tan response, Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Hair morphology traits, Hair greying, Skin, hair and eye pigmentation (multivariate alysis), Black vs. blond hair colour, Black vs. red hair colour, Skin pigmentation traits, Freckling, Tanning, Non-melanoma skin cancer, Skin colour saturation, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Freckles	rs2235874, rs7767355, rs9405192, rs2671427, rs74758148, rs12203592, rs3778607, rs12211228, rs9392504, rs4246064, rs62389423, rs62389424, rs143615986, rs1540771, rs12202284, rs12210050, rs974455, rs950039, rs13192740, rs6918152, rs71550013, rs9405909, rs7765892, rs139066356, rs115074082	Melanoma	rs62389423

123	173	rs9378896	Ease, Skin, Non-red, Red, CS	Low tan response, Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Hair morphology traits, Hair greying, Skin, hair and eye pigmentation (multivariate alysis), Black vs. blond hair colour, Black vs. red hair colour, Skin pigmentation traits, Freckling, Tanning, Non-melanoma skin cancer, Skin colour saturation, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Freckles	rs2235874, rs7767355, rs9405192, rs2671427, rs74758148, rs12203592, rs3778607, rs12211228, rs9392504, rs4246064, rs62389423, rs62389424, rs143615986, rs1540771, rs12202284, rs12210050, rs974455, rs950039, rs13192740, rs6918152, rs71550013, rs9405909, rs7765892, rs139066356, rs115074082	Melanoma	rs62389423

123	173	rs9378914	Ease, Skin, Non-red, Red, CS	Low tan response, Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Hair morphology traits, Hair greying, Skin, hair and eye pigmentation (multivariate alysis), Black vs. blond hair colour, Black vs. red hair colour, Skin pigmentation traits, Freckling, Tanning, Non-melanoma skin cancer, Skin colour saturation, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Freckles	rs2235874, rs7767355, rs9405192, rs2671427, rs74758148, rs12203592, rs3778607, rs12211228, rs9392504, rs4246064, rs62389423, rs62389424, rs143615986, rs1540771, rs12202284, rs12210050, rs974455, rs950039, rs13192740, rs6918152, rs71550013, rs9405909, rs7765892, rs139066356, rs115074082	Melanoma	rs62389423
124	174	rs240685	Ease	Hair colour, Low tan response	rs9405909, rs7765892, rs139066356, rs115074082, rs2143396, rs238027	Melanoma paper	rs12215602

124	174	rs10214796	Ease	Hair colour, Low tan response	rs9405909, rs7765892, rs139066356, rs115074082, rs2143396, rs238027	Melanoma paper	rs12215602
124	174	rs57354745	Ease	Hair colour, Low tan response	rs9405909, rs7765892, rs139066356, rs115074082, rs2143396, rs238027	Melanoma paper	rs12215602
125	175	rs9405194	Ease	Hair colour	rs7758776		
126	176	rs77571968	Non-red	Hair colour	rs78781079		
126	176	rs594329	Non-red	Hair colour	rs78781079		
127	177	rs9367849	CS				
128	178	rs115949579	Ease, Skin				
129	179	rs2223620	Non-red	Blond vs. brown/black hair colour, Hair colour	rs78287738, rs12208597, rs2223620		
129	179	rs4710940	Non-red	Blond vs. brown/black hair colour, Hair colour	rs78287738, rs12208597, rs2223620		
130	180	rs9368374	Non-red				

131	181	rs13193480	Ease, Skin, CS		
131	182	rs909728	Ease, Skin, CS		
131	181	rs2517674	Ease, Skin, CS		
132	185	rs9394026	Non-red	Hair colour, Blond vs. brown/black hair colour	rs2233981, rs17207524
132	183	rs9266756	Non-red	Hair colour, Blond vs. brown/black hair colour	rs2233981, rs17207524
132	184	rs34451818	Non-red	Hair colour, Blond vs. brown/black hair colour	rs2233981, rs17207524
132	183	rs17207524	Non-red	Hair colour, Blond vs. brown/black hair colour	rs2233981, rs17207524
133	186	rs511515	CS		
134	187	rs6938966	Ease, Skin	Hair colour	rs2496632
134	187	rs10947996	Ease, Skin	Hair colour	rs2496632
134	187	rs6933697	Ease, Skin	Hair colour	rs2496632

135	188	rs62400428	Non-red	Hair colour, Blond vs. brown/black hair colour	rs1928529, rs35735653, rs9349337
135	189	rs9349337	Non-red	Hair colour, Blond vs. brown/black hair colour	rs1928529, rs35735653, rs9349337
136	191	rs13201830	Skin, Red	Hair colour, Red vs. brown/black hair colour	rs116033650, rs9463733
136	190	rs2025753	Skin, Red	Hair colour, Red vs. brown/black hair colour	rs116033650, rs9463733
136	190	rs12529074	Skin, Red	Hair colour, Red vs. brown/black hair colour	rs116033650, rs9463733
137	193	rs9443928	Ease, Skin		
137	192	rs1034241	Ease, Skin		
138	194	rs1487441	CS		
139	195	rs10457158	Non-red	Hair colour	rs11961833
140	196	rs2127965	Non-red	Hair colour	rs2127965, rs11280053
140	196	rs2185710	Non-red	Hair colour	rs2127965, rs11280053
141	197	rs7740107	Non-red	Hair colour	rs6569648
142	198	rs6901548	Non-red	Hair colour	rs6901548, rs9492790

143	199	rs4896038	Non-red	Hair colour, Blond vs. brown/black hair colour	rs4896038, rs137912555, rs374522	Melanoma	rs228437
144	201	rs1546966	Ease, Skin				
144	200	rs1407373	Ease, Skin				
144	201	rs73783709	Ease, Skin				
144	201	rs77568325	Ease, Skin				
145	202	rs34286635	Non-red	Hair colour	rs13201081, rs4897010		
145	203	rs4897010	Non-red	Hair colour	rs13201081, rs4897010		
146	204	rs10434895	Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour	rs10434895, rs4869723, rs9397040		
147	205	rs7754951	Skin, Non-red	Blond vs. brown/black hair colour, Hair colour, Brown vs. black hair colour	rs3212308, rs9347258, rs923198, rs73013664		
147	205	rs3123135	Skin, Non-red	Blond vs. brown/black hair colour, Hair colour, Brown vs. black hair colour	rs3212308, rs9347258, rs923198, rs73013664		

147	205	rs3123139	Skin, Non-red	Blond vs. brown/black hair colour, Hair colour, Brown vs. black hair colour	rs3212308, rs9347258, rs923198, rs73013664
147	205	rs1471836	Skin, Non-red	Blond vs. brown/black hair colour, Hair colour, Brown vs. black hair colour	rs3212308, rs9347258, rs923198, rs73013664
147	205	rs9457478	Skin, Non-red	Blond vs. brown/black hair colour, Hair colour, Brown vs. black hair colour	rs3212308, rs9347258, rs923198, rs73013664
148	206	rs9365518	Non-red	Hair colour	rs9355416
149	207	rs62435865	Ease		
150	208	rs1043291	Non-red		
150	209	rs2266921	Non-red		
151	210	rs2282867	Non-red	Hair colour, Blond vs. brown/black hair colour	rs2282867, rs3213661, rs3801108

152	213	rs1721040	Ease, Skin, CS	Cutaneous malignt melanoma, Nevus count or cutaneous melanoma, melanoma, Melanoma paper	rs1636744, rs73069846, rs117132860
152	211	rs1721019	Ease, Skin, CS	Cutaneous malignt melanoma, Nevus count or cutaneous melanoma, melanoma, Melanoma paper	rs1636744, rs73069846, rs117132860

152	211	rs34585474	Ease, Skin, CS	Cutaneous malignt melanoma, Nevus count or cutaneous melanoma, melanoma, Melanoma paper	rs1636744, rs73069846, rs117132860
152	212	rs117132860	Ease, Skin, CS	Cutaneous malignt melanoma, Nevus count or cutaneous melanoma, melanoma, Melanoma paper	rs1636744, rs73069846, rs117132860

153	214	rs849138	Ease, Non-red	Blond vs. brown/black hair colour, Hair colour	rs864745, rs849133, rs849134
153	214	rs864745	Ease, Non-red	Blond vs. brown/black hair colour, Hair colour	rs864745, rs849133, rs849134
153	214	rs1635852	Ease, Non-red	Blond vs. brown/black hair colour, Hair colour	rs864745, rs849133, rs849134
154	215	rs10224680	Non-red	Hair colour, Blond vs. brown/black hair colour	rs4722843, rs28531809
155	216	rs117441897	Non-red	Hair colour	rs77969347, rs34764931
155	217	rs77969347	Non-red	Hair colour	rs77969347, rs34764931
156	220	rs6966398	Non-red	Hair colour	rs56225270, rs1464841, rs12702237
156	219	rs1464841	Non-red	Hair colour	rs56225270, rs1464841, rs12702237
156	218	rs11764140	Non-red	Hair colour	rs56225270, rs1464841, rs12702237
157	221	rs35332062	Skin		

158	224	rs80030895	Non-red	Hair colour, Blond vs. brown/black hair colour	rs12667582, rs80030895, rs77162945, rs2710956, rs1548881
158	222	rs59201311	Non-red	Hair colour, Blond vs. brown/black hair colour	rs12667582, rs80030895, rs77162945, rs2710956, rs1548881
158	223	rs73217177	Non-red	Hair colour, Blond vs. brown/black hair colour	rs12667582, rs80030895, rs77162945, rs2710956, rs1548881
158	222	rs2157725	Non-red	Hair colour, Blond vs. brown/black hair colour	rs12667582, rs80030895, rs77162945, rs2710956, rs1548881
159	226	rs314349	Skin, Non-red	Hair colour, Blond vs. brown/black hair colour	rs117326834, rs2075756, rs314349, rs12535629, rs80308281
159	225	rs12535629	Skin, Non-red	Hair colour, Blond vs. brown/black hair colour	rs117326834, rs2075756, rs314349, rs12535629, rs80308281

159	226	rs56134000	Skin, Non-red	Hair colour, Blond vs. brown/black hair colour	rs117326834, rs2075756, rs314349, rs12535629, rs80308281		
160	227	rs56269269	Ease, Skin				
160	227	rs7794285	Ease, Skin				
161	228	rs2529369	Non-red	Hair colour, Blond vs. brown/black hair colour	rs2529369		
162	229	rs17152960	Pigscore 2	Hair colour, Blond vs. brown/black hair colour	rs2529369		
163	230	rs7803075	Non-red	Hair colour, Brown vs. black hair colour, Blond vs. brown/black hair colour	rs12706959, rs62471615, rs58270997, rs12669378, rs10954300	Melanoma paper, Cutaneous malignt melanoma	rs7778378, rs4731742
163	230	rs12669378	Non-red	Hair colour, Brown vs. black hair colour, Blond vs. brown/black hair colour	rs12706959, rs62471615, rs58270997, rs12669378, rs10954300	Melanoma paper, Cutaneous malignt melanoma	rs7778378, rs4731742

163	230	rs10954300	Non-red	Hair colour, Brown vs. black hair colour, Blond vs. brown/black hair colour	rs12706959, rs62471615, rs58270997, rs12669378, rs10954300	Melanoma paper, Cutaneous malignt melanoma	rs7778378, rs4731742
164	231	rs1612590	Non-red				
165	232	rs9767875	Non-red	Hair colour	rs9690044		
166	233	rs56326046	Non-red	Hair colour	rs56326046		
167	234	rs1533059	Non-red, CS	Hair colour	rs12550199		
167	234	rs330926	Non-red, CS	Hair colour	rs12550199		
167	234	rs11779019	Non-red, CS	Hair colour	rs12550199		
168	235	rs2048651	Non-red	Hair colour, Blond vs. brown/black hair colour	rs17676443, rs7845221, rs2048651		
169	236	rs2595041	Non-red	Hair colour	rs2595041		
170	237	rs57128498	Non-red	Hair colour, Blond vs. brown/black hair colour	rs57128498		

171	238	rs2070711	Pigscore 1	Hair colour, Blond vs. brown/black hair colour	rs72110873, rs113060680, rs2070711
172	240	rs6473306	Non-red	Hair colour	rs774470
172	240	rs2600605	Non-red	Hair colour	rs774470
172	240	rs12674770	Non-red	Hair colour	rs774470
172	239	rs12545508	Non-red	Hair colour	rs774470
173	241	rs2446928	Non-red		
174	242	rs1389985	Non-red	Hair colour	rs424012
174	243	rs446454	Non-red	Hair colour	rs424012
174	243	rs429884	Non-red	Hair colour	rs424012
175	244	rs12543799	Ease, Skin, Non-red, CS		
175	247	rs6469606	Ease, Skin, Non-red, CS		
175	245	rs2721929	Ease, Skin, Non-red, CS		

175	244	rs2737207	Ease, Skin, Non-red, CS
175	246	rs2737212	Ease, Skin, Non-red, CS
175	244	rs2737220	Ease, Skin, Non-red, CS
175	244	rs6981915	Ease, Skin, Non-red, CS
175	247	rs1526460	Ease, Skin, Non-red, CS
175	247	rs7842594	Ease, Skin, Non-red, CS
176	248	rs2445922	Non-red

177	249	rs478882	Non-red	Brown vs. black hair colour, Hair colour, Blond vs. brown/black hair colour	rs573246, rs520015	rs478882,	Nevus count or cutaneous melanoma	rs600951
177	249	rs520015	Non-red	Brown vs. black hair colour, Hair colour, Blond vs. brown/black hair colour	rs573246, rs520015	rs478882,	Nevus count or cutaneous melanoma	rs600951
178	250	rs4601374	Skin, Non-red	Blond vs. brown/black hair colour, Hair colour	rs2093657, rs2233173, rs419595	rs10815302, rs7037587,		
178	250	rs7037587	Skin, Non-red	Blond vs. brown/black hair colour, Hair colour	rs2093657, rs2233173, rs419595	rs10815302, rs7037587,		
179	252	rs75123768	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Low tan response	rs13289810, rs13288130, rs10960765	rs1137134, rs1326797,	Melanoma paper	rs10960710

179	252	rs1925238	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Low tan response	rs13289810, rs13288130, rs10960765	rs1137134, rs1326797,	Melanoma paper	rs10960710
179	251	rs118087215	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Low tan response	rs13289810, rs13288130, rs10960765	rs1137134, rs1326797,	Melanoma paper	rs10960710
179	252	rs10960749	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Low tan response	rs13289810, rs13288130, rs10960765	rs1137134, rs1326797,	Melanoma paper	rs10960710
179	252	rs75630385	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Low tan response	rs13289810, rs13288130, rs10960765	rs1137134, rs1326797,	Melanoma paper	rs10960710
179	251	rs1137134	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Low tan response	rs13289810, rs13288130, rs10960765	rs1137134, rs1326797,	Melanoma paper	rs10960710
179	251	rs1326797	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Low tan response	rs13289810, rs13288130, rs10960765	rs1137134, rs1326797,	Melanoma paper	rs10960710
180	253	rs770197	Non-red	Hair colour	rs770199			

181	255	rs7858025	Ease, Skin, Non-red, CS	Hair colour, Freckles, Blond vs. brown/black hair colour, Low tan response, Brown vs. black hair colour, Freckling, Skin pigmentation traits	rs12115663, rs10810635, rs10962599, rs10962612, rs7865762, rs2153271, rs10810650, rs10810657, rs12350739
181	255	rs10810636	Ease, Skin, Non-red, CS	Hair colour, Freckles, Blond vs. brown/black hair colour, Low tan response, Brown vs. black hair colour, Freckling, Skin pigmentation traits	rs12115663, rs10810635, rs10962599, rs10962612, rs7865762, rs2153271, rs10810650, rs10810657, rs12350739
181	256	rs10962605	Ease, Skin, Non-red, CS	Hair colour, Freckles, Blond vs. brown/black hair colour, Low tan response, Brown vs. black hair colour, Freckling, Skin pigmentation traits	rs12115663, rs10810635, rs10962599, rs10962612, rs7865762, rs2153271, rs10810650, rs10810657, rs12350739
181	255	rs2254330	Ease, Skin, Non-red, CS	Hair colour, Freckles, Blond vs. brown/black hair colour, Low tan response, Brown vs. black hair colour, Freckling, Skin pigmentation traits	rs12115663, rs10810635, rs10962599, rs10962612, rs7865762, rs2153271, rs10810650, rs10810657, rs12350739

181	256	rs12350739	Ease, Skin, Non-red, CS	Hair colour, Freckles, Blond vs. brown/black hair colour, Low tan response, Brown vs. black hair colour, Freckling, Skin pigmentation traits	rs12115663, rs10810635, rs10962599, rs10962612, rs7865762, rs2153271, rs10810650, rs10810657, rs12350739		
181	254	rs62543565	Ease, Skin, Non-red, CS	Hair colour, Freckles, Blond vs. brown/black hair colour, Low tan response, Brown vs. black hair colour, Freckling, Skin pigmentation traits	rs12115663, rs10810635, rs10962599, rs10962612, rs7865762, rs2153271, rs10810650, rs10810657, rs12350739		
182	257	rs10811647	Non-red	Hair colour	rs10811650, rs9632885	Melanoma paper, Nevus count or cutaneous melanoma, Melanoma	rs871024, rs869329, rs201131773, rs7023329, rs11532907, rs2027939, rs55797833, rs3217986, rs79356439, rs75883022

183	258	rs10812605	Non-red	Hair colour	rs2484314		
183	258	rs10967976	Non-red	Hair colour	rs2484314		
184	259	rs11557154	Non-red	Hair colour	rs11557154		
185	260	rs1475675	Non-red				
186	261	rs3021523	Non-red	Hair colour	rs3021523, rs10818214		
186	261	rs35324451	Non-red	Hair colour	rs3021523, rs10818214		
187	262	rs1055187	Pigscore 1				
188	263	rs10739220	Non-red	Hair colour	rs10739220	Melanoma paper	rs10739220
189	264	rs540599	Non-red				
190	265	rs10985890	Skin				
191	267	rs57425397	Non-red	Hair colour	rs58979150, rs2787594		
191	271	rs969585	Non-red	Hair colour	rs58979150, rs2787594		
191	268	rs10818907	Non-red	Hair colour	rs58979150, rs2787594		
191	268	rs12555707	Non-red	Hair colour	rs58979150, rs2787594		
191	269	rs58979150	Non-red	Hair colour	rs58979150, rs2787594		
191	266	rs16936765	Non-red	Hair colour	rs58979150, rs2787594		
191	269	rs4838097	Non-red	Hair colour	rs58979150, rs2787594		
191	270	rs953470	Non-red	Hair colour	rs58979150, rs2787594		

191	267	rs10818930	Non-red	Hair colour	rs58979150, rs2787594		
192	272	rs76798990	Skin, Non-red				
192	272	rs7032500	Skin, Non-red				
193	273	rs1129614	Ease			Nevus count or cutaneous melanoma	rs45575338
194	274	rs17230340	Non-red				
195	275	rs12264698	Non-red	Hair colour	rs6481946, rs73262807		
195	275	rs1148246	Non-red	Hair colour	rs6481946, rs73262807		
196	276	rs2236295	Ease, Skin				
196	276	rs1888967	Ease, Skin				
197	277	rs4433500	Non-red	Hair colour	rs7074233		
198	278	rs703978	Non-red	Hair colour	rs703978		
199	279	rs55901776	Ease, Skin				
199	279	rs10788623	Ease, Skin				
199	279	rs10788627	Ease, Skin				

200	280	rs7086846	Ease, Skin		
200	280	rs56014906	Ease, Skin		
201	281	rs4918614	Non-red	Hair colour	rs7088364
202	282	rs10444039	Ease, Skin, CS		
202	282	rs35563099	Ease, Skin, CS		
202	282	rs7098111	Ease, Skin, CS		
203	283	rs4078279	Non-red	Brown vs. black hair colour, Hair colour, Blond vs. brown/black hair colour	rs4078279, rs11041426
203	283	rs11041426	Non-red	Brown vs. black hair colour, Hair colour, Blond vs. brown/black hair colour	rs4078279, rs11041426

204	287	rs75319234	Ease, Skin, Non-red, CS	Blond vs. brown/black hair colour, Hair colour	rs1531903, rs7108738, rs727324, rs117188914, rs61881743, rs117573112, rs7109376, rs61881798, rs7943712, rs4757417	Melanoma paper	rs7941496
204	287	rs7108738	Ease, Skin, Non-red, CS	Blond vs. brown/black hair colour, Hair colour	rs1531903, rs7108738, rs727324, rs117188914, rs61881743, rs117573112, rs7109376, rs61881798, rs7943712, rs4757417	Melanoma paper	rs7941496
204	291	rs1920672	Ease, Skin, Non-red, CS	Blond vs. brown/black hair colour, Hair colour	rs1531903, rs7108738, rs727324, rs117188914, rs61881743, rs117573112, rs7109376, rs61881798, rs7943712, rs4757417	Melanoma paper	rs7941496

204	289	rs4756838	Ease, Skin, Non-red, CS	Blond vs. brown/black hair colour, Hair colour	rs1531903, rs7108738, rs727324, rs117188914, rs61881743, rs117573112, rs7109376, rs61881798, rs7943712, rs4757417	Melanoma paper	rs7941496
204	290	rs10832552	Ease, Skin, Non-red, CS	Blond vs. brown/black hair colour, Hair colour	rs1531903, rs7108738, rs727324, rs117188914, rs61881743, rs117573112, rs7109376, rs61881798, rs7943712, rs4757417	Melanoma paper	rs7941496
204	286	rs77535014	Ease, Skin, Non-red, CS	Blond vs. brown/black hair colour, Hair colour	rs1531903, rs7108738, rs727324, rs117188914, rs61881743, rs117573112, rs7109376, rs61881798, rs7943712, rs4757417	Melanoma paper	rs7941496

204	289	rs10766301	Ease, Skin, Non-red, CS	Blond vs. brown/black hair colour, Hair colour	rs1531903, rs7108738, rs727324, rs117188914, rs61881743, rs117573112, rs7109376, rs61881798, rs7943712, rs4757417	Melanoma paper	rs7941496
204	291	rs7109206	Ease, Skin, Non-red, CS	Blond vs. brown/black hair colour, Hair colour	rs1531903, rs7108738, rs727324, rs117188914, rs61881743, rs117573112, rs7109376, rs61881798, rs7943712, rs4757417	Melanoma paper	rs7941496
204	286	rs12277740	Ease, Skin, Non-red, CS	Blond vs. brown/black hair colour, Hair colour	rs1531903, rs7108738, rs727324, rs117188914, rs61881743, rs117573112, rs7109376, rs61881798, rs7943712, rs4757417	Melanoma paper	rs7941496

204	285	rs297343	Ease, Skin, Non-red, CS	Blond vs. brown/black hair colour, Hair colour	rs1531903, rs7108738, rs727324, rs117188914, rs61881743, rs117573112, rs7109376, rs61881798, rs7943712, rs4757417	Melanoma paper	rs7941496
204	289	rs7929530	Ease, Skin, Non-red, CS	Blond vs. brown/black hair colour, Hair colour	rs1531903, rs7108738, rs727324, rs117188914, rs61881743, rs117573112, rs7109376, rs61881798, rs7943712, rs4757417	Melanoma paper	rs7941496
204	291	rs151060912	Ease, Skin, Non-red, CS	Blond vs. brown/black hair colour, Hair colour	rs1531903, rs7108738, rs727324, rs117188914, rs61881743, rs117573112, rs7109376, rs61881798, rs7943712, rs4757417	Melanoma paper	rs7941496

204	284	rs11023988	Ease, Skin, Non-red, CS	Blond vs. brown/black hair colour, Hair colour	rs1531903, rs7108738, rs727324, rs117188914, rs61881743, rs117573112, rs7109376, rs61881798, rs7943712, rs4757417	Melanoma paper	rs7941496
204	288	rs4578351	Ease, Skin, Non-red, CS	Blond vs. brown/black hair colour, Hair colour	rs1531903, rs7108738, rs727324, rs117188914, rs61881743, rs117573112, rs7109376, rs61881798, rs7943712, rs4757417	Melanoma paper	rs7941496
204	289	rs72632979	Ease, Skin, Non-red, CS	Blond vs. brown/black hair colour, Hair colour	rs1531903, rs7108738, rs727324, rs117188914, rs61881743, rs117573112, rs7109376, rs61881798, rs7943712, rs4757417	Melanoma paper	rs7941496
205	292	rs2049129	Skin, Non-red	Hair colour	rs12416821, rs12419588		

205	292	rs7478729	Skin, Non-red	Hair colour	rs12416821, rs12419588
206	293	rs11029693	Non-red		
207	294	rs4923447	Non-red	Hair colour, Blond vs. brown/black hair colour	rs1382470, rs58604758
208	295	rs145678014	Non-red	Hair colour	rs145678014
209	296	rs11037723	Non-red	Hair colour, Blond vs. brown/black hair colour	rs11037728, rs12421995, rs66716358, rs724411, rs357925
209	298	rs66716358	Non-red	Hair colour, Blond vs. brown/black hair colour	rs11037728, rs12421995, rs66716358, rs724411, rs357925
209	297	rs1902961	Non-red	Hair colour, Blond vs. brown/black hair colour	rs11037728, rs12421995, rs66716358, rs724411, rs357925
210	300	rs7395496	Ease, Non-red	Hair colour	rs2291120, rs60562402
210	299	rs2053979	Ease, Non-red	Hair colour	rs2291120, rs60562402

211	301	rs1938594	Pigscore 2				
212	303	rs7940665	Non-red	Hair colour, Blond vs. brown/black hair colour	rs61896141		
212	302	rs61897795	Non-red	Hair colour, Blond vs. brown/black hair colour	rs61896141		
213	304	rs10897275	Non-red	Hair colour, Blond vs. brown/black hair colour	rs61896141, rs9645690, rs10897275		
213	304	rs9645690	Non-red	Hair colour, Blond vs. brown/black hair colour	rs61896141, rs9645690, rs10897275		
214	305	rs11227247	Non-red	Hair colour, Blond vs. brown/black hair colour	rs12797706, rs56019505		
215	306	rs519380	CS	Hair colour	rs10896139		
216	307	rs72917317	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Low tan response, Brown vs. black hair colour, Blond vs. brown hair colour, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour	rs671191, rs118074499, rs72917317, rs7940105, rs72928978, rs3750965, rs35264875, rs34510004, rs3829241, rs1060435, rs72930659, rs72932540, rs34345560	Cutaneous malignt melanoma	rs2290419

216	307	rs150527451	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Low tan response, Brown vs. black hair colour, Blond vs. brown hair colour, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour	rs671191, rs118074499, rs72917317, rs7940105, rs72928978, rs3750965, rs35264875, rs34510004, rs3829241, rs1060435, rs72930659, rs72932540, rs34345560	Cutaneous malignt melanoma	rs2290419
216	307	rs72928978	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Low tan response, Brown vs. black hair colour, Blond vs. brown hair colour, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour	rs671191, rs118074499, rs72917317, rs7940105, rs72928978, rs3750965, rs35264875, rs34510004, rs3829241, rs1060435, rs72930659, rs72932540, rs34345560	Cutaneous malignt melanoma	rs2290419

216	307	rs3829241	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Low tan response, Brown vs. black hair colour, Blond vs. brown hair colour, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour	rs671191, rs118074499, rs72917317, rs7940105, rs72928978, rs3750965, rs35264875, rs34510004, rs3829241, rs1060435, rs72930659, rs72932540, rs34345560	Cutaneous malignt melanoma	rs2290419
217	308	rs652190	Red	Hair colour, Blond vs. brown/black hair colour, Low tan response, Brown vs. black hair colour, Blond vs. brown hair colour, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Red vs. brown/black hair colour	rs7940105, rs72917317, rs72928978, rs3750965, rs35264875, rs34510004, rs3829241, rs1060435, rs72930659, rs72932540, rs34345560, rs11263458, rs12806763, rs638640, rs118067673	Cutaneous malignt melanoma, Melanoma paper, Nevus count or cutaneous melanoma, Melanoma	rs2290419, rs4354713, rs498136
218	309	rs881361	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour	rs5792809, rs10899501, rs2292572		

218	309	rs11237489	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour	rs5792809, rs2292572	rs10899501,
218	309	rs10899501	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour	rs5792809, rs2292572	rs10899501,
219	310	rs592128	Pigscore 2	Hair colour	rs7925150	

220	312	rs117796400	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Skin, hair and eye pigmentation (multivariate alysis), Blond vs. brown/black hair colour, Tanning, Brown vs. black hair colour, Freckles, Skin pigmentation, Skin pigmentation traits	rs7925150, rs80150794, rs217120, rs7937452, rs75943957, rs12792431, rs188092667, rs7118677, rs71469216, rs148065054, rs10831496, rs598952, rs1042602, rs4753772, rs4121827, rs150319166, rs34749698, rs1393350, rs1126809, rs317194, rs11018623, rs12276991, rs61905505, rs17221124, rs78539082, rs10741336, rs1944068	Melanoma paper, melanoma, Nevus count or cutaneous melanoma	rs1126809, rs1847134, rs1393350, rs10830253

220	321	rs11021172	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Skin, hair and eye pigmentation (multivariate alysis), Blond vs. brown/black hair colour, Tanning, Brown vs. black hair colour, Freckles, Skin pigmentation, Skin pigmentation traits	rs7925150, rs80150794, rs217120, rs7937452, rs75943957, rs12792431, rs188092667, rs7118677, rs71469216, rs148065054, rs10831496, rs598952, rs1042602, rs4753772, rs4121827, rs150319166, rs34749698, rs1393350, rs1126809, rs317194, rs11018623, rs12276991, rs61905505, rs17221124, rs78539082, rs10741336, rs1944068	Melanoma paper, melanoma, Nevus count or cutaneous melanoma	rs1126809, rs1847134, rs1393350, rs10830253

220	312	rs2187128	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Skin, hair and eye pigmentation (multivariate alysis), Blond vs. brown/black hair colour, Tanning, Brown vs. black hair colour, Freckles, Skin pigmentation, Skin pigmentation traits	rs7925150, rs80150794, rs217120, rs7937452, rs75943957, rs12792431, rs188092667, rs7118677, rs71469216, rs148065054, rs10831496, rs598952, rs1042602, rs4753772, rs4121827, rs150319166, rs34749698, rs1393350, rs1126809, rs317194, rs11018623, rs12276991, rs61905505, rs17221124, rs78539082, rs10741336, rs1944068	Melanoma paper, melanoma, Nevus count or cutaneous melanoma	rs1126809, rs1847134, rs1393350, rs10830253

220	313	rs1042602	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Skin, hair and eye pigmentation (multivariate alysis), Blond vs. brown/black hair colour, Tanning, Brown vs. black hair colour, Freckles, Skin pigmentation, Skin pigmentation traits	rs7925150, rs80150794, rs217120, rs7937452, rs75943957, rs12792431, rs188092667, rs7118677, rs71469216, rs148065054, rs10831496, rs598952, rs1042602, rs4753772, rs4121827, rs150319166, rs34749698, rs1393350, rs1126809, rs317194, rs11018623, rs12276991, rs61905505, rs17221124, rs78539082, rs10741336, rs1944068	Melanoma paper, melanoma, Nevus count or cutaneous melanoma	rs1126809, rs1847134, rs1393350, rs10830253

220	311	rs1942493	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Skin, hair and eye pigmentation (multivariate alysis), Blond vs. brown/black hair colour, Tanning, Brown vs. black hair colour, Freckles, Skin pigmentation, Skin pigmentation traits	rs7925150, rs80150794, rs217120, rs7937452, rs75943957, rs12792431, rs188092667, rs7118677, rs71469216, rs148065054, rs10831496, rs598952, rs1042602, rs4753772, rs4121827, rs150319166, rs34749698, rs1393350, rs1126809, rs317194, rs11018623, rs12276991, rs61905505, rs17221124, rs78539082, rs10741336, rs1944068	Melanoma paper, melanoma, Nevus count or cutaneous melanoma	rs1126809, rs1847134, rs1393350, rs10830253

220	320	rs35300336	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Skin, hair and eye pigmentation (multivariate alysis), Blond vs. brown/black hair colour, Tanning, Brown vs. black hair colour, Freckles, Skin pigmentation, Skin pigmentation traits	rs7925150, rs80150794, rs217120, rs7937452, rs75943957, rs12792431, rs188092667, rs7118677, rs71469216, rs148065054, rs10831496, rs598952, rs1042602, rs4753772, rs4121827, rs150319166, rs34749698, rs1393350, rs1126809, rs317194, rs11018623, rs12276991, rs61905505, rs17221124, rs78539082, rs10741336, rs1944068	Melanoma paper, melanoma, Nevus count or cutaneous melanoma	rs1126809, rs1847134, rs1393350, rs10830253

220	316	rs4512823	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Skin, hair and eye pigmentation (multivariate alysis), Blond vs. brown/black hair colour, Tanning, Brown vs. black hair colour, Freckles, Skin pigmentation, Skin pigmentation traits	rs7925150, rs80150794, rs217120, rs7937452, rs75943957, rs12792431, rs188092667, rs7118677, rs71469216, rs148065054, rs10831496, rs598952, rs1042602, rs4753772, rs4121827, rs150319166, rs34749698, rs1393350, rs1126809, rs317194, rs11018623, rs12276991, rs61905505, rs17221124, rs78539082, rs10741336, rs1944068	Melanoma paper, melanoma, Nevus count or cutaneous melanoma	rs1126809, rs1847134, rs1393350, rs10830253

220	317	rs183139540	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Skin, hair and eye pigmentation (multivariate alysis), Blond vs. brown/black hair colour, Tanning, Brown vs. black hair colour, Freckles, Skin pigmentation, Skin pigmentation traits	rs7925150, rs80150794, rs217120, rs7937452, rs75943957, rs12792431, rs188092667, rs7118677, rs71469216, rs148065054, rs10831496, rs598952, rs1042602, rs4753772, rs4121827, rs150319166, rs34749698, rs1393350, rs1126809, rs317194, rs11018623, rs12276991, rs61905505, rs17221124, rs78539082, rs10741336, rs1944068	Melanoma paper, melanoma, Nevus count or cutaneous melanoma	rs1126809, rs1847134, rs1393350, rs10830253

220	322	rs144650507	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Skin, hair and eye pigmentation (multivariate alysis), Blond vs. brown/black hair colour, Tanning, Brown vs. black hair colour, Freckles, Skin pigmentation, Skin pigmentation traits	rs7925150, rs80150794, rs217120, rs7937452, rs75943957, rs12792431, rs188092667, rs7118677, rs71469216, rs148065054, rs10831496, rs598952, rs1042602, rs4753772, rs4121827, rs150319166, rs34749698, rs1393350, rs1126809, rs317194, rs11018623, rs12276991, rs61905505, rs17221124, rs78539082, rs10741336, rs1944068	Melanoma paper, melanoma, Nevus count or cutaneous melanoma	rs1126809, rs1847134, rs1393350, rs10830253

220	315	rs147546939	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Skin, hair and eye pigmentation (multivariate alysis), Blond vs. brown/black hair colour, Tanning, Brown vs. black hair colour, Freckles, Skin pigmentation, Skin pigmentation traits	rs7925150, rs80150794, rs217120, rs7937452, rs75943957, rs12792431, rs188092667, rs7118677, rs71469216, rs148065054, rs10831496, rs598952, rs1042602, rs4753772, rs4121827, rs150319166, rs34749698, rs1393350, rs1126809, rs317194, rs11018623, rs12276991, rs61905505, rs17221124, rs78539082, rs10741336, rs1944068	Melanoma paper, melanoma, Nevus count or cutaneous melanoma	rs1126809, rs1847134, rs1393350, rs10830253

220	312	rs1126809	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Skin, hair and eye pigmentation (multivariate alysis), Blond vs. brown/black hair colour, Tanning, Brown vs. black hair colour, Freckles, Skin pigmentation, Skin pigmentation traits	rs7925150, rs80150794, rs217120, rs7937452, rs75943957, rs12792431, rs188092667, rs7118677, rs71469216, rs148065054, rs10831496, rs598952, rs1042602, rs4753772, rs4121827, rs150319166, rs34749698, rs1393350, rs1126809, rs317194, rs11018623, rs12276991, rs61905505, rs17221124, rs78539082, rs10741336, rs1944068	Melanoma paper, melanoma, Nevus count or cutaneous melanoma	rs1126809, rs1847134, rs1393350, rs10830253

220	319	rs78043739	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Skin, hair and eye pigmentation (multivariate alysis), Blond vs. brown/black hair colour, Tanning, Brown vs. black hair colour, Freckles, Skin pigmentation, Skin pigmentation traits	rs7925150, rs80150794, rs217120, rs7937452, rs75943957, rs12792431, rs188092667, rs7118677, rs71469216, rs148065054, rs10831496, rs598952, rs1042602, rs4753772, rs4121827, rs150319166, rs34749698, rs1393350, rs1126809, rs317194, rs11018623, rs12276991, rs61905505, rs17221124, rs78539082, rs10741336, rs1944068	Melanoma paper, melanoma, Nevus count or cutaneous melanoma	rs1126809, rs1847134, rs1393350, rs10830253

220	313	rs77368047	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Skin, hair and eye pigmentation (multivariate alysis), Blond vs. brown/black hair colour, Tanning, Brown vs. black hair colour, Freckles, Skin pigmentation, Skin pigmentation traits	rs7925150, rs80150794, rs217120, rs7937452, rs75943957, rs12792431, rs188092667, rs7118677, rs71469216, rs148065054, rs10831496, rs598952, rs1042602, rs4753772, rs4121827, rs150319166, rs34749698, rs1393350, rs1126809, rs317194, rs11018623, rs12276991, rs61905505, rs17221124, rs78539082, rs10741336, rs1944068	Melanoma paper, melanoma, Nevus count or cutaneous melanoma	rs1126809, rs1847134, rs1393350, rs10830253

220	320	rs200522762	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Skin, hair and eye pigmentation (multivariate alysis), Blond vs. brown/black hair colour, Tanning, Brown vs. black hair colour, Freckles, Skin pigmentation, Skin pigmentation traits	rs7925150, rs80150794, rs217120, rs7937452, rs75943957, rs12792431, rs188092667, rs7118677, rs71469216, rs148065054, rs10831496, rs598952, rs1042602, rs4753772, rs4121827, rs150319166, rs34749698, rs1393350, rs1126809, rs317194, rs11018623, rs12276991, rs61905505, rs17221124, rs78539082, rs10741336, rs1944068	Melanoma paper, melanoma, Nevus count or cutaneous melanoma	rs1126809, rs1847134, rs1393350, rs10830253

220	312	rs2634309	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Skin, hair and eye pigmentation (multivariate alysis), Blond vs. brown/black hair colour, Tanning, Brown vs. black hair colour, Freckles, Skin pigmentation, Skin pigmentation traits	rs7925150, rs80150794, rs217120, rs7937452, rs75943957, rs12792431, rs188092667, rs7118677, rs71469216, rs148065054, rs10831496, rs598952, rs1042602, rs4753772, rs4121827, rs150319166, rs34749698, rs1393350, rs1126809, rs317194, rs11018623, rs12276991, rs61905505, rs17221124, rs78539082, rs10741336, rs1944068	Melanoma paper, melanoma, Nevus count or cutaneous melanoma	rs1126809, rs1847134, rs1393350, rs10830253

220	314	rs140100229	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Skin, hair and eye pigmentation (multivariate alysis), Blond vs. brown/black hair colour, Tanning, Brown vs. black hair colour, Freckles, Skin pigmentation, Skin pigmentation traits	rs7925150, rs80150794, rs217120, rs7937452, rs75943957, rs12792431, rs188092667, rs7118677, rs71469216, rs148065054, rs10831496, rs598952, rs1042602, rs4753772, rs4121827, rs150319166, rs34749698, rs1393350, rs1126809, rs317194, rs11018623, rs12276991, rs61905505, rs17221124, rs78539082, rs10741336, rs1944068	Melanoma paper, melanoma, Nevus count or cutaneous melanoma	rs1126809, rs1847134, rs1393350, rs10830253

220	314	rs10830452	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Skin, hair and eye pigmentation (multivariate alysis), Blond vs. brown/black hair colour, Tanning, Brown vs. black hair colour, Freckles, Skin pigmentation, Skin pigmentation traits	rs7925150, rs80150794, rs217120, rs7937452, rs75943957, rs12792431, rs188092667, rs7118677, rs71469216, rs148065054, rs10831496, rs598952, rs1042602, rs4753772, rs4121827, rs150319166, rs34749698, rs1393350, rs1126809, rs317194, rs11018623, rs12276991, rs61905505, rs17221124, rs78539082, rs10741336, rs1944068	Melanoma paper, melanoma, Nevus count or cutaneous melanoma	rs1126809, rs1847134, rs1393350, rs10830253

220	318	rs17201363	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Skin, hair and eye pigmentation (multivariate alysis), Blond vs. brown/black hair colour, Tanning, Brown vs. black hair colour, Freckles, Skin pigmentation, Skin pigmentation traits	rs7925150, rs80150794, rs217120, rs7937452, rs75943957, rs12792431, rs188092667, rs7118677, rs71469216, rs148065054, rs10831496, rs598952, rs1042602, rs4753772, rs4121827, rs150319166, rs34749698, rs1393350, rs1126809, rs317194, rs11018623, rs12276991, rs61905505, rs17221124, rs78539082, rs10741336, rs1944068	Melanoma paper, melanoma, Nevus count or cutaneous melanoma	rs1126809, rs1847134, rs1393350, rs10830253
221	323	rs492335	Non-red	Hair colour	rs10791037		
222	324	rs3764032	Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour	rs3764032, rs11834692		

223	325	rs10771034	Ease, Skin, Non-red	Hair colour	rs9971729
223	325	rs9971729	Ease, Skin, Non-red	Hair colour	rs9971729
224	326	rs6487406	Non-red	Hair colour	rs61914312
225	327	rs11049319	Non-red		
226	328	rs2405295	Ease, Skin	Skin pigmentation	rs1902910
226	328	rs10879095	Ease, Skin	Skin pigmentation	rs1902910
227	329	rs7979418	Skin, Non-red	Blond vs. brown/black hair colour, Hair colour	rs4768698, rs7968283
228	330	rs10875910	Non-red	Hair colour	rs10875912, rs10875910
229	331	rs1470321	Non-red	Blond vs. brown/black hair colour, Hair colour	rs535209331, rs116996456
230	332	rs11172124	Non-red	Hair colour	rs7965424
231	333	rs1147100	Non-red	Hair colour	rs2620709, rs7974210
231	333	rs1147094	Non-red	Hair colour	rs2620709, rs7974210

232	344	rs1818930	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Low tan response, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Blond vs. brown hair colour	rs1564120, rs7299005, rs148137046, rs184079650, rs10858628, rs35466244, rs34404753, rs7487365, rs7966318, rs2216153, rs7977437, rs7306001, rs115183712, rs35618688, rs1907703, rs2639099, rs34200313, rs12821256, rs146898186, rs12298351, rs12820818, rs58696617, rs7299139, rs143993772, rs10859040

232	337	rs12425342	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Low tan response, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Blond vs. brown hair colour	rs1564120, rs7299005, rs148137046, rs184079650, rs10858628, rs35466244, rs34404753, rs7487365, rs7966318, rs2216153, rs7977437, rs7306001, rs115183712, rs35618688, rs1907703, rs2639099, rs34200313, rs12821256, rs146898186, rs12298351, rs12820818, rs58696617, rs7299139, rs143993772, rs10859040

232	338	rs73200863	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Low tan response	rs10858628, rs35466244, rs34404753, rs7487365, rs7966318, rs2216153, rs7977437, rs7306001, rs115183712, rs35618688, rs1907703, rs2639099, rs34200313

232	337	rs10858711	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Low tan response, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Blond vs. brown hair colour	rs1564120, rs7299005, rs148137046, rs184079650, rs10858628, rs35466244, rs34404753, rs7487365, rs7966318, rs2216153, rs7977437, rs7306001, rs115183712, rs35618688, rs1907703, rs2639099, rs34200313, rs12821256, rs146898186, rs12298351, rs12820818, rs58696617, rs7299139, rs143993772, rs10859040

232	337	rs11104870	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Low tan response, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Blond vs. brown hair colour	rs1564120, rs7299005, rs148137046, rs184079650, rs10858628, rs35466244, rs34404753, rs7487365, rs7966318, rs2216153, rs7977437, rs7306001, rs115183712, rs35618688, rs1907703, rs2639099, rs34200313, rs12821256, rs146898186, rs12298351, rs12820818, rs58696617, rs7299139, rs143993772, rs10859040

232	338	rs11104887	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Low tan response, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Blond vs. brown hair colour	rs1564120, rs7299005, rs148137046, rs184079650, rs10858628, rs35466244, rs34404753, rs7487365, rs7966318, rs2216153, rs7977437, rs7306001, rs115183712, rs35618688, rs1907703, rs2639099, rs34200313, rs12821256, rs146898186, rs12298351, rs12820818, rs58696617, rs7299139, rs143993772, rs10859040

232	335	rs11104947	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Low tan response, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Blond vs. brown hair colour	rs1564120, rs7299005, rs148137046, rs184079650, rs10858628, rs35466244, rs34404753, rs7487365, rs7966318, rs2216153, rs7977437, rs7306001, rs115183712, rs35618688, rs1907703, rs2639099, rs34200313, rs12821256, rs146898186, rs12298351, rs12820818, rs58696617, rs7299139, rs143993772, rs10859040

232	340	rs4590952	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Low tan response, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Blond vs. brown hair colour	rs1564120, rs7299005, rs148137046, rs184079650, rs10858628, rs35466244, rs34404753, rs7487365, rs7966318, rs2216153, rs7977437, rs7306001, rs115183712, rs35618688, rs1907703, rs2639099, rs34200313, rs12821256, rs146898186, rs12298351, rs12820818, rs58696617, rs7299139, rs143993772, rs10859040

232	345	rs1798011	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Low tan response, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Blond vs. brown hair colour	rs1564120, rs7299005, rs148137046, rs184079650, rs10858628, rs35466244, rs34404753, rs7487365, rs7966318, rs2216153, rs7977437, rs7306001, rs115183712, rs35618688, rs1907703, rs2639099, rs34200313, rs12821256, rs146898186, rs12298351, rs12820818, rs58696617, rs7299139, rs143993772, rs10859040

232	339	rs10858760	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Low tan response, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Blond vs. brown hair colour	rs1564120, rs7299005, rs148137046, rs184079650, rs10858628, rs35466244, rs34404753, rs7487365, rs7966318, rs2216153, rs7977437, rs7306001, rs115183712, rs35618688, rs1907703, rs2639099, rs34200313, rs12821256, rs146898186, rs12298351, rs12820818, rs58696617, rs7299139, rs143993772, rs10859040

232	334	rs77081360	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Low tan response, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Blond vs. brown hair colour	rs1564120, rs7299005, rs148137046, rs184079650, rs10858628, rs35466244, rs34404753, rs7487365, rs7966318, rs2216153, rs7977437, rs7306001, rs115183712, rs35618688, rs1907703, rs2639099, rs34200313, rs12821256, rs146898186, rs12298351, rs12820818, rs58696617, rs7299139, rs143993772, rs10859040

232	338	rs2100883	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Low tan response, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Blond vs. brown hair colour	rs1564120, rs7299005, rs148137046, rs184079650, rs10858628, rs35466244, rs34404753, rs7487365, rs7966318, rs2216153, rs7977437, rs7306001, rs115183712, rs35618688, rs1907703, rs2639099, rs34200313, rs12821256, rs146898186, rs12298351, rs12820818, rs58696617, rs7299139, rs143993772, rs10859040

232	335	rs12821256	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Low tan response, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Blond vs. brown hair colour	rs1564120, rs7299005, rs148137046, rs184079650, rs10858628, rs35466244, rs34404753, rs7487365, rs7966318, rs2216153, rs7977437, rs7306001, rs115183712, rs35618688, rs1907703, rs2639099, rs34200313, rs12821256, rs146898186, rs12298351, rs12820818, rs58696617, rs7299139, rs143993772, rs10859040

232	346	rs9651934	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Low tan response, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Blond vs. brown hair colour	rs1564120, rs7299005, rs148137046, rs184079650, rs10858628, rs35466244, rs34404753, rs7487365, rs7966318, rs2216153, rs7977437, rs7306001, rs115183712, rs35618688, rs1907703, rs2639099, rs34200313, rs12821256, rs146898186, rs12298351, rs12820818, rs58696617, rs7299139, rs143993772, rs10859040

232	342	rs189785589	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Low tan response, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Blond vs. brown hair colour	rs1564120, rs7299005, rs148137046, rs184079650, rs10858628, rs35466244, rs34404753, rs7487365, rs7966318, rs2216153, rs7977437, rs7306001, rs115183712, rs35618688, rs1907703, rs2639099, rs34200313, rs12821256, rs146898186, rs12298351, rs12820818, rs58696617, rs7299139, rs143993772, rs10859040

232	337	rs2638470	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Low tan response, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Blond vs. brown hair colour	rs1564120, rs7299005, rs148137046, rs184079650, rs10858628, rs35466244, rs34404753, rs7487365, rs7966318, rs2216153, rs7977437, rs7306001, rs115183712, rs35618688, rs1907703, rs2639099, rs34200313, rs12821256, rs146898186, rs12298351, rs12820818, rs58696617, rs7299139, rs143993772, rs10859040

232	346	rs77495873	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Low tan response, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Blond vs. brown hair colour	rs1564120, rs7299005, rs148137046, rs184079650, rs10858628, rs35466244, rs34404753, rs7487365, rs7966318, rs2216153, rs7977437, rs7306001, rs115183712, rs35618688, rs1907703, rs2639099, rs34200313, rs12821256, rs146898186, rs12298351, rs12820818, rs58696617, rs7299139, rs143993772, rs10859040

232	339	rs10777186	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Low tan response, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Blond vs. brown hair colour	rs1564120, rs7299005, rs148137046, rs184079650, rs10858628, rs35466244, rs34404753, rs7487365, rs7966318, rs2216153, rs7977437, rs7306001, rs115183712, rs35618688, rs1907703, rs2639099, rs34200313, rs12821256, rs146898186, rs12298351, rs12820818, rs58696617, rs7299139, rs143993772, rs10859040

232	343	rs17782847	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Low tan response, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Blond vs. brown hair colour	rs1564120, rs7299005, rs148137046, rs184079650, rs10858628, rs35466244, rs34404753, rs7487365, rs7966318, rs2216153, rs7977437, rs7306001, rs115183712, rs35618688, rs1907703, rs2639099, rs34200313, rs12821256, rs146898186, rs12298351, rs12820818, rs58696617, rs7299139, rs143993772, rs10859040

232	341	rs2553101	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Low tan response, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Blond vs. brown hair colour	rs1564120, rs7299005, rs148137046, rs184079650, rs10858628, rs35466244, rs34404753, rs7487365, rs7966318, rs2216153, rs7977437, rs7306001, rs115183712, rs35618688, rs1907703, rs2639099, rs34200313, rs12821256, rs146898186, rs12298351, rs12820818, rs58696617, rs7299139, rs143993772, rs10859040

233	336	rs61924662	Non-red	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Low tan response, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Blond vs. brown hair colour	rs1564120, rs7299005, rs148137046, rs184079650, rs10858628, rs35466244, rs34404753, rs7487365, rs7966318, rs2216153, rs7977437, rs7306001, rs115183712, rs35618688, rs1907703, rs2639099, rs34200313, rs12821256, rs146898186, rs12298351, rs12820818, rs58696617, rs7299139, rs143993772, rs10859040		
234	347	rs12368980	Non-red	Hair colour	rs10859040		
235	348	rs3213737	CS			Melanoma paper	rs10859996

236	349	rs1265564	Non-red	Hair colour, Blond vs. brown/black hair colour	rs1265564, rs11066284		
236	349	rs73201772	Non-red	Hair colour, Blond vs. brown/black hair colour	rs1265564, rs11066284		
237	350	rs61939687	Ease, Skin, Non-red	Blond vs. brown/black hair colour	rs61939692, rs11068059	Melanoma paper	rs113469387
237	350	rs61939692	Ease, Skin, Non-red	Blond vs. brown/black hair colour	rs61939692, rs11068059	Melanoma paper	rs113469387
237	350	rs61935849	Ease, Skin, Non-red	Blond vs. brown/black hair colour	rs61939692, rs11068059	Melanoma paper	rs113469387
237	350	rs61937364	Ease, Skin, Non-red	Blond vs. brown/black hair colour	rs61939692, rs11068059	Melanoma paper	rs113469387
237	350	rs113159861	Ease, Skin, Non-red	Blond vs. brown/black hair colour	rs61939692, rs11068059	Melanoma paper	rs113469387
237	351	rs6490074	Ease, Skin, Non-red	Blond vs. brown/black hair colour	rs61939692, rs11068059	Melanoma paper	rs113469387
238	352	rs1716169	Non-red	Hair colour	rs28576953		
239	354	rs77717551	Ease, Skin, Non-red	Hair colour	rs77717551		

239	353	rs68099344	Ease, Skin, Non-red	Hair colour rs77717551
239	353	rs7304293	Ease, Skin, Non-red	Hair colour rs77717551
240	355	rs4760521	Ease	
241	356	rs9603422	Non-red	Hair colour rs9603422
242	357	rs765377	Skin, Non-red, Red	Hair colour rs188310489, rs2254690, rs7998826, rs1279403
242	357	rs1823554	Skin, Non-red, Red	Hair colour rs188310489, rs2254690, rs7998826, rs1279403
242	357	rs1766344	Skin, Non-red, Red	Hair colour rs188310489, rs2254690, rs7998826, rs1279403
242	357	rs1279403	Skin, Non-red, Red	Hair colour rs188310489, rs2254690, rs7998826, rs1279403
242	358	rs9544609	Skin, Non-red, Red	Hair colour rs188310489, rs2254690, rs7998826, rs1279403
243	359	rs9590030	Ease, Skin, Non-red	Hair colour rs717769

243	359	rs7337275	Ease, Skin, Non-red	Hair colour	rs717769		
243	359	rs6492711	Ease, Skin, Non-red	Hair colour	rs717769		
244	360	rs9515244	Non-red	Hair colour	rs2025905		
245	361	rs1278761	Ease, Skin, CS	Skin pigmentation	rs3024737, rs7326155	Melanoma paper	rs1278768
245	362	rs1765762	Ease, Skin, CS	Skin pigmentation	rs3024737, rs7326155	Melanoma paper	rs1278768
245	363	rs1278772	Ease, Skin, CS	Skin pigmentation	rs3024737, rs7326155	Melanoma paper	rs1278768
245	361	rs1278774	Ease, Skin, CS	Skin pigmentation	rs3024737, rs7326155	Melanoma paper	rs1278768
246	364	rs7999377	Non-red				
247	365	rs72681869	Skin, Non-red	Hair colour	rs72683923		
248	366	rs210381	Non-red	Hair colour, Blond vs. brown/black hair colour	rs210381		

248	367	rs56412931	Non-red	Hair colour, Blond vs. brown/black hair colour	rs210381		
249	368	rs10134019	Pigscore 2				
250	369	rs1951116	Non-red	Hair colour	rs34813910		
251	370	rs10873172	Non-red	Hair colour, Blond vs. brown/black hair colour	rs10873172	Nevus count or cutaneous melanoma	rs2357176
252	371	rs28649231	Non-red	Blond vs. brown/black hair colour, Hair colour	rs11158717, rs1980655, rs10139386, rs72731537, rs56232028		
253	372	rs12148044	Non-red	Blond vs. brown/black hair colour, Hair colour	rs72731537, rs56232028		

254	373	rs113005382	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Hair morphology traits, Blond vs. brown hair colour, Black vs. blond hair colour, Low tan response, Brown vs. black hair colour, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Skin pigmentation	rs536796893, rs75433889, rs6575242, rs12896399, rs941799, rs1885194, rs17184180, rs12883151, rs8014907, rs11621581, rs4904886, rs10133804, rs75263396, rs17783630, rs117324075
254	374	rs61977801	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Hair morphology traits, Blond vs. brown hair colour, Black vs. blond hair colour, Low tan response, Brown vs. black hair colour, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Skin pigmentation	rs536796893, rs75433889, rs6575242, rs12896399, rs941799, rs1885194, rs17184180, rs12883151, rs8014907, rs11621581, rs4904886, rs10133804, rs75263396, rs17783630, rs117324075

254	374	rs746586	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Hair morphology traits, Blond vs. brown hair colour, Black vs. blond hair colour, Low tan response, Brown vs. black hair colour, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Skin pigmentation	rs536796893, rs75433889, rs6575242, rs12896399, rs941799, rs1885194, rs17184180, rs12883151, rs8014907, rs11621581, rs4904886, rs10133804, rs75263396, rs17783630, rs117324075
254	374	rs941799	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Hair morphology traits, Blond vs. brown hair colour, Black vs. blond hair colour, Low tan response, Brown vs. black hair colour, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Skin pigmentation	rs536796893, rs75433889, rs6575242, rs12896399, rs941799, rs1885194, rs17184180, rs12883151, rs8014907, rs11621581, rs4904886, rs10133804, rs75263396, rs17783630, rs117324075

254	374	rs1885194	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Hair morphology traits, Blond vs. brown hair colour, Black vs. blond hair colour, Low tan response, Brown vs. black hair colour, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Skin pigmentation	rs536796893, rs75433889, rs6575242, rs12896399, rs941799, rs1885194, rs17184180, rs12883151, rs8014907, rs11621581, rs4904886, rs10133804, rs75263396, rs17783630, rs117324075
254	374	rs76519749	Ease, Skin, Non-red	Hair colour, Blond vs. brown/black hair colour, Hair morphology traits, Blond vs. brown hair colour, Black vs. blond hair colour, Low tan response, Brown vs. black hair colour, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Skin pigmentation	rs536796893, rs75433889, rs6575242, rs12896399, rs941799, rs1885194, rs17184180, rs12883151, rs8014907, rs11621581, rs4904886, rs10133804, rs75263396, rs17783630, rs117324075
255	375	rs3759579	Non-red, CS	Hair colour, Blond vs. brown/black hair colour	rs3825566, rs55859054, rs1744297

255	375	rs879552	Non-red, CS	Hair colour, Blond vs. brown/black hair colour	rs3825566, rs1744297	rs55859054,
255	375	rs7359090	Non-red, CS	Hair colour, Blond vs. brown/black hair colour	rs3825566, rs1744297	rs55859054,

256	387	rs146704333	Ease, Skin, Non-red, Red, CS	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Skin, hair and eye pigmentation (multivariate alysis), Skin pigmentation traits, Low tan response, Black vs. blond hair colour, Black vs. red hair colour, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Red vs non-red hair colour, Hair morphology traits, Red vs. brown/black hair colour, Tanning, Perceived skin darkness, Blond vs. brown hair colour, Skin colour saturation	rs1077037, rs139029488, rs75236462, rs8029469, rs924318, rs121918167, rs1448489, rs2703954, rs7169225, rs72625132, rs121918170, rs74653330, rs121918166, rs1800407, rs1800404, rs4778224, rs749846, rs116978932, rs11855019, rs7495174, rs1129038, rs12913832, rs117007668, rs75165924, rs4778249, rs144864205, rs79097182, rs1667394, rs1667392, rs77572354, rs8033165, rs56166703	Uveal melanoma, Cutaneous malignt melanoma, melanoma, Melanoma paper	rs11074306, rs4778138, rs141514981, rs12913832

256	384	rs35690998	Ease, Skin, Non-red, Red, CS	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Skin, hair and eye pigmentation (multivariate alysis), Skin pigmentation traits, Low tan response, Black vs. blond hair colour, Black vs. red hair colour, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Red vs non-red hair colour, Hair morphology traits, Red vs. brown/black hair colour, Tanning, Perceived skin darkness, Blond vs. brown hair colour, Skin colour saturation	rs1077037, rs139029488, rs75236462, rs8029469, rs924318, rs121918167, rs1448489, rs2703954, rs7169225, rs72625132, rs121918170, rs74653330, rs121918166, rs1800407, rs1800404, rs4778224, rs749846, rs116978932, rs11855019, rs7495174, rs1129038, rs12913832, rs117007668, rs75165924, rs4778249, rs144864205, rs79097182, rs1667394, rs1667392, rs77572354, rs8033165, rs56166703	Uveal melanoma, Cutaneous malignt melanoma, melanoma, Melanoma paper	rs11074306, rs4778138, rs141514981, rs12913832

256	378	rs78114576	Ease, Skin, Non-red, Red, CS	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Skin, hair and eye pigmentation (multivariate alysis), Skin pigmentation traits, Low tan response, Black vs. blond hair colour, Black vs. red hair colour, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Red vs non-red hair colour, Hair morphology traits, Red vs. brown/black hair colour, Tanning, Perceived skin darkness, Blond vs. brown hair colour, Skin colour saturation	rs1077037, rs139029488, rs75236462, rs8029469, rs924318, rs121918167, rs1448489, rs2703954, rs7169225, rs72625132, rs121918170, rs74653330, rs121918166, rs1800407, rs1800404, rs4778224, rs749846, rs116978932, rs11855019, rs7495174, rs1129038, rs12913832, rs117007668, rs75165924, rs4778249, rs144864205, rs79097182, rs1667394, rs1667392, rs77572354, rs8033165, rs56166703	Uveal melanoma, Cutaneous malignt melanoma, melanoma, Melanoma paper	rs11074306, rs4778138, rs141514981, rs12913832

256	387	rs142306097	Ease, Skin, Non-red, Red, CS	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Skin, hair and eye pigmentation (multivariate alysis), Skin pigmentation traits, Low tan response, Black vs. blond hair colour, Black vs. red hair colour, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Red vs non-red hair colour, Hair morphology traits, Red vs. brown/black hair colour, Tanning, Perceived skin darkness, Blond vs. brown hair colour, Skin colour saturation	rs1077037, rs139029488, rs75236462, rs8029469, rs924318, rs121918167, rs1448489, rs2703954, rs7169225, rs72625132, rs121918170, rs74653330, rs121918166, rs1800407, rs1800404, rs4778224, rs749846, rs116978932, rs11855019, rs7495174, rs1129038, rs12913832, rs117007668, rs75165924, rs4778249, rs144864205, rs79097182, rs1667394, rs1667392, rs77572354, rs8033165, rs56166703	Uveal melanoma, Cutaneous malignt melanoma, melanoma, Melanoma paper	rs11074306, rs4778138, rs141514981, rs12913832

256	380	rs75137181	Ease, Skin, Non-red, Red, CS	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Skin, hair and eye pigmentation (multivariate alysis), Skin pigmentation traits, Low tan response, Black vs. blond hair colour, Black vs. red hair colour, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Red vs non-red hair colour, Hair morphology traits, Red vs. brown/black hair colour, Tanning, Perceived skin darkness, Blond vs. brown hair colour, Skin colour saturation	rs1077037, rs139029488, rs75236462, rs8029469, rs924318, rs121918167, rs1448489, rs2703954, rs7169225, rs72625132, rs121918170, rs74653330, rs121918166, rs1800407, rs1800404, rs4778224, rs749846, rs116978932, rs11855019, rs7495174, rs1129038, rs12913832, rs117007668, rs75165924, rs4778249, rs144864205, rs79097182, rs1667394, rs1667392, rs77572354, rs8033165, rs56166703	Uveal melanoma, Cutaneous malignt melanoma, melanoma, Melanoma paper	rs11074306, rs4778138, rs141514981, rs12913832

256	383	rs35591793	Ease, Skin, Non-red, Red, CS	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Skin, hair and eye pigmentation (multivariate alysis), Skin pigmentation traits, Low tan response, Black vs. blond hair colour, Black vs. red hair colour, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Red vs non-red hair colour, Hair morphology traits, Red vs. brown/black hair colour, Tanning, Perceived skin darkness, Blond vs. brown hair colour, Skin colour saturation	rs1077037, rs139029488, rs75236462, rs8029469, rs924318, rs121918167, rs1448489, rs2703954, rs7169225, rs72625132, rs121918170, rs74653330, rs121918166, rs1800407, rs1800404, rs4778224, rs749846, rs116978932, rs11855019, rs7495174, rs1129038, rs12913832, rs117007668, rs75165924, rs4778249, rs144864205, rs79097182, rs1667394, rs1667392, rs77572354, rs8033165, rs56166703	Uveal melanoma, Cutaneous malignt melanoma, melanoma, Melanoma paper	rs11074306, rs4778138, rs141514981, rs12913832

256	376	rs139615925	Ease, Skin, Non-red, Red, CS	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Skin, hair and eye pigmentation (multivariate alysis), Skin pigmentation traits, Low tan response, Black vs. blond hair colour, Black vs. red hair colour, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Red vs non-red hair colour, Hair morphology traits, Red vs. brown/black hair colour, Tanning, Perceived skin darkness, Blond vs. brown hair colour, Skin colour saturation	rs1077037, rs139029488, rs75236462, rs8029469, rs924318, rs121918167, rs1448489, rs2703954, rs7169225, rs72625132, rs121918170, rs74653330, rs121918166, rs1800407, rs1800404, rs4778224, rs749846, rs116978932, rs11855019, rs7495174, rs1129038, rs12913832, rs117007668, rs75165924, rs4778249, rs144864205, rs79097182, rs1667394, rs1667392, rs77572354, rs8033165, rs56166703	Uveal melanoma, Cutaneous malignt melanoma, melanoma, Melanoma paper	rs11074306, rs4778138, rs141514981, rs12913832

256	376	rs12909167	Ease, Skin, Non-red, Red, CS	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Skin, hair and eye pigmentation (multivariate alysis), Skin pigmentation traits, Low tan response, Black vs. blond hair colour, Black vs. red hair colour, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Red vs non-red hair colour, Hair morphology traits, Red vs. brown/black hair colour, Tanning, Perceived skin darkness, Blond vs. brown hair colour, Skin colour saturation	rs1077037, rs139029488, rs75236462, rs8029469, rs924318, rs121918167, rs1448489, rs2703954, rs7169225, rs72625132, rs121918170, rs74653330, rs121918166, rs1800407, rs1800404, rs4778224, rs749846, rs116978932, rs11855019, rs7495174, rs1129038, rs12913832, rs117007668, rs75165924, rs4778249, rs144864205, rs79097182, rs1667394, rs1667392, rs77572354, rs8033165, rs56166703	Uveal melanoma, Cutaneous malignt melanoma, melanoma, Melanoma paper	rs11074306, rs4778138, rs141514981, rs12913832

256	382	rs72710600	Ease, Skin, Non-red, Red, CS	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Skin, hair and eye pigmentation (multivariate alysis), Skin pigmentation traits, Low tan response, Black vs. blond hair colour, Black vs. red hair colour, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Red vs non-red hair colour, Hair morphology traits, Red vs. brown/black hair colour, Tanning, Perceived skin darkness, Blond vs. brown hair colour, Skin colour saturation	rs1077037, rs139029488, rs75236462, rs8029469, rs924318, rs121918167, rs1448489, rs2703954, rs7169225, rs72625132, rs121918170, rs74653330, rs121918166, rs1800407, rs1800404, rs4778224, rs749846, rs116978932, rs11855019, rs7495174, rs1129038, rs12913832, rs117007668, rs75165924, rs4778249, rs144864205, rs79097182, rs1667394, rs1667392, rs77572354, rs8033165, rs56166703	Uveal melanoma, Cutaneous malignt melanoma, melanoma, Melanoma paper	rs11074306, rs4778138, rs141514981, rs12913832

256	380	rs17567007	Ease, Skin, Non-red, Red, CS	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Skin, hair and eye pigmentation (multivariate alysis), Skin pigmentation traits, Low tan response, Black vs. blond hair colour, Black vs. red hair colour, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Red vs non-red hair colour, Hair morphology traits, Red vs. brown/black hair colour, Tanning, Perceived skin darkness, Blond vs. brown hair colour, Skin colour saturation	rs1077037, rs139029488, rs75236462, rs8029469, rs924318, rs121918167, rs1448489, rs2703954, rs7169225, rs72625132, rs121918170, rs74653330, rs121918166, rs1800407, rs1800404, rs4778224, rs749846, rs116978932, rs11855019, rs7495174, rs1129038, rs12913832, rs117007668, rs75165924, rs4778249, rs144864205, rs79097182, rs1667394, rs1667392, rs77572354, rs8033165, rs56166703	Uveal melanoma, Cutaneous malignt melanoma, melanoma, Melanoma paper	rs11074306, rs4778138, rs141514981, rs12913832

256	386	rs141514981	Ease, Skin, Non-red, Red, CS	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Skin, hair and eye pigmentation (multivariate alysis), Skin pigmentation traits, Low tan response, Black vs. blond hair colour, Black vs. red hair colour, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Red vs non-red hair colour, Hair morphology traits, Red vs. brown/black hair colour, Tanning, Perceived skin darkness, Blond vs. brown hair colour, Skin colour saturation	rs1077037, rs139029488, rs75236462, rs8029469, rs924318, rs121918167, rs1448489, rs2703954, rs7169225, rs72625132, rs121918170, rs74653330, rs121918166, rs1800407, rs1800404, rs4778224, rs749846, rs116978932, rs11855019, rs7495174, rs1129038, rs12913832, rs117007668, rs75165924, rs4778249, rs144864205, rs79097182, rs1667394, rs1667392, rs77572354, rs8033165, rs56166703	Uveal melanoma, Cutaneous malignt melanoma, melanoma, Melanoma paper	rs11074306, rs4778138, rs141514981, rs12913832

256	377	rs1800407	Ease, Skin, Non-red, Red, CS	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Skin, hair and eye pigmentation (multivariate alysis), Skin pigmentation traits, Low tan response, Black vs. blond hair colour, Black vs. red hair colour, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Red vs non-red hair colour, Hair morphology traits, Red vs. brown/black hair colour, Tanning, Perceived skin darkness, Blond vs. brown hair colour, Skin colour saturation	rs1077037, rs139029488, rs75236462, rs8029469, rs924318, rs121918167, rs1448489, rs2703954, rs7169225, rs72625132, rs121918170, rs74653330, rs121918166, rs1800407, rs1800404, rs4778224, rs749846, rs116978932, rs11855019, rs7495174, rs1129038, rs12913832, rs117007668, rs75165924, rs4778249, rs144864205, rs79097182, rs1667394, rs1667392, rs77572354, rs8033165, rs56166703	Uveal melanoma, Cutaneous malignt melanoma, melanoma, Melanoma paper	rs11074306, rs4778138, rs141514981, rs12913832

256	377	rs117886461	Ease, Skin, Non-red, Red, CS	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Skin, hair and eye pigmentation (multivariate alysis), Skin pigmentation traits, Low tan response, Black vs. blond hair colour, Black vs. red hair colour, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Red vs non-red hair colour, Hair morphology traits, Red vs. brown/black hair colour, Tanning, Perceived skin darkness, Blond vs. brown hair colour, Skin colour saturation	rs1077037, rs139029488, rs75236462, rs8029469, rs924318, rs121918167, rs1448489, rs2703954, rs7169225, rs72625132, rs121918170, rs74653330, rs121918166, rs1800407, rs1800404, rs4778224, rs749846, rs116978932, rs11855019, rs7495174, rs1129038, rs12913832, rs117007668, rs75165924, rs4778249, rs144864205, rs79097182, rs1667394, rs1667392, rs77572354, rs8033165, rs56166703	Uveal melanoma, Cutaneous malignt melanoma, melanoma, Melanoma paper	rs11074306, rs4778138, rs141514981, rs12913832

256	382	rs4778136	Ease, Skin, Non-red, Red, CS	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Skin, hair and eye pigmentation (multivariate alysis), Skin pigmentation traits, Low tan response, Black vs. blond hair colour, Black vs. red hair colour, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Red vs non-red hair colour, Hair morphology traits, Red vs. brown/black hair colour, Tanning, Perceived skin darkness, Blond vs. brown hair colour, Skin colour saturation	rs1077037, rs139029488, rs75236462, rs8029469, rs924318, rs121918167, rs1448489, rs2703954, rs7169225, rs72625132, rs121918170, rs74653330, rs121918166, rs1800407, rs1800404, rs4778224, rs749846, rs116978932, rs11855019, rs7495174, rs1129038, rs12913832, rs117007668, rs75165924, rs4778249, rs144864205, rs79097182, rs1667394, rs1667392, rs77572354, rs8033165, rs56166703	Uveal melanoma, Cutaneous malignt melanoma, melanoma, Melanoma paper	rs11074306, rs4778138, rs141514981, rs12913832

256	386	rs77516282	Ease, Skin, Non-red, Red, CS	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Skin, hair and eye pigmentation (multivariate alysis), Skin pigmentation traits, Low tan response, Black vs. blond hair colour, Black vs. red hair colour, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Red vs non-red hair colour, Hair morphology traits, Red vs. brown/black hair colour, Tanning, Perceived skin darkness, Blond vs. brown hair colour, Skin colour saturation	rs1077037, rs139029488, rs75236462, rs8029469, rs924318, rs121918167, rs1448489, rs2703954, rs7169225, rs72625132, rs121918170, rs74653330, rs121918166, rs1800407, rs1800404, rs4778224, rs749846, rs116978932, rs11855019, rs7495174, rs1129038, rs12913832, rs117007668, rs75165924, rs4778249, rs144864205, rs79097182, rs1667394, rs1667392, rs77572354, rs8033165, rs56166703	Uveal melanoma, Cutaneous malignt melanoma, melanoma, Melanoma paper	rs11074306, rs4778138, rs141514981, rs12913832

256	381	rs746861	Ease, Skin, Non-red, Red, CS	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Skin, hair and eye pigmentation (multivariate alysis), Skin pigmentation traits, Low tan response, Black vs. blond hair colour, Black vs. red hair colour, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Red vs non-red hair colour, Hair morphology traits, Red vs. brown/black hair colour, Tanning, Perceived skin darkness, Blond vs. brown hair colour, Skin colour saturation	rs1077037, rs139029488, rs75236462, rs8029469, rs924318, rs121918167, rs1448489, rs2703954, rs7169225, rs72625132, rs121918170, rs74653330, rs121918166, rs1800407, rs1800404, rs4778224, rs749846, rs116978932, rs11855019, rs7495174, rs1129038, rs12913832, rs117007668, rs75165924, rs4778249, rs144864205, rs79097182, rs1667394, rs1667392, rs77572354, rs8033165, rs56166703	Uveal melanoma, Cutaneous malignt melanoma, melanoma, Melanoma paper	rs11074306, rs4778138, rs141514981, rs12913832

256	385	rs2122005	Ease, Skin, Non-red, Red, CS	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Skin, hair and eye pigmentation (multivariate alysis), Skin pigmentation traits, Low tan response, Black vs. blond hair colour, Black vs. red hair colour, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Red vs non-red hair colour, Hair morphology traits, Red vs. brown/black hair colour, Tanning, Perceived skin darkness, Blond vs. brown hair colour, Skin colour saturation	rs1077037, rs139029488, rs75236462, rs8029469, rs924318, rs121918167, rs1448489, rs2703954, rs7169225, rs72625132, rs121918170, rs74653330, rs121918166, rs1800407, rs1800404, rs4778224, rs749846, rs116978932, rs11855019, rs7495174, rs1129038, rs12913832, rs117007668, rs75165924, rs4778249, rs144864205, rs79097182, rs1667394, rs1667392, rs77572354, rs8033165, rs56166703	Uveal melanoma, Cutaneous malignt melanoma, melanoma, Melanoma paper	rs11074306, rs4778138, rs141514981, rs12913832

256	379	rs56108911	Ease, Skin, Non-red, Red, CS	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Skin, hair and eye pigmentation (multivariate alysis), Skin pigmentation traits, Low tan response, Black vs. blond hair colour, Black vs. red hair colour, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Red vs non-red hair colour, Hair morphology traits, Red vs. brown/black hair colour, Tanning, Perceived skin darkness, Blond vs. brown hair colour, Skin colour saturation	rs1077037, rs139029488, rs75236462, rs8029469, rs924318, rs121918167, rs1448489, rs2703954, rs7169225, rs72625132, rs121918170, rs74653330, rs121918166, rs1800407, rs1800404, rs4778224, rs749846, rs116978932, rs11855019, rs7495174, rs1129038, rs12913832, rs117007668, rs75165924, rs4778249, rs144864205, rs79097182, rs1667394, rs1667392, rs77572354, rs8033165, rs56166703	Uveal melanoma, Cutaneous malignt melanoma, melanoma, Melanoma paper	rs11074306, rs4778138, rs141514981, rs12913832

256	382	rs1129038	Ease, Skin, Non-red, Red, CS	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Skin, hair and eye pigmentation (multivariate alysis), Skin pigmentation traits, Low tan response, Black vs. blond hair colour, Black vs. red hair colour, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Red vs non-red hair colour, Hair morphology traits, Red vs. brown/black hair colour, Tanning, Perceived skin darkness, Blond vs. brown hair colour, Skin colour saturation	rs1077037, rs139029488, rs75236462, rs8029469, rs924318, rs121918167, rs1448489, rs2703954, rs7169225, rs72625132, rs121918170, rs74653330, rs121918166, rs1800407, rs1800404, rs4778224, rs749846, rs116978932, rs11855019, rs7495174, rs1129038, rs12913832, rs117007668, rs75165924, rs4778249, rs144864205, rs79097182, rs1667394, rs1667392, rs77572354, rs8033165, rs56166703	Uveal melanoma, Cutaneous malignt melanoma, melanoma, Melanoma paper	rs11074306, rs4778138, rs141514981, rs12913832

256	378	rs12913832	Ease, Skin, Non-red, Red, CS	Hair colour, Blond vs. brown/black hair colour, Brown vs. black hair colour, Skin, hair and eye pigmentation (multivariate alysis), Skin pigmentation traits, Low tan response, Black vs. blond hair colour, Black vs. red hair colour, Blond vs non-blond hair colour, Brown vs. non-brown hair colour, Light vs. dark hair colour, Red vs non-red hair colour, Hair morphology traits, Red vs. brown/black hair colour, Tanning, Perceived skin darkness, Blond vs. brown hair colour, Skin colour saturation	rs1077037, rs139029488, rs75236462, rs8029469, rs924318, rs121918167, rs1448489, rs2703954, rs7169225, rs72625132, rs121918170, rs74653330, rs121918166, rs1800407, rs1800404, rs4778224, rs749846, rs116978932, rs11855019, rs7495174, rs1129038, rs12913832, rs117007668, rs75165924, rs4778249, rs144864205, rs79097182, rs1667394, rs1667392, rs77572354, rs8033165, rs56166703	Uveal melanoma, Cutaneous malignt melanoma, melanoma, Melanoma paper	rs11074306, rs4778138, rs141514981, rs12913832
257	388	rs17228129	Ease, Skin				
257	388	rs28441327	Ease, Skin				

257	388	rs28456199	Ease, Skin		
257	388	rs4779826	Ease, Skin		
258	389	rs143692029	Non-red	Hair colour	rs143692029
259	390	rs11639111	Pigscore 2	Skin pigmentation	rs2899446, rs8033655
260	391	rs57249121	Skin	Skin pigmentation	rs2899446, rs8033655
260	391	rs11854435	Skin	Skin pigmentation	rs2899446, rs8033655
261	392	rs28653088	Non-red		
262	393	rs72744156	Skin, Non-red	Blond vs. brown/black hair colour, Hair colour, Brown vs. black hair colour	rs67093094, rs7179134, rs28648707, rs61219147
262	393	rs12324786	Skin, Non-red	Blond vs. brown/black hair colour, Hair colour, Brown vs. black hair colour	rs67093094, rs7179134, rs28648707, rs61219147
263	394	rs2455717	Ease, Skin	Hair colour, Blond vs. brown/black hair colour	rs8042597, rs8033380
263	394	rs4779035	Ease, Skin	Hair colour, Blond vs. brown/black hair colour	rs8042597, rs8033380
263	394	rs28371837	Ease, Skin	Hair colour, Blond vs. brown/black hair colour	rs8042597, rs8033380

264	395	rs1877024	Non-red				
265	396	rs8028550	Non-red	Hair colour	rs55767890		
265	396	rs17554929	Non-red	Hair colour	rs55767890		
266	398	rs34560261	Skin, Non-red	Hair colour	rs4932373		
266	397	rs12901150	Skin, Non-red	Hair colour	rs4932373		
267	399	rs2235578	Pigscore 2				
268	400	rs758045	Skin	Hair colour	rs4424915		
268	400	rs11639998	Skin	Hair colour	rs4424915		
269	401	rs7186212	Ease, CS				
269	401	rs6498485	Ease, CS				
269	401	rs1799798	Ease, CS				
269	401	rs3974346	Ease, CS				
270	402	rs76536391	Non-red	Hair colour	rs16953002	Melanoma, Nevus count or cutaneous melanoma	rs16953002, rs12596638

271	403	rs1801257	Non-red	
272	404	rs4788815	Ease, Skin, Hair colour Non-red	rs2008675
272	404	rs12935795	Ease, Skin, Hair colour Non-red	rs2008675
272	404	rs8054175	Ease, Skin, Hair colour Non-red	rs2008675
272	404	rs12708925	Ease, Skin, Hair colour Non-red	rs2008675
273	405	rs4238968	Ease	
273	405	rs6499612	Ease	
274	406	rs8044524	Pigscore 2	

275	413	rs9940427	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	408	rs112787727	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	408	rs117614496	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	410	rs116004099	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	425	rs75898184	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	427	rs117791738	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	424	rs9921065	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	408	rs113894462	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	426	rs112584165	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	409	rs72815590	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	434	rs12925392	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	433	rs149610032	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	411	rs150020387	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	418	rs117984432	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	430	rs9923267	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	425	rs191414071	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	415	rs116918202	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	420	rs143743429	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	423	rs4785577	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	424	rs59013041	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	432	rs369230	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	420	rs144988021	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	417	rs164737	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	416	rs139031001	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	417	rs77651727	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	431	rs11648089	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	422	rs61482648	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	431	rs35749174	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	418	rs117030214	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	409	rs11641963	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	408	rs72807518	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	420	rs56288641	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	420	rs79139787	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	422	rs371629999	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	426	rs75092208	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	409	rs12923751	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	422	rs34427386	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	428	rs34357723	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	416	rs72811597	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	420	rs7185897	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	421	rs141817469	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	425	rs73276549	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	424	rs8059075	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	422	rs11641201	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	418	rs1805005	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	422	rs1805006	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	414	rs1805007	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	430	rs1805008	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	429	rs1805009	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	408	rs142969375	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	408	rs150909008	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	422	rs117732683	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	409	rs11859970	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	430	rs7195043	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	431	rs77733403	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	412	rs117711828	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	419	rs137938320	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

275	407	rs9302777	Ease, Skin, Non-red, Red, CS	Hair colour, Low tan response, Hair morphology traits, Tanning, Skin pigmentation traits, Blond vs. brown/black hair colour, Brown vs. black hair colour, Black vs. red hair colour, Red vs. brown/black hair colour, Freckling, Blond vs. brown hair colour, Red vs non-red hair colour, Freckles, Non-melanoma skin cancer, Perceived skin darkness, Light vs. dark hair colour, Brown vs. non-brown hair colour, Skin colour saturation	rs142702217, rs188255148, rs883284, rs117463496, rs4782309, rs75410747, rs55924855, rs73252814, rs112584165, rs11862309, rs117156175, rs4785645, rs3114908, rs117646290, rs2965812, rs55637757, rs67689854, rs369230, rs154659, rs12930346, rs164745, rs11648089, rs35063026, rs116927526, rs258322, rs35432452, rs57119673, rs1800347, rs3743861, rs12931267, rs139414522, rs1006548, rs2238529, rs36233537, rs113271242, rs12932219, rs182948919, rs9939914,	Melanoma, Melanoma paper, Nevus count or cutaneous melanoma	rs258322, rs116927526, rs75570604, rs1805007, rs1805008, rs1805009, rs3212371, rs4785763

276	435	rs2746026	Non-red	Hair colour	rs68165736
277	436	rs9895999	Pigscore 2		
278	437	rs117827273	Non-red		
278	437	rs2314338	Non-red		
279	438	rs72828120	Non-red	Hair colour	rs117612447, rs117827273
279	438	rs117612447	Non-red	Hair colour	rs117612447, rs117827273
279	439	rs897418	Non-red	Hair colour	rs117612447, rs117827273
280	440	rs77471755	Pigscore 3		
281	441	rs11079764	Non-red		
282	442	rs2325752	Skin, Non-red	Hair colour	rs72833466, rs3874943
282	444	rs72833470	Skin, Non-red	Hair colour	rs72833466, rs3874943
282	442	rs7219107	Skin, Non-red	Hair colour	rs72833466, rs3874943

282	443	rs1962201	Skin, Non-red	Hair colour	rs72833466, rs3874943
282	442	rs16949720	Skin, Non-red	Hair colour	rs72833466, rs3874943
283	446	rs79571620	Skin, Non-red	Hair colour, Brown vs. black hair colour, Blond vs. brown/black hair colour	rs72833466, rs72833470, rs16949418, rs192123848, rs2957933, rs140756892
283	447	rs9303554	Skin, Non-red	Hair colour, Brown vs. black hair colour, Blond vs. brown/black hair colour	rs72833466, rs72833470, rs16949418, rs192123848, rs2957933, rs140756892
283	447	rs55788912	Skin, Non-red	Hair colour, Brown vs. black hair colour, Blond vs. brown/black hair colour	rs72833466, rs72833470, rs16949418, rs192123848, rs2957933, rs140756892

283	445	rs739992	Skin, Non-red	Hair colour, Brown vs. black hair colour, Blond vs. brown/black hair colour	rs72833466, rs72833470, rs16949418, rs192123848, rs2957933, rs140756892
283	445	rs739990	Skin, Non-red	Hair colour, Brown vs. black hair colour, Blond vs. brown/black hair colour	rs72833466, rs72833470, rs16949418, rs192123848, rs2957933, rs140756892
284	448	rs17833789	Non-red	Hair colour	rs62060349
284	448	rs62060349	Non-red	Hair colour	rs62060349
285	449	rs1133683	Non-red		
286	450	rs149789627	Skin		
287	452	rs56931525	Ease, Skin, Non-red, Red	Hair colour	rs34872037, rs7405453
287	453	rs34635363	Ease, Skin, Non-red, Red	Hair colour	rs34872037, rs7405453

287	451	rs71373084	Ease, Skin, Non-red, Red	Hair colour	rs34872037, rs7405453
287	452	rs9747347	Ease, Skin, Non-red, Red	Hair colour	rs34872037, rs7405453
287	452	rs6420484	Ease, Skin, Non-red, Red	Hair colour	rs34872037, rs7405453
287	452	rs34683731	Ease, Skin, Non-red, Red	Hair colour	rs34872037, rs7405453
288	454	rs75596709	Non-red		
289	455	rs7236870	Non-red		
290	456	rs9964220	Non-red		
291	457	rs4939832	Pigscore 2	Hair colour	rs9954676
292	458	rs12954491	Non-red		
293	459	rs12609746	Non-red		

294	461	rs77733715	Ease, Red	Nevus count or cutaneous melanoma, Melanoma paper	rs34466956, rs12984831
294	462	rs34878396	Ease, Red	Nevus count or cutaneous melanoma, Melanoma paper	rs34466956, rs12984831
294	460	rs8103610	Ease, Red	Nevus count or cutaneous melanoma, Melanoma paper	rs34466956, rs12984831

294	462	rs6510760	Ease, Red			Nevus count or cutaneous melanoma, Melanoma paper	rs34466956, rs12984831
295	463	rs350818	Non-red	Hair colour	rs28827913		
295	464	rs73234	Non-red	Hair colour	rs28827913		
296	465	rs2161568	Pigscore 1	Hair colour	rs73488486		
297	466	rs2163832	Non-red	Hair colour	rs8102380, rs11085749		
297	466	rs4804519	Non-red	Hair colour	rs8102380, rs11085749		
298	467	rs4808779	Pigscore 2	Hair colour	rs7251640		
299	468	rs16964556	Non-red				
300	469	rs1865048	Skin				
301	470	rs6051776	Ease				
301	470	rs151525	Ease				
302	471	rs2249048	Ease, Skin				
302	471	rs6137358	Ease, Skin				
302	471	rs11697152	Ease, Skin				

303	472	rs6113507	Non-red	Hair colour	rs6113507		
303	472	rs6113632	Non-red	Hair colour	rs6113507		
304	476	rs200948404	Ease, Skin, Non-red, Red, CS	Low tan response, Hair colour, Blond vs. brown/black hair colour, Red vs. brown/black hair colour, Skin colour saturation, Skin pigmentation, Tanning	rs77476644, rs80188179, rs1742982, rs17123518, rs6059655, rs55768819, rs4911268, rs71349703, rs117186940, rs6087480, rs291671, rs17401449, rs78382342, rs149389676, rs142564732, rs6142102, rs4911414, rs138420028	Melanoma paper, Melanoma	rs6059655

304	474	rs6059655	Ease, Skin, Non-red, Red, CS	Low tan response, Hair colour, Blond vs. brown/black hair colour, Red vs. brown/black hair colour, Skin colour saturation, Skin pigmentation, Tanning	rs77476644, rs80188179, rs1742982, rs17123518, rs6059655, rs55768819, rs4911268, rs71349703, rs117186940, rs6087480, rs291671, rs17401449, rs78382342, rs149389676, rs142564732, rs6142102, rs4911414, rs138420028	Melanoma paper, Melanoma	rs6059655
304	475	rs138420028	Ease, Skin, Non-red, Red, CS	Low tan response, Hair colour, Blond vs. brown/black hair colour, Red vs. brown/black hair colour, Skin colour saturation, Skin pigmentation, Tanning	rs77476644, rs80188179, rs1742982, rs17123518, rs6059655, rs55768819, rs4911268, rs71349703, rs117186940, rs6087480, rs291671, rs17401449, rs78382342, rs149389676, rs142564732, rs6142102, rs4911414, rs138420028	Melanoma paper, Melanoma	rs6059655

304	473	rs6087561	Ease, Skin, Non-red, Red, CS	Low tan response, Hair colour, Blond vs. brown/black hair colour, Red vs. brown/black hair colour, Skin colour saturation, Skin pigmentation, Tanning	rs77476644, rs80188179, rs1742982, rs17123518, rs6059655, rs55768819, rs4911268, rs71349703, rs117186940, rs6087480, rs291671, rs17401449, rs78382342, rs149389676, rs142564732, rs6142102, rs4911414, rs138420028	Melanoma paper, Melanoma	rs6059655
304	474	rs6120568	Ease, Skin, Non-red, Red, CS	Low tan response, Hair colour, Blond vs. brown/black hair colour, Red vs. brown/black hair colour, Skin colour saturation, Skin pigmentation, Tanning	rs77476644, rs80188179, rs1742982, rs17123518, rs6059655, rs55768819, rs4911268, rs71349703, rs117186940, rs6087480, rs291671, rs17401449, rs78382342, rs149389676, rs142564732, rs6142102, rs4911414, rs138420028	Melanoma paper, Melanoma	rs6059655

304	474	rs13042880	Ease, Skin, Non-red, Red, CS	Low tan response, Hair colour, Blond vs. brown/black hair colour, Red vs. brown/black hair colour, Skin colour saturation, Skin pigmentation, Tanning	rs77476644, rs80188179, rs1742982, rs17123518, rs6059655, rs55768819, rs4911268, rs71349703, rs117186940, rs6087480, rs291671, rs17401449, rs78382342, rs149389676, rs142564732, rs6142102, rs4911414, rs138420028	Melanoma paper, Melanoma	rs6059655
304	478	rs137942933	Ease, Skin, Non-red, Red, CS	Low tan response, Hair colour, Blond vs. brown/black hair colour, Red vs. brown/black hair colour, Skin colour saturation, Skin pigmentation, Tanning	rs77476644, rs80188179, rs1742982, rs17123518, rs6059655, rs55768819, rs4911268, rs71349703, rs117186940, rs6087480, rs291671, rs17401449, rs78382342, rs149389676, rs142564732, rs6142102, rs4911414, rs138420028	Melanoma paper, Melanoma	rs6059655

304	473	rs60095211	Ease, Skin, Non-red, Red, CS	Low tan response, Hair colour, Blond vs. brown/black hair colour, Red vs. brown/black hair colour, Skin colour saturation, Skin pigmentation, Tanning	rs77476644, rs80188179, rs1742982, rs17123518, rs6059655, rs55768819, rs4911268, rs71349703, rs117186940, rs6087480, rs291671, rs17401449, rs78382342, rs149389676, rs142564732, rs6142102, rs4911414, rs138420028	Melanoma paper, Melanoma	rs6059655
304	477	rs6088778	Ease, Skin, Non-red, Red, CS	Low tan response, Hair colour, Blond vs. brown/black hair colour, Red vs. brown/black hair colour, Skin colour saturation, Skin pigmentation, Tanning	rs77476644, rs80188179, rs1742982, rs17123518, rs6059655, rs55768819, rs4911268, rs71349703, rs117186940, rs6087480, rs291671, rs17401449, rs78382342, rs149389676, rs142564732, rs6142102, rs4911414, rs138420028	Melanoma paper, Melanoma	rs6059655

304	473	rs6060343	Ease, Skin, Non-red, Red, CS	Low tan response, Hair colour, Blond vs. brown/black hair colour, Red vs. brown/black hair colour, Skin colour saturation, Skin pigmentation, Tanning	rs77476644, rs80188179, rs1742982, rs17123518, rs6059655, rs55768819, rs4911268, rs71349703, rs117186940, rs6087480, rs291671, rs17401449, rs78382342, rs149389676, rs142564732, rs6142102, rs4911414, rs138420028	Melanoma paper, Melanoma	rs6059655
305	480	rs4519594	Ease, Skin, Non-red, Red, CS	Brown vs. black hair colour, Low tan response, Hair colour, Freckling	rs2424995, rs56238684, rs62211613, rs150385410, rs1885120, rs73095129, rs4911466, rs6087685, rs2425025, rs619865, rs62210588, rs6142422, rs112043138	Melanoma, Nevus count or cutaneous melanoma	rs910873, rs56238684, rs1885120

305	479	rs221314	Ease, Skin, Non-red, Red, CS	Brown vs. black hair colour, Low tan response, Hair colour, Freckling	rs2424995, rs56238684, rs62211613, rs150385410, rs1885120, rs73095129, rs4911466, rs6087685, rs2425025, rs619865, rs62210588, rs6142422, rs112043138	Melanoma, Nevus count or cutaneous melanoma	rs910873, rs56238684, rs1885120
306	481	rs6028446	Non-red	Hair colour, Skin pigmentation	rs186905936, rs755107, rs113020874		
307	482	rs73133429	Non-red	Hair colour	rs73132911		
307	482	rs55901013	Non-red	Hair colour	rs73132911		
307	482	rs73132906	Non-red	Hair colour	rs73132911		
308	483	rs6127868	Non-red	Hair colour	rs4811760		
309	484	rs6123874	Non-red	Hair colour	rs6128521		
309	484	rs12481673	Non-red	Hair colour	rs6128521		
309	484	rs3026622	Non-red	Hair colour	rs6128521		
310	485	rs2186299	Non-red	Hair colour	rs2830472, rs60841620		
310	485	rs2829780	Non-red	Hair colour	rs2830472, rs60841620		
311	486	rs229103	Non-red				

312	487	rs68088846	Non-red				
313	488	rs399907	Ease	Blond vs. brown/black hair colour, Hair colour, Red vs. brown/black hair colour	rs73220980, rs672948		
314	490	rs73220980	Skin, Non-red	Hair colour	rs235286, rs478075		
314	489	rs672948	Skin, Non-red	Hair colour	rs235286, rs478075		
315	491	rs165895	Non-red	Hair colour	rs165895		
316	492	rs13056416	Non-red	Hair colour	rs5752866		
317	493	rs137487	Non-red				
318	494	rs4465	Pigscore 2	Hair colour	rs138750		
319	496	rs4821721	Ease, Skin			Nevus count or cutaneous melanoma, melanoma, Melanoma	rs2005974, rs2284063, rs6001027, rs132985, rs132941
						paper	

319	496	rs132941	Ease, Skin	Nevus count or cutaneous melanoma, melanoma, Melanoma paper	rs2005974, rs2284063, rs6001027, rs132985, rs132941
319	496	rs4384	Ease, Skin	Nevus count or cutaneous melanoma, melanoma, Melanoma paper	rs2005974, rs2284063, rs6001027, rs132985, rs132941

319	495	rs34066050	Ease, Skin			Nevus count or cutaneous melanoma, melanoma, Melanoma paper	rs2005974, rs2284063, rs6001027, rs132985, rs132941
320	497	rs11089938	Non-red	Hair colour, Blond vs. brown/black hair colour	rs5757602, rs9611155		
320	497	rs9611155	Non-red	Hair colour, Blond vs. brown/black hair colour	rs5757602, rs9611155		
321	499	rs6007506	Ease, Skin, CS	Low tan response, Blond vs. brown/black hair colour, Hair colour	rs6007506, rs2294196, rs5766576, rs136047	Melanoma paper	rs5766565
321	498	rs5766565	Ease, Skin, CS	Low tan response, Blond vs. brown/black hair colour, Hair colour	rs6007506, rs2294196, rs5766576, rs136047	Melanoma paper	rs5766565

321	498	rs34980158	Ease, Skin, CS	Low tan response, Blond vs. brown/black hair colour, Hair colour	rs6007506, rs2294196, rs5766576, rs136047	Melanoma paper	rs5766565
322	500	rs79966207	Non-red	Hair colour, Blond vs. brown/black hair colour	rs79966207	Melanoma paper	rs79966207

Table E.1: Overlap independent signals for non-red hair colour, red hair, skin colour, tanning ability, childhood sunburns, pigscore 1, pigscore 2, pigscore 3 GWAS

Appendix F

Polygenic risk scores

F.1 Non-red hair PRS table

SNP ID	Chr:Pos	Effect Allele	PRS weighting
rs4918475	10:111771165	G	0.0124889
rs4918614	10:112704740	G	-0.0109508
rs10826545	10:29236395	G	0.0133452
rs17230340	10:32679078	C	-0.0151721
rs12264698	10:35464172	G	0.0163591
rs4433500	10:74068998	G	0.0128012
rs703978	10:80944147	C	0.0308436
rs492335	11:129601467	G	-0.0116104
rs75319234	11:15677913	A	0.0497639
rs7108738	11:15710084	T	-0.0244414
rs1920672	11:15858293	T	-0.011366
rs77535014	11:16189834	G	-0.0429225
rs297343	11:16354653	T	-0.016288
rs151060912	11:16500554	C	0.0447773
rs11023988	11:16586767	G	-0.0228624
rs2049129	11:18305333	T	-0.0138516
rs12420721	11:2237667	G	-0.0103086
rs11029693	11:26769140	G	0.0115429

rs4923447	11:27495259	C	-0.0234002
rs145678014	11:32927778	G	-0.030436
rs11037723	11:43922776	G	-0.0133459
rs66716358	11:44330610	C	0.0153586
rs1902961	11:44451792	G	-0.01347
rs7395496	11:47245803	A	-0.0184257
rs61897795	11:61618169	A	0.0232483
rs9645690	11:62206288	C	0.0220027
rs11227247	11:65422853	A	0.0265732
rs72928978	11:68831364	G	0.14901
rs3829241	11:68855363	G	0.0462023
rs11041426	11:7543519	G	0.0286847
rs1124590	11:77703983	C	0.0130236
rs11237489	11:78130325	G	-0.0452465
rs546225	11:87793317	G	-0.00813378
rs188703892	11:88756831	G	0.0359107
rs2187128	11:88879915	T	0.0462903
rs1042602	11:88911696	C	0.0521867
rs77368047	11:89204911	T	0.0417699
rs73201772	12:111905245	C	0.0159502
rs61937364	12:116683588	C	0.0198713
rs6490074	12:116948892	C	0.0120366
rs1716169	12:123716930	A	0.0127854
rs77717551	12:125368342	G	0.0214483
rs9971729	12:23979791	A	-0.0167111
rs6487406	12:24846522	T	0.0256683
rs11049319	12:28221404	T	-0.0135595
rs3764032	12:4317563	T	-0.052744
rs7979418	12:46774771	A	-0.0126191
rs10875910	12:49402393	G	-0.014156
rs1470321	12:54301044	G	0.0142836
rs11172124	12:57594955	G	0.0152392

rs1147094	12:65147371	T	0.0131905
rs1245811	12:79637453	T	-0.004953
rs10862001	12:80002106	G	0.0105212
rs1818930	12:85391419	G	0.0152219
rs12425342	12:85693252	G	0.0696063
rs73200863	12:88143958	G	0.0663579
rs10858711	12:88707968	C	-0.0912905
rs1798011	12:88956625	T	0.0394517
rs12821256	12:89328335	T	0.169611
rs9651934	12:89355709	C	-0.0406687
rs2638470	12:89675376	G	0.0271665
rs77495873	12:89684892	C	-0.0351508
rs10777186	12:89921277	G	-0.0182849
rs17782847	12:90078824	T	-0.0307009
rs12368980	12:91116069	C	-0.0196443
rs9515244	13:111187839	T	0.0113789
rs7999377	13:114825471	G	-0.0121264
rs9603422	13:39343822	C	-0.0275398
rs12430931	13:74095471	A	-0.0205678
rs1279403	13:78391757	C	0.0686138
rs9301719	13:91890457	A	-0.0147414
rs6492711	13:95196559	C	0.0388959
rs879552	14:104010198	G	-0.0136205
rs72681869	14:50655357	G	-0.0602237
rs210381	14:54107791	G	-0.0138639
rs56412931	14:54336842	C	0.0139373
rs1951116	14:60782189	G	0.0152544
rs10873172	14:64390030	G	-0.0185266
rs28649231	14:68423215	G	-0.0175597
rs12148044	14:69214219	G	-0.0207687
rs113005382	14:92652216	C	-0.0225753
rs1885194	14:92777462	T	0.154776

rs76519749	14:92815008	T	0.113534
rs4900133	14:92963513	T	-0.00822047
rs8018433	14:99170143	G	0.00854696
rs7151737	14:99718105	T	-0.00937641
rs2066703	15:25962058	G	0.0147133
rs117734725	15:28089081	A	-0.0203121
rs35591793	15:28130503	C	-0.0551288
rs12909167	15:28150846	G	-0.0303422
rs12913832	15:28365618	A	0.423616
rs4093386	15:29272905	A	-0.0103676
rs143692029	15:49189487	G	0.0438127
rs28653088	15:57542893	C	-0.0162745
rs8040384	15:73334233	G	-0.0100037
rs12324786	15:81526302	C	0.0434221
rs1877024	15:83964925	G	0.0206903
rs17554929	15:86050204	T	0.0110494
rs34560261	15:90734426	C	-0.0146584
rs3751656	15:90771704	G	-0.00982102
rs76536391	16:54019653	C	-0.0234891
rs1801257	16:57689385	C	0.0125834
rs12708925	16:72122706	A	0.0112279
rs9888791	16:81358164	G	0.00784137
rs3736239	16:83430201	C	0.00654503
rs7202966	16:84074278	G	0.0180209
rs116918202	16:89587080	C	0.136812
rs144988021	16:89655877	C	-0.0633354
rs117030214	16:89749165	T	0.0317044
rs371629999	16:89799124	G	0.0546141
rs73276549	16:89959387	A	0.10169
rs1805005	16:89985844	G	0.0578755
rs1805006	16:89985918	C	0.152777
rs2228479	16:89985940	G	-0.0358088

rs11547464	16:89986091	G	0.125518
rs1805007	16:89986117	C	0.214288
rs1805008	16:89986144	C	0.174076
rs885479	16:89986154	G	0.00622204
rs1805009	16:89986546	G	0.137588
rs2746026	17:18134354	C	-0.0136451
rs117827273	17:38279036	C	0.0161002
rs117612447	17:39551099	C	-0.0516712
rs897418	17:39641932	G	-0.022134
rs11079764	17:45309693	A	-0.0153842
rs72833470	17:45950721	A	0.0515985
rs7219107	17:45968591	G	0.0211956
rs16949720	17:46027257	G	-0.0254559
rs61397471	17:47956077	C	0.0126653
rs9303554	17:48008683	C	-0.0163254
rs55788912	17:48022018	G	-0.0245968
rs17833789	17:55230628	C	0.0213252
rs1133683	17:63533768	G	0.0123302
rs12952240	17:78340404	T	0.0166033
rs56931525	17:79385044	C	-0.0202764
rs11541225	17:79524903	T	0.0132591
rs71373084	17:79564930	C	0.0470231
rs34683731	17:79909537	G	0.0196559
rs75596709	17:80752132	G	0.0321975
rs7236870	18:42601019	A	0.0116166
rs9964220	18:43536260	C	-0.0199226
rs12954491	18:77305852	A	-0.0123154
rs4804519	19:10808770	C	0.0112459
rs12609746	19:1252827	T	0.0135771
rs16964556	19:31200501	A	0.0105915
rs350818	19:4086807	A	0.0130742
rs12722976	1:103345744	C	0.0114749

rs12375	1:10596341	C	0.0115553
rs56168930	1:11037447	A	0.0330375
rs1410860	1:153189978	A	0.0124402
rs4745	1:155106227	A	-0.0141164
rs848200	1:16293592	C	0.0193862
rs11585357	1:17601165	C	0.0199669
rs3851294	1:205130413	A	-0.058004
rs1501552	1:211183493	G	-0.0218003
rs28826269	1:211344388	G	-0.0199828
rs113614987	1:214597586	C	0.0112174
rs7533482	1:214673271	T	-0.0143586
rs59116340	1:227366626	A	-0.0159259
rs16857370	1:232618974	T	-0.0330528
rs12129097	1:236075564	C	-0.0311369
rs1947536	1:240395615	C	0.0104398
rs195700	1:24810069	A	-0.0105205
rs12077712	1:25463142	T	0.0107186
rs79598313	1:27284913	C	0.0381628
rs1937999	1:41898581	G	0.0111818
rs12568356	1:59053745	C	0.0127139
rs17377295	1:61710179	G	-0.0244758
rs34517439	1:78450517	C	0.0164372
rs80293268	1:8207579	G	0.138183
rs74865018	1:8208573	G	0.0453431
rs72635768	1:8322818	A	-0.0115386
rs12034421	1:85528006	C	0.0116013
rs236285	1:94132788	G	-0.0144944
rs6113507	20:22075642	G	-0.01169
rs6059655	20:32665748	A	-0.136961
rs4519594	20:34960201	T	-0.0275022
rs6028446	20:38101545	T	0.0109123
rs55901013	20:52642793	C	0.0612842

rs6127868	20:55409093	G	0.0284067
rs12481673	20:57845936	T	-0.0278387
rs2829780	21:26868270	C	-0.0312576
rs229103	21:28227149	G	-0.0117963
rs68088846	21:36208167	G	-0.0130246
rs73220980	21:44752768	G	0.0484281
rs672948	21:44793448	A	-0.0215473
rs165895	22:19953059	T	-0.015758
rs13056416	22:29525694	C	0.0188161
rs137487	22:33259104	A	0.0129699
rs9611155	22:39739187	C	-0.0132839
rs79966207	22:50722408	T	0.0168209
rs3749110	2:109605767	G	-0.0152687
rs72785456	2:11309644	A	0.027072
rs6716872	2:119538982	G	-0.0205953
rs114153232	2:121081603	G	0.0240747
rs6724774	2:135407285	C	-0.0181187
rs12714332	2:1642070	C	0.0136964
rs71421546	2:176991857	C	0.0326285
rs13431878	2:177582940	T	0.0158202
rs13030978	2:192117238	C	-0.0129593
rs12693954	2:202839971	T	-0.0172276
rs7583880	2:207999713	G	-0.011165
rs56020963	2:214082380	T	-0.0158721
rs10169459	2:222051419	T	-0.0496625
rs74773586	2:222085154	A	0.0112803
rs17349283	2:222089797	A	0.0423894
rs7570475	2:222220804	G	0.0172304
rs13029372	2:222348434	A	-0.0203804
rs744174	2:223024442	A	-0.0150973
rs116254882	2:223025055	G	-0.0520561
rs12618431	2:223110512	A	-0.0198919

rs2303948	2:223159119	C	-0.0179702
rs17266608	2:225929038	T	0.0106794
rs812383	2:233737132	C	0.0111093
rs9287636	2:239680992	G	-0.0170601
rs4075018	2:239949681	G	0.0165265
rs11683301	2:25539985	T	-0.0162833
rs1260326	2:27730940	T	-0.0124865
rs71443018	2:28613302	G	0.114681
rs55870117	2:28647650	G	-0.0206211
rs4952542	2:42148519	T	0.0262589
rs57696714	2:43626521	T	0.0199373
rs2706764	2:70545164	C	0.0126288
rs7608166	2:88581126	A	0.0295164
rs2437689	3:11660766	C	0.0158351
rs9847240	3:122526816	G	0.0272605
rs115182912	3:125993274	G	0.0419582
rs2715610	3:133607321	A	-0.0107395
rs2293252	3:138123854	C	-0.0123538
rs6440003	3:141094209	G	0.0368757
rs73869619	3:141179873	T	-0.0464917
rs3804772	3:141634056	G	0.0231994
rs323610	3:151904190	T	-0.0116192
rs1713843	3:153786394	C	0.0148427
rs34128525	3:171037521	G	-0.0124957
rs833268	3:181517982	G	-0.0121059
rs698083	3:186997742	T	0.0130122
rs62289589	3:189214006	C	-0.0100364
rs9809528	3:250758	A	0.0114483
rs754821	3:66789001	A	-0.012164
rs2014520	3:69797413	G	0.0171452
rs704278	3:71710269	G	0.0156434
rs9809116	3:72397279	A	-0.014176

rs586936	3:73862616	G	-0.0151105
rs9998015	4:105816898	T	0.0201095
rs17429682	4:106028823	G	-0.0123791
rs7673917	4:109012183	T	-0.0719431
rs17038688	4:109076822	C	-0.0127473
rs220625	4:109252929	C	0.0259
rs219493	4:109350880	T	0.0239933
rs11731416	4:109478108	C	-0.020615
rs1584590	4:139446206	T	0.0153277
rs1996020	4:145511194	A	0.0148835
rs72719803	4:149805677	C	0.0251045
rs4696396	4:154000968	C	-0.0108014
rs10015223	4:4388874	C	-0.0298674
rs6554121	4:54509016	C	-0.0132082
rs1107674	4:57838690	T	-0.0133604
rs17804499	4:74442349	G	0.022683
rs1874202	4:75328479	C	-0.0242025
rs11942966	4:77490144	A	-0.0118136
rs371273	4:79281682	A	0.0313695
rs1458046	4:81199966	G	0.0240845
rs72661730	4:81661426	G	0.0393305
rs17392334	4:86602918	C	-0.0170715
rs28483422	4:86929983	T	0.0174814
rs163888	5:106889936	C	-0.0104892
rs9326880	5:112541296	G	-0.0139143
rs10519488	5:116231256	A	-0.0295518
rs10479082	5:133852468	G	0.0146354
rs2936938	5:173831110	C	-0.0156771
rs4242182	5:174156168	T	-0.0305105
rs17714046	5:180661980	T	-0.0278712
rs13184208	5:26227296	C	-0.00698857
rs16891982	5:33951693	C	0.393619

rs13289	5:33986409	C	-0.0113175
rs6881866	5:34700000	A	-0.01284
rs77325285	5:53018404	G	-0.0398646
rs6875907	5:53112624	C	0.0327607
rs3846492	5:53484470	G	-0.0126596
rs7714232	5:56011357	A	-0.0222837
rs1835873	5:57135050	C	0.0118061
rs4282273	5:66789202	C	-0.0116674
rs259035	5:79695370	T	-0.0222404
rs11957689	5:90263347	G	-0.0337392
rs1651282	5:9547021	C	0.0104786
rs2545729	5:98125568	A	0.0131008
rs594329	6:10580294	C	0.0187277
rs1508355	6:106326339	T	-0.00952116
rs10457158	6:107426530	G	0.0131257
rs1062034	6:109005588	A	0.0104145
rs2127965	6:113304903	A	0.0150497
rs7740107	6:130374461	T	0.0119929
rs6901548	6:131388483	T	-0.01435
rs4896038	6:134609291	A	0.0148943
rs137912555	6:134961601	T	0.032729
rs34286635	6:148570402	A	0.0146398
rs4897010	6:148737041	C	0.0140414
rs10434895	6:151577739	A	0.0226854
rs57304695	6:159182635	C	0.00792163
rs9457478	6:159248431	C	0.0415351
rs9365518	6:163391444	C	0.0104837
rs61376093	6:192288	T	0.0572761
rs2223620	6:20627777	C	-0.0197734
rs9368374	6:21920773	A	-0.0122174
rs9501926	6:3126911	T	8.05E-05
rs9266756	6:31351541	A	-0.0263029

rs34451818	6:31583841	C	-0.0263155
rs17207524	6:31726850	G	0.0222876
rs7745968	6:3467745	G	-0.00953608
rs71550004	6:376962	C	0.082694
rs12203592	6:396321	C	-0.322913
rs62400428	6:45739785	C	-0.0141789
rs9349337	6:45901916	G	-0.0170402
rs60729976	6:466518	G	0.0363969
rs395180	6:4768691	T	-0.00457857
rs9378896	6:521147	T	-0.088581
rs190352051	6:7108116	C	0.029146
rs10498672	6:7797840	C	0.00184179
rs314349	7:100401825	T	-0.0162811
rs12535629	7:100451732	C	0.0209057
rs2529369	7:105416560	C	-0.0178656
rs7803075	7:130742066	A	0.0366374
rs2282867	7:14015318	T	-0.0167272
rs1612590	7:144076764	A	-0.0100594
rs9767875	7:155091406	G	-0.0166157
rs56326046	7:156151175	A	0.0104262
rs1043291	7:2577781	T	-0.010672
rs864745	7:28180556	T	-0.0218211
rs2266921	7:2822933	C	0.0141619
rs10224680	7:28818245	T	-0.0174742
rs7784327	7:35824415	C	0.0103834
rs117441897	7:41327697	A	-0.0252988
rs77969347	7:41408001	A	0.0392845
rs6960634	7:42499768	C	-0.0106481
rs1464841	7:46801428	T	-0.0176344
rs11764140	7:46905525	C	0.0136912
rs80030895	7:90867230	C	-0.0395039
rs73217177	7:90913120	T	0.0283901

rs2157725	7:91209001	C	0.0114881
rs2446928	8:101076864	A	0.0105895
rs1389985	8:108994382	G	0.0249194
rs446454	8:109092699	C	-0.018623
rs1526460	8:116823453	G	-0.0167163
rs2445922	8:119133623	C	-0.0172567
rs2048651	8:22600953	T	0.0129215
rs2595041	8:29852712	T	0.0117674
rs57128498	8:38568215	C	0.0222206
rs6473306	8:82704598	T	0.0153332
rs2600605	8:82720760	G	0.0303745
rs330926	8:9000965	C	-0.0116455
rs3021523	9:100616583	T	0.0207145
rs10739220	9:109054417	C	-0.0160777
rs540599	9:116387254	C	-0.0151371
rs10985112	9:123731408	G	0.0164433
rs57425397	9:126705122	G	-0.0270787
rs969585	9:126768582	T	0.0277606
rs58979150	9:126808006	C	-0.0689052
rs16936765	9:126808021	C	0.0493388
rs953470	9:126984382	A	0.0150885
rs10818930	9:126991185	T	0.0361643
rs1326797	9:12716762	T	0.0436314
rs7032500	9:129822536	G	0.0114543
rs770197	9:15866367	C	0.0159713
rs7858025	9:16560854	A	0.0213729
rs2254330	9:16802973	A	-0.0266045
rs12350739	9:16885017	G	0.0349393
rs7851902	9:19298581	A	-0.0107196
rs1536523	9:19827538	T	-0.0106577
rs520015	9:211762	C	-0.0233756
rs10811647	9:22065002	C	-0.0135091

rs10812605	9:27510360	C	0.0109636
rs11557154	9:34107505	C	0.0178001
rs7037587	9:5896161	T	0.0144527
rs1475675	9:82289001	T	-0.0199103

Table F.1: Variants and corresponding weighting used in non-red hair PRS

F.2 Red hair PRS table

SNP ID	Chr:Pos	Effect Allele	PRS weighting
rs652190	11:69388143	G	0.092068567816393
rs4512823	11:88966584	A	0.076618389845287
rs9544609	13:78392656	C	0.083743543432859
rs1129038	15:28356859	C	-0.228531791844379
rs9928829	16:83613777	A	-0.067620599197698
rs117614496	16:88895529	C	-1.31930883633927
rs117791738	16:89029902	C	-0.257622424780464
rs9921065	16:89066873	A	-0.675067605493479
rs112584165	16:89085743	A	-1.50551492953155
rs12925392	16:89133302	A	0.366398516410895
rs117984432	16:89454991	T	-0.458280613102224
rs191414071	16:89568288	C	-1.40026347648073
rs164737	16:89666511	A	-0.611161745818654
rs61482648	16:89715564	A	-0.913432321025675
rs117030214	16:89749165	T	-0.870872914345581
rs34357723	16:89886519	C	-2.1709342174339
rs1805005	16:89985844	G	1.12033790244137
rs1805006	16:89985918	C	-0.998244013532363
rs2228479	16:89985940	G	2.16061591646883
rs11547464	16:89986091	G	-1.37166193907558
rs1805007	16:89986117	C	-2.28017801708322
rs1805008	16:89986144	C	-1.34625859559667
rs885479	16:89986154	G	1.78447130832052
rs1805009	16:89986546	G	-1.70669248133833
rs11859970	16:90018389	T	0.709187838553977
rs9747347	17:79606820	T	0.116261879284445
rs34878396	19:3546264	C	0.340891030782924
rs6059655	20:32665748	A	0.610384944986995
rs4519594	20:34960201	T	0.141534283893119

rs60780889	5:173976464	C	-0.121425536515297
rs12203592	6:396321	C	-0.142422227798653
rs2025753	6:51697567	T	0.134102479523353
rs12529074	6:51713402	G	0.140961222950641

Table F.2: Variants and corresponding weighting used in red hair PRS

F.3 Skin colour PRS table

SNP ID	Chr:Pos	Effect Allele	PRS weighting
rs7086846	10:102102612	G	-0.00907469
rs1407696	10:112638028	T	-0.00784483
rs35563099	10:119572403	C	-0.0286417
rs2236295	10:64564892	G	-0.00855942
rs55901776	10:82089493	T	0.0079929
rs10766301	11:16217413	C	0.011244
rs72632979	11:16615883	A	-0.0156084
rs7478729	11:18319915	G	0.00902784
rs150527451	11:68817897	G	0.0636416
rs3829241	11:68855363	G	0.0176973
rs881361	11:78128047	C	-0.0182549
rs11021172	11:88484644	C	0.0418997
rs144650507	11:89005318	C	-0.068772
rs1126809	11:89017961	G	0.0718351
rs10830452	11:89966202	A	-0.00819661
rs113159861	12:116745298	G	0.0135548
rs7304293	12:125389272	T	-0.0100969
rs10771034	12:23979199	T	-0.00872492
rs2405295	12:41097386	G	-0.0096057
rs7979418	12:46774771	A	-0.00787618
rs12425342	12:85693252	G	0.021716
rs11104887	12:88847783	C	-0.0210587
rs11104947	12:88942980	G	0.0455125
rs12821256	12:89328335	T	0.0347045
rs61924662	12:90485617	G	0.00972695
rs1278761	13:113532990	T	0.0179839
rs1823554	13:78382705	A	0.00889803
rs9590030	13:95171058	G	0.020757
rs8004474	14:102756498	T	0.00910866

rs72681869	14:50655357	G	-0.0429606
rs941799	14:92776825	C	0.0430806
rs76519749	14:92815008	T	0.0283699
rs17116015	15:25954518	C	0.0118683
rs35591793	15:28130503	C	-0.0293736
rs17567007	15:28201539	C	-0.0164001
rs117886461	15:28230378	G	0.127005
rs746861	15:28266235	T	-0.0200838
rs12913832	15:28365618	A	0.201014
rs60837134	15:29273260	C	-0.0027098
rs77512536	15:31387008	G	-0.0132682
rs28456199	15:31395538	G	-0.0208606
rs57249121	15:50276245	T	-0.00994071
rs72744156	15:81526279	C	0.0110745
rs28371837	15:83383187	T	0.010498
rs12901150	15:91116386	T	-0.00908155
rs758045	16:4434198	T	0.0090629
rs8054175	16:71990424	G	0.0120628
rs731957	16:85492882	C	-0.00528032
rs72812225	16:87785911	T	-0.0126209
rs113894462	16:89067671	G	0.1037
rs11648089	16:89713938	T	-0.0933675
rs56288641	16:89777078	G	0.177892
rs371629999	16:89799124	G	0.0836207
rs75092208	16:89808241	T	-0.097026
rs144937287	16:89971469	C	0.0215835
rs1805005	16:89985844	G	0.0196333
rs1805006	16:89985918	C	0.166466
rs2228479	16:89985940	G	0.0119492
rs11547464	16:89986091	G	0.163544
rs1805007	16:89986117	C	0.227012
rs1805008	16:89986144	C	0.160931

rs885479	16:89986154	G	-0.00196462
rs1805009	16:89986546	G	0.201471
rs117711828	16:90107861	G	0.0662157
rs2325752	17:45921906	T	0.0120803
rs739992	17:48430291	C	0.00904658
rs149789627	17:73426048	G	-0.0137019
rs6420484	17:79612397	A	-0.0131845
rs1865048	19:39437559	C	-0.0082478
rs535930	1:110724488	G	0.0156665
rs3894771	1:150789961	A	-0.00957156
rs4656781	1:170510306	C	-0.00902332
rs4656781	1:170510306	C	-0.00902332
rs3851294	1:205130413	A	-0.025019
rs3845302	1:25259884	C	0.00824906
rs6696511	1:42110888	T	0.00915781
rs670318	1:63727542	T	-0.0238871
rs1613999	1:66895085	T	-0.0152516
rs80293268	1:8207579	G	0.0311559
rs2249048	20:21457672	C	0.00804572
rs6059655	20:32665748	A	-0.130898
rs6120568	20:32793500	T	-0.0375183
rs13042880	20:32944549	C	-0.0186604
rs6060343	20:33865588	G	-0.0258583
rs4519594	20:34960201	T	-0.0262161
rs6025355	20:37017831	G	0.00932864
rs672948	21:44793448	A	-0.0108721
rs34066050	22:38612604	G	0.0109169
rs5766565	22:45622684	A	-0.0184586
rs3731976	2:172378893	G	-0.00878313
rs7591631	2:175292326	C	-0.00755724
rs62186153	2:222034085	A	0.0151769
rs10192020	2:222072535	G	0.0107757

rs887829	2:234668570	C	-0.0112595
rs780093	2:27742603	T	-0.00950479
rs71443018	2:28613302	G	0.0228634
rs7584123	2:33001100	T	-0.00807964
rs10169939	2:38291967	T	-0.0134402
rs10185673	2:43302571	G	0.00901539
rs6440008	3:141154542	T	0.0107724
rs4683632	3:141576863	A	0.00879031
rs9867857	3:156491160	C	0.0119568
rs1345417	3:181511951	C	-0.0103053
rs11921010	3:85434260	T	0.00752477
rs6819541	4:75341885	A	-0.00778934
rs285894	5:113996889	G	-0.00755604
rs329120	5:133861756	C	0.00766898
rs251466	5:149195603	C	-0.0186082
rs923520	5:173975145	T	-0.00894683
rs76277456	5:31103664	T	0.000191237
rs4866399	5:33887419	G	-0.0104558
rs79014396	5:33936721	C	-0.0751096
rs16891982	5:33951693	C	0.405692
rs6875907	5:53112624	C	0.00992193
rs7720119	5:59018067	C	0.0113077
rs2662225	5:66321379	G	-0.00924273
rs6887203	5:90277037	A	-0.00970301
rs423358	6:1251437	T	0.00590535
rs1407373	6:146847519	G	-0.0317252
rs3123139	6:159242809	A	0.0127972
rs115949579	6:19538436	C	0.0228865
rs909728	6:29719561	T	0.0079022
rs12203592	6:396321	C	0.0902685
rs10947996	6:41892809	G	-0.00871642
rs13201830	6:51696355	A	-0.00890831

rs17134857	6:566509	G	-0.00763922
rs9443928	6:82225675	G	0.00872232
rs56134000	7:100499789	T	0.012232
rs7794285	7:104849737	A	-0.0117846
rs1721019	7:17004301	C	0.0105806
rs34585474	7:17041434	C	0.0115067
rs117132860	7:17134708	G	0.0599109
rs35332062	7:73012042	G	-0.011159
rs6469606	8:116601902	C	-0.0184577
rs6981915	8:116823046	T	0.0301342
rs10985890	9:125884052	C	-0.0157095
rs75630385	9:12675342	A	-0.0478312
rs1137134	9:12712157	G	0.0292361
rs7020076	9:12924863	A	0.00737894
rs76798990	9:129801021	T	0.0120782
rs10962494	9:16574838	T	0.00451376
rs2254330	9:16802973	A	-0.0294543
rs12350739	9:16885017	G	0.0436998
rs3126954	9:25113256	G	0.0120425
rs4601374	9:5874777	G	0.00888966

Table F.3: Variants and corresponding weighting used in skin colour PRS

F.4 Tanning ability PRS table

SNP ID	Chr:Pos	Effect Allele	PRS weighting
rs7086846	10:102102612	G	0.0148764
rs7098111	10:119573178	C	0.05771
rs729465	10:119576277	T	0.00840287
rs3902976	10:134564070	G	-0.022114
rs1129614	10:5808086	G	0.0188853
rs2236295	10:64564892	G	0.0172134
rs10788623	10:82199415	G	-0.0136178
rs10766301	11:16217413	C	-0.0155047
rs72632979	11:16615883	A	0.0248906
rs2053979	11:47439444	A	-0.0135684
rs72928978	11:68831364	G	-0.0859449
rs3829241	11:68855363	G	-0.0203475
rs10899501	11:78131408	C	0.0179166
rs117796400	11:88303869	G	0.0382223
rs11021172	11:88484644	C	-0.081385
rs183139540	11:89005267	G	0.167995
rs1126809	11:89017961	G	-0.132163
rs78043739	11:89180697	G	0.0576366
rs200522762	11:89323846	T	-0.0449229
rs140100229	11:89822066	C	0.0552804
rs518328	11:94665400	A	0.00817585
rs7310615	12:111865049	C	-0.0125467
rs61935849	12:116637641	A	-0.0261492
rs68099344	12:125387877	A	-0.0199568
rs4760521	12:129306308	G	-0.0139926
rs10771034	12:23979199	T	0.0143292
rs10879095	12:41123651	G	0.0141225
rs11104870	12:88829294	C	0.019846
rs11104947	12:88942980	G	-0.112623

rs77081360	12:88962365	C	-0.0734268
rs4408360	12:90394861	A	0.0128821
rs12865907	13:107895045	T	-0.0123711
rs1278761	13:113532990	T	-0.0365957
rs9528082	13:60793386	G	0.0127125
rs9590030	13:95171058	G	-0.0223874
rs1885194	14:92777462	T	-0.0337858
rs8033747	15:27680487	G	-0.0167336
rs78114576	15:28062125	G	-0.0546942
rs142306097	15:28103388	T	0.0619886
rs75137181	15:28114816	G	0.0509105
rs139615925	15:28143305	T	-0.0469435
rs72710600	15:28157441	T	-0.0688763
rs141514981	15:28206623	C	-0.0860262
rs117886461	15:28230378	G	-0.260305
rs77516282	15:28264167	G	0.0707306
rs12913832	15:28365618	A	-0.158113
rs17228129	15:31388787	G	0.0266387
rs11639071	15:49749726	A	-0.0127548
rs9302151	15:50290361	T	0.0123848
rs2455717	15:82427605	A	0.0132847
rs8028550	15:85535071	T	-0.0122692
rs12901150	15:91116386	T	0.0144172
rs6498485	16:14010994	G	0.0182705
rs4788815	16:71634811	A	-0.0163471
rs4238968	16:73143479	C	-0.0149959
rs74633564	16:81768470	C	0.0177147
rs2270500	16:85826114	A	0.00853875
rs75898184	16:88929011	A	-0.197374
rs72815590	16:89121692	C	0.0942502
rs149610032	16:89368895	G	0.123683
rs150020387	16:89370630	G	0.0892222

rs143743429	16:89592248	G	-0.292752
rs59013041	16:89631186	A	0.131852
rs139031001	16:89675362	C	0.121548
rs11641963	16:89749244	C	0.167065
rs371629999	16:89799124	G	-0.158277
rs73276549	16:89959387	A	-0.197396
rs1805005	16:89985844	G	-0.0364647
rs1805006	16:89985918	C	-0.355558
rs2228479	16:89985940	G	-0.0262257
rs11547464	16:89986091	G	-0.316383
rs1805007	16:89986117	C	-0.444685
rs1805008	16:89986144	C	-0.315606
rs885479	16:89986154	G	0.0149795
rs1805009	16:89986546	G	-0.408045
rs150909008	16:90004693	G	-0.0285404
rs117711828	16:90107861	G	-0.131693
rs34635363	17:79549250	G	-0.0222677
rs77733715	19:3537184	A	-0.0842058
rs6510760	19:3565253	G	-0.0290997
rs6689641	1:110720400	A	-0.0135609
rs3894771	1:150789961	A	0.0212938
rs76798800	1:154994978	G	-0.0206176
rs13374238	1:166307175	G	-0.0163532
rs17350312	1:170377400	T	-0.012202
rs17350312	1:170377400	T	-0.012202
rs3851294	1:205130413	A	0.0475615
rs35845538	1:205138321	G	-0.0269986
rs12727947	1:212417130	T	0.0167566
rs4393146	1:42118766	A	-0.0157746
rs670318	1:63727542	T	0.0491019
rs1613999	1:66895085	T	0.0337458
rs6137358	20:21465380	A	-0.0144249

rs6059655	20:32665748	A	0.277202
rs138420028	20:32752091	C	0.0444303
rs6120568	20:32793500	T	0.0699159
rs13042880	20:32944549	C	0.0391094
rs6060343	20:33865588	G	0.0593954
rs4519594	20:34960201	T	0.0597669
rs151525	20:3522015	C	-0.0165726
rs138284316	20:36038982	G	-0.0782392
rs150077408	20:36591620	G	0.00542163
rs6029393	20:39494767	G	0.00613479
rs73112279	20:41964714	C	0.00677937
rs399907	21:43429646	G	0.0180441
rs4821721	22:38244216	C	0.0239314
rs4384	22:38572440	G	0.0183167
rs6007506	22:45622014	C	0.0350977
rs35380972	2:172381948	A	0.018748
rs116078354	2:203001079	T	-0.0185705
rs1541966	2:221964207	T	-0.0140438
rs1430663	2:223121674	T	0.0176812
rs13428823	2:25373298	G	-0.0172546
rs780094	2:27741237	T	0.0180476
rs1800440	2:38298139	T	0.0358255
rs162561	2:38298877	T	0.0349641
rs12612692	2:43105088	T	-0.0142777
rs36027953	3:148894357	T	0.02708
rs9818780	3:156492758	T	-0.0218865
rs7616157	3:63340880	T	-0.0142034
rs2874270	3:70013349	A	0.013205
rs7617323	3:85625852	G	-0.0184578
rs9821965	3:98698414	A	0.0141216
rs4689314	4:4346065	T	-0.0284848
rs251468	5:149194485	C	0.0365879

rs6883391	5:173973627	C	0.0201838
rs28362590	5:176731452	G	0.0165644
rs76277456	5:31103664	T	-0.0137386
rs10941108	5:33867419	T	0.0109081
rs16891982	5:33951693	C	-0.45817
rs72738461	5:39354913	A	0.0123742
rs11954036	5:59028853	T	-0.0228528
rs74852363	5:60696108	C	0.0212086
rs469394	5:66347990	T	0.0135695
rs240683	6:1133692	C	0.00414012
rs10214796	6:1142731	G	-0.0206671
rs423358	6:1251437	T	-0.0078791
rs1546966	6:146386018	G	0.0118971
rs73783709	6:146864384	T	0.0445835
rs2472788	6:14735129	A	-0.015283
rs62435865	6:164516218	T	-0.0239107
rs61376093	6:192288	T	0.0565427
rs115949579	6:19538436	C	-0.0372403
rs6914598	6:21163919	T	-0.0128597
rs2517674	6:29935806	G	-0.017217
rs9405194	6:3355595	G	-0.0149599
rs7753579	6:375759	T	-0.0813511
rs12203592	6:396321	C	-0.280662
rs115467869	6:4080299	G	-0.01451
rs6938966	6:41760990	T	-0.019717
rs2432790	6:5393201	G	-0.0221605
rs115631737	6:7127350	G	0.0221544
rs9502564	6:7230680	G	-0.00897222
rs1034241	6:82473953	T	-0.024033
rs12215011	6:9348912	G	0.00540519
rs56269269	7:104846634	G	0.0138647
rs1721040	7:16990804	T	-0.0228225

rs34585474	7:17041434	C	-0.0311051
rs117132860	7:17134708	G	-0.147771
rs849138	7:28177338	G	-0.0172294
rs17172691	7:46398911	G	-0.0202923
rs2721929	8:116613258	G	0.0427519
rs7842594	8:116837897	G	-0.0465726
rs78268832	8:42001544	G	-0.019794
rs6996198	8:65463442	C	-0.0172217
rs373184	9:10058873	C	-0.0120516
rs10960749	9:12671566	G	-0.0440204
rs10810636	9:16799109	A	-0.0560382
rs12350739	9:16885017	G	-0.0745272
rs7868612	9:20616618	G	0.00991378
rs6475567	9:21733644	T	-0.00822887
rs7469146	9:27541041	C	-0.0119083

Table F.4: Variants and corresponding weighting used in tanning ability PRS

F.5 Pigscore 1 PRS table

SNP ID	Chr:Pos	Effect Allele	PRS weighting
rs56014906	10:102103508	A	-0.0234134
rs35563099	10:119572403	C	-0.0623144
rs729465	10:119576277	T	-0.0141647
rs1148246	10:35496626	C	-0.0275143
rs1888967	10:64579460	T	-0.0232142
rs4433500	10:74068998	G	0.0247034
rs703978	10:80944147	C	0.030727
rs2923099	11:10369340	C	-0.0188624
rs7108738	11:15710084	T	-0.0337487
rs4756838	11:15918766	C	0.0365455
rs12277740	11:16305351	G	-0.0636843
rs7929530	11:16381813	C	0.0307627
rs72632979	11:16615883	A	-0.0490441
rs7478729	11:18319915	G	0.0282396
rs66716358	11:44330610	C	0.0206727
rs4755834	11:44469765	C	-0.0192302
rs10897275	11:62203865	G	0.0216892
rs479844	11:65551957	A	-0.0178781
rs72917317	11:68817441	T	0.238868
rs3829241	11:68855363	G	0.0662971
rs4078279	11:7541511	C	0.0259185
rs12577359	11:77723797	C	0.0190172
rs10899501	11:78131408	C	-0.0668783
rs117796400	11:88303869	G	-0.064377
rs1942493	11:88920675	C	-0.181719
rs183139540	11:89005267	G	-0.238081
rs1126809	11:89017961	G	0.179548
rs200522762	11:89323846	T	0.0670438
rs2634309	11:89421092	C	-0.0407436

rs17201363	11:90432150	C	-0.0756123
rs1265564	12:111708458	A	-0.0228511
rs61939692	12:116535976	G	0.0347289
rs9971729	12:23979791	A	-0.0282213
rs7979418	12:46774771	A	-0.0204762
rs12425342	12:85693252	G	0.0679809
rs73200863	12:88143958	G	0.0610774
rs10858711	12:88707968	C	-0.0766175
rs10858760	12:88957352	A	0.085062
rs12821256	12:89328335	T	0.163476
rs11611632	12:89445539	C	-0.00923236
rs1278761	13:113532990	T	0.0437619
rs9603422	13:39343822	C	-0.0354897
rs1766344	13:78385068	A	0.0555086
rs7337275	13:95171411	A	0.062265
rs7359090	14:104184322	A	-0.0237246
rs72681869	14:50655357	G	-0.0983729
rs10873172	14:64390030	G	-0.0221911
rs941799	14:92776825	C	0.182979
rs76519749	14:92815008	T	0.126499
rs4905808	14:99692271	T	-0.0166725
rs146704333	15:28014947	C	-0.130311
rs35690998	15:28038935	C	-0.112995
rs17567007	15:28201539	C	-0.0461893
rs117886461	15:28230378	G	0.464168
rs746861	15:28266235	T	-0.0493756
rs56108911	15:28271247	G	-0.0822616
rs12913832	15:28365618	A	0.649624
rs11852567	15:29292973	G	-0.0125409
rs4779826	15:31410084	A	0.0205301
rs12324786	15:81526302	C	0.0459529
rs4779035	15:83260436	T	0.0267067

rs11639998	16:4527109	T	0.0235095
rs12935795	16:71987452	G	0.0325681
rs62048536	16:85432741	C	-0.0173566
rs118129610	16:88190981	C	0.0389281
rs8044675	16:88230086	G	0.020795
rs9940427	16:88711819	G	-0.134736
rs112787727	16:88739665	G	0.352191
rs116004099	16:88896100	T	-0.14871
rs9923267	16:89494760	C	0.0551605
rs4785577	16:89618530	G	-0.304483
rs72807518	16:89754335	G	0.433672
rs34427386	16:89866177	G	0.292347
rs7185897	16:89896481	C	0.11753
rs1805005	16:89985844	G	-0.017967
rs1805006	16:89985918	C	0.714641
rs2228479	16:89985940	G	-0.141299
rs11547464	16:89986091	G	0.771814
rs1805007	16:89986117	C	1.13166
rs1805008	16:89986144	C	0.75573
rs885479	16:89986154	G	-0.126847
rs1805009	16:89986546	G	0.992328
rs137938320	16:90135457	C	0.949971
rs72833470	17:45950721	A	0.0601479
rs890432	17:46629232	G	-0.0137454
rs739990	17:48431032	G	0.0227111
rs62060349	17:55231168	T	0.0208342
rs9747347	17:79606820	T	-0.0795377
rs9964220	18:43536260	C	-0.031791
rs77733715	19:3537184	A	0.126906
rs2161568	19:7574918	G	-0.0353921
rs12081181	1:11034165	T	0.0300423
rs6689641	1:110720400	A	0.0247652

rs9733	1:150618961	T	0.0209222
rs76798800	1:154994978	G	0.0271361
rs2235741	1:16243482	C	0.0314523
rs11584287	1:17607501	G	0.0319681
rs3851294	1:205130413	A	-0.101321
rs351376	1:212426172	C	0.0231038
rs4660121	1:236045229	G	-0.0332402
rs6702659	1:42112236	T	0.0225203
rs670318	1:63727542	T	-0.0622656
rs1613999	1:66895085	T	-0.030716
rs80293268	1:8207579	G	0.139302
rs6113632	20:22305785	A	-0.0211072
rs6059655	20:32665748	A	-0.507173
rs138420028	20:32752091	C	-0.0812904
rs6087561	20:32783827	T	-0.113974
rs13042880	20:32944549	C	-0.0699321
rs6088778	20:33836110	T	-0.0698705
rs6060343	20:33865588	G	-0.0966195
rs4519594	20:34960201	T	-0.104973
rs221314	20:35298668	G	-0.0447719
rs6101622	20:38233064	C	0.0177074
rs6102033	20:39184691	C	-0.0133934
rs73132906	20:52659595	G	0.0686765
rs6127868	20:55409093	G	0.028096
rs6123874	20:57843358	G	-0.0333367
rs399907	21:43429646	G	-0.0217415
rs672948	21:44793448	A	-0.0420852
rs132941	22:38545942	T	-0.0206829
rs11089938	22:39662703	C	0.022346
rs5766565	22:45622684	A	-0.0463692
rs6716872	2:119538982	G	-0.0247724
rs13410301	2:135353183	C	-0.0212526

rs35380972	2:172381948	A	-0.0227058
rs10169459	2:222051419	T	-0.0520252
rs17349283	2:222089797	A	0.0487733
rs1430663	2:223121674	T	-0.0357833
rs10932952	2:223145672	T	-0.0213847
rs887829	2:234668570	C	-0.0234986
rs76645364	2:25393388	A	-0.0425831
rs13036246	2:25532969	C	-0.0207117
rs112272576	2:26337154	T	-0.0320321
rs780094	2:27741237	T	-0.0276752
rs71443018	2:28613302	G	0.116419
rs1800440	2:38298139	T	-0.0422257
rs162561	2:38298877	T	-0.0410157
rs1962386	2:42140464	A	0.0300768
rs10185673	2:43302571	G	0.0205443
rs6730279	2:88556133	T	0.0320043
rs9847240	3:122526816	G	0.0294369
rs6440008	3:141154542	T	0.0377898
rs3804772	3:141634056	G	0.0365791
rs9867857	3:156491160	C	0.0282574
rs75114713	3:69870391	C	0.0505704
rs17006579	3:69934529	A	-0.0232723
rs999870	4:109246930	A	0.0317418
rs7696175	4:38820986	T	0.0197622
rs4689314	4:4346065	T	0.0476675
rs1874202	4:75328479	C	-0.0271727
rs436034	4:79256036	T	0.0281324
rs72661730	4:81661426	G	0.0319456
rs17392334	4:86602918	C	-0.0222867
rs10479082	5:133852468	G	0.0255132
rs251466	5:149195603	C	-0.0413186
rs4242182	5:174156168	T	-0.0358123

rs76277456	5:31103664	T	0.0267652
rs4866399	5:33887419	G	-0.0286971
rs116125333	5:33930012	T	-0.0721164
rs79014396	5:33936721	C	-0.181006
rs16891982	5:33951693	C	0.988295
rs427527	5:52902307	A	-0.023675
rs6875907	5:53112624	C	0.0443655
rs7720119	5:59018067	C	0.0346669
rs62370543	5:59133111	C	-0.0558973
rs2662225	5:66321379	G	-0.0254227
rs60325490	5:90277797	T	-0.0361121
rs77571968	6:10545282	A	0.0337772
rs2185710	6:113377048	A	0.0204452
rs57354745	6:1144979	C	0.021336
rs1407373	6:146847519	G	-0.0676183
rs10434895	6:151577739	A	0.0239942
rs3123135	6:159234178	A	-0.0306848
rs1471836	6:159247859	T	0.0567965
rs115949579	6:19538436	C	0.063705
rs4710940	6:20658012	A	0.0226796
rs10498701	6:21144599	G	0.0207196
rs9394026	6:30982544	G	-0.0274571
rs10947996	6:41892809	G	-0.0289245
rs62389424	6:422631	C	0.0243746
rs9392618	6:515393	C	-0.0302578
rs13201830	6:51696355	A	-0.0344639
rs12535629	7:100451732	C	0.031618
rs7794285	7:104849737	A	-0.0232945
rs2529369	7:105416560	C	-0.0323461
rs10954300	7:130761235	G	0.0397973
rs1721040	7:16990804	T	0.0269731
rs34585474	7:17041434	C	0.0271936

rs117132860	7:17134708	G	0.152559
rs1035049	7:17735392	G	0.017692
rs35332062	7:73012042	G	-0.0337327
rs59201311	7:90905531	C	-0.0350778
rs12543799	8:116555360	C	-0.0443443
rs7842594	8:116837897	G	0.0800309
rs2048651	8:22600953	T	0.0246651
rs2070711	8:42045933	G	0.0338033
rs12545508	8:82774442	C	0.0342955
rs11779019	8:9229139	T	-0.0276529
rs35324451	9:100660136	A	0.0237445
rs1055187	9:102733298	T	0.0232502
rs75123768	9:12596242	T	-0.133163
rs16936765	9:126808021	C	0.0490695
rs4838097	9:126814579	T	-0.0608355
rs10818930	9:126991185	T	0.0320756
rs1137134	9:12712157	G	0.0887477
rs4961671	9:16211366	C	0.0130915
rs10962605	9:16799623	A	0.0858734
rs12350739	9:16885017	G	0.114878
rs1536523	9:19827538	T	-0.0229201
rs10967976	9:27544943	G	0.0215016
rs4601374	9:5874777	G	0.0314786
rs1412401	9:9508427	G	0.0146204

Table F.5: Variants and corresponding weighting used in pigscore 1 PRS

Appendix G

Joint-analysis

G.1 Independent signals table

locus	CHR	BP	RS	P ease ME	P hair ME	P mel	P pigscore 1 ME	P red ME	P skin ME	P ease MT	P hair MT	P pigscore 1 MT	P red MT	P skin MT
1	1	8204596	rs78924983	-	-	-	-	-	-	-	6.46E-29	6.03E-10	-	-
2	1	11037434	rs112115136	-	-	-	-	-	-	-	4.75E-11	-	-	-
3	1	24770594	rs2294524	-	-	-	-	-	-	-	1.2E-09	-	-	-
4	1	25262022	rs1005734	-	-	-	-	-	-	1.7E-09	-	-	-	2.71E-08
5	1	63727542	rs670318	2.25E-08	-	1.21E-08	-	-	-	-	-	-	-	-
6	1	66895085	rs1613999	-	-	-	-	-	-	1.49E-17	-	5.43E-09	-	9.69E-12
7	1	78086718	rs71641308	-	-	-	-	-	-	-	4.85E-08	-	-	-
8	1	110720400	rs6689641	-	-	-	-	-	-	-	-	-	-	1.4E-10

9	1	120466108	rs2793830	-	-	-	-	-	-	7.25E-10	-	-	1.01E-08	2.59E-09
10	1	150800117	rs2089081	-	-	-	-	-	-	-	-	-	-	2.66E-18
10	1	150860471	rs7412746	-	-	-	-	-	-	-	1.22E-08	-	4.29E-14	-
10	1	150938571	rs8444	8.22E-28	-	3.89E-14	-	-	1.45E-18	6.56E-23	-	1.85E-15	-	-
11	1	154994978	rs76798800	3.98E-26	-	3.86E-15	-	5.13E-10	3.55E-14	1.29E-22	5.96E-12	8.65E-18	2.35E-15	5.71E-16
12	1	204521577	rs12119098	-	-	-	-	-	-	-	-	-	8.16E-09	-
12	1	204618861	rs12125207	-	-	-	-	-	-	-	4.96E-08	-	-	-
13	1	205130413	rs3851294	-	-	-	-	-	-	1.76E-24	-	5.14E-30	-	8.88E-22
13	1	205181062	rs2369633	1.4E-38	1.41E-82	8.13E-09	1.41E-82	-	8.45E-34	-	6.52E-32	-	3.26E-09	-

14	1	214673271	rs7533482	-	-	-	-	-	-	-	8.43E-09	-	-	-
15	1	226603635	rs2695237	-	-	2.12E-18	-	-	-	-	-	-	1.81E-14	-
15	1	226608104	rs1858550	-	-	-	-	-	-	-	1.93E-10	-	-	-
16	2	25778637	rs12473635	-	-	2.86E-09	-	-	-	-	-	-	1.01E-08	-
17	2	27741237	rs780094	-	-	-	-	-	-	4.24E-09	-	4.22E-08	-	4.86E-08
18	2	28600617	rs6547849	-	-	-	-	-	-	-	1.03E-12	-	-	-
19	2	38298139	rs1800440	1.64E-45	-	4.84E-14	-	-	1.25E-29	1.72E-32	1.05E-13	1.26E-22	1.92E-17	1.62E-24
20	2	119601637	rs34897436	-	-	-	-	-	-	-	1.22E-08	-	-	-
21	2	202122995	rs3769823	-	-	-	-	-	-	-	-	4.01E-08	1.19E-10	-

21	2	202143928	rs10931936	-	-	1.71E-11	-	-	-	-	5.84E-10	-	-	-
22	2	222051419	rs10169459	-	-	-	-	-	-	-	1.43E-22	3.87E-09	-	-
23	2	222254215	rs34737625	-	-	-	-	-	-	-	1.37E-11	-	-	-
24	2	240074584	rs12617027	-	-	-	-	-	-	7.93E-09	-	-	-	-
24	2	240081540	rs3791512	1.36E-08	-	-	-	-	-	-	-	-	-	-
25	3	69936430	rs141211247	-	-	-	-	-	-	2.71E-10	-	2.64E-11	-	-
25	3	70397087	rs189911042	-	-	-	-	-	-	-	-	-	-	7.28E-10
26	3	69950451	rs183783391	-	-	-	-	-	-	-	2.57E-12	-	1.96E-10	-
26	3	70014091	rs149617956	3.27E-24	4.3E-31	2.06E-24	4.3E-31	6.02E-15	1.21E-20	-	-	-	-	-

27	3	122535198	rs13314128	-	-	-	-	-	-	-	1.19E-08	-	-	-
28	3	141133450	rs1991431	-	-	-	-	-	-	-	6.05E-20	-	-	-
28	3	141140968	rs9846396	-	-	-	-	-	-	-	-	6.57E-09	-	-
29	3	156491160	rs9867857	-	-	-	-	-	-	-	-	-	-	4.07E-14
29	3	156492758	rs9818780	-	-	-	-	-	-	2.8E-15	-	3.46E-12	-	-
30	3	169492101	rs10936599	-	-	-	-	-	-	-	8.15E-09	-	6.3E-09	-
30	3	169493283	rs3950296	-	-	8.51E-11	-	-	-	-	-	-	-	-
31	4	75341885	rs6819541	-	-	-	-	-	-	-	1.28E-08	-	-	-
32	4	79283091	rs67458307	-	-	-	-	-	-	-	1.21E-08	-	-	-

33	4	81661426	rs72661730	-	-	-	-	-	-	-	1.11E-13	-	-	-
34	5	1323212	rs13178866	-	6.52E-24	7.85E-44	6.52E-24	-	-	3.73E-31	6.02E-39	2.43E-27	2.5E-35	2.84E-22
35	5	33946571	rs35407	-	-	-	-	-	-	-	-	-	1.25E-147	-
35	5	33951693	rs16891982	0	0	1.44E-112	0	8.28E-39	0	-	-	-	-	-
35	5	33955673	rs35391	-	-	-	-	-	-	-	3.37E-289	-	-	-
35	5	33960698	rs34675	-	-	-	-	-	-	-	-	2.59E-301	-	-
35	5	33962877	rs1423676	-	-	-	-	-	-	-	-	-	-	5.75E-171
35	5	33963333	rs26721	-	-	-	-	-	-	3.66E-300	-	-	-	-
36	5	53112624	rs6875907	-	-	-	-	-	-	-	-	1.92E-08	-	-
36	5	53127645	rs12518752	-	-	-	-	-	-	-	2E-14	-	-	-

37	5	56007339	rs7709971	-	-	-	-	-	-	-	4.74E-10	-	-	-
38	5	59018442	rs11738977	-	-	-	-	-	-	1.24E-09	-	1.2E-08	-	-
39	5	90262612	rs12523094	5.37E-12	1.37E-78	5.85E-13	1.37E-78	-	4.39E-20	-	-	-	5.29E-12	-
39	5	90263863	rs6879563	-	-	-	-	-	-	2.67E-13	2.08E-36	1.72E-19	-	6.02E-18
40	5	149191111	rs4235745	-	-	-	-	-	-	-	-	9.92E-26	-	3.04E-31
40	5	149196090	rs251465	-	-	-	-	-	-	5.77E-35	-	-	-	-
40	5	149211868	rs32578	2.39E-45	-	6.91E-16	-	-	1.73E-39	-	7.01E-11	-	1.1E-23	-
41	5	176731452	rs28362590	-	-	-	-	-	-	9.43E-09	-	-	-	-
42	6	237332	rs4959612	-	-	-	-	-	-	4.17E-19	-	-	-	-

43	6	396321	rs12203592	-	-	1.91E-13	-	-	-	-	-	1.35E-37	4.29E-102	2.41E-239
43	6	475489	rs12210050	-	-	-	-	-	-	5.8E-189	1.36E-136	-	-	-
44	6	836203	rs11243206	-	-	-	-	-	-	1.01E-08	-	-	-	-
45	6	1144979	rs57354745	-	-	-	-	-	-	-	-	-	-	3.07E-13
45	6	1145265	rs12215602	5.34E-22	-	1.78E-08	-	-	4.22E-15	1.09E-16	-	4.3E-10	3.36E-10	-
46	6	20640316	rs9465851	-	-	-	-	-	-	-	1.96E-09	-	-	-
47	6	21163919	rs6914598	8.31E-18	-	1.18E-18	-	3.46E-15	-	9.67E-20	6.72E-11	1.25E-16	4.3E-17	1.8E-14
48	6	22710500	rs16886790	-	-	-	-	-	-	-	-	-	2.62E-08	-
48	6	22719379	rs72834823	-	-	1.04E-12	-	-	-	-	-	-	-	-

49	6	41833823	rs9357370	-	-	-	-	-	-	-	-	3.34E-08	-	-
49	6	41913778	rs4714520	-	-	-	-	-	-	1.27E-09	-	-	-	-
50	6	51694906	rs10948654	-	-	-	-	-	-	-	-	8.02E-11	-	4.11E-08
51	6	91005743	rs6908626	-	-	3.92E-09	-	-	-	-	-	-	-	-
52	6	151579432	rs4869723	-	-	-	-	-	-	7.23E-10	5.18E-20	3.25E-11	-	3.86E-10
53	6	159233043	rs9347258	-	-	-	-	-	-	-	5.95E-24	-	-	-
53	6	159235343	rs7761544	-	-	-	-	-	-	-	-	2.68E-11	-	-
54	7	16867772	rs1589886	-	-	-	-	-	-	-	-	-	3.84E-08	-
55	7	16996528	rs73069846	-	-	2.51E-08	-	-	-	-	-	-	5.18E-10	-

56	7	17134708	rs117132860	2.85E-110	-	3.83E-21	-	-	7.51E-63	4.05E-69	1.94E-14	4.34E-39	1.26E-26	9.28E-46
57	7	17571392	rs76633003	-	-	-	-	-	-	1.22E-09	-	-	-	-
57	7	17578637	rs12699867	-	-	-	-	-	-	-	-	8.13E-09	-	-
57	7	17579065	rs35257307	-	-	-	-	-	-	-	-	-	-	6.92E-09
58	7	22115454	rs12539524	-	-	-	-	-	-	-	2.62E-08	-	-	-
58	7	22121163	rs12533015	-	-	2.36E-08	-	-	-	-	-	-	4.61E-08	-
59	7	28189411	rs1635852	-	-	-	-	-	-	3.07E-09	-	-	-	-
60	7	124396645	rs4731207	-	-	1.99E-15	-	-	-	1.08E-09	1E-10	-	1.35E-12	1.19E-09
60	7	124436139	rs28523990	-	-	-	-	-	-	-	-	3.92E-10	-	-

61	7	130738666	rs7778378	-	-	1.86E-08	-	-	-	-	-	-	-	-
61	7	130742066	rs7803075	-	1.26E-76	-	1.26E-76	-	3.1E-09	-	1.02E-30	1.29E-15	-	9.17E-10
62	8	21951009	rs6994183	-	-	1.29E-09	-	-	-	-	-	-	-	-
63	8	72864240	rs13263376	3.62E-10	-	3.24E-12	-	-	-	3.15E-12	-	1.43E-08	2.04E-11	1.92E-10
64	8	116610180	rs2737205	-	-	-	-	-	-	1.38E-11	-	-	-	-
65	9	205964	rs478882	-	-	-	-	-	-	-	7.42E-19	-	-	-
66	9	12582787	rs10429629	-	-	-	-	-	-	-	-	-	2.44E-13	-
66	9	12587153	rs10960710	-	-	8.41E-13	-	-	-	-	-	-	-	-
66	9	12600284	rs10809803	2.47E-70	1.32E-109	-	1.32E-109	-	5.18E-91	-	-	-	-	-

66	9	12668805	rs10960748	-	-	-	-	-	-	2.58E-45	-	3.1E-54	-	-
66	9	12696499	rs2762461	-	-	-	-	-	-	-	4.38E-45	-	-	1.28E-52
67	9	16780344	rs12344799	-	-	-	-	-	-	-	-	-	2.06E-09	-
68	9	16885017	rs12350739	-	-	-	-	-	-	1.48E-65	4.97E-17	2.3E-56	-	2.16E-64
69	9	21803880	rs871024	-	-	1.24E-64	-	-	-	2.4E-28	8.12E-47	1.98E-29	9.02E-52	3.55E-30
70	9	22003223	rs3217992	-	-	-	-	-	-	-	-	5.54E-13	-	-
70	9	22042086	rs12352425	-	-	-	-	-	-	1.07E-10	-	-	-	-
70	9	22060833	rs77283072	-	-	-	-	-	-	-	-	-	-	6.81E-13
71	9	109054417	rs10739220	-	1.35E-26	4.4E-18	1.35E-26	-	-	-	4.04E-24	-	-	-

71	9	109060830	rs10739221	-	-	-	-	-	-	1.15E-11	-	1.84E-15	1.38E-14	1.76E-14
72	9	110711586	rs1339759	-	-	1.87E-17	-	-	-	-	-	-	-	-
72	9	110712028	rs10979125	-	-	-	-	-	-	-	8.49E-11	-	1.31E-15	-
73	9	126808006	rs58979150	-	-	-	-	-	-	-	1.78E-23	-	-	-
73	9	126814579	rs4838097	-	-	-	-	-	-	-	-	1.14E-08	-	-
74	9	126991185	rs10818930	-	-	-	-	-	-	-	2.14E-08	-	-	-
75	9	134457580	rs3780269	-	-	4.43E-08	-	-	-	-	-	-	-	-
76	10	80944095	rs703979	-	-	-	-	-	-	-	1.35E-11	-	-	-
77	10	105694301	rs7902587	-	-	1.38E-22	-	-	-	1.32E-13	2.83E-15	2.18E-11	3.36E-18	3.08E-15

78	10	111920995	rs12260509	-	-	-	-	-	-	-	6.15E-09	-	-	-
79	10	119572403	rs35563099	-	-	-	-	-	-	-	-	-	-	2.28E-27
79	10	119573178	rs7098111	-	-	-	-	-	-	6.38E-35	-	2.61E-20	5.35E-10	-
80	11	16041305	rs7941496	2.03E-15	2.97E-09	1.75E-09	2.97E-09	-	1.5E-20	-	2.01E-10	1.4E-14	1.25E-12	2.62E-16
80	11	16220882	rs6486277	-	-	-	-	-	-	9.14E-14	-	-	-	-
81	11	16587580	rs4578351	-	-	-	-	-	-	-	-	3.75E-09	-	-
82	11	62206288	rs9645690	-	-	-	-	-	-	-	1.5E-14	-	-	-
83	11	65561369	rs12797706	-	-	-	-	-	-	-	3.8E-10	-	-	-
84	11	68831364	rs72928978	-	-	-	-	-	-	1.13E-46	2.52E-111	1.14E-93	-	4.7E-66

85	11	69380898	rs4354713	-	2.4E-08	2.6E-21	2.4E-08	3.6E-29	-	4.73E-08	-	-	-	-
85	11	69386962	rs660963	-	-	-	-	-	-	-	-	-	1.19E-18	-
86	11	78126162	rs11237486	-	-	-	-	-	-	-	8.01E-20	2.3E-15	-	9.23E-11
87	11	89011046	rs1393350	-	-	-	-	-	-	4.92E-290	-	-	-	6.46E-260
87	11	89017961	rs1126809	0	5.55E-200	3.36E-71	5.55E-200	-	0	-	3.51E-127	3.24E-274	3.36E-133	-
88	11	90901865	rs10830671	-	-	-	-	-	-	-	-	2.32E-08	-	-
89	11	91616691	rs12225068	-	-	-	-	-	-	-	-	-	-	3.29E-08
90	11	108168995	rs4988023	-	-	-	-	-	-	1.04E-14	-	5.71E-13	-	-
90	11	108175462	rs1801516	-	-	2.33E-21	-	-	-	-	1.23E-16	-	4.3E-19	3.78E-12

91	11	120195702	rs12290699	-	-	4.17E-08	-	-	-	-	-	-	1.35E-08	-
92	12	4317819	rs11834692	-	-	-	-	-	-	-	2.28E-09	-	-	-
93	12	13060788	rs1861183	-	-	-	-	-	-	-	6.87E-10	-	-	-
93	12	13070752	rs1056927	-	-	1.01E-14	-	-	-	-	-	-	-	-
93	12	13072952	rs1684387	-	-	-	-	-	-	-	-	-	4.1E-14	-
94	12	17275460	rs4237963	-	-	1.99E-10	-	-	-	-	-	-	-	-
95	12	86604785	rs113913676	-	-	-	-	-	-	-	-	1.4E-08	-	-
95	12	89235648	rs12822439	-	-	-	-	-	-	-	-	-	-	1.62E-16
95	12	89328335	rs12821256	-	-	-	-	-	-	-	6.15E-121	2.13E-36	-	-

96	12	88790241	rs77404262	-	-	-	-	-	-	1.29E-10	-	-	-	-
97	12	96374750	rs2270318	-	-	-	-	-	-	-	-	-	2.67E-17	-
97	12	96378807	rs10859996	-	-	9.35E-10	-	-	-	-	-	-	-	-
98	12	116580291	rs113469387	2.61E-19	3.17E-16	7.05E-09	3.17E-16	-	1.23E-14	3.93E-16	1.22E-13	-	3.49E-08	-
98	12	116675052	rs61935859	-	-	-	-	-	-	-	-	-	-	1.09E-13
98	12	116702506	rs61937385	-	-	-	-	-	-	-	-	1.83E-11	-	-
99	13	78391757	rs1279403	-	-	-	-	-	-	-	1.1E-44	-	-	-
99	13	78408695	rs1933432	-	-	-	-	-	-	-	-	1.2E-12	-	-
100	13	95170420	rs4470024	-	-	-	-	-	-	-	8.08E-18	2.43E-15	-	3.2E-16

101	13	113533594	rs1278763	4.32E-65	-	-	-	-	5.43E-50	-	-	-	-	-
101	13	113533653	rs1765762	-	-	-	-	-	-	1.42E-38	-	4.61E-28	-	1.56E-31
101	13	113539523	rs1063098	-	-	-	-	-	-	-	6.79E-11	-	-	-
102	13	113535949	rs1278768	-	-	5.02E-12	-	-	-	-	-	-	-	-
103	13	113539894	rs1046793	-	-	-	-	-	-	-	-	-	6.31E-18	-
104	14	64390030	rs10873172	-	1.46E-27	-	1.46E-27	-	1.74E-08	-	-	-	-	-
104	14	64400120	rs12881652	-	-	-	-	-	-	-	1.54E-14	1.05E-09	-	1.8E-08
105	14	69226931	rs11625064	-	-	-	-	-	-	-	1.27E-11	-	-	-
106	14	92651679	rs150212431	-	-	-	-	-	-	-	3.28E-08	-	-	-

107	14	92777462	rs1885194	-	-	-	-	-	-	-	4.13E-274	5.61E-129	-	-
107	14	92787761	rs35983729	-	-	-	-	-	-	-	-	-	-	6.52E-76
108	14	92795912	rs4904871	-	-	-	-	-	-	1.97E-23	-	-	-	-
109	14	92954757	rs7493127	-	-	-	-	-	-	-	1.74E-09	-	-	-
110	14	93014929	rs61975764	-	-	-	-	-	-	-	4.62E-08	-	-	-
111	14	103878882	rs4906321	-	-	-	-	-	-	-	-	2.89E-08	3.18E-08	-
111	14	103923475	rs2273699	-	-	-	-	-	-	-	1.25E-09	-	-	-
112	15	28080692	rs11852452	-	-	-	-	-	-	-	-	-	3.89E-08	-
113	15	28206623	rs141514981	1.59E-47	5.09E-117	-	5.09E-117	-	1.13E-10	-	-	-	-	-

114	15	28365618	rs12913832	0	0	7.54E-23	0	2.1E-76	0	9.6E-283	-	-	1.87E-48	-
114	15	28370389	rs61585051	-	-	-	-	-	-	-	1.32992097016e-310		-	-
114	15	28460620	rs61460557	-	-	-	-	-	-	-	-	2.18E-307	-	-
114	15	28507825	rs10162958	-	-	-	-	-	-	-	-	-	-	0
115	15	29021242	rs4966335	-	-	-	-	-	-	2.94E-11	-	-	-	-
115	15	29033909	rs56166703	-	-	-	-	-	-	-	3.09E-49	8.02E-34	-	1.54E-31
116	15	29209775	rs60959391	-	-	-	-	-	-	-	-	1.72E-08	-	-
116	15	29215186	rs72625147	-	-	-	-	-	-	-	3.04E-10	-	-	-
117	15	29399597	rs902274	-	-	-	-	-	-	-	4.19E-08	-	-	-
118	15	33277710	rs117648907	-	-	1.77E-12	-	-	-	1.3E-08	-	-	9.16E-11	3.89E-10

119	15	81533762	rs61219147	-	-	-	-	-	-	-	1.95E-18	1.1E-08	-	-
120	16	4431202	rs758044	-	-	-	-	-	-	-	-	-	-	3.78E-08
121	16	54118132	rs62034121	-	-	4.08E-09	-	-	-	-	1.17E-09	-	-	-
122	16	68688677	rs12923069	-	-	-	-	-	-	7.38E-11	2.82E-11	1.63E-08	-	5.37E-09
122	16	68822971	rs4420522	-	-	4.25E-14	-	-	-	-	-	-	7.39E-11	-
123	16	71856611	rs4788445	-	-	-	-	-	-	-	-	2.69E-08	-	-
124	16	82188801	rs2911423	-	-	-	-	-	-	-	-	-	6.64E-09	-
124	16	82217153	rs2967383	-	-	7.96E-10	-	-	-	-	-	-	-	1.76E-09
125	16	89654800	rs35251956	-	-	-	-	-	-	1.19774333566771e-310			-	-

125	16	89674320	rs376402	-	-	-	-	-	-	-	-	-	-	1.59E-306
125	16	89691769	rs12930346	-	-	-	-	-	-	-	-	1.14E-306	-	-
125	16	89859753	rs12921383	-	-	-	-	-	-	-	-	-	6.19E-294	-
125	16	89986117	rs1805007	0	0	1.95E-195	0	0	0	-	-	-	-	-
125	16	90024202	rs8049897	-	-	-	-	-	-	-	1.76E-304	-	-	-
126	16	89743627	rs116927526	1.12E-160	1.81E-43	-	1.81E-43	4.99e-308	1.94E-139	-	-	-	-	-
127	16	89985918	rs1805006	-	3.04E-93	-	3.04E-93	-	-	-	-	-	-	-
128	16	89986144	rs1805008	0	0	-	0	0	0	-	-	-	-	-
129	16	89986546	rs1805009	0	3.11E-133	-	3.11E-133	0	0	-	-	-	-	-
130	17	7571752	rs78378222	-	-	3.33E-10	-	-	-	-	-	-	2.89E-10	-

131	17	45938105	rs72833461	-	-	-	-	-	-	-	-	2.63E-14	-	3.6E-08
131	17	45941289	rs112427233	-	-	-	-	-	-	-	4.65E-27	-	-	-
132	17	79530993	rs8070929	-	-	-	-	-	-	-	6.81E-20	1.52E-19	-	-
133	19	3353622	rs34466956	-	-	-	-	-	-	-	-	-	3E-08	-
134	19	3540539	rs12984831	3.86E-15	-	3.86E-10	-	-	3.15E-10	-	-	-	-	-
134	19	3542983	rs12608592	-	-	-	-	-	-	9.83E-16	5.84E-09	6.67E-13	5.85E-10	3.06E-12
135	20	25590744	rs242132	-	-	-	-	-	-	1.54E-11	-	1.21E-12	-	6.1E-10
136	20	31187504	rs751199	-	-	2.69E-09	-	-	-	-	-	-	-	-
137	20	32665748	rs6059655	0	0	2.79E-96	0	0	0	-	1.88E-261	-	1.28E-165	-
137	20	33642396	rs62213709	-	-	-	-	-	-	7.36e-308	-	3.27E-306	-	-

137	20	33943897	rs62211528	-	-	-	-	-	-	-	-	-	-	0
138	20	57843358	rs6123874	-	-	-	-	-	-	-	1.22E-10	-	-	-
139	20	62291767	rs143190905	-	-	6.54E-13	-	-	-	-	1.48E-08	-	2.88E-11	-
140	21	42743327	rs443099	-	-	-	-	-	-	-	-	-	-	8.22E-27
140	21	42743496	rs408825	5.98E-21	-	4.1E-31	-	-	1.27E-17	1.33E-28	-	-	7.32E-34	-
141	21	42745414	rs416981	-	-	-	-	-	-	-	4.03E-24	2.74E-24	-	-
142	21	44781984	rs559480	-	-	-	-	-	-	-	-	2.85E-08	-	-
143	22	38545942	rs132941	-	-	2.24E-22	-	-	-	8.08E-29	3.28E-18	4.3E-23	-	-
143	22	38602140	rs11914181	6.63E-29	-	-	-	-	6.02E-30	-	-	-	4.85E-22	6.41E-29
144	22	45622684	rs5766565	2.44E-54	4.92E-10	5.93E-10	4.92E-10	-	7.84E-47	8.83E-32	1.16E-10	4.93E-25	1.83E-15	3.97E-28

145	22	50722408	rs79966207	-	1.75E-19	9.56E-09	1.75E-19	-	-	-	7.24E-14	-	-	-	

Table G.1: Overlap in independent signals identified in joint-analyses of pigmentation and melanoma.

G.2 Pathway information

locus	CHR	BP	RS	Previous melanoma associations	Previous melanoma SNPs	Previous nevus associations	Previous nevus SNPs	Previous telomere associations	Previous telomere SNPs
1	1	8204596	rs78924983	-	-	-	-	-	-
2	1	11037434	rs112115136	-	-	-	-	-	-
3	1	24770594	rs2294524	-	-	-	-	-	-
4	1	25262022	rs1005734	-	-	-	-	-	-
5	1	63727542	rs670318	-	-	-	-	-	-
6	1	66895085	rs1613999	-	-	-	-	-	-
7	1	78086718	rs71641308	-	-	-	-	-	-
8	1	110720400	rs6689641	-	-	-	-	-	-
9	1	120466108	rs2793830	-	-	Nevus count or cutaneous melanoma	rs2453042	-	-
10	1	150800117	rs2089081	Nevus count or cutaneous melanoma, Melanoma	rs72704658, rs7412746, rs1722784	Nevus count or cutaneous melanoma	rs72704658, rs7412746	-	-

10	1	150860471	rs7412746	Nevus count or cutaneous melanoma, Melanoma	rs72704658, rs7412746, rs1722784	Nevus count or cutaneous melanoma	rs72704658, rs7412746	-	-
10	1	150938571	rs8444	Nevus count or cutaneous melanoma, Melanoma	rs72704658, rs7412746, rs1722784	Nevus count or cutaneous melanoma	rs72704658, rs7412746	-	-
11	1	154994978	rs76798800	-	-	Nevus count or cutaneous melanoma	rs76798800	-	-
12	1	204521577	rs12119098	-	-	-	-	-	-
12	1	204618861	rs12125207	-	-	-	-	-	-
13	1	205130413	rs3851294	-	-	Nevus count or cutaneous melanoma	rs11240396	-	-
13	1	205181062	rs2369633	-	-	Nevus count or cutaneous melanoma	rs11240396	-	-

14	1	214673271	rs7533482	-	-	-	-	-	-
15	1	226603635	rs2695237	Melanoma, Nevus count or cutaneous melanoma	rs3219090, rs2695237	Nevus count or cutaneous melanoma	rs747657, rs2695237	Leukocyte telomere length, Telomere length	rs3219104, rs1151814
15	1	226608104	rs1858550	Melanoma, Nevus count or cutaneous melanoma	rs3219090, rs2695237	Nevus count or cutaneous melanoma	rs747657, rs2695237	Leukocyte telomere length, Telomere length	rs3219104, rs1151814
16	2	25778637	rs12473635	-	-	Nevus count or cutaneous melanoma	rs61712781	-	-
17	2	27741237	rs780094	-	-	-	-	-	-
18	2	28600617	rs6547849	-	-	-	-	-	-

19	2	38298139	rs1800440	Cutaneous malig-nt melanoma, Melanoma, Nevus count or cutaneous melanoma	rs6750047, rs4670813	Nevus count or cutaneous melanoma, Nevus count	rs1800440, rs4670813	-	-
20	2	119601637	rs34897436	-	-	-	-	-	-
21	2	202122995	rs3769823	-	-	-	-	-	-
21	2	202143928	rs10931936	-	-	-	-	-	-
22	2	222051419	rs10169459	-	-	-	-	-	-
23	2	222254215	rs34737625	-	-	-	-	-	-
24	2	240074584	rs12617027	-	-	-	-	-	-
24	2	240081540	rs3791512	-	-	-	-	-	-
25	3	69936430	rs141211247	-	-	Nevus count or cutaneous melanoma	rs2320172	-	-

25	3	70397087	rs189911042	-	-	Nevus count or cutaneous melanoma	rs2320172	-	-
26	3	69950451	rs183783391	-	-	Nevus count or cutaneous melanoma	rs2320172	-	-
26	3	70014091	rs149617956	-	-	Nevus count or cutaneous melanoma	rs2320172	-	-
27	3	122535198	rs13314128	-	-	-	-	-	-
28	3	141133450	rs1991431	-	-	-	-	-	-
28	3	141140968	rs9846396	-	-	-	-	-	-
29	3	156491160	rs9867857	-	-	-	-	-	-
29	3	156492758	rs9818780	-	-	-	-	-	-

30	3	169492101	rs10936599	Melanoma, Nevus count or cutaneous melanoma, Cutaneous malig-nt melanoma	rs13097028, rs12696304	Nevus count, Nevus count or cutaneous melanoma	rs12696304, rs10936600	Telomere length, Leuko-cyte telomere length	rs12638862, rs12696304, rs2293607, rs10936599, rs1317082, rs10936600, rs10936601
30	3	169493283	rs3950296	Melanoma, Nevus count or cutaneous melanoma, Cutaneous malig-nt melanoma	rs13097028, rs12696304	Nevus count, Nevus count or cutaneous melanoma	rs12696304, rs10936600	Telomere length, Leuko-cyte telomere length	rs12638862, rs12696304, rs2293607, rs10936599, rs1317082, rs10936600, rs10936601
31	4	75341885	rs6819541	-	-	-	-	-	-
32	4	79283091	rs67458307	-	-	-	-	-	-
33	4	81661426	rs72661730	-	-	-	-	-	-

34	5	1323212	rs13178866	Melanoma, Nevus count or cutaneous melanoma, Uveal melanoma	rs139996880, rs380286, rs401681, rs452932	Nevus count or cutaneous melanoma	rs380286, rs2447853	Telomere length, Leukocyte telomere length	rs7705526, rs2736100, rs2853677
35	5	33946571	rs35407	Melanoma	rs35407, rs16891982, rs35390	Nevus count or cutaneous melanoma	rs185146	-	-
35	5	33951693	rs16891982	Melanoma	rs35407, rs16891982, rs35390	Nevus count or cutaneous melanoma	rs185146	-	-
35	5	33955673	rs35391	Melanoma	rs35407, rs16891982, rs35390	Nevus count or cutaneous melanoma	rs185146	-	-
35	5	33960698	rs34675	Melanoma	rs35407, rs16891982, rs35390	Nevus count or cutaneous melanoma	rs185146	-	-

35	5	33962877	rs1423676	Melanoma	rs35407, rs16891982, rs35390	Nevus count or cutaneous melanoma	rs185146	-	-
35	5	33963333	rs26721	Melanoma	rs35407, rs16891982, rs35390	Nevus count or cutaneous melanoma	rs185146	-	-
36	5	53112624	rs6875907	-	-	-	-	-	-
36	5	53127645	rs12518752	-	-	-	-	-	-
37	5	56007339	rs7709971	-	-	-	-	-	-
38	5	59018442	rs11738977	-	-	-	-	-	-
39	5	90262612	rs12523094	-	-	-	-	-	-
39	5	90263863	rs6879563	-	-	-	-	-	-
40	5	149191111	rs4235745	-	-	-	-	-	-
40	5	149196090	rs251465	-	-	-	-	-	-
40	5	149211868	rs32578	-	-	-	-	-	-
41	5	176731452	rs28362590	-	-	-	-	-	-
42	6	237332	rs4959612	Melanoma	rs62389423	Nevus count	rs12203592	-	-
43	6	396321	rs12203592	Melanoma	rs62389423	Nevus count	rs12203592	-	-
43	6	475489	rs12210050	Melanoma	rs62389423	Nevus count	rs12203592	-	-

44	6	836203	rs11243206	Melanoma	rs62389423	Nevus count, Nevus count or cutaneous melanoma	rs12203592, rs10214796	-	-
45	6	1144979	rs57354745	-	-	Nevus count or cutaneous melanoma	rs10214796	-	-
45	6	1145265	rs12215602	-	-	Nevus count or cutaneous melanoma	rs10214796	-	-
46	6	20640316	rs9465851	-	-	-	-	-	-
47	6	21163919	rs6914598	Cutaneous malig-nt melanoma, Melanoma	rs6914598	Nevus count or cutaneous melanoma	rs6914598	-	-
48	6	22710500	rs16886790	-	-	Nevus count or cutaneous melanoma	rs16886790	-	-

48	6	22719379	rs72834823	-	-	Nevus count or cutaneous melanoma	rs16886790	-	-
49	6	41833823	rs9357370	-	-	-	-	-	-
49	6	41913778	rs4714520	-	-	-	-	-	-
50	6	51694906	rs10948654	-	-	-	-	-	-
51	6	91005743	rs6908626	-	-	-	-	-	-
52	6	151579432	rs4869723	-	-	Nevus count or cutaneous melanoma	rs10434896	-	-
53	6	159233043	rs9347258	-	-	-	-	-	-
53	6	159235343	rs7761544	-	-	-	-	-	-
54	7	16867772	rs1589886	Cutaneous malig-nt melanoma, Nevus count or cutaneous melanoma, Melanoma	rs1636744	Nevus count or cutaneous melanoma	rs1636744, rs117132860	-	-

55	7	16996528	rs73069846	Cutaneous malig-nt melanoma, Nevus count or cutaneous melanoma, Melanoma	rs1636744	Nevus count or cutaneous melanoma	rs1636744, rs117132860	-	-
56	7	17134708	rs117132860	Cutaneous malig-nt melanoma, Nevus count or cutaneous melanoma, Melanoma	rs1636744	Nevus count or cutaneous melanoma	rs1636744, rs117132860	-	-
57	7	17571392	rs76633003	-	-	Nevus count or cutaneous melanoma	rs117132860	-	-
57	7	17578637	rs12699867	-	-	Nevus count or cutaneous melanoma	rs117132860	-	-

57	7	17579065	rs35257307	-	-	Nevus count or cutaneous melanoma	rs117132860	-	-
58	7	22115454	rs12539524	-	-	-	-	-	-
58	7	22121163	rs12533015	-	-	-	-	-	-
59	7	28189411	rs1635852	-	-	-	-	-	-
60	7	124396645	rs4731207	Cutaneous malig-nt melanoma	rs4731207	Nevus count or cutaneous melanoma	rs113394869	Leukocyte telomere length	rs59294613
60	7	124436139	rs28523990	Cutaneous malig-nt melanoma	rs4731207	Nevus count or cutaneous melanoma	rs113394869	Leukocyte telomere length	rs59294613
61	7	130738666	rs7778378	Cutaneous malig-nt melanoma	rs4731742	Nevus count or cutaneous melanoma	rs78706986	-	-
61	7	130742066	rs7803075	Cutaneous malig-nt melanoma	rs4731742	Nevus count or cutaneous melanoma	rs78706986	-	-

62	8	21951009	rs6994183	-	-	Nevus count or cutaneous melanoma	rs56953221	-	-
63	8	72864240	rs13263376	-	-	-	-	Leukocyte telomere length	rs28365964
64	8	116610180	rs2737205	-	-	-	-	-	-
65	9	205964	rs478882	Nevus count or cutaneous melanoma	rs600951	Nevus count, Nevus count or cutaneous melanoma	rs600951, rs593179	-	-
66	9	12582787	rs10429629	-	-	Nevus count or cutaneous melanoma	rs13283146	-	-
66	9	12587153	rs10960710	-	-	Nevus count or cutaneous melanoma	rs13283146	-	-

66	9	12600284	rs10809803	-	-	Nevus count or cutaneous melanoma	rs13283146	-	-
66	9	12668805	rs10960748	-	-	Nevus count or cutaneous melanoma	rs13283146	-	-
66	9	12696499	rs2762461	-	-	Nevus count or cutaneous melanoma	rs13283146	-	-
67	9	16780344	rs12344799	-	-	-	-	-	-
68	9	16885017	rs12350739	-	-	-	-	-	-
69	9	21803880	rs871024	Nevus count or cutaneous melanoma, Melanoma	rs869329, rs201131773, rs7023329, rs11532907	Nevus count or cutaneous melanoma, Nevus count	rs869330, rs869329, rs11532907	-	-
70	9	22003223	rs3217992	Nevus count or cutaneous melanoma, Melanoma	rs869329, rs201131773, rs7023329, rs11532907	Nevus count or cutaneous melanoma, Nevus count	rs869330, rs869329, rs11532907	-	-

70	9	22042086	rs12352425	Nevus count or cutaneous melanoma, Melanoma	rs869329, rs201131773, rs7023329, rs11532907	Nevus count or cutaneous melanoma, Nevus count	rs869330, rs869329, rs11532907	-	-
70	9	22060833	rs77283072	Nevus count or cutaneous melanoma, Melanoma	rs869329, rs201131773, rs7023329, rs11532907	Nevus count or cutaneous melanoma, Nevus count	rs869330, rs869329, rs11532907	-	-
71	9	109054417	rs10739220	-	-	-	-	-	-
71	9	109060830	rs10739221	-	-	-	-	-	-
72	9	110711586	rs1339759	-	-	-	-	-	-
72	9	110712028	rs10979125	-	-	-	-	-	-
73	9	126808006	rs58979150	-	-	-	-	-	-
73	9	126814579	rs4838097	-	-	-	-	-	-
74	9	126991185	rs10818930	-	-	-	-	-	-
75	9	134457580	rs3780269	-	-	-	-	-	-
76	10	80944095	rs703979	-	-	-	-	-	-
77	10	105694301	rs7902587	Melanoma	rs17119461	-	-	-	-
78	10	111920995	rs12260509	-	-	-	-	-	-

79	10	119572403	rs35563099	-	-	-	-	-	-
79	10	119573178	rs7098111	-	-	-	-	-	-
80	11	16041305	rs7941496	-	-	Nevus count or cutaneous melanoma	rs2054095	-	-
80	11	16220882	rs6486277	-	-	Nevus count or cutaneous melanoma	rs2054095	-	-
81	11	16587580	rs4578351	-	-	Nevus count or cutaneous melanoma	rs2054095	-	-
82	11	62206288	rs9645690	-	-	-	-	-	-
83	11	65561369	rs12797706	-	-	-	-	-	-
84	11	68831364	rs72928978	Cutaneous malig-nt melanoma	rs2290419	-	-	-	-

85	11	69380898	rs4354713	Cutaneous malig-nt melanoma, Nevus count or cutaneous melanoma, Melanoma	rs2290419, rs498136	Nevus count or cutaneous melanoma	rs498136, rs9651783	-	-
85	11	69386962	rs660963	Cutaneous malig-nt melanoma, Nevus count or cutaneous melanoma, Melanoma	rs2290419, rs498136	Nevus count or cutaneous melanoma	rs498136, rs9651783	-	-
86	11	78126162	rs11237486	-	-	-	-	-	-
87	11	89011046	rs1393350	Melanoma, Nevus count or cutaneous melanoma	rs1847134, rs1393350, rs10830253	Nevus count or cutaneous melanoma	rs1126809, rs10830253	-	-

87	11	89017961	rs1126809	Melanoma, Nevus count or cutaneous melanoma	rs1847134, rs1393350, rs10830253	Nevus count or cutaneous melanoma	rs1126809, rs10830253	-	-
88	11	90901865	rs10830671	-	-	-	-	-	-
89	11	91616691	rs12225068	-	-	-	-	-	-
90	11	108168995	rs4988023	Melanoma, Nevus count or cutaneous melanoma	rs1801516, rs73008229	Nevus count or cutaneous melanoma	rs1801516, rs73008229	Leukocyte telomere length	rs228595, rs227080
90	11	108175462	rs1801516	Melanoma, Nevus count or cutaneous melanoma	rs1801516, rs73008229	Nevus count or cutaneous melanoma	rs1801516, rs73008229	Leukocyte telomere length	rs228595, rs227080
91	11	120195702	rs12290699	-	-	Nevus count or cutaneous melanoma	rs7102217	-	-
92	12	4317819	rs11834692	-	-	-	-	-	-

93	12	13060788	rs1861183	Nevus count or cutaneous melanoma, Cutaneous malig-nt melanoma	rs1148732, rs1640875, rs2111398	Nevus count or cutaneous melanoma, Nevus count	rs1061036, rs1148732, rs1640875	-	-
93	12	13070752	rs1056927	Nevus count or cutaneous melanoma, Cutaneous malig-nt melanoma	rs1148732, rs1640875, rs2111398	Nevus count or cutaneous melanoma, Nevus count	rs1061036, rs1148732, rs1640875	-	-
93	12	13072952	rs1684387	Nevus count or cutaneous melanoma, Cutaneous malig-nt melanoma	rs1148732, rs1640875, rs2111398	Nevus count or cutaneous melanoma, Nevus count	rs1061036, rs1148732, rs1640875	-	-
94	12	17275460	rs4237963	-	-	-	-	-	-

95	12	86604785	rs113913676	-	-	Nevus count or cutaneous melanoma, Nevus count	rs7487314, rs7313352	-	-
95	12	89235648	rs12822439	-	-	Nevus count or cutaneous melanoma, Nevus count	rs7487314, rs7313352	-	-
95	12	89328335	rs12821256	-	-	Nevus count or cutaneous melanoma, Nevus count	rs7487314, rs7313352	-	-
96	12	88790241	rs77404262	-	-	Nevus count or cutaneous melanoma, Nevus count	rs7487314, rs7313352	-	-
97	12	96374750	rs2270318	-	-	-	-	-	-
97	12	96378807	rs10859996	-	-	-	-	-	-

98	12	116580291	rs113469387	-	-	Nevus count or cutaneous melanoma	rs113469387	-	-
98	12	116675052	rs61935859	-	-	Nevus count or cutaneous melanoma	rs113469387	-	-
98	12	116702506	rs61937385	-	-	Nevus count or cutaneous melanoma	rs113469387	-	-
99	13	78391757	rs1279403	-	-	-	-	-	-
99	13	78408695	rs1933432	-	-	-	-	-	-
100	13	95170420	rs4470024	-	-	-	-	-	-
101	13	113533594	rs1278763	-	-	-	-	-	-
101	13	113533653	rs1765762	-	-	-	-	-	-
101	13	113539523	rs1063098	-	-	-	-	-	-
102	13	113535949	rs1278768	-	-	-	-	-	-
103	13	113539894	rs1046793	-	-	-	-	-	-

104	14	64390030	rs10873172	Nevus count or cutaneous melanoma	rs2357176	Nevus count or cutaneous melanoma	rs10873172, rs2357176	-	-
104	14	64400120	rs12881652	Nevus count or cutaneous melanoma	rs2357176	Nevus count or cutaneous melanoma	rs10873172, rs2357176	-	-
105	14	69226931	rs11625064	-	-	Nevus count or cutaneous melanoma	rs11625064	-	-
106	14	92651679	rs150212431	-	-	-	-	-	-
107	14	92777462	rs1885194	-	-	-	-	-	-
107	14	92787761	rs35983729	-	-	-	-	-	-
108	14	92795912	rs4904871	-	-	-	-	-	-
109	14	92954757	rs7493127	-	-	-	-	-	-
110	14	93014929	rs61975764	-	-	-	-	-	-
111	14	103878882	rs4906321	-	-	-	-	-	-
111	14	103923475	rs2273699	-	-	-	-	-	-

112	15	28080692	rs11852452	Uveal melanoma, Cutaneous malig-nt melanoma, Melanoma	rs11074306, rs4778138	Nevus count or cutaneous melanoma	rs12913832	-	-
113	15	28206623	rs141514981	Uveal melanoma, Cutaneous malig-nt melanoma, Melanoma	rs11074306, rs4778138	Nevus count or cutaneous melanoma	rs12913832	-	-
114	15	28365618	rs12913832	Cutaneous malig-nt melanoma, Melanoma	rs4778138	Nevus count or cutaneous melanoma	rs12913832	-	-
114	15	28370389	rs61585051	Cutaneous malig-nt melanoma, Melanoma	rs4778138	Nevus count or cutaneous melanoma	rs12913832	-	-

114	15	28460620	rs61460557	Cutaneous malig-nt melanoma, Melanoma	rs4778138	Nevus count or cutaneous melanoma	rs12913832	-	-
114	15	28507825	rs10162958	Cutaneous malig-nt melanoma, Melanoma	rs4778138	Nevus count or cutaneous melanoma	rs12913832	-	-
115	15	29021242	rs4966335	-	-	-	-	-	-
115	15	29033909	rs56166703	-	-	-	-	-	-
116	15	29209775	rs60959391	-	-	-	-	-	-
116	15	29215186	rs72625147	-	-	-	-	-	-
117	15	29399597	rs902274	-	-	-	-	-	-
118	15	33277710	rs117648907	Nevus count or cutaneous melanoma	rs117648907	Nevus count or cutaneous melanoma, Nevus count	rs150962800, rs117648907	-	-
119	15	81533762	rs61219147	-	-	-	-	-	-
120	16	4431202	rs758044	-	-	-	-	-	-

121	16	54118132	rs62034121	Melanoma, Nevus count or cutaneous melanoma	rs16953002, rs12596638	Nevus count or cutaneous melanoma	rs12596638, rs62034121	-	-
122	16	68688677	rs12923069	Cutaneous malig-nt melanoma	rs35158985	Nevus count or cutaneous melanoma	rs8046299	-	-
122	16	68822971	rs4420522	Cutaneous malig-nt melanoma	rs35158985	Nevus count or cutaneous melanoma	rs8046299	-	-
123	16	71856611	rs4788445	-	-	-	-	-	-
124	16	82188801	rs2911423	-	-	Nevus count or cutaneous melanoma	rs6564982	Leukocyte telomere length	rs7194734, rs2967374
124	16	82217153	rs2967383	-	-	Nevus count or cutaneous melanoma	rs6564982	Leukocyte telomere length	rs7194734, rs2967374

125	16	89654800	rs35251956	Melanoma, Nevus count or cutaneous melanoma	rs258322, rs75570604, rs1805007, rs4785763	Nevus count or cutaneous melanoma	rs75570604	-	-
125	16	89674320	rs376402	Melanoma, Nevus count or cutaneous melanoma	rs258322, rs75570604, rs1805007, rs4785763	Nevus count or cutaneous melanoma	rs75570604	-	-
125	16	89691769	rs12930346	Melanoma, Nevus count or cutaneous melanoma	rs258322, rs75570604, rs1805007, rs4785763	Nevus count or cutaneous melanoma	rs75570604	-	-
125	16	89859753	rs12921383	Melanoma, Nevus count or cutaneous melanoma	rs258322, rs75570604, rs1805007, rs4785763	Nevus count or cutaneous melanoma	rs75570604	-	-

125	16	89986117	rs1805007	Melanoma, Nevus count or cutaneous melanoma	rs258322, rs75570604, rs1805007, rs4785763	Nevus count or cutaneous melanoma	rs75570604	-	-
125	16	90024202	rs8049897	Melanoma, Nevus count or cutaneous melanoma	rs258322, rs75570604, rs1805007, rs4785763	Nevus count or cutaneous melanoma	rs75570604	-	-
126	16	89743627	rs116927526	Melanoma, Nevus count or cutaneous melanoma	rs258322, rs75570604, rs1805007, rs4785763	Nevus count or cutaneous melanoma	rs75570604	-	-
127	16	89985918	rs1805006	Melanoma, Nevus count or cutaneous melanoma	rs258322, rs75570604, rs1805007, rs4785763	Nevus count or cutaneous melanoma	rs75570604	-	-

128	16	89986144	rs1805008	Melanoma, Nevus count or cutaneous melanoma	rs258322, rs75570604, rs1805007, rs4785763	Nevus count or cutaneous melanoma	rs75570604	-	-
129	16	89986546	rs1805009	Melanoma, Nevus count or cutaneous melanoma	rs258322, rs75570604, rs1805007, rs4785763	Nevus count or cutaneous melanoma	rs75570604	-	-
130	17	7571752	rs78378222	-	-	Nevus count or cutaneous melanoma	rs78378222	-	-
131	17	45938105	rs72833461	-	-	-	-	-	-
131	17	45941289	rs112427233	-	-	-	-	-	-
132	17	79530993	rs8070929	-	-	-	-	-	-
133	19	3353622	rs34466956	Nevus count or cutaneous melanoma	rs34466956	Nevus count or cutaneous melanoma	rs34466956	-	-

134	19	3540539	rs12984831	Nevus count or cutaneous melanoma	rs34466956	Nevus count or cutaneous melanoma	rs34466956	Leukocyte telomere length	rs143276018
134	19	3542983	rs12608592	Nevus count or cutaneous melanoma	rs34466956	Nevus count or cutaneous melanoma	rs34466956	Leukocyte telomere length	rs143276018
135	20	25590744	rs242132	-	-	-	-	telomere length	rs73598373, rs67679582, rs8126111, rs80105553
136	20	31187504	rs751199	-	-	-	-	-	-
137	20	32665748	rs6059655	Melanoma	rs6059655	Nevus count or cutaneous melanoma	rs6059655	-	-
137	20	33642396	rs62213709	Melanoma	rs6059655	Nevus count or cutaneous melanoma	rs6059655	-	-

137	20	33943897	rs62211528	Melanoma	rs6059655	Nevus count or cutaneous melanoma	rs6059655	-	-
138	20	57843358	rs6123874	-	-	-	-	-	-
139	20	62291767	rs143190905	-	-	-	-	-	-
140	21	42743327	rs443099	-	-	-	-	-	-
140	21	42743496	rs408825	-	-	-	-	-	-
141	21	42745414	rs416981	-	-	-	-	-	-
142	21	44781984	rs559480	-	-	-	-	Leukocyte telomere length	rs7276273
143	22	38545942	rs132941	Nevus count or cutaneous melanoma, Melanoma	rs2005974, rs2284063, rs6001027, rs132985	Nevus count or cutaneous melanoma, Nevus count	rs2005974, rs132985, rs3761444	-	-
143	22	38602140	rs11914181	Nevus count or cutaneous melanoma, Melanoma	rs2005974, rs2284063, rs6001027, rs132985	Nevus count or cutaneous melanoma, Nevus count	rs2005974, rs132985, rs3761444	-	-

144	22	45622684	rs5766565	-	-	Nevus count or cutaneous melanoma	rs5766565	-	-
145	22	50722408	rs79966207	-	-	Nevus count or cutaneous melanoma	rs79966207	-	-

Table G.2: Previously identified associations with independent signals from joint-analysis for melanoma risk pathways