Genomics in Myeloma: Evaluating Technologies for a Consistent, Quality Assured Assessment of the Genetic Aberrations Associated with Myeloma, with Reference to Myeloma Bone Disease.

Polly Jane Talley

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

The University of Sheffield
Department of Oncology & Metabolism.

December 2020

Abstract / Summary

Plasma cell myeloma (PCM), also known as multiple myeloma (MM) or myeloma, is a cancer of the bone marrow. It is a neoplastic disorder characterised by an abnormal clonal proliferation of plasma cells in the bone marrow (BM) and the consequent overproduction of circulating monoclonal immunoglobulin (paraprotein). Myeloma is defined by the presence of ≥10% clonal plasma cells in the bone marrow, a paraprotein and the presence of end organ damage, including hypercalcaemia, renal insufficiency, anaemia and bone lesions.

Over the past two decades the treatment of myeloma has seen huge advances leading to significantly improved outcomes, specifically as a result of the introduction of new classes of therapeutic agents including immunomodulatory agents (IMiDs), proteasome inhibitors, monoclonal antibodies and improved stem cell transplantation techniques. It is usually possible to reduce disease load substantially and induce a remission phase in the patient's myeloma. However, complete elimination of residual disease is not possible, and the disease almost invariably begins to recrudesce resulting in relapse, and so, in the majority of cases, myeloma remains an incurable disease.

Bone disease is seen in approximately 70% of patients at diagnosis, but with a range of severity. It is estimated that over 90% of patients exhibit evidence of bone disease at some stage throughout the course of their disease.

Myeloma genetics is intrinsically complex and highly heterogeneous; there is no single, discrete genetic aberration that causes the typical phenotype seen in myeloma, rather a range of characteristic aberrations in key regions of the genome. Clonal evolution drives a tendency for myeloma to accumulate genetic aberrations over time. Despite this complexity, genetic analysis offers an opportunity to categorise the disease, offer prognosis based on these categorisations and potentially apply a personalised medicine approach to therapy. Access to diagnostic genetic testing in myeloma is *ad hoc* and highly dependent upon whether patients are entered into national trials and local commissioning arrangements, even then the techniques employed and the exact nature of the testing vary dramatically.

This project assessed the role of five different genetic techniques used in the diagnosis of myeloma; cytogenetic analysis, fluorescence in situ hybridisation (FISH), multiplex ligation-dependant probe amplification (MLPA), DNA array and targeted next generation sequencing

(NGS). These data, alongside current recommendations, were used to design and propose appropriate diagnostic testing strategies in myeloma, and to inform national best practice.

Alongside this work, and through collaboration with UKNEQAS (GenQA), a quality assessment and assurance programme was introduced for the genetic testing of myeloma patients, with the aim of not only providing quality assurance in this area, but in order to create a more consistent and equitable service through provision of an educational component to the scheme.

The final aim related to myeloma bone disease was to explore, using targeted NGS, a possible association between high risk variants in bone related genes and bone phenotype, which if proven, would allow a more personalised and targeted approach to the treatment of bone manifestations associated with myeloma.

Contents

Tables and Figures

Publication and Presentation List

During my time working on this project, I have completed a number of publications, spoken and poster presentations which are listed here:

Publication list:

Talley PJ, Chantry AD & Buckle CH. Genetics in Myeloma: Genetic Technologies and their Application to Screening Approaches in Myeloma.

Br Med Bull 2015; 113(1): 15-30

Rack KA, van den Berg E, Haferlach C, Beverloo HB, Costa D, Espinet B, Foot N, Jeffries S, Martin K, O'Connor S, Schoumans J, Talley PJ, Telford N, Stioui S, Zemanova Z & Hastings RJ. European recommendations and quality assurance for cytogenomic analysis of haematological neoplasms.

Leukemia 33:1851-1867, 2019

Presentation lists:

Identification, Validation and Translation of Therapeutic Targets in Myeloma Tumour Biology Using Conventional and Novel Genetic Technologies - Poster presentation for Mellanby Centre Research Day, Sheffield, December 2014

Clinical and Genetic Aspects of Myeloma - Spoken presentation at the UK Cancer Cytogenetic Group, Newcastle, March 2015

Results from the First Pilot CEQAS Scheme for Myeloma Genetic Diagnosis - Poster presented at the 15th International Myeloma Working Group (IMWG) Workshop, Rome, September 2015

MLPA in Plasma Cell Myeloma - Spoken presentation for the Mellanby Centre Seminar Series, Sheffield, February 2016

Route to Impact: PhD Studies Evaluating Genetic Technologies for Myeloma Testing Leads to the Introduction of a European External Quality Assessment Scheme (EQA) and Best Practice Guidelines - Poster presentation at the University of Sheffield, Medical School Research Day, Sheffield, June 2016

Myeloma Genetic Diagnosis for HODS - spoken presentation for the Haemato-oncology Diagnostic Service (HODS) educational meeting, Sheffield, September 2016

Genetic Diagnosis in Myeloma - techniques and application - spoken presentation as part of the confirmation review, Sheffield, October 2016

Myeloma Testing Information for HODS - spoken presentation for the Haemato-oncology Diagnostic Service (HODS) educational meeting, Sheffield, February 2017

Diagnostic and Prognostic Markers in Myeloma - spoken presentation at the HMDS Study Day, Leeds, June 2018

Introduction of a European Quality Assessment Scheme for Genetic Testing in Myeloma: Progress and Update - spoken presentation at the UK Myeloma Forum Autumn Meeting, London, November 2018

Abbreviations

ACGS	Association for Clinical Genetic Science
ALL	Acute Lymphoblastic Leukaemia
BM	Bone Marrow
BWA	Burrows-Wheeler Aligner (NGS alignment software)
CEQAS	Cytogenetic External Quality Assessment Scheme
CGH	Comparative Genome Hybridisation
Chas	Chromosome Analysis Suite (Affymetrix array analysis software)
DNA	Deoxyribonucleic Acid
FISH	Fluorescence in situ Hybridisation
GATK	Genome Analysis Toolkit (NGS analysis software)
GWAS	Genome Wide Association Studies
Ig	Immunoglobulin
ISCN	International System for Cytogenetic Nomenclature
LLR	Leukaemia and Lymphoma Research
LOH	Loss of Heterozygosity
MACS	Magnetic-Activated Cell Sorting
MGUS	Monoclonal Gammopathy of Undetermined Significance
MLPA	Multiplex Ligation-dependant Probe Amplification
MM	Multiple Myeloma
NGS	Next Generation Sequencing
PCL	Plasma Cell Leukaemia
PCR	Polymerase Chain Reaction
SBS	Sequencing By Synthesis
SDGS	Sheffield Diagnostic Genetics Service
SITraN	Sheffield Institute of Translational Neuroscience
SM	Smouldering Myeloma
SMaRT	Sheffield Myeloma Research Team
UPD	Uniparental Disomy
aUPD	Acquired Uniparental Disomy
VDJ	Variable / Diversity / Junctional Segments
WES	Whole Exome Sequencing
WGS	Whole Genome Sequencing

Chapter 1: Introduction

Plasma Cell Myeloma

Plasma cell myeloma is a neoplastic disorder characterised by an abnormal monoclonal proliferation of plasma cells in the bone marrow and overproduction of circulating monoclonal immunoglobulin product (paraprotein). Neoplastic plasma cells secrete osteoclast activating factors and osteoblast deactivating factors leading to destructive, osteolytic bone disease. Malignant plasma cells accumulate within the bone marrow effectively crowding out normal haemopoietic tissue leading to bone marrow failure manifesting as anaemia, thrombocytopaenia and leucopenia as well as impaired cellular and antibody-mediated immunity.

Plasma cell myeloma is defined by the presence of ≥10% clonal plasma cells in the bone marrow, increased paraprotein levels and the presence of end organ damage[1,2], which can be summarised by the acronym, CRAB.

- Hyper**C**alcaemia
- **R**enal insufficiency
- **A**naemia
- **B**one lesions

Hypercalcaemia results from bone destruction and is seen in 20% of patients at diagnosis[3]. Renal insufficiency, and ultimate failure, is due to tubular damage resulting from proteinuria. Anaemia is observed in approximately 67% of patients, and results from effacement of BM by plasma cells and renal damage resulting in loss of erythropoietin. Osteolytic disease is seen in 70% of patients with MM[1,4].

Plasma cell neoplasms progress through distinct clinical phases: monoclonal gammopathy of undetermined significance (MGUS), asymptomatic myeloma (AS) or smouldering myeloma (SM), plasma cell myeloma (MM), progressing to plasma cell leukaemia (PCL). Other clinical phases following treatment can include myeloma in remission (termed 'plateau phase'), relapsed myeloma and refractory myeloma (where disease is non-responsive to treatment).

Myeloma is the 19th most common cancer in the UK accounting for 2% of all new cases reported[5]. This equates to approximately 6,000 new patients being diagnosed with myeloma in the UK per year[5]. MM is a cancer of older adults: the majority of cases occur in patients over the age of 65. 26% of patients are aged between 65-74 years at diagnosis, and 44% are over 75. MM is diagnosed in younger people, but only 2% of cases are in patients below 40 years of age[5]. MM is not a disease that is seen in children[1]. Myeloma is likely to increase in incidence in the UK in line with the aging population, and over the last decade (2007-2017), myeloma incidence rates have increased by 15%, and future projections suggest a further rise of 11% by 2035[5]. Data suggests that myeloma is more common in men than in women (ratio 1.33:1), and more common in black people than in people of Asian or Caucasian descent (ratio ~2:1)[1,5]. There is a 2.3 fold increased risk to first degree family members of patients with MM, suggesting some inherited component[1,6]. Approximately 14% of myeloma cases in the UK are considered to be preventable and are linked with obesity[5].

Approximately 15% of MM patients have no symptoms at presentation, and a high paraprotein level may be discovered following routine screening. Approximately 40% present with more substantial morbidity including anaemia, renal failure and skeletal disease including pathological fractures, hypercalcaemia, spinal cord compression or generalised bone loss (osteoporosis). Less critical symptoms can include backache, bone pain, anaemia and tiredness. Diagnosis is dependent on results from a number of clinical tests, including full blood count and chemical analysis, serum and urine electrophoresis, bone marrow morphology, radiography and genetic analysis[1].

Treatment strategies are conventionally divided into intensive and non-intensive regimens. The former feature chemotherapy using a combination of steroids (e.g. dexamethasone), alkylating agents (e.g. cyclophosphamide) and immunomodulatory agents (e.g. thalidomide or lenalidamide) and /or a proteasome inhibitor (e.g. bortezomib (Velcade) or carfilzomib) for younger (<60 years), fitter patients, followed by autologous stem cell transplantation. High dose melphalan (i.e. 200mg/m^2) is used as a consolidation therapy prior to autologous stem-cell transplantation, further reducing plasma cell load within the bone marrow. Non-

intensive regimens comprise similar combinations used at attenuated doses and without autologous stem cell transplantation[4,7].

Immunomodulatory drugs, such as lenalidomide (and other thalidomide analogues) have been shown to have clinical efficacy in the treatment of myeloma. These immunomodulatory drugs have been shown to bind to and inhibit the cereblon ubiquitin ligase. It has been demonstrated that lenalidomide-bound cereblon, acquires the ability to target two specific lymphoid transcription factors, IKZF1 and IZKF3, both known to play a central role in B and T cell biology, for selective ubiquitination and degradation[8,9].

Proteasome inhibitors are also extremely effective in myeloma treatment. The proteasome is an intracellular enzyme complex that breaks down damaged proteins. The exquisite sensitivity of myeloma to proteasome inhibitors remains largely unexplained, although the drugs are thought to act in a multi-faceted and extensive mechanistic fashion. They have been thought to stimulate apoptotic pathways, inhibit the NF-κB pathway, down regulate expression of genes associated with DNA repair and induce an endoplasmic reticulum stress response[10-12].

Anti-CD38 monoclonal antibodies, such as daratumumab (Darzalex) and isatuximab (Sarclisa), are now in routine use for the treatment of myeloma. Daratumumab has been NICE approved for use in second line therapy alongside Velcade and dexamethasone, and for fourth line therapy as a monotherapy. The Cassiopeia trial which is currently on going is likely to suggest that Daratumumab in combination with Velcade, thalidomide and dexamethasone (VTD) is more effective that VTD alone[13]. Isatuximab, an alternative anti-CD38 monoclonal antibody, has recently been approved by NICE in combination with pomalidomide and dexamethasone as a fourth line treatment for patients with myeloma not yet treated with an anti CD38 monoclonal antibody[14].

Shorter survival times correlate with higher clinical stage at diagnosis, renal insufficiency, degree of marrow replacement, increased proliferative activity and certain karyotypic abnormalities[15]. Although significant advances have been made in MM treatment over the past two decades, and it is now regarded as highly treatable, myeloma remains almost always incurable. The only potentially curative option

available for the treatment of myeloma is allogeneic stem cell transplantation. However, although this intensive procedure can achieve long term remission, it is associated with high treatment related mortality and risk of relapse, and tends to be considered only in a small minority of patients[16].

Of patients diagnosed with myeloma (between 2013-2017), 82.7% of patients survived for one year or more, 52.3% of patients survived their disease for five years or more, and it is predicted that 29.1% will survive ten years or more[5]. These figures demonstrate a dramatic improvement in survival rates; in the 1970's, only 5% of patients were expected to survive their disease beyond ten years, and now approaching one third of patients will[5]. Importantly, there remain a subset of patients for whom current treatment modalities are not effective, with approximately 20% of clinical responses considered sub-optimal and 5% non-responsive.

Plasma Cell Biology

The human immune system has evolved in order to confer resistance to infection. Plasma cells are part of the 'adaptive' component and are required to produce antibodies in response to antigenic insult[17-19]. The immunoglobulin (Ig) antibody molecule is composed of two heavy chain and two light chain proteins. These are encoded by the *IGH* gene for the heavy chain, located on chromosome 14, and the *IGK* and *IGL* genes for the light chain, kappa and lambda, located on chromosomes 2 and 22 respectively[18]. Variable gene segments at these loci undergo irreversible rearrangement, at the DNA level, and this creates individual B cells with specificity for a single antigen. In summary, the several stages of B cell development each represent a change to the genomic DNA involving the variable (V), diversity (D) and junctional (J) gene segments of the immunoglobulin genes. These stages can be divided into three processes, all of which generate double stranded DNA breaks: VDJ recombination, somatic hypermutation and IgH-switch recombination[20,21].

Following maturation, plasma cells which have undergone the final stages of development home to the bone marrow[20]. These cells are long-lived, terminally

differentiated and non-dividing. They are highly dependent on the BM microenvironment where their longevity is favoured by survival factors found in permissive niches[22,23]. Plasma cells interact with the BM microenvironment via a number of complex interactions, which are crucial to tumour survival and disease progression[22].

Whilst key to the creation of a diverse adaptive immune repertoire these complex developmental processes imply inherent genome instability, and it is this required instability that may facilitate oncogenic transformation, resulting in neoplastic disease such as myeloma.

Genetic Complexity of Plasma Cell Neoplasms

Myeloma is a genetically complex disorder characterised by multiple genetic changes, affecting different pathways, that have the ability to deregulate plasma cell biology leading to a broadly similar phenotypic manifestation of disease: genetic heterogeneity[22,24]. Myeloma genetic abnormalities include hyperdiploidy, the non-random gain of a specific subset of chromosomes, rearrangements of the *IGH* gene resulting in aberrant expression of a number of different oncogenes, loss and gain of specific chromosomal regions, gene mutation thought to target specific pathways, specific gene variants thought to predispose a subset of patients to myeloma and epigenetic modification of the DNA. As well as this complex range of genetic change, there is evidence of clonal heterogeneity. This genetic complexity and heterogeneity, in part, is likely to have hindered the development of effective treatments. A deeper and clearer understanding of genetic abnormalities associated with myeloma, and their role in specific pathways, may offer new routes for drug development and the possibility of a more personalised medicine approach.

Hyperdiploidy vs Hypodiploidy

From a genetic perspective, myeloma can be divided into those with and without a hyperdiploid karyotype[25,26]. Hyperdiploidy is seen in approximately 30-40% of patients. Chromosome number ranges from 48-75, median 53, and the chromosome

gains are non-random, often involving the odd numbered chromosomes[4,25]. Chromosome 9 is the most common gain (seen in 48% of hyperdiploid cases), followed by chromosome 15 (47%), 19 (46%), 5 (38%), 3 (36%), 11 (33%), 7 (26%), 21 (23%), 18 (11%) and 17 (5%)[27,28]. These changes are not seen as serial gains, but rather as the result of a single catastrophic mitotic event[22]. It is thought that compensatory mechanisms exist in cancer cells allowing the stress of aneuploidy to be tolerated[29] and, indeed, are considered relatively stable. Hyperdiploidy in myeloma has been associated with a better overall survival when compared to the non-hyperdiploid group[30]. Within the hyperdiploid group, trisomy 3 and/or 5 have been associated with an improved survival, and indeed are thought to overcome the poor prognosis associated with t(4;14), whereas trisomy 21 has a significantly worse prognosis[28]. This associated good prognosis is likely to be related to the chromosome excess and the aneuploidy stress in the cells, which has been demonstrated to be uniquely chemosensitive. This phenomenon is also seen in ALL patients with hyperdiploidy (of a different set of chromosomes), where the prognosis and response to therapy is also good[29].

The non-hyperdiploid group include karyotypes with hypodiploid, pseudodiploid or near tetraploid chromosome number[30]. The near tetraploid groups appear to be a doubling (i.e. 4n) of the hypodiploid and pseudodiploid cell lines. The non-hyperdiploid group are typically associated with *IGH* translocations, although *IGH* rearrangements are also present in approximately 10% of the hyperdiploid group[25].

Translocations

Non-hyperdiploid patients are frequently (55-70%) associated with rearrangements of *IGH* on chromosome 14[22,31,32]. *IGH* rearrangements are considered promiscuous due to their many partner genes. They are usually simple reciprocal translocations juxtaposing the *IGH* enhancers to an oncogene[25]. This gives rise to abnormal expression of the oncogene and contributes to the myelomagenic effect. Detailed in table 1.1, the five main translocation partners are: t(4;14), t(11;14), t(6;14), t(14;16) and t(14;20). Together these are seen in approximately 40% of patients. These rearrangements are thought to be associated with upregulation of one of the cyclin

D genes. t(11;14) and t(6;14) directly deregulate *CCND1* and *CCND3* respectively, t(14;16) upregulates *CCND2* as *MAF* directly binds to the *CCND2* promoter and t(4;14) also upregulates *CCND2* via *FGFR3* and *MMSET*, although the exact mechanism is unknown[22]. Such translocations are mediated by errors in DNA modification associated with B cell maturation, following plasma cell-antigen interactions. *IGH* switching is considered the most common, whilst somatic hypermutation and VDJ recombination are likely to be less frequent events[17,18].

Translocation	Gene(s)	Frequency	Prognosis	Gene Role / Function
t(4;14)(p16.3;q32)	*FGFR3* *MMSET*	15%	Poor	Bone development & maintenance Oncogene, over expression results in proliferation and anti-apoptotic effects
t(11;14)(q13;q32)	*CCND1*	15%	Good	Cell cycle G1/S transition
t(14;16)(q32;q23)	*MAF*	5%	Poor	Oncogene, enhances tumour stroma interactions
t(6;14)(p21;q32)	*CCND3*	3%	Good	Cell cycle G1/S transition
t(14;20)(q32;q11)	*MAFB*	2%	Poor	Regulation of lineage-specific haematopoiesis

Table 1.1: The five main translocations associated with *IGH* in myeloma

The t(4;14)(p16.3;q32) translocation is cytogenetically cryptic[25], and therefore FISH (or another non-karyotypic technique) is required in order to determine the presence of this rearrangement. The t(4;14) rearrangement involves the *IGH* gene and two protein coding genes located at 4p16.3, multiple myeloma SET domain, *MMSET* (also known as the Wolf Hirschhorn candidate 1, *WHSC1* or *NSD2*) and the fibroblast growth factor receptor 3 (*FGFR3*), an oncogenic receptor tyrosine kinase[25,33,34]. In the balanced translocation, both *FGFR3* and *MMSET* are juxtaposed to *IGH* enhancers. *FGFR3* is over expressed from the derivative chromosome 14 and *MMSET* is overexpressed from the derivative chromosome 4. 25% of cases are seen in an unbalanced form, with loss of the derivative chromosome 14, which is associated with loss of the aberrant *FGFR3* expression[25,26]. This translocation is universally associated with a poor prognosis[34-37] [38]. The upregulation of *FGFR3*

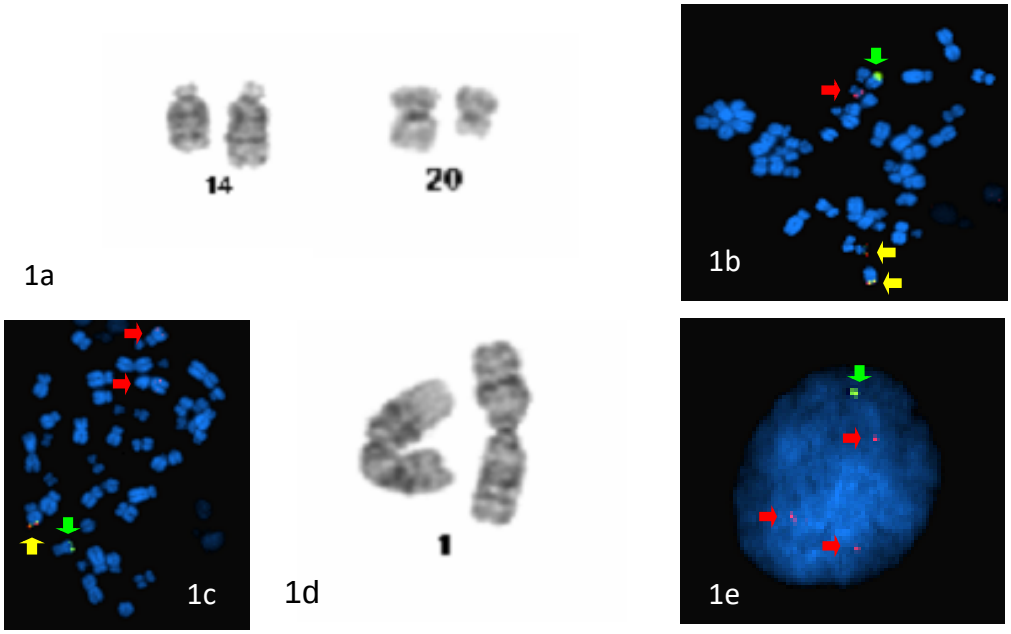
associated with t(4;14) induces CCL3 expression, which has receptors expressed on osteoblasts, osteoclasts and plasma cells, and may be the pathway through which the adverse effect is mediated, resulting in differential effects on osteolytic disease and cell migration[39]. It is seen in MGUS (approximately 3% of cases), but more highly associated with SM and MM[40-42]. Data suggests that these patients benefit from bortezomib therapy and in addition TKI-258 (Dovitinib), an FGFR3 inhibitor, is also undergoing clinical evaluation[26,43].

t(11;14)(q13;q32) involves translocation of the *CCND1* gene located at 11q13, where it comes under the regulatory influence of *IGH* at 14q32[44]. This translocation is balanced in the majority of cases and associated with a favourable prognosis[38]. Additional copies of the derivative chromosome 14 have been seen and are thought to be associated with progressive disease[44].

The t(6;14)(q21;q32) involves the gene *CCND3* at 6q21, resulting in its upregulation. It is often seen on a backdrop of a complex karyotype, and the derivative chromosome 14 can be present in multiple copies. This is a rare event and has only been associated with approximately 3% of MM patients [22].

t(14;16)(q32;q23) and t(14;20)(q32;q11) both show juxtaposition of the *IGH* gene to a *MAF* family gene, *MAF* and *MAFB* respectively [45,46]. They have been described in 5-7% of MM. *MAF* rearrangements are considered to be mediated by a fragile site in the *WWOX* gene on the long arm of chromosome 16[20,26]. This results in *MAF* coming under the regulatory influence of *IGH*, which in turn has the effect of upregulating *CCND2* [22]. *MAF* rearrangements have been associated with a more aggressive clinical course and a poor prognosis[26,38]. Data on *MAFB* rearrangements associated with t(14;20) (see figure 1a & 1b) are not robust due to their rarity; although a similar clinical course would be predicted [26].

Translocations not involving *IGH* do occur, but are considered unusual and usually associated with progressive disease. *CCND3-MAF, MAF-FGFR3/MMSET, CCND3-FGFR3/MMSET* rearrangements have all been described [22].



1a & b) The partial karyotype shows the normal chromosome 14 and the normal chromosome 20 on the left of each pair. Additional material can be seen on the long arm of the derivative chromosome 14, from the derivative chromosome 20 depicting the t(14;20)(q32;q11). The FISH image using the Cytocell probe set for *IGH* and *MAFB*, confirms that the rearrangement involves the juxtaposition of *IGH* and *MAFB*.
1c) FISH image depicting the Vysis *MYC* breakapart probe. There is an intact copy of the probe on one chromosome 8. The second fusion has broken apart, but also demonstrates associated complexity with duplication of the centromeric part of the probe and involvement of three chromosomes.
1d & e) The partial karyotype shows a normal chromosome 1 on the left with an isochromosome 1q on the right. The isochromosome shows two copies of the long arm mirror imaged around the centromere with effective loss of the short arm of chromosome 1. The FISH image shows the Cytocell *CDKN2C/CKS1B* probe which highlights a gene of interest on the long arm in red and a gene of interest on the short arm of chromosome 1. A three red and one green signal pattern confirms the gain of 1q and the loss of 1p.
Coloured arrows demonstrate the probe positions, green and red arrows show green and red signals respectively, yellow arrows depict fusion signals.

Figure 1.1: Partial karyotypes and FISH images of genetic abnormalities associated with MM.

MYC translocations are seen in approximately 15-20% of patients with myeloma at presentation[25,47,48] and up to 45% of patients with advanced disease[25,26]. These translocations are not thought to be initiating events, but late events associated with increased proliferation and stromal independent plasma cells, and in turn associated with a poor prognosis[25,38]. These are frequently seen as non-reciprocal translocation events involving more than one chromosome and associated regions of amplification and duplication[25] (figure 1c). The t(8;14)(q24;q32) accounts for only 25% of *MYC*

rearrangement [20,48], and recent studies have demonstrated that *MYC* is able to recruit active super-enhancers from highly expressed genes associated with B cell, plasma cell or myeloma development[49]. Examples of this include enhancers associated with *CCND1, XBP1, KRAS, FAM46C* and *CHST15* [49].

Deletions and Duplications

Myeloma genetics includes copy number variation as a common and recurrent finding[22,26]. Theoretically, and simplistically, deletions are likely to involve loss of tumour suppressor genes at those sites, and duplications are likely to be associated with overexpression of oncogenes within the region. Regions of recurrent deletion and/or duplication include 1p, 1q, 9q, 11q, 12p, 13q, 15q, 16q, 17p, 19q and 22q[20,22,26,50], and a number of these are described further.

Deletion of chromosome 13 was the first chromosome abnormality to be associated with a poor prognosis in myeloma[51]. Chromosome 13 abnormalities are seen in 50% of myeloma patients[26], of these 85% are associated with monosomy for chromosome 13, and the remaining 15% associated with an interstitial deletion, involving the 13q14 region known to be associated with the *RB1* gene[48,52]. More recently the poor prognosis associated with deletion/monosomy of chromosome 13 has been contested in the literature. Its close association with the presence of t(4;14), suggest that the statistically poor prognosis may be a result of association alone[26,53].

Deletion of the short arm of chromosome 17, the site of the *TP53* tumour suppressor gene remains the most important prognostic factor in myeloma genetics[22,26] conferring an extremely poor prognosis[54]. Deletion of 17p is seen in 7-10%[22,25,37] and has been associated with shorter survival, aggressive disease and increased hypercalcaemia and extramedullary disease[26].

Chromosome 1 rearrangements are the most common aberrations in myeloma[25]. These usually result in deletions of the short arm of chromosome 1 and duplications of the long arm of chromosome 1[25,26]. Short arm deletions have been shown to span a region from 1p13~1p31 and have been associated with a poor prognosis[25,55,56].

Genes that have been associated with this deletion include *CDKN2C, FAF1* and *FAM46C*[22,56]. Rearrangements of the long arm of chromosome 1 are often complex[25,57] and have been associated with pericentromeric instability. They are seen in 40% of myeloma patients at diagnosis and 70% of cases at relapse[25]. 'Jumping' translocations can be associated with chromosome 1q duplication[25]; this term describes rare chromosomal events in which the same donor chromosome segment is translocated onto two or more recipient chromosome sites, such that it can be present in different guises in different cells. Implicated genes include *CKS1B* and *ANP32E*[25]. Duplication of 1q is also associated with a poor prognosis, although the intrinsic relationship with the presence of deletion 1p, creates difficulties in assessing these abnormalities separately[25]. A common manifestation of del1p/dup1q is the isochromosome 1q (Figure 1.1d), in which the short arm is lost and the long arm is duplicated and mirrored around the centromere.

Uniparental disomy (UPD) represents a further mechanism by which alleles can be lost. UPD is term used to describe loss of a chromosome, or chromosome part, which is then duplicated from the second allele. This results in loss of heterozygosity (LOH) without the associated loss of copy number. UPD can be seen associated with cancer genotypes and is termed 'acquired UPD' (aUPD) in this scenario. Walker *et al* (2006) have demonstrated that aUPD / LOH is prevalent in MM with a median number of three aUPD regions per sample. The size ranges from 677kb to whole chromosomes, but tends to implicate relatively small regions at 1p, 1q, 6q, 8p, 13q and 16q[58]. These areas of aUPD have the potential to highlight areas containing genes important in myeloma pathology.

Mutations and Implicated Pathways

Whole exome sequencing (WES) studies have suggested that an average of 35 gene mutations are present per MM patient sample. This figure sits part way between the suggested mutation level of eight in the more genetically simple haematological malignancies and approximately 540 mutations detected in the genetically complex epithelial tumours[59]. Determination of these multiple mutations, and general genetic heterogeneity seen in myeloma, support the hypothesis that a pathological

requirement of myeloma development is pathway deregulation, rather than specific gene rearrangement or modification. It is likely that multiple genetic abnormalities seen converge to result in a more simplified effect, of targeting a smaller number of specific functional pathways.

Although the presence of copy number abnormalities and recurrent translocations in myeloma have been known for some time, the presence and pattern of gene mutations associated with myeloma has only developed with the more novel technologies such as massively parallel sequencing. The first genome sequencing of myeloma was described by Chapman *et al* in 2011[59], and since then a number of publications have described a group of significantly mutated genes[38,60-62].

Fifteen significantly mutated genes were described by Walker *et al*, with similar sets described by both Lohr *et al* and Bolli *et al*[38,61,62]. The dominant mutations include those from the RAS/MAPK pathway including *NRAS* (G12, G13, Q61 & Y64 variant hotspots), *KRAS* (G12, G13, Q61 & Y64 variant hotspots), *BRAF* (most commonly V600), *NF1* and *RASA2* genes making up 43% of patients[38,60,62,63] The NF-κB pathway is next most commonly affected which includes the genes *TRAF3*, *CYLD* and *LTB*, affecting 17% of patients[38]. These mutations are considered prognostically neutral but could still be therapeutic targets. Mutations in *CCND1* and DNA repair pathway genes such as *TP53, ATM, ATR* and *ZNFHX4* are considered poor prognostic markers, whereas *IRF4* and *EGR1* are considered favourable[38].

Keats *et al*, and Annunziata *et al*, have also demonstrated constitutive activation of the NF-κB pathway in 50% of MM cases using gene expression profiling[64,65]. This is not caused by a single mutation, but rather a collection of gene mutations and deletions. Genes implicated include *BIRC2* and *BIRC3* at 11q, *CYLD* at 16q and *TRAF3* at 14q32. Increased NF-κB nuclear activity is thought to have an anti-apoptotic effect[22]. In addition, data suggests that low-level *TRAF3* expression is associated with a better response to bortezomib[26].

Other affected pathways include the WNT signalling pathway, RANK/RANKL/OPG, PI3K, Notch and JAK pathways, some of which are discussed further as part of the section on bone involvement.

Involvement of *IKZF1* and *IKZF3* has become important. Recent work has shown that cereblon ubiquitin ligase is inhibited by the immunomodulatory drugs and lenalidamide-bound cereblon targets IKZF1 and IZKF3 for degradation disrupting their central roles in the lymphoid transcription pathway[8,9].

A number of mutational groups have been associated with previously described genetic sub-groups, and these have been described as oncogenic dependencies. Mutations in *FGFR3, DIS3* and *PRKD2* have been associated with t(4;14), mutations in *CCND1* and *IRF4* with t(11;14) and mutations in *MAF, BRAF, DIS3* and *ATM* with t(14;16), and then finally *FAM46C* mutations with hyperdiploidy[66-68]. Of interest, increased age has also been associated with hyperdiploidy in myeloma[68].

It is highly likely that these clinically relevant mutations will become incorporated into the diagnostic testing and risk stratification of myeloma patients.

Epigenetic Changes

Epigenetic factors are also involved in the aetiology of MM, with changes to both DNA glycosylation/acetylation and histone modification playing a part in modulating gene expression. Global DNA hypomethylation and specific gene hypermethylation have been reported in association with the transformation of MGUS to MM[22]. 15% of t(4;14) patients show a gene-specific hypermethylation pattern. Overexpression of *MMSET* leads to histone modification which in turn promotes cell survival and cell cycle progression[22]. Dysregulation of miRNA has been associated with a particular gene cluster on chromosome 13 implicated in the MGUS to MM transition[69].

Genetic Predisposition

Relatives of myeloma patients have a 2-4 fold increase in their risk of developing myeloma providing good evidence for an inherited component of the genetic aetiology of myeloma.

An inherited genetic variation at 2p23.3, 3p22.1 and 7p15.3 has been associated with a genetic predisposition to MGUS[70,71], with genome wide association studies

(GWAS) demonstrating further regions of common variation at 3q26.2, 6p21.33, 17p11.2 and 22q13.1[72]. Chubb *et al* (2013) suggested that these seven loci were likely to account for only 13% of the familial risk of myeloma suggesting that many further regions would be involved[72]. Since 2013, a further 16 regions have been proposed[73]. A more recent GWAS published in 2019 by Went *et al*, assessed data on 7319 myeloma cases against 244,385 controls across different tissue types. 108 genes were identified at 13 regions, all of which were within 1Mb of known myeloma risk variants. Of the 108 genes, 94 located at eight regions had not previously been considered candidate genes[74]. Of particular interest were a number of genes from the *APOBEC* family and *RNF40* with known roles in myelomagenesis, and *QPRT* which has been implicated in other malignancies[74].

Intra-tumour Clonal Heterogeneity

Whole exome sequencing, cytogenetics and copy number analysis have been used to show that myeloma cases can have a complex subclonal structure with a high level of clonal genetic heterogeneity suggesting that myeloma populations are not homogeneous, but made up of different populations with different clonally related and unrelated changes. Intraclonal heterogeneity is present at all disease stages, and clonal progression is a key feature of transformation[75]. Bolli *et al* have demonstrated, using serial sampling, that clonal evolution is diverse and includes both linear and branching evolution[62]. Further to that, Keats *et al* describe the existence of three temporal myeloma types; genetically stable, linearly evolving or heterogeneous clonal mixtures with clones shifting in dominance over time[76]. They also suggested that few changes were seen in standard risk patients, with significantly more changes being associated with high risk patients[76]. The t(4;14) is usually considered a primary change, but this has also been shown to be present in low level silent sub-clones both at diagnosis and relapse, suggestive of the movement of the dominant nature of specific clones throughout the disease course, a phenomenon described as clonal tides[77].

Although myeloma is a disease of the bone marrow where a rapid movement and dissemination of more dominant clones may be expected, the distribution of disease

is not homogeneous either. Multi-region sequencing has also shown that spatial heterogeneity is present in approximately 75% of patients, with dissemination of new clones happening more freely in the early stages, but regional evolution commonly seen in advanced disease[78].

There are important implications to therapy as these clones can demonstrate a differential response to therapy; reducing the level of one clone, may result in expansion of a more prognostically detrimental refractory clone. The process of branched evolution, intra-tumour and spatial heterogeneity provide further ambiguity, since a single genetic picture obtained at a single timepoint and single site is likely to under-represent the complexity of the underlying disease[79]. It remains likely that combination therapies targeting a number of coexisting disease subclones will be particularly important in myeloma as a whole, and specifically high risk myeloma[75,76,78].

It is recognised that the prognostic associations of specific genomic changes may be treatment dependant, and where these have been specified in the text, they have been referenced appropriately demonstrating the context of the association.

Recent review articles from Anwar *et al* (2019) and Pawlyn *et al* (2019) describe the current potential of personalised medicine in myeloma based on determining genomic aberrations contributing to disease progression and drug resistance. The t(11:14) has been shown to respond specifically to treatment with venetoclax and an *FGFR3* inhibitor (AZD4547) has been shown to be effective in myeloma associated with t(4;14). *BRAF* mutations can be targeted by vemurafenib (alone or in combination), JAK2 pathway activation can be targeted by ruxolitinib, and the NF-kB pathway activation has been shown to respond to danfin combined with bortezomib. Other potential targets include *IDH1/2* mutations and IDH inhibitors, *ATM/ATR* mutations and PARP inhibitor treatment, and *MDM2* overexpression, *CCND1* and *MYC* rearrangement[80,81].

Pawlyn *et al* also describe targeting a high-risk markers patient cohort using intensified therapeutic approaches agnostic to molecular lesions; combinations of quadruplet or even quintuplet regimens and/or novel immunotherapy approaches such as bispecific antibodies, antibody–drug conjugates, and CAR T cells[80,81].

Microenvironment and Bone Involvement in Myeloma

There are complex interactions between the plasma cells and their supporting microenvironment, these interactions are crucial to the tumour survival and progression of disease [22]. The marrow microenvironment is made up of extracellullar matrix and five types of bone marrow stromal cells; fibroblastic stromal cells, osteoblasts, osteoclasts, vascular endothelial cells and lymphocytes[25,82]. The cross-talk between myeloma cells and the microenvironment can either be directly or indirectly by secretion of adhesion molecules, cytokines and growth factors. The interaction between myeloma cells and the bone microenvironment results in the uncoupling of the normal bone remodelling process leading to osteoclast activation and osteoblast suppression. It is this decoupling process that is thought to be pivotal in the pathogenesis of the osteolytic lesions and bone damage associated with myeloma[20,83].

A number of intra- and intercellular signalling pathways have been implicated in this complex process, and over recent years osteocytes have emerged as key regulators associated with the bone lesions seen in myeloma[39]. These signalling pathways include RANK/RANKL/OPG, WNT, TNF and Notch, and a number of specific signalling molecules have also been implicated such as DKK1, sclerostin, periostin, osteopontin, TGFβ[84,85] and activin A[86] amongst others. It is thought that a better understanding of these biological interactions will offer promising future therapeutic strategies[39,83].

The RANK/RANKL/OPG signalling pathway is a crucial regulatory system of bone remodelling. Deregulation of this pathway has been well described in myeloma, and the resulting increase in RANKL and decreased OPG expression, ultimately increases bone resorption[4,39,87]. A promising agent targeting this pathway is denosumab, which is able to prevent RANK activation, and subsequent osteoclast activation[39].

The WNT signalling pathway drives osteoblast differentiation. WNT pathway inhibitors associated with myeloma, such as sclerostin and Dickkopf-1, therefore result in the inhibition of osteoblast differentiation, and hence reduced numbers of mature osteoblasts leading to reduced bone formation. Antibodies to these WNT

inhibitors such as Romosozumab, an anti-sclerostin antibody[39], and anti-DKK1 antibodies[39,87] are demonstrating some encouraging results.

Similarly, the Notch pathway is actively implicated in MM-induced osteoclastogenesis, and offers a further treatment targetable area for future development[39].

Bisphosphonates are an important class of drugs used to treat bone manifestations in MM. They are potent inducers of osteoclast apoptosis, thereby reducing elevated bone resorption (anti-resorptive therapy) associated with MM[88-90]. Long-term or high-dose use of bisphosphonates is associated with some adverse effects. Whilst anti-resorptive therapies are widely used in the treatment of myeloma, anabolic therapies, those that promote bone formation by enhancing osteoblast activity are less widely used. In osteoporosis, the most common bone disease worldwide, a number of anabolic therapies are available; parathyroid hormone (PTH) and its analogues, teriparatide and abaloparatide, and romosozumab (a sclerostin neutralising antibody). A number of more experimental agents[91,92] are also being considered as suitable anabolic therapies[88,90]. It is likely that these anabolic therapies used in osteoporosis will impact on myeloma treatment also.

In early MGUS and myeloma, the plasma cells are highly stromal dependent. As the disease progresses and becomes more aggressive, the plasma cells are able to develop stromal independence. Many of these cellular interactions are also governed by genetic changes affecting the normal regulation of these interactions and relationships[22,89].

These relationships between myeloma plasma cells, the microenvironments and osteoclast/osteoblast related pathways where disruption can result in the imbalance between osteoblast and osteoclast function become important in this study in informing the 'osteome' work described in chapter 4.

Genetic Screening Methodologies

Genetic analysis of BM samples in myeloma patients has proven to be useful, offering both diagnostic and prognostic information to the clinician and patient[4].

Since myeloma is a genetically heterogeneous disease, it is likely that genetic abnormalities seen in myeloma patients will increasingly be used to inform treatment decisions. Patients exhibiting a t(4;14) are more likely to be offered more aggressive therapy, including transplantation, and *BRAF* inhibitors are being considered in those patients shown to have *BRAF* mutations.

A range of genetic methodologies, conventional and novel, have been used in the genetic analysis of myeloma patients. Each technology offers pros and cons, and the different data capture from each methodology is complex and varied. Currently there is not a single technique able to assess the full range of genetic change seen within the myeloma genome. This introduction covers the five genetic technologies chosen for evaluation as part of this study, two well established techniques in myeloma analysis (cytogenetics and FISH) and three more novel and less used techniques in the analysis of myeloma (MLPA, DNA microarray and NGS).

Cytogenetic Analysis

Cytogenetic analysis was the first methodology to be widely adopted in the analysis of myeloma, and relies of the ability to capture cells, specifically plasma cells, at metaphase of the cell division process. Robust techniques are employed in many laboratories to manipulate the cell cycle, often utilising colchicine, or one of its analogues, to disrupt formation of the spindle apparatus during cell division. At metaphase the chromosomes have condensed and are preparing to arrange themselves on the equator of the spindle apparatus, before anaphase proceeds to split the two chromatids into separate daughter cells. Disruption of the spindle apparatus arrests the cell cycle at metaphase. Harvesting samples at this stage and treating chromosomes with trypsin and stain, to produce the characteristic banding pattern, allows chromosomes to be karyotyped and analysed.

Plasma cells are mature cells and have been shown to cycle slowly, often adaptations to the culturing process, allowing for longer term culturing (4-6days), are introduced for the assessment of myeloma to increase the chances of capturing these cells in division[24,42]. Cytogenetic analysis offers a full genome screen, but at low resolution.

The rate of plasma cell capture is poor, even with enhanced culturing, and therefore the abnormality rate is low.

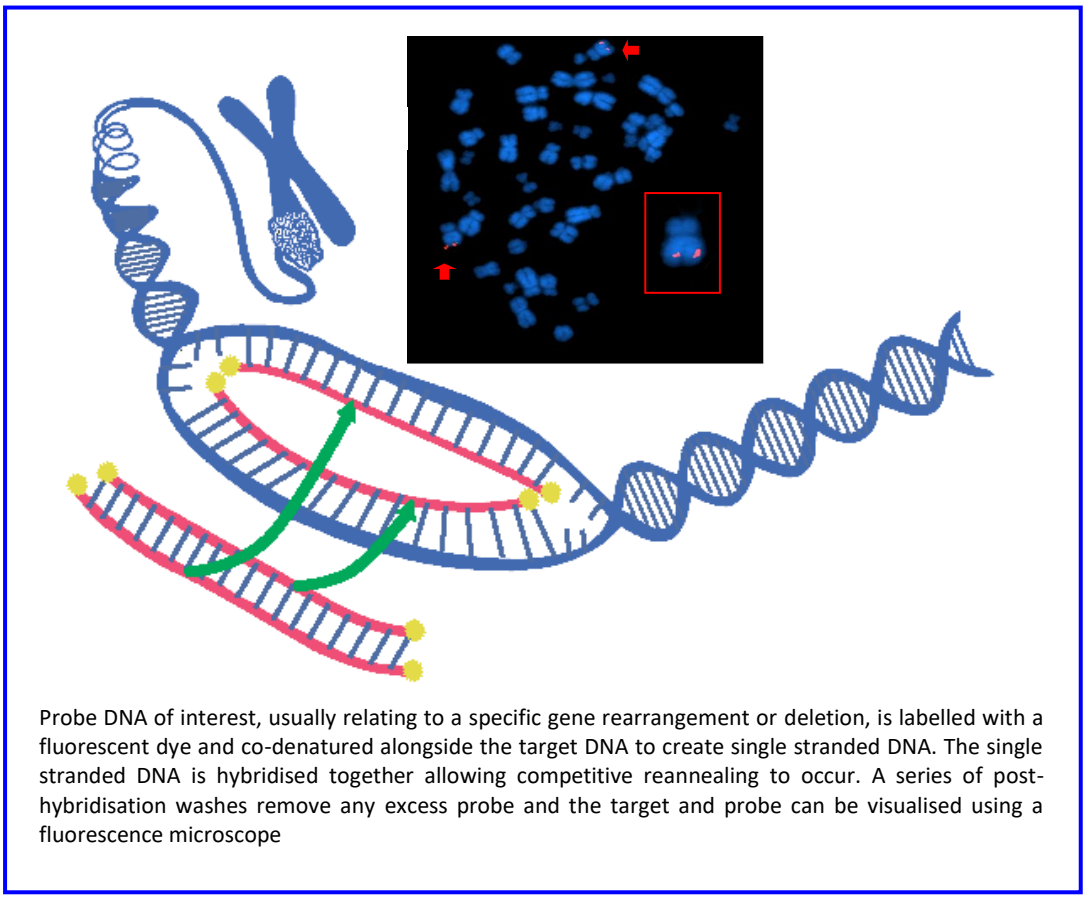
Fluorescence in situ Hybridisation (FISH)

FISH, currently the most widely adopted technique in the genetic analysis of myeloma, is the process of hybridising a fluorescently labelled target DNA (probe) from a known region of interest, to metaphase or interphase cells to allow enumeration or rearrangement detection, see figure 1.2. Probes can be created in-house, but more usually, in diagnostic laboratories, are bought in from commercial companies.

Probes come in a variety of sizes, designs and formats and it is important to have a good understanding of these before analysis is undertaken. The ability to utilise interphase cells, eliminates the problems associated with metaphase capture for cytogenetic analysis. The abnormality rate is much improved, but is dependent upon the probe set used. Each probe 'asks a question', with a yes or no answer. The more probes used, the more questions asked and answered. Small probe panels are inexpensive and efficient, but extended FISH panels, whilst providing more information, quickly become expensive and extremely time-consuming.

Probe design depends on the question being asked, when looking for loss and gain of a probe, it is extremely useful to have a control probe region in order to determine whether loss is true deletion or lack of hybridisation. Rearrangement detection probes are usually designed as breakapart probes or dual colour dual fusion probes. Breakapart probes usually have a red and a green probe flanking either side of the breakpoints of the gene in question. They appear as a fusion (i.e. yellow) signal in a normal scenario and a single red and green when a rearrangement is present. Dual colour, dual fusion probes usually have a single colour probe (red) spanning one gene and a single colour probe (green) spanning the second gene. A normal pattern would be 2R2G, and a rearrangement results in the probes across the genes splitting and rejoining with the opposite coloured probe resulting in a 2F1R1G signal pattern. More complex patterns can be seen when rearrangements are unbalanced or when

additional genes not covered by the probed regions are involved. The specific patterns associated with the probe sets utilised here are referred to further in the results section.



Probe DNA of interest, usually relating to a specific gene rearrangement or deletion, is labelled with a fluorescent dye and co-denatured alongside the target DNA to create single stranded DNA. The single stranded DNA is hybridised together allowing competitive reannealing to occur. A series of post-hybridisation washes remove any excess probe and the target and probe can be visualised using a fluorescence microscope

Figure 1.2: Schematic diagram to demonstrate the process of FISH.

Multiplex Ligation-dependant Probe Amplification (MLPA)

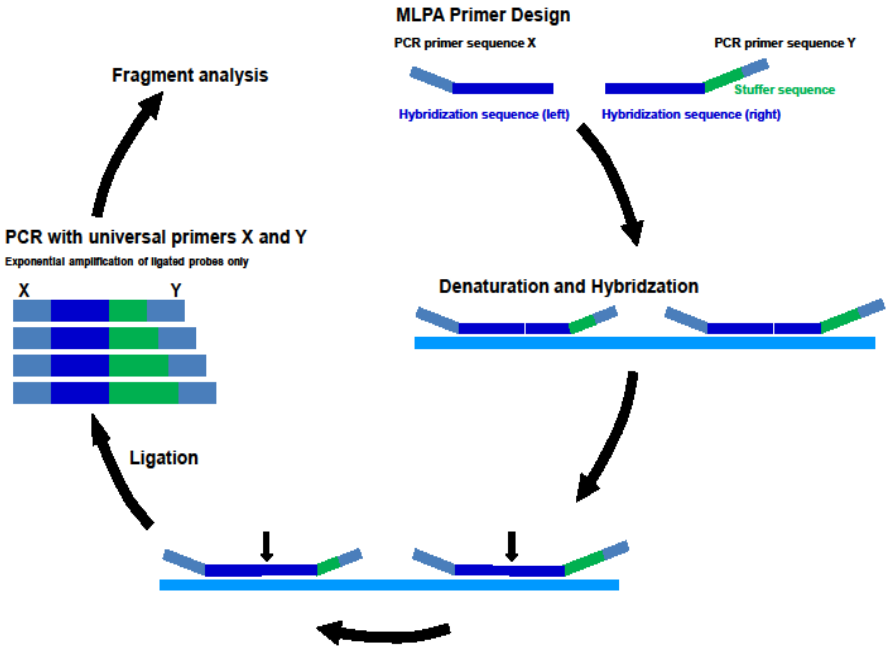
MLPA is a multiplex polymerase chain reaction (PCR) methodology, which allows simultaneous detection of deletions or duplications of up to 50 targeted gene regions. This is an attractive technique in myeloma as a large number of copy number changes including hyperdiploidy have prognostic implications in myeloma. As shown in the schematic in figure 1.3, MLPA probes are designed in pairs with the same two primer sequences present in each of the pairs. These primers are attached

to the region of interest. One of the MLPA probe pair incorporates a stuffer sequence, creating a known total probe length.

The DNA is denatured and hybridised with the MLPA probes. When the DNA target is present, with no alterations, the probe pairs lie immediately adjacent to each other and can ligate. When there is a deletion, or in some cases a mutation, the probes cannot hybridise to the DNA correctly and probe ligation does not occur. PCR is used to directly amplify the ligated probe pairs and not the target sequence.

As only ligated probes will be exponentially amplified, the number of products is a direct measure of the number of target sequences, when compared to a control. Separation of the products is carried out using electrophoresis and the known probe length allows identification of the specific target region.

MLPA offers a cost effective, multiplex technique, but the information output only relates to the loss or gain of the specific regions targeted. When mutation is present, again preventing the MLPA probe ligation, it is not possible to distinguish this from a deletion event.



Paired probes around an area of interest are labelled with primer sequence. A stuffer sequence of known length is also included to allow for probe identification. The probes are annealed to the DNA of interest. When no DNA alteration is present the probes lie adjacent to each other allowing probe ligation to occur. Ligation is not possible where deletions or mutation are present.Ligated probes are amplified and represent a measure of the number of copies. The products are separated by electrophoresis based on their size.

Figure 1.3: Schematic diagram to demonstrate the process of MLPA

DNA Microarray / Comparative Genome Hybridisation (CGH)

DNA microarrays employ the same technology as comparative genome hybridisation (CGH), and offer a genome wide copy number analysis (and with the use of single nucleotide polymorphisms (SNPs) allow assessment of Loss of heterozygosity (LOH)) at an excellent resolution level. Both DNA of interest and control DNA are prepared and fluorescently labelled; control DNA in red and DNA of interest in green. The two DNA samples competitively hybridise to an array chip containing thousands of complementary DNA oligonucleotides. In the case of equal copy numbers in the two DNA samples the competitive hybridisation will be equal. Increased copy number in the DNA of interest will show an excess of green fluorescence and conversely when deletions are present in the DNA of interest the competition will be less effective

and the region will demonstrate an excess of red fluorescence. The level of fluorescence is measured across the chip for each oligonucleotide and collated to demonstrate regions of loss and gain across the genome.

DNA microarrays are designed to provide a genome screen with a backbone level of resolution across the genome. Areas of interest (for example oncology related probes) can be designed to have enhanced resolution. Many types of array exist, but we propose the use of the Affymetrix CytoScan HD Array. These arrays utilise slightly different chemistries, in which the DNA of interest is compared to an internal normal reference set. This array allows interrogation of the entire genome using more than 2.6 million markers for copy-number analysis and approximately 750,000 SNPs. The higher probe density areas include 533 cancer genes, 100% of which are covered with 25 markers per 100kb region. The array would allow DNA copy number to be determined at a high sensitivity and has the ability to detect LOH and regions of aUPD.

SNP array and CGH are both high resolution, whole genome techniques, but can only assess loss and gain of genetic material; balanced rearrangements resulting in no loss or gain of genetic material and mutations will not be detected.

Next Generation Sequencing (NGS)

NGS has evolved dramatically over the last few years. It offers the power of simultaneous sequencing of multiple DNA templates, without the requirement of the same number of target-specific DNA primers[93]. NGS has the ability to detect small nucleotide variants, and small insertion and deletion (indels) events across the whole genome or sets of targeted genes as required. The commercial platforms have evolved to provide confidence in the robustness of this approach to provide an accurate and comprehensive means of mutational analysis. There are a range of NGS platforms currently available, each of which employs different sequencing chemistries, but with similar resultant output[94].

Current methodologies start with the production of short DNA fragments with the addition of non-specific adaptor DNA sequences, this can be either by chemical

binding or ligation[95]. Adaptor sequences allow DNA fragments to be captured and tethered to a solid surface (microchip). Subsequent PCR is then carried out by implementing a bridging PCR reaction with Illumina technology. This allows single DNA fragments to be amplified non-specifically from the same clone, making the signals in the later stages to be more easily readable. See figure 1.4 for a schematic representation of the NGS process.

The Illumina chemistry utilises sequencing primers that are complementary to the ligated adapter sequences. The PCR reaction mix is made up of chain terminating deoxyribonucleotides, labelled with a specific fluorescent tag corresponding to the base type. During the extension phase, termed sequencing by synthesis (SBS), each individual base is recorded by its fluorescence. These fluorescent tags are then cleaved, removing the chain terminating effect and allowing further extension to occur. The sequence builds up sequentially over multiple cycles (of flowing each of the four bases).

Once the sequence fragments are created, they are aligned to the known reference sequence using complex bioinformatic techniques. Longer reads make this process more robust. The computational technology is able to measure the number of fragments over a given region (read depth or coverage) and highlight (call) sites of discordance to the reference sequence (variance)[93].

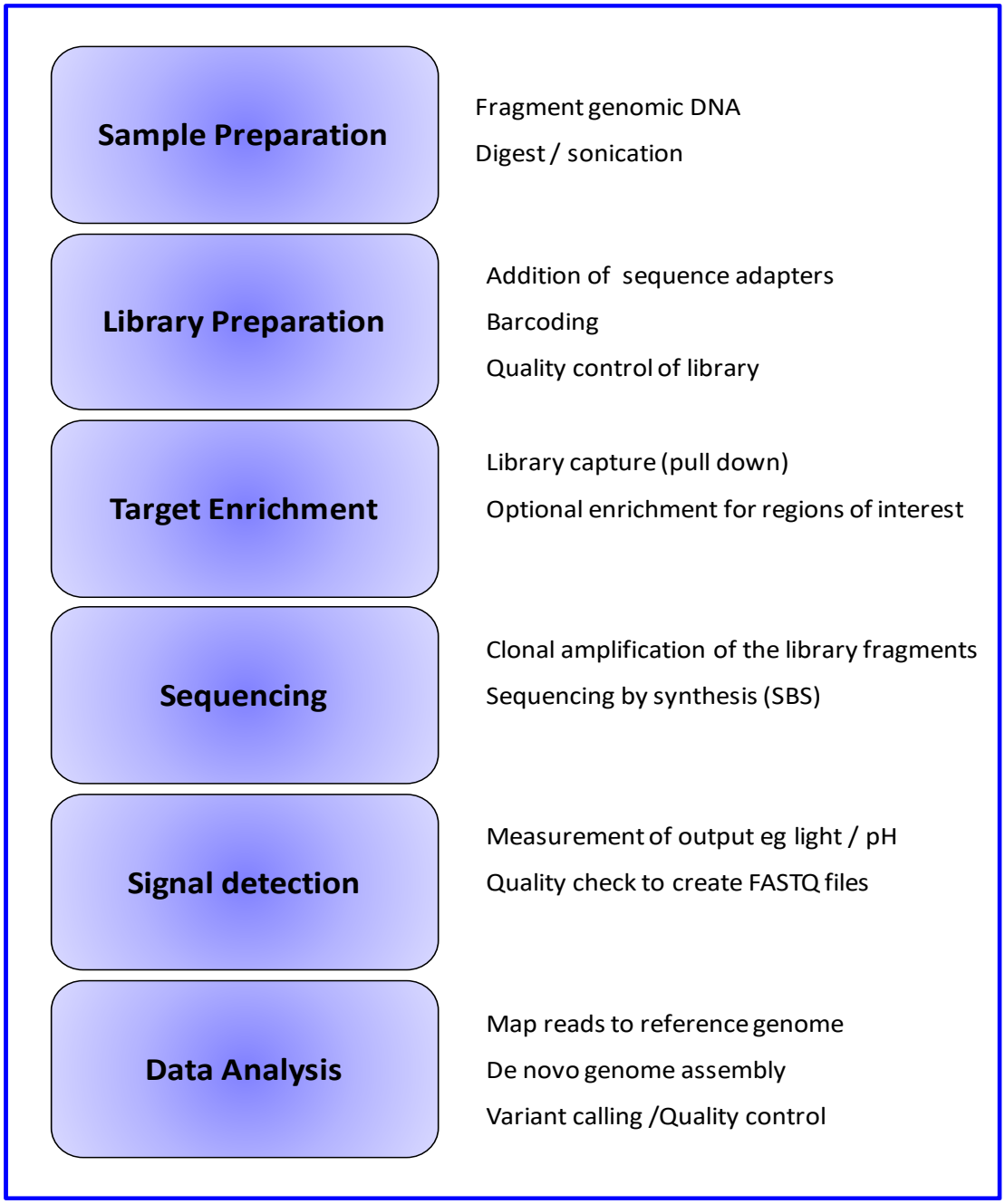


Figure 1.4: Schematic representation of Illumina based Next Generation Sequencing (NGS) Process.

In 2011, Chapman *et al* applied both whole genome sequencing (WGS) and WES technologies to a cohort of 38 myeloma patients. This was the first use of this technology in myeloma and a number of novel mutations in genes were reported, as well as providing additional evidence of the importance of the NF-κB pathway[59]. The involvement of *BRAF* in 4% of myeloma patients was also determined in these studies[96]. WGS has also been assessed in a linear fashion across a single patient's

disease course. The patient was shown to have t(4;14), and 10 other single nucleotide variants, seen at all timepoints. Other changes were shown to appear and disappear with time. Five new events were seen at the final timepoint of PCL and these changes have been postulated as leukaemic transformation events[97]. Since then many studies using this technology have been reported, and although this is not currently being utilised in the diagnostic and clinical setting for myeloma patients as yet, it is clear that this is the direction of travel based on its use in other haematological malignancies[98,99].

Project Aims and Objectives:

Plasma cell myeloma is a neoplastic disorder characterised by an abnormal monoclonal proliferation of plasma cells in the bone marrow and overproduction of circulating monoclonal immunoglobulin (paraprotein). Despite huge advances in the treatment of myeloma over the past two decades, and although myeloma is considered highly treatable, it remains an incurable disease in most patients.

Myeloma genetics is intrinsically complex and highly heterogeneous, but offers an opportunity to categorise the disease, offer prognosis based on these categorisations and potentially apply a personalised medicine approach. Relating genetics to potential treatment pathways, and a deeper understanding of myeloma genomics will become crucial for improved myeloma outcomes.

Access to diagnostic genetic testing is currently *ad hoc* and highly dependent upon local commissioning arrangements and whether patients are entered into national trials, even then the techniques employed and the exact nature of the testing vary dramatically. At the outset of this project there were no best practice guidelines for the genetic testing of myeloma and no quality assessment schemes associated with this testing.

My roles within Sheffield Diagnostic Genetics Service (SDGS), and more recently the Haematological Malignancy Diagnostic Service (HMDS), drive a strong interest in translational research. I am keen to ensure that the results obtained as part of this project are used directly to influence, not only the diagnostic services in both SDGS and HMDS, but best practice and the quality of genetic testing in myeloma across the UK.

Myeloma patients exhibit osteolytic disease in 70% of cases, and treatment with bisphosphonates can slow the progression of this bone disease. In early stages of the disease it is not always possible to determine who will go on to have severe bone disease. We would like to explore DNA changes in the osteome in relation to patient bone phenotypes to determine if an association might exist. Assessment of the 'osteome' related to the myeloma bone phenotypes would contribute to the research interests of the SMaRT (Sheffield Myeloma Research Team).

The aims of the project are to explore a number of different objectives:

Objective 1: To compare five different genetic technologies, readily available in a diagnostic genetic laboratory and their ability to identify genetic abnormalities and signatures associated with plasma cell neoplasms.

Patient bone marrow samples will be collected across all disease stages. We will perform karyotyping, FISH, MLPA, DNA array analysis & targeted NGS in order to establish genetic signatures and to assess their ability to do this in an effective and efficient manner, within the financial and time constraints required of a diagnostic testing scenario. A number of genetic changes have been shown to provide valuable information about disease prognosis and are beginning to guide treatment decisions.

Objective 2: To explore and assess possible or potential relationships between the genetic signatures associated with bone related genes, the 'osteome', and the likelihood or extent of bone damage associated with the patient's myeloma disease.

Within the targeted NGS panel, we aim to include genes associated with the 'osteome'; genes involved in the development, influence, maintenance and destruction of bone. Assessing these genetic findings against the clinical bone manifestations may allow correlations to be made. Treating the bone manifestations of myeloma is challenging; current treatments are able to offer protection and slowing of bone degradation, but can rarely reverse the process. Highlighting a cohort of patients who may be considered at risk of a severe bone phenotype associated with their myeloma may offer the opportunity to treat more proactively and slow the advancement of the bone manifestations. The use of bone anabolics is a specific interest of the research group, and this work may inform and link into that interest.

Objective 3: To generate recommendations for an all-encompassing diagnostic genetic panel for use in a diagnostic genetic laboratory, to introduce a quality assessment scheme for myeloma genetic testing and consider the requirement for

best practice recommendations to create a more consistent, equitable harmonisation of diagnostic genetic testing for myeloma within the UK.

Translation of myeloma research into diagnostic genetic testing for NHS patients is lagging far behind the work reported from the research and trial groups. Comparing the different genetic technologies associated with objective 1, will allow us to make informed decisions in proposing diagnostic genetic testing strategies for myeloma, that provide the clinical requirements of this service. Through links with UK NEQAS (GenQA), we aim to create a quality assessment programme for myeloma genetic testing including an educational component to the scheme. Understanding myeloma diagnostic genetic testing at the outset and towards the end of the project will allow the effects of the scheme to be assessed. Again, through links within the profession (ACGS, the new genomic laboratory hub (GLH) structure and GenQA), we would aim to influence best practice in this area.

Chapter 2: Material and Methods

This part-time PhD project forms part of a three year specialist programme grant, and a further two year extension funded by Bloodwise; Novel Targets and Therapeutic Combinations in Myeloma. Principle Investigator: Dr Andrew Chantry (LLR ref: Chantry 12053). Funding has been made available through this specialist programme grant, and through a further application made to the Sheffield Blood Cancer Charity.

Patient Cohort and Recruitment

Patients were recruited and consented to the study by Dr Chantry via patient clinics held at the Royal Hallamshire Hospital, Sheffield. Appropriate ethical approval governing collection of primary human material (serum, urine and bone marrow) from patients with myeloma was applied for as part of the Specialist Programme Grant. The ethics statement and letter are included as appendix 1 (REC reference: 05/Q2305/96) demonstrating approval from the South Sheffield Research Ethics Committee in August 2005 and subsequent ratification by the NHS Health Research Authority, National Research Ethics Committee Yorkshire and the Humber - Sheffield in November 2012 & 2016.

From January 2014, patient bone marrow (BM) samples were sent to Sheffield Diagnostic Genetic Service (SDGS) as part of the study. The CD138 positive separation process was introduced into SDGS at the end of March 2014, and it was samples from this point that were included as part of this PhD project.

A total of 101 bone marrow samples from patients with MGUS, plasma cell myeloma or plasma cell leukaemia were collected and anonymised for genetic analysis as part of this study. All samples except one (a case of plasma cell leukaemia) had a CD138+ve cell separation. Karyotyping was performed on 91 patient samples. All patients were processed for FISH and DNA extraction of the CD138+ve cell separation. MLPA was performed on 45 patient samples, 36 samples were processed for Affymetrix DNA array, and 24 samples were assessed using the bespoke NGS

gene panel. The numbers of samples processed was limited by the quantity and quality of the sample and by the funding available.

MACS CD138 Positive Cell Separation

The CD138 separation methodology utilises the Miltenyi Magnetic Activated Cell Separation (MACS) methodology and was undertaken according to the protocol supplied by the manufacturer www.miltenyibiotec.com/protocols. The CD138 antigen (aka syndecan-1) is primarily expressed on normal or malignant plasma cells in the bone marrow. It is not expressed on naïve B cells, germinal centre B cells, memory cells, T cells or monocytes.

Cell filtering: 1ml whole bone marrow sample was initially diluted in equal volume of running buffer (PBS, 2mM EDTA, 0.5% fetal calf serum) and filtered using the MACS pre-separation filter supplied by Miltenyi Biotec. A further 1ml of running buffer was used to wash the filter. This excludes large clumps and debris within the sample prior to separation and ensures that as much of the sample is utilised as possible.

Cell labelling: CD138 positive plasma cells were labelled using 50ul of CD138 microbeads per 1ml of sample. The beads and the sample were mixed well before incubating for 15 minutes on ice. The cells were washed by adding 5-10ml of running buffer and centrifuged at 1360 rpm for 10 minutes. The supernatant was removed before re-suspending in 3mls of running buffer.

Cell separation: 3ml of running buffer was applied to the column by way of preparation before the cell suspension was then loaded into a MACS column secured in the magnetic field of the MACS separator. The magnetically labelled CD138 positive cells were retained within the column whilst a series of washes 3x 3ml running buffer encourages the unlabelled cells to pass through the column.

CD138 positive cell capture: Following removal of the column from the magnetic field, the beads and the CD138 positive cell fraction were eluted using 5ml of Miltenyi elution buffer. The sample was centrifuged at 1300rpm, the supernatant removed and resuspended into 500ul of buffer. The sample was labelled and stored.

The CD138 positive cell selection was assessed for purity using flow cytometric methods, to obtain a percentage of plasma cells within the sample, and then the sample was split equally, for both DNA extraction and FISH analysis.

DNA Extraction

The Chemagen Magnetic Separation Module is employed at SDGS. This is an automated methodology to extract DNA from whole BM or CD138 positive cell populations. Chemagen separation kits are bought in and contain a lysis buffer, magnetic beads and binding buffer, a series of wash buffers and an elution buffer. The automated extraction occurred in two stages; the first stage involves mixing and lysing of cells in order to release the DNA and the second stage involves binding DNA to magnetic beads, in order to undergo a series of washes before subsequent elution from the beads into 300µl 1xTE buffer.

The protocol was followed according to the manufacturer's instructions www.chemagen.com/chemagic-kits.

Following extraction, a 260/280 ratio qualitative measurement was undertaken to determine the DNA yield and check for protein contamination, with 1.7-2.1 considered an acceptable range and the quantity measured and recorded using NanoDrop (model no: 8000, ThermoFisher).

Further DNA quantitation was undertaken for the array studies and the NGS studies using the Qubit 2.0 fluorometer. This system uses target-specific fluorescence which only emits light when specifically bound to double stranded DNA (dsDNA), this results in an increased accuracy in measurement of DNA concentration as the nanodrop measures both dsDNA and single stranded DNA (ssDNA), and RNA and free nucleotides.

Conventional Culturing and Karyotyping

The first method of the five methods to be assessed and compared as part of this project is conventional karyotyping. In order to prepare chromosomes for

karyotyping, the patient BM aspirate sample was treated according to the following methodology.

Culture set up: The patient BM sample was assessed for white cell count in order to determine the amount of sample to be added to 10mls of bone marrow culture medium (McCoys medium with added fetal calf serum, L glutamine, penicillin and streptomycin), resulting in a final concentration of 10^6 cells/ml. Two cultures were set up and cultured at 37^{o}C for 4 and 6 days respectively. This is much longer in duration than culturing for other leukaemic samples and is based on protocols described by Ross *et al*[42]. Plasma cells have a slow cell cycling time and this additional time in culture and the use of two culture times aims to capture as many plasma cells as possible.

Culture harvest: Colcemid was added to the cultures 20 hours prior to the harvesting. Colcemid is a colchicine analogue which prevents microtubule formation of the spindle apparatus, this in turn has the effect of arresting cells at the metaphase stage of cell division. Cultures were centrifuged at 1200rpm for 10 minutes, the supernatant was removed, and the cells re-suspended in hypotonic solution (0.075M KCl). The cells were incubated at 37^{o}C for 10mins before a further centrifugation step. Fixation using approximately 2-5mls 3:1 methanol:acetic acid was performed dropwise to avoid clumping of the cells, and centrifuged once again. The fixation step was repeated 3 times, or until the sample was clear, before being stored in the freezer.

Slide preparation and banding: Chromosomes were prepared and G-banded using standard laboratory techniques[100]. Two drops of fixed bone marrow suspension were dropped onto pre-washed, wet glass slides, and allowed to air dry. The slides were baked on a 60^{o}C hotplate for 10 minutes. The slide was flooded with trypsin working solution (10x trypsin stock solution dissolved into 49.5mls of a 0.9% Sodium Chloride solution) for 10-15 seconds before rinsing using Gurrs solution and then staining using 0.4% Giemsa stain for 90 seconds. Trypsin treatment produces the characteristic banding pattern on the chromosomes.

Karyotype analysis: The slides were scanned using the CytoVysion metaphase finder and analysed using CytoVysion software. Twenty cells were fully analysed where

possible (10 cells minimum) unless an abnormal clone was detected where two cells with the same abnormalities would be considered clonal. Where only normal cells were detected all cells captured (in the region of 200 cells dependant on the culturing success) were viewed to assess for gross abnormality.

See appendix 2 for reagent list associated with culturing bone marrow.

Fluorescence in situ Hybridisation (FISH)

A minimum probe set of *IGH/FGFR3, IGH/MAF* and *TP53* has been suggested for the analysis of myeloma samples[42]. The FISH strategy employed at SDGS is considered a two-step process. Step 1 involves the *IGH*, *TP53* and *CDKN2C/CKS1B* probes, and step 2 is initiated if an *IGH* rearrangement is detected at step 1. The aim of step 2 is to elucidate the *IGH* partner and includes *IGH/CCND1, IGH/FGFR3* and *IGH/MAF* probes (the most common low risk, and the two most common high risk rearrangements). The *IGH* breakapart probe was obtained from Vysis (Abbott Molecular, Des Plaines, IL) and *TP53, CDKN2C/CKS1B, IGH/FGFR3, IGH/MAF* and *IGH/CCND1* were obtained from Cytocell (Cambridge, UK). Figure 2.1 shows the probe set employed in step 1.

FISH was performed by using standard methods on interphase cells and in accordance with the manufacturer's instructions for co-denaturation using ULTRAhyb (Life technologies) probe buffer:

https://www.abbottmolecular.com/products/oncology/fish/hematology-probes.html, &

http://www.cytocell.co.uk/products/aquarius/haematology-probes/

Slide preparation: Slides for FISH processing were prepared as described in the chromosome karyotyping section above up until the baking step.

Slide pretreatment: The slides were incubated in 2xSCC at 75°C for 15 minutes, and then transferred to a pepsin solution (4mg/ml in 0.9% NaCl, pH 1.5) for digestion for 15mins at 37°C. The slides were then washed in 2xSSC for 5mins at room temperature, before an ethanol series (70%, 95% and 100%) for 2mins each was used to dehydrate the cell preparations.

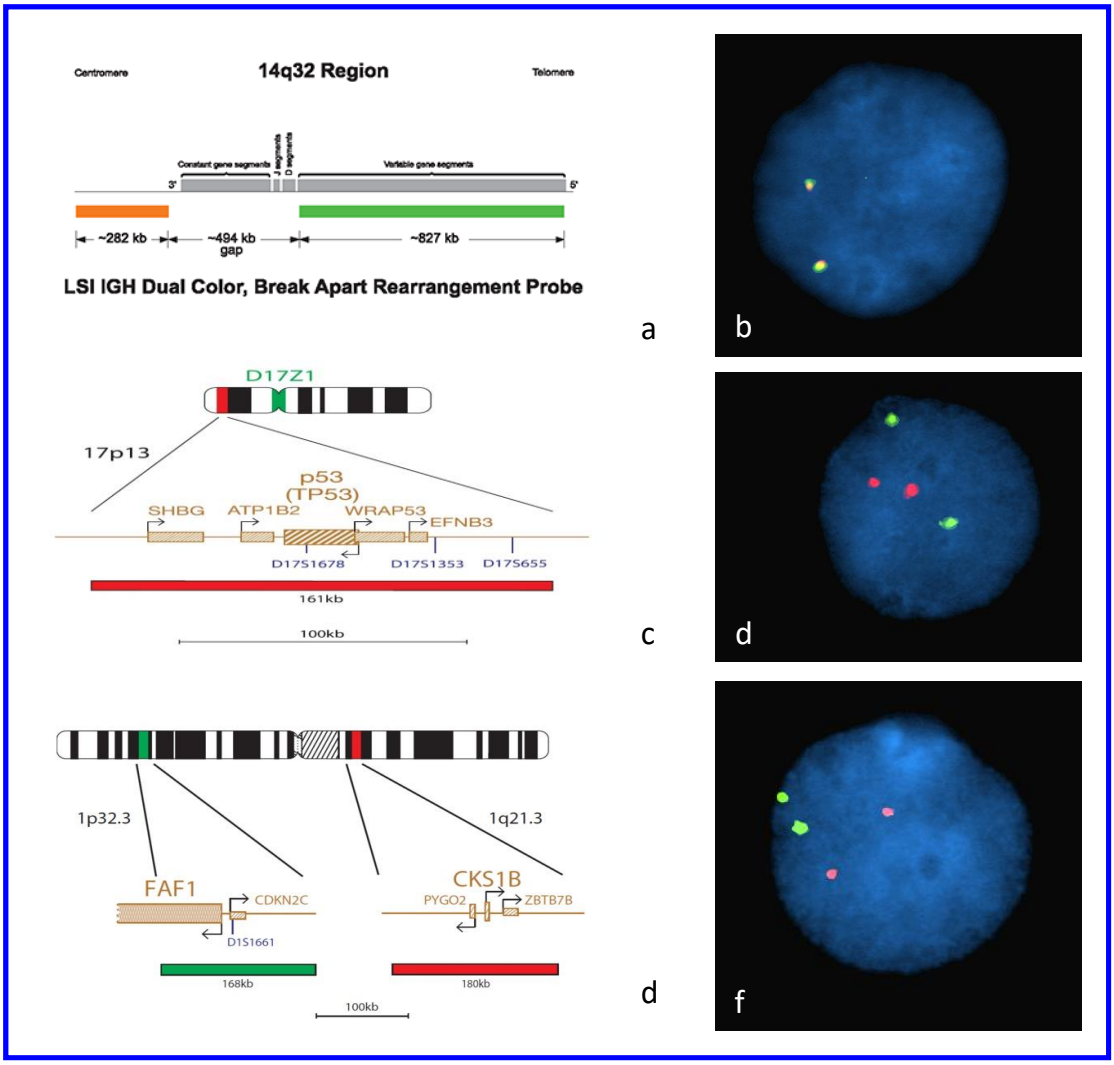


Figure 2.1: Representation of the FISH probes used in Step 1 of the described myeloma FISH strategy. This includes the Vysis (Abbott) *IGH* breakapart probe (a) and the Cytocell *TP53*/17cen (c) and *CDKN2C/CKS1B* (1p/1q) (e) probes. FISH images (b, d & f) demonstrate the normal expected pattern for each of these probes.

Probe preparation: Probes were prepared according to the manufacturer's instructions, using hybridisation buffer containing dextran sulphate, formamide and SCC (pH 7.0), added to a coverslip which was overlayed onto the slide and sealed using rubber cement.

Denaturation and hybridisation: Co-denaturation of both the target DNA and the probes was performed on a PCR thermocycler adapted for slides, at 72°C for 2mins and then hybridised at 37°C for 16 hours (overnight).

Post hybridisation washes: The following morning a series of washes were performed to remove excess probe. The rubber cement seal and the coverslip were removed, and the slides washed for 2mins in 0.4xSSCT (T denoting the addition of

the detergent tween) at 75°C. The slides were then transferred to 2xSSC for 2mins at room temperature, before using the same ethanol series to dehydrate the slides before airdrying. The slides were mounted using DAPI (4′,6-diamidino-2-phenylindole) counterstain in vectashield antifade.

See appendix 2 for reagent list associated with FISH.

FISH analysis: The slides were then analysed using a fluorescence microscope set up with the appropriate triple, dual, and single band pass filters to allow visualisation of the individual fluorophores associated with each probe. At least 200 interphase nuclei were scored for each probe set per sample. It is important at this stage to understand the probe set being employed, and what both expected normal and abnormal populations would demonstrate in terms of signal patterns.

Multiplex Ligation-dependant Probe Amplification (MLPA)

The MRC Holland MLPA probe mix for myeloma (SALSA MLPA probemix P425-B1 Multiple Myeloma) contains 46 probes for the following chromosomal regions, 1p32-p12, 1q21-q23, 5q31, chr 9, 12p13, 13q14, 14q32, chr 15, 16q12-q23 and 17p13, all thought to have diagnostic / prognostic relevance for myeloma patients, and 11 reference probes for regions thought to be relatively stable in myeloma genomes.

The SDGS in-house protocol is based on the one-tube protocol supplied by the manufacturer http://www.mrc-holland.com but has been validated in house to use half the volumes. DNA is required at 50ng/µl diluted in 1xTE buffer. The protocol was undertaken as follows:

DNA denaturation: 2.5µl of DNA was incubated on a thermocycler for 5 minutes at 98°C

Hybridisation of probes to sample DNA: The sample was cooled to room temperature, and 1.5µl of the 1:1 probemix:MLPA buffer was added and mixed. This was incubated for 1 minute at 95°C and then hybridised for 16 hours at 60°C overnight.

Ligation of hybridised probes: The thermocycler temperature was reduced to 54°C, and 16µl of ligase master mix was added and incubated for 15mins at 54°C. The ligase enzyme was then thermally inactivated by heating to 98°C for 5 minutes.

PCR amplification of ligated probes: The samples were cooled to room temperature, and 5µl of the polymerase master mix was added and gently mixed. The PCR thermocycler program was set to 35 cycles of 95°C for 30 seconds, 60°C for 30 seconds and 72°C for 60 seconds. At the end of 35 cycles the reaction was held at 72°C for 20 minutes and then cooled to 15°C for 10 minutes.

Capillary electrophoresis of PCR products: The PCR product was then processed using the 3730 GeneMapper.

Analysis: Results were transferred from GeneMapper to the Coffalyser analytical software supplied by MRC Holland, to assess the control quality and determine the relative size of the fluorescent peaks within each sample, which were compared to the reference samples. The results were expressed as ratios for the likelihood of a probe being either deleted or duplicated.

Affymetrix DNA Array – CytoScan HD

The Affymetrix arrays were run using the equipment and hardware available at Sheffield Institute of Translational Neuroscience (SITraN). Dr Paul Heath and his team were available for help and guidance throughout this process.

The Affymetrix CytoScan High Density (HD) Array was chosen as the array of choice for the analysis of the myeloma patient samples in this PhD project. The CytoScan HD array offers a method to detect high resolution copy number changes across the human genome and utilises single nucleotide polymorphisms (SNPs) to determine loss of heterozygosity (LOH). This high density array has >2.4 million markers; 1.7 million unique non-polymorphic probes designed to offer coverage of RefSeq and OMIM genes relevant to both constitutional and cancer, and 750,000 SNPs taken from dbSNP and designed to offer maximum genomic coverage, genotyping accuracy

and optimised for LOH detection. CytoScan HD Array can reliably detect copy number changes of 25-50kb across the genome with high specificity.

The CytoScan processing protocol was run over four days and can be divided into the following stages:

Genomic DNA requirements: 5µl of DNA was required at a concentration of 50ng/µl (250ng in total)

Restriction enzyme digestion: A digestion master mix, made up with Nsp1 enzyme was prepared, added to the samples and incubated at 37°C for 2 hours before thermally arresting the reaction at 65°C for 20 minutes, and holding at 4°C.

Ligation: A ligation master mix was made up including the supplied ligase buffer, the Nsp1 adaptor and DNA ligase. 5.25µl was added to the samples and incubated for 3 hours at 16°C. The reaction was thermally arrested at 70°C for 20 minutes, and then held at 4°C.

PCR step: The samples were diluted (1:3) in nuclease-free water. A PCR master mix was made up with PCR primers, DNA polymerase, dNTP mixture, GC melt reagent and a supplied PCR buffer. 10µl of the sample dilution was then added to 90µl of PCR master mix in quadruplicate. The thermal cycler lid was preheated and following 3 minutes incubation at 94°C, was then run for 30 cycles of 94°C for 30 seconds, 60°C for 45 seconds and 68°C for 15 seconds. At the end of 35 cycles the reaction was completed at 68°C for 7 minutes and was then held at 4°C.

PCR product purification: 3µl of the PCR product was run on a 2% agarose gel at 5V/cm for 45 minutes to ensure that the product was between 150-2000bp. A purification step was performed in which magnetic purification beads were used to attract the DNA whilst a series of washes were performed. A final elution step removed the DNA from the magnetic beads.

Product quantitation: The sample concentration was determined using nanodrop and a 260/280 ratio qualification measurement was undertaken. This quality control step required the 260/280 ratio to be between 1.8 and 2.0 and the concentration to be >2.5µg/ηl in order to proceed.

Product fragmentation: A fragmentation master mix, using the fragmentation reagent supplied with the kit, was made up. 10µl of the master mix was added to 45µl of the purified PCR product and incubated for 35 minutes at 37°C. The reaction was thermally arrested at 95°C for 15 minutes, and then held at 4°C. The fragmented PCR product was run on a 4% TBE gel at 5V/cm for 45 minutes to ensure fragment distribution was between 25-125bp.

Product labelling: A labelling master mix was made up using the DNA labelling reagent supplied with the kit. 19.5µl of the master mix was added to 51µl of the fragmented product and incubated for 4 hours at 37°C. The reaction was thermally arrested at 95°C for 15 minutes, and then held at 4°C.

Target array hybridisation: A hybridisation master mix was made according to the manufacturer's instructions and 190µl was added to each sample. This was heated to 95°C for 10 minutes in order to create single stranded DNA and then held at 49°C. 200µl of the sample and hybridisation master mix was loaded into the DNA array cassettes ensuring even flooding of the array and sealed. The cassettes were loaded into the hybridisation oven and incubated for 16-18 hours at 50°C, turning at 60rpm.

Washing, staining and scanning the arrays: The array cassettes were loaded into the fluidics stations and automatically processed to wash and stain the arrays according to the preloaded software. Once the wash and stain process was complete the arrays were scanned using the CytoScan software.

Analysis: The scanner records a high-resolution image. This is converted to a .CEL file, which is a text (ASCII) file containing the intensity of each probe, the standard deviation, the number of pixels used in the calculation, and a flag marking any results as an outlier. Analysis of the files produced by the scanner was carried out using the Chromosome Analysis Suite (ChAS) software supplied by Affymetrix. The .CEL file was read by the ChAS analysis workflow and converted to produce a cyhd.cychp file, and this cyhd.cychp file was further processed to provide data for interpretation. Three array QC metrics are provided for the CytoScan arrays, however only the MAPD (Median of the Absolute values of all Pairwise Differences) should be considered for non-constitutional samples and was therefore used as the quality metric for this analysis. The $\log_2$ratio, smooth signal, and LOH tracks were

used for analysis. Table 2.1 details the cut-off levels set for each of the abnormality types in order to be called by the software.

Abnormality type	Marker count	Size (kbp)
Gain	10	100
Mosaic Gain	10	100
Loss	10	100
Mosaic Loss	10	100
Loss of Heterozygosity (LOH)	-	10,000

Table 2.1: Details the cut off levels set for each of the different abnormality types in order for the Affymetrix Chas array software to call an abnormality.

Targeted Next Generation Sequencing (NGS) Gene Panel

The wet work for the NGS was completed through the research group at SDGS, funded by the PhD project. Whilst the panel design was entirely my work, and much discussion was had with the NGS team, the wet work and the initial bioinformatic pipeline was completed by Dr Elsie Place and Dr Matt Parker. Analysis of the pipeline output was again my work.

NGS Panel Design

NGS panels are often utilised as a way of applying NGS technology in a more manageable way, from both a technological, analytical and financial point of view. WGS and WES are often excluded based on the current costs. NGS panels involve targeting smaller regions of interest, and targeting the NGS to the exonic (or even mutation hotspot) regions of genes.

For this PhD project, a custom Agilent SureSelect hybridisation probe set was designed. The hybridisation probes are biotinylated RNA sequences complementary to the regions of interest. When the probes are incubated with the fragmented

sample DNA, the probes have the effect of 'pulling down' and enriching the regions of interest. The probe design was completed as a single panel of 139 genes, but was essentially made up of two panels; the SMaRT NGS Myeloma Gene Mutation Panel which included 79 genes known to be involved in myelomagenesis, and the SMaRT NGS Osteome Probe Panel which had 60 bone related genes. The 'osteome' NGS panel was required to address a further objective of the PhD project, and playing to the strength of the SMaRT. The genes included in the panels are laid out in tables 2.2 and 2.3.

SMaRT NGS Myeloma Gene Mutation Panel									
ANP32E	ATM	ATR	ATRIP	B2M	BCL2	BCL6	BCL7A	BIRC2	BIRC3
BRAF	CCND1	CCND2	CCND3	CDKN2A	CDKN2C	CHD4	CKS1B	CRBN	CUL4A
CUL4B	CYLD	DDB1	DIS3	DNAH5	EGR1	ERBB4	FAF1	FAM46C	FAT1
FAT3	FAT4	FGFR3	FLT3	HIST1H1E	HOXA9	IKZF1	IKZF2	IKZF3	IRF4
KDM6A	KDM6B	KMT2A	KRAS	LRRK2	LTB	LTBR	MAF	MAFB	MAPK1
MAX	MRE11A	MYC	NCKAP5	NF1	NFKB1	NFKB2	NR3C1	NRAS	NRM
PARP1	PIK3CA	PRDM1	PRKD2	PSMB5	PSMG2	RASA2	RB1	RET	ROBO1
ROS1	SF3B1	SP140	TP53	TRAF3	WHSC1	WWOX	XBP1	ZFHX4	

Table 2.2: Detailing the genes included in the bespoke NGS panel covering known mutations associated with myeloma.

SMaRT NGS Osteome Probe Panel									
ACVR1	AKT1	CCL3	CCL4	CCR1	CCR3	CDH2	CER1	CSF2	CTNNB1
DCN	DKK1	DKK2	FRZB	FZD1	FZD2	FZD3	FZD4	HGF	ICAM1
JAG1	KREMEN1	KREMEN2	LRP5	LRP6	MET	MMP9	NOG	NOTCH1	POSTN
RUNX2	SDC1	SFRP1	SFRP2	SFRP4	SMAD2	SMAD3	SMAD4	SMAD7	SOST
TAZ	TGFB1	TGFB2	TGFB3	TNF	TNFRSF11A	TNFRSF11B	TNFRSF13B	TNFSF11	TNFSF12
TNFSF13	TNFSF13B	VCAM1	VEGFA	WIF1	WNT10A	WNT10B	WNT3	WNT3A	WNT5A

Table 2.3: Detailing the genes included in the bespoke NGS panel covering the 'osteome', - bone related genes.

The genes selected for inclusion in these panels were based on a number of papers, and also from experience of the SMaRT. The composition of the osteome NGS panel will be discussed further in the materials and methods section of Chapter 4. The myeloma gene NGS panel was initially based on three papers which had been published at the time of NGS panel design detailing abnormalities which were considered to be of importance in the diagnosis of myeloma:

- Walker *et al* (2015). Mutational spectrum, copy number changes, and outcome: Results of a sequencing study of patients with newly diagnosed myeloma. J Clin Oncol 33:3911-20[38].
- Lohr *et al* (2014). Widespread heterogeneity in multiple myeloma: Implications for targeted therapy. Cancer Cell 25:91-101[61]
- Bolli *et al* (2014). Heterogeneity of genomic evolution and mutational profiles in multiple myeloma. Nature Communications 5:2997[62]

These papers covered mutations that were recurrently seen in myeloma, those considered to be associated with prognostic impact and those considered to be predictive for response to current or potential future therapies. Gene names were used according to the HUGO gene nomenclature committee[101].

Walker *et al,* described 13 significantly mutated genes which included, *KRAS, NRAS, TRAF3, TP53, FAM46C, DIS3, BRAF, LTB, CYLD, RB1, HIST1H1E, IRF4* and *MAX*[38]. Lohr *et al* described a similar set of 11 significantly mutated genes which did not include

LTB, HIST1H1E, IRF4 or *MAX*, but did include *PRDM1* and *ACTG1*[61], and *Bolli et al* described seven highly recurrent *genes KRAS, NRAS, BRAF, TP53, FAM46C, SP140* and *LTB*[62].

Genes associated with the RAS/MAPK pathway are the most frequently mutated in all series, seen in approximately 43.2% of patients, these genes as well as *MAPK1* were included in the SMaRT myeloma NGS panel[62]. The NF-κB pathway is the next most frequently affected, seen in 17% of patients, the *NFKB1, NFKB2 and LTBR* genes were included. *TP53* variants were seen in 11% of myeloma cases, other genes associated with delivering an apoptotic pathway, *ATM*, *ATR, LRRK2, MRE11A* and *FAF1* were also included.

IRF4 is a key plasma cell survival gene and downstream of cereblon (*CRBN*), the target of the IMiD group of drugs used for myeloma patients. *EGR1* is another IMiD target gene, and the Ikaros group of genes (*IKZF1, IKZF2* and *IKZF3*) have also shown association with cereblon and these were also included in the panel. A further paper from Kortüm *et al* published in 2015[102] also described an NGS panel which included additional genes associated with specific treatment types. As well as the genes already described, they included *CUL4A, CUL4B* and *DDB1* known to be associated with IMiD resistance, *PSMG2* and *PSMB5* associated with proteosome inhibitors and *NR3C1* associated with glucocorticoid therapies, as well as *DNAH5* and *XBP1* which were added to our panel.

Walker *at al* also described a number of significantly mutated genes associated with specific cytogenetic sub-groups, for example *FGFR3* was only mutated in t(4;14) cases, and a further five genes were associated with this group; *PRDM1, BCL7A, ATRIP, NRM* and *PRKD2. CCND1* mutation was highly associated with the t(11;14) group and *EGR1* was associated with the hyperdiploid group[38]. The other common *IGH* gene partners were also added in at this stage; *MAF, MAFB, CCND3, WHSC1* as well as *CCND2,* although less common, and then finally *MYC.*

CHD4 is known to interact with *ZFHX4* to modulate *TP53* function and mutations were seen in both genes. These were also shown to be associated with a negative survival impact as were *CCND1, NCKAP5, ATM* and *ATR. IRF4* and *EGR1* were shown to have a positive prognostic effect[38].

Bolli *et al* go on to describe a further set of genes that may be associated with myeloma using less stringent cut-offs. These included: *ROBO1, EGR1, FAT3* a transmembrane protein which is part of the cadherin family of genes, *NFKB1, NFKBIA, CYLD, TRAF3,* and some rarer mutations associated with other known cancer genes such as *RAG2, SF3B1, PIK3CA, PTEN, KDM6A, CDKN2C* and *SETD2*[62]. Finally, a set of known genes associated with other lymphoid diseases and cancer have been implicated such as *BCL2, BCL6, ERBB4, FLT3, HOXA9, KMT2A, PARP1, RET* and *ROS1,* some of which were included in our panel along with closely related genes e.g. *FAT1, FAT4, CDKN2A, KDM6B.*

B2M, beta2 microglobulin is known to be increased in patients with myeloma and this gene was also added to the NGS panel. A number of genes known to be associated with copy number change in myeloma were also included in the panel, some previously associated with mutations using NGS, and these included: *CDKN2C, CKS1B, BIRC2, BIRC3, CYLD* and *WWOX* [22].

The probe set covers all exonic coding regions, from all transcripts (if multiple transcripts were present) and including +/-25bp into the intron/exon boundaries. The National Centre for Biotechnology (NCBI) website was used to obtain the reference sequence accession numbers for the genes involved:

www.ncbi.nlm.nih.gov

A BED file was downloaded based on these accession numbers containing the genomic coordinates for the start and finish +/-25bp for all exons, of all known transcripts of the genes involved. These co-ordinates were then uploaded to the panel design module of the Agilent web-based SureDesign software:

https://earray.chem.agilent.com/suredesign/

Selection parameters were applied; Density was set to 5x, meaning that the overlapping, staggered probes covered the region five deep in a tiled manner, Masking was set at moderately stringent, this hides repetitive regions so that probes are not designed to these regions, and finally Boosting was set to Max performance, which ensures that probes with a higher GC content are replicated by a higher factor.

The software provides tiled probes that can be downloaded again as a BED file, which gives information on regions that have been missed on the first probe selection. These missed regions were then resubmitted at lower stringency selection parameters to ensure all regions were adequately covered. These additional probes were added to the original probe set and assessed using the UCSC (University of California Santa Cruz) genome browser alignment tool (BLAT):

https://genome.ucsc.edu/

This gives a quality score, based on how many other regions of the genome the probes map to, and any probes that had a score of >40 were removed.

Once the design of the probe panel was complete, it was ordered through the website for production.

NGS Processing and Sequencing

Twenty four samples were processed and these included both DNA from the CD138+ve cell selection, i.e. DNA from the myeloma tumour cells, and matched blood samples, representing the germline patient DNA.

Library preparation: The blood sample libraries were prepared with the Agilent SureSelect QXT kit according to the manufacturer's protocol using 50ηg input genomic DNA (gDNA). Briefly, the gDNA is enzymatically fragmented and adapter tagged in a single 10 minute reaction. The samples were purified using AMPure XP beads, before amplifying through 8 cycles of PCR. The PCR primers overlap with the ligated adapter sequences and, at this stage, are unindexed.

The CD138 positive cell selection DNAs from the patient BM samples were prepared with the Agilent SureSelect XT HS according to the manufacturer's protocol using 32-200ηg input DNA. The starting DNA was fragmented to ~200bp using the Covaris E220 instrument. Adapters were ligated to the fragments using a 30min incubation step at 20°C, followed by an A-tailing step by way of end repair. Again, the samples were purified using AMPure XP beads, before amplifying the adapter-ligated library. Amplification of the samples uses the adaptor sequences ligated to the DNA

fragments and involves the use of one universal primer and one indexed primer, which is unique to each sample. The number of PCR cycles required is dependent upon the starting amount of DNA, in this case 8-12 cycles were used, based on the variable (but around 50ηg) DNA input.

For both sets of libraries, a further purification step is required post amplification, and before assessment of the quality and quantity.

Pre-hybridisation quality assessment: The resulting DNA libraries were assessed for quality and quantity using the D1000 ScreenTape and reagents on an Agilent TapeStation 2200, and concentrations measured using Qubit high sensitivity ds DNA assay. The samples should ideally demonstrate a peak fragment size of 200-400bp in size.

Hybridisation and capture: The libraries from both the blood sample DNA, and the bone marrow CD138+ve cell selection DNA, were hybridised to the target-specific probes designed as part of the bespoke NGS panel. Hybridisation occurred on a thermocycler using the manufacturer's suggested run cycle. Capture of the regions of interest was carried out using streptavidin-coated magnetic beads. The samples were incubated with the magnetic beads for 30mins at room temperature, and a magnetic separator device designed for 96 well plates, effectively 'pulls down' the beads. A series of washes were performed, before the bead-bound, target enriched DNA was further amplified using a PCR step. At this stage, for the germline samples (blood samples), this PCR step also served to introduce sample indexes in a similar way to that described for the CD138+ve libraries. A further purification step was completed.

Post hybridisation quality assessment: As described for the pre-hybridisation quality assessment the resulting post capture DNA libraries were analysed on the Agilent TapeStation 2200 using the high sensitivity ScreenTape. A peak fragment size of 200-400bp was expected following this step also.

Sample pooling and sequencing: The DNA libraries are denatured and passed through an Illumina eight lane flow cell; these flow cells are lined with millions of oligonucleotides, which are complementary to the adapters added to the DNA

libraries, allowing binding to occur. Amplification occurs to generate amplicon clusters in a process named parallel bridge amplification. Further denaturation leaves only the amplicon strands as extensions to the oligonucleotides adhered to the flow cell. These strands are then 'sequenced by synthesis' (SBS); SBS involves the addition of a complementary terminating fluorescent nucleotide in each cycle of sequencing, at which stage the fluorescence of each cluster is read. The termination molecule and the fluorescence is cleaved before moving on to the next cycle. This was repeated for 108 cycles in forward and reverse directions, resulting in 216 total cycles, with each cluster generating one forward and one reverse sequence read.

The pooling was performed such that blood sample DNA libraries were allocated approximately 1/100th of a lane each, and the bone marrow sample CD138+ve cell selection DNA libraries were allocated approximately 1/14th of a lane each. This was calculated based on the approximate read depth required for each sample type, 30x read depth for the germline samples and 500x read depth for the myeloma disease samples.

A paired end 2x 100bp sequencing run was performed on a HiSeq 2500 using TruSeq (v3) reagents, with an 8bp index 1 read and a 10bp index 2 read. Index Sequences, pooling data and final library sizes are provided in appendix 3.

Bioinformatics: Raw FASTQ DNA sequence files obtained from the Illumina HiSeq platform are analysed through a bioinformatics pipeline developed at SDGS. This pipeline conforms to the Broad Institute best practice guidance with some additional customised data filtering and QC check steps. The pipeline aligns reads to the hg19 (GRCh37) build human genome using Burrows-Wheeler Alignment (BWA)[103]. Variants were called using the Genome Analysis Toolkit (GATK) software[104], and were then analysed using a variety of bioinformatics tools; mutect2, vardict and strelka[104-106] to assess the likelihood of associated pathogenicity. A consensus set based on calls from all three tools was created. This list of variants with further annotation including ClinVar and ExAC allele frequencies as well as links to the Myeloma or Osteome gene panel, was provided from the SDGS service for further analysis, and interpretation. ClinVar is a freely accessible, public archive of the relationships of human variants and phenotypes with supporting evidence. ExAC (the Exome

Aggregation Consortium) database covers a number of large-scale sequencing projects with the aim of aggregating and harmonising exome sequencing data.

A full list of the bioinformatics reference files, and software tools applied to the NGS data are detailed in appendix 4.

Genetic Data Analysis

The output from the five genetic methodologies described is broad, and direct comparison can be difficult, as the data are not presented in the same way. It can relate to individual pieces of genomic information about specific genes or regions to whole genome wide analyses, and can represent analysis on many different levels of resolution from base pair level to megabase level.

Cytogenetics provides a karyotype image and an International System for Cytogenetic Nomenclature (ISCN) description. ISCN is the coded written description of the karyotype. FISH provides information on the specific regions assessed; this can be rearrangement in the case of *IGH* and its partner genes, loss/deletion or gain/duplication in the case of 1p/1q and *TP53* FISH. FISH results can be represented in ISCN, but this is often considered complex and difficult to interpret and therefore summary statements are often used. MLPA will also offer information on loss or gain of the regions included in the panel, these results can be written using ISCN, but again is not considered as user friendly as written summary statements. Array analysis again, highlights areas of loss and gain but on a whole genome wide level as well as providing evidence of loss of heterozygosity (LOH). Once again, ISCN nomenclature is available for array results, but is not always the method of choice to display these results. Finally, NGS results in lists of variants for analysis and interpretation of pathogenicity, usually variants are written according to Human Genome Variation Society (HGVS)[107]. HGVS nomenclature, is equivalent to ISCN, when reporting and exchanging information about variants in DNA, RNA and protein sequence and serves as an international standard.

Chapter 3: Genetic Results from the Patient Cohort

Patient Cohort

The rationale for this part of the project was to assess five different genetic technologies on a set of patient samples received as part of the research project. The five technologies included cytogenetic analysis (karyotyping), FISH, MLPA, DNA arrays and NGS, all technologies available within an NHS diagnostic laboratory. This covers Objective 1, which aims to establish genetic signatures for a set of patients in order to assess the ability of each technology to do this in an effective and efficient manner, within the requirements of a diagnostic testing scenario. The results from this part of the project will also contribute towards Objective 3, in which a best practice testing strategy will be proposed, alongside setting up a quality assessment scheme for the genetic testing of myeloma.

Since January 2014, just over 100 patient bone marrow samples were received in SDGS, following consent taken by Dr Andrew Chantry, for diagnostic analysis and use in this research project. Twelve patients were shown to have, either an unrelated neoplastic condition or a normal bone marrow, and these patients were excluded from the research project analysis cohort.

All patient BM samples received in medium were set up for long term cell culture for karyotypic analysis on the whole bone marrow sample.

MACS separation was introduced as a validated technique in March 2014 in SDGS, and only samples that had been through this process were used as part of the patient cohort for this project. One exception to this was patient #113, who was diagnosed with plasma cell leukaemia and showed an extremely high level of plasma cells (80%), and therefore did not require the plasma cell selection process in order to target and analyse the plasma cells.

CD138 positive cells were processed for both FISH and DNA extraction, where enough material was available. FISH processing and analysis were prioritised where there was not sufficient material, as this makes up the standard of care (SOC) analysis for diagnostic workup of myeloma patients at diagnosis. FISH for *IGH, TP53*

and *CDKN2C/CKS1B* (including further follow up, using *IGH/CCND1, IGH/FGFR3* and *IGH/MAF* probes, if *IGH* is rearranged) was performed.

MLPA was performed on 45 patients who were selected based on those who had enough DNA at sufficient quality to perform the analysis.

The Affymetrix array analysis demands good quality DNA, at a quantity of 250ng at a concentration of 50ng/µl. Patients, again, were selected based on the quality and quantity of DNA available, and was completed on a cohort of 36 patients.

Finally, the NGS sequencing was performed on 24 patients, and the cohort was chosen not only by the availability of sufficient DNA from the CD138+ve cell selection, but by the availability of a germline sample. In the case of myeloma, the patient's blood is not involved in the disease and therefore this is a suitable sample type for germline analysis. Although blood samples were taken alongside the bone marrow for most patients, following consent for the project, this sample was not available for all patients. The NGS arm of the project not only assessed the feasibility of using NGS technology in the diagnosis of myeloma, but also explored a possible correlation or predisposition to an associated bone phenotype in myeloma through abnormalities that may or may not be found in the 'osteome'; bone related genes. It was important therefore, that this cohort represented myeloma/MGUS with and without bone involvement, and ideally split 50:50. This arm of the project is further explained and discussed in Chapter 4. The NGS processing and analysis has a substantial cost associated with it, and the number of patients processed for NGS were, in part, limited by funds available for the project.

The quality and quantity of patient samples varied dramatically across the full patient cohort. Although the ideal would have been to carry out the full genomic analysis on a single set of patients, very few patients were likely to have had enough genetic material, and the funds and the time for the project did not allow for all 100 patients to have all five technologies applied. However, over the whole cohort, we have data on a subset of patients for each technology described here.

The patient cohort included samples from 101 patients in total, 91 were karyotyped, 101 were FISHed, 45 had MLPA analysis, 36 were processed for the Affymetrix DNA

array and 24 had the NGS bespoke gene panel analysis. Table 3.1 shows the number of cases that had which tests applied. Four cases had only FISH analysis, and a single case had all five genetic technologies. The majority of patient cases, 68/101 (67.3%) had either three or four of the genetic technologies applied.

Test(s) applied	Number of Cases
FISH	4
Cytogenetics & FISH	23
FISH & MLPA	5
Cytogenetics, FISH & MLPA	21
Cytogenetics, FISH & Array	13
Cytogenetics, FISH & NGS	5
FISH, MLPA & NGS	1
Cytogenetics, FISH, MLPA & Array	11
Cytogenetics, FISH, MLPA & NGS	6
Cytogenetics, FISH, Array & NGS	11
Cytogenetics, FISH, MLPA, Array & NGS	1

Table 3.1 Demonstrating the number of cases that had the described tests applied. Four cases had a single FISH test and only case had all five technologies applied, however the vast majority, 68/101 patient samples had either three or four technologies applied.

Table 3.2 shows all 101 patients in the sample cohort used as part of this research project demonstrating the genomic analysis applied to each patient, the diagnosis at

the study timepoint and the percentage of plasma cells seen. This table also includes a basic result of the test, although these results will be discussed in further detail within each specific technique results section. Failed tests are coloured purple, NAD (no abnormality detected) tests are coloured green, and all abnormal tests are coloured blue.

Of the 101 patient samples, 89 (88.1%) showed abnormality (or a failed / unknown result) in at least one of the techniques applied. Twelve of the total 101 cases showed normal results across the full range of tests applied, however, it should be noted that nine of those 12 patients only had 1 or 2 test types, and therefore abnormality may have been detected using the alternative testing technologies.

Sample No:	Cytogenetics	FISH	MLPA	Array	NGS	Diagnosis	Plasma cells (%)
#113	Complex	Abnormal	Abnormal	Complex		Plasma cell myeloma	63
#119	NAD	NAD		Hyperdiploid		Plasma cell myeloma	36
#134	Complex	Abnormal	Abnormal	Complex		Plasma cell myeloma	70
#136	Complex	Abnormal	Failed		Abnormal	Plasma cell myeloma	70
#138	NAD	Abnormal	Failed			Plasma cell myeloma	60
#139	NAD	Abnormal				Plasma cell myeloma	70
#141	NAD	Abnormal	Abnormal			Plasma cell myeloma	
#140	NAD	Abnormal	NAD		Abnormal	Plasma cell myeloma	15
#142	NAD	NAD	Abnormal			MGUS	8
#144	Missing Y	NAD	Abnormal	Complex	Abnormal	Plasma cell myeloma	70
#146	NAD	Abnormal	Abnormal	Complex		Plasma cell myeloma	18
#147	NAD	NAD	Abnormal	Hyperdiploid		Plasma cell myeloma	60
#106	Hyperdiploid	NAD	Hyperdiploid			Plasma cell myeloma	54
#51	NAD	NAD	Hyperdiploid	Hyperdiploid		Plasma cell myeloma	55
#149	Failed	Abnormal				Plasma cell myeloma	90
#138	NAD	NAD		Hyperdiploid		Plasma cell myeloma	75
#150	Hyperdiploid	NAD				Plasma cell myeloma	60
#88	NAD	NAD			Abnormal	Plasma cell myeloma	50
#151	NAD	NAD	Failed	Hyperdiploid		MGUS	7
#152	NAD	Abnormal				Asymptomatic Myeloma	10

#156	NAD	Abnormal				Plasma cell myeloma	30
#157	NAD	Abnormal		Abnormal		Plasma cell myeloma	70
#158	NAD	NAD		Complex		Plasma cell myeloma	25
#159	NAD	Abnormal	Failed	NAD		Plasma cell myeloma	20
#160	Missing Y	NAD				MGUS	7
#161		Abnormal	Failed		NAD	Plasma cell myeloma	40
#162	NAD	NAD			NAD	Plasma cell myeloma	35
#165	Complex	Abnormal	Abnormal			Plasma cell myeloma	70
#166	Failed	NAD				Plasma cell myeloma	60
#168	Hyperdiploid	Abnormal	Hyperdiploid			Plasma cell myeloma	70
#170	Failed	Abnormal	Abnormal			Plasma cell myeloma	80
#44	Complex	NAD	Failed		Abnormal	Plasma cell myeloma	56
#171	Failed	Abnormal		Hyperdiploid		Plasma cell myeloma	25
#175	Hyperdiploid	Abnormal	Hyperdiploid			Plasma cell myeloma	70
#176	NAD	NAD				Plasma cell myeloma	20
#178	Complex	Abnormal	Abnormal			Plasma Cell Leukaemia	95
#179	Failed	NAD		Hyperdiploid	NAD	Plasma cell myeloma	30
#180	Failed	NAD	Failed		NAD	Plasma cell myeloma	20
#184	Failed	NAD		Complex	NAD	Plasma cell myeloma	21
#187	NAD	Abnormal	Failed			Plasma cell myeloma	15
#188	NAD	Abnormal	Abnormal			MGUS	4
#186	NAD	NAD				Plasma cell myeloma	20

#190	NAD	Abnormal		NAD		Plasma cell myeloma	24
#189	Hyperdiploid	NAD	Failed	Failed		Plasma cell myeloma	59
#84	Complex	Abnormal				Relapsed myeloma	95
#192	Failed	NAD				Plasma cell myeloma	11
#191	NAD	Abnormal	Failed		Abnormal	Plasma cell myeloma	25
#196	NAD	Abnormal	Abnormal			Plasma cell myeloma	16
#197	NAD	Abnormal	Abnormal			Plasma cell myeloma	7
#198	NAD	Abnormal			Abnormal	Plasma cell myeloma	62
#200	Missing X	Abnormal			Abnormal	Plasma cell myeloma	50
#199	Failed	NAD		Failed	Abnormal	Plasma cell myeloma	60
#201	NAD	Abnormal		Complex		Plasma cell myeloma	60
#202	NAD	Abnormal		NAD	NAD	Plasma cell myeloma	20
#203	NAD	Abnormal	Hyperdiploid	Hyperdiploid		Plasma cell myeloma	70
#204	NAD	Abnormal		Hyperdiploid		Plasma cell myeloma	59
#205	NAD	NAD	Abnormal			Plasma cell myeloma	18
#207	NAD	NAD		Hyperdiploid		Plasma cell myeloma	73
#209	Failed	Abnormal		Hyperdiploid	NAD	Plasma cell myeloma	14
#208	NAD	Abnormal		Abnormal	Abnormal	Plasma cell myeloma	12
#210	Complex	Abnormal	Failed	Failed		Plasma cell myeloma	40
#213	Missing X	Abnormal				MGUS	8
#146 (2)	NAD	Abnormal				Plasma cell myeloma	12
#215	NAD	NAD		Failed	NAD	Plasma cell myeloma	36

#216	NAD	NAD		NAD		Plasma cell myeloma	50
#218	NAD	NAD	Failed			Relapsed myeloma	16
#219	NAD	NAD				Plasma cell myeloma	
#220	Failed	NAD				Plasma cell myeloma	15
#217	NAD	Abnormal	Failed		NAD	Plasma cell myeloma	24
#222	Failed	NAD	Hyperdiploid	Hyperdiploid		Plasma cell myeloma	52
#226	NAD	Abnormal		Hyperdiploid		Plasma cell myeloma	80
#225	NAD	Abnormal		Failed		Plasma cell myeloma	21
#224	Failed	NAD		Hyperdiploid		Plasma cell myeloma	80
#227	NAD	NAD				MGUS	8
#229	NAD	NAD				Asymptomatic Myeloma	15
#230	Failed	Abnormal		Hyperdiploid	Abnormal	Plasma cell myeloma	85
#231	NAD	NAD				MGUS	7
#232	NAD	Abnormal		NAD	NAD	Plasma cell myeloma	17
#233	NAD	Abnormal		Hyperdiploid	Abnormal	Plasma cell myeloma	45
#235	NAD	NAD			Abnormal	Plasma cell myeloma	42
#234	NAD	NAD		Hyperdiploid	Failed	Plasma cell myeloma	80
#236	NAD	Abnormal				MGUS	7
#237	Complex	NAD				Plasma cell myeloma	50
#238	NAD	NAD	NAD			Plasma cell myeloma	35
#239	Triploid	Abnormal				Plasma cell myeloma post treatment	6

#240	NAD	NAD				MGUS	5
#241	NAD	NAD				Plasma cell myeloma	55
#243	NAD	NAD	Hyperdiploid			Plasma cell myeloma	65
#244	NAD	Abnormal	Abnormal			MGUS	7
#61		Abnormal	Abnormal			Plasma cell myeloma	39
#245	NAD	Abnormal	Abnormal			Plasma cell myeloma	60
#247	NAD	NAD	Hyperdiploid			Plasma cell myeloma	28
#248	NAD	Abnormal	Abnormal			Plasma cell myeloma	20
#251		Abnormal	Abnormal			Plasma cell myeloma	65
#250		Abnormal	Failed			Plasma cell myeloma	26
#253		Abnormal				Plasma cell myeloma	45
#252		NAD				Plasma cell myeloma	30
#256		Abnormal	Failed			Plasma cell myeloma	90
#260		Abnormal				Plasma cell myeloma	37
#263		NAD	Failed			Plasma cell myeloma	30
#268		NAD				Plasma cell myeloma	18
TOTALS	91	101	45	36	24		

Table 3.2: Total number of samples processed as part of the research project cohort. This indicates the number of samples processed for each technology and the technologies applied to each patient in the cohort. Each test has a basic result recorded, NAD (no abnormality detected) is coloured green, failed tests are coloured purple, and any test with an abnormal result is coloured blue. The table also details the working diagnosis of the patient along with the percentage of plasma cells associated with the sample (where this figure is known).

Conventional Culturing and Karyotyping

Whole bone marrow samples from 91 patients were processed for long term culturing for cytogenetic analysis, and of these 91 patients, 78 yielded sufficient metaphases for a cytogenetic result (85.7% success rate). Cytogenetic analysis of plasma cell myeloma is notoriously difficult and relies of the ability to capture cells, specifically plasma cells, at metaphase of cell division. Plasma cells are mature cells and cycle slowly. Although robust processes are used routinely to analyse chromosomes in leukaemia samples, myeloma bone marrow samples do not yield good preparations for karyotyping using standard techniques. Adaptations to the culturing process were employed, allowing for longer term culturing (4-6 days), which are recommended for myeloma chromosome analysis[42]. The success rate was based on the presence of at least 10 metaphase cells for karyotype analysis. This compares to a success rate of 97.4% for standard karyotyping for leukaemia samples, using the 2018 figures from the SDGS routine service. This figure is known to be low for myeloma, and success rates of only 30-40% are reported[25,108].

Of the 78 bone marrow samples successfully cultured, 59/78 (75.6%) showed a normal karyotype, and 19/78 (24.4%) demonstrated an abnormal karyotype. These figures are presented in table 3.3.

	Total samples cultured	Successful karyotypes	Normal karyotypes	Abnormal karyotypes
Number of samples	91	78 (85.7%)	59 (75.6%)	19 (24.4%)

Table 3.3: Details the total number of samples cultured, those with successful karyotypes further divided into those with abnormal and normal karyotypes.

The rate of plasma cell capture is poor, even with enhanced culturing, and therefore the abnormality rate is low. Cytogenetic abnormalities are reported in the literature in approximately 30% of myeloma patients[42], using long term culturing. The low abnormality rate in our cohort (24.4% compared to the reported 30%) is likely to reflect the composition of our cohort, which is known to include a proportion of patients classified as MGUS, rather than myeloma. MGUS patients show a lower

level of abnormality, but also, by definition have <10% plasma cells, increasing the difficulty in capturing these cells in metaphase. The abnormality rate seen karyotypically in myeloma both in the literature and in our cohort is very much lower than those rates reported by FISH or MLPA technologies which do not rely on the ability to capture abnormal plasma cells in division.

Cytogenetic analysis requires the individual karyotyping of a number of cells, this time-consuming step contributes to the slow and expensive nature of karyotyping. Cytogenetic analysis is also highly dependent upon the skill and experience of the analyser. A standard diagnostic analysis for the majority of leukaemia types involves the analysis of 20 cells, but screening of 100 cells is recommended in order to use karyotyping as a diagnostic tool for myeloma[42].

The 19 abnormal karyotypes are listed in table 3.4 using the International System for Cytogenetic Nomenclature (ISCN)[109].

Patient No.	Karyotype
#113	44,X,-Y,i(1)(q10),add(2)(q31),add(3)(p1?3),del(6)(q21),add(8)(q24),-13, t(14;20)(q32;q11)[cp10]
#134	44~46,XY,add(1)(q42),?del(2)(q32),del(5)(q31),?add(6)(q21),del(7)(p12),-11, del(11)(q14.2),-13,del(17)(p12),add(17)(p13),der(21)t(1;21)(q21;p11), 2~7mar[cp10]
#136	43,X,-X,-13,der(1)t(1;19),-20,+mar,inc[cp2]/ 46,XX[2]
#144	45,X,-Y[3]/ 46,XY[8]
#106	55,XY,+del(3)(p21),+5,+7,+9,+9,+11,+15,-16,+19,+21,+22[2]/ 46,XY[10]
#150	53,XX,add(1)(p3?6),+3,+5,+7,-8,+9,+add(11)(p1),+?13,+15,+19,+19,-21[4]/ 46,XX[6]
#160	45,X,-Y[3]/ 46,XY[16]
#165	43~44,X,-Y,+add(1)(p1),add(3)(q27),add(8)(q24),+9,-13,+mar[cp4]/ 46,XY[1]
#168	59~60,XY,+add(1)(p1)x2,+2,+3,+6,+7,+9,+9,+add(11)(q23),add(14)(q32),+18,

	+18,+20[cp2]/ 46,XY[12]
#44	50~53,XX,+3,+add(3)(q27),+del(5)(q31),+7,+9,+10,-18,+22,+1~5mar[cp8]/ 46,XX[2]
#175	56,XY,+3,+5,+5,+7,+9,+9,+11,+15,der(17)t(1;17)(q11;p13),+19,+21[cp9]/ 46,XY[5]
#178	33~41,X,-X,-1,add(1)(p1),add(1)(q1),-2,-4,add(5)(q31),del(6)(q?21), der(7)add(7)(p22)add(7)(q36),add(8)(p1),-9,-10,-12,-13,-13,add(15)(q22),- 16,-16,-17,del(18)(p1),-21,-22,+2~7mar[cp9]/ 46,XX[1]
#189	54,X,-Y,?del(1p),+3,+5,-6,+9,-10,+11,-14,+15,-18,+19,+21,+6mar[1]/ 46,XY[38]
#84	50~54,XY,-1,+3,add(5)(q3?5),+7,der(8)t(8;22)(q24;q11),+9,+9,+11,+15,-16, add(17)(p1),+19,+19,+21,-22,+1~4mar[cp10]
#200	45,X,-X[7]/ 46,XX[3]
#210	41~46,XX,del(1)(p32),-3,-8,-13,+14,add(14)(q32)x2,-15,-16,add(19)(q1),-20, +r,+2~5mar,inc[cp6]/ 46,XX[5]
#234	45,X,-X[7]/ 46,XX[3]
#237	48~53,XX,+?add(3)(q21q21),+7,+add(9)(p1),?del(13)(q12q14),+15,-19,-20, +4~5mar[cp4]/ 46,XX[8]
#239	66,XXY,del(1)(p3?4),der(2)t(1;2)(p21;q21)x2,del(3)(p1),-4,+5,-6,-7,-9,+11, del(13)(q1)x2,-14,-16,-17,-19,-20,-21,-21,-21,-22,+10~12mar,inc[cp4]/ 46,XY[7]

Table 3.4: Details the 19 abnormal karyotypes detected in the 91 patients whose bone marrow samples were cultured and processed for karyotyping.

Two of the 19 abnormal cases have a missing chromosome Y karyotype, and 2/19 show a missing chromosome X cell line. Loss of a sex chromosome, specifically loss of chromosome Y is a known age related effect[110]. Although this abnormality has been associated with haematological disease, it is also seen in older patients with no known disorder, and therefore when seen as sole change is not considered consistent with neoplastic change. In essence, these should be considered normal with regard to the patient's myeloma.

The remaining 15 cases can be sub-divided into hyperdiploid cases 7/15 (46.6%) & non-hyperdiploid cases 8/15 (53.3%). This reflects the reported incidence of hyperdiploid vs non-hyperdiploid cases in the literature[25,26].

The hyperdiploid cases include #106, #150, #168, #44, #175, #189 and #84. Of the eight non-hyperdiploid cases, three showed a cytogenetically detectable rearrangement of chromosome 14 at the site of *IGH* (#113, #168 and #210). Only one of those three demonstrated a known and detectable partner chromosome, #113 in which a t(14;20) was detected. It should be noted that t(4;14) rearrangements are not detectable cytogenetically as the regions are small, approximately the same size and with similar banding patterns. The t(4;14) is one of the two most common *IGH* rearrangements, along with t(11;14), and the cytogenetically cryptic nature of this abnormality demonstrates the need to employ a different technology in order to detect the t(4;14).

Four cytogenetically abnormal cases have been selected to demonstrate the range of karyotypes seen. Figure 3.1 shows a representative karyotype from patient #113 demonstrating complex abnormalities including a t(14;20) *IGH-MAFB* rearrangement. Figure 3.2 shows a second complex karyotype from patient #134 who was shown to have a cytogenetically cryptic t(4;14), but also has abnormalities of chromosome 1q and loss of *TP53* on chromosome 17. Figure 3.3 shows a karyotype from patient #106 demonstrating a hyperdiploid karyotype with classic chromosome gains. Finally figure 3.4 shows a karyotype from patient #175, again with a hyperdiploid karyotype, but with a further abnormality of chromosomes 1 and 17 resulting in loss of *TP53* on the short arm of chromosome 17 and gain of the long arm of chromosome 1. Patients #113 and #175 will be discussed in further detail as part of the case reports section.

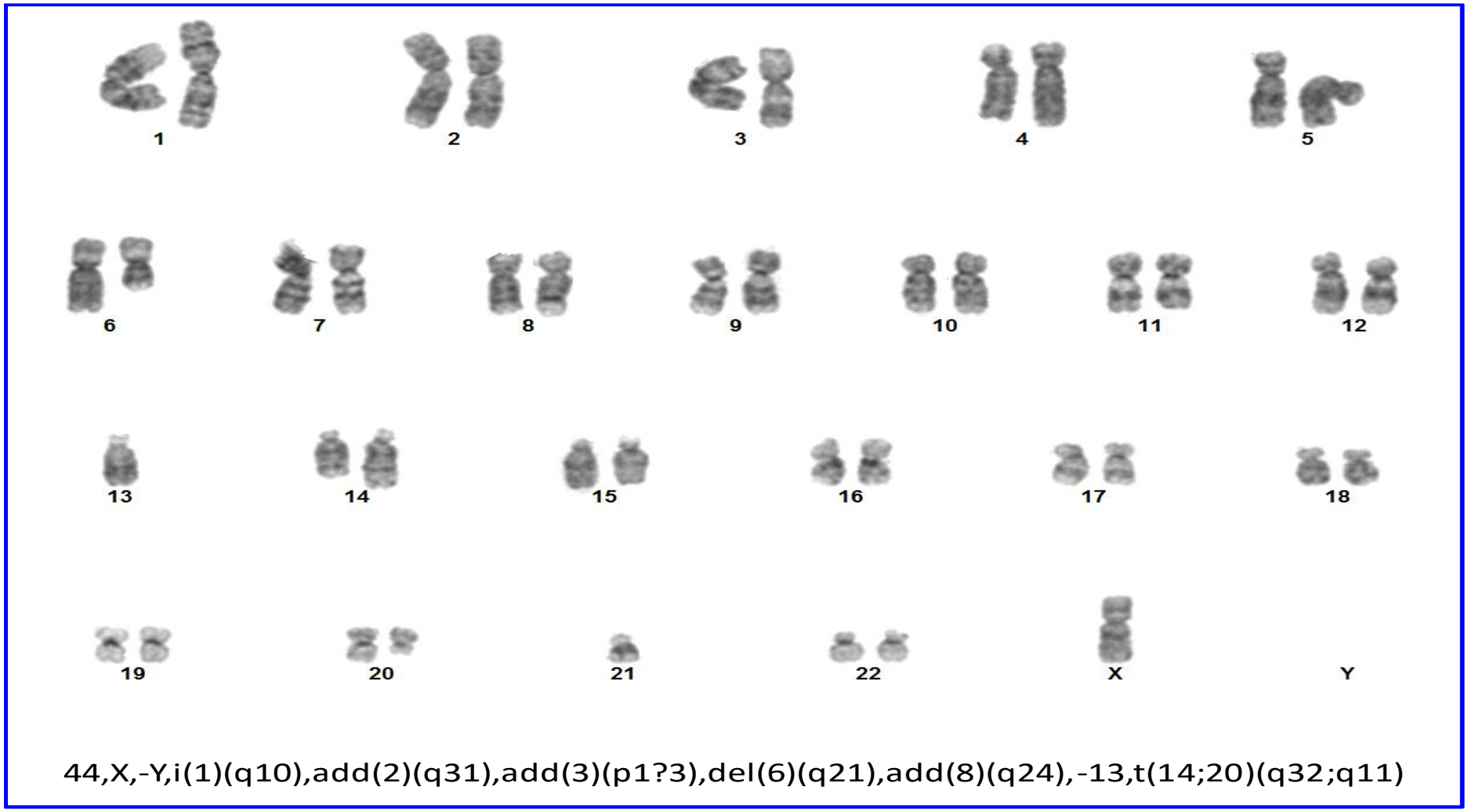


Figure 3.1: Demonstrating a representative karyotype from patient #113 with abnormalities including a t(14;20) *IGH-MAFB* rearrangement. This case is discussed further as a case report.

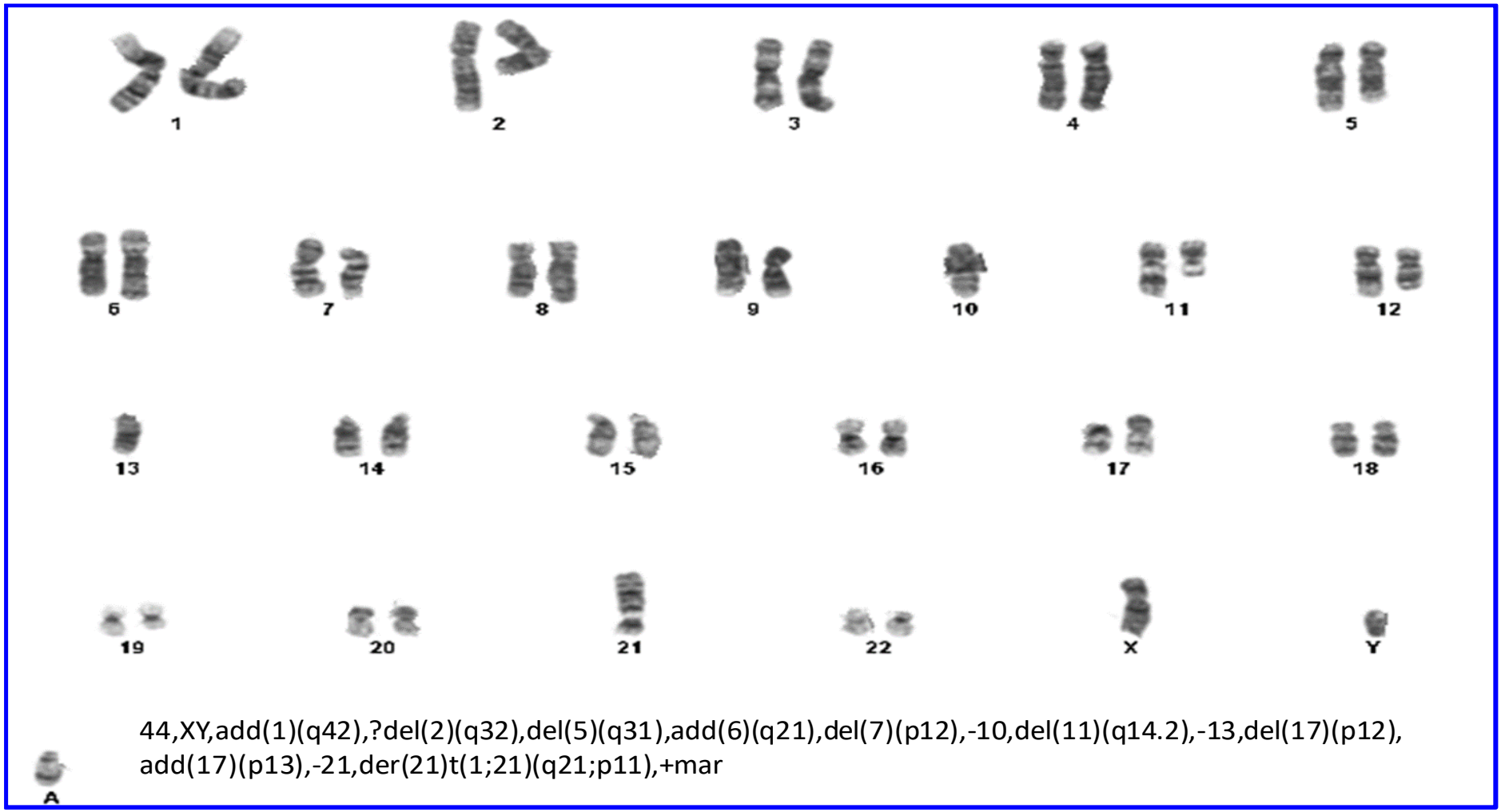


Figure 3.2: Demonstrating a representative karyotype from patient #134 with a complex karyotype including rearrangements of chromosomes 1, 2, 5, 6, 7, 11, 17, 21 and a marker chromosome. This patient also has a cytogenetically cryptic t(4;14) translocation. Loss of chromosomes 10, 13 and 21 are also seen.

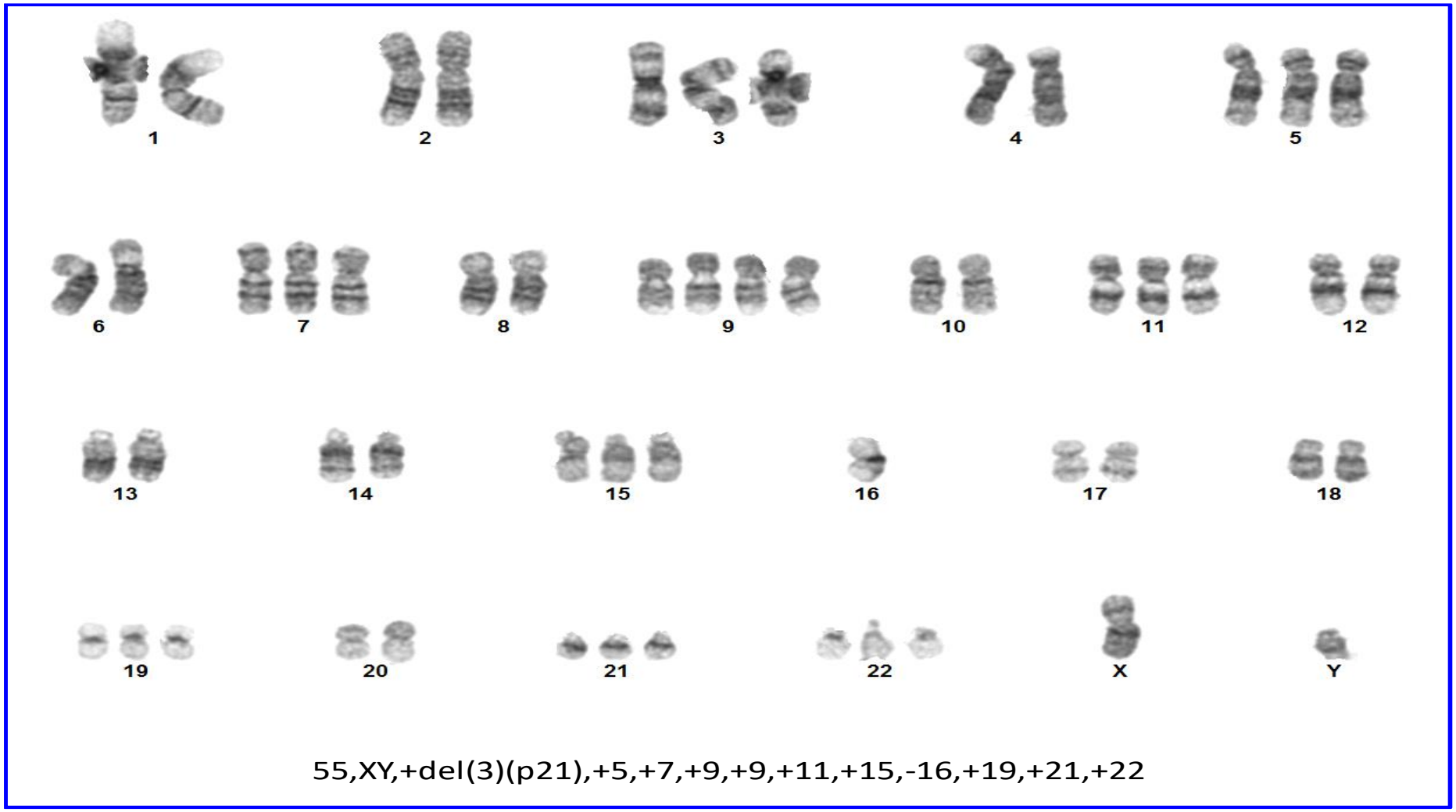


Figure 3.3: Demonstrating a representative karyotype from patient #106 with a hyperdiploid karyotype. This shows classic gains of chromosomes associated with myeloma, in this case, gains of chromosomes 3, 5, 7, 9, 11, 15, 19, 21, and 22. There is also a structural rearrangement of the short arm of chromosome 3 and missing chromosome 16.

Figure 3.4: Demonstrating a representative karyotype from patient #175 with a hyperdiploid karyotype, but with an additional abnormality of chromosomes 1 and 17 resulting in loss of *TP53* on the short arm of chromosome 17 and gain of the long arm of chromosome 1. This case is discussed further as a case report.

Cytogenetic analysis offers a full genome screen and can detect hyperdiploidy and abnormalities involving the *IGH* locus on chromosome 14, however, the resolution of cytogenetic analysis is low. The example karyotypes demonstrate the resolution level that can be obtained using cytogenetic analysis. A fine chromosome band is considered to be around 5Mb, and the quality of chromosomes and banding seen in leukaemia preparations make these fine bands difficult to assess. Therefore, changes involving approximately 10Mb are likely to be the best level of detection for chromosome analysis in myeloma.

Fluorescence in situ Hybridisation (FISH)

FISH was performed on 101 patients with a CD138 positive separated fraction. As described in the methods, FISH was undertaken as a two-step process as part of this project.

Step 1 involves the Abbott Vysis *IGH* breakapart probe, Cytocell *TP53* (including a control region at the centromere of chromosome 17) and Cytocell *CDKN2C/CKS1B* probes. Probe maps for each of these three probes are shown in the materials and methods chapter, figure 2.1. The *IGH* breakapart probe is designed with a red and a green probe flanking the *IGH* region. In a normal cell with no *IGH* rearrangement this would be seen as two fusion (2F) signals, fusion signals describe visualisation of red and green at the same region which is often seen as yellow. When a rearrangement of *IGH* is seen, there is a break between the red and the green probes and these are then visualised as separate signals, a standard rearrangement pattern would be one fusion signal, one red signal and one green signal (1F1R1G). The *TP53*/17cen probe is designed to have a red probe covering the critical region of *TP53* and a green probe at the centromere of chromosome 17. The control probe provides reassurance that the probe is working correctly, but also allows the number of chromosome 17s to be enumerated. A normal cell would show two red signals and two green signals (2R2G). Deletion of *TP53* would usually be seen as loss of one of the red signals (1R2G), but relative deletions can also be seen with three copies of chromosome 17, demonstrated by the green centromere probe, and loss of one of the *TP53* signals

leaving two (2R3G). Occasionally, trisomy for chromosome 17 is seen, showing three red signals and three green signals (3R3G). The *CDKN2C/CKS1B* probe is designed to have a green probe covering the critical region of *CDKN2C* at 1p21.3 and a red probe covering the critical region of *CKS1B* at 1q21.3. A normal cell would show two red signals and two green signals (2R2G). The expected abnormal pattern would be either loss of 1p (2R1G) or gain of 1q (3R2G), or these can be seen together (3R1G). Sometimes further additional copies of 1q are gained (e.g. 4R1G), and amplification of *CKS1B* is also seen with five or more copies of the probe (e.g. 5R2G).

Step 2 is initiated if an *IGH* rearrangement is detected at step 1 using the *IGH* breakapart probe. The aim of step 2 is to elucidate the *IGH* partner gene and the FISH panel used here includes *IGH/CCND1, IGH/FGFR3* and *IGH/MAF* probes. This covers the two most common rearrangements *IGH-CCND1* and *IGH-FGRF3* each seen in approximately 15% of patients, and *IGH-MAF* which is the next most common poor prognostic translocation seen in approximately 5% of patients. NICE guidelines recommend, as a minimum, the exclusion of *IGH-FGFR3* and *IGH-MAF* as the two most common poor risk rearrangements, along with *TP53* deletion status[111]. These three probe sets used in step 2 are described as 'dual colour, dual fusion'. Each probe set has a green probe covering the *IGH* gene and a red probe covering the potential *IGH* partner gene; *FGFR3, CCND1* or *MAF*. In a normal cell, the probes would demonstrate two green signals on each of the normal *IGH* genes, and two red signals on each of the normal *IGH* partner genes (2R2G). In a cell with an *IGH* rearrangement, for example *IGH-CCND1*, both the abnormal *IGH* and *CCND1* split to give three signals, and two of each fuse giving a 2F1R1G pattern. The fusion signals represent the derivative chromosomes 11 and 14 in this case, and the single red signal marks the normal *CCND1* gene and the single green probe marks the normal *IGH* gene. Myeloma genetics can be extremely complex and non-standard rearrangements can be seen along with loss and gain, FISH probes can show non-standard signal patterns reflecting these complex rearrangements.

Chapter 5 describes the introduction of a quality control assessment scheme for myeloma genetic analysis, and demonstrates a number of alternative FISH panels utilised by different laboratories. We have described a two-step FISH process here,

which covered the NICE guideline requirements at the time as well as the 1p/1q assessment. *IGH-CCND1* is seen in approximately 15% of myeloma patients. *IGH-MAFB* is a marker of poor prognosis, but seen at a very low frequency, 1-2% of myeloma patients. I*GH-CCND1* was included in this panel based on the frequency of occurrence, but currently *IGH-MAFB* is not included although there is a strong argument for inclusion based on the prognosis attached to this rearrangement.

At this stage it should be noted that initially in this project, FISH was performed on whole BM samples whilst the CD138+ve selection process was undergoing validation. Of the first 10 patient samples that had FISH processed on whole BM samples an abnormality rate of 0% was seen. This compares to a 54.4% abnormality rate in cases with a preceding plasma cell separation. Plasma cells can be present at levels as low as 10% in myeloma patients and less in MGUS patients, if the abnormality is only present in a proportion of those plasma cells, for example 10%, then only one cell in a standard analysis of 100 cells would show abnormality. This demonstrates the importance of ensuring that a separation step is completed and offers good evidence for why this step is considered best practice in the analysis of myeloma bone marrow samples. FISH results from the whole bone marrow samples were excluded from this study and the FISH results presented here are solely from the patients FISHed following plasma cell separation.

FISH is generally a robust technique and no cases failed using all FISH probes. Two cases failed to give a result for the *IGH* breakapart probe, but all cases gave a result with the *TP53* and *CDKN2C/CKS1B* probes. Considering numbers of FISH hybridisation performed in this project, 301/303 (99.3%) achieved a result, demonstrating an excellent success rate.

Table 3.5 details the FISH results for all probes used across the 101 patients. Boyd *et al*[112], describe a worsening prognosis based on the presence of three poor risk abnormalities; poor risk *IGH* rearrangements (*IGH-FGFR3, IGH-MAF and IGH-MAFB*), *TP53* deletion and gain of chromosome 1q. Patients in our cohort have been scored based on the number of these adverse risk abnormalities described in Boyd's prognostic schema, and also scored for the number of abnormalities detected using the FISH panel presented here (see table 3.5). Where these numbers are shown in

blue, the *IGH* partner gene is not known, and definitive proof of a poor risk *IGH* rearrangement is unknown. Table 3.6 presents the abnormalities seen for each of the probe sets again across the 101 patients tested.

Sample No:	TP53	CDKN2C/CKS1B	IGH	*IGH* Rearrangement Partner	Further information	No. of adverse abnormalities	Total no. of abnormalities
#113	Normal	Loss of 1p & gain of 1q	Rearranged	*IGH-MAFB* rearrangement	*MYC* rearrangement	2	4
#119	Normal	Not enough material to complete	Normal			0	0
#134	Deleted	Normal 1p, gain of 1q	Rearranged	*IGH-FGFR3* rearrangement		3	3
#136	Normal	Normal 1p, gain of 1q	Rearranged	*IGH-MAF* rearrangement		2	2
#138	Normal	Normal 1p, gain of 1q	Normal			1	1
#139	Normal	Normal 1p, gain of 1q	Normal			1	1
#141	Normal	Normal	Rearranged	*IGH-CCND1* rearrangement		0	1
#140	Normal	Normal	Rearranged	*IGH-CCND1* rearrangement		0	1
#142	Normal	Not enough material to complete	Normal			0	0
#144	Normal	Not enough material to complete	Normal			0	0
#146	Normal	Normal 1p, gain of 1q	Rearranged	*IGH-CCND1* rearrangement		1	2
#147	Normal	Normal	Normal			0	0
#106	Normal	Normal	Normal			0	0
#51	Normal	Not enough material to complete	Normal			0	0
#149	Deleted	Normal	Rearranged	No partner detected using *FGFR3*, *CCND1* or *MAF*		1	2
#138	Normal	Normal	Normal			0	0

#150	Normal	Normal	Normal			0	0
#88	Normal	Normal	Normal			0	0
#151	Gain	Gain of 1p & 1q	Normal		Extra copies of chromosome 1 & 17	0	0
#152	Normal	Normal	Rearranged	*IGH-CCND1* rearrangement		0	1
#156	Normal	Loss of 1p & normal 1q	Normal			0	1
#157	Normal	Normal	Rearranged	*IGH-CCND1* rearrangement		0	1
#158	Normal	Normal	Failed			0	0
#159	Deleted	Normal	Normal			1	1
#160	Normal	Normal	Normal			0	0
#161	Abnormal	Normal	Rearranged	No partner detected using *FGFR3*, *CCND1* or *MAF*	Loss of 17 centromere	1	2
#162	Normal	Normal	Normal			0	0
#165	Normal	Normal 1p, gain of 1q	Normal		*MYC* rearrangement	1	2
#166	Normal	Normal	Normal			0	0
#168	Gain	Normal 1p, gain of 1q	Rearranged	No partner detected using *FGFR3*, *CCND1* or *MAF*	Extra copies of chromosome 17	1	3
#170	Normal	Normal 1p, gain of 1q	Rearranged	No partner detected using *FGFR3*, *CCND1* or *MAF*		1	2
#44	Normal	Normal	Normal			0	0
#171	Deleted	Normal 1p, gain of 1q	Normal			2	2
#175	Deleted	Normal 1p, gain of 1q	Normal			2	2

#176	Normal	Normal	Normal			0	0
#178	Deleted	Loss of 1p & gain of 1q	Rearranged	*IGH-FGFR3* rearrangement, with loss of *FGFR3*		3	4
#179	Normal	Normal	Normal			0	0
#180	Normal	Normal	Normal			0	0
#184	Normal	Normal	Normal			0	0
#187	Normal	Normal	Rearranged	*IGH-FGFR3* rearrangement		1	1
#188	Normal	Normal 1p, gain of 1q	Rearranged	No partner detected using *FGFR3*, *CCND1* or *MAF* on whole BM only		1	2
#186	Normal	Normal	Normal			0	0
#190	Normal	Normal	Rearranged	*IGH-CCND1* rearrangement		0	1
#189	Normal	Normal	Normal			0	0
#84	Relative deletion	Loss of 1p & gain of 1q	Normal		*MYC* rearrangement	1	4
#192	Normal	Normal	Normal			0	0
#191	Normal	Normal	Rearranged	No partner detected using *FGFR3*, *CCND1* or *MAF* on whole BM only		0	1
#196	Normal	Normal 1p, gain of 1q	Normal		Loss of one *IGH* signal	1	2
#197	Normal	Normal	Rearranged	No partner detected using *FGFR3*, *CCND1* or *MAF* on whole BM only		0	1
#198	Normal	Normal 1p, gain of 1q	Normal		Loss of one *IGH* signal	1	2
#200	Normal	Normal 1p, gain of 1q	Rearranged	No partner detected using *FGFR3*, *CCND1* or *MAF*		1	2
#199	Normal	Normal	Normal			0	0

#201	Normal	Normal	Normal		Loss of one *IGH* signal	0	1
#202	Normal	Loss of 1p & normal 1q	Failed			0	1
#203	Relative deletion	Normal	Normal			0	1
#204	Normal	Loss of 1p & gain of 1q	Normal			1	2
#205	Normal	Normal	Normal			0	0
#207	Normal	Normal	Normal			0	0
#209	Gain	Normal	Normal		Extra copies of chromosome 17	0	1
#208	Gain	Normal 1p, gain of 1q	Rearranged	*IGH-MAF* rearrangement	Extra copies of chromosome 17	2	3
#210	Deleted	Normal	Normal			1	1
#213	Normal	Normal	Rearranged	No further material available to assess *IGH* partner genes		0	1
#146 (2)	Normal	Normal 1p, gain of 1q	Rearranged	No further material available to assess *IGH* partner genes		1	2
#215	Normal	Normal	Normal			0	0
#216	Normal	Normal	Normal			0	0
#218	Normal	Normal	Normal			0	0
#219	Normal	Normal	Normal			0	0
#220	Normal	Normal	Normal			0	0
#217	Normal	Loss of 1p & normal 1q	Normal			0	1
#222	Normal	Normal	Normal			0	0

#226	Normal	Normal	Rearranged	*IGH-CCND1* rearrangement	0	1
#225	Normal	Normal 1p, gain of 1q	Normal		1	1
#224	Normal	Normal	Normal		0	0
#227	Normal	Normal	Normal		0	0
#229	Normal	Normal	Normal		0	0
#230	Normal	Normal	Rearranged	No partner detected using *FGFR3*, *CCND1* or *MAF*	0	1
#231	Normal	Normal	Normal		0	0
#232	Normal	Normal 1p, gain of 1q	Rearranged	No partner detected using *FGFR3*, *CCND1* or *MAF* on whole BM only	1	2
#233	Normal	Normal 1p, gain of 1q	Normal		1	1
#235	Normal	Normal	Normal		0	0
#234	Normal	Normal	Normal		0	0
#236	Normal	Normal	Rearranged	No partner detected using *FGFR3*, *CCND1* or *MAF* on whole BM only	0	1
#237	Normal	Normal	Normal		0	0
#238	Normal	Normal	Normal		0	0
#239	Normal	Normal 1p, gain of 1q	Normal		1	1
#240	Normal	Normal	Normal		0	0
#241	Normal	Normal	Normal		0	0
#243	Normal	Normal	Normal		0	0

#244	Normal	Normal 1p, gain of 1q	Normal		Additional copies of 1, 14 & 17	1	3
#61	Normal	Normal	Rearranged	*IGH-FGFR3* rearranged		1	1
#245	Normal	Normal	Rearranged	*IGH-FGFR3* rearrangement		1	1
#247	Normal	Normal	Normal			0	0
#248	Normal	Normal 1p, gain of 1q	Rearranged	No partner detected using *FGFR3, CCND1* or *MAF* on whole BM only		1	2
#251	Normal	Normal	Rearranged	*IGH-CCND1* rearrangement		0	1
#250	Normal	Normal 1p, gain of 1q	Rearranged	No partner detected using *FGFR3, CCND1* or *MAF* on whole BM only		1	2
#253	Normal	Normal	Rearranged	No further material available to assess *IGH* partner genes		0	1
#252	Normal	Normal	Normal			0	0
#256	Normal	Normal 1p, gain of 1q	Normal			1	1
#260	Normal	Normal 1p, gain of 1q	Rearranged	*IGH-FGFR3* rearranged		2	2
#263	Normal	Normal	Normal			0	0
#268	Normal	Normal	Normal			0	0

Table 3.5: Details the results seen from the FISH performed on the project cohort. The first three results columns shows the results from FISH step 1*; IGH, CDKN2C/CKS1B* (1p/1q) and *TP53*. Column four and five details the results of FISH step 2 if completed, ie information on the *IGH* partner gene and any other comments on the FISH performed. The final two columns record the total number of abnormalities seen, and the number of adverse abnormalities seen based on Boyd *et als* model. Blue font shows those where the full results aren't available, usually the partner gene information for *IGH*, and therefore the lowest number of abnormalities are scored.

TP53	**Normal**	**Failed**	**Deleted**	**Gain**	**Relative deletion**	**Abnormal**				**Notes**
No.	87	0	7	4	2	1				'Abnormal' case showed loss of the 17 centromere
IGH	**Normal**	**Failed**	**Rearranged**	***FGFR3***	***CCND1***	***MAF***	***MAFB***	**Other**	**Loss**	
No.	66	2	33	6	8	2	1	16	3	'Other' includes those where *IGH* partner not found using standard 3 probes (3 had no more material, 7 were completed on whole BM). Cases of loss were not associated with rearrangement.
1p/1q	**Normal**	**Failed**	**1q gain**	**1p del**	**1q gain & 1p del**	**Other**	**Not enough material**			
No.	64	0	25	3	4	1	4			'Other' case was both 1p & 1q gain (trisomy 1)
MYC	**Normal**	**Failed**	**Rearranged**							
No.	N/A	0	3							Rearranged in three cases, but only completed when cytogenetic analysis suggested abnormality

Table 3.6: Demonstrating the abnormalities detected by FISH by category; *TP53, IGH, CDKN2C/CKS1B* (1p/1q) and *MYC.*

Of the 101 cases FISHed, 46 patients (45.5%) of patients showed no abnormality with this panel of probes, and 55 patients (54.4%) demonstrated abnormality. Based on the literature, 60% is the approximate FISH abnormality rate expected utilising the probe panel described here. An abnormality rate of 90% has been reported in myeloma samples where extended FISH panels are employed[42], particularly those including hyperdiploidy assessment.

Fifty-five patients (54.4%) showed one or more abnormalities. Within our cohort, 27/55 abnormal patients showed a single poor risk abnormality, 6/55 patients showed two poor risk abnormalities and 2/55 patients showed three poor risk abnormalities according to the Boyd *et al* prognostic schema[112]. The presence of this set of abnormalities confers an accumulative risk in terms of prognosis. Patients harbouring two or three poor risk abnormalities are referred to as 'double' or 'triple hit' myeloma respectively. As an overall cohort, we have the full data to complete this assessment on 79 patients, based on this 6.2% and 2.1% of patients could be considered double hit and triple hit myeloma respectively[112]. Currently this assessment of prognosis is not used in treatment decision making in the UK, although it may be that this becomes increasingly utilised with time and as more trials data corroborate these findings.

Step 1 only was performed on 68/101 (67.3%) cases. Sixty-six cases showed no evidence of an *IGH* rearrangement, and two cases failed to give a result with the *IGH* breakapart probe. Thirty-three (32.7%) cases showed an *IGH* rearrangement and were FISHed according to step 2 to elucidate the partner gene for *IGH* where possible. *IGH* rearrangements are reportedly seen in approximately 60% (range 55-70%) of myeloma patients[22,33,34]. *IGH* rearrangement is lower in our cohort than the expected rate, although this could be explained by our cohort including patients with a range of plasma cell neoplasms, not simply myeloma. Of note a similar study by Rack *et al*[113], comparing cytogenetics, FISH and array reported *IGH* rearrangement in 34% of their cohort very similar to abnormality rate detected here. Of the 33 rearranged cases, a partner gene was demonstrated in 17 (51.5%) cases, eight cases (24.2%) with *IGH-CCND1*, six cases (18.2%) with *IGH-FGFR3* and two cases (6%) with *IGH-MAF.* It should be noted that *MAFB* is not part of the standard step 2 FISH panel,

but was completed following detection of a t(14;20) by cytogenetics and therefore *IGH-MAFB* was detected in a single case. Sixteen cases did not demonstrate an *IGH* partner gene using the probe sets employed. Based on reported data, it would be expected that partner genes would be detected in more cases, however, three patients had no remaining material to test, and seven cases had step 2 completed on the whole bone marrow as insufficient CD138+ cell separated cells were available. FISH on whole bone marrow in myeloma has been demonstrated to be a very much less sensitive test, and therefore partner gene detection may have failed based on the test sensitivity in these cases. The relative levels of each abnormality are approximately in line with the expected frequencies with *IGH-FGFR3* and *IGH-CCND1* being the most common rearrangements, and *IGH-MAF* being very much rarer. Loss of *IGH* was seen in 3/101 (3.0%) cases, and these cases were not associated with rearrangement of *IGH*.

Figure 3.5 shows the probe map for the Abbott Vysis *IGH/MAFB* probe, (the *IGH* region on chromosome 14 labelled in green in figure 3.5a and the *MAFB* region on chromosome 20 labelled in red in figure 3.5c) used to detect the *IGH-MAFB* rearrangement in patient #113. Figure 3.5b shows the 2F1R1G signal pattern in an interphase cell whilst figure 3.5d shows the same rearrangement pattern in a metaphase image.

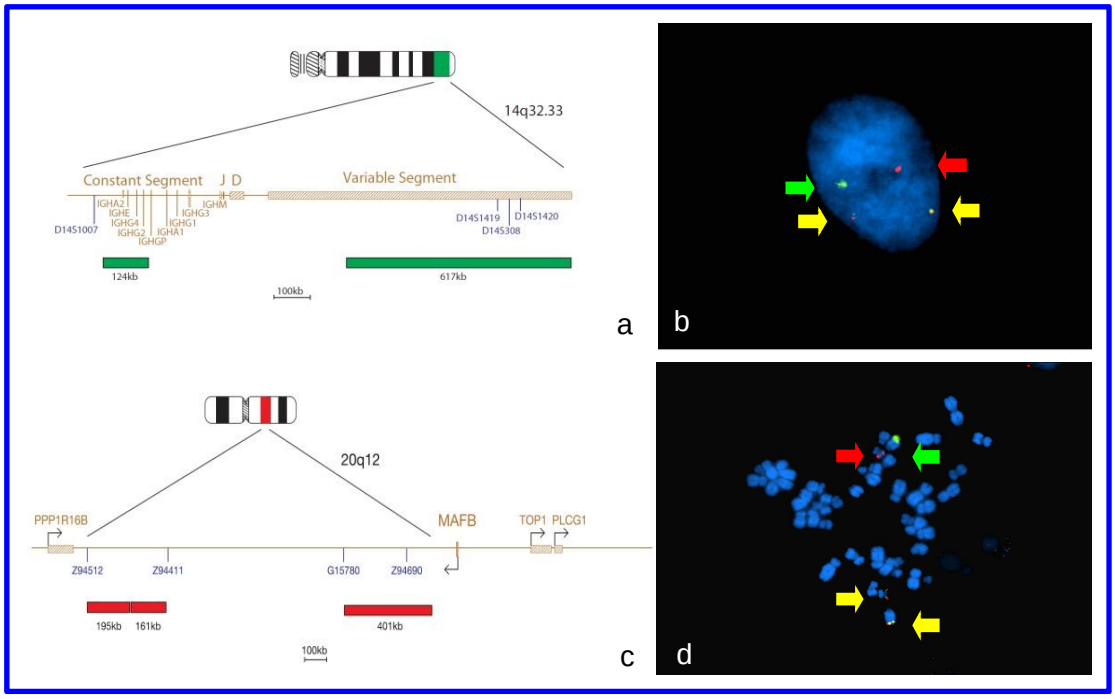


Figure 3.5: To demonstrate an example of an *IGH* rearrangement detected by FISH. The schematic demonstrates the Vysis *IGH/MAFB* dual colour dual fusion probe map covering *IGH* (a) and *MAFB* (c), and shows both an interphase (b) and metaphase (d) cell with the two fusion (yellow arrows), one red, one green (2F1R1G) pattern expected in a rearranged case.

TP53 deletion was seen in 7/101 (6.9%) cases, fitting well with the level reported in the literature of approximately 7-10%[22,25,37]. Relative deletion of *TP53* is seen as a 2R3G signal pattern and was seen in 2/101 cases (2.0%). Relative deletion describes those cases where there is evidence of three copies of chromosome 17 as assessed by the centromere probe, with loss of *TP53* relative to the baseline of three. This still leaves two functional copies of *TP53*. It is not totally clear from the literature how these patients should be classified, i.e. are they *TP53* deleted or not? However, it seems likely that as two functional copies of *TP53* remain, these patients are not likely to respond in the same way as *TP53* deleted patients. Gain of *TP53* was seen in four cases (4.0%), and a single case showed an abnormal pattern 2R1G, suggesting loss of the 17 centromere, but with two normal copies of *TP53*.

The *CDKN2C/CKS1B* (1p/1q) probe was not completed on four cases where insufficient material was available. In this situation, *IGH* and *TP53* FISH were prioritised based on the original NICE guidelines which include these two probes as a minimum diagnostic set. Of the remaining 97 cases FISHed with 1p/1q, 33/97

(34.0%) cases demonstrated an abnormality of chromosome 1. Of these 33 cases, 25 (25.8% of total cases FISHed) showed gain of chromosome 1q, three cases (3.1%) showed loss of chromosome 1p, and a further four cases (4.1%) demonstrated both gain of 1q and loss of 1p. Gain of 1q is reported to be seen in 40% of myeloma patients at diagnosis[25]. The figures seen in our cohort are lower, again likely to be related to the composition of our cohort as previously discussed, however, they do proportionally reflect those reported in the literature[22]. A common manifestation of concurrent del1p/dup1q is the isochromosome 1q in which the short arm is lost and the long arm is duplicated and mirrored around the centromere. Although we do not have definitive evidence that this is the mechanism seen in all four cases in the cohort, cytogenetic analysis in one patient has confirmed the presence of an isochromosome 1q. A single case showed both gain of both 1p and 1q and this would suggest trisomy for chromosome 1. Gain of 1q and loss of 1p have both been associated with a poor prognosis[55,57], but are not recognised by all myeloma risk stratifications[114-116]. Boyd *et al* include gain of 1q in their prognostic schema and based on this we have included *CKS1B* (1q) in the FISH panel.

Figure 3.6 shows examples of *CDKN2C*/*CKS1B* and *TP53* probe signals seen in the patient cohort. The first three images (a-c) show the *CDKN2C*/*CKS1B* 1p(green)/1q(red) probe; image a shows gain of 1q, image b shows loss of 1p, and image c shows concurrent loss and gain. Images d-f show the *TP53*(red)/17centromere(green) probe; image d shows a normal pattern of two of each signal, image e shows loss of *TP53* and image f shows loss of *TP53*, but on a background of three copies of chromosome 17 centromere.

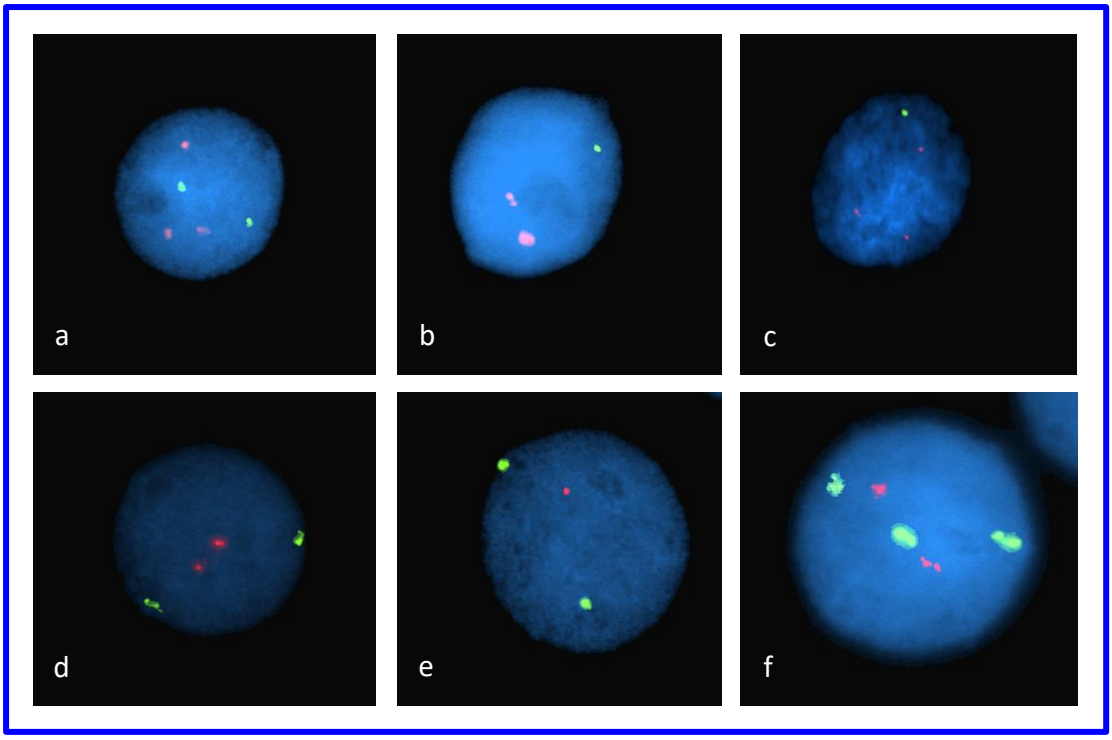


Figure 3.6: To demonstrate examples of probe signals seen in the patient cohort. Images a, b & c show the *CDKN2C* (1p in green) and *CKS1B* (1q in red). Image a shows gain of 1q (*CKS1B*) 3R2G, image b shows loss of 1p (*CDKN2C*) 2R1G and image c shows concurrent loss of 1p and gain of 1q 3R1G. Images d, e & f shown the TP53 (red) and 17 centromere (green) probe. Image d shows a normal pattern with 2R2G, e shows a deletion of TP53 1R2G and image f shows a relative deletion of *TP53*, loss of a red signal against a background of three copies of 17 centromere, 2R3G.

Finally, three cases were FISHed with a *MYC* breakapart probe; all of these showed rearrangement. This FISH test was only undertaken when the karyotype analysis was abnormal, and where chromosome 8 demonstrated abnormality in the region of *MYC* at 8q24. *MYC* translocations are seen in approximately 15% of patients with myeloma at presentation[25,47,48] and up to 45% of patients with advanced disease [25,26]. These translocations are not thought to be initiating events, but late events associated with increased proliferation and stromal independent plasma cells[25]. Recent studies have demonstrated that *MYC* is able to recruit active super-enhancers from highly expressed genes associated with B cell, plasma cell or myeloma development[49]. The current literature suggests that *MYC* rearrangement is highly prognostic[49], this may simply be related to the association with advanced disease, but also provides evidence for including *MYC* in a diagnostic myeloma FISH panel.

Of the 55 cases with FISH abnormalities, 44 (80.0%) were successfully karyotyped. Seven patients (12.7%) were not processed for cytogenetic analysis, four patients (7.3%) failed their cytogenetic analysis. Of those karyotyped, 11 patients (20.0%) also demonstrated abnormal cytogenetics, and 33 samples (60.0%) were cytogenetically normal. This provides confirmation that normal non-neoplastic populations had been karyotyped in these cases, in turn providing further evidence that karyotyping myeloma patients is not a good analytical tool in the vast majority (approaching 80% in our cohort) of cases.

Conversely, of the 19 patients with a karyotypic abnormality, eight patients did not demonstrate abnormality with this FISH panel. Three of the cases showed missing sex chromosomes (two with missing Y and one with missing X), and the remaining five cases would be considered hyperdiploid. Even though one case demonstrated complex abnormalities alongside the hyperdiploidy, the abnormalities would not have been detected on the FISH panel employed here. This highlights the limitations of the current FISH panel and demonstrates the need to extend this panel, either with additional probes or with additional techniques to increase the detection rate further. Hyperdiploidy can be confidently assessed using a 3-chromosome combination of FISH probes (chromosomes 5, 9, 15). The presence of two of the three chromosomes is a highly specific indicator of hyperdiploidy[117]. Of note hyperdiploidy would also be detected on an MLPA or array analysis. Within this particular cohort, none of these five patients had MLPA or array analysis and therefore we have not been able to confirm hyperdiploidy in these cases using these alternative technologies.

Myeloma genetics can be extremely complex and non-standard rearrangements can be seen along with loss and gain. This complexity is also represented in the FISH probes which can generate non-standard signal patterns, and reflects a strong need to ensure robust training of personnel involved in this analysis to ensure that non-standard patterns are understood and correctly interpreted.

Multiplex Ligation-dependant Probe Amplification (MLPA)

Forty-five cases were processed for MLPA using the DNA extracted from the CD138+ve plasma cell fraction. Data were transferred from GeneMapper to the Coffalyser analytical software supplied by MRC Holland. Coffalyser involves a series of steps to assess the QC of the controls and samples to ensure robustness of data. For each run the control samples require assessment, and any controls showing deletion or gain should be removed from the comparative analysis and a minimum of four controls should be taken forward to the final analysis. For each sample, the DNA Denaturation (DD) fragments and the four quality fragments (Q-fragments) require assessment before the individual samples can be deemed of sufficient quality for analysis. The DD fragments at 88 and 96 nucleotide (nt) length, detect fragments in exceptionally high GC regions, these regions can be difficult to denature and when these fragments appear low it indicates incomplete denaturation which in turn can result in false positive deletions. The Q-fragments at 64, 70, 76 and 82 nt in length are complete fragments which do not require ligation in order to be amplified during the PCR step, high level Q-fragments indicate poor quality or low quantity DNA. These quality parameters are assessed as part of the FMRS (Fragment MLPA Reaction Score) bars in the software, and only these with ≥2 bars should be processed. Those with 0 or 1 bar should be considered a failed sample. Other samples which pass this initial QC step, can still fail on the analysis, some samples show such wide standard deviation bars that it cannot be considered robust enough data to report on, so a further set of samples fail at the final processing and analysis point.

The 'comparative analysis' step involves a complex algorithm which assesses the relative size of the fluorescent peaks within each sample and compares and normalises them against the reference samples and the other samples within the run. The results were expressed as ratios (also known as dosage quotients) for the likelihood of a probe being either deleted or duplicated. A score of 1.0 represents two copies of a probe and was considered normal, scores of 0.5 (0.35-0.65) and 1.5 (1.35-1.65) represent deletion and duplication of a probe respectively, and scores of 0 (<0.15) and 2 (>1.85) denote no copies or four or more copies respectively. The

Coffalyser software demonstrates the data in a number of ways; the data are presented as a table with each probe alongside the height, area, ratio and standard deviation figures. The ratios are highlighted in blue (gain) or red (loss) when an abnormal ratio is detected. These results are also available as electropherograms detailing the peak heights for each probe. The DD- and Q-fragments can also be assessed (left hand side of this figure). Finally, the results are also represented graphically with a ratio of one used as the normalised line with standard deviation bars, regions of gain appear as blue data points above the standard deviation line and regions of loss are seen as red data points below the standard deviation line. These ratio graphs demonstrate the losses and gains pictorially making the assessment relatively simple at this level. Examples of both the electropherogram and the ratio graphs are shown in figure 3.7 and 3.8 respectively, for patient #238 with a normal MLPA result, but for further results the ratio graphs will be shown.

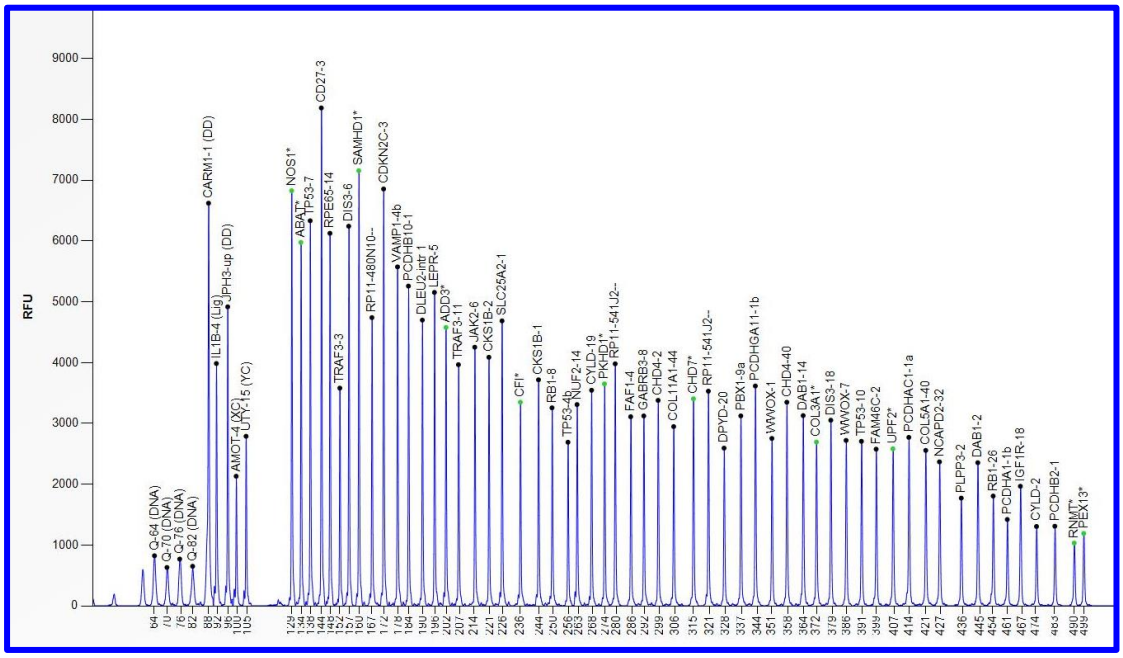


Figure 3.7: To show the electropherograms produced as part of the MLPA analysis by the Coffalyser software.

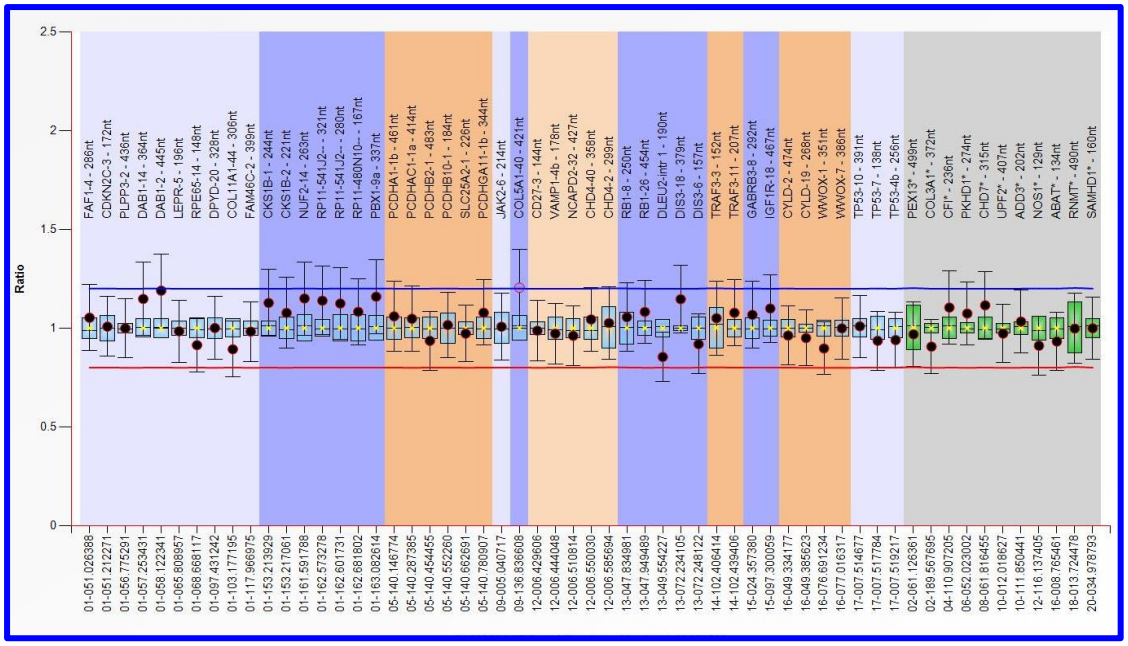


Figure 3.8: To show the ratio graphs produced as part of the MLPA analysis by the Coffalyser software.

Of the 45 cases processed for MLPA analysis, 16 cases (35.6%) failed to give a result. This is a high level of failure. In part, this may be due to poor quality DNA or lack of experience in using this technique, however, the MLPA process can be highly sensitive to contaminants in the system that can affect the PCR, more so than standard PCR. Following discussions with a laboratory that use this technique routinely for myeloma analysis, they too report similar levels of failure and indeed now set strict thresholds for processing lower quality samples. MLPA is still considered a useful and accurate test on samples with high quality DNA, but routine practice includes an option to FISH samples considered too poor for MLPA.

If MLPA were to be included as part of a diagnostic testing strategy, then contingency testing would have to be considered for failed samples. This is likely to be additional FISH hybridisations for the regions considered prognostic (i.e. *TP53*, 1p/1q, and possibly the three chromosomes associated with hyperdiploidy assessment). This adds further steps to a testing strategy and increases the expense.

Of the remaining 29 cases, 27 cases (93.1%) showed abnormal results (see table 3.7). The table details those cases with successful MLPA analysis, with results described for each individual chromosomal region, a final comments column has been used to ascribe hyperdiploidy status or NAD (no abnormality detected). Of the successful 29 cases, 27 cases (93.1%) were shown to have an abnormal result. This is a very high

abnormality rate. Of note, MLPA has been incorporated into the genetic analysis of myeloma in a number of current UK myeloma trials.

The most common change detected by MLPA in our cohort was loss of chromosome 13 seen in 15/27 cases (55.6%). Gain of regions of chromosome 1q were seen in 12 cases (44.4%), and loss of regions of chromosome 1p were seen in eight cases (29.6%). Only two patients (7.4%) showed loss of *TP53* on the short arm of chromosome 17. Nine of the abnormal cases (33.3%) show hyperdiploidy as determined by the presence of two of the three chromosomes highly associated with gain in hyperdiploidy (chromosomes 5, 9 and 15)[118].

Sample No:	1p	1q	5	9	12	13	14	15	16	17	Comments
#188	DAB1 Gain	Gain									
#196		*CKS1B & RP11* Gain		*COL5A 1* Gain							
#197						Loss		*IGF1R* Gain			
#205										Low level loss	
#222			Gain	Gain				Gain			Hyperdiploid
#238											NAD
#243	*DPYD, COL11A1 & FAM46C* Loss		Gain	Gain				*GABRB3* Gain	*WWOX* Loss		Hyperdiploid
#61	*DPYD, COL11A1 & FAM46C* Loss					Low level loss					
#245	*DPYD, COL11A1 & FAM46C* Loss			*COL5A 1* Gain		Loss					
#247			Low level gain					Gain			Hyperdiploid
#248		Gain				Loss			*CYLD* Loss		
#251						Loss	Loss				
#113	Loss	Gain				Loss					
#140											NAD

#144					Loss	Loss		Loss		
#106			Gain	Gain			Gain	Loss		Hyperdiploid
#51			Gain	Gain		Loss	Gain	*WWOX* Loss		Hyperdiploid
#141								Loss		
#142			*PCDHB2* Loss							
#147	*FAM46C* Loss			*JAK2* Gain			Gain			Hyperdiploid
#134		Gain				Loss				
#146		Gain	Loss			Loss				
#165	*DPYD & FAM46C* Loss	Gain				Loss				
#168		Gain	Gain	Gain			Gain			Hyperdiploid
#170		Gain		Gain		Loss		*WWOX* Gain		
#175	*FAM46C* Gain	Gain	Gain	Gain			Gain			Hyperdiploid
#178	Loss	Gain			Loss	Loss		Loss	Loss	
#203	*DPYD & FAM46C* Loss		Gain	Gain		*DIS3* Loss	Gain			Hyperdiploid
#244		Gain	*PCDHA1* Gain	*COL5A 1* Gain		Loss				

Table 3.7: Detailing the results seen in the cohort of patients who had successful MLPA analysis. The results are presented by the chromosomal regions in which the 46 MLPA probes lie. Where gain or loss is recorded, this includes all probes within the chromosomal region, where specific probes are recorded smaller regions of loss or gain are involved. The final column records hyperdiploidy where two or more of the chromosomal gains associated with hyperdiploidy are present and NAD (no abnormality detected) where appropriate.

Of the 27 abnormal cases using MLPA, 14 cases were shown to be normal using the described FISH panel. This demonstrates that MLPA is able to offer additional information in some circumstances. Six of the 14 cases were shown to be hyperdiploid by MLPA and the current FISH panel design would not detect this. The remaining eight cases demonstrated abnormalities such as loss of chromosomes 12, 13, 14 and 16 not present on the FISH panel. Of these eight, there were four cases of 1p loss. The FISH panel includes a probe for 1p (*CDKN2C*), however all four cases demonstrated a partial loss not covered by this specific FISH probe.

Of the two cases shown to be normal by MLPA, one was also normal by both cytogenetics and FISH, and did not have array or NGS analysis. The second case was shown to have an *IGH-CCND1* rearrangement by FISH. This highlights the potential of using a combinatorial testing strategy of MLPA and FISH as the testing regimes offer strengths in assessing different abnormality types associated with myeloma.

In the 29 cases where copy number was assessed by both FISH and MLPA, (ie. deletion and duplication of chromosome 1p/1q and deletion of chromosome 17), data were checked for concordance. There were no cases of discordance for the 1p/1q except in those cases described above where the FISH probes employed did not cover the region of loss. However, chromosome 17 demonstrated considerable discordance. There were five discordant cases within the cohort of 29 (17.2%). The first case showed what was considered a possible low-level loss of *TP53* by MLPA, however, this was not confirmed by FISH. The remaining four cases were normal by MLPA, but abnormal by FISH. Two of these cases demonstrated deletion by FISH at 8% and 15%, the level of resolution for MLPA is considered to be around 15-25%[119]. Low level abnormal populations, therefore, can be problematic for MLPA, and in the case of multiple myeloma, plasma cell purification to a minimum level of 50% purification would be a prerequisite. All samples within our cohort were completed on DNA extracted from CD138+ve selected cells, and showed plasma cell purification levels exceeding 50% in each case. Even with this quality assessment in place, a small number of cases demonstrated low level abnormality. Small abnormal populations are widely seen and reported in myeloma patients, and the issue of mosaicism will always be a consideration in the genomic analysis of these patients.

The third case showed relative deletion of *TP53*, two copies of *TP53* against a background level of three copies of chromosome 17. The MLPA panel only assesses *TP53* so would not detect a change here, two being the normal and expected copy number for *TP53*. Relative deletion of *TP53* is a rare, but recurrent finding in myeloma. It will not be detected by all techniques and consequently has not been consistently assessed in trials or defined as part of the poor prognostic group. The final case demonstrated gain of chromosome 17 in 72% by FISH, it would be expected that MLPA would detect this abnormality. Considering the design and resolution of both FISH probes and the MLPA panel, this final case is the only true discordant case giving a discordance rate of 1/29 (3.4%) in our cohort.

A technological understanding of the MLPA technique is a prerequisite for those staff analysing the results of MLPA, however, the analytical software (Coffalyser) supplied by MRC Holland is straightforward to manage and use.

By way of example, the ratio graphs are shown for three cases in figures 3.9, 3.10 and 3.11, these demonstrate the ease with which the results can be interpreted once the analytical process has been completed and the ratio graphs are represented pictorially. Each probe is labelled with a number on the X axis, which is prefixed with the chromosome number. Above the graph the probes are labelled with the gene name, the exon number and the nucleotide length. The ratio for each probe in the myeloma MLPA probe set is represented by a circle with corresponding standard deviation bars. The rectangles show the same for the full run result, and are coloured blue for the 46 myeloma probes of interest, and green for the 11 control probes. The ratios represent loss and gain as described previously, and the ratio circles are coloured blue for gain and red for loss.

Figure 3.9 shows the MLPA result for patient #140. This patient shows a normal result for all probes within the MLPA set with all ratios at approximately 1.0, and most within the range bars set as acceptable standard deviation.

Figure 3.10 shows the MLPA result for patient #106. This patient has hyperdiploidy with gains of chromosomes 5, 9 and 15. The ratios for the probes on chromosomes 5 and 15 are at approximately 1.3-1.4 and suggest one additional copy of each chromosome. The ratios for the probes on chromosome 9 are approximately 1.7-1.8

suggestive of two additional copies, i.e. four copies of chromosome 9. This was confirmed by the cytogenetic analysis for this patient. There is also loss of the probes associated with chromosome 16.

As the final example, figure 3.11 shows the MLPA result for patient #248. This shows gain of the probes associated with the long arm of chromosome 1, loss of the probes on chromosome 13, and loss of the *CYLD* markers on chromosome 16. There is also evidence of loss of three probes within the control set. The control probes have been selected from 'quiet' areas of the myeloma genome, however given the highly complex nature of myeloma genomics, it would not be possible to select areas that are never involved in change. Involvement of the control probes is a phenomenon seen in a number of our samples. Whilst the results are clear from the histogram, abnormalities in the control probes disturb the internal algorithms and can cause interpretative problems. The Coffalyser software has been designed to manage this scenario, but of note this probe set is much more 'messy' than the MLPA CLL probe set as an example, where the data are clean with little deviation from the expected ratio and this is partly attributed to the less complex nature of CLL genetics and the fact that the control probes are rarely affected.

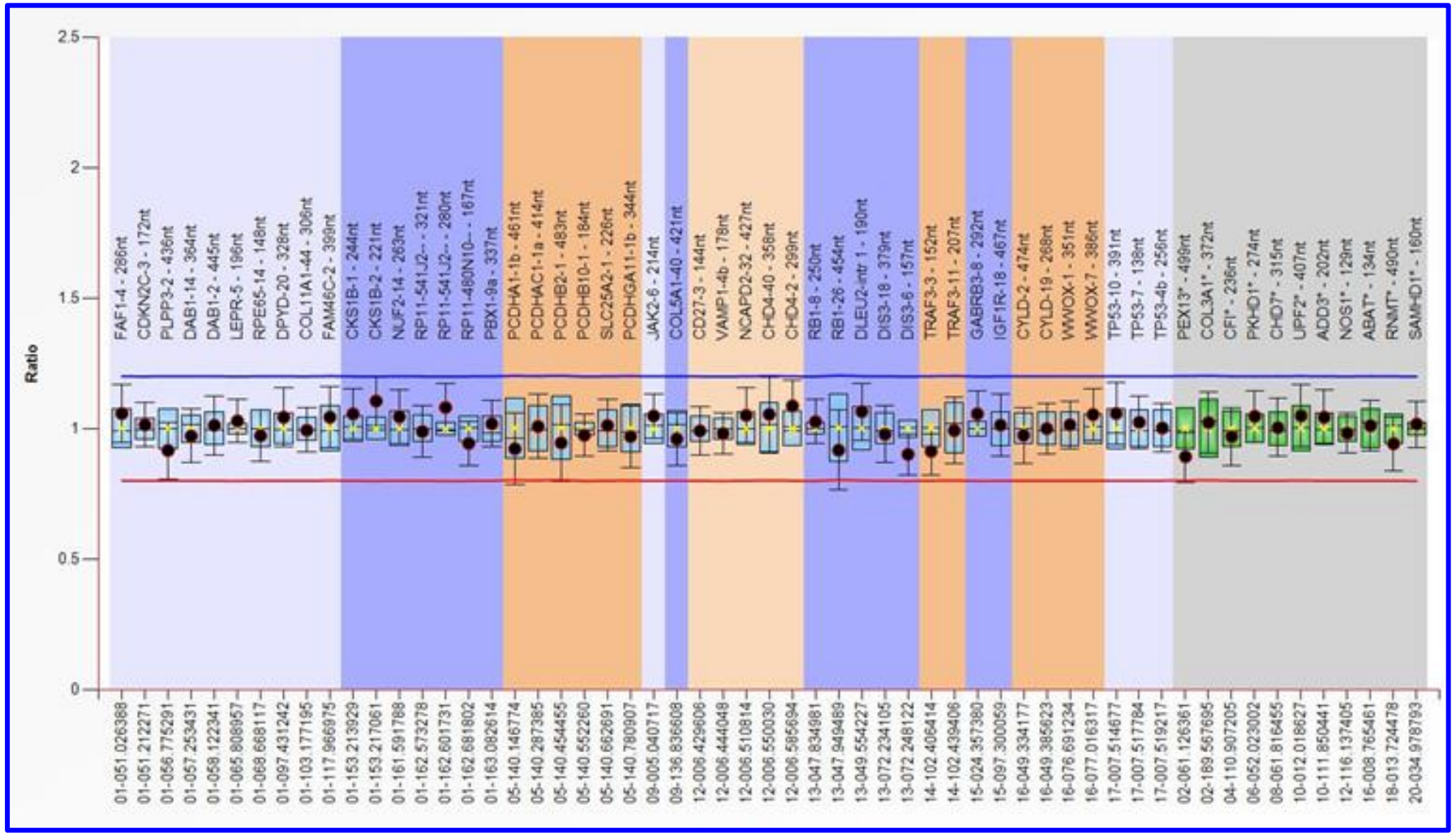


Figure 3.9: To show the MLPA results for patient #140 who has a normal MLPA profile.

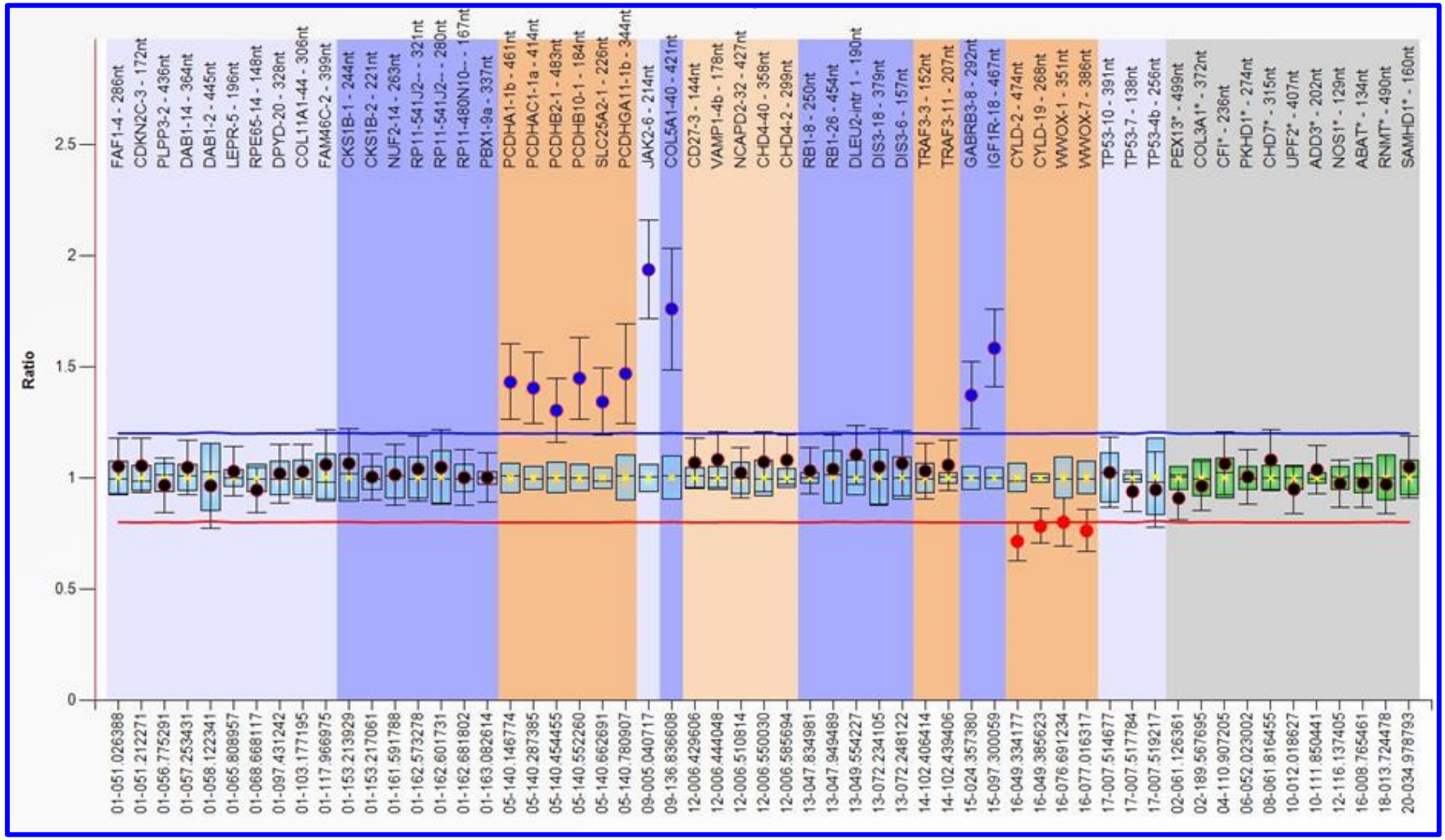


Figure 3.10: To show the MLPA results for patient #106, this shows gain of all probes associated with chromosomes 5, 9 and 15 and loss of probes associated with chromosome 16.

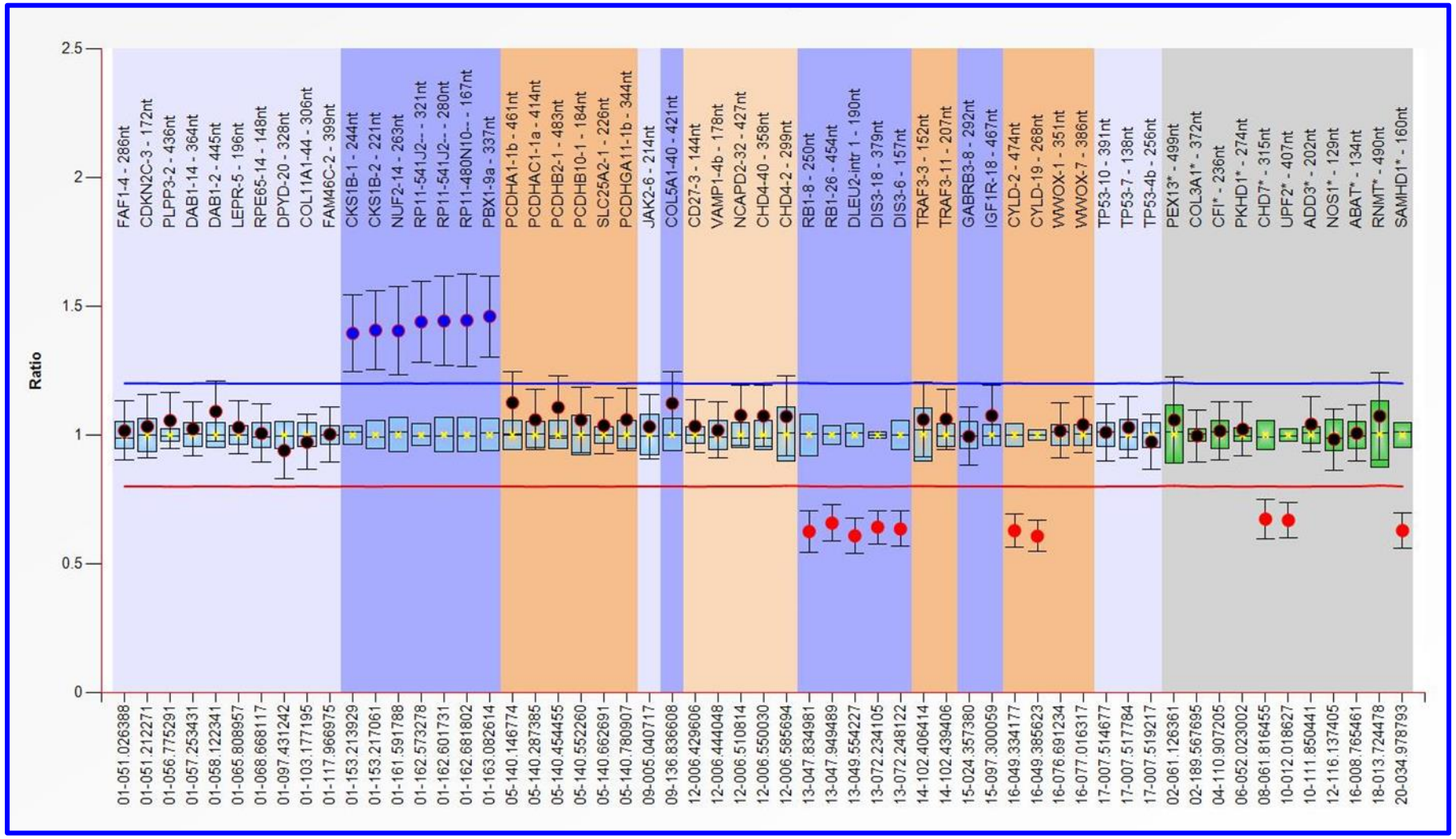


Figure 3.11 : To show the MLPA results for patient #248, the result shows gain of the probes associated with the long arm of chromosome 1, loss of probes on chromosome 13 and loss of the *CYLD* markers on chromosome 16. There is also evidence of loss of three probes within the control set.

Affymetrix CytoScan HD DNA Array

Many types of DNA array exist, but for the purposes of this project we have employed the Affymetrix CytoScan HD Array. Affymetrix arrays compare the DNA of interest to an internal normal reference set. This allows interrogation of the entire genome for copy number change, using more than 2.6 million markers for copy-number analysis and approximately 750,000 SNPs. The higher probe density areas include 533 cancer genes, 100% of which are covered with 25 markers per 100kb region. The array has the ability to detect loss of heterozygosity (LOH) and regions of acquired uniparental disomy (aUPD). Although SNP incorporation allows for extended analysis (of LOH), DNA array technology remains predominantly a method for detecting copy number changes. The nature of these arrays means the quantity of information created is huge and the extent of analysis required is also great.

Array analysis is not designed to detect balanced rearrangements or translocations, or mutations at a base pair level. However, in the context of myeloma, a number of translocations are seen in an unbalanced form which would be detected. It is also recognised, in other leukaemia types as well as myeloma, that small regions of copy number change can be generated at and around rearrangement breakpoints[120], detailed analysis of these regions can alert to the presence of rearrangements involving those genes. It should be noted that although array analysis may provide this information in a number of cases, this is not definitive evidence and should not be considered to exclude a rearrangement and indeed would require confirmation with an alternative technique in order to clinically report the finding.

DNA from 36 myeloma patients were processed using the Affymetrix CytoScan HD Array, DNA from CD138+ve cell selections was used in the assessment which allows targeting of tumour DNA.

Of this cohort, five samples failed to give a result. These failures were absolute and demonstrated large regions of the array with no image consistent with the probes simply not being present. The damage was integral and internal to the array cartridge, and therefore not considered to be damage that could have occurred during the processing, but more likely to be a result of damage at the time of

production. The company were contacted and informed, and confirmed that the error was in array manufacture.

Of the remaining 31 samples in which analysis was possible, the array quality was measured using MAPD (Median of the Absolute values of all Pairwise Differences). MAPD is a per-microarray estimate of the variability. It measures the variability in the log2 ratios by looking at the paired difference of all of the probes and taking median values, this can be considered a standard deviation measure. Cases were required to have a MAPD value of ≤0.25 to be considered good quality. Of the 31 cases, 23 cases (74.2%) met this QC score, whilst eight of the 31 cases (25.8%) samples had MAPD figures above this cut-off value. Five of these eight cases were marked as borderline quality (see table 3.8), and the analysis was considered acceptable. Three of these eight cases were considered poor quality, of these, one case was suggestive of hyperdiploidy, one showed no clear abnormalities, and the third showed some abnormalities, however, these findings should not be considered reliable in this context, and indeed the third case was not consistent with the FISH findings.

Analysis is undertaken using Chas software supplied from Affymetrix. The software is excellent but requires extensive training. The data sets from the arrays are huge and the analysis is time consuming and therefore, expensive. Whilst options are included in the software for reporting the data, it remains problematical to ensure that a full and comprehensive analysis has been undertaken, but within the confines of reporting relevant information to the clinician which will inform patient management. Much of the data and information gained from this analysis would not inform clinical management.

Of the 28 cases with an acceptable MAPD QC score, 24 (85.7%) cases showed an abnormal result, and four (14.3%) patients showed a no abnormality detected (NAD) result. This is a high abnormality rate, but not quite as high as that seen with the MLPA analysis. Given the MLPA analysis is panel based, (i.e. only assessing a small number of regions), genome-wide array analysis might have been expected to have the highest abnormality rate, although this technique also does not have the ability to detect balanced *IGH* rearrangements. This may be related to the specific cohort of

patients, but may also be related to the presence of low level disease. Array analysis assesses pooled DNA, and although the DNA was extracted from CD138+ve cell selections, it is thought that abnormalities present at 30% or less may not be detected using array analysis[113]. Myeloma is known to heterogeneous, and even with selected cell populations, some abnormalities may still be present at a low level and therefore will be missed using this type of analysis.

The array results are presented in Table 3.8. Of the 24 abnormal results, 16 patients (57.1% of the total analysed) showed a pattern consistent with hyperdiploidy, slightly higher than the expected reported figures.

The remaining eight cases (28.6% of the total cases analysed) showed losses, gains and LOH regions, but not associated with hyperdiploidy. Gain of 1q was seen in seven (25.0%) cases, loss of 1p was seen in five (17.9%) cases, loss of chromosome 13 was seen in 11 (39.3%) and loss of chromosome 17 was detected in a single case (3.6%). These figures relate to the expected abnormality rates described in the introduction.

Trial number	MAPD Score	Quality Comment	Affymetrix CytoScan HD Array Result
#51 (2)	0.171	Good	Hyperdiploid, with loss of 13, 16q & 20q
#113	0.294	Borderline	Gain of 1q, loss of 1p, 2q, 3p, 13, LOH of 16
#119	0.219	Good	Hyperdiploid, gain of 1q and 2p, loss of 13 and X
#134	0.168	Good	Gain of 1q, 8q, 3p, loss of 2q, 6q 12q & 13
#138 (2)	0.192	Good	Hyperdiploid, with loss of 6q
#144	0.169	Good	Gains of 15q, 18 and 19. Loss of 1p, 4, 8, 10p, 12, 13, 14, 15, 16, 17, 20, 21 & 22,
#146	0.185	Good	Gain of 1q, loss of 6q, 12q, 13,
#147	0.181	Good	Hyperdiploid, with loss of 1p, 8p, 2q, 6q
#151	0.210	Good	Hyperdiploid, with loss of regions on 8p, 13, 20p & 21
#157	0.247	Good	Gain of 18
#158	0.525	Poor	Failed, but suggestive of loss of 1p & gain of 1q (not consistent with FISH), gain of 19q & 20q, loss of 13, 14 & 20p.
#159	0.647	Poor	Failed, no clear abnormalities
#171	0.259	Borderline	Hyperdiploid with 1q gain
#179	0.247	Good	Hyperdiploid
#184	0.261	Borderline	Loss of chromosome 13 & 14

#190	0.274	Borderline	NAD
#201	0.215	Good	Gain of 11p, and loss of chromosome 8, 13, 14.
#202	0.216	Good	NAD
#203	0.199	Good	Hyperdiploid
#204	0.220	Good	Hyperdiploid with 1q gain, and 8p loss
#207	0.222	Good	Hyperdiploid with some double gains, loss of 8p, and LOH for chromosome 1
#209	0.227	Good	Hyperdiploid with loss of 8p
#208	0.251	Borderline	Gain of 1q
#216	0.242	Good	NAD
#222	0.240	Good	Hyperdiploid, LOH 16q
#226	0.239	Good	Hyperdiploid, with gain of 1q & X, & loss of 1p, 6q, 13.
#224	0.171	Good	Hyperdiploid
#230	0.200	Good	Hyperdiploid, loss and gain of chromosome 6 and LOH of chromosome 12
#232	0.242	Good	NAD
#233	0.235	Good	Hyperdiploid, with gain of 1q, and loss of 1p and 16q
#234	0.480	Poor	Failed, but suggestive of hyperdiploidy

Table 3.8: Table to show result of the array analysis excluding those five cases considered an outright fail.

The abnormalities detected by array analysis were not entirely concordant with the cytogenetic, FISH and MLPA results. Of the 28 cases processed for array and considered to have either good or borderline quality parameters, 19 (67.9%) cases demonstrated total concordance with the other genomics tests performed. There were nine (32.1%) cases, however, that demonstrated some discordance. Of these nine cases, five cases showed abnormalities by FISH that were seen in <30% of cells. As discussed, these low-level clones will not be detected by the DNA array. The remaining four included a relative deletion of chromosome 17, a 1p gain seen in 40% of cells by FISH, gains of chromosome 1 and 17 seen in 60% of cells by FISH, and finally a deletion of 17p seen in 58% of cells by FISH. DNA microarrays are designed to offer a dramatically better resolution than karyotyping and indeed FISH, and it would be expected that the array would detect this level of clone size. However, based on experience within a diagnostic laboratory using arrays alongside other techniques, the concordance is not always perfect. This view is also reported by Rack *et al* [113], who state 'we also observed some discordance relating to array sensitivity and to accurate assignment of ploidy group'. They reported that arrays fail to detect aberrations at low level, but specifically below 20% clonality[113].

The information is presented in multiple ways for array analysis but perhaps the simplest pictorial representation is the karyoview. This shows the chromosome karyogram with regions of gain marked in blue and regions of loss marked in red. Regions of LOH are represented in purple. More detailed analysis is possible by chromosome which gives further information on the B-alleles, copy number and the smoothed logR ratio. These images can be further enhanced to assess specific regions when required.

Figure 3.12, 3.13 and 3.14 show images from three of the array cases. The first case (patient #202) demonstrates a normal array. The second case (patient #138) shows a hyperdiploid karyotype with classic gains of whole chromosomes 3, 5, 7, 9 (two additional copies), 11, 15, 19 and 21 and loss of a region on the telomeric long arm of chromosome 6. The final example (patient #146) shows a gain of 1q, and loss of regions of 6q, 12q and 13, and a likely smaller region of loss on 5q. Expanded images of chromosome 11 have shown a 515kb region of gain at the site of *CCND1*. FISH

confirmed an *IGH-CCND1* rearrangement, and this pattern of gain (and/or loss) of regions around the breakpoints has been previously demonstrated by array in a number of 'balanced' cancer related translocations[120].

These traces are not entirely clean which may reflect poor quality DNA in some cases and DNA arrays are known to be sensitive to DNA quality. The project demonstrates optimistic results relating to the utility of this technique and the higher level of resolution, they also show that further work would be required in this area to optimise the quality of the output and to fully utilise the Chas software to analyse these cases in order to be used as a robust diagnostic tool in the analysis of myeloma.

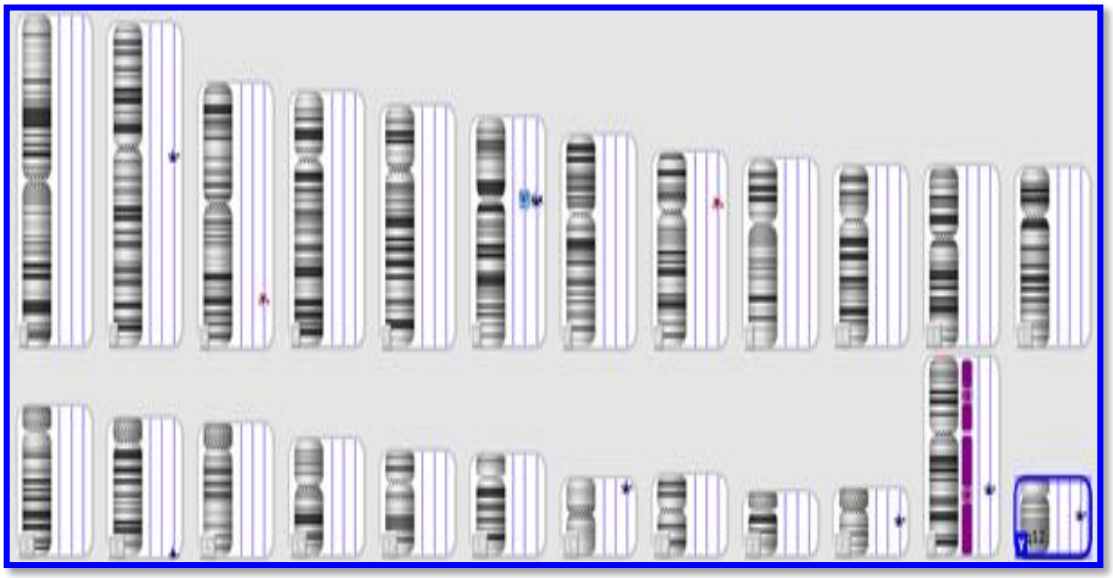


Figure 3.12: Demonstrates the array result for patient #202 showing a normal karyoview image with no clear abnormalities detected. The patient is male, and chromosome X shows as a region of LOH in this scenario.

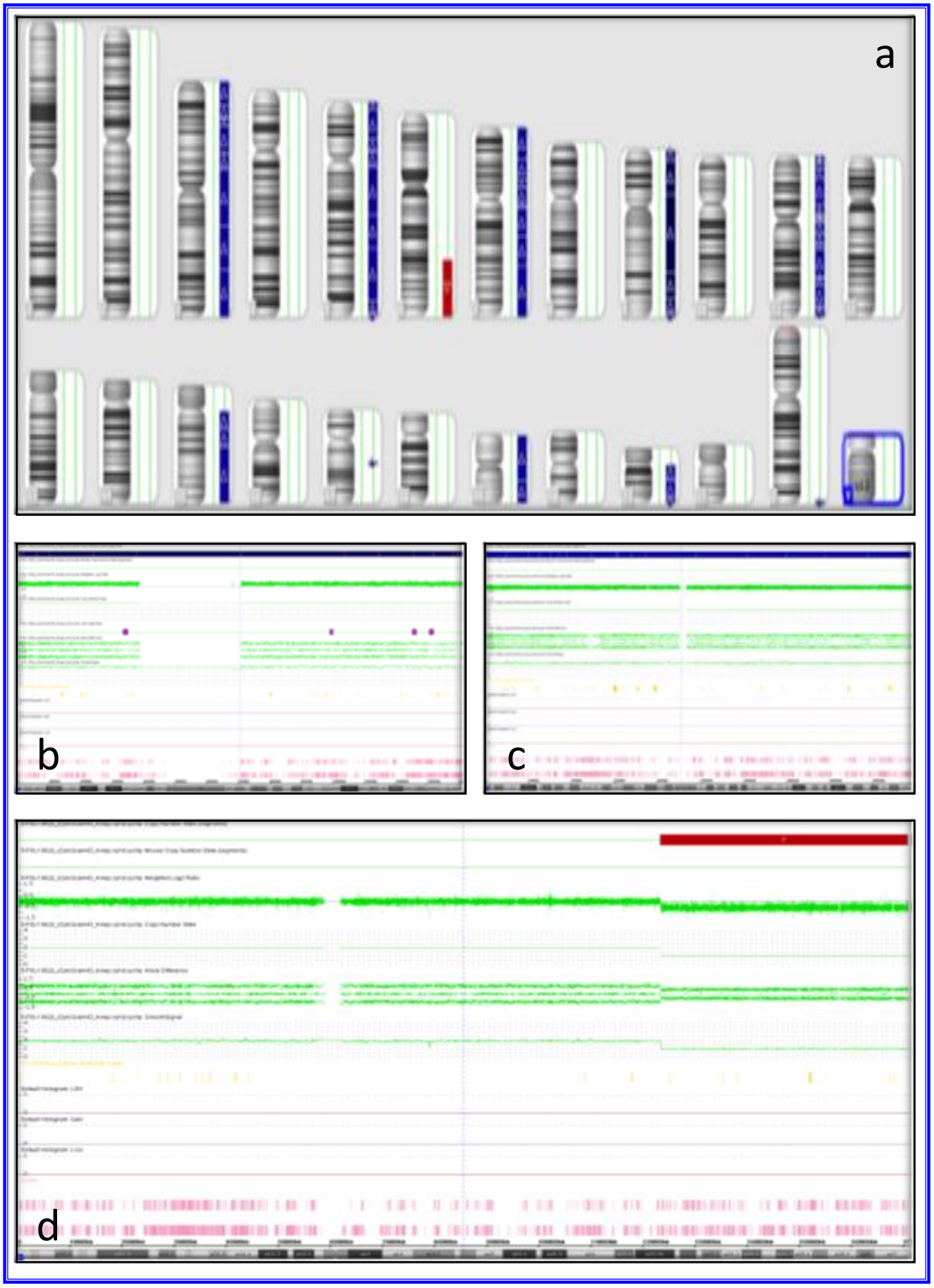


Figure 3.13: Demonstrates the array result for patient #138. The karyoview (a) demonstrates gain of chromosomes 3, 5, 7, 9, 11, 15, 19 and 21 as shown by the blue bars, and loss of the terminal region of the long arm of chromosome 6. Chromosome 9 has a darker blue bar suggesting an increase in copy number. The second two images (b & c) show a chromosome view of chromosome 3 and 9 in order to compare the B- alleles bars (second set of green bars); this confirms the presence of three copies of chromosome 3 and four copies of chromosome 9. The final image (d) in this set shows an enhanced view of chromosome 6 showing the copy number (top green bar) and B-allele bars (second set of green bars). This indicates the presence of two chromosome 6's up to 6q22, when the pattern changes to show a single copy and a telomeric deletion (indicated by the red bar).

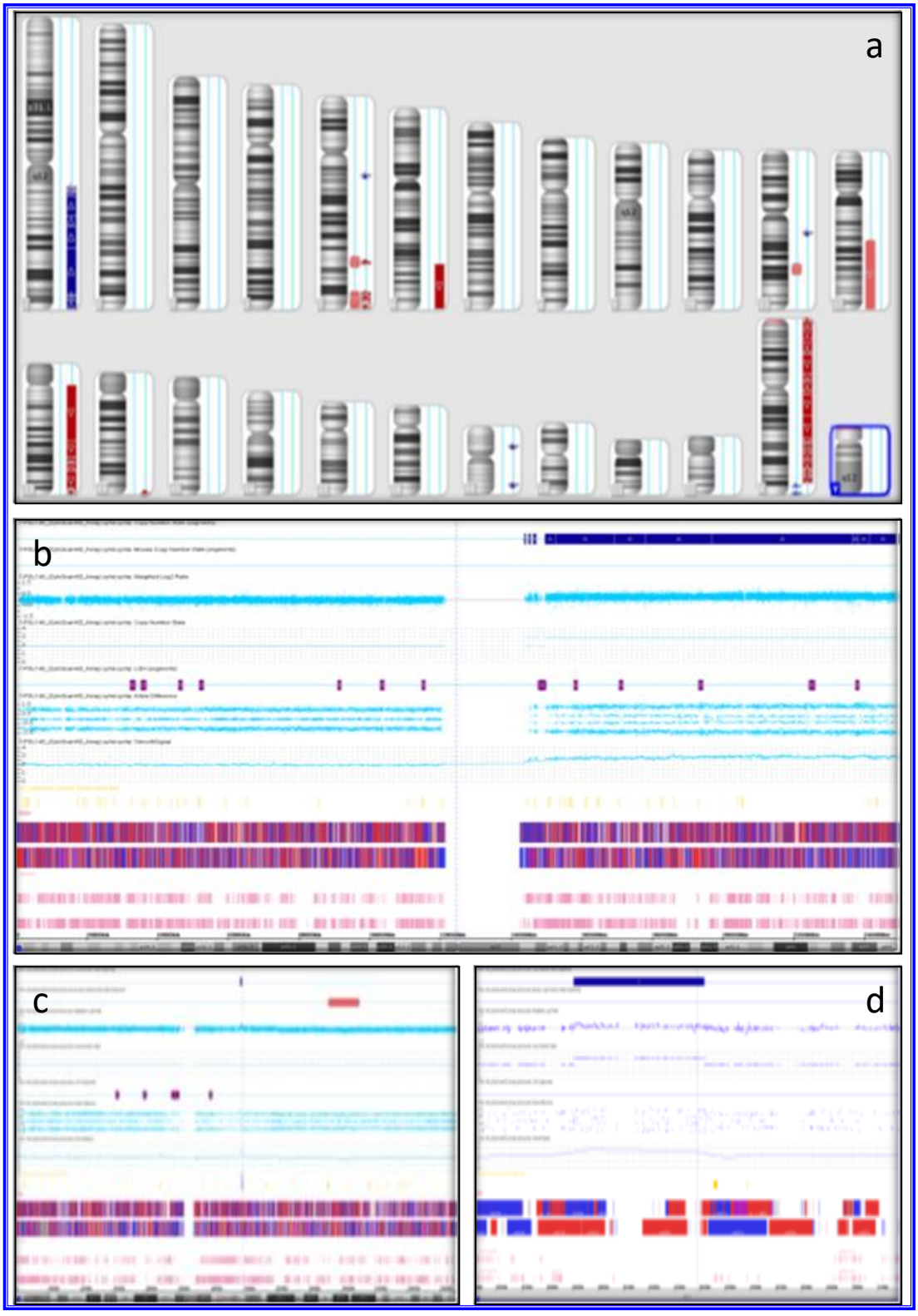


Figure 3.14: Demonstrates the array result for patient #146. The karyoview image (a) shows a gain of 1q (represented by a blue bar), and loss of regions of 6q, 12q and 13, and a likely smaller region of loss on 5q (represented by a red bar). The second image (b) within this group shows a chromosomal view of chromosome 1 depicting the gain of the long arm as shown by the copy number bar (top blue bar) and the B-alleles (second set of blue bars). Images c & d show chromosome 11; c shows the whole chromosome with a tiny blue bar (at the top of the image) at the region of *CCND1*, and d shows an expanded version of this demonstrating a 515kb gain. This patient was shown to have an *IGH-CCND1* rearrangement using FISH probes, some patients are reported to have both losses and gains around the region of the translocation break as demonstrated in this patient.

Targeted Next Generation Sequencing (NGS) Gene Panel

NGS technology has superseded Sanger sequencing for all but small target regions within the research arena, and indeed implementation of NGS technology into diagnostic laboratories is moving forward at a pace. There remain issues and concerns when transitioning this technology to the diagnostic arena with the expense, the magnitude of data, the time-consuming analysis and interpretation, the minimal analytical tools, and the issues of consent and ethics.

There are number of ways of applying NGS technology in a more manageable way, from technological, analytical and financial points of view. Instead of working with whole genome sequencing (WGS) or whole exome sequencing (WES), NGS panels offer a chance to harness the accuracy and depth of NGS analysis in a more targeted fashion. They require selection of a set of genes to create a panel design. Analysis is clearly limited to the set of genes that are involved in the panel, and this analysis strategy loses the functionality of novel gene discovery.

The targeted NGS Myeloma Gene Panel was designed as a single panel of 139 genes. However, this was essentially made up of two panels; the SMaRT NGS Myeloma Gene Mutation Panel which included 79 genes known to be involved in myelomagenesis, and the SMaRT NGS Osteome Probe Panel which had 60 bone related genes known to be involved in the development, influence, maintenance and destruction of bone. The results presented in this chapter will describe the performance of the NGS panel as a whole, but will discuss only those results associated with variants / mutations in the first part of the panel, those genes known to be involved in the pathogenesis of myeloma. The assessment of the osteome has been designed as an innovative aspect of the project to explore and assess possible or potential relationships between the genetic signatures associated with bone related genes, the 'osteome', and the likelihood or extent of bone damage associated with the patient's myeloma disease. This covers objective two of this project and will be described and discussed in Chapter 4.

DNA from 24 patients with myeloma or MGUS were processed using the bespoke targeted NGS myeloma gene panel. The samples were processed as pairs; DNA from

the CD138+ve plasma cell selections representing the tumour (myeloma) DNA, and DNA from blood samples from the same patients representing their constitutional (germline) DNA. Of the 24 patients processed, one sample (4.2%), patient #234, failed to give a result. However, 23/24 samples have given successful sequencing and therefore a success rate of 95.8% has been achieved for this bespoke myeloma NGS panel.

Of the 24 cases processed, the raw DNA sequence data taken from the Illumina HiSeq were analysed using the analysis pipeline described in chapter 2 - materials and methods NGS section. This pipeline conforms to the Broad Institute best practice guidance with some additional customised data filtering and QC check steps. The pipeline aligns reads to the hg19 (GRCh37) build human genome using Burrows-Wheeler Alignment (BWA). Variants were called using the Genome Analysis Toolkit (GATK) software and were then analysed using a variety of bioinformatics tools; mutect2, vardict and strelka to assess the likelihood of associated pathogenicity. A consensus set based on calls from all three tools was created based on the most popular call. This produced a list of variants with further annotation including ClinVar and ExAC allele frequencies, and information as to whether the gene was on either the Myeloma or Osteome gene panel for each patient in the cohort.

Figure 3.15 shows the number of paired and mapped reads for each of the patient samples. The first graph within the figure demonstrates the blood (germline) samples, which show a very consistent level of mapped and paired reads across each patient. An exception to this is patient #234 whose sample has failed the NGS process and does not have any mapped or paired reads. The second graph within the figure covers the CD138+ve (tumour) samples, these levels are very much more erratic likely to be associated with the variable quality of the DNA and disease stage but show a general increase in the number of mapped and paired reads compared to those in the blood samples.

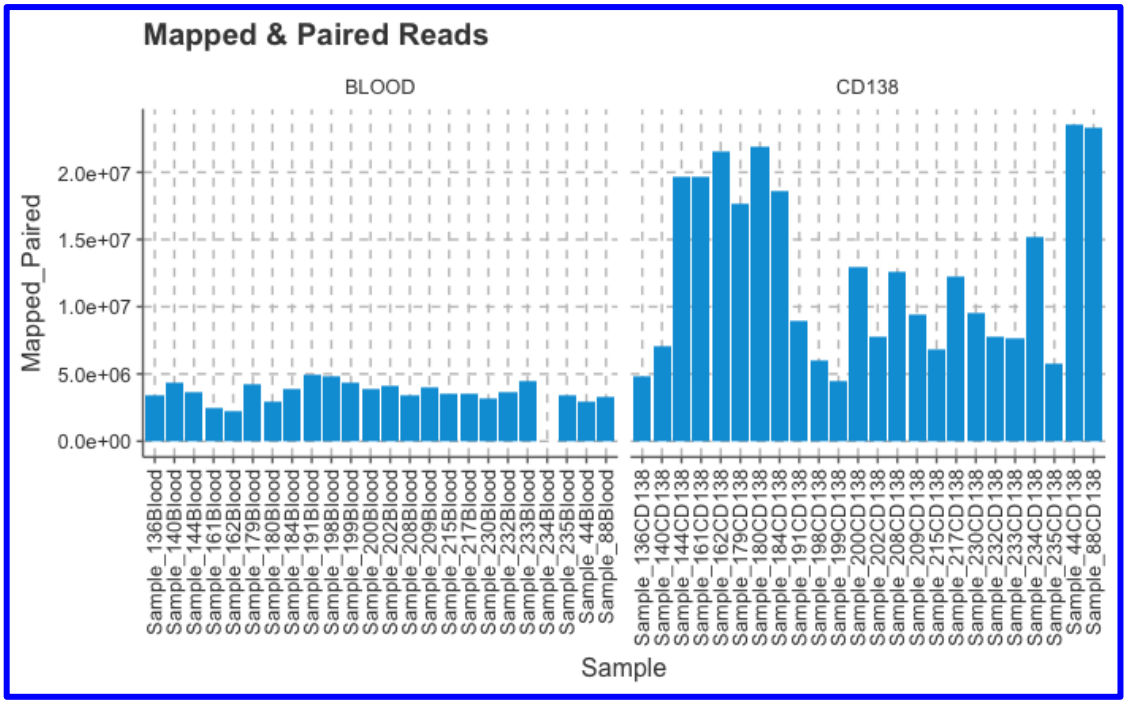


Figure 3.15: Demonstrates the number of paired and mapped reads for both blood (germline) and CD138+ve (tumour) samples for each of the 24 patients.

Duplicate reads arise following PCR of the same DNA molecule and are often defined as reads with an identical start and stop point. However, reads with identical start and stop points can arise from independent DNA molecules so this rule cannot identify true duplicates in all cases. In order to overcome problems associated with duplicate reads it is possible to introduce unique molecular identifiers (UMIs) to essentially barcode specific reads. This allows a separation of technical duplicates based on PCR from the same molecule, from true biological duplicates.

A proportion of our samples were prepared using UMIs, this included all of the bone marrow (tumour) samples and five of the blood (germline) samples. Samples prepared with UMIs allow the technical duplicates to be removed using bioinformatic methods leaving only the true biological duplicates. Figure 3.16 and 3.17 shows the number of duplicates and then duplicates as a percentage of the total reads. These graphs were created following the bioinformatic removal of the technical duplicates using the UMI information, and therefore show no duplicates. There is a high level of duplicate reads in the blood (germline) samples, with most samples showing between 30-45% duplicated reads.

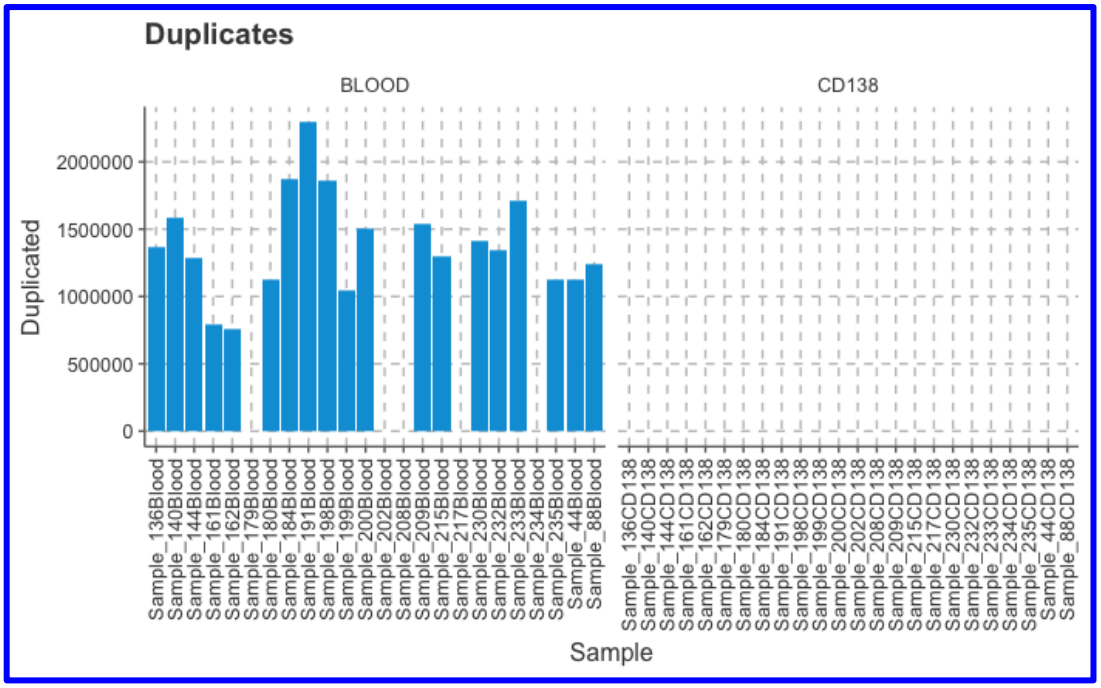


Figure 3.16: Demonstrating the number of duplicate reads per sample. A number of samples were prepared using UMIs and these samples show no duplication.

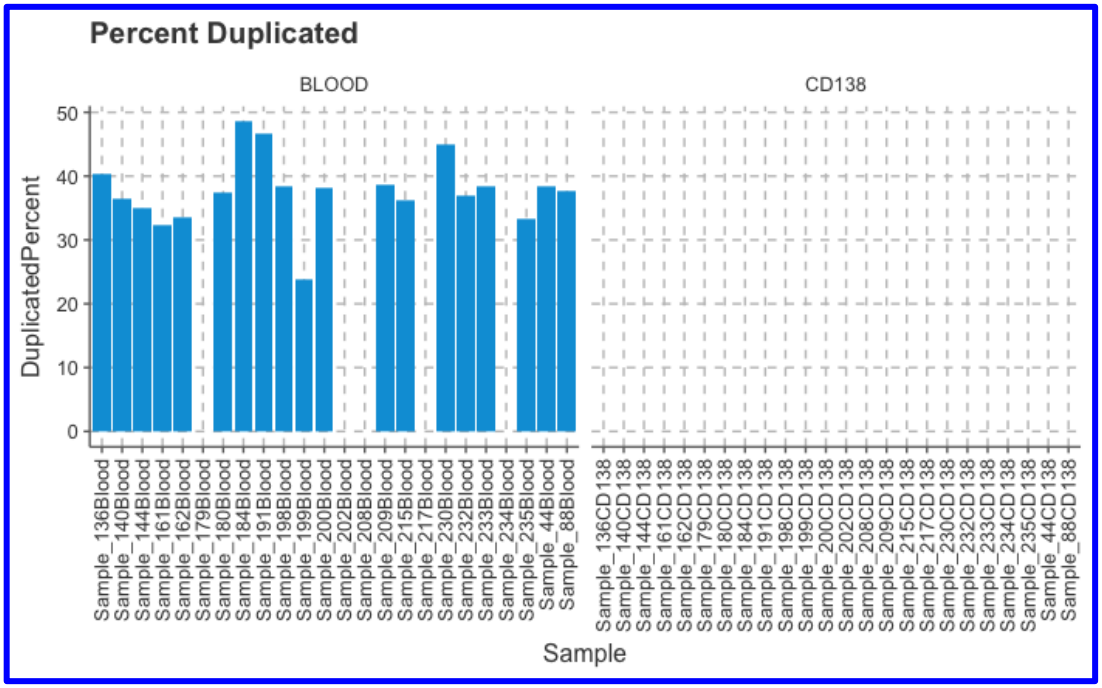


Figure 3.17: Demonstrates the number of duplicate reads expressed as a percentage of the overall reads. The samples prepared using UMIs show no evidence of duplicated reads.

Figure 3.18 shows a graph of the bone marrow samples and the five blood samples before the UMI data were utilised to remove the technical duplicates, allowing the original duplication rate to be seen.

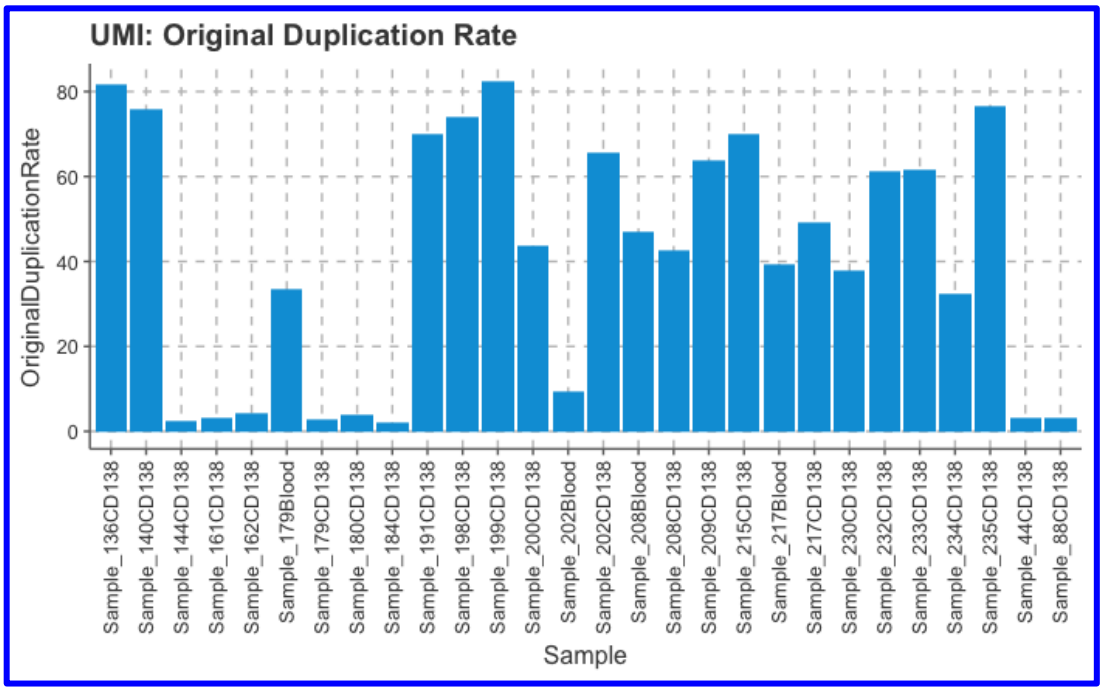


Figure 3.18: To show the original level of duplication seen in those samples processed using UMIs. This includes all of the CD138+ve tumour samples and the few blood (germline) samples also processed with UMIs.

Figure 3.19 shows a graph of the percentage of bases off target. This gives an indication of the number of bases sequenced that are not within the target regions described within the designed NGS panel. These values show low levels of off target sequencing (<1% for all cases). Although there is a level of sequencing outside of the targeted regions of the panels which can be considered wasted sequencing data, this is low and does not affect the coverage seen for the targeted regions of this panel.

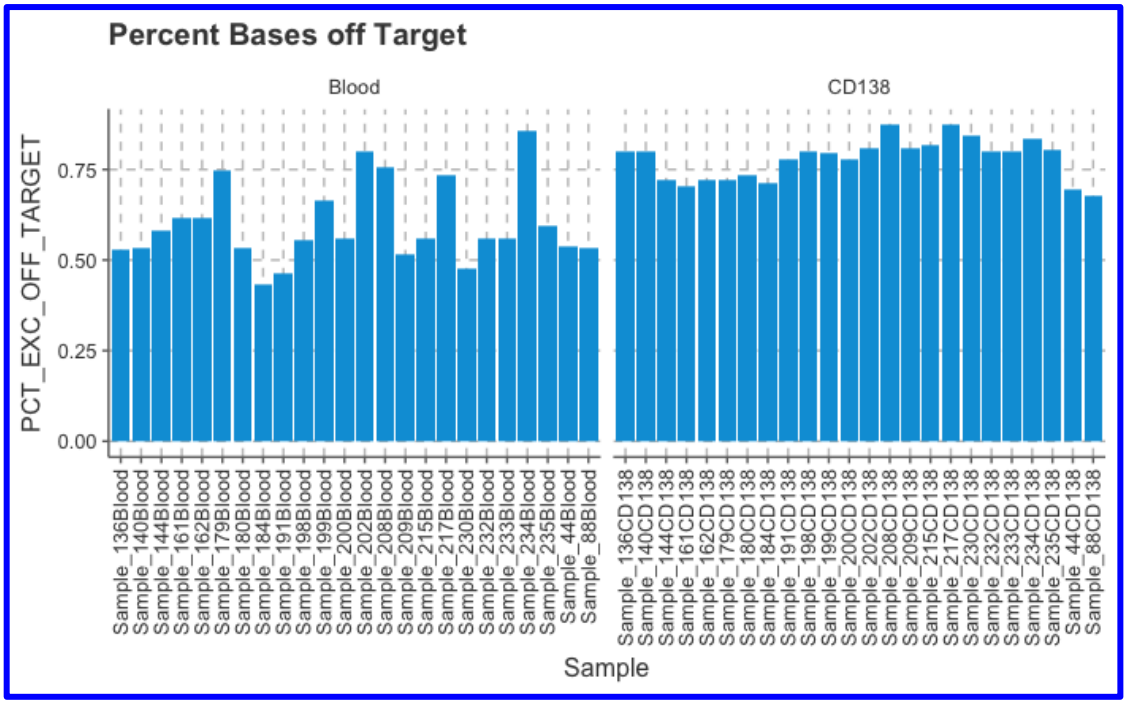


Figure 3.19: Demonstrating the levels of 'off target' sequencing for both blood (germline) samples and the CD138+ve bone marrow (tumour) samples.

The term 'coverage' is defined as the number of times a particular base in the reference genome is sequenced. The mean coverage was calculated using reads with a mapping quality of ≥30, bases with a quality of ≥30, and overlapping regions of the mate pairs were excluded. These figures relate to the Phred scores, which are a linear representation of the probability of a base being called incorrectly. A Phred score of 30 represents 99.9% accuracy (i.e. 1 in 1000 chance that a base call is incorrect). Figure 3.20 shows the mean coverage for each sample, both for the blood (germline) samples and for the CD138+ve (tumour) samples. When designing the experiment, coverage was considered and a higher proportion of the sequencing run was dedicated to the CD138+ve cell selection (tumour) DNA samples to facilitate the higher coverage we intended for the tumour samples aiming for a mean coverage level of 30x for the blood (germline) samples and 500x for the CD138+ve (tumour) samples. A level of 100x coverage is marked on the graph for reference.

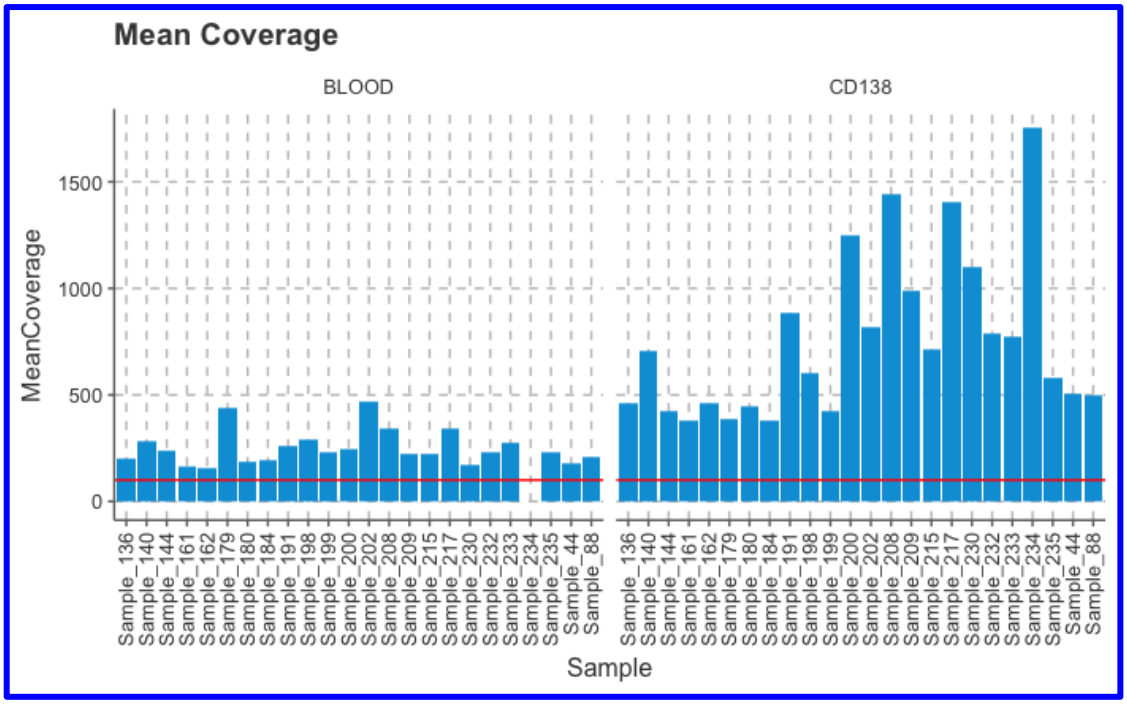


Figure 3.20: Demonstrates the mean coverage for each sample for both the germline (blood) and tumour (CD138+). Coverage is defined as the number of times a particular base in the reference genome is sequenced. The red line indicates 100x coverage.

The graph shows that coverage levels for the blood (germline) samples exceeded the 100x reference and well over the 30x coverage aimed for. The mean coverage levels for the CD138+ve (tumour) samples were all above 350x coverage, the majority (16/24) were above the 500x mean coverage aimed for with a number of samples dramatically above this level. The germline graph confirms the failure of sample #234, which shows zero coverage. Although this sample has failed the NGS sequencing, 23 of 24 samples have shown successful sequencing, resulting in a 95.8% success rate for this bespoke myeloma NGS panel.

Gaps in coverage are defined as bases which have <30x coverage, this again was calculated using reads with a mapping quality of ≥30 and bases with a quality of ≥30. Figure 3.21 demonstrates three samples with substantial gaps in coverage. Sample #234 is the sample that failed the NGS panel and has 100% gaps in sequencing. The other two samples #162 and #44 have shown 42.8% and 12.8% gaps in sequencing respectively, and the sequencing results from these samples require interpretation with caution.

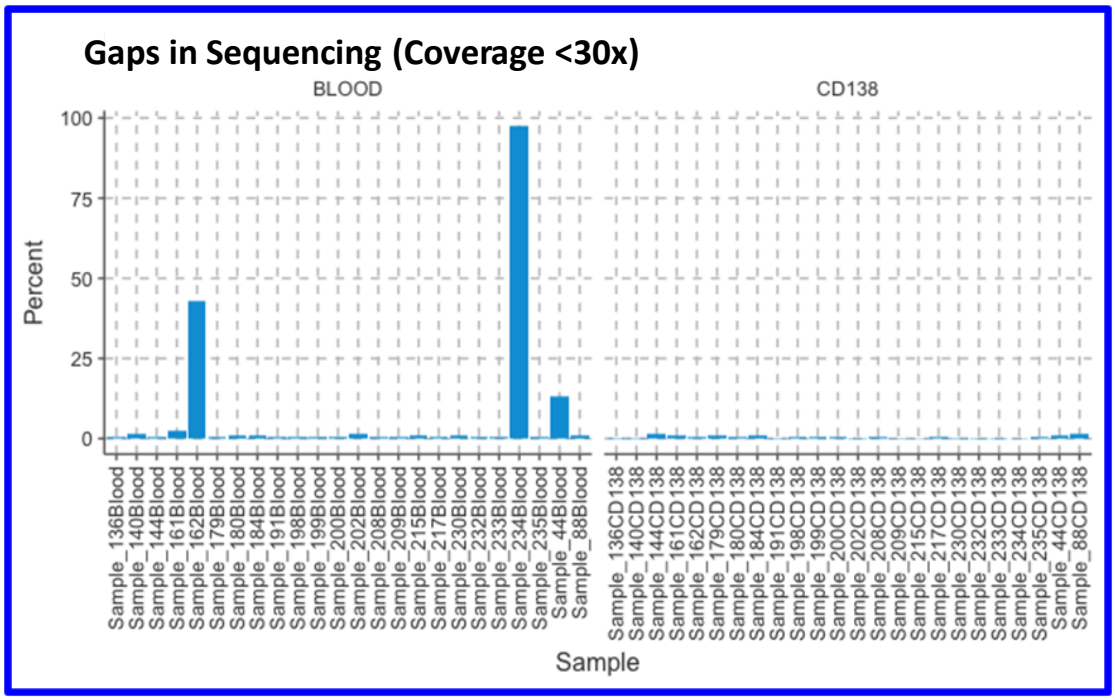


Figure 3.21: Demonstrating the percentage of gaps in sequencing seen across both germline and tumour samples. Sample #245 failed sequencing and showed 100% gaps in sequence. Samples #162 and #44 have shown 42.8% and 12.8% gaps in sequencing respectively.

The failed sample has skewed the scale in figure 3.21, therefore the following figure 3.22 shows the same information, but with the scale zoomed to demonstrate the gaps in sequencing percentage for the remaining samples. All remaining samples show <2.5% gaps in sequencing with the vast majority of samples showing <1% gaps in sequencing. Overall, the gaps in coverage are lower in the tumour (CD138+ve) samples explained in part by the mean coverage, which is higher in the tumour samples.

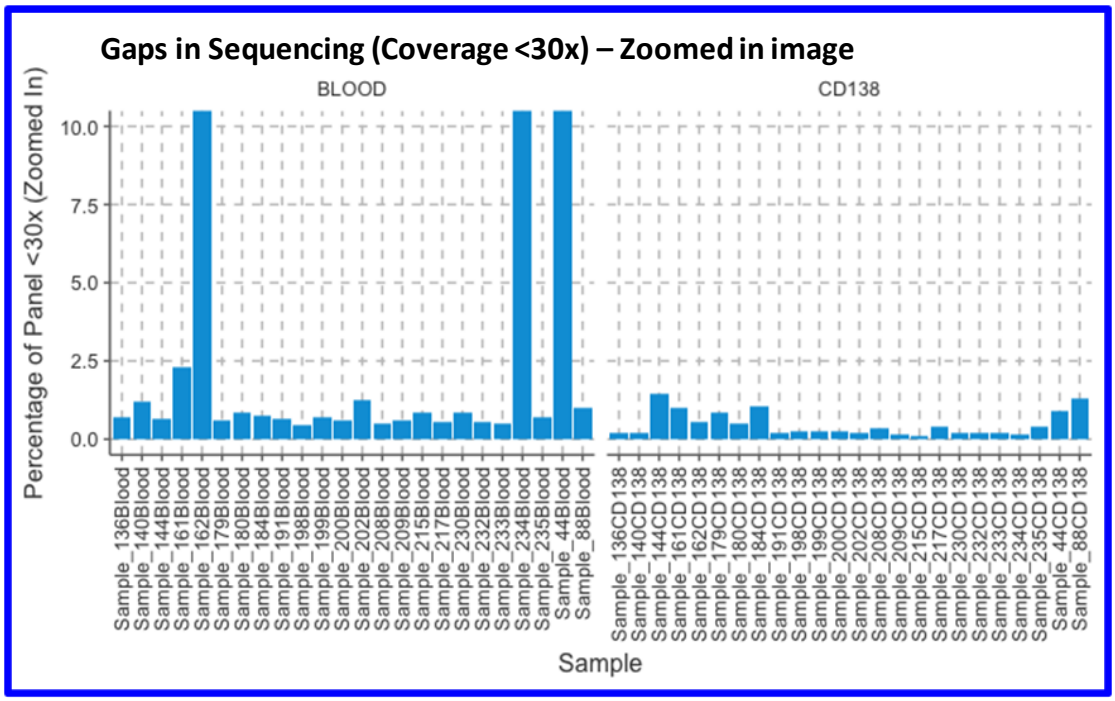


Figure 3.22: Demonstrating expansion of the lower end of the previous graph to show the levels of gaps in sequencing seen in the remaining samples (too small to see on previous graphs scale). All remaining samples show <2.5% gaps in sequencing with the vast majority of samples showing <1% gaps in sequencing.

The quality of the NGS data can be measured in a number of ways, and we have assessed the number of paired and mapped reads, the coverage and the gaps in coverage. Paired and mapped reads for each of the patient samples demonstrate a very consistent level across each patient in the blood (germline) samples, except for patient #234, which shows no evidence of paired or mapped reads. The number of paired and mapped reads within the CD138+ve (tumour) samples, are very much more erratic, and this is likely to be associated with the variable quality and quantity of the patient DNA and disease stage. However, in general there is an increase in the number of mapped and paired reads associated with the tumour samples compared to the germline samples. Dedicating a higher proportion of the sequencing capacity to the CD138+ve cell selection (tumour) DNA samples has facilitated the higher coverage we intended for the tumour samples, which in turn has produced an increase in the number of paired and mapped reads and the coverage, and a lower level in the gaps in coverage in the tumour samples. These quality assessments give a degree of confidence in the data.

Somatic variants have been derived by undertaking a further analytical step in which the germline variants have been subtracted from the list of variants seen in the tumour. This step is completed bioinformatically and effectively gives a set of variants only associated with the tumour sample. Figure 3.23 demonstrates the number of somatic variants recorded for each sample. Each of the 23 successfully processed samples has showed <10 variants, with the exception of patient #233 which has a dramatically higher level of variants, approaching 150. This patient has been discussed in greater depth in the Case Scenarios section within this chapter.

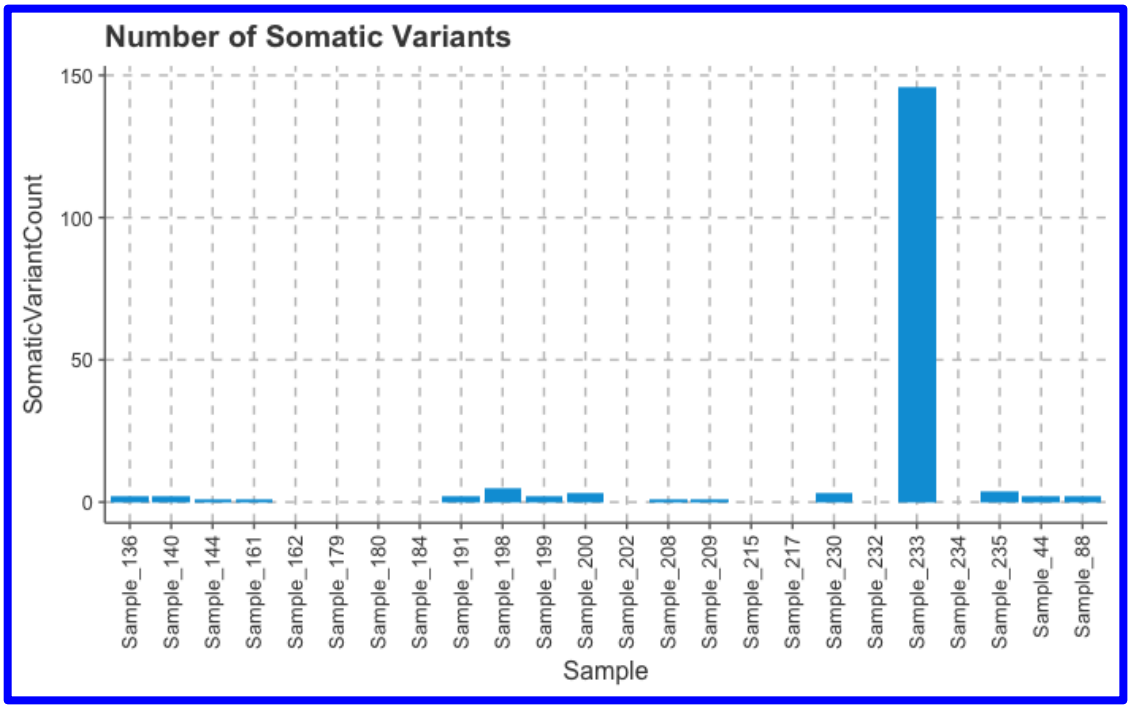


Figure 3.23: Shows the number of somatic variants for each sample sequenced, demonstrating that each sample has less than 10 somatic variants per sample with the exception of sample #233 that has approaching 150 somatic variants.

The somatic variant numbers demonstrated in figure 3.23, were called following analysis through the pipeline and included some duplicate variants where the bioinformatic pipeline had assessed the same variant in different transcripts. Further manual analysis was undertaken to remove these duplicates and then to assess the pathogenic status of each variant. Those variants likely to result in a protein change and associated with a high level of pathogenicity were considered of potential clinical utility. Assessing pathogenicity can be considered quite subjective and can be

extremely time-consuming if each variant is fully assessed individually; the pathogenicity in these cases was based on the bioinformatic pipeline calls using the consensus of three variant calling databases.

Table 3.9 shows the somatic variants considered to potentially pathogenic for each patient following the removal of duplicate variants. The table details the gene in which the variant was seen, with both the genomic level and protein level HGVS (Human Genome Variation Society) variant description nomenclature[107], as well as the likely effect of the variant and the likely pathogenic status based on the consensus from the pathogenicity assessment using mutect2, verdict and strelka. Of the 22 patients, an average of 1.32 somatic variants (range 0-4) were seen per patient. Those variants likely to result in a protein change and associated with a high level of pathogenicity are highlighted in the darker blue.

Patient	Variant Numbers	Gene	Effect	hgvs.c	hgvs.p	Status
#44	2	*FAM46C*	missense variant	c.731T>G	p.Leu244Arg	Strong Somatic
		FAT1	synonymous variant	c.12915G>A	p.Ala4305Ala	Strong Somatic
#88	2	*DNAH5*	intron variant	c.8010+67T>A		Strong Somatic
		ZFHX4	missense variant	c.6848A>C	p.Lys2283Thr	Strong Somatic
#136	2	*NFKB2*	missense variant	c.1405G>A	p.Asp469Asn	Strong Somatic
		DIS3	missense variant	c.2458C>T	p.Arg820Trp	Strong Somatic
#140	2	*NRAS*	missense variant	c.182A>G	p.Gln61Arg	Strong Somatic
		EGR1	missense variant	c.89A>G	p.Asp30Gly	Strong Somatic
#144	1	*NRAS*	missense variant	c.190T>G	p.Tyr64Asp	Strong Somatic
#161	1	*LTB*	splice donor variant & intron variant	c.280+2T>C		Strong Somatic
#162	0					
#179	0					
#180	0					
#184	0					
#191	2	*CCND1*	missense variant	c.174G>C	p.Lys58Asn	Strong Somatic
		KMT2A	intron variant	c.11430-12C>A		Likely Somatic
#198	4	*FAT4*	splice donor variant & intron variant	c.12816+1G>A		Likely Somatic
		FAT3	missense variant	c.2389T>C	p.Tyr797His	Likely Somatic
		TRAF3	frameshift variant	c.1142_1158delTGGAGTCCCAGCTGAGC	p.Leu381fs	N/A
		NF1	frameshift variant	c.2033dupC	p.Ile679fs	Strong Somatic
#199	2	*NRAS*	missense variant	c.181C>A	p.Gln61Lys	Strong Somatic
		FAM46C	conservative inframe deletion	c.1054_1056delAAC	p.Asn352del	Strong Somatic
#200	3	*BRAF*	missense variant	c.1780G>A	p.Asp594Asn	Likely Somatic
		CDKN2A	missense variant	c.35C>T	p.Ser12Leu	Likely Somatic
		KMT2A	missense variant	c.368G>A	p.Gly123Asp	Likely Somatic
#202	0					

#208	1	*ERBB4*	missense variant	c.1361G>T	p.Ser454Ile	Strong Somatic
#215	0					
#217	0					
#230	3	*KRAS*	missense variant	c.183A>C	p.Gln61His	Strong Somatic
		CUL4A	sequence feature	c.849-27A>G		Strong Somatic
		PSMG2	splice region variant & intron variant	c.57+3G>T		Strong Somatic
#232	0					
#234	Sample Failed NGS Analysis					
#235	4	*ATM*	stop gained	c.5623C>T	p.Arg1875*	Likely Somatic
		KRAS	missense variant	c.64C>A	p.Gln22Lys	Likely Somatic
		KRAS	missense variant	c.57G>C	p.Leu19Phe	Strong Somatic
		MAF	synonymous variant	c.168C>T	p.Ser56Ser	Likely Somatic
#233	89	Complete list of variants not included here, see table 3.10.				

Table 3.9: Detailing the somatic variants seen in patients assessed using the bespoke NGS myeloma panel.

Table 3.10 presents a simplified list of those variants considered of potential clinical utility. Patient #234 has been excluded as this sample failed NGS, and patient #233 has also been excluded from this part of the analysis given the exceptionally high variant level observed in this sample.

Patient	Variant Numbers	Gene	Effect	hgvs.p
#44	1	*FAM46C*	missense variant	p.Leu244Arg
#88	1	*ZFHX4*	missense variant	p.Lys2283Thr
#136	2	*NFKB2*	missense variant	p.Asp469Asn
		DIS3	missense variant	p.Arg820Trp
#140	2	*NRAS*	missense variant	p.Gln61Arg
		EGR1	missense variant	p.Asp30Gly
#144	1	*NRAS*	missense variant	p.Tyr64Asp
#161	0			
#162	0			
#179	0			
#180	0			
#184	0			
#191	1	*CCND1*	missense variant	p.Lys58Asn
#198	2	*FAT3*	missense variant	p.Tyr797His
		NF1	frameshift variant	p.Ile679fs
#199	2	*NRAS*	missense variant	p.Gln61Lys
		FAM46C	conservative inframe deletion	p.Asn352del
#200	3	*BRAF*	missense variant	p.Asp594Asn
		CDKN2A	missense variant	p.Ser12Leu
		KMT2A	missense variant	p.Gly123Asp
#202	0			
#208	1	*ERBB4*	missense variant	p.Ser454Ile
#209	0			
#215	0			
#217	0			
#230	1	*KRAS*	missense variant	p.Gln61His
#232	0			
#235	3	*ATM*	stop gained	p.Arg1875*
		KRAS	missense variant	p.Gln22Lys
		KRAS	missense variant	p.Leu19Phe

Table 3.10: Detailing the variants seen in patients assessed by NGS considered to be pathogenic or resulting in a protein change.

Based on the analysis of the gaps in sequencing, both patients #44 and #162 were shown to have poorer quality sequencing. The gaps in sequencing amount 12.8% in #44, but 42.8% in #162. It is possible in these cases, and particularly in patient #162, that variants may not have been detected.

As detailed in table 3.10, ten cases (45.5%) showed no evidence of variants considered to be pathogenic. Twelve cases demonstrated the presence of 20 pathogenic variants; two cases showed three variants, four cases showed two variants and six cases showed a single variant in the genes included in the NGS myeloma gene panel. This gives an average of 0.9 (range 0-3) variants per patient sample. The 20 pathogenic variants seen were within 15 different genes listed here:

> *ATM, BRAF, CCND1, CDKN2C, DIS3, EGR1, ERBB4, FAM46C, FAT3, KMT2A, KRAS, NF1, NFKB2, NRAS* and *ZFHX4*

A single pathogenic variant was detected in each gene except *NRAS, KRAS* and *FAM46C. NRAS* and *KRAS* show three different variants and *FAM46C* demonstrated two variants.

The MAPK pathway is dysregulated in approximately 55% of MM patients. Genes which have been associated with MAPK-signalling include *KRAS, NRAS, BRAF* and more recently *NF1* and *RASA2*[62]. In our cohort, eight of the 20 (40.0%) variants were seen in these genes, three in *NRAS*, three in *KRAS*, and one in each of *BRAF* and *NF1*, and these relate to seven of the 22 (31.8%) patients.

In a diagnostic context, the variants produced as a result of NGS or WGS are analysed by gathering data, often using Alamut, from a number of different sources. This assesses the type of variant, population frequencies, COSMIC data, ClinVAR data and dbSNP data to more thoroughly assess the likelihood of being pathogenic or disease causing. For the final 20 variants, this analysis was completed. The information gathered through Alamut and comparison with known hotspot data from Lohr *et al* [61] are presented in appendix 3. This analysis would effectively downgrade a number of variants from the assessment produced by the bioinformatics analysis. Of the 20 variants considered pathogenic based on the bioinformatics assessment, 11 would be considered pathogenic following this more in depth analysis, five of those being previously reported hotspots in myeloma: the

NRAS Q61 (seen twice in our data) and Y64 hotspots, *BRAF* D594 and the *KRAS* Q61 hotspot. The remaining nine variants would be considered variants of unknown significance (VUS). This would suggest that pathogenicity callers used as part of the bioinformatics pipeline variant assessment are overcalling the level of pathogenicity.

Patient #233 showed dramatically more variants than the other samples. This number was approaching 150 overall, but with removal of duplicates numbered 89 variants. Those variants likely to result in a protein change and associated with a high level of pathogenicity were extracted and presented in table 3.11.

Patient	Variant Number	Gene	Effect	hgvs.c	hgvs.p
#233	45	*VCAM1*	splice region variant & intron variant	c.662-7C>T	
		ANP32E	disruptive inframe deletion	c.576_581delAGAGGA	p.Glu193_Glu194del
		ANP32E	missense variant	c.332C>A	p.Ala111Glu
		PARP1	missense variant	c.2285T>C	p.Val762Ala
		NCKAP5	missense variant	c.2930T>C	p.Ile977Thr
		NCKAP5	missense variant	c.2809G>A	p.Val937Ile
		FRZB	missense variant	c.598C>T	p.Arg200Trp
		ROBO1	splice region variant & intron variant	c.4283-6C>T	
		FAT4	missense variant	c.9246A>T	p.Glu3082Asp
		FAT4	missense variant	c.9643C>T	p.His3215Tyr
		SFRP2	splice region variant & intron variant	c.502+6C>T	
		FAT1	missense variant	c.11155G>A	p.Val3719Met
		FAT1	missense variant	c.10660T>G	p.Ser3554Ala
		FAT1	missense variant	c.385G>C	p.Val129Leu
		DNAH5	missense variant	c.12401C>T	p.Ala4134Val
		DNAH5	missense variant	c.8586G>T	p.Leu2862Phe
		DNAH5	missense variant	c.2296A>T	p.Ile766Leu
		CCND3	missense variant	c.775T>G	p.Ser259Ala
		CCND3	missense variant	c.395T>G	p.Leu132Arg
		ROS1	missense variant	c.500G>A	p.Arg167Gln
		LRP5	splice region variant & intron variant	c.884-4T>C	
		LRP5	splice region variant & intron variant	c.2318+6T>C	
		FAT3	missense variant	c.8983C>A	p.Gln2995Lys

		FAT3	missense variant	c.10552G>T	p.Val3518Leu
		FAT3	splice region variant & intron variant	c.12938-6A>T	
		ATM	splice region variant & intron variant	c.5497-8T>C	
		CCND2	splice region variant & synonymous variant	c.570C>G	p.Thr190Thr
		CHD4	splice region variant & intron variant	c.4371-5C>T	
		CHD4	missense variant	c.417G>T	p.Glu139Asp
		KRAS	missense variant	c.34G>A	p.Gly12Ser
		FLT3	splice region variant & intron variant	c.1310-3T>C	
		FLT3	missense variant	c.680C>T	p.Thr227Met
		TNFSF11	missense variant	c.107C>G	p.Pro36Arg
		PSMB5	missense variant	c.70C>T	p.Arg24Cys
		KREMEN2	missense variant	c.225C>A	p.Ser75Arg
		KREMEN2	missense variant	c.234C>G	p.His78Gln
		KREMEN2	missense variant	c.1222G>C	p.Ala408Pro
		WWOX	missense variant	c.535G>A	p.Ala179Thr
		KDM6B	splice region variant & intron variant	c.456+6A>G	
		CDH2	missense variant	c.352G>A	p.Ala118Thr
		SMAD7	splice region variant & intron variant	c.667+8G>A	
		TGFB1	missense variant	c.29C>T	p.Pro10Leu
		MMP9	missense variant	c.836A>G	p.Gln279Arg
		KREMEN1	missense variant	c.203G>A	p.Arg68Gln
		KREMEN1	missense variant	c.874C>G	p.Leu292Val

Table 3.11: Somatic variants for #233, only including those considered pathogenic and resulting a protein change

There were 45 pathogenic variants in this patient, a significant increase on the other 22 patients which showed an average of 0.91 (range 0-3) potentially pathogenic variants. This patient will be discussed further as part of the case studies section in this chapter. The 45 pathogenic variants are shown in table 3.11, which details the gene in which the variant was seen, the effect of the variant, both the genomic level and protein level HGVS (Human Genome Variation Society) variant description nomenclature[107].

The *KRAS* G12 variant has been highlighted in blue, demonstrating the presence of this as one of the hotspots seen in the Lohr *et al* [61] data.

Overall, the bespoke myeloma NGS assay worked well and picked up known, recognised and relevant mutations associated with these genes in myeloma. Currently these abnormalities do not contribute to the diagnosis of myeloma or change clinical management or treatment in myeloma patients, but current literature and trials data suggests that inclusion of this data is consistent with the direction of travel[38,61,62,67], and will direct specific treatment strategies.

Case Studies

Three cases studies have been prepared with the aim of demonstrating some of the more interesting abnormal cases seen as part of the study, but to also highlight different aspects detected using the different comparative technologies. The first case study is illuminating in showing the strengths of the different techniques demonstrating the understanding obtained from the cumulative effect of these genetic technologies and their output. The second case describes a rare but interesting finding of a jumping translocation, and the third explores the hypermutation profile detected as part of the NGS studies.

Patient #113 – Complex genetics associated with myeloma patient at progression to plasma cell leukaemia

Patient #113 is a male, born in 1945, and was 67 years old when diagnosed in November 2012 with plasma cell myeloma. He had a 70g/L IgG paraprotein with 80% plasma cells. His first line treatment was cyclophosphamide, thalidomide, dexamethasone (CTD) and warfarin. He suffered a thigh bleed on warfarin and was switched to bortezomib (Velcade), cyclophosphamide, dexamethasone (VCD). Following nine courses of VCD, patient #113 showed a partial response to therapy (paraprotein level reduced to 9g/l) and proceeded to high dose melphalan and an autologous transplant in July 2013. In October 2013, at day 100 post transplant, patient #113 was shown to have 80% plasma cells and relapsed disease. Third line treatment included lenalidomide (Revlimid) and dexamethasone, which gave a short lived partial response before rapid progression to plasma cell leukaemia with hepatic plasmacytomas. The bone marrow sample used as part of this study was taken at this stage, in December 2013. Patient #113 died later that month.

The bone marrow sample was processed for cytogenetic analysis, FISH, MLPA and array analysis. The cytogenetic karyotype was:

44,X,-Y,i(1)(q10),add(2)(q31),?add(3)(p1?3),del(6)(q21),add(8)(q24),-13,t(14;20)(q32;q11)

This shows a complex karyotype (see figure 3.24) with a translocation between chromosomes 14 and 20, a recognised *IGH* rearrangement in MM. The cytogenetics

also showed an 8q rearrangement suggesting *MYC* involvement, an isochromosome 1q resulting in the loss of the short arm of chromosome 1 and gain of the long arm of chromosome 1, monosomy for chromosome 13 and loss of the Y chromosome. Further abnormalities of chromosomes 2, 3 and 6 were also identified, but their exact nature could not be determined by cytogenetics alone.

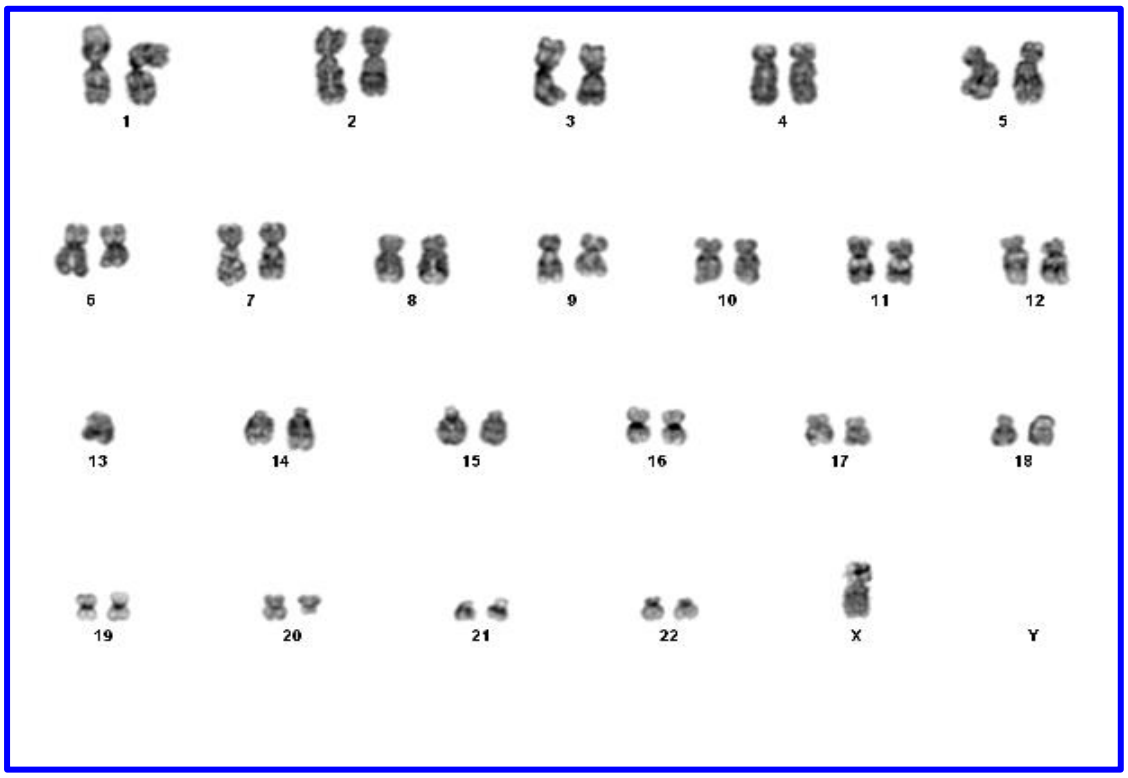


Figure 3.24: Demonstrates a representative karyotype from patient #113. It shows a translocation between chromosome 14 and 20, the presence of the isochromosome 1q, loss of chromosome 13, abnormality of the long arm of chromosome 8, and additional abnormalities of chromosomes 2, 3 and 6.

FISH using the Vysis Abbott *IGH* breakapart probe demonstrated a rearrangement. Figure 3.25 shows the *IGH* probe map (top) with both red and green probes either side of the *IGH* breakpoint. This shows a fusion signal (i.e. yellow) when *IGH* is not rearranged and splits into its component red and green signals when a rearrangement is present.

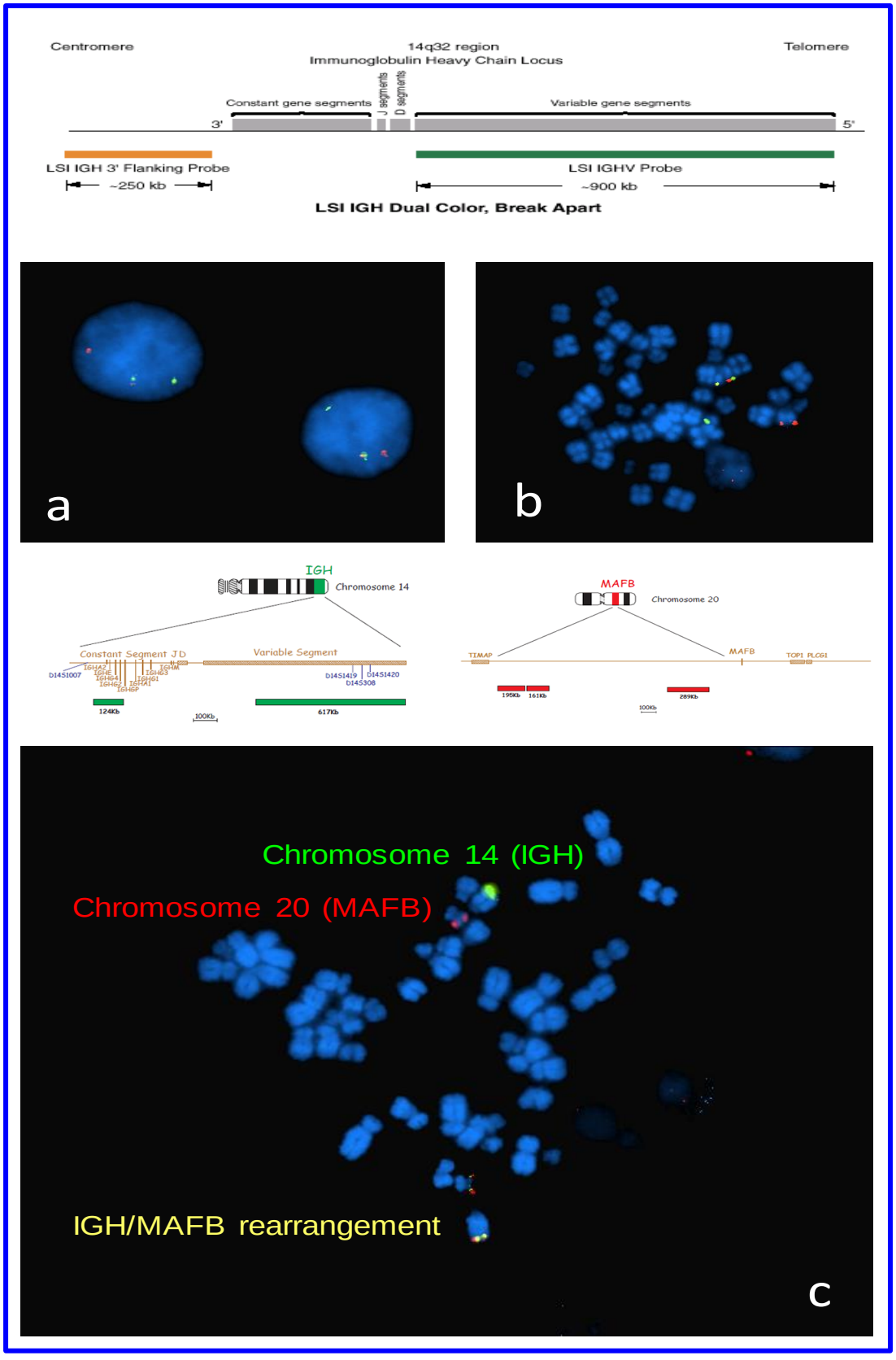


Figure 3.25: FISH images from patient #113 using the Vysis Abbott *IGH* breakapart and *IGH/MAFB* probes. The *IGH* probe map demonstrates the red and green probes flanking the *IGH* region of chromosome 14. Image a and b show an abnormal rearranged pattern in two interphase cells and a metaphase cell. The second probe map shows the *IGH* gene covered using a green probe and the *MAFB* gene covered using a red probe. Image c shows a red signal on chromosome 20, and green signal on chromosome 14 and two fusion (yellow) signals on the derivative 14 and derivative 20 chromosomes.

Figure 3.25a shows a IF1R1G rearranged signal pattern in two interphase cells, and figure 3.25b shows this same pattern in an interphase cell allowing the position of these signals to been seen. The interphase cell shows the normal fusion signal on a D group chromosome consistent with the chromosome 14 and shows the red signal on a large D group chromosome likely to be the derivative 14 and the green signal on the small derivative chromosome 20.

To confirm the involvement of chromosome 20, the Abbott Vysis *IGH/MAFB* probe was used. The second probe in figure 3.25 map shows the *IGH* gene covered using a green probe and the *MAFB* gene covered using a red probe. Figure 3.25c shows a red signal on chromosome 20, and green signal on chromosome 14 and two fusion (yellow) signals on the derivative 14 and derivative 20 chromosomes, confirming the presence of an *IGH-MAFB* rearrangement.

The step 1 FISH assessment also included probes for *CDKN2C/CKS1B* (1p/1q) and the *TP53*/17cen. The *CDKN2C/CKS1B* (1p/1q) FISH confirmed the loss of the short arm and gain of the long arm as expected from the cytogenetic analysis. Figure 3.26 shows the chromosome 1 array profile (top left) with loss of 1p in red and gain of 1q in blue, 3.26a shows the normal chromosome 1 (left) and the isochromosome 1q (right) karyotypically and 3.26b shows an interphase FISH image with three copies of the red *CKS1B* probe on 1q and one copy of the green *CDKN2C* probe on 1p. The TP53/17cen probe is shown on an interphase FISH image in figure 3.26c in which there are two copies of *TP53* in red and two copies of 17cen in green, i.e. a normal profile for this probe.

Figure 3.26 also shows the array profile (bottom left) and the karyotypic image (figure 3.26e) of both chromosome 2s, the array profile would suggest both loss and gain of the long arm of chromosome 2 confirming the complex nature of the rearrangement seen karyotypically. *MYC* FISH was also completed following the finding of the karyotypic abnormality of the long arm of chromosome 8 and confirmed a rearrangement of *MYC.* Figure 3.26 also shows the array profile of chromosome 8 (bottom right) and the karyotypic image of chromosome 8 (figure 3.26f); the array profile suggested gain of the long arm of chromosome 8, and the *MYC* FISH (figure 3.26d) confirmed a non-standard *MYC* rearrangement with a

1F2R1G pattern suggesting rearrangement, with duplication of the derivative chromosome.

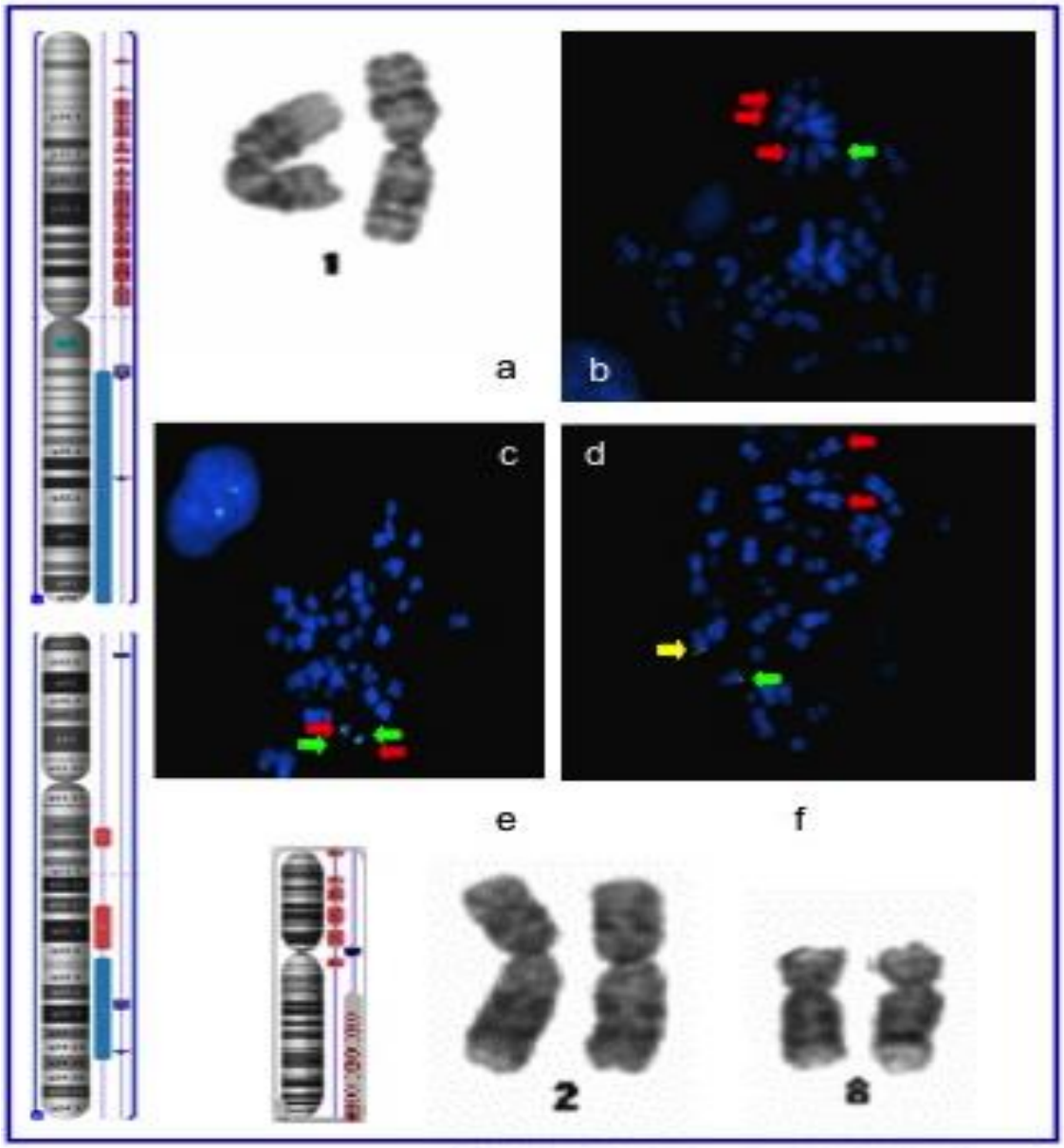


Figure 3.26: Shows cytogenetic, FISH and array results from patient #113. The isochromosome 1 is shown karyotypically in image a, with the array profile (top left) and the FISH (image b) confirming loss of the short arm and gain of the long arm. Image c shows a normal profile using the TP53/17cen probe. Image e shows a complex rearrangement of chromosome 2 and this is confirmed by the array profile (bottom left). Image f shown the karyotypic rearrangement of chromosome 8, with the array profile (bottom right) and the MYC FISH (image d) confirming a non-standard 1F2R1G signal pattern consistent with MYC rearrangement and duplication of the derivative chromosome.

MLPA analysis (figure 3.27), also acts as confirmation of both the loss and gain of chromosome 1p and 1q respectively and monosomy for chromosome 13. The other regions included in the MLPA probe list are considered within normal limits. However, one of the control probes which lies on the long arm of chromosome 2 (involving gene *COL3A1*) also demonstrates loss. The control genes have been chosen in 'quiet' areas of the myeloma genome due to their rarity of involvement in myeloma karyotypes. However, as many myeloma karyotypes demonstrate extreme complexity, the control areas are frequently involved. The demonstration of loss in the region of chromosome 2q provides additional information to the rearrangement of chromosome 2 seen cytogenetically.

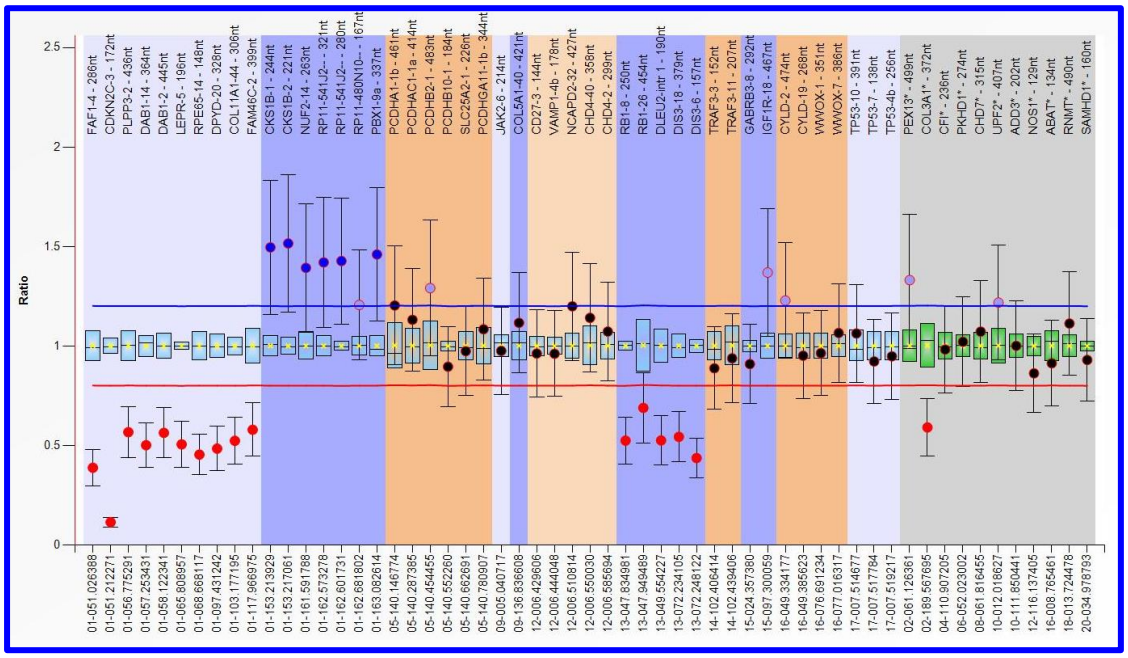


Figure 3.27: Shows the MLPA graphical representation from patient #113. The quality is poor demonstrated by the very large confidence limit bars, but still clearly demonstrates loss (in red) of the short arm of chromosome 1, gain (in blue) of the long arm of chromosome 1, and loss of chromosome 13. Loss of the control marker on chromosome 2 is also noted.

The array had a poor MAPD QC value of 0.294 (values of <0.25 are considered to be of high enough quality for a diagnostic report), however, some valuable information was still gained from the analysis. The array was able to confirm the isochromosome 1q with both loss of 1p and gain of 1q, and loss of chromosome 13 (figure 3.28). It

also showed loss of a much larger region of chromosome 2 than was determined by cytogenetics, and highlighted a more complex rearrangement of *MYC* with gain around the region of 8q24. When revisiting the karyotype, it looks possible that the deletion of the long arm of chromosome 2 could be replaced at least in part by the derivative chromosome 8 involving the *MYC* rearrangement.

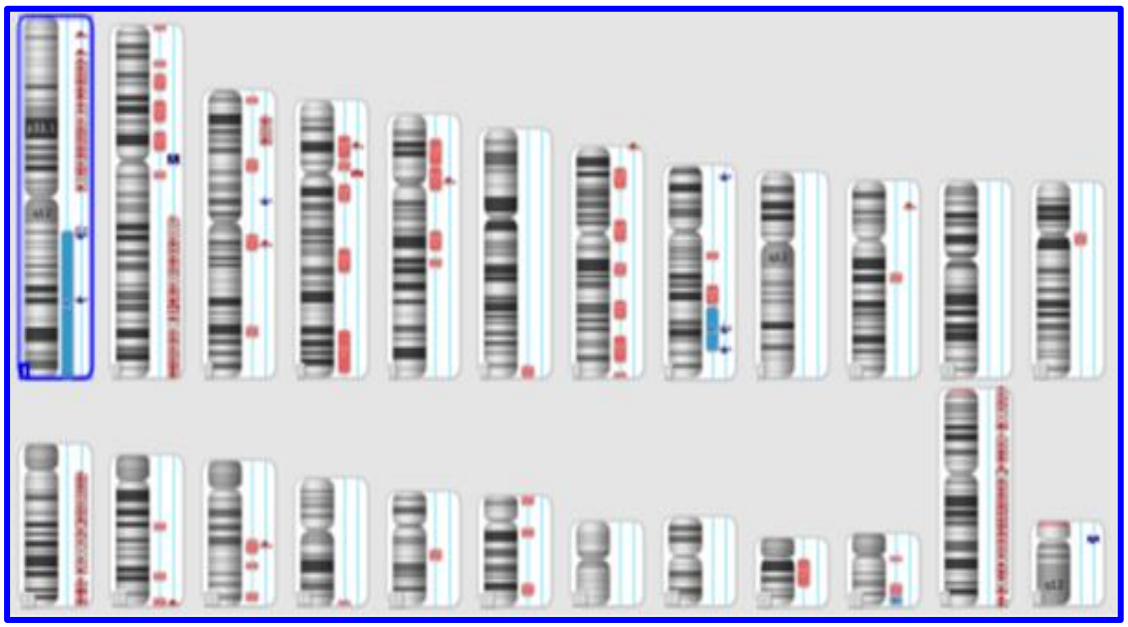


Figure 3.28: Shows array results from patient #113. The QC for this patient was poor, but losses of 1p and 13 were confirmed as were gains of the long arm of chromosome 1.

In summary, patient #113 has a plasma cell leukaemia with an *IGH-MAFB* rearrangement. Although *IGH* rearrangements are seen in approximately 60% of myeloma patients, the *IGH-MAFB* rearrangement is rare, seen in only ~2% of patients, and it is associated with a poor prognosis. The rearrangement was seen karyotypically and confirmed to involve the *IGH* and *MAFB* genes using FISH probes. Loss of 1p and gain of 1q (as an isochromosome in our case) was detected cytogenetically and then confirmed by FISH, MLPA and array. Gain of 1q is also associated with poor prognosis, and has been seen in higher frequencies in patients with disease progression. Monosomy 13 was determined by cytogenetics and confirmed by MLPA and array. In terms of prognosis, monosomy 13 is no longer considered an independent prognostic factor in myeloma, its original poor prognostic status is thought to be due to its strong association with t(4;14). The

rearrangement of chromosome 8 at the site of the *MYC* gene was detected cytogenetically and then confirmed using FISH probes. The FISH signal pattern suggested a non-standard rearrangement, which was also confirmed with the array. When the karyotype was revisited with the FISH and array information, a more complex rearrangement of *MYC* involving the long arm of chromosome 2 was suspected.

This case report shows the strengths of each technique in the setting of a complex genomic picture. Each of the additional techniques is able to add information to create a deeper and clearer understanding of the abnormalities seen. This has been useful in demonstrating the cumulative effect of these genetic technologies and their output.

Patient #175 – Myeloma patient with hyperdiploidy and a possible jumping translocation

Patient #175 is an Afro Caribbean male who was 58 years old when he presented in June 2015 with an IgG kappa monoclone of 62.6g/L, immunoparesis, haemoglobin (Hb) of 110g/L, normal creatinine and no evidence of renal failure. His bone marrow plasma cells were at a level of 70%. Multiple lytic lesions were detected in his ribs, sternum, pelvis and spine, and he was diagnosed at this stage with plasma cell myeloma.

He was treated with cyclophosphamide, thalidomide, dexamethasone (CTD) off trial and completed 5 courses. At this stage he demonstrated a VGPR (very good partial response) with his IgG monoclone dropping to 4.3g/L. In January 2016, he received an autologous stem cell transplant (ASCT) following induction chemotherapy. In March 2018, our patient relapsed and his second line treatment included bortezomib (Velcade) and dexamethasone, and he received a second ASCT in June 2018. The six cycles of treatment were completed in October 2018 and he is currently in remission and clinically well.

The bone marrow sample involved in this study was taken at diagnosis (June 2015) and was processed for cytogenetic analysis, FISH and MLPA. The cytogenetic karyotype was:

56,XY,+3,+5,+5,+7,+9,+9,+11,+15,der(17)t(1;17)(q11;p13),+19,+21[cp9]/

46,XY[5]

This demonstrated a male hyperdiploid karyotype with whole chromosome gains of 3, 5, 7, 9, 11, 15, 19 and 21. The majority of chromosomes show a single extra copy, but both chromosomes 5 and 9 were present as two additional copies, four copies per cell in total. There was also an unbalanced translocation between chromosomes 1 and 17; this presented as two normal copies of chromosome 1 with an additional copy of the long arm of chromosome 1 on the short arm of chromosome 17, resulting in an overall the gain of the long arm of chromosome 1 and loss of the short arm of chromosome 17 (See figure 3.29).

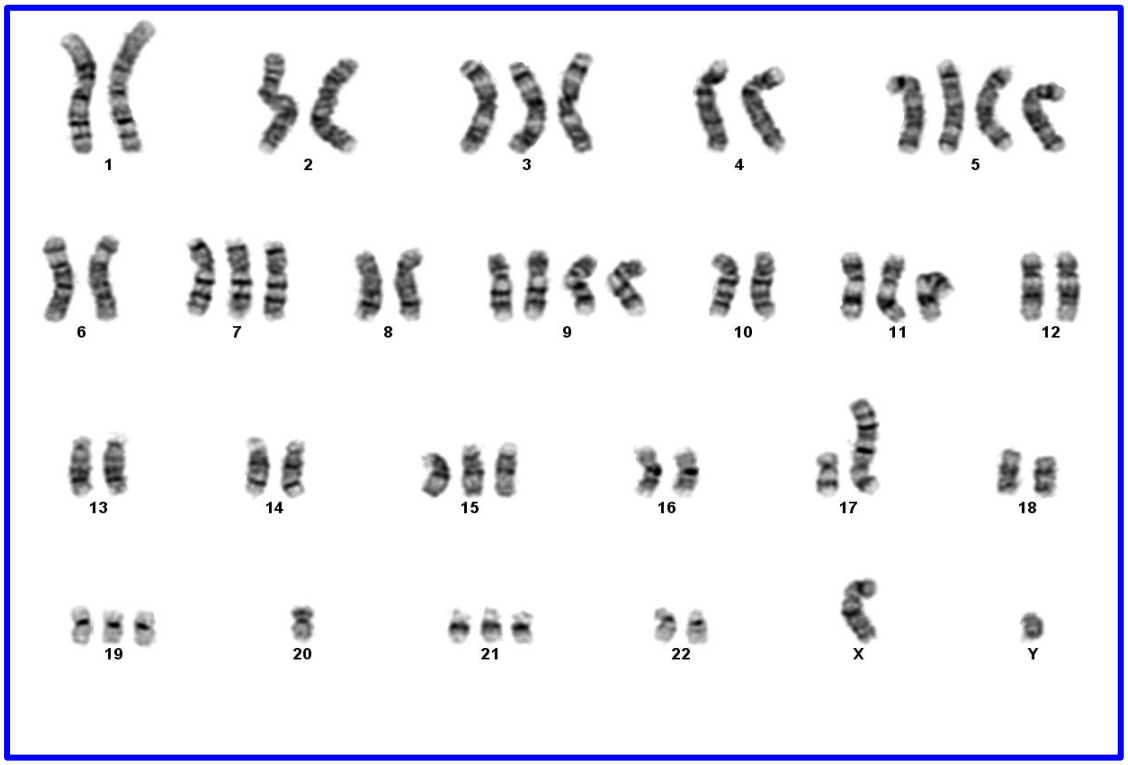


Figure 3.29: Demonstrating a representative karyotype from patient #175. This shows a hyperdiploid karyotype with a single additional copy of chromosomes 3, 7, 11, 15, 19 and 21, and two additional copies of chromosomes 5 and 9. There is also an unbalanced translocation between chromosomes 1 and 17. It should be noted that this cell also showed loss of chromosome 20 which was not shown to be clonal.

The ‘cp’ as part of the karyotype ISCN indicates that this is a composite karyotype made up of nine cells that were fully analysed. In order to be part of a composite karyotype and considered clonal, structural abnormalities or chromosomal gains must be present in two or more cells, and chromosome losses must be present in three or more cells. Three further abnormalities of chromosome 1 were seen in which the long arm of chromosome 1 was not present on the short arm of chromosome 17. The first cell was poor but suggested the presence of three copies of chromosome 1, a single copy of chromosome 2 and two copies of a derivative chromosome 2 from a t(1;2) in which the long arm of chromosome 1 appeared on the short arm of chromosome 2. The second cell appeared to have two additional copies of the long arm of chromosome 1q on an unknown chromosome, this could have been chromosome 19, but did not look like a straightforward deletion of the short arm of chromosome 1p. The final cell had a marker chromosome in which the long arm of chromosome 1 seemed to be involved. These changes are demonstrated in figure 3.30 and were seen with the same additional chromosomes associated with the patient’s hyperdiploidy, but with two normal copies of chromosome 17. As only individual cells were seen with these abnormalities, they could not be included in the composite karyotype. Although these abnormalities cannot be considered clonal, they do provide evidence for a level of fragility in the chromosomal make up resulting in this pattern of non-clonal changes. The movement of the long arm of chromosome 1q also raises the possibility of a ‘jumping translocation’. Jumping translocations, as referred to in the introduction, are rare chromosomal events in which the same donor chromosome segment is translocated onto two or more recipient chromosomal sites[121]. In this case the long arm of chromosome 1 is ‘jumping’ from the short arm of chromosome 17, to the short arm of chromosome 2, and to two further unknown, undetermined chromosome regions.

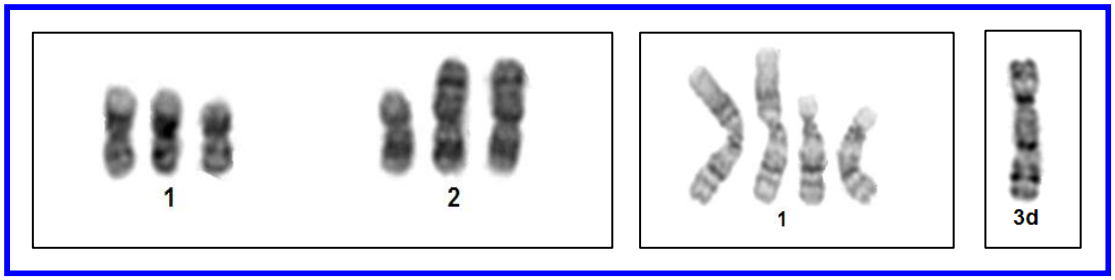


Figure 3.30: Partial karyotypes from three cells demonstrating a possible jumping translocation in which the long arm of chromosome 1 appears to be present on the short arm of chromosome 2 and present in two copies alongside three copies of chromosome 1 in the first cell, then in two further cells the long arm of chromosome 1 is present on two unrecognisable chromosomes, the first again present in duplicate. These unknown regions were considered to be chromosome 19 or 22 in the second cell and possibly chromosome 6 in the third cell. These cells also showed the same additional chromosomes as part of the hyperdiploid karyotype but were present with tow normal copies of the short arm of chromosome 17.

FISH analysis showed no evidence of rearrangement of the *IGH* gene, but demonstrated duplication of the *CKS1B* gene on the long arm of chromosome 1 in 72% of cells scored and a low level deletion of the *TP53* gene on the short arm of chromosome 17 in 8% of cells scored. The predominant karyotypic evidence of the unbalanced t(1;17) would suggest the same level of gain of chromosome 1q and loss of chromosome 17p. This is not the pattern seen by FISH, which provides evidence to support the likely presence of a jumping translocation involving the long arm of chromosome 1. The long arm of chromosome 1 is present in a number of different guises not all involving the short arm of chromosome 17. These rearrangements always result in the gain of 1q but not always the loss of the short arm of chromosome 17.

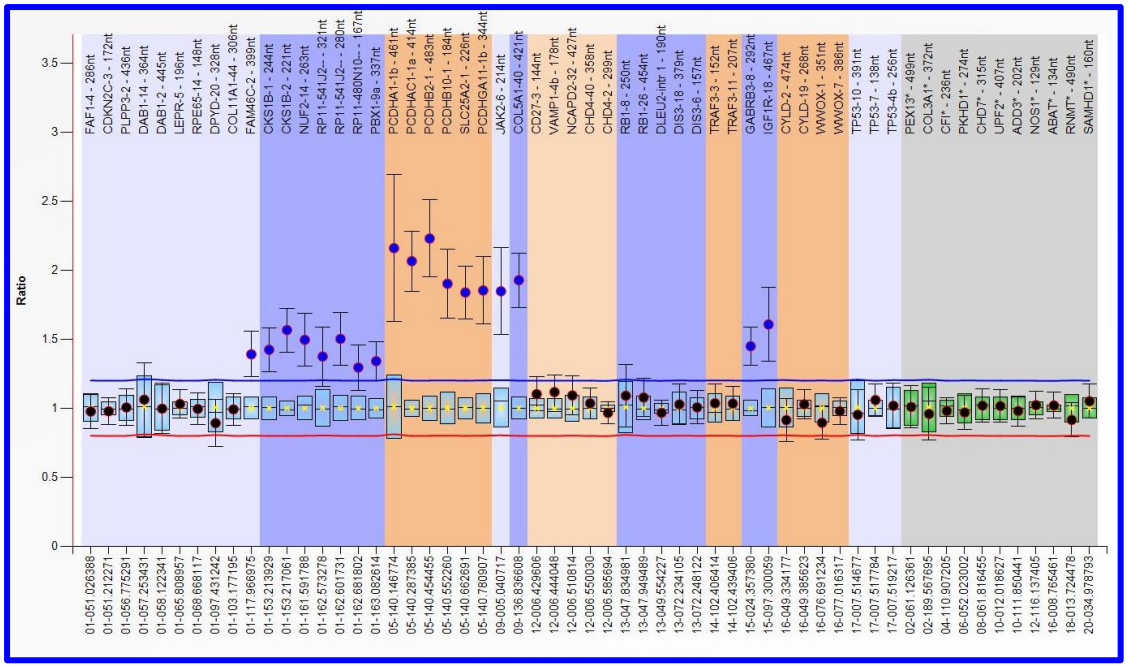


Figure 3.31: MLPA ratio chart demonstrating gain of regions of chromosomes 5, 9 and 15, the long arm of chromosome 1 and the marker in the *FAM46C* gene on the short arm of chromosome 1. It should be noted that the ratio levels are higher for both chromosomes 5 and 9 consistent with the presence of two additional copies of both chromosomes 5 and 9 seen karyotypically.

The MLPA analysis completed on this case is shown in figure 3.31, and shows gain of regions of 1q, chromosome 5, 9 and 15 and a single marker in the *FAM46C* gene on the short arm of chromosome 1p. The gain of chromosomes 5, 9 and 15 is reflected in the karyotype in which additional copies of these chromosomes are present. It should be noted that the ratio levels of the markers associated with chromosome 5 and 9 are increased in comparison to chromosome 15, confirming the presence of four copies each of chromosome 5 and 9 seen karyotypically compared to the three copies of chromosome 15. Gain of 1q and *FAM46C* are seen, but there is no evidence of loss of the short arm of chromosome 17 using the markers associated with *TP53*. Loss of *TP53* was seen at a level of 8% using FISH probes for the *TP53* region. Even though both FISH and MLPA analysis have been performed on CD138+ve cell selections, MLPA does have difficulty in detecting low level abnormalities (i.e. <15-20%), and in this case has shown no change in the regions associated with *TP53* on the short arm of chromosome 17, not even a slight deviation from the normal.

In conclusion, patient #175 demonstrates a plasma cell myeloma patient with a hyperdiploid karyotype. Hyperdiploidy is seen is approaching 50% of myeloma

patients. This patient also shows an unbalanced rearrangement of chromosomes 1 and 17, which results in a loss of the short arm of chromosome 17 and the long arm of chromosome 1, both of which are considered poor risk markers according to Boyd *et al* [112], and would be classified as a 'double hit' myeloma in this context.

Although this abnormality was the only clonal abnormality of chromosome 1 detected by karyotyping, there was some non-clonal evidence of other abnormalities involving chromosome 1q which was considered likely to be a 'jumping translocation' involving 1q. The FISH pattern showed gain of *CKS1B* on the long arm of chromosome 1 at a high level, but loss of *TP53* in only 8% providing further evidence for the presence of a jumping translocation. MLPA also confirmed the hyperdiploidy and gain of chromosome 1q, but did not show any evidence of *TP53* loss, highlighting one of the downfalls of this technique, i.e. the inability to detect low level rearrangements.

From a clinical perspective, this patient would have demonstrated a 1q gain only using standard of care genetic diagnosis, although this is currently considered a poor prognostic marker it does not change clinical management. Jumping translocations are well documented, but a rare and interesting finding, however, this abnormality has no real clinical impact. This patient is currently in a good remission following two rounds of treatment and two autologous stem cell transplants.

Patient #233 – Hypermutation detected by NGS in myeloma patient

Clinically, patient #233 is a 72yr old female who was diagnosed with smouldering myeloma in August 2016. At this time point she had an IgG kappa monoclone of 21.1g/L, her creatinine was 64, haemoglobin (Hb) 110, and her adjusted Calcium (adCa) level (this is the value corrected against albumin levels) was 2.4mmol/L. Her bone marrow plasma cells were at 45%. Cytogenetics showed a normal female karyotype, and FISH demonstrated gain of *CKS1B* on chromosome 1q. Array analysis subsequently showed a hyperdiploid karyotype with whole chromosome gains of 5, 9, 11, 15 and 19, with regions of gain seen on chromosome X and confirmed the FISH finding of gain of the long arm of chromosome 1. Regions of loss were seen on the

short arm of chromosome 1, not including the *CDKN2C* and therefore not seen by FISH, and the whole of the long arm of chromosome 16. Her sample was not processed for MLPA.

She was subsequently diagnosed with lung cancer following referral for a lump noted on the scan in October 2016. Following this diagnosis, the patient had a lumpectomy in February of 2017, but she did not receive chemotherapy at this stage. Therefore, at the time of the sample assessed as part of this study (August 2016), the patient had had no previous exposure to chemotherapeutic agents. In May 2019, the patient's myeloma demonstrated signs of progression with her Hb dropping to 88, and the requirement for a vertebraplasty procedure in response to skeletal damage detected at follow up scan. Treatment with Velcade, Thalidomide and Dexamethasone (VTD) was commenced at the end of May 2019.

Patient #233 demonstrates an anomaly in the NGS data and showed dramatically more variants compared to the remaining samples in the patient cohort. Figure 3.32 shows the number of somatic variants seen in all 24 patient samples. The variant number in patient #233 was approaching 150 in comparison to less than 10 variants in the remaining patients in the cohort.

Following the step of removing duplicates in the NGS data, the variants in patient #233 numbered 89, this compared to an average of 1.32 variants (range 0-4) in the remaining 22 patients (one patient failed NGS).

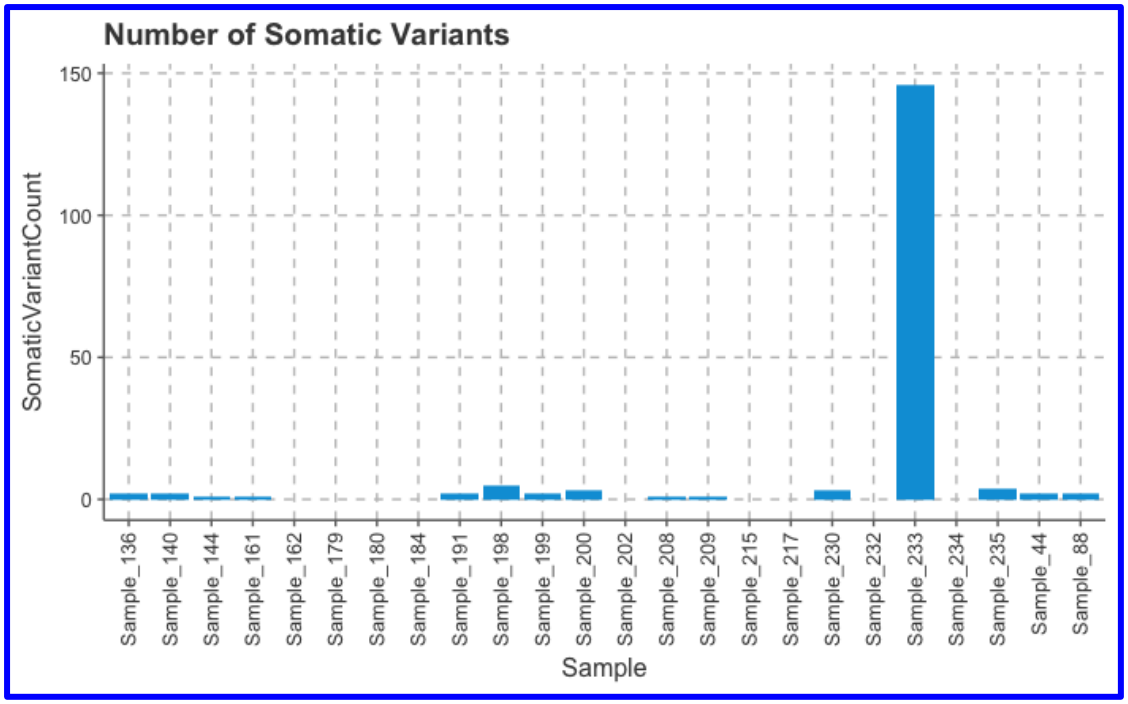


Figure 3.32: Graph to demonstrate the number of somatic variants detected across 24 patients in the NGS patient cohort using the bespoke myeloma / osteome NGS panel. This shows a low level of variants across 23 patients, but a dramatically increased level in patient #233

Following the pathogenic assessment of the variants, the variants likely to result in a protein change associated with a high level of pathogenicity were presented in table 3.11 for patient #233 and in table 3.10 for the remaining patients in the cohort. There were 45 pathogenic variants seen in patient #233, which is a significant increase based on the other 22 patients who showed an average of 0.91 (range 0-3) potentially pathogenic variants. It is acknowledged that the cohort number for the NGS processing is small, however, the number of mutations across the cohort is extremely consistent with the exception of patient #233.

Hypermutation is the process of producing an unusually high number of mutations or changes[122]. Patient #233 can be described as hypermutated.

There are number of explanations for this anomaly including the NGS processing and analysis itself. A remarkably similar case was reported in a myeloma patient assessed as part of the 100K genome project. However, when explored in more depth, the team involved with this patient considered that the high level of mutations detected were artefactual based on the poor quality of the starting material, as many variants were seen at extremely low allele frequencies. Based on this assessment we

analysed our patient; the patient sample quality was good and the sequencing quality metrics were also considered good for this patient. We assessed the allele frequencies for the somatic variants detected in this patient, which are shown graphically in figure 3.33. Although the graph can be seen to have a range of allele frequencies, the majority of the somatic mutations are present in approximately 50% of the sequencing reads. It is difficult to be sure if this is a true finding or an artefactual finding based on the bioinformatics alone, but it is unlikely that the mutations would be seen at such a high level if this result was not true.

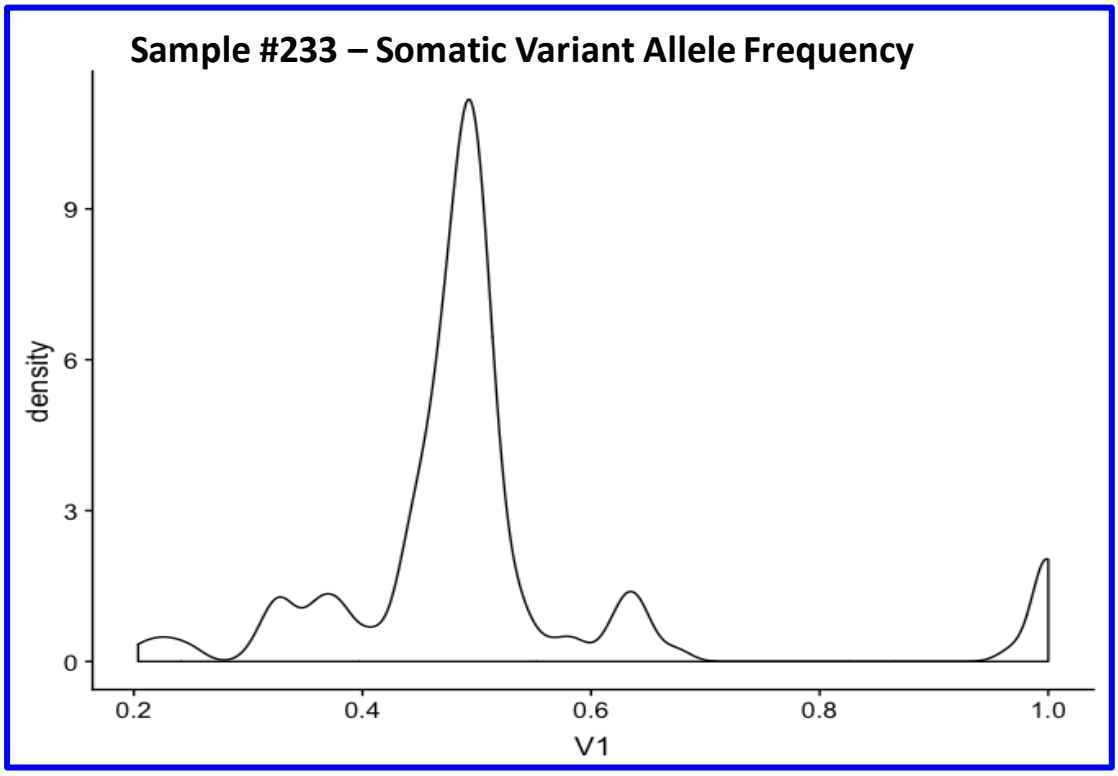


Figure 3.33: Graph to demonstrate the allele frequencies for the somatic variants detected in patient sample #233, showing an allele frequency peak at approximately 50%.

Patient #233 is known to have lung cancer, and exposure to the types of chemotherapy used for lung cancer could cause this picture of mutation. However, revisiting the patient notes and chemotherapy records, confirmed that the patient's initial lung lump was seen in October 2016 whilst the patient's myeloma treatment was ongoing, but the lumpectomy was performed in February 2017. At the time of sampling (August 2016), there was no record of chemotherapy for either the lung

cancer or the myeloma, or indeed any other malignancy, and therefore this pattern of mutation cannot be related to chemotherapy.

Having excluded treatment response or poor-quality sequencing data as reasons for the hypermutation, this has to be considered a true finding for patient #233.

The mutation count has been shown to be highly dependent on the tumour type, and based on a WES study from Chapman *et al*, average numbers of gene mutation have been shown to range from eight mutations in the more simple leukaemia's to approximately 540 mutations in the genetically more complex epithelial tumours[59]. Chapman *et al* demonstrated that myeloma patients fell between these two extremes with 35 being the average number of gene mutations per patient. In general, only modest variations in the number of somatic mutations were seen in the majority of patients within a tumour type, however, there is evidence to suggest that outliers exist within many cancer types in which dramatically higher mutation levels are seen, and this has been described as hypermutation[19,122].

Hypermutation has been shown to be caused by environmental factors such as UV light, particularly in association with malignant melanoma, and tobacco smoke, particularly in association with lung and larynx tumours. Mutations in genes associated with DNA replication repair (*POLE* and *POLD1*) or DNA mismatch repair (*MLH1, MSH2, MSH6* and *PMS2*) have also been associated with specific cancer types[122]. It may have been prudent to assess these genes in our patient, however, additional material was not available to complete this testing.

In conclusion, patient #233 has myeloma with a hyperdiploid karyotype with a gain of the long arm of chromosome 1. Although hyperdiploidy is usually considered a standard risk feature, gain of 1q is a high-risk marker. This patient was processed through the bespoke NGS panel designed as part of this project and showed a hypermutated profile with an approximately 45-fold increase in the number of variants seen when compared to the remaining patients in the cohort. Further examination, and exclusion of sequencing artefacts and treatment causes, suggested this was likely to be a true finding. Hypermutation is a rare finding but these outliers with dramatically increased variant numbers have been described in a number of cancer types. Potential causes of hypermutation were discussed.

Although the expectation from this genetic profile would be a poor outlook, this patient continues to be treated for both her myeloma and concurrent lung cancer and is clinically well.

Proposed Diagnostic Strategies and Best Practice Recommendations

In conclusion, the five genetic technologies described and explored here for the genomic assessment of myeloma offer range of both advantages and disadvantages, all of which should be considered in the context of providing a diagnostic testing strategy for NHS patients, which has to balance the clinical utility of the information provided with the time and cost of testing.

Traditional cytogenetics i.e. karyotyping offers a full genomic screen which allows the simultaneous assessment of copy number, including hyperdiploidy, and structural rearrangements specifically for the *IGH* gene on chromosome 14, however the resolution is low, a high level of skill and expertise is required and it is time consuming to carry out the level of analysis required which all contribute to the slow and expensive nature of this technology. Added to this, it is notoriously difficult technically, to capture plasma cells at the metaphase stage of division, analysis is therefore not targeted to the plasma cells and abnormalities are only reported in approx. 30% of patients in the literature. Although, when the abnormal cells are captured, much information can be gleaned from this analysis, the fact that it is only applicable to such a small proportion of patients, excludes the possibility of detecting smaller abnormalities including deletions, and cannot detect mutations effectively exclude this technology as a robust method for assessing the genetics of myeloma patients.

FISH can be targeted to interphase cells, and specifically can be used on samples that have been preselected to enhance the number of plasma cells, overcoming many of the difficulties seen with cytogenetic analysis. Many more cells can be reasonably analysed when compared to karyotyping, 50-200 cells not being an unreasonable analysis level, and probes are available that can assess both copy number and rearrangement detection covering the main abnormalities seen in myeloma. FISH does have limitations such as probe size and false positive rates. Probe sizes range from 150kb-1Mb, and therefore small deletions or mutations within genes or within the probe target cannot be detected. Specific probes have an associated false positive rate, dependant on the probe set formation. This is often low and quoted as 1-4% by the manufacturers. FISH probes are designed to specific regions, and

although probes are available to cover most abnormalities required in myeloma, each requires a specific probe and a single assay. Myeloma panels can, therefore, be extremely large. FISH processing is not conducive to high throughput processing, and the expense of the probes and the time-consuming nature of the analysis can be prohibitive. However, FISH is currently the most widely used methodology in the genetic assessment of myeloma and is currently considered gold standard. Much of the trials data has been generated using FISH and therefore the literature supports the use of FISH also.

MLPA offers an excellent multiplex methodology, providing information on the loss and gain of multiple regions (up to 50 targets), 46 in the case of the myeloma probe set from MRC Holland in a single assay. The probes employed in this technology are around 60 nucleotides in length and therefore deletions of a single exon can be detected, much smaller than conventional FISH techniques. The technique can also detect mutations in the targeted regions but is unable to offer distinction between deletion and mutation. Due to its multiplex nature it is cost efficient. MLPA has previously been described in studies of multiple myeloma with excellent results[118], however, MLPA is often considered a difficult technique and is highly sensitive to contaminants in the system. It requires a high level of technical skill and good quality DNA, and the failure rate is considered quite high, therefore a strategy including MLPA would have to have contingency in-built to cover failed samples. MLPA is also a copy number detection technology only and therefore only covers a subset of those abnormalities seen in myeloma, so there is the requirement of being used in conjunction with a technology able to determine *IGH* rearrangement[118].

DNA arrays offer a high resolution copy number whole genome screen which can be targeted at DNA from the CD138+ve plasma cell selections. The array has background resolution across the whole genome, but has particularly good coverage (25 markers per 100kb region) over key areas which include 533 cancer genes, and in the case of the array employed here, includes LOH detection. Balanced translocations, rearrangements and mutations are not reliably detected so again there is the requirement of being used in conjunction with a technology able to determine *IGH* rearrangement[113]. Low level abnormalities are also not reliably

detected (below the level of approximately 30%). Array processing is time consuming and expensive, and it is not considered a high throughput technology. The analysis is also time consuming, and the data provided were highly detailed, but in fact provides substantially more information than is currently clinical useful.

Next Generation Sequencing (NGS) can be applied in a number of ways from the assessment of the whole genome, whole exome or targeted panel testing. We assessed a bespoke panel of myeloma genes designed as part of this project. The panel performed well providing good, robust data for the genes chosen. The panel was able to show a number of abnormalities consistent with those reported in the literature, and in particular detected those mutations associated with the MAPK and NF-κB pathways known to be associated with myeloma is a proportion of cases. The NGS panel has to be considered as additional testing as it does not provide information on copy number changes or rearrangement detection. As NGS becomes more sophisticated, it can be envisaged that an NGS technology could be designed using both DNA and RNA to assess copy number and rearrangement detection for those changes commonly seen in myeloma alongside a specific panel of known variants / mutations. There are currently difficulties associated with detecting *IGH* rearrangements using some types of NGS, including targeted single nucleotide variant panel testing, whole exome sequencing and indeed RNAseq, however WGS even at low level coverage could pick these rearrangements up.

Table 3.12 shows a tabulated comparison of these technologies across a range of parameters from sample required, type of abnormalities detected through to costs and failure rates.

	Whole Genome Screen?	Targeted to plasma cells	Resolution level	Copy No. (inc. hyperdip)	Rearrangement detected	Mutations (SNV)	Minimum sample	Cost	Time taken	Expertise	Failure rate	Additional information
Cytogenetics (karyotyping)	Yes	No, even long-term culturing cannot target plasma cells specifically	Poor (5-10Mb)	Yes	Yes It should be noted that t(4;14) is cytogenetically cryptic	No	BM - 0.5ml	Processing: cheap Analysis time and skill: expensive	Processing: 4-6 days Analysis: 1-6 hours per sample	High level skill and expertise required	5-15% fail to produce metaphases ~70% fail to show abnormality	Also requires cells to be in division (metaphase) for analysis). Small no. of cells analysed
FISH	No. FISH probes target specific regions	Yes, if CD138+ve cell selection undertaken	Medium (probe sizes range from 150kb-1Mb) Small deletions within a probe could be missed	Yes	Yes	No	BM - 0.5ml	Processing: expensive for large probe panels Analysis time and skill: expensive	Processing: 1-2 days Analysis: 1-2 hours (depending on panel size)	SOP based, and even complex analysis can be learnt relatively fast.	Very low 0.07% in this project	Can be targeted to interphase cells. 50-200 cells can be analysed with ease. Much trial data relates to FISH
MLPA	No. Assesses 10 regions of interest	Yes, if CD138+ve cell selection undertaken	Excellent - deletions of a single exon can be detected Abnormalities at <20% may not be detected	Yes	No	No. Distinction between mutation & deletion not possible	2.5µl DNA (50ng/µl)	Cost efficient when batched	Processing: 3 days Analysis: ~20mins per sample	Technique highly sensitive to contaminants. Analysis software easy to use.	Very high: approx 35% in this study	Requires high quality DNA
Array	Yes	Yes, if CD138+ve cell selection undertaken	Good (25 markers / 100kb in cancer genes) Abnormalities at <30% may not be detected	Yes	No. Some evidence of copy number changes around breakpoints, but not robust	No	5µl DNA (50ng/µl)	Processing: expensive Analysis time and skill: expensive	Processing: 3-4 days Analysis: 1-2 hours per sample	Methodology is involved but basic techniques. Analysis is complex even with software	High: approx 10% in this study	Can detect regions of LOH. More data provided than clinically useful.
NGS (Panel sequencing)	No. 139 genes assessed in this bespoke panel	Yes, if CD138+ve cell selection undertaken	Excellent - single base changes can be detected	No	No	Yes	32-200ŋg DNA	Processing: expensive Analysis time and skill: expensive	Processing: 3-4 days Analysis: 1-2 hours per sample	High level skill and expertise required. Requirement for bioinformatics support for most.	2-5%	Requires both tumour and germline DNA samples

Table 3.12: Tabulated comparison of the five techniques undertaken as part of this study.

Using the information gathered through the assessment of these five technologies in the context of detecting genetic abnormalities associated with myeloma and understanding the pros and cons of what each technology can offer, we wanted to propose an appropriate diagnostic strategy. Currently, I would consider that a single technology cannot provide comprehensive analysis of myeloma genetics, however, FISH and MLPA technologies used in tandem provide both a comprehensive and cost-effective genetic testing strategy for myeloma. With regard to future proofing, there is likely to be a need for mutation analysis in addition to FISH and MLPA. I would also consider that as technologies advance and mature, NGS may be able to offer a dual DNA and RNA solution that can assess copy number, both ploidy and gains and losses known to have clinical significance, a number of rearrangements and the mutational spectrum in a single assay, creating a simple bespoke myeloma genetic assessment tool. Assessment of *IGH* rearrangement may remain technically difficult for NGS based analysis, so FISH is likely to play a role in diagnostic myeloma testing strategies in the foreseeable future.

As well as considering appropriate technologies for assessment, we need to consider the targets to include in the assessment. Currently NICE (National Institute for Clinical Excellence) guidelines for myeloma[111], the revised International Staging System for myeloma[114] and the IMWG (International Myeloma Working Group) consensus for risk stratification[116], state the importance of genetics in providing powerful prognostic and potential treatment guiding information. Based on these it should be considered that the targets include those regions with prognostic or therapeutic importance. These could be divided into mandatory regions for testing; including those regions prescribed by NICE and the IMWG, and those used to risk stratify patients. Other regions could be considered recommended where the information does not direct treatment, but provides a level of prognostic information, and then finally regions that are not currently required for testing. My views on this are described below:

Mandatory regions for testing:

Adverse *IGH* rearrangements

Specifically this relates to the adverse *IGH* rearrangements seen in >5% of cases i.e. t(4;14) *IGH-FGFR3* and t(14;16) *IGH-MAF*. If an *IGH* breakapart probe is used, then the sample must be further assessed (either simultaneously or sequentially) for both *IGH-FGFR3* and *IGH-MAF*. *IGH/FGFR3* dual colour dual fusion probes can also be considered for the front-line screen. The *IGH-MAFB* rearrangement is not currently considered in this group due its very low frequency, and I have included this in the 'recommended' regions for testing.

TP53 deletion

Deletion of chromosome 17p (including the *TP53* gene) remains the most significant prognostic marker and must be included as part of a genetic testing strategy.

1q gain and 1p loss

Both 1q gain and 1p loss are prognostically significant, and I, therefore, consider that they should be included as part of a diagnostic panel. The adverse prognostic impact of 1q amplification (i.e. >4 copies) has been reported to be stronger than gain of 1q (i.e. 3 copies) and therefore, the number of additional copies of the long arm of chromosome 1 should be reported. Loss of 1p is currently considered to be poor prognosis but less powerful than the prognosis associated with 1q gain/amplification. However, the FISH testing for these regions utilises the same probe in the majority of cases and therefore can be considered together when testing.

Recommended regions for testing:

Other *IGH* rearrangements

An extended panel to assess *IGH* rearrangement partners may include testing for t(11;14) *IGH-CCND1*, t(14;20) *IGH-MAFB* and t(6;14) *IGH-CCND3*. *IGH-MAFB* is considered a high risk *IGH* rearrangement but is seen at a very low frequency (1-2%) in myeloma patients.

Hyperdiploidy

Hyperdiploidy is generally considered a standard risk prognostic marker and can be confidently assessed using a 3-chromosome combination of probes (chromosomes 5, 9 & 15). The presence of two of the three chromosomes is a highly specific indicator of hyperdiploidy.

***MYC* rearrangement**

MYC rearrangements have recently been shown to be a marker of poor prognosis irrespective of the fusion partner, and could be considered as part of a genetic testing panel[38,49]. *MYC* rearrangements are seen in 15% of diagnostic patients and are associated with disease progression. Recent UK myeloma trials suggest that *MYC* rearrangement can be considered as strongly prognostic as *TP53* deletion, and based on this, I would suggest that this region is likely to become a mandatory region for testing in the near future.

Regions of testing that are not currently required:

Chromosome 13 deletion

Chromosome 13 deletion / monosomy, and loss of other loci such as 12p, 8p, 14q, 16q and 22q are not considered to be prognostically or therapeutically useful and are, therefore, not required to be part of a genetic testing strategy. It should be noted, however, that a number of these regions are included in the commercially available myeloma MLPA assay.

Gene mutation

Mutational assessment in myeloma currently remains in the research / trials arena. A number of gene mutations have been associated with favourable and negative impacts on survival, and others may confer specific treatment regimens. I would consider that this area of testing is likely to change in the very near future and is likely to be considered as part of the recommended or mandatory regions for testing.

In 2018, NHS England reconfigured the genomics services creating seven genomics laboratory hubs (GLH) across England resulting from a collaboration of a number of the existing NHS diagnostic testing laboratories. This aim of this reconfiguration, as well as efficiency savings and financial savings is to introduce an element of consistency to the diagnostic genomics testing across England to remove the so called 'postcode lottery'. As part of this reconfiguration a Genomics Test Directory has been produced which details the genomic testing that NHSE consider should be available, and in turn what they will consider funding.

Myeloma genomics testing is recognised as part of the new NHSE test directory and covers an extensive list of abnormalities for assessment. Detailed below are the suggested regions for testing for myeloma from the most current Test Directory published in January 2019:

FISH/RT-PCR for rearrangement detection of:

IGH-FGFR3 t(4;14)

IGH-CCND3 t(6;14)

IGH-CCND1 t(11;14)

IGH-MAF t(14;16)

IGH-MAFB t(14;20)

The test directory states that these abnormalities could also be detected using an NGS fusion panel even though *IGH* rearrangement is technically difficult using NGS panels. There is also a FISH test listed for *IGH* rearrangement testing to cover the use of a breakapart probe.

FISH for copy number changes of:

Hyperdidploidy

CDKN2C (deletion of 1p)

CKS1B (gain of 1q)

TP53 deletion

The test directory currently makes no reference to MLPA or indeed array analysis as an alternative methodology for copy number detection, however recent discussions have concluded that other technologies considered equivalent are valid alternatives.

Multi-target NGS panel for:

KRAS, NRAS, BRAF, TP53, DIS3, FAM46C & IRF4

The test directory has been contributed to by a range of personnel both as experts in their field, and scientists and clinicians from both diagnostic laboratory and research settings. However, it is also acknowledged that there are errors and inconsistencies in the test directory and that it has not been through the rigorous assessment initially hoped for. It is recognised that the test directory will 'direct' the testing that will be available and more importantly determine testing that will be funded from NHSE. In the last few months, three 'expert working groups' have been set up to review and manage changes to the test directories. These three groups will cover one each of the three test directory areas; rare disease, cancer and a newly formed pharmacogenomics section. I am now part of this cancer expert working group.

These changes can be seen as removing a degree of autonomy over the testing strategies employed by different laboratories. The test directory proposals are not entirely consistent with the proposals suggested here, although it is mostly consistent. The test directory tests are considered mandatory in part, although NHSE recognise that laboratories are not currently reaching the levels set. The test directory goes as far as including a mutation panel for myeloma when this is not employed by any diagnostic centres in the UK. There is no real consensus for the genes to be included in a mutation panel, although a panel of genes has been proposed by Walker *et al* which would include 13 significantly mutated genes which included, *KRAS, NRAS, TRAF3, TP53, FAM46C, DIS3, BRAF, LTB, CYLD, RB1, HIST1H1E, IRF4* and *MAX*[38], the test directory includes a set of seven genes excluding some genes that have been associated with a poorer outcome. The test directory also covers the five *IGH* rearrangements irrespective of their prognosis or their frequency, but does not include assessment of the *MYC* gene which is known to be associated with disease progression.

With regard to the technologies and the targets for assessment that we have discussed here, the following process detailed in figure 3.34 would be an example of how to implement this proposed strategy:

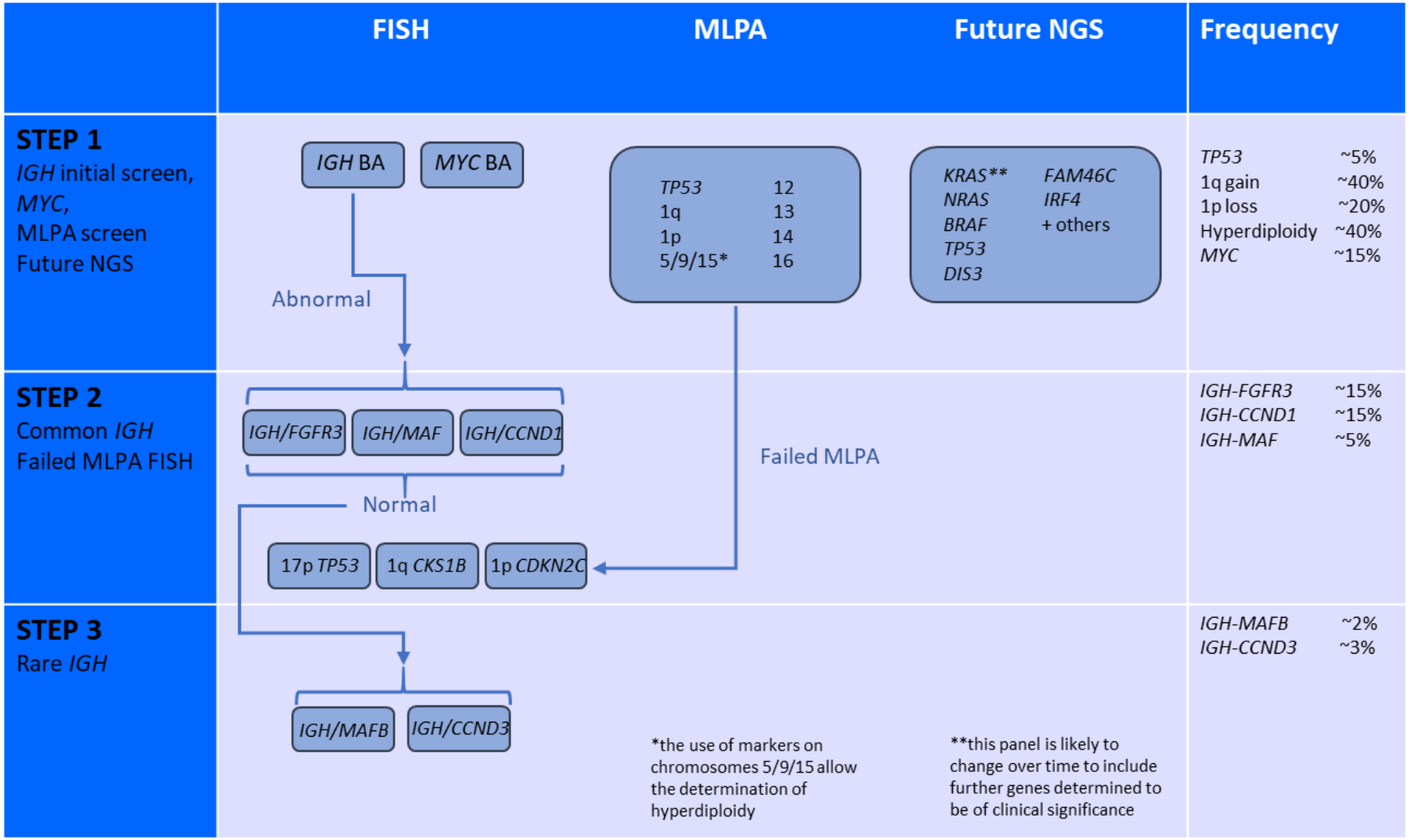


Figure 3.34: To describe a proposed genetic testing strategy for myeloma at diagnosis.

Whilst completing this work, and through my involvement with GenQA, I have been part of a group who have written and published 'European recommendations and quality assurance for cytogenomic analysis of haematological neoplasms' in Leukemia in 2019[123]. I specifically wrote the section devoted to myeloma. This raised the need for specific myeloma laboratory guidelines, which are currently being prepared and we hope to publish. These guidelines would cover the targets required as discussed in the general guidelines, but also cover in more detail the technologies that could be employed. There are disease specific concerns that could also be covered, such as whether to limit testing to myeloma patients or whether MGUS patients would also benefit, whether follow up genetic testing is appropriate or whether this would be better managed with flow cytometry, the age of patients to be tested, the requirement for testing specific cell selections, and understanding the consequences and how to manage reporting if this is not possible. The recommendations also cover how to report these, sometimes, complex results and discusses appropriate turnaround times and quality assessment.

Whilst the proposed testing strategy is appropriate for the current timepoint both from the perspective of the diagnostic laboratories and the way in which genomic information is used in the treatment and clinical management within UK guidelines, it is perhaps useful to consider a more forward thinking and future approach to the genomic diagnosis of myeloma. Gene expression signatures (sometimes known as the SKY92 signature based on the assessment of a 92 gene classifier developed using RNAseq data) have recently been shown to robustly identify a group of high-risk patients with myeloma[124]. The clinical validity was validated in two independent UK trial datasets as well as multiple other international datasets. Data from MRC Myeloma IX and XI furthermore demonstrated that the test identifies 10% of patients with short progression free and overall survival that were not identified by genetic tests described on the current testing strategy. This could inform treatment intensification decisions based on current trials but would require NICE approval to be included in standard clinical practice.

It could be considered that testing for these gene expression profiles (SKY92) alongside low coverage WGS and WES could provide information on *IGH*

rearrangements & copy number changes (from low coverage WGS), mutational profiling (from the WES or RNAseq) and high-risk gene expression profile (from the SKY92 expression data) would provide a future proof risk assessment genomic strategy. Although highly informative, this would require a step change for the diagnostic genetic laboratories, as currently neither gene expression profiling nor WGS is offered outside of the research realm.

In summary, this chapter describes the introduction and validation of a CD138+ve magnetic cell separation assay for all samples received from myeloma patients. The patient cohort included samples from 101 patients in total, 91 were karyotyped, 101 were FISHed using a small panel, 45 had MLPA analysis, 36 were processed using the Affymetrix DNA array. A bespoke NGS gene panel analysis was designed covering both a myeloma gene panel and the 'osteome' panel which covered a number of bone related genes, and 24 paired patient tumour sample and germline blood were assessed using this technology.

This chapter describes the results seen using the five different technologies and assesses the pros and cons for each technology. This information has been used alongside information taken from the literature and the new NHSE test directory, to devise a proposed genetic testing strategy for myeloma patients informing best practice recommendations. This demonstrates the importance of translational research in direct patient impact in the shaping of services offered to patients through our NHS system. The results pertaining to the osteome work will be presented in Chapter 4.

Chapter 4: NGS Analysis of the Osteome in Myeloma

Introduction

Bone disease is a key feature of plasma cell myeloma, and is seen in 67% of patients presenting with plasma cell myeloma, and up to 90% of patients with myeloma exhibit bone related complications at some stage of their disease[125,126].

Osteoclasts and osteoblasts are cells which work in tandem in the bone marrow and are responsible for the constant renewal and remodelling of the skeleton. Osteoclasts are responsible for resorbing and breaking down the bone matrix, and osteoblasts rebuild the bone matrix. Osteocytes are mature osteoblasts which become trapped in the matrix they have rebuilt. Neoplastic plasma cells secrete osteoclast activating factors and osteoblast deactivating factors, resulting in increased numbers and activity of the osteoclasts and a decrease in the number and activity of the osteoblasts. This imbalance between the osteoblasts and osteoclasts disrupts the normal balanced process of physiological bone remodelling, leading to accelerated osteoporosis and the development of lytic lesions and osteolytic bone disease. The resulting bone lesions are considered a marker of end organ damage, and therefore a criterion for the definition of a diagnosis of plasma cell myeloma[1,2].

Consequences of osteolytic bone disease are often seen as the presenting features of myeloma, and skeletal disease can manifest in a number of ways; this can be as generalised bone loss (osteoporosis) or focal lytic lesions. These can both lead to pathological fractures, spinal cord compression or collapse and hypercalcaemia. Less critical symptoms such as backache or bone pain are also seen. Myeloma bone disease can cause severe morbidity in myeloma patients leading to distressing pain, loss of mobility and disfigurement. A number of these lesions are shown in figure 4.1.

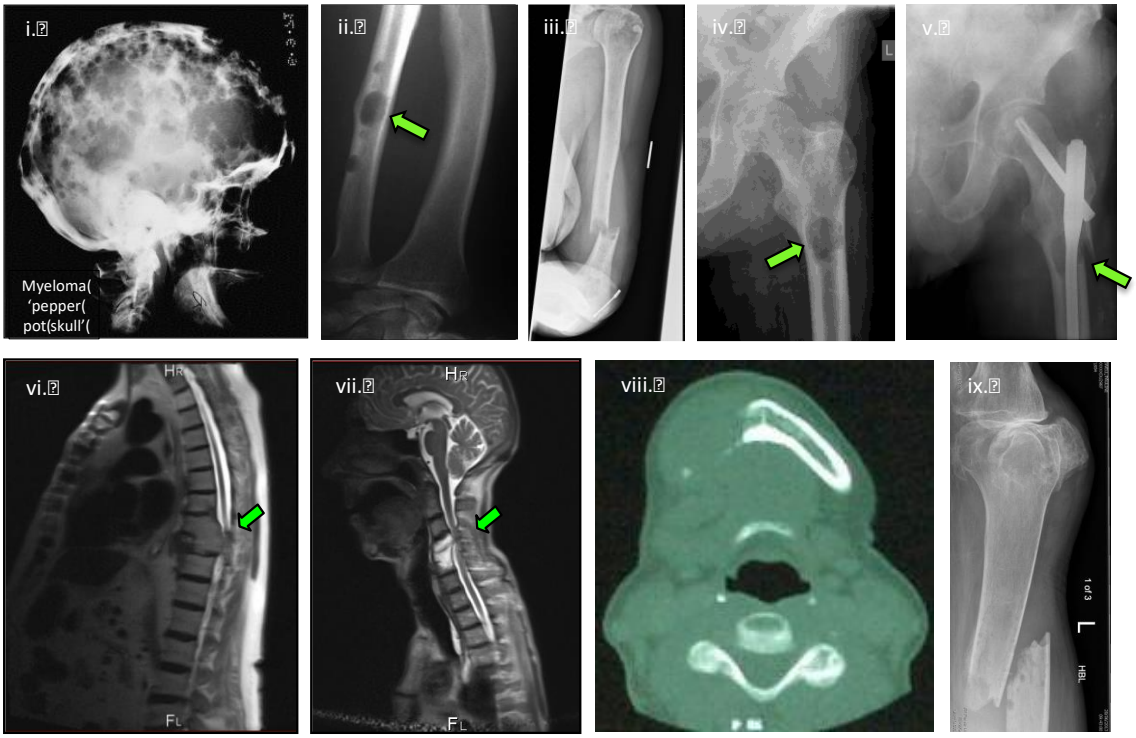


Figure 4.1: Radiological manifestations of myeloma bone disease. Image i) shows an X-ray of a classical 'pepper pot' skull in myeloma. Images ii) & iii) show X-ray images of lytic lesions and a pathological fracture respectively in the left arm. iv) & v) show a large lytic lesion in the patients left femur, which is then subsequently pinned. Images vi) and vii) demonstrate an osseous plasmacytoma leading to vertebrae destruction and spinal cord compression. Image viii) shows and MRI of a large plasmacytoma in the lower mandible destroying a large portion of the mandible and image ix) shows a fracture of the femur. Images used thanks to Dr Andrew Chantry.

The aim of treatment for myeloma can be considered threefold:

- To reduce the myeloma tumour burden or the bulk of disease
- To prevent and treat bone and tissue damage
- To improve quality of life and survival

With regard to the bone disease, treatment remains a substantial clinical problem with frequent catastrophic manifestations. Current strategies for managing and treating myeloma bone disease achieve some success in the prevention of subsequent skeletal events, but do virtually nothing to repair the damage that may already be present. Bisphosphonates are an important class of drugs, which are used to treat bone manifestations in myeloma, in fact myeloma bone disease treatment continues to rely almost exclusively on bisphosphonate therapy. Bisphosphonates are absorbed by macrophages (the cells from which osteoclasts are derived) and by mature osteoclasts. They are potent inducers of osteoclast apoptosis, thereby

reducing elevated bone resorption associated with myeloma. The most commonly used bisphosphonate in myeloma is zoledronic acid. A series of trials were undertaken using different bisphosphonates, the MRC Myeloma IX trial[127] showed significant benefit of treating with zoledronic acid, measured by the number of skeletal related events and overall survival. Bisphosphonates are also the most commonly prescribed drugs used to treat osteoporosis. There are advantages and disadvantages in using bisphosphonates; it is a relatively cheap treatment, but requires intravenous administration. It can cause renal problems; renal function is already compromised in myeloma patients so requires careful monitoring. There are side effects including flu like symptoms, bone, joint or muscle pain and a risk of osteonecrosis of the jaw (seen in 3.5% in MRC Myeloma IX trial)[127,128] which can be extremely difficult to manage clinically. There remains little consensus in the management of bisphosphonate treatment; some patients are treated early to prevent bone manifestations, some receive indefinite treatment based on the proven survival advantage, others have breaks in treatment coinciding with remission (quiescent disease states) and some receive very little treatment based on the side effect concerns[89].

Treatment of myeloma bone disease has relied almost exclusively on the use of bisphosphonates which work in an anti-resorptive manner (by inhibiting osteoclasts), but attention has begun to be focussed on anabolic therapies which act to promote bone formation through enhancement of osteoblastic activity. Anabolic therapies are widely used in osteoporosis, with a range of therapies available such as parathyroid hormone (PTH) and its analogues, romosozumab (a sclerostin neutralising antibody) and a number of newer, more experimental drugs[88,90]. Currently, these treatments are less frequently used for bone disease in myeloma, although it is likely that their use will impact myeloma management also.

A number of additional targets associated with signalling pathways including RANK/RANKL/OPG, WNT, TNF and Notch, and specific signalling molecules such as DKK1, sclerostin, periostin, osteopontin, TGFβ[84,85] and activin A[86] amongst others, known to be implicated in myeloma and its associated bone disease are also being explored[39,83].

A promising agent targeting the RANK/RANKL/OPG pathway is denosumab, which is able to prevent RANK activation, and subsequent osteoclast activation[39]. Inhibition of the deregulated WNT pathway is also demonstrating some encouraging results with the use of WNT pathway inhibitors such as anti-sclerostin[39], anti-DKK1[39,87].

Objective two of this thesis was to explore and assess the hypothesis that relationships exist between the genetic mutations associated with bone related genes (the osteome) and the likelihood or extent of myeloma related bone damage. The ability to highlight a cohort of patients who may be considered at risk of a severe bone phenotype associated with their myeloma, or to predict those patients likely to benefit from treatment, may allow the opportunity to treat more proactively with early or extended bisphosphonate treatment to slow the advancement of the bone manifestations, or may identify patients who would benefit from treatment with anabolic agents. It may also be possible to speculate that a more personalised medicine approach could be adopted matching specific novel anabolic agents to the abnormalities detected in the osteome panel.

Materials and Methods

The materials and methods for this chapter have previously been described in Chapter 2 as part of the Next Generation Sequencing section. Although this chapter aims to assess different questions and relates to the osteome genes included in the panel, the NGS panel was designed as a single entity and processed as a single panel, using the same cohort of patients and under exactly the same conditions.

The bone related genes included in the panel were derived from a process of literature review and discussion with the SMaRT (Sheffield Myeloma Research Team) as bone disease in myeloma patients is a particular interest of this group. A starting point for selecting genes to include in the osteome NGS panel was a review paper covering bone manifestations in myeloma published by Walker *et al*[89] from the SMaRT group.

SMaRT NGS Osteome Probe Panel									
ACVR1	AKT1	CCL3	CCL4	CCR1	CCR3	CDH2	CER1	CSF2	CTNNB1
DCN	DKK1	DKK2	FRZB	FZD1	FZD2	FZD3	FZD4	HGF	ICAM1
JAG1	KREMEN1	KREMEN2	LRP5	LRP6	MET	MMP9	NOG	NOTCH1	POSTN
RUNX2	SDC1	SFRP1	SFRP2	SFRP4	SMAD2	SMAD3	SMAD4	SMAD7	SOST
TAZ	TGFB1	TGFB2	TGFB3	TNF	TNFRSF11A	TNFRSF11B	TNFRSF13B	TNFSF11	TNFSF12
TNFSF13	TNFSF13B	VCAM1	VEGFA	WIF1	WNT10A	WNT10B	WNT3	WNT3A	WNT5A

Table 4.1: Detailing the genes included in the bespoke NGS panel covering the 'osteome' - bone related genes.

Walker *et al* describe a series of osteoclast activating factors, osteoblast inhibitory factors and adhesion molecules associated with myeloma bone disease[89]. A number of osteoclast activating factors have been described and identified including interleukin-6 / 1β / 3, macrophage inhibitory protein 1α (HUGO (Human Genome Organisation) gene name *CCL3*), tumour necrosis factor α (*TNF*), hepatocyte growth factor (*HGF*) and vascular endothelial growth factor (*VEGFA*). CCL3 is a chemotactic cytokine which can attract amongst other cells, osteoclasts and induces activity via ATK and MAPK pathways. TNF is potent upregulator of interleukin 6, it is able to induce osteoclast development in conjunction with RANKL and can prevent osteoblast maturation via downregulation of *RUNX2* and *TAZ*. Osteoblasts produce interleukin 3 and granulocyte-macrophage colony stimulating factor (*CSF2*). IL3 is thought to increase osteoclastogenesis through interaction with activin A (*ACVR1*). CCL4 is also a chemotactic cytokine closely related to CCL3; the *CCL4* gene produces the macrophage inhibitory protein 1β. These cytokines work through interaction with a number of chemokine receptors including CCR1 and CCR3[89]. Many of these activating factors are associated with the NFκB pathway, specifically the receptor activator of NFκB (*RANK*, also known as *TNFRSF11A*), its ligand (*RANKL*, or *TNFSF11*) and its inhibitor, osteoprogerin (*TNFRSF11B*). These genes as well as other members of the TNF family of genes; *TNFRSF13B, TNFSF12, TNFSF13* & *TNFSF13B* were also included in the gene panel.

Conversely, a series of factors have been shown to be osteoblast inhibitory factors, which include hepatocyte growth factor (*HGF*), transforming growth factor β1 (*TGFB1*) and the WNT signalling inhibitors; Dickkopf 1 (*DDK1*), soluble frizzled-related protein 3 (*FRZB*) and sclerostin (*SOST*). These have been shown to be produced by bone marrow stromal cells and some malignant plasma cells. Plasma cells also express the *HGF* receptor (*MET*) which binds to *HGF* driving proliferation of malignant plasma cells and inhibition of apoptosis via the RAS pathway. *MET* also regulates osteoclast development and inhibition of osteoblast differentiation. There are two other closely related members of the transforming growth factor β family (*TGFB2* and *TGFB3*). These TGFB family proteins tightly regulate the SMAD proteins which act as transcription factors: *SMAD2, SMAD3, SMAD4* and *SMAD7*[89]. Decorin (*DCN*) is also known to interact directly with the TGFB1 protein, and all these genes were also included as part of the osteome panel.

The WNT signalling pathway influences osteoblastogenesis; the WNT proteins bind to cell surface receptors made up of a complex of LRP5/LRP6 and frizzled transmembrane proteins (*FZD1-4*), and this complex in turn induces a cellular cascade preventing phosphorylation of beta-catenin (*CTNNB1*) and preventing its breakdown. There are a number of endogenous inhibitors of this pathway which include the secreted frizzled proteins (*SFRP1, SFRP2 & SFRP4*), WNT inhibitory factor 1 (*WIF1*) and the dickkopfs (*DKK1* and *DKK2*). For the dickkopfs to work in an inhibitory fashion, *DKK1* protein binds *LRP6* and *KREMEN1 & KREMEN2* proteins to create a complex that causes internalisation of the WNT receptor. Other members of the WNT signalling pathway were included; *WNT10A, WNT10B, WNT3, WNT3A* and *WNT5A*. Cerberus (*CER1*) is a cytokine associated with the WNT pathway and has the effect of inhibiting the bone morphogenic protein (BMP) activity.

Malignant plasma cells are known to occupy the same bone marrow niches that normal quiescent state plasma cells occupy before differentiation. The normal interactions between the quiescent plasma cells and the bone marrow niches are likely to be mediated by chemotactic factors and adhesion molecules. Many of these factors have also been implicated in aspects of myelomagenesis, as well as disease proliferation, treatment resistance and, through their association

with osteoblast and osteoclast activity, myeloma-related osteolytic disease. These adhesion molecules include integrins (for example, vascular cell adhesion molecule 1, *VCAM1* and intercellular adhesion molecule 1, *ICAM1*) and syndecan 1 (*SDC1*), which have been shown to have a role in malignant plasma cell proliferation and cell adhesion-mediated drug resistance[129]. There are also studies assessing the role of other adhesion molecules such as Notch-1 (*NOTCH1*), Jagged-1 (*JAG1*) and N-cadherin (*CDH2*). Expression of these genes has been shown to be greatly increased in patients with myeloma and indeed MGUS, and their role in proliferation in early myelomagenesis has been suggested[89,130]. The serine-threonine protein kinase encoded by the *AKT1* gene is also known to have a role in cell migration and was also included in this gene panel.

Finally, three genes with specific bone or tissue related functions were included. Periostin is a protein encoded by the *POSTN* gene. Periostin binds to integrins which support adhesion and migration of epithelial cells. This protein plays a role in cancer, when periostin binds to the integrins on cancer cells, activation of specific pathways leads to increased cell survival and promotes invasion, metatasis and angiogenesis[131]. The *NOG* gene encodes the Noggin protein, which is involved in the development of many tissues, NOG is required for osteoclast accumulation in normal skeletal development. NOG has also been shown to mediate breast cancer metastatic bone colonization by osteoclast differentiation[132]. The *MMP9* gene encodes for one of the matrix metalloproteinase family. It has a normal function in tissue remodelling through the breakdown of the extracellular matrix. It has also been implicated in cancer pathology involved metastasis and tissue invasion[133].

DNA samples (from both CD138+ve cells from the bone marrow and from blood lymphocytes) from 24 patients were processed and analysed using the NGS panel. Of those patients, five patients were asymptomatic myeloma (aMM), 18 patients were myeloma, and one patient was a relapsed myeloma (rMM), as detailed in table 4.2. Twelve patients demonstrated bone lesions, and twelve patients did not.

Total	aMM	MM	rMM	No bone lesions	Bone lesions
24	5	18	1	12	12

Table 4.2: Demonstrating the samples processed and analysed using the NGS osteome panel. Twenty four patients were processed; five patients were asymptomatic myeloma (aMM), 18 were newly diagnosed myeloma (MM) and one patient had relapsed myeloma (rMM). The cohort was split evenly into those that had bone lesions and those that did not.

Table 4.3 details the type of bone lesions seen in those patients with bone lesions, and these range in number from one lesion to multiple and include lesions of the vertebrae, ribs, femur, hips and pelvis, as well as cord compression. The table also indicates the gender, ethnicity, diagnosis and the percentage of plasma cells recorded at the time of sampling.

Sample No:	Gender	Ethnicity	Diagnosis	No. of Bone Lesions	Lesions	Plasma cells
#136	Female	British	Myeloma	1	Wedge fracture T9	70%
#140	Male	British	Myeloma	3	Lesion at L5, T12 and in the 9th rib	15%
#144	Male	British	Myeloma	multiple	Multiple lytic lesions in the spine and pelvis	70%
#88	Male	British	Myeloma	0		40%
#161	Male	British	Myeloma	0		75%
#162	Female	British	Myeloma	1	Cord compression at T4	35%
#44	Female	British	rMM	2	Sternal lytic lesion and one in rib	56%
#179	Male	British	Myeloma	0	2 tiny queries on the R hip	35%
#180	Female	British	Myeloma	>5	3/5th rib. T5. L1 & L2 and pelvic bone	20%
#184	Female	British	Myeloma	2	?Small lucencies	15%
#191	Female	Caribbean	Myeloma	2		25%

#198	Male	British	Myeloma	0		70%
#200	Female	British	Myeloma	1	Left 8th rib plasmacytoma	50%
#199	Male	British	Myeloma	2	T1/T2 collapse	60%
#202	Male	British	Myeloma	Multiple	L3 / pelvis / ribs	20%
#209	Female	Caribbean	aMM	0		14%
#208	Female	British	Myeloma	0		12%
#215	Male	African	aMM	0		36%
#217	Male	British	Myeloma	0		24%
#230	Male	British	Myeloma	0		45%
#232	Male	British	aMM	0		17%
#233	Female	British	aMM	0		45%
#235	Male	British	aMM	0		42%
#234	Female	British	Myeloma	Multiple	2cm femur lesion	80%

Table 4.3: Samples processed using the osteome NGS panel with additional details of the gender, the ethnicity, the disease classification, and the percentage of plasma cells. The number and type of bone lesions seen are also detailed demonstrating an even split (twelve samples each) in the bone lesion vs non-bone lesion cohorts.

Results

As previously described, the NGS panel applied in this project was designed as a single panel of 139 genes, but is essentially split into two functional panels from an analytical point of view; the myeloma gene panel with 79 genes and the osteome bone gene panel with 60 genes. The samples were processed as pairs, DNA from the bone marrow CD138+ve cell selected sample to represent the tumour (myeloma) and DNA from a blood sample, which represents the patients' germline DNA. The analysis process is exactly as described in Chapter 3, in which a bioinformatics pipeline was employed to assess the variants seen and assign a likelihood of pathogenicity based on a consensus of the three main tools applied. Lists of both

somatic and germline variants were produced, which included further annotation with ClinVar and ExAC allele frequencies and could be divided into the two individual panels. Details relating to the performance and quality of the panel, such as the mapped and paired reads, coverage and gaps in coverage, are presented in Chapter 3. The variants associated with the osteome panel are presented here.

The first part of the analysis assessed the variants observed in the germline samples in order to pick up potential variants in bone genes that are part of the constitutional genetic make-up and therefore potentially providing a predisposition to bone related manifestations associated with myeloma. The numbers of variants called in the germline samples following analysis through the bioinformatics pipeline were large and additional manual selection analysis was required to prioritise variants for analysis. Table 4.4 details the number of variants at each stage of the process of variant selection. Column two shows the total number of germline variants called by the bioinformatic pipeline which includes single nucleotide variants (SNVs) events, that could be either synonymous or non-synonymous, and small insertion or deletion (indels) events. The average number of germline variants was 188.7 per patient (range 172-200). The first filtering step involved selecting variants in those genes included in the osteome panel and this was recorded in column three 'osteome only variant count'. The average number of germline variants seen within the osteome genes was 66.1 per patient (range 59-73). The next selection step excluded those variants seen outside of the coding regions of the osteome genes and the variant number was recorded in column four 'coding only'. The average number of variants seen in coding regions of the osteome panel genes was 18.3 per patient (range 10-26). The numbers of variants per individual at each stage was remarkably similar with a small range in number. Then finally the ExAC population frequencies were used to exclude those mutations seen at a high level in the population and therefore most likely to be considered polymorphisms. Population frequencies differ across different populations explaining the need to know the ethnicity of the patient. Additionally, ethnicity has a particular bearing on myeloma analysis as it is known to be a disease with a higher frequency in black people than in people of Asian or Caucasian descent (ratio ~2:1)[1,5]. The African/American ExAC population frequency

was used in the three patients (#191, #209 and #215 marked with an * in the ExAC frequency column) known to fit this ethnicity group. The average number of variants per patient following the full additional filtering is 3.0 (range 0-6).

Following this described series of filtering, the variants could then be classified according to their pathogenicity. This was based on the use of three bioinformatic tools designed to assess the likelihood of associated pathogenicity; mutect2, strelka and vardict. Those variants considered 'pathogenic', 'risk' or 'conflicting' (where the evidence seen included either 'pathogenic' or 'risk' information) have been highlighted in blue. Variants were also classified as 'benign' or 'likely benign', or as VUS (variant of unknown significance). NA (not applicable) was also used where there was not enough evidence to classify the variants, and in this scenario were considered to be equivalent to the VUS category.

Nine variants were classified as Pathogenic, Risk or Conflicting (described here as a group of risk variants) and these were seen across six genes. Three variants were described in the gene, *WNT10A*, two variants in *MET* and one variant in each of *ICAM1, FZD4, TGFB1* and *NOG*. These variants are detailed in table 4.5.

Patient	Total Variant Count	Osteome Only Variant Count	Coding Only	ExAC freq ≤0.05	Gene Involved	Pathogenicity classification
#44	193	66	13	1	*FRZB*	NA
#88	178	59	26	6	*FZD4*	Pathogenic
					TNFSF11	Likely Benign
					CCL3	NA
					CDH2	NA
					MMP9	Likely Benign
					MMP9	NA
#136	189	66	21	2	*DKK2*	NA
					SMAD3	Likely Benign
#140	183	63	10	3	*TNFRSF11A*	VUS
					ICAM1	Risk
					TGFB1	Benign
#144	186	64	17	4	*SDC1*	NA
					WNT10B	NA
					HGF	Likely Benign
					LRP5	Likely Benign
#161	194	69	19	2	*NOTCH1*	Benign
					POSTN	NA
#162	179	68	18	3	*LRP5*	Likely benign
					MMP9	Likely benign
					MMP9	NA
#179	191	67	13	1	*SDC1*	NA
#180	183	65	20	5	*TGFB2*	VUS
					WNT10A	VUS
					NOTCH1	NA
					CDH2	NA
					LRP5	Likely benign
#184	193	69	20	0		
#191	196	68	18	3*	*NOTCH1*	Benign
					LRP6	NA
					TNFRSF13B	NA
#198	200	71	18	4	*WNT10A*	VUS
					HGF	Likely benign
					NOTCH1	Likely benign
					POSTN	NA
#199	183	60	15	1	*CDH2*	NA
#200	180	62	19	3	*TGFB2*	Conflicting
					CDH2	NA
					ICAM1	NA

#202	189	69	22	5	*CCR3*	NA
					CCR3	NA
					FZD1	NA
					MET	Conflicting
					SOST	VUS
#208	200	73	20	2	*MET*	Likely benign
					FZD4	Likely benign
#209	198	65	19	2*	*MET*	NA
					TNFRSF11A	NA
#215	199	63	21	4*	*VEGFRA*	NA
					TNFSF13B	NA
					SMAD7	NA
					JAG1	Benign
#217	195	67	21	3	*FRZB*	NA
					WNT10A	Conflicting
					MET	Conflicting
#230	190	73	20	2	*LRP5*	Likely benign
					FZD4	Likely benign
#232	185	69	19	5	*WNT10A*	Conflicting
					NOG	Conflicting
					TGFB1	Benign
					POSTN	NA
					CDH2	NA
#233	172	61	18	6	*FRZB*	NA
					CDH2	NA
					WNT10A	Conflicting
					TGFB1	Benign
					FZD4	Benign
					NOTCH1	Benign
#234	Sample Failed NGS Analysis					
#235	185	64	15	2	*HGF*	Benign
					MET	VUS

Table 4.4: Detailing the germline results seen as part of the NGS osteome panel. Twenty-four samples were processed and 23 gave a result. Multiple variants were detected in the germline samples and these were manually filtered to give variants seen in the coding regions of the osteome genes with a population frequency of less than 0.05. The appropriate population frequency was applied, three patients were assessed using the African/American population frequency, marked with an *. The variants following this additional filtering were then assigned a pathogenicity classification using bioinformatic tools; Pathogenic, Risk, Conflicting, VUS (variant of unknown significance), Likely Benign, Benign and NA (not applicable). The NA group was used where there was not enough evidence to assign a specific category and for the purposes of this analysis these were considered in the same category as the VUSs. Those falling into the Pathogenic, Risk or Conflicting category were highlighted in blue

Of the 23 patients with results from the NGS panel, the nine variants classified as risk variants in the osteome genes were seen in seven patients (30.4%). A single patient showed no variants. Of this list, a number of potentially druggable targets were noted, including *TGFB1, TGFB2, DKK2, SDC1* and *HGF*, raising the potential as discussed previously of a more personalised medicine approach to treatment.

Of the seven patients with risk variants, two patients showed two variants and five patients showed a single variant and are detailed in table 4.5. Of these seven patients, five patients were classified as myeloma and two patients were asymptomatic myeloma. Three patients were part of the patient cohort with bone lesions and four were patients with no bone lesions.

Excluding those variants classified as 'conflicting' within the appropriate population frequency assessment, left a group of only two variants assessed as Pathogenic or Risk (considered here as high-risk variants). These two high risk variants were seen in two genes; *ICAM1* and *FZD4* and seen in two patients (#88 and #140). These patients were both classified as myeloma, but one patient belonged to the patient cohort with bone lesions and the second patient was part of the non-bone lesion cohort. These two high risk variants were assessed using the tools available through Alamut to provide a more in depth and what would be considered as a 'diagnostic' interpretation level; both are considered variants of unknown significance (VUS), once again demonstrating that the variant calling tools used within the bioinformatics pipeline err towards the more pathogenic end of the scale. The *FZD4* variant has a number of indications suggesting pathogenicity, but the *ICAM1* variant has a number of indications suggesting that the change could be considered a polymorphism.

Patient Number	Gene Involved	Pathogenicity classification	HGVS.c / HGVS.p	Number of Bone Lesions	Disease Classification
#88	*FZD4*	Pathogenic	c.766A>G p.Ile256Val	0	MM
#140	*ICAM1*	Risk	c.167A>T p.Lys56Met	3	MM
#200	*TGFB2*	Conflicting		1	MM
#202	*MET*	Conflicting		Multiple	MM
#217	*WNT10A*	Conflicting		0	MM
	MET	Conflicting			
#232	*WNT10A*	Conflicting		0	aMM
	NOG	Conflicting			
#233	*WNT10A*	Conflicting		0	aMM

Table 4.5: Detailing those patients with risk variants, the gene involved, the disease classification and the number of bone lesions seen in the patient (i.e. whether they belong to the bone lesion vs no bone lesions cohort).

For completeness, somatic changes seen in the osteome NGS panel genes were also assessed. These represent additional changes seen in the osteome present as part of the tumour, i.e. only seen in the DNA of the plasma cells. These changes are detailed in table 4.6.

Patient	Variant Numbers	Gene	Effect	HGVS.c / HGVS.p
#161	1	*TNF*	downstream gene variant	c.*4020A>G
#209	1	*NOTCH1*	synonymous variant	c.2703G>A p.Gly901Gly

Table 4.6: Detailing those patients with somatic variants detected in the osteome genes, the number of variants, the gene involved, the likely effect and HGVS nomenclature for both the DNA and protein change (where appropriate).

Two variants were detected, one in each of two patients in the cohort. The variants were seen in the genes *TNF* and *NOTCH1*. Neither variant is considered pathogenic; the *TNF* variant is a downstream variant resulting in no protein change, and the *NOTCH1* variant is synonymous, meaning that the base substitution results in no change to the protein, in this case a glycine is replaced with a glycine. Both patients #161 and #209 belong to the non-bone lesion cohort.

Discussion

Bone disease associated with myeloma is a key feature of the disease. It is associated with 67% of patients at diagnosis, but up to 90% of patients have some bone related complication through the course of their disease[125,126]. Treatment of these bone manifestations in myeloma remains a clinical problem. Bisphosphonates are an important class of drugs used for the treatment of myeloma bone disease, and although some success is seen in the prevention and protection from further skeletal events, little can be done to repair or reverse the already existing bone damage. This poses the question of whether bisphosphonates should or could be used as a preventative measure, i.e. treating patients before the bone disease is present or as early as possible following evidence of bone disease. However, although bisphosphonate treatment has the advantage of being cheap, it has a number of disadvantages including the requirement for intravenous administration, and side effects including renal damage (which is particularly problematical in myeloma patients who already have issues related to renal function), flu-like symptoms, bone, joint or muscle pain and a risk of osteonecrosis of the jaw which can be extremely difficult to manage clinically. Guidelines for the use of bisphosphonates are currently vague, and therefore patients can receive indefinite bisphosphonate treatment, but others receive very little, in part as a result of these disadvantages. Anabolic agents act to promote bone formation through enhancement of osteoblastic activity and have been shown to have a positive effect in osteoporotic and myeloma patients. However, although anabolic therapies are widely used in osteoporosis, with a range of therapies available such as parathyroid hormone (PTH) and its analogues, romosozumab (a sclerostin neutralising antibody) and a number of newer, more experimental drugs[88,90], currently, these treatments are less frequently used in myeloma management.

Objective two of this thesis was to explore and assess the hypothesis that relationships exist between the genetic mutations associated with bone related genes and the likelihood or extent of myeloma related bone damage. This proposed relationship could even go so far as to predict those patients who would be more likely to benefit from early or extended bisphosphonate treatment or the use of

more novel anabolic agents for their osteolytic disease. Highlighting a cohort of patients who may be considered at risk of a severe bone phenotype associated with their myeloma at diagnosis may offer the opportunity to treat more proactively and slow the advancement of the bone manifestations.

The genes included as part of the osteome NGS panel were derived from a process of research and discussion with the SMaRT (Sheffield Myeloma Research Team) as bone disease in myeloma patients is a particular interest of this research group. Although the NGS Myeloma and Osteome panels have been presented as separate panels for the purpose of data analysis and results discussion, the NGS panel was created as a single panel of 139 genes, 60 of these being genes associated with the osteome. Twenty-four patients were processed and analysed using the NGS panel. Of those patients, five patients were classified as asymptomatic myeloma (aMM), 18 patients had myeloma and one patient had relapsed myeloma (rMM). Twelve patients demonstrated bone lesions, and 12 patients did not. The bone lesions ranged in number from one lesion to multiple and include lesions of the vertebrae, ribs, femur, hips and pelvis, as well as cord compression.

The samples were processed as pairs, DNA from the bone marrow CD138+ve cell selected sample to represent the tumour (myeloma) and DNA from a blood sample, which represents the patient's germline sample. The quality metrics associated with the NGS run are reported and discussed in Chapter 3, as the osteome NGS panel was simply part of the full NGS panel the quality metrics for this part of the work are exactly the same as those reported previously. A single patient failed the NGS assessment completely, therefore results were available on 23/24 patients. This patient was part of the bone lesion cohort and showed multiple lesions including a 2cm femur lesion.

For analysis of the variants associated with the osteome panel, both the germline samples and the CD138+ selected plasma cells (the tumour samples) were assessed. Somatic changes of the osteome panel associated with the tumour plasma cells were assessed, based on the hypothesis that variants seen in these cells may affect genes and proteins directly involved in the interaction and communication between the plasma cells and the bone marrow niche, resulting in a more severe bone

phenotype. Only two somatic variants were detected involving the genes *TNF* and *NOTCH1,* one in each of two patients. Neither variant would be considered pathogenic; the change detected in *TNF* was a downstream variant, and the *NOTCH1* variant resulted in a synonymous change, so the protein consequences of the variants would be nil. Both variants were detected in patients in the non-bone phenotype cohort. Based on this experiment no association with somatic changes in the osteome panel was detected.

Germline changes in the osteome panel were assessed based on the hypothesis that variants in the genes as part of the patient's constitutional genetic make-up may predispose patients with myeloma to a more severe bone phenotype. As such high numbers of variants were called in the germline samples using the bioinformatics pipeline, further manual assessment of the variants was performed. Analysis was limited to those variants seen in the 60 osteome genes, only those seen in coding regions of these genes and then utilising the ExAC population frequencies, only those variants at a population frequency of <0.05% using the appropriate population frequency data. The average number of variants per patient following the full additional filtering is 3.0 (range 0-6).

The bioinformatic pipeline used a variety of tools to assess the pathogenicity. Those variants considered 'pathogenic', 'risk' or 'conflicting' were considered a risk group, and those where the variants were 'pathogenic' or 'risk' only were considered the high risk group. Variants were also classified as 'benign' or 'likely benign', or as VUS (variant of unknown significance) or NA (not applicable) where the variants did not have evidence to classify them as disease or predisposition causing, these were considered low risk variants.

Nine germline variants were classified as part of the risk group, these variants were seen in six genes, three variants in *WNT10A*, two variants in *MET* and one variant in each of *ICAM1, FZD4, TGFB1* and *NOG.* Of the 23 patients with NGS panel results, the nine variants classified as risk variants in the osteome genes were seen in seven patients (30.4%). A single patient showed no variants, and the remaining 15 patients showed low risk variants. Of the seven patients with risk variants, two patients showed two variants and five patients showed a single variant. Of these seven

patients, five patients were classified as myeloma and two patients were asymptomatic myeloma. Three patients were part of the patient cohort with bone lesions and four were patient with no bone lesions. Although it is recognised that the sample numbers are low, no association of the number of risk variants was seen with the cohort of patients with bone lesions.

The high-risk variants were also assessed to ensure that the variants with conflicting data were not skewing the analysis. Once appropriate populations frequencies were applied only two variants were classified as either 'pathogenic' or 'risk'. These two high risk variants were seen in two genes; *ICAM1* and *FZD4* and were seen in two patients (#88 and #140).

Plasma cells are part of the adaptive component of our immune system and are required to produce antibodies in response to antigenic assault. Normal quiescent plasma cells occupy niches in the bone marrow until this antigenic assault occurs. Interactions between the niche and the quiescent plasma cells are likely to be mediated by chemotactic factors and adhesion molecules. The *ICAM1* gene encodes an intercellular adhesion protein known to be part of these interactions. A number of adhesion molecules have been implicated in proliferation, bone disease and even treatment resistance. *ICAM1* is thought to be part of this group[129].

FZD4 is one of the frizzled transmembrane proteins that make up a cell surface complex that WNT proteins are able to bind to. It is part of a complex cascade event that influences the production of osteoblasts[89].

Both patients carrying the pathogenic variants were classified as myeloma, but one patient belonged to the patient cohort with bone lesions and the second patient was part of the non-bone lesion cohort. We were not able to demonstrate any association of these high-risk variants with the cohort of patients known to have bone lesions.

Again, based on the large numbers of variants, the assessment of pathogenic vs benign status for each variant was taken from the bioinformatic pipeline which utilised three tools (mutect2, vardict and strelka) to gather information on this status. Within a diagnostic setting, a more rigorous assessment is completed, and it

would have been an ideal scenario to have assessed each variant to that level, but within the project timescales this would not have been possible. A more in depth analysis akin to that utilised in diagnostic laboratories was undertaken on these two high risk variants using the tools available through Alamut, and both the *FZD4* and the *ICAM1* variants would be considered variants of unknown significance (VUS). The information for the *FZD4* variant would suggest it more likely to be pathogenic. The amino acid is highly conserved with evidence from ClinVar, dbSNP and Uniprot suggesting pathogenicity, but without enough evidence to definitively state it is a pathogenic variant. Conversely, the *ICAM1* variant involves a weakly conserved amino acid, with some evidence suggesting that this could be a polymorphism, but again without the strength of evidence to call this. This information can be found in appendix 6.

As for the somatic variant calling associated with the myeloma gene panel described in chapter 3, it seems that the pathogenicity calling tools incorporated in the bioinformatics pipeline show a leaning towards over calling the pathogenicity status, however, it does not exclude the possibility that important variants have been classified at a lower level or within the NA category based on insufficient information available within these three databases.

It is recognised that the number of samples used in this part of the study is small. We processed 48 samples which equates to 24 patients with both tumour bone marrow samples and control (germline) blood samples. These patients were split into two cohorts, those with bone lesions and those with no bone lesions resulting in 12 patients per cohort. We had one sample fail, belonging to the bone lesion cohort. A sample size of 24 has 71% power to detect the presence of a somatic variant which is present in the tumour population at a frequency of 5%. Thus, our observations of a number of unique variants in the group of 24 is consistent with this expectation. However, the power to detect a difference in frequency between those with and without bone lesions will be much lower than this. For example, to detect a doubling in frequency from 5% to 10%, with 80% power requires over 470 cases in each group (Cox, personal communication).

The sample of 24 has good power (91%) to detect germline variants present in the population at an allele frequency of 5%, but the sample of 24 only has 21% power to detect rare mutations (for example allele frequencies of around 0.5%) (Cox, personal communication). Therefore, we were not able to detect any rare germline mutations that might be expected to have a major effect on risk of bone lesions. More common variants (polymorphisms) can also affect risk of myeloma and might also be expected to affect risk of bone lesions, but these effects tend to be small, with relative risks (odds ratios) of 1.4 or less. So although a larger number of genes were associated with variants in the germline samples, assessing associations that may provide a predisposition to a worse bone phenotype would really require extensive GWAS studies, like those described by Houlston *et al*[74], who assessed over 7,000 myeloma patients and over 200,000 control samples. Those sample numbers and levels of analysis are beyond the scope of a project like this.

The sample number was limited in part by the funding we had available for the project as NGS processing and bioinformatic analysis remains a costly process. It was also limited by the number of samples that had sufficient bone marrow sample DNA following the previous testing as part of the project and those patients who also had a blood sample taken at the time of the bone marrow sample. Although samples taken to be used as part of the SMaRT patient research cohort did request a blood sample, many did not have this germline sample. This is, in part, due to fact that the blood sample does not represent the patient's disease and therefore the need to take this sample was not recognised by all involved in sample collection.

The design of the NGS panel was limited by the number of genes that could be included based on sequencing capacity and based on the cost of the panel. A number of genes that could have been included were not part of this analysis. With the hindsight of analysis and increased understanding of the pathways involved, other genes that could have been included are the interleukins, *IL6, IL1B* and *IL3*, Sclerostin domain containing 1 (*SOSTDC1*) gene and the insulin-like growth factor 1 (*IGF1*).

Patients in this study were assigned to the bone lesion or non-bone lesion cohorts based on an assessment and review of the patient notes. Whilst some notes were

extremely ordered and detailed and the data were easy to access, others were more difficult and required more clinical interpretation and lacked some detail. The bone lesion information was taken from the timepoint at which the bone marrow sample was taken. Whilst this would be considered an appropriate time period for assessment of bone lesions, a number of patients with no recorded bone lesions may have gone on to develop a more aggressive bone phenotype as their disease progressed. Therefore, this analysis is only an assessment of an association at the time of sampling and may misrepresent the myeloma patients who have the most severe bone manifestations. Initially we had considered assessing three groups of bone lesions; none vs 1-3 bone lesions vs multiple ≥3 bone lesions. This reduced the cohort numbers further and it was decided to limit the analysis to two disease cohorts; no bone lesions vs bone lesions.

In conclusion, the work presented as part of this chapter has allowed some initial work to be completed to explore the presence of variants in bone related genes in both the germline and somatic samples processed using the NGS osteome panel. It is understood that the power of the current study is not strong enough to show positive or negative relationships, however, this exploratory work has not been able to demonstrate any evidence to support a hypothesis that relationships exist between the genetic mutations in bone related genes and the likelihood or extent of myeloma related bone damage. Within the somatic changes seen, no variants were considered pathogenic, and those germline variants associated with risk (according to the bioinformatic pathogenicity callers), and indeed high risk, were seen in both the bone lesion and the non-bone lesion cohorts.

Based on the discussion above, there may many be reasons why an association was not seen in this study. This could be based on sample size, the genes and indeed the number of genes chosen as part of the osteome cohort, the level of filtering required to make a manageable data set, the depth of analysis undertaken on each variant, and the selection of the bone lesion vs non-bone lesion cohort.

We had considered this work could direct treatment to a cohort of patients likely to have a poorer bone phenotype, as long term treatment with bisphosphonates has some drawbacks, but perhaps this opens up the discussion about the more

widespread use of anabolic therapies in myeloma which have shown the ability to stimulate bone formation and therefore the ability to counteract the bone devastation caused by the disease in many patients.

Further work that may be considered within the SMaRT, may be to repeat this analysis adjusting a number of these factors; a larger patient cohort to meet the requirements of the power calculations, with follow up data on the bone lesions over the course of their disease, an increased number of genes, and a deeper level of variant assessment looking at variants in all gene regions. A similar project could also be undertaken using RNAseq or gene expression profiling instead of the osteome gene panel.

The RNAseq assessment allows analysis of the gene expression levels. It covers all genes in single sequencing event and therefore genes cannot be missed out. The analysis can start in a more selective manner looking at specific genes i.e. the osteome but could be extended to include other genes if a negative result is seen initially. There are some downsides to using RNAseq, as this analysis can be hampered by 3' bias. This is a phenomenon seen when cDNA fragments are sequenced; the number of reads corresponding to each transcript is proportional to the number of cDNA fragments rather than the number of transcripts. Since longer transcripts are generally sheared into more fragments, more reads are assigned to longer transcripts than shorter transcripts. Therefore, when assessing differential expression analysis, the differentially expressed genes are more likely to be enriched for longer transcripts resulting in this bias.

Recent work has shown the utility of gene expression signatures in robustly identifying a group of high risk myeloma patients[124]. Assessing gene expression rather than changes in the DNA, may prove a more sensitive assessment of any relationship between bone genes and the status of the bone phenotype associated with the disease. It may also be prudent to consider which cells are being assessed, i.e. the germline analysis will examine a predisposition to a more severe bone phenotype, but it may be interesting to assess somatic changes in cells of the bone marrow niche where you might expect bone related genes to be expressed, and not specifically the plasma cells associated with the bone marrow disease.

Chapter 5: Myeloma Quality Assessment Scheme

UK NEQAS and GenQA Introduction

The previous chapters have described extensive assessment of five genetic technologies suitable for the assessment of myeloma, and how these could be translated and utilised in a diagnostic NHS setting. The testing has to demonstrate clinical utility, be comprehensive, but also fit into a reasonable time frame and financial envelope to be able to justify the testing. Alongside the assessment of technologies has to be consideration of quality assessment and best practice. Many NHS diagnostic genetic tests have External Quality Assessment (EQA) schemes available that can be accessed by laboratories; completion of these schemes provides a certification that the test meets a suitable standard. The EQA schemes are often used to consider best practice, and to offer standardisation across the cohort involved, and this educative element is extremely valuable to participants. At the outset of the project, there was no EQA scheme available for myeloma, and it was considered a crucial aspect of the project to assess this need and work towards provision of an EQA scheme.

UK NEQAS (UK National External Quality Assessment Scheme) is a charitable, not-for profit consortium of external quality assessment providers for pathology services. UK NEQAS offers an independent, international service managed by experts in specific pathology fields and assured via specialist advisory groups (SAGs) made up of medical and scientific experts, and Boards of Trustees.

The aim of the individual EQAs is to improve patient care through monitoring the quality of tests and their reporting. Test results should be comparable, safe and clinically useful to the patient irrespective of where or when the test is performed. UK NEQAS also provides an educational component; the schemes are continually improved to fit the tests provided and EQA data are continually improved to support and train participating laboratories and to use the experience to embed and inform best practice and critical thinking.

https://ukneqas.org.uk/about/educate-me-about-uk-neqas/

In 2014, Cytogenomic External Quality Assessment Service (CEQAS) was established. This was a result of a merger of Cytogenetic European Quality Assessment (CEQA) and UK NEQAS for Clinical Cytogenetics, the two largest internationally operating EQA schemes at the time. In 2018, a further merger of CEQAS and the UKNEQAS for molecular genetics resulted in GenQA (Genomics Quality Assessment). GenQA is accredited through United Kingdom Accreditation Service (UKAS) as a proficiency testing provider and remains a member of the UK NEQAS Consortium. From the website https://www.genqa.org/ the aims of GenQA are to:

> Provide professionally lead and scientifically based EQAs with an educational objective
>
> Assess technical, analytical and interpretative performance of a laboratory
>
> Help the laboratory appraise its performance and monitor improvements externally
>
> Achieve this through continuous operation, frequent distributions of samples and performance feedback
>
> Produce reports which are designed to be clear, informative, intelligible and structured to assist interpretation and use by different levels of laboratory staff.

GenQA Scheme Design and Implementation

I sit on the Haematological-Oncology (HaemOnc) disease group Specialist Advisory Group (SAG), and this involvement with GenQA provided an ideal platform from which to introduce and implement a scheme for myeloma genetic diagnosis. At the point of myeloma EQA scheme inception, genetic diagnosis for myeloma patients was considered *ad hoc* and often available only in centres with a specialist or research interest, and frequently dependent on whether the patient was enrolled in national trials which included myeloma genetics as part of trial data collation. An EQA scheme for myeloma genetics was missing from the repertoire of any of the NEQAS organisations, and GenQA offered a good fit from which to launch this.

In conjunction with Ros Hastings (GenQA Director) and the GenQA SAG, a small team was devised to make up the myeloma EQA assessment panel. I was part of this team, and the following people were chosen for their interest, expertise and background:

Ros Hastings - Director of GenQA. Ros was keen to set up a myeloma EQA, initially as a pilot EQA, and was able to provide the infrastructure required to launch the scheme through GenQA.

Katrina Rack - Deputy Director of GenQA, and lead for the oncology schemes. Katrina was taken on by GenQA in 2015. At this point she became involved in the assessment of the myeloma EQA, but Ros remained involved from scheme management point of view.

Nicola Foot - Deputy Head of the Oncology Section, Cytogenetics Laboratory, Guy's and St Thomas' NHS Foundation Trust, London, UK. Nicky is a clinical scientist in a busy HaemOnc department and has many years of cytogenetic experience and offers a myeloma service for a large number of patients.

Jacqueline Schoumans - Head of Oncology Genetics, Medical Genetics Service, Lausanne University Hospital, Switzerland. Jacqueline has many years' experience in HaemOnc. The Lausanne laboratory offers an array service for their myeloma patients, but Jacqueline also has extensive FISH experience. She was also involved in writing the current laboratory guidelines for the use of arrays in HaemOnc disorders.

Sheila O'Connor - Haematology Clinical Scientist, Haematological Diagnostic Malignancy Service (HMDS), St James's University Hospital, Leeds, UK. Sheila is a haematologist by training with extensive experience in FISH. Sheila works in HMDS, the largest integrated HaemOnc specialist centre in the UK.

The pilot EQA scheme for myeloma was launched through GenQA in 2014 and included two cases of myeloma. These are described in detail below to illustrate the format of the EQA scheme. The first case in the 2014 pilot, was presented to the laboratories as fixed cells (bone marrow samples treated with 3:1 methanol:acetic acid), a referral form describing the case scenario and immunophenotyping results. Laboratories were expected to process the FISH for this case according to their

current in-house myeloma testing strategy and complete the analysis of the FISH testing.

The second case was available for analysis on the GenQA website. It included a referral form describing the patient case scenario, immunophenotyping results and morphology results and a number of FISH probes which could be selected for analysis. Ten probes were provided, six of those as a set of FISH images, and four probes sets had results given as text only.

The six probe sets that had images provided were:

Cytocell *IGH* Breakapart probe

Abbott *IGH/CCND1* Dual colour dual fusion probe

Abbott *IGH/FRGR3* Dual colour dual fusion probe

Abbott *IGH/MAF* Dual colour dual fusion probe

Cytocell *TP53*/17cen deletion probe

Cytocell 13q14.3/13q34 deletion probe

Images covering approximately 20 cells were provided for analysis and participants were instructed to assume that each cell represented 10 cells in a standard analysis. Examples of the images provided for each of these probe sets are demonstrated in figure 5.1.

These remaining probes sets listed below, were supplied with result given as text, e.g. *MYC* showed no evidence of rearrangement in 200 cells scored.

Abbott *MYC* Breakapart probe

Cytocell *CKS1B/CDKN2C* (1p/1q) Amplification/Deletion probe

Cytocell *IGH/CCND3* plus Dual colour dual fusion

Cytocell *IGH/MAFB* Dual colour dual fusion probe

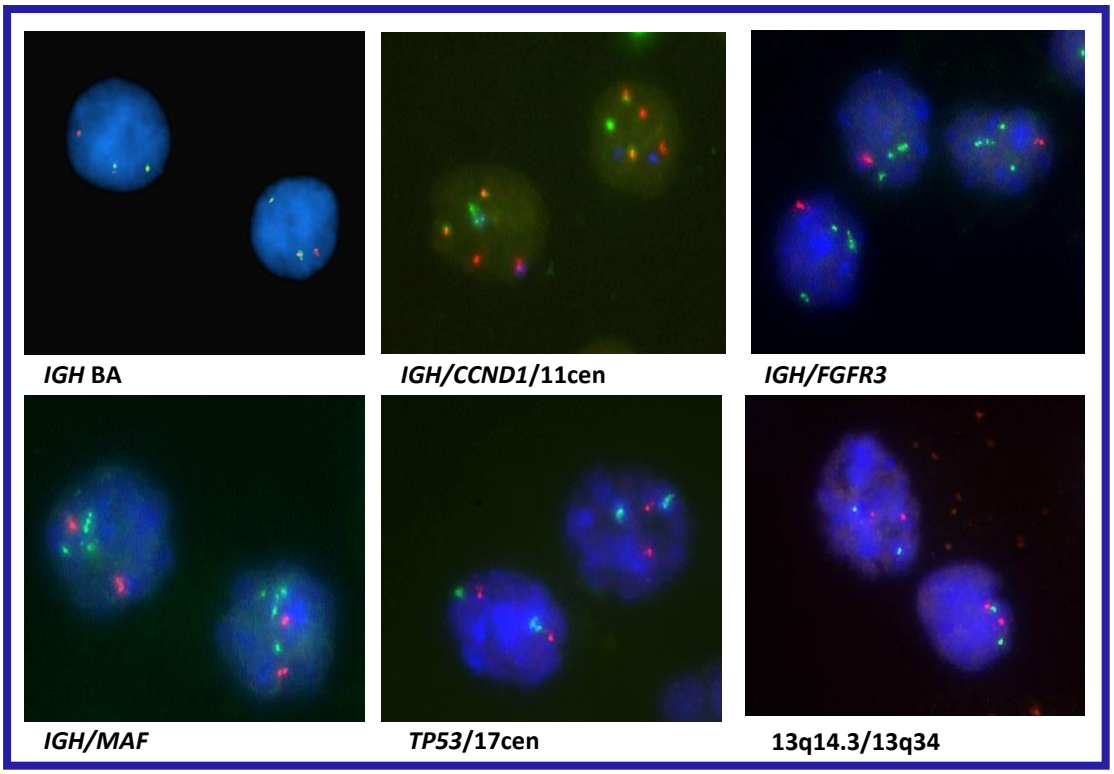


Figure 5.1: Demonstrating example images for the online EQA case presented as part of the pilot EQA for Myeloma 2014

Participants were expected to analyse the probe sets with FISH images, or with text-based results according to the standard strategy employed in their laboratory. For both cases, following FISH analysis, submission was in the form of an interpretative genetic report which was uploaded via the GenQA website. Table 5.1 shows the patient information for case 1 and 2, presented as part of the pilot Myeloma EQA including the expected, validated results. It should be noted that the names and dates of birth are fictitious, allowing the reports to be written and formatted according to the standard laboratory protocols, but with no concerns over data-breach.

Case	Name	DoB	Referral Reason	Validated Results
1	Sara ECRU	18/07/1950	Cough, ascites, pleural effusion. Confirmed Myeloma	*IGH* rearranged plus extra copies of *IGH.* *MAF-IGH* rearrangement. No *FGFR3-IGH* or *CCNDI-IGH* rearrangement. Monosomy for 13q14 & 13q34

				Normal for *TP53* No abnormality for 5, 9 & 15 centromeres. Sub-population with *MYC* rearrangement.
2	Seth EBONE	18/08/1951	Confirmed Myeloma. FISH please for prognostic indicators	*IGH* rearranged plus extra copies of *IGH.* *CCND1-IGH* rearrangement**.** No *MAF-IGH or FGFR3-IGH* rearrangement**.** Normal for 13q14, 13q34, *MYC* & *TP53* No abnormality for 5,9 & 15 centromeres.

Table 5.1: Detailing the validated results of the two cases presented as part of the pilot EQA for Myeloma 2014

Both cases were validated by four assessors independently and without prior knowledge of the FISH results; all results were consistent and confirmed. Assessment of the laboratory reports submitted to the GenQA website was performed by the same panel of four assessors and the scheme organiser, against pre-specified criteria, (see table 5.2), relevant to these cases and in-line with other schemes offered by GenQA. These criteria looked at the areas of analysis, interpretation and clerical accuracy. The 'analysis' category includes assessing the result obtained from the laboratory, how this is expressed either as a summary statement or in correct ISCN (International System for Cytogenetic Nomenclature), the nomenclature devised for describing cytogenetic and FISH abnormalities, and a clear written description of the correct findings with no ambiguity. The 'interpretation' assessment was based on providing a correct interpretation of the abnormalities found, including a statement of the genes involved and a prognosis comment. Finally the 'clerical accuracy' category covers a range of points relating the way the report is structured, and assesses the report against guidelines[134,135] produced describing requirements of a genetic report, i.e. the correct patient identifying information, the inclusion of the sample type and the probes types used, including manufacturer, and the overall report quality. Poor performance is defined as a critical analytical error or a critical interpretative error. Whilst scoring is completed in all scheme rounds, performance status is not assigned to the individual laboratories in the pilot rounds. Performance status is awarded once full scheme status is obtained.

Case	Category	Criteria	Marks
1	Analysis	Result is correct	2.0 marks
		Result is correctly given in ISCN or as a summary statement	0.5 marks
		Clear written description (marks deducted if written description is misleading or incorrect)	0.5 marks
		TP53 & *IGH* (*FGFR3, MAF* if *IGH* rearranged) FISH undertaken	0.5 marks
	Interpretation	Interpretation present and correct, comprising of:	2.0 marks
		Stating the gene fusion or genes involved	0.5 marks
		The prognosis is given	0.5 marks
	Clerical Accuracy	Correct patient name	0.5 marks
		Correct date of birth	0.5 marks
		Correct sample type	0.5 marks
		Probe manufacturer given if applicable	0.5 marks
		Two patient identifiers present	0.5 marks
		Absence of multiple typographical errors	0.5 marks
2	Analysis	Result is correct	2.0 marks
		Result is correctly given in ISCN or as a summary statement	0.5 marks
		Clear written description (marks deducted if written description is misleading or incorrect)	0.5 marks
		TP53 & *IGH* (*FGFR3, MAF* if *IGH* rearranged) FISH undertaken	0.5 marks
	Interpretation	Interpretation present and correct, comprising of:	2.0 marks
		Stating the gene fusion or genes involved	0.5 marks
		The prognosis is given	0.5 marks
	Clerical Accuracy	Correct patient name	0.5 marks
		Correct date of birth	0.5 marks
		Correct sample type	0.5 marks
		Probe manufacturer given if applicable	0.5 marks
		Two patient identifiers present	0.5 marks
		Absence of multiple typographical errors	0.5 marks

Table 5.2: Detailing the pre-specified assessment criteria of the two cases presented as part of the pilot EQA for Myeloma 2014

As part of each scheme, an individual laboratory report (ILR) is provided to each participating laboratory which details specific comments and observations under the three described areas based on their submission. In addition to the ILR, a scheme letter is produced which aims to describe the cases chosen, expected results, assessment criteria, the range of submissions seen and educational / learning points that have arisen throughout the assessment. This provides a valuable teaching and learning tool associated with the EQA scheme, and has been used to encourage and direct best practice.

GenQA Scheme Results and Progress

The format of the scheme has remained the same for subsequent years, year 2 (2015) was also run as a pilot and then in year three the myeloma EQA scheme was awarded full scheme status following ratification at the HaemOnc Specialist Advisory Group (SAG). The scheme has completed seven years and is now part of the substantive GenQA scheme repertoire. The cases included in the schemes for the first five years are summarised in table 5.3.

Scheme Year	EQA Case / Name	EQA Result
2014 **Year 1 Pilot**	Case 1: Sara ECRU	*MAF-IGH* rearrangement. Chromosome 13 monosomy
	Case 2: Seth EBONE	*CCND1-IGH* rearrangement.
2015 **Year 2 Pilot**	Case 1: Edgar MANN	*MAF-IGH* rearrangement. Chromosome 13 monosomy *MYC* deletion, no rearrangement. *IGH-MYC* rearrangement (with dual fusion probe)
	Case 2: Mia CANNE	*FGFR3-IGH* rearrangement. Loss of *DLEU* (13q14)
2016 **Year 3**	Case 1:Marc SYLVAN	No *IGH* rearrangement Gain of *CKS1B* (1q21.3) Loss of *CDKN2C* (arrays only)

	Case 2: Greta MERTENS	*IGH-FGFR3* rearrangement detected Gain of *MYC,* but no rearrangement Gain of *TP53* (17p13) Loss of *DLEU* (13q14) Gain of CKS1B (1q21.3) – 4 copies
2017 **Year 4**	Case 1: Roger WEST	No IGH rearrangement Gain of *CKS1B* (1q21.3) (3-4 copies) Gain of 5, 9 and 15 (3 copies) Chromosome 13 monosomy
	Case 2: Maria KAAS	*IGH-FGFR3* rearrangement detected Gain of *CKS1B* (1q21.3) *MYC* rearrangement detected Loss of *DLEU* (13q14)
2018 **Year 5**	Case 1: Jake BIRCH	*IGH-CCND1* rearrangement (non-standard pattern) Gain of *CKS1B* (1q21.3) Loss of *TP53* (17p13) Complex *MYC* rearrangement (by array)
	Case 2: Samuel OKEKE	*IGH-MAF* rearrangement Loss of *CDKN2C* (1p32.2)

Table 5.3: Detailing the validated results of the two cases presented as part of the Myeloma EQA scheme from 2014-2018

At scheme introduction in 2014, 39 laboratories participated; of those 33 (84.6%) obtained the correct results and produced satisfactory reports.

Participation has increased and plateaued over the period of five years; 39, 56, 65, 61 and 64 participants per year (2014-2018) respectively, shown in figure 5.2. This is in part due to expansion in the number of laboratories accessing the GenQA services, but also reflects an increase in UK laboratories offering this testing. This is likely to be due in part to this scheme being offered, but also as a result of myeloma being higher on the national genomics radar, which culminated in the publication of NICE guidance early in 2016[111].

Based on the scores, poor performance was recorded as 15.4%, 10.7%, 3.1%, 14.8% and 4.7% of participants over 2014, 2015, 2016, 2017 and 2018 respectively. See figure 5.2.

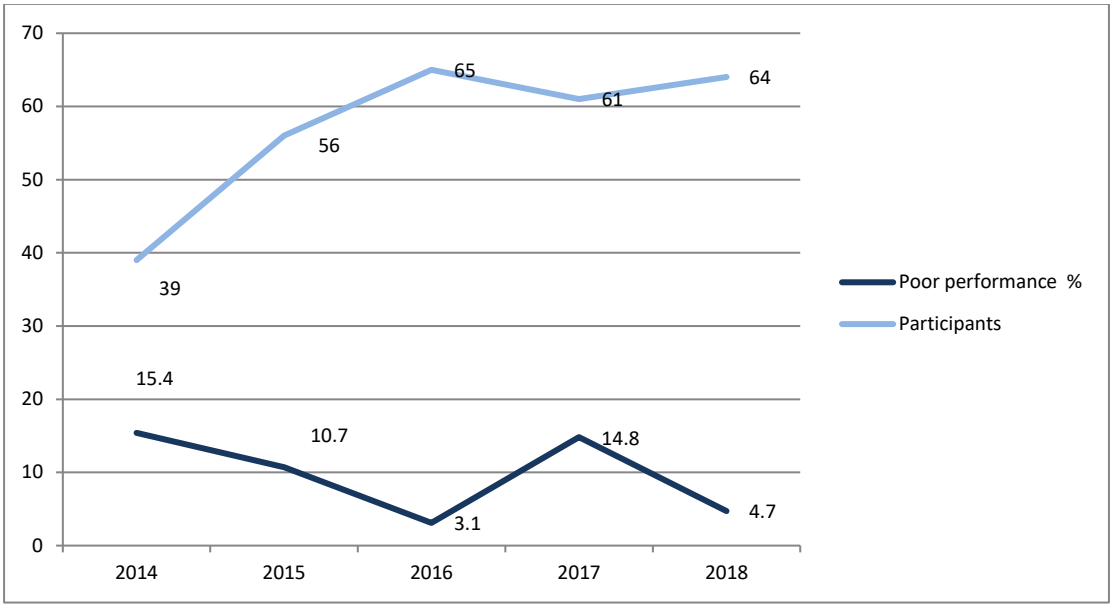


Figure 5.2: Graph to show the number of participants and the % of poor performance over the five years of the Myeloma EQA 2014-2018

Poor performance has fluctuated, but has, in general, shown a downward trend. This is likely to be related to the educational component of the scheme, and due to the increased profile of this testing with publication of the NICE guidelines on Myeloma: Diagnosis and Management[111]. The myeloma EQA scheme run in 2017 showed an increase in the poor performance; this peak in poor performance was due to a change in testing strategy employed by some laboratories, with penalties being applied to the use of specific probes in an inappropriate context. In more detail, assessors expected that the *IGH-FGFR3* rearrangement and the *TP53* deletion status should be determined as a minimum. *IGH-FGFR3* rearrangement can be excluded using an *IGH* breakapart probe or the *IGH/FGFR3* dual colour dual fusion probes. In this scheme round, other *IGH* dual fusion probes, such as *IGH/CCND1* or *IGH/ MAF* dual fusion probes were being used to detect rearrangements of *IGH* as a first line test. Whilst in the great majority of cases, these probes will pick up an *IGH-FGFR3* rearrangement, often seen as an additional copy of the *IGH* green signal, there are rare cases and unusual rearrangement patterns, where the third green signal could

be lost. This signal loss can be due to either loss of the whole chromosome or partial deletion of chromosome 14, a not uncommon finding, or loss of the derivative chromosome 14 involved in the rearrangement; both situations could result in a normal signal pattern using *IGH/CCND1* or *IGH/MAF* dual fusion probes. Therefore, it was deemed that if participants had used a different *IGH* dual fusion probe to detect an *IGH* rearrangement, other than the *IGH-FGFR3* considered the minimum rearrangement to determine, then in these, albeit rare scenarios, an *IGH* rearrangement had not been conclusively ruled out. In these cases, laboratories were assigned a poor performance. This was described in the educational scheme letter and there were no laboratories employing this testing strategy the following year, demonstrating a change in practice.

In general, case 1 in which the sample is sent out to laboratories to complete their own FISH testing, results in more poor performance. However, this case also allows assessment of a number of different criteria; the testing strategy employed, the technical ability to perform the FISH processing and analytical ability of participating laboratories. When presented with a range of FISH probes to analyse on the website, many participants analyse everything available when this does not reflect the strategies undertaken on a real-life sample. The images provided are usually clear and unambiguous when chosen for the website, and again this may not reflect the quality of the FISH produced in individual laboratories.

Managing the samples for each GenQA round can be challenging; providing images for cases is relatively straightforward, but as the number of participants rise, increased starting material is required for the wet work round. Approximately 200µl of fixed cells are provided to each participating laboratory, so in the most recent round a total of 13ml of fixed cells were required. Currently these samples are accessed through two large laboratories who have been willing to prepare and supply samples when diagnostic samples are seen with very high plasma cell percentages. The expense of separating CD138 positive cells has financially excluded this step as a possibility in preparing EQA samples, and therefore accessing samples from myeloma patients with high plasma cell content, best reflects the samples processed routinely through diagnostic genetic laboratories. It is easier to manage an

EQA scheme based on online images alone due to the challenging issues described in access, expense, and preparation of the wet samples. However, it is not possible to fully and accurately assess the laboratories testing strategies and technical FISH ability without providing the wet sample. For the foreseeable future, we would choose to use both methodologies over the two cases provided.

A further challenge seen in the provision of an EQA for myeloma, is the scoring and assessment that is applied when assigning prognosis in the reports. Interpretation and prognostic reporting is expected, but given that genomic testing in myeloma is a rapidly changing field, it is recognised that prognostic variables may lose or gain significance as a consequence of changes in patient management and clinical trial outcomes, and indeed are different amongst publications and recommendations internationally. When assessing this EQA round we have aimed to use the consensus of the different current recommendations to inform the minimum genetic tests expected. This is particularly relevant to 1q gain and 1p loss, where the prognostic significance has no clear consensus within international publications and guidelines. We have made this clear via the scheme letter that although omission of 1q/1p testing has not been penalised, that this may change in future EQA rounds if further guidance becomes available.

CEQAS Scheme Survey

Inaugural Scheme Survey 2014

Alongside the first pilot myeloma EQA scheme, a survey of current diagnostic genetic practice was conducted with the aim of gauging the breadth of testing available at the time of scheme inception.

The survey was short and simple to encourage participation and focussed on three main areas:

- The type of lab and location
- The referral type, samples & turnaround times (TAT)

- The techniques employed and the gene regions examined

The survey questions are included as appendix 7.

At scheme introduction, Autumn 2014, 39 participants took part from across Europe, and 18/39 (46%) completed the survey. Laboratories that completed the survey were all considered diagnostic laboratories, although approximately 13% were also considered research laboratories in part. The majority of laboratories were considered to be from the public sector (88%), with participation from a single private laboratory and a single laboratory from a private/public partnership. The majority, 10/17 laboratories (59%), were from the UK, but with laboratories from the following countries also participating: Italy, Belgium, Sweden, Switzerland and Greece, one laboratory did not supply this information. See figures 5.3 and 5.4.

Laboratories processed between 25 and 400 samples per year as detailed in graph 5.5. Where ranges were given, average figures were used and where specific figures were given they were attributed to the nearest group. The same applies to the turn around times (TATs) given in the survey shown in graph 5.6. Based on the averages, an overall average of 18.7 days TAT was calculated.

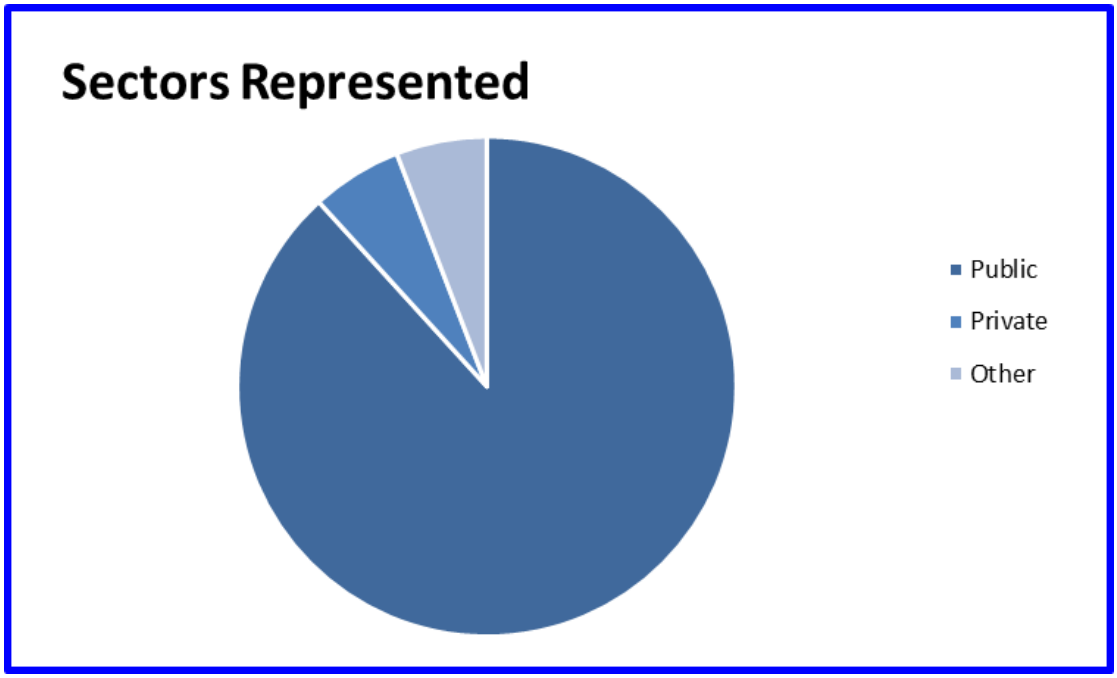


Figure 5.3: Demonstrating the range of public, private and public/private partnership laboratories participating in the GenQA survey in 2014.

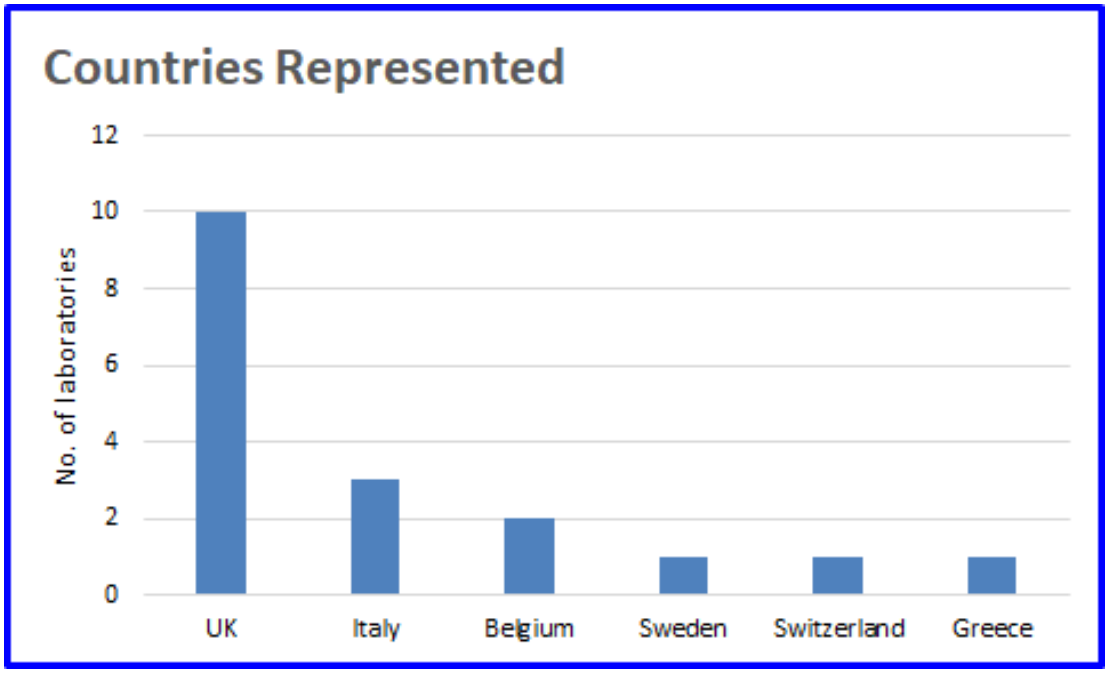


Figure 5.4: Demonstrating the range of countries participating in the GenQA survey in 2014.

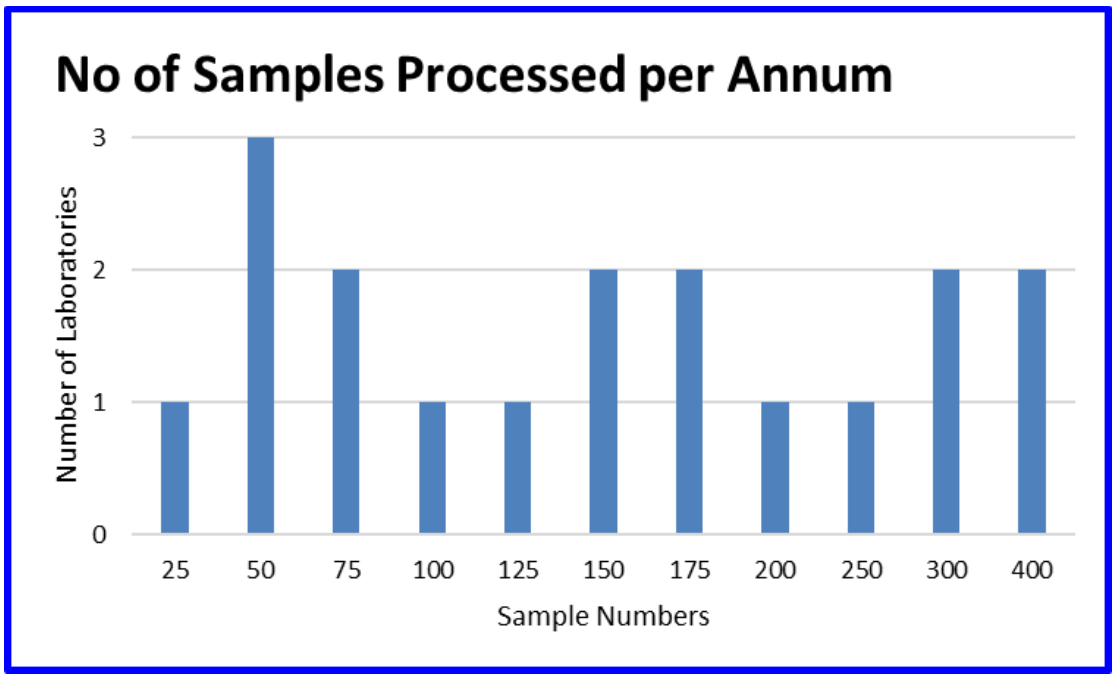


Figure 5.5: Demonstrating the number of samples processed per annum.

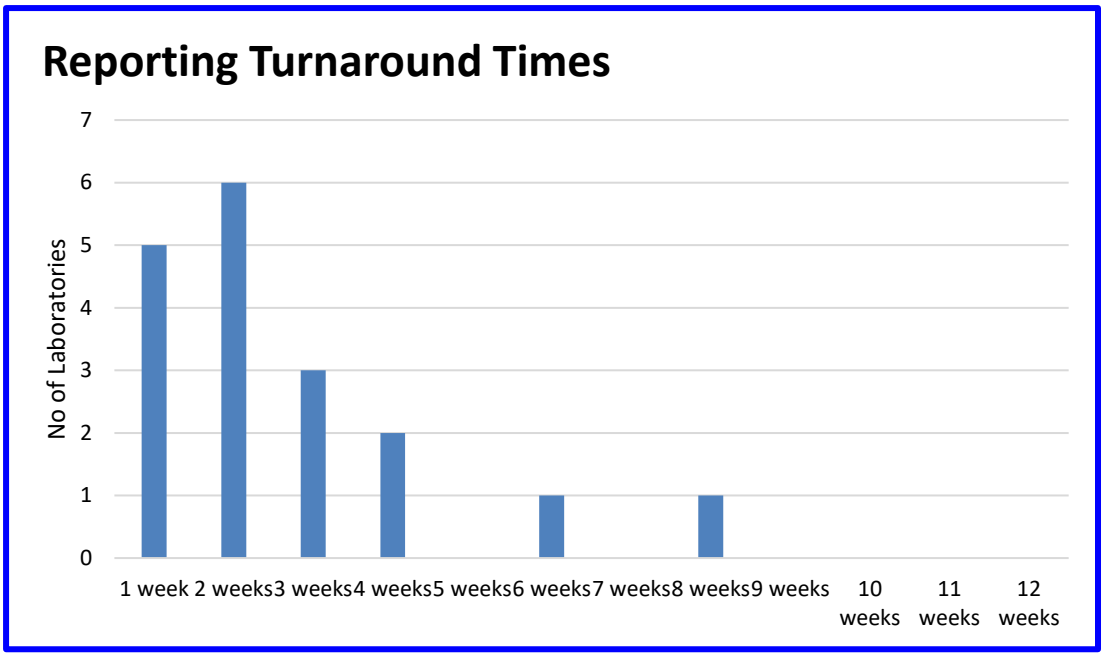


Figure 5.6: Showing the range of TATs reported by laboratories in the survey performed in 2014.

Sixteen of the 18 (89%) laboratories employed a plasma cell separation method, all of whom utilised the CD138 postive methodology as described in the materials and methods chapter. Of the 16 laboratories who separate the plasma cells, 6 (38%) assess the purity of the sample. Five of the 16 labs utilise morphological methods and 1 lab uses flow cytometry to assess the plasma cell purity. Only a single laboratory (1/18) utilises the fluorescent cytoplasmic immunoglobulin (cIg) method.

The following chart (figure 5.7) shows the methodologies employed for myeloma analysis. The majority of laboratories offered FISH testing as a sole technique, one laboratory offered FISH and DNA arrays, one laboratory offered karyotyping and FISH, and one laboratory offered a combination of all three methods.

Only the two laboratories employing arrays for myeloma analysis completed the section on DNA requirements as it was not applicable to the other respondents. No laboratories participating in the survey are routinely using MLPA for analysis of myeloma although two laboratories report that assessment and validation is currently being considered.

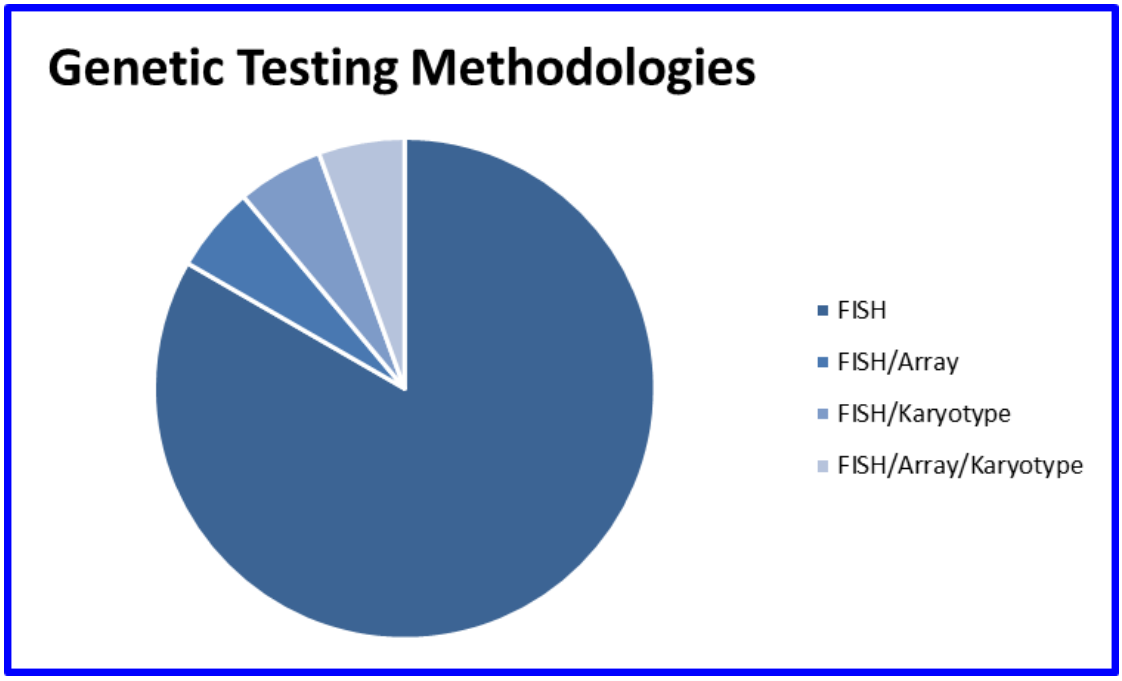


Figure 5.7: Demonstrating the methodologies employed for myeloma genetics analysis based of the survey results received in 2014.

One of the cases provided to the GenQA participants in 2014 was a cell suspension sample for FISH processing in-house. Based on the testing methodology employed for the GenQA case we can infer the testing provided by the 39 laboratories that took part in the pilot scheme. This information is provided in table 5.4.

All 39 labs offered *IGH* testing, and 37/39 (95%) offered *FGFR3* and *MAF* testing as a follow-on test if the *IGH* proved abnormal. Of the remaining five common partner genes, *CCND1* testing was offered by 18/39 (46%), *MAFB* testing by 6/39 (15%) and *CCND3* testing by only 1/39 (3%) laboratories. 15/18 labs describe a multistep FISH process with additional probes added based on the outcome of an initial *IGH* assessment.

TP53 testing, which can be considered one of the most important prognostic indicators in myeloma, was only offered by 37/39 (95%) labs. Chromosome 1 abnormality testing was offered by 30/39 (77%) laboratories and 13q abnormalities by 24/39 (62%). A minority of labs offered hyperdiploidy assessment, 8/39 (21%), and *MYC* testing 3/39 (8%).

Chromosome/Gene regions tested for	Number of laboratories
TP53	37
IGH	39
IGH-FGFR3	37
IGH-MAF	37
IGH-MAFB	6
IGH-CCND3	1
IGH-CCND1	18
1p/1q	26
1q only	2
13q	24
MYC	3
Hyperdiploid assessment	8

Table 5.4: Probes used by the laboratories who took part in the GenQA myeloma pilot scheme in 2014

Eight laboratories routinely accept referrals of MGUS and eight laboratories do not. Two labs report that they will analyse MGUS samples if specifically requested.

Only 3/18 laboratories employ an age limit to the requests for myeloma FISH. The age limit is <70 years for 1 of the 3 laboratories and <75 years in the remaining two laboratories.

The inaugural scheme survey, provided evidence for the *ad hoc* nature of genetic testing in myeloma, although the majority of laboratories did carry out the essential tests.

2018 Scheme Survey

The survey was repeated alongside the 2018 scheme in order to reassess the findings and to determine changes in practice seen over this period. The questions remained similar and can found in appendix 7. Thirty one of the 65 (48%) laboratories enrolled in the 2018 myeloma EQA scheme completed the survey.

Laboratories that completed the survey fall into two of the following three categories (diagnostic, research or both): 29/31 (94%) are considered diagnostic and 2/31 (6%) are considered both diagnostic and research. Most laboratories (61%) are part of a public hospital, 26% are part of a public university and 10% are considered private laboratories, and single laboratory described themselves as 'other', see figure 5.8.

Countries represented by the survey contributors are shown in figure 5.9, and whilst the majority are UK laboratories, the spread of countries covered is far greater than was seen in the original survey.

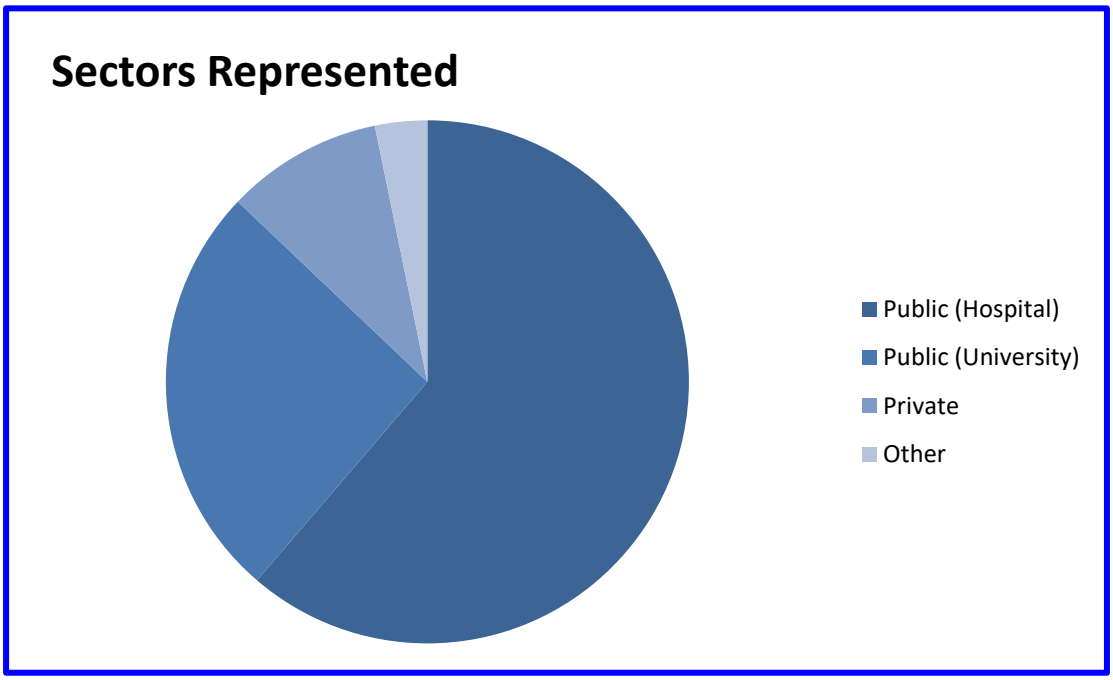


Figure 5.8: Demonstrating the range of public (hospital or university based), private and public/private partnership laboratories participating in the GenQA survey in 2018.

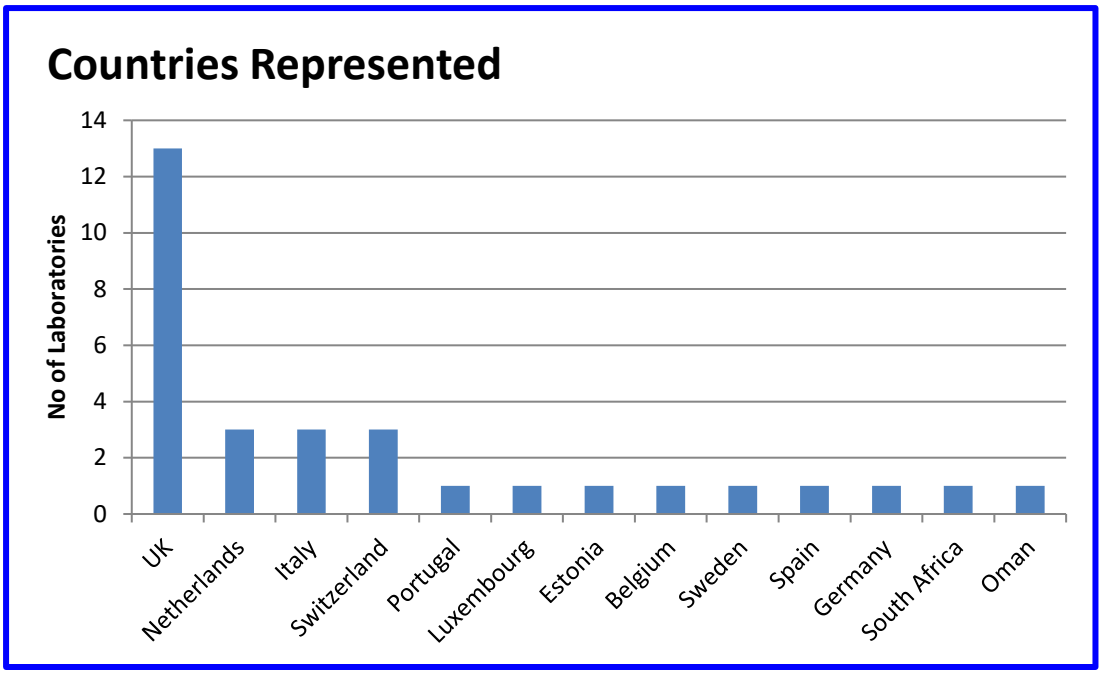


Figure 5.9: Demonstrating the range of countries participating in the GenQA survey in 2018.

Laboratories processed between 26 and 650 samples per year, with an average of 228 samples/year, and this related to populations served ranging from 230,000 - 6,000,000 with an average population of 2.5 million. The average TAT was 21.5 days, but with a huge range from 4 to 120 days (see figure 5.10). If the two outliers were removed the mean TAT was 16.8 days. This compares to a mean TAT of 18.7 days from the previous 2014 survey, although it is acknowledged that this figure is based on averages assumed from numerical ranges supplied in both survey responses. The TAT is affected by the testing strategies employed; many laboratories are employing sequential testing strategies, which, whilst offering a cost saving overall (as some tests may not be required), does result in an extended time to complete the testing.

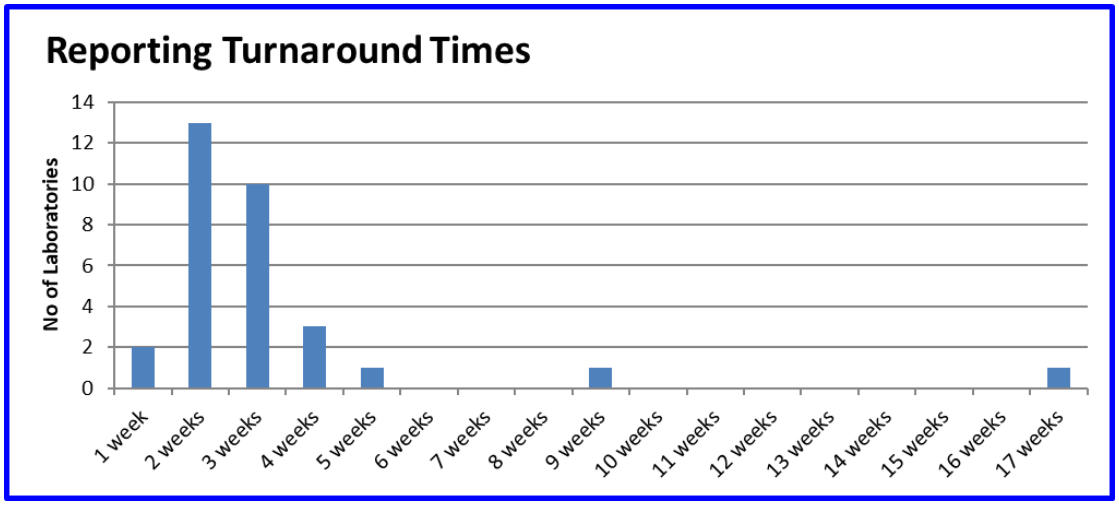


Figure 5.10: Showing the range of TATs reported by laboratories in the survey performed in 2018.

The mean abnormality rate was 72% but with a range of 30-100%, and the failure rates were reported as between 0-32% with a mean of 6%.

20/31 (65%) laboratories routinely accept referrals of MGUS and 11/31 laboratories do not. This compares to 50% of laboratories accepting MGUS referrals in 2014. Only 2/31 (6%) laboratories employ an age limit to the requests for myeloma FISH, and the age limit is 75 years and 85 years for those two laboratories. This compares to 3/18 (17%) seen in the survey results from 2014.

29/31 (94%) laboratories employ a plasma cell separation method, all of whom separate using CD138 postivity. Three laboratories also utilise CD38+ separations, and one laboratory describes the use of CD19+ and CD56+ separations in a proportion of cases. 23/28 (82%) of laboratories who described the technique used employed a magnetic cell sorting strategy, one described a FACS based methodology and 4/28 used 'other' to describe their strategy.

Of the 29 laboratories who separate plasma cells, 11 (40%) assess the purity of the sample. Six labs utilise morphological methods and five laboratories use flow cytometry to assess the plasma cell purity. This has not changed dramatically since 2014. Only a single laboratory (3%) utilises the fluorescent cytoplasmic immunoglobulin (cIg) method, and this is the same as 2014.

The following chart (figure 5.11) shows the methodologies employed for myeloma genetic analysis. The majority, 22/31 (71%) use FISH only as their testing strategy.

One laboratory used array analysis only, and 4/31 (13%) laboratories employed both FISH and array. One laboratory completed FISH and karyotyping and a further laboratory used this strategy with *TP53* sequencing in addition. Two laboratories undertook a combination of FISH, array and *TP53* sequencing.

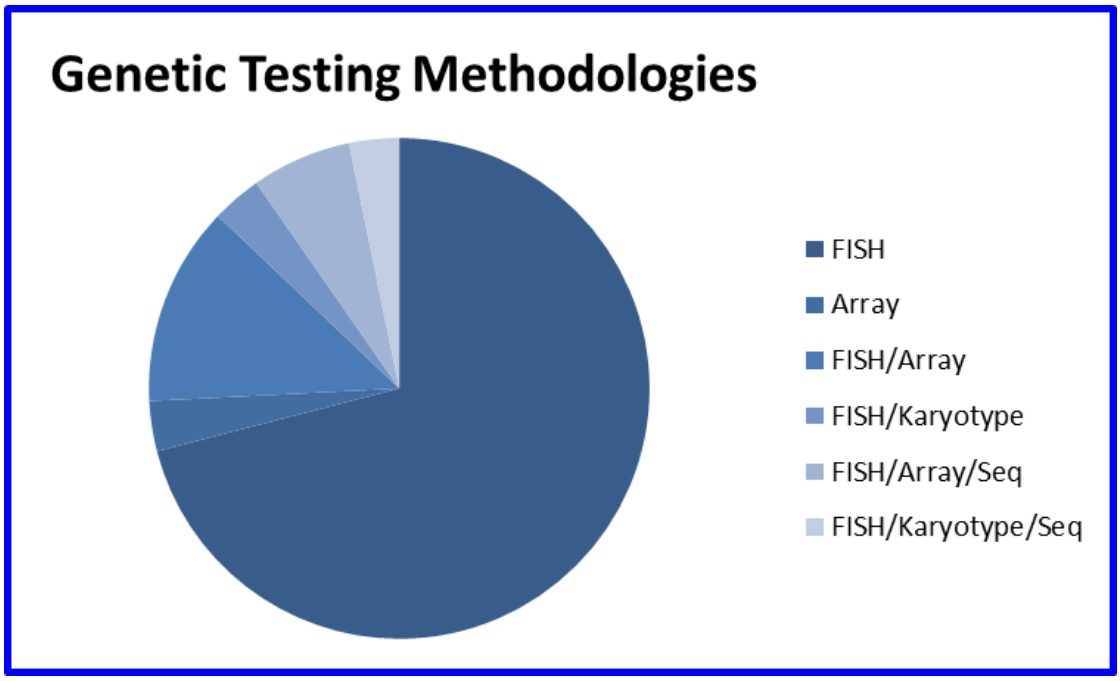


Figure 5.11: Demonstrating the methodologies employed for myeloma genetics analysis based of the survey results received in 2018.

Based on both the survey and case 1 in which the wet work was completed in house, we were able to determine the chromosomes or gene regions tested for, and the testing strategies undertaken by different laboratories. These results in the table are based on the 63 participants of the 2018 myeloma EQA scheme, and the survey results received.

The regions tested for varied between laboratories and is shown in table 5.5.

Chromosome/Gene regions tested for	Number of laboratories
TP53	63
IGH	63
1p/1q	61
13q	19
Hyperdiploidy	17
MYC	5

Table 5.5: Probes used by the laboratories who took part in the GenQA myeloma EQA scheme in 2018

All 63 laboratories performed *IGH* rearrangement testing, although different strategies were employed. These individual differing strategies are outlined in table 5.6. Overall 28 laboratories used a combination of an *IGH* breakapart probe and dual fusion translocation probes and 35 laboratories used dual fusion translocation probes alone. It is recognised that the *IGH* testing strategy may be sequential or simultaneous and that the tests performed are therefore case dependent. Of the 63 participants, all laboratories assessed *TP53* status; 61 (97%) laboratories assessed 1p/1q copy number; 19 (30%) laboratories assessed chromosome 13 deletion/loss; 17 (27%) laboratories tested for hyperdiploidy and five (8%) laboratories tested for *MYC* gene rearrangement.

Testing Panel*	Number of laboratories
IGH-FGFR3	1
t(11;14) only	3
IGH-FGFR3 **and** ***IGH-MAF***	10
IGH-FGFR3**, and t(11;14)**	4
IGH-FGFR3, IGH-MAF **and t(11;14)**	11
IGH-FGFR3**,** ***IGH-MAF*** **and** ***IGH-MAFB***	4
IGH-FGFR3**,** ***IGH-MAF*****,** ***IGH-MAFB*** **and t(11;14)**	2
IGH **BA and probes specific for** ***IGH-FGFR3*****,** ***IGH-MAF*** **and t(11;14)**	9
IGH **BA and probes specific for** ***IGH-FGFR3*****, and** ***IGH-MAF***	6
IGH **BA and probes specific for t(11;14)**	5
IGH **BA and probes specific for** ***IGH-FGFR3*** **and t(11;14)**	4
IGH **BA and probes specific for** ***IGH-FGFR3*****,** ***IGH-MAF*****, t(11;14) and** ***IGH-MAFB***	2
IGH **BA and probes specific for** ***IGH-FGFR3*****,** ***IGH-MAF*** **and** ***IGH-MAFB***	2

Table 5.6: Table detailing the FISH panels employed by laboratories to determine the IGH rearrangements in the 2018 myeloma EQA scheme. * different probes sets are used for the detection of the t(11;14), but have been considered a single test

Forty laboratories undertook testing for the t(11;14) rearrangement. Different approaches were taken: a single laboratory used a *CCND1* break apart probe, 32 laboratories used an *IGH-CCND1* dual fusion probe and 7 laboratories used an *IGH-*

MYEOV dual fusion probe. In table 5.6, these different probe combinations have been considered as a single test for the purposes of understanding the testing panels employed.

All laboratories included *IGH* and *TP53* in their assessment of myeloma genetics. This demonstrates a shift from the survey performed in 2014 where all laboratories performed *IGH* assessment, but only 95% included *TP53* deletion testing. *TP53* and *IGH,* leading to an assessment of *FGFR3-IGH* and *MAF-IGH*, are considered a minimum testing set according to Ross *et al,* 2012.

A range of 100-200 cells were scored per FISH set, and the cut-off was usually given as 10% although some labs use a probe specific cut off and this ranged from 2-30%.

Only the laboratories employing arrays for myeloma analysis completed the section on DNA requirements as it was not applicable to the other respondents. No laboratories participating in the survey routinely use MLPA for analysis of myeloma, although it is known that at least one UK laboratory uses MLPA assessment for their copy number determination.

Twenty laboratories stated that they were not currently considering NGS for their myeloma testing, whilst eight laboratories are either using NGS or considering it. Of those eight laboratories, three are using NGS for *TP53* testing only, one for *BRAF* testing only, and a single lab described a panel of *KRAS, NRAS, BRAF, IRF4, TP53, DIS3* and *FAM46C* which was being validated. The remaining three labs did describe the NGS testing that was being considered.

Survey Conclusions

In conclusion, these surveys have demonstrated a change over the period of implementation of the myeloma EQA scheme. In 2014, although all laboratories assessed *IGH* rearrangement in myeloma referrals, 5% of laboratories did not include *TP53* testing in their panel. In 2018, the picture has changed with 100% of labs now including tests that meet the requirements laid out in both the NICE guidance and the BSH guidance to assess patient for both *IGH-FGFR3* and *IGH-MAF* rearrangements and *TP53* deletion testing. However, although the essential tests are

being provided by all of the laboratories enrolled in the scheme, an inequitable service is provided due to the variability of testing offered and indeed the quality of service provided.

In 2014, 89% of labs reported completing their testing on separated plasma cells, which rose to 94% in 2018; however, a proportion of laboratories are completing their testing on whole bone marrow which would not be considered best practice. Average TATs have improved from 18.7 days to 16.8 days. In general this fits with the diagnostic pathway for myeloma, but may need further improvement if treatments are to be based on these findings.

FISH is by far the methodology of choice currently, but with a minority of laboratories including MLPA and array in their current testing strategies. The 2018 survey showed that four labs were considering including mutational analysis NGS panels as part of their genetic testing, although no labs are offering this at present.

Diaceutics Study

During the period of this PhD study, I also contributed to a data collection exercise run by Diaceutics on behalf of Takeda Oncology. Diaceutics are a commercial company that provide consulting services to Pharmaceutical companies; they provide data research and understanding of the diganostic pathways and markets, and create plans for the integration of the Pharma companies therapies into the diagnostic and therapeutic pathway. This study was published in the Myeloma Hub Connect in January 2018[136].

Their data are extremely interesting and complements information we have gathered through the GenQA surveys. The Diaceutics data are limited to the UK and Ireland, but extends the data collection and information to cover populations served, and an understanding of the funding streams covering this testing and the future requirements of a testing strategy.

Their study concludes, as ours does, that the genetic testing in the UK is variable with regards to the methodologies employed. They state that the 'NICE and BSH guidance to test for the high-risk abnormalities del(17p), t(4;14) and t(14;16) are

being followed, but a significant proportion of patients are not screened for cytogenetic abnormalities at diagnosis'. They go on to suggest that inequity of access to testing is due to the lack of central funding, and this limitiation on funding means that local funding has to be sourced or that haematologists have to be selective in the patients they choose to test. They also recognise that the TATs in this area of testing are relatively high, affected by the sequential testing methodologies employed. A change in this TAT would result in the need for parallel testing and an increased financial burden. Finally, they conclude that future genetic testing for myeloma is likely to increase with the introduction of further targeted therapies, highlighting the potential use of Vemurafenib and Venetoclax in *BRAF* mutated and t(11;14) myeloma respectively, and indeed touch on the need to complete testing at more than one time point in a patients disease.

NHS England Rationalisation of Genomics Services

At this stage, by way of background understanding, it becomes ncessary to discuss National changes we are seeing as part of NHS Englands reconfiguration of genomics services. Genomic laboratories were invited to tender for the provision of genomics services as seven regions covering England, not as the 20+ laboratories currently offering genomics services. In October 2018, the seven genomics laboratory hubs (GLH) were announced following the complex tender process.

As part of this reconfiguration process, a National Test Directory has been produced detailing the testing repertoire to be covered by each GLH, and ultimately funded by NHS England from April 2021. For myeloma this testing covers:

Structural *IGH* rearrangement testing: t(4;14) *IGH-FGFR3*, t(6;14) *IGH-CCND3*, t(11;14) *IGH-CCND1*, t(14;16) *IGH-MAF* and t(14;20) *IGH-MAFB,* and this can be offered as part of an NGS panel

Hyperdiploidy copy number changes

del(1p) copy number detection

Gain(1q) copy number detection

del(17p) *TP53* copy number detection

Small variant detection; to include *KRAS, NRAS, BRAF, TP53, DIS3, FAM46C* and *IRF4*

Currently the copy number changes are recorded as FISH testing, but we are aware of a small number of laboratories offering array or MLPA testing for these tests. As these technologies may represent a cost saving over large panels of FISH tests, proposals to include this as alternative test have been accepted.

The small variant detection is listed as a multi-target NGS panel, and information from the most recent GenQA survey shows only a single laboratory considering offering NGS panel testing covering these genes.

The reconfiguration of genomics services remains in flux with many unanswered questions in relation to the finances, the operational functions of the main central laboratories and indeed the 'spoke' laboratories, the IT infrastructure, quality processes, research and development, and staffing. However, from a testing perspective, and specifically for myeloma, this could drive the standardisation of testing, and indeed financial standardisation in turn leading to equity of access to diagnostic testing.

The most current test directory published in August 2020 describes the following testing myeloma:

FISH/RT-PCR for rearrangement detection of:

IGH-FGFR3 t(4;14)

IGH-CCND3 t(6;14)

IGH-CCND1 t(11;14)

IGH-MAF t(14;16)

IGH-MAFB t(14;20)

The test directory states that these abnormalities could also be detected using an NGS fusion panel. As *IGH* rearrangements are not standard fusion rearrangements these remain difficult to detect with panel based NGS, and would require the application of low coverage WGS. There is also a FISH test listed for *IGH*

rearrangement testing to cover the use of a breakapart probe, and MYC rearrangement by FISH has been added to this iteration.

FISH for copy number changes of:

Hyperdidploidy

CDKN2C (deletion of 1p)

CKS1B (gain of 1q)

TP53 deletion

The test directory now makes reference to MLPA and NGS (but not array analysis) as suitable alternative methodologies for copy number detection.

Multi-target NGS panel for:

KRAS

NRAS

BRAF

TP53

DIS3

FAM46C

IRF4

Best Practice Recommendations

It was clear from the survey carried out in 2014, that there was a need for standardisation and best practice recommendations. As part of the GenQA consortium, we considered that generalised guidelines were required for the whole of haemato-oncology diagnostic testing and led by Katrina Rack, we have completed and published recommendation in Leukemia. My role in this was to complete the myeloma recommendations section.

- Rack *et al* (2019). European recommendations and quality assurance for cytogenomic analysis of haematological neoplasms. Leukemia 33(8):1851-1867[123].

These recommendations are brief, and production of myeloma specific guidelines including additional detail are also in production, but have not yet been published. These guidelines would cover the genomic targets required as discussed in the general guidelines, but also cover in more detail the technologies that could be employed. There are disease specific concerns that could also be covered, such as whether to limit testing to myeloma patients or whether MGUS patients would also benefit, whether follow up genetic testing is appropriate or whether this would be better managed with flow cytometry, consideration of the age of patients to be tested, the requirement for testing specific cell selections, and understanding the consequences and how to manage reporting if this is not possible. The recommendations also cover how to report these, sometimes, complex results and discusses appropriate turnaround times and quality assessment.

GenQA Scheme Concluding Remarks

To summarise this chapter, we have been able to describe the successful introduction and implementation of an EQA for genetic diagnosis in myeloma through GenQA and my involvement on the Haematological-Oncology SAG. This ran for two years as a pilot scheme and has now run for a further five years with full EQA scheme status. The involvement of the EQA scheme into the overall project was considered crucial; suggesting potential genomic testing strategies for myeloma in a diagnostic NHS laboratory, has a requirement of an EQA scheme to complement the service. The EQA scheme validates the testing strategies. The EQA scheme goes further allowing the educational component to guide participants, and reduce the *ad hoc* nature of testing, it also offers a spring board from which to access expertise required to produce and recommend best practice.

The EQA participation has gradually increased over the scheme period from 39 in 2014 to 64 in 2018, showing a 64% increase in participation from 2014 to 2018. 2018 saw participation from all UK laboratories offering myeloma testing, and a number of European and International laboratories Although the performance status over the five years has fluctuated, there is an overall downward trend in poor performance

from 15.4% in 2014 to 4.7% in 2018, demonstrating a positive effect of the educational component provided by the scheme and increased awareness of the requirements of a diagnostic genomics strategy for myeloma over this time period.

As part of this EQA scheme, a survey was issued in 2014 alongside the pilot, and the further similar survey was issued in 2018 alongside the fifth EQA scheme. The inaugural scheme survey confirmed the suspicion that the genetic diagnosis in myeloma at this time was indeed *ad hoc*, and demonstrated that the quality and extent of testing provided was inequitable. This highlighted the requirement, not only for the educational component provided by the EQA scheme, but for the production of best practice recommendations in this area. Broad best practice recommendations covering genomic testing for all haem-onc disease groups have been produced and published[123], and further work to produce more in depth myeloma specific best practice guidelines are also currently being produced.

Over the period that the EQA scheme has run, we have seen a degree of harmonsation of genetic testing, with a more consistent, equitable approach to the strategies employed by diagnostic genetic laboratories, demonstrating the impact that EQA can have through participation and education. Although all laboratories are now offering what are considered to be the essential tests in the genomic testing of myeloma, not all laboratories are carrying out this analysis on plasma cell populations, (which is considered best practice) and the extent and quality of testing can still be considered inequitable. The majority of laboratories only employ FISH methodologies, but the 2018 scheme and survey show a small number of laboratories using both MLPA and array, and indeed four laboratories describing the future validation of NGS technologies as part of their testing regimes.

A similar study undertaken by a commercial company has confirmed very similar findings to our own, but has gone further to suggest that based on population sizes and patient numbers receiving testing, a number of myeloma patients are still not receiving any kind of genomics testing as part of their diagnosis. This study suggests this is related to the funding of this testing; direct NHS funding is currently not available, so laboratories are offering this as cost per test or have found innovative local funding sources. With the imminent changes in NHS Englands reconfiguration

of genomics services, we know that myeloma testing does appear on the National Test Directory which in turn means that this test should be funded nationally from April 2021.

Chapter 6: Discussion

During the time of this project I have held positions within both Sheffield Diagnostic Genetics Service (SDGS) at Sheffield Children's Hospital NHS Trust and at the Haematological Malignancy Diagnostic Service (HMDS) based at St James's Hospital Cancer Centre within Leeds Teaching Hospitals Trust. These roles have driven a strong interest in translational research, and I have been keen to ensure that the results from this work directly influence the strategy for genetic testing for myeloma patients, inform best practice and result in setting up of a quality assessment scheme across the UK.

The project aims to explore a number of different objectives:

Objective 1 was to compare five different genetic technologies, readily available in a diagnostic genetic laboratory, and their ability to identify genetic signatures associated with plasma cell neoplasms. Karyotyping, FISH, MLPA, DNA array analysis & targeted NGS were applied to a cohort of patients in order to assess the results obtained, and their ability to do this in an effective and efficient manner, within the time constraints required of a diagnostic testing scenario. A number of genetic changes have been shown to provide valuable information about disease prognosis and are beginning to guide treatment decisions.

Objective 2 was to explore a possible or potential relationship between the genetic signatures associated with bone related genes, the 'osteome', and the likelihood or extent of bone damage associated with the patient's myeloma disease. This objective worked to the strengths of the research group and their interest in myeloma bone disease. A targeted NGS panel was created including genes associated with the 'osteome'; genes involved in the development, influence, maintenance and destruction of bone. Treating the bone manifestations of myeloma is challenging; current treatments are able to offer protection and slowing of bone degradation but can rarely reverse the process. If correlations between the mutations and variants in these genes and the extent of bone disease can be made, it may be possible to identify patients considered at increased risk of a severe bone phenotype associated with their myeloma. This, in turn, may offer the opportunity to

treat this group of patients more proactively, slowing the advancement of the bone manifestations or open up the discussion about the use of anabolic treatments in myeloma and the potential opportunity to adopt a more personalised medicine approach.

Finally, objective 3 involves using the evidence gathered in objective 1 to devise an all-encompassing diagnostic genetic testing panel for use in a diagnostic genetic laboratory considering and proposing best practice guidelines to create a more consistent, equitable harmonisation of diagnostic genetic testing for MM within the UK. Quality assessment schemes exist for the majority of genetic tests completed as part of a diagnostic laboratory. At the start of this project, this was not available for myeloma genetic testing, and this objective includes the introduction, through links with NEQAS, of a quality assessment scheme for the genetic testing for myeloma patients.

Discussion of the Genetic Results and Techniques Employed

As described, the initial part of this project has focussed on collection and processing of patient bone marrow samples and exploring the genetic signatures associated with plasma cells neoplasms, consistent with the first objective described in the project aims. Five main genetic technologies have been explored with respect to analysis of myeloma patient samples; cytogenetics (karyotyping), FISH, MLPA, DNA arrays and NGS, all of which are available within a diagnostic genetic laboratory.

Over the period of this project, bone marrow samples were received from 113 patients with a plasma cell neoplasm, all consented as part of the research project being conducted by Dr Andrew Chantry. Twelve patients were removed from the cohort following diagnosis of an unrelated condition or following confirmation that no plasma cell neoplasm was present.

Of the remaining 101 patients, 91 patient samples had their marrows cultured and processed for karyotyping, 101 of these samples were FISHed, 45 patients had MLPA analysis, 36 patient samples were analysed using the Affymetrix DNA array, and 24

patient samples were assessed using a bespoke NGS panel designed to assess both genes associated with myeloma, and a set of genes, considered the osteome, which are known to be involved with bone and myeloma interactions. The ideal scenario would have been to assess all samples using all five technologies, however the quality and quantity of bone marrow sample, and the CD138+ve cell selections from the bone marrow varied dramatically. Therefore, many samples were insufficient to analyse in parallel using all five platforms. In addition to the sample limitations, genetic technologies are expensive, and the number of samples processed for each technology was carefully planned to fit within the finances available to fund this project.

At the beginning of the project a number of samples were processed for FISH using the whole marrow, and it became clear very quickly that samples must be selected in order to assess the CD138+ve plasma cells only. The final project cohort only included samples with CD138+ve cell selections, except for a single patient with plasma cell leukaemia who demonstrated an extremely high level of plasma cells as part of his disease. This must be considered part of the best practice recommendations, and indeed it is recognised throughout the literature, that analysis of whole bone marrow samples is not suitable. The disease can be present in 10% of cells and the abnormalities can also present at a low level within the plasma cells. Therefore, to increase the chance of detecting genetic abnormalities a method of plasma cell selection must be employed.

The five genetic technologies explored demonstrate a range of results and these are explored and described in depth in chapter 3, the advantages and disadvantages of these technologies are also discussed. In summary, conventional G banded cytogenetic analysis (karyotyping) offers the benefit of whole genome analysis, albeit at a low resolution, which allows the simultaneous assessment of rearrangements, including the *IGH* rearrangements (except t(4;14) which is cytogenetically cryptic) and copy number changes including ploidy changes. Small changes (<5-10Mb) and mutations cannot be detected using this technology. The requirement for cells in metaphase can be extremely problematic due to the low proliferative rate of the terminally-differentiated mature plasma cell and is recognised as technically

challenging. Although adaptations to the culturing process are possible to increase the level of dividing plasma cells, it cannot create the equivalent of a cell selection process, and therefore cytogenetic abnormalities are only reported in approximately 30% of myeloma patients[42]. Our project cohort showed a failure rate of 14.3%, and an abnormality rate of 24.4%. Karyotyping requires a high level of skill and training, carries a degree of analytical subjectivity and is time-consuming which contribute to the slow and expensive nature of this technology. The cumulative effect of these factors excludes karyotyping as a robust method for the genetic assessment of myeloma patient samples.

FISH has the ability to utilise interphase cells, eliminating the problems associated with metaphase capture. FISH allows the analysis of increased cell numbers; 50-200 cells not being an unreasonable analysis level. FISH can also be performed on bone marrow samples that have undergone a CD138+ve cell selection process, each of these aspects offer a distinct advantage over karyotypic analysis. FISH is generally considered a robust technique, and in this project, the failure rate was just 0.7%.

We have described a two-step FISH process in this project, including *IGH* rearrangement, *TP53* and 1p/1q in step 1 and then assessment of the partner gene (*FGFR3, MAF* or *CCND1*) in step 2 if *IGH* was shown to be rearranged. The panel does not currently include hyperdiploidy assessment.

Of the 101 cases processed using FISH in this study, 55 patients (54.4%) demonstrated abnormality. Based on the literature, 60% is the approximate FISH abnormality rate expected utilising the probe panel described here[25,42,113]. An abnormality rate of 90% has been reported in myeloma samples where extended FISH panels including hyperdiploidy assessment are employed[42]. Our abnormality rate (54.4%) is likely to reflect our patient cohort which included a range of plasma cell neoplasms, not simply myeloma.

FISH does have limitations such as probe size and false positive rates. Probe sizes range from 150kb-1Mb, and therefore small deletions or mutations within genes or within the probe target cannot be detected. Specific probes have an associated false positive rate, dependant on the probe set formation. This is often low and quoted as 1-4% by the manufacturers. Myeloma genetics can be extremely complex, and this

complexity is represented in the FISH probes which can generate non-standard signal patterns. This reflects a strong need to ensure robust training of personnel involved in this analysis to ensure that non-standard patterns are understood and correctly interpreted.

Although the FISH technique itself is not considered expensive, FISH processing is not conducive to high throughput processing, and the application of multiple probe sets and the time consuming nature of the test and analysis can become costly. Currently, FISH remains the most widely used methodology in the genetic assessment of myeloma.

Within our cohort, 11/14 (78%) cases with FISH abnormalities were cytogenetically normal, confirming that normal non-neoplastic populations had been analysed cytogenetically in these cases. Conversely those patients with abnormal karyotypes and a normal FISH profile, were all hyperdiploid cases and would have been detected with an extended panel, or a further technique to assess hyperdiploidy. Therefore, based on clinical utility, FISH offers a much superior test compared to karyotyping.

In conclusion, small FISH panels offer an efficient and accurate way to detect rearrangements and copy number changes in myeloma. However, FISH can be an expensive and time-consuming technique when multiple probes are used within panels, and has limitations based on probe size and an associated false positive rate, as well as a need to have well trained personnel who fully understand the genomic basis of the signal patterns allowing correct and robust interpretation to be made.

The MRC Holland MLPA kit offers a multiplex technique designed to detect gain and loss of ~60 nucleotide length probes. It is considered a cost-effective technique and can determine both hyperdiploidy and the prognostic copy number changes seen in myeloma using DNA from selected cell populations. Abnormalities were detected at a high level (93.1%) in our cohort. As the probes are so small, MLPA can pick up much smaller deletions than FISH, and the ease of analysis is also attractive. MLPA is part of the analysis undertaken as part of the UK myeloma trials.

However, MLPA can only detect copy number changes, and only in the regions targeted by the probe set. It cannot distinguish between mutation and deletion, as

both scenarios result in non-amplification. Therefore, MLPA would have to be considered alongside a technique to determine *IGH* rearrangement status[118]. Low level abnormal populations can also be problematic, and in the case of multiple myeloma, plasma cell purification to a minimum level of 50% purification would be a prerequisite.

We demonstrated a high failure rate for the technique. This is likely to be related to the requirement for high quality DNA and the MLPA process can be highly sensitive to contaminants in the system. If this technique were to be included in a testing strategy, contingency would be required for those samples with poor quality DNA.

The Affymetrix CytoScan HD DNA array allows whole genome copy-number analysis and LOH detection at a high level of resolution which can targeted at DNA from CD138+ve cell selections. However, balanced translocations and rearrangements, and mutations at a base pair level are not detectable. A number of rearrangements do show small regions of loss and/or gain at, or around, the breakpoint regions, but this cannot be considered definitive evidence of rearrangement. Therefore, if arrays were to be utilised as part of a genomic testing strategy for myeloma, then this would need to be alongside a technique for *IGH* rearrangement detection.

This specific array methodology is highly sensitive to DNA quality, and has shown a high level of variability in the quality of the results. 85.7% of cases in this cohort showed an abnormal result, lower than the reported abnormality rate in the MLPA cohort. Array analysis requires the abnormality to be present in approximately 30% of cells, and even processing DNA extracted from selected CD138+ve cells, some abnormalities will be missed. This has created a degree of discordance with the array results and other techniques used in this study.

Array processing is time consuming and expensive, and it is not considered a high throughput technology. Whilst the arrays do provide highly detailed, good quality information at a whole genome level and potentially more than offers true clinical utility. This, in turn requires extensive, time-consuming analysis resulting in an ultimately expensive technology based on the cost of the arrays and the personnel time.

Implementation of NGS technology into NHS diagnostic laboratories is moving at a very fast pace. This has been encouraged, in part, by the NHS England changes of the delivery of genomics services. There remain issues and concerns with the expense, the magnitude of data, the time-consuming analysis and interpretation, the minimal analytical tools and the issues of consent and ethics relating to the transition towards this technology.

There are number of ways of applying NGS technology in a more manageable way, from both a technological, analytical and financial point of view. NGS panels offer a chance to harness the accuracy and depth of NGS analysis in a more targeted fashion by selecting a set of genes than, for example, WGS or WES. As part of this project, a targeted panel was created to include myeloma genes which could be considered part of a diagnostic testing strategy and osteome genes designed to address the second hypothesis; to assess whether specific bone related genetic signatures can indicate the extent or severity of bone involvement in MM, allowing proactive preventative and protective treatment. Using this approach, the analysis is clearly limited to the set of genes that are involved in the panel and loses the functionality of novel gene discovery.

The panel performed well, providing good, robust data for the genes chosen. The panel was able to show a number of abnormalities consistent with those reported in the literature, and in particular detected those mutations associated with the MAPK and NFκB pathways known to be associated with myeloma is a proportion of cases. The NGS panel has to be considered as additional testing as it does not provide information on copy number changes or rearrangement detection.

This information is not currently used to direct treatment or clinical management of myeloma patients, but this is likely to be incorporated into clinical decision making in the future. As NGS becomes more sophisticated, it can be envisaged that an NGS panel could be designed using both DNA and RNA to assess copy number, rearrangement detection for those commonly seen in myeloma and a specific panel of mutations. However, currently this is not available within the context of NHS diagnostic testing.

In conclusion, the five genetic technologies described and explored here for the genomic assessment of myeloma offer a range of both advantages and disadvantages, all of which should be considered in the context of providing a diagnostic testing strategy for NHS patients, which has to balance the clinical utility of the information provided and the cost and turnaround times of the testing.

Proposed Diagnostic Strategies and Best Practice Recommendations

Based on the assessment of the five technologies described here, understanding the pros and cons, and the limitations of each, it is clear that a single technology solution is not possible for all aspects of the genetic assessment of myeloma. It is also clear that all analysis should be targeted to selected cell populations in order to capture abnormalities present in the plasma cells. FISH is still considered by many to be the gold standard, and indeed, based on the information collected as part of the quality assessment chapter (Chapter 5), FISH is overwhelmingly the most common strategy. However, FISH is not conducive to high throughput processing, and as FISH panels become larger, the expense and the time-consuming nature of the processing and analysis becomes difficult to manage within a busy diagnostics laboratory. We have proposed that the tandem use of FISH and MLPA provide a more cost-effective testing strategy. FISH would be used for the assessment of the *IGH* rearrangements, whereas the copy number targets can be assessed in a higher throughput and multiplexed MLPA assay. It is understood from the assessment of these technologies, that MLPA does have a relatively high failure rate, and a contingency FISH pathway for these failed MLPA samples would be required.

Karyotyping has essentially been excluded from the proposed strategies, as it cannot be targeted to plasma cells (robustly), the abnormality rate is low (so clearly cannot detect the abnormal population in the majority of cases), the resolution is poor and the expertise and time required for analysis makes this technique expensive. Array analysis has also been excluded from our proposed strategies; the technique cannot detect *IGH* rearrangements (although some cases do show copy number changes around the breakpoints) and essentially assess copy number only, and compared to

MLPA, which does similar, is time consuming costly and provides a level of resolution and detail that is not clinically required.

Our proposed testing strategies, also, exclude the assessment of the mutational status of myeloma related genes. Currently, there is no real clinical need for this, although many publications suggest both a prognostic and therapeutic benefit. With regard to future proofing a testing strategy, the NGS mutational assessment would be a requirement. Currently this would mean adding a third technology to the strategy, although as NGS technologies mature, it is possible that a dual DNA and RNA NGS testing strategy could provide the full rearrangement, copy number and mutational assessment required in a bespoke NGS panel. However, the *IGH* rearrangements currently remain difficult to determine using an NGS strategy, therefore it may be that FISH is the most appropriate technique for this assessment for the foreseeable future.

Whilst the proposed testing strategy is currently appropriate based on available technologies in diagnostic laboratories and the clinical utility of myeloma genomics, a more forward thinking and future approach to the genomic diagnosis of myeloma was also proposed considering the use of low coverage WGS, WES and expression profiling[124].

As well as considering the technologies that would be suitable for myeloma genetic testing, it is important to assess the regions that should be targeted by such strategies focussing on those regions associated with a prognostic or therapeutic indication. As part of chapter 3, and indeed through the quality assessment work as part of chapter 5, I have presented my views on the target regions for inclusion ranking them from mandatory, highly recommended and recommended, and then regions that are not currently required, some of which however, are likely to become future requirements. Changes in technology, therapeutic advancements and increased knowledge from trials and publications are likely to mean this remains a dynamic area with a requirement for regular review.

During the period of this PhD project we have seen quite dramatic changes in the way in which genomics services are managed. NHS England have undertaken a reconfiguration process resulting in genomics services being offered by seven

genomics laboratory hubs (GLHs) across England. This process has had the aim of creating efficiency and financial savings, but also introducing consistency to the services offered. This has been managed through a procurement process against a specified test directory of services for which delivery will be mandatory. It is recognised that the initial test directory did not include all required tests, and steps are now in place in which an expert group of clinicians and scientists have been assigned to each of three test directories; rare disease, cancer and pharmacogenomics, in order to manage the changes and amendments to those test directories going forward. I am now part of the cancer test directory expert group.

Currently, for myeloma the test directory includes assessment of the *IGH* rearrangements (including five partners *FGFR3, MAF, MAFB, CCND1* and *CCND3*), deletion of 1p (*CDKN2C*), gain of 1q (*CKS1B*), *TP53* deletion, *MYC* rearrangement and hyperdiploidy assessment. It also includes suggests a multi-target NGS panel to include variant detection of seven genes: *KRAS, NRAS, BRAF, TP53, DIS3, FAM46C* and *IRF4*. Although the current test directory does not make any reference to this for myeloma, it has since been recognised that a level of 'essential' and 'desirable' should be added to the test directory, and we are likely to see this for myeloma in future iterations. Whilst the test directory indications do not entirely support the proposed testing strategy from this project, it does give some indication of what testing would be commissioned from NHS England.

As part of the GenQA consortium, from the initial survey described in chapter 5, and with the changes proposed by NHS England, it was considered that generalised guidelines were required for the whole of haemato-oncology diagnostic testing. Led by Katrina Rack (GenQA), we have completed and published recommendations in Leukemia. My role in this was to complete the myeloma recommendations section. These recommendations are brief, and production of myeloma specific guidelines including additional detail are also in production, but have not yet been published. These guidelines would cover the genomic targets required as discussed in the general guidelines, but also cover in more detail the technologies that could be employed. There are disease specific concerns that could also be covered, such as whether to limit testing to myeloma patients or whether MGUS patients would also

benefit, whether follow up genetic testing is appropriate or whether this would be better managed with flow cytometry, consideration of the age of patients to be tested, the requirement for testing specific cell selections, and understanding the consequences and how to manage reporting if cell selection is not possible. The recommendations also cover reporting results and discusses appropriate turnaround times and quality assessment.

Case Report Discussion

Three cases were presented to demonstrate a range of abnormalities and clinical scenarios that have made up the patient cohort in this study.

Patient #113 had plasma cell leukaemia and showed a complex karyotype with an *IGH-MAFB* rearrangement. Although *IGH* rearrangements are seen in approximately 60% of myeloma patients, the *IGH-MAFB* rearrangement is rare, seen in only ~2% of patients and considered one of the poor prognostic *IGH* rearrangements. The t(14;20) translocation was seen karyotypically and then confirmed by FISH. Loss of 1p, gain of 1q and loss of chromosome 13 were seen karyotypically and confirmed by MLPA and array. A rearrangement of chromosome 8 seen cytogenetically raised the suspicion of a *MYC* rearrangement, which was confirmed by FISH. Array analysis provided additional information to support a more complex non-standard rearrangement involving chromosome 2 and duplication of chromosome 8. This case report demonstrates the additive value of utilising the techniques described in this project, providing practical evidence of the strengths and weaknesses of each technique. Each additional technique provides additional information to the overall genomics picture creating a deeper and clearer understanding of the abnormalities present. The time-consuming nature, and cost of this strategy for each case precludes diagnostic use but demonstrates the cumulative effect of each technology and provides information for both the first and third hypotheses.

Patient #175 describes a patient with plasma cell myeloma and a hyperdiploid karyotype. Hyperdiploidy is seen in approaching 50% of myeloma patients. This patient also had an unbalanced rearrangement of chromosomes 1 and 17, which

results in loss of the short arm of chromosome 17 and gain of the long arm of chromosome 1, both of which are considered poor risk markers according to Boyd *et al*, and would be classified as a 'double hit' myeloma in this context[112]. However, perhaps more interestingly, this patient also shows evidence (non-clonally by cytogenetic analysis) of a possible jumping translocation in which the long arm of chromosome 1 was shown to be associated with a number of different chromosomal partners. The FISH pattern offered further evidence of the clonal nature of this phenomenon with gain of *CKS1B* on the long arm of chromosome 1 seen at a dramatically higher level that the loss of *TP53*. MLPA also confirmed the hyperdiploidy and gain of chromosome 1q, but did not show any evidence of *TP53* loss, highlighting one of the downfalls of this technique, i.e. the inability to detect low level rearrangements. Jumping translocations are well documented, but a rare and interesting finding, however, this abnormality has no real clinical impact.

Finally, the third case report described patient #233 who presented with plasma cell myeloma with a hyperdiploid karyotype with a gain of the long arm of chromosome 1. A few months following the myeloma diagnosis, this patient was also diagnosed with lung cancer. NGS panel analysis of this patient showed a hypermutated profile with an approximately 50-fold increase in the number of variants seen when compared to the remaining patients in the cohort. Further examination, and exclusion of possible sequencing artefacts and treatment causes, suggested this was likely to be a true finding. Hypermutation is a rare finding, but small numbers of cases with dramatically increased variant numbers have now been described in a number of cancer types. The expectation from this type of genetic profile would be a poor outlook, however, this patient continues to be treated for both her myeloma and concurrent lung cancer and is clinically well.

NGS Analysis of the Osteome

Bone disease associated with myeloma is seen in approximately two thirds of patients at diagnosis, and is involved in up to 90% of patients throughout the disease course. This bone phenotype can, however, be extremely variable, as can the

treatment of these bone manifestations. Bisphosphonates are readily used, but although prevention and protection from further skeletal events has been shown, there is little evidence of reversing or repairing existing bone disease. Anabolic therapies currently used in osteoporosis may have a role in myeloma and have been shown to stimulate bone formation. Objective two of the thesis aimed to explore the hypothesis that a relationship between genetic mutations associated with the osteome and the likelihood or extent of the bone disease associated with the patient's myeloma exists. A relationship of this type could potentially highlight a cohort of patients that could be treated more proactively or intensively to prevent or slow the bone disease associated with the myeloma. A number of genes associated with the osteome were identified and added to the bespoke NGS panel created as part of this project. The cohort assessed using the NGS panel was made up of 24 patients, 12 with bone lesions and 12 who demonstrated no evidence of bone lesions.

Excluding one patient, in which the NGS failed completely, high risk germline variants were detected in six different genes present in seven of the remaining 23 patients. Three of the seven patients belonged to the patient cohort with bone lesions and four of the patients demonstrated no bone lesions. Two somatic variants were detected in the CD138+ve cell selection DNA, one each in two patients, but neither variant was considered pathogenic. It is acknowledged that the sample numbers were extremely low, but these results provide no evidence to suggest an association or relationship of risk variants and bone damage. High risk variants in the germline samples were present in both the bone lesion and non-bone lesion cohorts.

The discussion in chapter 4 details reasons why an association was not seen in this study; the number of patients included in this analysis cohort, the genes and indeed the number of genes chosen as part of the osteome cohort, the level of filtering required to make a manageable data set, the depth of analysis undertaken on each variant, the tissues available for analysis and the selection of the bone lesion vs non-bone lesion cohort. Repeating this analysis adjusting a number of these factors; a larger patient cohort, with follow up data on the bone lesions over the course of their disease, an increased number of genes, and a deeper level of variant

assessment looking at variants in all gene regions would address a number of issues seen in this part of the analysis. A similar project could also be undertaken using RNAseq instead of the osteome gene panel. The RNAseq assessment allows analysis of the gene expression levels, it covers all genes in single sequencing event.

A larger number of genes were associated with variants in the germline samples, and assessing associations that may provide a predisposition to a worse bone phenotype would really require extensive GWAS studies, like those described by Houlston *et al*[74], who assessed over 7,000 myeloma patients and over 200,000 control samples. Those sample numbers and levels of analysis are beyond the scope of a project like this.

In conclusion, the work presented as part of chapter 4 on the osteome has allowed some initial work to be completed to explore the presence of variants in bone related genes in both the germline and somatic samples processed using the NGS osteome panel. The power of the current study is not strong enough to show positive or negative relationships, however, this exploratory work has not been able to demonstrate any evidence to support a hypothesis that relationships exist between the genetic mutations in bone related genes and the likelihood or extent of myeloma related bone damage. Within the somatic changes seen, no variants were considered pathogenic, and those germline variants associated with risk, and indeed high risk, were seen in both the bone lesion and the non-bone lesion cohorts. We had considered this work could direct treatment to a cohort of patients likely to have a poorer bone phenotype, as long term treatment with bisphosphonates has some drawbacks, but perhaps this opens up the discussion about the more widespread use of anabolic therapies in myeloma which have shown the ability to stimulate bone formation and therefore the ability to counteract the bone devastation caused by the disease in many patients.

Quality Assessment Scheme for Myeloma

Objective three describes the more translational aspects of this project by way of proposing strategies to provide an all-encompassing diagnostic genetic testing panel,

introducing a quality assessment scheme for myeloma genetic testing in collaboration with colleagues within NEQAS (GenQA), both with the aim of influencing best practice through GenQA and the Association for Clinical Genetic Science (ACGS) to encourage a more consistent, equitable and harmonised approach to diagnostic testing.

Chapter 5 describes the successful introduction and implementation of an EQA scheme for genetic diagnosis in myeloma through GenQA and my involvement on the Haematological-Oncology SAG, and demonstrates the impact that more translational research like this can have. The EQA scheme ran for two years as a pilot scheme and has now run for a further five years with full EQA scheme status. As well as providing a forum for validation of the genetic services aimed at providing myeloma genetic testing, the EQA scheme goes further providing an educational component to guide participants, which in turn has the effect of reducing the *ad hoc* nature of genetic testing.

The EQA scheme was first offered as a pilot scheme in 2014. The EQA scheme participation has gradually increased over the scheme period from 39 in 2014 to 64 in 2018, showing a 64% increase in participation from 2014 to 2018. 2018 saw participation from all UK laboratories offering myeloma testing, and a number of European and International laboratories. Although the performance status over the five years has fluctuated, there is an overall downward trend in poor performance from 15.4% in 2014 to 4.7% in 2018, demonstrating a positive effect of the educational component provided by the scheme and increased awareness of the requirements of a diagnostic genomics strategy for myeloma over this time period. It should be noted that the scheme has continued in 2019, and will be performed again in the Autumn of 2020.

As part of this EQA scheme, a survey was issued in 2014 alongside the pilot, and then a further similar survey was issued in 2018 alongside the fifth EQA scheme. The inaugural scheme survey confirmed the suspicion that the genetic diagnosis in myeloma at this time was indeed *ad hoc*, and demonstrated that the quality and extent of testing provided was inequitable. This highlighted the requirement, not only for the educational component provided by the EQA scheme, but for the

production of best practice recommendations in this area. These have now been published as described.

Over the period that the EQA scheme has run, we have seen a degree of harmonisation of genetic testing, with a more consistent, equitable approach to the strategies employed by diagnostic genetic laboratories, demonstrating the impact that EQA can have through participation and education. Although all laboratories are now offering what are considered to be the essential tests in the genomic testing of myeloma, not all laboratories are carrying out this analysis on plasma cell populations, (which is considered best practice) and the extent and quality of testing can still be considered inequitable. The majority of laboratories only employ FISH methodologies, but the 2018 scheme and survey show a small number of laboratories using both MLPA and array, and indeed four laboratories describing the future validation of NGS technologies as part of their testing regimes in line with the chnanges proposed by the new NHS England test directory.

Conclusions and Further work

At the outset, this project set out three objectives:

Objective 1: To compare five different genetic technologies, readily available in a diagnostic genetic laboratory, and their ability to identify genetic signatures associated with plasma cell neoplasms.

Objective 2: To explore and assess possible or potential relationships between the genetic signatures associated with bone related genes, the 'osteome', and the likelihood or extent of bone damage associated with the patient's myeloma disease.

Objective 3: To propose testing strategies for an all-encompassing diagnostic genetic panel for use in a diagnostic genetic laboratory, to introduce a quality assessment scheme for myeloma genetic testing and consider the requirement for best practice guidelines to create a more consistent, equitable harmonisation of diagnostic genetic testing for MM within the UK.

Approximately 100 myeloma patient samples were processed for genetic testing. Resources and indeed sample quantities were not available for all samples to be processed with all technologies; the majority of samples were processed for both karyotyping and FISH, approximately half had MLPA analysis, one third had DNA array analysis and one quarter were processed for NGS analysis. The results from this work are presented to complete objective one, including a number of interesting cases reports. A number of genetic changes have been shown to provide valuable information about disease prognosis and are beginning to guide treatment decisions. The results from these techniques were assessed based on the quality, effectiveness and utility of the diagnostic information provided within the financial and time constraints required of a diagnostic testing scenario. Understanding the techniques and the results provided allowed diagnostic testing strategy proposals to be presented, as well as looking to the future changes we are likely to see in this area, covering parts of objective three.

The NGS work completed as part of the technology platform assessment was extended to include a number of 'osteome' genes. Chapter 4 focussed on this work, specifically addressing objective two aiming to explore the possibility of association with high risk variants in these bone genes and more severe bone phenotypes seen in patients with myeloma. Looking at these variants within both a somatic and germline context showed no evidence of association, although it is recognised that the sample numbers and experimental design was unlikely to demonstrate true association. It was considered that understanding the likelihood of whether a myeloma patient would exhibit a severe bone phenotype at the start of their disease could help target bisphosphonate therapy which is associated with some side effects and therefore treatment is often limited. This does open up the discussion of the blanket use of anabolic therapies in myeloma which have been shown stimulate bone formation in patients with other bone disorders.

Objective three involved working with colleagues in NEQAS (GenQA) to introduce and implement an EQA scheme for genetic testing in myeloma. This was successfully implemented in 2014 as a pilot EQA initially, and has continued to run as part of the establishment schemes on a yearly basis since 2016. This included an educational

component to the scheme. Assessing the role and performance of the different genetic technologies associated with objective one, provided valuable information to guide the translational aspects of the project described in objective three. This allowed informed decisions to be made based on this information and on clinical guidelines to propose appropriate diagnostic genetic testing panel so for myeloma that not only provide clinical utility, but that also offer a diagnostic and financial fit in line with current NHS services. A survey was undertaken alongside the scheme in 2014 and then repeated in 2018. We were able to demonstrate the impact that participation in the scheme, including the educational component, had on the strategies and testing programmes undertaken in the UK, with all laboratories conforming to the guidelines based on results from the 2018 survey. Collaboration with our NEQAS (GenQA) colleagues has also resulted in the production of best practice recommendations which we published recently in Leukemia.

A number of areas of this project would benefit from further work. Objective two focussed on changes seen in the osteome, genes associated with bone formation and maintenance. The numbers of patients assessed did not offer the required power to assess a true relationship in this area. We looked at variants in somatic tissue and in the germline. The somatic tissue showed no pathogenic variants, and although a small number of potentially pathogenic variants were detected in the germline samples, they were seen equally in both bone and non-bone lesion cohorts. In order to determine germline variants that may carry a predisposition to a more severe bone phenotype, GWAS studies would be more appropriate requiring dramatically increased sample sizes, beyond the scope of this project. It may also be interesting to consider a similar project using RNAseq instead of the osteome gene panel. The RNAseq assessment allows analysis of the gene expression levels, and it covers all genes in single sequencing event.

The work to introduce a NEQAS (GenQA) EQA scheme for genetic testing in myeloma has proven a success, and it is important to maintain this scheme allowing it to flex with the changes we are likely to see in both technology and breadth of testing in the coming years.

Finally, although publication of best practice recommendations provides an excellent outcome for this project, the current guidelines are broad and cover genetic testing and reporting for a range of leukaemia types. There are more specific and more detailed considerations associated with genetic testing and reporting in myeloma, and I would consider that more in depth myeloma specific recommendations are required. Work towards this goal is ongoing.

Acknowledgements

Dr Andrew Chantry, Supervisor

Dr Clive Buckle and Professor Angela Cox, Co-supervisors

Sheffield Myeloma Research Team (SMaRT), University of Sheffield.

Staff at Sheffield Diagnostic Genetics Service (SDGS)

Haematological Malignancy Diagnostic Service (HMDS)

Dr Paul Heath, Catherine Gelsthorpe & Mike Spiller, Sheffield Institute for Translational Neuroscience (SITraN)

Dr Elsie Place, Research NGS lead at SDGS

Dr Matt Parker, Bioinformatics support, Sheffield Institute for Translational Neuroscience (SITraN)

Dr Ros Hastings, Katrina Rack, Genomics Quality Assessment (GenQA)

Ryan, Hannah, Mum, and Dad - my family, for their encouragement, support, patience and understanding throughout this project - it would not have been possible without them.

Appendices:

Appendix 1: Ethics Statement & Letter

Ethics Statement:

All procedures involving animals has been approved by the Home Office (PPL 70/8799) and the University of Sheffield's Animal Ethics Committee. Patient cells will be acquired with appropriate ethical permission (REC reference: 05/Q2305/96). All participants will provide written consent to participate in this study. Original consent forms will be stored in a secure location, and patient demographics and disease features entered into an encrypted database governed by the Research and Development Service Sheffield Teaching Hospitals, NHS Foundation Trust UK. This consent procedure was approved by the South Sheffield Research Ethics Committee in August 2005 and subsequently ratified by the NHS Health Research Authority, National Research Ethics Committee Yorkshire and the Humber - Sheffield in November 2012 & 2016.

Copy of Ethics Letter, held by Dr Chantry, Programme Grant PI, and PhD project supervisor:

NHS
Health Research Authority
Yorkshire & The Humber - Sheffield Research Ethics Committee

Room 001
Jarrow Business Centre
Rolling Mill Road
Jarrow
Tyne & Wear
NE32 3DT

Tel: 0207 104 8282

21 February 2017

Dr Andrew Chantry
Senior Clinical Lecturer
University of Sheffield
Department of Oncology
University of Sheffield Medical School
Beech Hill Road, Sheffield
S10 2RX

Dear Dr Chantry

Study title:	**Studies of the mechanisms of myeloma bone disease.**
REC reference:	**05/Q2305/96**
IRAS project ID;	

This study was given a favourable ethical opinion by the Committee on 17 August 2005.

Research Ethics Committees are required to keep a favourable opinion under review in the light of progress reports and any developments in the study. You should submit a progress report for the study 12 months after the date on which the favourable opinion was given, and then annually thereafter. Our records indicate that a progress report is overdue. It would be appreciated if you could complete and submit the report by no later than one month from the date of this letter.

Guidance on progress reports and a copy of the standard NRES progress report form is available from the Health Research Authority website.

The Health Research Authority website also provides guidance on declaring the end of the study.

If you fail to submit regular progress reports – which is a condition of the favourable ethical opinion – the REC may wish to consider suspending or terminating its opinion.

05/Q2305/96:	**Please quote this number on all correspondence**

Yours sincerely

Kerry Dunbar
REC Assistant

E-mail: nrescommittee.yorkandhumber-sheffield@nhs.net

Appendix 2: Reagent Recipes Associated with Culturing of Bone Marrow & FISH

Bone Marrow Medium

250mls McCoys 5A

1.5ml 200mM L-Glutamine

2.5ml Pen/Strep Antibiotic Mixture

5mls 1M Hepes solution

Filter into this 20mls Fetal Calf Serum

Colcemid

In ready diluted vials 10µg/ml

Use PBS to dilute to concentration required (10mls colcemid to 10mls PBS)

KCl Sweller (Potassium Chloride)

Stock Solution (0.075M)

Dissolve 11.2g of Potassium Chloride power into 2 litres of Deionised Distilled Water (DDW)
Autoclave and store in fridge

Banding Trypsin

2.4g Trypsin 1:250 diluted in 2 Litres Sorensons

Mix together on stirrer until just dissolved, and pot into 100mls aliquots

Gurrs

Add 1 Gurrs tablet to 1 litre DDW.

Mix together well and pH to 6.8

Hoescht 33258

Stock Solution (50µg/ml)

Use Oxford Series-S weighing scales with glass doors and wear face mask.

Add 0.025g of Bisbenzimide (light sensitive) to 500mls Gurrs buffer and mix.

Wrap bottle in foil to avoid exposure to light.

20xSSC (VYSIS)

Dissolve 20xSSC powder (Abbott-Vysis/30-805850/66g) into 200ml DDW

pH to 5.3

2xSSC

900ml DDW and 100ml 20xSSC (Fridge F1)

pH to 7.0

Pepsin

Mix 250mg Pepsin 3,300 U/mg (Sigma/P6887/5g) into 2.5ml DDW

Aliquot into eppendorfs and store at -20^{o}C

Ethanol series

(Hayman Absolute Alcohol A.R. Quality/2.5L)

100% ~ 500ml Ethanol

95% ~ 475ml Ethanol + 25ml DDW

70% ~ 350ml Ethanol + 150ml DDW

Hybridisation Mixture

5ml Formamide (BDH Analar/103266T/2.5L)

10µl Tween 20 (Sigma/P7949/500ml)

1ml 20xSSC

4ml DDW

1g Dextran Sulphate (Sigma/D8906/10g) - add slowly to dissolve

Aliquot into eppendorfs and store at -20^{o}C

0.4xSSCT

970ml DDW and 20ml 20xSSC

3ml NP40 (Abbott-Vysis/30-804820) or Tween 20 (Sigma/P7949/500ml)

Mix together and make final volume up to 1000ml

pH to 7.0 and store in Fridge

Appendix 3: NGS Index Sequences, Final Library sizes & Pooling Information

XT HS libraries (= all CD138 samples + 4 blood libraries):								
Ref No.	SureSelect XT HS index primer	Index sequence (i7)	i5	Peak size	Average size	Conc. (ng/ul)	Qubit conc (ng/ul)	Molarity (nM, from Qubit and average size)
#144	A03	AGCAGGAA	NNNNNNNNNN	294	318	18.8	17.8	84.8
#88	B03	AGCCATGC	NNNNNNNNNN	280	317	18.5	15.1	72.2
#161	C03	TGGCTTCA	NNNNNNNNNN	277	309	15.3	13	63.7
#162	D03	CATCAAGT	NNNNNNNNNN	279	308	16.8	12.9	63.5
#44	E03	CTAAGGTC	NNNNNNNNNN	269	299	17	13.6	68.9
#179	F03	AGTGGTCA	NNNNNNNNNN	283	320	19	16.6	78.6
#180	G03	AGATCGCA	NNNNNNNNNN	287	322	19	14.7	69.2
#184	H03	ATCCTGTA	NNNNNNNNNN	280	312	36.1	30.4	147.6
#136	A01	GTCTGTCA	NNNNNNNNNN	356	392	1.5	1.69	6.5
#140	B01	TGAAGAGA	NNNNNNNNNN	301	374	1.51	1.39	5.6
#191	C01	TTCACGCA	NNNNNNNNNN	323	362	2.12	1.89	7.9
#198	D01	AACGTGAT	NNNNNNNNNN	293	385	1.35	1.42	5.6
#200	E01	ACCACTGT	NNNNNNNNNN	325	372	2.16	2.08	8.5
#199	F01	ACCTCCAA	NNNNNNNNNN	306	369	0.714	0.766	3.1
#202	G01	ATTGAGGA	NNNNNNNNNN	336	380	1.66	1.69	6.7
#209	H01	ACACAGAA	NNNNNNNNNN	290	374	1.53	1.65	6.7
#232	A02	GCGAGTAA	NNNNNNNNNN	298	373	1.37	1.32	5.4
#233	B02	GTCGTAGA	NNNNNNNNNN	336	377	1.68	1.63	6.6
#235	C02	GTGTTCTA	NNNNNNNNNN	337	367	0.907	0.805	3.3
GIAB	A04	CCGTGAGA	NNNNNNNNNN	403	422	1.52	2.5	9.0
#215	F02	TGGTGGTA	NNNNNNNNNN	330	385	4.33	3.37	13.3
#230	G02	ACTATGCA	NNNNNNNNNN	340	386	3.21	3.4	13.3
#234	D02	TATCAGCA	NNNNNNNNNN	337	373	3.54	3.03	12.3
#208	B04	GACTAGTA	NNNNNNNNNN	365	404	4.01	3.12	11.7
#217	C04	GATAGACA	NNNNNNNNNN	356	402	3.53	2.79	10.5
Blood libraries prepared with XT HS								
#202	E04	GGTGCGAA	NNNNNNNNNN	318	350	3.89	3.43	14.8
#179	F04	AACAACCA	NNNNNNNNNN	301	338	1.95	1.58	7.1
#208	G04	CGGATTGC	NNNNNNNNNN	299	353	1.59	1.13	4.9
#217	H04	AGTCACTA	NNNNNNNNNN	289	351	1.35	1.02	4.4

QXT libraries (= 21 blood sample libraries):										
Ref No.	QXT P5 index primer	P5 index sequence	QXT P7 index primer	P7 index sequence	Peak size	Average size	Conc. (ng/ul)	Region molarity (nM)	Qubit conc (ng/ul)	Molarity (nM, from Qubit and average size)
#136	P5 i14	CTCTCTAT	P7 i1	TAAGGCGA	385	441	1.29	4.84	1.31	4.5
#140	P5 i14	CTCTCTAT	P7 i2	CGTACTAG	335	389	2.1	8.9	1.71	6.7
#198	P5 i13	TAGATCGC	P7 i4	TCCTGAGC	384	445	12.3	45.8	12.7	43.2
#200	P5 i13	TAGATCGC	P7 i5	GTAGAGGA	409	457	6.21	22.4	9.82	32.6
#199	P5 i13	TAGATCGC	P7 i6	TAGGCATG	363	423	10.1	39.5	14	50.1
#234	P5 i14	CTCTCTAT	P7 i5	GTAGAGGA	415	459	6.45	23.3	7.05	23.3
#232	P5 i13	TAGATCGC	P7 i1	TAAGGCGA	357	422	4.94	19.4	5.56	20.0
#233	P5 i13	TAGATCGC	P7 i2	CGTACTAG	393	447	7.27	27	9.41	31.9
#209	P5 i14	CTCTCTAT	P7 i8	CAGAGAGG	348	409	4.44		5.58	20.7
#235	P5 i13	TAGATCGC	P7 i3	AGGCAGAA	403	450	5.92		6.99	23.5
#144	P5 i13	TAGATCGC	P7 i4	TCCTGAGC	406	449	6.44		7.02	23.7
#88	P5 i13	TAGATCGC	P7 i5	GTAGAGGA	345	412	4.79		8.72	32.1
#44	P5 i13	TAGATCGC	P7 i8	CAGAGAGG	381	439	9.75		11.6	40.0
#180	P5 i14	CTCTCTAT	P7 i2	CGTACTAG	344	395	7.04		7.22	27.7
#184	P5 i14	CTCTCTAT	P7 i3	AGGCAGAA	334	395	5.46		6.04	23.2
#215	P5 i14	CTCTCTAT	P7 i4	TCCTGAGC	376	432	6.7		6.97	24.4
#230	P5 i14	CTCTCTAT	P7 i7	CTCTCTAC	355	418	3.72		5.67	20.6
#191	P5 i13	TAGATCGC	P7 i3	AGGCAGAA	358	430	20.4		19.2	67.7
#161	P5 i13	TAGATCGC	P7 i6	TAGGCATG	463	483	11.6		13	40.8
#162	P5 i13	TAGATCGC	P7 i7	CTCTCTAC	457	486	16.5		21.7	67.7
GIAB	P5 i14	CTCTCTAT	P7 i8	CAGAGAGG	428	463	10.5		9.39	30.7
GIAB = Genome in a bottle reference set										

NGS lane summary					
		Pooling fraction	Average size	Qubit conc (ng/ul)	**Qubit molarity**
Lane 1 (XT HS)	13 x XT HS CD138 libraries	1/14 each	363	0.673	**2.8**
	4 x XT HS blood libraries	1/4 x 1/14 each			
Lane 2 (XT HS)	14 x XT HS CD138 libraries	1/14 each	380	0.578	**2.3**
Lane 3 (QXT)	External NGS Run	33/100	435	0.547	**1.9**
	External NGS Run	33/100			
	10 x QXT Germline blood libraries	1/100 each			
	External NGS Run	17/100			
Lane 4 (QXT)	External NGS Run	33/100	433	0.544	**1.9**
	External NGS Run	33/100			
	11 x QXT Germline blood libraries	1/100 each			
	External NGS Run	16/100			

Appendix 4: Bioinformatic Reference Files & Software Tools applied to the NGS data

Reference Files:

Genome	Resource	Version
GRCh37	seq	broad-20120813
GRCh37	GA4GH_problem_regions	20180317
GRCh37	capture_regions	20161202
GRCh37	MIG	20150730
GRCh37	prioritize	20160215
GRCh37	dbsnp	150-20170710-1
GRCh37	hapmap	3.3
GRCh37	1000g_omni_snps	2.5
GRCh37	ACMG56_genes	20160810
GRCh37	1000g_snps	2.8
GRCh37	mills_indels	2.8
GRCh37	clinvar	20170905
GRCh37	cosmic	68-20180114
GRCh37	ancestral	20141010
GRCh37	qsignature	20140703
GRCh37	genesplicer	2004.04.03
GRCh37	effects_transcript	16/03/2017
GRCh37	vcfanno	20171008
GRCh37	viral	2017.02.04
GRCh37	transcripts	01/12/2015
GRCh37	RADAR	v2-20180202
GRCh37	srnaseq	20180122
GRCh37	giab-NA12878	v3_3_2
GRCh37	giab-NA24385	v3_3_2-sv_v0.5.0
GRCh37	giab-NA24631	v3_3_2
GRCh37	dream-syn3	04/08/2014
GRCh37	dream-syn4	11/06/2016
GRCh37	giab-NA12878-NA24385-somatic	v3_2_2
GRCh37	twobit	broad-20120813

Software Tools:

Tool	Version	Tool	Version
bamtools	2.4.0	mirdeep2	2.0.0.7
bcbio-nextgen	1.0.9	mutect	1.1.5
bcbio-variation	0.2.6	novoalign	3.07.00
bcftools	1.7	novosort	V3.00.02
bedtools	2.27.1	oncofuse	1.1.1
biobambam	2.0.87	phylowgs	20180317
bioconductor-bubbletree	2.8.0	picard	2.18.4
bowtie2	2.2.8	platypus-variant	0.8.1.1
break-point-inspector	1.5	preseq	2.0.3
bwa	0.7.17	qualimap	2.2.2a
chanjo		rna-star	
cnvkit	0.9.3	rtg-tools	3.9
cufflinks	2.2.1	sailfish	0.10.1
cutadapt	1.16	salmon	0.9.1
fastqc	0.11.7	sambamba	0.6.6
featurecounts	1.4.4	samblaster	0.1.24
freebayes	1.1.0.46	samtools	1.7
gatk	3.8	scalpel	0.5.3
gatk4	4.0.3.0	seqbuster	3.1
gemini	0.20.1	snpeff	4.3.1t
grabix	0.1.8	vardict	2018.04.27
hisat2	2.1.0	vardict-java	1.5.1
htseq	0.9.1	variant-effect-predictor	
lumpy-sv	0.2.14a	varscan	2.4.3
manta	1.4.0	vcflib	1.0.0_rc1
metasv	0.4.0	vt	2015.11.10
		wham	1.7.0.311

Appendix 5: Table Demonstrating Alamut Evidence for Variants

Patient	Variant Numbers	Gene	Effect	HGVS.p	HGVS.c	Alamut
#44	1	*FAM46C*	missense variant	p.Leu244Arg L244R	c.731T>G	Not in Lohrs hotspot list, but in the same region Transversion from T to G in exon 2. Missense substitution: Leu244 is changed to Arg. Nothing on google **Additional information** • Highly conserved nucleotide (phyloP: 4.97 [-14.1;6.4]) • Highly conserved amino acid, up to C. elegans (considering 12 species) • Moderate physicochemical difference between Leu and Arg (Grantham dist.: 102 [0-215]) • This variant is in protein domain: Domain of unknown function DUF1693 • Align GVGD (v2007): C65 (GV: 0.00 - GD: 101.88) • SIFT (v6.2.0): Deleterious (score: 0, median: 3.56) • MutationTaster (v2013): disease causing (prob: 1) PATHOGENIC
#88	1	*ZFHX4*	missense variant	p.Lys2283Thr K2283T	c.6848A>C	Transversion from A to C in exon 10. Missense substitution: Lys2283 is changed to Thr. Nothing on google **Additional information** • Highly conserved nucleotide (phyloP: 4.56 [-14.1;6.4]) • Highly conserved amino acid, up to Chicken (considering 9 species) • Moderate physicochemical difference between Lys and Thr (Grantham dist.: 78 [0-215]) • This variant is in protein domain: Homeobox • Align GVGD (v2007): C65 (GV: 0.00 - GD: 77.74) • SIFT (v6.2.0): Deleterious (score: 0, median: 4.32) • MutationTaster (v2013): disease causing (prob: 1) VUS
#136	2	*NFKB2*	missense variant	p.Asp469Asn D469N	c.1405G>A	Transition from G to A in exon 14. Missense substitution: Asp469 is changed to Asn. Nothing on google **Additional information** • Highly conserved nucleotide (phyloP: 4.89 [-14.1;6.4]) • Highly conserved amino acid, up to Zebrafish (considering 12 species) • Small physicochemical difference between Asp and Asn (Grantham dist.: 23 [0-215]) • Align GVGD (v2007): C0 (GV: 226.93 - GD: 0.00)

					• SIFT (v6.2.0): Tolerated (score: 0.06, median: 3.15) • MutationTaster (v2013): disease causing (prob: 1) VUS
	DIS3	missense variant	p.Arg820Trp R820W	c.2458C>T	Not in Lohrs hotspot list Transition from C to T in exon 18. Missense substitution: Arg820 is changed to Trp. COSM3469577 flagged as a SNP (low frequency) **This variant is known to dbSNP** (151): rs372878316 (validated dbSNP entry). **This variant is known to ESP** (ESP6500SIV2): Eur. Am.: A=0.01% - Afr. Am.: A=0.00% **This variant is known to gnomAD** (2.1) <Exomes+Genomes>: ALL:0.0032% - AMR:0.014% - ASJ:0.0097% - SAS:0.0033% - NFE:0.0016% **Additional information** • Weakly conserved nucleotide (phyloP: 1.42 [-14.1;6.4]) • Highly conserved amino acid, up to C. elegans (considering 13 species) • Moderate physicochemical difference between Arg and Trp (Grantham dist.: 101 [0-215]) • Align GVGD (v2007): C35 (GV: 26.00 - GD: 95.78) • SIFT (v6.2.0): Deleterious (score: 0, median: 3.38) • MutationTaster (v2013): disease causing (prob: 1) VUS/Benign

#140	2	*NRAS*	missense variant	p.Gln61Arg Q61R	c.182A>G	Known G61R hotspot seen in Lohr supplementary data Transition from A to G in exon 3. Missense substitution: Gln61 is changed to Arg. **This variant is known to ClinVar** (July-2020 RCV000431883.1 (Likely pathogenic - Multiple myeloma) **This variant is reported as possibly pathogenic by Uniprot** (view report). **This variant is known to dbSNP** (151): rs11554290 (not validated dbSNP entry - Clinical significance: CLIN_likely_pathogenic,CLIN_pathogenic). **Additional information** • Highly conserved nucleotide (phyloP: 4.89 [-14.1;6.4]) • Highly conserved amino acid, up to Baker's yeast (considering 13 species) • Small physicochemical difference between Gln and Arg (Grantham dist.: 43 [0-215]) • This variant is in protein domains: ○ Small GTPase superfamily ○ Mitochondrial Rho-like ○ Small GTPase superfamily, ARF/SAR type ○ Elongation factor, GTP-binding domain ○ Small GTP-binding protein domain ○ Small GTPase superfamily, Ras type ○ Small GTPase superfamily, Rab type ○ Small GTPase superfamily, Rho type ○ P-loop containing nucleoside triphosphate hydrolase • Align GVGD (v2007): C0 (GV: 223.30 - GD: 39.51) • SIFT (v6.2.0): Deleterious (score: 0, median: 4.32) • MutationTaster (v2013): disease causing (prob: 1) PATHOGENIC
		EGR1	missense variant	p.Asp30Gly D30G	c.89A>G	Not in Lohrs hotspot list, but in the same region Transition from A to G in exon 1. Missense substitution: Asp30 is changed to Gly **Additional information** • Moderately conserved nucleotide (phyloP: 3.19 [-14.1;6.4]) • Highly conserved amino acid, up to Frog (considering 11 species) • Moderate physicochemical difference between Asp and Gly (Grantham dist.: 94 [0-215]) • Align GVGD (v2007): C0 (GV: 213.16 - GD: 64.73) • SIFT (v6.2.0): Deleterious (score: 0, median: 4.32) • MutationTaster (v2013): disease causing (prob: 1) PATHOGENIC

#144	1	*NRAS*	missense variant	p.Tyr64Asp Y64D	c.190T>G	Known Y64D hotspot seen in Lohr supplementary data Transversion from T to G in exon 3. Missense substitution: Tyr64 is changed to Asp. **This variant is known to dbSNP** (151): rs752508313 (validated dbSNP entry). **This variant is known to gnomAD** (2.1) <Exomes>: ALL:0% (**Filter**: AC0;RF) **Additional information** • Highly conserved nucleotide (phyloP: 4.89 [-14.1;6.4]) • Highly conserved amino acid, up to Baker's yeast (considering 13 species) • Large physicochemical difference between Tyr and Asp (Grantham dist.: 160 [0-215]) • This variant is in protein domains: ○ Small GTPase superfamily ○ Mitochondrial Rho-like ○ Small GTPase superfamily, ARF/SAR type ○ Elongation factor, GTP-binding domain ○ Small GTP-binding protein domain ○ Small GTPase superfamily, Ras type ○ Small GTPase superfamily, Rab type ○ Small GTPase superfamily, Rho type ○ P-loop containing nucleoside triphosphate hydrolase • Align GVGD (v2007): C0 (GV: 247.85 - GD: 110.05) • SIFT (v6.2.0): Deleterious (score: 0, median: 4.32) • MutationTaster (v2013): disease causing (prob: 1) PATHOGENIC
#191	1	*CCND1*	missense variant	p.Lys58Asn K58N	c.174G>C	Transversion from G to C in exon 1. Missense substitution: Lys58 is changed to Asn. **Additional information** • Weakly conserved nucleotide (phyloP: 0.37 [-14.1;6.4]) • Moderately conserved amino acid (considering 13 species) • Moderate physicochemical difference between Lys and Asn (Grantham dist.: 94 [0-215]) • This variant is in protein domains: ○ Cyclin, N-terminal ○ Cyclin-like ○ Cyclin A/B/D/E • Align GVGD (v2007): C0 (GV: 247.33 - GD: 0.00) • SIFT (v6.2.0): Tolerated (score: 0.35, median: 2.81) • MutationTaster (v2013): disease causing (prob: 1) VUS

#198	2	*FAT3*	missense variant	p.Tyr797His Y797H	c.2389T>C	Transition from T to C in exon 1. Missense substitution: Tyr797 is changed to His. **Additional information** • Moderately conserved nucleotide (phyloP: 3.35 [-14.1;6.4]) • Highly conserved amino acid, up to Chicken (considering 7 species) • Moderate physicochemical difference between Tyr and His (Grantham dist.: 83 [0-215]) • This variant is in protein domain: Cadherin • Align GVGD (v2007): C65 (GV: 0.00 - GD: 83.33) • SIFT (v6.2.0): Deleterious (score: 0, median: 4.32) MutationTaster (v2013): disease causing (prob: 0.996) VUS
		NF1	frameshift variant	p.Ile679fs	c.2033dupC	Duplication (1 bp) in exon 18. This variation creates a frame shift starting at codon Ile679. The new reading frame ends in a STOP codon at position 21. **This variant is known to ClinVar** (July-2020): RCV000130078.2 (Pathogenic* - Hereditary cancer-predisposing syndrome), RCV000204850.9 (Pathogenic** - Neurofibromatosis), RCV000265986.2 (Pathogenic* - not provided), RCV001009578.1 (Pathogenic* - Neurofibromatosis). **This variant is known to dbSNP** (151): rs1232596244 (validated dbSNP entry). **This variant is known to gnomAD** (2.1) <Exomes>: ALL:0.0016% - SAS:0.0065% - NFE:0.00090% - OTH:0.016% (**Filter**: RF) PATHOGENIC

#199	2	*NRAS*	missense variant	p.Gln61Lys Q61K	c.181C>A	Known G61K hotspot seen in Lohr supplementary data Transversion from C to A in exon 3. Missense substitution: Gln61 is changed to Lys. **This variant is known to ClinVar** (July-2020): RCV000444882.1 (Likely pathogenic - Multiple myeloma), **This variant is reported as possibly pathogenic by Uniprot** (view report). **This variant is known to dbSNP** (151): rs121913254 (validated dbSNP entry - Clinical significance: CLIN_not_provided,CLIN_likely_pathogenic,CLIN_pathogenic,CLIN_drug_response). **This variant is known to gnomAD** (2.1) <Exomes>: ALL:0% (**Filter**: AC0) **Additional information** • Highly conserved nucleotide (phyloP: 5.94 [-14.1;6.4]) • Highly conserved amino acid, up to Baker's yeast (considering 13 species) • Small physicochemical difference between Gln and Lys (Grantham dist.: 53 [0-215]) • This variant is in protein domains: ○ Small GTPase superfamily ○ Mitochondrial Rho-like ○ Small GTPase superfamily, ARF/SAR type ○ Elongation factor, GTP-binding domain ○ Small GTP-binding protein domain ○ Small GTPase superfamily, Ras type ○ Small GTPase superfamily, Rab type ○ Small GTPase superfamily, Rho type ○ P-loop containing nucleoside triphosphate hydrolase • Align GVGD (v2007): C0 (GV: 223.30 - GD: 36.80) • SIFT (v6.2.0): Deleterious (score: 0, median: 4.32) • MutationTaster (v2013): disease causing (prob: 1) PATHOGENIC
		FAM46C	conservative inframe deletion	p.Asn352del	c.1054_1056delAAC	Deletion (3 bps) in exon 2. This variation leads to the loss of residue Asn352. Towards the end of exon 2 in a two exon gene. ?VUS

#200	3	*BRAF*	missense variant	p.Asp594Asn D594N	c.1780G>A	D594G seen in Lohr supplementary data (very similar) Transition from G to A in exon 15. Missense substitution: Asp594 is changed to Asn. **This variant is known to ClinVar** (July-2020): RCV000432575.1 (Likely pathogenic - Multiple myeloma), **This variant is known to dbSNP** (151): rs397516896 (not validated dbSNP entry - Clinical significance: CLIN_likely_pathogenic,CLIN_pathogenic). **Additional information** • Highly conserved nucleotide (phyloP: 6.18 [-14.1;6.4]) • Highly conserved amino acid, up to C. elegans (considering 12 species) • Small physicochemical difference between Asp and Asn (Grantham dist.: 23 [0-215]) • Align GVGD (v2007): C15 (GV: 0.00 - GD: 23.01) • SIFT (v6.2.0): Deleterious (score: 0, median: 3.42) PATHOGENIC
		CDKN2A	missense variant	p.Ser12Leu S12L	c.35C>T	Transition from C to T in exon 1. Missense substitution: Ser12 is changed to Leu. **This variant is known to ClinVar** (July-2020): RCV000232824.1 (Uncertain significance* - Hereditary cutaneous melanoma). **This variant is known to dbSNP** (151): rs141798398 (not validated dbSNP entry - Clinical significance: CLIN_uncertain_significance). **Additional information** • Weakly conserved nucleotide (phyloP: 0.29 [-14.1;6.4]) • Weakly conserved amino acid (considering 12 species) • Large physicochemical difference between Ser and Leu (Grantham dist.: 145 [0-215]) • This variant is in protein domain: Ankyrin repeat-containing domain • Align GVGD (v2007): C0 (GV: 353.86 - GD: 0.00) • SIFT (v6.2.0): Tolerated (score: 0.56, median: 3.43) • MutationTaster (v2013): polymorphism (prob: 1) VUS - weak

		KMT2A	missense variant	p.Gly123Asp G123D	c.368G>A	Transition from G to A in exon 1. Missense substitution: Gly123 is changed to Asp. **This variant is known to gnomAD** (2.1) <Exomes>: ALL:0.00099% - NFE:0.0012% - FIN:0.0053% (**Filter**: RF) **Additional information** • Weakly conserved nucleotide (phyloP: 1.42 [-14.1;6.4]) • Weakly conserved amino acid (considering 13 species) • Moderate physicochemical difference between Gly and Asp (Grantham dist.: 94 [0-215]) • This variant is in protein domain: Methyltransferase, trithorax • Align GVGD (v2007): C0 (GV: 353.86 - GD: 0.00) • SIFT (v6.2.0): Tolerated (score: 0.67, median: 3.60) • MutationTaster (v2013): disease causing (prob: 0.999) VUS
#208	1	*ERBB4*	missense variant	p.Ser454Ile S454I	c.1361G>T	Transversion from G to T in exon 12. Missense substitution: Ser454 is changed to Ile. **Additional information** • Highly conserved nucleotide (phyloP: 4.48 [-14.1;6.4]) • Highly conserved amino acid, up to Zebrafish (considering 14 species) • Large physicochemical difference between Ser and Ile (Grantham dist.: 142 [0-215]) • This variant is in protein domains: ○ EGF receptor, L domain ○ Tyrosine protein kinase, EGF/ERB/XmrK receptor • Align GVGD (v2007): C65 (GV: 0.00 - GD: 141.80) • SIFT (v6.2.0): Deleterious (score: 0, median: 4.32) • MutationTaster (v2013): disease causing (prob: 1) VUS

#230	1	*KRAS*	missense variant	p.Gln61His Q61H	c.183A>C	Known Q61H hotspot seen in Lohr supplementary data Transversion from A to C in exon 3. Missense substitution: Gln61 is changed to His. **This variant is known to ClinVar** (July-2020): RCV000154530.3 (Pathogenic - Non-small cell lung cancer), RCV000424748.1 (Likely pathogenic - Acute myeloid leukemia), RCV001004043.1 (Likely pathogenic - Juvenile myelomonocytic leukemia), RCV000444370.1 (Pathogenic - Neoplasm of the large intestine). **This variant is known to dbSNP** (151): rs17851045 (not validated dbSNP entry - Clinical significance: CLIN_likely_pathogenic,CLIN_pathogenic). **This variant is known to gnomAD** (2.1) <Exomes>: ALL:0.00040% - ASJ:0.0099% **Additional information** • Moderately conserved nucleotide (phyloP: 3.43 [-14.1;6.4]) • Highly conserved amino acid, up to Zebrafish (considering 9 species) • Small physicochemical difference between Gln and His (Grantham dist.: 24 [0-215]) • This variant is in protein domains: ○ Small GTPase superfamily ○ Mitochondrial Rho-like ○ Small GTPase superfamily, ARF/SAR type ○ Elongation factor, GTP-binding domain ○ Small GTP-binding protein domain ○ Small GTPase superfamily, Ras type ○ Small GTPase superfamily, Rab type ○ Small GTPase superfamily, Rho type ○ Ran GTPase ○ P-loop containing nucleoside triphosphate hydrolase • Align GVGD (v2007): C15 (GV: 0.00 - GD: 24.08) • SIFT (v6.2.0): Deleterious (score: 0, median: 4.32) PATHOGENIC
#235	3	*ATM*	stop gained	p.Arg1875*	c.5623C>T	Transition from C to T in exon 38. Nonsense substitution. The reading frame is interrupted by a premature STOP codon. The mRNA produced might be targeted for nonsense mediated decay (NMD). **This variant is known to ClinVar** (July-2020): RCV000236653.4 (Pathogenic* - not provided), RCV000493350.5 (Pathogenic** - Hereditary cancer-predisposing syndrome), RCV000540911.5 (Pathogenic* - Ataxia-telangiectasia syndrome). **This variant is known to dbSNP** (151): rs376603775 (validated dbSNP entry - Clinical significance: CLIN_likely_benign,CLIN_pathogenic). **This variant is known to ESP** (ESP6500SIV2): Eur. Am.: T=0.01% - Afr. Am.: T=0.00% **This variant is known to gnomAD** (2.1) <Exomes>: ALL:0.0020% - NFE:0.0035% - FIN:0.0047% PATHOGENIC

	KRAS	missense variant	p.Gln22Lys Q22K	c.64C>A	Transversion from C to A in exon 2. Missense substitution: Gln22 is changed to Lys. **This variant is known to ClinVar** (July-2020): RCV000433522.1 (Likely pathogenic - Neoplasm of the large intestine), RCV001078206.1 (Likely pathogenic* - Epidermal nevus syndrome). **This variant is known to dbSNP** (151): rs121913236 (not validated dbSNP entry - Clinical significance: CLIN_likely_pathogenic,CLIN_pathogenic). **Additional information** • Highly conserved nucleotide (phyloP: 5.86 [-14.1;6.4]) • Highly conserved amino acid, up to Zebrafish (considering 9 species) • Small physicochemical difference between Gln and Lys (Grantham dist.: 53 [0-215]) • This variant is in protein domains: ○ Small GTPase superfamily ○ Mitochondrial Rho-like ○ Small GTPase superfamily, ARF/SAR type ○ Small GTP-binding protein domain ○ Small GTPase superfamily, Ras type ○ Small GTPase superfamily, Rab type ○ Small GTPase superfamily, Rho type ○ Ran GTPase ○ P-loop containing nucleoside triphosphate hydrolase • Align GVGD (v2007): C45 (GV: 0.00 - GD: 53.23) • SIFT (v6.2.0): Deleterious (score: 0, median: 4.32) PATHOGENIC

	KRAS	missense variant	p.Leu19Phe L19F	c.57G>C	Transversion from G to C in exon 2. Missense substitution: Leu19 is changed to Phe. **This variant is known to ClinVar** (July-2020): RCV000441871.1 (Likely pathogenic - Angiosarcoma), RCV000201922.3 (Pathogenic - OCULOECTODERMAL SYNDROME). **This variant is known to dbSNP** (151): rs121913538 (validated dbSNP entry - Clinical significance: CLIN_uncertain_significance,CLIN_likely_pathogenic). **This variant is known to gnomAD** (2.1) <Exomes>: ALL:0% (**Filter**: AC0) **Additional information** • Moderately conserved nucleotide (phyloP: 2.47 [-14.1;6.4]) • Highly conserved amino acid, up to Zebrafish (considering 9 species) • Small physicochemical difference between Leu and Phe (Grantham dist.: 22 [0-215]) • This variant is in protein domains: ○ Small GTPase superfamily ○ Mitochondrial Rho-like ○ Small GTPase superfamily, ARF/SAR type ○ Small GTP-binding protein domain ○ Small GTPase superfamily, Ras type ○ Small GTPase superfamily, Rab type ○ Small GTPase superfamily, Rho type ○ Ran GTPase ○ P-loop containing nucleoside triphosphate hydrolase • Align GVGD (v2007): C15 (GV: 0.00 - GD: 21.82) • SIFT (v6.2.0): Deleterious (score: 0, median: 4.32) PATHOGENIC

Appendix 6: Table Demonstrating Alamut Evidence for Osteome Variants

Gene	Effect	HGVS.p	HGVS.c	Alamut information
FZD4	missense variant	p.Ile256Val	c.766A>G	Transition from A to G in exon 2. Missense substitution: Ile256 is changed to Val. **This variant is known to ClinVar** (July-2020): RCV000005824.2 (Pathogenic - Retinopathy of prematurity). **This variant is reported as possibly pathogenic by Uniprot** **This variant is known to dbSNP** (151): rs104894223 (validated dbSNP entry - Clinical significance: CLIN_pathogenic). **This variant is known to ESP** (ESP6500SIV2): Eur. Am.: C=0.20% - Afr. Am.: C=0.00% **This variant is known to gnomAD** (2.1) <Exomes+Genomes>: ALL:0.052% - AFR:0.0040% - AMR:0.051% - EAS:0.0050% - NFE:0.088% - FIN:0.040% - OTH:0.055% **Additional information** • Highly conserved nucleotide (phyloP: 5.05 [-14.1;6.4]) • Highly conserved amino acid, up to Tetraodon (considering 11 species) • Small physicochemical difference between Ile and Val (Grantham dist.: 29 [0-215]) • This variant is in protein domains: Frizzled protein / GPCR, family 2-like • Align GVGD (v2007): C25 (GV: 0.00 - GD: 28.68) • SIFT (v6.2.0): Deleterious (score: 0, median: 4.32 • MutationTaster (v2013): disease causing (prob: 1) **VUS**

ICAM1	missense variant	p.Lys56Met	c.167A>T	Transversion from A to T in exon 2. Missense substitution: Lys56 is changed to Met. **This variant is known to ClinVar** (July-2020): RCV000015768.3 (Risk factor - Malaria). **This variant is known to dbSNP** (151): rs5491 (MAF/MinorAlleleCount: T=0.084/36 - Clinical significance: CLIN_risk_factor). **This variant is known to ESP** (ESP6500SIV2): Eur. Am.: T=0.30% - Afr. Am.: T=21.20% **This variant is known to gnomAD** (2.1) <Exomes+Genomes>: ALL:3.28% - AFR:22.60% - AMR:1.20% - ASJ:1.37% - EAS:5.36% - SAS:1.61% - NFE:0.31% - FIN:3.97% - OTH:1.76% **Additional information** • Not conserved nucleotide (phyloP: -1.09 [-14.1;6.4]) • Weakly conserved amino acid (considering 16 species) • Moderate physicochemical difference between Lys and Met (Grantham dist.: 95 [0-215]) • This variant is in protein domains: Intercellular adhesion molecule, N-terminal / Immunoglobulin subtype • Align GVGD (v2007): C0 (GV: 353.86 - GD: 0.00) • SIFT (v6.2.0): Tolerated (score: 0.11, median: 3.18) • MutationTaster (v2013): polymorphism (prob: 1) **VUS**

Appendix 7: GenQA Survey Questions

Genetic Testing in Myeloma Survey 2018

Your laboratory details

Which category does your laboratory belong too?

Diagnostic / Research / Diagnostic and Research / Commercial

What sector does your laboratory belong too?

Public (Hospital, NHS) / Public (University) / Public (other) / Private

What country is your laboratory based in?

What is your GenQA laboratory code?

Myeloma testing

How many Myeloma samples do you analyse per year?

What is the population covered by the laboratory (geographically & numerically)?

What is the average turnaround time for your Myeloma service?

What is the abnormality rate for those patients tested?

What is the failure rate?

Do you offer testing for MGUS patients?

Do you have an age limit for testing Myeloma / MGUS patients?

If so, what is this?

Do you offer genetic testing for follow up samples?

Methodology

Do you employ a method of plasma cell enrichment?

If so, what cell marker is this separation based on?

What method is used?

Do you assess the purity of your samples following plasma cell separation?

If so how?

Do you employ fluorescent cytoplasmic immunoglobulin (cIg) staining?

Which method do you routinely use for Myeloma analysis?

Microarray / SNP array / MLPA / PCR / FISH / Other (specify)

Describe your testing strategy.

What regions are tested for?

If FISH is undertaken, how many cells are scored?

What are the cut off values employed for reporting positivity?

Please select and complete for all that you use

Method	Manufacturer	Version/kit/	FISH probe(s)	Comments

DNA (complete only if applicable)

What is the minimum amount of DNA you require?

Do you require control DNA?

What procedure / methods do you use for extracting DNA

Storage

Are sample stored once the genetic testing is complete?

If so, what samples are stored?

And for what period?

Further comments

Please add any further comments you have regarding your myeloma genetic service

Bibliography

1. Swerdlow SH, Cancer IAfRo, Organization WH: WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues, International Agency for Research on Cancer, 2008

2. Kyle RA, Rajkumar SV: Multiple myeloma. N Engl J Med 351:1860-73, 2004

3. Oyajobi BO: Multiple myeloma/hypercalcemia. Arthritis Res Ther 9 Suppl 1:S4, 2007

4. Palumbo A, Anderson K: Multiple myeloma. N Engl J Med 364:1046-60, 2011

5. Cancer Research UK - Myeloma Statistics, 2017

6. Frank C, Fallah M, Ji J, et al: The population impact of familial cancer, a major cause of cancer. Int J Cancer 134:1899-906, 2014

7. Kyle RA, Elveback LR: Management and prognosis of multiple myeloma. Mayo Clin Proc 51:751-60, 1976

8. Kronke J, Udeshi ND, Narla A, et al: Lenalidomide causes selective degradation of IKZF1 and IKZF3 in multiple myeloma cells. Science 343:301-5, 2014

9. Lu G, Middleton RE, Sun H, et al: The myeloma drug lenalidomide promotes the cereblon-dependent destruction of Ikaros proteins. Science 343:305-9, 2014

10. Tagoug I, Jordheim LP, Herveau S, et al: Therapeutic enhancement of ER stress by insulin-like growth factor I sensitizes myeloma cells to proteasomal inhibitors. Clin Cancer Res 19:3556-66, 2013

11. Vincenz L, Jager R, O'Dwyer M, et al: Endoplasmic reticulum stress and the unfolded protein response: targeting the Achilles heel of multiple myeloma. Mol Cancer Ther 12:831-43, 2013

12. White-Gilbertson S, Hua Y, Liu B: The role of endoplasmic reticulum stress in maintaining and targeting multiple myeloma: a double-edged sword of adaptation and apoptosis. Front Genet 4:109, 2013

13. Moreau P, Attal M, Hulin C, et al: Bortezomib, thalidomide, and dexamethasone with or without daratumumab before and after autologous stem-cell transplantation for newly diagnosed multiple myeloma (CASSIOPEIA): a randomised, open-label, phase 3 study. Lancet 394:29-38, 2019

14. Attal M, Richardson PG, Rajkumar SV, et al: Isatuximab plus pomalidomide and low-dose dexamethasone versus pomalidomide and low-dose dexamethasone in patients with relapsed and refractory multiple myeloma (ICARIA-MM): a randomised, multicentre, open-label, phase 3 study. Lancet 394:2096-2107, 2019

15. Fletcher CDM: Who Classification of Tumours of Soft Tissue and Bone (ed 4th edition), World Health Organization, 2013

16. Bird JM, Owen RG, D'Sa S, et al: Guidelines for the diagnosis and management of multiple myeloma 2011. Br J Haematol 154:32-75, 2011

17. Kuppers R, Klein U, Hansmann ML, et al: Cellular origin of human B-cell lymphomas. N Engl J Med 341:1520-9, 1999

18. LeBien TW, Tedder TF: B lymphocytes: how they develop and function. Blood 112:1570-80, 2008

19. Pasqualucci L, Neumeister P, Goossens T, et al: Hypermutation of multiple proto-oncogenes in B-cell diffuse large-cell lymphomas. Nature 412:341-6, 2001

20. Kuehl WM, Bergsagel PL: Multiple myeloma: evolving genetic events and host interactions. Nat Rev Cancer 2:175-87, 2002

21. Fairfax KA, Kallies A, Nutt SL, et al: Plasma cell development: from B-cell subsets to long-term survival niches. Semin Immunol 20:49-58, 2008

22. Morgan GJ, Walker BA, Davies FE: The genetic architecture of multiple myeloma. Nat Rev Cancer 12:335-48, 2012

23. Shapiro-Shelef M, Calame K: Regulation of plasma-cell development. Nat Rev Immunol 5:230-42, 2005

24. Saxe D, Seo EJ, Bergeron MB, et al: Recent advances in cytogenetic characterization of multiple myeloma. Int J Lab Hematol 41:5-14, 2019

25. Sawyer JR: The prognostic significance of cytogenetics and molecular profiling in multiple myeloma. Cancer Genet 204:3-12, 2011

26. Fonseca R, Bergsagel PL, Drach J, et al: International Myeloma Working Group molecular classification of multiple myeloma: spotlight review. Leukemia 23:2210-21, 2009

27. Bergsagel PL, Chesi M: V. Molecular classification and risk stratification of myeloma. Hematol Oncol 31 Suppl 1:38-41, 2013

28. Chretien ML, Corre J, Lauwers-Cances V, et al: Understanding the role of hyperdiploidy in myeloma prognosis: which trisomies really matter? Blood 126:2713-9, 2015

29. Carroll WL: Safety in numbers: hyperdiploidy and prognosis. Blood 121:2374-6, 2013

30. Smadja NV: Hypodiploidy is a major prognostic factor in multiple myeloma. Blood 98:2229-2238, 2001

31. Quesnel B: Multiple myeloma: all roads lead to cyclin D. Blood 106:1-2, 2005

32. Bergsagel PL, Kuehl WM, Zhan F, et al: Cyclin D dysregulation: an early and unifying pathogenic event in multiple myeloma. Blood 106:296-303, 2005

33. Chesi M, Nardini E, Brents LA, et al: Frequent translocation t(4;14)(p16.3;q32.3) in multiple myeloma is associated with increased expression and activating mutations of fibroblast growth factor receptor 3. Nat Genet 16:260-4, 1997

34. Chesi M, Nardini E, Lim RS, et al: The t(4;14) translocation in myeloma dysregulates both FGFR3 and a novel gene, MMSET, resulting in IgH/MMSET hybrid transcripts. Blood 92:3025-34, 1998

35. Binder M, Rajkumar SV, Ketterling RP, et al: Occurrence and prognostic significance of cytogenetic evolution in patients with multiple myeloma. Blood Cancer J 6:e401, 2016

36. Walker BA, Wardell CP, Melchor L, et al: Intraclonal heterogeneity and distinct molecular mechanisms characterize the development of t(4;14) and t(11;14) myeloma. Blood 120:1077-86, 2012

37. Avet-Loiseau H, Attal M, Moreau P, et al: Genetic abnormalities and survival in multiple myeloma: the experience of the Intergroupe Francophone du Myelome. Blood 109:3489-95, 2007

38. Walker BA, Boyle EM, Wardell CP, et al: Mutational Spectrum, Copy Number Changes, and Outcome: Results of a Sequencing Study of Patients With Newly Diagnosed Myeloma. J Clin Oncol 33:3911-20, 2015

39. Terpos E, Ntanasis-Stathopoulos I, Gavriatopoulou M, et al: Pathogenesis of bone disease in multiple myeloma: from bench to bedside. Blood Cancer J 8:7, 2018

40. Lopez-Corral L, Gutierrez NC, Vidriales MB, et al: The progression from MGUS to smoldering myeloma and eventually to multiple myeloma involves a clonal expansion of genetically abnormal plasma cells. Clin Cancer Res 17:1692-700, 2011

41. Leone PE, Walker BA, Jenner MW, et al: Deletions of CDKN2C in multiple myeloma: biological and clinical implications. Clin Cancer Res 14:6033-41, 2008

42. Ross FM, Avet-Loiseau H, Ameye G, et al: Report from the European Myeloma Network on interphase FISH in multiple myeloma and related disorders. Haematologica 97:1272-7, 2012

43. Scheid C, Reece D, Beksac M, et al: Phase 2 study of dovitinib in patients with relapsed or refractory multiple myeloma with or without t(4;14) translocation. Eur J Haematol 95:316-24, 2015

44. An G, Xu Y, Shi L, et al: t(11;14) multiple myeloma: a subtype associated with distinct immunological features, immunophenotypic characteristics but divergent outcome. Leuk Res 37:1251-7, 2013

45. Chesi M, Bergsagel PL, Shonukan OO, et al: Frequent dysregulation of the c-maf proto-oncogene at 16q23 by translocation to an Ig locus in multiple myeloma. Blood 91:4457-63, 1998

46. Ross FM, Chiecchio L, Dagrada G, et al: The t(14;20) is a poor prognostic factor in myeloma but is associated with long-term stable disease in monoclonal gammopathies of undetermined significance. Haematologica 95:1221-5, 2010

47. Walker BA, Leone PE, Chiecchio L, et al: A compendium of myeloma-associated chromosomal copy number abnormalities and their prognostic value. Blood 116:e56-65, 2010

48. Avet-Loiseau H, Li C, Magrangeas F, et al: Prognostic significance of copy-number alterations in multiple myeloma. J Clin Oncol 27:4585-90, 2009

49. Walker BA, Wardell CP, Brioli A, et al: Translocations at 8q24 juxtapose MYC with genes that harbor superenhancers resulting in overexpression and poor prognosis in myeloma patients. Blood Cancer J 4:e191, 2014

50. Cremer FW, Bila J, Buck I, et al: Delineation of distinct subgroups of multiple myeloma and a model for clonal evolution based on interphase cytogenetics. Genes Chromosomes Cancer 44:194-203, 2005

51. Tricot G, Barlogie B, Jagannath S, et al: Poor prognosis in multiple myeloma is associated only with partial or complete deletions of chromosome 13 or abnormalities involving 11q and not with other karyotype abnormalities. Blood 86:4250-6, 1995

52. Avet-Loiseau H, Daviet A, Saunier S, et al: Chromosome 13 abnormalities in multiple myeloma are mostly monosomy 13. British journal of Haematology 111:1116-1117, 2000

53. Gutierrez NC, Castellanos MV, Martin ML, et al: Prognostic and biological implications of genetic abnormalities in multiple myeloma undergoing autologous stem cell transplantation: t(4;14) is the most relevant adverse prognostic factor, whereas RB deletion as a unique abnormality is not associated with adverse prognosis. Leukemia 21:143-50, 2007

54. Drach J, Ackermann J, Fritz E, et al: Presence of a p53 gene deletion in patients with multiple myeloma predicts for short survival after conventional-dose chemotherapy. Blood 92:802-9, 1998

55. Chang H, Ning Y, Qi X, et al: Chromosome 1p21 deletion is a novel prognostic marker in patients with multiple myeloma. Br J Haematol 139:51-4, 2007

56. Boyd KD, Ross FM, Walker BA, et al: Mapping of chromosome 1p deletions in myeloma identifies FAM46C at 1p12 and CDKN2C at 1p32.3 as being genes in regions associated with adverse survival. Clin Cancer Res 17:7776-84, 2011

57. Wu KL, Beverloo B, Lokhorst HM, et al: Abnormalities of chromosome 1p/q are highly associated with chromosome 13/13q deletions and are an adverse prognostic factor for the outcome of high-dose chemotherapy in patients with multiple myeloma. Br J Haematol 136:615-23, 2007

58. Walker BA, Leone PE, Jenner MW, et al: Integration of global SNP-based mapping and expression arrays reveals key regions, mechanisms, and genes important in the pathogenesis of multiple myeloma. Blood 108:1733-43, 2006

59. Chapman MA, Lawrence MS, Keats JJ, et al: Initial genome sequencing and analysis of multiple myeloma. Nature 471:467-72, 2011

60. Weaver CJ, Tariman JD: Multiple Myeloma Genomics: A Systematic Review. Semin Oncol Nurs 33:237-253, 2017

61. Lohr JG, Stojanov P, Carter SL, et al: Widespread genetic heterogeneity in multiple myeloma: implications for targeted therapy. Cancer Cell 25:91-101, 2014

62. Bolli N, Avet-Loiseau H, Wedge DC, et al: Heterogeneity of genomic evolution and mutational profiles in multiple myeloma. Nat Commun 5:2997, 2014

63. Lionetti M, Barbieri M, Todoerti K, et al: Molecular spectrum of BRAF, NRAS and KRAS gene mutations in plasma cell dyscrasias: implication for MEK-ERK pathway activation. Oncotarget 6:24205-17, 2015

64. Keats JJ, Fonseca R, Chesi M, et al: Promiscuous mutations activate the noncanonical NF-kappaB pathway in multiple myeloma. Cancer Cell 12:131-44, 2007

65. Annunziata CM, Davis RE, Demchenko Y, et al: Frequent engagement of the classical and alternative NF-kappaB pathways by diverse genetic abnormalities in multiple myeloma. Cancer Cell 12:115-30, 2007

66. Maura F, Bolli N, Angelopoulos N, et al: Genomic landscape and chronological reconstruction of driver events in multiple myeloma. Nat Commun 10:3835, 2019

67. Walker BA, Mavrommatis K, Wardell CP, et al: Identification of novel mutational drivers reveals oncogene dependencies in multiple myeloma. Blood 132:587-597, 2018

68. Hoang PH, Cornish AJ, Dobbins SE, et al: Mutational processes contributing to the development of multiple myeloma. Blood Cancer J 9:60, 2019

69. Dimopoulos K, Gimsing P, Gronbaek K: Aberrant microRNA expression in multiple myeloma. Eur J Haematol 91:95-105, 2013

70. Broderick P, Chubb D, Johnson DC, et al: Common variation at 3p22.1 and 7p15.3 influences multiple myeloma risk. Nat Genet 44:58-61, 2012

71. Koura DT, Langston AA: Inherited predisposition to multiple myeloma. Ther Adv Hematol 4:291-7, 2013

72. Chubb D, Weinhold N, Broderick P, et al: Common variation at 3q26.2, 6p21.33, 17p11.2 and 22q13.1 influences multiple myeloma risk. Nat Genet 45:1221-5, 2013

73. Went M, Sud A, Forsti A, et al: Identification of multiple risk loci and regulatory mechanisms influencing susceptibility to multiple myeloma. Nat Commun 9:3707, 2018

74. Went M, Kinnersley B, Sud A, et al: Transcriptome-wide association study of multiple myeloma identifies candidate susceptibility genes. Hum Genomics 13:37, 2019

75. Walker BA, Wardell CP, Melchor L, et al: Intraclonal heterogeneity is a critical early event in the development of myeloma and precedes the development of clinical symptoms. Leukemia 28:384-390, 2014

76. Keats JJ, Chesi M, Egan JB, et al: Clonal competition with alternating dominance in multiple myeloma. Blood 120:1067-76, 2012

77. Hebraud B, Caillot D, Corre J, et al: The translocation t(4;14) can be present only in minor subclones in multiple myeloma. Clin Cancer Res 19:4634-7, 2013

78. Rasche L, Chavan SS, Stephens OW, et al: Spatial genomic heterogeneity in multiple myeloma revealed by multi-region sequencing. Nat Commun 8:268, 2017

79. Gerlinger M, Rowan AJ, Horswell S, et al: Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med 366:883-92, 2012

80. Anwer F, Gee KM, Iftikhar A, et al: Future of Personalized Therapy Targeting Aberrant Signaling Pathways in Multiple Myeloma. Clin Lymphoma Myeloma Leuk 19:397-405, 2019

81. Pawlyn C, Davies FE: Toward personalized treatment in multiple myeloma based on molecular characteristics. Blood 133:660-675, 2019

82. Uchiyama H, Barut BA, Mohrbacher AF, et al: Adhesion of human myeloma-derived cell lines to bone marrow stromal cells stimulates interleukin-6 secretion. Blood 82:3712-20, 1993

83. Giuliani N, Rizzoli V, Roodman GD: Multiple myeloma bone disease: Pathophysiology of osteoblast inhibition. Blood 108:3992-6, 2006

84. Paton-Hough J, Tazzyman S, Evans H, et al: Preventing and Repairing Myeloma Bone Disease by Combining Conventional Antiresorptive Treatment With a Bone Anabolic Agent in Murine Models. J Bone Miner Res 34:783-796, 2019

85. Green AC, Lath D, Hudson K, et al: TGFbeta Inhibition Stimulates Collagen Maturation to Enhance Bone Repair and Fracture Resistance in a Murine Myeloma Model. J Bone Miner Res 34:2311-2326, 2019

86. Chantry AD, Heath D, Mulivor AW, et al: Inhibiting activin-A signaling stimulates bone formation and prevents cancer-induced bone destruction in vivo. J Bone Miner Res 25:2633-46, 2010

87. Heath DJ, Chantry AD, Buckle CH, et al: Inhibiting Dickkopf-1 (Dkk1) removes suppression of bone formation and prevents the development of osteolytic bone disease in multiple myeloma. J Bone Miner Res 24:425-36, 2009

88. Russow G, Jahn D, Appelt J, et al: Anabolic Therapies in Osteoporosis and Bone Regeneration. Int J Mol Sci 20, 2018

89. Walker RE, Lawson MA, Buckle CH, et al: Myeloma bone disease: pathogenesis, current treatments and future targets. Br Med Bull 111:117-38, 2014

90. Lane NE, Kelman A: A review of anabolic therapies for osteoporosis. Arthritis Res Ther 5:214-22, 2003

91. Ring ES, Lawson MA, Snowden JA, et al: New agents in the Treatment of Myeloma Bone Disease. Calcif Tissue Int 102:196-209, 2018

92. Lath DL, Buckle CH, Evans HR, et al: ARQ-197, a small-molecule inhibitor of c-Met, reduces tumour burden and prevents myeloma-induced bone disease in vivo. PLoS One 13:e0199517, 2018

93. Shendure J, Ji H: Next-generation DNA sequencing. Nat Biotechnol 26:1135-45, 2008

94. Bentley DR: Whole-genome re-sequencing. Curr Opin Genet Dev 16:545-52, 2006

95. Raffan E, Semple RK: Next generation sequencing--implications for clinical practice. Br Med Bull 99:53-71, 2011

96. Braggio E, Fonseca R: Unraveling the multiple myeloma genome in the next-generation sequencing era: challenges to translating knowledge into the clinic. Expert Rev Hematol 4:579-81, 2011

97. Egan JB, Shi CX, Tembe W, et al: Whole-genome sequencing of multiple myeloma from diagnosis to plasma cell leukemia reveals genomic initiating events, evolution, and clonal tides. Blood 120:1060-6, 2012

98. Kohlmann A, Grossmann V, Nadarajah N, et al: Next-generation sequencing - feasibility and practicality in haematology. Br J Haematol 160:736-53, 2013

99. Matthijs G, Souche E, Alders M, et al: Guidelines for diagnostic next-generation sequencing. Eur J Hum Genet 24:2-5, 2016

100. DE R: Human Cytogenetics: malignancy and acquired abnormalities (ed Third Edition), Oxford University Press, 2001

101. HUGO: The resource for approved human gene nomenclature,

102. Kortum KM, Langer C, Monge J, et al: Targeted sequencing using a 47 gene multiple myeloma mutation panel (M(3) P) in -17p high risk disease. Br J Haematol 168:507-10, 2015

103. Li H, Durbin R: Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754-60, 2009

104. McKenna A, Hanna M, Banks E, et al: The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 20:1297-303, 2010

105. Lai Z, Markovets A, Ahdesmaki M, et al: VarDict: a novel and versatile variant caller for next-generation sequencing in cancer research. Nucleic Acids Res 44:e108, 2016

106. Saunders CT, Wong WS, Swamy S, et al: Strelka: accurate somatic small-variant calling from sequenced tumor-normal sample pairs. Bioinformatics 28:1811-7, 2012

107. den Dunnen JT, Dalgleish R, Maglott DR, et al: HGVS Recommendations for the Description of Sequence Variants: 2016 Update. Hum Mutat 37:564-9, 2016

108. Smadja NV, Louvet C, Isnard F, et al: Cytogenetic study in multiple myeloma at diagnosis: comparison of two techniques. Br J Haematol 90:619-24, 1995

109. McGowan-Jordan J, Simons, A., Schmid, M. : ISCN 2016: An International System for Human Cytogenomic Nomenclature. . 2016

110. Pierre RV, Hoagland HC: Age-associated aneuploidy: loss of Y chromosome from human bone marrow cells with aging. Cancer 30:889-94, 1972

111. NICE GUIDELINES Myeloma: diagnosis and management, (ed http://nice.org.uk/guidance/ng35). National Institute for Health and Care Excellence, National Institute for Health and Care Excellence, 2016, pp 25

112. Boyd KD, Ross FM, Chiecchio L, et al: A novel prognostic model in myeloma based on co-segregating adverse FISH lesions and the ISS: analysis of patients treated in the MRC Myeloma IX trial. Leukemia 26:349-55, 2012

113. Rack K, Vidrequin S, Dargent JL: Genomic profiling of myeloma: the best approach, a comparison of cytogenetics, FISH and array-CGH of 112 myeloma cases. J Clin Pathol 69:82-6, 2016

114. Palumbo A, Avet-Loiseau H, Oliva S, et al: Revised International Staging System for Multiple Myeloma: A Report From International Myeloma Working Group. J Clin Oncol 33:2863-9, 2015

115. Rajkumar SV, Dimopoulos MA, Palumbo A, et al: International Myeloma Working Group updated criteria for the diagnosis of multiple myeloma. The Lancet Oncology 15:e538-e548, 2014

116. Chng WJ, Dispenzieri A, Chim CS, et al: IMWG consensus on risk stratification in multiple myeloma. Leukemia 28:269-77, 2014

117. Chng WJ, Van Wier SA, Ahmann GJ, et al: A validated FISH trisomy index demonstrates the hyperdiploid and nonhyperdiploid dichotomy in MGUS. Blood 106:2156-61, 2005

118. Alpar D, de Jong D, Holczer-Nagy Z, et al: Multiplex ligation-dependent probe amplification and fluorescence in situ hybridization are complementary techniques to detect cytogenetic abnormalities in multiple myeloma. Genes Chromosomes Cancer 52:785-93, 2013

119. Stuppia L, Antonucci I, Palka G, et al: Use of the MLPA assay in the molecular diagnosis of gene copy number alterations in human genetic diseases. Int J Mol Sci 13:3245-76, 2012

120. Huntly BJ, Bench A, Green AR: Double jeopardy from a single translocation: deletions of the derivative chromosome 9 in chronic myeloid leukemia. Blood 102:1160-8, 2003

121. Sawyer JR, Tian E, Thomas E, et al: Evidence for a novel mechanism for gene amplification in multiple myeloma: 1q12 pericentromeric heterochromatin mediates breakage-fusion-bridge cycles of a 1q12 approximately 23 amplicon. Br J Haematol 147:484-94, 2009

122. Campbell BB, Light N, Fabrizio D, et al: Comprehensive Analysis of Hypermutation in Human Cancer. Cell 171:1042-1056 e10, 2017

123. Rack KA, van den Berg E, Haferlach C, et al: European recommendations and quality assurance for cytogenomic analysis of haematological neoplasms. Leukemia 33:1851-1867, 2019

124. Kuiper R, Broyl A, de Knegt Y, et al: A gene expression signature for high-risk multiple myeloma. Leukemia 26:2406-13, 2012

125. Coleman RE: Skeletal complications of malignancy. Cancer 80:1588-94, 1997

126. Kariyawasan CC, Hughes DA, Jayatillake MM, et al: Multiple myeloma: causes and consequences of delay in diagnosis. QJM 100:635-40, 2007

127. Morgan GJ, Davies FE, Gregory WM, et al: First-line treatment with zoledronic acid as compared with clodronic acid in multiple myeloma (MRC Myeloma IX): a randomised controlled trial. Lancet 376:1989-99, 2010

128. Morgan GJ, Davies FE, Gregory WM, et al: Long-term follow-up of MRC Myeloma IX trial: Survival outcomes with bisphosphonate and thalidomide treatment. Clin Cancer Res 19:6030-8, 2013

129. Katz BZ: Adhesion molecules--The lifelines of multiple myeloma cells. Semin Cancer Biol 20:186-95, 2010

130. Skrtic A, Korac P, Kristo DR, et al: Immunohistochemical analysis of NOTCH1 and JAGGED1 expression in multiple myeloma and monoclonal gammopathy of undetermined significance. Hum Pathol 41:1702-10, 2010

131. Morra L, Moch H: Periostin expression and epithelial-mesenchymal transition in cancer: a review and an update. Virchows Arch 459:465-75, 2011

132. Tarragona M, Pavlovic M, Arnal-Estape A, et al: Identification of NOG as a specific breast cancer bone metastasis-supporting gene. J Biol Chem 287:21346-55, 2012

133. Huang H: Matrix Metalloproteinase-9 (MMP-9) as a Cancer Biomarker and MMP-9 Biosensors: Recent Advances. Sensors (Basel) 18, 2018

134. Claustres M, Kozich V, Dequeker E, et al: Recommendations for reporting results of diagnostic genetic testing (biochemical, cytogenetic and molecular genetic). Eur J Hum Genet 22:160-70, 2014

135. Hastings R, Howell R, Betts D, et al: Guidelines and Quality Assurance for Acquired Cytogenetics 2013

136. Pratt GA, N: Multiple myeloma cytogenetic testing in the UK and Ireland: guidance vs practice. Myeloma Hub Connect, 2018