Irina-Alexandra Bacila

Understanding the Metabolic Problems in Patients with Congenital Adrenal Hyperplasia

Thesis submitted to The University of Sheffield

for the degree of Doctor of Philosophy

August 2022

TABLE OF CONTENTS

Acknowledgements

This thesis and the wonderful experience of the research work that led to it would not have been possible without my academic supervisor Nils Krone, who gave me the opportunity to work within his research team, supporting and guiding me throughout the process. My second supervisor Vincent Cunliffe made me welcome in his laboratory and patiently helped me develop my skills and knowledge in molecular biology. I am also grateful to the other members of the Krone/Cunliffe laboratory team Lara Oberski, Nan Li, James Oakes and Jake Pavely, who provided invaluable help and teaching, while creating a very pleasant working environment. My thanks also go to my colleague Neil Lawrence, whose passion for numbers materialised as very helpful statistical advice for the clinical research in my project. A great number of people are to be thanked for technical support, including Maggie Glover and the aquarium staff. Finally, I would like to thank my wonderful family for their continual support, warm encouragement, and unfaltering faith in my potential.

List of publications

Bacila, I. A., et al. 2019. "Update on adrenal steroid hormone biosynthesis and clinical implications." *Arch Dis Child* **104**(12): 1223-1228. DOI: 10.1136/archdischild-2017-313873

Bacila, I., et al 2019. "Measurement of salivary adrenal-specific androgens as biomarkers of therapy control in 21-hydroxylase deficiency." *J Clin Endocrinol Metab* **104**(12): 6417-6429. DOI: 10.1210/jc.2019-00031

Bacila, I., et al 2021. "Interrenal development and function in zebrafish." *Mol Cell Endocrinol* **535**: 111372. DOI: 10.1016/j.mce.2021.111372

Bacila, I., et al 2022. "Health status of Children and young persons with congenital adrenal hyperplasia in the UK (CAH-UK): a cross-sectional multi-centre study." *Eur J Endocrinol* **187**(4):543-553. DOI: 10.1530/EJE-21-1109

List of abbreviations

11OHA: 11-hydroxyandrostenedione
11KT: 11-ketotestosterone
17OHP: 17-hydroxyprogesterone
21OHD: 21-hydroxylase deficiency
A4: androstenedione
acrpl*:* adiponectin gene
ACTH: Adrenocorticotropic hormone
AI: adrenal insufficiency
BMI: body mass index
BSA: Body surface area
CAH: congenital adrenal hyperplasia
cDNA: complementary DNA
cebpα: CCAAT/enhancer-binding protein gene
CRH: corticotrophin releasing hormone
CRISPR/Cas9: clustered regularly interspaced short palindromic repeats/ Cas9 endonuclease
CYP: cytochrome P450
CYP11A1 or P450scc: cholesterol side-chain cleavage enzyme
CYP11B1 or P450c11: 11beta-hydroxylase
CYP11B2 or P450c11AS: aldosterone synthase
Cyp11c1: zebrafish 11beta-hydroxylase
CYP17A1 or P450c17: 17 alpha-hydroxylase
CYP19A1 or P450aro: P450 aromatase
CYP21A2 or P450c21: 21-hydroxylase
DAX1: dosage-sensitive sex reversal, adrenal hypoplasia critical region, chromosome X, gene 1
DEGs: differentially expressed genes
DHEA: dehydroepiandrosterone
DHEAS: dehydroepiandrosterone sulphate
DHT: dihydrotestosterone
dpf: days post fertilisation
dβh*:* dopamine-beta-hydroxylase
EDTA: Ethylenediaminetetraacetic acid
ef1a*:* elongation factor 1-alpha
ELISA: enzyme-linked immunosorbent assay
fabp11a: fatty acid binding protein 11a gene
fasn*:* fatty acid synthase gene
FDX1: ferredoxin
FDXR: ferredoxin reductase
ff1b*: ftz-fl* gene

fkbp5: FK506 binding protein 5
GC: glucocorticoid
gck: glucokinase gene
GLUT4: insulin-regulated glucose transporter
GO: gene ontology
GR: glucocorticoid receptor
GSEA: gene set enrichment analysis
hadha: Hydroxyacyl-CoA Dehydrogenase gene
HC: hydrocortisone
HDL: high density lipoproteins
HE: Haematoxylin and eosine
hmgcr: 3-hydroxy-3-methylglutaryl coenzyme A reductase gene
HOMA-IR: Homeostatic Model of Insulin Resistance
HPA: hypothalamic-pituitary-adrenal axis
HPC: high performance computing
hpf: hours post fertilisation
HPI: hypothalamic-pituitary-interrenal axis
HSD/KSR: hydroxysteroid dehydrogenases/ketosteroid reductases
HSD11B1: 11-hydroxysteroid dehydrogenase type 1
HSD11B2: 11-hydroxysteroid dehydrogenase type 2
HSD17B3: 17-hydroxysteroid dehydrogenase type 3
HSD3B2: 3-hydroxysteroid dehydrogenase type 2
HSD17B5 or AKR1C3: 17-hydroxysteroid dehydrogenase type 5 or aldo-keto-reductase C3
insa: insulin gene
irs2a: insulin receptor substrate 2a gene
JAK-STAT: Janus kinase - signal transducer and activator of transcription (pathway)
LC-MS/MS: liquid chromatography tandem mass spectrometry
LDL: low density lipoproteins
lep: leptin gene
LFC: log fold change
lpl; lipoprotein lipase gene
LXR: liver X receptor
MAPK: mitogen activated protein kinase (pathway)
MC: mineralocorticoid
Mc2r: zebrafish melanocortin receptor 2
MR: mineralocorticoid receptor
MRAP: melanocortin receptor 2 accessory protein
mz: mass ratio
mTOR: mechanistic target of rapamycin
NAD: nicotinamide adenine dinucleotide
NADPH: nicotinamide adenine dinucleotide phosphate (reduced form)
NAFLD: non-alcoholic fatty liver disease
NES: normalised enrichment score
NR3C1: Nuclear Receptor Subfamily 3 Group C Member 1 (GR)
NR3C2: Nuclear Receptor Subfamily 3 Group C Member 2 (MR)
PAP1: phosphatidic acid phosphatase 1
PBS: phosphate buffered saline
PCA: principal component analysis
PEPCK: Phosphoenolpyruvate carboxykinase
PCN: probabilistic quotient normalisation
PCOS: polycystic ovary syndrome
PKB/Akt: phosphatidylinositol 3-kinase / protein kinase B
POMC: proopiomelanocortin
POR: P450 oxidoreductase
pparg: peroxisome-proliferator activated receptor gamma gene
Prox1: encoding a prospero-related homeodomain protein
PVN: paraventricular nucleus
qPCR: quantitative polymerase chain reaction
rt: retention time
RT-qPCR: reverse transcription - quantitative polymerase chain reaction

rx3: retinal homebox gene 3
SDS: standard deviation score
***SF1/Ad4BP (NR5A1)*:** steroidogenic factor 1 human gene
SREBP: sterol-regulatory-element binding protein
srebf1: sterol regulatory element-binding protein 1 gene
StAR: steroidogenic acute regulatory protein
SULT2A: sulphonotransferase
T: testosterone
TALEN: transcription activator-like effector nuclease
VLDL: very low-density lipoproteins
UPLC-MS: Ultra-high-performance liquid-chromatography – mass spectrometry
WISH: whole mount *in situ* hybridisation
WT: wild type (zebrafish)
Wt1: Wilms tumour 1 transcription factor

List of figures

List of tables

Summary

Introduction. The prevalence of metabolic disease is increased in congenital adrenal hyperplasia (CAH) due to 21-hydroxylase deficiency (21OHD), however, there is limited knowledge regarding the mechanisms through which it occurs. **Aim and objectives:** (1) Explore the health status of CAH juvenile patients, insisting on early signs of metabolic disease; (2) Assess the metabolic phenotype and transcriptomic analysis of a 21-hydroxylase deficient zebrafish model; (3) Compare the clinical and zebrafish data to identify evolutionarily conserved and distinct aspects of the molecular mechanisms underlying the pathophysiology of CAH. **Methodology.** Clinical and biological data from a UK-wide multicentre study involving 107 children with 21OHD were analysed. The phenotype and transcriptome of *cyp21a2-/-* zebrafish mutants were compared to wild-type siblings. **Results.** There was increased weight gain in CAH patients compared to controls, a small number of patients also having abnormal lipid profiles. Leptin, adiponectin and lipid metabolites correlated with the glucocorticoid dose, body mass index, Homeostatic Model Assessment for Insulin Resistance and plasma androgens. Adult *cyp21a2-/-* zebrafish mutants were larger, with more body fat compared to controls. Differential gene expression and overrepresentation analysis of gene ontology (GO) terms showed significant dysregulation of several metabolic processes in larvae and adult livers, with marked downregulation of genes involved in mitochondrial oxidative phosphorylation and energy homeostasis. **Conclusions.** In CAH, metabolic comorbidities develop as a combined effect of GC replacement and GC deficiency.

Although frank metabolic disease is rare in children with CAH, subtle changes are already present in the blood metabolites and may serve as early markers of metabolic risk. Insights from *cyp21a2-/-* zebrafish complete the clinical findings, indicating that cortisol deficiency has complex metabolic effects, likely centred around the suppression of ATP synthesis and energy homeostasis. Better treatment and monitoring strategies are needed in CAH, targeting the development of metabolic disease at an early stage.

1. Introduction

Glucocorticoid (GC) deficiency can be caused by a variety of conditions encountered in human pathology. Due to the broad involvement of GC in multiple physiological functions and most importantly in energy regulation, severe GC deficiency is incompatible with life. The introduction of GC therapy in the 1950's, while saving many lives, has added complexity to a challenging clinical problem. This relates largely to the fact that synthetic steroid replacement cannot mimic the complex physiological functions of the hypothalamo-pituitary-adrenal axis, which regulates GC synthesis, release and turnover in response to a wide range of external and internal stimuli.

Congenital adrenal hyperplasia (CAH) due to 21-hydroxylase deficiency (21OHD) is the most common metabolic cause of inherited GC deficiency, occurring in 1 in 18,000 live births in the UK (El-Maouche, et al., 2017). Patients require life-long treatment with synthetic GC, often in supraphysiological doses. Synthetic GCs have been shown to be associated with a wide variety of side-effects, incurring significant metabolic and cardio-vascular comorbidities and limiting the duration and quality of life in treated patients (Bancos, et al., 2015). Thus, metabolic complications experienced by GC deficient patients on replacement therapy, such as obesity and insulin resistance, have been traditionally attributed to GC overexposure. By contrast, the pathophysiological consequences of intermittent daily GC deficiency (that the patients experience throughout their lives) on glucose and fat metabolism remain unclear.

Recent evidence (Torky, et al., 2021) showed that metabolic morbidities in patients with CAH start during childhood, before puberty. Importantly, the incidence of the metabolic problems was not associated with the GC type, dose or regime (circadian vs reverse

circadian) (Torky, et al., 2021). Moreover, animal research using zebrafish has provided *in vivo* evidence of the impact of GC deficiency on several metabolic pathways in the absence of GC treatment (Weger, et al., 2018). Preliminary transcriptomic analysis conducted by our laboratory employing 21-hydroxylase (Cyp21a2) deficient zebrafish larvae found significant dysregulations of energy pathways that persist even after hydrocortisone treatment (unpublished data). These findings would suggest that several common comorbidities prevalent in patients with CAH may in part relate to the effect of GC deficiency on different metabolic pathways. Thus, my project originates from the hypothesis that better understanding of the pathophysiology of metabolic disease in CAH will help improve monitoring strategies in guiding steroid replacement, leading to better health outcomes.

1.1 Human adrenal steroid biosynthesis

(Modified from published review (Bacila, et al. 2019))

Steroidogenesis represents a complex process through which cholesterol is converted into active steroid hormones, by undergoing multiple chemical reactions that are regulated by several enzymes and cofactors. While steroidogenesis takes place primarily in the adrenal glands and gonads, other tissues are also known to have steroidogenic capacity, in particular, the placenta and the brain (Miller and Auchus, 2011; Miller, 2017). Other structures, such as the adipose tissue and the liver, were also found to express several steroid-converting enzymes, which demonstrates their role in the steroid metabolism (Miller, 2017). However, there are site-specific characteristics of steroid hormone synthesis, caused by differences in the expression of steroidogenic enzymes and co-factors (Miller and Auchus, 2011). Thus, the biosynthesis of gluco- and

mineralocorticoids takes place mainly in the adrenal, while sex hormones (androgens, oestrogens, progesterone) are produced by both adrenals and gonads.

There are two major groups of **steroidogenic enzymes**: cytochrome P450 (CYP) and hydroxysteroid dehydrogenases/ketosteroid reductases (HSD/KSR) (Miller and Auchus, 2011). **CYP enzymes** catalyse unidirectional oxidative reactions by activating molecular oxygen, due to the presence of a *haeme* prosthetic group in their structure. There are two types of CYP enzymes: type 1, present in the mitochondria and using ferredoxin (FDX1) also termed adrenodoxin and ferredoxin reductase (FDXR) as a cofactor, and type 2, found in the endoplasmic reticulum that are dependent on cofactor P450 oxidoreductase (POR). The CYP enzymes important for human steroidogenesis are cholesterol side-chain cleavage enzyme (CYP11A1 or P450scc), 11 beta-hydroxylase (CYP11B1 or P450c11), aldosterone synthase (CYP11B2 or P450c11AS), 17 alpha-hydroxylase (CYP17A1 or P450c17), 21-hydroxylase (CYP21A2 or P450c21) and P450 aromatase (CYP19A1 or P450aro). **HSD/KSR enzymes** are able *in vitro* to catalyse bidirectionally oxidative-reductive reactions, however, *in vivo* they act preferentially towards oxidation or reduction (Miller and Auchus, 2011). Depending on the type of reactions they catalyse, HSD enzymes are classified into dehydrogenases (using $NAD^{+)}$ and reductases (using NADPH). Several HSDs are required for steroid synthesis: the 3-hydroxysteroid dehydrogenase type 2 (HSD3B2), the 11-hydroxysteroid dehydrogenase type 1 and type 2 (HSD11B1 and HSD11B2), and a series of 17-hydroxysteroid dehydrogenases (HSD17) (Miller and Auchus, 2011).

Unlike other types of hormones, such as peptides, steroids are stored in very small amounts, their availability relying mainly on immediate synthesis, a function that becomes time-critical in the stress response. Steroidogenic cells are capable of *de novo* steroid synthesis from cholesterol. A key role in the initiation of steroidogenesis is played by the Steroidogenic Acute Regulatory protein (StAR), directing the flow of cholesterol from the outer to the inner mitochondrial membrane. Here, it is converted to pregnenolone by CYP11A1, the side-chain cleavage enzyme, which represents the first step of steroid biosynthesis (**Figure 1.1**). Briefly, adrenal steroid hormone biosynthesis follows three major pathways, mineralocorticoid, glucocorticoid and androgen hormone synthesis. Each pathway takes place preferentially in a different layer of the adrenal cortex: *zona glomerulosa, zona fasciculata* and *zona reticularis,* respectively, which have specific profiles of enzyme expression. Mineralocorticoid synthesis involves the conversion of pregnenolone to aldosterone, following a sequence of reactions catalysed by HSD3B2, CYP21A2 and CYP11B2 (Turcu and Auchus, 2015). Glucocorticoid synthesis begins with the 17-hydroxylation of pregnenolone by CYP17A1 to 17-hydroxypregnenolone, which is then converted to 17-hydroxyprogesterone (17OHP) by HSD3B2. Under the action of CYP21A2, 17OHP is hydroxylated to 11-deoxycortisol, then CYP11B1 completes the process leading to cortisol. Finally, for the synthesis of androgens in the *zona reticularis,* 17-hydroxypregnenolone is converted to dehydroepiandrosterone (DHEA) by CYP17A1 in the presence of co-factor cytochrome b5, expressed only in the *zona reticularis.* DHEA is as weak androgen, its main role being that of a precursor for more potent androgens (Schiffer, et al., 2018), being converted to androstenedione by HSD3B2. However, DHEA is also partly used by sulphotransferase (SULT2A) to form DHEA sulphate (DHEAS), a

process that regulates the amount of DHEA available for androgen synthesis (Noordam, et al., 2009). Conversion of androstenedione to testosterone in the adrenals is catalysed by 17-hydroxysteroid dehydrogenase type 5 or aldo-keto-reductase C3 (HSD17B5 or AKR1C3) (Nakamura Y, 2009). Once in the circulation testosterone is activated by 5α-reductase dihydrotestosterone (DHT) in target tissues (Schiffer, et al., 2018).

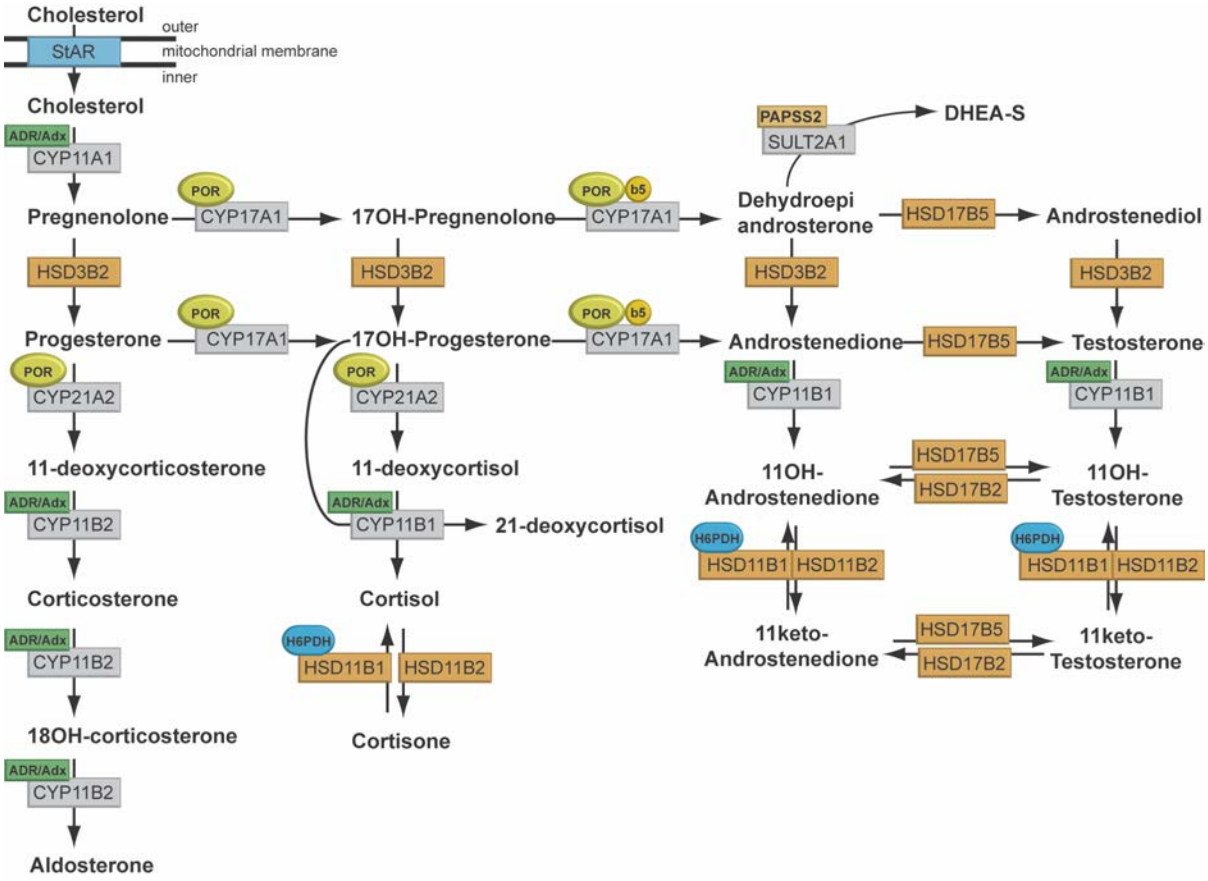


Figure 1.1. Diagram of adrenal steroid biosynthesis
Grey and orange boxes are used to indicate CYP and HSD steroidogenic enzymes, green boxes correspond to adrenodoxin/adrenodoxin reductase, orange box to 3-phosphoadenosine-5-phosphosulfate synthase type 2, yellow ovals to P450 oxidoreductase, orange circles to cytochrome b5 and blue ovals to the coenzyme hexose-6-phosphate dehydrogenase. (Modified from Bacila, et al. 2019)

The adrenal GC synthesis is regulated by an anatomical and functional entity called the hypothalamic – pituitary – adrenal (HPA) axis, or the stress axis **(Figure 1.2).** Centrally, the paraventricular nuclei in the hypothalamus produce corticotropin -

releasing hormone (CRH) which reaches the anterior pituitary, stimulating the release of adrenocorticotropic hormone (ACTH) and regulates the synthesis and posttranslational changes of its precursors, proopiomelanocortin (POMC) (Liyanarachchi, et al., 2017). In turn, ACTH activates the melanocortin 2 receptor (MC2R) in the adrenal cortex, inducing GC (cortisol) synthesis. An essential component of the HPA, MC2R is a G-protein-coupled receptor that is functionally expressed only in the adrenal glands. Its function requires the presence of the melanocortin receptor 2 accessory protein (MRAP), which has essential roles in the assembly of MC2R in the endoplasmic reticulum, its trafficking to the plasma membrane, where it is activated by ACTH, and the MC2R signalling (Novoselova, et al., 2013; Berruien and Smith, 2020). Mutations affecting the MC2R or MRAP lead to severe familial glucocorticoid deficiency (Novoselova, et al., 2013).

The majority of GC actions on different organs and systems is regulated through the activation of the GC receptor (NR3C1), although cortisol also binds to the mineralocorticoid receptor (NR3C2), which plays an important role in the control of the stress axis (Harris, et al., 2013). The HPA axis holds a pivotal role in regulating the management of energy and maintaining body homeostasis in both predictable and unpredictable situations (Gans, et al., 2021). The secretion of both CRH and ACTH follows a circadian pattern, leading to diurnal fluctuation in the cortisol release which has complex impact on multiple functions and processes, including energy management and metabolism. Cortisol exerts negative feedback on the CRH and ACTH secretion from the hypothalamus and pituitary, respectively (Liyanarachchi, et al., 2017). The function of the HPA axis is influenced by multiple factors, including age, gender, ethnicity, diet, stress

and illness (Liyanarachchi, et al., 2017). In its inactivated state, GR is present in the cytoplasm, however, upon binding with the GC it migrates to the nucleus. Here, it has extensive roles in regulating transcription, either by directly binding to the DNA or by linking to other transcription factors (Dinarello, et al., 2021). Thus, GC act to induce or repress the expression of a very wide range of genes, which explains its vast an complex involvement in so many physiological processes and functions.

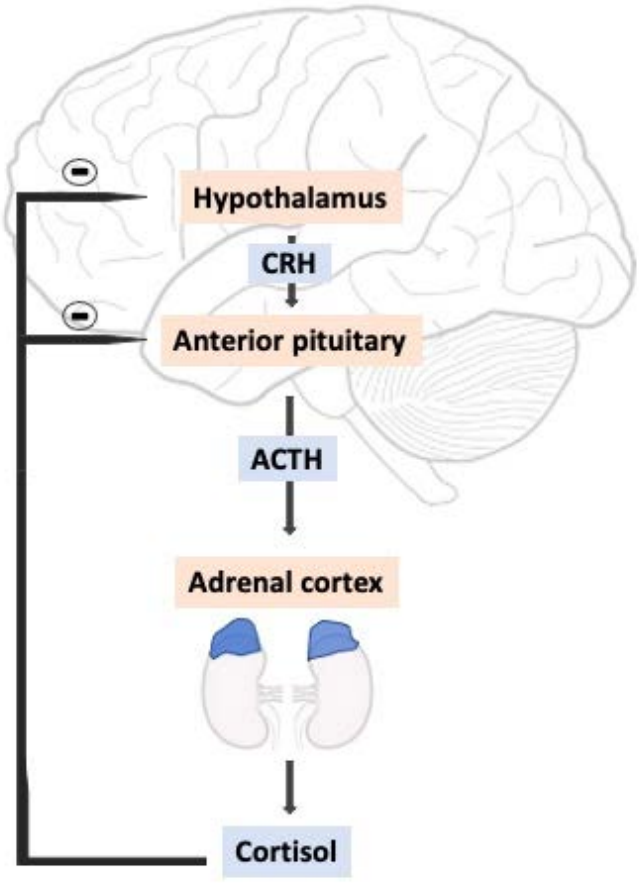


Figure 1.2. Diagram of the hypothalamic – pituitary – adrenal axis
CRH: corticotropin - releasing hormone, ACTH: adrenocorticotropic hormone. (Brain and kidney diagrams used as free public domain from http://www.clker.com)

1.2 The roles of glucocorticoids in human metabolism

Glucocorticoids have considerable impacts on the regulation of many metabolic pathways, although most of the mechanisms through which they contribute to certain physiological and pathological processes remain to be established. In glucose metabolism, GC increase blood glucose, by stimulating gluconeogenesis in the liver and decreasing glucose uptake in the adipose tissue and skeletal muscle (Kuo, et al., 2015), antagonising insulin through these functions. Moreover, GC interfere directly with components of the insulin signalling pathway (glycogen synthase kinase-3, glycogen synthase, glucose transporter type 4 [GLUT4] translocation), further inhibiting insulin-mediated glucose uptake (Hwang and Weiss, 2014). Additionally, GC suppress the post-insulin receptor cascade involving the phosphatidylinositol 3-kinase/protein kinase B (PKB/Akt) and mechanistic target of rapamycin (mTOR) pathways, causing an inhibition of protein synthesis and increased protein degradation (Hwang and Weiss, 2014). GC are known to regulate the functions of the pancreatic cells in secreting insulin and glucagon (Kuo, et al., 2015). They also increase glycogen storage in the liver, however, in the skeletal muscle they inhibit insulin-stimulated glycogen synthesis (Kuo, et al., 2015). The implications of GC in lipid metabolism consist of promoting the differentiation of adipocytes and increasing the uptake and turnover of fatty acids in the adipose tissue. In the liver, GC stimulate phosphoenolpyruvate carboxykinase (PEPCK) leading to glycerol synthesis. Additionally, they induce the activity of lipoprotein lipase, thus increasing the release of fatty acids in the blood, which interferes with glucose utilisation and results in insulin resistance (Macfarlane, et al., 2008). Many studies support the role of GC in modulating central structures involved in food intake and satiety response, as well as their interrelation to adipokines in regulating adipose tissue distribution and functions, however, the mechanisms of action and pathological implications of these aspects are not fully understood (Dallman, et al., 2004; Macfarlane, et al., 2008).

1.3 Congenital adrenal hyperplasia

(Modified from published review (Bacila, et al. 2019))

The term “congenital adrenal hyperplasia” (CAH) encompasses a number of autosomal recessive inborn errors of steroidogenesis caused by deficiencies of different enzymes or co-factors. All these conditions are characterised by impaired cortisol synthesis and some also associate genital ambiguity (**Table 1**). Seven such conditions are currently included in CAH, caused by deficiencies of the following enzymes or co-factors: 21-hydroxylase (CYP21A2, P450c21), 17 alpha-hydroxylase (CYP17A1, P450c17), 11 beta-hydroxylase (CYP11B1, P450c11), 3-hydroxysteroid dehydrogenase type 2 (HSD3B2), P450 oxidoreductase (POR), StAR (lipoid CAH) and P450 side-chain cleavage enzyme (CYP11A1, P450scc). Of these, 21OHD is by far the most common form, accounting for 95% of the cases (El-Maouche, et al., 2017). Co-factor P450 oxidoreductase deficiency presents with impaired function of both CYP17A1 and CYP21A2, and is characterised clinically by Antley-Bixler-like bone malformations. (Miller and Auchus, 2011; El-Maouche, et al., 2017; Krone, et al., 2012)

Table 1. Clinical features of inborn conditions in the CAH group

Enzyme/co-factor	DSD	Affected organ	Deficiency	Excess
CYP21A2	**46,XX**	**Adrenal**	**MC, GC**	**SexH**
CYP11B1	**46,XX**	**Adrenal**	**GC**	**MC, SexH**
CYP17A1	**46,XY**	**Adrenal, gonad**	**GC, SexH**	**MC**
HSD3B2	**46,XY (46,XX)**	**Adrenal, gonad**	**MC, GC, SexH**	
POR	**46,XY + 46,XX**	**Adrenal, gonad, liver**	**GC, SexH**	**(MC)**

StAR	**46,XY**	**Adrenal, gonad**	**MC, GC, SexH**	
CYP11A1	**46,XY**	**Adrenal, gonad**	**MC, GC, SexH**	

CYP21A2: 21-hydroxylase, CYP11B1: 11-hydroxylase, CYP17A1: 17-hydroxylase, HSD3B2: 3-hydroxysteroid dehydrogenase type 2, POR: P450 oxidoreductase, StAR: steroidogenic acute regulatory protein, CYP11A1: P450 side-chain cleavage enzyme, DSD: disorder of sex development, MC: mineralocorticoid, GC: glucocorticoid, SexH: sex steroid hormone, (MC): variable MC excess. (Modified from Bacila, et al. 2019)

Steroid **21-hydroxylase deficiency** (21OHD) leads to GC and mineralocorticoids (MC) deficiency, as well as an excessive accumulation of up-stream precursors of the reactions catalysed by the enzyme, in particular 17OHP, leading to hyperandrogenism. In the absence of cortisol, which normally inhibits ACTH secretion from the pituitary through negative feedback, excessive ACTH will further amplify the altered steroid synthesis in the adrenal. There is a continuum in the severity of presentation in 21OHD, dictated by the amount of residual CYP21A2. However, a clinical classification is used in practice, dividing the condition into classic and non-classic 21OHD. In turn, classic 21OHD is frequently subdivided in salt wasting and simple virilising forms (Nermoen, et al., 2017); the difference between the two subclasses is due to a residual 21-hydroxylase activity of 1-2% in the simple virilising form, preventing the occurrence of the adrenal crisis present in the salt wasting form (El-Maouche, et al., 2017). Cortisol deficiency drives androgen overproduction antenatally, leading to virilisation of the external genitalia in females, which are usually diagnosed immediately after birth due to ambiguous genitalia. In males the condition manifests with adrenal crisis in the neonatal life, in the case of the salt wasting form, or precocious puberty in less severe forms. Both sexes can present advanced skeletal growth, with early closure of the growth plates and reduced adult

height. Patients with non-classic 21OHD have up to 50% residual *in vitro* enzyme activity and thus present milder phenotypes, where the hyperandrogenism is less manifest or asymptomatic (El-Maouche, et al., 2017).

The management of classic 21OHD requires chronic replacement treatment with synthetic GC and, in 75-80% cases, MC. While in other forms of adrenal insufficiency the GC therapy aims to merely replace the absent cortisol, in 21OHD patients usually require higher doses in order to normalise the excessive ACTH secretion and hyperandrogenism. The hormonal therapy remains a significant challenge in the management of CAH, due to the difficulty in mimicking the physiological circadian pattern of the HPA axis, in the context of variable pharmacokinetics of synthetic GC among individuals and in the absence of reliable monitoring strategies (Dauber, et al., 2010). Consequently, throughout a day, patients experience episodes of GC excess alternating with periods of GC deficiency which can occasionally last many hours (**Figure 1.3**). It is known that patients with CAH have increased prevalence of metabolic and cardio-vascular comorbidities (Reisch, et al., 2011; Mooij, et al., 2017), however, it is not clear to what extent the regular exposure to periods of GC deficiency contributes to these problems.

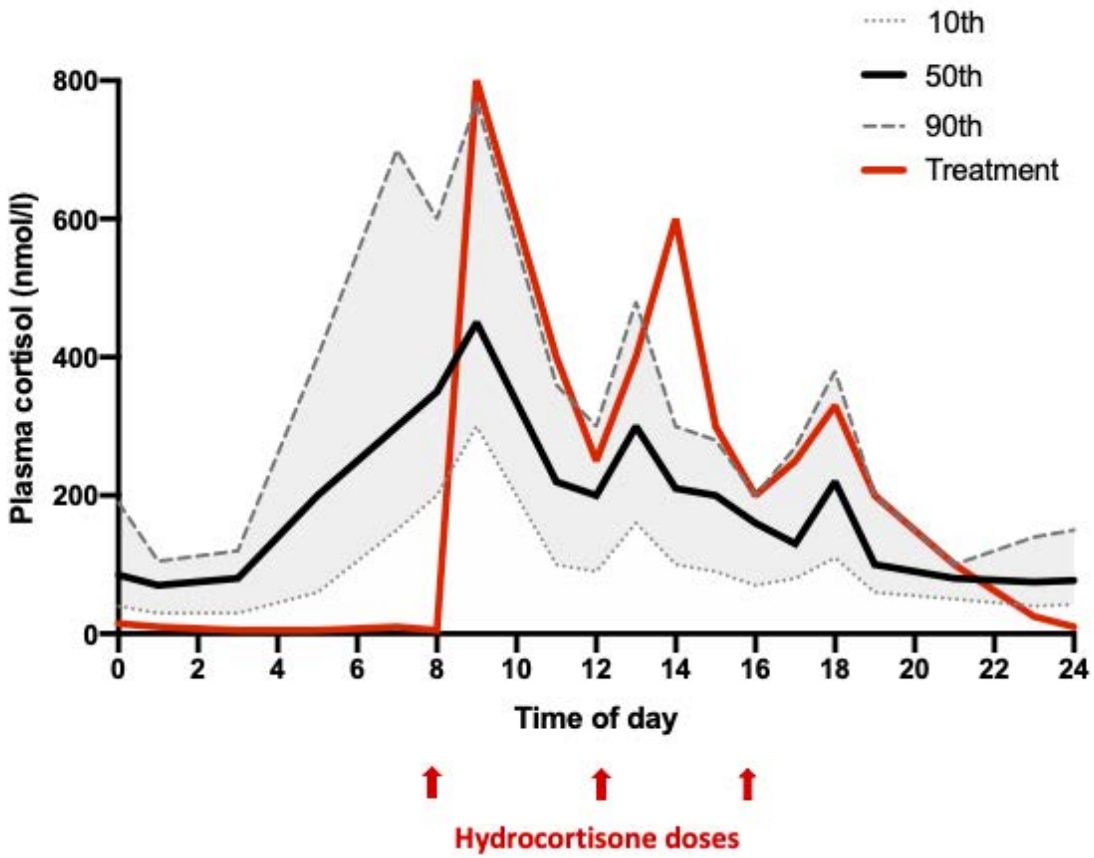


Figure 1.3. Blood cortisol profile over 24 hours
Normal subjects are represented by the black and dotted grey lines (corresponding to the 50th, 10th and 90th percentiles respectively). The red line corresponds to the cortisol pattern of a patient treated with three daily doses of oral hydrocortisone. (Modified from Hindmarsh and Honour. 2020)

1.4 Current evidence on the metabolic effects of glucocorticoid deficiency

To establish what is currently known on the metabolic effects of GC deficiency, I undertook a literature review using the PubMed, Medline and Web of Science databases, which showed that the evidence on the topic is relatively sparse and heterogeneous. A number of studies, involving either human participants or animal models, addressed the topic of GC deficiency in relation to common metabolic pathways (Kawai, 1977; Malerbi, et al., 1988; Parry-Billings, et al., 1990; Blair, et al., 1995; Bishayi and Ghosh, 2003; Poyrazoglu, et al., 2003; Wu, et al., 2004; Christiansen, et al., 2007; Zhang, et al., 2008; Kumari, et al., 2010; Maripuu, et al., 2016; Priyadarshini and Anuradha, 2017; Weger, et

al., 2018), while others explored the cardio-vascular metabolic effects of GC deficiency (van der Sluis, et al., 2012) or the impact of GC deficiency on centres of food intake and satiety control (Müller, et al., 2000; Makimura, et al., 2003; Uchoa, et al., 2009; Uchoa, et al., 2012). A summary of the findings reported by these studies is presented in this subchapter.

1.4.1 Animal studies

1.4.1.1 Studies that explored specifically the effect of glucocorticoid deficiency on metabolism

Several articles reported animal studies *in vivo* or *in vitro* that explored the impact of GC deficiency on metabolic pathways, using different methods. The majority of studies induced GC deficiency by unilateral or bilateral adrenalectomy (Kawai, 1977; Blair, et al.,1995; Bishayi and Ghosh, 2003; Sherwin and Sacca, 1984; Wu, et al., 2004; an der Sluis, et al., 2012) or by using enzyme deficient mutants (Weger, et al., 2018), while in some studies GC receptor (GR) antagonists were used to reproduce GC resistance (Priyadarshini and Anuradha, 2017; Zhang, et al., 2008). The exposure to GC deficiency or resistance was also variable, the experiments being conducted between two (Kawai, 1977) and 30 days (van der Sluis, et al., 2012) after the adrenalectomy and nine hours (Zhang, et al., 2008) to 18 days (Priyadarshini and Anuradha, 2017) after the GR blockade.

Glucose metabolism and insulin signalling pathways. In keeping with the known antagonistic effect of GC on insulin, an *in vivo* study reported increased insulin sensitivity in rats two days after bilateral adrenalectomy (Kawai, 1977). Interestingly, following glucose loading there was reduced secretion of insulin from the pancreatic cells in GC

deficient rats, as well as reduced glucose metabolism in adipocytes under the effect of insulin. These changes were corrected by administration of dexamethasone, implying that GC may enhance the pancreas sensitivity to insulinogenic stimuli and insulin actions in the adipose tissue. Improved insulin sensitivity associated with GC deficiency was also reported by a more recent study (Blair, et al., 1995) exploring the effects of GC deficiency on tissue specific glucose uptake. GC deficient rats were also found to have decreased liver glycogen and glucose production one week after bilateral adrenalectomy. As the glucose uptake was reduced in the adipose tissue, heart and skeletal muscle, the findings support a positive effect of GC deficiency on insulin sensitivity due to a decrease in hepatic glucose production. Similar results were reported by a study that explored the metabolic consequences of GC insufficiency after unilateral adrenalectomy (Bishayi and Ghosh, 2003). Interestingly, the authors reported raised blood glucose concentrations in GC insufficient rats at 10, 20 and 30 days post adrenalectomy. Insulin levels were not measured, which limits the value of the findings regarding the relationship between GC deficiency and insulin sensitivity. However, as the study involved unilateral adrenalectomy, consideration must be given to the potential contribution of the epinephrine secreted by the remaining adrenal in raising the blood glucose level (Sherwin and Sacca, 1984). These findings contrast with those of a more recent publication (Priyadarshini and Anuradha, 2017) where the authors used a GR antagonist to block GC actions in mice. The results only showed reduced insulin resistance in obese mice following a high fructose diet, but not in controls; this introduces the controversial topic of the direct role of GR in the pathogenesis of obesity. One study conducted on rats with bilateral adrenalectomy explored specifically the effects of GC deficiency on the insulin signalling pathways related to protein and glycogen metabolism in the heart muscle (Wu, et al., 2004), by quantifying the phosphorylation of enzymes involved in pathways of

lipolysis, protein and glycogen synthesis, as well as the activity of glycogen synthase and glycogen content in cardiomyocytes. The findings indicated a suppressive effect of GC deficiency on insulin signalling pathways of protein production and lipolysis in the myocardium, as well as decreased glycogen content through abolished insulin induced glycogen synthesis.

Lipid metabolism. GC-deficient rats were found to have reduced body mass and adipose tissue after bilateral adrenalectomy (Blair, et al., 1995). By contrast, a more recent study (Priyadarshini and Anuradha, 2017) using a GR antagonist in mice failed to demonstrate significant impact on body weight or the expression of lipogenic and lipolytic enzymes in the liver. However, GR blockade reduced lipid profile abnormalities and insulin resistance induced by a high fructose diet, suggesting a potential role of the GR in the pathogenesis of metabolic syndrome and obesity. An *in vitro* study (Zhang, et al., 2008) on murine and human adipose cells focused on the role of GC in the expression of lipin-1 gene encoding the phosphatidic acid phosphatase-1 (PAP1) enzyme, known to enhance adipocyte differentiation, triglyceride storage and insulin sensitivity. The study demonstrated the expression of lipin-1 is induced by GC and blocked by GR antagonists, indicating that GC deficiency could potentially induce insulin resistance and abnormal adipose differentiation through downregulation of the lipin-1 gene.

Protein metabolism. One study explored the impact of GC deficiency specifically on protein metabolism (Parry-Billings, et al., 1990), demonstrating decreased release of glutamine from the skeletal muscle in rats treated with GR antagonist. As glutamine is a conditionally essential amino acid with wide use in protein and lipid synthesis and a cellular source of energy, such results would indicate a negative impact of GC deficiency

on protein metabolism. Of note, there was large variability of results among control groups, limiting the reliability of the findings. Nevertheless, the dysregulation of glutamine deficiency was also demonstrated by a more recent study on zebrafish (Weger, et al., 2018). To explore the metabolic effects of GC deficiency in fdx1b-/- mutant fish, the authors combined a bioinformatics approach and biomolecular technology, including transcriptomic, metabolomic profiles and RNA-sequencing as well as enrichment analysis. The results demonstrated that systemic GC deficiency induced extensive dysregulations of multiple metabolic pathways including glutathione and purine metabolism. Moreover, blood amino acid profiles from patients with adrenal insufficiency showed similar abnormalities to the results obtained in zebrafish, supporting the relevance of the study to human pathology.

One study examined the effect of blocking the GR on protein metabolism in the skeletal muscle of a porcine model (Du, et al., 2021), the result consisting of reduced expression of protein degradation-related genes and increased protein deposition in the skeletal muscle.

1.4.1.2 Studies that explored the effect of glucocorticoid deficiency on food intake regulation

GC are known to play an important role in caloric intake and even food preference (Macfarlane, et al., 2008). However, their actions on energy balance in different pathological conditions varies depending on context (acute or chronic) and on the interactions with other hormones, such as insulin, catecholamines and leptin (Dallman, et al., 2004). Thus, it can be speculated that GC deficiency can further influence metabolic processes due to its impact on food intake and satiety response.

The effect of GC deficiency on the regulation of food intake and satiety response was explored by an *in vivo* study (Makimura, et al., 2003) on adrenalectomised mice who were found to lack the appropriate upregulation of orexigenic factors (Agouti related protein) and downregulation of anorexigenic neuropeptides (POMC) following fasting. These results support the role of GC in the hypothalamic response to fasting and are complemented by more recent studies on rats with adrenalectomy (Uchoa, et al., 2009; Uchoa, et al., 2012). Thus, it was shown that GC deficiency causes hypophagia associated with a higher activation of central neural structures involved in the satiety response including nucleus of the solitary tract and paraventricular nucleus (PVN) of the hypothalamus (Uchoa, Sabino et al., 2009). Moreover, in fasting animals, adrenalectomy associated decreased expression of orexigenic neuropeptides (Neuropeptide Y, Agouti related protein and orexin A) as well as upregulation of POMC (Uchoa, et al., 2012). Interestingly, GC deficient rats were found to have lower post-feeding leptin and insulin levels despite a more marked rise in blood glucose levels, as well as absent downregulation of the orexigenic neuropeptides after feeding. This led the authors to conclude that GC could play a role in the expression of hypothalamic neuropeptides induced by food intake. Interestingly, a study exploring the impact of corticotrophin releasing hormone (CRH) receptor deficiency in transgenic mice (Müller, et al., 2000) found that, while causing marked GC deficiency, it did not impact on body weight and quantity of food ingested. However, CRH receptor deficiency resulted in abnormal circadian feeding pattern (diurnal), which was reversed following GC treatment.

1.4.1.3 Studies that explored specifically cardio-vascular effects of glucocorticoid deficiency

One publication focused on the cardio-vascular effects of GC deficiency, reporting an *in vivo* study that explored the development of atherosclerosis in LDL-receptor deficient mice following adrenalectomy (van der Sluis, et al., 2012). GC-deficient mice developed a higher degree of aortic root atherosclerosis compared to controls, an effect that was reversed following treatment with GC. However, the mechanism through which GC deficiency increased the risk of atherosclerosis appeared to relate to its pro-inflammatory consequences rather than the metabolic ones, based on the high macrophage content of the plaques and the low plasma lipid levels found after adrenalectomy.

1.4.2 Clinical studies

A small number of studies have addressed the topic of metabolic effects of GC deficiency in human subjects. In a group of six patients with adrenal insufficiency (AI), it was shown that 72 hours without medication caused a decrease in insulin levels, with no impact on blood glucose or hepatic glucose production, supporting the theory that GC deficiency increases insulin sensitivity (Malerbi, et al., 1988). This concept was further developed by a more recent crossover study (Christiansen, et al. 2007) involving seven females with adrenal insufficiency (AI) that showed that 24 hours of GC withdrawal caused increased insulin sensitivity, reduced protein loss in the skeletal muscle and decreased energy expenditure. A third study (Poyrazoglu, et al., 2003) explored the impact of GC deficiency on leptin levels in relation to BMI and androgen status in 11 children with 21OHD before and after commencing replacement therapy. Secreted by white adipose tissue, leptin is known to have important roles in human metabolism, by regulating food intake and energy expenditure. Although the study found significantly lower leptin levels in GC deficient

patients compared to both medicated patients and healthy controls, the three groups were comparable with regards to BMI's. Interestingly, leptin correlated positively with the pre-dose cortisol concentration and negatively with testosterone, the authors suggesting its potential use as a marker for monitoring treatment.

Two studies assessed the effects of low cortisol concentrations on metabolism in individuals with hypocortisolism, who did however not have a clinical diagnosis of GC deficiency. A cross-sectional study involving a large cohort (2915 men and 1041 women) identified a non-linear association between obesity and salivary cortisol (Kumari, et al., 2010). Interestingly, obese individuals presented flatter slopes of cortisol secretion caused by lower waking cortisol and higher evening cortisol. Thus, the circadian pattern of the GC synthesis also appears to have an impact on the way in which fat metabolism is affected, adding further challenges to replacement strategies. The second study (Maripuu, et al., 2016) explored the relationship between hypocortisolism and metabolic cardio-vascular risk factors by analysing the metabolic profiles of 245 patients with affective disorders and 258 healthy controls. The authors reported an increased rate of obesity, dyslipidaemia and metabolic syndrome associated with hypocortisolism in both mental health patients and controls.

A recent review of the published cases of human GC resistance syndrome (Vitellius and Lombes, 2020), identified that 18 out of 33 patients were obese. The authors commented that the mechanisms for this associated morbidity are not completely understood, raising the possibility of depressed GR signalling by cortisol, but also taking into consideration the effect the transactivation of the mineralocorticoid receptor (MR) by cortisol and its potential effects on adipocyte differentiation and function. Previous evidence from *in vitro*

research on human adipocytes (Lee and Fried, 2014) would support the former hypothesis, having shown that GR knockdown caused reduced differentiation of preadipocytes, as well as reduced expression of leptin and adiponectin. These effects were not found in the MR knockdown cells (Lee and Fried, 2014).

1.4.3 Interpretation and summary of published evidence

Taken together, the findings resulted from animal studies confirm a number of aspects that are in keeping with what is generally known about the GC actions, namely that the GC deficiency leads to decreased liver glycogen storage and blood glucose levels, in association with increased insulin sensitivity (Kawai, 1977; Blair, et al., 1995; Bishayi and Ghosh, 2003) (**Figure 1.4**). It appears that the effect on insulin sensitivity relates mainly to decreased hepatic glucose production rather than an increased uptake of glucose in the adipose tissue or skeletal muscle (Kawai, 1977), suggesting a potential need for GC to enhance insulin actions in these structures. This point is further supported by results describing the effect on insulin secretion and on insulin signalling pathways of protein and lipid metabolism in the heart (Wu, et al., 2004). The impact of GC deficiency on adipose differentiation through more recently identified pathways, such as the lipin-1 gene (Zhang, et al., 2008), as well as the dysregulations of glutamine and glutathione metabolism (Parry-Billings, et al., 1990; Weger, et al., 2018) are significant and warrant further research in relation to the wide potential implications in human pathology. Interestingly, a cardio-vascular risk of developing atherosclerosis was also identified by one of the studies (van der Sluis, et al., 2012), though not apparently related to dyslipidaemia but rather as a consequence of the pro-inflammatory effects of GC deficiency; this aspect highlights the intricate mechanisms of pathology related to this group of steroids, specifically, the interplay between their metabolic and immune regulation roles. Finally, the adjacent effect

of blunting the central responses to fasting and satiety (Müller, et al., 2000; Uchoa, et al., 2009; Uchoa, et al., 2012) has important implications for metabolic processes related to GC deficiency. However, certain limitations of the research studies discussed need to be considered, including the diverse palette of techniques and the large variation in the duration of the induced GC deficiency, usually consisting of maximum 30 days, which limits the relevance of the findings in relation to the impact of prolonged GC deficiency in the metabolism of chronic patients.

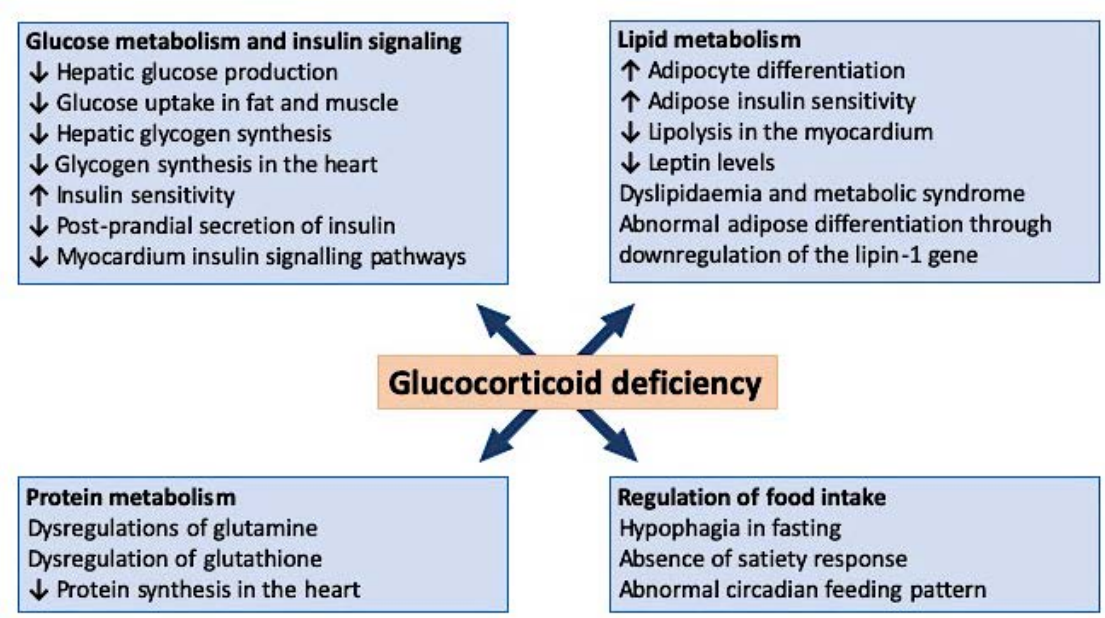


Figure 1.4. Summary of the effects of GC deficiency on metabolic pathways
Downward arrows indicate metabolic processes that are inhibited by GC deficiency; upward arrows indicate processes that are stimulated exacerbated in GC deficiency.

Importantly, different animal models of GC deficiency were used with their respective merits and limitations. Thus, the bilateral adrenalectomy used in the majority of the publications (Kawai, 1977; Blair, et al., 1995; Bishayi and Ghosh, 2003; Makimura, et al., 2003; Wu, et al., 2004; Uchoa, et al., 2009; Uchoa, et al., 2012; van der Sluis, et al., 2012), while effectively inducing severe GC deficiency, also deprives the research organism of other factors and hormones that have metabolic roles including MC and

catecholamines, thus producing confounding effects. This impediment is avoided in some of the studies that used GR antagonists (Parry-Billings, et al., 1990; Priyadarshini and Anuradha, 2017) which, however, limit the effects of GC deficiency to those that are GR-mediated. In that respect, a potential alternative is presented by the study involving GC deficient zebrafish mutants (Weger, et al., 2018), in addition to the novel available techniques of molecular biology and bioinformatics that yielded very extensive results suggesting the wide metabolic implications of congenital GC deficiency. Of course, this study provided a very general perspective on the pathology of GC deficiency, incurring the need for more in-depth exploration of its implications.

The limited number of clinical studies related to the topic is not surprising given the ethical considerations entailed by the potential withholding of replacement treatment in GC deficiency. Consequently, the three studies discussed had the disadvantage of using small numbers of patients and only explored the metabolic impact of short periods of GC deficiency. This can probably account for the limited results yielded, which mainly related to the reduction in insulin resistance associated with GC deficiency and confirmed the conclusions of animal studies. Probably the most notable finding is the low concentrations of serum leptin in untreated patients in the absence of increased body mass index (BMI) (Poyrazoglu, et al., 2003). Thus, this study would support the role of cortisol in regulating leptin secretion directly. However, given the pleiotropic effect of both hormones on multiple metabolic pathways, the underlying mechanisms for their interrelation remain unclear. A clear disadvantage of the research described in these publications is the short duration of the GC deficiency, chosen to prioritise patient safety; this limits the relevance of the findings with relation to the topic of this review, namely the effect of chronic GC deficiency on metabolism. By contrast, the two large cohort studies (Kumari, et al., 2010;

Maripuu, et al., 2016) provided the benefit of identifying metabolic consequences of long-standing hypocortisolism, albeit non-clinical and undiagnosed, suggesting that abnormal patterns of cortisol secretion led to obesity and metabolic syndrome. While these publications did not include participants suffering from clinically significant GC deficiency, they obtained clear correlations between hypocortisolism and metabolic abnormalities, in particular related to fat metabolism and obesity, which were made robust and reliable due to the large study samples that enhanced the power of the statistical analysis. Such conclusions can be translated to clinical GC deficiency, with the expectation that metabolic consequences may be even more pronounced in related pathology.

1.5 Zebrafish as a research model of steroid hormone deficiency

(Modified from published review (Bacila, et al. 2021))

Over the last two decades zebrafish have become an increasingly popular animal model for translational research. Their small size, easy bread and maintenance, as well as their high fecundity and short generation time (3 months), are among the main advantages that zebrafish has over other animal models. Females can lay as many as 100-200 eggs in a clutch, all of which are fertilised externally, and embryos progress fast to larvae which are free feeding at 5 days post fertilisation (dpf). Another important quality is that both zebrafish embryos and larvae are transparent, making them a convenient model for live imaging as a means to explore the development and physiological functioning of internal tissues and organs (McGonnell and Fowkes, 2006).

The organisation and function of the stress-regulating system is conserved to a high degree between zebrafish and mammals, including humans (Steenbergen, et al., 2011).

In the same way as in humans, zebrafish are diurnal organisms and cortisol is the principal active GC (Alsop and Vijayan, 2009). The zebrafish and human genomes also have surprising similarities, including the genetic regulation of steroid hormone synthesis and signalling (Tokarz, et al., 2013). However, following a genome duplication believed to have taken place around 450 million years back in the evolution of teleost fish (the infraclass to which zebrafish belongs), in zebrafish some steroidogenic genes are found as pairs of paralogues (Glasauer and Neuhauss, 2014). Many aspects related to the differences in expression and function of several such paralogues remain to be explored. The interrenal gland represents the zebrafish equivalent of the adrenals, being located in the head kidney. Alongside the gonads and the brain (Weger, et al., 2018), the interrenal gland is an important site of steroid synthesis (Tokarz, et al., 2013). Over the last 20 years, there has been an increasing research focus on the zebrafish interrenal gland, providing useful evidence on their development and function, as well as the roles of steroidogenic enzymes.

1.5.1 Specification and morphogenesis of the zebrafish interrenal gland

The interrenal gland develops alongside the renal primordia, originating from the intermediate mesoderm (Hsu, et al., 2004). The initial stage in its organogenesis is the differentiation of cell clusters in the pronephric primordium, which can be detected through the expression of *ff1b* (*ftz-fl* gene, zebrafish orthologue of mammalian steroidogenic factor 1) from 22 – 27 hours post fertilisation (hpf) onwards, ventral to the third somite (**Figure 1.5**) (Liu, et al., 2003; To, et al., 2007). The two clusters of primordial interrenal cells proceed to separate from the renal primordium and meet left to the midline where they

form a single discreet organ between 24 - 30 hpf (Chai, et al., 2003; To, et al., 2007). After centralisation, the interrenal tissue starts expressing the steroidogenic acute regulatory protein, cytochrome P450 side-chain cleavage enzyme *(cyp11a2)* and 3ß-hydroxysteroid dehydrogenase *(hsd3b)*, which marks its capacity to synthesise steroid hormones (Chai, et al., 2003; To, et al., 2007). Zebrafish also have the equivalent of the adrenal medulla, the chromaffin primordia, which are cells that express dopamine-ß-hydroxylase (*dβh)* within the bilateral primordial interrenal structures. They also converge centrally at 2 dph and fuse into one domain at 3 dpf (To, et al., 2007)

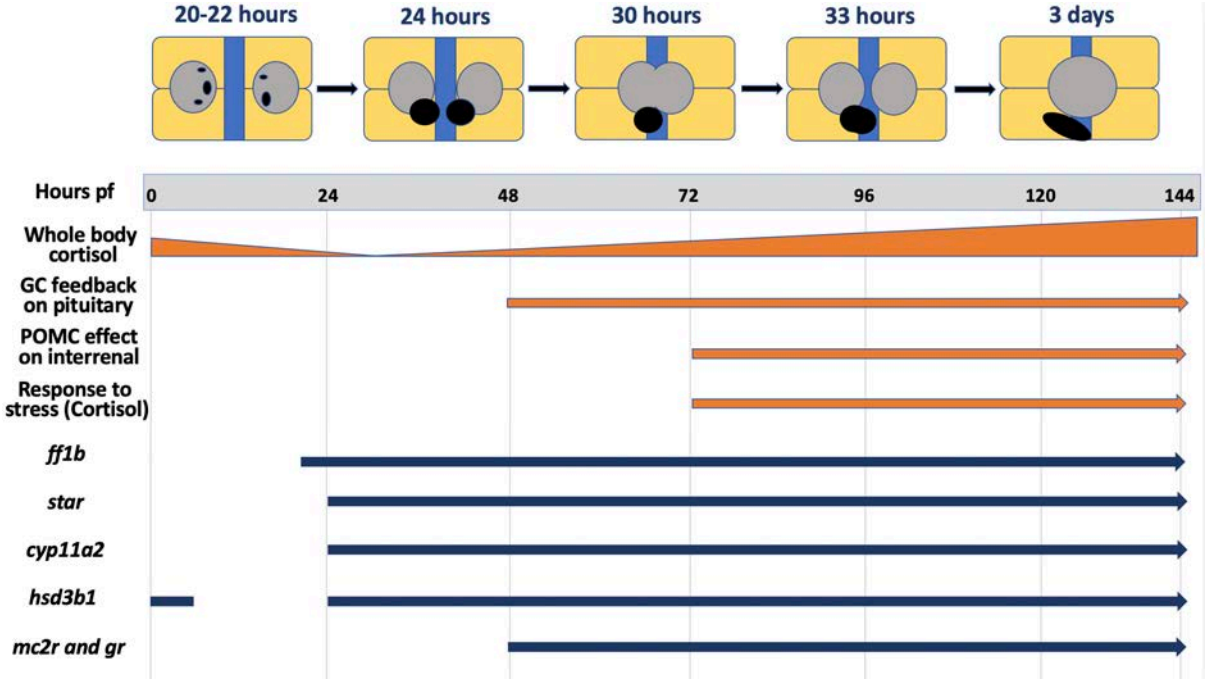


Figure 1.5. Organogenesis and functional development of the interrenal gland
A. Specification, central migration and fusion of the kidney (Greytak, Champlin et al.) and interrenal primordium. **B.** Development of the stress response (blue) and expression relevant genes (orange) over the first 6 dpf. (Modified from Bacila, et al. 2021)

The development of the interrenal gland is regulated by several factors. *ff1b* represents the zebrafish orthologue of the gene encoding SF1/Ad4BP/NR5A1 (steroidogenic factor 1) in mammals, a nuclear receptor essential for the interrenal specification and steroid

synthesis (Hsu, et al., 2004). The interrenal, gonads and hypothalamus express *ff1b* (Hsu, et al., 2004) which binds to the zebrafish equivalent of the CYP11A1 promoter (Quek and Chan, 2009). Another crucial element for the interrenal gland development is the expression of Wilms tumour 1 (*wt1*) in the primordial kidneys, *wt1*-deficient mutants having been shown to have reduced interrenal size and *ff1b* down-regulation (Hsu, et al., 2004). The prospero-related homeodomain protein (PROX1) also has a zebrafish orthologue in the interrenal gland, which is expressed until 3 dpf, modulating the *ff1b* expression. *ff1b* is also regulated by the *dax1*, ortholog of the mammalian dosage-sensitive sex reversal, adrenal hypoplasia critical region, on chromosome X, gene 1 (*DAX1*), expressed in the interrenal tissue at 31 hpf. *Dax1* is needed to initiate steroid synthesis as it controls the expression of *cyp11a* and *star* (Zhao, et al., 2006).

In the same way in which the stress axis functions in mammals, in zebrafish the interrenal steroidogenesis is regulated by the hypothalamic-pituitary-interrenal (HPI) axis, which develops in the early larval stages and becomes functional after 48 hpf. This is demonstrated by an initial reduction in whole body cortisol (likely of maternal origin) between 8 and 36 hpf, followed by a rise between 48 and 120 hpf (Alsop and Vijayan, 2008; Wilson, et al., 2013). The pituitary gland starts to control the function of the interrenal, by regulating the expression of steroidogenic genes (*cyp11a*, *mc2r, star*), from 2 dpf onwards (To, et al., 2007). Thus, research work using a selection of mutants lacking different types of pituitary cells showed that at 5 dpf, corticotrophs able to express pro-opiomelanocortin (*pomc*) are crucial for the normal development of the interrenal gland. Their product, the adrenocorticotropic hormone (ACTH), transactivates the melanocortin

receptor 2 (Mc2r). The negative feed-back of exogenous GC on the pituitary gland is manifested by 2 dpf and the suppression of *pomc* expression and inhibition of interrenal proliferation by exposure to dexamethasone occurs at 3 dpf (To, et al., 2007).

1.5.2 Genes involved in zebrafish steroidogenesis

Like in mammals, the presence of the steroidogenic acute regulatory protein (StAR) is required in zebrafish to regulate the transport of cholesterol from the outer to the inner mitochondrial membrane, allowing the initiation of interrenal steroidogenesis (**Figure 1.6**). The zebrafish steroidogenic structures (gonads and interrenal) express the *star* gene which is similar to the mammalian homologue (Bauer, et al., 2000). Cytochrome P450 side-chain cleavage enzyme (Cyp11a) is another essential enzyme, as it catalyses the first step of steroid synthesis by converting cholesterol to pregnenolone. Zebrafish have two *cyp11a* genes (*cyp11a1* and *cyp11a2)*, which are situated on chromosome 25, the sequence of the amino acids being 80% identical in the corresponding encoded proteins (Goldstone, et al., 2010; Parajes, et al., 2013). *Cyp11a1* contains a promoter that binds to *ff1b*, in the same way in which in mammals the CYP11A1 promoter is activated by SF1 (Quek and Chan, 2009). There are two phases in the zebrafish development during which *cyp11a1* is expressed: early embryogenesis (from zygote stage to 24 hpf), and after 14 dpf, corresponding to the timing of sexual differentiation (Hsu, et al., 2002; Hu, et al., 2004; Goldstone, et al., 2010; Parajes, et al., 2013; Weger, et al., 2018). *Cyp11a2* was found to be expressed at 4dpf (dome stage), however, the expression is low for the first 24 hpf, to increase after 30 hpf, peaking at 120 hpf (Weger, et al., 2018). Thus, it was hypothesised that there is a transition from *cyp11a1* to *cyp11a2* corresponding to the initiation of cortisol synthesis. This would assign Cyp11a2 as the functional equivalent of

the mammalian cytochrome P450 side-chain cleavage enzyme (Parajes, et al., 2013). In the adult fish, the expression of *cyp11a2* was demonstrated in the interrenal, gonads and brain (Parajes, et al., 2013; Weger, et al., 2018).

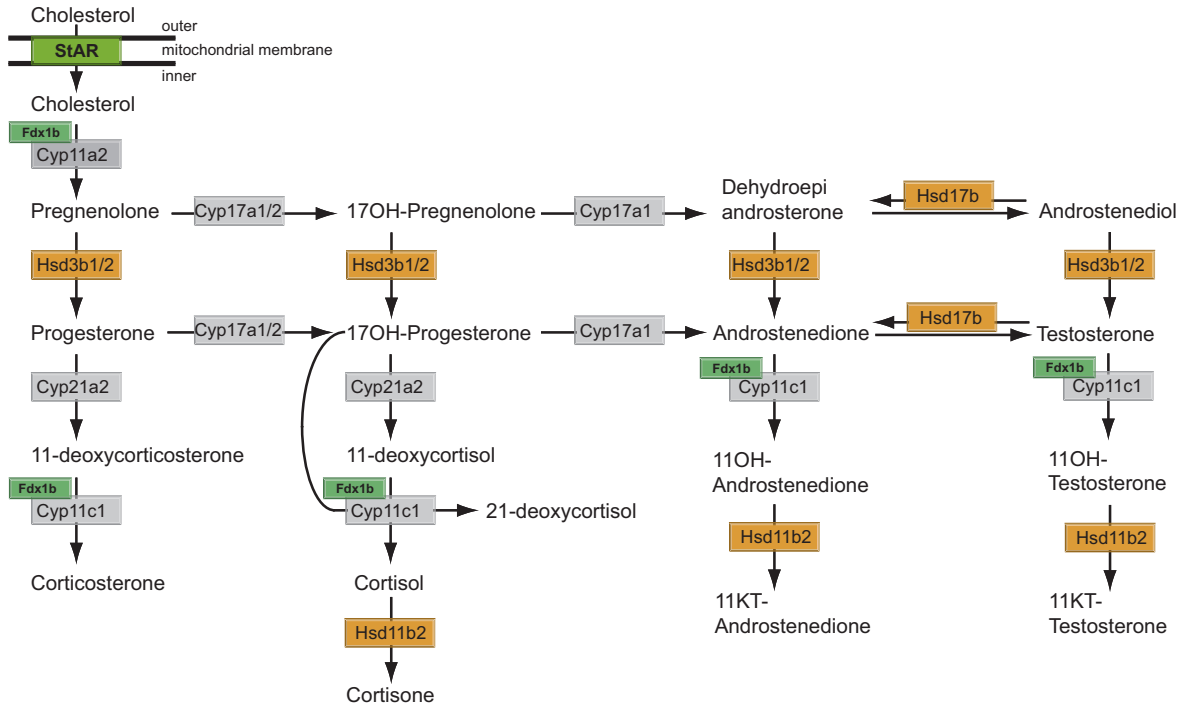


Figure 1.6. Schematic diagram of interrenal steroid biosynthesis
The first and second vertical pathways correspond to mineralocorticoid and glucocorticoid synthesis, respectively. The next two columns to the right represent androgen synthesis. The arrows indicate chemical reactions, and the adjacent rectangles specify the enzymes and co-factors involved in catalysing them. Grey boxes indicate cytochrome P450 enzymes, orange boxes show hydroxysteroid dehydrogenase and green boxes correspond to the co-factor ferredoxin. (Modified from Bacila, et al. 2021)

The zebrafish equivalent of the two human genes (*HSD3B1* and *HSD3B2*) is *hsd3b1*, situated on chromosome 9. There are two peaks in the *hsd3b1* expression: at zygote stage and between 30 to 120 hpf (Weger, et al., 2018). Similar to the human *HSD3B2*, in the adult fish, *hsd3b1* is expressed in the interrenal gland and gonads (Lin, et al., 2015; Weger, et al., 2018).

In zebrafish, 21-hydroxylase is encoded by the *cyp21a2* gene (Weger, et al., 2018), found on chromosome 16, its protein product sharing 40.71% homology with its human CYP21A2 orthologue (Eachus, et al., 2017). During larval development, the interrenal tissue expresses *cyp21a2* from 28 hpf, the transcripts levels increasing afterwards, with a peak at 120 hpf (Eachus, et al., 2017; Weger, et al., 2018). Cyp11c1 is the zebrafish equivalent to human 11ß-hydroxylase (CYP11B), involved in the synthesis of corticosterone and cortisol. The *cyp11c1* gene is expressed at 8 hpf before hatch, however, transcript levels rise again after 24 hpf peaking between 72-120 hpf (Alsop and Vijayan, 2008; Wilson, et al., 2013; Weger, et al., 2018). Its involvement in steroid synthesis was demonstrated by *in situ* hybridisation showing it to be expressed only in the interrenal gland at 2 dpf (Zhang, et al., 2020).

The gene coding 17a-hydroxylase (CYP17A1) *has* two isomers in zebrafish, *cyp17a1* and *cyp17a2* (Goldstone, et al., 2010; Pallan, et al., 2015; Gonzalez, et al., 2018; Weger, et al., 2018). Gene phylogeny analysis showed that *cyp17a2* was the result of a duplication of *cyp17a1* (Zhou, et al., 2007). However, unlike Cyp17a1 which has both 17α-hydroxylation and 17α,20-lyase activity, Cyp17a2 lacks the 17α,20-lyase capacity, required for the synthesis of sex hormones (Zhou, et al., 2007). The zebrafish interrenal gland was found to only express *cyp17a2 (*at 120 hpf) and not *cyp17a1* (Weger, et al., 2018), indicating that the interrenal is not involved in androgen synthesis.

The equivalent of the co-factor ferredoxin (FDX1) in zebrafish is also coded by two genes *fdx1* and *fdx1b*. However, Fdx1b was the only one shown to play a part in steroid

synthesis as co-factor Cyp11a1, Cyp11a2and Cyp11c1. It is believed that Fdx1 is likely to be involved in the early development of zebrafish (Griffin, et al., 2016). The expression of *fdx1b* appears to peak at two stages in the embryogenesis: at 5-8 hpf and after 30 hpf (Griffin, et al., 2016; Weger, et al., 2018). Transcripts of *fdx1b* were found only in the interrenal at 120 hpf, and limited to steroidogenic tissues in adults (brain, gonads, interrenal) (Griffin, et al., 2016; Weger, et al., 2018), further supporting its involvement in steroid synthesis. By contrast, *fdx1* was expressed in multiple organs (gonads, kidney, interrenal, intestines, heart, brain) (Griffin, et al., 2016).

1.5.3 The study of steroidogenic enzyme deficiencies in the zebrafish model

Several steroidogenic enzyme deficiencies have been explored in zebrafish, leading to valuable animal models of altered steroid synthesis. Morpholino knockdown studies confirmed the role of *cyp11a1* in the early development and that of *cyp11a2* in the initiation and maintenance of steroidogenesis, with morphants being profoundly cortisol deficient (Parajes, et al., 2013). In a more recent study, the clustered regularly interspaced short palindromic repeats/ Cas9 nuclease (CRISPR/Cas9) approach was used to produce a zebrafish lacking *cyp11a2.* Mutant adults were exclusively males with a female external phenotype and disorganised testicular histology, who were infertile and lacked breeding behaviours (Li, et al., 2020). Steroid measurements confirmed severe GC and androgen deficiency in the *cyp11a2-/-* fish. Gene expression analysis showed upregulation of the HPI axis, a feature consistent with adrenal insufficiency in humans (Charmandari, et al., 2014). Of note, neither of the two studies (Parajes, et al., 2013; Li, et al., 2020) explored the metabolic phenotype of the cyp11a2-/- fish. A zebrafish model

of 21OHD (*cyp21a2-/-*) has also been produced, using the TALEN mutagenesis (Eachus, et al., 2017). The mutant larvae had reduced cortisol, impaired visual background adaptation and downregulation of GC responsive genes. Moreover, 21-deoxycortisol, a disease specific marker in 21OHD, was raised compared to wild type larvae. Mutant larvae also presented interrenal hyperplasia and upregulation of the *pomca* gene (Eachus, et al., 2017). The authors reported the downregulation of phosphoenolpyruvate carboxykinase 1 (*pck1)*, an enzyme involved in gluconeogenesis, no other aspects were explored in relation to the metabolic profile of *cyp21a2* mutants. Similar findings were reported in a morpholino-induced mutant with *hsd3b1* knockdown (Lin, et al., 2015). The mutant fish had interrenal hyperplasia, upregulation of *pomca*, low cortisol and increased pigmentation, findings which are consistent with human HSD3B2 deficiency (Lin, et al., 2015).

While in humans 11β-hydroxylase deficiency presents with hypocortisolism, increased MC activity and hyperandrogenism, the phenotype of *cyp11c1*-deficient zebrafish mutants is profoundly affected by the reduced 11-ketotestosterone concentrations causing gonadal impairment (Oakes, et al., 2020; Zhang, et al., 2020). All adult fish had external female sexual characteristics; however, they developed either testes or ovaries (Zhang et al., 2020; Oakes et al., 2020). Fertility, the structure of the gonads and breeding behaviours were also impaired, however, they could be partially saved by administration of cortisol in females and 11-ketoandrostenedione in males (Zhang, et al., 2020). There was also interrenal hyperplasia, correctable with cortisol. A *fdx1b*-/- zebrafish presented a similar phenotype, strengthening the concept that testicular development is

independent of androgens, but requires 11-ketotestosterone for normal structure and function (Oakes, et al., 2019). The *fdx1b*-/- fish also had reduced cortisol, upregulation of the *star, cyp11a2* and *cyp11c1* genes, as well as increased body weight and body length (Oakes, et al., 2019). While in human pathology no cases of ferredoxin reductase deficiency have so far been reported, the *fdx1b* deficient zebrafish proved to be a very useful model to study the impact of severe GC on different body system and functions, including metabolism (Weger, et al., 2018).

Finally, a zebrafish homeodomain-containing transcription factor (*rx3)* mutant, lacking pituitary corticotropes, was used as a model of secondary adrenal insufficiency, exploring related metabolomic and transcriptomic changes (Weger, et al., 2016). There were important dysregulations of the genes that have diurnal expression, of which many were involved in metabolic pathways, in particular glutamine and glyoxylate metabolism, as well as the citrate and urea cycles.

Thus, zebrafish equivalents to humans were established for several steroidogenic genes and their respective enzymes and co-factors. Their involvement in the early stages of zebrafish development, combined with the multiple practical advantages of using zebrafish as a research model, provide ample translational applications in the study of adrenal disease. Subsequently, several models of specific enzyme deficiencies have been established, many of which present striking pathognomonic similarities to human conditions associated with GC insufficiency. A very important advantage is that many of these enzyme deficiencies appear to be compatible with life in zebrafish in the absence

of GC replacement, which allows the opportunity to explore in a dynamic manner the effects of GC deficiency on the genotype and phenotype throughout life and even permits the study of transgenerational emergence of adaptive mechanisms (Vera-Chang MN et al., 2018; Vera-Chang MN et al., 2019).

1.5.4 Glucose and fat metabolism in zebrafish

Zebrafish possess the same main organs and tissues that are involved in fat and glucose metabolism and energy regulation in mammals, namely, gastro-intestinal tract, liver, pancreas, adipose tissue and skeletal muscles (Zang, et al., 2018). The regulatory mechanisms governing functions such as appetite and satiety control, lipid differentiation, lipid storage and insulin signalling are well conserved (Elo, et al., 2007; Nishio, et al., 2008; Flynn, et al., 2009). Moreover, there is also resemblance to mammals in relation to the maintenance of energy homeostasis, including the melanocortin system, the effects of the Agouti-related protein and adiponectin.

Glucose homeostasis in zebrafish. The morphology and function of the zebrafish pancreas, with both endocrine and exocrine components, is similar to that of mammals (Tehrani and Lin, 2011). The endocrine pancreas includes α-cells secreting glucagon and **β**-cells secreting insulin (Argenton, et al., 1999). While formation of the islets of Langerhans begins at 24 hpf, the expression of insulin starts earlier in the embryogenesis, as 12-somite stage (Biemar, et al., 2001). Insulin was shown to regulate the uptake of glucose in the skeletal muscle in zebrafish, indicating the skeletal muscle as a site of glycolysis and providing the opportunity to develop a model of insulin resistance

(Maddison, et al., 2015). The liver was also shown to play a major role in zebrafish glucose metabolism, being an important site of gluconeogenesis (Jurczyk, et al., 2011).

Phosphoenolpyruvate carboxykinase (PEPCK), an enzyme which catalyses an early rate-limiting step of gluconeogenesis, was shown to be regulated by insulin and glucagon in zebrafish larvae, similarly to mammals (Elo, et al., 2007). It is expressed in several structures in adult zebrafish, including the digestive system. Other genes involved in gluconeogenesis (glucose 6 phosphatase, fructose 1,6 biphosphatase) and glycolysis (glucokinase, pyruvate kinase) have also been identified in zebrafish, although there is limited evidence available on their expression and roles.

Interestingly, in zebrafish, leptin appears to have a role in glucose homeostasis and regulation of **β**-cells development and function, but it does not act as an adipostatic factor as it does in mammals (Michel, et al., 2016). Thus, the main features identified in zebrafish mutants lacking the leptin receptor consisted of increased blood glucose levels, **β**-cell hyperplasia and upregulation of the insulin and glucagon gene. Hepatic glucose metabolism was also affected with upregulation of phosphoenolpyruvate carboxykinase, glucose-6-phosphatase and pyruvate kinase. (Michel, et al., 2016)

Fat metabolism in zebrafish. In zebrafish the adipose tissue develops from pluripotent cells from the mesenchyme, which differentiate into preadipocytes and then into adipocytes (Den Broeder, et al., 2015). The nuclear peroxisome-proliferator activated receptor gamma (Pparg) and CCAAT/enhancer-binding protein (Cebp*α*) are essential regulators of the initial steps of adipocyte differentiation in zebrafish, their genes being expressed in the fat tissue and liver (Imrie and Sadler, 2010). The fatty acid binding

protein 11a (Fabp11a) is involved in the final steps of the process. Exogenous nutrition was shown to be required for adipocyte development, as visceral adipocytes are formed through the accumulation of neutral lipid droplets (Flynn, et al., 2009). During the first 5 days of development, nutrition is provided by the yolk sac, containing fat-soluble vitamins, cholesterol (Anderson, et al., 2011) and triacylglycerol, and from 5-6 dpf zebrafish larvae become self-feeding. Adipocytes become visible in the visceral cavity from 8-12 dpf, their amount correlating with the size rather than the age of the larvae (Flynn, et al., 2009; Imrie and Sadler, 2010). Subcutaneous and cranial fat develop at 20 dpf and 22 dpf, respectively in a size-dependent manner (Imrie and Sadler, 2010).

Before being absorbed in the intestine, lipids are processed by digestive enzymes and bile. As in humans, bile promotes the formation of micelles, being produced by hepatocytes and released in the intestinal lumen following food consumption. Pancreatic lipases break down triacylglycerol and phospholipids to free fatty acids, mono- and di-acylglycerols, which are absorbed by enterocytes. Thus, following food consumption, lipid drops accumulate in the enterocytes. From here, fats are either burned in the mitochondria or peroxisomes through oxidative processes or packed into chylomicrons that are then released into the lymph or blood. The lipoproteins contained in the zebrafish chylomicrons are closely similar to those of mammals and it is believed that many of the mechanisms involved in lipoprotein production are conserved. (Anderson, et al., 2011)

Because they feed on a high-cholesterol diet, zebrafish are hyperlipidaemic and hypercholesterolaemic in comparison to mammals. They also have a high percentage of unsaturated fatty acids and HDL is the highest fraction of circulating cholesterol. In females, the lipoprotein level is also influenced by the concentration of circulating

vitellogenin, a VLDL synthesised by the liver under oestrogen stimulation. (Hölttä-Vuori, et al., 2010)

Two main regulators of cholesterol metabolism are the sterol-regulatory-element binding protein (SREBP) and liver X receptor (LXR) systems, acting to increase and decrease, respectively, cellular cholesterol. The regulation of lipid synthesis through the SREBP pathway is conserved in zebrafish. It is also the pathway that is activated causing lipogenesis in alcohol-induced fatty liver disease in humans (Hölttä-Vuori, et al., 2010). Moreover, orthologues were found in zebrafish for several genes involved in cholesterol synthesis (*hmgcr*, coding the equivalent of β-Hydroxy β-methylglutaryl-CoA that catalyses a rate-limiting step in the synthesis of cholesterol) and fatty acids oxidation (*hadha,* coding hydroxyacyl-CoA dehydrogenase), related to defects of lipid metabolism in mammals.

In humans, adiponectin was shown to enhance beta-oxidation of fatty acids, decrease plasma triglycerides and increase insulin sensitivity, having important roles in body-weight regulation and energy homeostasis. Two adiponectin genes were identified in zebrafish, one of which was only expressed in the kidney, while the other was found in several organs including gut, liver and adipose tissue. Adiponectin was downregulated in fasted fish suggesting that the pathway is regulated by food deprivation (Nishio, et al., 2008), which would suggest similar roles to those observed in humans.

In conclusion, as well as presenting important logistic advantages as a research animal model in Molecular Biology, zebrafish resembles key aspects of human development, regulation, and pathophysiology of both the stress axis and metabolic processes. This makes it a convenient model for exploring the effects of GC deficiency on metabolism.

However, the potential caveats related to its use in this thesis will also have to be taken into consideration, including the potential differences in the expression and action of steroidogenic gene paralogues, the metabolic implications of being an ectotherm organism and the potential distinctions in the roles of adipokines compared to human pathology.

1.6. Aim and Objectives

This project is based on the hypothesis that GC deficiency contributes to the dysregulation of glucose and fat metabolism associated with metabolic morbidity in CAH, rather than it being exclusively caused due to over-exposure to synthetic corticosteroids. Thus, the aim is to improve the understanding of the molecular mechanisms involved in the pathology of metabolic disease in 21OHD and clarify their interrelation with GC deficiency. The overarching goal in conducting this work was to provide information that will help develop novel biomarkers for developing metabolic disease in CAH and help improve therapy monitoring in clinical practice. The project combined basic science with clinical research, and it was intended that the information derived from the transcriptomic and metabolomic analysis of a *cyp21a2*-deficient zebrafish model would help comprehend the molecular processes that lead to metabolic complications found in patients with CAH. The objectives set in planning this project were:

1. To define alterations of metabolism in relation to hormonal status in patients with CAH by:

1a. Defining the health status of children and young people with 21OHD and in particular, the onset of metabolic disease during childhood

1b. Analysing the Cytokine/Adipokine profile of patients with 21OHD in relation to the hormonal profile and clinical outcomes

1c. Identifying patterns in blood metabolomes for patients with 21OHD in relation to the hormonal profile and clinical outcomes

2. To explore the metabolic profile of *cyp21a2-/-* zebrafish by analysing their phenotype and transcriptome:

2a. Phenotypic characteristics including weight, length, blood glucose levels, macroscopic fat distribution, liver size and histology

2b. The expression of genes involved in glucose and fat metabolism through qPCR

2c. Transcriptomic analysis focused on metabolic processes and functions, exploring the differential gene expression between *cyp21a2*-/- and wild type controls in larvae and adult livers

3. Compare the clinical and zebrafish data to identify evolutionarily conserved and distinct aspects of the molecular mechanisms underlying the pathophysiology of CAH

A national cohort of 107 children and young persons with 21OHD had already been established, through a multicentre study (CAH-UK, Clinical Trials Registration Number: SCH/15/088), providing a resource for clinical data and biomaterial available for analysis. A first objective was to identify health problems that are linked to metabolic pathology, including overweight and obesity, insulin resistance as well as metabolic syndrome. Following that, a more in-depth analysis of the metabolic status of patients with CAH included metabolomic and cytokine/adipokine analysis. Since alterations of GC and androgen metabolism have been associated with adverse metabolic constellations in

common obesity and polycystic ovarian syndrome, measurement of steroid hormone status in our patient cohort will serve as the basis for further analysis of metabolic alterations in CAH. The project sought to establish the interrelation of dysregulated glucose metabolism with atypically altered adipokine synthesis and metabolism in CAH patients and its interdependency with hormonal status and hormone replacement.

By exploring the differential gene expression between *cyp21a2*-/- fish and WT siblings, objectives 2b and 2c are focused on identifying the molecular mechanisms involved in fat and glucose metabolism that are disrupted as a consequence of GC deficiency. Our laboratory team has already established a zebrafish model of 21OHD that was shown to possess the phenotypical characteristics of human CAH (Eachus, et al., 2017). While proven to be severely cortisol-deficient, the *cyp21a2*-/- fish live to adulthood and can reproduce, thus representing a very convenient model for studying the impact of cortisol deficiency, without having to consider the overlapping effects of synthetic steroids, as is the case in most patients with CAH. A previous study conducted using the *cyp21a2*-/- fish model, has shown that mutant larvae had upregulation of the HPI axis and interrenal hyperplasia (Eachus, et al., 2017). Steroid measurement also showed that *cyp21a2*-/- larvae were cortisol-deficient (also demonstrated by the downregulation of glucocorticoid-responsive genes *pck1* and *fkbp5*), having increased 17-hydroxyprogesterone and 21-deoxycortisol; however, androgen precursors were not detected (Eachus, et al., 2017). This provides the advantage of allowing us to focus only on GC deficiency, without the interference of hyperandrogenism. The project hypothesis was that the impact of GC deficiency on metabolic processes are likely to change throughout different stages of development. Thus, transcriptomic analysis was conducted in larvae and adult fish in order to study the dynamics of these dysregulations. The liver was chosen for the

phenotypical and genotypical analysis in adult fish due to its important role in fat and glucose metabolism, with several pathways that were shown to be similar between zebrafish and humans (Michel, et al., 2016). To provide more insights into the effects of cortisol deficiency in zebrafish, the *cyp21a2*-/- fish were compared with another cortisol-deficient zebrafish line established in our laboratory, namely *cyp11a2*-/- (side cleavage enzyme deficiency). Mutants from this line were shown to be profoundly glucocorticoid and androgen deficient, with adult fish only developing as infertile males with female secondary sex characteristics (Li, et al., 2020).

Overall, this project intended to provide significant new insights into the pathophysiology of CAH and GC deficiency with significant potential to form the basis of optimising patient care.

2. Methodology

2.1 Clinical research: The metabolic profile of patients with CAH in relation to the hormonal profile

(*Modified from paper accepted for publication (Bacila, Lawrence et al., 2022))*

A cross-sectional cohort study was conducted, including 14 tertiary UK centres (**Table 2.1**). A total of 107 patients with 21OHD and 83 age-, sex- matched controls were recruited (Clinical Trials Registration Number: SCH/15/088).

Table 2.1. List of collaborating clinical centres

1	Sheffield Children's Hospital and NIHR Clinical Research Facility, Sheffield Teaching Hospital, Sheffield
2	Birmingham Children's Hospital, Birmingham
3	Royal Hospital for Sick Children, Yorkhill, Glasgow
4	Great Ormond Street Hospital, London
5	Addenbrooke's Clinical Research Centre, Cambridge
6	Bristol Royal Hospital for Children, Bristol
7	Royal Manchester Children's Hospital, Manchester
8	Oxford Children's Hospital, Oxford
9	Alder Hey Children's Hospital, Liverpool
10	Great North Children's Hospital, Newcastle
11	Leeds General Infirmary, Leeds
12	University Hospital Southampton, Southampton
13	The Royal London Hospital, London
14	Nottingham University Hospitals NHS Trust, Nottingham

The recruitment was based on set inclusion and exclusion criteria (**Table 2.2**) and took place between September 2016 and December 2018. Patients were recruited by their local tertiary centres and provided with information leaflets. Advertisements were also placed in the UK-CAH support group newsletter. For the recruitment of controls, local advertisements were placed through associated universities, NHS trusts and GP surgeries.

Table 2.2. Inclusion and exclusion criteria in the CAH-UK cohort

Patients	
Inclusion criteria	**Exclusion criteria**
• Known diagnosis of 21OHD confirmed by hormonal or genetic testing • Age between eight and 18 years • Capacity to consent/assent and provide a signed and dated informed consent	• Pregnancy
Controls	
Inclusion criteria	**Exclusion criteria**
• Age between eight and 18 years • Capacity to consent/assent and provide a signed and dated informed consent	• Past or present history of an endocrinopathy (all stages) • Type 1, diabetes, type 2 diabetes, insulin resistance • Known conditions of lipid/ cholesterol metabolism • Presence of any psychiatric disorder, current or past use of psychiatric medication • Glucocorticoid use within the last 6 months • Diagnosed learning difficulties and/or full-scale IQ <70 • Medication known to affect steroid metabolism • Pregnancy

2.1.1 Data collection

The study visits were organised by each tertiary centre, where patients presented between 8:00 and 9:00 am, in a fasted state. They were advised to take the usual morning GC dose. The study visit included a physical examination with measurement of height, weight, waist and hip circumference, four-limb blood pressure, assessment of pubertal development, as well as collection of blood samples. The medical history was taken from the notes, patients and parents, and consisted of details regarding the initial presentation, history of adrenal crisis, genital surgery, current medication and the most recent bone age. No formal assessment of treatment compliance was undertaken, however, the time of the last GC dose, as reported by patients and/or parents was documented.

The body surface area (BSA) was determined using the Mosteller formula (square root of the height (cm) multiplied by the weight (kg) divided by 3600) and mid-parental height was established by Tanner's formula (girls: [(father's height – 13 cm) + mother's height]/2 ; boys: [(mother's height + 13 cm) + father's height]/2). In keeping with the World Health Organisation definition, participants' weight status was classified based on BMI (overweight > +1SD, obesity > +2 SD). A systolic and/or diastolic pressure above the 95th percentile for age in all four limb measurements was interpreted as high blood pressure (Rao, 2016).

The biochemical profiles included electrolytes, urea, lipids, glucose, and plasma renin activity. They were measured locally, in the centres' NHS laboratories in keeping with local operating procedures. The interpretation of results was based on the local normal ranges.

Centralised sample analysis. For the measurement of blood steroid hormones (including 17-hydroxyprogesterone, androstenedione, testosterone, 11-hydroxyandrostenedione and 11-ketotestosterone) the blood samples collected locally were sent to the Biochemistry Department, University Hospital of South Manchester, Manchester where they were analysed by liquid chromatography tandem mass spectrometry (LC-MS/MS). The results for 17-hydroxyprogesterone and androstenedione were interpreted in accordance with published recommended ranges (Merke and Bornstein, 2005; Kushnir, Blamires et al., 2010). The measurement of blood insulin was conducted using enzyme-linked immunosorbent assay (ELISA) at the Clinical Chemistry Department, Sheffield Teaching Hospital Trust. The presence of Insulin resistance was defined by the Homeostatic Model of Insulin Resistance (HOMA-IR) (formula: fasting insulin (μU/L) x fasting glucose (nmol/L)/22.5), using a cut-off 1.68 for normal-weight and 3.42 for overweight individuals based on previously published data from children (Shashaj, Luciano et al., 2016). Measurement of blood cytokines and adipokines was conducted in collaboration with the Department Infection, Immunity & Cardiovascular Disease, University of Sheffield (Professor Allan Lawrie), using a Luminex Multi-Analyte Profiling platform (Thermo Fisher, United States), a technology that allows the measurement of multiple proteins in a single well. The initial project aimed to include the measurement of several cytokines (resistin, visfatin, interleukines 1b, 6, 8 and 10), however, the application of the multiplex assay was unsuccessful yielding unmeasurable results in more that 50% of participants. Consequently, only the analysis of adiponectin and leptin was included in the result section.

Metabolomic analysis. The blood sample collection was standardised (sample collection protocols included in the **Appendix**). The samples were stored frozen at -80^{o}C and

shipped on dry ice to the central laboratory for analysis. Blood metabolomic analysis of patients with CAH was carried out at the School of Biosciences, University of Birmingham (Professor Warwick Dunn) using ultraperformance liquid chromatography – mass spectrometry (UPLC-MS). Four chemical assays were applied (HILIC negative, HILIC positive, LIPIDS negative and LIPIDS positive), yielding a total of 12,001 compounds characterised by a retention time (rt), mass to charge ratio (mz) and chromatographic peak area (the quantified feature). The identity of the compounds was established based on the rt and mz using an "in-house" annotation system. The data were normalised using probabilistic quotient normalisation (PCN) and missing values imputation was conducted using the k-nearest neighbour algorithm. The comparison between different GC dose groups was conducted using the Kruskal-Wallis H and the compounds identified to be significantly different between the subgroups were selected for the remaining analysis, using the chromatographic peak area, expressed as milli-Absorbance Units (mAU*sec) as the measured outcome.

2.1.2 Statistical analysis

The clinical study was powered to explore the correlations between steroid hormones measured in plasma and saliva and continuous clinical variables (anthropometrics, GC dose). Thus, it aimed to recruit a minimum of 60 patients with 21OHD, a sample size considered to allow the detection of a simple correlation (r=0.5) between hormone concentrations in plasma and saliva as well as the proposed validation of novel biomarkers, using a two-sided test, 5% significance level test (α=0.05) with power 80% (β=0.2). Demographic data and the information related to medication (treatment doses, timing of doses) was processed using descriptive statistics. Differences between groups of participants were explored by analysis of variance (Chi-squared,

ANOVA, Mann Whitney U and Fisher's exact test) and analysis of covariance (ANCOVA or Quade test), with a statistical threshold of significance (*p* value) of < 0.05. Correlations and regressions were used to study the relationship between different variables. The statistical work and computation were carried out using R, SPSS Version 25 and GraphPad Prism 7. The weight, height and BMI were expressed as standard deviation scores, calculated with the use of the Growth Analyser RCT Software, the chosen reference normative values being the WHO data for The United Kingdom of Great Britain and Northern Ireland.

2.2 Basic science research: The metabolic profile and transcriptome of cyp21a2 mutant zebrafish

2.2.1 Zebrafish husbandry

All the experiments involving live zebrafish were approved by the United Kingdom Home Office and conducted in adherence to the local policies and the Animals Scientific Procedures Act 1986. The zebrafish were housed in the aquariums of the Zebrafish Facility of the MRC Centre for Developmental and Biomedical Genetics, University of Sheffield. Adult zebrafish were maintained in a recirculating system at 28.5^{o}C on a 10-hour dark: 14-hour light photoperiod (with the period of light commencing at 08:00 and ending at 22:00). Fish were fed twice daily (08:00 and 14:00-16:00), a combination of live artemia, flake and pellet. Embryos were obtained by natural spawning, following in crossing of *cyp21a2*-/- or wild type siblings, and incubated in E3 medium (containing 5 mmol/l NaCl, 0.17 mmol/l KCl, 0.33 mmol/l $CaCl_2$, 0.33 mmol/l $MgSO_4$) containing 2 µg/ml gentamicin, at 28.5^{o}C. Developmental stages were assessed based on hours post fertilisation (hpf) and days post fertilisations (dpf). Adult fish were culled for experiments at 18 months (*cyp21a2*) and 12 months (*cyp11a2*), being humanely euthanised by

immersion in anaesthetic agent (tricaine mesylate), followed by destruction of the brain by piercing the skull with a needle.

Of note, neither the *cyp21a2-/-* nor the *cyp11a2-/-* fish received glucocorticoid or mineralocorticoid replacement treatment at any stage of the project.

2.2.2 Genotyping *cyp21a2* mutants

The mutant zebrafish line used in the experiments had been previously generated by our group using the TALEN approach (Eachus, et al., 2017), causing a 14-bp deletion (c.del 211-224) in exon 2, leading to a frameshift with a premature stop at amino acid 96 (p.P70 fs26X). Mutant *cyp21a2-/-* fish and wild type siblings were obtained by in-crossing heterozygotes carrying the mutation and the progeny underwent genotyping following fin clipping biopsy, conducted by the aquarium technician. Due to the long time that had elapsed between the initial genotyping and the time adult fish were collected for experiments (18 months), consideration was given to potential inter-tank contamination. Thus, on culling the fish, a small portion of the caudal fin was collected to confirm the genotype. The material was placed in 100 µl NaOH (50 mM) and genomic DNA was isolated by heating at 98^{o}C for 10 minutes, then cooling and adding 10 µl Tris pH 8. End-point PCR amplification of the targeted genomic region was carried out in a 20 µl reaction with 0.5 µl (10 µM) of each primer (forward: CTCTCGTGGGCTAAACAAGC and reverse: ACATGTATCCACCATTTGCG) and 1 µl genomic DNA in FIREPol® Master Mix (Solis Biodyne, Tartu, Estonia). The PCR program consisted of an initial activation at 94°C (2 minutes) followed by 36 cycles of denaturation at 94°C (30 seconds) - annealing at 58°C (30 seconds) - elongation at 72°C (30 seconds), and finally elongation at 72°C for 10 minutes. The PCR product (10 µl) was digested by adding 2 µl mix made of BseYI (0.25

µl), 10X buffer 3.1 (1.2 µl) and water (0.55 µl). The digests were analysed in comparison to the non-digested PCR product on a 1% agarose gel, with the mutation coinciding with the absence of cleavage of the 179-bp product into 103-bp and 76-bp products (**Figure 2.1**).

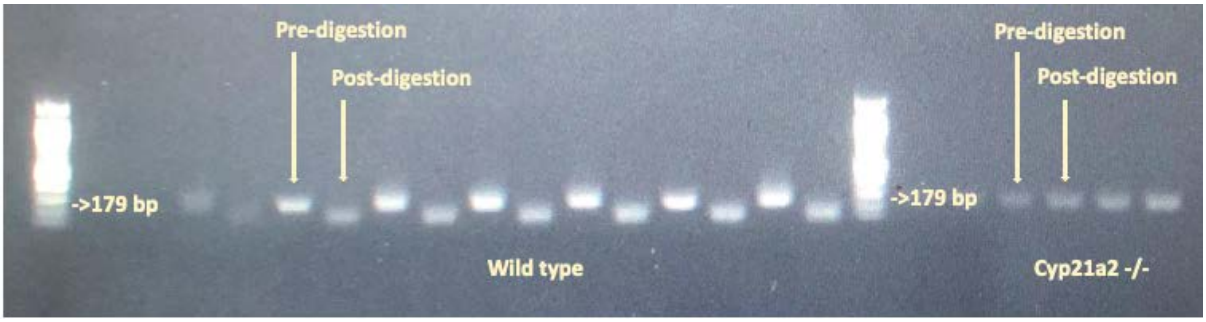


Figure 2.1. Genotyping results for the cyp21a2 fish, gel electrophoresis
The pairs of bands show the PCR product before and after BseYI digestion. The first row corresponds to wild type fish, showing a smaller DNA fragment after cleavage. The second row corresponds to *cyp21a2*-/- where the cleavage did not take place. The two rows are separated by a DNA ladder.

The *cyp11a2* fish, generated in our laboratory through the CRISPR/Cas9 approach (Li, et al., 2020), were genotyped by the aquarium technician.

2.2.3 Morphological analysis and blood sugar measurement

Adult fish were humanely euthanised, their length and weight were measured and recorded, then they were photographed under the dissecting microscope. Following tail ablation one drop of blood (0.5 µl) was expressed and blood sugar measured using a True Metrix glucometer (Trividia Health, UK). Subsequently, fish were dissected to expose the internal organs which were also photographed. Organs and tissues (including gonads, kidneys, muscle, brain, liver and gastro-intestinal tract) were collected and frozen at -80^{o}C for RNA extraction. Due to the observation that livers were larger in females compared to males, the weight of frozen livers was measured as well. All fish were culled and dissected in the morning between 09:00 and 12:00, having been fed at 08:00.

2.2.4 Haematoxylin and eosin (H&E) staining

The principle behind this technique is to combine haematoxylin, a basic dye that colours basophilic structures blue-purple contrast on basophilic structures such as chromatin, and helps primarily the visualisation of the nuclei, with the acidic dye eosin which stains cytoplasm, collagen, muscle, and red blood cells in red, pink and orange. The result is the convenient visualisation of the basic cell components and their organisation within tissues. Adult fish were humanely euthanised and fixed in 4% paraformaldehyde for 4 days at 4^{o}C. After a 20-minute wash in phosphate-buffered saline (PBS), fish were placed in 0.25 M EDTA pH8 for 4 days to be decalcified, then transferred in 70% ethanol and stored at 4^{o}C. The head, the caudal and anal fin were removed, and the fish were transferred to a tissue processor (Leica TP2010) for dehydration and paraffin infiltration. Once embedded in paraffin wax, 5 µm sections were cut through the liver. Then, samples were dewaxed and rehydrated through a series of ethanol baths at decreasing concentrations. Subsequently, samples were stained with haematoxylin for 1 minute, washed in tap water and dehydrated by transfer through a series of ethanol at increasing concentrations. Then, the sections were stained with 1% eosin in 95% ethanol for 30 seconds, washed in absolute ethanol, transferred to xylene and mounted using DPX mountant. A number of three individuals per genotype (WT and homozygous) and gender were used for histology sections for each line. In the case of the *cyp11a2* line, due to limited availability of adult fish, heterozygous fish (*cyp11a2*+/-) were used instead of WT. However, there were no identifiable distinctions in the external phenotype of *cyp11a2*+/- in comparison to WT.

2.2.5 Gene expression analysis by reverse transcription - quantitative PCR (RT -qPCR)

RT - qPCR combines the synthesis of complementary DNA (cDNA) from RNA by reverse transcriptase with the amplification of specific DNA targets by polymerase chain reaction (PCR). The PCR is a quantitative (qPCR), as the amplification is monitored in real time based on fluorescence, ultimately providing a quantitative value of gene expression, represented by the number of cycles used. Total RNA was extracted from 5 dpf larvae (clusters of 30 larvae) and adult livers using Trizol (Ambion, United States). 1 µg of RNA was used to synthesise 10 µl cDNA using the Superscript II Reverse Transcriptase kit (Thermo Fisher, United States). The qPCR primers used were either taken from previous publications or designed using NCBI Primer-BLAST (**Table 2.3**). Primer efficiency was established for each primer, only primers with amplification efficiency between 90-110% being used. To measure gene expression, comparing *cyp21a2* with wild-type siblings, 10 µl reactions (5 µl SYBR Green Master Mix [Promega, USA] 1 µl forward and reverse primers (10µM), 3 µl water and 1 µl cDNA) were run on a 7900HT Fast Real-Time PCR System. The following settings were used: 95°C (5 minutes) followed by 40 cycles of denaturation at 95°C (15 seconds) - annealing at 60°C (60 seconds) - elongation at 95°C (15 seconds). Five biological replicates with three technical replicates for each were used for all the qPCR experiments. Quantitative gene expression was calculated using the Delta-Delta Ct Method, with elongation factor 1 alpha (ef1a) used as housekeeping gene. Fold changes in gene expression were expressed in relation to the gene expression in WT. For all biological samples qPCR was initially conducted for *fkbp5,* an important gene involved in the response to cortisol, as a means to ensure the cortisol presence/deficiency.

The qPCR experiments were conducted twice (2020 and 2021) using the same methodology but different biological samples, in order to check the validity and reliability of the results. On both occasions 5 independent larval samples were used, for both mutants and controls, each resulting from pooling 30 individual 5 dpf larvae. The first qPCR experiment used 5 adult livers (2 males and 3 females) for both *cyp21a2-/-* and wild-type fish, due to limited sample availability. The second time (2021), the experiment was conducted using 5 male and 5 female livers for both *cyp21a2-/-* and controls. However, one of the WT female livers had significantly low expression of FK506 binding protein 5 (*fkbp5)* and was consequently removed from the analysis of the glucose metabolism genes as an outlier. For the experiments exploring the expression of fat metabolism genes another complementary DNA (cDNA) sample was obtained from a different WT female.

2.2.5.1 Primer design and establishing primer efficacy

The qPCR primers used were either taken from previous publications (**Table 2.3**) or designed using NCBI Primer-BLAST. The search parameters used in NCBI Primer-BLAST included a PCR product size between 70 and 200 base pairs, the optimal melting temperature of 62^{o}C with a range 59 – 65^{o}C, with exon junction span. The selected primers had a length between 18 – 22 nucleotides and GC content between 50 – 60%. The first step in establishing primer efficacy consisted of conducting end-point PCR using cDNA from wild type fish (PCR settings: activation at 94°C (2 minutes) followed by 36 cycles of denaturation at 94°C (30 seconds) - annealing at 58°C (30 seconds) - elongation at 72°C (30 seconds), and, finally, elongation at 72°C for 10 minutes). In the case of primers that did not yield a unique clear band on agarose gel electrophoresis, the annealing temperature of 58^{o}C was considered suboptimal and a gradient end-point PCR

was conducted, with annealing temperatures ranging between 50 – 60ºC by 2ºC increments. In all cases the best electrophoresis bands were obtained for annealing temperatures of 58 – 60ºC, consequently a melting temperature of 60ºC was set for all qPCR experiments. The next step involved establishing primer efficacy by qPCR using a series of cDNA dilutions (1:1 – 1:4 – 1:16 – 1:64 – 1:256). The PCR efficacy was calculated based on the slope of the linear correlation between the number of cycles and the cDNA dilution (Pfaffl. 2001).

Table 2.3. qPCR Primers

Gene	Metabolic process	Primers	Efficacy (%)	Source
acrpl	Adipocyte endocrine function	F: ATGTGGATGGGCGGTGTTCC R: TCCAGGGTCGCCTTTCTCAC	103.7	This thesis
cebpa	Adipocyte differentiation	F: GCGCAAACCACCATGCATCT R: GGCTATGTTGTTGCGCTCCC	97	This thesis
ef1a	Housekeeping gene	F: GTGGCTGGAGACAGCAAGA R: AGAGATCTGACCAGGGTGGTT	99	(Oakes, et al., 2019)
fabp11a	Lipid storage	F: GGTTGACAAATTCGTAGGAACGT R: AACCCACACCTATAGCCTTCATG	96.2	(Landgraf, et al., 2017)
fasn	Lipid storage	F: GAGAAAGCTTGCCAAACAGG R: GAGGGTCTTGCAGGAGACAG	95.1	(Lin, et al., 2017)
gck	Glycolysis	F: GCTGTGAAGTCGGCATGATA R: CTTCAACCAGCTCCACCTTAC	95.4	This thesis
hadhaa	ß-oxidation of fatty acids	F: CAGGCCTGTACGACAGCAGA R: GGGGTGCCCAGAACTGTCTT	100.6	This thesis
hadhb	ß-oxidation of fatty acids	F: GACCAACATACGCCACCCCA R: TGAGCGAACCAGTCAGAGCC	95.0	This thesis
hmgcra	Cholesterol synthesis	F: CTGAGGCTCTGGTGGACGTG R: GATAGCAGCTACGATGTTGGCG	98.6	(Lin, et al., 2017)
hmgcrb	Cholesterol synthesis	F: CCTGTTAGCCGTCAGTGGA R: TCTTTGACCACTCGTGCCG	98.5	(Lin, et al., 2017)
insa	Insulin signalling	F: GTTGGTCGTGTCCAGTGTAAGCACTAA R: CCACCTCAGTTTCCTGGGCAGATTTA	89.7	(Jurczyk, et al., 2011)
irs2a	Insulin signalling	F: TTCATCCACCACCACAGTTG R: GACTCATACTCCTCATCAGAACC	89.8	(Meng, et al., 2017)

lep	Adipocyte endocrine function	F: CTGGGAGAATCAGGCAGGGG R: GCAATACCTTTGAGATGTGAGGACT	94.5	This thesis
lpl	Lipolysis	F: CTGAGGGCTCTCGTTCATAAAGA R: AATCCATCAAAGACTGTAACTTCAATACA	97.3	(Landgraf, et al., 2017)
pck1	Gluconeogenesis	F: ATCACGCATCGCTAAAGAGG R: CCGCTGCGAAATACTTCTTC	98.0	(Seiliez, et al., 2013)
pck2	Gluconeogenesis	F: TGCCTGGATGAAATTTGACA R: GGCATGAGGGTTGGTTTTTA	96.4	(Seiliez, et al., 2013)
pkl	Glycolysis	F: TCCTGGAGCATCTGTGTCTG R: GTCTGGCGATGTTCATTCCT	92.7	(Rocha, et al., 2015)
pkm	Glycolysis	F: TGGGCTTATTAAGGGCAGTG R: TGCACCACCTTTGTGATGTT	93.5	(Rocha, et al., 2015)
pparg	Adipocyte differentiation	F: GCTGCACAGGCGCTTCA R: CTCCAGCTCCTCCAGTTCCA	100.5	(Landgraf, et al., 2017)
srebf1	Lipid storage	F: TGCCCACTCTTCTGGTGTGG R: CTGAGACGCCTGCGCAAAAT	93.5	This thesis

(Abbreviations: ***acrpl****:* adiponectin*,* ***cebpa****:* CCAAT/enhancer-binding protein alpha*,* ***ef1a****:* elongation factor 1-alpha*,* ***fabp11a****:* fatty acid binding protein 11a*,* ***fasn****:* fatty acid synthase*,* ***gck****:* glucokinase*,* ***hadhaa****:* hydroxyacyl-CoA dehydrogenase trifunctional multienzyme complex subunit alpha a*,* ***hadhb****:* hydroxyacyl-CoA dehydrogenase trifunctional multienzyme complex subunit beta*,* ***hmgcra****:* 3-hydroxy-3-methylglutaryl-CoA reductase a*,* ***hmgcrb****:* 3-hydroxy-3-methylglutaryl -CoA reductase b*,* ***insa****:* insulin*,* ***irs2a****:* insulin receptor substrate 2a*,* ***lep****:* leptin*,* ***lpl****;* lipoprotein lipase*,* ***pck1****:* phosphoenolpyruvate carboxykinase 1*,* ***pck2****:* phosphoenolpyruvate carboxykinase 2*,* ***pkl****:* liver pyruvate kinase*,* ***pkm****:* muscle pyruvate kinase*,* ***pparg****;* peroxisome proliferator activated receptor gamma*,* ***srebf1****:* sterol regulatory element-binding protein 1

2.2.6 Steroid measurement from adult fish

The steroid measurement was performed in collaboration with K.H. Storbeck, (Department of Biochemistry, Stellenbosch University, Stellenbosch, Matieland, South Africa). The whole bodies of adult 180 dpf cyp21a2-/- and WT siblings (3 males and 3 females each) were snap frozen on dry ice. The samples were homogenised using a Mikro-Dismembrator S (Sartorius, Gottingen, Germany) and then freeze-dried. Approximately 100 mg of the dried samples were transferred to a glass test tube and resuspended in 900 µL MilliQ water and 100 µL MilliQ water containing internal standard.

The steroids were extracted twice using 3 ml methanol. Following centrifugation for 5 min at low speed the methanol fractions for each sample were pooled and dried under a stream of nitrogen at 45°C. The dried residue was resuspended in 150 µL 50% (v/v) methanol prior to analysis. Steroids were separated and quantified using an Acquity UPLC System (Waters, Milford, CT, USA) coupled to a Xevo TQ-S tandem mass spectrometer (Waters, Milford, CT, USA) as previously described (O'Reilly et al. 2017).

2.2.7 Statistical analysis

The sample size of zebrafish was powered for gene expression in adult fish. Based on our preliminary data, there is typically at least a 75% difference for gene expression between wild-type (WT) and mutants for classical markers. To detect at least 10% differences for gene involved in metabolism between groups, 5 - 7 animals per group are required (power=0.8, α=0.05). All the statistical analysis was conducted using GraphPad Prism (GraphPad Software, San Diego, California, United States). Due to the small number of samples, data normality was assessed using the Shapiro-Wilk test. Normally distributed data were analysed using unpaired *t*-tests, with the Welch's correction for unequal variances. Non-normally distributed data was compared between groups using the Mann-Whitney test. Statistical significance was considered a *p* value below 0.05. Graphics were produced using GraphPad Prism and Adobe Illustrator (Adobe Inc., United States).

2.2.8 RNA sequencing and transcriptomic analysis of zebrafish larvae and adult livers

2.2.8.1 RNA sequencing

RNA was extracted from zebrafish larvae and adult livers using the Qiagen RNAeasy Kit (Qiagen, Germany). Specifically, total larval RNA was obtained from pooled 5 dpf larvae (n=30 larvae per sample) resulted from genotyped homozygous fish (cyp21a2-/- mutants and WT siblings). Four WT and three mutant larval RNA samples were sent for sequencing. Livers were harvested and RNA extracted from 18-month-old mutants and WT fish (6 males and 6 females for each group). The quality of the RNA extracted was determined by a 2100 Bioanalyzer (Agilent, Sheffield Institute for Translational Neuroscience, University of Sheffield). Larval RNA samples were processed in the Institute of Cancer and Genomic Sciences, University of Birmingham, library preparation being performed using NEBNext® Ultra™ II Directional mRNA Library Prep Kit for Illumina®, followed by high-output 150 cycles, paired-end sequencing on a NextSeq 500 system. Liver RNA samples were sent to the Deep Sequencing Facility in Dresden Genome Center where mRNA-Isolation via poly-dT enrichment and strand-specific RNA-Seq library preparation was followed by paired-end sequencing using a NovaSeq 6000 S4 Reagent Kit v1.5 (200 cycles) 4-lane XP mode.

2.2.8.2 Pre-processing of RNA sequencing raw data

The raw RNA sequencing data was pre-processed using the University of Sheffield's latest central High-Performance Computing (HPC) cluster ShARC. The fastq.gz files that were produced by RNA sequencing were first subjected to quality control using FastQC, which was also applied at each step of the subsequent pre-processing sequence. Next, rRNAs was removed with SortmeRNA version 2.1 (Kopylova, et al., 2012), then the files were trimmed with Trimmomatic version 0.39 (Bolger, et al., 2014) (ILLUMINACLIP:Trimmomatic 0.39/adapters/TruSeq3- SE.fa:2:30:10 LEADING:3

TRAILING:3 SLIDINGWINDOW :4:15 MINLEN:36). The Phred score threshold set when running trimmomatic was Phred 33. FastQC quality control was also conducted after this step, individual quality control assessment being combined with MultiQC (Ewels, Magnusson et al., 2016). Finally, the mapping/alignment of the reads to the reference zebrafish genome (GRCz11) was performed with STAR (Dobin and Gingeras, 2015). The code used for the pre-processing pathway can be found in the **Appendix**.

2.2.8.3 Differential gene expression analysis following RNA sequencing

FeatureCounts was used for gene count quantification from the reads of RNA-seq data for each gene for each sample (Liao, et al., 2014). The bioinformatic analysis was performed in R (version 4.2.0) using the RStudio. Differential expression analysis was conducted with the 'DESeq2' package. The threshold of significance was set at the adjusted p-value < 0.05. Principal component analysis and heat map production was performed in R studio by the 'plotPCA' and 'pheatmap' packages. The code used for the pre-processing pathway can be found in the **Appendix**.

2.2.8.4 Gene ontology (GO) term overrepresentation

Differentially expressed genes (DEGs) were identified by using the 'DESeq2' package in R, with a threshold of padj < 0.05. Statistical overrepresentation analysis of gene ontology (GO) terms with DEGs was conducted using two different bioinformatics tools. First, GO enrichment was performed with GOrilla (Eden et al., 2009), entering as background list all genes identified by RNA sequencing and ranked using the 'DESeq2' package in R, setting a threshold $p < 10^{-5}$ and using REVIGO as a visualisation tool. Secondly, the 'clusterProfiler' Bioconductor package was used in R, analysing a list of significantly

differentially expressed genes defined by p-adj < 0.05 against all annotated genes, setting a q-value of < 0.05 to identify overexpression of biological processes.

2.2.8.5 Gene set enrichment analysis

Gene set enrichment analysis was conducted in R using the 'gseGO' function of the 'clusterProfiler' package. The input gene set consisted of the ranked list of all annotated genes following RNA sequencing and this was analysed against the genome wide annotation for zebrafish available on the 'Bioconductor' website ('org.Dr.eg.db'), using the Ensembl ID as gene identifier. The graphic visualisation of results was achieved using the 'ridgeplot' function in R and PraphPad Prism.

2.2.8.6 Exploring associations with human pathology

The association of differentially expressed genes with human disease were explored using the DisGeNET database (where the human orthologs of the DEG's were identified) and the R package "disgenet2r". The graphics were produced using R, GraphPad Prism, REVIGO and the Venn diagram tool on the Bioinformatics and Evolutional Genomics site (www.bioinformatics.psb.urgent.be/webtools).

3. The metabolic profile of patients with CAH in relation to hormonal profiles

Clinical research was carried out using clinical and biological data (steroid and metabolic profiles, adipokines and metabolomics) from a UK-wide multicentre cohort study (CAH-UK) involving 107 children with 21OHD, as well as 83 healthy controls. The aim was to obtain an overview of the health status of children and young people with CAH within the UK, focusing on early signs of metabolic disease, in relation to the GC replacement therapy. Clinical data, including anthropometric measurements and treatment doses, as well as hormonal, lipid and adipokine profiles were analysed. A significant number of children with CAH were found to be overweight and obese compared to healthy controls. Moreover, a small number of patients had raised total cholesterol, low HDL, raised LDL and triglycerides. Leptin and adiponectin correlated with the patients' BMI, HOMA-IR and plasma androgens.

3.1 Results

3.1.1 Study participants

(Modified from published paper (Bacila, et al., 2022))

The study recruited a total of 107 patients with 21OHD that underwent the study visit**.** The median age for patients was 12.4 years with interquartile range of 10.1 – 15.1 years and 55.1% were females. Controls were well matched with the patients' group for age and sex, however not for ethnicity, as there were fewer Asian Pakistani controls compared to patients ($p<0.001$) (**Table 3.1**). The majority of patients (62.2%) had been diagnosed in the first month of life, commonly presenting with ambiguous genitalia (32.7%) or salt losing crisis (25.2%).

3.1.2 Glucocorticoid replacement therapy

(Modified from published paper (Bacila, et al., 2022))

Most of the patients received GC replacement with hydrocortisone according to a circadian administration regime. However, there was large variation in the timing of administration and in the relative GC dose, with a third of patients exceeding the recommended dose range.

Hydrocortisone (HC) was used by 94.3% of patients for cortisol replacement, the remaining receiving prednisolone. The majority of patients on HC received three (65.3%) or four daily doses (24.8%); a smaller number received two (7 patients), five (2 patients) and six doses (1 patient).

Table 3.1. Demographic characteristics of participants

	Patients	Controls	Statistical difference
Number	n=107	n=83	
Age (years)	12.4(10.1-15.1)	13.3(10.3-15.3)	p=0.550
< 12 years	51(47.7%)	36(43.4%)	p=0.556
12-18 years	56(52.3%)	47(56.6%)	
Girls	59(55.1%)	45(54.2%)	p=0.899
Boys	48(44.9%)	38(45.8%)	
White	79(73.8%)	72(85.7%)	p<0.001
Asian Pakistani	20(18.7%)	0	
Asian other	0	6(7.2%)	
Black ethnicity	1(0.9%)	2(2.4%)	
Mixed ethnicity	6(5.6%)	2(2.4%)	
Other ethnicities	1(0.9%)	1(1.2%)	

(Statistical test: Chi-squared)

Administration times for the HC doses varied widely among patients (**Figure 3.1**). A circadian regime, with the morning dose being the highest dose of the day, was used in 66.3% of patients on HC, while only 8.4% received a reverse circadian regime. Prednisolone was commonly given in two daily doses (7:00 - 8:00 and 21:00 – 22:00), one patient received one daily dose.

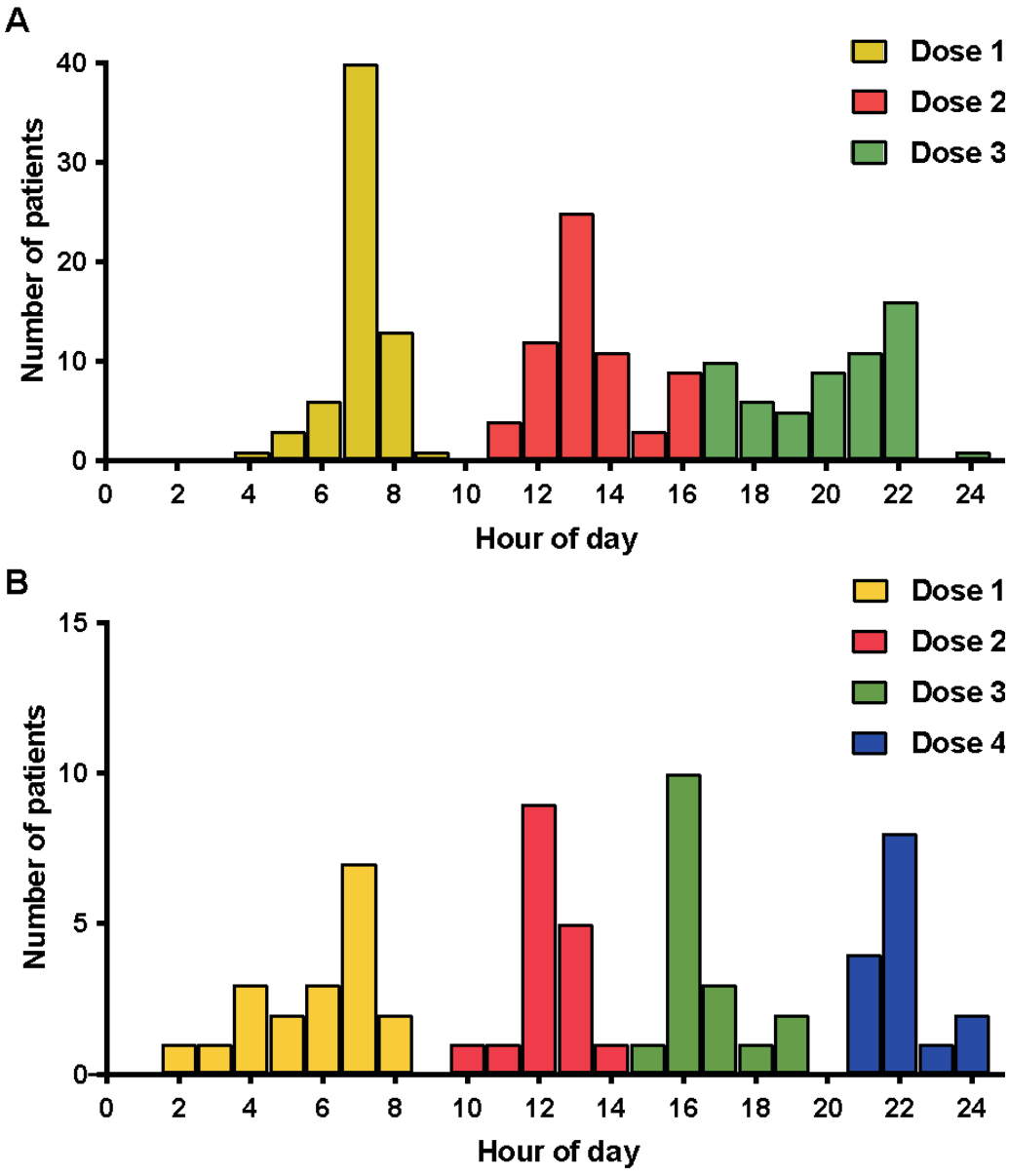


Figure 3.1. Administration times for hydrocortisone doses
Times of administration of hydrocortisone doses for children on three daily doses (**A**) and children on four daily doses regimes (**B**). Each bar represents the number of patient visits recording a dose given at that time; the different patterns correspond to the order of the doses throughout the day.

Regarding the GC replacement doses, the overall mean relative daily dose was 13.3 (±3.7) mg/m^2/day HC – equivalent. Comparing mean ranks between patient subgroups, no significant difference was found between males and females, however, doses were higher in patients aged 12-18 years compared to younger patients (p = 0.025*)* (**Figure 3.2A-C**). A GC dose within the recommended range (10-15 mg/m^2/day) (Speiser, Arlt et al., 2018), was found in only 43.9% of the patients, while 22.4% of daily doses were lower than 10 mg/m^2/day and 33.6% higher than 15 mg/m^2/day. Doses lower than 10 mg/m^2/day were more frequently used in females compared to males (p=0.036); these patients were evenly distributed among centres. No relationship was found between the timing of the doses and the adherence to the recommended dose range.

3.1.3 Mineralocorticoid replacement therapy

(Modified from published paper (Bacila, et al., 2022))

Mineralocorticoid therapy with fludrocortisone was used for 82 patients (76.6%). This was administered once daily in 85.4% of patients, most frequently at 8:00 am. The total daily doses ranged between 15 – 350 µg/ day, and the relative doses between 5.0 – 228.6 µg/m^2/day with a mean of 105.0 (±50.2) µg/m^2/day, with no marked differences between age and sex groups (**Figure 3.2D, E**).

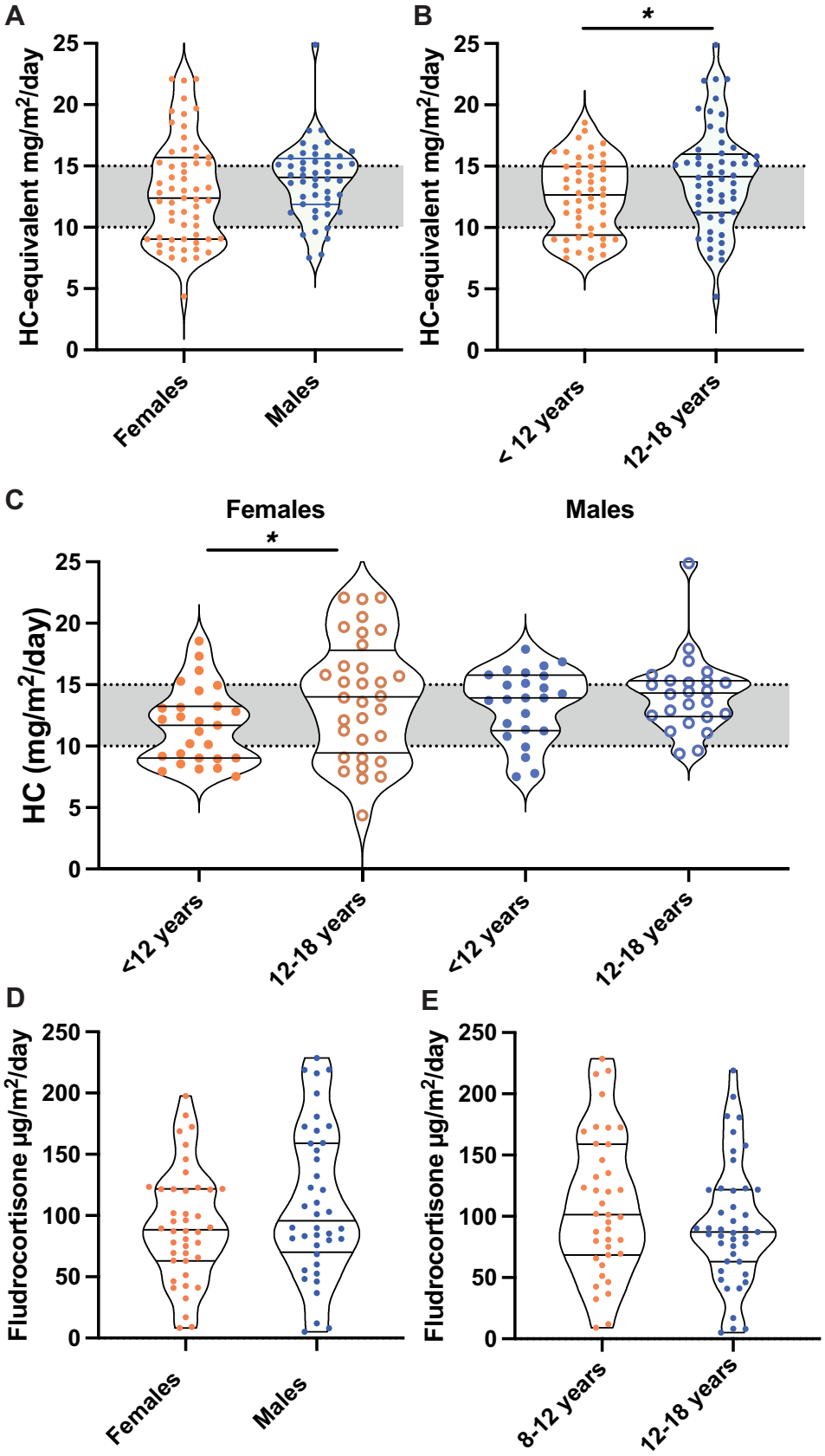


Figure 3.2. Daily replacement doses

Daily glucocorticoid doses for gender (**A**) and age (**B**) groups, as well as subgroups of age and gender (**C**). The shaded area corresponds to the recommended glucocorticoid dose range of 10-15 mg/m^2/day hydrocortisone equivalent. Daily mineralocorticoid doses for gender (**D**) and age (**E**) groups, expressed as fludrocortisone µg/m2/day. The horizontal lines correspond to the median, 25th and 75th quartiles. (Statistical test: Mann Whitney U; * indicates statistical significance)

3.1.4 Biochemical markers of control

(Modified from published paper (Bacila, et al., 2022))

The conventional hormonal markers of disease control were frequently abnormal, with androstenedione being raised in over half of the patients (**Figure 3.3**), especially in the case of male patients. The extended steroid profiles showed significantly raised concentrations of androgens and androgen precursors in patients compared to controls for all subgroups of age and sex, with the exception of testosterone.

Androstenedione concentrations were above the normal range in 67.8% and suppressed in 4.4% of patients. Plasma 17-hydroxyprogesterone (17OHP) was below the recommended range suggesting suppression in 13% of patients (**Figure 3.3**). Hormone concentrations were similar between most sex and age subgroups; however, androstenedione was above range more frequently in males compared to females (79.5% vs 56.5%, p=0.04). There was no correlation between plasma androstenedione or 17OHP and the relative GC dose.

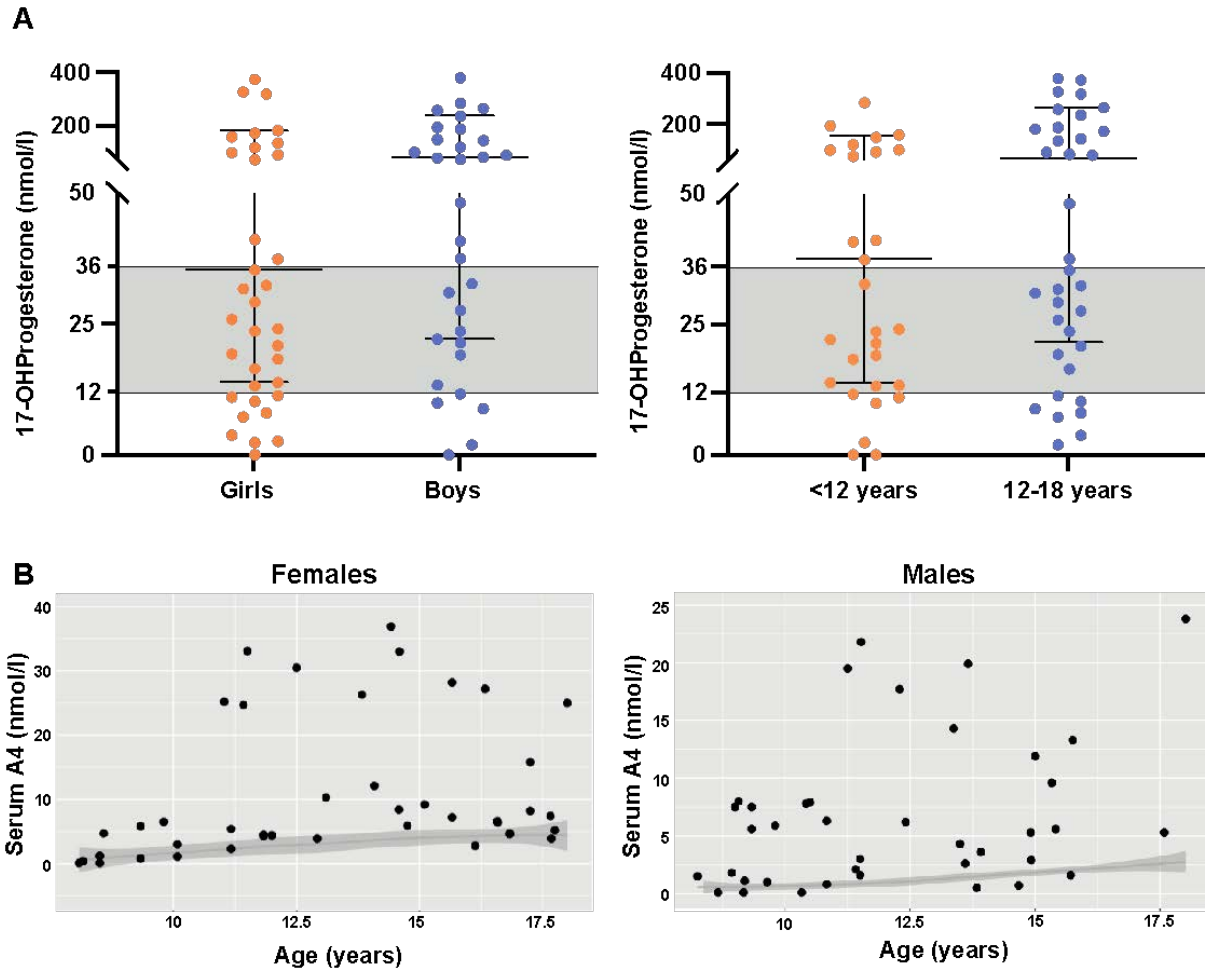


Figure 3.3. Serum biomarkers
(**A**) Concentration of serum 17-hydroxyprogesterone in CAH patients for gender and age groups. The shaded areas correspond to the recommended range for hormone concentrations in CAH. (**B**) Concentration of serum androstenedione in gender groups; the grey ribbons correspond to the normal ranges for age and gender. A4: androstenedione

3.1.4.1 Extended steroid profiles

(Modified from published paper (Bacila, et al., 2019))

Plasma androstenedione, 17-hydroxyprogesterone, 11-hydroxyandrostenedione and 11-ketotestosterone were higher in patients with 21OHD compared to controls ($p<0.001$). This finding was consistent across age and sex subgroup. Plasma testosterone concentrations were overall similar for patients and controls, however, a subgroup analysis showed that female patients had raised testosterone compared to controls ($p<0.001$). Similarly, testosterone was also significantly higher in patients younger than

12 years (p=0.008) in comparison to matched controls. Male patients showed an opposite trend, having lower plasma testosterone than controls (p=0.01), the difference being more marked in pubertal boys (p=0.001) (**Figure 3.4**). In patients with CAH due to 21OHD, strong (r_s*=0.6–0.8*) and very strong (r_s*>0.8*) correlation correlations were detected between plasma 17-hydroxyprogesterone, androstenedione, testosterone and the 11-oxygenated androgens (11-hydroxyandrostenedione and 11-ketotestosterone), with the strongest correlation found between androstenedione and 11-hydroxyandrostenedione (r_s*=0.895, p<0.001*). These correlations also existed in controls, although they were weaker (r_s*=0.5–0.7*), the strongest relationship still being found between androstenedione and 11-hydroxyandrostenedione (r_s*=0.690, p<0.001*).

Cortisol concentrations ranged from 14.7 to 827.3 nmol/l in CAH patients, with a median of 287.3 nmol/l, which was comparable to the values found in healthy controls. (**Figure 3.5**). A cortisol concentration below 100 mmol/l was found in 20 patients (19%), of which only five were treated with prednisolone. There was no correlation between plasma cortisol and any of the five adrenal steroids measured in patients with CAH. This was not the case in healthy controls where plasma cortisol correlated with 17-hydroxyprogesterone (r_s= 0.346, p = 0.006*)*, androstenedione (r_s= 0.302, p = 0.017), testosterone (r_s= 0.227, p = 0.079), 11-hydroxyandrostenedione (r_s= 0.642, p < 0.001) and 11-ketotestosterone (r_s= 0.116, p = 0.370).

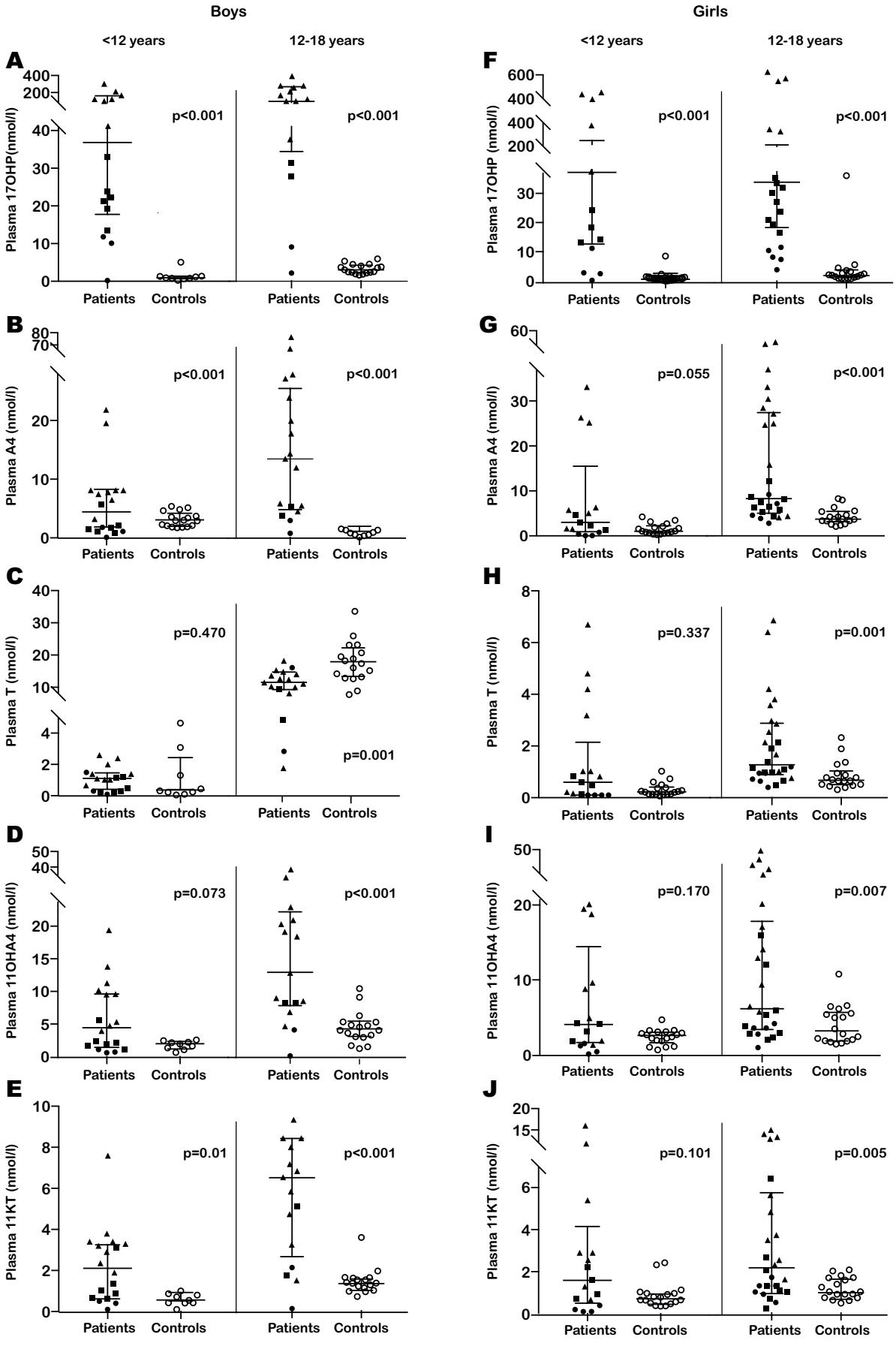

Boys
Girls
<12 years
12-18 years
A
Plasma 17OHP(nmol/l)
p<0.001
p<0.001
B
Plasma A4 (nmol/l)
p<0.001
p<0.001
C
Plasma T (nmol/l)
p=0.470
p=0.001
D
Plasma 11OHA4 (nmol/l)
p=0.073
p<0.001
E
Plasma 11KT (nmol/l)
p=0.01
p<0.001
F
Plasma 17OHP (nmol/l)
p<0.001
p<0.001
G
Plasma A4 (nmol/l)
p=0.055
p<0.001
H
Plasma T (nmol/l)
p=0.337
p=0.001
I
Plasma 11OHA4 (nmol/l)
p=0.170
p=0.007
J
Plasma 11KT (nmol/l)
p=0.101
p=0.005
Patients
Controls

Figure 3.4. Plasma steroid concentrations
Comparison of plasma steroid concentrations between patients and controls. Patient values are represented by coloured shapes, classified according to 17-hydroxyprogesterone concentrations as low (black circles), normal (black squares) and high (black triangles); control values are represented by clear circles. The left column represents the results obtained for boys younger than 12 years of age (left pair of scatter columns) and 12 to 18 years of age (right pair of scatter columns) for 17-hydroxyprogesterone (**A**), androstenedione (**B**), testosterone (C), 11-hydroxyandrostenedione (**D**) and 11-ketotestosterone (**E**). The right column represents the results obtained for girls younger than 12 years of age (left pair of scatter columns) and 12 to 18 years of age (right pair of scatter columns) for 17-hydroxyprogesterone (**F**), androstenedione (**G**), testosterone (H), 11-hydroxyandrostenedione (**I**) and 11-ketotestosterone (**J**). The horizontal bars within the scatter columns correspond to the median and interquartile range. (17OHP: 17-hydroxyprogesteone, A4: androstenedione, T: testosterone, 11OHA4: 11-hydroxyandrostenedione, 11KT: 11-ketotestosterone). (Statistical test: Mann Whitney U; * indicates statistical significance) (Modified from Bacila, et al., 2019)

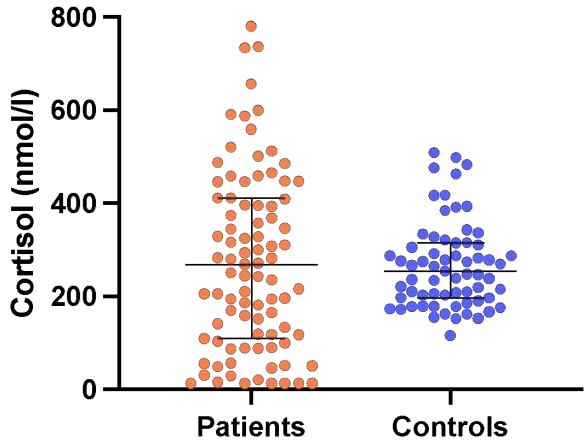


Figure 3.5. Plasma cortisol concentrations in patients and controls
The horizontal bars within the scatter columns correspond to the median and interquartile range.

3.1.5 Clinical and anthropometric characteristics
(Modified from published paper (Bacila, et al., 2022))

Patients with CAH had advanced bone age and their height-SDS decreased with age, pubertal patients being shorter compared to healthy controls. A higher prevalence of overweight and obesity was found in CAH patients compared to controls.

Assessment of pubertal development by Tanner staging found advanced pubarche in females (p=0.002) and genital development males (p=0.021) with CAH compared to

controls. Patients with CAH had overall increased height-SDS in relation to the mid-parental height-SDS (p=0.005). However, there was a significant decrease in the height-delta SDS (representing the difference between patient-height SDS and mid-parental-height SDS) in relation to age, as shown by linear regression (**Figure 3.6.A**). Compared to healthy controls, 8-12-year-old patients were taller (p=0.011) and patients aged 12-18 years were shorter (*p*=0.037) (**Table 3.2**). The weight SDS were higher in CAH patients (p=0.009), as was the case for the BMI (p=0.004) and hip circumference (p=0.006), compared to controls (**Figure 3.6.B**). Analysing subgroups of participants, the BMI was found to be higher in female patients (p=0.004) and patients aged 12-18 years (p=0.006). The hip circumference was higher in patients compared to controls only for males aged 8-12 years (p=0.028). In females, waist-to-hip ratio was significantly higher in patients compared to controls (p=0.01), however, there was no difference in males. Importantly, while in controls males has higher waist-to-hip ratio than females (p=0.019), in patients there was no significant difference between sexes. There was increased prevalence in patients of overweight (26.4% vs 10.8% in controls) and obesity (22.6% vs 10.8% in controls) (p<0.001) also found in all subgroups of age and sex. Patient BMI-SDS did not correlate with GC doses or plasma androstenedione and 17OHP.

Bone age results from 85 patients showed a difference between bone age and chronological age of 1.9 (±2.2) years. Advanced bone age by more than 1.5 years was found in 43 patients and delayed bone age by over 1.5 years in 2 patients. There was a significant difference between age groups, as children aged 8-12 years had a more advanced bone age compared to patients aged 12-18 years (2.4 (±2.4) years vs. 0.9(±1.4) years, p=0.004). There was no significant difference in this parameter between sex groups and no correlation between bone age and relative GC doses.

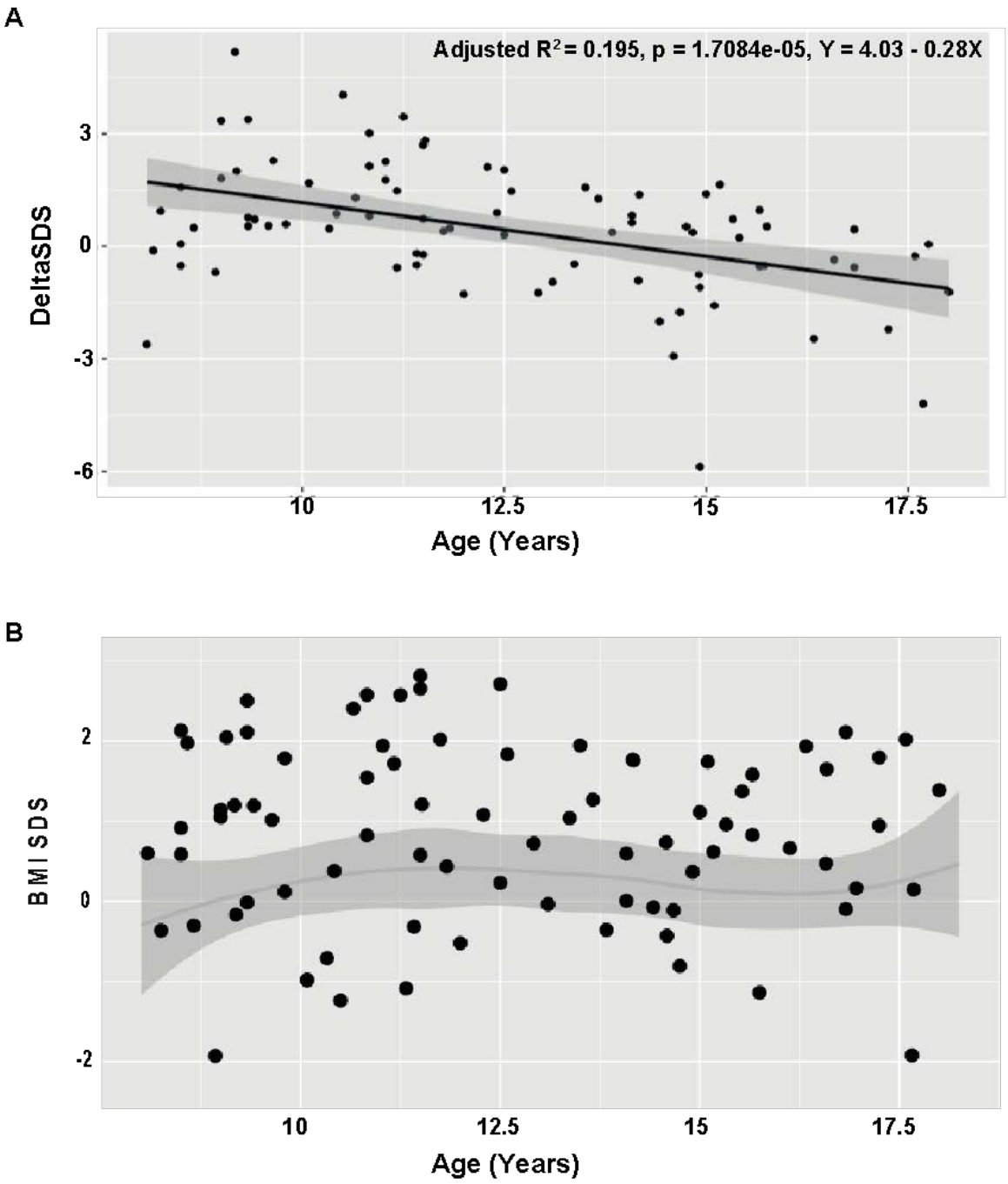


Figure 3.6. Anthropometric characteristics in patients with CAH
A. Linear regression showing the relationship between height-delta SDS (difference between height SDS and mid-parental-height SDS in patients with CAH). **B.** Patients' BMI SDS for age and gender (males: dark dots, females: clear dots). The shaded area represents a regression estimate based on the BMI SDS of the 83 healthy controls.

Table 3.2. Anthropometric data in patients and controls

Variable	Patients	Controls	Patients vs Controls*	Subgroup comparison
Height				
All	**0.3 (-0.7 - 1.2)** n=107	**0.1 (-0.4 - 0.8)** n=83	***p=0.658***	
Males	**0.4 (-0.7 - 2.2)** n=48	**0.0 (-0.6 - 1.5)** n=38	***p=0.267***	
Females	**0.1 (-0.7 - 0.7)** n=59	**0.1 (-0.3 - 0.7)** n=45	***p=0.800***	***p=0.022[a]*** ***p=0.188[b]*** ***p=0.269[c]*** ***p=0.062[d]***
Age 8-12 y	**0.7 (0.1 - 2.3)** n=51	**0.1 (-0.3 - 1.2)** n=36	***p=0.011***	
Age 12-18y	**-0.2 (-1.0 - 0.4)** n=56	**0.1 (-0.5 - 0.6)** n=47	***p=0.037***	
Weight				
All	**0.8 (-0.0 - 1.9)** n=107	**0.2 (-0.4 - 1.0)** n=83	***p=0.009***	
Males	**1.2 (0.1 – 2.0)** n=48	**0.3 (-0.2 - 1.2)** n=38	***p=0.078***	
Females	**0.4 (-0.0 - 1.7)** n=59	**0.1 (-0.5 - 0.7)** n=45	***p=0.026***	***p=0.095[a]*** ***p=0.783[b]*** ***p=0.129[c]*** ***p=0.117[d]***
Age 8-12 y	**1.2 (-0.4 - 1.6)** n=51	**0.5 (-0.4 - 1.4)** n=36	***p=0.019***	
Age 12-18y	**0.6 (-0.4 - 1.6)** n=56	**0.1 (-0.3 - 0.6)** n=47	***p=0.225***	
BMI				
All	**0.9 (-0.0 - 1.9)** n=107	**0.2 (-0.3 - 0.8)** n=83	***p=0.004***	
Males	**1.1 (-0.1 - 1.9)** n=48	**0.3 (-0.1 - 1.0)** n=38	***p=0.076***	
Females	**0.7 (-0.0 - 1.9)** n=59	**0.0 (-0.4 - 0.6)** n=45	***p=0.004***	***p=0.434[a]*** ***p=0.149[b]*** ***p=0.187[c]*** ***p=0.006[d]***
Age 8-12 y	**1.0 (-0.2 - 2.1)** n=51	**0.4 (-0.3 - 1.0)** n=36	***p=0.100***	
Age 12-18y	**0.9 (-0.0 - 1.7)** n=56	**0.1 (-0.3 - 0.8)** n=47	***p=0.006***	
Waist circumference				
All	**0.8 (0.2 - 1.9)** n=95	**0.7 (-0.1 - 1.4)** n=78	***p=0.102***	***p=0.672[a]*** ***p=0.448[b]*** ***p=0.280[c]*** ***p=0.236[d]***
Males	**0.9 (0.2 - 1.9)** n=46	**0.7 (-0.1 - 1.4)** n=34	***p=0.475***	
Females	**1.0 (0.2 - 1.9)**	**0.6 (-0.1 - 1.2)**	***p=0.116***	

Age 8-12 y **Age 12-18y**	n=49 **1.0 (0.2 - 1.9)** n=44 **0.8 (0.2 - 1.8)** n=51	n=44 **0.9 (-0.2 - 1.4)** n=33 **0.6 (-0.1 - 1.1)** n=45	***p=0.399*** ***p=0.168***	
<u>Hip circumference</u> **All** **Males** **Females** **Age 8-12 y** **Age 12-18y**	**0.9 (-0.5 - 1.7)** n=97 **1.3 (-0.2 - 2.3)** n=47 **0.6 (-0.6 - 1.5)** n=50 **1.2 (-0.0 - 2.1)** n=44 **0.7 (-0.7 - 1.6)** n=53	**0.0 (-1.0 - 0.8)** n=78 **0.1 (-0.6 - 1.2)** n=34 **0.0 (-1.0 - 0.8)** n=44 **0.0 (-0.6 - 1.1)** n=33 **0.0 (-1.0 - 0.7)** n=47	***p=0.006*** ***p=0.020*** ***p=0.144*** ***p=0.020*** ***p=0.118***	***p=0.028[a]*** ***p=0.173[b]*** ***p=0.189[c]*** ***p=0.571[d]***

The variables are expressed as median with interquartile ranges of z-scores. ([a] Males age 8-12 years, [b] Males age 12-18 years, [c] Females age 8-12 years, [d] Females age 12-18 years) (Statistical test: Mann Whitney U)

Raised blood pressure was found in five patients (three females, two males) and one control (female). Their replacement doses were mostly in keeping with recommendation, ranging for GC between 9-15 mg/m^2/day and for MC between 60-200 µg/m^2/day; one was not taking fludrocortisone treatment. Of these patients, one was overweight, and one had cushingoid features. Analysis of covariance (Quade test) used to compare mean absolute systolic and diastolic blood pressures, conducted in different sex groups and using age as covariate, found no significant difference between patients and controls.

3.1.6 Biochemical markers of metabolic risk

(Modified from published paper (Bacila, et al., 2022))

Most of the patients had normal non-steroid serum biochemistry (**Table 3.3**). Raised creatinine was found in 10% of patients and a third of the patients had raised plasma renin. Altered lipid profiles were found in a modest number of patients, consisting of

raised total cholesterol (3.9%), low HDL (2.9%), raised LDL (7.8%) and triglycerides (4.9%). Using Analysis of covariance (ANCOVA) to compare absolute values of biochemical markers, using age as covariate, showed no significant difference between treatment groups (HC-equivalent in mg/m^2/day < 10 mg/m^2/day, 10-15 mg/m^2/day and > 15 mg/m^2/day) and groups of disease control based on the 17OHP values (below, within and above the 12-36 mmol/l target range). The only exceptions were LDL and total cholesterol, where levels were found to be higher in the group of overtreatment (high 17OHP), and lower in that of poor control (low 17OHP), ANCOVA results: $F(2, 77)=7.368$, $p=0.001$ for LDL cholesterol and $F(2, 77)=5.888$, $p=0.004$ for total cholesterol. There was no significant difference in LDL and total cholesterol between patients treated with circadian versus revers circadian GC replacement regimes. A significant number of CAH patients had a HOMA-IR suggestive of insulin resistance (53.8%), however, this prevalence was similar to that found in controls (56.3%).

Table 3.3. The results of biochemical investigations in patients with CAH

Test	Normal	High	Low	Result not available
Sodium	88.2%			11.8%
Potassium	86.3%		1%	12.7%
Urea	86.3%	2%		11.8%
Creatinine	77.5%	8.8%	2%	11.8%
Plasma renin	17.6%	30.4%	3.9%	48%
Total cholester	73.5%	3.9%	3.9%	18.6%
HDL	72.5%		2.9%	24.5%
LDL	61.8%	7.8%		30.4%

Triglycerides	77.5%	4.9%	1%	16.7%
Glucose	82.4%			17.6%

The results are interpreted in relation to the normal ranges of the respective local laboratories.

3.1.7 Adipokine analysis in CAH patients

In patients with CAH both leptin and adiponectin correlated with weight and BMI. Leptin was higher in patients compared to controls only in male participants. A significant relationship was found in patients between leptin and insulin, as well as leptin and the first GC dose of the day, but not the total daily dose. No relationship was found between adiponectin and insulin, however, adiponectin increased significantly with the decrease of plasma androgens.

A positive correlation between leptin and BMI was found in both patients (regression equation: *leptin(ng/dl) = 4810 + 3939 x BMI-SDS*, $p < 0.001$) and controls (regression equation: *leptin(ng/dl) = 4444 + 2301 x BMI-SDS*, $p < 0.001$). Leptin corrected for BMI also correlated positively with waist SDS in patients ($r_s = 0.381, p < 0.001$) and controls ($r_s = 0.400, p < 0.001$). Adiponectin decreased with increasing BMI in patients (regression equation: *adiponectin(pg/dl) = 98 – 11 x BMI-SDS*, $p = 0.008$) (**Figure 3.7**), but the relationship was not significant in controls. In CAH patients, adiponectin correlated negatively with both the waist circumference ($r_s = -0.477, p < 0.001$) and the hip circumference SDS ($r_s = -0.504, p < 0.001$); in controls these correlations were weaker ($r_s = -0.250, p = 0.033$ for waist circumference and $r_s = -0.256, p = 0.029$ for hip circumference SDS).

Within the patient group, there was no difference between age subgroups and although regression analysis showed an increase in leptin corrected for BMI with age, this was not statistically significant; similar findings with regards to leptin versus age were found in controls. Leptin corrected for BMI was significantly higher in females than males in patients ($p = 0.001$) and controls ($p < 0.001$) (**Figure 3.8**). Of note, there was no significant difference in weight, BMI SDS or waist over hip circumference SDS ratio between males and females.

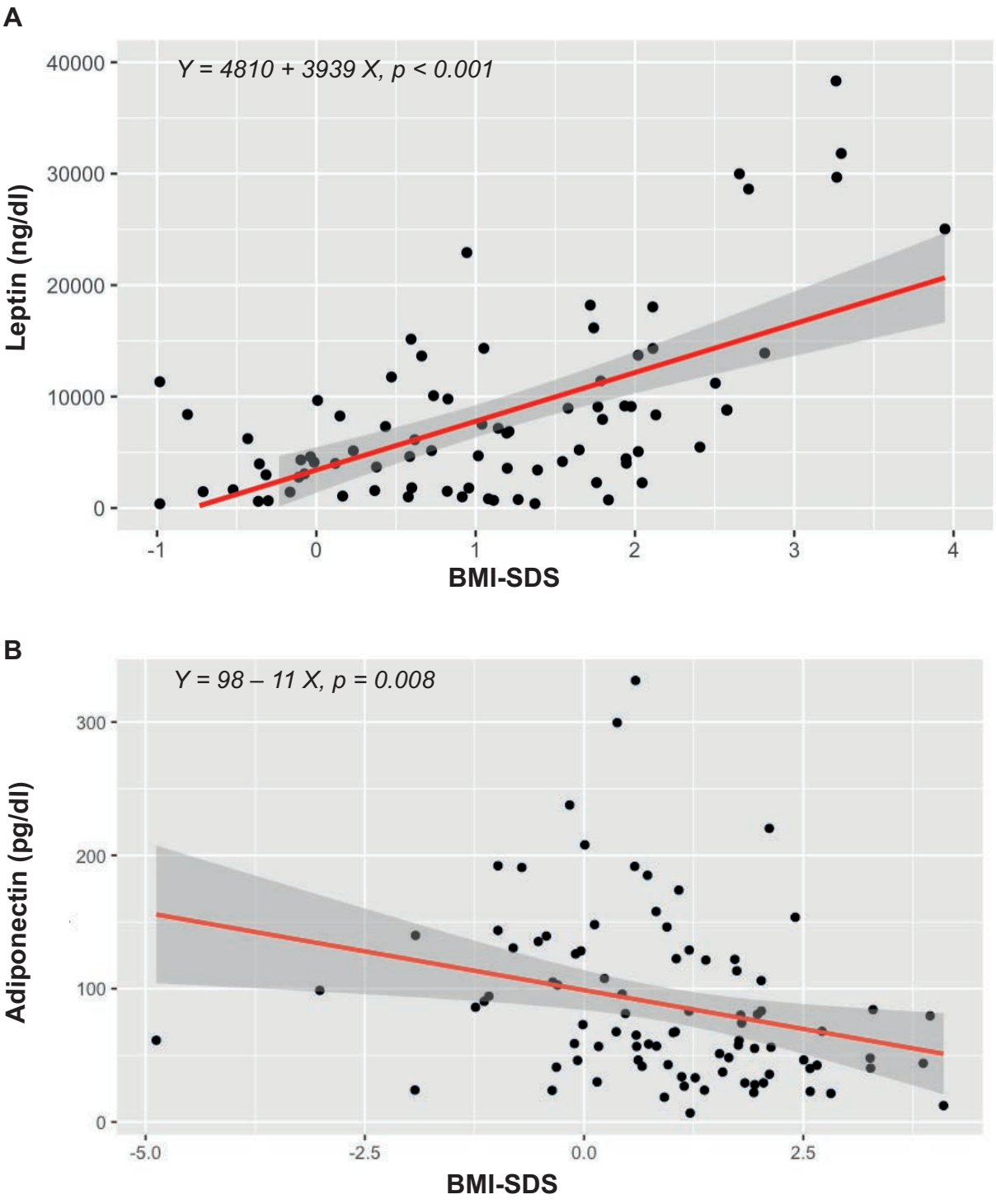


Figure 3.7. Leptin and adiponectin vs BMI
Linear regression showing a statistically significant relationship between leptin and BMI-SDS (**A**) and a negative relationship between adiponectin and BMI-SDS (**B**) in patients.

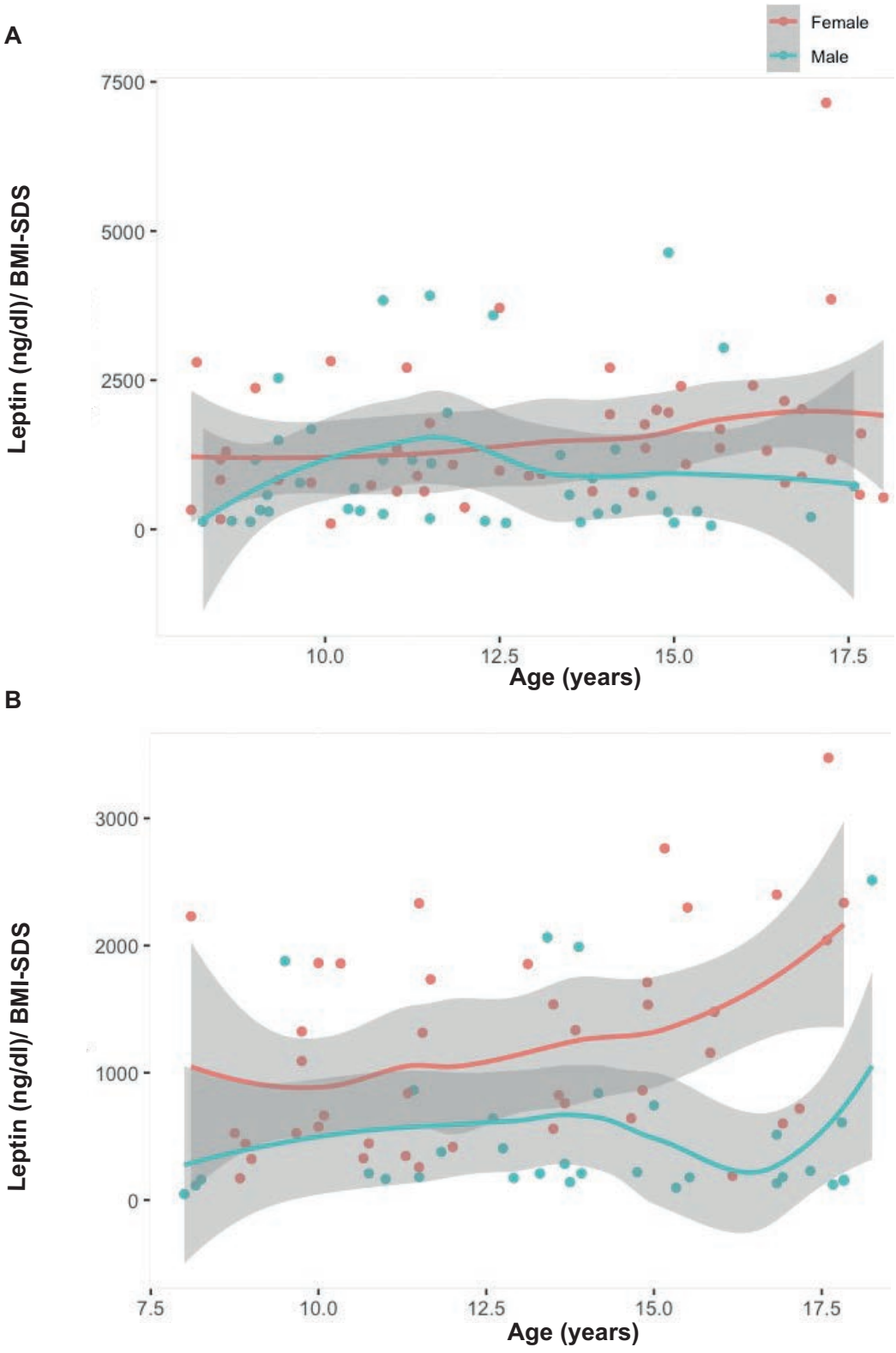


Figure 3.8. Leptin in sex groups
Comparison of serum leptin (corrected for BMI-SDS) between males (blue) and females (red) in patients (**A**) and controls (**B**). The horizontal axis displays the ages of the participants. The dots represent individual values, the grey ribbons represent the 95% confidence interval of the non-linear regression (coloured lines).

Patients had significantly higher leptin (corrected for BMI) compared to controls (p=0.027), however, when comparing sex subgroups, the difference was only significant in boys (p=0.033) (**Figure 3.9**).

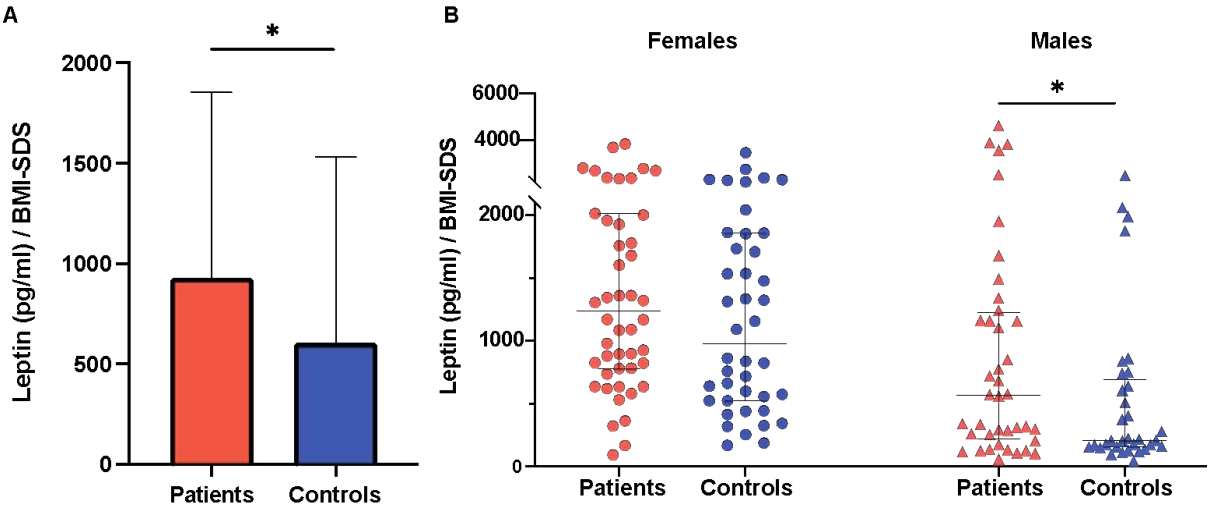


Figure 3.9. Leptin in patients vs controls
Comparison of the plasma leptin (corrected for BMI SDS) between patients (orange) and controls (blue) in all participants (**Panel A**) and sex subgroups (**Panel B**); the data is expressed as median with interquartile range (error bars). (Statistical test: ANOVA; * indicates statistical significance)

In CAH patients, leptin levels (corrected for BMI-SDS) decreased significantly with the increase of the first daily relative HC dose, irrespective of the time elapsed between the administration and sample collection. This was demonstrated by the regression analysis, with the equation: leptin(ng/dl)/BMI-SDS = 2225 – 174 x HC dose (HC-equivalent in mg/m^2/dose), *p=0.02* (**Figure 3.10**). No relationship was found between leptin and the daily relative GC dose.

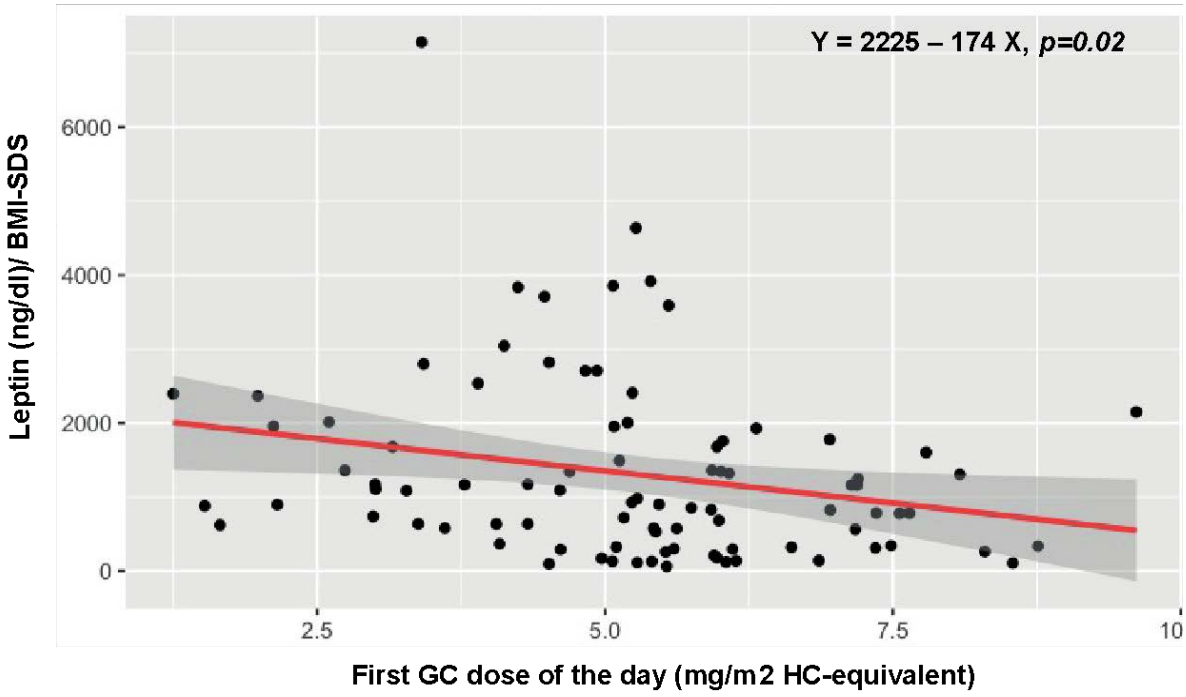


Figure 3.10. Leptin vs GC dose
Linear regression showing a statistically significant relationship between leptin corrected for BMI SDS and the first GC dose of the day, expressed in mg/m^2 HC-equivalent.

There was a weak positive correlation in patients between leptin and insulin (r_s = *0.315, p = 0.008*) and leptin and HOMA-IR (r_s = *0.302, p = 0.010*), however, the relationship was stronger in controls (r_s = *0.584, p < 0.001* for insulin and r_s = *0.496, p < 0.001* for HOMA-IR) **(Figure 3.11**). Regression analysis using patient data showed a significant increase in leptin corrected for BMI-SDS with insulin and decrease with the first relative GC dose: leptin/BMI-SDS = 1826 + 27 x insulin – 169 x dose (*p = 0.014*). Of the measured plasma androgens, leptin was only found to correlate with testosterone, increasing as testosterone decreased (r_s = *-0.297, p = 0.016*); subgroup analysis showed that this was only the case for male patients (r_s = *-0.379, p = 0.036*).

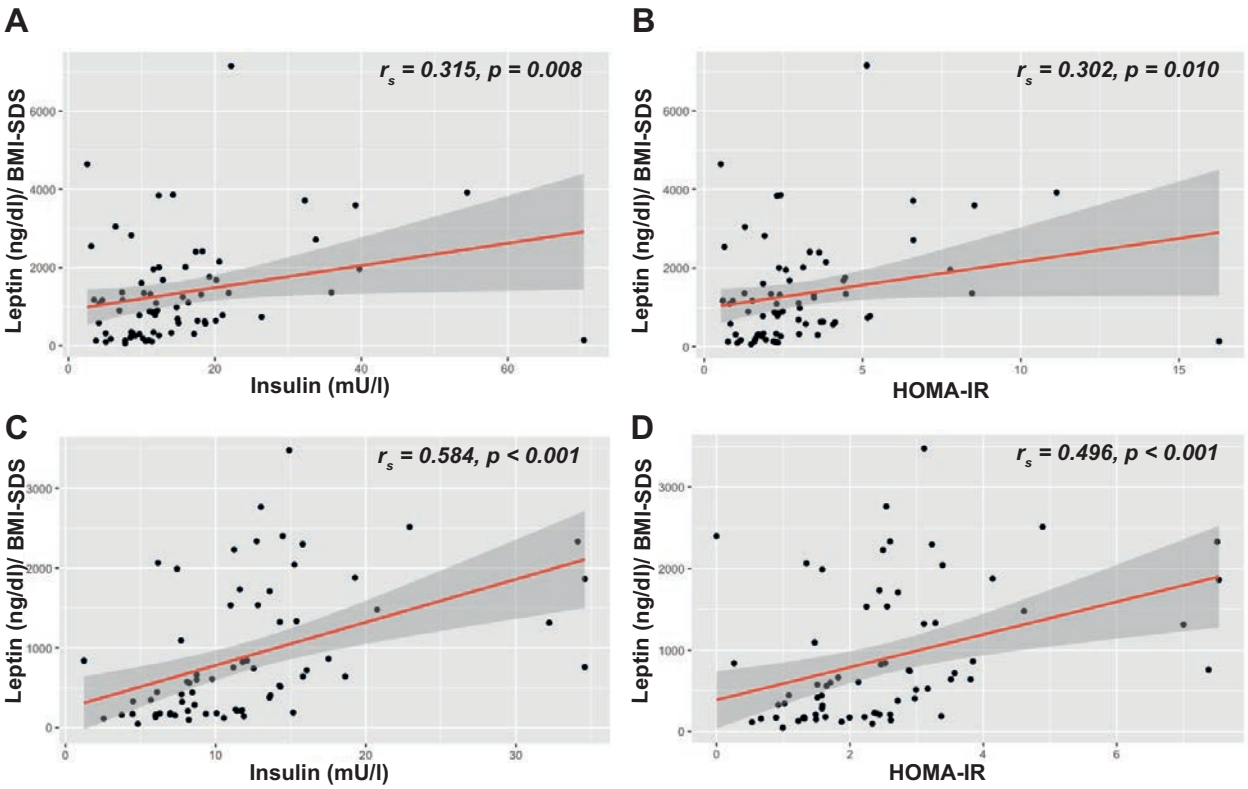


Figure 3.11. Leptin and insulin sensitivity
Correlations between leptin (corrected for BMI SDS) and insulin and HOMAIR in patients (**A, B**) and controls (**C, D**). The central red line represents the linear regression between the variables with its 95% confidence interval (grey ribbons).

For adiponectin there was no difference between patients and controls in the measured levels, with or without correction for BMI. In patients, no association was found between adiponectin and the GC dose, and the negative correlation with insulin and HOMA-IR was not significant. By contrast, in controls there was a significant negative correlation between adiponectin and insulin (r_s = *-0.332, p = 0.005*) and HOMA-IR (r_s = *-0.391, p < 0.001*).

A negative correlation was found between adiponectin and 17OHP (r_s = *-0.346, p = 0.004*), androstenedione (r_s = *-0.344, p = 0.004*), testosterone (r_s = *-0.331, p = 0.006*), 11-hydroxyandrostenedione (r_s = *-0.326, p = 0.006*) and 11-ketotestosterone (r_s = *-0.285, p*

= *0.019*) (**Figure 3.12**); no such relationship was found in controls, except for a weak negative correlation with testosterone (r_s = *-0.257, p = 0.047*).

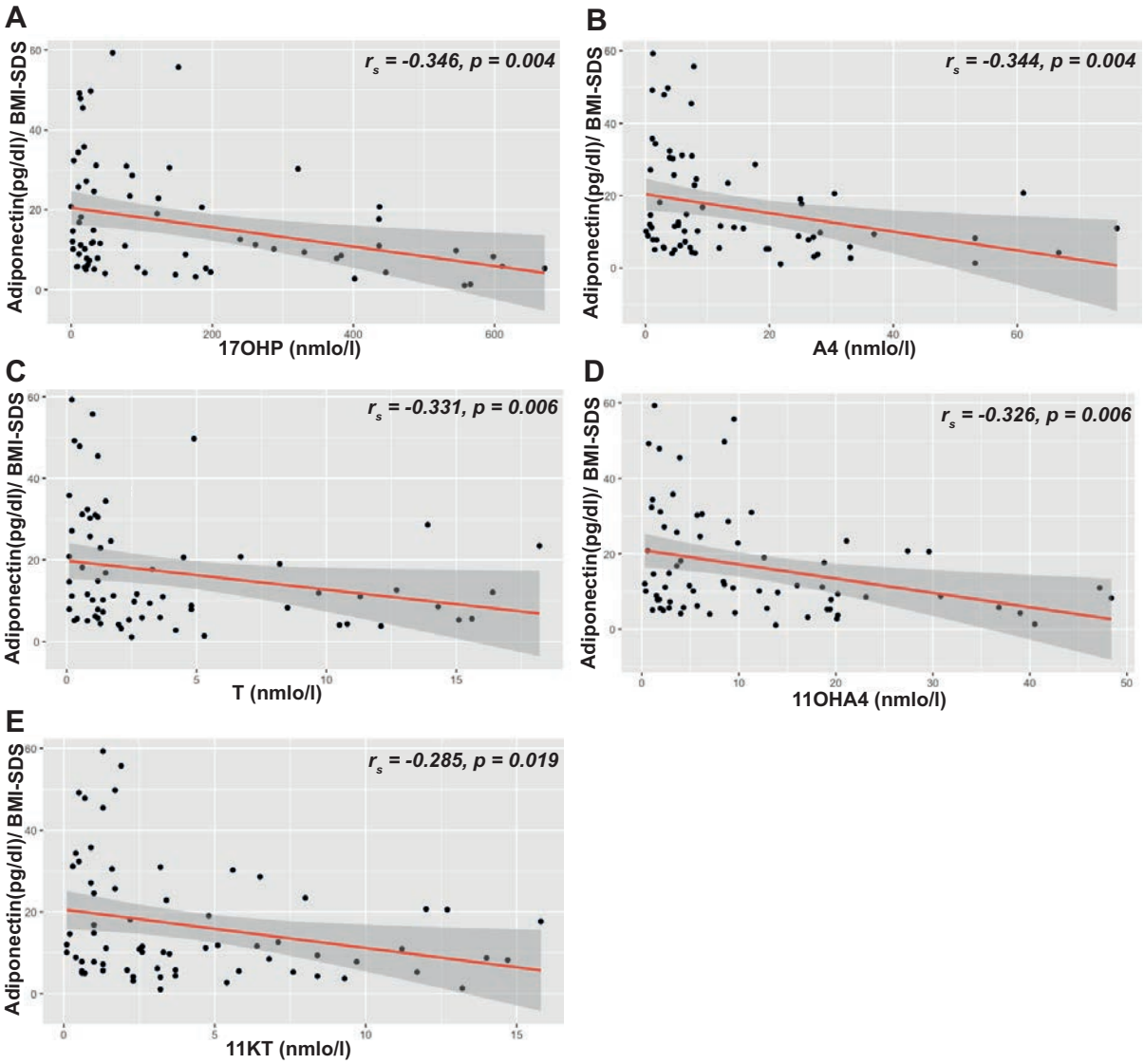


Figure 3.12. Adiponectin vs plasma androgens
Correlations between adiponectin (corrected for BMI SDS) and plasma androgens in patients: 17-hydroxyprogesterone (**A),** androstenedione **(B**), testosterone (**C)**, 11-hydroxyandrostenedione **(D**), and 11-ketotestosterone (**E**). The central red line represents the linear regression between the variables with its 95% confidence interval (grey ribbons). (17OHP: 17-hydroxyprogesterone; A4: androstenedione; T: testosterone; 11OHA4: 11-hydroxyandrostenedione; 11KT: 11-ketotestosterone).

3.2 Discussion

(Modified from published paper (Bacila, et al., 2022))

While in adults with CAH the raised prevalence of comorbidities including metabolic and cardiovascular disease was clearly demonstrated (Merke and Auchus, 2020), the current study indicates that in children growth and weight gain problems are by far the most frequent health problems. However, some of the results indicate that metabolic comorbidities have their onset during the childhood and teenage years, in keeping with more recently published evidence (Torky, et al., 2021).

Children with CAH were found to have altered growth patterns in comparison to healthy individuals, consisting of accelerated growth before puberty and reduced height in postpubertal patients. This was confirmed by the age-dependent fluctuation observed when patient height was analysed against the respective mid-parental height, showing younger patients to be above the expected height SDS, while data from patients reaching the end of puberty indicated reduced final height. A similar relationship was noted when patients' height SDS were compared to those of controls, as in the younger group of participants (8-12 years) patients were taller than controls with a more advanced bone age and higher Tanner scores. This is a consequence of the higher levels of adrenal androgens commonly found in patients with CAH causing an early growth spurt. Height-SDS of patients aged 12-18 years was however significantly lower compared to controls in this age group. The results resonate with the published UK evidence from adults with CAH who had reduced height (Arlt, et al., 2010). Importantly, in a significant number of patients the recommended dose range of 10-15mg/m^2/day HC (Speiser, et al., 2018) was exceeded. The reduced height shown in older patients with CAH may partly relate to the growth-suppressive effect of GC. Although in the present study the relative GC doses did not correlate with height SDS, previous evidence from a longitudinal study showed that

using doses over 17 mg HC/m^2/day was associated with increased prevalence of short stature (Bonfig, et al., 2007).

Overweight and obesity were more prevalent in CAH patients than in the cohort. Of note, the prevalence of overweight and obesity in the control group was lower than the one reported by national data (namely, between 11% and 29% in year-6 children, with variations based on the socio-economic status) (NHSDigital, 2017). However, these results are in keeping with evidence from adults with CAH in the UK (Arlt, et al., 2010) and with the results found in paediatric patients reported by studies conducted in other countries (Völkl, et al., 2006; Ariyawatkul, et al., 2017). Previous evidence also demonstrated an increased percentage of fat mass in children with CAH, which correlated with BMI and HOMA-IR (Mooij, et al., 2010). Cushingoid features were only found in one patient from the cohort, significantly less compared to adult data reporting them in 63% of females and 38% of males (Arlt, et al., 2010). This might be indicative of increased awareness of GC toxicity, but it could also relate in some measure to a shorter duration of the exposure to synthetic steroids in children.

The finding of raised blood pressure in five of the patients was based on single time-point assessments, and thus, was not sufficient for a diagnosis of hypertension (Rao, 2016). This number of patients was too modest to identify any trend relating to medication or weight. The evidence on the effects of CAH on blood pressure is sparse and yielded contradicting results, ranging from indicating increased prevalence of raised blood pressure in children (Ariyawatkul, et al., 2017) to reports of hypotension in normal-weight patients with this condition (Völkl, et al., 2006).

The number of patients with altered lipid profiles is modest compared to findings from adults (Arlt, et al., 2010). However, the fact that CAH patients in the cohort had significantly raised prevalence of overweight and obesity would indicate that they have an increased risk of developing metabolic syndrome. Recent evidence of increased prevalence of the metabolic syndrome before puberty in patients with CAH (Torky, et al., 2021) further supports this theory. In the present study, analysing absolute values of biochemical markers, there appeared to be an association between supressed plasma 17OHP and higher levels of LDL and total cholesterol. This would suggest indirectly that higher doses of GC may lead to deranged lipid profile, while a direct association was not confirmed as cholesterol levels were comparable between patients with high and low doses of GC. Potential explanations for this discrepancy may be the small group of patients with supressed 17OHP (n=10), as well as potential lack of compliance to the GC replacement therapy. Insulin resistance, based on the HOMA-IR calculation, was present in as high as 53.8% of patients, much higher than findings reported by previous evidence in adults with CAH of 28-36% (Arlt, et al., 2010). This may relate to the much lower HOMA-IR threshold of 1.68 recommended in children (Shashaj, et al., 2016), compared to the 2.5 used in adults (Arlt, et al., 2010). The published evidence on the prevalence of insulin resistance in children with CAH produced variable results (Charmandari, et al., 2002; Völkl, et al., 2009; Ariyawatkul, et al., 2017) and it is difficult to draw clear conclusions. This study could not confirm an association between CAH and insulin resistance during childhood, and it is likely that the high HOMA-IR found in participants is linked to the increasing frequency of prevalence of the metabolic syndrome in children in general (Weihe and Weihrauch-Blüher., 2019).

The GC replacement therapy consisted mostly of HC administered according to a circadian three or four daily doses regime. This is in keeping with the international recommendations (Speiser, et al., 2018) based on the fast clearance of HC (Charmandari, et al., 2001), the aim being to mimic the physiological cortisol release from the adrenals. Although a study comparing the efficacy of circadian and reverse circadian GC regimes produced inconclusive results (German, et al., 2008), it was shown that a raised plasma cortisol in the evening was associated with higher risk of metabolic complications in adrenal insufficiency (Plat, et al., 1999). The GC dose was in adherence to the recommended range of 10 - 15mg/m^2/day in less than 50% of the patients, 22.4% using lower and 33.6% higher doses. Comparing age subgroups, patients aged 12-18 years received higher doses, however, the androstenedione and 17OHP concentrations were similar between age groups. This could be due to either an increased requirement for HC, or reduced compliance to the treatment with the increasing age. Importantly, the study used no formal strategy for assessing compliance to treatment, an aspect that needs to be taken into consideration when interpreting the results reporting the relationship between GC replacement and other variables, such as plasma steroids and anthropometric data.

The limitations of the conventional hormonal biomarkers used in CAH are well known (Dauber, et al., 2010) and there is a need for monitoring tools that are superior to those which are currently available. The extended steroid profiles in the CAH-UK patients confirm this statement, showing variability in the concentrations of 17OHP and androstenedione, the most common biomarkers of control used in clinical practice, with no relationship to the HC dose identified. The findings also show the gonadal testosterone suppression by adrenal androgens in pubertal males, which was previously described in

adults with CAH (Engels, et al., 2018), confirming that testosterone is only relevant as a biomarker for CAH in girls and pre-pubertal boys. Of note, the measurements of plasma steroids were conducted using samples collected between 8:00 and 10:00 at variable times after the morning GC dose, an aspect likely to reduce their value as monitoring markers. The results also showed raised concentrations of 11-hydroxyandrostenedione and 11-ketotestosterone in children and young people with CAH. This is in keeping with previous evidence from adults with CAH (Turcu, et al., 2016; Kamrath, et al., 2018) and also from women with polycystic ovary syndrome (PCOS) (O'Reilly MW, 2017), highlighting the practical potential of 11-oxygenated androgens as biomarkers of disease control. Importantly, although the plasma concentrations were higher in patients compared to healthy controls for 17OHP, androstenedione, 11-hydroxyandrostenedione and 11-ketotestosterone, there was still consistent overlap in their values between the two groups, especially in the case of the two 11-oxygenated androgens. A closer look showed that the overlap corresponded to patients with normal or low 17OHP. This would indicate that 11-oxygenated androgens may be useful in identifying CAH patients who are under-treated, also highlighting the need for using hormonal profiles when assessing disease control. There was no correlation found between plasma cortisol and androgens in CAH patients, although this was present in healthy controls. This aspect is in keeping with the results of an earlier study where it was attributed to the high variability in HC pharmacokinetics and HPA response between patients (Sarafoglou, et al., 2015), however, consideration must also be given to variable compliance in taking the GC medication, especially in view of the high number of patients with low cortisol concentrations.

Overall, these results highlight abnormal patterns of growth and increased weight in patients, which can be attributed to GC treatment and hyperandrogenism. The findings also suggest that the onset of the metabolic syndrome and cardio-vascular pathology in CAH may occur during childhood. It has been previously hypothesised that metabolic changes in CAH may be partly caused by an altered leptin axis (Charmandari, et al., 2002; Saygili, et al., 2005; Völkl, et al., 2009). High adiponectin levels have also been reported in children with CAH (Volkl, et al., 2009) making published data on adipokines and insulin resistance in CAH all the more confusing, as adiponectin counteracts insulin resistance.

The **adipokine analysis** provided further insight regarding the developing risk for metabolic disease in CAH patients during childhood. Leptin is an established regulator of energy homeostasis, neuroendocrine and metabolic functions (Obradovic, al., 2021). In patients with CAH leptin was shown to correlate positively with fat mass, BMI and cardio-vascular risk in both adults (Borges, et al., 2021) and children (Charmandari, et al., 2002; Zurita-Cruz, et al., 2018). The raised levels of leptin and insulin found in patients with CAH were attributed to the GC replacement therapy and to the long-term impairment of the adrenal medulla observed in these patients, which may lead to a reduction in the beta-adrenergic suppressive effect on leptin synthesis (Charmandari, et al., 2002). The correlation between leptin and BMI was also found in the participants of the CAH-UK study, both patients and controls. However, leptin was only significantly raised in male patients compared to controls in this study. It is important to note that there was a significant difference in leptin between males and females in both patients and controls. This difference was reported before by evidence, females being known to have higher leptin (Saad, et al., 1997). Possible explanations offered for this difference included

variations in hypothalamic regulation, leptin sensitivity, sex hormones profiles and fat disposition. Regarding the latter, the theory is supported by evidence showing that subcutaneous fat secreted more leptin compared to visceral fat (Russell, et al., 1998; Van Harmelen, et al., 1998). In support of this argument, in the study male patients had higher hip circumference SDS compared to controls and were also found to have higher leptin levels. This may also account for the fact that the difference in the leptin level between males and females was less marked in the patient group compared controls. This sexual dysmorphism in leptin regulation may be the reason why the differences were not consistent between patients and controls in the CAH-UK study. Another important aspect to be considered relates to possible leptin resistance that was described before in patients with CAH. A study involving children and adolescents with CAH found no difference in serum leptin compared to healthy controls (Völkl, et al., 2006), however, in a later study, the same researchers reported reduced leptin receptors in patients, concluding a dysregulation of the leptin axis with leptin resistance that is related to the increased risk of overweight in CAH (Völkl, et al., 2009). Similar to previously published evidence, in the CAH-UK study leptin correlated positively with insulin and HOMA-IR in both patients and controls, suggesting that it could potentially be used as an indicator of increased risk of metabolic disease. Moreover, leptin decreased significantly with the increase of the first GC dose of the day. These findings are in keeping with previous evidence on the impact of insulin and GC on leptin synthesis (Russell, et al., 1998). The negative relationship between leptin and testosterone was described by a previous study that associated raised leptin with good androgen control in CAH (Oliveira, et al., 2013). However, in the CAH-UK study this association was only found in male patients and there was no relationship found between leptin and the other, more reliable androgen markers of control.

Another adipocyte derived hormone, adiponectin, is known to mediate the metabolic function of many organs and tissues (Wang and Scherer, 2016) and have insulin sensitising actions (Yadav, et al., 2013). Reduced adiponectin levels are associated with increased fat mass and insulin resistance (Yadav, et al., 2013). In patients with CAH adiponectin was reported to be increased in both children (Völkl, et al., 2009) and adults (Mooij, et al., 2011). By contrast, in the CAH-UK study patients and controls had comparable adiponectin levels, despite the fact that patients had a higher prevalence of obesity and overweight. Adiponectin correlated negatively with BMI in CAH patients, an aspect that had been reported before by the literature (Völkl, et al., 2009). However, no such relationship was found in controls. An association was reported in patients with CAH between high serum testosterone, metabolic disorder indexes and low adiponectin (Zhang, et al., 2010). Moreover, in patients with non-classic CAH on long term GC replacement adiponectin was reported to be low while the waist to hip ratio was high and HDL levels were low (Delai, et al., 2022). In the present study the number of CAH patients with abnormal lipid profiles was too limited to allow exploring the association with adiponectin, however, adiponectin decreased significantly with increasing waist and hip SDS, which supports its potential use as a marker of metabolic risk. It is important to highlight the relationship found between adiponectin and the extended androgen profile in patients with CAH. The negative correlations, which, although weak, were consistent for all measured androgens, would suggest that high adiponectin could be interpreted as a marker of good androgen control, as well as an indicator of low-fat mass. Similar relationships were previously reported with regards to adiponectin and DHEAS and testosterone in CAH (Völkl, et al., 2009). Of note, there is evidence demonstrating the direct effect of androgens in the regulation of adiponectin secretion from adipose tissue (O'Reilly et al. 2014), however, the study of their relationship in CAH is relatively limited.

The novelty of the current study lies in establishing an association between adiponectin, which is a marker of metabolic risk and an expanded steroid profile of androgen control, including two 11-oxygenated androgens, which currently are gaining popularity as potential superior markers of disease control in CAH (Turcu, et al., 2017; Bacila, et al., 2019). As these associations were not present in the control group, it appears that the relationship may be specific to CAH or hyperandrogenic conditions, a theory that further highlights the need for further research focused on defining markers of metabolic risk in these conditions.

4. Plasma metabolomes and their association with steroid replacement in patients with CAH

The metabolomic analysis we chose to employ uses high-throughput biochemical assay technology to identify changes in the small molecular weight compounds involved in metabolic processes. It was used in this study to further explore the impact of hypocortisolism and synthetic GC replacement on metabolic pathways. The analysis focused on the metabolites that varied among patients based on steroid replacement doses. There was a marked predominance of classes of compounds involved in lipid metabolism, including glycerophospholipids, lysophospholipids, sphingolipids, fatty acids and triglycerides. These metabolites were influenced not only by dose but also by the timing of GC administration and some of them also correlated consistently with other important variables, including the BMI and leptin.

4.1 Results

4.1.1 Associations between the glucocorticoid treatment and plasma metabolites in CAH patients

The analysis was focused on the metabolites that were different between treatment subgroups: patients receiving relative hydrocortisone doses of < 10 mg/m^2/day, 10-15 mg/m^2/day and > 15 mg/m^2/day. A total 237 compounds were found to be significantly different between at least two of these dose subgroups, following nonparametric analysis of variance by Kruskal-Wallis H and Mann-Whitney U tests. Of these, the majority belonged to lipid metabolism, in particular glycerophospholipids, lysophospholipids, sphingolipids, triglycerides and fatty acid metabolism (**Table 4.1**).

Table 4.1. Groups of metabolites

Class of metabolites	Number of compounds
Alpha hydroxy acids - lactic acid	1
Amino acids	4
Bile acids	1
Cholesterol and derivates	3
Components of the respiratory chain	2
Fatty acid esters	2
Fatty acid metabolism	12
Gluten digestion	2
Glycerophospholipids	78
Indolyl carboxylic acids	1
Leukotrienes and metabolites	2
Lipoxygenase pathway - hydroxyeicosatetraenoic acid, lipoxin C4	2
Lysophospholipids	41
N-acylamide	1
Product of protein catabolism	5
Propanoate metabolism	1
Pyrimidine metabolism	2
Pyrrolidines	1
Quinolone derivates	2
Retinoids	2
Sphingolipids	36
Steroid glucuronide conjugates	1
Steroid glycoside	1
Sterol	1

Triglycerides	24
Tryptophan metabolism	2
Urea cycle - Ornithine	1
Vitamin D metabolism	4
Xanthophylls	1

Summary of the metabolites that were different between subgroups of patients based on the relative GC dose.

Glycerophospholipids were the largest group of metabolites that made the distinction among different dose subgroups. Among them, phosphatidylcholine metabolites were the most common form (32%), however, other types of glycerophospholipids were also well represented including phosphatidylethanolamine (18%), phosphatidylserine (19%), phosphatidylinositol (19%), phosphatidic acid (6%), phosphatidylglycerol (10%). Analysing the compounds that showed the highest statistical difference in the distinction between treatment subgroups, it was found that in many cases the metabolites decreased in patients with higher doses (**Figure 4.1**). These results were based on the group comparison using the Kruskal-Wallis H and Mann-Whitney U tests. The most significant difference was found for a phosphatidylcholine, PI (20:4(5Z,8Z,11Z,14Z)/22:6 (4Z,7Z,10Z,13Z,16Z,19Z), Kruskal-Wallis H $p = 0.007$), the levels decreased significantly in patients with higher doses. Further regression analysis only identified a linear relationship between the GC dose and the measured metabolites in a limited number of cases. By contrast, for the majority of the glycerophospholipid compounds the variations in dose induced a “U shape” or “inverted U shape” effect, with the group of patients that received a dose within the recommended range of 10-15 mg/m^2/day hydrocortisone having lower or, respectively, higher metabolite levels compared to both groups outside the dose range (**Figure 4.1**).

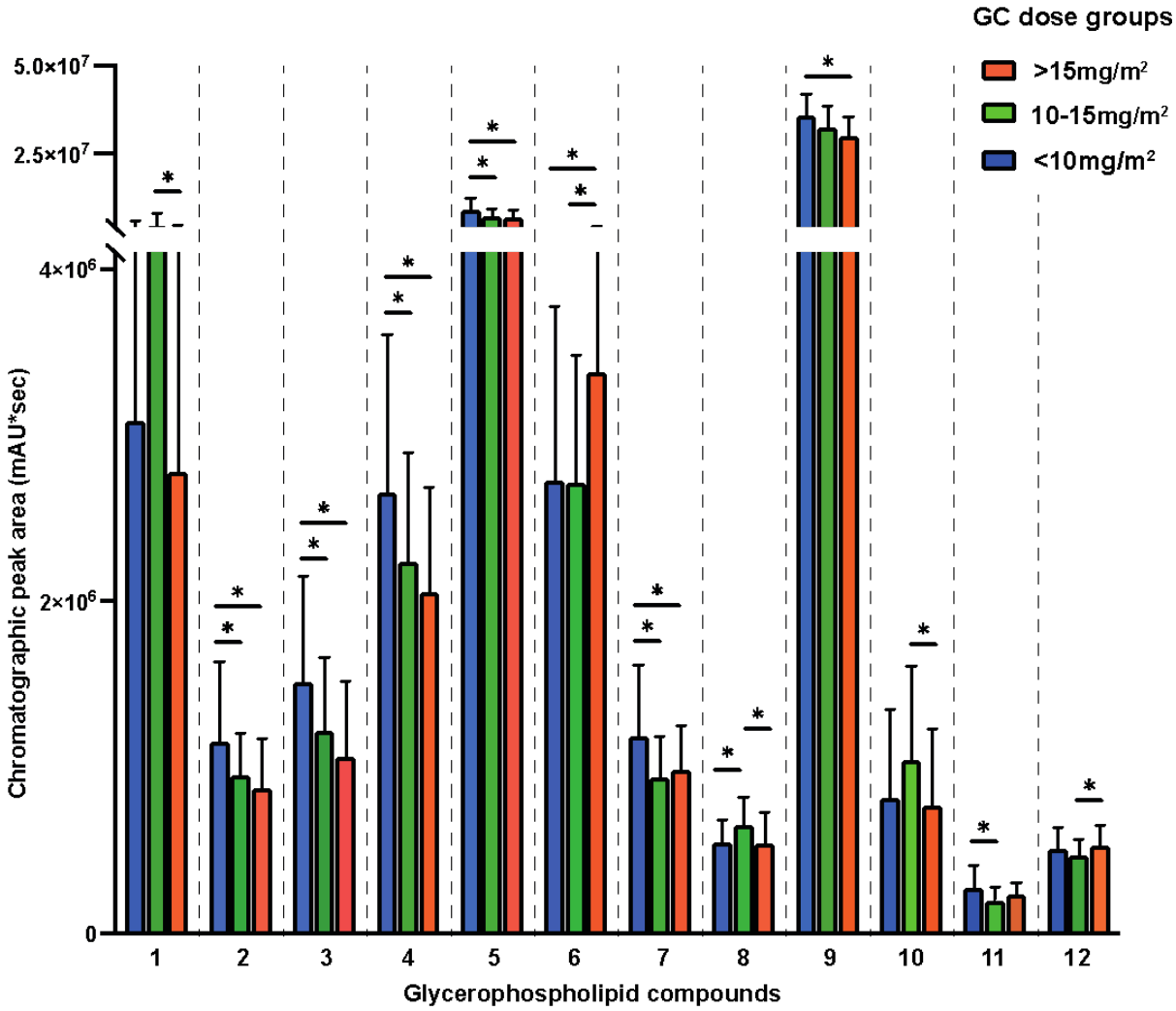


Figure 4.1. Glycerophospholipids
Comparison of glycerophospholipid compounds between patient subgroups based on the relative daily GC dose: under 10 mg/m^2/day (blue, n=20), 10-15 mg/m^2/day (green, n=42) and over 15 mg/m^2/day (orange, n=36). This is a selection of the glycerophospholipid metabolites which varied most significantly between dose subgroups: PC(0:0/18:1(6Z)) (1); PI(13:0/22:1 (11Z)) (2); PI(18:0/22:6 (4Z,7Z,10Z,13Z,16Z,19Z)) (3); PI(20:4(5Z,8Z,11Z,14Z) /22:6(4Z,7Z,10Z,13Z,16Z,19Z)) (4); PS(O-16:0/16:0)||PS(O-18:0/14:0) (5); PS(O-18:0/20:3(8Z,11Z,14Z)) (6); PG(20:1(11Z)/22:4(7Z,10Z,13Z,16Z)) (7); PG(O-18:0/22:6(4Z,7Z,10Z,13Z,16Z,19Z)) (8); CL(74:0) (9); PIM2(17:0/16:0) (10); PIM2(18:0/18:0) (11); PIM2(18:2(9Z,12Z)/18:1(9Z)) (12). (Overall comparison among the three groups was initially carried out by Kruskal-Wallis H test, followed by Mann-Whitney U test for significant differences between each two groups, * indicating statistical significance for $p<0.05$)

The majority of the **lysophospholipids** identified were lysophosphatidylcholine species (46%), with a smaller number belonging to the other forms: lysophosphatidylethanolamine (12%), lysophosphatidic acid (5%),

lysophosphatidylserine (5%), lysophosphatidylinositol (2%), and lysophosphatidylglycerol (10%). All the compounds pertaining to the lysophospholipids class presented the same type of variation among the three dose subgroups, “inverted U shape”, where the patients receiving doses within the recommended range had higher levels compared to patients on lower or higher doses (**Figure 4.2A, B**).

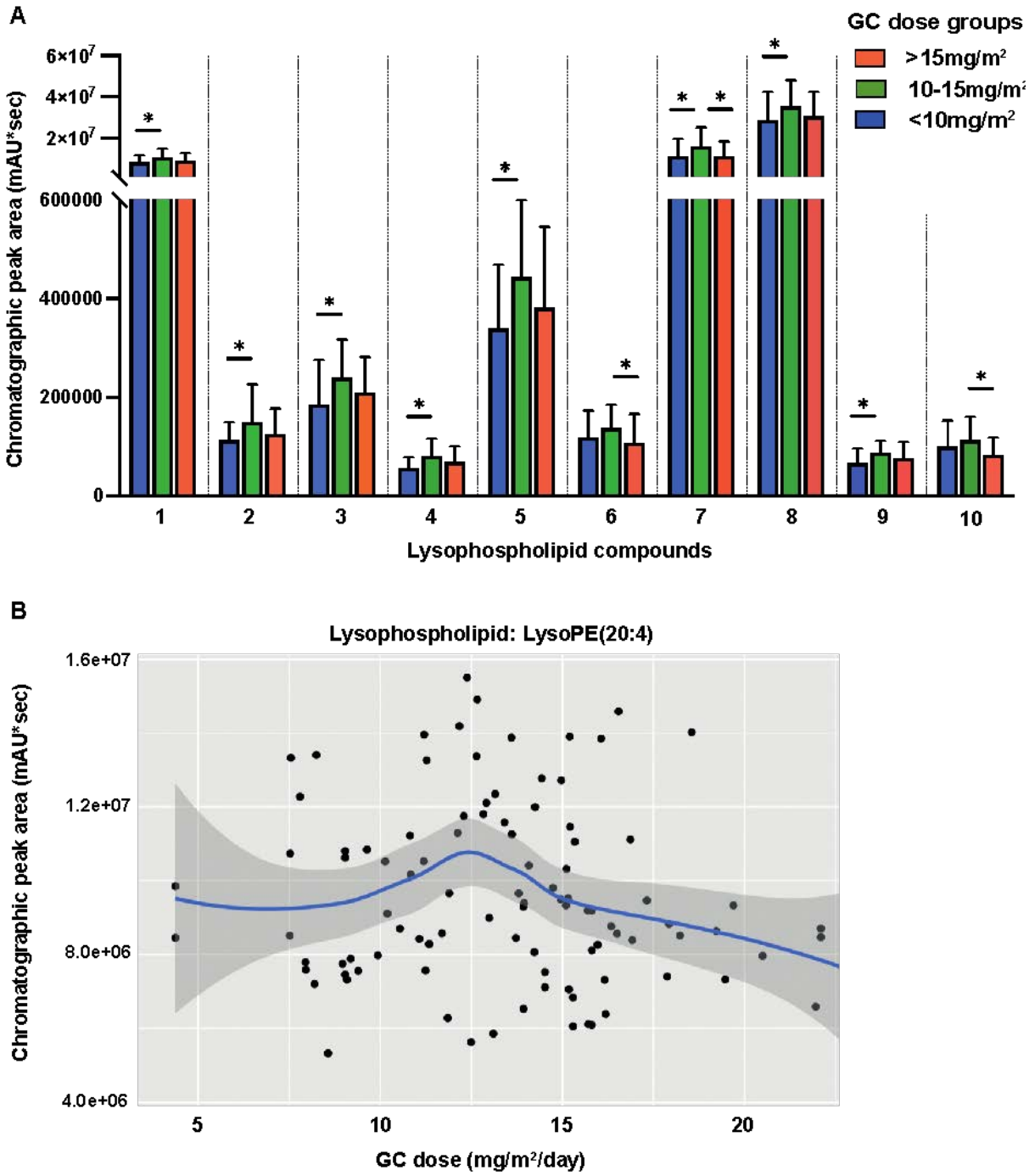


Figure 4.2. Lysophospholipid

A. Comparison of lysophospholipid compounds between patient subgroups based on the relative daily GC dose: under 10 mg/m^2/day (blue, n=20), 10-15 mg/m^2/day (green, n=42) and over 15 mg/m^2/day (orange, n=36). This is a selection of the metabolites that varied most significantly between dose subgroups: LysoPC(15:1) (1); LysoPC(17:1) (2); LysoPC(20:4) (3); LysoPE(20:5) (4); LysoPG(22:0) (5); LysoPA(P-20:0) (6); LysoPC(18:1) (7); LysoPC(22:5) (8); LysoPE(20:4) (9); LysoPC(P-19:1) (10). **B**. Scatterplot of the relationship between independent lysophospholipid LysoPE(20:4) and the relative GC dose, demonstrating the "inverted U shape", dose-dependent effect. The blue line and shaded area represent the smooth line (with 95% CI) of the relationship. (Overall comparison among the three groups was initially carried out by Kruskal-Wallis H test, followed by Mann-Whitney U test for significant differences between each two groups, * indicating statistical significance for $p<0.05$)

A very similar relationship between subgroups was noted for most of the metabolites belonging to sphingolipids, fatty acid metabolism, and triglycerides (**Figures 4.3 A-C**). For fatty acid metabolism, the strongest relationship with the GC dose was found for N-nervonoyl taurine (r_s = *-0.301, p = 0.002*), and for sphingolipids, one of the ceramides (LacCer(d16:0/22:0)) (r_s = *-0.305, p = 0.002*), both showing a consistent decrease with the increasing relative hydrocortisone dose in all three subgroups (**Figure 4.3D, E**).

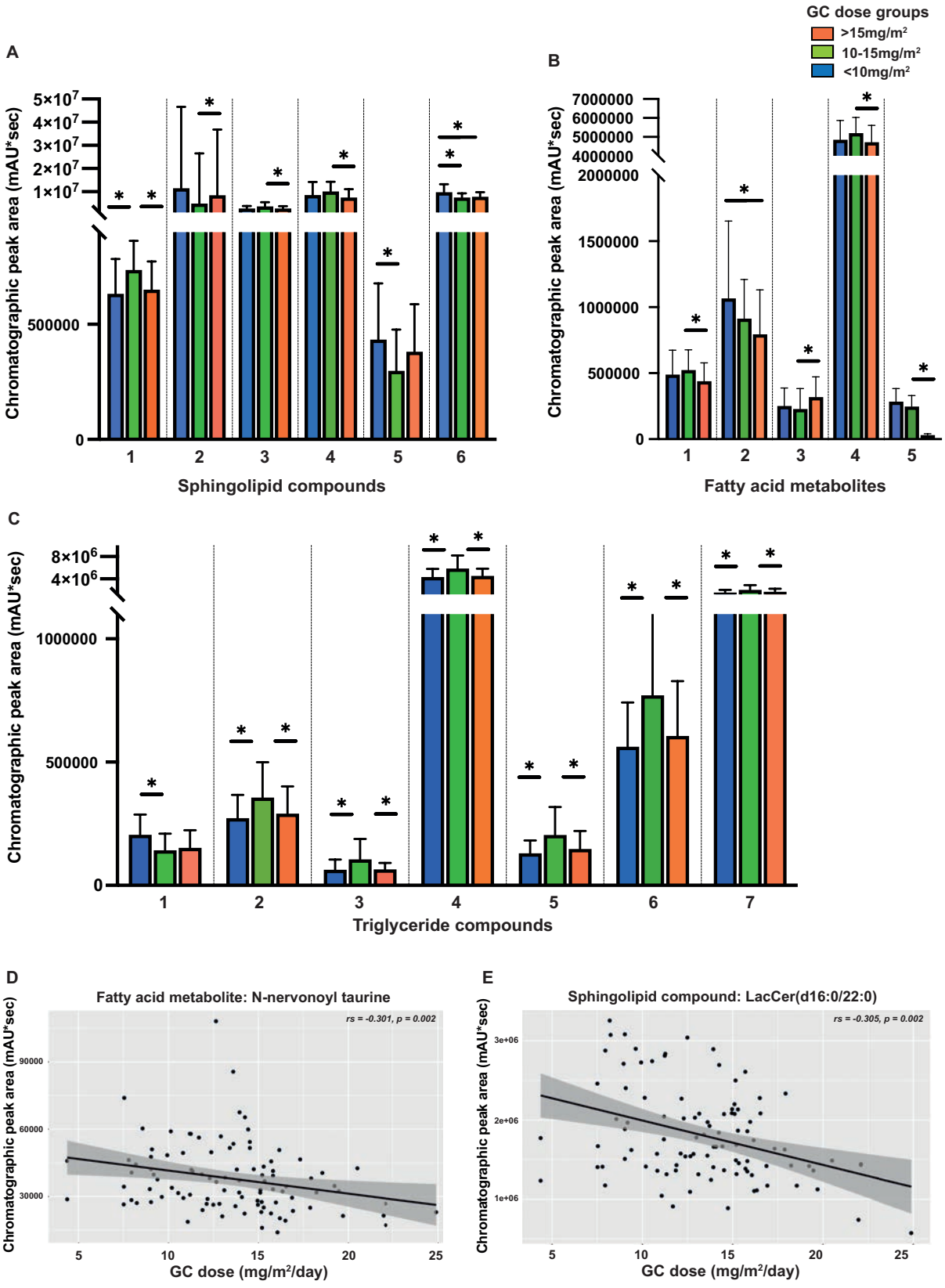


Figure 4.3. Sphingolipids, fatty acids, and triglycerides
Comparison of sphingolipids (**A**), fatty acid (**B**) and triglycerides (**C**) compounds between patient subgroups based on the relative daily GC dose: under 10 mg/m^2/day (blue, n=20), 10-15 mg/m^2/day (green, n=42) and over 15 mg/m^2/day (orange, n=36). This is a selection

of the metabolites that varied most significantly between dose subgroups: **sphingolipids** (**A**): GM3(d18:1/22:0) (1); omega-linoleoyloxy-Cer(t18:0/33:0) (2); omega-linoleoyloxy-Cer (t18:1(6OH)/26:0) (3); GM4(d18:1/16:0) (4); Cer(d14:1(4E)/ 26:0(2OH)) (5); Cer(d16:0/24:0) (6); **fatty acid metabolism** (**B**): Tetradecadiencarnitine (1); Cervonyl carnitine (2); Octadecadienoic acid (3); LysoPE(20:5) (4); Dodecaprenyl diphosphate (5); Trans-retinyl oleate (6); **triglyceride compounds** (**C**): TG(14:1(9Z)/15:0/18:4(6Z,9Z,12Z,15Z)) (1); TG(18:0/22:0/ 22:3(10Z,13Z,16Z)) (2); TG(18:0/22:0/22:5(7Z,10Z,13Z,16Z,19Z))[iso6] (3); TG(59:5) (4); TG(62:3) (5); TG(62:3) (6); TG(63:7) (7). **D, E**: Correlations of relative GC daily dose between specific fatty acid compound N-nervonoyl taurine (**D**) and sphingolipid compound, ceramide LacCer(d16:0/22:0) (**E**). The lines indicate the linear regression with 95% confidence interval (grey ribbons) between the variables. (Overall comparison among the three groups was initially carried out by Kruskal-Wallis H test, followed by Mann-Whitney U test for significant differences between each two groups, * indicating statistical significance for $p<0.05$)

Regarding the GC administration regime, the number of daily doses caused no significant variation of the selected metabolites. A small number of lipid metabolites were significantly different between patients who received a circadian vs a non-circadian GC administration regime. These included octadecadienoic acid ($p = 0.021$), several glycerophospholipids: PC(20:5(5Z,8Z,11Z,14Z,17Z)/22:6 (4Z,7Z,10Z,13Z,16Z,19Z)) ($p = 0.034$), PC(O-18:0/22:2(13Z,16Z)) ($p = 0.027$), PC(14:0/22:5 (4Z,7Z,10Z,13Z,16Z)) ($p = 0.007$), PI-Cer(d20:1/14:0) (p = 0.011), PI(O-16:0/18:2(9Z,12Z)) ($p = 0.036$), PS(12:0/21:0) ($p = 0.006$), PG(17:0/22:4 (7Z,10Z,13Z,16Z)) ($p = 0.026$), LacCer(d16:0/22:0) ($p = 0.001$), SM(d18:0/20:2 (11Z,14Z)) ($p = 0.023$), and SM(d18:1/16:1(9Z)) ($p = 0.033$). The majority of these metabolites were significantly lower in patients receiving a circadian GC regime (**Figure 4.4**).

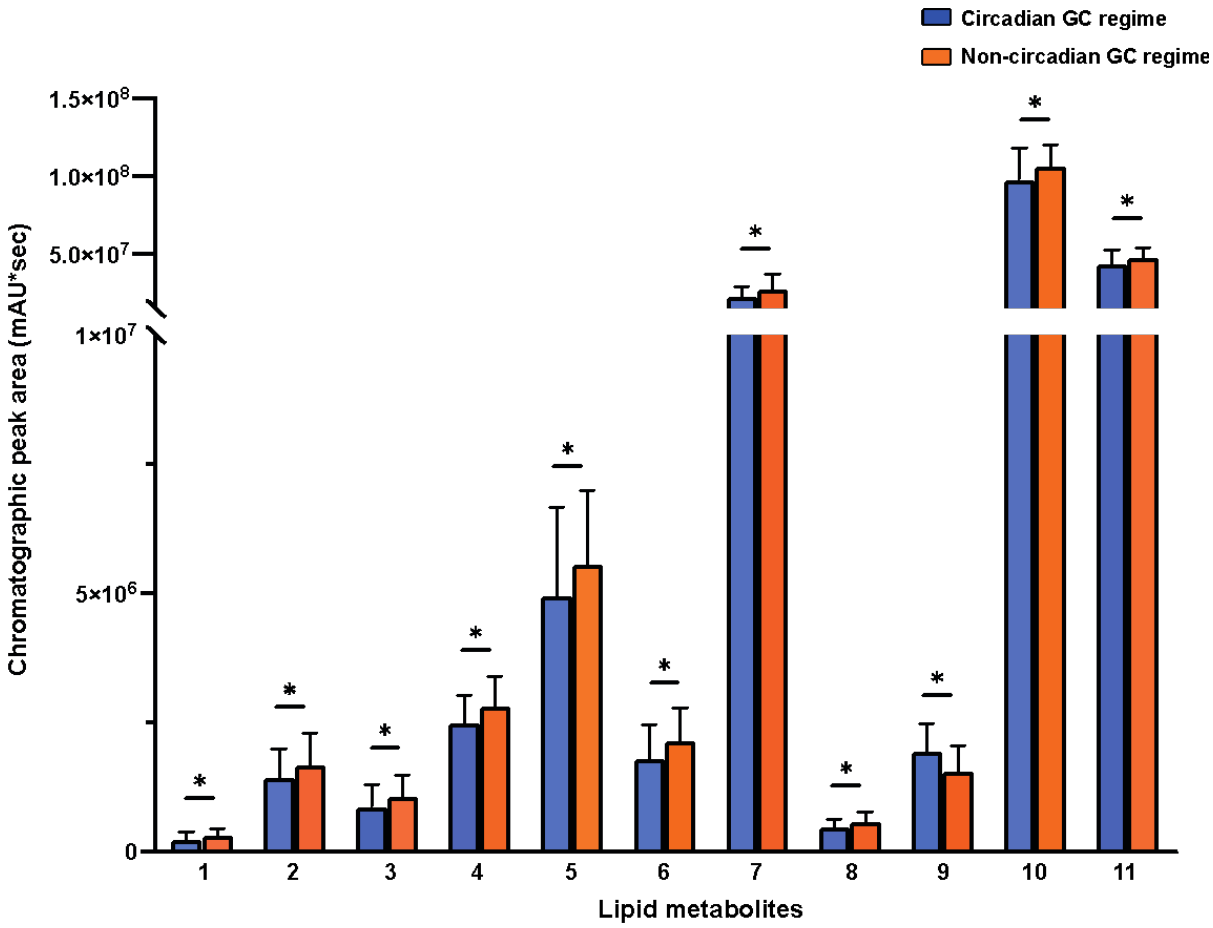


Figure 4.4. GC replacement regime and lipid metabolites
Lipid metabolites that were significantly different between patients receiving a circadian GC regime (blue, n=62) versus a non-circadian one (orange, n=36): PC(20:5(5Z,8Z,11Z,14Z,17Z)/22:6(4Z,7Z,10Z,13Z, 16Z,19Z)) (1), PC(O-18:0/22:2 (13Z,16Z)) (1), PC(14:0/22:5 (4Z,7Z,10Z,13Z,16Z)) (2), PI-Cer(d20:1/14:0) (3), PI(O-16:0/18:2(9Z,12Z)) (4), PS(12:0/21:0) (5), PG(17:0/22:4 (7Z,10Z,13Z,16Z)) (6), LacCer(d16:0/22:0) (7), SM(d18:0/20:2 (11Z,14Z)) (8), and SM(d18:1/16:1(9Z)) (9). (Comparison by Mann-Whitney U test, * indicating statistical significance for $p<0.05$)

The correlations of lipid metabolites with the time elapsed between the last GC dose and the samples collection were weak but there was a tendency for lysophospholipids to decrease and glycerophospholipids and sphingolipids to increase with the time elapsed from the GC dose, indicating that hydrocortisone acts to increase the lysophospholipids and reduce glycerophospholipids (**Figure 4.5**).

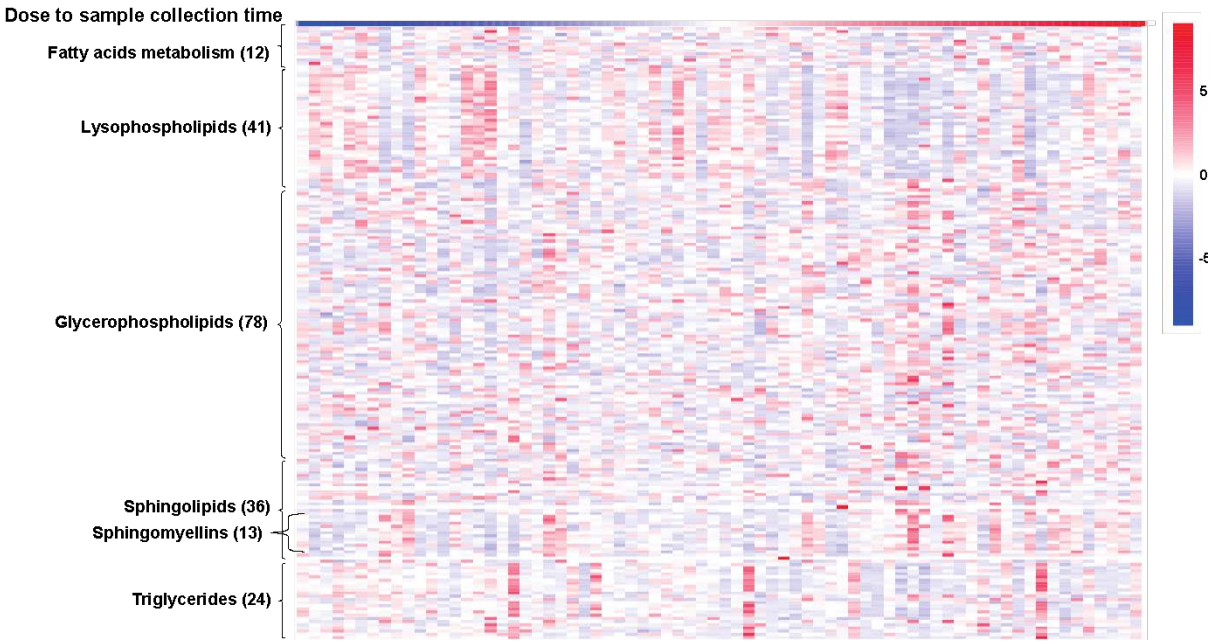


Figure 4.5. Lipid metabolites vs time from GC dose
Heatmap showing the lipid metabolites in relation to the time elapsed from the administration of the hydrocortisone dose and the sample collection. The first row shows the dose-to-blood collection time (minutes) in ascending order from left to right. The subsequent rows correspond to the measurements of individual metabolites; the classes to which the metabolites belong are marked on the left (n = number of compounds in each class). The columns correspond to individual patients.

Interestingly, for the lipid metabolites that were significantly associated with both the mineralocorticoid and glucocorticoid doses, the direction of the correlation found with the fludrocortisone dose for BSA, was opposite to that found for the relative GC dose in most cases (**Table 4.2**). Of note, there was negative correlation between the relative GC and fludrocortisone daily doses (GC (mg/m^2/day) = 15.7 – 0.02 FC (µg/m^2/day), $p = 0.016$)

Table 4.2. Metabolites correlations with GC and FC dose

Compound name	Spearman correlations metabolites vs dose			Metabolite class
	Treatment	r_s *value*	*p value*	
Galbeta-Cer(d18:1/24:1(15Z))	Hydrocortisone(mg/m^2/day)	0.211	0.042	**Sphingolipids**
	Fludrocortisone (µg/m^2/day)	-0.309	0.008	
SM(d18:1/12:0)	Hydrocortisone(mg/m^2/day)	-0.203	0.043	**Sphingomyelins**
	Fludrocortisone (µg/m^2/day)	0.264	0.020	
	Hydrocortisone(mg/m^2/day)	-0.257	0.039	

Hexenoyl carnitine	Fludrocortisone (μg/m^2/day)	-0.307	0.029	**Fatty acid metabolism**
LysoPC(11:0)	Hydrocortisone(mg/m^2/day)	-0.199	0.047	
	Fludrocortisone (μg/m^2/day)	0.233	0.040	
LysoPS(18:0)	Hydrocortisone(mg/m^2/day)	0.205	0.044	**Lysophospholipid**
	Fludrocortisone (μg/m^2/day)	-0.249	0.031	
Methylmalonylcarnitine	Hydrocortisone(mg/m^2/day)	-0.234	0.025	
	Fludrocortisone (μg/m^2/day)	0.300	0.011	
PC(12:0/18:2(9Z,12Z))	Hydrocortisone(mg/m^2/day)	-0.207	0.049	**Glycerophospholipid**
	Fludrocortisone (μg/m^2/day)	0.241	0.044	
PC(13:0/22:4(7Z,10Z,13Z,16Z))	Hydrocortisone(mg/m^2/day)	-0.212	0.034	
	Fludrocortisone (μg/m^2/day)	0.269	0.018	
PE-Cer(d14:1(4E)/23:0)	Hydrocortisone(mg/m^2/day)	0.216	0.035	
	Fludrocortisone (μg/m^2/day)	-0.250	0.032	
PI(12:0/20:0)	Hydrocortisone(mg/m^2/day)	-0.228	0.026	
	Fludrocortisone (μg/m^2/day)	0.299	0.009	
PI(13:0/22:1(11Z))	Hydrocortisone(mg/m^2/day)	-0.241	0.016	
	Fludrocortisone (μg/m^2/day)	0.229	0.046	
PS(O-16:0/20:3(8Z,11Z,14Z))	Hydrocortisone(mg/m^2/day)	-0.222	0.027	
	Fludrocortisone (μg/m^2/day)	0.307	0.007	
TG(12:0/15:0/22:4(7Z,10Z,13Z,16Z))[iso6]	Hydrocortisone(mg/m^2/day)	0.232	0.027	**Triglycerides**
	Fludrocortisone (μg/m^2/day)	-0.290	0.014	
TG(12:0/21:0/22:1(11Z))[iso6]	Hydrocortisone(mg/m^2/day)	-0.202	0.049	
	Fludrocortisone (μg/m^2/day)	0.277	0.016	
TG(12:0/22:0/22:5(7Z,10Z,13Z,16Z,19Z))[iso6]	Hydrocortisone(mg/m^2/day)	-0.214	0.036	
	Fludrocortisone (μg/m^2/day)	0.275	0.016	

Comparison of the correlations between fat metabolites vs. relative hydrocortisone dose and fat metabolites vs fludrocortisone dose.

Compared to lipid metabolites, a much smaller number of compounds pertaining to other metabolic pathways were identified. The effect of the relative GC dose on protein catabolism, amino acid metabolism and gluten digestion was statistically weaker and more variable, although there was a tendency for a positive relationship, the metabolites increasing with the daily dose (**Figure 4.6A, B**). The most consistent linear relationship was found for Alanylglycine ($r_s = -0.263$, $p = 0.009$). Two products of gluten digestion were also identified and found to increase with the dose (**Figure 4.6C**).

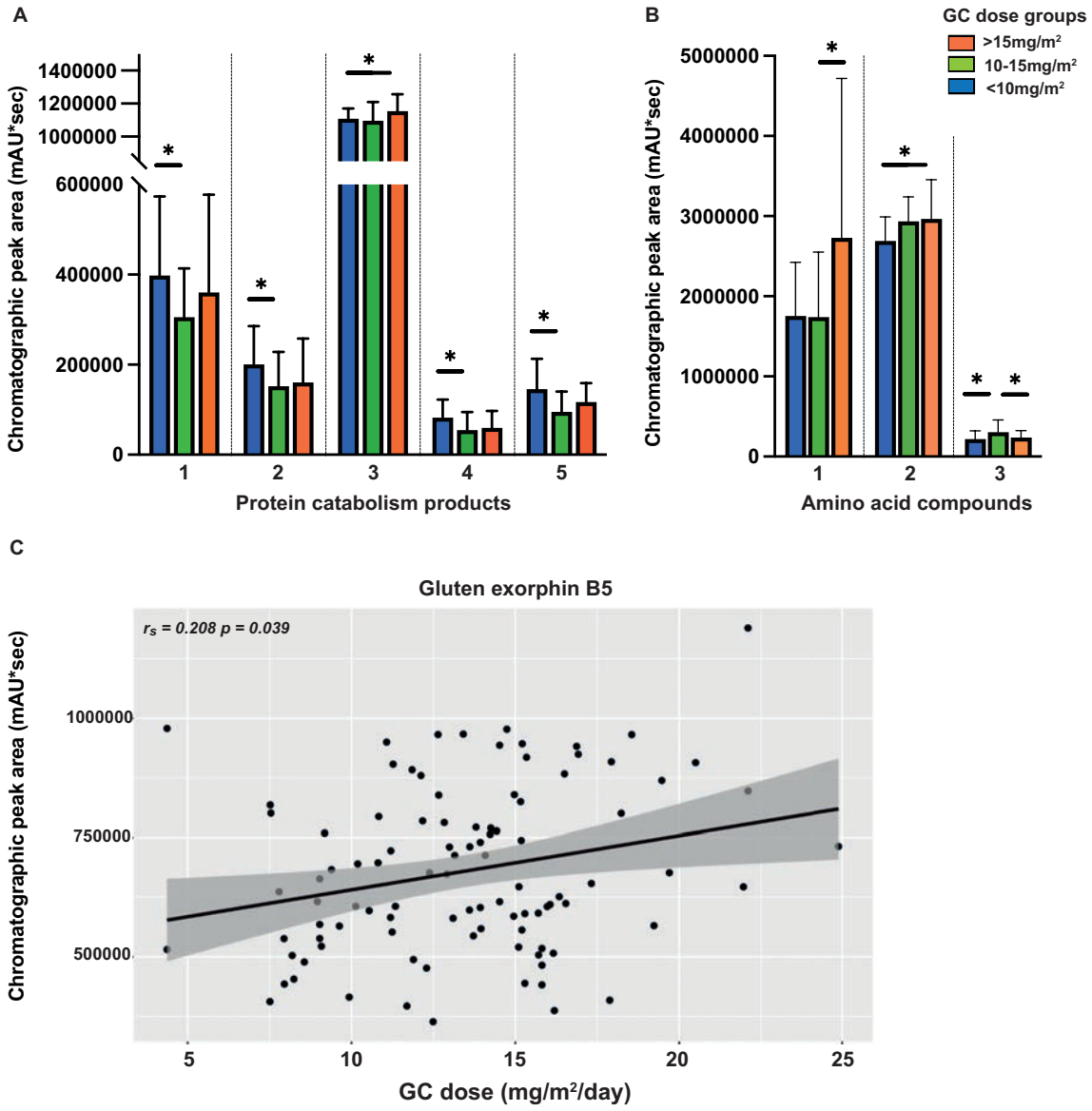


Figure 4.6. Protein and amino acid metabolites
Comparison of protein catabolism product (**A**) and amino acids (**B**) between patient subgroups based on the relative daily GC dose: under 10 mg/m^2/day (blue, n=20), 10-

15 mg/m^2/day (green, n=42) and over 15 mg/m^2/day (orange, n=36). **Protein catabolism compounds** (**A**): 2-oxoarginine (1); alanyl-gamma-glutamate (2); alanylglycine (3); asparaginyl-tryptophan (4); phenylalanylglycine (5); **amino acids**(**B**): hepteneoylglycine (1); imidazolone (2); serine (3). **C**: Correlations of relative GC daily dose between specific gluten digestion product, Gluten exorphin B5. The line indicates the linear regression with 95% confidence interval (grey ribbons) between the variables. (Overall comparison among the three groups was initially carried out by Kruskal-Wallis H test, followed by Mann-Whitney U test for significant differences between each two groups, * indicating statistical significance for $p<0.05$)

4.1.2 Associations between plasma metabolites and other variables

The majority of the identified metabolites were similar between sexes and had no significant relationship with age. The most consistent association with the BMI was found for lysophospholipids which decreased with the increasing weight (**Figure 4.7A**). Similarly, lysophospholipids had the most significant correlation with plasma leptin, with lower values being found in patients with high leptin. The strongest correlation was found for a lysophosphatidylcholine (LysoPC(P-17:0)) (r_s = *-0.504, p < 0.001* for BMI SDS and r_s = *-0.455, p < 0.001* for leptin) (**Figure 4.7B, C**).

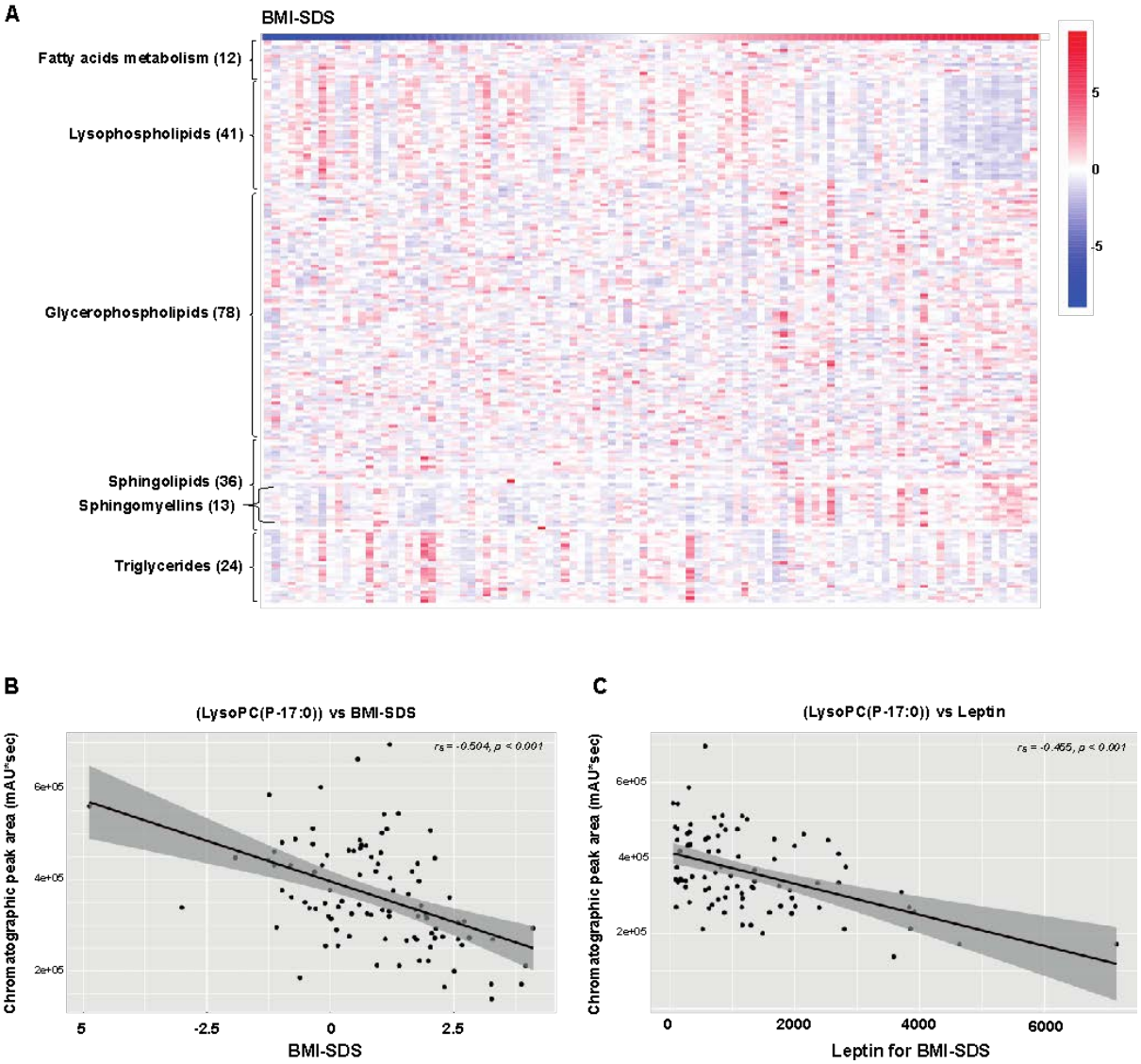


Figure 4.7. Lipid metabolites vs BMI
A. Heatmap showing the lipid metabolites in relation to the patients' weight gain, represented by the BMI SDS. The first row shows the BMI SDS in ascending order from left to right. The subsequent rows correspond to the measurements of individual metabolites; the classes to which the metabolites belong are marked on the left (n = number of compounds in each class). The columns correspond to individual patients. **B, C**: Correlations of the individual lysophospholipid: lysophosphatidylcholine (LysoPC (P-17:0)) with BMI-SDS (**B**) and leptin (**C**).

Comparing the lysophospholipid levels between patient subgroups based on the BMI-SDS (normal weight, overweigh and obese), all members of the group were similar between normal weight and overweight participants, but significantly lower in obese patients compared to the other groups (a selection of three compounds within the

lysophospholipid group is shown in **Figure 4.8).** For the other classes of metabolites, the relationship with the BMI and leptin was more variable, although there was a tendency for patients with high BMI SDS to have higher glycerophospholipids and sphingolipids (in particular sphingomyelin) species. There were no consistent associations between the lipid metabolites and insulin resistance based on HOMAIR.

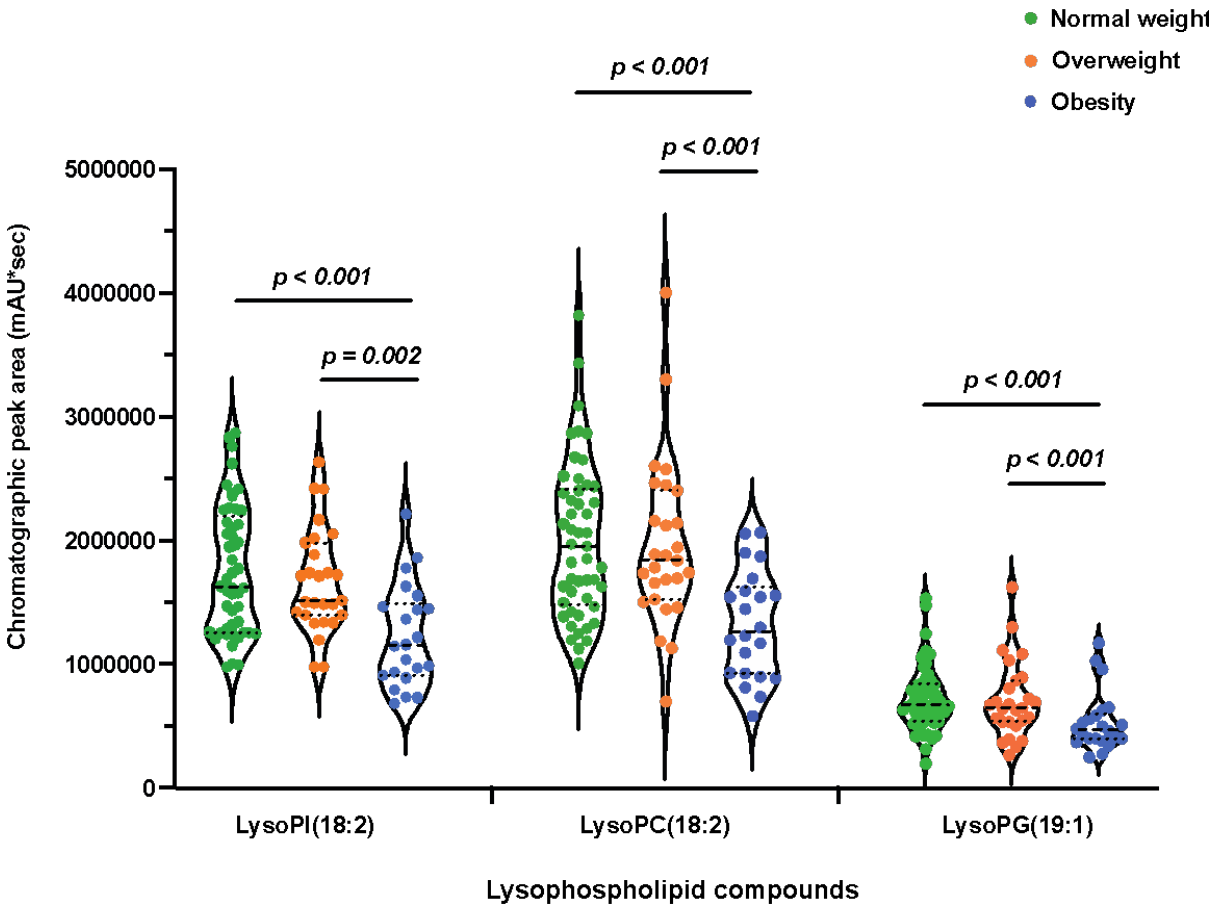


Figure 4.8. Variations in lysophospholipids between BMI subgroups
Comparison of three lysophospholipids (LysoPI(18:1); LysoPC(18:2) and LysoPG(19:1)) between patient subgroups based on BMI-SDS: normal weight (green, n=49), overweight (orange, n=27) and obesity (blue, n=21). The values of *p* are indicated above the groups. The lines within the violin plots correspond to the median with interquartile range. (Overall comparison among the three groups was initially carried out by Kruskal-Wallis H test, followed by Mann-Whitney U test for significant differences between each two groups, * indicating statistical significance for p<0.05)

The correlations between plasma hormonal biomarkers (17-hydroxyprogesterone, androstenedione, testosterone, 11-hydroxyandrostenedione and 11-ketotestosterone) and metabolites were inconsistent, although there was a tendency towards lower

lysophospholipids in patients with suppressed androgens (**Figure 4.9**). There were very few weak correlations between plasma metabolites and plasma cortisol in CAH patients.

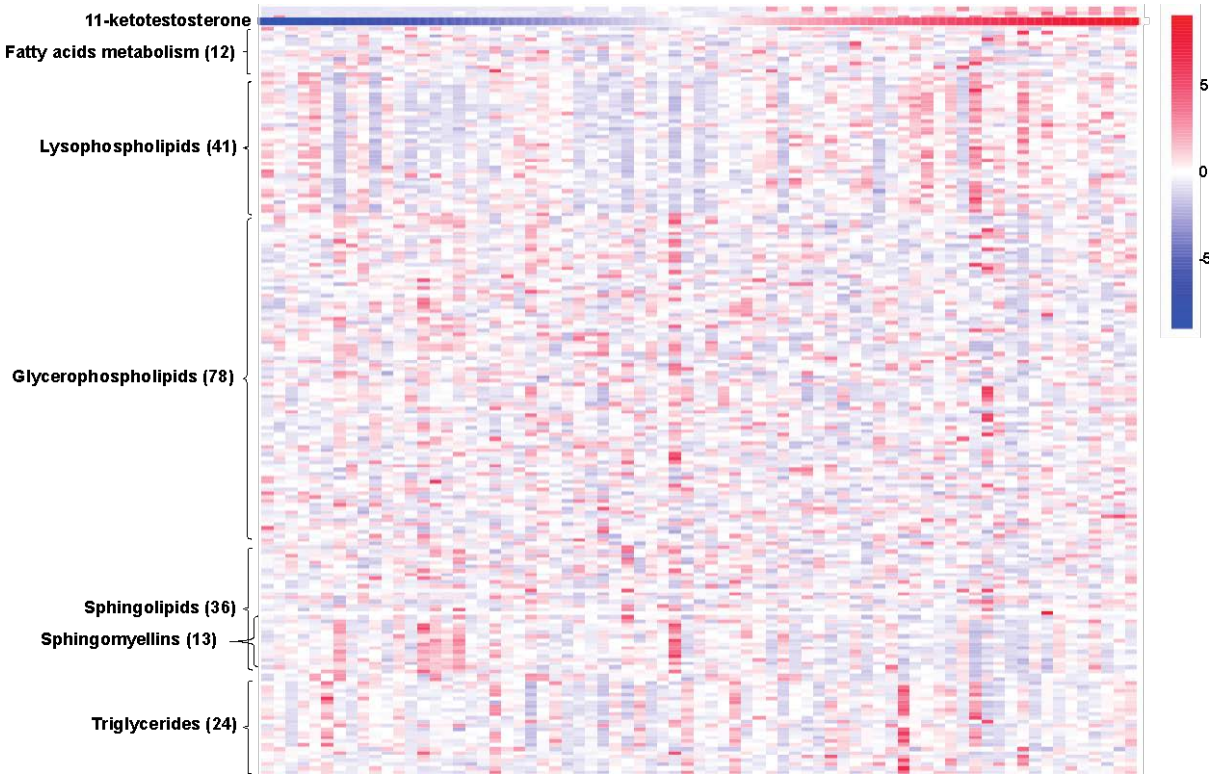


Figure 4.9. Lipid metabolites vs plasma androgens
Heatmap showing the lipid metabolites in relation to the patients' plasma androgen concentration. The columns correspond to individual patients, ordered based on the plasma 11-ketotestosterone concentration, ascending from left to right. The first row shows the 11-ketotestosterone in ascending order from left to right. The rows above 11-ketotestosterone correspond to the other hormones (17-hydroxyprogesterone, androstenedione, testosterone and 11-hydroxyandrostenedione). The subsequent rows correspond to the measurements of individual metabolites; the classes to which the metabolites belong are marked on the left (n = number of compounds in each class).

4.2 Discussion

The development of high-throughput biochemical assay has provided a whole new dimension for exploring the mechanisms of metabolic disease, namely the analysis of the metabolome, which consists of a constellation of small molecular weight compounds involved in biological functions. However, the application of metabolomic analysis in the research of adrenal disease has been limited to date. The only exception is the steroid

metabolome which has been more extensively explored in congenital adrenal hyperplasia in plasma and urine, and consequently disease-specific profiles and normative values are now available (Storbeck, et al., 2019; Kamrath, et al., 2020). The published evidence exploring the wider metabolome in patients with CAH is otherwise sparse and limited to adult studies. Thus, several metabolites were reported to show distinction between CAH patients and healthy controls, especially purine and pyrimidine metabolites, branched amino acids, and tricarboxylic acid cycle metabolites (Nguyen, et al., 2012). A study conducted in adult patients with 21OHD receiving treatment with prednisolone found that the daily GC dose influenced a number of metabolites related to fatty acids, bile acids, and amino acid metabolism (Alwashih, et al., 2017). In the current study analysing data from the CAH-UK cohort of children with CAH, the main focus of the metabolomic analysis was on variations induced by the hydrocortisone treatment, as a way of exploring the impact of cortisol deficit and replacement on metabolism. There was a marked predominance of the compounds involved in several processes of lipid metabolism that were influenced by different aspects of the GC replacement therapy including the daily dose and timing of administration. In contrast to the study conducted on adults (Alwashih, et al., 2017), the data did not highlight the compounds related to amino acid metabolism, however, this may be explained by age-related variations and the difference in the type of synthetic steroid used, which was prednisolone in adults, while the CAH-UK patients were treated with hydrocortisone. Of course, a study caveat lies in the absence of formal information on the patients' compliance to taking the GC medication, an aspect that needs to be considered when interpreting the results.

Lipid metabolites composed the majority of metabolic compounds found to fluctuate with the GC dose, which resonates with previous evidence from adults (Alwashih, et al., 2017).

Among these, glycerophospholipids were predominant. Alongside sphingolipids, glycerophospholipids are an essential component of the neuronal membrane, being also involved in the regulation of signal transduction (Frisardi, et al., 2011; Hishikawa, et al., 2014). Moreover, they have important roles in the dendrite elongation, further impacting the development and function of the brain (Ziegler and Tavosanis, 2019). The predominant type of glycerophospholipids identified in this patient cohort to be influenced by the GC dose were phosphatidylcholines, which is not surprising considering that they are some of the most common glycerophospholipids in mammals (Ziegler and Tavosanis, 2019). Phosphatidylcholines are essential for the synthesis and secretion of lipoproteins thus having an important role in the transport and clearance of VLDL and HDL (Cole, et al., 2012). Consequently, impaired synthesis of phosphatidylcholines plays a part in the pathogenesis of cardiovascular and fatty liver disease (Cole, et al., 2012). In the present study, CAH patients were found to have in general lower levels of glycerophospholipids associated with higher GC daily doses. However, for many compounds a “U shape” or “inverted U shape” dose-depended effect was noted, indicating that deranged phospholipid metabolism may be caused by both lower and higher doses. This strengthens the concept that there are phases in the progression of GC deficiency: hypercortisolism causing reduced GR signalling, adequate replacement, and GR resistance caused by chronic use/ overtreatment with synthetic GCs which action an autoregulatory loop causing suppression of the GR expression (Vandevyver, et al., 2014). Glycerophospholipids were also the metabolites most frequently influenced by the timing of the use of circadian versus non-circadian regimes. This resonates with earlier evidence from healthy adults that showed that the metabolic effects of synthetic GCs are influenced by the timing of the dose, evening administrations associating an increased risk of insulin resistance (Plat, et al., 1999). Thus, these findings help to highlight the importance of

achieving the optimal replacement regime for each patient in order to reduce the risk of developing metabolic disease.

Lysophospholipids were another well represented group of compounds among the metabolites detected to be influenced by the GC dose. Compared between dose subgroups, all the lysophospholipids identified exhibited that “inverted U shape” distribution, being invariably lower in patients with high or low doses compared to those within the recommended dose range, again an indication that cortisol deficiency and toxicity may have similar impact on the metabolism. Lysophospholipids were also the metabolites that correlated best with the BMI, being lower in overweight patients. Lysophospholipids are known to have complex signalling roles and are important precursors for membrane synthesis (Tan, et al., 2020). Lysophosphatidylcholines, the most abundant lysophospholipids in human blood (and in the present results) act as bioactive lipid metabolites that bind to cell-specific G-protein-coupled receptors and were shown to play an important role in the regulation of the vascular inflammatory response and pathogenesis of arteriosclerosis (Li et al., 2016; Knuplez and Marsche, 2020). Correlations were also reported between lysophosphatidylcholines and insulin sensitivity in the muscle (Li et al., 2016). The concentration of lysophosphatidylcholine in the plasma relies on the transacylation of a fatty acid from phosphatidylcholine to free cholesterol (Tan, et al., 2020) and it is reasonable to assume that variations in the administration of GC will lead to the dysregulation of multiple metabolic pathways that will be interlinked. The present findings support this concept, showing that several fatty acid and triglyceride metabolites varied between dose subgroups. This is in keeping with published evidence on the effects of GC on fatty acid metabolism (Macfarlane, et al., 2008). Endogenous and synthetic GC were shown to regulate lipolysis and *de novo* lipogenesis, promoting hepatic VLDL secretion and also potentiating the role of insulin in these processes; however, the

actions of GC on fat metabolism appear to be highly variable depending on whether the administration is acute or chronic (Macfarlane, et al., 2008). This further strengthens the notion that the optimal replacement regime should involve more than just establishing the correct dose and must combine a personalised dose administration programme with lifestyle measures and adequate monitoring strategies.

Another important group of metabolites affected by the GC dose were the sphingolipid compounds. Sphingolipids represent one of the major classes of bioactive lipids in eukaryotes, known to have extensive functions in all essential aspects of cell biology including cell proliferation, differentiation, growth, apoptosis, membrane organisation and dynamics (Hannun and Obeid, 2018). Consequently, their involvement in pathophysiology is equally wide, ranging from inflammation and immune response, neoplastic pathology, to neurodegenerative conditions, metabolic and cardiovascular disease (Hannun and Obeid, 2018). Their inclusion among the metabolites identified in this study further underlines the impact of cortisol deficiency and GC therapy on the health outcomes of patients with CAH.

The metabolomic analysis provided much more modest insights regarding the effects of the GC dose on other major metabolic pathways, such as protein and carbohydrate metabolism. A small number of metabolites related to protein metabolism was identified, where the observed dose effect indicated increased catabolism in patients receiving high or low doses of GC in contrast to those within the recommended 10-15 mg/m^2/day hydrocortisone daily range. To some extent, these findings are in keeping with the known actions of cortisol in reducing protein synthesis through the suppression of the mTOR pathway (Shah, et al., 2002; Jellyman, et al., 2012; Schakman, et al., 2013) and increasing protein catabolism by stimulating the ribosomal ubiquitin system and

lysosomal autophagy (Schakman, et al., 2013). The implication that low cortisol concentrations would have similar effects is interesting, however, the overall small number of protein metabolites makes it difficult to interpret the results, especially in the absence of published evidence on the topic.

Several lipid metabolites were found to have opposite relationships with the relative fludrocortisone dose compared to those with the relative GC dose. This may represent an indirect effect of the negative correlation existent in the cohort between the mineralo- and glucocorticoid doses. On that topic, previous studies have debated the benefits of using fludrocortisone to reduce GC overexposure, based on the transactivation of the GR by mineralocorticoids (Ekman, et al., 2015; Puglisi, et al., 2021). However, it is now known that aldosterone plays a role in the regulation of adiposity and lipid metabolism. Published evidence indicated an association between raised aldosterone concentrations and increased fat mass, suggesting that aldosterone modulates insulin sensitivity, the release of free fatty acid and adipokines from adipocytes, contributing to the pathogenesis of metabolic syndrome (Even, et al., 2014; Feliciano Pereira, et al., 2014). Moreover, it was hypothesised that aldosterone, together with cortisol, is involved in potentiating a cross-talk between adipose tissue and blood vessels through the activation of gluco- and mineralocorticoid receptors which contributes to the development of the cardiometabolic syndrome (Even, et al., 2014). This could be an explanation for the finding of such discrepant effects on lipid metabolites between the fludrocortisone and the hydrocortisone dose, suggesting that mineralocorticoid as well as GC replacement must be taken into account when monitoring the development of metabolic complications in patients with CAH.

When exploring associations between plasma metabolites and other measured outcomes, the most notable one was with weight gain expressed as BMI-SDS. The correlations found with leptin were almost superposable to those with the BMI, which is not surprising given the known association between them. Lysophospholipids, which were found to be lower in patients subgroups receiving daily GC doses outside the recommended range (lower or higher), correlated negatively with the BMI. Thus, an association is indicated between suboptimal replacement doses, increased body weight and deranged lipid metabolism. Given the important role played by lysophospholipids, including involvement in cellular signalling, it is likely that in the long run their dysregulation will contribute to the development of metabolic and cardiovascular disease in CAH. The relationships with the weight gain status of other groups of lipid compounds were weaker and less consistent. However, overall, these associations further demonstrate the fact that metabolic disease in CAH due to 21OHD has its onset during childhood and is linked to the quality of hormone replacement. The relationship between the lipid metabolites and other variables such as HOMA-IR and plasma androgens, when present, were weaker and less consistent. Lysophospholipids had a tendency towards lower levels in patients with suppressed androgens. This would indicate the downregulation of the pathways in which they are involved in patients that are overtreated with GC, however, as hormonal results were based on one-time measurements at variable times after the GC dose, the relevance of plasma androgens as markers of disease control in the CAH-UK cohort is limited. On the other hand, based on evidence from animal and human studies, it is known that androgens have a direct effect on lipid metabolism, being regulators of adipocyte differentiation, adiponectin secretion from adipose tissue, lipolysis, lipogenesis and insulin signalling (O'Reilly, et al., 2014; Lopes,

et al., 2021). Thus, it is to be expected that hyperandrogenism will contribute to fat metabolism dysregulations in undertreated CAH patients.

Overall, the metabolomic analysis showed that in children with CAH the GC regime impacts on several groups of compounds related to lipid metabolism. These metabolites varied in relation to the GC dose and timing of the administration. The most important groups of compounds were glycerophospholipids, lysophospholipids and sphingolipids, which are known to have wide roles in essential biological processes and are involved in the pathogenesis of metabolic and cardiovascular disease. The variation of these metabolites with the GC dose and their relationship with the patients' BMI indicate the importance of aiming to optimise hormone replacement and develop adequate monitoring strategies in CAH in order to reduce the development of metabolic comorbidities.

5. Metabolic analysis of *cyp21a2* function in adult zebrafish

It is important to note that mutant fish from both lines used (*cyp21a2* and *cyp11a2*), were able to survive to adulthood, despite being cortisol deficient, in the absence of glucocorticoid replacement. To help explore the molecular mechanisms involved in the development of metabolic problems observed in patients with CAH, the clinical findings were complemented by research using a zebrafish model of 21OHD. The metabolic phenotype of adult *cyp21a2-/-* was determined and the differential expression of specific genes involved in glucose and fat metabolism was measured in mutant larvae and adult livers in comparison to wild type (WT) controls. This chapter will present evidence that adult *cyp21a2-/-* zebrafish mutants have higher weight and length, as well as increased visceral and subcutaneous fat compared to WT siblings. There was dysregulation of several genes involved in glucose and fat metabolism, indicating downregulation of gluconeogenesis, insulin signalling and adipocyte differentiation in *cyp21a2-/-* deficient larvae and adult livers.

5.1 Cortisol-deficient adult *cyp21a2* mutant zebrafish are larger and have increased fat mass compared to wild type siblings

5.1.1 Results

5.1.1.1 Morphological characterisation of *cyp21a2* deficient 18 months old zebrafish

There were no significant differences between *cyp21a2* fish and WT regarding the secondary sexual characteristics, which is a distinction from human 21OHD, where females are virilised, and males present with premature adrenarche. In zebrafish, androgens are not produced in the interrenal gland, consequently, *cyp21a2* were not

expected to have androgen excess. However, mutant fish were larger compared to controls and had more visceral and subcutaneous fat on dissection.

The mutant *cyp21a2-/-* fish developed as both male and female adults that were capable of reproduction. The secondary sexual characteristics, including the pigmentation of the fins and body, body shape and the prominence of the genital papilla, were no different between mutant and WT siblings from the same generation (**Figure 5.1**).

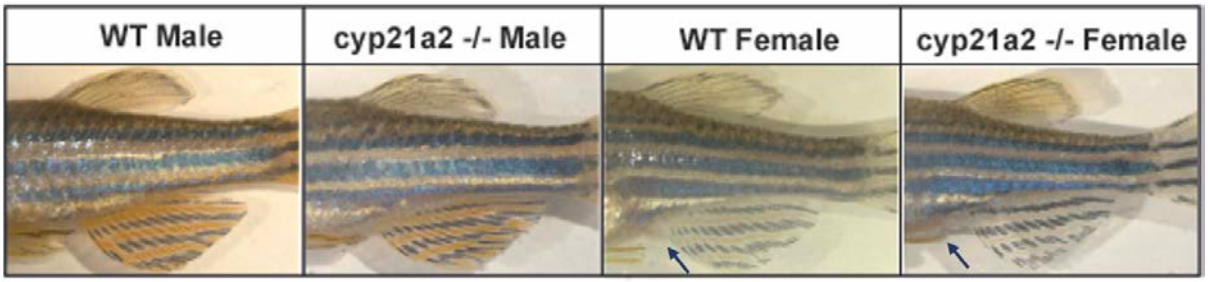


Figure 5.1. External phenotypes of *cyp21a2* mutant and wild-type sibling fish
Secondary sex characteristics are similar between adult *cyp21a2* mutants and WT for both genders. WT adult male zebrafish exhibit strong stripes of golden pigment in the anal fin, interposed with the black stripes; this feature is absent in females. Female WT fish have a visible genital papilla anterior to the anal fin (black arrows). These features were reported to be altered, as abnormal external sex characteristics, in other mutant zebrafish lines of deranged steroid synthesis, however, in both male and female adult cyp21a2-/- fish they were comparable to WT controls.

However, the weight and length measurements showed *cyp21a2*-/- fish to be both heavier and longer compared to WT (**Figure 5.2A, C**). This feature was very consistent in males, being present in two different populations of fish for both weight ($p = 0.003$ for the first and $p < 0.0001$ for the second population) and length ($p = 0.001$ for the first and $p < 0.0001$ for the second population). However, in females, the differences in weight and length were only significant in one population (data not shown). As female weight fluctuates significantly in relation to the reproductive cycle and subsequent number of oocytes, the findings relating to males were more consistent and are much more relevant in this context. In the *cyp11a2* line, mutants were also found to have higher weights ($p <$

0.0001 for the first, p = 0.036 for the second and p = 0.0003 for the third population) and lengths (p < 0.0001 for the first, p = 0.012 for the second and p = 0.022 for the third population) compared to controls (**Figure 5.2B, D**).

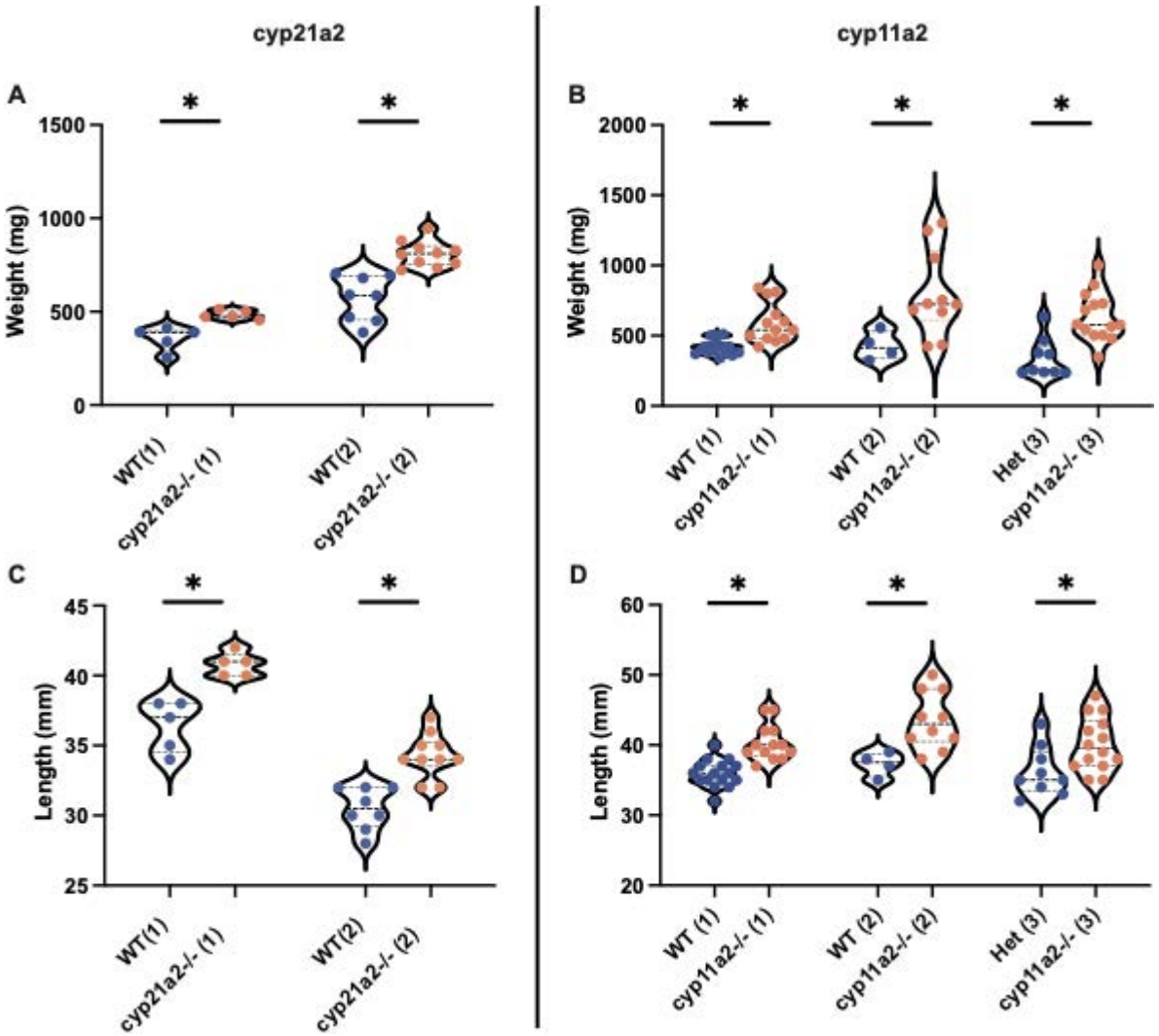


Figure 5.2. Weight and length measurements in cyp21a2 and cyp11a2 fish
Male zebrafish, WT (blue) and mutant (orange) fish of the c*yp21a2* (**A and C**) and *cyp11a2* (**B and D**) lines over different populations. The *cyp21a2* fish were 18 months old in both populations. The *cyp11a2* fish were 7 months – 1, 12 month – 2 and 21 months – 3. (* indicates statistical significance, p<0.05). The horizontal lines represent median with interquartile range. (N= 4 – 14 fish. For the *cyp21a2* line, for each population, the sample size was powered for gene expression analysis, aiming for minimum 5-7 individuals per group, the same fish being used for experiments and measurements. Of the *cyp11a2* line, a similar approach was taken; the second population had a smaller number of WT fish and it was not used of gene expression analysis in this project, however the results of the body measurements were displayed to highlight that the increased body size is a trend observed in separate mutant populations).

On dissecting the *cyp21a2* fish, it was observed that homozygotes had a tendency towards increased visceral fat deposition compared to WT, with more extensive fat deposits around the organs of the gastrointestinal tract. This feature was much more marked in males compared to females. Interestingly, it was noted that even in WT fish, males had a tendency to have more visceral fat compared to females. All mutant fish, both male and females, displayed an increased amount of fat surrounding the brain and subcutaneous fat compared to control siblings. The presence of increased fat in mutants was also noted in the *cyp11a2-/-* mutant fish, where it was even more marked compared to the Cyp21a2-deficient fish (**Figure 5.3**).

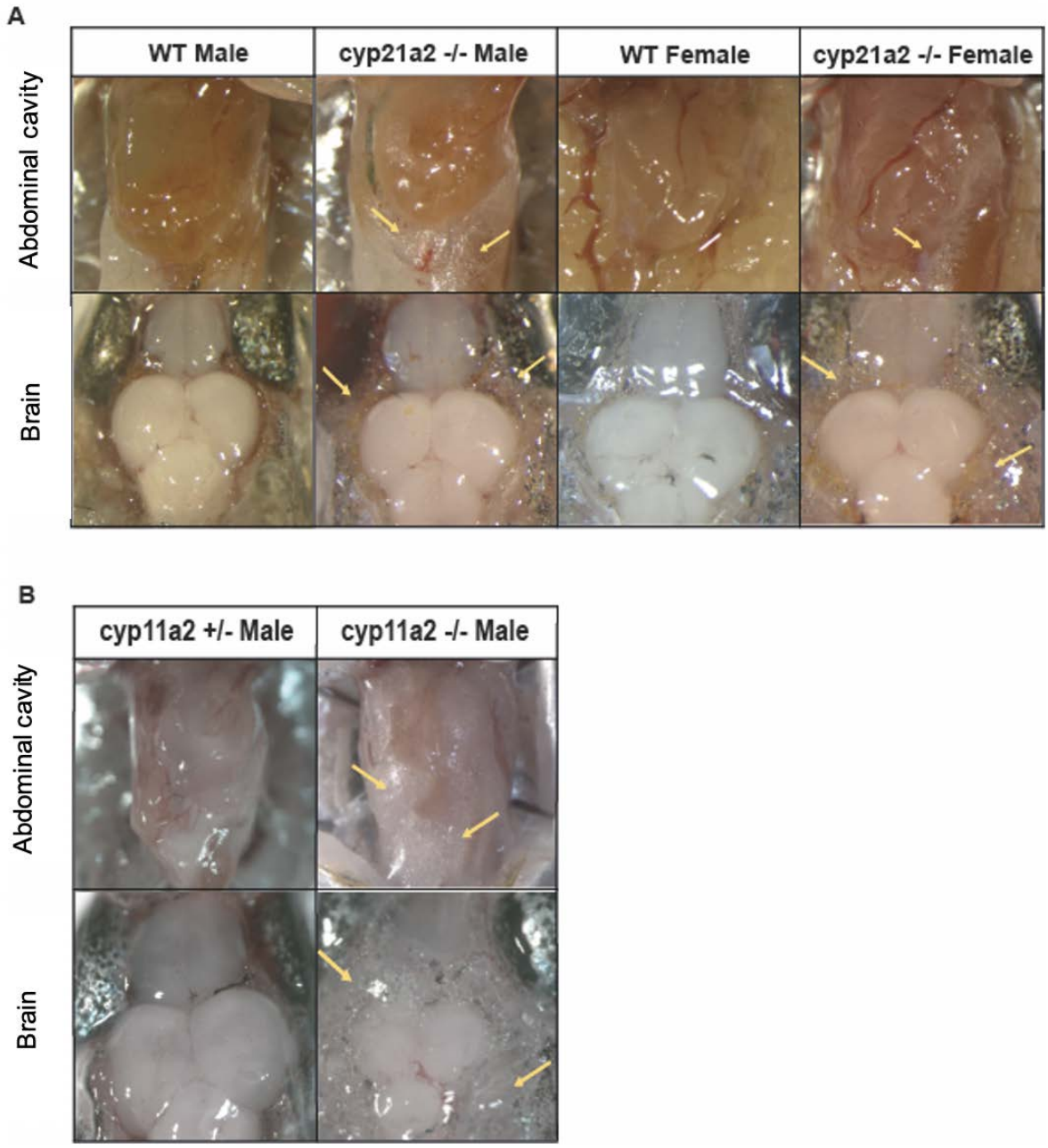


Figure 5.3. Dissection images
Dissection images from the *cyp21a2* (**A**) and *cyp11a2* (**B**) fish. The upper pictures show the visualisation of visceral fat on inspecting the gastrointestinal tract. The lower images were obtained on dissecting the skull, visualising the brain and the subcutaneous fat around it. The visceral and subcutaneous fat is indicated by yellow arrows.

Female zebrafish had significantly larger livers compared to males (**Figure 5.4A**). This was found in both WT ($p = 0.0001$) and mutant fish ($p = 0.004$) with no significant difference between WT and mutants pertaining to the same sex. These differences were

maintained when using the liver weight to body length ratio, as means to remove body size as a potential confounder for organ size (**Figure 5.4 B**).

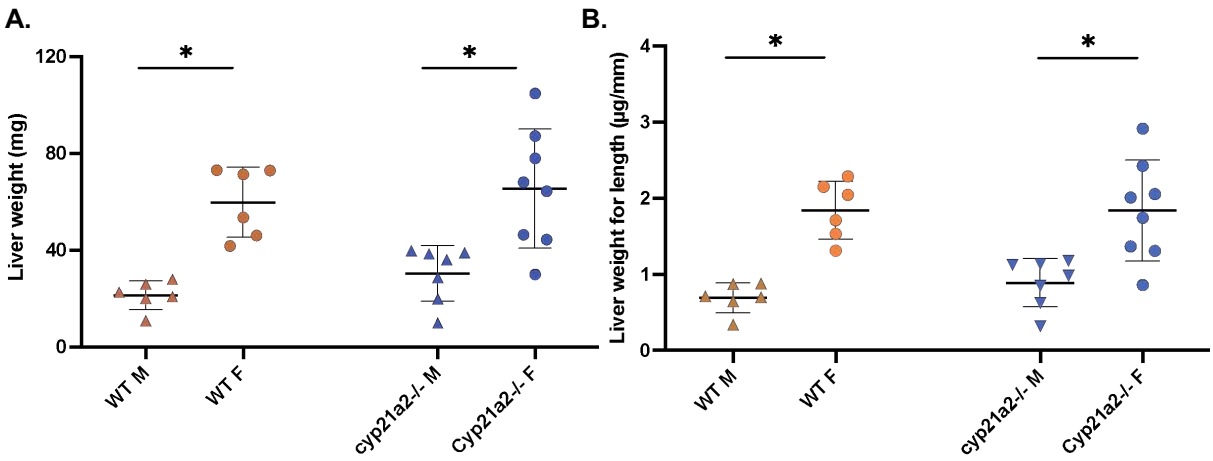


Figure 5.4. Liver weight in Cyp21a2-deficient fish
A. Liver weight in WT (orange) and *cyp21a2*-/- (blue), males (M, triangles) and females (F, circles). (* indicates statistical significance). **B.** Liver weight to body length ratio compared between the same groups. The horizontal lines represent the median with interquartile range. n= 6 - 8 fish

5.1.1.2 Blood glucose measurements

The blood glucose concentrations (measured between 1-3 hours after feeding) varied widely in WT and mutants of both sexes. The levels were similar statistically between mutants and their control siblings in both *cyp21a2* and *cyp11a2* (**Figure 5.5**).

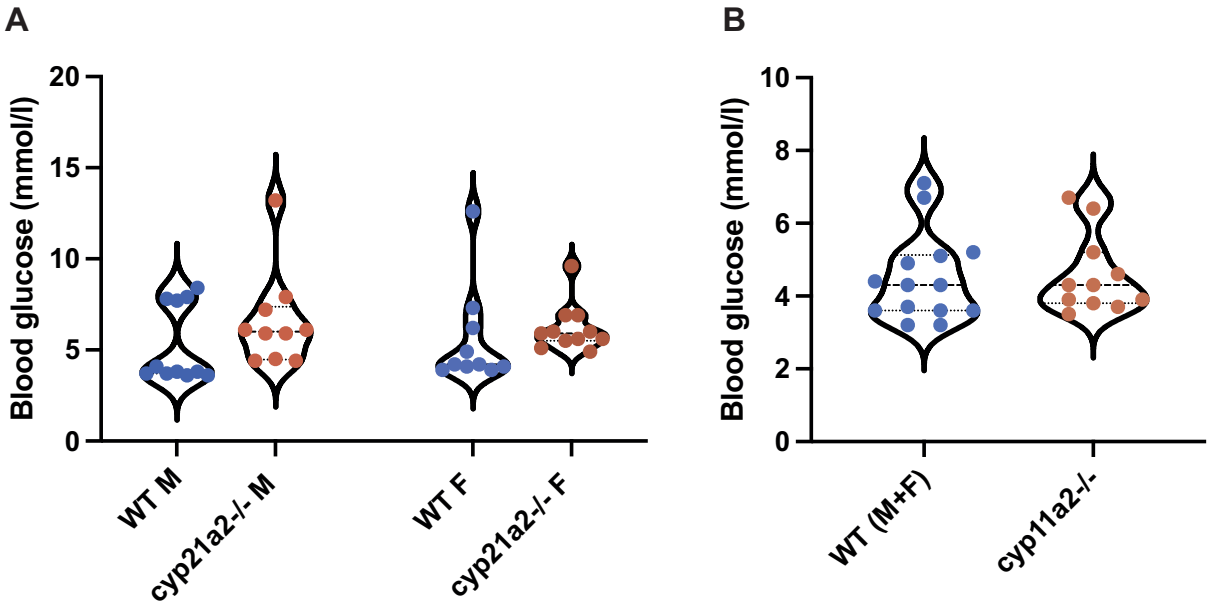


Figure 5.5. Blood glucose in zebrafish
Blood glucose concentrations in WT (blue) and mutant (orange) fish of the *cyp21a2* (**A**) and c*yp11a2* (**B**). The horizontal lines represent the median with interquartile range. n= 10 - 14 fish

5.1.1.3 Liver histology sections

The study focused specifically on the adult zebrafish liver because of the important roles it has in metabolism and energy homeostasis in both humans and zebrafish. Liver histology was similar between *cyp21a2-/-* fish and wild type siblings when using H-E staining, however, the hepatocytes of *cyp11a2-/-* mutants presented a large number of clear inclusions.

The histological appearance of HE stained livers was different between males and females in WT and mutant fish, with females having slightly more basophilic hepatocytes (**Figure 5.6A**), a feature that has been previously described and is associated with the increased vitellogenin content of female livers (Menke, et al., 2011).

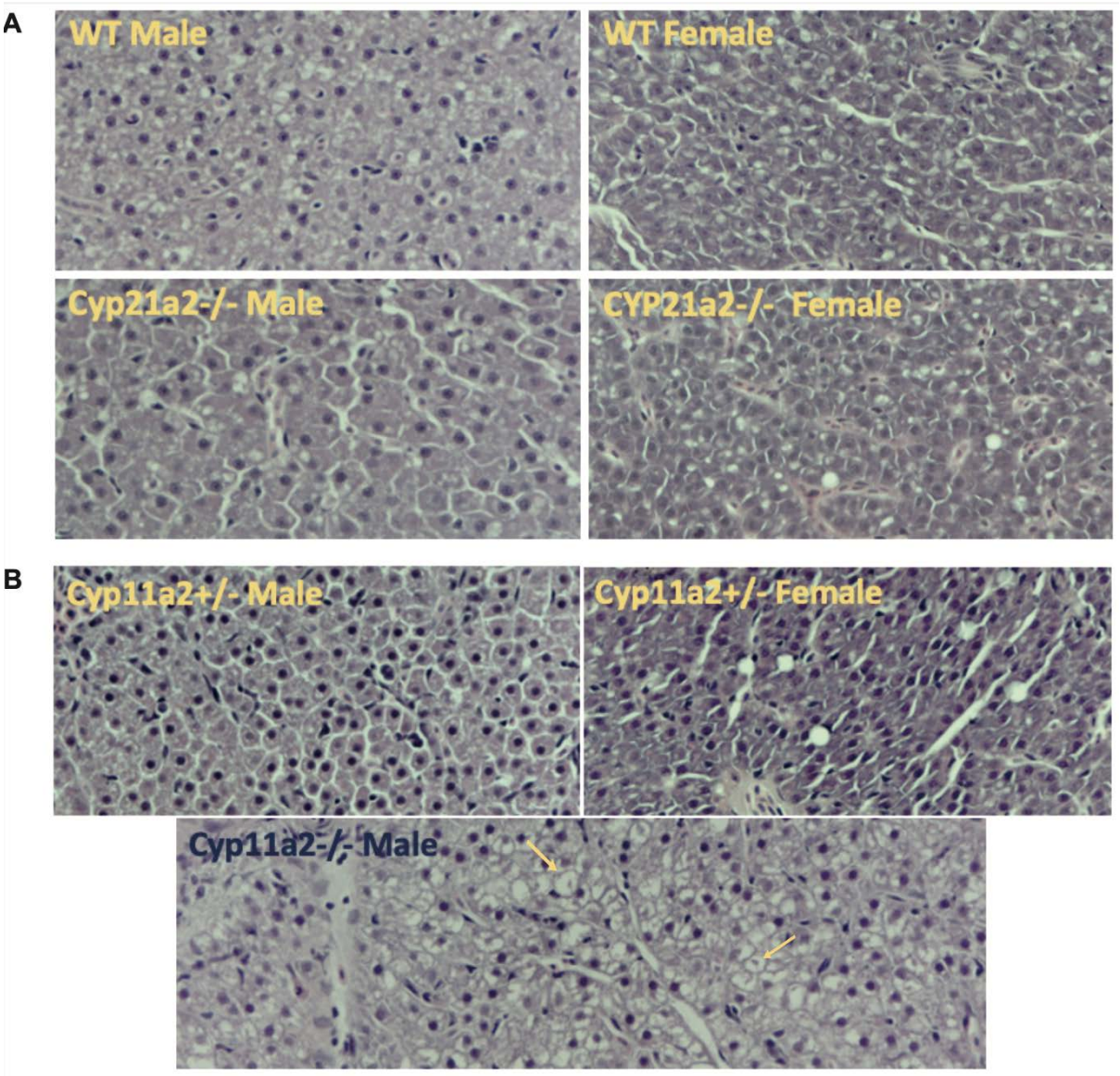


Figure 5.6. Histology sections liver
Histology sections with HE staining of adult livers (x40) in *cyp21a2* (**A**) and *cyp11a2* (**B**) adult fish. The upper sections correspond to WT (*cyp21a2*) and heterozygous (c*yp11a2*) siblings and the lower images to mutant fish. The yellow arrows indicate the clear inclusions observed in the cyp11a2-/- fish hepatocytes.

There was no marked difference in the histological appearance of livers from WT and *cyp21a2-/-* fish pertaining to the same sex (**Figure 5.6A**). However, *cyp11a2-/-* mutant hepatocytes displayed large clear inclusions in comparison to both male and female heterozygous controls (**Figure 5.6B**).

5.1.1.4 Whole body steroids in adult cyp21a2 fish

Because this project aimed to explore the metabolic abnormalities developed as a consequence of deranged steroid synthesis and in particular hypocortisolism, it was important to define the steroid profile of the adult *cyp21a2-/-* fish model. The steroids detected and measured were progesterone, 17OHP, cortisol, androstenedione and testosterone; corticosterone, 11-deoxycortisol, DHEA, 11-hydroxyandrostenedione and 11-ketotestosterone could not be detected by the technique used.

Mutant *cyp21a2-/-* fish had significantly lower cortisol ($p < 0.001$), confirming systemic GC deficiency. Overall, 17OHP was not statistically different between mutants and WT, however, that was due to one outlier among the WT, the removal of which led to results that were consistent with human 21ODH presenting with excessive 17OHP ($p < 0.001$). By contrast, androstenedione was significantly lower in *cyp21a2-/-* fish compared to WT siblings, which is the opposite to what is seen in CAH patients ($p=0.005$). There was no distinction in testosterone and progesterone based on genotype in adults (**Figure 5.7**). There were no significant differences between sexes for any of the steroids measured.

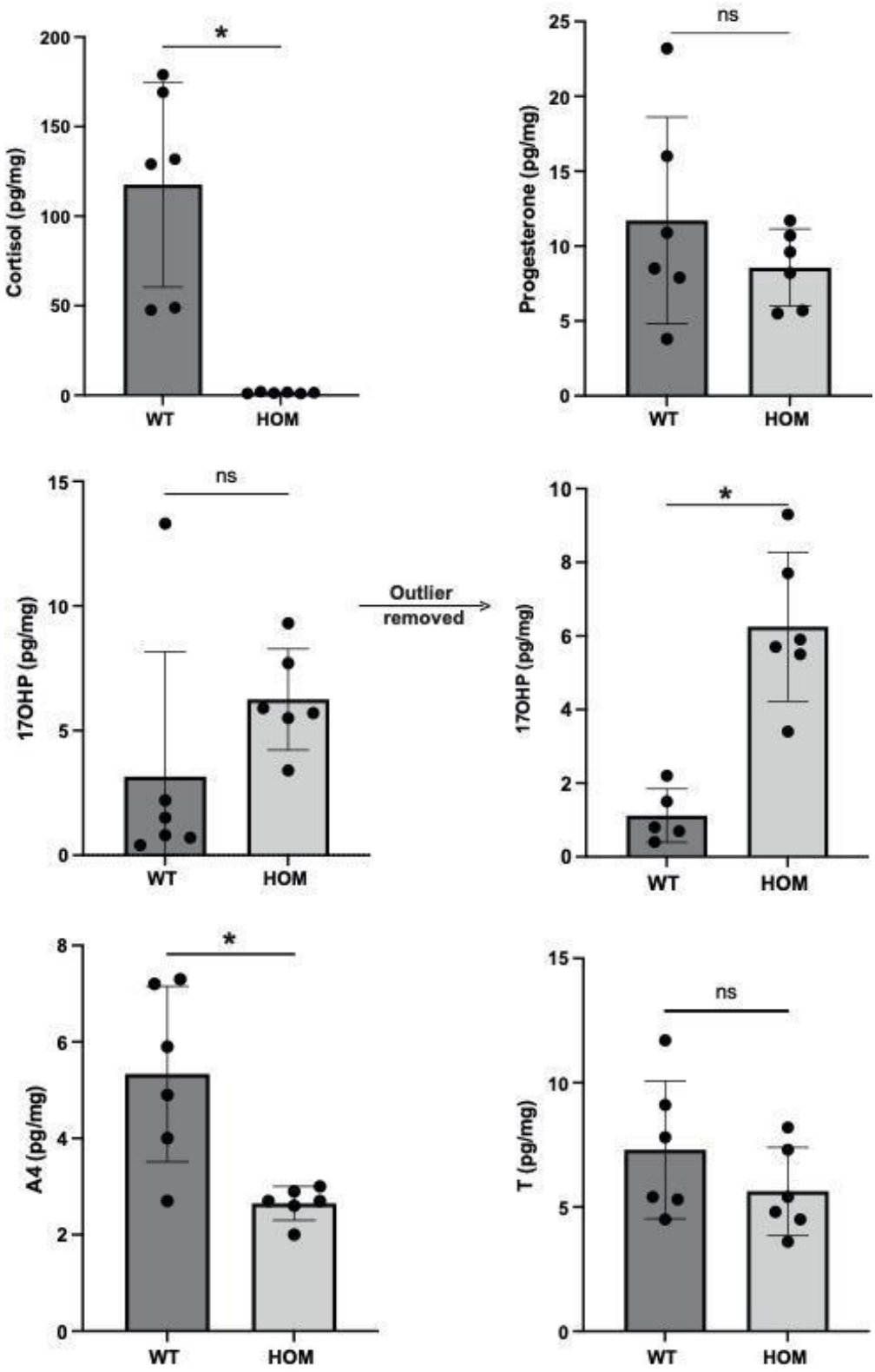


Figure 5.7. Whole body steroid measurements in *cyp21a2* fish
Comparison of whole-body steroids between WT (dark grey) and mutant (light grey) The horizontal lines represent median with interquartile range. (*indicates statistical significance). n= 6 fish per group (17OHP: 17-hydroxyprogesterone, A4: androstenedione, T: testosterone, WT: wild type, HOM: homozygous *cyp21a2-/-*).

5.1.2 Discussion

The findings indicate that adult (18 months) *cyp21a2*-/- mutant fish are heavier and longer compared to WT siblings. However, they are fertile and display normal sexual characteristics. Moreover, the macroscopic findings on dissection showed a tendency for increased visceral and subcutaneous fat in *cyp21a2*-/- mutant fish. Several mechanisms and pathways should be discussed as possible determinants of these findings. Firstly, we must question the possibility of a direct effect of cortisol deficiency on metabolism, causing the raised body size in mutant fish. This is supported by evidence from several animal studies, as well as a more modest number of human studies, that were discussed in the introduction, indicating an impact of cortisol deficiency on multiple metabolic pathways. Importantly, increased body size was also found in other cortisol deficient zebrafish lines that were established in our laboratory, including *cyp11a2-/-* (Li, et al., 2020), *cyp11c1-/-* (Oakes, et al., 2020), *fdx1b-/-* (Oakes, et al., 2019), supporting a role for cortisol in the regulation of metabolism. However, some of these lines (*cyp11a2-/-, cyp11c1-/-, fdx1b-/-)* are also androgen-deficient. There is some evidence indicating that hyperandrogenism promotes somatic growth in the presence of cortisol deficiency in *pomca-/-* zebrafish mutants (Shi, et al., 2020), but no evidence was identified on the effects of hyperandrogenism on growth. In humans with 21OHD, cortisol and aldosterone deficiencies are accompanied by androgen excess, including raised androstenedione, testosterone, as well as the newly established adrenal-specific 11-oxygenated androgens. However, the steroid profile of adult *cyp21a2-/-* zebrafish determined by whole body measurements showed androstenedione to be reduced significantly in mutants while testosterone was similar to WT. The inability to measure 11-oxygenated androgens, including 11-hydroxyandrostenedione and 11-ketotestosterone, is likely

related to a technical problem, as previous studies were able to quantify them in both WT and steroid deficient fish (Oakes, et al. 2019, Li, et al. 2020). However, 11-oxygenated androgens, especially 11-ketotestosterone, were shown to play a role in the acquisition of secondary sexual characteristics in zebrafish males (Oakes, et al., 2019; Oakes, et al., 2020). There is limited evidence on the impact of androgen excess in female zebrafish, however, one study reported an effect on ovary maturation and the maintenance of secondary sexual characteristics (Yu, et al., 2020). Since there were no significant differences between *cyp21a2*-/- mutant and wild-type fish regarding the secondary sexual characteristics, it is possible that strong androgens and especially those belonging 11-oxygentaded C19 group may be normal or not significantly altered in mutant *cyp21a2-/-* adult fish. Overall, this would suggest that the large phenotype observed in mutants would be related to the cortisol deficiency, which was plainly demonstrated by the steroid profiles. On the other hand, the incidental finding of female livers being larger than those of males in both wild types and mutants would support the theory that there are physiological differences in the liver function between males and females.

It is also important to consider the interplay between growth hormone (GH) and the stress axis, which has been previously recognised and described in paediatric pathology (Stratakis., 2006), although its complexities are not fully understood. In vitro evidence using sheep pituitary glands showed that cortisol inhibits the secretion of GH, possibly through actioning calcium channels (Sartin, et al., 1994). Evidence from zebrafish demonstrated the existence of an anatomical and functional circuit comprising POMC – somatostatin – GH, with POMC neurons exercising an inhibitory effect on the pituitary secretion of GH (Löhr, et al., 2018). Therefore, future research is required to help define the role of GH as a regulator of somatic growth in the GC deficient *cyp21a2-/-* zebrafish

model, the differential expression of genes involved in the GH regulation in adult brain - growth hormone (*gh*), proopiomelanocortin (*pomc*), growth hormone releasing hormone (*ghrh*), somatostatin 1 (*sst1*), insulin-like growth factor 1 (*igf-1*), igf-1 binding protein (*igfbp-1*).

Finally, the effect of GC and more importantly in this case, their deficiency, on the regulation of food intake and satiety was demonstrated in murine models (Makimura, et al., 2003; Uchoa, et al., 2009; Uchoa, et al., 2012). Thus, GC deficiency was associated with a deranged satiety response and abnormal circadian feeding pattern. The abnormal phenotype with raised body mass and adipose tissue in the *cyp21a2-/-* mutant fish, which by comparison appeared to be even more marked in the severely cortisol and androgen deficient *cyp11a2-/-* line, raises the question of a possible impact of altered steroidogenesis on feeding behaviours in zebrafish. Previous evidence has demonstrated reduced food intake in response to stress in zebrafish, associated with overexpression of *pomc* (Cortés, et al., 2018). On a related note, Cyp21a2-deficient larvae were shown to have increased expression of *pomc* in response to cortisol deficiency (Eachus, et al., 2017). While including feeding experiments in this thesis was not logistically possible, exploring the differential expression of orexigenic (neuropeptide y, agouti related protein, ghrelin) and anorexigenic (pomc) peptides in adult brain for *cyp21a2-/-* zebrafish could be included in future work seeking further insights into the regulation of metabolic processes in cortisol deficient zebrafish.

The histological sections of adult livers with HE staining showed no marked differences between *cyp21a2-/-* fish and their WT siblings. By contrast, the hepatocytes of *cyp11a2-/-* presented a significant number of large clear inclusions, which dramatically

distinguished them from both male and female livers of *cyp11a2+/-* fish. It can be speculated that these inclusions correspond to fat or glycogen deposits. Both alternatives would have implications related to the metabolic abnormalities that occur in the liver of adult zebrafish with absent P450 side chain cleavage enzyme. Moreover, both hepatic steatosis and glycogenosis have been previously described in human pathology in relation to dysregulation of insulin signalling pathways (Julián, et al., 2015). Thus, these findings are relevant from a basic science and translational perspective, warranting further work to explore the nature of the hepatocyte inclusions and the mechanism involved in their development.

Overall, the steroid profiles of *cyp21a2-/-* mutant adult fish confirmed profound cortisol deficiency, as well as the build-up of androgen precursor 17OHP, which is consistent with human 21OHD. However, the hyperandrogenism we see in CAH patients was not identified in the zebrafish model, rather the opposite was observed, as androstenedione was raised. It could be speculated that this could be due to a preferential deviation of the steroid synthesis pathway towards 21-deoxycortisol as a result of a compensatory upregulation of *cyp11c1*. Previous evidence using the *fdx1b-/-* zebrafish model would support this possibility (Oakes, et al. 2019). Unfortunately, the other androgens, including corticosterone, 11-deoxycortisol, DHEA, 11-hydroxyandrostenedione and 11-ketotestosterone could not be detected, most likely due to a technical problem. However, it is possible that levels of 11-oxygenatded androgens are similar between *cyp21a2-/-* and WT adult zebrafish, based on the fact that mutant fish are fertile and develop normal secondary sexual characteristics, which have been shown to be dependent on 11-ketotestosterone (Oakes, et al., 2019; Oakes, et al., 2020). The advantage of using the *cyp21a2-/-* zebrafish line is that it allows us to study the isolated effects of

hypocortisolism, without the interference of androgen excess, thus resonating with the aim of this project. However, from a translational perspective, the *cyp21a2-/-* zebrafish is not a complete model of human 21OHD, where hyperandrogenism is an important factor in the disease physiopathology and a major determinant of the steroid replacement therapy, an aspect that will have to be considered when interpreting the results.

5.2 Cyp21a2 is required to promote expression of genes involved in gluconeogenesis in zebrafish larvae and adult livers

In order to explore the impact of cortisol deficiency on some of the most important metabolic pathways, RT - qPCR was used to determine the differential expression of specific metabolic genes in the *cyp21a2-/-* zebrafish larvae and adult livers. To ensure the reliability of results, two separate experiments were conducted by the same researcher, using the same methodology but different biological replicates. The most consistent finding was the downregulation of gluconeogenesis, identified in both larvae and adult livers. The other metabolic processes showed inconsistent dysregulation between the two different stages of development.

5.2.1 Results

As a first step, the expression of *fkbp5* was compared between mutant and WT fish. Because *fkbp5* plays an important role in GR signalling (Gans, et al., 2021), it is one the GC responsive genes most commonly used to ascertain GC deficiency. For all samples included in the analysis, there was significant down-regulation of *fkbp5* (**Figure 5.10**). On replicating the experiments (2021), one of the female livers showed significantly low expression of *fkbp5* and was consequently removed from the analysis of the glucose metabolism genes as an outlier. For the experiments exploring the expression of fat

metabolism genes another cDNA sample was obtained from a different WT female with normal expression of *fkbp5*.

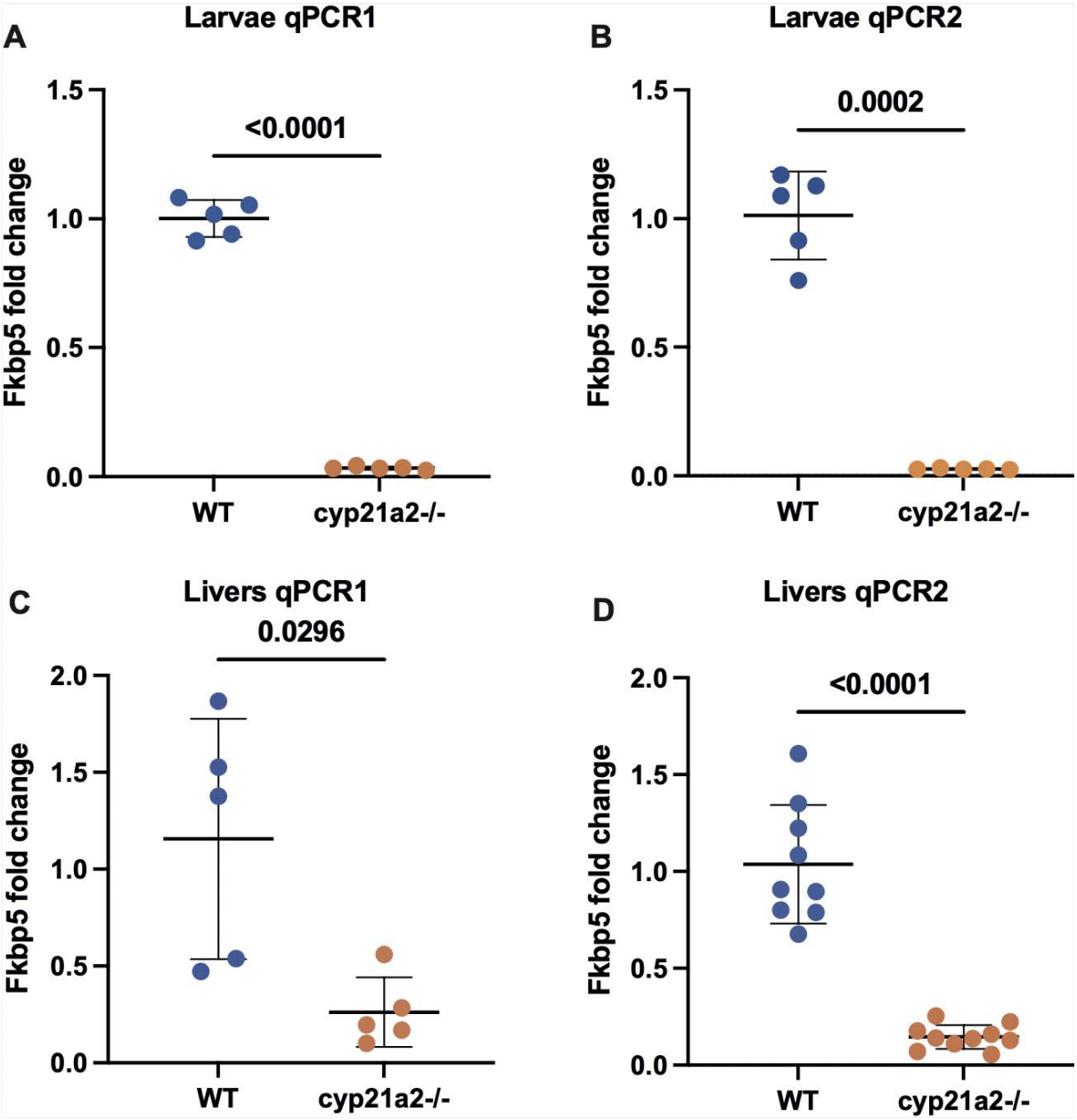


Figure 5.8. Expression of fkbp5 in mutant larvae and livers
The relative expression of *fkbp5* by qPCR showed significant downregulation in the *cyp21a2-/-* fish for larvae and adult livers in both experiments (**A**. 1st experiment in larvae, n=5 samples, *p<0.0001*; **B**. 2nd experiment in larvae, n=5 samples, *p=0.0002;* **C**. 1st experiment in livers, n=5, *p=0.029;* **D**. 2nd experiment in livers, n=9 for WT and n=10 for mutant livers, *p<0.0001*) (* indicates statistical significance). The error bars mark the standard deviation.

The **qPCR results for genes involved in glucose metabolism** indicated in the first experiment (2020) down-regulation of gluconeogenesis in the *cyp21a2-/-* larvae, as shown by the significantly low expression of *pck1* ($p < 0.001$) and *pck2* ($p = 0.008$) in mutants compared to WT siblings. The insulin *insa* gene was also downregulated ($p = 0.001$); however, there was no significant dysregulation of glycolysis or the *irs2* genes. (**Figure 5.11A**). The replication of the experiment in 2021 found that the only significant difference was the downregulation of *pck1* ($p < 0.001$) (**Figure 5.11C**).

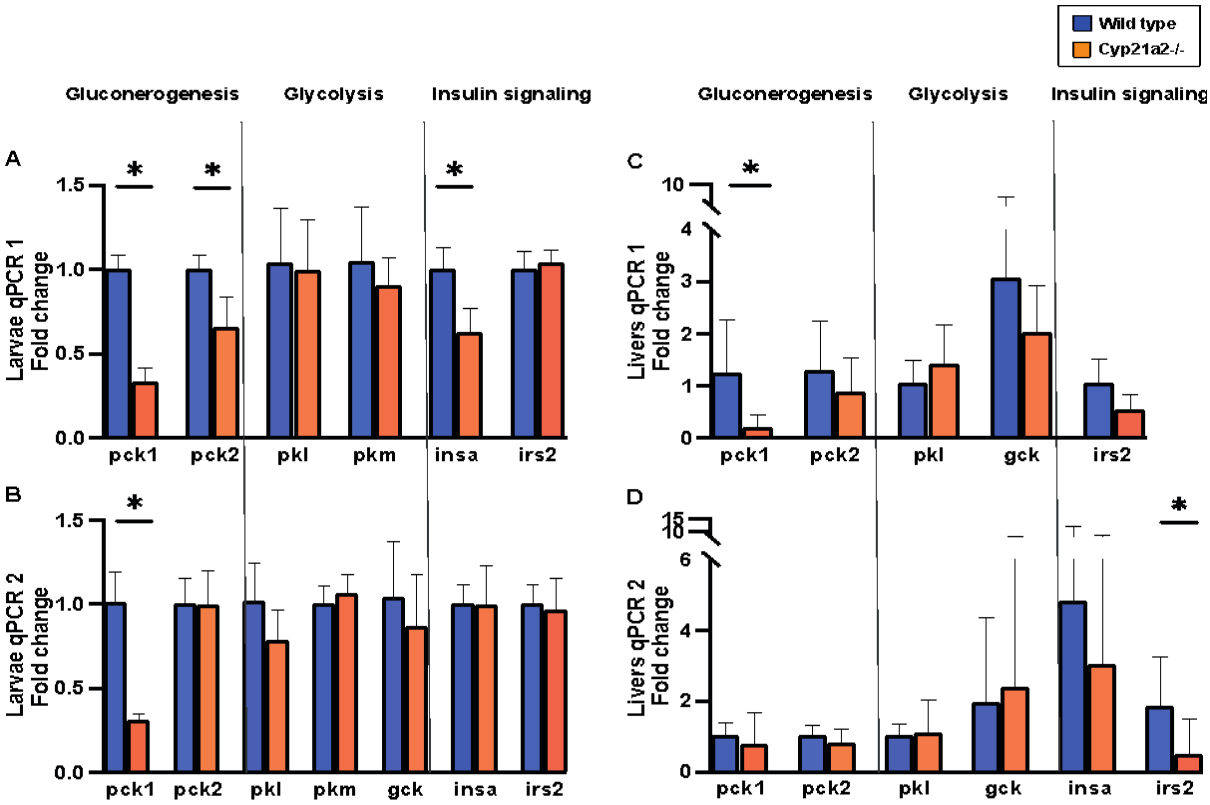


Figure 5.9. Glucose metabolism qPCR
Differential expression of genes involved in glucose metabolism in larvae (left: **A** and **B**, corresponding to the two separate experiments using different biological replicates, n=5 samples for each group.) and adult livers (right **C** and **D**, corresponding to the two separate experiments using different biological replicates, for **C**: n=5 livers for each group, for **D**: n=9 for WT and n=10 for mutant livers). The upper graphs represent the results of the first experiments and the lower ones, those following the repeated experiment. Fold changes in *cyp21a2-/-* fish (blue) were expressed in relation to WT (orange). The dotted lines separate genes based on the metabolic processes they regulate, which are indicated above. (* indicates statistical significance). The error bars mark the standard deviation.

The suppression of gluconeogenesis was also found in adult liver samples during the first qPCR experiment (*pck1*, p = 0.047), with no other significant differences found in the expression of other genes pertaining to glucose metabolism (**Figure 5.11B**). However, on replicating the experiment using a larger number of samples (5 males and 5 females of each genotype), there was no statistical significance in the differential expression of *pck1* (**Figure 5.11D**). The sex subgroup analysis showed that this was caused by the male mutants, where there was a tendency towards upregulation of *pck1,* which was significantly downregulated in female mutants (p < 0.001). The second experiment also showed the *irs2* gene to be significantly downregulated in mutants (p = 0.046); this was due to the female samples (p = 0.002), while in males there was no significant difference between mutants and WT**.**

Regarding **genes involved in fat metabolism,** differential expression analysis by qPCR showed variations between larvae and adult livers. In larvae there was downregulation of genes involved in adipocyte differentiation (*cebpa,* p=0.009) and upregulation of genes involved in lipid storage (*fasn,* p=0.008 and *fabp11a,* p=0.001), cholesterol synthesis (*hmgcrb,* p=0.011) and leptin (*lep,* p=0.005) (**Figure 5.12A**). However, these results were not replicated on repeating the experiment, when adipocyte differentiation was found to be upregulated (*cebpa,* p=0.004), as was the beta-oxidation of fatty acids (*hadhaa,* p=0.042). For the genes corresponding to the remaining processes there was no significant dysregulation, although there was a trend towards downregulation for lipid storage and upregulation for cholesterol synthesis and leptin (**Figure 5.12B**).

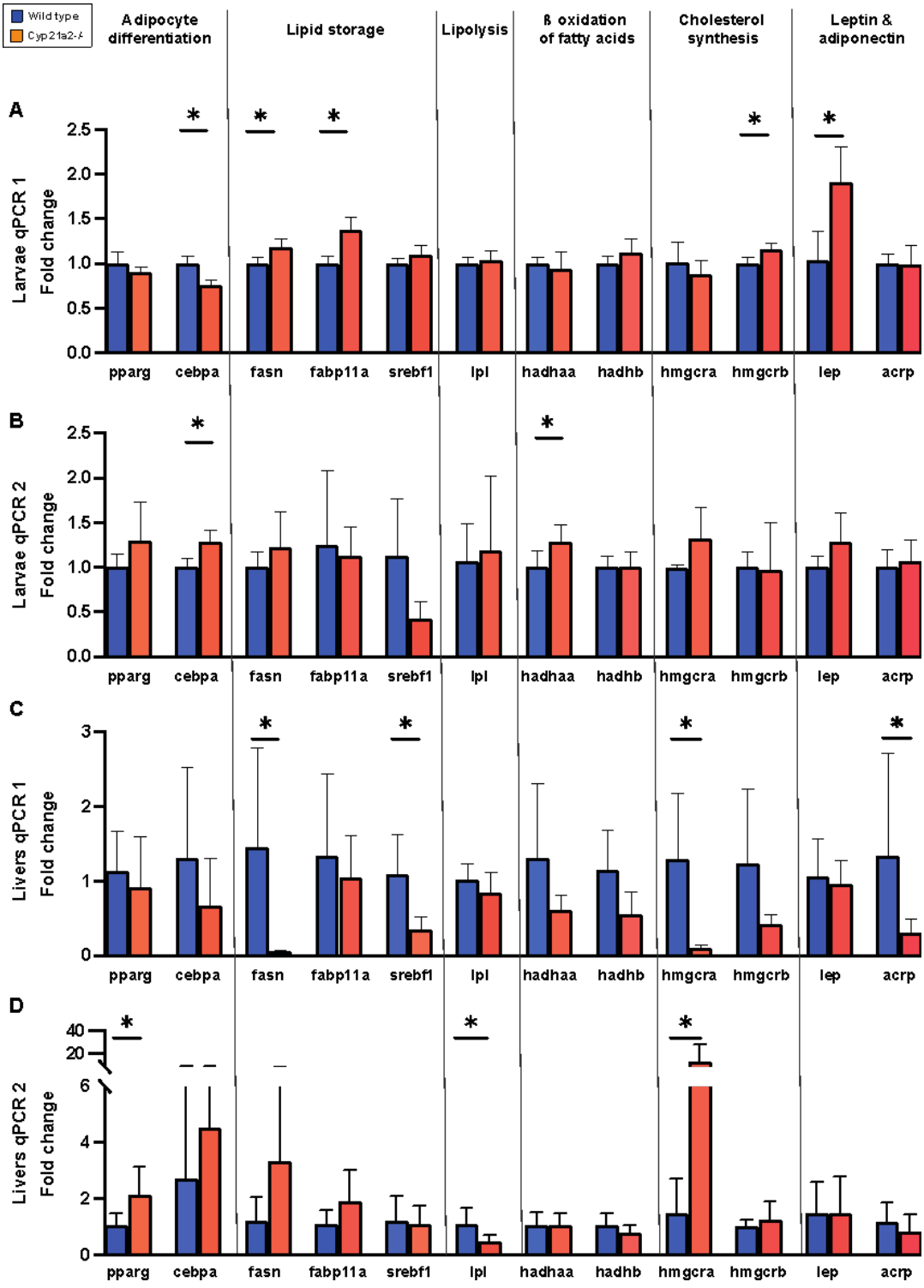


Figure 5.10. Fat metabolism qPCR

Differential expression of genes involved in fat metabolism in larvae (up: **A** and **B**, corresponding to the two separate experiments using different biological replicates, n=5 samples for each group) and adult livers (down: **C** and **D**, corresponding to the two separate experiments using different biological replicates, for **C**: n=5 livers for each group, for **D**: n=9 for WT and n=10 for mutant livers). Fold changes in *cyp21a2*-/- fish (blue) were expressed in relation to WT (orange). The dotted lines separate genes based on the

metabolic processes they regulate, which are indicated above. (* indicates statistical significance). The error bars mark the standard deviation.

The first experiment conducted on adult livers, showed lipid storage to be downregulated in mutants (*fasn,* p=0.047 and *srebf1,* p=0.032), as was cholesterol synthesis (*hmgcr,* p=0.038). Adiponectin was upregulated (p=0.015) and there was a trend towards downregulation for the genes involved in adipocyte differentiation, lipolysis and beta-oxidation of fatty acids, however, there was no statistical significance in these findings. (**Figure 5.12C**). The replication of the experiment yielded almost opposite results, showing upregulation of adipocyte differentiation (*pparg,* p=0.009) and cholesterol synthesis (*hmgcra,* p=0.011) and downregulation of lipolysis (*lpl,* p=0.004). There was a trend towards upregulation for lipid storage with no statistical significance (**Figure 5.12D**). There were no marked differences between males and females regarding the dysregulation of hepatic fat metabolism based on the second qPCR experiment.

5.2.2 Discussion

The study of differential gene expression by qPCR was marked by several inconsistencies when attempting to replicate the experiments, especially in the case of genes involved in the fat metabolism. Notably, comparing the qPCR results between the first and second experiments, in both larvae and livers, the general distribution of the data was overall similar, despite the inconsistencies found in the statistical significance. However, the variations seen in the results make it difficult to draw clear conclusions. Possible causes related to the methodology include differences in the sample collection and technical errors. As the procedures related to collecting and sorting the fish embryos, as well as collection on day 5 pf are standardised within the zebrafish facility, this represents an unlikely source of inconsistency. Furthermore, the collection of the adult

livers was conducted by the same researcher at similar times and using the same technique between the two time points, as were the RNA extraction and cDNA synthesis. Technical errors (such as calculation of reagents, pipetting) remain a possibility, however, they were monitored by the use in each qPCR experiment of three technical replicates, which monitored the quality of pipetting. Assuming that no errors related to technique or methodology occurred, the alternative interpretation would be that there are natural inter-individual variations in the timing and direction of the metabolic dysregulations caused by GC deficiency. This is further supported by the significant variance of gene expression observed in adult fish within each group of individuals pertaining to the same genotype and sex.

For genes related to **glucose metabolism** the main and most consistent finding was the downregulation of gluconeogenesis in *cyp21a2-/-* larvae. Previous research evidence from adrenalectomized mice (Blair, et al., 1995; Bishayi and Ghosh, 2003) and data from patients with adrenal insufficiency (Christiansen, Det al., 2007) also demonstrated suppressed gluconeogenesis. The downregulation of gluconeogenesis was also found in the adult livers, however, in the second experiment, this was only true for females. Similarly, *irs2* was only downregulated in the livers of female *cyp21a2-/-* mutants, raising the question of potential insulin resistance, although there was no upregulation of insulin which would be expected in that case. It is also important to note that no significant differences were found in blood sugar between mutants and WT, however, the glucose measurement were not conducted in a fasted state, taking place between 1-3 hours after feeding. Moreover, these results would indicate a sex-dependent metabolic response to GC deficiency which, coupled with the findings related to liver size, would support the hypothesis that there are differences in the hepatic metabolic processes between males

and females in zebrafish. The expression of *pck2* in mutant livers was comparable to WT. *Pck2* is the mitochondrial isomer of *pck1*, which is a cytosolic enzyme. While *pck1* is regulated mainly by endocrine factors, in particular, insulin, glucocorticoids and cAMP (O'Brien, et al., 1990), it is known that the roles of *pck2* go beyond gluconeogenesis and its expression is constitutive (Modaressi, et al., 1998). This is a likely reason for its variable expression between different biological samples.

In relation to **fat metabolism**, reduced adipogenesis was described by previous animal studies as a consequence of GC deficiency (Morton, et al., 2004). Importantly, it was shown that in human preadipocytes exposure to GC promotes differentiation, increasing their sensitivity to insulin (Tomlinson, et al., 2010); this raises the question of whether GC deficiency present at larval stage may cause abnormal adipose differentiation and subsequent insulin resistance at a later stage in the developing zebrafish organism. Unfortunately, the findings were not consistent between the two experiments, showing many conflicting results.

A trend towards upregulation of lipolysis, beta-oxidation of fatty acids, cholesterol synthesis and leptin were the only common findings between the two experiments in larvae, although in most cases lacking statistical significance. Arguably, the upregulation of lipolysis and beta-oxidation of fatty acids could be a compensatory mechanism related to energy metabolism in the context of reduced gluconeogenesis. Similarly, dysregulations of glucose metabolism may be primarily responsible for the upregulated leptin, given the previous evidence regarding its primary role in the glucose homeostasis in zebrafish (Michel, et al., 2016). The conflicting results between the two liver qPCR experiments render a holistic interpretation impossible. The second experiment used a

larger number of liver samples compared to the first one (10 mutants and 10 WT vs 5 mutants and 5 WT), thus making it likely to have yielded more robust findings. They would indicate an upregulation of adipocyte differentiation and cholesterol synthesis, as well as downregulation of lipolysis. These findings are in keeping with the increased fat mass found around the viscera and in the subcutaneous tissue in the mutant adults. This also resonates with human research involving large cohorts of individuals with non-clinical hypocortisolism who were found to have high prevalence of obesity and metabolic syndrome, in strong correlation with the cortisol deficiency (Kumari, et al., 2010; Maripuu, et al., 2016).

Overall, there were clear limitations to the use of qPCR as a method for exploring differential expression of metabolic processes in this project. A time-consuming technique, it yielded very few consistent results, indicating the need for a more robust and comprehensive approach to exploring metabolic dysregulations in the *cyp21a2-/-* zebrafish. This was addressed by the RNA sequencing based transcriptomic analysis described in the next chapter.

6. Transcriptomic analysis of *cyp21a2* function in zebrafish

As presented in the previous section, RT-qPCR provided limited insight into the dysregulations of metabolic pathways induced in zebrafish larvae and adult livers as a result of cortisol deficiency. The variability found on replicating the qPCR experiments prevented drawing robust conclusions, and consequently, a more robust and unbiased strategy was sought to explore the molecular mechanisms through which 21-hydroxylase deficiency impacts on metabolism. Thus, differential gene expression analysis following RNA sequencing was used to identify changes in the transcriptome of *cyp21a2-/-* larvae and adult livers. In mutant larvae the main findings consisted of the upregulation of genes involved in the cell cycle and cell proliferation, while genes involved in basic cell housekeeping, including protein synthesis and energy homeostasis, were downregulated. Genes associated with carbohydrate metabolism, including gluconeogenesis, glycolysis and glycogenesis, were consistently downregulated. In adult livers, processes related to mitochondrial organisation and provision of energy through ATP metabolism were significantly suppressed in *cyp21a2-/-* fish, as were genes implicated in protein synthesis and proteolysis. By contrast, there was a significant upregulation of genes involved in the inflammatory response, including components of the complement cascade and cytokine signalling pathways.

6.1 *Cyp21a2* deficiency causes marked downregulation of energy homeostasis and metabolic processes in larvae and adult livers

6.1.1 Results

6.1.1.1 Quality control of samples in larvae and livers

Overall, 14 files were examined corresponding to pair-end reads from seven samples (four wild type and three *cyp21a2-/-* mutant larvae). There were no major differences

between the quality of the RNA-sequencing data, assessed by FastQC and MultiQC, before and after the pre-processing sequence that included removal of rRNA and trimming. All samples had good per base sequence quality and per sequence quality scores, consistent with high quality data, however consistently failed the GC content and duplication percentage metrics. The figures shown below correspond to the final quality check.

For larvae, the sequencing depth ranged between 14.2 and 18.2 million mapped reads (**Table 6.1**); all the *cyp21a2-/-* larval samples had higher sequencing depth and a higher percentage of duplications compared to the wild type samples (**Figure 6.1A**). All samples had high per base sequence quality and per sequence quality scores. The per base sequence quality refers to the probability that the bases are correctly identified within the sequence, and it is expressed by the Phred score (**Figure 6.1B**). All samples had a Phred score above 30, corresponding to 1 error in every 1000 base calls or an accuracy of 99.9%. The histogram of the per sequence quality, corresponding to the distribution of Phred scores at read level, skewed to the right indicating that higher reads had higher Phred scores, an indicator of high-quality data (**Figure 6.1 C**). Similarly, the per base N content, indicating the number of indistinguishable bases was low, further confirming data quality (**Figure 6.1D**). All samples failed for per base sequence content, due to the bias in the first 12 bases, which is a characteristic feature of Illumina RNA sequencing (**Figure 6.1E, F**). All samples failed the per sequence GC content due to a consistent left skewing of the GC content (**Figure 6.1G**). The sequence duplication level was also failed by all samples, with mutant samples having higher duplication percentages, and a marked peak of level 4 duplications (**Figure 6.1H**). Before the pre-processing of raw data, six of the 14 files failed the overexpressed sequences module, however this was due mainly to the

presence of sequencing adapters which were removed by trimming data with Trimmomatic.

Table 6.1. Number of mapped reads for each larvae sample

Sample	Wild type 1	Wild type 2	Wild type 3	Wild type 4	Mutant 1	Mutant 2	Mutant 3
Million mapped reads	11.585	11.436	12.199	12.332	14.408	13.199	14.457

(Expressed as million reads)

The liver samples (six males and six females from both wild type and mutant adult fish) also met the criteria for high-quality sequencing data with regards to the per base sequence quality and per sequence quality scores, as well as the low per base N content. Adapter content and overexpressed sequences were to a high extent resolved by data trimming with Trimmomatic. The sequencing depth was much higher compared to the larval samples, ranging between 32 and 48 million reads (**Table 6.2**), with the exception of one mutant female sample (F6), which had relatively low sequencing depth (13 million reads) and was excluded from the analysis. There was also a higher duplication rate (amounting to 87%), however, due to the high sequencing depth, the number of unique reads was comparable to the larval samples. Similar to larval data, all samples failed for per base sequence content due to the bias in the first 12 bases and failed or triggered warning in the CG content module due to left skewing (**Figure 6.2)**.

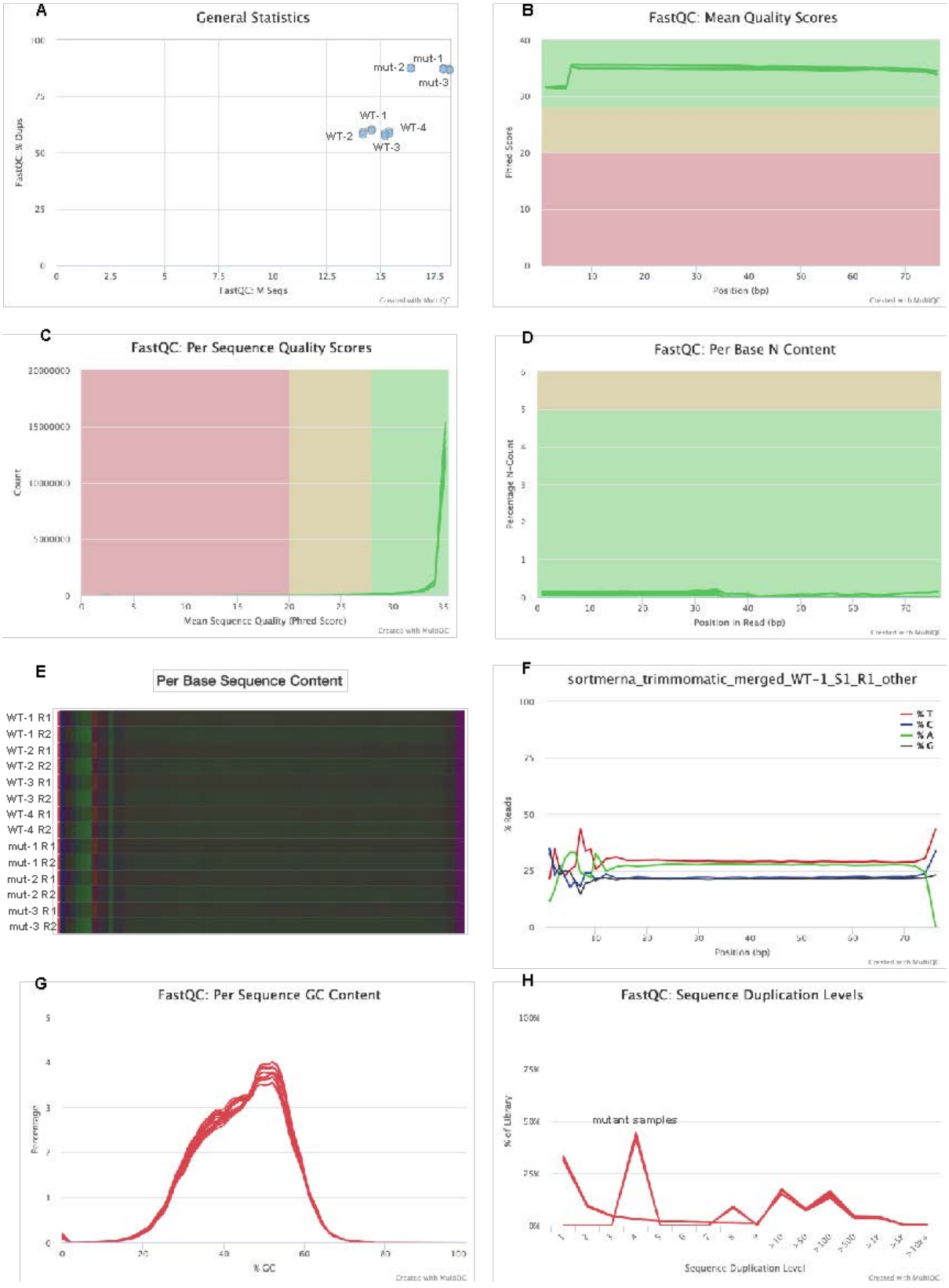


Figure 6.1. Quality of sequencing data in larvae

A. General statistics of the sequencing depth (expressed on the horizontal axis as million reads) plotted against the percentage of sequence duplication (on the vertical axis), showing that mutant samples (mut) had higher sequencing depth and more duplications compared to wild type (WT). **B**. Per base mean quality expressed as the Phred score within the read, showing consistently scores above 30, suggesting high-quality data. **C**. The per sequence quality histogram was shifted to the right towards higher Phred scores. **D**. All samples had very low per base N content (indistinguishable bases). **E**. The heatmap shows all samples failed the per sequence base content; this was due to the bias in the first 12 bases characteristic to Illumina sequencing (**F**), demonstrated in an individual sample (WT-1, R1). **G**. All samples failed the per sequence GC content check due to a consistent skewing to the left of the GC content curve. **H**. All samples failed the sequence duplication level module, with mutant samples presenting 4-level duplication for up to 40% of the library.

Table 6.2. Number of mapped reads for each liver sample

Sample	Wild type 1	Wild type 2	Wild type 3	Wild type 4	Wild type 5	Wild type 6
Million mapped rea	28.333	31.308	31.855	30.101	28.402	30.060
Sample	**Mutant 1**	**Mutant 2**	**Mutant 3**	**Mutant 4**	**Mutant 5**	
Million mapped rea	41.1833	32.726	30.536	31.788	30.230	

(Expressed as million reads)

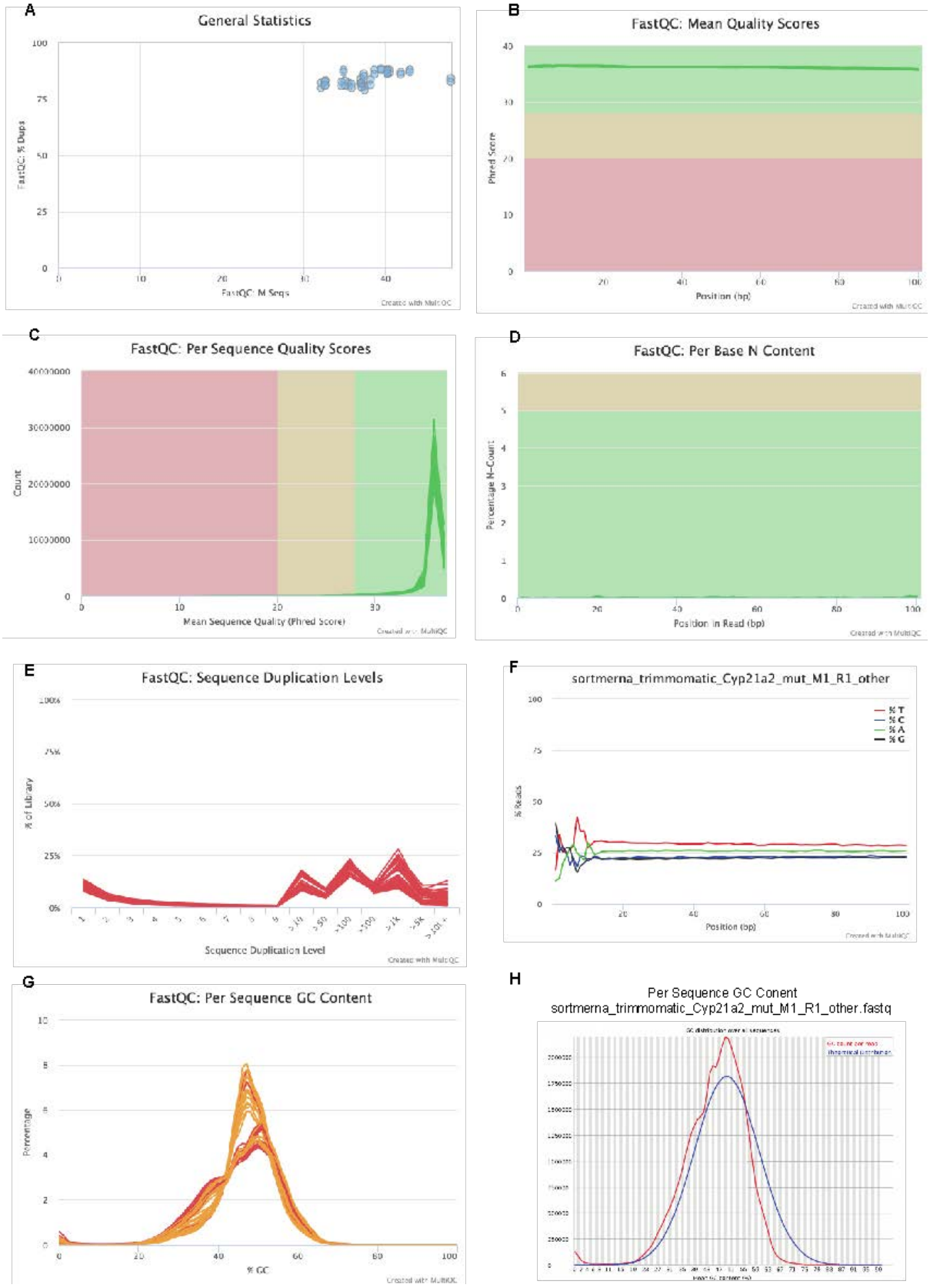


Figure 6.2. Quality of sequencing data in adult livers
A. General statistics of the sequencing depth (expressed on the horizontal axis as million reads) plotted against the percentage of sequence duplication (on the vertical axis). **B**.

Per base mean quality expressed as the Phred score within the read, showing consistently scores above 30, suggesting high-quality data. **C**. The per sequence quality histogram was shifted to the right towards higher Phred scores. **D**. All samples had very low per base N content (indistinguishable bases). **E**. All samples failed the sequence duplication level module. **F**. Individual per sequence base content (Mutant Male 1) failed due to the bias in the first 12 bases characteristic to Illumina sequencing. **G**. All samples alarmed or failed the per sequence GC content check due to a consistent skewing to the left of the GC content curve. **H**. Per sequence GC content in an individual sample (Mutant Male 1).

6.1.1.2 Transcriptomic profile of *cyp21a2-/-* larvae

Principal component analysis (PCA) showed clear distinction between wild-type and mutant larvae, with obvious clustering according to the first principal component, which accounted for 80% of variance (**Figure 6.3.A**). The cluster dendrogram (**Figure 6.3.B**) and heatmap of normalised counts (**Figure 6.3.C**) demonstrated good homogeneity of the group of wild type samples, while the third of the mutant samples ("cyp3") was different from the other two. This could not be explained by differences in sequencing depth as the size factors were comparable among the three mutant samples (**Figure 6.3.D**).

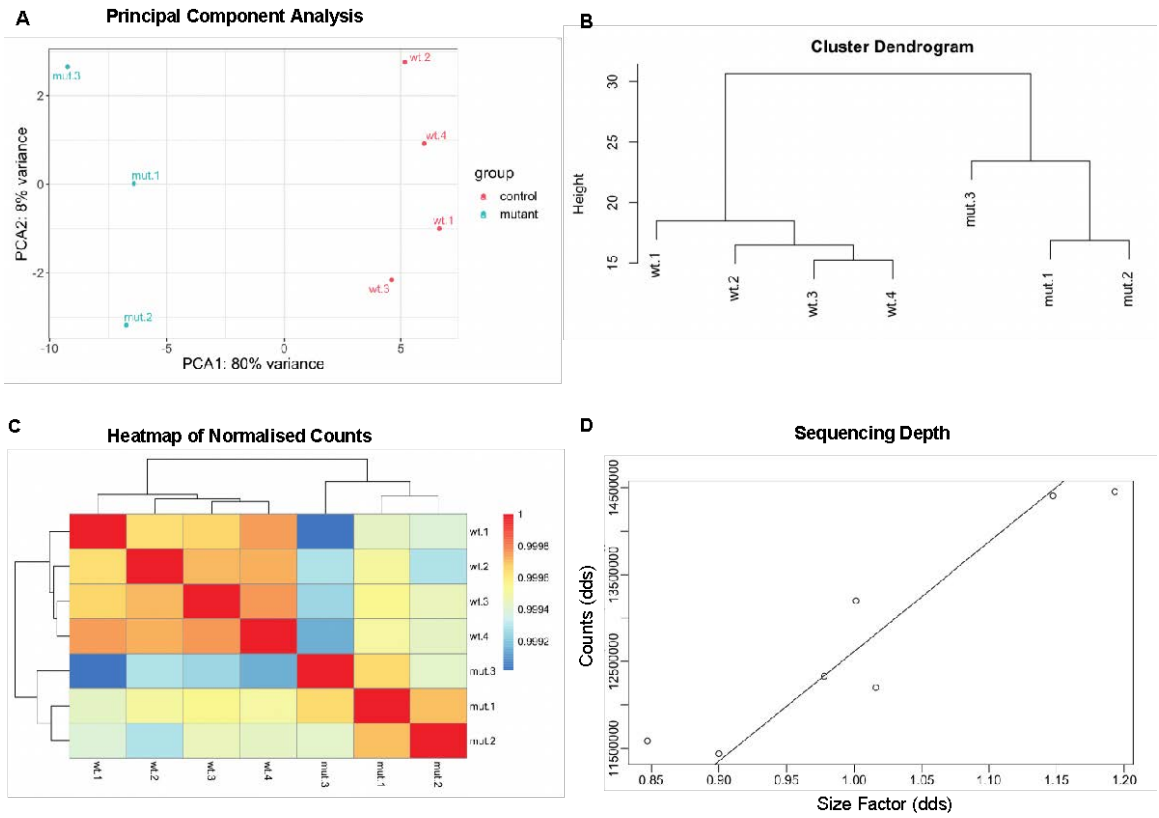
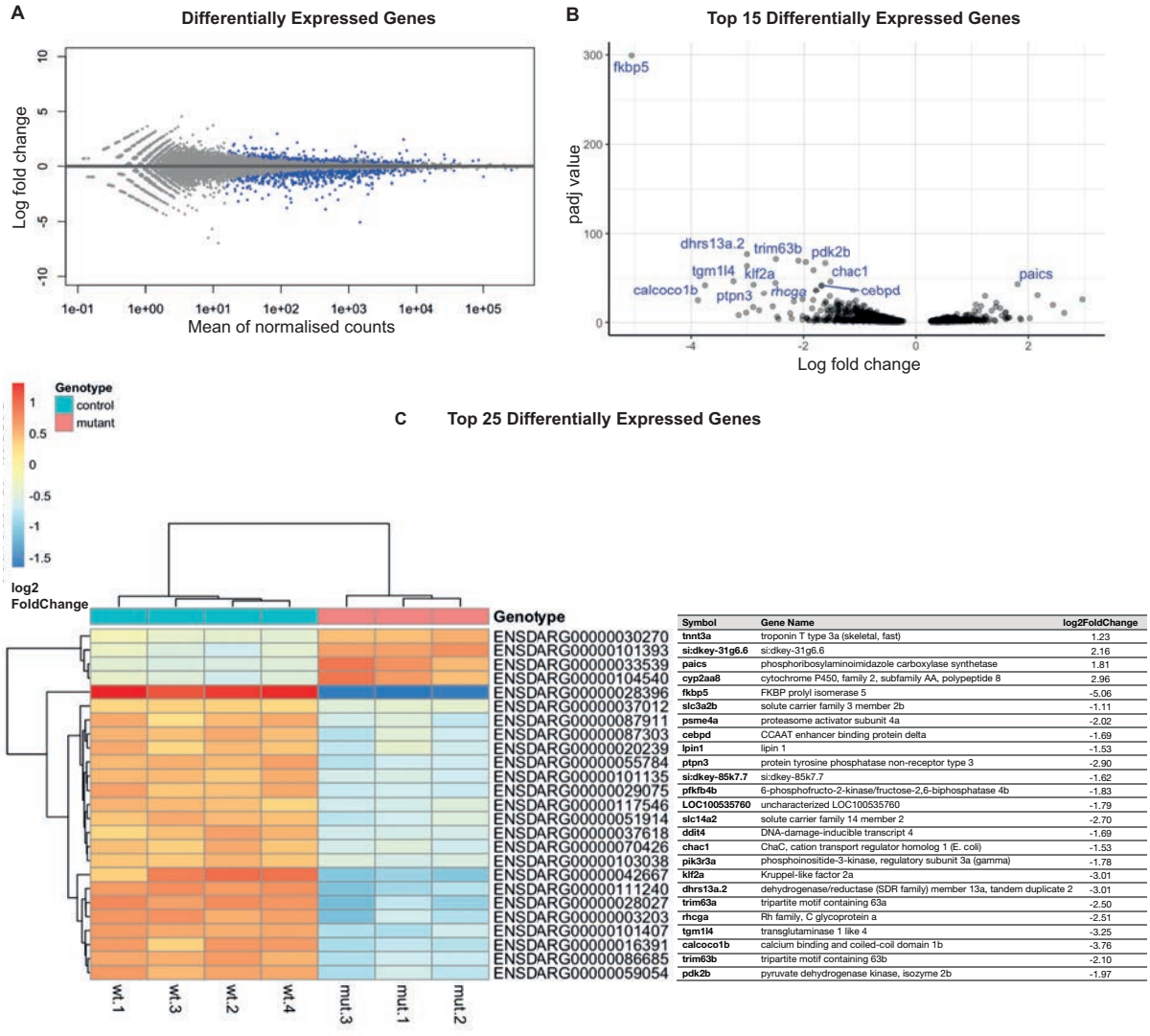


Figure 6.3. Sample variation in *cyp21a2-/-* mutant larvae
Inter-sample variations in differential gene expression in wild type and *cyp21a2-/-* mutant larvae. **A.** Principal Component Analysis; **B**. Cluster dendrogram; **C**. Heatmap of normalised counts; **D**. Sequencing depth expressed by plotting column counts for each sample against the size factor. (mut: *cyp21a2-/-* mutants, wt: wild type).

Differential gene expression (DGE) analysis identified 398 genes that were significantly upregulated (log2 fold change (LFC) > 0, *padj* < *0.05*) and 698 genes significantly downregulated (LFC < 0, *padj* < *0.05*) (**Figure 6.4.A**). The most dysregulated genes are shown in the volcano plot (**Figure 6.4.B**), with *fkbp5* being distinctly the most downregulated gene (LFC = -5.05, padj = 5.92e-300). There was a very marked distinction between wild type and mutant larvae in the expression of the top 25 DEGs, among which the majority were downregulated in *cyp21a2-/-* larvae (**Figure 6.4.C**).



Symbol	Gene Name	log2FoldChange
tnnt3a	troponin T type 3a (skeletal, fast)	1.23
si:dkey-31g6.6	si:dkey-31g6.6	2.16
paics	phosphoribosylaminoimidazole carboxylase synthetase	1.81
cyp2aa8	cytochrome P450, family 2, subfamily AA, polypeptide 8	2.96
fkbp5	FKBP prolyl isomerase 5	-5.06
slc3a2b	solute carrier family 3 member 2b	-1.11
psme4a	proteasome activator subunit 4a	-2.02
cebpd	CCAAT enhancer binding protein delta	-1.69
lpin1	lipin 1	-1.53
ptpn3	protein tyrosine phosphatase non-receptor type 3	-2.90
si:dkey-85k7.7	si:dkey-85k7.7	-1.62
pfkfb4b	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 4b	-1.83
LOC100535760	uncharacterized LOC100535760	-1.79
slc14a2	solute carrier family 14 member 2	-2.70
ddit4	DNA-damage-inducible transcript 4	-1.69
chac1	ChaC, cation transport regulator homolog 1 (E. coli)	-1.53
pik3r3a	phosphoinositide-3-kinase, regulatory subunit 3a (gamma)	-1.78
klf2a	Kruppel-like factor 2a	-3.01
dhrs13a.2	dehydrogenase/reductase (SDR family) member 13a, tandem duplicate 2	-3.01
trim63a	tripartite motif containing 63a	-2.50
rhcga	Rh family, C glycoprotein a	-2.51
tgm1l4	transglutaminase 1 like 4	-3.25
calcoco1b	calcium binding and coiled-coil domain 1b	-3.76
trim63b	tripartite motif containing 63b	-2.10
pdk2b	pyruvate dehydrogenase kinase, isozyme 2b	-1.97

Figure 6.4. Differential gene expression in *cyp21a2-/-* mutant larvae
A. MA plot showing differentially expressed genes; the statistical significance (*p-adj* < *0.05*) is indicated by the blue dots while differentially expressed genes without statistical significance are coloured grey. **B.** Volcano plot showing the top 25 DEGs in *cyp21a2-/-* larvae, with statistical significance (*p-adj)* represented on the vertical axis and the LFC on the horizontal one. **C**. Heatmap and table indicating the top 25 DEGs in *cyp21a2-/-* larvae.

6.1.1.2.1 Comparison between the differential expression of essential metabolic genes explored by qPCR and RNA sequencing in *cyp21a2-/-* larvae

Of the genes chosen to explore glucose and fat metabolism by qPCR of larval RNA samples, the most consistent finding across the two qPCR experiments and the RNA sequencing analysis was the downregulation of *pck1,* a cortisol response gene involved in gluconeogenesis. For the rest of the genes involved in metabolism, the results were

inconsistent, not only between qPCR and RNA sequencing but also between the two qPCR experiments. These inconsistencies related to differences in the direction of the dysregulation as well as their statistical significance (**Table 6.3**). Importantly, the lipoprotein lipase gene (*lpl*) was very significantly downregulated in the RNA sequencing analysis, which was not found using qPCR.

Table 6.3. RNA sequencing vs qPCR results in *cyp21a2* larvae

Genes involved in glucose metabolism						
	1st qPCR experiment		**2nd qPCR experiment**		**RNA sequencing**	
Gene	**Mean fold change**	***p* value**	**Fold change**	***p* value**	**Log2 fold change**	***p-adj***
pck1	0.33	< 0.001*	0.31	< 0.001*	-0.75	3.57e-05*
pck2	0.66	0.008*	1.00	0.946	0.19	0.586
pkl	0.99	0.823	0.78	0.110	0.16	0.440
pkm	0.95	0.150	1.07	0.309	-0.02	0.938
insa	0.63	0.001*	1.00	0.983	0.15	n/a
irs2	1.04	0.481	0.97	0.746	-0.53	0.030*
Genes involved in fat metabolism						
	1st qPCR experiment		**2nd qPCR experiment**		**RNA sequencing**	
Gene	**Mean fold change**	***p* value**	**Fold change**	***p* value**	**Log2 fold change**	***p-adj***
pparg	0.90	0.150	1.29	0.226	0.03	n/a
cebpa	0.76	0.007*	1.28	0.007*	0.18	0.582
fasn	1.18	0.008*	1.22	0.323	-0.04	0.890
fabp11a	1.38	0.001	1.12	0.766	0.13	0.584

srebf1	1.10	0.098	0.42	0.046*	0.14	0.902
lpl	1.04	0.458	1.19	0.781	-0.64	1.75e-07*
hadhaa	0.94	0.560	1.28	0.042*	0.13	0.734
hadhb	1.12	0.198	1.01	0.968	-0.08	0.766
hmgcra	0.88	0.299	1.32	0.097	0.56	0.026*
hmgcrb	1.15	0.011*	0.97	0.876	0.03	0.937
lep	1.83	0.007*	1.24	0.111	0.87	n/a
acrp	0.96	0.841	1.04	0.690	0.10	n/a

The results of the qPCR experiments are expressed as the mean of the relative fold change (with 1 being the threshold for up- vs downregulation) and the p value corresponding to the t test (n=5 WT and 5 mutant samples). The results of the RNA sequencing data are expressed as log2 fold change (the threshold for up- vs downregulation is 0) and the adjusted p value (n=4 WT and 3 mutant samples). * Indicates statistical significance.

6.1.1.2.2 Gene ontology enrichment analysis

Gene ontology (GO) enrichment analysis is a technique used to study biological pathways and processes that are significantly dysregulated in a set of genes provided, using an arbitrary established cut-off for statistical significance (in this case 10^{-5}). GO enrichment was initially performed using the GOrilla (Gene Ontology Enrichment Analysis and Visualisation). Of the 12 GO terms identified, the majority pertained to metabolic processes, while three related to cell cycle processes (**Table 6.4 and Figure 6.5.A**). Thus, GO terms corresponding to the most dysregulated metabolic processes were small molecule metabolism, organic acid metabolism, carboxylic acid metabolism, oxoacid metabolism, drug metabolic process, catabolic process, alpha-amino acid metabolism, carboxylic acid biosynthesis and metabolic process. Similar results were obtained when conducting gene ontology analysis in R using the “clusterProfiler” package, however, here

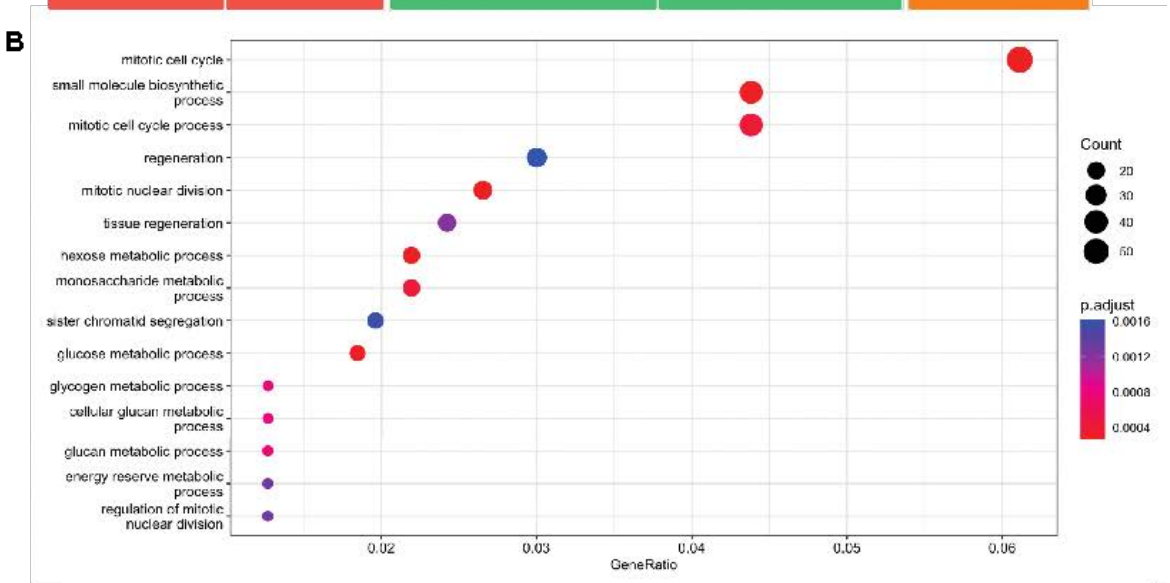
the overexpressed biological processes were more equally divided between those pertaining to cell cycle and metabolism (**Figure 6.5.B**).

Table 6.4. Results of the gene ontology analysis for biological processes in GOrilla for cyp21a2 larvae

GO Term	Description	Enrichment (N,B,n,b)	p-value	FDR q-value
GO:0044281	Small molecule metabolic process	1.7 (14317, 920, 1205, 132)	6.58e-09	5.65e-05
GO:0006082	Organic acid metabolic process	1.95 (14317, 506, 1205, 83)	2.44e-08	1.05e-04
GO:1903047	Mitotic cell cycle process	2.35 (14317, 217, 1429, 51)	2.87e-08	8.20e-05
GO:0019752	Carboxylic acid metabolic process	2 (14317, 445, 1205, 75)	4.03e-08	8.64e-05
GO:0022402	Cell cycle process	2.02 (14317, 338, 1387, 66)	1.76e-07	3.02e-04
GO:0043436	Oxoacid metabolic process	1.93 (14317, 469, 1205, 76)	1.78e-07	2.55e-04
GO:0017144	Drug metabolic process	2.27 (14317, 276, 1074, 47)	6.62e-07	8.12e-04
GO:0051726	Regulation of cell cycle	2.19 (14317, 293, 1115, 50)	7.99e-07	8.57e-04
GO:0009056	Catabolic process	2.01 (14317, 962, 444, 60)	1.98e-06	1.89e-03
GO:1901605	Alpha-amino acid metabolic process	2.61 (14317, 141, 1205, 31)	2.67e-06	2.29e-03
GO:0046394	Carboxylic acid biosynthetic process	2.42 (14317, 168, 1162, 33)	8.07e-06	6.29e-03
GO:0008152	Metabolic process	1.16 (14317, 4856, 1623, 641)	8.39e-06	6.00e-03

The **p-value** is the enrichment p-value computed according to the mHG or HG model. The **FDR q-value** is the correction of the p-value for multiple testing using the Benjamini and Hochberg (1995) method. The **Enrichment = (b/n) / (B/N) (**N - total number of genes, B - total number of genes associated with a specific GO term, n – a flexible cut-off, being an automatically determined number of genes in the input list or 'target set', b - is the number of genes in the target set that are associated with the GO term).


A
metabolic process
xenobiotic metabolic process
regulation of cell cycle
alpha-amino acid metabolic process
organic acid metabolic process
carboxylic acid biosynthetic process
small molecule metabolic process
mitotic cell cycle process
cell cycle process
carboxylic acid metabolic process
oxoacid metabolic process
catabolic process


B
mitotic cell cycle
small molecule biosynthetic process
mitotic cell cycle process
regeneration
mitotic nuclear division
tissue regeneration
hexose metabolic process
monosaccharide metabolic process
sister chromatid segregation
glucose metabolic process
glycogen metabolic process
cellular glucan metabolic process
glucan metabolic process
energy reserve metabolic process
regulation of mitotic nuclear division
0.02
0.03
0.04
0.05
0.06
GeneRatio
Count
20
30
40
50
p.adjust
0.0016
0.0012
0.0008
0.0004

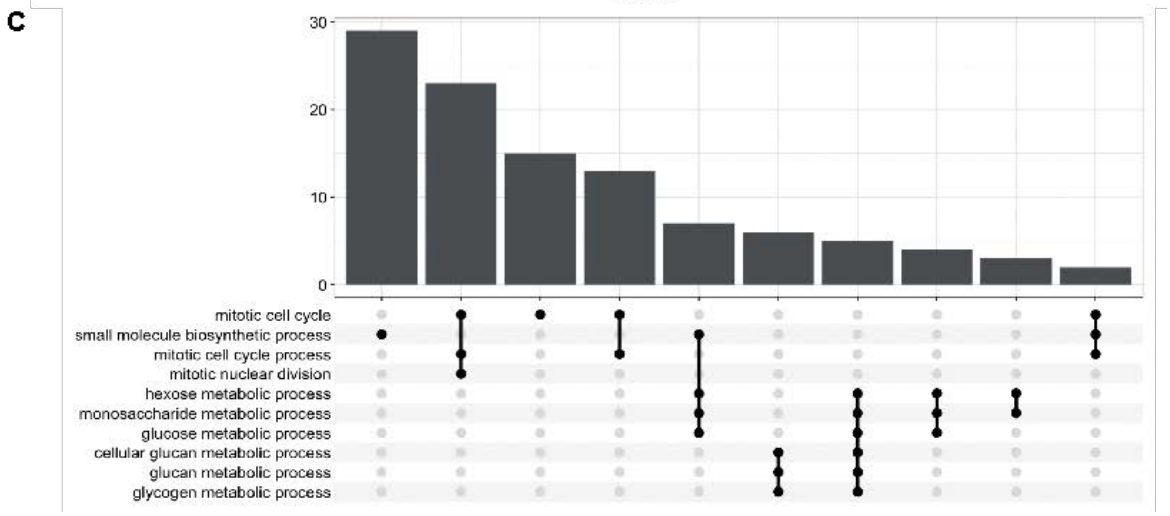

C
30
20
10
0
mitotic cell cycle
small molecule biosynthetic process
mitotic cell cycle process
mitotic nuclear division
hexose metabolic process
monosaccharide metabolic process
glucose metabolic process
cellular glucan metabolic process
glucan metabolic process
glycogen metabolic process

Figure 6.5. Gene ontology over-expression analysis in *cyp21a2-/-* larvae
A. Tree map produced in REVIGO showing a graphic representation of the biological processes that are identified to be dysregulated in *cyp21a2-/-* larvae following the gene ontology analysis conducted in GOrilla. The size of the rectangles is given by the p-value. **B**. Enrichment plot produced in R using the "clusterProfiler" package. The horizontal axis shows the ratio between the number of DEGs in a category and all DGEs, the size of the dots represents the number of DEGs in a category and the colour, the *p-adj* value. **C.** Upset plot showing the overlap of differentially expressed genes between the overexpressed biological processes identified by the GO enrichment analysis. The dotted vertical lines indicate the processes that are included in the analysis and the height of the bars corresponds to the number of overlapping genes.

To explore the association between genes involved in the different overexpressed biological processes, the "upsetplot" function was used in R "clusterProfiler". This showed important overlap in the differentially expressed genes involved in the mitotic cell cycle and mitotic nuclear division processes, with more modest overall in the metabolic processes (**Figure 6.5.C**).

Small molecule metabolism was the most significantly dysregulated metabolic process identified by both methods. Network visualisation in STRING of the genes pertaining to this GO term that were found in the analysed set showed that they were relatively evenly divided between lipid, carbohydrate, amino acid, and aromatic compound (including nucleotide biosynthesis) metabolism (**Figure 6.6**).

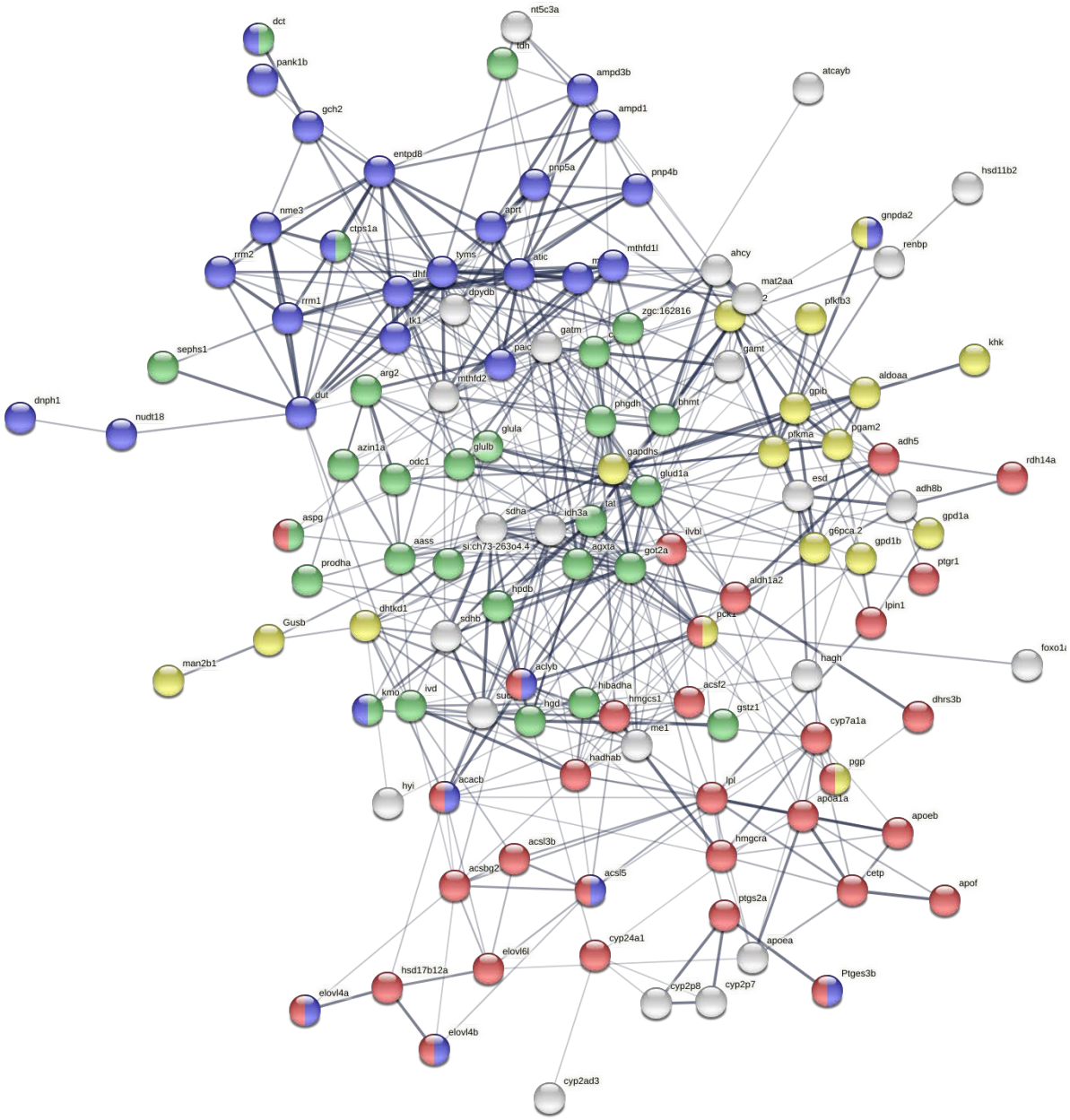


Figure 6.6. Small molecule metabolism in *cyp21a2-/-* larvae
Visualisation in STRING of small molecule metabolism gene set identified following GO over-representation analysis in GOrilla. The connecting lines indicate interactions between proteins, the width of the lines being proportional to the confidence or the strength of data supporting the interaction. The colours of nodes the indicate the main sub-processes to which they belong: lipid metabolism (red, n=35), carbohydrate metabolism (yellow, n=17), alpha amino acid metabolism (green, n=29), aromatic compound biosynthesis, including nucleotide biosynthesis (blue, n=32), and also response to insulin (pink, n=5).

6.1.1.2.3 Gene set enrichment analysis

Gene set enrichment analysis (GSEA) represents another way of exploring enriched pathways and biological processes for a given gene set. The main distinction to the GO overexpression analysis consists in the absence of an established cut-off for statistical significance in GSEA, where the whole list of ranked genes is analysed against pre-existing organism specific genome wide annotation. In this case, the genes were ranked based on their log2FoldChange, thus the enrichment score related to the degree of dysregulation, with negative scores corresponding to downregulated and positive scores to upregulated processes. The GSEA results were to an extent different from those of the GO enrichment analysis, with processes related to cell division, proliferation and the functioning of cell organelles being found to be the most dysregulated (**Figure 6.7 and 6.8**). Thus, the most upregulated processes included those involved in mitosis and meiosis, such as DNA replication and chromosome segregation and separation, nuclear division, as well as eye development and sensory perception. The most significant downregulation was found in the ribosome synthesis and function, autophagy including organelle disassembly, as well as response to stimuli such as peptide hormones, starvation, temperature, reactive oxygen species, and the positive regulation of white cell proliferation.

The GSEA results relate to biological processes that are overexpressed in mutants, without providing numeric outputs related to the expression of individual genes. However, the genes in the target set associated with each GO term are specified and they were analysed against the results of the differential gene expression (DGE) analysis. Consequently, in the following presentation of the GSEA results, the values provided in

brackets for specific genes were taken from the differential gene expression (DGE) analysis.

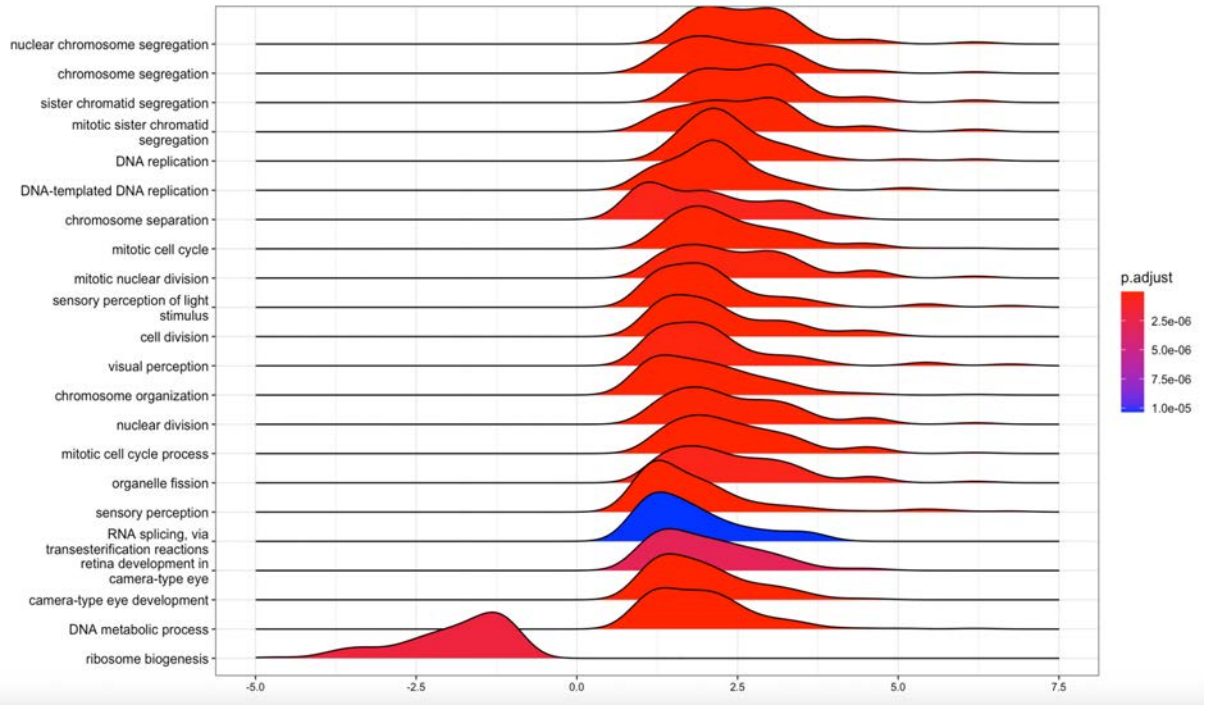


Figure 6.7. Gene set enrichment analysis in *cyp21a2-/-* larvae
The top 20 dysregulated biological processes found by the GSEA. Ridgeplot produced in R using the clusterProfiler' package. The processes were selected automatically based on the normalised enrichment score (showed on the horizontal axis) and the size of the gene set, ie. sets with small sizes were excluded despite having higher enrichment scores. The x axis indicates the normalised enrichment score (NES), related to the degree of upregulation (NES > 0) or downregulation (NES < 0) of each biological process. The shape of the ridges corresponds to the distribution of the genes including in the set, in relation to the enrichment score and the colour indicates the adjusted p-value, decreasing from red to blue.

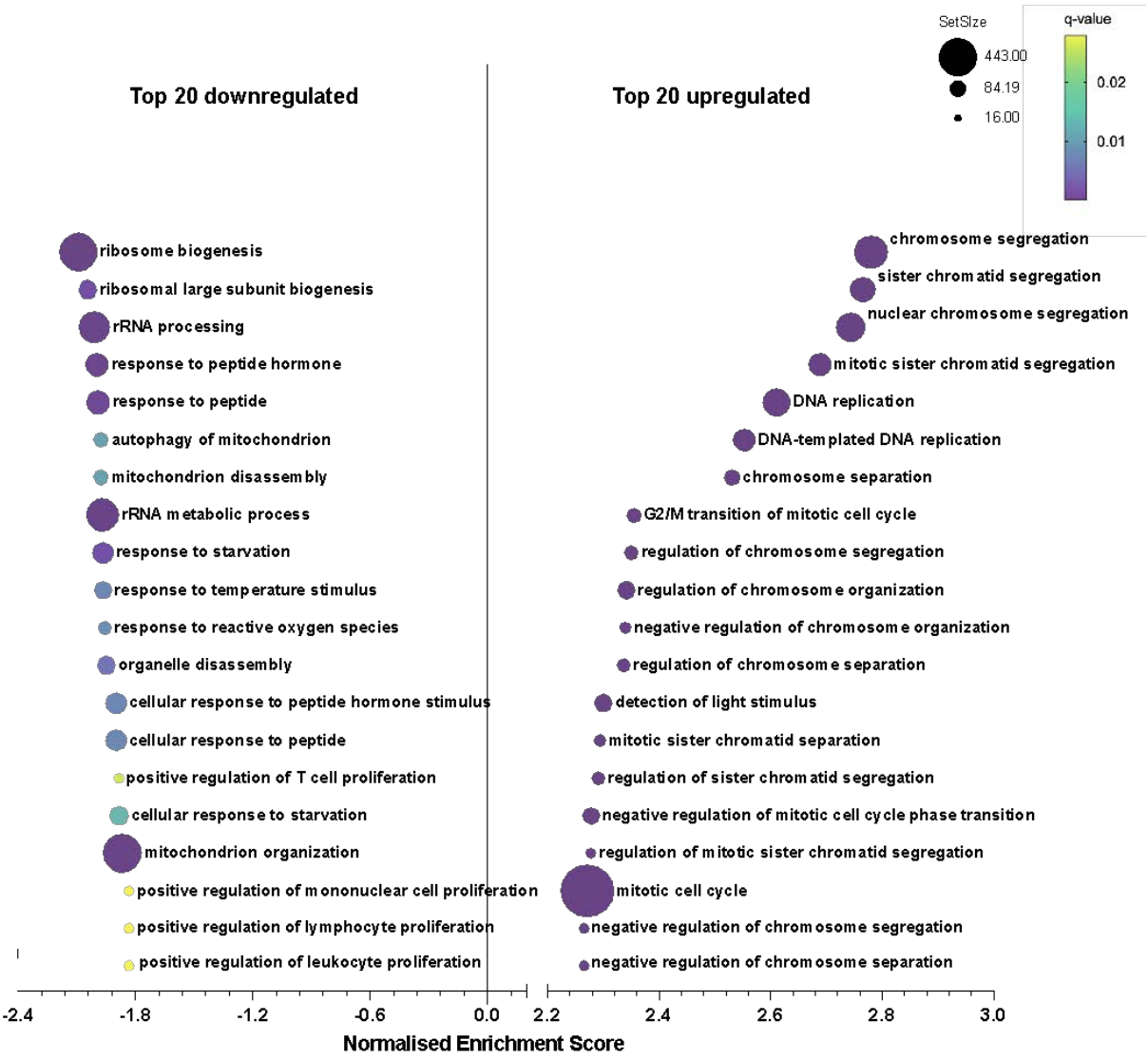


Figure 6.8. Gene set enrichment analysis in cyp21a2-/- larvae (2)
The top 20 up- and down-regulated biological processes found by the GSEA. Bubbleplot produced GraphPad Prism showing the 20 most downregulated (left) and upregulated (right) biological processes in *cyp21a2* larvae. The selection was based exclusively on the normalised enrichment score (showed on the horizontal axis). The size of the dots indicates the size of the gene set and the colour corresponds to the q-value, decreasing from yellow to violet.

The most down-regulated processes represented subclasses of two large groups, ribosomal biosynthesis and function and mitochondrion organisation. For ribosomal biosynthesis important sub-processes included the biosynthesis of the large and small ribosomal units, with a more modest number of genes being involved in translation (**Figure 6.9**). Although not identified amongst the gene list included in the sets related to

the ribosome function and regulation, the eukaryotic translation initiation factor 4E binding protein 3 (*eif4ebp3*) was significantly downregulated in the DGE analysis (log2Fc = -0.95, $p < 0.00001$).

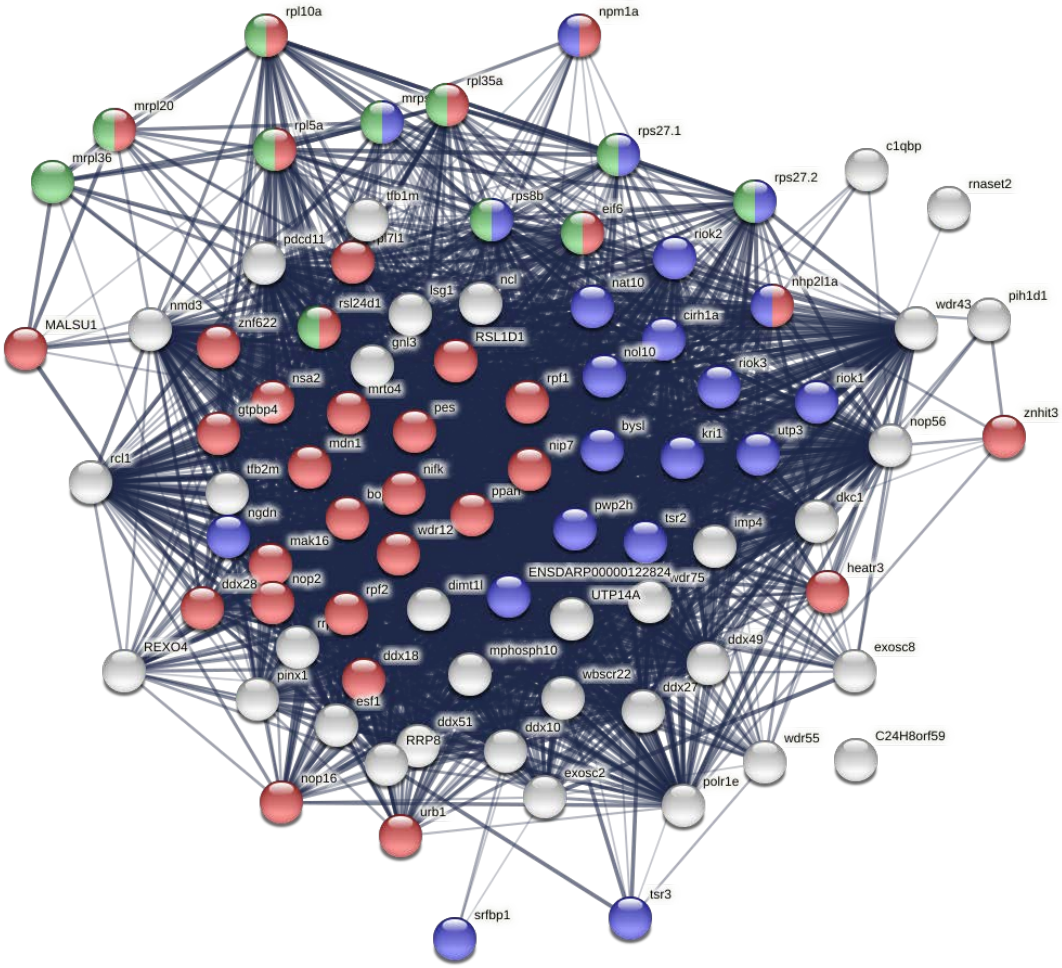


Figure 6.9. Ribosome biosynthesis in *cyp21a2-/-* larvae
Visualisation in STRING of dysregulated genes involved in ribosome biosynthesis identified by GSEA. The connecting lines indicate interactions between proteins, the width of the lines being proportional to the confidence or the strength of data supporting the interaction. The colours of the nodes indicate the main sub-processes to which they belong: large unit biosynthesis (red, n=32), small unit biosynthesis (blue, n=21), and translation (green, n=11).

Exploring the dysregulated genes involved in mitochondrion organisation by network visualisation in STRING (**Figure 6.10**), it was found that the downregulation affected different sub-processes, including autophagy and catabolism, membrane organisation

and protein transport, respiratory chain complex assembly and oxidation-reduction process. Of note, the cytoprotective protein bcl2, previously described to interact with mitochondrial GR, thus being involved in the GC role in regulating cell plasticity (DU, 2009), was not included in the gene set following GSEA; however, the initial DGE analysis identified three members of its family to be significantly downregulated: the bcl2 interacting protein 4 (*bnip4,* log2FC = -1.01, p < 0.00001), MCL1 apoptosis regulator, bcl2 family member a (*mcl1a*, log2FC = -0.63, p < 0.00001) and MCL1 apoptosis regulator, bcl2 family member b (*mcl1b*, log2FC = -0.53, p = 0.00002).

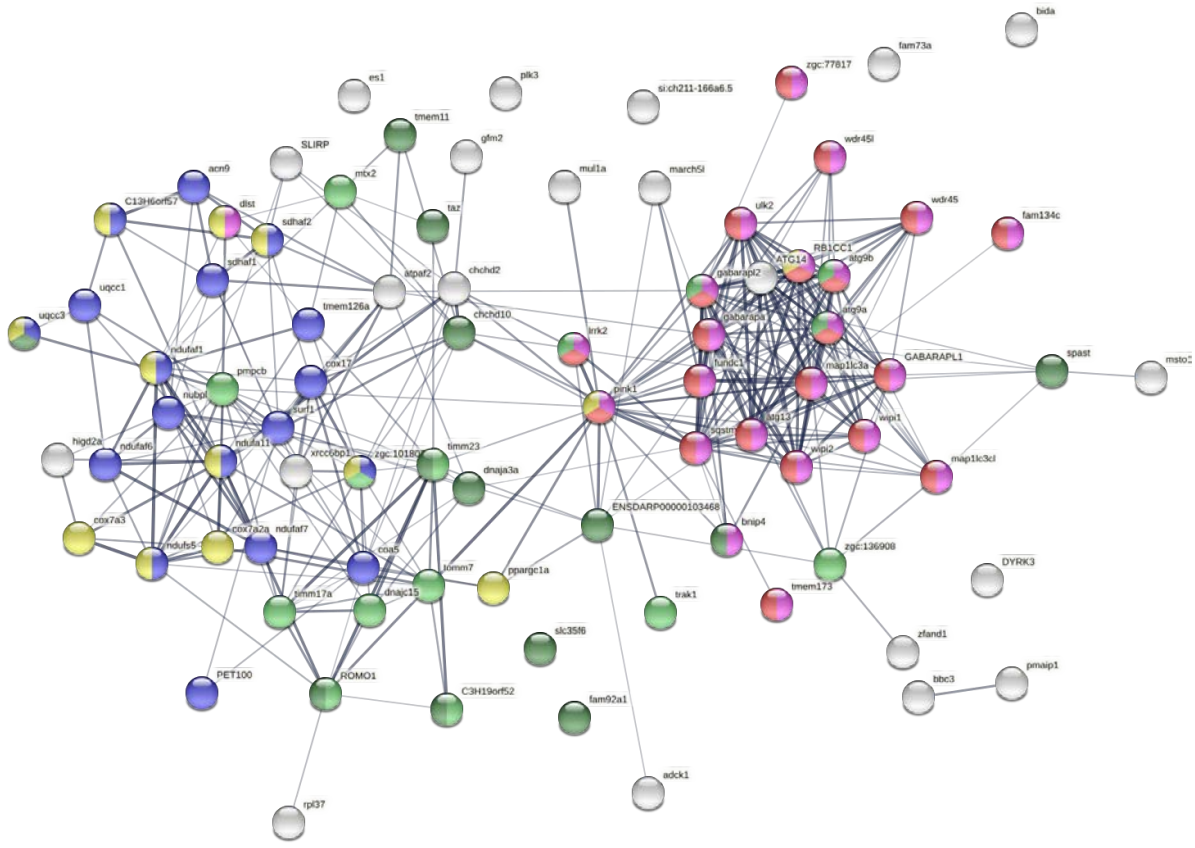


Figure 6.10. Mitochondrion organisation in *cyp21a2-/-* larvae
Visualisation in STRING of dysregulated genes involved in mitochondrion organisation identified by GSEA. The connecting lines indicate interactions between proteins, the width of the lines being proportional to the confidence or the strength of data supporting the interaction. The colours of the nodes indicate the main sub-processes to which they belong: autophagy (red, n=21), cellular catabolic process (pink, n=23), protein transport (light green, n=15), membrane organisation (dark green, n=13), respiratory chain complex assembly (blue, n=18), and oxidation-reduction process (yellow, n=13).

Significantly upregulated GO terms were dominated by the mitotic cell cycle and subjacent processes. Within this group there was upregulation of multiple interrelated genes involved in positive and negative regulation of cell cycle and DNA repair (**Figure 6.11**).

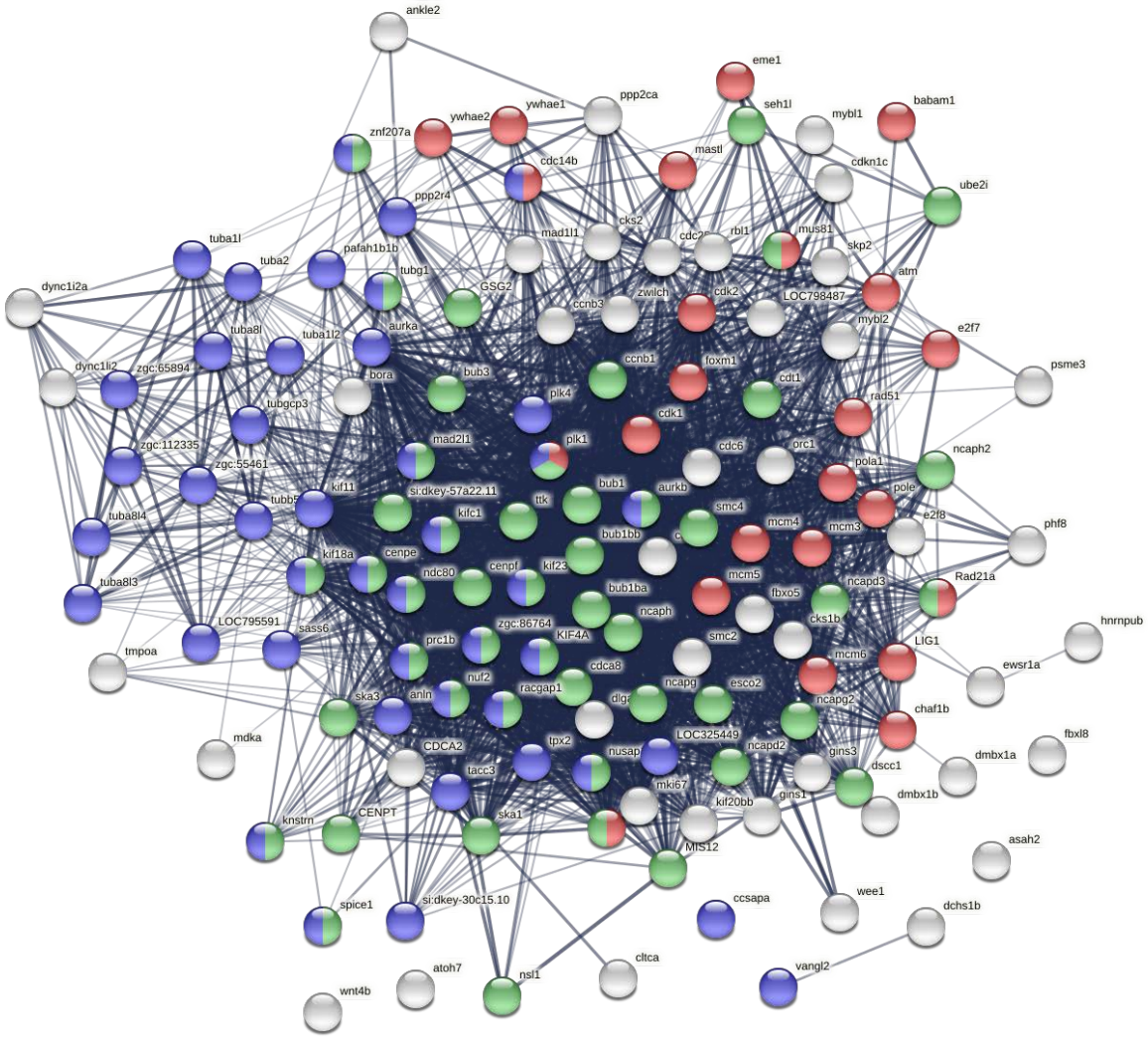


Figure 6.11. Mitotic cell cycle in *cyp21a2-/-* larvae
Visualisation in STRING of dysregulated genes involved in mitotic cell cycle identified by GSEA. The connecting lines indicate interactions between proteins, the width of the lines being proportional to the confidence or the strength of data supporting the interaction. The colours of the nodes indicate sub-processes to which they belong: cytoskeleton organisation (blue, n=44), chromosome segregation (green, n=48), and cellular response to stress (red, n=24).

From the GSEA output list the enriched gene sets related to metabolic processes and functions were manually selected (**Figure 6.12**). Many of the processes included in this list overlapped in the composition of their gene sets and were consequently grouped accordingly. Thus, the down-regulated processes appeared to be mainly related to response to peptide hormones (including insulin), response to nutrients and starvation, carbohydrate metabolism, and energy homeostasis including ATP metabolism and generation of energy by oxidation of organic compounds. The list of upregulated terms was more heterogeneous, consisting of metabolic processes of different compounds: cholesterol, alcohol, amine, isoprenoid, folic acid, tetrahydrofolate, steroids, organic hydroxy compounds, deoxyribonucleotides and carbohydrate derivates.

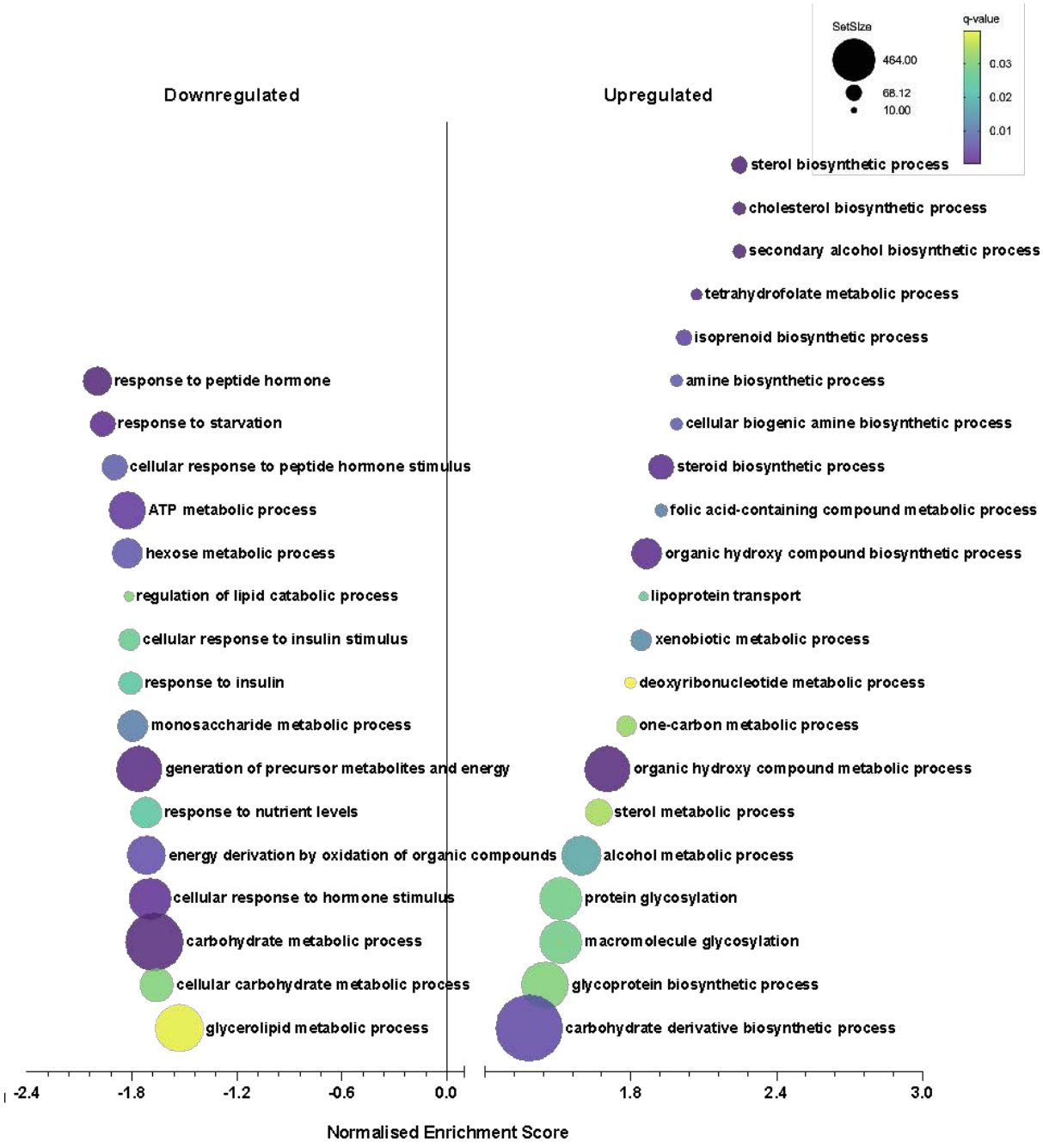


Figure 6.12. Dysregulated metabolic processes in *cyp21a2-/-* larvae
The processes presented are a manual selection of the top 20 up- and down-regulated metabolic processes from the list of GSEA of results. Downregulated gene sets are indicated on the left side and upregulated ones on the right side. The horizontal line indicates the normalised enrichment score. The size of the dots indicates the size of the gene set, and the colour corresponds to the q-value, decreasing from yellow to violet.

Extensive gene dysregulations were found at various steps involved in the energy homeostasis and oxidation of organic compounds. Processes that were downregulated included carbohydrate, in particular glycogen, metabolism, the tricarboxylic acid cycle and

the electron transport chain (**Figure 6.13**). Both glycogenesis and glycogenolysis were downregulated as shown by the expression genes involved in their regulation: glycogenin1 (*gyg1b*, log2FC = -0.47, p < 0.00001), glycogen synthase (*gys1*, log2FC = -0.69, p < 0.00001), phosphorylase kinase (*phkg1a*, log2FC = -0.66, p = 0.001), protein phosphatase 1 (*ppp1r3ab*, log2FC = -1.27, p < 0.00001), muscle glycogen phosphorylase (*pygma*, log2FC = -0.90, p < 0.00001), as were other genes involved in the catabolism of carbohydrates: phosphofructokinase (*pfkma*, log2FC = -0.93, p < 0.00001), aldolase a (*aldoaa*, log2FC = -0.39, p < 0.00001), TP53 induced glycolysis regulatory phosphatase a (*tigara*, log2fc = -1.01, p = 0.005). There was also downregulation of genes involved in different steps of the oxidative phosphorylation such as glyceraldehyde-3-phosphate dehydrogenase (*gapdh*, log2FC = -0.41, p = 0.002), isocitrate dehydrogenase, succinate dehydrogenase, NADH:ubiquinone oxidoreductase.

The dysregulation of carbohydrate metabolism included downregulation of certain processes such as glycogen synthesis, glycolysis, glucose and fructose metabolism, insulin signalling pathway (**Figure 6.14.A**)., however, some pathways were upregulated, including protein glycosylation, proteoglycan biosynthesis, nucleotide biosynthesis (**Figure 6.14.B**). The gene coding phosphoenolpyruvate carboxykinase 1, *pck1*, essential in the initiation of gluconeogenesis was significantly downregulated (log2FC = -0.75, p < 0.00001). A small number of genes involved in ATP synthesis were also included in the upregulated group, such as the ATP synthase membrane subunit (*atp5mc1b*, log2FC = 0.31, p = 0.013).

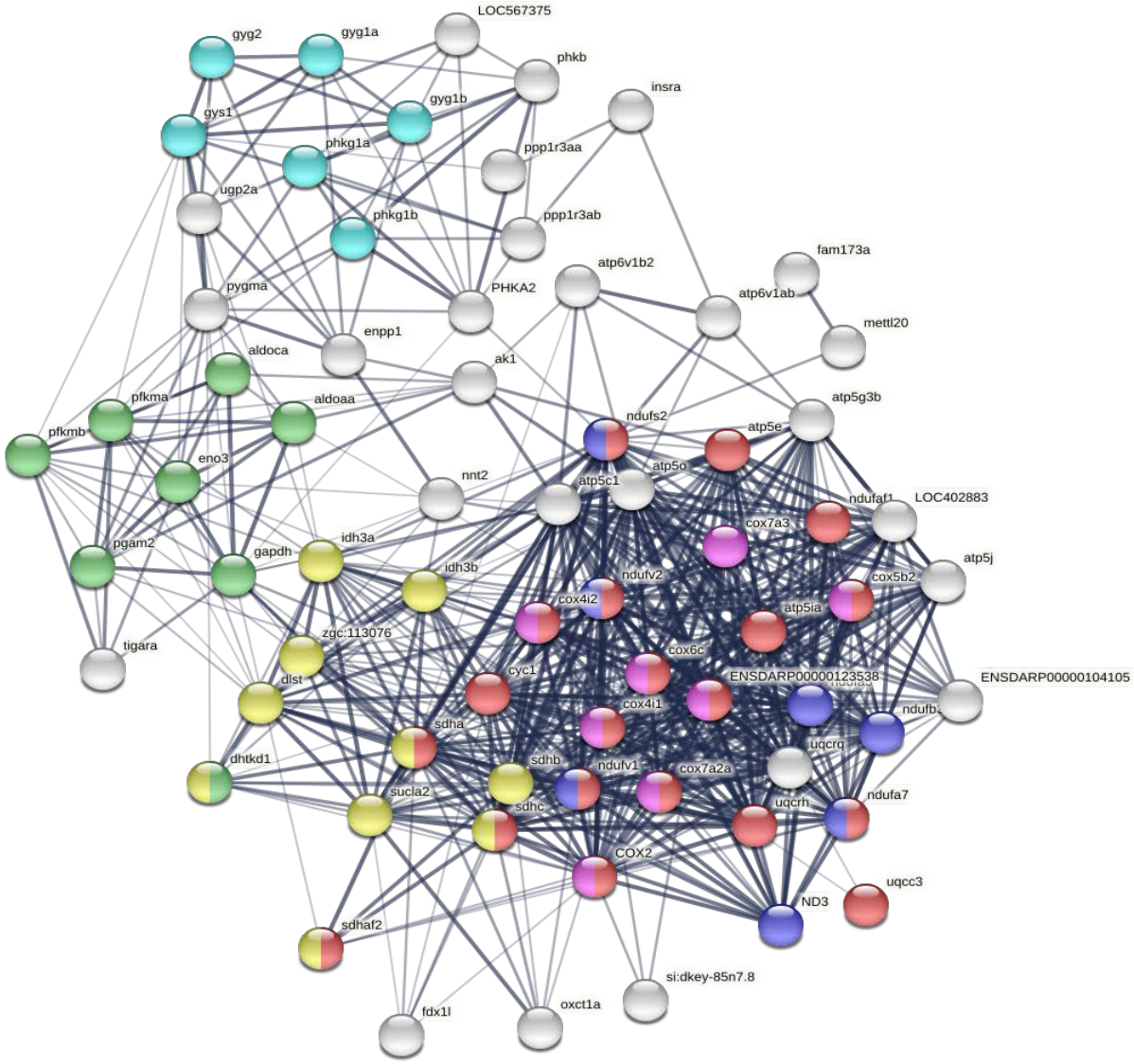
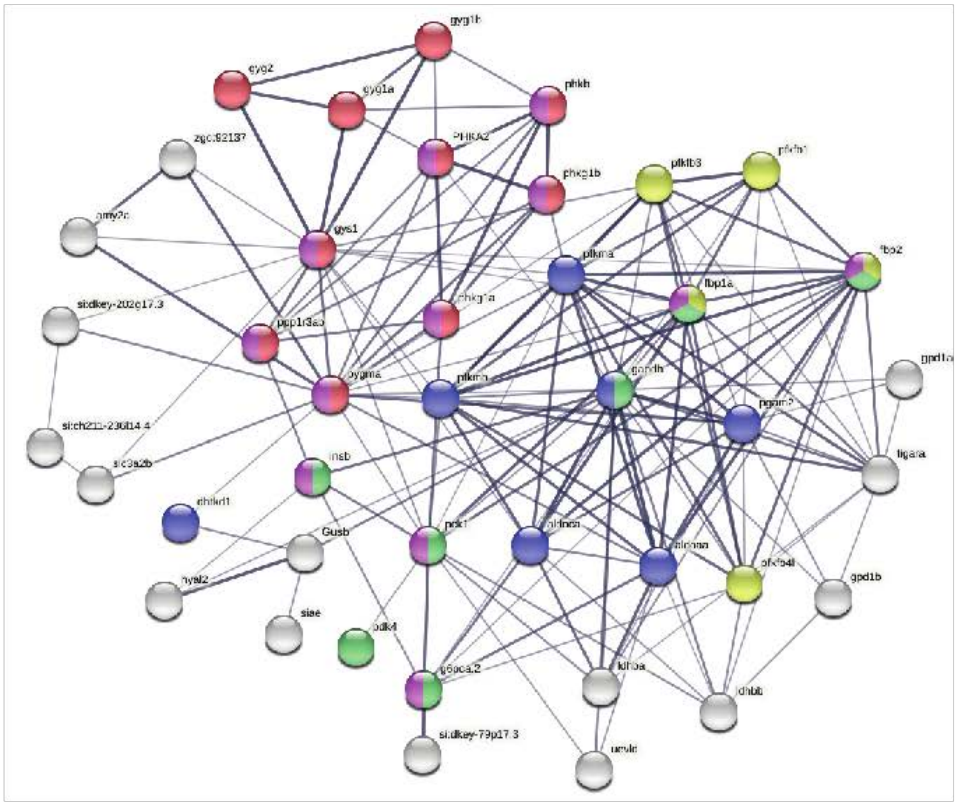


Figure 6.13. ATP metabolism in cyp21a2-/- larvae
Visualisation in STRING of dysregulated genes involved in larvae ATP metabolism and synthesis of energy derivatives identified by GSEA. The connecting lines indicate interactions between proteins, the width of the lines being proportional to the confidence or the strength of data supporting the interaction. The colours of the nodes indicate the main sub-processes to which they belong: all oxidative phosphorylation (red, n=20), NADH dehydrogenase (ubiquinone) activity (dark blue, n=7), cytochrome-c oxidase activity (pink, n=8), tricarboxylic acid cycle (yellow, n=10), glycogen metabolism (light blue, n=6) and glycolytic process (green, n=8).

A



B

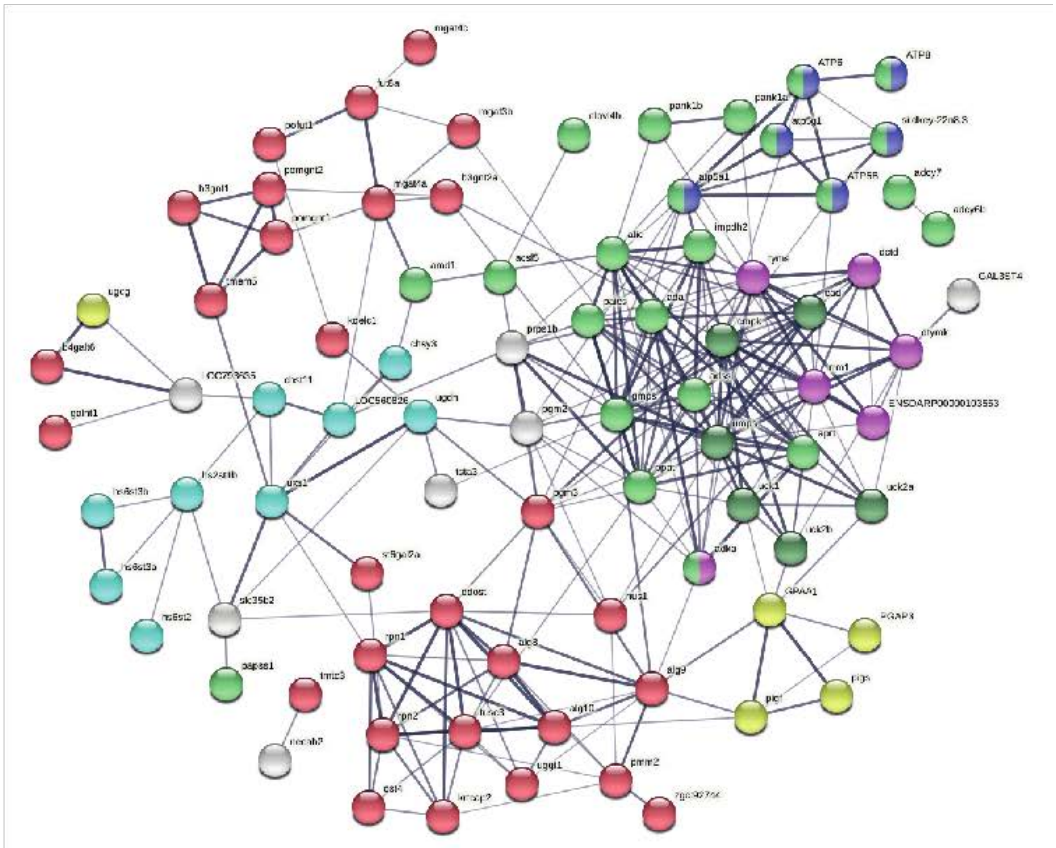


Figure 6.14. Carbohydrate metabolism in *cyp21a2-/-* larvae

Visualisation in STRING of dysregulated genes involved in downregulated (**A**) and upregulated (**B**) processes related to carbohydrate metabolism in larvae. The connecting lines indicate interactions between proteins, the width of the lines being proportional to the confidence or the strength of data supporting the interaction. The colours of the nodes indicate the main sub-processes to which they belong: **A**. Down-regulated processes: glycogen synthesis (red, n=10), glycolysis (blue, n=7), glucose metabolism (green, n=9), fructose metabolism (yellow, n=5), insulin signalling pathway (purple, n=12); **B**. Upregulated processes: protein glycosylation (red, n=34), protein lipidation (yellow, n=5), proteoglycan biosynthesis (light blue, n=10), ATP synthesis (dark blue, n=6), purine nucleotide metabolism (light green, n=28), pyrimidine nucleotide metabolism (dark green, n=6), and deoxyribonucleotide biosynthesis (pink, n=6).

From the group of DEGs identified to be involved in the response to peptide hormones, eight were found to play a role in the cellular response to insulin (**Figure 6.15A**). Both paralogues of the insulin receptor substrate 2 were downregulated (*irs2a,* log2FC = -0.53, p = 0.030 and *irs2b,* log2FC = -0.97, p < 0.00001), as was phosphoinosistide-3-kinase (*pik3r1,* log2FC = -0.41, p < 0.001). The lipin 1 gene (*lpin1*) encoding a phosphatidic acid phosphohydrolase enzyme involved in insulin signalling and lipid metabolism, was also severely downregulated in *cyp21a2-/-* larvae (log2FC = -1.53, p < 0.00001). An important number of genes were involved in the regulation of nuclear transcription such as nuclear receptors *nr1d1, nr4a1*and signal transducers and activators of transcription *stat1a, stat1b, stat2*, the latter being involved also in the cytokine signalling pathway. Analysis of the results in STRING showed that the insulin and cytokine pathways are interlinked with the growth hormone pathway, the growth hormone receptor gene *ghrb* being also significantly downregulated (log2FC = -0.90, p < 0.00001). Interestingly, within the set of genes involved in the response to nutrients and starvation, the main sub-process was autophagy (composed of genes that are also involved in the response to stress), while glucose, lipid and pyruvate metabolism were more modestly represented (**Figure 6.15.B**).

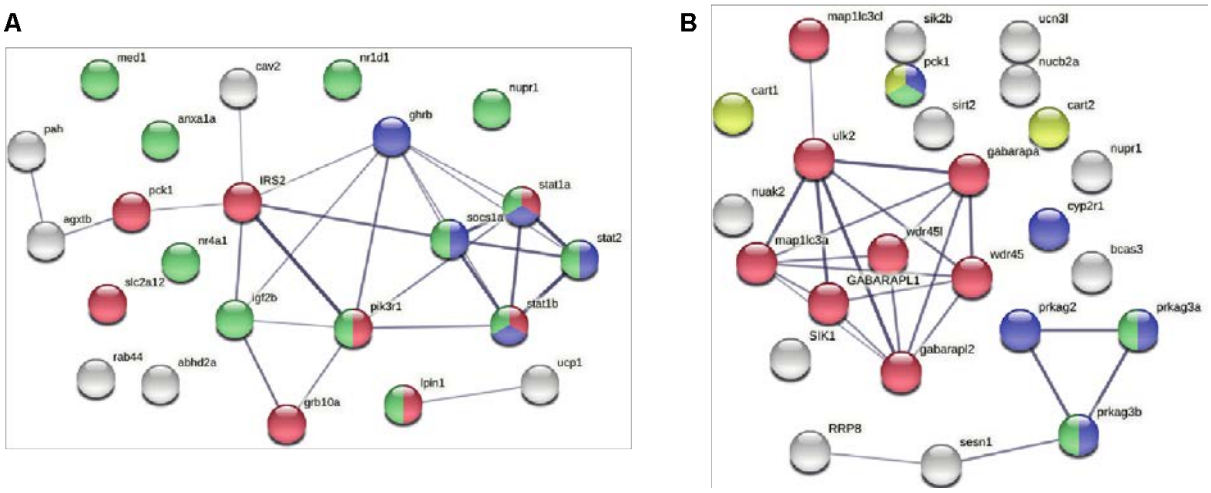


Figure 6.15. Response to peptide hormones and to nutrients in *cyp21a2-/-* larvae
Visualisation in STRING of dysregulated genes involved in different overexpressed metabolic processes identified by GSEA. The connecting lines indicate interactions between proteins, the width of the lines being proportional to the confidence or the strength of data supporting the interaction. **A.** Response to peptide hormones: insulin signalling (red, n=8), regulation of gene expression (green, n=12), and cytokine signalling (blue, n=4). **B.** Response to nutrients and starvation: autophagy (red, n=8), lipid metabolism (blue, n=5), glucose metabolism (yellow, n=3), pyruvate metabolism (green, n=3).

The main upregulated metabolic processes identified in the *cyp21a-/-* larvae were organic hydroxy compound metabolism, carbohydrate derivate biosynthesis and glycoprotein biosynthesis. The main sub-classes identified within the organic hydroxy compound metabolism were steroids, inositol phosphate, sphingolipid, melanin. A significant number of dysregulated enzymes with oxidoreductase activity were also identified within this process. (**Figure 6.16**) Importantly several genes involved in the cholesterol metabolism were significantly upregulated including apolipoprotein A-1a (*apoa1a*, log2FC = 2.43, p < 0.00001), 3-hydroxy-3-methylglutaryl-CoA reductase a (*hmgcra*, log2FC = 0.53, p = 0.026), and cytochrome b5 reductase (*cyb5r2*, log2FC = 0.67, p = 0.027). The GO term also included a number of enzymes with oxidoreductase activity, such as fatty acid hydroxylase domain containing 2 (*faxdc2*), alcohol dehydrogenase 8b (*adh8b*), alcohol dehydrogenase 5 (*adh5*).

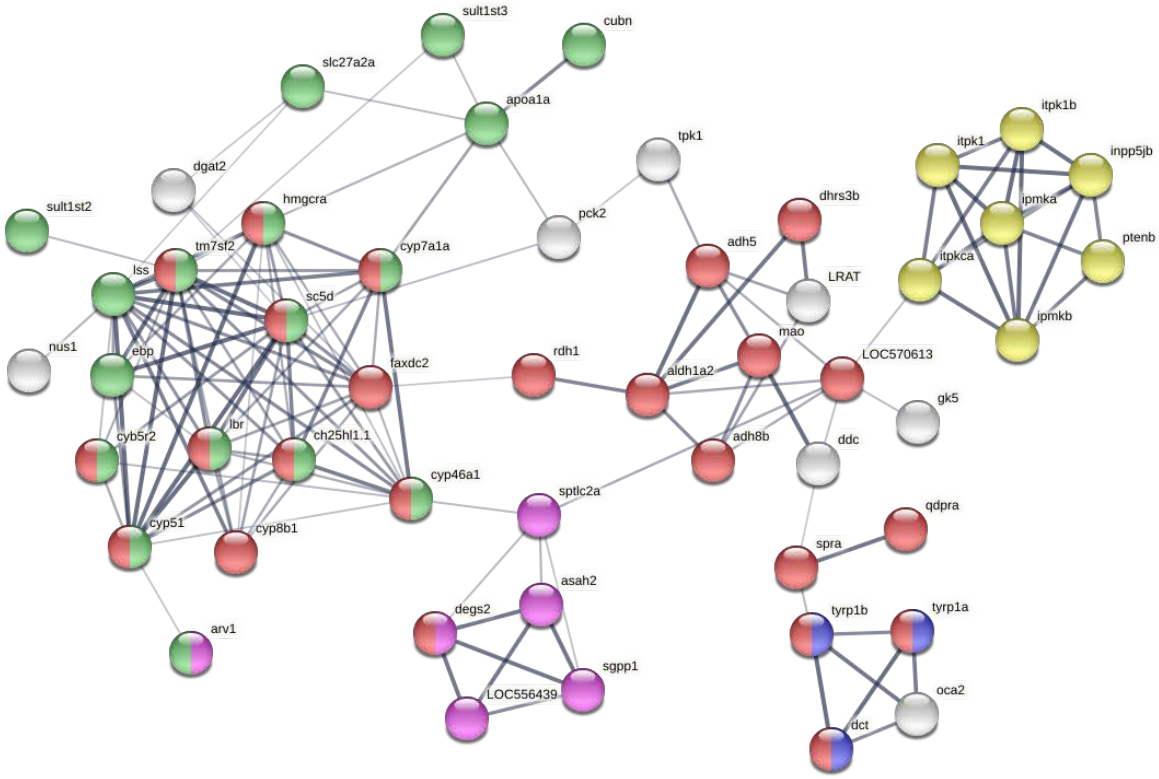


Figure 6.16. Organic hydroxy compound metabolism in *cyp21a2-/-* larvae
Visualisation in STRING of dysregulated genes involved in organic hydroxy compound metabolic processes identified by GSEA. The connecting lines indicate interactions between proteins, the width of the lines being proportional to the confidence or the strength of data supporting the interaction. The colours of the nodes indicate the main sub-processes to which they belong to the metabolism of steroids (green, n=19), inositol phosphate (yellow, n=9), melanin (blue, n=3), sphingolipid (pink, n=6); oxidoreductases (red, n=25).

6.1.1.2.4 Associations between differentially expressed genes in *cyp21a2-/-* larvae and human disease

The association between transcriptomic dysregulation in the mutant larvae and human pathology was explored using the DGEs in larvae to generate a list of human orthologues (**Figure 6.17**). The predominant pathology was cancer, including neoplastic conditions of variable histology and localisation. However, the findings also included a number of metabolic and cardiovascular problems, such as diabetes, metabolic syndrome, arteriosclerosis and coronary heart disease.

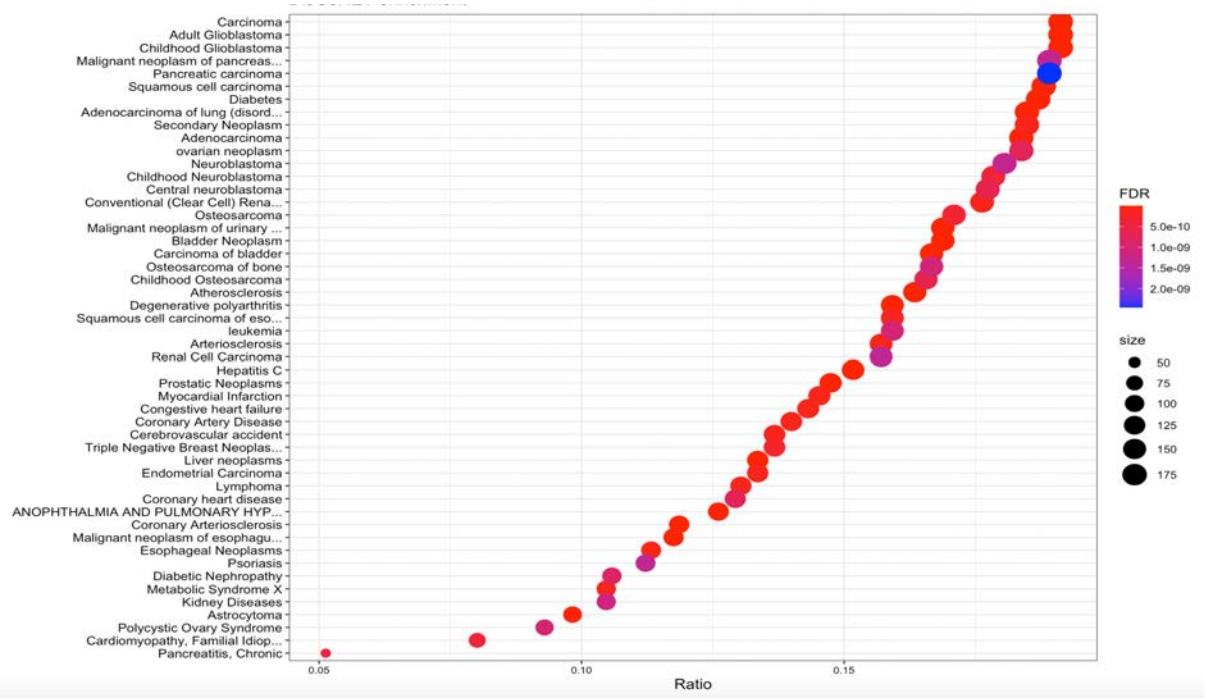


Figure 6.17. Associations between larval *cyp21a2-/-* transcriptome and human disease
Enrichment plot showing the associations between the transcriptomic changes in *cyp21a2-/-* larvae and human diseases, produced in R using the "disgenet2r" package. The horizontal axis shows the ratio between the number of DEGs in a category and all DGEs, the size of the dots represents the number of DEGs in a category and the colour, the *p-adj* value.

6.1.1.3 Transcriptomic profile of cyp21a2-/- adult livers

When including all liver samples in the PCA, there was no consistent differentiation between wild types and mutants, however, there was clear clustering based on sex (**Figure 6.18.A**), except for a female mutant sample ("mutf5") which clustered with the male samples. This was further confirmed by the cluster dendrogram (**Figure 6.18.B**) and heatmap of normalised counts (**Figure 6.18.C**). Examining the sequencing depths of the samples, the size factors were fairly similar, with the exception of one of the mutant male samples ("mutm6"), which had a very low size factor and was consequently excluded from the subsequent analysis (**Figure 6.18.D**).

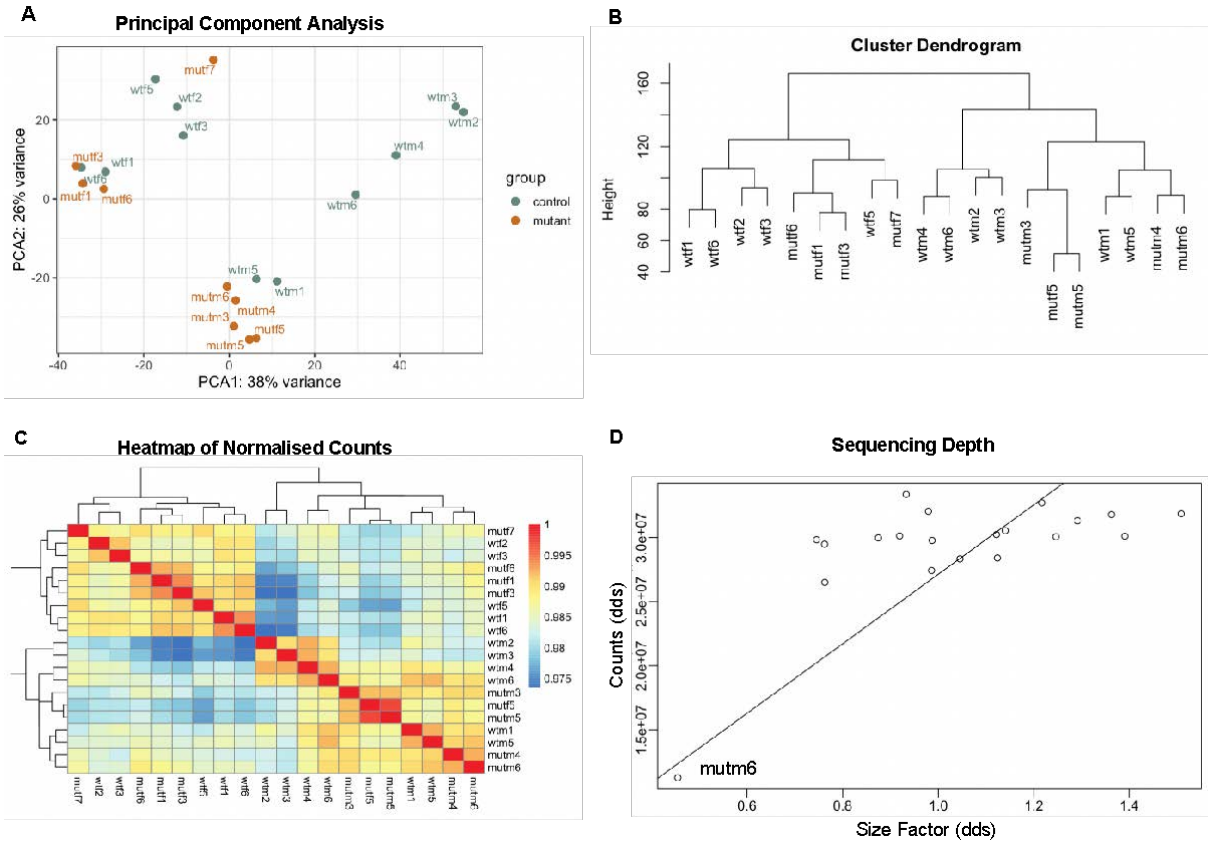


Figure 6.18. Inter-sample variations in *cyp21a2* livers
A. Principal Component Analysis; **B**. Cluster dendrogram; **C**. Heatmap of normalised counts; **D**. Sequencing depth expressed by plotting column counts for each sample against the size factor. (mutm and mutf: *cyp21a2* mutants male and female respectively, wtm and wtf: wild type male and female respectively).

Due to the evident difference in differential gene expression between males and females, which appeared to surpass distinctions induced by genotype, it was decided to continue the differential gene expression analysis exploring dysregulations in the cyp21a2-/- livers focusing on the male samples. This was based on the fact that while in male livers there was a satisfactory level of clustering between wild types and mutants, in females no such distinction could be found (**Figure 6.18**).

After excluding the mutant sample that had a significantly low sequencing depth ("mutm6"), differential gene expression was explored between male livers, comparing

mutants against wild types. The PCA showed satisfactory clustering of the mutant samples, although the wild type samples were more widely spread (**Figure 6.19**).

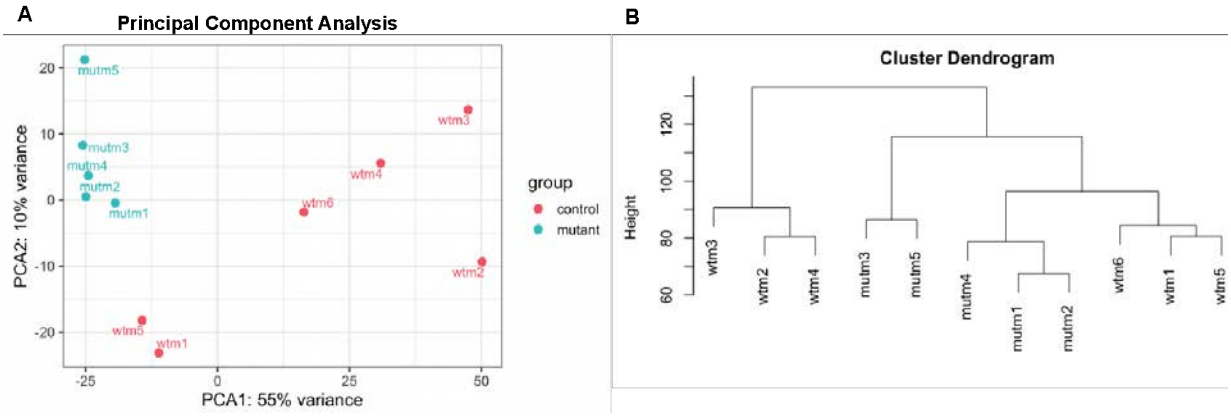
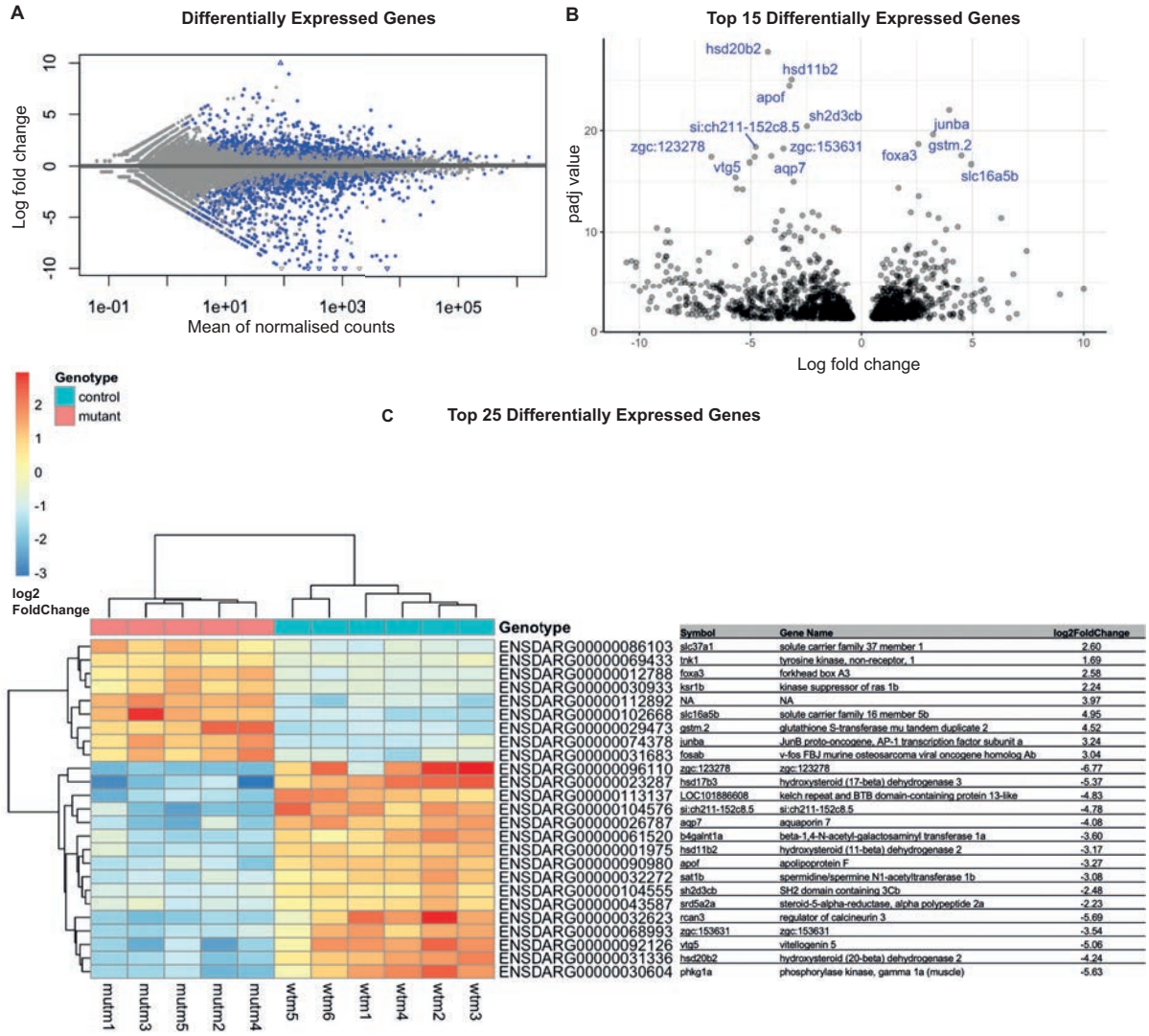


Figure 6.19. Inter-sample variations in *cyp21a2* male livers
A. Principal Component Analysis; **B**. Cluster dendrogram; (mutm: *cyp21a2* mutants male, wtm: wild type male).

Differential gene expression (DGE) analysis identified 717 genes that were significantly upregulated (log2 fold change (LFC) > 0, $p < 0.05$) and 960 genes significantly downregulated (LFC < 0) in male mutant livers (**Figure 6.20A**). The most dysregulated genes are shown in the volcano plot (**Figure 6.20B**), with *hsd20b2* being the most significantly downregulated gene (LFC = -4.24, p-adj = 1.52e-28). There was good distinction between WT and mutant samples in the expression of the top 25 DEGs, among which the majority were downregulated in *cyp21a2-/-* livers (**Figure 6.20C**).



	Symbol	Gene Name	log2FoldChange
ENSDARG00000086103	slc37a1	solute carrier family 37 member 1	2.60
ENSDARG00000069433	tnk1	tyrosine kinase, non-receptor, 1	1.69
ENSDARG00000012788	foxa3	forkhead box A3	2.58
ENSDARG00000030933	ksr1b	kinase suppressor of ras 1b	2.24
ENSDARG00000112892	NA	NA	3.97
ENSDARG00000102668	slc16a5b	solute carrier family 16 member 5b	4.95
ENSDARG00000029473	gstm.2	glutathione S-transferase mu tandem duplicate 2	4.52
ENSDARG00000074378	junba	JunB proto-oncogene, AP-1 transcription factor subunit a	3.24
ENSDARG00000031683	fosab	v-fos FBJ murine osteosarcoma viral oncogene homolog Ab	3.04
ENSDARG00000096110	zgc:123278	zgc:123278	-6.77
ENSDARG00000023287	hsd17b3	hydroxysteroid (17-beta) dehydrogenase 3	-5.37
ENSDARG00000113137	LOC101886608	kelch repeat and BTB domain-containing protein 13-like	-4.83
ENSDARG00000104576	si:ch211-152c8.5	si:ch211-152c8.5	-4.78
ENSDARG00000026787	aqp7	aquaporin 7	-4.08
ENSDARG00000061520	b4galnt1a	beta-1,4-N-acetyl-galactosaminyl transferase 1a	-3.60
ENSDARG00000001975	hsd11b2	hydroxysteroid (11-beta) dehydrogenase 2	-3.17
ENSDARG00000090980	apof	apolipoprotein F	-3.27
ENSDARG00000032272	sat1b	spermidine/spermine N1-acetyltransferase 1b	-3.08
ENSDARG00000104555	sh2d3cb	SH2 domain containing 3Cb	-2.48
ENSDARG00000043587	srd5a2a	steroid-5-alpha-reductase, alpha polypeptide 2a	-2.23
ENSDARG00000032623	rcan3	regulator of calcineurin 3	-5.69
ENSDARG00000068993	zgc:153631	zgc:153631	-3.54
ENSDARG00000092126	vtg5	vitellogenin 5	-5.06
ENSDARG00000031336	hsd20b2	hydroxysteroid (20-beta) dehydrogenase 2	-4.24
ENSDARG00000030604	phkg1a	phosphorylase kinase, gamma 1a (muscle)	-5.63

Figure 6.20. Differentially expressed genes in the *cyp21a2-/-* adult male livers
A. MA plot showing differentially expressed genes; the statistical significance (*p-adj* < *0.05*) is indicated by the blue dots while differentially expressed genes without statistical significance are coloured grey. **B.** Volcano plot showing the top 25 DEGs in mutant livers, with statistical significance (*p-adj)* represented on the vertical axis and the LFC on the horizontal one. **C**. Heatmap and table indicating the top 25 DEGs in mutant livers.

Comparing the group of significantly dysregulated genes found in the *cyp21a2-/-* male livers against those identified in mutant larvae, a total of 167 genes were found to be significantly dysregulated in both larvae and adult livers as a result of the 21-hydroxylase deficiency. Of these, 20 were up-regulated and 89 down-regulated in both larvae and

livers, for the remaining 58 genes the direction of the dysregulation being different between larvae and livers (**Figure 6.21**).

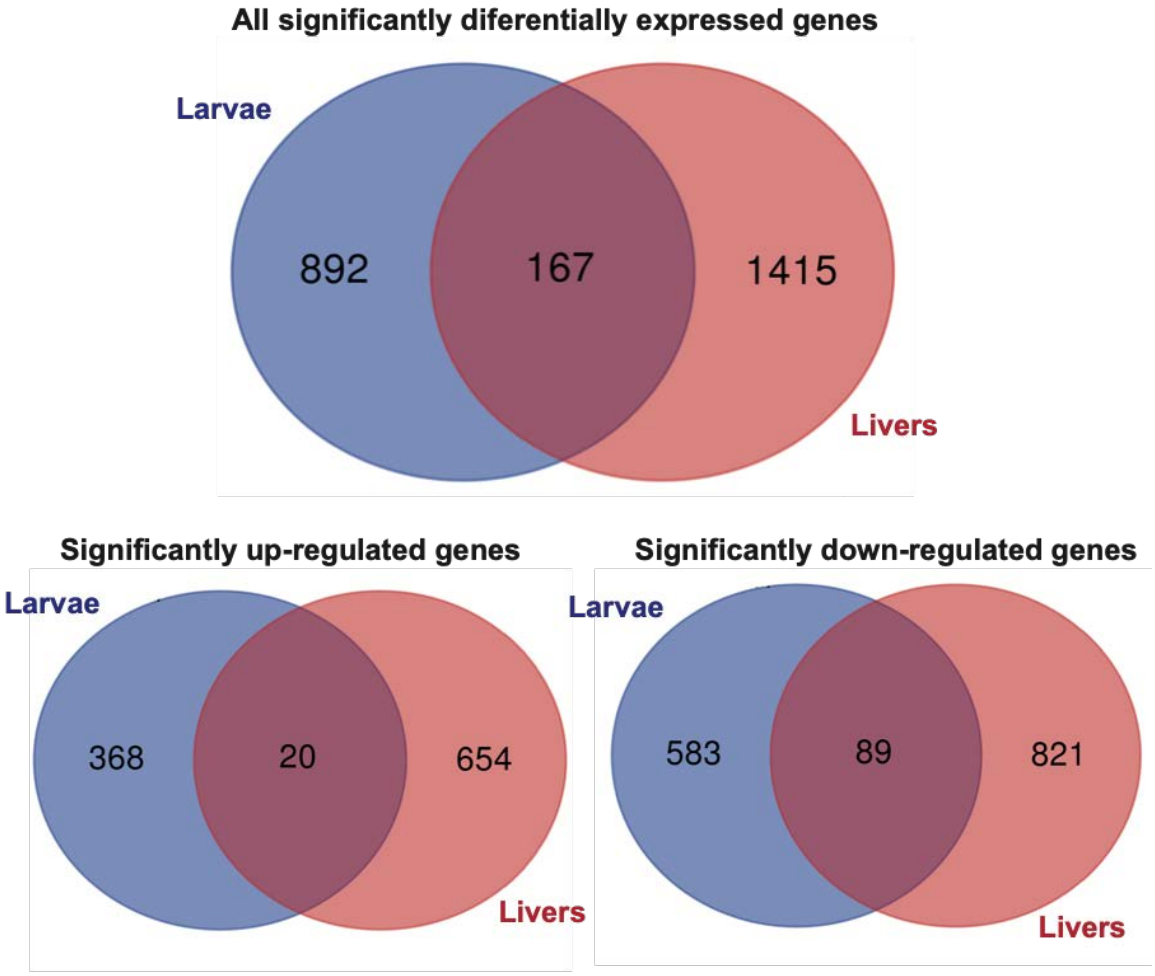


Figure 6.21. Overlap between differentially expressed genes found in *cyp21a2-/-* larvae and adult livers
The upper Venn diagram shows all the genes that were differentially expressed in mutant larvae (blue) and adult male livers (red), while the lower diagrams show up- and down-regulated genes.

6.1.1.3.1 Comparison between the differential expression of essential metabolic genes explored by qPCR and RNA sequencing in *cyp21a2-/-* male livers

Exploring the differential gene expression of the specific metabolic genes targeted by the qPCR experiment, the results of the liver RNA sequencing showed several discrepancies to those of the qPCR (**Table 6.5**). Due to the limited number of samples in the first qPCR experiment, only the second one (including 5 male liver samples) was included in the

comparison. For glucose metabolism, the only significant result was the downregulation of *pck2*, which was also found to be downregulated by qPCR though not significantly. Similar to the qPCR results for male livers, *pck1* was not significantly downregulated in the RNA sequencing results. For the lipid metabolism, the direction of the dysregulations was more consistent between DGE and qPCR, with RNA sequencing showing marked upregulation of the two genes explored for adipocyte differentiation (*pparg* and *cebpa*) and one involved in cholesterol synthesis (*hmgcra),* as well as downregulation of beta-oxidation of fatty acids (*hadhb*) (**Table 6.5**).

Table 6.5. RNA sequencing vs qPCR results in cyp21a2 livers

Genes involved in glucose metabolism				
	qPCR experiment		**RNA sequencing**	
Gene	**Mean fold change**	***p* value**	**Log2 fold change**	***p-adj***
pck1	1.55	0.086	-0.38	0.747
pck2	0.77	0.228	-0.73	0.002*
pkl	0.69	0.465	-0.07	0.944
gck	0.47	0.150	2.63	0.050
insa	3.40	0.841	0.28	0.954
irs2	1.03	0.209	-2.12	0.438
Genes involved in fat metabolism				
	qPCR experiment		**RNA sequencing**	
Gene	**Mean fold change**	***p* value**	**Log2 fold change**	***p-adj***
pparg	1.86	0.214	1.08	0.007*
cebpa	4.94	0.096	3.42	3.54e-07*

fasn	3.20	0.377	1.05	0.138
fabp11a	1.65	0.353	n/a	n/a
srebf1	1.00	0.615	1.46	0.057
lpl	0.42	0.031*	-0.67	0.267
hadhaa	0.81	0.356	-0.39	0.350
hadhb	0.61	0.138	-0.91	0.009*
hmgcra	20.8	0.078	3.82	5.79e-11*
hmgcrb	1.66	0.095	0.54	0.222
lep	0.90	0.342	0.62	n/a
acrp	0.67	0.152	-2.12	n/a

The results of the qPCR experiments are expressed as means of the relative fold change (with 1 being the threshold for up- vs downregulation) and the p value corresponding to the t test (n = 5 WT and 5 mutant samples). The results of the RNA sequencing data are expressed as log2 fold change (the threshold for up- vs downregulation is 0) and the adjusted p value (n = 6 WT and 5 mutant samples). * Indicates statistical significance.

6.1.1.3.2 Gene ontology enrichment analysis

GO enrichment analysis performed in GOrilla identified 12 GO terms, of which six were processes related to the metabolism of steroids, lipids, organic hydroxy compounds, chitin and small molecule metabolism, while the remaining six included ions and organic acid transport, and multicellular organism response to stress. (**Table 6.6 and Figure 6.22.A**). Of note, three of the identified GO terms were associated with only two or three genes: multicellular organism response to stress (*hsd11b2, hsd20b2*) and chitin metabolic/catabolic process (*chia.1, chia.2, chia.3*). Given the small total numbers of genes associated with these GO terms, the enrichment values were accordingly very high.

These results were comparable to those of the gene ontology analysis conducted in R using the "clusterProfiler" package, which also identified the response to biotic stimuli as overexpressed biological processes (**Figure 6.22.B, C**).

Table 6.6. Results of the gene ontology analysis for biological processes in GOrilla in *cyp21a2* livers

GO Term	Description	Enrichment (N,B,n,b)	p-value	FDR q-value
GO:0015711	Organic anion transport	2.75 (13643, 160, 1551, 50)	7.28e-11	6.21e-07
GO:0006820	Anion transport	2.34 (13643, 244, 1551, 65)	2.16e-10	9.23e-07
GO:0008202	Steroid metabolic process	7.16 (13643, 88, 368, 17)	2.73e-09	7.78e-06
GO:0046942	Carboxylic acid transport	3.07 (13643, 103, 1551, 36)	4.42e-09	9.43e-06
GO:0015849	Organic acid transport	3.02 (13643, 105, 1551, 36)	8.07e-09	1.38e-05
GO:0006629	Lipid metabolic process	2.59 (13643, 547, 327, 34)	3.64e-06	5.17e-03
GO:0033555	Multicellular organism response to stress	1,240.27 (13643, 11, 2, 2)	3.79e-06	4.63e-03
GO:0044281	Small molecule metabolic process	1.56 (13643, 891, 1175, 120)	4.46e-06	4.75e-03
GO:0006811	Ion transport	1.94 (13643, 676, 624, 60)	6.42e-06	6.08e-03
GO:1901615	Organic hydroxy compound metab. proc.	3.77 (13643, 172, 379, 18)	7.66e-06	6.54e-03
GO:0006032	Chitin catabolic process	110.02 (13643, 6, 62, 3)	8.30e-06	6.44e-03
GO:0006030	Chitin metabolic process	110.02 (13643, 6, 62, 3)	8.30e-06	5.90e-03

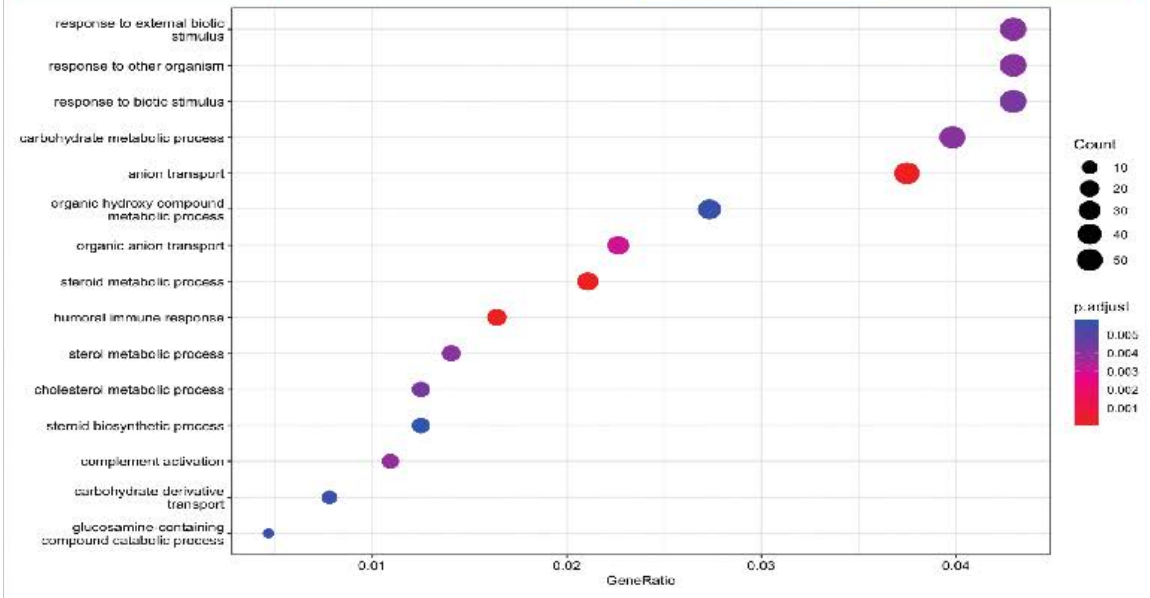
Biological Process GO term analysis of genes exhibiting differential expression in adult male WT and *cyp21a2-/-* mutant livers, using GOrilla. The **p-value** is the enrichment p-

value computed according to the mHG or HG model. The **FDR q-value** is the correction of the p-value for multiple testing using the Benjamini and Hochberg (1995) method. The **Enrichment = (b/n) / (B/N) (**N - total number of genes, B - total number of genes associated with a specific GO term, n – a flexible cut-off, being an automatically determined number of genes in the input list or 'target set', b - is the number of genes in the target set that are associated with the GO term).

A


anion transport
organic anion transport
response to other organism
defense response
organic acid transport
response to biotic stimulus
steroid biosynthetic process
steroid metabolic process
small molecule metabolic process
lipid metabolic process
lipid biosynthetic process

B


response to external biotic stimulus
response to other organism
response to biotic stimulus
carbohydrate metabolic process
anion transport
organic hydroxy compound metabolic process
organic anion transport
steroid metabolic process
humoral immune response
sterol metabolic process
cholesterol metabolic process
steroid biosynthetic process
complement activation
carbohydrate derivative transport
glucosamine containing compound catabolic process
0.01
0.02
0.03
0.04
GeneRatio
Count
10
20
30
40
50
p.adjust
0.005
0.004
0.003
0.002
0.001

C

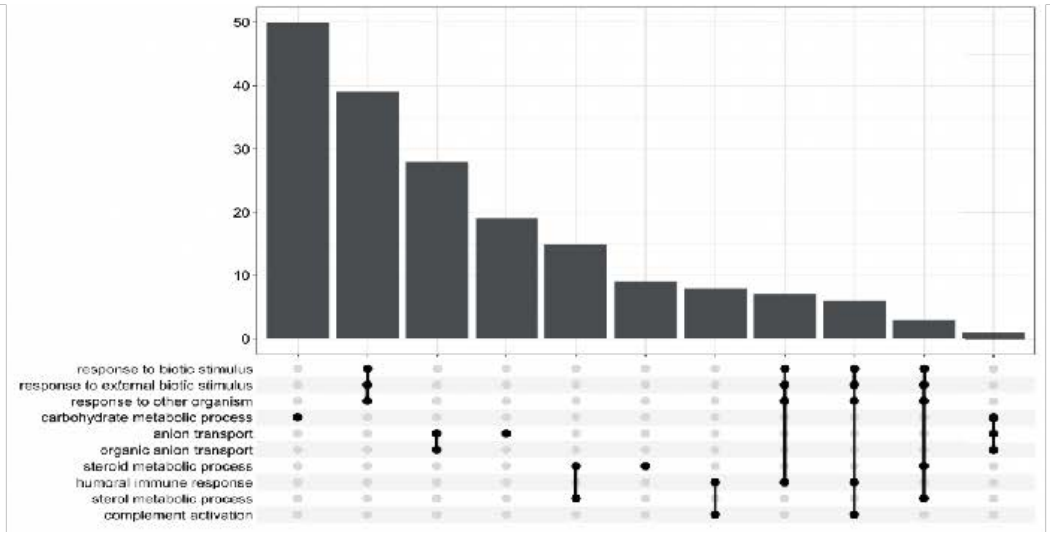

50
40
30
20
10
0
response to biotic stimulus
response to external biotic stimulus
response to other organism
carbohydrate metabolic process
anion transport
organic anion transport
steroid metabolic process
humoral immune response
sterol metabolic process
complement activation

Figure 6.22. Gene ontology over-expression analysis in *cyp21a2-/-* male livers
A. Tree map produced in REVIGO showing a graphic representation of the biological processes that are identified to be dysregulated in *cyp21a2-/-* livers following the gene ontology analysis conducted in GOrilla, the size of the rectangles being determined by the p-value. **B**. Enrichment plot produced in R using the "clusterProfiler" package. The horizontal axis shows the ratio between the number of DEGs in a category and all DGEs, the size of the dots represents the number of DEGs in a category and the colour, the *p-adj* value. **C.** Upset plot showing the overlap of differentially expressed genes between the overexpressed biological processes identified by the GO enrichment analysis. The dotted vertical lines indicate the processes that are included in the analysis and the height of the bars corresponds to the number of overlapping genes.

6.1.1.3.3 Gene set enrichment analysis in *cyp21a2-/-* male livers

The results of the GSEA were overall consistent with the GO enrichment analysis, providing further information regarding the nature of the dysregulation of overexpressed metabolic processes in adult male *cyp21a2-/-* livers (**Figures 6.23 and 6.24**). Thus, processes pertaining to the defence response, immune/inflammatory response and response to biotic stimuli were significantly upregulated, as were certain biological processes including cholesterol biosynthesis. By contrast and similar to larvae, significant downregulation was found for processes involved in the normal cellular housekeeping and function, including energy derivation and oxidative phosphorylation, ribosome synthesis and function, cellular respiration, ion transport.

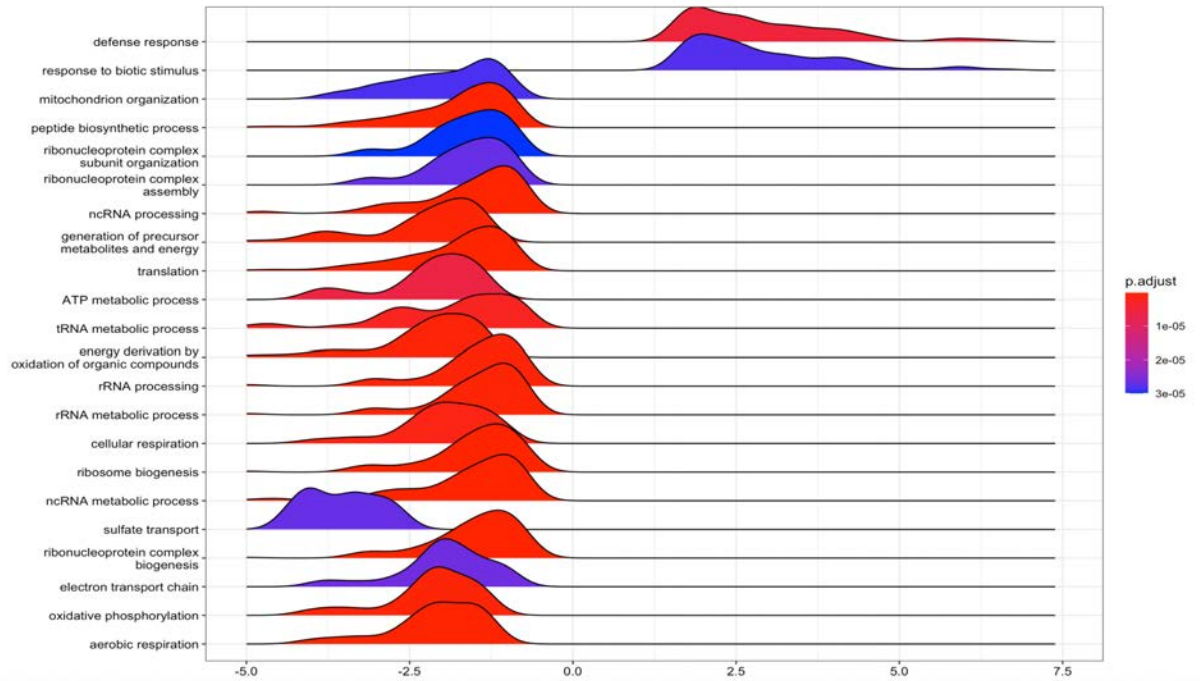


Figure 6.23. Gene set enrichment analysis in *cyp21a2-/-* male livers
The top 20 dysregulated biological processes found by the GSEA. Ridgeplot produced in R using the clusterProfiler' package. The processes were selected automatically based on the normalised enrichment score (showed on the horizontal axis) and the size of the gene set, ie. sets with small sizes were excluded despite having higher enrichment scores. The x axis indicates the normalised enrichment score (NES), related to the degree of upregulation (NES > 0) or downregulation (NES < 0) of each biological process. The shape of the ridges corresponds to the distribution of the genes included in the set, in relation to the enrichment score and the colour indicates the adjusted p-value, decreasing from red to blue.

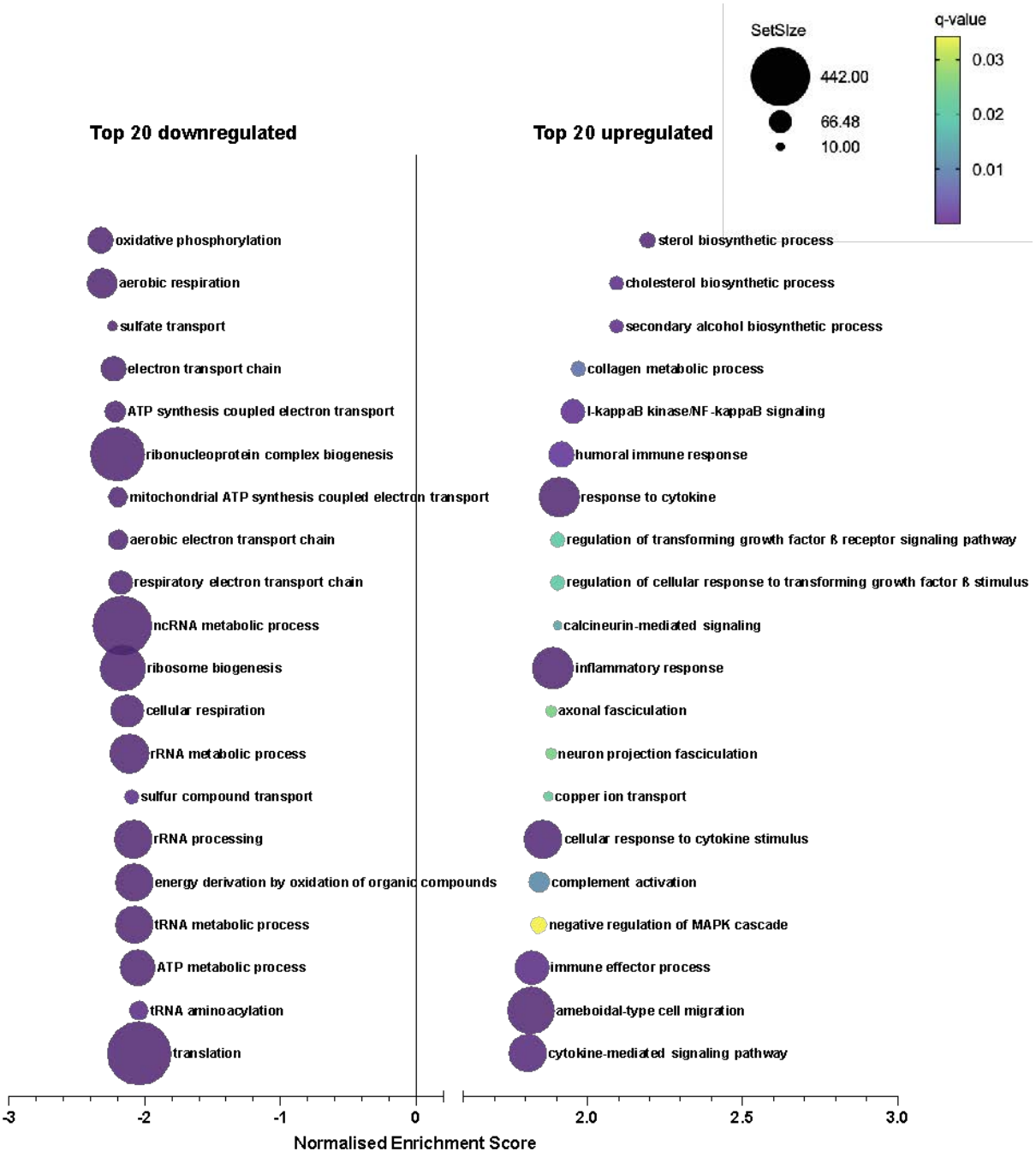


Figure 6.24. Gene set enrichment analysis in *cyp21a2-/-* male livers (2)
The top 20 dysregulated biological processes found by the GSEA. Bubbleplot produced GraphPad Prism showing the 20 most downregulated (left) and upregulated (right) biological processes in *cyp21a2* male livers. The selection was based exclusively on the normalised enrichment score (showed on the horizontal axis). The size of the dots indicates the size of the gene set, and the colour corresponds to the q-value, decreasing from yellow to violet.

As specified in the GSEA analysis of larval samples, the values provided in brackets for specific genes were taken from the differential gene expression (DGE) analysis.

Several processes found within the top 20 downregulated GO terms related to energy homeostasis, in particular ATP metabolism (**Figure 6.25**). The mitochondrial oxidative phosphorylation system was supressed, through the downregulation of multiple components including several subunits of the mitochondrial respiratory complex I, of which the most significant were NADH:ubiquinone oxidoreductase subunit A4 (*nudfa4*, log2FC = -4.75, p = 0.0009) and NADH dehydrogenase subunit 6 (*ND6*, log2FC = -1.07, p = 0.041), the cytochrome-c oxidoreductase complex, most significantly cytochrome c oxidase subunit 7A2a (*cox7a2a*, log2FC = -0.71, p = 0.0001) and cytochrome c oxidase subunit 5Aa (*cox5aa*, log2FC = -0.71, p = 0.042), as well as the tricarboxylic acid cycle. Additionally, several glycolytic enzymes were also downregulated, such as aldolase b, fructose-bisphosphate (*aldob*, log2FC = -1.69, p = 0.032) and TP53 induced glycolysis regulatory phosphatase a (*tigara*, log2FC = -0.79, p = 0.024), as were different factors involved in glycogen metabolism such as the protein phosphatase 1, regulatory subunit 3Ca (*ppp1r3ca*, log2FC = -3.89, p < 0.0001), phosphorylase kinase, gamma 1a (*phkg1a*, log2FC = -5.40, p < 0.0001), phosphorylase kinase, gamma 1b (*phkg1b*, log2FC = -1.67, p = 0.0008), and phosphorylase kinase, alpha 2 (*phka2*, log2FC = -1.78, p = 0.023).

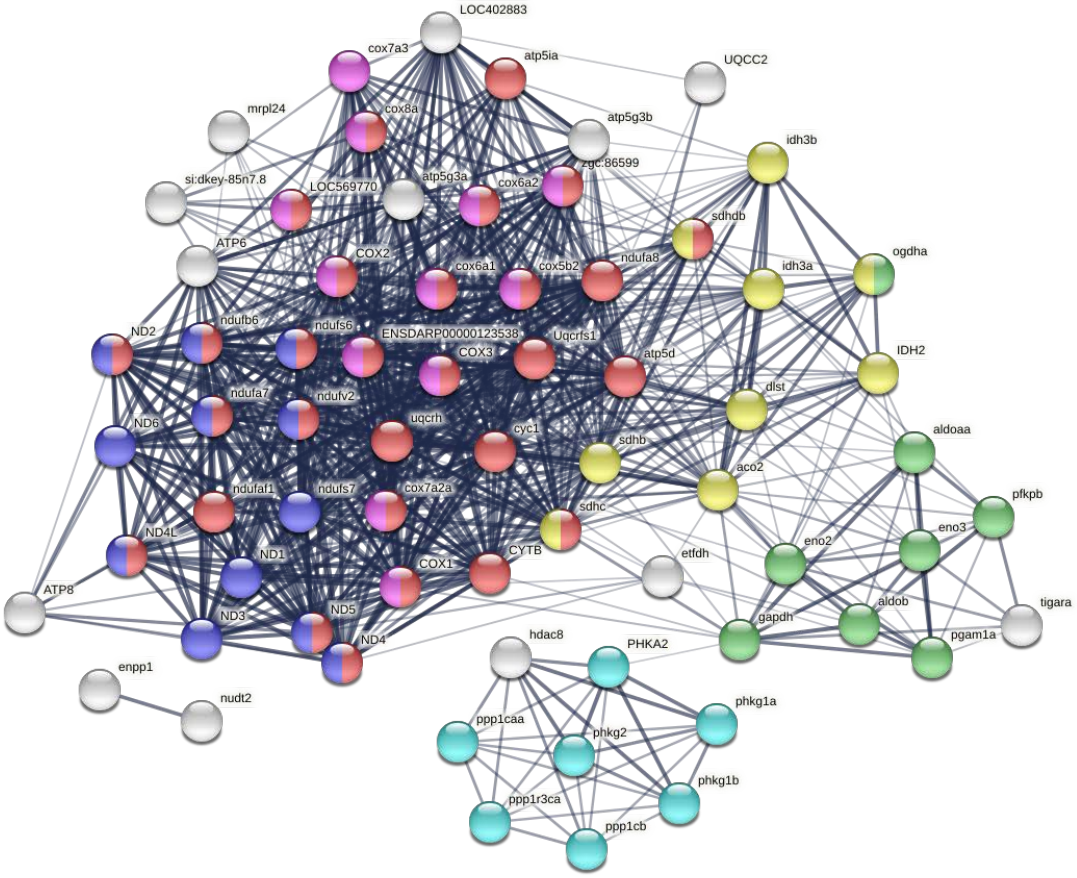


Figure 6.25. ATP metabolism in *cyp21a2-/-* male livers
Visualisation in STRING of interactions between proteins encoded by genes involved in ATP metabolism that are differentially expressed in adult male WT and *cyp21a2-/-* mutant livers, identified by GSEA. The connecting lines indicate interactions between genes, the width of the lines being proportional to the confidence or the strength of data supporting the interaction. The colours of the nodes indicate the main sub-processes to which they belong: all oxidative phosphorylation (red, n=29), NADH dehydrogenase (ubiquinone) activity (dark blue, n=12), cytochrome-c oxidase activity (pink, n=12), tricarboxylic acid cycle (yellow, n=9), glycogen metabolism (light blue, n=7) and glycolytic process (green, n=8).

Among the top down-regulated biological processes, many were related to the ribosomal function, ribonucleoprotein biosynthesis, the GO term “translation” being among the most significantly down-regulated and including many of the other processes identified (**Figure 6.26**). Thus, multiple eukaryotic translation initiation factors were downregulated, in particular eukaryotic translation initiation factor 4E binding protein 3 (*eif4ebp3*, log2FC =

-2.59, p = 0.001) and eukaryotic translation initiation factor 4 gamma, 3b (*eif4g3b*, log2FC = -1.1.7, p = 0.015). Similarly, several members of the tRNA synthase family were downregulated causing reduced tRNA aminoacylation, the most significantly affected being arginyl-tRNA synthetase 1 and Sep (O-phosphoserine) tRNA:Sec (selenocysteine) tRNA synthase (*sepsecs*, log2FC = -0.82, p = 0.022). There was also downregulation of multiple mitochondrial ribosome protein genes, involving both the large and small subunit.

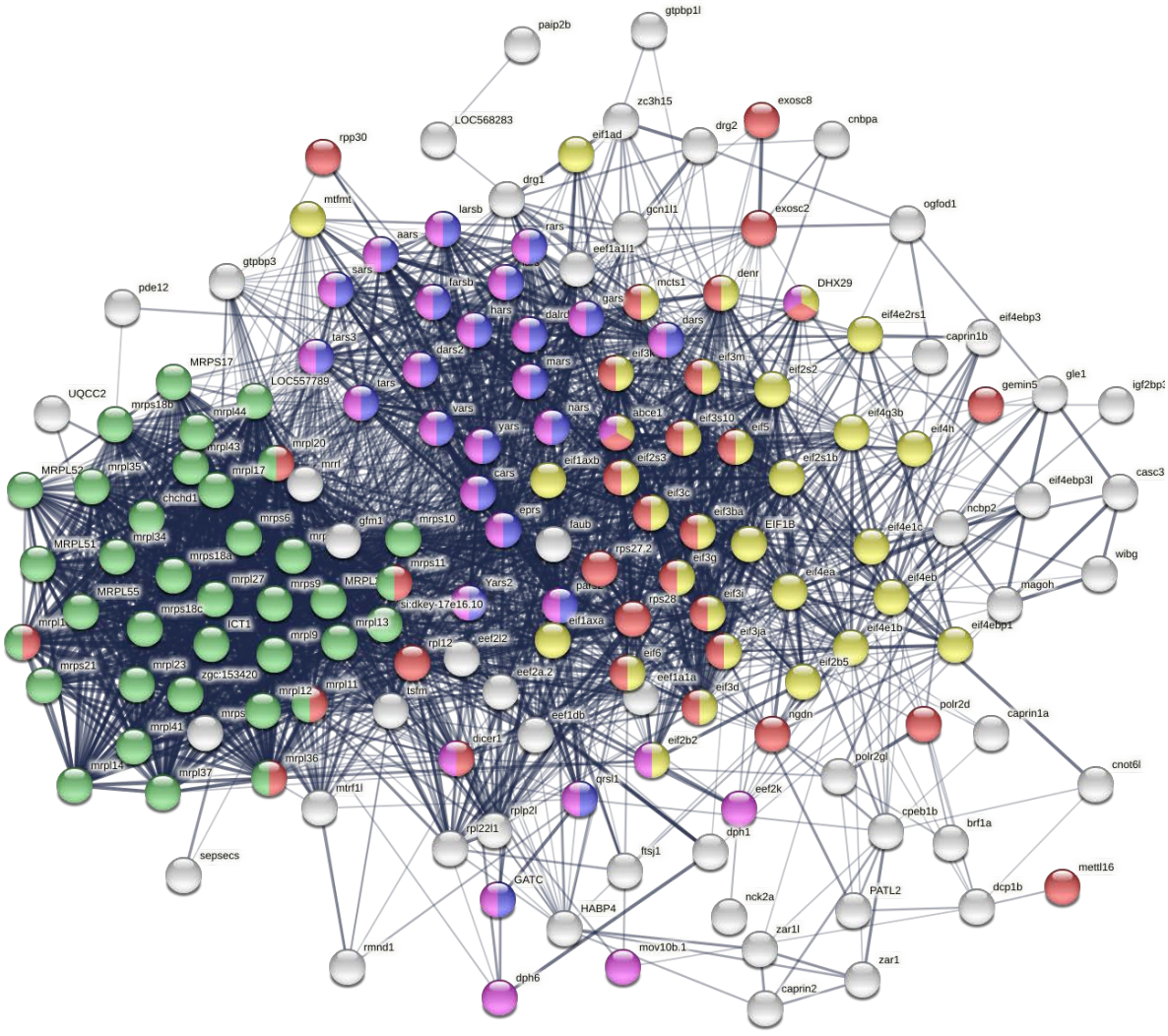


Figure 6.26. Protein translation in *cyp21a2-/-* male livers
Visualisation in STRING of interactions between proteins encoded by dysregulated genes involved in protein translation that are differentially expressed in adult male WT and *cyp21a2-/-* mutant livers, identified by GSEA. The connecting lines indicate interactions between genes, the width of the lines being proportional to the confidence or the strength of data supporting the interaction. The colours of the nodes indicate the main sub-processes to which they belong: ribonucleoprotein complex biogenesis (red, n=32), tRNA

aminoacylation (blue, n=23), translation initiation (yellow, n=33), ATP binding (pink, n=30), mitochondrial ribosome function (green, n=36).

Of the most upregulated over-represented biological processes, a significant number pertained to the inflammatory response (**Figure 6.27**). Several components of the complement system were overexpressed (*c3a.2*, log2FC = 1.28, p = 0.014; *c3a.3*, log2FC = 1.01, p = 0.001; *c3a.3*, log2FC = 1.22, p < 0.00; *c3b.1*, log2FC = 0.89, p = 0.017; tyty *c3b.2*, log2FC = 1.30, p < 0.001; *c4*, log2FC = 1.08, p = 0.009). The cytokine-mediated signalling pathway was also upregulated, with the most significant dysregulation found for interleukin 12 receptor, beta 2a, like (*il12rb2l*, log2FC = 1.81, p = 0.002), tumor necrosis factor receptor superfamily, member 1a (*tnfrsf1a*, log2FC = 1.98, p < 0.001). Other significantly upregulated genes involved in the inflammatory response included MYD88 innate immune signal transduction adaptor (*myd88*, log2FC = 0.94, p = 0.033), MAPK activated protein kinase 2a (*mapkapk2a*, log2FC = 1.96, p < 0.001). Several other members of the mitogen-activated protein kinase (MAPK) signalling pathway were also upregulated, including dual specificity phosphatase 8a (*dusp8a*, log2FC = 2.63, p < 0.001) and dual specificity phosphatase 4 (*dusp4*, log2FC = 2.01, p = 0.024). The signal transducer and activator of transcription gene 3 (*stat3*) and 5a (stat5a) were also upregulated (*stat3,* log2FC = 1.89, p < 0.001; *stat5a,* log2FC = 1.55, p = 0.114*)*, as was the toll-like receptor 21 gene (*tlr21*, log2FC = 2.20, p = 0.038). Genes of the Janus-kinases group (*jak*) were also upregulated. Of note, the negative regulator of the Janus kinase - signal transducer and activator of transcription (JAK-STAT) signalling pathway, suppressor of cytokine signalling 1a (*socs1a*) was downregulated (log2FC = -1.55, p = 0.332).

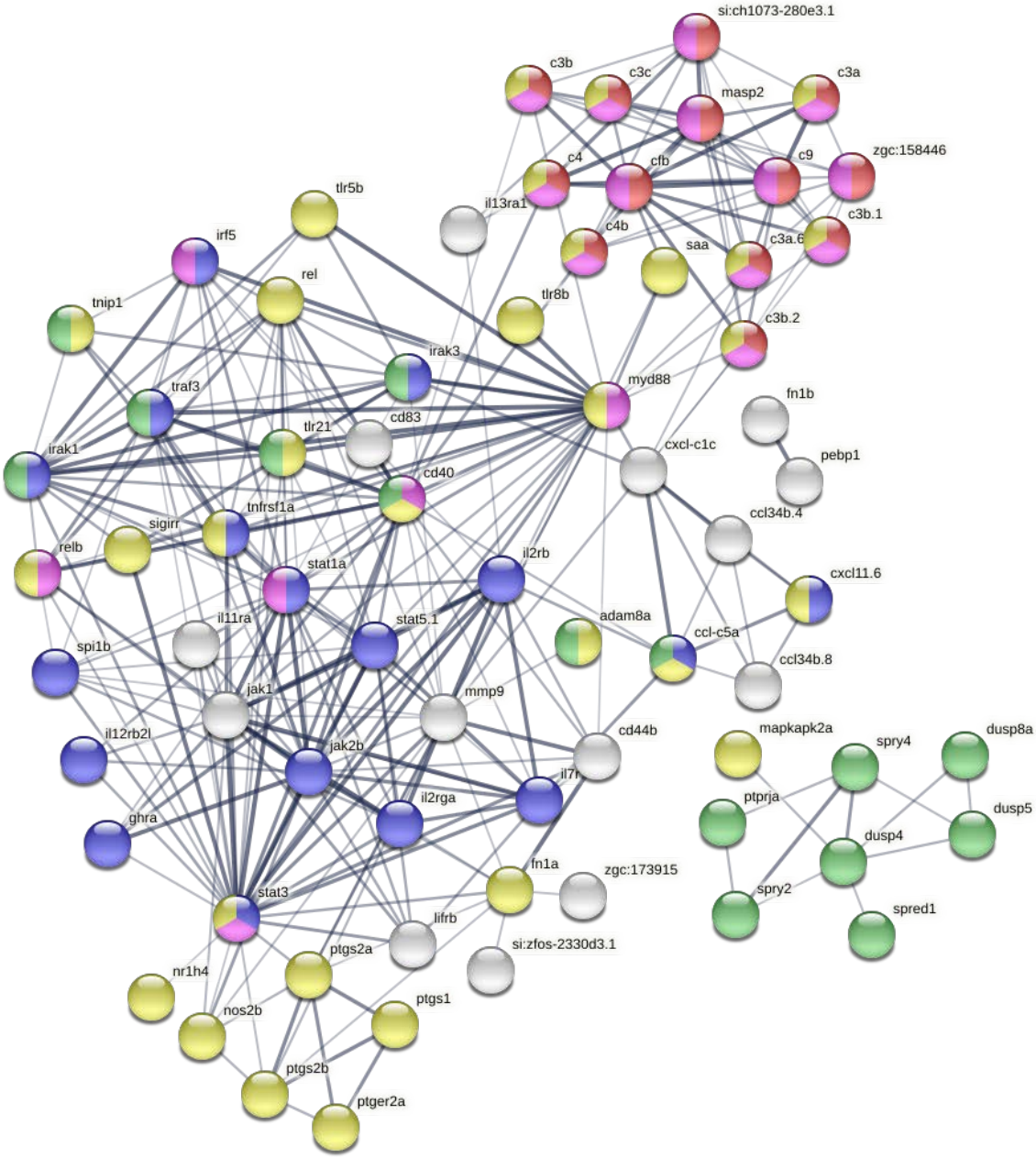


Figure 6.27. Immune response in *cyp21a2-/-* male livers
Visualisation in STRING of interactions between proteins encoded by dysregulated genes involved in immune responses and MAPK signalling, that are differentially expressed in adult male WT and *cyp21a2-/-* mutant livers, identified by GSEA. The connecting lines indicate interactions between genes, the width of the lines being proportional to the confidence or the strength of data supporting the interaction. The colours of the nodes indicate the main sub-processes to which they belong: immune effector process (pink, n=21), complement activation (red, n=13), inflammatory response (yellow, n=36), cytokine mediated signalling pathway (blue, n=18) and regulation of MAPK cascade (green, n=17).

Because the focus of this project was on metabolic disease induced by GC deficiency,the study analysed specifically metabolic processes identified by the GSEA (**Figure 6.28**), the list of upregulated processes was very limited containing only sterol metabolism (including cholesterol biosynthesis), collagen biosynthesis and phospholipid dephosphorylation. The processes that were most significantly downregulated were those involved in the energy homeostasis including mitochondrial respiratory chain assembly, ATP metabolism, generation of precursor metabolism and energy, tricarboxylic acid cycle, beta-oxidation of fatty acids, as well as peptide synthesis, lipid, protein and amino-acid catabolism.

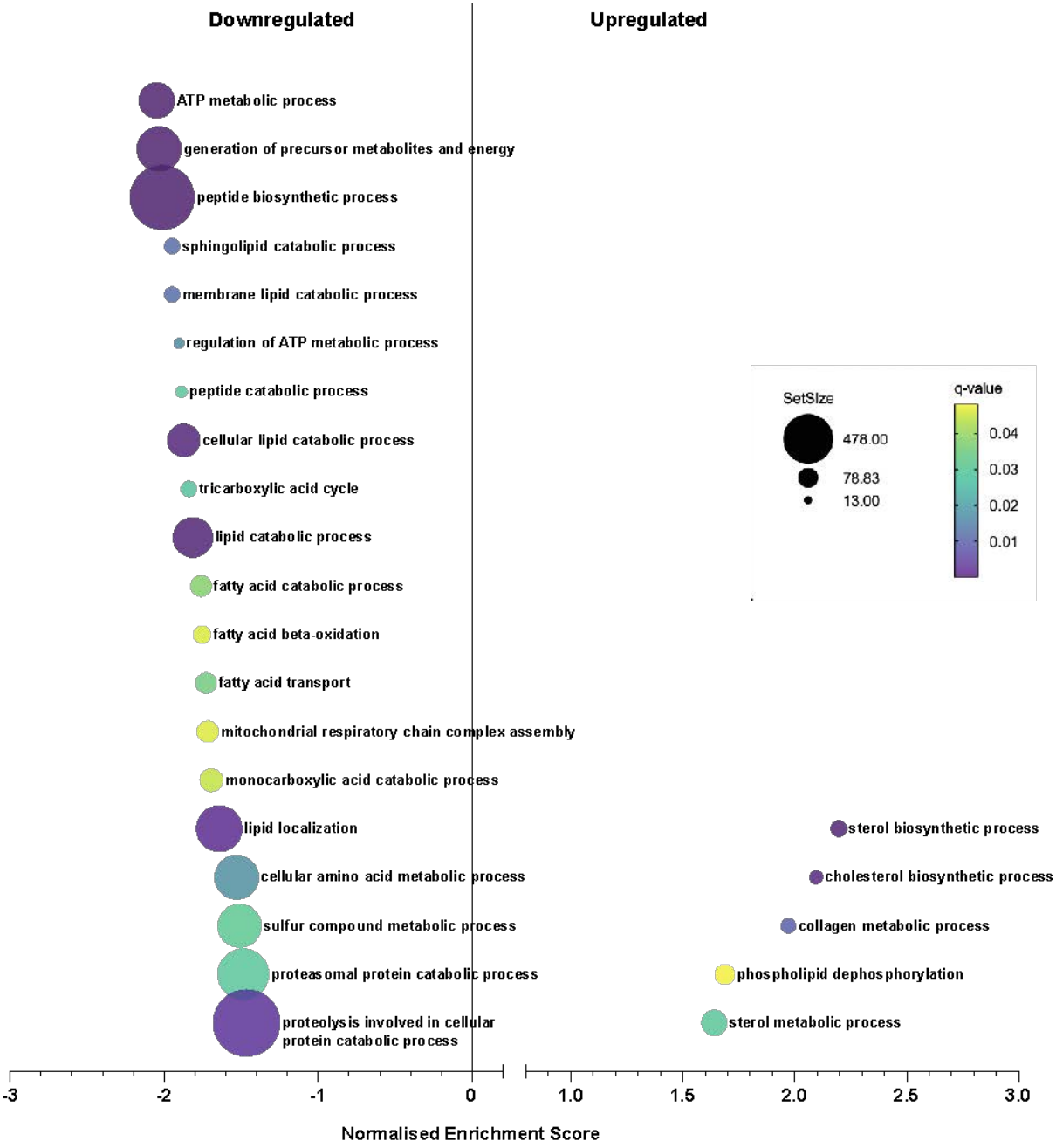


Figure 6.28. Dysregulated metabolic processes in *cyp21a2-/-* male livers
The processes presented are a manual selection of the top up- and down-regulated metabolic processes from the list of GSEA of results. Downregulated gene sets are indicated on the left side and upregulated ones on the right side. The horizontal line indicates the normalised enrichment score. The size of the dots indicates the size of the gene set, and the colour corresponds to the q-value, decreasing from yellow to violet.

The genes corresponding to peptide biosynthesis were all included in the “translation” GO term, thus they were already presented above (**Figure 6.26**), as were the ATP metabolism and generation of precursor metabolism and energy (**Figure 6.27**), being

among the top 20 down-regulated biological processes in mutant livers. Additionally, there was significant downregulation of several genes and pathways involved in lipid metabolism (**Figure 6.29**). Thus, for lipid catabolism, phospholipase D1b (*pld1b*, log2FC = -2.44, p = 0.003), glycerophosphodiester phosphodiesterase domain containing 1 (*gdpd1*, log2FC = -1.96, p < 0.001), N-acylsphingosine amidohydrolase 2 (*asah2*, log2FC = -6.51, p = 0.009), including the important regulators of fatty acid beta-oxidation, acyl-CoA synthetase long chain family member 5 (*acsl5*, log2FC = -8.35, p < 0.001) and carnitine palmitoyltransferase 1B (muscle) (*cpt1b*, log2FC = -2.91, p < 0.001), were amongst the most significantly downregulated genes. Downregulated genes involved in lipid localisation included the CD36 molecule (thrombospondin receptor) (*cd36*, log2FC = -9.69, p < 0.001), apolipoprotein A-IV a (*apoa4*, log2FC = -7.83, p = 0.003), fatty acid binding protein 2, intestinal (*fabp2*, log2FC = -9.10, p < 0.001), vitellogenin family, the most significantly downregulated being vitellogenin 5 (*vtg5*, log2FC = -4.80, p < 0.001), oxysterol binding protein-like 3a (*osbpl3a*, log2FC = -1.12, p = 0.024), fatty acid binding protein 1b, tandem duplicate 1 (*fabp1b.1*, log2FC = -4.80, p < 0.001), annexin A2b (*anxa2b*, log2FC = -4.86, p < 0.001). The number of suppressed genes involved in regulating lipid biosynthesis was modest.

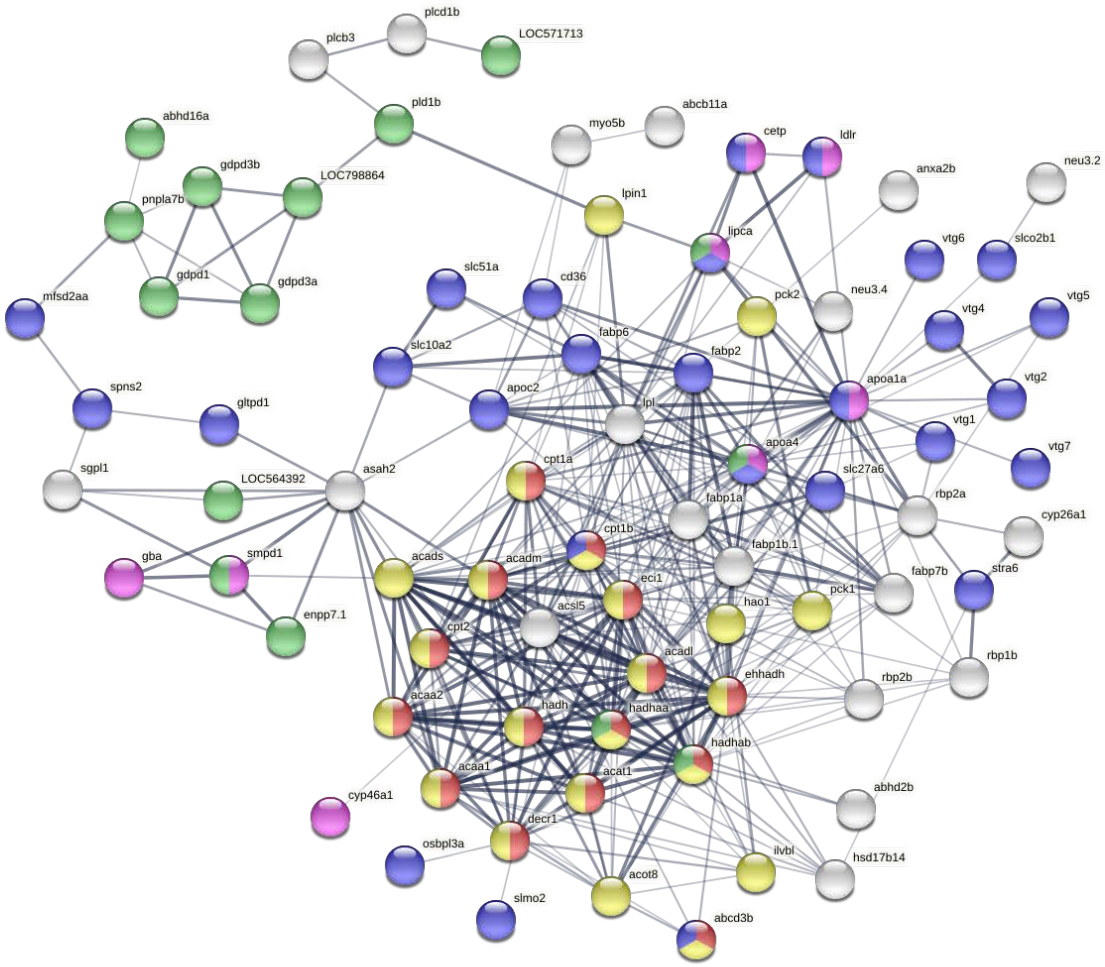


Figure 6.29. Lipid metabolism in *cyp21a2-/-* male livers
Visualisation in STRING of interactions between proteins encoded by dysregulated genes involved in lipid metabolism, that are differentially expressed in adult male WT and *cyp21a2-/-* livers, identified by GSEA. The connecting lines indicate interactions between genes, the width of the lines being proportional to the confidence or the strength of data supporting the interaction. The colours of the nodes indicate the main sub-processes: lipid transport (blue, n=37), cholesterol metabolism (pink, n=9), fatty acid catabolism (yellow, n=22) including fatty acid ß-oxidation (red, n=15), and phospholipid metabolism (green, n=17).

Protein catabolism was another importantly suppressed process, with reduced expression of genes involved in the regulation of several sub-processes: synthesis of proteasome subunits: proteasome 26S subunit, non-ATPase 11b (*psmd11b*, log2FC = -1.17, p = 0.004), lysosome functions: cathepsin Z (*ctsz*, log2FC = -0.90, p = 0.003), protein deubiquitation: ubiquitin carboxyl-terminal esterase L1 (ubiquitin thiolesterase)

(*uchl1*, log2FC = -2.68, p < 0.001), protein folding (pink, n=7) and protein ubiquitation: speckle type BTB/POZ protein like b (*spoplb*, log2FC = -3.16, p < 0.001) (**Figure 5.30**).

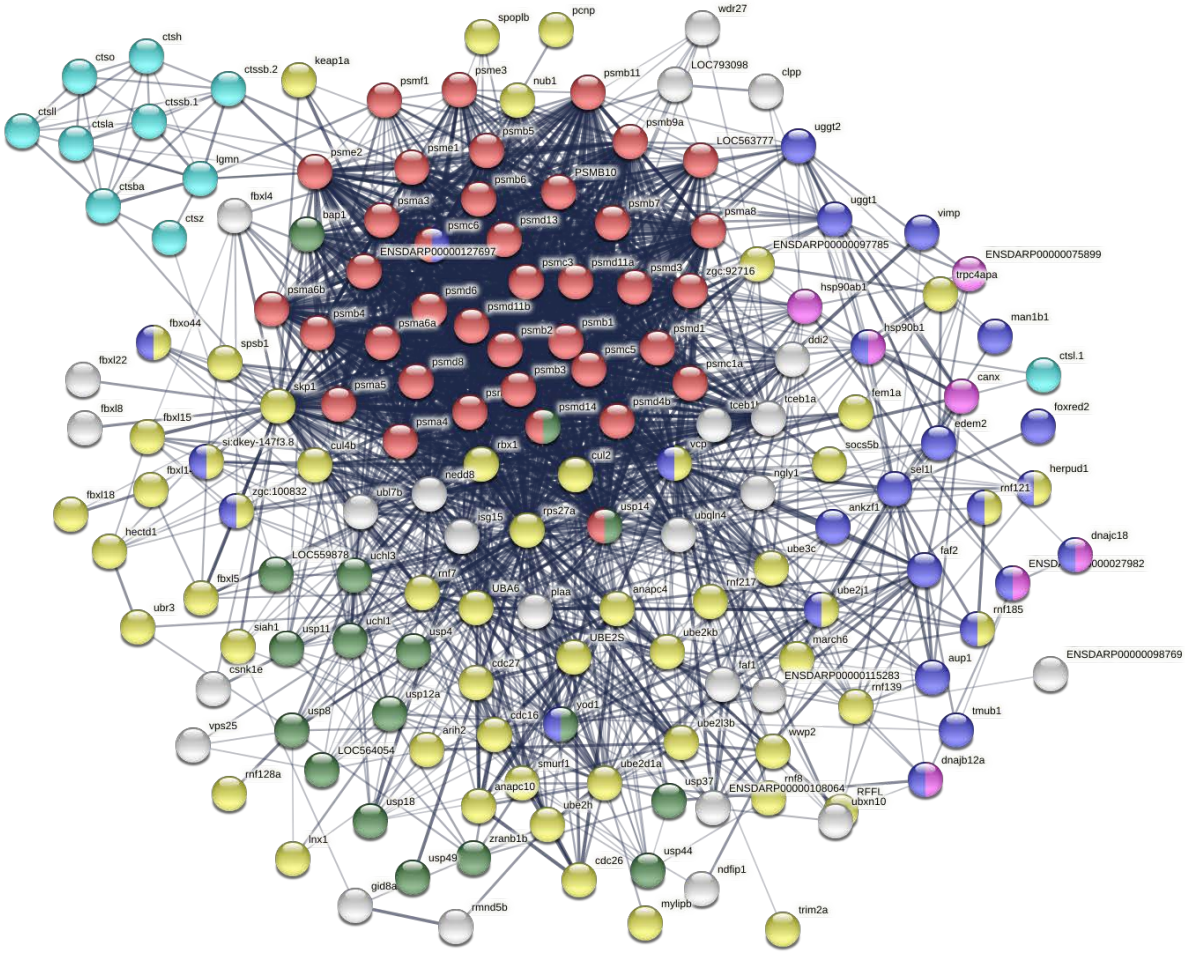


Figure 6.30. Protein catabolism in *cyp21a2-/-* male livers
Visualisation in STRING of interactions between proteins encoded by dysregulated genes involved in cellular protein catabolism, that are differentially expressed in adult male WT and *cyp21a2-/-* mutant livers, identified by GSEA. The connecting lines indicate interactions between genes, the width of the lines being proportional to the confidence or the strength of data supporting the interaction. The colours of the nodes indicate the main sub-processes to which they belong: proteasome assembly (red, n=35), protein deubiquitation (green, n=17), protein folding (pink, n=7), protein ubiquitation (yellow, n=56), lysosome functions (light blue, n=10), and ERAD pathway (dark blue, n=25).

6.1.1.3.4 Associations between genes exhibiting differential expression in adult male wild type and *cyp21a2-/-* mutant livers and human disease

The association between transcriptomic dysregulation in the mutant livers and human pathology was explored using the DGEs in livers to generate a list of human disease-associated orthologues (**Figure 6.31**). The predominant pathology was cardiovascular disease and metabolic diseases; however, inflammatory bowel disease and cancer were also among the most common human conditions linked to the list of differentially expressed genes.

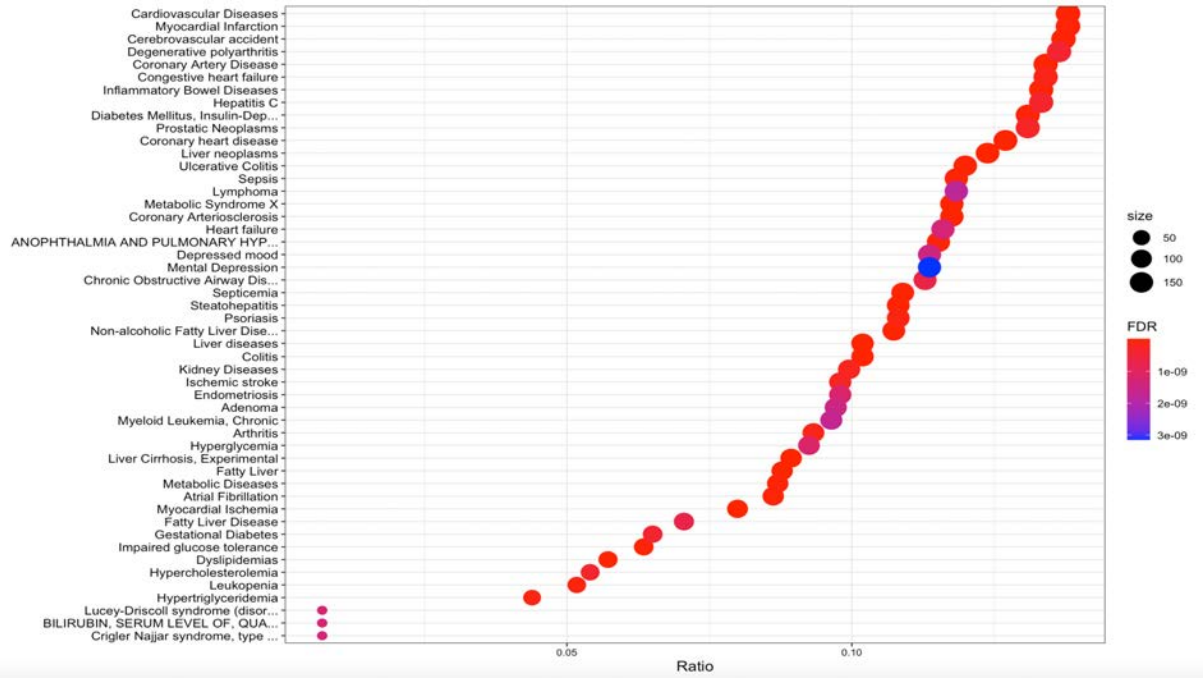


Figure 6.31. Associations between liver *cyp21a2-/-* transcriptome and human disease
Enrichment plot showing the associations between the transcriptomic changes in *cyp21a2-/-* adult liver and human diseases, produced in R using the "disgenet2r" package. The horizontal axis shows the ratio between the number of DEGs in a category and all DGEs, the size of the dots represents the number of DEGs in a category and the colour, the *p-adj* value.

6.1.2 Transcriptomic analysis of *cyp11a2-/-* adult livers

In order to find if the results of the transcriptomic analysis of the *cyp21a2-/-* are consistent between different zebrafish models of cortisol deficiency, RNA sequencing and transcriptomic analysis was repeated using a *cyp11a2-/-* zebrafish. This mutant line

corresponds to side chain cleavage enzyme deficiency, blocking the first step of the steroid synthesis and causing steroid deficiency (Li, Oakes et al., 2020).

The results of the sample quality check were very similar in *cyp11a2* compared to *cyp21a2* mutants, with good per base sequence quality and per sequence quality scores, as well as failing per base sequence content due to the bias in the first 12 bases, and the CG content module due to left skewing (**Figure 6.32)**.

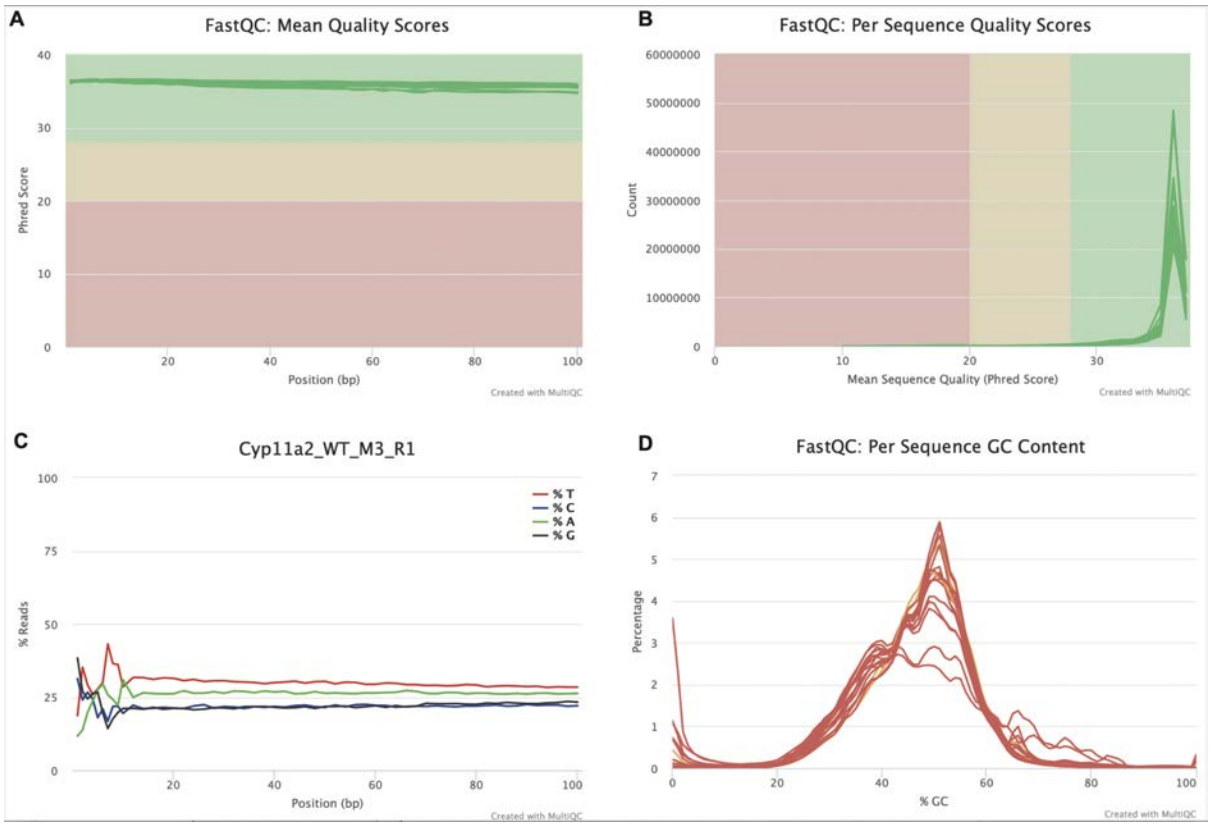


Figure 6.32. Quality of sequencing data in *cyp11a2* male adult livers
A. Per base mean quality expressed as the Phred score within the read, showing consistently scores above 30, suggesting high-quality data. **B**. The per sequence quality histogram was shifted to the right towards higher Phred scores. **C**. Individual per sequence base content (WIld type Male 3) failed due to the bias in the first 12 bases characteristic to Illumina sequencing. **D**. All samples alarmed or failed the per sequence GC content check due to a consistent skewing to the left of the GC content curve.

Principal component analysis (PCA) (**Figure 6.33.A**) and cluster dendrogram (**Figure 6.33.B**) showed good distinction between WT and mutant livers, with obvious clustering according to the first principal component, which accounted for 58% of variance.

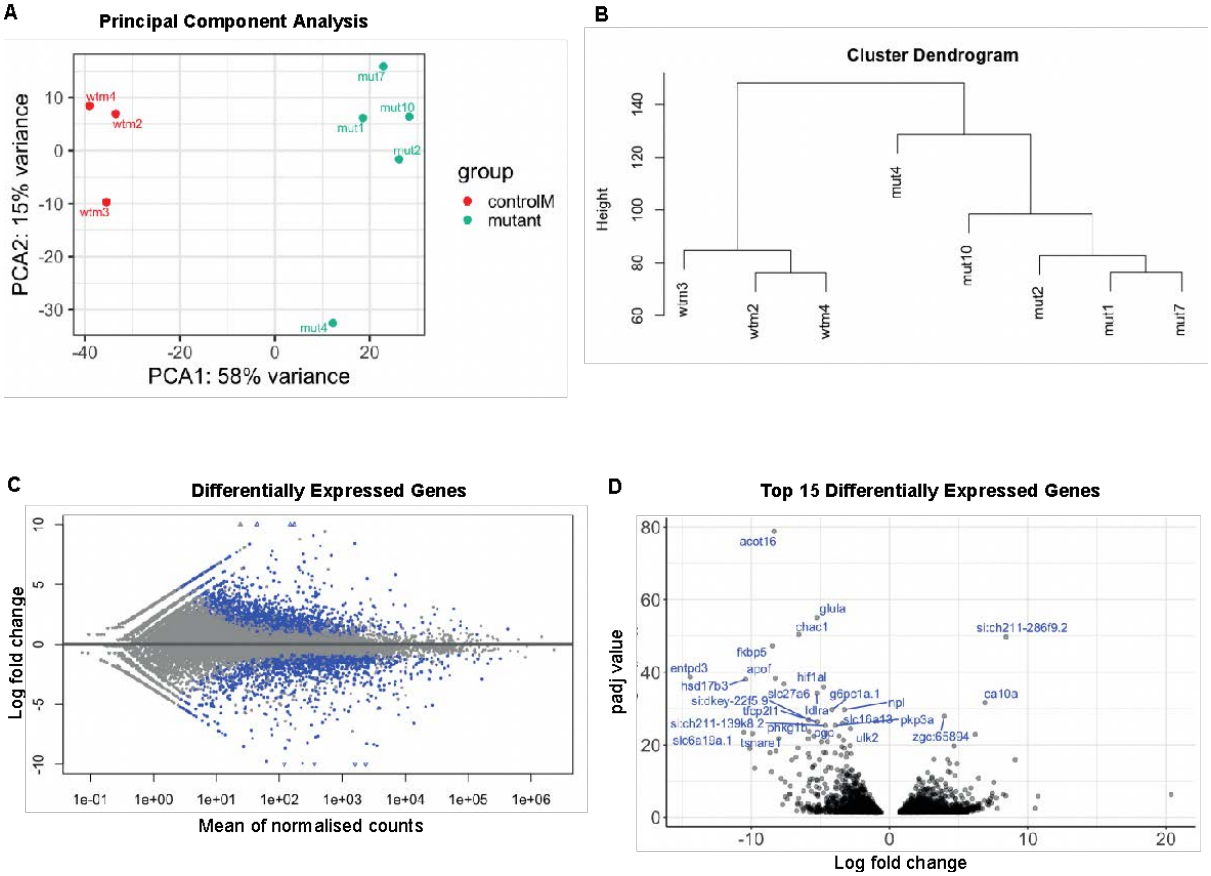


Figure 6.33. Inter-sample variations in the *cyp11a2* adult livers
A. Principal Component Analysis; **B**. Cluster dendrogram. **C**. MA plot showing differentially expressed genes; the statistical significance (*p-adj < 0.05*) is indicated by the blue dots while differentially expressed genes without statistical significance are coloured grey. **D.** Volcano plot showing the top 25 DEGs in mutant livers, with statistical significance (*p-adj)* represented on the vertical axis and the LFC on the horizontal one.

Differential gene expression (DGE) analysis identified 1653 genes that were significantly upregulated (log2 fold change (LFC) > 0, *p < 0.05*) and 1451 genes significantly downregulated (LFC < 0) (**Figure 6.33.D**). The most dysregulated genes are shown in the volcano plot (**Figure 6.33.C**), with *fkbp5* being within the five most downregulated genes (LFC = -8.46, padj = 5.946e-48).

GO enrichment was initially performed using the GOrilla (Gene Ontology Enrichment Analysis and Visualisation), setting the statistical significance threshold as $p < 0.001$. In the top first 15 terms identified, 14 pertained to metabolic compound metabolism or transport (**Table 6.7**).

Table 6.7. Results of the gene ontology analysis for biological processes in GOrilla in *cyp11a2* livers

GO Term	Description	Enrichment (N,B,n,b)	p-value	FDR q-value
GO:0015711	Organic anion transport	3.9 (13653, 160, 568, 26)	1.50E-08	1.28E-04
GO:0046942	Carboxylic acid transport	4.4 (13653, 103, 623, 21)	1.06E-07	4.51E-04
GO:0015849	Organic acid transport	4.3 (13653, 105, 623, 21)	1.46E-07	4.16E-04
GO:0006820	Anion transport	3.0 (13653, 244, 569, 31)	2.30E-07	4.91E-04
GO:0008152	Metabolic process	1.1 (13653, 4709, 1569, 621)	9.58E-05	1.64E-01
GO:0015718	Monocarboxylic acid transport	4.7 (13653, 43, 737, 11)	1.18E-04	1.68E-01
GO:0044281	Small molecule metabolic process	9.1 (13653, 891, 10, 6)	1.39E-04	1.69E-01
GO:0098656	Anion transmembrane transport	3.6 (13653, 98, 569, 15)	1.47E-04	1.57E-01
GO:0048511	Rhythmic process	2.4 (13653, 85, 1550, 24)	1.82E-04	1.73E-01
GO:0006114	Glycerol biosynthetic process	21.6 (13653, 3, 630, 3)	1.90E-04	1.62E-01
GO:0016042	Lipid catabolic process	2.7 (13653, 124, 888,22)	1.90E-04	1.35E-01
GO:0042398	Cellular modified amino acid biosynthetic process	5.6 (13653, 23, 847, 8)	3.05E-04	2.00E-01

GO:0015760	Glucose-6-phosphate transport	62.0 (13653, 2, 220, 2)	4.03E-04	2.45E-01
GO:1901575	Organic substance catabolic process	1.4 (13653, 791, 1514, 123)	4.79E-04	1.95E-01
GO:0044237	Cellular metabolic process	1.1 (13653, 4223, 1569, 555)	6.37E-04	2.47E-01

The **p-value** is the enrichment p-value computed according to the mHG or HG model. The **FDR q-value** is the correction of the p-value for multiple testing using the Benjamini and Hochberg (1995) method. The **Enrichment = (b/n) / (B/N) (**N - total number of genes, B - total number of genes associated with a specific GO term, n – a flexible cut-off, being an automatically determined number of genes in the input list or 'target set', b - is the number of genes in the target set that are associated with the GO term).

The gene ontology analysis in R using "clusterProfiler", produced more heterogeneous results, with a predominance of processes involved in the mitochondrial function and provision of energy derivates (**Figure 6.34**).

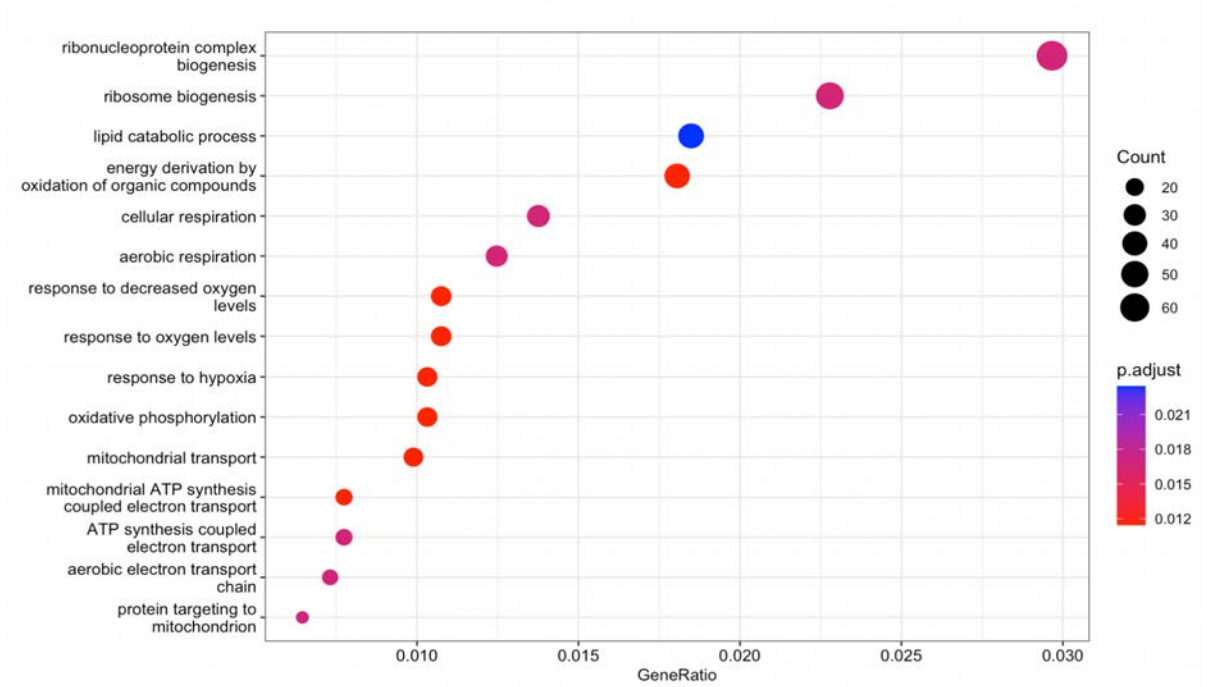


Figure 6.34. Gene ontology over-expression analysis in *cyp11a2-/-* livers
Enrichment plot produced in R using the "clusterProfiler" package. The horizontal axis shows the ratio between the number of DEGs in a category and all DGEs, the size of the dots represents the number of DEGs in a category and the colour, the *p-adj* value.

The GSEA results were similar to those of the GO enrichment analysis (**Figure 6.35**). The most downregulated GO terms were those related to ribosome synthesis and function, oxidative phosphorylation and energy homeostasis, as well as a number of metabolic processes such as fatty acid catabolism and small molecule catabolism. The most significant downregulation was found in processes related to the cell cycle, cell proliferation and differentiation.

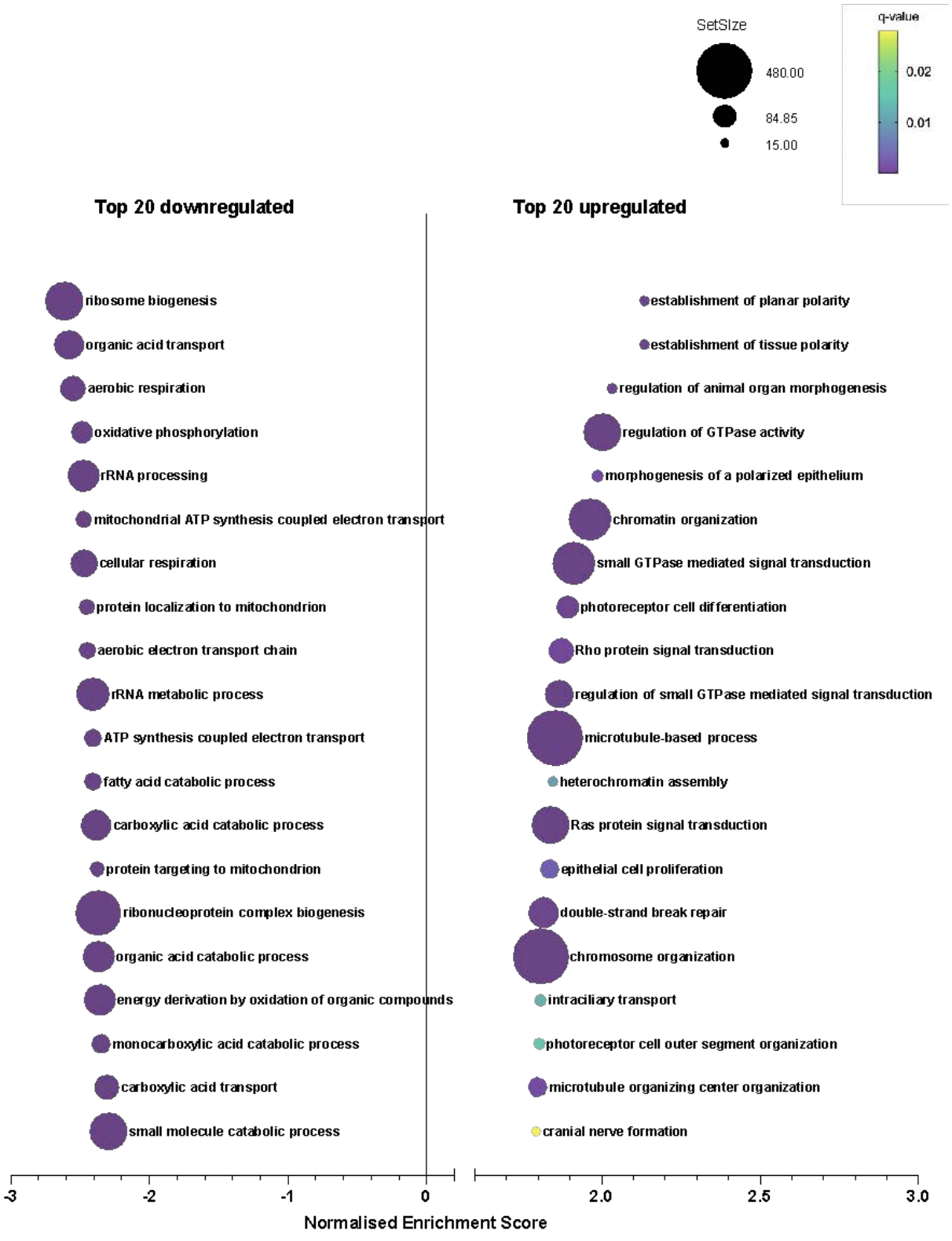


Figure 6.35. Gene set enrichment analysis in *cyp11a2-/-* adult livers
Bubbleplot produced GraphPad Prism showing the 20 most downregulated (left) and upregulated (right) biological processes in *cyp11a2* male livers. The selection was based exclusively on the normalised enrichment score (showed on the horizontal axis). The size of the dots indicates the size of the gene set, and the colour corresponds to the q-value, decreasing from yellow to violet.

The downregulation of different processes involved in the provision of energy derivates (**Figure 6.36**) was similar to that seen in the *cyp21a2* livers. The mitochondrial oxidative phosphorylation system was supressed, including several subunits of the mitochondrial respiratory complex I, of which the most significant were NADH:ubiquinone oxidoreductase complex assembly factor 1 (*nudfaf1*, log2FC = -1.26, p = 0.004) and NADH:ubiquinone oxidoreductase subunit S6 (*ndufs6*, log2FC = -1.52, p = 0.001), core subunit S2 (*nudfs2*, log2FC = -1.09, p = 0.013), subunit B9 (*nudfb9*, log2FC = -1.10, p = 0.007 subunit B2 (*nudfb2*, log2FC = -1.65, p = 0.001), the cytochrome-c oxidoreductase complex, most significantly cytochrome c oxidase subunit 7A2a (*cox7a2a*, log2FC = -1.21, p = 0.013), cytochrome c oxidase subunit 5Aa (*cox5aa*, log2FC = -1.02, p = 0.032), cytochrome c oxidase subunit 5Ba (*cox5ba*, log2FC = -1.44, p = 0.008), cytochrome c oxidase subunit 6A1 (*cox6a1*, log2FC = -1.03, p = 0.032), as well as the tricarboxylic acid cycle, including aconitase 2, mitochondrial (*aco2*, log2FC = -1.13, p = 0.001) . Additionally, several glycolytic enzymes were also downregulated, such as aldolase b, fructose-bisphosphate (*aldob*, log2FC = -1.69, p = 0.032) and TP53 induced glycolysis regulatory phosphatase a (*tigara*, log2FC = -0.79, p = 0.024), as were different factors involved in glycogen metabolism such as phosphorylase kinase, gamma 1a (*phkg1a*, log2FC = -5.10, p < 0.0001), phosphorylase kinase, gamma 1b (*phkg1b*, log2FC = -5.81, p < 0.001), and phosphorylase kinase, alpha 2 (*phka2*, log2FC = -2.34, p = 0.003).

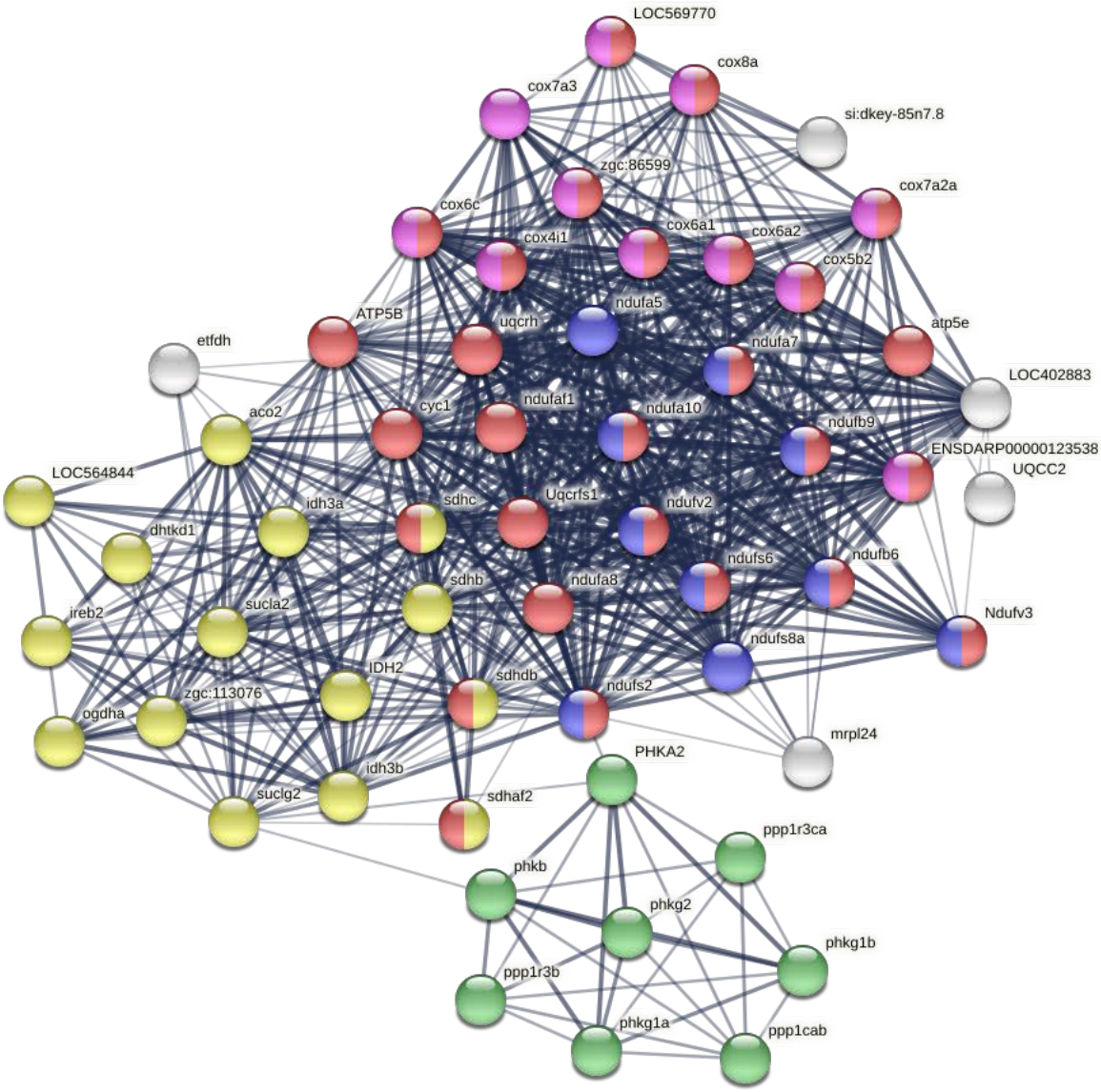


Figure 6.36. ATP metabolism in *cyp11a2-/-* livers
Visualisation in STRING of dysregulated genes involved in liver ATP metabolism identified by GSEA. The connecting lines indicate interactions between genes, the width of the lines being proportional to the confidence or the strength of data supporting the interaction. The colours of the nodes indicate the main sub-processes to which they belong: all oxidative phosphorylation (red, n=28), NADH dehydrogenase (ubiquinone) activity (blue, n=10), cytochrome-c oxidase activity (pink, n=11), tricarboxylic acid cycle (yellow, n=15), and glycolytic process (green, n=8).

The association between transcriptomic dysregulation in the mutant livers and human pathology was explored using the DGEs in livers to generate a list of human orthologues (**Figure 6.37**). The associated human conditions included several metabolic and liver

diseases, such as metabolic syndrome, fatty liver and steatohepatitis, as well as neoplasia.

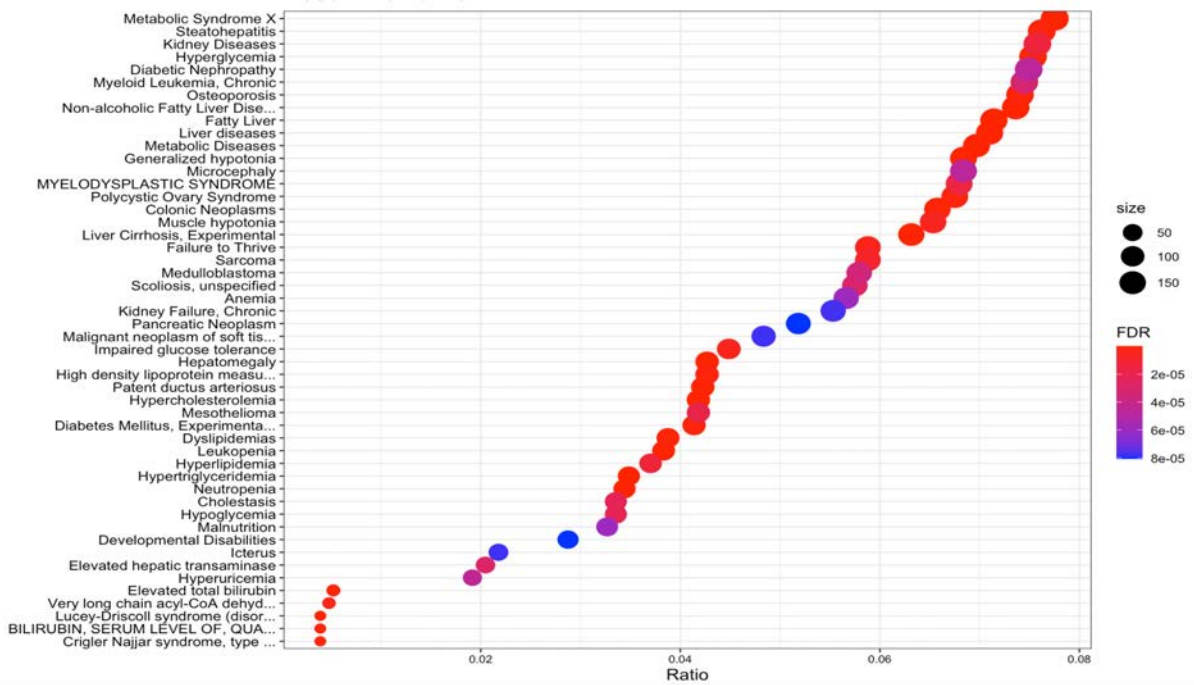


Figure 6.37. Associations between liver *cyp11a2-/-* transcriptome and human disease

Enrichment plot showing the associations between the transcriptomic changes in *cyp11a2-/-* adult liver and human diseases, produced in R using the "disgenet2r" package. The horizontal axis shows the ratio between the number of DEGs in a category and all DGEs, the size of the dots represents the number of DEGs in a category and the colour, the *p-adj* value.

6. 2 Discussion

6.2.1 The quality of the RNA sequencing data

Overall, the quality of the sequencing data was high, based on the per base sequence quality and per sequence quality scores. The initial problem of overexpressed adapter sequences was effectively resolved by trimming of the data using Trimmomatic. Two of the quality metrics were consistently failed, respectively the GC content and the sequence duplication level. The left skewing of the GC content can be related to high content of rRNA, however, the samples were subjected to removal of rRNA by SortmeRNA. Another explanation could be sample contamination (bacterial or fungal material), however, as the

left skewing was consistent among all samples, it is more likely that the high GC content is related to the library preparation process, in particular the PCR amplification (Browne, et al., 2020). High sequence duplication level may be similarly caused by the PCR amplification involved in the library preparation protocol; however, it can also be triggered by the difference between over- and under-expressed genes. It is generally recommended that sequence duplications should not be removed as it may reduce the power of the analysis (Parekh, et al., 2016). Thus, the results of the quality checks by FastQC and MultiQC were interpreted as validating the sequencing data and allowing the progress to the next steps in data analysis.

6.2.2 Differential gene expression results and comparison to qPCR

There was a very clear distinction in gene expression between wild type and mutant larval samples, with over 1000 genes found to be significantly differentially expressed in mutants. Almost two thirds of the DEG's were downregulated in cortisol-deficient larvae, the most significant one being FK506 binding protein 5 (*fkbp5*). *Fkbp5* plays a major role in the GR signalling dynamics, being part of the intracellular negative feed-back loop that regulates GR function (Gans, et al., 2021). Thus, GR activates *fkbp5* transcription, which in turn sequesters GR in the cytoplasm, causing resistance to the GC action. Consequently, *fkbp5* is frequently used as a marker of cortisol action/absence in gene expression experiments, as demonstrated by the results presented herein.

In livers of wild-type and *cyp21a2-/-* mutant adults, sex-specific differential gene expression obscured any gene expression differences related to genotype, thus leading to the decision to focus on male livers for the analysis. The DGE between sexes identified by the transcriptomic analysis is in keeping with the finding of larger livers in females,

further supporting the hypothesis that metabolic pathways and functions are regulated differently between male and female zebrafish. The decision to focus primarily on males was due to the greater variability of female metabolic homeostasis, due to the substantial metabolic demands of egg production. Even between males, in the overall DGE study the principal component analysis and cluster dendrogram showed a less clear distinction between mutants and wild type livers, in comparison to the results from larvae. As sample quality and sequencing depth were similar, the differences were likely induced by inter-individual variability, which was also evident in the qPCR results.

The comparison of targeted gene expression between the results of the qPCR experiments and the gene expression analysis following RNA sequencing showed wide inconsistencies for both glucose and fat metabolism. There was however a consensus regarding the suppression of gluconeogenesis, explored here by targeting the glucocorticoid-responsive gene *pck1*, which is in keeping with previous studies conducted on adrenalectomised murine models (Kawai, 1977; Bishayi and Ghosh, 2003). Interestingly, this result contrasted with the findings of both qPCR and RNA sequencing in male adult livers, although the downregulation was present in females. Since other endocrine factors including insulin and cAMP influence the expression of *pck1*, it may be argued such differences are possibly due to the dynamic nature of the metabolic dysregulations in a developing organism and to sex-related differences. Overall, the main benefit of comparing the results of the qPCR experiments to the gene expression analysis following RNA sequencing is that is highlights the limited insight obtained by exploring a small number of “representative’ genes, compared to the rich and complex information provided by the gene ontology analysis offered by the latter. Consequently, qPCR may

serve as preliminary exploratory work, to be followed by the robust and comprehensive RNA sequencing-based Gene Ontology analysis.

6.2.3 Dysregulation of biological processes in *cyp21a2* deficient zebrafish

Two different strategies were used to explore the biological processes that were dysregulated in zebrafish as a consequence of cortisol deficiency, namely gene ontology over-representation analysis and gene set enrichment analysis. This two-strategy approach was used as a way to identify robust findings. It is important to acknowledge that there is an ever-developing range of bioinformatic analysis and functional genomics techniques intended for the interpretation of large data generated by high-throughput sequencing. They all share the challenge of transforming an enormous amount of raw data into results that are statistically sound and contextually pertinent. While there are recognised advantages and disadvantages to all such techniques, some of the published literature (Tarca, et al., 2013) appears to favour GSEA, commenting on the drawbacks of over-representation analysis, of which the most important are the need to establish externally a threshold for statistical significance and the commonly violated assumption of independence among genes which may lead to false positive results. By contrast, GSEA, belonging to the functional class scoring group of methods, avoids the exclusion of genes, deriving the statistical significance of differential gene expression internally. Thus, in this project, while both methods were used for insight into the biological processes that were dysregulated in mutants, it was the findings of the GSEA that were used for further analysis.

6.2.3.1 Transcriptomic analysis of *cyp21a2-/-* larvae

Both methods used for the gene ontology over-representation analysis placed metabolism among the most significantly dysregulated biological processes, alongside the regulation of cell cycle, and in particular mitosis. Small molecule metabolism was the most importantly affected metabolic process in both types of GO analysis. Small molecule metabolites are the basic precursors and key products of metabolic processes (Hu, et al., 2020). The genes pertaining to this GO term that were dysregulated in *cyp21a2-/-* larvae were relatively evenly distributed among lipid, carbohydrate and amino acid metabolism, suggesting a wide metabolic impact of cortisol deficiency affecting multiple interrelated pathways.

GSEA provided further insight by indicating which over-represented processes were up- or downregulated. Holistically, in larvae, there was a clear trend towards the upregulation of genes involved in cell division, in particular mitosis, with the inclusion of biological processes such as chromosome and chromatid segregation and separation, DNA replication, suggesting elevated cell proliferation and growth. By contrast, the list of the most severely downregulated processes was dominated by those involved in basic cell housekeeping functions such as protein synthesis (ribosome biogenesis and, rRNA metabolism), energy homeostasis (mitochondrion organisation), response to external stimuli (response to hormones, temperature stimulus, reactive oxygen species), consistent with an expectation of a rising demand for these cellular activities in stressful situations.

There is limited published evidence regarding the impact of cortisol insufficiency on the **cell cycle**. It was found that in mammalian breast epithelium, endogenous GC induced

the cell cycle inhibitor, cyclin-dependent kinase (Cdk) inhibitor p21(Cip1), and suppressed the activator of the MAPK pathway, serine/threonine kinase 39 (Stk 39), thus downregulating cell proliferation (Hoijman, et al., 2012). The role of GC in the regulation of the cell cycle is further supported by the more abundant evidence on the effects of external GC in neoplastic conditions. Thus, it is known that GC treatment suppresses proliferation of several types of neoplastic cells, in general by producing a G1/G0 arrest of the cell cycle (Goya, et al., 1993; Greenberg, et al., 2002; Mattern, et al., 2007). The GSEA of the 5 dpf cyp21a2-/- larvae suggest that cortisol deficiency may alter cell proliferation, warranting further studies to monitor cell proliferation rates in this mutant line.

The results of the GSEA highlighted the downregulation of **mitochondrial organisation** as one of the most prominent effects of cortisol deficiency in larvae. This result is in keeping with the major role that GC are known to play in energy homeostasis, especially in the regulation of the response to stress. Due to the crucial role played by the mitochondria in energy homeostasis, such extensive dysregulations of its function at such an early stage of development are likely to have long term impact, including maladaptive effects on metabolic pathways. The interrelations between the stress axis and the function of the mitochondria have been explored by research, with a particular focus on defining the effects of psychosocial stress (Picard and McEwen, 2018). It is well established that GC are involved in the regulation of the mitochondrial function due to the translocation of the GR to the mitochondria (Du, et al., 2009). The impact of acute and chronic stress, as well as the effects of GC treatment on mitochondrial function has been repeatedly explored by animal and in vitro studies (Du, et al., 2009; Psarra and Sekeris, 2011; Hunter, et al., 2016). Thus, it was shown that in the brain, within the mitochondria, GR

forms a complex with the antiapoptotic protein Bcl2, regulating cell plasticity and resilience (Du, et al., 2009). Importantly, the genes coding members of the bcl2 protein family were significantly downregulated in the study, suggesting that apoptosis may be defective as a result of cortisol deficiency. It was also demonstrated that in hepatocytes, GC regulate mitochondrial transcription both directly through the mitochondrial GR and by inducing nuclear transcription of genes encoding mitochondrial transcription factors (Psarra and Sekeris, 2011). While there is limited evidence on the impact of cortisol deficiency on mitochondrial function, it was shown that in adrenalectomized rats, acute stress-induced mtRNA regulation was abolished (Hunter, et al., 2016). Moreover, it is suspected that there is a dose dependent character to the GC actions on mitochondrial function with an “inverted U-shape response” (Du, et al., 2009) which would indicate that cortisol deficiency and excess GC could elicit similar effects. The results showed that there was a suppression of the mitochondrial organisation that extended to a number of functions, including apoptosis, respiratory chain assembly and oxidation-reduction process, and protein transport, which would translate into poor efficiency in generating energy and defective programmed cell death.

A significant impact of cortisol deficiency found in *cyp21a2-/-* mutant larvae was the downregulation of **ribosome biogenesis and rRNA processing**. There is very limited evidence on the effects of GC on ribosome biosynthesis, although in vivo research showed that GC reduce the synthesis of ribosomal proteins and impair transcription of rRNA in lymphosarcoma cells (Meyuhas, et al., 1990). Since the introduction of synthetic GC in the treatment of various conditions, their effects on protein synthesis have been explored with keen interest. Given the wide distribution and actions of the GR, it is perhaps not surprising that GC regulates protein synthesis in a variety of different ways.

Thus, cortisol was found to the downregulate ribosomal protein synthesis by suppressing the mTOR pathway and especially protein S6 kinase beta-1 (S6K1) in vitro (Shah, et al., 2002) and in vivo in sheep foetuses (Jellyman, et al., 2012). Similarly, in muscle, GC reduced protein synthesis, again by the regulation of the mTOR pathway with downstream suppression of the ribosomal S6K and eukaryotic translation initiation factor 4E-binding protein 1 (4E-BP1) (Schakman, et al., 2013; Xu, et al., 2022). Similar suppressive effects of GC on protein synthesis were reported by in vitro evidence using skin fibroblasts (Smith, 1988) and in vivo murine models in the liver (Kim and Kim, 1975). By contrast, GC were found to increase protein synthesis in the heart in murine models (Clark, et al., 1986; Savary, et al., 1998). The impact of cortisol deficiency on protein synthesis was less extensively explored, although in zebrafish, an impaired HPA axis due to mutations of POMC was shown to increase protein synthesis in the muscle (Shi, et al., 2020). Thus, the transcriptomic results from the *cyp21a2-/-* fish are not in keeping with the majority of published studies, indicating that cortisol deficient larvae (and adult livers) have reduced protein synthesis through a suppression of the ribosome synthesis and function. This could be explained by differences in methodology, as the majority of the previous studies focused on exploring targeted gene expression or protein phosphorylation, with limited evidence on the GC effects on the genome. Moreover, most of these studies analysed the effects of synthetic exogenous GC on protein synthesis, rather than the effects of GC deficiency, and we must consider again the possibility of a "inverted U-shape response" depending on the dose and timing of GC administration.

There were very few processes pertaining to **lipid metabolism** that were identified by the GSEA to be dysregulated in larvae. The most significant findings relate to the upregulation of cholesterol/sterol synthesis (*hmgcra*), lipid transport (*apoa1a*) and protein lipidation.

The reduced impact of cortisol deficiency on lipid metabolism in 5 dpf larvae may be explained by residual fat deposits in the yolk, which supply lipids to the developing larva. Thus, the regulation of these processes by cortisol may still be at a more incipient stage, while cholesterol synthesis may be upregulated as a compensatory mechanism to the cortisol absence in the mutant progenitors and larvae.

By contrast, **carbohydrate metabolism** was well represented among the list of downregulated biological processes. The findings indicate reduced energy production from carbohydrates in hypocortisolism due to 21-hydroxylase deficiency, which appeared to be related to an overall suppression of carbohydrate metabolism, including gluconeogenesis, glycolysis and glycogenesis. This is in keeping with previously published studies on murine models having undergone adrenalectomy (Kawai, 1977; Blair, et al., 1995; Bishayi and Ghosh, 2003), and may explain why no abnormalities were found when measuring blood glucose in mutant fish compared to wild type. The sense of the reaction flow between gluconeogenesis and glycolysis is driven by the level of glycemia and by substrate availability, such as pyruvate and lactate (Hers, 1990; Petersen, et al., 2017). Arguably, in the context of reduced mitochondrial organisation and function, there is reduced consumption of such substrates, and this may be an explanation for the downregulation of these processes. There was downregulation of the insulin signalling pathway, which has been shown to play an important role in glucose metabolism, lipid catabolism and protein synthesis in zebrafish (Yang, et al., 2018). Additionally, the response to both nutrients and starvation were blunted in the zebrafish mutants, in agreement to previous studies on GC deficient rats (Müller, et al., 2000; Makimura, et al., 2003; Uchoa, et al., 2012). On the other hand, there was extensive upregulation of protein glycosylation. In human pathology, disorders of glycosylation

represent a large group of conditions which are heterogeneous in presentation, frequently involving neurological, musculoskeletal and cardiovascular impairment (Ondruskova, et al., 2021). In zebrafish, glycosylation has been shown to play an important part in embryogenesis with disorders in growth, development and organogenesis in mutants with impaired protein glycosylation pathways (Flanagan-Steet and Steet, 2013).

Overall, the GSEA results on *cyp21a2-/-* larvae would suggest a tendency in the mutant cortisol deficient larvae towards uncontrolled cell proliferation to the detriment of cellular health. Thus, it is not surprising that when exploring the association between the larval DEGs and human disease, cancer was found to be the most frequent pathology.

6.2.3.2 Transcriptomic analysis of *cyp21a2* male adult livers

The study aimed to observe if any parallels could be drawn between the transcriptomic dysregulations identified in the adult *cyp21a2-/-* livers and those found in the 5 dpf larvae, in order to understand the dynamics of the impact of cortisol deficiency in the zebrafish development. It is important to note that this analysis was rendered much richer by using the GO analysis and looking at whole biological processes that were over-expressed, rather than just comparing the DGE of separate genes, based on an arbitrary set threshold of statistical significance. This was evidenced from the relatively small number of differentially expressed genes that were common between mutant larvae and livers. By contrast, the GO over-expression analysis and the GSEA provided much more complex insights, by taking into account whole groups of genes and their interactions within molecular pathways. This analytic approach showed clear similarities between larvae and male adult livers in the dysregulation of biological processes caused by 21-hydroxylase deficiency. In livers, there was again a very marked tendency towards

reduced expression of genes involved in mitochondrial organisation and provision of energy through ATP metabolism. These processes appeared to be even more markedly downregulated in livers compared to larvae, with extensive suppression of several complexes of the respiratory chain. Moreover, there was a clear interlink between altered expression of the mitochondrial function and production of energy and that of ribosome biosynthesis and protein synthesis, with expression of genes encoding mitochondrial ribosome proteins also being widely suppressed.

Similar to larvae, genes encoding proteins involved in **ribosome biogenesis** were downregulated in male adult livers. The GSEA results also showed a significant and important downregulation of the ribosomal function, in particular translation, which included mitochondrial ribosomes. It is important to highlight that several steps of protein synthesis were downregulated, including the initiation of translation by several eukaryotic translation initiation factors which are components of the mTOR pathway (Schakman, et al., 2013; Xu, et al., 2022). There was also downstream downregulation of the tRNA-synthetase group, which catalyse the tRNA aminoacylation. Importantly, this step is ATP-dependent and ATP metabolism was another group of severely down-regulated biological processes in the liver. It was found that **mitochondrial oxidative phosphorylation** was suppressed through the downregulation of several components including several subunits of the mitochondrial respiratory complex I (NADH:ubiquinone oxidoreductase subunit A4, NADH dehydrogenase subunit 6) and the cytochrome-c oxidoreductase complex (cytochrome c oxidase subunit 7A2a, cytochrome c oxidase subunit 5Aa). Overall, these findings would support the interrelation between the suppression of the respiratory chain in the mitochondria and that of ribosomal protein synthesis. Arguably, this may primarily occur through abnormal mitochondrial organisation and reduced ATP

synthesis in the context of GC deficiency, since it was found that GC impact on the mitochondrial function both indirectly by controlling the expression of mitochondrial components in the nucleus and directly by regulating mitochondrial gene expression (Kokkinopoulou and Moutsatsou, 2021; Psarra and Sekeris., 2011)). However, another potential mechanism may be via the mTOR signalling pathway, which is influenced by GC and stimulates the oxidative phosphorylation in the mitochondria (Cunningham, et al., 2007).

Glucocorticoids are known to have a stimulating effect on **proteolysis** in the muscle (Schakman, et al., 2013), activating lysosomal proteolysis and the ubiquitin-proteasome system. The findings in adult livers are in keeping with this evidence, as cortisol deficiency in mutants was associated with deep downregulation of these processes, the ubiquitin-proteasome system being by far the most extensively affected (**Figure 6.29**). Proteolysis via the ubiquitin-proteasome system is a crucial process needed for amino acid homeostasis and for the regulation of multiple cell functions such as apoptosis, regulation of oncogenesis, metabolism and survival (Rousseau and Bertolotti, 2018). Proteasome assembly, protein ubiquitylation and deubiquitylation are essential sub-processes of proteolysis and they were all significantly downregulated in the mutant livers. Of note, ATP is required in order to initiate protein ubiquitylation, thus, the proteolysis dysregulation may well be linked with the suppressed oxidative phosphorylation within the mitochondria. Other potential mechanisms would be the direct effect of reduced cortisol on the mTORC and MAPK pathways, both playing a role in the regulation of proteasomal proteolysis (Rousseau and Bertolotti, 2018) and both proven to be influenced by GC (Shah, et al., 2002; Jellyman, et al., 2012; Bläuer, et al., 2021).

The **proinflammatory effects** of cortisol deficiency in *cyp21a2-/-* mutant adult livers were supported by the marked upregulation of several systems involved in the inflammatory response, including the complement cascade and cytokine signalling pathways. As in mammals, the roles played by cytokines and the complement system in humoral immunity have previously been described in zebrafish (Lieschke, 2001; Zhang and Cui, 2014). Of note, the JAK-STAT signalling pathway was also upregulated in the *cyp21a2-/-* mutant livers, while its suppressor, *socs1a,* was downregulated. While well-known to play a crucial role in the pathogenesis of cancer and inflammatory disorders (Yu, et al., 2014), there is increasing evidence (mainly from murine models) that JAK-STAT signalling is an important regulator of metabolic function in several organs, including the liver (Dodington, et al., 2018). Moreover, previous evidence from zebrafish has shown that suppression of the *socs1a* gene, which is the main negative regulator of the JAK-STAT pathway, led to steatohepatitis and insulin resistance (Dai, et al., 2015). The MAPK pathway was also upregulated in the *cyp21a2-/-* mutant livers. This is in keeping with previous evidence showing that stress-induced cortisol inhibits MAPK in cows (Dong, et al., 2018). The importance of the MAPK pathway in the pathology of liver metabolism has been repeatedly proven in murine models and human research (Lawan and Bennett, 2017). By contrast, there is limited evidence related to the effects of MAPK within the zebrafish liver, although it was shown that treatment with triclosan cause weight gain, fat accumulation and liver injury through the activation of the MAPK pathway (Liu, et al., 2019).

There was clear downregulation in mutant livers of the **beta-oxidation of fatty acids** (FA), which is likely to have its explanation in the overall suppressed mitochondrial function. Moreover, genes involved in the lipid transport and localisation, including fatty acids (*cd36*, *fabp2, fabp1b.1*) and cholesterol (*osbpl3a, anxa2b*) were also severely

downregulated, indicating an overall imbalance in the lipid management within the liver. Based on published evidence, a parallel can be drawn between these findings and the mechanisms described in the pathogenesis of human non-alcoholic fatty liver disease (NAFLD) (Friedman, et al., 2018; Barbier-Torres, et al., 2020). Of note, this was coupled with the upregulation of the inflammatory response, including cytokine and MAPK signalling, which in human liver disease contribute to the development of hepatic steatosis in the context of lipid accumulation in the liver (Friedman, et al., 2018; Barbier-Torres, et al., 2020). Importantly, NAFLD, steatohepatitis and even hepatic cirrhosis were among the human conditions associated with the liver DGE gene list from *cyp21a2-/-* adult fish livers. The liver microscopy with HE staining could not demonstrate lipid accumulation in the hepatocytes in the *cyp21a2* mutants and specific lipid staining on cryosections was not undertaken due to logistic impediments, thus, further work is required to explore his concept in the zebrafish model.

The analysis of the association between the list of DGE's in mutant adult livers and human pathology showed a marked predominance of cardio-vascular disease in the pathogenesis of which arteriosclerosis plays a central role. Given that inflammation and abnormal lipid homeostasis are major determinants in the development of arteriosclerosis (Golia, et al., 2014), these findings are not surprising.

6.2.4 Differential gene expression in the *cyp11a2* adult livers

The transcriptomic analysis of adult livers from the *cyp11a2* mutant line was used as a control, to explore if there is consistency in the dysregulations caused by GC deficiency due to a different enzymatic defect. It is important to acknowledge that the *cyp11a2*-/- fish

constitute a model for human side chain cleavage enzyme deficiency, one of the most severe forms of CAH, where the first step of steroid biosynthesis is arrested. Thus, in *cyp11a2-/-* fish the synthesis of sex hormones is impaired alongside the gluco- and mineralocorticoid deficiency. Consequently, we must consider that androgen deficiency is likely to contribute to the transcriptomic changes observed in this mutant line. The findings were, however, very consistent with those from the *cyp21a2-/-* adult fish livers, especially regarding the downregulation of the ribosome biogenesis and function, mitochondrial organisation and provision of energy derivatives, as well as the suppression of several metabolic processes, including lipid catabolism. Examining more closely the differential expression of the genes involved in oxidative phosphorylation and ATP metabolism, there was a wider range of genes that were more significantly downregulated in the *cyp11a2* mutants, where the steroid deficiency is more severe. As previously discussed, the effects of GC deficiency on the mitochondrial function is a likely result of a decreased transactivation of the mitochondrial GR, thus it is not surprising that the dysregulation of related processes should be more prominent in the context of more severe steroid deficiency. Interestingly, the biological processes that were more significantly upregulated in this line were those involved in cell cycle and cell proliferation, bearing more similarities to the GSEA results from *cyp21a2-/-* larvae. Overall, in *cyp11a2-/-* fish, the effects of severe steroid deficiency on the adult liver transcriptome combined the suppression of energy provision/homeostasis and several metabolic processes with the upregulation of cell proliferation. These results are similar to those obtained in the *cyp21a2-/-* larvae and adult livers, as are the associations found between the transcriptomic analysis of *cyp11a2* livers and human condition, where a tendency was found towards metabolic liver disease.

6.2.5 Overall conclusions from the transcriptomic analysis

The results of the transcriptomic analysis offer a holistic view on the effects of GC deficiency on zebrafish metabolism and development, which is in keeping with existing evidence from previous studies that explored more focused aspects of the roles of GC.

In mutant larvae the findings indicate that GC deficiency leads to an imbalance between the basic cell housekeeping, which is downregulated, and cell proliferation, which is activated. Disorders of development and neoplastic conditions are the most logical consequences of such dysregulations. However, *cyp21a2-/-* larvae were found to be morphologically similar to controls (Eachus, et al., 2017), as are adult individuals. Thus, it is possible that compensatory mechanisms develop, which may also explain the viability of this zebrafish model in the context of severe cortisol deficiency. Importantly, small molecule metabolism was among the most dysregulated GO terms indicating impairment of several basic metabolic processes.

The transcriptomic profile of adult *cyp21a2-/-* livers presented some similarities but was not identical to that of larvae, indicating some developmental and/or tissue-specific effects of cortisol deficiency. The most prominent feature was the liver-specific suppression of energy provision through suppression of mitochondrial function. Since ATP is crucial for the normal progress of multiple biological processes, it is likely that suppression of the other downregulated functions, such as protein synthesis, proteolysis, fatty acid beta-oxidation, is secondary to that of the oxidative phosphorylation. This theory is further supported by the presence of the GR in the mitochondria, suggesting a direct regulation of its functions by GC. Lipid metabolism and lipid transport were also impaired significantly in adult livers, which is the likely reason why adult mutant fish had more fat

mass compared to wild type siblings. By contrast, there was marked upregulation of the inflammatory response, including major pathways such as the cytokine cascade, JAK-STAT and MAPK. Importantly, a combined effect of deranged lipid metabolism and accelerated inflammatory response constitutes a “steppingstone” in the pathophysiology of metabolic and cardiovascular disease, including arteriosclerosis and steatohepatitis (Golia, et al., 2014; Friedman, et al., 2018; Barbier-Torres, et al., 2020). Thus, these findings support the contribution of cortisol deficiency to the development of such comorbidities in patients with CAH independent of the use of synthetic GC. The transcriptomic results from the *cyp11a2-/-* adult livers confirm those described here for the *cyp21a2-/-* mutant adult livers, indicating that the dysregulation of energy homeostasis and multiple metabolic processes is consistent among different steroid deficient zebrafish lines.

7. Summary and final conclusions

Congenital adrenal hyperplasia is associated with metabolic and cardiovascular comorbidities in adult life, however, there is limited knowledge about the onset and presentation of metabolic problems during childhood. Moreover, the molecular mechanisms through which metabolic disease develops in CAH are poorly understood. It had been assumed that the metabolic and cardiovascular comorbidities are side effects of chronic use of synthetic GC. However, the individual daily cortisol profile of patients with CAH would suggest a much more complex pathophysiology. Over the day, patients alternate between high peaks and long periods of cortisol deficiency, and the hypothesis was that glucocorticoid deficiency and toxicity each play a part in the development of metabolic disease in CAH. This study aimed to improve the understanding of the metabolic complications associated with congenital adrenal hyperplasia by exploring their onset during childhood in relation to GC therapy and hormonal profiles and by using a zebrafish model of 21OHD which allowed us to observe the effects of cortisol deficiency, in the absence of replacement treatment and androgen excess, on the metabolic phenotype and transcriptome.

7.1 Conclusions from the clinical research on patients with 21-hydroxylase deficiency

The children and young persons with CAH in the study had increased prevalence of growth and weight gain problems compared to healthy controls. There was a clear trend towards an early growth spurt and reduced final height in relation to the genetic potential. This resonates with existing evidence from adults with CAH and it is likely to be

the effect of two factors: hyperandrogenism (demonstrated by higher Tanner scores in patients compared to controls) and the growth-suppressive effect of synthetic GC (since a third of the patients exceeded the recommended dose range of 10-15mg/m^2/day HC). This would overall indicate suboptimal treatment which failed to achieve androgen suppression despite using high doses that increase the risk of steroid toxicity, which is a known challenge for hormone replacement in CAH. To some extent, this is also due to a lack of reliable biochemical markers of control, as the most common hormones used for monitoring treatment, 17OHP and androstenedione, did not correlate to the treatment doses or to the anthropometric measurements. The current development of biochemical assays and growing evidence of the use of more disease-specific androgens, the 11-oxygenated C19 steroids, provide a prospect of developing better monitoring strategies.

Patients with CAH had raised prevalence of overweight and obesity, however, in the absence of cushingoid features (found in only one patient), this cannot be with certainty attributed to GC excess. The results could not confirm the raised risk of insulin resistance based on the HOMA-IR in patients, as the prevalence was similar to that of controls. **Several patients had altered lipid profiles**, which would indicate that at least for some CAH patients, metabolic problems have their onset during childhood. Although the number of patients with abnormal lipid profiles was too small to allow further analysis of variance in relation to treatment, hormonal profile and clinical outcomes, the adipokine and metabolomic analysis provided further insights into these aspects.

Leptin, a known regulator of energy homeostasis, neuroendocrine and metabolic functions, increased with the BMI and the waist circumference in CAH patients. Furthermore, leptin was only significantly raised in male patients compared to controls, supporting the theory that the sex-specific differences in the leptin axis are at least partially explained by the distribution of body fat. Leptin correlated positively with insulin and HOMA-IR in CAH patients and decreased as the relative GC dose increased, which is in keeping with previous evidence on the regulation of the leptin axis. Adiponectin, another fat derived hormone, decreased with the increase in BMI-SDS and with the waist to hip ratio, which indicates its potential use as a marker of metabolic risk. It also correlated negatively with the plasma androgens, including the 11-oxygenated C19 steroids, in patients with CAH but not in controls. **Thus, the findings confirmed in children with CAH the known relationship between leptin and adiponectin, GC replacement treatment, hyperandrogenism, weight gain and fat distribution.** This suggests that adipokines may be used as markers of metabolic risk when monitoring treatment.

Several classes of compounds pertaining to lipid metabolism (glycerophospholipids, lysophospholipids, sphingolipids, fatty acids and triglycerides) fluctuated significantly with the daily GC dose and timing of the administration in children with CAH. These metabolites are known to have wide implications in essential biological processes including lipid synthesis and clearance, inflammatory response, and regulation of cell proliferation. Many of them act as bioactive metabolites, being involved in the regulation of signal transduction and dysregulation of

their synthesis are linked to several metabolic and cardiovascular conditions. Importantly, for many of these species, the relationship with the GC dose was not linear but rather followed a “U shape” or “inverted U shape” distribution, suggesting that over- and undertreatment may lead to similar dysregulations in lipid metabolism. **Lipid metabolites, and in particular lysophospholipids, correlated significantly with the patients’ BMI, distinguishing obese individuals from the rest, further confirming that the onset of metabolic and cardiovascular problems in CAH is related to weight gain and takes place in childhood.** As it was the case for adiponectin, lipid metabolites also correlated to plasma androgens (17OHP, androstenedione, testosterone, 11-hydroxyandrostenedione and 11-ketotestosterone). Given the existing evidence on the roles played by androgens in lipid metabolism and insulin sensitivity, this would further suggest that hyperandrogenism contributes directly to development of metabolic disease in CAH.

7.2 Conclusions from research on the cyp21a2-/- zebrafish model

The project also included translational research using a zebrafish model of 21OHD, *cyp21a2-/-*. Mutant fish were found to be severely cortisol deficient in both larval and adult stage, however, they survived and were able to reproduce in the absence of replacement treatment. This provided an in vivo model of cortisol deficiency in which the metabolic consequences of hypercortisolism were explored without the need to consider the added effects of synthetic steroids or androgen excess.

Adult ***cyp21a2-/-*** **fish were heavier and longer compared to WT siblings, having a tendency for increased visceral and subcutaneous fat**. These characteristics were shared by other cortisol deficient zebrafish lines that were established in our laboratory, including *cyp11a2-/-, cyp11c1-/-, fdx1b-/-.* The potential causes for this phenotype could be the direct effect of cortisol deficiency on metabolism, however, other factors should also be considered, such as the interplay between growth hormone (GH) and the stress axis, and the impact of cortisol deficiency on the regulation of food intake and satiety. The steroid profiles of mutant adult fish confirmed profound cortisol deficiency and increased 17OHP, in keeping with human 21OHD. However, ***cyp21a2-/-*** **adult fish do not present hyperandrogenism**, which distinguishes them from humans with 21OHD. It is important to take this difference into account when interpreting these findings on the *cyp21a2-/-* zebrafish model, especially since hyperandrogenism may contribute directly to the metabolic problems observed in CAH patients.

The most consistent finding when exploring differential gene expression by qPCR in zebrafish was the **downregulation of gluconeogenesis in *cyp21a2-/-* larvae and adult livers,** in the context of normal glycolysis**.** This supports the previously advanced hypothesis that cortisol deficiency increases insulin sensitivity through reduced glucose production. The **upregulation of genes involved in adipocyte differentiation and cholesterol synthesis, and downregulation of lipolysis, found in adult livers** are in keeping with the increased fat mass found around the viscera and in the subcutaneous tissue in the mutant adults. However, overall using qPCR to study differential expression of a small number of targeted genes involved in metabolic processes yielded limited

insight, the variable results obtained when replicating experiments, and frequently lacking statistical significance, making it difficult to draw robust conclusions.

The transcriptomic analysis of the *cyp21a2-/-* larvae and adult livers brought significant depth and complexity to the insights regarding the metabolic dysregulations and their dynamic progression in the development of the zebrafish model. Different analytical methods of gene ontology analysis yielded similar results, placing metabolic processes among the most significantly overrepresented GO terms (**Table 7.1**). **Small molecule metabolism, involving basic precursors and key products of metabolic processes, was among the most importantly dysregulated metabolic processes in both larvae and adult livers and genes involved in lipid, carbohydrate and amino acid metabolism were evenly affected**.

In larvae, GSEA showed the loss of *cyp21a2* function caused a clear trend towards the upregulation of cell division, in particular mitosis, and the suppression of processes involved in basic cell housekeeping functions such as energy homeostasis. **One of the most prominent effects of cortisol deficiency in larvae is the suppression of the mitochondrial functions that included apoptosis, respiratory chain assembly and oxidation-reduction processes**. These findings resonate with previous evidence of the fine interplay between the stress axis and the function of the mitochondria, where an important part is played by the translocation of the GR to the mitochondria.

Table 7.1 Key dysregulated biological processes in *cyp21a2* larvae and livers

Larvae		
GOrilla GO analysis	**Gene set enrichment analysis (GSEA)**	
	Key downregulated processes	**Key upregulated processes**
Small molecule metabolic process Organic acid metabolic process Mitotic cell cycle process Carboxylic acid metabolic process Cell cycle process Oxoacid metabolic process Drug metabolic process Regulation of cell cycle Catabolic process Alpha - amino acid metabolic process Carboxylic acid biosynthetic process Metabolic process	Ribosome biogenesis rRNA processing Response to peptide rRNA metabolic process Response to starvation Response to temperature stimulus Response to reactive oxygen species Organelle disassembly Mitochondrion organisation	Chromosome segregation DNA replication Chromosome separation Regulation of chromosome organisation Detection of light stimulus Mitotic cell cycle
Adult livers		
GOrilla GO analysis	**Gene set enrichment analysis (GSEA)**	
	Key downregulated processes	**Key upregulated processes**
Organic anion transport Steroid metabolic process Carboxylic acid transport Organic acid transport Lipid metabolic process Multicellular organism response to stress Small molecule metabolic process Ion transport Organic hydroxy compound metabolic process	Oxidative phosphorylation Aerobic respiration Ribonucleoprotein complex biogenesis ncRNA metabolic process Ribosome biogenesis rRNA metabolic process Energy derivation by oxidation of organic compounds tRNA metabolic process ATP metabolic process Translation	Sterol biosynthetic process Response to cytokine Inflammatory response Cellular response to cytokine stimulus Immune effector process Cytokine-mediated signalling pathway

Summary of the key biological processes that were identified to be over-expressed in larvae and livers by the gene ontology (GO) overexpression analysis (using GOrilla) and GSEA). For the Gorilla GO analysis all processes identified to be significantly over-expressed ($p < 10^{-5}$) were included. For the GSEA, the table provides a selection from the top 20 down-regulated and up-regulated processes, based on the statistical significance (q value), log2FoldChange and size of the gene set.

The transcriptomic analysis in larvae helped extend the limited findings of the qPCR experiments in relation to glucose metabolism, demonstrating overall **suppression of carbohydrate metabolism, including gluconeogenesis, glycolysis and glycogenesis**. Additionally, there was downregulation of the insulin signalling pathway, and of the response to nutrients and starvation.

The transcriptomic changes found in the livers of *cyp21a2-/-* adult fish were similar to those found in larvae, with some distinctions that would suggest a dynamic character in the dysregulation of biological processes during development and growth. **The inhibition of the mitochondrial organisation and provision of energy through ATP metabolism appeared to be even more markedly downregulated in livers with extensive suppression of several complexes of the respiratory chain**. Other processes significantly downregulated in adult livers were protein synthesis, proteolysis and beta-oxidation of fatty acids. As all these pathways contain essential steps that are ATP-dependent, it is possible that the suppression of oxidative phosphorylation in the mitochondria through reduced transactivation of the mitochondrial GR may represent a crucial link in the metabolic dysregulations induced by cortisol deficiency. **The downregulation of lipid metabolic processes in adult livers was associated with the upregulation of the inflammatory response**, including the complement cascade and cytokine signalling pathways, both important contributing factors to the development of fatty liver disease and hepatic steatosis in humans. Additionally, the JAK-STAT and

MAPK pathways, known to hold essential roles in the pathogenesis of inflammatory and neoplastic conditions were also significantly upregulated in mutant livers.

7.3 Combining basic science with clinical research

Overall, the findings from zebrafish complement those from patients with CAH in the understanding of metabolic problems associated with 21OHD. Zebrafish is a convenient and reliable animal model, known to have many similarities to humans in the functionality of the stress axis, factors involved in steroid synthesis and various metabolic functions. The *cyp21a2-/-* mutant line allowed us to study the effects of cortisol deficiency at different stages of development. The increased body size and body fat compared to wild type controls resonate with the finding that patients with CAH have an increased prevalence of overweight and obesity. Arguably, in patients we have to consider the contribution of chronic synthetic steroid use to weight gain, however, the fact that weight did not correlate the dose and patients did not have cushingoid features indicates a more complex mechanism. In the absence of any cortisol replacement, the *cyp21a2-/-* zebrafish showed extensive dysregulations of many metabolic processes in the transcriptome of larvae and adult livers. The marked downregulation of genes involved in mitochondrial organisation, synthesis of ATP and provision of energy derivatives was the most significant finding at both stages of development. This finding is in keeping with the interplay between the mitochondria, on which cortisol acts via GR signalling, and the stress axis which is functional in zebrafish at 5 dpf. An emerging hypothesis from the transcriptomic analysis of *cyp21a2-/-* zebrafish is that cortisol deficiency impairs ATP synthesis in the mitochondria with a downstream effect on several metabolic processes that have ATP-

dependent steps, such as beta-oxidation of fatty acids, lipid transport, protein synthesis and proteolysis. The metabolic dysregulations observed in the

zebrafish model resonate with the results of the metabolomic analysis from the patient CAH cohort and from previous studies, showing abnormalities of several classes of metabolites, including fatty acids and amino acids. Many groups of compounds produced by fat metabolism fluctuated with the dose and timing of GC replacement. Among them, glycerophospholipids, lysophospholipids and sphingolipids, which are known to have wide implications in essential biological functions and to be involved in the pathophysiology of metabolic and cardiovascular disease. Many of these metabolites fluctuated with the GC dose in a non-linear fashion, but rather following a “U shape” or “inverted U shape trend, indicating that low and higher doses may impact on metabolic processes in a similar way. A small number of patients had deranged lipid profiles, however, the plasma metabolomes that fluctuated with the GC dose also correlated with the body weight, leptin and plasma androgens.

Taken together, these findings indicate that in patients with CAH, both 21OHD cortisol deficiency and GC replacement contribute to the development of metabolic disease in these patients. Although in children and young persons with CAH, frank marks of metabolic disease are unlikely, subtle changes are already present in the blood metabolites, which correlate with body weight and may serve to inform early of increased metabolic risks. At a molecular level, the effects of cortisol deficiency are wide and complex, and likely centred around the depression of ATP synthesis and energy provision. Further complexity is rendered by the physiological circadian rhythm of cortisol

secretion, which makes it difficult to salvage these dysregulations by administration of synthetic steroids. However, it is evident that the GC dose and administration regime influence the regulation of metabolic processes. Consequently, there is room for improvement in clinical practice, where we should aim to achieve the optimal replacement regime, complemented by improved monitoring strategies that will target early signs of increased metabolic risk, in order to help improve long term health outcomes in patients.

7.4 Limitations and other considerations

One limitation of the CAH-UK cohort study related to the characteristics of the control group which was not matched to patients ethnically and socio-economically. Moreover, controls had reduced prevalence of overweight and obesity in comparison to the rates reported by WHO for the wider UK paediatric population. However, the finding of increased weight gain in CAH is consistent with previous evidence, including data from adult patients. Plasma metabolomic measurements were not conducted in the control group, thus we could not compare the metabolomes of patients with those of healthy children. Previous evidence on that topic exists, showing an abnormal metabolomic profile in CAH (Nguyen, et al., 2012). Moreover, the limited availability of resources had to be taken into consideration and balanced against the main focus of the project, which was to explore variations of plasma metabolites in patients with impaired cortisol synthesis treated with synthetic GC.

Another study caveat was the use of one-time measurements of plasma androgens and especially the timing of sample collection at variable intervals after the first daily dose of

hydrocortisone. This fact was due to logistical reasons since many patients receive their first dose at very early hours of the morning before 06:00 and this reduced the reliability of androgens concentrations as markers of disease control. Moreover, the conventional markers, 17OHP, androstenedione and testosterone, are known to fluctuate widely with the dose in a manner that is variable among different individuals. This highlights the way forward and the need for future research that will explore the application of novel biomarkers, such as the 11-oxygenated androgens, and of hormonal profiles in monitoring variations in plasma metabolomes, adipokines/cytokines, as early markers of metabolic risk.

A significant distinction between the *cyp21a2-/-* zebrafish model and human 21OHD is that hyperandrogenism is not present in mutant fish as it is in CAH patients. In the paediatric CAH cohort plasma androgens correlated with fat metabolites and adiponectin and it was demonstrated that androgens are directly involved in the regulation of several metabolic processes. Thus, when translating the zebrafish findings into insight on the metabolic effects of GC deficiency in CAH, we need to take into consideration the fact that the pathophysiology of metabolic problems in patients is made more complex by hyperandrogenism.

The results raise several questions that have remained unanswered due to time constraints, and which will require further research work. The HE liver sections did not show any significant differences between *cyp21a2-/-* adult fish and wild type siblings, while the livers of mutants from the *cyp11a2-/-* line presented clear inclusions, not found

in controls. This aspect needs to be further investigated by cryosections of livers followed by fat-specific staining. Moreover, the finding of increased body fat on dissections, while marked and consistent was not confirmed by any quantification techniques. To achieve this, several fat quantification techniques are available, including microCT (Babaei, et al., 2016) and adipocyte labelling with fluorescent dies *in vivo* (Minchin and Rawls, 2017). The differences in the size of the livers and gene expression results found between adult males and females was an unexpected, interesting finding, further supporting the involvement of sex hormones in regulating metabolic processes, a topic that can be further explored using zebrafish as *in vivo* model.

While the results indicate a marked direct impact of GC deficiency and of synthetic steroids on metabolic processes, we need to consider the interplay of the stress axis and other systems and pathways. Thus, the role of the growth hormone (GH) as a regulator of somatic growth in the GC deficient *cyp21a2* zebrafish model will have to be defined, in particular the differential expression of essential genes such as *gh* (growth hormone)*, pomc* (proopiomelanocortin)*, ghrh* (growth hormone releasing hormone)*, sst1* (somatostatin 1)*, igf-1* (insulin-like growth factor 1)*, igfbp-1* (igf-1 binding protein) - will be undertaken through qPCR and transcriptomic analysis.

Additionally, the effect of GC the regulation of food intake and satiety was demonstrated in murine models and the present transcriptomic results from *cyp21a2-/-* showed dysregulation of the response to nutrients and starvation. Consequently, another step further will involve feeding experiments as well as exploring the differential expression of

orexigenic (neuropeptide y, agouti related protein, ghrelin) and anorexigenic (pomc) peptides in adult brain for *cyp21a2*-/- zebrafish. This approach will test the hypothesis that GC deficiency impacts on metabolism by changing the zebrafish feeding behaviours.

Finally, the next step in using the zebrafish model of 21OHD would involve treatment experiments, exploring the impact of the exposure to synthetic GC on the metabolism of the *cyp21a2-/-* mutants at different stages of development. This will help to answer questions regarding the possibility to salvage the expression of metabolic genes and processes by different types of steroids, doses and times of administration. The insight derived from such a project will help to direct future clinical research aiming to optimise treatment and monitoring strategies in CAH.

Appendix

1.Blood sample for metabolomic analysis

Blood collection general points

Blood needs to be collected in the appropriate collection tube for serum and plasma. Blood must be packed well in wet ice after collection until it arrives at the laboratory for processing and must be stored on ice during processing in the laboratory. Ideally, sample transport and processing should be performed within 30 minutes. The involvement of research nurses or clinical fellows is advisable if possible. Avoid any exposure of samples to ambient temperatures after collection. The required serum collection tubes and cryovials will be send to the study sites at the start of the study.

Serum collection (if collecting for metabolomics analysis only)

1. Use high quality plastics so that there is no leakage of plasticisers and contaminants into blood samples. We recommend the use of the consumables described below as these have been applied for a number of years and we trust them regarding the absence of contamination. DO NOT USE GEL-BASED CLOTTING TUBES.
2. Draw blood (typically 10ml) into suitable serum collection tubes (Greiner Vacuettes, cat. no. 455092) and allow to clot for a minimum of 1 hour at 4°C on ice. The clotting time (time from sample collection to centrifugation) must be recorded.
3. Prepare the serum fraction by centrifugation of the blood collection tube at 2500 x *g* for 15 minutes at 4°C.
4. Divide samples immediately into aliquots (0.5ml) in cryovials (Greiner, cat. no. 122261/122263) and freeze by placing in a -80°C freezer.
5. Collect enough blood (1-2 ml) so that there is sufficient serum for at least 2 aliquots. Store samples at -80°C and transport on dry ice.

Plasma collection (if collecting for proteomics and metabolomics studies)

1. Use high quality plastics so that there is no leakage of plasticisers and contaminants into blood samples. We recommend the use of the consumables described below as these have been applied for a number of years and we trust them regarding the absence of contamination.
2. Draw blood into Lithium heparin plasma collection tubes (Greiner, cat. no. 455084) and mix it with the anticoagulant by inverting the tube three times. Prepare the plasma fraction immediately by centrifugation at 3,000 x *g* for 20 min at 4°C.
3. Divide samples immediately into aliquots (0.5ml) in cryovials (e.g. Greiner, cat. no. 122261/122263) and freeze by placing in a -80°C freezer.

The samples will be stores at -80°C and transported on dry ice to the <Address study coordinator>,(Study coordinator: <name and email.) where samples will be stored at -80 Celsius until send in batches for further analysis of metabolomics analysis by LC-MS/MS.

2. Code for RNA sequencing analysis

1. Pre-processing of raw data

#Connection to server and choosing directory

```
ssh -X md1iab@sharc.shef.ac.uk
cd /mnt/fastdata/md1iab
```

#Make directory (eg “larvae”)

```
mkdir larvae
cd /mnt/fastdata/md1iab/larvae
```

#Upload raw data fastq.gz files to the appropriate directory (eg for larvae “larvae”: /mnt/fastdata/md1iab/larvae) - Unzip fastq.gz files

```
gunzip *fastqc.gz
```

#If applicable merge the files from the different runs, keeping the R1 & R2 separate (eg below for WT-1 sample)

```
cat WT-1_S1_L001_R1_001.fastq WT-1_S1_L002_R1_001.fastq WT-1_S1_L003_R1_001.fastq WT-1_S1_L004_R1_001.fastq > merged_WT-1_S1_R1_001.fastq

cat WT-1_S1_L001_R2_001.fastq WT-1_S1_L002_R2_001.fastq WT-1_S1_L003_R2_001.fastq WT-1_S1_L004_R2_001.fastq > merged_WT-1_S1_R2_001.fastq
```

#Check quality of raw data by FASTQC (this will be repeated at each step – after SortmeRNA and Trimmomatic)

```
wget http://www.bioinformatics.babraham.ac.uk/projects/fastqc/fastqc_v0.11.9.zip

unzip fastqc_v0.11.9.zip

qrsh -l rmem=10G

module load apps/java
```

```
java -version
cd /mnt/fastdata/md1iab/larvae/FastQC
chmod 755 fastqc
export PATH=$PATH:/fastdata/md1iab/larvae/FastQC
fastqc /fastdata/md1iab/larvae/merged*
mkdir QC_Reports
mv *.html QC_Reports
mv *.zip QC_Reports
cd QC_reports
mkdir RawQC
mv *.html RawQC
mv *.zip RawQC
# MultiQC
cd /mnt/fastdata/md1iab/larvae
wget https://repo.anaconda.com/miniconda/Miniconda2-latest-Linux-x86_64.sh
bash Miniconda2-latest-Linux-x86_64.sh
conda create --name py3.7 python=3.7
conda activate py3.7
conda install -c bioconda -c conda-forge multiqc
cd /mnt/fastdata/md1iab/larvae/QC_Reports/RawQC/
multiqc .

# Applying SortmeRNA
cd /mnt/fastdata/md1iab/larvae
wget https://github.com/biocore/sortmerna/archive/2.1.zip
unzip 2.1.zip
```

```
cd sortmerna-2.1/

qrsh -l rmem=10G

module load apps/java

bash ./build.sh

make install

export PATH=$PATH:/fastdata/md1iab/larvae/sortmerna-2.1

#indexing part
./indexdb_rna --ref /fastdata/md1iab/larvae/sortmerna-2.1/rRNA_databases/rfam-5.8s-database-id98.fasta,/fastdata/md1iab/larvae/sortmerna-2.1/index/rfam-5.8s-db:/fastdata/md1iab/larvae/sortmerna-2.1/rRNA_databases/rfam-5s-database-id98.fasta,/fastdata/md1iab/larvae/sortmerna-2.1/index/rfam-5s-db:/fastdata/md1iab/larvae/sortmerna-2.1/rRNA_databases/silva-arc-16s-id95.fasta,/fastdata/md1iab/larvae/sortmerna-2.1/index/silva-arc-16s-db:/fastdata/md1iab/larvae/sortmerna-2.1/rRNA_databases/silva-arc-23s-id98.fasta,/fastdata/md1iab/larvae/sortmerna-2.1/index/silva-arc-23s-db:/fastdata/md1iab/larvae/sortmerna-2.1/rRNA_databases/silva-bac-16s-id90.fasta,/fastdata/md1iab/larvae/sortmerna-2.1/index/silva-bac-16s-db:/fastdata/md1iab/larvae/sortmerna-2.1/rRNA_databases/silva-bac-23s-id98.fasta,/fastdata/md1iab/larvae/sortmerna-2.1/index/silva-bac-23s-db:/fastdata/md1iab/larvae/sortmerna-2.1/rRNA_databases/silva-euk-18s-id95.fasta,/fastdata/md1iab/larvae/sortmerna-2.1/index/silva-euk-18s-db:/fastdata/md1iab/larvae/sortmerna-2.1/rRNA_databases/silva-euk-28s-id98.fasta,/fastdata/md1iab/larvae/sortmerna-2.1/index/silva-euk-28s-db


#Merge paired reads (inside the /sortmerna-2.1/scripts)
bash ./merge-paired-reads.sh R1.fq readR2.fq read-interleaved.fq

cd scripts/

#example for WT-1
bash ./merge-paired-reads.sh /fastdata/md1iab/larvae/merged_WT-1_S1_R1_001.fastq/fastdata/md1iab/larvae/merged_WT-1_S1_R2_001.fastq merged_WT-1_S1_001.interleaved.fastq

mv *interleaved.fastq /mnt/fastdata/md1iab/larvae/sortmerna-2.1

#SortmeRNA for Paired-end data example for WT-1:
./sortmerna --ref /fastdata/md1iab/larvae/sortmerna-2.1/rRNA_databases/rfam-5.8s-database-id98.fasta,/fastdata/md1iab/larvae/sortmerna-2.1/index/rfam-5.8s-db:/fastdata/md1iab/larvae/sortmerna-2.1/rRNA_databases/rfam-5s-database-id98.fasta,/fastdata/md1iab/larvae/sortmerna-2.1/index/rfam-5s-db:/fastdata/md1iab/larvae/sortmerna-2.1/rRNA_databases/silva-arc-16s-
```

id95.fasta,/fastdata/md1iab/larvae/sortmerna-2.1/index/silva-arc-16s-db:/fastdata/md1iab/larvae/sortmerna-2.1/rRNA_databases/silva-arc-23s-id98.fasta,/fastdata/md1iab/larvae/sortmerna-2.1/index/silva-arc-23s-db:/fastdata/md1iab/larvae/sortmerna-2.1/rRNA_databases/silva-bac-16s-id90.fasta,/fastdata/md1iab/larvae/sortmerna-2.1/index/silva-bac-16s-db:/fastdata/md1iab/larvae/sortmerna-2.1/rRNA_databases/silva-bac-23s-id98.fasta,/fastdata/md1iab/larvae/sortmerna-2.1/index/silva-bac-23s-db:/fastdata/md1iab/larvae/sortmerna-2.1/rRNA_databases/silva-euk-18s-id95.fasta,/fastdata/md1iab/larvae/sortmerna-2.1/index/silva-euk-18s-db:/fastdata/md1iab/larvae/sortmerna-2.1/rRNA_databases/silva-euk-28s-id98.fasta,/fastdata/md1iab/larvae/sortmerna-2.1/index/silva-euk-28s-db --reads /fastdata/mdp19lo/Paired-end/GR_data/merged_WT-1_S1_001.interleaved.fq --aligned merged_WT-1_S1_001.interleaved.fq _rRNA --other merged_WT-1_S1_001.interleaved.fq -sortmeRNA --fastx --num_alignments 1 --log --sam -v --paired-in

cd sortmerna-2.1

#Run all samples as batch jobs using nano
nano BatchJob1.sh

#When the batch job opens as separate window add
#$ #WT sortmerna script
#$ #!/bin/bash
#$ -l h_rt=90:00:00
#$ -l rmem=32G
#$ -m bea
#$ -M md1iab@sheffield.ac.uk
#$ -j y

Then enter the sortmerna command for all the intended files – aim for no more than 3-4 files per batch job otherwise it may not complete. Example for WT-1 below
./sortmerna --ref /fastdata/md1iab/larvae/sortmerna-2.1/rRNA_databases/rfam-5.8s-database-id98.fasta,/fastdata/md1iab/larvae/sortmerna-2.1/index/rfam-5.8s-db:/fastdata/md1iab/larvae/sortmerna-2.1/rRNA_databases/rfam-5s-database-id98.fasta,/fastdata/md1iab/larvae/sortmerna-2.1/index/rfam-5s-db:/fastdata/md1iab/larvae/sortmerna-2.1/rRNA_databases/silva-arc-16s-id95.fasta,/fastdata/md1iab/larvae/sortmerna-2.1/index/silva-arc-16s-db:/fastdata/md1iab/larvae/sortmerna-2.1/rRNA_databases/silva-arc-23s-id98.fasta,/fastdata/md1iab/larvae/sortmerna-2.1/index/silva-arc-23s-db:/fastdata/md1iab/larvae/sortmerna-2.1/rRNA_databases/silva-bac-16s-id90.fasta,/fastdata/md1iab/larvae/sortmerna-2.1/index/silva-bac-16s-db:/fastdata/md1iab/larvae/sortmerna-2.1/rRNA_databases/silva-bac-23s-id98.fasta,/fastdata/md1iab/larvae/sortmerna-2.1/index/silva-bac-23s-db:/fastdata/md1iab/larvae/sortmerna-2.1/rRNA_databases/silva-euk-18s-id95.fasta,/fastdata/md1iab/larvae/sortmerna-2.1/index/silva-euk-18s-db:/fastdata/md1iab/larvae/sortmerna-2.1/rRNA_databases/silva-euk-28s-id98.fasta,/fastdata/md1iab/larvae/sortmerna-2.1/index/silva-euk-28s-db --reads /fastdata/md1iab/larvae/sortmerna-2.1/merged_WT-1_S1_001.interleaved.fastq --

aligned merged_WT-1_S1_001.interleaved_rRNA --other merged_WT-1_S1_001.interleaved_sortmeRNA --fastx --num_alignments 1 --log --sam -v --paired_in

Submit batch job
qsub BatchJob1.sh

Applying Trimmomatic

cd /mnt/fastdata/md1iab/larvae
mkdir Trimmomatic
mv sortmerna* /mnt/fastdata/md1iab/larvae/Trimmomatic
cd /mnt/fastdata/md1iab/larvae/Trimmomatic

unmerge paried reads (from scripts)
cd /mnt/fastdata/md1iab/larvae/sortmerna-2.1/scripts

bash ./unmerge-paired-reads.sh /mnt/fastdata/md1iab/larvae/Trimmomatic/merged_WT-1_S1_001.interleaved_sortmeRNA.fastq sortmerna_merged_WT-1_S1_R1_other.fastq sortmerna_merged_WT-1_S1_R2_other.fastq

mv sortmerna* /mnt/fastdata/md1iab/larvae/Trimmomatic

wget http://www.usadellab.org/cms/uploads/supplementary/Trimmomatic/Trimmomatic-0.39.zip

unzip Trimmomatic-0.39.zip

qrsh -l rmem=10G

module load apps/java

Code for trimmomatic – template
java -jar Trimmomatic-0.39/trimmomatic-0.39.jar PE -phread33 read_other_R1.fq.gz read_other_R2.fq.gz read_sortmerna_trimmomaticR1.fq.gz read_sortmerna_unpaired_R1.fq.gzread_sortmerna_trimmomaticR2.fq.gz read_sortmerna_unpaired_R2.fq.gz ILLUMINACLIP:"TruSeq3-PE-2.fa":2:30:10

Example for WT-1
java -jar Trimmomatic-0.39/trimmomatic-0.39.jar PE -phred33 sortmerna_merged_WT-1_S1_R1_other.fastq.gz sortmerna_merged_WT-1_S1_R2_other.fastq.gz sortmerna_trimmomatic_merged_WT-1_S1_R1_other.fastq.gz unpaired_trimmomatic_merged_WT-1_S1_R1_other.fastq.gz sortmerna_trimmomatic_merged_WT-1_S1_R2_other.fastq.gz unpaired_trimmomatic_merged_WT-1_S1_R2_other.fastq.gz ILLUMINACLIP:Trimmomatic-0.39/adapters/TruSeq3-PE-2.fa:2:30:10

#Run all samples as batch jobs using nano

Applying STAR

git clone https://github.com/alexdobin/STAR.git

cd STAR/source

make STAR

export
PATH=$PATH:/fastdata/md1iab/larvae/Trimmomatic/STAR/bin/Linux_x86_64_static/

cd /mnt/fastdata/md1iab/larvae/Trimmomatic/STAR

mkdir genome

wget ftp://ftp.ensembl.org/pub/release-
100/fasta/danio_rerio/dna/Danio_rerio.GRCz11.dna.primary_assembly.fa.gz

wget ftp://ftp.ensembl.org/pub/release-
103/gtf/danio_rerio/Danio_rerio.GRCz11.103.gtf.gz

gunzip *.gz

qrsh -l rmem=32G

#Run STAR in genome folder
cd /mnt/fastdata/md1iab/larvae/Trimmomatic/STAR/genome

export
PATH=$PATH:/mnt/fastdata/md1iab/larvae/Trimmomatic/STAR/bin/Linux_x86_64_stati
c/

export
PATH=$PATH:/fastdata/md1iab/larvae/Trimmomatic/STAR/bin/MacOSX_x86_64/

#Still in genome folder
STAR --runThreadN 8 --runMode genomeGenerate --genomeDir
/fastdata/md1iab/larvae/Trimmomatic STAR/genome --genomeFastaFiles
/fastdata/md1iab/larvae/Trimmomatic/STAR/Danio_rerio.GRCz11.dna.primary_assembl
y.fa --sjdbGTFfile
/fastdata/md1iab/larvae/Trimmomatic/STAR/Danio_rerio.GRCz11.103.gtf --
sjdbOverhang 100

#Running STAR on samples as batch jobs using nano.
nano BatchJob.sh
qsub BatchJob.sh

#$ #!/bin/bash
#$ -l h_rt=94:00:00

```
#$ -l rmem=32G
#$ -m bea
#$ -M i.bacila@sheffield.ac.uk
#$ -j y

export
PATH=$PATH:/fastdata/md1iab/larvae/Trimmomatic/STAR/bin/Linux_x86_64_static/
export PATH=$PATH:/fastdata/md1iab/Trimmomatic/STAR/genome

#Template for running STAR on samples (example WT-1 below)
STAR --runThreadN 1 --genomeDir /fastdata/md1iab/Trimmomatic/STAR/genome/ --
sjdbGTFfile /fastdata/md1iab/Trimmomatic/STAR/Danio_rerio.GRCz11.104.gtf --
readFilesIn /fastdata/md1iab/Trimmomatic/STAR
/sortmerna_trimmomatic_merged_WT-1_S1_R1_other.fastq
/fastdata/md1iab/Trimmomatic/STAR /sortmerna_trimmomatic_merged_WT-
1_S1_R2_other..fastq --outSAMtype BAM SortedByCoordinate Unsorted --
outFileNamePrefix /fastdata/md1iab/Trimmomatic/STAR
/sortmerna_trimmomatic_merged_WT-1_S1 --outSAMunmapped Within --quantMode
TranscriptomeSAM GeneCounts --twopassMode Basic
```

3.R code for differential gene expression

```
setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)


install.packages("BiocManager")
BiocManager::install("DESeq2")
install.packages("GenomicRanges")
library(GenomicRanges)
library(DESeq2)


# Read in the tables - counts, metadata
# Counts
larvaecyp21vswt <- read.table("/Users/irina/Desktop/larvaecyp21vswt.txt")
metadatacyp21larvae <-
read.table("/Users/irina/Desktop/metadatalarvaecyp21vswt.txt")


# Get the summary of the number of reads in each column
colSums(larvaecyp21vswt)


# Sort it and divide by a million
sort(colSums(larvaecyp21vswt))/1e6
```

```
# Construct a DESeq data set from a counts matrix
ddslarv21 <- DESeqDataSetFromMatrix(countData = larvaecyp21vswt,colData =
metadatacyp21larvae,design = ~Genotype)

levels(ddslarv21$Genotype)


ddslarv21 <- DESeq(ddslarv21)


# Run DESeq2 on the object dds
larvresults21 <- results(ddslarv21, alpha = 0.05)
larvresults21sig <- subset(larvresults21, padj<0.05)

head(larvresults21)
head(larvresults21sig)


#Summary of results
summary(larvresults21)


#Plotting size factors
library(rafalib)
mypar()
ddslarv21 <- estimateSizeFactors(ddslarv21)
sizeFactors(ddslarv21)
plot(sizeFactors(ddslarv21), colSums(counts(ddslarv21)))
abline(lm(colSums(counts(ddslarv21)) ~ sizeFactors(ddslarv21) + 0))


log.norm.counts <- log2(counts(ddslarv21, normalized=TRUE)+ 1)
rs <- rowSums(counts(ddslarv21))
mypar(1,2)
boxplot(log2(counts(ddslarv21)[rs > 0,]+1)) #not normalised
boxplot(log.norm.counts[rs > 0,]) #normalised


#PCA plot
library(ggplot2)
rldlarv21 <- rlog(ddslarv21, blind = TRUE)
plotPCA(rldlarv21, intgroup="Genotype") + theme_bw()


pca_data <- plotPCA(rldlarv21, intgroup="Genotype", returnData = TRUE)
library(ggrepel)
ggplot(pca_data, aes(x = PC1, y = PC2, col = group, label = name)) + geom_point() +
geom_text_repel() + theme_bw() + theme(text = element_text(size=15)) + labs(y=
"PCA2: 8% variance", x = "PCA1: 80% variance")
```

```
MyLivDataPCA <- plotPCA(rldlarv21, intgroup=("Genotype"), returnData=TRUE)
MyLivDataPCA


#Making a cluster dendrogram
library(rafalib)
plot(hclust(dist(t(assay(rldlarv21)))), labels = colData(rldlarv21)$name)


#Making a heatmap
larv_logmatrix <- assay(rldlarv21)
library(pheatmap)
larv_cor <- cor(larv_logmatrix)
pheatmap(larv_cor)


# Another box plot
logu.norm.counts <- log2(counts(ddslarv21, normalized=TRUE)+1)
rs <- rowSums(counts(ddslarv21))
boxplot(logu.norm.counts)


#Making a MA plot
plotMA (larvresults21, ylim = c(-10, 10))


#Only significant ones
plotMA (larvresults21sig, ylim = c(-13,10))


library(vsn)
mypar(1,2)
meanSdPlot(logu.norm.counts, ranks=FALSE, ylim=c(0,3))
meanSdPlot(assay(rldlarv21), ranks=FALSE, ylim=c(0,3), main="rlog")


#Volcano plot without annotations
library(dplyr)
library(tibble)
#Save the shrunk significant results (lfc) as data frame
larvresShrink <- lfcShrink(ddslarv21, coef="Genotype_mutant_vs_control",
type="apeglm")

larvresShrink.df <-as.data.frame(larvresShrink)

# Create a volcano plot
larvresShrink.df %>%
  ggplot(aes(x = log2FoldChange, y = -log10(padj))) + geom_point()
```

```
#Making a volcano plot and colouring sig genes
larvresShrink.df %>%
    ggplot(aes(x = log2FoldChange, y = -log10(padj), col = padj < 0.05)) + geom_point()
+ theme_bw() +  theme(axis.line = element_line(colour = "black"),
    panel.border = element_blank(),
    panel.background = element_blank())


#Sorting results
# (i) Sorting all genes

larvcyp21results_sorted <- larvresults21[order(larvresults21$padj),]

# (ii) Filtering out those that are significantly expressed
larvcyp21results_sorted_Sig <- subset(larvcyp21results_sorted, padj<0.05)


#Adding annotation
library("zebrafish.db")

# (i) Write ordered genes as a data frame
larvcyp21results_sorted_df <- as.data.frame(larvcyp21results_sorted) %>%
  rownames_to_column("ENSEMBL")

# (i) Build the names for all the genes
larvcyp21results_sorted_AnnotatedEnsembl <-
AnnotationDbi::select(org.Dr.eg.db,keys=larvcyp21results_sorted_df$"ENSEMBL",
keytype="ENSEMBL", columns=c("SYMBOL","GENENAME"))

# (ii) Write significant *and* ordered genes as a data frame
larvcyp21results_sorted_Sig_df <- as.data.frame(larvcyp21results_sorted_Sig) %>%
  rownames_to_column("ENSEMBL")

# (ii) Build the names for significant *and* ordered genes
larvcyp21results_sorted_Sig_Annotated_Ensembl <-
AnnotationDbi::select(org.Dr.eg.db,keys=larvcyp21results_sorted_Sig_df $"ENSEMBL",
keytype="ENSEMBL", columns=c("SYMBOL","GENENAME"))


# (ii) Identify the duplicated entries using the duplicated function
dup_ids_larvcyp21results_sorted_Sig_Annotated_Ensembl <-
larvcyp21results_sorted_Sig_Annotated_Ensembl$ENSEMBL[duplicated(larvcyp21resu
lts_sorted_Sig_Annotated_Ensembl$ENSEMBL)]
filter(larvcyp21results_sorted_Sig_Annotated_Ensembl, ENSEMBL %in%
dup_ids_larvcyp21results_sorted_Sig_Annotated_Ensembl) %>%
  arrange(ENSEMBL) %>% head
```

```
# (ii) Discard the duplicated entries - the first occurence of the duplicated ID remains included in the table
larvcyp21results_sorted_Sig_Annotated_Ensembl_without_dupl <-
AnnotationDbi::select(org.Dr.eg.db,keys=larvcyp21results_sorted_Sig_df$ENSEMBL,
            columns=c("ENSEMBL","SYMBOL","GENENAME","ENTREZID"),
            keytype="ENSEMBL") %>%
  filter(!duplicated(ENSEMBL))


# (ii) Attach the annotation info
larvcyp21results_sorted_Sig_results.annotated_joined <-
left_join(larvcyp21results_sorted_Sig_df,
larvcyp21results_sorted_Sig_Annotated_Ensembl_without_dupl,by="ENSEMBL")

#Write it as CSV
write.csv(as.data.frame(larvcyp21results_sorted_Sig_results.annotated_joined), file =
"/Users/irina/Desktop/larvcyp21results_sorted_Sig_results.annotated_joined.csv")


#Adding volcano plot - showing those that are SIGNIFICANT only
library(ggrepel)
top_fifteen_genes_larvcyp21results_sorted_Sig_results.annotated_joined <-
larvcyp21results_sorted_Sig_results.annotated_joined$ENSEMBL[1:15]


larvcyp21results_sorted_Sig_results.annotated_joined %>%
  mutate(Label = ifelse(ENSEMBL %in%
top_fifteen_genes_larvcyp21results_sorted_Sig_results.annotated_joined, SYMBOL,
"")) %>%
  ggplot(aes(x = log2FoldChange, y = -log10(padj), label=Label)) +
geom_point(alpha=0.4) + geom_text_repel(col="blue") + theme_bw() +  theme(axis.line
= element_line(colour = "black"),
    panel.border = element_blank(),
    panel.background = element_blank())


#Plots for individual genes of interest

# pck1 (example)
SingleGenePlot <- plotCounts(ddslarv21, gene="ENSDARG00000013522",
intgroup="Genotype", returnData=TRUE)

ggplot(SingleGenePlot, aes(x=Genotype, y=count, col=Genotype)) +
  geom_point(position = position_jitter(width=.1,height=0)) +
  scale_y_log10()+theme_bw() +
    theme(axis.line = element_line(colour = "black"),
    panel.grid.major = element_blank(),
    panel.border = element_blank(),
    panel.background = element_blank())
```

SingleGenePlot

```
#Making a heatmap
library(pheatmap)
larvHMtopgenes <- head(rownames(larvcyp21results_sorted_Sig), 25)
larvHMtopgenes

matlarv21<- assay(rldlarv21)[larvHMtopgenes,]
matlarv <- matlarv21-rowMeans(matlarv21)

df_larv_heatmap <- as.data.frame(colData(rldlarv21)[,c("Genotype")])
colnames(df_larv_heatmap) = "Genotype"
rownames(df_larv_heatmap) = colnames(matlarv)
pheatmap(matlarv, annotation_col = df_larv_heatmap)


#Using GOrilla to produce a list of GO terms in order to input it in REVIGO - running as 'single ranked list of genes'.

# (i) Identify the duplicated entries using the duplicated function
dup_ids_larvcyp21results_sorted_AnnotatedEnsembl <-
larvcyp21results_sorted_AnnotatedEnsembl$ENSEMBL[duplicated(larvcyp21results_so
rted_AnnotatedEnsembl$ENSEMBL)]
filter(larvcyp21results_sorted_AnnotatedEnsembl, ENSEMBL %in%
dup_ids_larvcyp21results_sorted_AnnotatedEnsembl) %>%
  arrange(ENSEMBL) %>% head

# (i) Discard the duplicated entries - the first occurence of the duplicated ID remains included in the table
larvcyp21results_sorted_AnnotatedEnsembl_without_dupl <-
AnnotationDbi::select(org.Dr.eg.db,keys=larvcyp21results_sorted_df$ENSEMBL,
            columns=c("ENSEMBL","SYMBOL","GENENAME","ENTREZID"),
            keytype="ENSEMBL") %>%
  filter(!duplicated(ENSEMBL))

# (i) Attach the annotation info
larvcyp21results_sorted_AnnotatedEnsembl.annotated_joined <-
left_join(larvcyp21results_sorted_df,
larvcyp21results_sorted_AnnotatedEnsembl_without_dupl,by="ENSEMBL")

# (i) Write it as CSV
write.csv(as.data.frame(larvcyp21results_sorted_AnnotatedEnsembl.annotated_joined),
file =
"/Users/irina/Desktop/larvcyp21results_sorted_AnnotatedEnsembl.annotated_joined.csv
")

# GO overexpression analysis using Cluster profiler
BiocManager::install("clusterProfiler")
```

```
library(clusterProfiler)

install.packages("ggupset")
library(ggupset)

# Create the universe
larvcyp21universe <- larvcyp21results_sorted_AnnotatedEnsembl.annotated_joined
%>%  pull("ENSEMBL")

# Create the significant DEGs variable
larvcyp21sigGenes <- larvcyp21results_sorted_Sig_results.annotated_joined %>%
  filter(!is.na("ENSEMBL")) %>%
  pull("ENSEMBL")

enrich_gocyp21 <- enrichGO(
  gene = larvcyp21sigGenes,
  OrgDb = "org.Dr.eg.db",
  keyType = "ENSEMBL",
  ont = "BP",
  universe = fcyp11universe,
  qvalueCutoff = 0.05,
  readable = TRUE
)

# Turn it into data frame
data_for_enrich_gocyp21_plot <- enrich_ gocyp21 %>% data.frame
write_csv(as.data.frame(data_for_enrich_ gocyp21_plot), file = "data_for_enrich_
gocyp21_plot.csv")

# Make the dotplot
dotplot(enrich_ gocyp21, showCategory = 15, font.size = 10)

# Make the upsetplot
enrichplot::upsetplot(enrich_ gocyp21, font.size = 18)

#GSEA analysis - does not impose arbitrary cut-offs on the dataset
install.packages("ggridges")
library("ggridges")

#Setting up GSEA
larvcyp21ranked_genes <-
larvcyp21results_sorted_AnnotatedEnsembl.annotated_joined %>%
  arrange(desc(stat)) %>%
  filter(!is.na(stat))

larvcyp21gene_list <- pull(larvcyp21ranked_genes, stat)

names(larvcyp21gene_list) <- pull(larvcyp21ranked_genes, "ENSEMBL")

larvcyp21gse_GO <- gseGO(geneList = larvcyp21gene_list,
```

```
            OrgDb = org.Dr.eg.db,
            ont = "BP", keyType = "ENSEMBL")


# Produce a data frame
larvcyp21GSE_GO <- larvcyp21gse_GO %>% as.data.frame
write_csv(as.data.frame(larvcyp21GSE_GO), file = " larvcyp21GSE_GO.csv")

# Make a ridge plot
library(ggridges)
ridgeplot(larvcyp21gse_GO, showCategory = 22) + theme(axis.text.x =
element_text(size = 8), axis.text.y = element_text(size = 10)) +  xlim(-5,7.5)

# Explore correlations with human disease

# Desgenet2r installation
# https://www.disgenet.org/static/disgenet2r/disgenet2r.html

library(devtools)
library(SPARQL)
install_bitbucket("ibi_group/disgenet2r")
library(disgenet2r)

disgenet_api_key <- get_disgenet_api_key(
              email = "i.bacila@sheffield.ac.uk",
              password = "*************" )
disgenet_api_key
Sys.setenv(DISGENET_API_KEY= "*******************************")

#Read in the table with human orthologues, downloaded as TSV - specify the separator
OrthologuesLarv <- read.table("Human_Orthologue_conversionLarv.txt",
header=TRUE, fill = TRUE, sep = "\t")
head(OrthologuesLarv)
head(myGenelistLarv)

res_enrichlarv <- disease_enrichment(entities = myGenelistLarv, universe=
"DISGENET", vocabulary = "HGNC", verbose = "TRUE", database = "ALL")


plot(res_enrichlarv, class = "Enrichment", count =2,  cutoff= 0.05) + theme(axis.text.x =
element_text(size = 8), axis.text.y = element_text(size = 10))
```

References

"WHO Child Growth Standards based on length/height, weight and age." (2006). *Acta Paediatr Suppl* **450**: 76-85.

Alsop, D. and M. M. Vijayan (2008). "Development of the corticosteroid stress axis and receptor expression in zebrafish." *Am J Physiol Regul Integr Comp Physiol* **294**(3): R711-719.

Alsop, D. and M. M. Vijayan (2009). "Molecular programming of the corticosteroid stress axis during zebrafish development." *Comp Biochem Physiol A Mol Integr Physiol* **153**(1): 49-54.

Alwashih, M. A., D. G. Watson, R. Andrew, R. H. Stimson, M. Alossaimi, G. Blackburn and B. R. Walker (2017). "Plasma metabolomic profile varies with glucocorticoid dose in patients with congenital adrenal hyperplasia." *Scientific Reports* **7**(1):17092.

Anderson, J. L., J. D. Carten and S. A. Farber (2011). "Zebrafish lipid metabolism: from mediating early patterning to the metabolism of dietary fat and cholesterol." *Methods Cell Biol* **101**: 111-141.

Argenton, F., E. Zecchin and M. Bortolussi (1999). "Early appearance of pancreatic hormone-expressing cells in the zebrafish embryo." *Mech Dev* **87**(1-2): 217-221.

Ariyawatkul, K., S. Tepmongkol, S. Aroonparkmongkol and T. Sahakitrungruang (2017). "Cardio-metabolic risk factors in youth with classical 21-hydroxylase deficiency." *Eur J Pediatr* **176**(4): 537-545.

Arlt, W., D. S. Willis, S. H. Wild, N. Krone, E. J. Doherty, S. Hahner, T. S. Han, P. V. Carroll, G. S. Conway, D. A. Rees, R. H. Stimson, B. R. Walker, J. M. C. Connell, R. J. Ross and T. U. K. C. A. H. A. S. Executive (2010). "Health status of adults with congenital adrenal hyperplasia: A cohort study of 203 patients." *J Clin Endocrinol Metab* **95**(11): 5110-5121.

Babaei, F., T. L. Hong, K. Yeung, S. H. Cheng and Y. W. Lam (2016). "Contrast-enhanced x-ray micro-computed tomography as a versatile method for anatomical studies of adult zebrafish." *Zebrafish* **13**(4): 310-316.

Bacila, I. A., C. Elder and N. Krone (2019). "Update on adrenal steroid hormone biosynthesis and clinical implications." *Arch Dis Child* **104**(12): 1223-1228.

Bacila, I., J. Adaway, J. Hawley, S. Mahdi, R. Krone, L. Patel, S. Alvi, T. Randell, E. Gevers, M. Dattani, T. Cheetham, A. Kyriakou, L. Schiffer, F. Ryan, E. Crowne, J. H. Davies, S. F. Ahmed, B. Keevil and N. Krone (2019). "Measurement of salivary adrenal-specific androgens as biomarkers of therapy control in 21-hydroxylase deficiency." *J Clin Endocrinol Metab* **104**(12): 6417-6429.

Bacila, I., V. T. Cunliffe and N. P. Krone (2021). "Interrenal development and function in zebrafish." *Mol Cell Endocrinol* **535**: 111372.

Bacila, I., N. R. Lawrence, S. Mahdi, S. Alvi, T. D. Cheetham, E. Crowne, U. Das, M. T. Dattani, J. H. Davies, E. Gevers, R. E. Krone, A. Kyriakou, L. Patel, T. Randell, F. Ryan, B. Keevil, S. F. Ahmed and N. P. Krone (2022). "Health status of Children and young persons with congenital adrenal hyperplasia in the UK (CAH-UK): a cross-sectional multi-centre study." *Eur J Endocrinol* **187**(4):543-553.

Bancos, I., S. Hahner, J. Tomlinson and W. Arlt (2015). "Diagnosis and management of adrenal insufficiency." *Lancet Diabetes & Endocrinology* **3**(3): 216-226.

Barbier-Torres, L., K. A. Fortner, P. Iruzubieta, T. C. Delgado, E. Giddings, Y. Chen, D. Champagne, D. Fernández-Ramos, D. Mestre, B. Gomez-Santos, M. Varela-Rey, V. G. de Juan, P. Fernández-Tussy, I. Zubiete-Franco, C. García-Monzón, Á. González-Rodríguez, D. Oza, F. Valença-Pereira, Q. Fang, J. Crespo, P. Aspichueta, F. Tremblay, B. C. Christensen, J. Anguita, M. L. Martínez-Chantar and M. Rincón (2020). "Silencing hepatic MCJ attenuates non-alcoholic fatty liver disease (NAFLD) by increasing mitochondrial fatty acid oxidation." *Nat Commun* **11**(1): 3360.

Bauer, M. P., J. T. Bridgham, D. M. Langenau, A. L. Johnson and F. W. Goetz (2000). "Conservation of steroidogenic acute regulatory protein structure and expression in vertebrates." *Mol Cell Endocrinol* **168**(1-2): 119-125.

Berruien, N.N.A., C.L. Smith (2020). “Emerging roles of melanocortin receptor accessory proteins (MRAP and MRAP2) in physiology and pathophysiology. *Gene* **5**;757:144949.

Biemar, F., F. Argenton, R. Schmidtke, S. Epperlein, B. Peers and W. Driever (2001). "Pancreas development in zebrafish: early dispersed appearance of endocrine hormone expressing cells and their convergence to form the definitive islet." *Dev Biol* **230**(2): 189-203.

Bishayi, B. and S. Ghosh (2003). "Metabolic and immunological responses associated with in vivo glucocorticoid depletion by adrenalectomy in mature Swiss albino rats." *Life Sci* **73**(24): 3159-3174.

Blair, S. C., I. D. Caterson and G. J. Cooney (1995). "Glucocorticoid deprivation alters in vivo glucose uptake by muscle and adipose tissues of GTG-obese mice." *Am J Physiol Endocrinol* **269**(5): E927-E933.

Bläuer, M., J. Sand and J. Laukkarinen (2021). "Regulation of p38 MAPK and glucocorticoid receptor activation by hydrocortisone in mono-and co-cultured pancreatic acinar and stellate cells." *Pancreatology* **21**(2): 384-389.

Bolger, A. M., M. Lohse and B. Usadel (2014). "Trimmomatic: a flexible trimmer for Illumina sequence data." *Bioinformatics* **30**(15): 2114-2120.

Bonfig, W., S. Bechtold, H. Schmidt, D. Knorr and H. P. Schwarz (2007). "Reduced final height outcome in congenital adrenal hyperplasia under prednisone treatment: Deceleration of growth velocity during puberty." *J Clin Endocrinol Metab* **92**(5): 1635-1639.

Borges, J. H., D. M. de Oliveira, S. H. V. de Lemos-Marini, B. Geloneze, E. M. Gonçalves and G. Guerra-Júnior (2021). "Fat distribution and lipid profile of young adults with congenital adrenal hyperplasia due to 21-hydroxylase enzyme deficiency." *Lipids* **56**(1): 101-110.

Browne, P. D., T. K. Nielsen, W. Kot, A. Aggerholm, M. T. P. Gilbert, L. Puetz, M. Rasmussen, A. Zervas and L. H. Hansen (2020). "GC bias affects genomic and metagenomic reconstructions, underrepresenting GC-poor organisms." *GigaScience* **9**(2).

Chai, C., Y. W. Liu and W. K. Chan (2003). "Ff1b is required for the development of steroidogenic component of the zebrafish interrenal organ." *Dev Biol* **260**(1): 226-244.

Charmandari, E., A. Johnston, C. Brook and P. Hindmarsh (2001). "Bioavailability of oral hydrocortisone in patients with congenital adrenal hyperplasia due to 21-hydroxylase deficiency." *J Endocrinol* **169**(1):65-70

Charmandari, E., M. Weise, S. R. Bornstein, G. Eisenhofer, M. F. Keil, G. P. Chrousos and D. P. Merke (2002). "Children with classic congenital adrenal hyperplasia have elevated serum leptin concentrations and insulin resistance: potential clinical implications." *J Clin Endocrinol Metab* **87**(5): 2114-2120.

Charmandari, E., N. C. Nicolaides and G. P. Chrousos (2014). "Adrenal insufficiency." *Lancet* **383**(9935): 2152-2167.

Christiansen, J. J., C. B. Djurhuus, C. H. Gravholt, P. Iversen, J. S. Christiansen, O. Schmitz, J. Weeke, J. O. Jørgensen and N. Møller (2007). "Effects of cortisol on carbohydrate, lipid, and protein metabolism: studies of acute cortisol withdrawal in adrenocortical failure." *J Clin Endocrinol Metab* **92**(9): 3553-3559.

Clark, A. F., G. N. DeMartino and K. Wildenthal (1986). "Effects of glucocorticoid treatment on cardiac protein synthesis and degradation." *Am J Physiol* **250**(6 Pt 1): C821-827.

Cole, L. K., J. E. Vance and D. E. Vance (2012). "Phosphatidylcholine biosynthesis and lipoprotein metabolism." *Biochim Biophys Acta* **1821**(5): 754-761.

Cortés, R., M. Teles, M. Oliveira, C. Fierro-Castro, L. Tort and J. M. Cerdá-Reverter (2018). "Effects of acute handling stress on short-term central expression of orexigenic/anorexigenic genes in zebrafish." *Fish Physiol Biochem* **44**(1): 257-272.

Cunningham, J. T., J. T. Rodgers, D. H. Arlow, F. Vazquez, V. K. Mootha and P. Puigserver (2007). "mTOR controls mitochondrial oxidative function through a YY1-PGC-1alpha transcriptional complex." *Nature* **450**(7170): 736-740.

Dai, Z., H. Wang, X. Jin, H. Wang, J. He, M. Liu, Z. Yin, Y. Sun and Q. Lou (2015). "Depletion of suppressor of cytokine signaling-1a causes hepatic steatosis and insulin resistance in zebrafish." *Am J Physiol Endocrinol Metab* **308**(10): E849-859.

Dallman, M. F., S. E. la Fleur, N. C. Pecoraro, F. Gomez, H. Houshyar and S. F. Akana (2004). "Minireview: Glucocorticoids—food intake, abdominal obesity, and wealthy nations in 2004." *Endocrinology* **145**(6): 2633-2638.

Dauber, A., M. Kellogg and J. A. Majzoub (2010). "Monitoring of therapy in congenital adrenal hyperplasia." *Clin Chem* **56**(8): 1245-1251.

Delai, A., P. M. Gomes, M. C. Foss-Freitas, J. Elias, S. R. Antonini, M. Castro, A. C. Moreira and L. M. Mermejo (2022). "Hyperinsulinemic-euglycemic clamp strengthens the insulin resistance in nonclassical congenital adrenal hyperplasia." *J Clin Endocrinol Metab* **107**(3): e1106-e1116.

Den Broeder, M. J., V. A. Kopylova, L. M. Kamminga and J. Legler (2015). "Zebrafish as a model to study the role of peroxisome proliferating-activated receptors in adipogenesis and obesity." *PPAR Res* **2015**: 358029.

Dinarello A., Licciardello G., Fontana C.M., Tiso N., Argenton F., Dalla Valle L (2020). "Glucocorticoid receptor activities in the zebrafish model: a review". *J Endocrinol* **247**(3):R63-R82.

Dobin, A. and T. R. Gingeras (2015). "Mapping RNA-seq reads with STAR." *Curr Protoc Bioinformatics* **51**: 11.14.11-11.14.19.

Dodington, D. W., H. R. Desai and M. Woo (2018). "JAK/STAT - Emerging players in metabolism." *Trends Endocrinol Metab* **29**(1): 55-65.

Dong, J., J. Li, L. Cui, Y. Wang, J. Lin, Y. Qu and H. Wang (2018). "Cortisol modulates inflammatory responses in LPS-stimulated RAW264.7 cells via the NF-κB and MAPK pathways." *BMC Vet Res* **14**(1): 30.

Du, J., Y. Wang, R. Hunter, Y. Wei, R. Blumenthal, C. Falke, R. Khairova, R. Zhou, P. Yuan, R. Machado-Vieira, B. S. McEwen and H. K. Manji (2009). "Dynamic regulation of mitochondrial function by glucocorticoids." *Proc Natl Acad Sci U S A* **106**(9): 3543-3548.

Du, X. L., W. J. Xu, J. L. Shi, K. Guo, C. T. Guo, R. Zheng, S. W. Jiang and J. Chai (2021). "Glucocorticoid receptor alpha targets SLC2A4 to regulate protein synthesis and breakdown in porcine skeletal muscle cells." *Biomolecules* **11**(5).

Eachus, H., A. Zaucker, J. A. Oakes, A. Griffin, M. Weger, T. Guran, A. Taylor, A. Harris, A. Greenfield, J. L. Quanson, K. H. Storbeck, V. T. Cunliffe, F. Muller and N. Krone (2017). "Genetic disruption of 21-hydroxylase in zebrafish causes interrenal hyperplasia." *Endocrinology* **158**(12): 4165-4173.

Ekman, B., M. Quinkler, B. A. Jones, C. Marelli, R. Murray, P. Zelissen and J. Wahlberg (2015). "Fludrocortisone therapy in patients with primary adrenal insufficiency: relationships with different hydrocortisone doses." *Endocrine Abstracts* (2015) **37** EP71.

El-Maouche, D., W. Arlt and D. P. Merke (2017). "Congenital adrenal hyperplasia." *The Lancet* **390**(10108): 2194-2210.

Elo, B., C. M. Villano, D. Govorko and L. A. White (2007). "Larval zebrafish as a model for glucose metabolism: expression of phosphoenolpyruvate carboxykinase as a marker for exposure to anti-diabetic compounds." *J Mol Endocrinol* **38**(3-4): 433-440.

Engels, M., K. Gehrmann, H. Falhammar, E. A. Webb, A. Nordenstrom, F. C. Sweep, P. N. Span, A. E. van Herwaarden, J. Rohayem, A. Richter-Unruh, C. Bouvattier, B. Kohler, B. B. Kortmann, W. Arlt, N. Roeleveld, N. Reisch, N. Stikkelbroeck, H. L.

Claahsen-van der Grinten and L. G. dsd (2018). "Gonadal function in adult male patients with congenital adrenal hyperplasia." *Eur J Endocrinol* **178**(3): 285-294.

Even, S. E., M. G. Dulak-Lis, R. M. Touyz and A. Nguyen Dinh Cat (2014). "Crosstalk between adipose tissue and blood vessels in cardiometabolic syndrome: implication of steroid hormone receptors (MR/GR)." *Horm Mol Biol Clin Investig* **19**(2): 89-101.

Ewels, P., M. Magnusson, S. Lundin and M. Käller (2016). "MultiQC: summarize analysis results for multiple tools and samples in a single report." *Bioinformatics* **32**(19): 3047-3048.

Feliciano Pereira, P., S. Eloiza Priore and J. Bressan (2014). "Aldosterone: a cardiometabolic risk hormone?" *Nutr Hosp* **30**(6): 1191-1202.

Flanagan-Steet, H. R. and R. Steet (2013). ""Casting" light on the role of glycosylation during embryonic development: insights from zebrafish." *Glycoconj J* **30**(1): 33-40.

Flynn, E. J., 3rd, C. M. Trent and J. F. Rawls (2009). "Ontogeny and nutritional control of adipogenesis in zebrafish (Danio rerio)." *J Lipid Res* **50**(8): 1641-1652.

Friedman, S. L., B. A. Neuschwander-Tetri, M. Rinella and A. J. Sanyal (2018). "Mechanisms of NAFLD development and therapeutic strategies." *Nat Med* **24**(7): 908-922.

Frisardi, V., F. Panza, D. Seripa, T. Farooqui and A. A. Farooqui (2011). "Glycerophospholipids and glycerophospholipid-derived lipid mediators: a complex meshwork in Alzheimer's disease pathology." *Prog Lipid Res* **50**(4): 313-330.

Gans, I. M., J. Grendler, R. Babich, N. Jayasundara and J. A. Coffman (2021). "Glucocorticoid-responsive transcription factor krüppel-like factor 9 regulates fkbp5 and metabolism." *Front Cell Dev Biol* **9**: 727037.

German, A., S. Suraiya, Y. Tenenbaum-Rakover, I. Koren, G. Pillar and Z. Hochberg (2008). "Control of childhood congenital adrenal hyperplasia and sleep activity and quality with morning or evening glucocorticoid therapy." *J Clin Endocrinol Metab* **93**(12): 4707-4710.

Glasauer, S. M. and S. C. Neuhauss (2014). "Whole-genome duplication in teleost fishes and its evolutionary consequences." *Mol Genet Genomics* **289**(6): 1045-1060.

Goldstone, J. V., A. G. McArthur, A. Kubota, J. Zanette, T. Parente, M. E. Jönsson, D. R. Nelson and J. J. Stegeman (2010). "Identification and developmental expression of the full complement of Cytochrome P450 genes in Zebrafish." *BMC Genomics* **11**: 643.

Golia, E., G. Limongelli, F. Natale, F. Fimiani, V. Maddaloni, I. Pariggiano, R. Bianchi, M. Crisci, L. D'Acierno, R. Giordano, G. Di Palma, M. Conte, P. Golino, M. G. Russo, R. Calabrò and P. Calabrò (2014). "Inflammation and cardiovascular disease: from pathogenesis to therapeutic target." *Curr Atheroscler Rep* **16**(9): 435.

Gonzalez, E., K. M. Johnson, P. S. Pallan, T. T. N. Phan, W. Zhang, L. Lei, Z. Wawrzak, F. K. Yoshimoto, M. Egli and F. P. Guengerich (2018). "Inherent steroid 17α,20-lyase activity in defunct cytochrome P450 17A enzymes." *J Biol Chem* **293**(2): 541-556.

Goya, L., A. C. Maiyar, Y. Ge and G. L. Firestone (1993). "Glucocorticoids induce a

G1/G0 cell cycle arrest of Con8 rat mammary tumor cells that is synchronously reversed by steroid withdrawal or addition of transforming growth factor-alpha." *Mol Endocrinol* **7**(9): 1121-1132.

Greenberg, A. K., J. Hu, S. Basu, J. Hay, J. Reibman, T. A. Yie, K. M. Tchou-Wong, W. N. Rom and T. C. Lee (2002). "Glucocorticoids inhibit lung cancer cell growth through both the extracellular signal-related kinase pathway and cell cycle regulators." *Am J Respir Cell Mol Biol* **27**(3): 320-328.

Greytak, S. R., D. Champlin and G. V. Callard (2005). "Isolation and characterization of two cytochrome P450 aromatase forms in killifish (Fundulus heteroclitus): differential expression in fish from polluted and unpolluted environments." *Aquat Toxicol* **71**(4): 371-389.

Griffin, A., S. Parajes, M. Weger, A. Zaucker, A. E. Taylor, D. M. O'Neil, F. Müller and N. Krone (2016). "Ferredoxin 1b (Fdx1b) Is the essential mitochondrial redox partner for cortisol biosynthesis in zebrafish." *Endocrinology* **157**(3): 1122-1134.

Hampl, R., L. Foretová, J. Sulcová and L. Stárka (1990). "Daily profiles of salivary cortisol in hydrocortisone treated children with congenital adrenal hyperplasia." *Eur J Pediatr* **149**(4): 232-234.

Hannun, Y. A. and L. M. Obeid (2018). "Sphingolipids and their metabolism in physiology and disease." *Nat Rev Mol Cell Biol* **19**(3): 175-191.

Harris, A. P., M. C. Holmes, E. R. de Kloet, K. E. Chapman and J. R. Seckl (2013). "Mineralocorticoid and glucocorticoid receptor balance in control of HPA axis and behaviour." *Psychoneuroendocrinology* **38**(5): 648-658.

Hers, H. G. (1990). "Mechanisms of blood glucose homeostasis." *J Inherit Metab Dis* **13**(4): 395-410.

Hindmarsh, P. C. and J. W. Honour (2020). "Would cortisol measurements be a better gauge of hydrocortisone replacement therapy? Congenital Adrenal Hyperplasia as an Exemplar." *Int J Endocrinol* **2020**: 2470956.

Hishikawa, D., T. Hashidate, T. Shimizu and H. Shindou (2014). "Diversity and function of membrane glycerophospholipids generated by the remodeling pathway in mammalian cells." *J Lipid Res* **55**(5): 799-807.

Hoijman, E., L. Rocha-Viegas, S. G. Kalko, N. Rubinstein, M. Morales-Ruiz, E. B. Joffé, E. C. Kordon and A. Pecci (2012). "Glucocorticoid alternative effects on proliferating and differentiated mammary epithelium are associated to opposite regulation of cell-cycle inhibitor expression." *J Cell Physiol* **227**(4): 1721-1730.

Hölttä-Vuori, M., V. T. Salo, L. Nyberg, C. Brackmann, A. Enejder, P. Panula and E. Ikonen (2010). "Zebrafish: gaining popularity in lipid research." *Biochem J* **429**(2): 235-242.

Hsu, H. J., G. Lin and B. C. Chung (2004). "Parallel early development of zebrafish interrenal glands and pronephros: differential control by wt1 and ff1b." *Endocr Res* **30**(4): 803.

Hsu, H. J., P. Hsiao, M. W. Kuo and B. C. Chung (2002). "Expression of zebrafish cyp11a1 as a maternal transcript and in yolk syncytial layer." *Gene Expr Patterns* **2**(3-4): 219-222.

Hu, L., J. Liu, W. Zhang, T. Wang, N. Zhang, Y. H. Lee and H. Lu (2020). "Functional metabolomics decipher biochemical functions and associated mechanisms underline small-molecule metabolism." *Mass Spectrom Rev* **39**(5-6): 417-433.

Hu, M. C., H. J. Hsu, I. C. Guo and B. C. Chung (2004). "Function of Cyp11a1 in animal models." *Mol Cell Endocrinol* **215**(1-2): 95-100.

Hunter, R. G., M. Seligsohn, T. G. Rubin, B. B. Griffiths, Y. Ozdemir, D. W. Pfaff, N. A. Datson and B. S. McEwen (2016). "Stress and corticosteroids regulate rat hippocampal mitochondrial DNA gene expression via the glucocorticoid receptor." *Proc Natl Acad Sci U S A* **113**(32): 9099-9104.

Hwang, J. L. and R. E. Weiss (2014). "Steroid-induced diabetes: a clinical and molecular approach to understanding and treatment." *Diabetes Metab J* **30**(2): 96-102.

Imrie, D. and K. C. Sadler (2010). "White adipose tissue development in zebrafish is regulated by both developmental time and fish size." *Dev Dyn* **239**(11): 3013-3023.

Jellyman, J. K., M. S. Martin-Gronert, R. L. Cripps, D. A. Giussani, S. E. Ozanne, Q. W. Shen, M. Du, A. L. Fowden and A. J. Forhead (2012). "Effects of cortisol and dexamethasone on insulin signalling pathways in skeletal muscle of the ovine fetus during late gestation." *PLoS One* **7**(12): e52363.

Julián, M. T., N. Alonso, I. Ojanguren, E. Pizarro, E. Ballestar and M. Puig-Domingo (2015). "Hepatic glycogenosis: An underdiagnosed complication of diabetes mellitus?" *World J Diabetes* **6**(2): 321-325.

Jurczyk, A., N. Roy, R. Bajwa, P. Gut, K. Lipson, C. X. Yang, L. Covassin, W. J. Racki, A. A. Rossini, N. Phillips, D. Y. R. Stainier, D. L. Greiner, M. A. Brehm, R. Bortell and P. diIorio (2011). "Dynamic glucoregulation and mammalian-like responses to metabolic and developmental disruption in zebrafish." *Gen Comp Endocrinol* **170**(2): 334-345.

Kamrath, C., L. Wettstaedt, C. Boettcher, M. F. Hartmann and S. A. Wudy (2018). "Androgen excess is due to elevated 11-oxygenated androgens in treated children with congenital adrenal hyperplasia." *J Steroid Biochem Mol Biol* **178**: 221-228.

Kamrath, C., M. F. Hartmann, J. Pons-Kühnemann and S. A. Wudy (2020). "Urinary GC-MS steroid metabotyping in treated children with congenital adrenal hyperplasia." *Metabolism* **112**: 154354.

Kawai, A., Kuzuya,N. (1977). "On the role of glucocorticoid in glucose-induced insulin secretion." *Horm Metab Res* **9**(05): 361-365.

Kim, Y. S. and Y. Kim (1975). "Glucocorticoid inhibition of protein synthesis in vivo and in vitro." *J Biol Chem***250**(6): 2293-2298.

Knuplez, E. and G. Marsche (2020). "An updated review of pro- and anti-inflammatory properties of plasma lysophosphatidylcholines in the vascular system." *Int J Mol Sci* **21**(12).

Kokkinopoulou, I. and P. Moutsatsou (2021). "Mitochondrial Glucocorticoid Receptors and Their Actions." *Int J Mol Sci* **22**(11):6054

Kopylova, E., L. Noé and H. Touzet (2012). "SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data." *Bioinformatics* **28**(24): 3211-3217.

Krone, N., N. Reisch, J. Idkowiak, V. Dhir, H. E. Ivison, B. A. Hughes, I. T. Rose, D. M. O'Neil, R. Vijzelaar, M. J. Smith, F. MacDonald, T. R. Cole, N. Adolphs, J. S. Barton, E. M. Blair, S. R. Braddock, F. Collins, D. L. Cragun, M. T. Dattani, R. Day, S. Dougan, M. Feist, M. E. Gottschalk, J. W. Gregory, M. Haim, R. Harrison, A. H. Olney, B. P. Hauffa, P. C. Hindmarsh, R. J. Hopkin, P. E. Jira, M. Kempers, M. N. Kerstens, M. M. Khalifa, B. Köhler, D. Maiter, S. Nielsen, S. M. O'Riordan, C. L. Roth, K. P. Shane, M. Silink, N. M. Stikkelbroeck, E. Sweeney, M. Szarras-Czapnik, J. R. Waterson, L. Williamson, M. F. Hartmann, N. F. Taylor, S. A. Wudy, E. M. Malunowicz, C. H. Shackleton and W. Arlt (2012). "Genotype-phenotype analysis in congenital adrenal hyperplasia due to P450 oxidoreductase deficiency." *J Clin Endocrinol Metab* **97**(2): E257-267.

Kumari, M., T. Chandola, E. Brunner and M. Kivimaki (2010). "A nonlinear relationship of generalized and central obesity with diurnal cortisol secretion in the whitehall II study." *J Clin Endocrinol Metab* **95**(9): 4415-4423.

Kuo, T., A. McQueen, T.-C. Chen and J.-C. Wang (2015). “Regulation of glucose homeostasis by glucocorticoids.” *Adv Exp Med Biol* **872**:99-126

Kushnir, M. M., T. Blamires, A. L. Rockwood, W. L. Roberts, B. Yue, E. Erdogan, A. M. Bunker and A. W. Meikle (2010). "Liquid chromatography-tandem mass spectrometry assay for androstenedione, dehydroepiandrosterone, and testosterone with pediatric and adult reference intervals." *Clin Chem* **56**(7): 1138-1147.

Landgraf, K., S. Schuster, A. Meusel, A. Garten, T. Riemer, D. Schleinitz, W. Kiess and A. Körner (2017). "Short-term overfeeding of zebrafish with normal or high-fat diet as a model for the development of metabolically healthy versus unhealthy obesity." *BMC Physiology* **17**(1): 4.

Lawan, A. and A. M. Bennett (2017). "Mitogen-activated protein kinase regulation in hepatic metabolism." *Trends Endocrinol Metab* **28**(12): 868-878.

Lee, M. J. and S. K. Fried (2014). "The glucocorticoid receptor, not the mineralocorticoid receptor, plays the dominant role in adipogenesis and adipokine production in human adipocytes." *Int J Obes (Lond)* **38**(9): 1228-1233.

Li, N., J. A. Oakes, K. H. Storbeck, V. T. Cunliffe and N. P. Krone (2020). "The P450 side-chain cleavage enzyme Cyp11a2 facilitates steroidogenesis in zebrafish." *J Endocrinol* **244**(2): 309-321.

Li, Y. F., R. S. Li, S. B. Samuel, R. Cueto, X. Y. Li, H. Wang and X. F. Yang (2016). "Lysophospholipids and their G protein-coupled receptors in atherosclerosis." *Front Biosci (Landmark Ed)* **21**(1): 70-88.

Lieschke, G. J. (2001). "Zebrafish--an emerging genetic model for the study of cytokines and hematopoiesis in the era of functional genomics." *Int J Hematol* **73**(1): 23-

31.

Lin, H., Z. Zhou, W. Zhong, P. Huang, N. Ma, Y. Zhang, C. Zhou, Y. Lai, S. Huang, H. An, X. Sun, L. Gao and Z. Lv (2017). "Naringenin inhibits alcoholic injury by improving lipid metabolism and reducing apoptosis in zebrafish larvae." *Oncology Reports* **38**(5): 2877-2884.

Lin, J. C., S. Hu, P. H. Ho, H. J. Hsu, J. H. Postlethwait and B. C. Chung (2015). "Two zebrafish hsd3b genes are distinct in function, expression, and evolution." *Endocrinology* **156**(8): 2854-2862.

Liu, M., W. Ai, L. Sun, F. Fang, X. Wang, S. Chen and H. Wang (2019). "Triclosan-induced liver injury in zebrafish (Danio rerio) via regulating MAPK/p53 signaling pathway." *Comp Biochem Physiol C Toxicol Pharmacol* **222**: 108-117.

Liu, Y. W., W. Gao, H. L. Teh, J. H. Tan and W. K. Chan (2003). "Prox1 is a novel coregulator of Ff1b and is involved in the embryonic development of the zebra fish interrenal primordium." *Mol Cell Biol* **23**(20): 7243-7255.

Liyanarachchi, K., R. Ross and M. Debono (2017). "Human studies on hypothalamo-pituitary-adrenal (HPA) axis." *Best Pract Res Clin Endocrinol Metab* **31**(5): 459-473.

Löhr, H., S. Hess, M. M. A. Pereira, P. Reinoß, S. Leibold, C. Schenkel, C. M. Wunderlich, P. Kloppenburg, J. C. Brüning and M. Hammerschmidt (2018). "Diet-induced growth Is regulated via acquired leptin resistance and engages a pomc-somatostatin-growth hormone circuit." *Cell Rep* **23**(6): 1728-1741.

Lopes, C., E. Rocha, I. L. Pereira and T. V. Madureira (2021). "Deciphering influences of testosterone and dihydrotestosterone on lipid metabolism genes using brown trout primary hepatocytes." *Aquat Toxicol* **235**: 105819.

Macfarlane, D. P., S. Forbes and B. R. Walker (2008). "Glucocorticoids and fatty acid metabolism in humans: fuelling fat redistribution in the metabolic syndrome." *J Endocrinol* **197**(2): 189-204.

Maddison, L. A., K. E. Joest, R. M. Kammeyer and W. Chen (2015). "Skeletal muscle insulin resistance in zebrafish induces alterations in beta-cell number and glucose tolerance in an age- and diet-dependent manner." *Am J Physiol Endocrinol* **308**(8): E662-E669.

Makimura, H., T. M. Mizuno, F. Isoda, J. Beasley, J. H. Silverstein and C. V. Mobbs (2003). "Role of glucocorticoids in mediating effects of fasting and diabetes on hypothalamic gene expression." *BMC Physiology* **3**(1): 5.

Malerbi, D., B. Liberman, A. Giurnofilho, D. Giannellaneto and B. L. Wajchenberg (1988). "Glucocorticoids and glucose-metabolism - hepatic glucose-production in untreated addisonian patients and on 2 different levels of glucocorticoid administration." *Clin Endocrinol* **28**(4): 415-422.

Maripuu, M., M. Wikgren, P. Karling, R. Adolfsson and K. F. Norrback (2016). "Relative hypocortisolism is associated with obesity and the metabolic syndrome in recurrent affective disorders." *J Affect Disord* **204**: 187-196.

Mattern, J., M. W. Büchler and I. Herr (2007). "Cell cycle arrest by glucocorticoids may protect normal tissue and solid tumors from cancer therapy." *Cancer Biol Ther* **6**(9): 1345-1354.

McGonnell, I. M. and R. C. Fowkes (2006). "Fishing for gene function – endocrine modelling in the zebrafish." *Journal of Endocrinology* **189**(3): 425.

Meng, X. H., B. Chen and J. P. Zhang (2017). "Intracellular insulin and impaired autophagy in a zebrafish model and a cell model of type 2 diabetes." *Intl J Biol Sci* **13**(8): 985-995.

Menke, A. L., J. M. Spitsbergen, A. P. Wolterbeek and R. A. Woutersen (2011). "Normal anatomy and histology of the adult zebrafish." *Toxicol Pathol* **39**(5): 759-775.

Merke, D. P. and R. J. Auchus (2020). "Congenital adrenal hyperplasia due to 21-hydroxylase deficiency." *N Engl J Med* **383**(13): 1248-1261.

Merke, D. P. and S. R. Bornstein (2005). "Congenital adrenal hyperplasia." *Lancet* **365**(9477): 2125-2136.

Meyuhas, O., V. Baldin, G. Bouche and F. Amalric (1990). "Glucocorticoids repress ribosome biosynthesis in lymphosarcoma cells by affecting gene expression at the level of transcription, posttranscription and translation." *Biochim Biophys Acta* **1049**(1): 38-44.

Michel, M., P. Page-McCaw, W. Chen and R. Cone (2016). "Leptin signalling regulates glucose homeostasis, but not adipostasis, in the zebrafish." *Proc Natl Acad Sci* **113**: 201513212.

Miller, W. L. (2017). "Disorders in the initial steps of steroid hormone synthesis." *J Steroid Biochem Mol Biol* **165**(Pt A): 18-37.

Miller, W. L. (2017). "Steroidogenesis: unanswered questions." *Trends Endocrinol Metab* **28**(11): 771-793.

Miller, W. L. and R. J. Auchus (2011). "The molecular biology, biochemistry, and physiology of human steroidogenesis and its disorders." *Endocr Rev* **32**(1): 81-151.

Minchin, J. E. and J. F. Rawls (2017). "In vivo imaging and quantification of regional adiposity in zebrafish." *Methods Cell Biol* **138**: 3-27.

Modaressi, S., K. Brechtel, B. Christ and K. Jungermann (1998). "Human mitochondrial phosphoenolpyruvate carboxykinase 2 gene. Structure, chromosomal localization and tissue-specific expression." *Biochem J* **333 (Pt 2)**(Pt 2): 359-366.

Mooij, C. F., E. A. Webb, H. L. Claahsen van der Grinten and N. Krone (2017). "Cardiovascular health, growth and gonadal function in children and adolescents with congenital adrenal hyperplasia." *Arch Dis Child* **102**(6): 578-584.

Mooij, C. F., J. M. Kroese, F. C. Sweep, A. R. Hermus and C. J. Tack (2011). "Adult patients with congenital adrenal hyperplasia have elevated blood pressure but otherwise a normal cardiovascular risk profile." *PLoS One* **6**(9): e24204.

Mooij, C. F., J. M. Kroese, H. L. Claahsen-van der Grinten, C. J. Tack and A. R. M. M. Hermus (2010). "Unfavourable trends in cardiovascular and metabolic risk in paediatric and adult patients with congenital adrenal hyperplasia?" *Clinl Endocrinol* **73**(2): 137-146.

Morton, N. M., J. M. Paterson, H. Masuzaki, M. C. Holmes, B. Staels, C. Fievet, B. R. Walker, J. S. Flier, J. J. Mullins and J. R. Seckl (2004). "Novel adipose tissue–mediated resistance to diet-induced visceral obesity in 11β-hydroxysteroid dehydrogenase type 1–deficient mice." *Diabetes* **53**(4): 931-938.

Müller, M. B., M. E. Keck, S. Zimmermann, F. Holsboer and W. Wurst (2000). "Disruption of feeding behavior in CRH receptor I-deficient mice is dependent on glucocorticoids." *NeuroReport* **11**(9): 1963-1966.

Nakamura Y, H. P., Casson P,et al. (2009). "Type 5 17beta-hydroxysteroid dehydrogenase (AKR1C3) contributes to testosterone production in the adrenal reticularis." *J Clin Endocrinol Metab* **94**: 2192-2198

Nermoen, I., E. S. Husebye, A. G. Myhre and K. Løvås (2017). "Classic congenital adrenal hyperplasia." *Tidsskr Nor Laegeforen* **137**(7): 540-543.

Nguyen, A. T., A. Emelyanov, C. H. Koh, J. M. Spitsbergen, S. Parinov and Z. Gong (2012). "An inducible kras(V12) transgenic zebrafish model for liver tumorigenesis and chemical drug screening." *Dis Model Mech* **5**(1): 63-72.

NHSDigital. National Child Measurement Programme - England, 2016/2017. 2017 [Available from: https://digital.nhs.uk/data-and-information/publications/statistical/nationa l-child-measurement-programme/2016-17-school-year.]

Nishio, S., Y. Gibert, L. Bernard, F. Brunet, G. Triqueneaux and V. Laudet (2008). "Adiponectin and adiponectin receptor genes are coexpressed during zebrafish embryogenesis and regulated by food deprivation." *Dev Dyn* **237**(6): 1682-1690.

Noordam, C., V. Dhir, J. C. McNelis, F. Schlereth, N. A. Hanley, N. Krone, J. A. Smeitink, R. Smeets, F. Sweep, H. L. Claahsen-van der Grinten and W. Arlt (2009). "Inactivating PAPSS2 mutations in a patient with premature pubarche." *N Eng J Med* **360**(22): 2310-2318.

Novoselova, T.V., D. Jackson, D.C. Campbell, A.J. Clark, L.F (2013). “Melanocortin receptor accessory proteins in adrenal gland physiology and beyond.” *J Endocrinol* **217**(1):R1-11.

O'Brien, R. M., P. C. Lucas, C. D. Forest, M. A. Magnuson and D. K. Granner (1990). "Identification of a sequence in the PEPCK gene that mediates a negative effect of insulin on transcription." *Science* **249**(4968): 533-537.

O'Reilly, M. W., P. J. House and J. W. Tomlinson (2014). "Understanding androgen action in adipose tissue." *J Steroid Biochem Mol Biol* **143**: 277-284.

O’Reilly MW, K. P., Jenkinson C, et al (2017). "11-Oxygenated C19 steroids are the predominant androgens in polycystic ovary syndrome." *J Clin Endocrinol Metab* **102**: 840-848.

Oakes, J. A., L. Barnard, K. H. Storbeck, V. T. Cunliffe and N. P. Krone (2020). "11β-Hydroxylase loss disrupts steroidogenesis and reproductive function in zebrafish." *J Endocrinol* **247**(2): 197-212.

Oakes, J. A., N. Li, B. R. C. Wistow, A. Griffin, L. Barnard, K. H. Storbeck, V. T. Cunliffe and N. P. Krone (2019). "Ferredoxin 1b deficiency leads to testis disorganization, impaired spermatogenesis, and feminization in zebrafish." *Endocrinology* **160**(10): 2401-2416.

Obradovic, M., E. Sudar-Milovanovic, S. Soskic, M. Essack, S. Arya, A. J. Stewart, T. Gojobori and E. R. Isenovic (2021). "Leptin and obesity: role and clinical implication." *Front Endocrinol (Lausanne)* **12**: 585887.

Oliveira, L. M., J. A. Faria, Jr., D. Nunes-Silva, R. Lago and M. B. Toralles (2013). "[Elevated levels of leptin and LDL-cholesterol in patients with well controlled congenital adrenal hyperplasia]." *Arq Bras Endocrinol Metabol* **57**(5): 354-359.

Ondruskova, N., A. Cechova, H. Hansikova, T. Honzik and J. Jaeken (2021). "Congenital disorders of glycosylation: Still "hot" in 2020." *Biochim Biophys Acta Gen Subj* **1865**(1): 129751.

Pallan, P. S., L. D. Nagy, L. Lei, E. Gonzalez, V. M. Kramlinger, C. M. Azumaya, Z. Wawrzak, M. R. Waterman, F. P. Guengerich and M. Egli (2015). "Structural and kinetic basis of steroid 17α,20-lyase activity in teleost fish cytochrome P450 17A1 and its absence in cytochrome P450 17A2." *J Biol Chem* **290**(6): 3248-3268.

Parajes, S., A. Griffin, A. E. Taylor, I. T. Rose, I. Miguel-Escalada, Y. Hadzhiev, W. Arlt, C. Shackleton, F. Muller and N. Krone (2013). "Redefining the initiation and maintenance of zebrafish interrenal steroidogenesis by characterizing the key enzyme cyp11a2." *Endocrinology* **154**(8): 2702-2711.

Parekh, S., C. Ziegenhain, B. Vieth, W. Enard and I. Hellmann (2016). "The impact of amplification on differential expression analyses by RNA-seq." *Scientific Reports* **6**(1): 25533.

Parry-Billings, M., B. Leighton, G. D. Dimitriadis, J. Bond and E. A. Newsholme (1990). "Effects of physiological and pathological levels of glucocorticoids on skeletal-muscle glutamine-metabolism in the rat." *Biochem Pharmacol* **40**(5): 1145-1148.

Petersen, M. C., D. F. Vatner and G. I. Shulman (2017). "Regulation of hepatic glucose metabolism in health and disease." *Nat Rev Endocrinol* **13**(10): 572-587.

Pfaffl, M. W. (2001). "A new mathematical model for relative quantification in real-time RT-PCR." *Nucleic Acids Res* **29**(9): e45.

Picard, M. and B. S. McEwen (2018). "Psychological stress and mitochondria: a systematic review." *Psychosom Med* **80**(2): 141-153.

Plat, L., R. Leproult, M. L'Hermite-Baleriaux, F. Fery, J. Mockel, K. S. Polonsky and E. Van Cauter (1999). "Metabolic effects of short-term elevations of plasma cortisol are more pronounced in the evening than in the morning." *J Clin Endocrinol Metab* **84**(9): 3082-3092.

Poyrazoglu, S., H. Gunoz and F. Darendeliler (2003). "Serum leptin levels in patients with 21-hydroxylase deficiency before and after treatment." *Turk J Pediatr* **45**(1): 33-38.

Priyadarshini, E. and C. V. Anuradha (2017). "Glucocorticoid antagonism reduces

insulin resistance and associated lipid abnormalities in high-fructose-fed mice." *Can J Diabetes* **41**(1): 41-51.

Psarra, A. M. and C. E. Sekeris (2011). "Glucocorticoids induce mitochondrial gene transcription in HepG2 cells: role of the mitochondrial glucocorticoid receptor." *Biochim Biophys Acta* **1813**(10): 1814-1821.

Puglisi, S., A. Rossini, I. Tabaro, S. Cannavò, F. Ferrau, M. Ragonese, G. Borretta, M. Pellegrino, F. Dughera, A. Parisi, A. Latina, A. Pia, M. Terzolo and G. Reimondo (2021). "What factors have impact on glucocorticoid replacement in adrenal insufficiency: a real-life study." *J Endocrinol Invest* **44**(4): 865-872.

Quek, S. I. and W. K. Chan (2009). "Transcriptional activation of zebrafish cyp11a1 promoter is dependent on the nuclear receptor Ff1b." *J Mol Endocrinol* **43**(3): 121-130.

Rao, G. (2016). "Diagnosis, epidemiology, and management of hypertension in children." *Pediatrics* **138**(2).

Reisch, N., W. Arlt and N. Krone (2011). "Health problems in congenital adrenal hyperplasia due to 21-hydroxylase deficiency." *Horm Res Paediatr* **76**(2): 73-85.

Rocha, F., J. Dias, S. Engrola, P. Gavaia, I. Geurden, M. T. Dinis and S. Panserat (2015). "Glucose metabolism and gene expression in juvenile zebrafish (Danio rerio) challenged with a high carbohydrate diet: effects of an acute glucose stimulus during late embryonic life." *Br J Nutr* **113**(3): 403-413.

Rousseau, A. and A. Bertolotti (2018). "Regulation of proteasome assembly and activity in health and disease." *Nat Rev Mol Cell Biol* **19**(11): 697-712.

Russell, C. D., R. N. Petersen, S. P. Rao, M. R. Ricci, A. Prasad, Y. Zhang, R. E. Brolin and S. K. Fried (1998). "Leptin expression in adipose tissue from obese humans: depot-specific regulation by insulin and dexamethasone." *Am J Physiol* **275**(3): E507-515.

Saad, M. F., S. Damani, R. L. Gingerich, M. G. Riad-Gabriel, A. Khan, R. Boyadjian, S. D. Jinagouda, K. el-Tawil, R. K. Rude and V. Kamdar (1997). "Sexual dimorphism in plasma leptin concentration." *J Clin Endocrinol Metab* **82**(2): 579-584.

Sarafoglou, K., C. L. Zimmerman, M. T. Gonzalez-Bolanos, B. A. Willis and R. Brundage (2015). "Interrelationships among cortisol, 17-hydroxyprogesterone, and androstenendione exposures in the management of children with congenital adrenal hyperplasia." *J Investig Med* **63**(1): 35-41.

Sartin, J. L., R. J. Kemppainen, E. S. Coleman, B. Steele and J. C. Williams (1994). "Cortisol inhibition of growth hormone-releasing hormone-stimulated growth hormone release from cultured sheep pituitary cells." *J Endocrinol* **141**(3): 517-525.

Savary, I., E. Debras, D. Dardevet, C. Sornet, P. Capitan, J. Prugnaud, P. P. Mirand and J. Grizard (1998). "Effect of glucocorticoid excess on skeletal muscle and heart protein synthesis in adult and old rats." *Br J Nutr***79**(3): 297-304.

Saygili, F., A. Oge and C. Yilmaz (2005). "Hyperinsulinemia and insulin insensitivity in women with nonclassical congenital adrenal hyperplasia due to 21-hydroxylase deficiency: the relationship between serum leptin levels and chronic hyperinsulinemia."

Horm Res **63**(6): 270-274.

Schakman, O., S. Kalista, C. Barbé, A. Loumaye and J. P. Thissen (2013). "Glucocorticoid-induced skeletal muscle atrophy." *Int J Biochem Cell Biol* **45**(10): 2163-2172.

Schiffer, L., W. Arlt and K. H. Storbeck (2018). "Intracrine androgen biosynthesis, metabolism and action revisited." *Mol Cell Endocrinol* **465**: 4-26.

Seiliez, I., F. Médale, P. Aguirre, M. Larquier, L. Lanneretonne, H. Alami-Durante, S. Panserat and S. Skiba-Cassy (2013). "Postprandial regulation of growth- and metabolism-related factors in zebrafish." *Zebrafish* **10**(2): 237-248.

Shah, O. J., J. A. Iniguez-Lluhi, A. Romanelli, S. R. Kimball and L. S. Jefferson (2002). "The activated glucocorticoid receptor modulates presumptive autoregulation of ribosomal protein S6 protein kinase, p70 S6K." *J Biol Chem* **277**(4): 2525-2533.

Shashaj, B., R. Luciano, B. Contoli, G. S. Morino, M. R. Spreghini, C. Rustico, R. W. Sforza, B. Dallapiccola and M. Manco (2016). "Reference ranges of HOMA-IR in normal-weight and obese young Caucasians." *Acta Diabetol* **53**(2): 251-260.

Sherwin, R. S. and L. Sacca (1984). "Effect of epinephrine on glucose metabolism in humans: contribution of the liver." *Am J Physiol Endocrinol* **247**(2): E157-E165.

Shi, C., Y. Lu, G. Zhai, J. Huang, G. Shang, Q. Lou, D. Li, X. Jin, J. He, Z. Du, J. Gui and Z. Yin (2020). "Hyperandrogenism in POMCa-deficient zebrafish enhances somatic growth without increasing adiposity." *J Mol Cell Biol* **12**(4): 291-304.

Shi, W. J., G. Y. Huang, Y. X. Jiang, D. D. Ma, H. X. Chen, M. Z. Huang, L. P. Hou, L. Xie and G. G. Ying (2020). "Medroxyprogesterone acetate affects eye growth and the transcription of associated genes in zebrafish." *Ecotoxicol Environ Saf* **193**: 110371.

Smith, T. J. (1988). "Glucocorticoid regulation of glycosaminoglycan synthesis in cultured human skin fibroblasts: evidence for a receptor-mediated mechanism involving effects on specific de novo protein synthesis." *Metabolism* **37**(2): 179-184.

Speiser, P. W., W. Arlt, R. J. Auchus, L. S. Baskin, G. S. Conway, D. P. Merke, H. F. L. Meyer-Bahlburg, W. L. Miller, M. H. Murad, S. E. Oberfield and P. C. White (2018). "Congenital adrenal hyperplasia due to steroid 21-hydroxylase deficiency: an endocrine society clinical practice guideline." *J Clin Endocrinol Metab* **103**(11): 4043-4088.

Steenbergen, P. J., M. K. Richardson and D. L. Champagne (2011). "The use of the zebrafish model in stress research." *Prog Neuropsychopharmacol Biol Psychiatry* **35**(6): 1432-1451.

Storbeck, K. H., L. Schiffer, E. S. Baranowski, V. Chortis, A. Prete, L. Barnard, L. C. Gilligan, A. E. Taylor, J. Idkowiak, W. Arlt and C. H. L. Shackleton (2019). "Steroid metabolome analysis in disorders of adrenal steroid biosynthesis and metabolism." *Endocr Rev* **40**(6): 1605-1625.

Stratakis, C. A. (2006). "Cortisol and growth hormone: clinical implications of a complex, dynamic relationship." *Pediatr Endocrinol Rev* **3 Suppl 2**: 333-338.

Tan, S. T., T. Ramesh, X. R. Toh and L. N. Nguyen (2020). "Emerging roles of

lysophospholipids in health and disease." *Prog Lipid Res* **80**: 101068.

Tarca, A. L., G. Bhatti and R. Romero (2013). "A comparison of gene set analysis methods in terms of sensitivity, prioritization and specificity." *PLoS One* **8**(11): e79217.

Tehrani, Z. and S. Lin (2011). "Endocrine pancreas development in zebrafish." *Cell Cycle* **10**(20): 3466-3472.

To, T. T., S. Hahner, G. Nica, K. B. Rohr, M. Hammerschmidt, C. Winkler and B. Allolio (2007). "Pituitary-interrenal interaction in zebrafish interrenal organ development." *Mol Endocrinol* **21**(2): 472-485.

Tokarz, J., G. Möller, M. H. de Angelis and J. Adamski (2013). "Zebrafish and steroids: what do we know and what do we need to know?" *J Steroid Biochem Mol Biol* **137**: 165-173.

Tomlinson, J. J., A. Boudreau, D. Wu, H. Abdou Salem, A. Carrigan, A. Gagnon, A. J. Mears, A. Sorisky, E. Atlas and R. J. Haché (2010). "Insulin sensitization of human preadipocytes through glucocorticoid hormone induction of forkhead transcription factors." *Mol Endocrinol* **24**(1): 104-113.

Torky, A., N. Sinaii, S. Jha, J. Desai, D. El-Maouche, A. Mallappa and D. P. Merke (2021). "Cardiovascular disease risk factors and metabolic morbidity in a longitudinal study of congenital adrenal hyperplasia." *J Clin Endocrinol Metab*.

Turcu, A. F. and R. J. Auchus (2015). "Adrenal steroidogenesis and congenital adrenal hyperplasia." *Endocrinol Metab Clin North Am* **44**(2): 275-296.

Turcu, A. F., A. Mallappa, M. S. Elman, N. A. Avila, J. Marko, H. Rao, A. Tsodikov, R. J. Auchus and D. P. Merke (2017). "11-oxygenated androgens are biomarkers of adrenal volume and testicular adrenal rest tumors in 21-hydroxylase deficiency." *J Clin Endocrinol Metab* **102**(8): 2701-2710.

Turcu, A. F., A. T. Nanba, R. Chomic, S. K. Upadhyay, T. J. Giordano, J. J. Shields, D. P. Merke, W. E. Rainey and R. J. Auchus (2016). "Adrenal-derived 11-oxygenated 19-carbon steroids are the dominant androgens in classic 21-hydroxylase deficiency." *Eur J Endocrinol* **174**(5): 601-609.

Uchoa, E. T., H. A. C. Sabino, S. G. Ruginsk, J. Antunes-Rodrigues and L. L. K. Elias (2009). "Hypophagia induced by glucocorticoid deficiency is associated with an increased activation of satiety-related responses." *J Appl Physiol* **106**(2): 596-604.

Uchoa, E. T., L. Silva, M. de Castro, J. Antunes-Rodrigues and L. L. K. Elias (2012). "Glucocorticoids are required for meal-induced changes in the expression of hypothalamic neuropeptides." *Neuropeptides* **46**(3): 119-124.

van der Sluis, R. J., G. H. van Puijvelde, T. J. C. Van Berkel and M. Hoekstra (2012). "Adrenalectomy stimulates the formation of initial atherosclerotic lesions: Reversal by adrenal transplantation." *Atherosclerosis* **221**(1): 76-83.

Van Harmelen, V., S. Reynisdottir, P. Eriksson, A. Thörne, J. Hoffstedt, F. Lönnqvist and P. Arner (1998). "Leptin secretion from subcutaneous and visceral adipose tissue in women." *Diabetes* **47**(6): 913-917.

Vandevyver, S., L. Dejager and C. Libert (2014). "Comprehensive overview of the structure and regulation of the glucocorticoid receptor." *Endocr Rev* **35**(4): 671-693.

Vera-Chang MN, St-Jacques AD, Gagné R, Martyniuk CJ, Yauk CL, Moon TW, Trudeau VL (2018). “Transgenerational hypocortisolism and behavioral disruption are induced by the antidepressant fluoxetine in male zebrafish Danio rerio.” *Proc Natl Acad Sci U S A* **115**(52) E12435-e12442.

Vera-Chang MN, Moon TW, Trudeau VL (20190. “Cortisol disruption and transgenerational alteration in the expression of stress-related genes in zebrafish larvae following fluoxetine exposure.” *Toxicol Appl Pharmacol* **382**:114742.

Vitellius, G. and M. Lombes (2020). "Genetics in endocrinology: Glucocorticoid resistance syndrome." *Eur J Endocrinol* **182**(2): R15-r27.

Volkl, T. M. K., D. Simm, A. Korner, W. Kiess, J. Kratzsch and H. G. Dorr (2009). "Adiponectin levels are high in children with classic congenital adrenal hyperplasia (CAH) due to 21-hydroxylase deficiency." *Acta Pædiatrica* **98**(5): 885-891.

Völkl, T. M. K., D. Simm, C. Beier and H. G. Dörr (2006). "Obesity among children and adolescents with classic congenital adrenal hyperplasia due to 21-hydroxylase deficiency." *Pediatrics* **117**(1): e98.

Völkl, T. M., D. Simm, A. Körner, W. Kiess, J. Kratzsch and H. G. Dörr (2009). "Adiponectin levels are high in children with classic congenital adrenal hyperplasia (CAH) due to 21-hydroxylase deficiency." *Acta Paediatr* **98**(5): 885-891.

Völkl, T. M., D. Simm, A. Körner, W. Rascher, W. Kiess, J. Kratzsch and H. G. Dörr (2009). "Does an altered leptin axis play a role in obesity among children and adolescents with classical congenital adrenal hyperplasia due to 21-hydroxylase deficiency?" *Eur J Endocrinol* **160**(2): 239-247.

Völkl, T. M., D. Simm, J. Dötsch, W. Rascher and H. G. Dörr (2006). "Altered 24-hour blood pressure profiles in children and adolescents with classical congenital adrenal hyperplasia due to 21-hydroxylase deficiency." *J Clin Endocrinol Metab* **91**(12): 4888-4895.

Wang, Z. V. and P. E. Scherer (2016). "Adiponectin, the past two decades." *J Mol Cell Biol* **8**(2): 93-100.

Weger, B. D., M. Weger, B. Görling, A. Schink, C. Gobet, C. Keime, G. Poschet, B. Jost, N. Krone, R. Hell, F. Gachon, B. Luy and T. Dickmeis (2016). "Extensive regulation of diurnal transcription and metabolism by glucocorticoids." *PLoS Genet* **12**(12): e1006512.

Weger, M., B. D. Weger, B. Gorling, G. Poschet, M. Yildiz, R. Hell, B. Luy, T. Akcay, T. Guran, T. Dickmeis, F. Muller and N. Krone (2018). "Glucocorticoid deficiency causes transcriptional and post-transcriptional reprogramming of glutamine metabolism." *Ebiomedicine* **36**: 376-389.

Weger, M., N. Diotel, B. D. Weger, T. Beil, A. Zaucker, H. L. Eachus, J. A. Oakes, J. L. do Rego, K. H. Storbeck, P. Gut, U. Strähle, S. Rastegar, F. Müller and N. Krone

(2018). "Expression and activity profiling of the steroidogenic enzymes of glucocorticoid biosynthesis and the fdx1 co-factors in zebrafish." *J Neuroendocrinol* **30**(4): e12586.

Weihe, P. and S. Weihrauch-Blüher (2019). "Metabolic syndrome in children and adolescents: diagnostic criteria, therapeutic options and perspectives." *Curr Obes Rep* **8**(4): 472-479.

Wilson, K. S., G. Matrone, D. E. Livingstone, E. A. Al-Dujaili, J. J. Mullins, C. S. Tucker, P. W. Hadoke, C. J. Kenyon and M. A. Denvir (2013). "Physiological roles of glucocorticoids during early embryonic development of the zebrafish (Danio rerio)." *J Physiol* **591**(24): 6209-6220.

Wu, Y. S., E. J. Barrett, W. Long and Z. Q. Liu (2004). "Glucocorticoids differentially modulate insulin-mediated protein and glycogen synthetic signaling downstream of protein kinase B in rat myocardium." *Endocrinology* **145**(3): 1161-1166.

Xu, W. J., K. Guo, J. L. Shi, C. T. Guo, J. L. Xu, R. Zheng, S. W. Jiang and J. Chai (2022). "Glucocorticoid regulates the synthesis of porcine muscle protein through m(6)a modified amino acid transporter SLC7A7." *Int J Mol Sci* **23**(2).

Yadav, A., M. A. Kataria, V. Saini and A. Yadav (2013). "Role of leptin and adiponectin in insulin resistance." *Clin Chim Acta* **417**: 80-84.

Yang, B. Y., G. Zhai, Y. L. Gong, J. Z. Su, X. Y. Peng, G. H. Shang, D. Han, J. Y. Jin, H. K. Liu, Z. Y. Du, Z. Yin and S. Q. Xie (2018). "Different physiological roles of insulin receptors in mediating nutrient metabolism in zebrafish." *Am J Physiol Endocrinol Metab* **315**(1): E38-e51.

Yu, G., X. Liu, D. Zhang, J. Wang, G. Ouyang, Z. Chen and W. Xiao (2020). "Zebrafish Nedd8 facilitates ovarian development and the maintenance of female secondary sexual characteristics via suppression of androgen receptor activity." *Development* **147**(18).

Yu, H., H. Lee, A. Herrmann, R. Buettner and R. Jove (2014). "Revisiting STAT3 signalling in cancer: new and unexpected biological functions." *Nat Rev Cancer* **14**(11): 736-746.

Zang, L., L. A. Maddison and W. Chen (2018). "Zebrafish as a model for obesity and diabetes." *Front Cell Dev Biol* **6**: 91.

Zhang, H. J., J. Yang, M. N. Zhang, C. Q. Liu, M. Xu, X. J. Li, S. Y. Yang and X. Y. Li (2010). "Metabolic disorders in newly diagnosed young adult female patients with simple virilizing 21-hydroxylase deficiency." *Endocrine* **38**(2): 260-265.

Zhang, P., L. O'Loughlin, D. N. Brindley and K. Reue (2008). "Regulation of lipin-1 gene expression by glucocorticoids during adipogenesis." *J Lipid Res* **49**(7): 1519-1528.

Zhang, Q., D. Ye, H. Wang, Y. Wang, W. Hu and Y. Sun (2020). "Zebrafish cyp11c1 knockout reveals the roles of 11-ketotestosterone and cortisol in sexual development and reproduction." *Endocrinology* **161**(6).

Zhang, S. and P. Cui (2014). "Complement system in zebrafish." *Dev Comp Immunol* **46**(1): 3-10.

Zhao, Y., Z. Yang, J. K. Phelan, D. A. Wheeler, S. Lin and E. R. McCabe (2006).

"Zebrafish dax1 is required for development of the interrenal organ, the adrenal cortex equivalent." *Mol Endocrinol* **20**(11): 2630-2640.

Zhou, L. Y., D. S. Wang, T. Kobayashi, A. Yano, B. Paul-Prasanth, A. Suzuki, F. Sakai and Y. Nagahama (2007). "A novel type of P450c17 lacking the lyase activity is responsible for C21-steroid biosynthesis in the fish ovary and head kidney." *Endocrinology* **148**(9): 4282-4291.

Ziegler, A. B. and G. Tavosanis (2019). "Glycerophospholipids - Emerging players in neuronal dendrite branching and outgrowth." *Dev Biol* **451**(1): 25-34.

Zurita-Cruz, J. N., M. Villasís-Keever, L. Damasio-Santana, L. Manuel-Apolinar, R. Ferrusca-Ceja, E. Nishimura-Meguro, A. J. Rivera-Hernández and E. Garrido-Magaña (2018). "[Association of leptin with cardiometabolic factors in schoolchildren and adolescents with congenital adrenal hyperplasia]." *Gac Med Mex* **154**(2): 202-208.