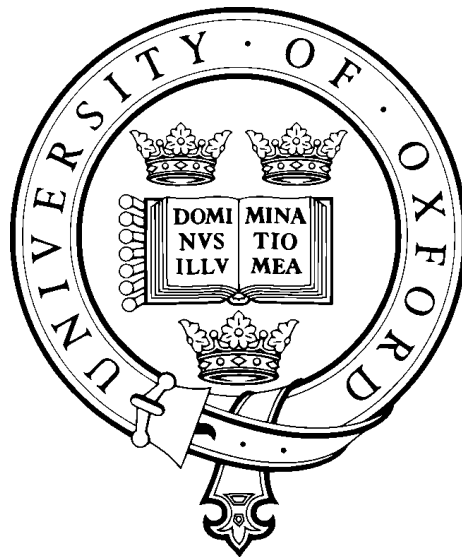


**Investigation into the
pathogenesis and behaviour of
two disorders of cholesterol
homeostasis**



Joanna L. Cross

Linacre College

Department of Pharmacology

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2016

Acknowledgements

First and foremost, I would like to thank my mentors, Prof. Frances Platt in Oxford and Dr. Forbes Porter at the NIH. Without their continued support and guidance, this thesis would not have been possible. I have learnt so much under their leadership and have developed into the scientist I am today, furnished with multiple skills to take me forward to the next stage of my career.

Additionally, these four years would not have been possible without the members of both laboratories; Allie, James, Paul, Danielle, Dave, David, Claire and Lauren in the UK and Celine, Antony, Katy, Jackie, Emily, Chris, Alex, Stephanie, Nicole, and Lee-Ann in the USA. A special mention must go to Dr. Ian Williams, who was my mentor from my first foray into research in 2011 and continued to help and guide me throughout my time in the UK and USA. To all of you, I thank you for your help, friendship and the many excellent times that were had at both the University Club and Rock Bottom!

Lastly, I must thank all my friends and family for their continued support. To all my friends, of whom there are too many to name individually, my family, Linda and Peter and Adam, for multiple cups of tea, wine and for reading through my thesis, and to my wonderful boyfriend, Ryan, who put up with the hours I had to spend in the lab and at my computer and was always there to help me relax at the end of a hard day.

Thank you all!

Abstract

Cholesterol is essential for many life processes, including correct development, fluidity of cell membranes, production of steroid hormones and bile acids, and is a major component of myelin. Smith-Lemli-Opitz Syndrome (SLOS) and Niemann-Pick disease type C (NPC) are devastating diseases and both involve dysregulation of cholesterol homeostasis. SLOS is caused by a defect in 7-dehydrocholesterol reductase (*DHCR7*), resulting in increased levels of 7-dehydrocholesterol (7DHC) and decreased levels of cholesterol. On the other hand, NPC is caused by defects in *NPC1* or *NPC2*. These genes encode two lysosomal proteins that are responsible for the transport of cholesterol and other lipids out of lysosomes. Consequently, defects in these proteins results in accumulation of unesterified cholesterol within late endosomes/lysosomes. The severity ranges of both disorders are broad, and no or limited therapeutic options are available.

This thesis aimed to establish the incidence and mechanisms behind SLOS and NPC in order to aid development of novel therapeutic interventions. Using a bioinformatics approach, **Chapter 2** showed that the estimated incidences of classical SLOS and NPC were similar to clinical reports. However, the analysis suggested that a late onset form of NPC1 could be more prevalent. **Chapter 3** investigated the behavioural phenotype of the SLOS *Dhcr7*^{T93M/Δ3-5} mouse model and it was found that some parallels could be drawn to the behaviour observed in SLOS patients. It also highlighted some defects in neuro-anatomy that could potentially explain certain cognitive defects. **Chapter 4** explored the suggestion of oxidative stress in the *Dhcr7*^{T93M/Δ3-5} mouse model. However, this study did

not support a role for oxidative stress in this model. Genetic insights were generated via an RNAseq study on SLOS and NPC patient fibroblasts, which suggested that the WNT signalling pathway could play a role in the pathogenesis of SLOS. This pathway was also highlighted when the cells were treated with miglustat, the only approved therapy for NPC. However, the main pathway apparently affected by this drug in both SLOS and NPC was cell proliferation.

Table of Contents

ACKNOWLEDGEMENTS	I
ABSTRACT	II
ABBREVIATIONS	VIII
CHAPTER 1: INTRODUCTION	1
1.1 CHOLESTEROL METABOLISM	1
1.1.i: <i>The cholesterol biosynthesis pathway.....</i>	<i>2</i>
1.1.ii: <i>Regulation of cholesterol biosynthesis</i>	<i>6</i>
1.1.iii: <i>Cholesterol in development.....</i>	<i>8</i>
1.1.iv: <i>Further roles of cholesterol in the cell</i>	<i>11</i>
1.1.v: <i>Cholesterol and disease.....</i>	<i>14</i>
1.2 SMITH-LEMLI-OPITZ SYNDROME.....	16
1.2.i: <i>Biochemical findings in Smith-Lemli-Opitz Syndrome</i>	<i>16</i>
1.2.ii: <i>Smith-Lemli-Opitz Syndrome phenotype.....</i>	<i>18</i>
1.2.iii: <i>Altered development in Smith-Lemli-Opitz Syndrome.....</i>	<i>20</i>
1.2.iv: <i>Altered biosynthesis and metabolic pathways in Smith-Lemli-Opitz Syndrome.....</i>	<i>21</i>
1.3 NIEMANN-PICK DISEASE TYPE C DISEASE	24
1.3.i: <i>Structures of NPC1 and NPC2</i>	<i>25</i>
1.3.iii: <i>Niemann-Pick type C disease phenotype and diagnosis</i>	<i>30</i>
1.3.iv: <i>Altered metabolism in Niemann-Pick type C disease</i>	<i>33</i>
1.3.iv: <i>Neurological defects in Niemann-Pick type C disease.....</i>	<i>34</i>
1.4 ANIMAL MODELS	35
1.4.i <i>Smith-Lemli-Opitz Syndrome animal models.....</i>	<i>35</i>

1.4.ii Niemann-Pick type C disease animal models.....	40
1.5 THERAPEUTIC INTERVENTION.....	44
1.5.i: Smith-Lemli-Opitz Syndrome.....	44
1.5.ii: Niemann-Pick type C disease	46
1.6 AIMS OF THESIS	51
 CHAPTER 2: DETERMINING ALLELIC FREQUENCY OF SMITH-LEMLI-OPITZ SYNDROME AND NIEMANN-PICK DISEASE TYPE C VIA USE OF MASSIVELY PARALLEL SEQUENCING DATA SETS	
53	53
2.1 INTRODUCTION	53
2.1.i: Smith-Lemli-Opitz Syndrome.....	53
2.1.ii: Niemann-Pick type C disease	55
2.1.iii: Aims of Chapter 2.....	57
2.2 MATERIALS AND METHODS.....	58
2.2.i Data sets	58
2.2.ii Variant calls and annotation	59
2.2.iii Determination of pathogenicity	59
2.2.iv Molecular analysis	60
2.2.v Determination of predicted disease incidence	61
2.3 RESULTS.....	63
2.3.i Smith-Lemli-Opitz Syndrome	63
2.3.ii Niemann-Pick type C1 disease and Niemann-Pick type C2 disease.....	70
2.4 DISCUSSION	80
 CHAPTER 3: ANALYSIS OF THE BEHAVIOUR AND NEUROLOGICAL STRUCTURE OF <i>DHCR7</i>^{T93M/Δ3-5} MICE	
84	84
3.1: INTRODUCTION.....	84
3.1.i: Aims of Chapter 3	88
3.2: MATERIALS AND METHODS	89

3.2.i: Mice	89
3.2.ii: Burrowing test	89
3.2.iii: Nestlet test.....	90
3.2.iv: Marble burying test.....	92
3.2.v: Elevated plus maze	93
3.2.vi: Open field.....	94
3.2.vii: Histology.....	95
3.2.viii: Statistical analysis	96
3.3: RESULTS.....	97
3.3.i: Burrowing test.....	97
3.3.ii: Nestlet shredding test	101
3.3.iii: Marble Burying.....	105
3.3.iv: Elevated plus maze.....	107
3.3.v: Open field	111
3.3.vi: Histological analysis.....	113
3.4: DISCUSSION	118

CHAPTER 4: ANALYSIS OF RNA EXPRESSION IN SMITH-LEMLI-OPITZ

SYNDROME AND NIEMANN-PICK TYPE C DISEASE	125
4.1. INTRODUCTION	125
4.1.i: Aims of Chapter 4	129
4.2: MATERIALS AND METHODS	130
4.2.i: RNA isolation.....	130
4.2.ii: cDNA synthesis and profiler array	130
4.2.iii: Taqman validation.....	131
4.2.iv: Western blot.....	132
4.2.v: Cell culture for RNAseq	133
4.2.vi: Collecting cell pellets.....	134

4.2.vii: Immunofluorescence.....	135
4.2.viii: FACS staining protocol.....	136
4.2.ix: RNAseq.....	137
4.3: RESULTS.....	139
4.3.i: RT ² Oxidative Stress Array	139
4.3.ii: Validation of significant genes.....	141
4.3.iii: Protein analysis of EPX	143
4.3.iv: Superoxide dismutase: further investigation.....	144
4.3.v: RNAseq	146
4.3.vi: RNAseq – miglustat study.....	147
4.4.vii: RNAseq - results.....	150
4.4: DISCUSSION	156
4.4.i: Oxidative Stress Array.....	156
4.4.ii: RNAseq – Miglustat study.....	158
CHAPTER 5: CONCLUDING REMARKS	164
FUTURE DIRECTION	170
REFERENCES.....	172
APPENDICES	203
DHCR7 SUMMARY TABLE	203
NPC1 SUMMARY TABLE	206
NPC2 SUMMARY TABLE	210
RESEARCH PUBLICATIONS	211

Abbreviations

129/SvHsd: 129 inbred mouse strain

3 α HSD: 3 α -hydroxysterol dehydrogenase

7DHC: 7-dehydrocholesterol

8DHC: 8-dehydrocholesterol

AA: African American

ACTH: Adrenocorticotrophic Hormone

ADAM19: ADAM Metallopeptidase Domain 19

BBB: Blood-Brain Barrier

BIRC5: Baculoviral IAP Repeat Containing 5

BL6/129: Black 6/129 mixed mouse strain

C-SHH: Sonic Hedgehog C-terminal

C57BK/6: C57 Black 6 inbred mouse strain

CALB2: Calbindin 2

CCL5: Chemokine (C-C motif) Ligand 5

CD: Cyclodextrin

CEP55: Centrosomal Protein 55

CNS: Central Nervous System

CPZ: Carboxypeptidase Z

CRABP2: Cellular Retinoic Acid Binding Protein 2

CSNBs: Coding Single Nucleotide Base Variants

DES: Desmin

DHCR24: 24-dehydrocholesterol reductase

DHCR7: 7-dehydrocholesterol reductase

DIS: Dispatched

DMEM: Dulbecco's Modified Eagle Medium

DNA: Deoxyribonucleic Acid

EA: European American

EPX: Eosinophil Peroxidase

ER: Endoplasmic Reticulum

ERT: Enzyme Replacement Therapy

ESP: National Heart, Lung and Blood Institute GO Exome Sequencing Project

FACS: Fluorescence Activated Cell Sorting

FAM64A: Family with Sequence Similarity 64 Member A

FDR: False Discovery Rate

FOXO3: Forkhead Box O3

FOXO4: Forkhead Box O4

FOXS1: Forkhead Box S1

FZD8: Frizzled class Receptor 8

GRID1: Glutamate Ionotropic Receptor Delta type subunit 1

GSEA: Gene Set Enrichment Analysis

GSL: Glycosphingolipid

HGMD: Human Gene Mutation Database ©

HMG-CoA: 3-hydroxy-3-methylglutaryl-coenzyme A

HMGR: HMG-CoA reductase

HMGS: HMG-CoA synthase

HPLC: High Performance Lipid Chromatography

HTR1B: 5-hydroxytryptamine receptor 1B

HTR2A: 5-hydroxytryptamine receptor 2A

IQGAP3: IQ Motif Containing GTPase Activating Protein 3

KIF18B: Kinesin Family Member 18B

KIF20A: Kinesin Family Member 20B

KIFC1: Kinesin Family Member C1

KRT15: Keratin 15

KRT19: Keratin 19

KRT34: Keratin 34

LAMP3: Lysosomal Associated Member Protein 3

LDL: Low-density Lipoprotein

LEF1: Lymphoid Enhancer Binding Factor 1

LPDS: Lipo-deficient Serum

MBP2: Myelin-basic Protein 2

MC4R: Melanocortin 4 Receptor

MKI67: Marker of Proliferation Ki-67

MRI: Magnetic Resonance Imaging

MYPN: Myopalladin

N-SHH: Sonic Hedgehog N-terminal

NGFR: Nerve Growth Factor Receptor

NIH: National Institutes of Health

NP: Niemann-Pick disease

NPA: Niemann-Pick type A disease

NPB: Niemann-Pick type B disease

NPC: Niemann-Pick type C disease

NPC1: Niemann-Pick type C1

NPC2: Niemann-Pick type C2

NPD: Niemann-Pick type D disease

PCs: Purkinje Cells

PRR13: Proline Rich 13

PTC: Patched

RNA: Ribonucleic Acid

ROS: Reactive Oxygen Species

RPMI: Roswell Park Memorial Media

RRM2: Ribonucleotide Reductase Regulatory subunit 2

SCAP: SREBP Cleavage-activating Protein

SERCA: Sarco/Endoplasmic Reticulum Calcium ATPase

SEM: Standard Error of Mean

SHH: Sonic Hedgehog

SLOS: Smith-Lemli-Opitz Syndrome

SMO: Smoothened

SNAG1: Sorting Nexin 1

SOD1: Superoxide Dismutase 1

SOD2: Superoxide Dismutase 2

SOD3: Superoxide Dismutase 3

SQS: Squalene Synthase

SRE: Sterol Regulating Element

SREBP: Sterol Regulating Element Binding Protein

SRT: Substrate Reduction Therapy

SSD: Sterol-sensing Domain

TM4SF: Transmembrane 4 Superfamily, isoform B

VSGP: Vertical Supranuclear Gaze Palsy

WNT10B: Wingless-Type Integration Site Family Member 10B

WNT2: Wingless-Type Integration Site Family Member 2

WNT4: Wingless-Type Integration Site Family Member 4

WNT9A: Wingless-Type Integration Site Family Member 9A

ZPA: Zone of Polarising Activity

Chapter 1: Introduction

*“Equipped with his five senses, man explores the universe around him and
calls the adventure Science.”*

Edwin Powell Hubble

1.1 Cholesterol metabolism

Since the late 1940s, cholesterol has been associated with heart disease [1] and therefore cholesterol, particularly low-density lipoprotein (LDL) cholesterol, has been considered by many to be “bad” [2]. However, cholesterol is crucial for the proper functioning of the body [3], being both an important component of cell membranes [4] and a precursor in a number of biosynthetic pathways [3]. Consequently, errors at any point of its biosynthesis or metabolism can have devastating consequences [5,6]. In healthy adults, cholesterol is present at 150-200 mg/dL in blood [7] and this level is maintained both by dietary intake and *de novo* synthesis [8], also known as cholesterologenesis. About two thirds of the total body cholesterol is derived from *de novo* synthesis [9] and this contributes around 500 - 1000 mg/day [9,10].

1.1.i: The cholesterol biosynthesis pathway

Cholesterol is a 27-carbon molecule (**Figure 1.1A**) and its biosynthesis is a multi-step process. The cholesterol biosynthetic pathway begins with the conversion of Acetyl-CoA and Acetoacetyl-CoA to 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA) via HMG-CoA synthase (HMGS) [11] (**Figure 1.1B**). The next step converts HMG-CoA to mevalonate via HMG-CoA reductase (HMGR), the rate-limiting step of cholesterol biosynthesis [10]. HMGR is found in membranes of the endoplasmic reticulum (ER) [12] and the N-terminal domain comprises alternating hydrophobic and hydrophilic amino acids which transverse the membrane 7 or 8 times, depending on the model [13,14] (**Figure 1.1C**). The C-terminal domain is responsible for the catalytic activity of the enzyme [11]. This approximately 53kDa soluble fragment can be released from the protein via proteolytic cleavage and retains all of the catalytic ability [15].

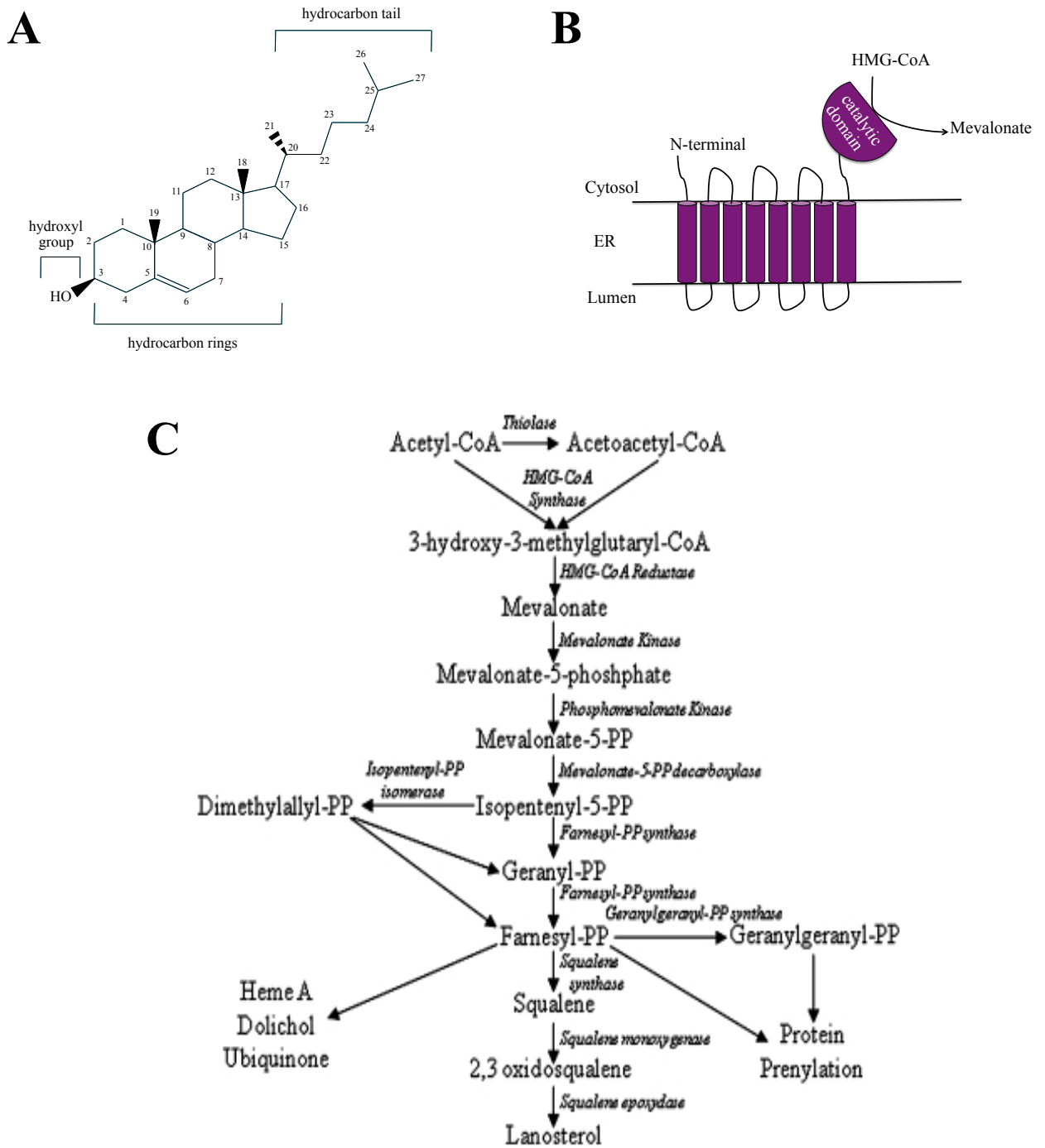


Figure 1.1: Cholesterol biosynthesis: pre-squalene pathway

Figure 1.1 (previous page): Cholesterol biosynthesis: pre-squalene pathway**A) Molecular structure of cholesterol**

Cholesterol is a 27-carbon molecule comprising a hydroxyl group attached to C-3. The middle section is composed of hydrocarbon rings and attached to carbon 17 is a hydrophobic hydrocarbon tail.

B) Diagram showing HMG-CoA reductase.

This model shows the N-terminal domain, which comprises the sterol sensing domain, transversing the membrane 8 times. The catalytic domain, which is responsible for the reduction of HMG-CoA to mevalonate, can be seen at the C-terminal end in the cytosol.

C) The pre-squalene pathway of cholesterol biosynthesis [5].

Squalene represents the first committed step to cholesterol biosynthesis. In the prior steps, other metabolites can be produced, such as Heme A, Dolichol and Ubiquinone. Farnesyl-PP and Geranylgeranyl-PP can additionally be used for protein prenylation, a process that is involved in localising proteins to cell membranes.

The post-squalene pathway completes the conversion to cholesterol (**Figure 1.2**). This can occur via two pathways - the Bloch [16] or the Kandutsch-Russell [17] pathway. Both pathways involve the same catalytic steps but 24-dehydrocholesterol reductase (DHCR24) reduces the C24 double bond at different stages. In the Kandutsch-Russell pathway DHCR24 acts on lanosterol, whereas in the Bloch pathway the reduction occurs in the final step, catalysing the conversion of desmosterol to cholesterol [18].

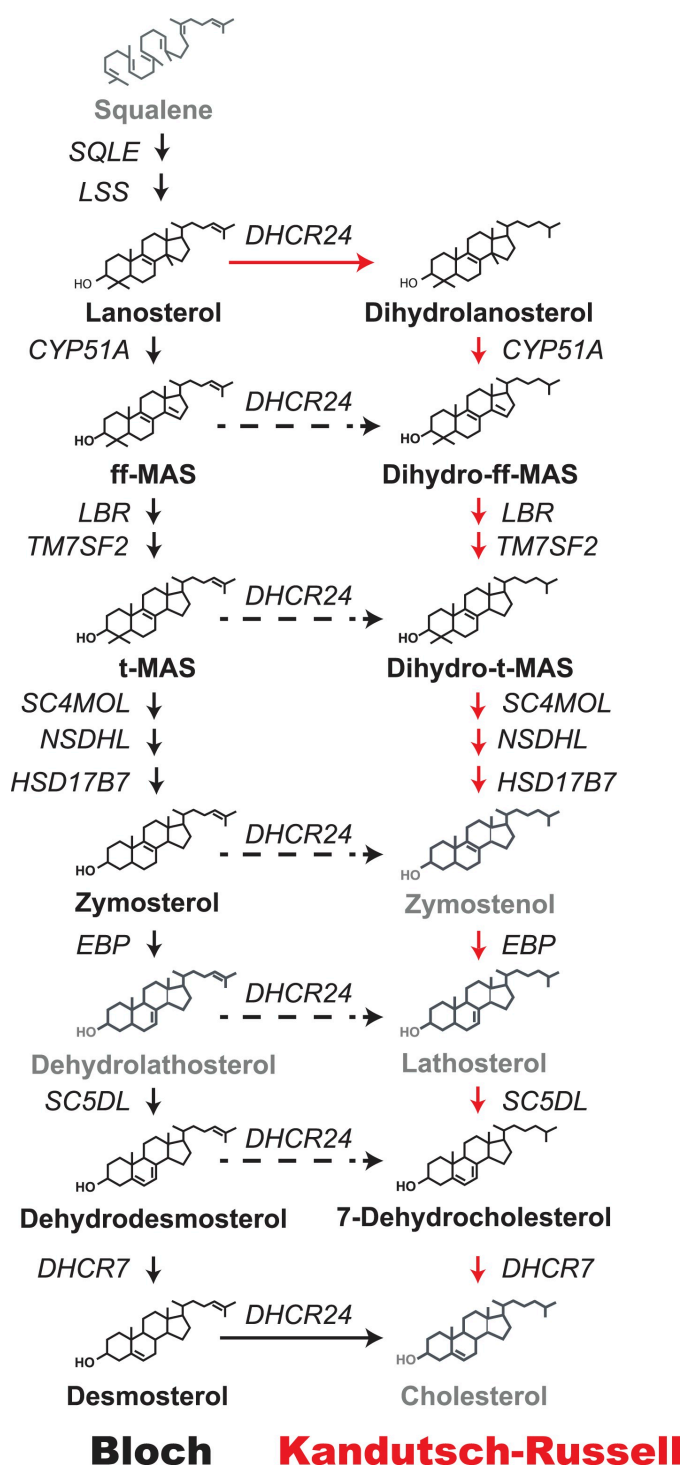


Figure 1.2: The post-squalene pathway of cholesterol synthesis [18]

The pathway is split into the Bloch pathway (left, black arrows) or the Kandutsch-Russell (right, red arrows). DHCR24 reduces the C24 of lanosterol to proceed through the Kandutsch-Russell pathway. In contrast, the C24 is reduced in the final step of the Bloch pathway, converting desmosterol to cholesterol. DHCR24 may additionally be able to convert other intermediates of the Bloch pathway.

1.1.ii: Regulation of cholesterol biosynthesis

Regulation of cholesterologenesis is key to keeping cholesterol levels in the body consistent. This is achieved by feedback mechanisms, which detect both the cellular requirement for cholesterol and the dietary intake and, subsequently, cause synthesis to then be up- or down-regulated. It was discovered that if cells contained the full-length HMGR protein and were exposed to sterols, HMGR activity was repressed and degradation of HMGR increased [8,19]. On the other hand, if cells only contained the catalytic domain, degradation of HMGR was reduced to one fifth of the rate of the full-length protein and this did not increase in the presence of sterols [20]. This led to the suggestion that the transmembrane domain is involved in the degradation process and detection of sterols [20]. Several studies ([21–23]) have tested this by replacing or removing the transmembrane domains of HMGR and incubating the cells with sterols. These studies confirmed the sterol-sensing function of this region, particularly four phenylalanine residues present in transmembrane domain 6 [23].

Other proteins involved in the regulation of cholesterol synthesis also contain sterol-sensing domains (SSDs), for example, the N-terminal domain of SREBP cleavage-activating protein (SCAP) [24]. SCAP is a membrane-bound glycoprotein which complexes with sterol regulating element binding protein (SREBP). These two proteins work together to regulate the transcription of genes involved in cholesterol synthesis (**Figure 1.3**). If sterol levels are low, SCAP enables SREBP to translocate from the ER to the Golgi [25,26]. This results in proteolytic cleavage at sites on SREBP, releasing the active form of the protein [25]. The active SREBP can then enter the nucleus and bind to the Sterol Regulating Element (SRE) in the promoter region of specific genes

involved in cholesterol synthesis; LDL endocytosis, fatty acid and phospholipid synthesis and glucose metabolism [25,27]. However, if cholesterol levels are high, SREBP is prevented from binding to SCAP, impeding gene transcription.

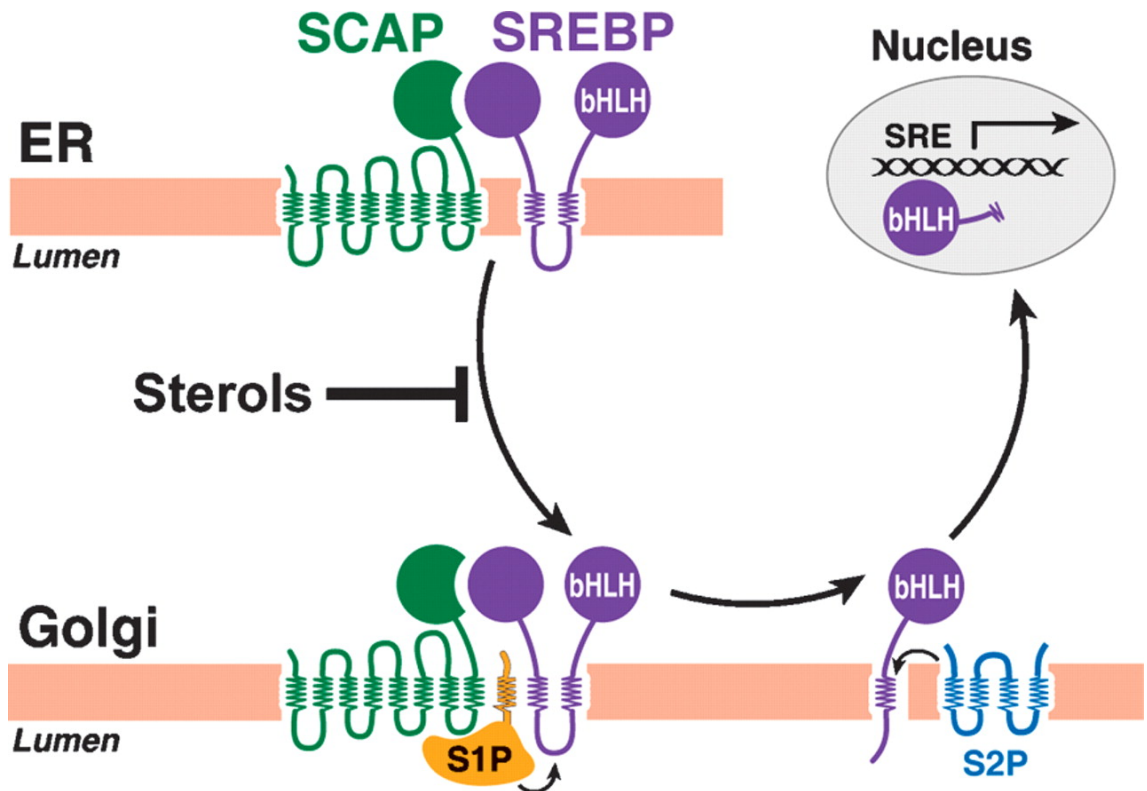


Figure 1.3: Model of the SREBP/SCAP pathway [28].

SCAP senses the presence of high sterol conditions via the SSD and SREBP is not translocated to the Golgi. Under low sterol conditions this translocation can occur. Once at the Golgi, SREBP is subjected to proteolytic cleavage by Site-1 Protease (S1P), which releases the active form, the N-terminal bHLH-Zip domain. The bHLH-Zip domain is further cleaved and released to the nucleus where it binds to the Sterol Regulating Element to enable transcription of genes involved in cholesterol synthesis.

1.1.iii: Cholesterol in development

Regulation of cholesterol is essential for correct development [29]. Morphogens are secreted signalling molecules, which determine cellular patterning via a concentration gradient [30] and an example of a family of these secreted proteins are Hedgehog proteins, which rely on cholesterol for correct function [31]. While *Drosophila* have only one hedgehog protein, mice have three; Indian hedgehog, Desert hedgehog and, the most common of the three, Sonic Hedgehog (SHH) [32]. The full-length hedgehog proteins are ~45kDa and undergo an intracellular auto-proteolytic mechanism involving cholesterol [33]. The cleavage reaction occurs in two steps; the first step produces a thioester intermediate, which can then subsequently be attacked by cholesterol [33]. This forms an ester-linkage to the C-terminal end of the ~20kDa N-terminal fragment (N-SHH), thereby releasing the ~25kDa C-terminal fragment (C-SHH) [34].

In the developing foetus, the differentiation of neural progenitor cells in the neural tube is determined by the gradients of multiple molecules [35], including SHH (**Figure 1.4**). SHH is expressed from the posterior ventral floor plate, with decreasing concentration towards the anterior end. This concentration gradient determines if the cells differentiate into one of four types of interneurons (V0 – V3) or lower motor neurons [36]. SHH is also expressed in the zone of polarizing activity (ZPA) in the posterior section of the limb bud [37]. Higher doses of SHH close to the ZPA induces the development of posterior digits whereas lower concentrations induce development of anterior digits [38]. However, in this instance, it is not only the concentration of SHH that determines the cell fate, but also length of exposure time [38]. Precise formation

of the SHH gradient is critical and therefore the diffusion of SHH involves interactions with multiple proteins to enable tight regulation. It has been shown that the attachment of cholesterol to SHH enables the correct range and shape of the gradient but is not required for signalling activity [31]. Palmitoylation of SHH at the N-terminus also is important and is subsequently added to the fragment by an acyltransferase, Skinny Hedgehog [31].

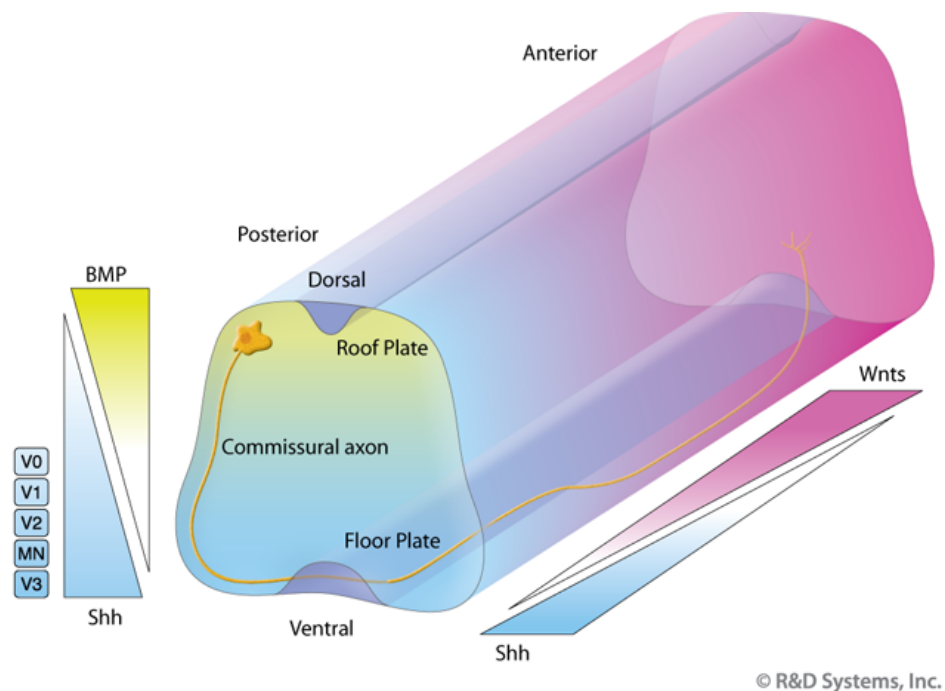


Figure 1.4: Expression of Sonic hedgehog in the neural tube [39].

Diagram showing the distribution of sonic hedgehog, BMP and WNT throughout the neural tube [39].

Three other proteins, Dispatched (DIS), Smoothed (SMO) and Patched (PTC) are also involved in the regulation of SHH [40]. DIS contains a SSD and aids the movement of the active signalling domain of SHH across the cell membrane [31,41]. This domain can then inhibit PTC, a transmembrane protein that constitutively inhibits SMO, a

second transmembrane protein [42]. However this inhibition is relieved when SHH binds to PTC, allowing signal transduction to occur [43] (**Figure 1.5**). PTC is also a transcriptional target of the SHH pathway therefore providing a feedback mechanism [31]. The interaction of SHH with PTC may occur in an area known as lipid rafts, which are lipid-rich microdomains in membranes that are thought to have the function of concentrating signalling components to a specific area [31].

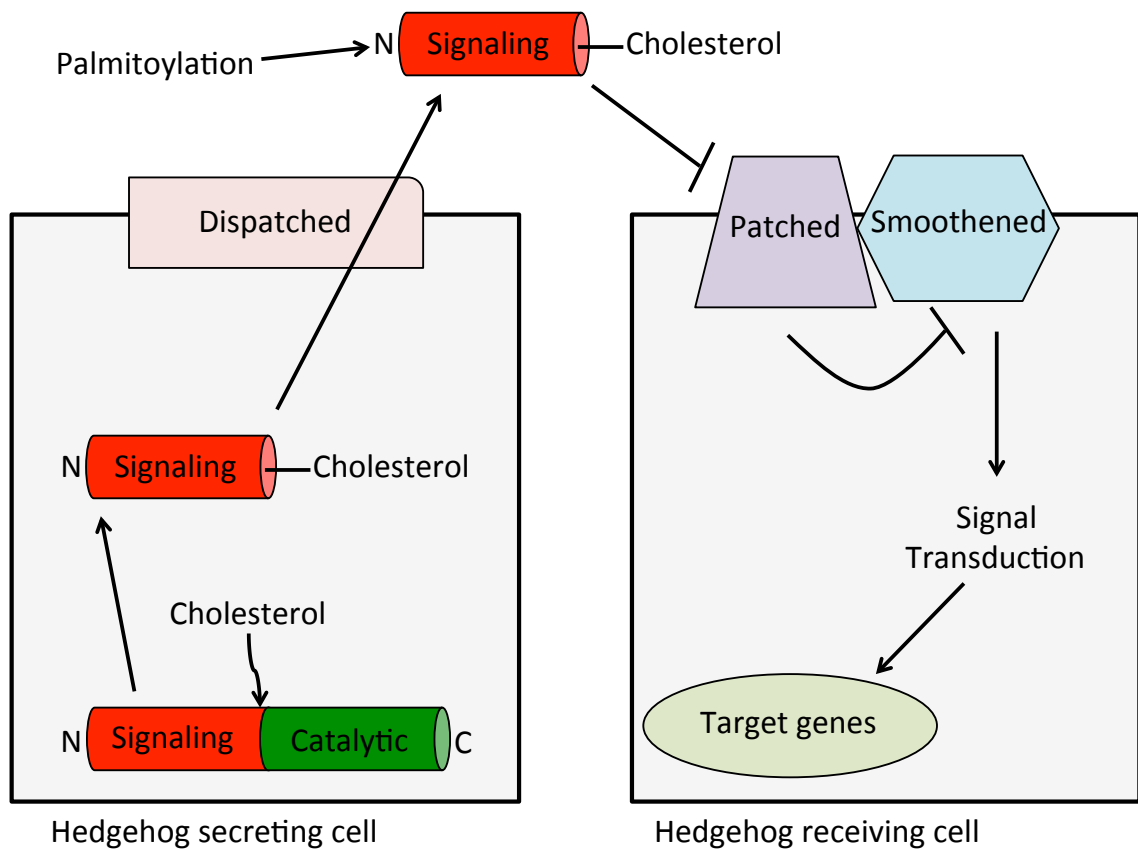


Figure 1.5: Model depicting the secretion of sonic hedgehog.

Cholesterol attacks the hedgehog protein in the hedgehog secreting cell to separate the signaling and catalytic domains. The signaling domain moves out of the cell with the help of Dispatched. Patched constitutively inhibits Smoothed, but when a palmitoylated signaling hedgehog domain is present, Patched is inhibited. This allows Smoothed to confer signal transduction to target genes in the nucleus involved in cholesterol homeostasis.

1.1.iv: Further roles of cholesterol in the cell

In addition to its role in development, cholesterol has important functions in cell membrane structure and fluidity [44], gene transcription [8], enzyme function and protein degradation [25]. Cholesterol is also particularly abundant in the central nervous system (CNS), in particular in myelin, which contains 70-80% of the total CNS cholesterol [45,46]. As cholesterol cannot pass through the blood-brain barrier (BBB), cells in the CNS rely on *de novo* synthesis [25]. Not only is cholesterol itself used in multiple ways but it also has many metabolites such as bile acids, oxysterols and steroid hormones [25].

Oxysterols are oxygenated derivatives of cholesterol and can be formed through Reactive Oxygen Species (ROS) or enzymatic reactions (via P450 enzymes) [47]. Oxysterols have important roles in cholesterol metabolism and regulation but although there is evidence to suggest that oxysterols can repress cholesterol synthesis through SCAP and HMGCR, not all data agrees with this [48]. Oxysterols can also aid the flux of cholesterol across membranes and excretion from the body [49]. For example, cholesterol cannot easily pass through membranes or the BBB but the additional hydroxyl groups in oxysterols can facilitate this transport [49].

Cholesterol exits the body largely as bile acids, which are converted from oxysterols by the family of P450 enzymes [50]. This conversion takes place mainly in the liver where the incorporation of the oxygen atom into cholesterol directs it towards degradation and excretion [49].

Cytochrome P450 enzymes contain a haem group and are named due to the characteristic absorption band at 450nm when this group binds to carbon monoxide [49]. Although a large family, certain enzymes have particular prominence in cholesterol metabolism; four in cholesterol homeostasis (CYP46A1, CYP7A1, CYP27A1 and CYP11A1) [51] and one in cholesterol biosynthesis (CYP51) [52] (**Figure 1.6**).

The four enzymes involved in homeostasis have specific functions. CYP46A1 is produced at low levels in the testis and liver, where it plays roles in signalling and regulation, but its main function is in the CNS [49]. Cholesterol cannot cross the BBB, therefore CYP46A1 converts CNS cholesterol to (24S)-hydroxycholesterol [51]. (24S)-hydroxycholesterol is membrane-permeable and can freely move across the BBB to be degraded in the liver or converted into bile acids [49]. The conversion into bile acids is split into two pathways; neutral and acidic. CYP7A1 is responsible for the catalysis of the first step of the neutral pathway, converting cholesterol to 7 α -hydroxycholesterol in the liver. The acidic pathway can also result in the production of oxysterols and the first step is the production of 27-hydroxycholesterol by CYP27A1 [51].

Unlike the other three P450 enzymes involved in cholesterol metabolism, CYP11A1 is not involved in bile acid production but is the facilitator into the steroidogenic pathway, converting cholesterol into pregnenolone, the precursor to all steroid hormones [49].

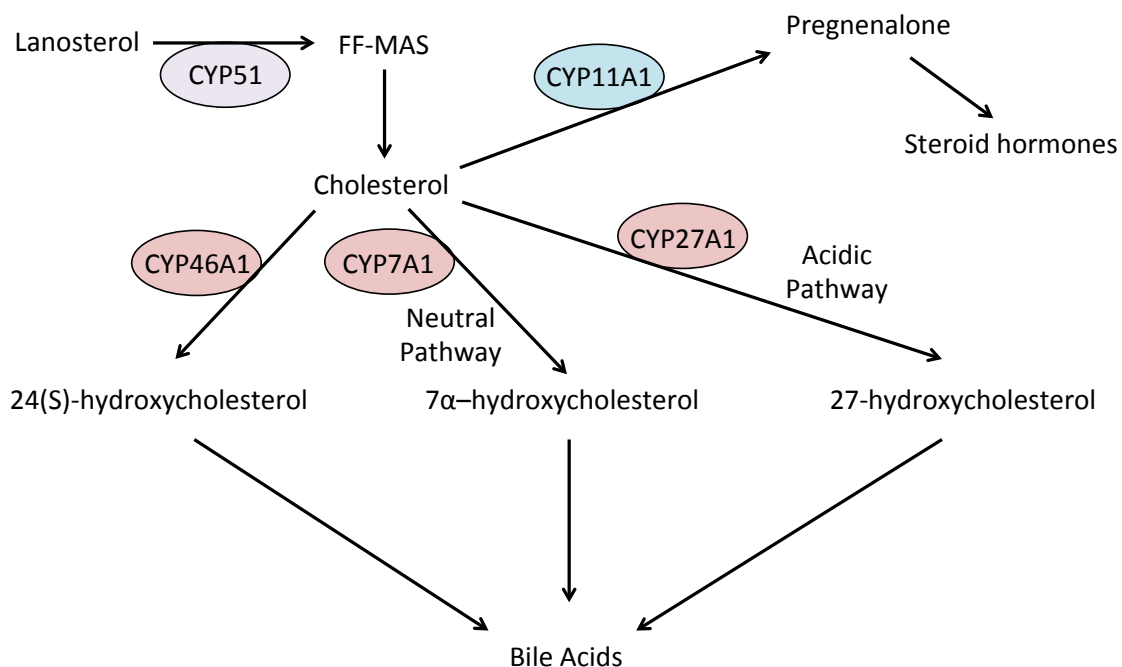


Figure 1.6: Diagram showing the 5 main P450 enzymes involved in cholesterol metabolism and synthesis.

CYP51 converts Lanosterol to FF-MAS, which is subsequently converted to cholesterol. CYP11A catalyzes the reaction of cholesterol to pregnenolone, the first step in steroid hormone synthesis. CYP46A1 converts cholesterol to 24(S)-hydroxycholesterol, which is a key step in the efflux of cholesterol from the CNS. 24(S)-hydroxycholesterol is then directed to the liver and excretion from the body in the form of Bile Acids. CYP7A1 and CYP27A1 catalyze the first step into the neutral and acidic pathways to bile acid synthesis respectively.

1.1.v: Cholesterol and disease

As cholesterol is such an important molecule, genetic errors in the cholesterologenesis pathway can be catastrophic. Prior to birth, the foetus relies on both maternal transport of molecules across the placenta and endogenous synthesis of cholesterol [5]. However, there is a limit to the maternal protection due to the inability of small molecules to pass through the placental and blood-brain barriers and therefore defects in cholesterol biosynthesis can cause severe developmental abnormalities [5]. Since the discovery in 1993 that Smith-Lemli-Opitz Syndrome is caused by a defect in sterol metabolism [53], many other disorders of sterol biosynthesis have been discovered, for example desmosterolosis, lathosterolosis and CHILD syndrome [5,54] (**Figure 1.7**). Although these disorders share some similar characteristics, each disorder has a distinct pathophysiology [55]. This led to the conclusions that it is not only the decrease of cholesterol that is important but also the production of the individual biochemical metabolites, which are produced as a consequence of each genetic defect [5].

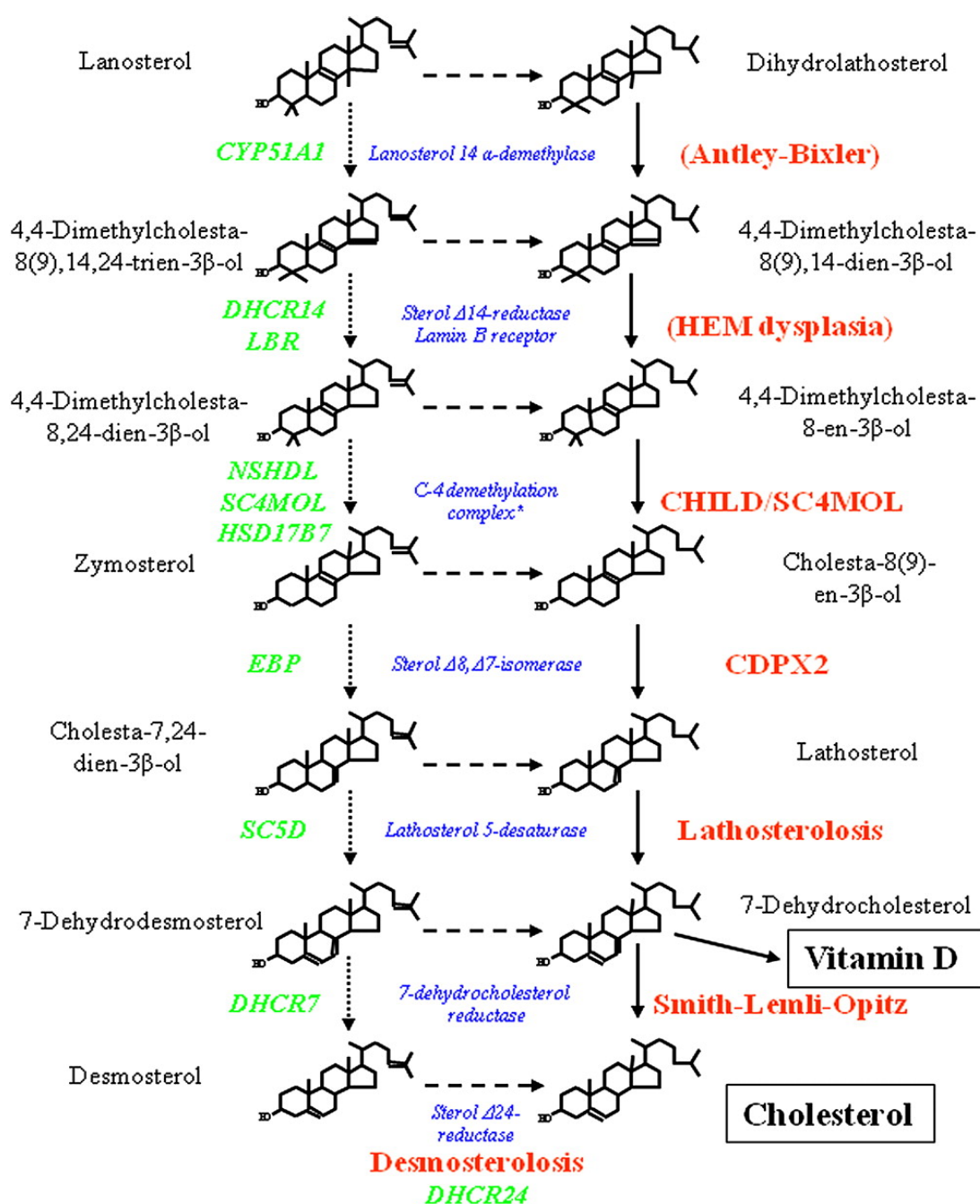


Figure 1.7: Diagram depicting the disorders corresponding to genetic deficiencies in the enzymes of the post-squalene pathway of cholesterol biosynthesis. [5]

Blue; enzymes responsible for the catalysis at each step. Red; Disorders related to the Kandutsch-Russell pathway. Green: Disorders related to the Bloch pathway.

1.2 Smith-Lemli-Opitz Syndrome

Smith-Lemli-Opitz Syndrome (SLOS) is an inborn error of cholesterol biosynthesis caused by inherited defects in 7-dehydrocholesterol reductase (DHCR7) and was first described by Drs. Smith, Lemli and Opitz in 1964 [56]. DHCR7 reduces the 7-8 double bond of 7-dehydrocholesterol (7DHC) to form cholesterol in the final step of the Kandutsch-Russell pathway of cholesterol biosynthesis [57]. Additionally, DHCR7 also reduces the 7-8 double bond of 7-dehydrodesmosterol to form desmosterol. Consequently, desmosterol is generally absent in SLOS patients [58].

1.2.i: Biochemical findings in Smith-Lemli-Opitz Syndrome

Biochemically, this enzymatic defect causes a shift in sterol levels: an increase in 7DHC and its isomer 8DHC, and a decrease in cholesterol [53]. Abnormal 7DHC levels are considered to be anything greater than 2 µg/ml in plasma [56], although levels ranging from 2.7 – 470 µg/ml (10 – 2000 fold above normal) have been found in SLOS patients [59]. Diagnosis is primarily based on identification of these abnormal sterol levels, which can be detected prenatally in amniotic fluid after the 14th week of gestation [56]. Earlier sampling in the 10th - 12th weeks is possible from chorionic villus sample, but as this method increases risk of abortion, it is not typically used at this stage [59].

It is widely agreed that the altered sterol levels are hallmarks of SLOS, but it is unclear if it is the elevated 7DHC levels, decreased sterol levels or a mixture of the two, that causes the pathogenic defects [60]. As mentioned previously, lipid rafts are important

components of cell membranes that have roles in signal transduction. These domains are rich in lipids, including cholesterol, and alteration in the composition could affect the signalling properties. Keller *et al* showed that 7DHC is able to incorporate into lipid rafts *in vitro* with the consequence of altering the protein composition [61]. Changes in signalling have also been observed in a mouse model of SLOS where the decreased stability of lipid rafts caused by the sterol change led to increased degranulation of mast cells and an enhanced cytokine response [62]. Cholesterol is not just important in lipid rafts but in cell membranes as a whole and it is known that incorporation of 7DHC into cell membranes alters properties such as organisation and dynamics [63].

The alteration of cell composition could explain why vitamin D levels are not significantly increased in SLOS, which is surprising as 7DHC is the precursor to vitamin D [64]. Cell membranes play a role in stabilizing pre-vitamin D into vitamin D. The change in the lipid composition could affect the availability of biosynthesis-enhancing sites in cell membranes therefore leading to a decreased or unchanged level of vitamin D produced in SLOS patients [64]. Another suggestion is that as vitamin D biosynthesis involves exposure to sunlight, the photosensitivity of many SLOS patients prevents this from occurring [64].

1.2.ii: Smith-Lemli-Opitz Syndrome phenotype

DHCR7 is not always completely inactive as a result of different mutations, which causes a broad phenotypic severity spectrum [65,66]. Where there is little or no residual enzyme activity, severe malformations are present and the patients die *in utero* or very soon after birth [66]. The mutations found in *DHCR7*, and the resulting carrier frequency and incidence of SLOS, will be discussed further in **Chapter 2**. As the amount of residual enzyme activity increases, the phenotype becomes milder and patients are able to live into adulthood with only minor physical anomalies (**Figure 1.8A**). Due to the variation in severity, physical anomalies tend to be variable but easily recognisable [56]. The most common physical feature is 2,3-toe syndactyly. Other characteristic physical features include postaxial polydactyly of the hands or feet and facial features such as microcephaly, a short upturned nose, ptosis, epicanthus and retrognathia [60] (**Figure 1.8B**). The CNS also shows malformations including septum pellucidum abnormalities, agenesis of the corpus callosum and holoprosencephaly [67,68]. Other phenotypic manifestations include photosensitivity and sleep-cycle disturbance [60].

Patients also have a characteristic behavioural phenotype including mild-moderate intellectual disability, autism-spectrum disorder traits, anxiety, hyper-activity, and increased aggression and self-injury related to the menstrual cycle [69]. These will be discussed further in **Chapter 3**.

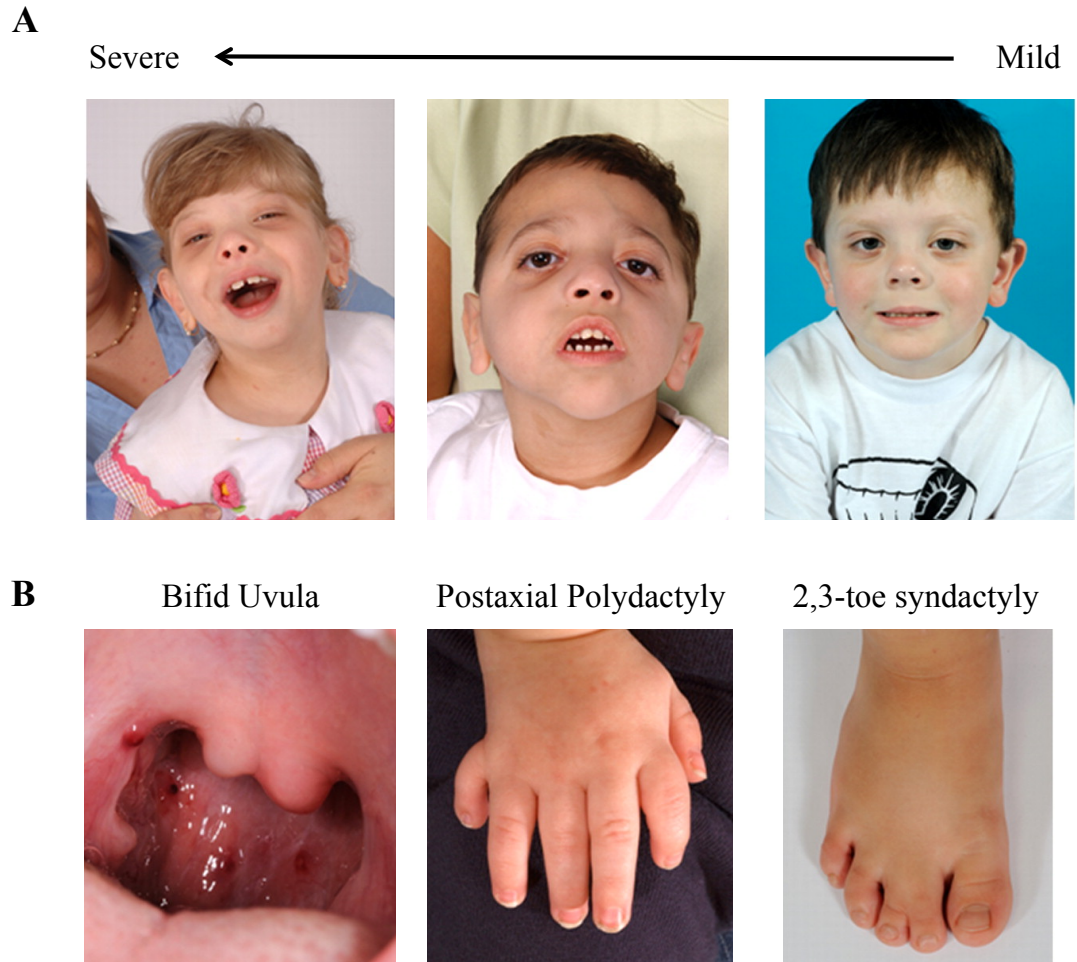


Figure 1.8: Physical characteristics of SLOS [5]

- A) Profile images of the phenotypic range in SLOS patients; severe (left image), classical (middle image) and mild (right image).
- B) Examples of the most common physical malformations; bifid uvula (left image), postaxial polydactyly (middle image), 2,3-toe syndactyly (right image)

1.2.iii: Altered development in Smith-Lemli-Opitz Syndrome

It is known that embryonic development relies heavily on cholesterol, for example, in the hedgehog-signalling pathway [29], but currently it is not known how the increased levels of 7DHC affect development. The majority of cholesterol is synthesised by the developing foetus, but maternal-foetal cholesterol transport also occurs through the yolk sac and placenta during development [70]. Although cholesterol cannot pass the BBB, the maternally derived cholesterol may aid the development of an SLOS foetus [71].

As discussed earlier in the chapter, cholesterol is important for the correct functioning of SHH during development. Two synthetic inhibitors of DHCR7, AY9944 [72] and BM15766 [73,74], can cause similar defects in the SHH pathway to mutations in *DHCR7*. Additionally, cholesterol was synthetically lowered in chick embryos using cyclodextrin, resulting in a phenotype similar to that seen in a mouse embryo homozygous for mutations in SHH [75]. Interestingly it was observed that in mouse embryonic fibroblasts, using synthetic inhibitors was more effecting than only depleting cholesterol. This could be because autoprocessing of SHH can still be achieved via addition of 7DHC, but the lack of cholesterol may affect the targeting via SSD [76].

Hormones are also critical in development and the adrenal glands produce three main types of steroid hormones: mineralocorticoids, glucocorticoids and androgens. As cholesterol is a precursor to these hormones, adrenal insufficiency has been studied in patients with SLOS, but there are some discrepancies between reports. Andersson *et al* reported a case of adrenal insufficiency [77] whereas Bianconi *et al* did not find evidence of insufficiency, although it was noted that there was an increase in baseline

adrenocorticotrophic hormone (ACTH) values. These elevated baseline levels were suggested to be a consequence of the alteration of cell membranes due to the biochemical defect instead of adrenal insufficiency. [78]. The inconsistencies in reports could be due to the severities of patients studied. The patient with adrenal insufficiency died at a young age and therefore had a severe phenotype [77] whereas the patients studied in the Bianconi study were of mild-moderate severity [78], and were on cholesterol supplementation, which could have corrected any insufficiency present [78]. Additionally, DHC-derivatives of hormones, such as pregnenalone and allopregnenalone, have been found in SLOS patients [79], although the biological consequence of this is yet to be fully determined [56].

1.2.iv: Altered biosynthesis and metabolic pathways in Smith-Lemli-Opitz Syndrome

7DHC is one of the most susceptible molecules to free-radical oxidation and DHC-derived oxysterols, and the derived bile acids, have also been detected in SLOS patients [80]. This will be discussed further in **Chapter 4**. As cholesterol levels are lower than usual in SLOS patients, it could be expected that bile acid production would also be lower. However, this is not seen in all patients, generally only being observed in severe cases [59]. CYP7A1 and CYP27A1 are important enzymes in the synthesis of bile acids and it has been shown that both 7- and 8-DHC are competitive inhibitors of these enzymes, causing 7- and 8-DHC bile acid derivatives to be formed [81].

Cell culture can also reveal insights into the mechanisms involved in SLOS. In control fibroblasts, cholesterol is transported to lysosomes through LDL binding to LDL receptors. Cholesterol esters are released from LDL in the lysosomes and LDL apolipoproteins are degraded and released [82]. Fibroblast cell lines from SLOS patients can be grown in lipo-deficient serum (LPDS) to force cholesterol synthesis and therefore the accumulation of 7DHC [82]; biochemical analysis showed that 7DHC in SLOS fibroblasts increased from 2.5 µg/ml at 5 days in LPDS to over 5 µg/ml at 7 days in LPDS [82].

After 4 days in LPDS, SLOS fibroblasts showed an increase in filipin staining, a polyene auto-fluorescing antibiotic that binds to unesterified sterols, compared to controls [82] (**Figure 1.9**). Increased 7DHC levels have been shown to cause abnormal effects in membrane properties, but if the increased filipin staining was due to 7DHC accumulation, it would be expected that the staining would continue to increase over time. However, the opposite occurs and after 7 days the filipin staining is back to control levels [82]. This suggests that it is unesterified cholesterol that accumulates. Using electron microscopy it can be seen that these lysosomal inclusion bodies are similar to those seen in Niemann-Pick type C disease (NPC) and it can be hypothesised that NPC1 may be inhibited in SLOS due to dysregulation in the SSD of the protein [82]. Additionally, SLOS cells showed decreased LDL degradation, possibly due to the accumulation of 7DHC.

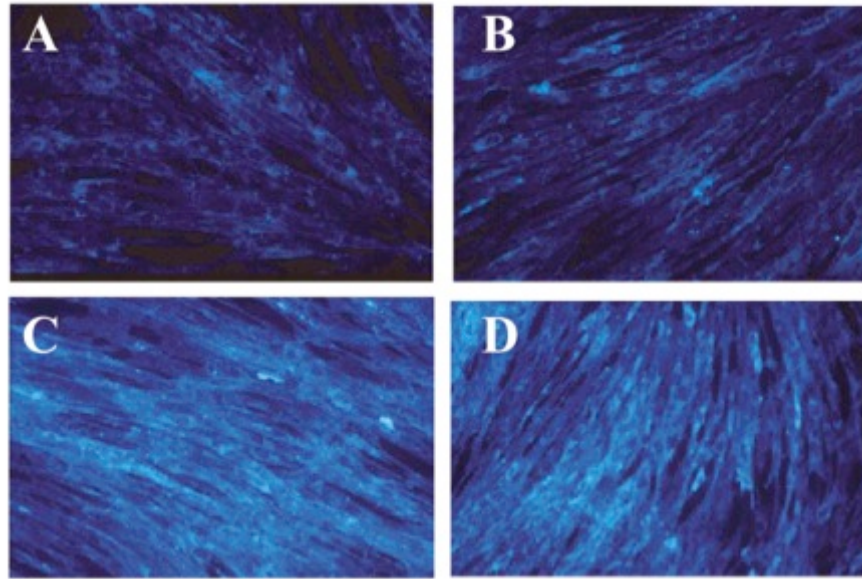


Figure 1.9: Filipin staining of fibroblasts [82].

Fibroblasts were grown in LPDS for four days and stained with filipin. The filipin staining is increased in the SLOS cells compared to control and heterozygotes.

- A) Control fibroblasts
- B) Heterozygous fibroblasts
- C) A2SLOS fibroblasts
- D) B3SLOS fibroblasts

1.3 Niemann-Pick disease type C disease

A second disorder, Niemann-Pick type C (NPC), does not have defects in cholesterologenesis but in the transport of cholesterol. Albert Niemann first described NPC in 1914, when he identified a young child who had an enlarged liver and spleen, an impaired nervous system and died by 3 years of age [83]. Ludwig Pick continued this work by studying tissues of deceased affected infants and in the late 1920s, the existence of a new disease, Niemann-Pick Disease (NP) was confirmed [84].

Although NP disease is characterised by the accumulation of lipids, in 1961 Crocker and Farber classified the disease into four sub-types according to age of onset and organs affected; Niemann-Pick type A (NPA), Niemann-Pick type B (NPB), Niemann-Pick type C (NPC) and Niemann-Pick type D (NPD) [85]. These classifications were further refined by the discovery that both NPA and NPB, but not NPC or NPD, had a deficiency in acid sphingomyelinase activity [86]. On the other hand, NPC and NPD were characterised by defects in the transport of LDL, leading to accumulation of un-esterified (free) cholesterol. NPD is derived from patients with Nova-Scotia ancestry but allelic to type C therefore the two sub-types were combined to NPC [84,87].

Free cholesterol is produced when LDL is endocytosed and hydrolysed in the late endosomes/lysosomes. In healthy cells, cholesterol is then transported out of the endosomes to the plasma membrane and endoplasmic reticulum (ER). However, this efflux is impaired in NPC cells causing cholesterol to accumulate in the lysosomes [84]. Cholesterol cannot cross the BBB but is synthesized by oligodendroglial cells, astrocytes and glia [88]. Therefore the pathogenesis of brain defects is more complex.

1.3.i: Structures of NPC1 and NPC2

Cholesterol movement from the lysosome to ER requires two proteins, NPC1 and NPC2. Mutations in *NPC1* are responsible for approximately 95% of clinical cases, while the remaining 5% of cases are caused by mutations in *NPC2* [89]. The mutations and incidence of NPC will be discussed further in **Chapter 2**. The same phenotype can be observed from homozygous loss-of-function mutations of *NPC1* or *NPC2*, which suggests that both proteins have important roles in the egress of cholesterol from the lysosomes [90]. One model of how cholesterol is passed between NPC1 and NPC2 is shown in **Figure 1.10**.

NPC1 is a large transmembrane protein and comprises 1252 amino acids, 13 transmembrane domains [91] and several functional domains [92,93] (**Figure 1.11A**):

- A cleaved N-terminal signal sequence, which targets the protein for insertion into the membranes (amino acids 1 – 24)
- Luminal loop-1 (amino acids 25 – 264)
- Luminal loop-2 (amino acids 371 – 615)
- A SSD in transmembrane helices 3 – 7, with homology to the SSDs in HMGR, SCAP and PTC. (amino acids 616 – 791)
- Luminal loop-3 (amino acids 855 – 1098)
- A C-terminal sequence which targets the protein to the lysosomes (amino acids 1225 – 1278)

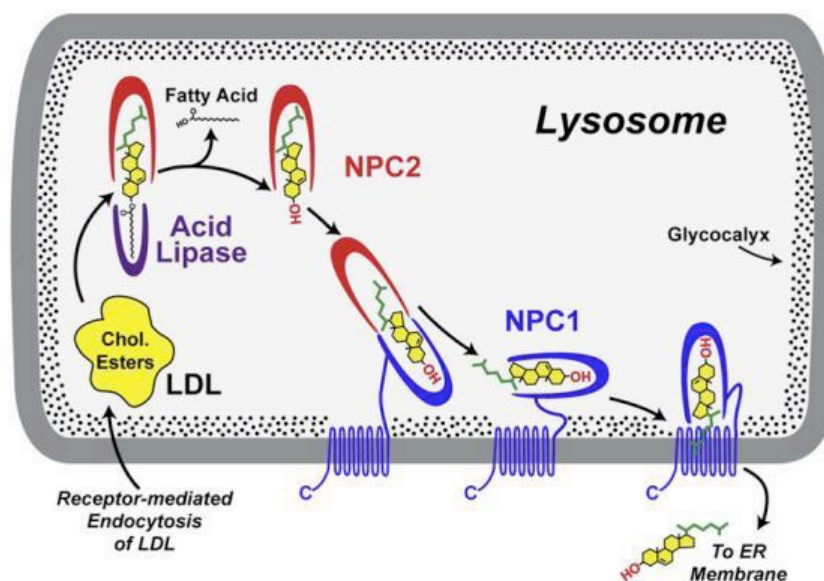


Figure 1.10: Proposed model for the role of NPC1 and NPC2 in de-esterification and trafficking cycle of cholesterol in lysosomes (modified from [94]).

In the model proposed above, NPC2 binds to cholesterol and holds it while acid lipase cleaves lipids bound to the 3-beta hydroxyl of cholesterol. NPC2 then transfers the free cholesterol to the N-terminal domain of NPC1. However, the directionality of cholesterol transfer between these two proteins has not been determined experimentally. The cholesterol is then transported out of the lysosomes to the ER.

The sterol binding domain of NPC1 was discovered when Infante *et al.* were searching for a membrane-bound protein that bound oxysterols [92]. They found a sterol binding site on the luminal loop-1 of human NPC1, which could also bind cholesterol but only in low detergent conditions [92]. Recently, the majority of the full-length protein has been crystallised, incorporating 12 of the 13 transmembrane domains and a large extension, comprising the luminal loops and C-terminal domain [93] (**Figure 1.11B**). This study found that the transmembrane domains 3 – 5 define a large hydrophobic cavity, large enough for 1 molecule of cholesterol, which is accessible from both the

lipid bilayer and lysosome lumen. Although the exact role for the SSD of NPC1 has not been fully elucidated [91], the accessibility of the SSD suggests that this allows it to sense the cholesterol concentration of the lysosome lipid bilayer [93].

In contrast, the NPC2 protein is a small soluble protein of 132 amino acids and is transported to the lysosomes via the mannose-6-phosphate receptor. NPC2 binds cholesterol and likely transfers it from inner lysosomal vesicles to the limiting membrane of the lysosome [95]. The crystal structure of NPC2 has been evaluated, both bound to cholesterol-3-O-sulphate [96] (**Figure 1.11C**) and in its unbound form [96,97]. NPC2 is comprised almost exclusively from β -strands [97] with a deep hydrophobic binding pocket. This pocket is too small for cholesterol in the unbound state, but expands to incorporate the hydrocarbon section of cholesterol [96].

Cholesterol-3-O-sulphate was used in this study as NPC2 had a higher affinity for this substrate over cholesterol. Cholesterol-3-O-sulphate binds to NPC2 with the isooctyl chain buried deep in the pocket and the sulphate group positioned at the mouth of the tunnel. As the sulphate moiety does not interact with the NPC2 protein structure, the reason for the preference for NPC2 to bind this substrate over cholesterol is uncertain [96]. A suggestion is that cholesterol-3-O-sulphate has greater solubility and consequently, more substrate is available to the protein. However, it is presumed cholesterol binds in an analogous manner [96].

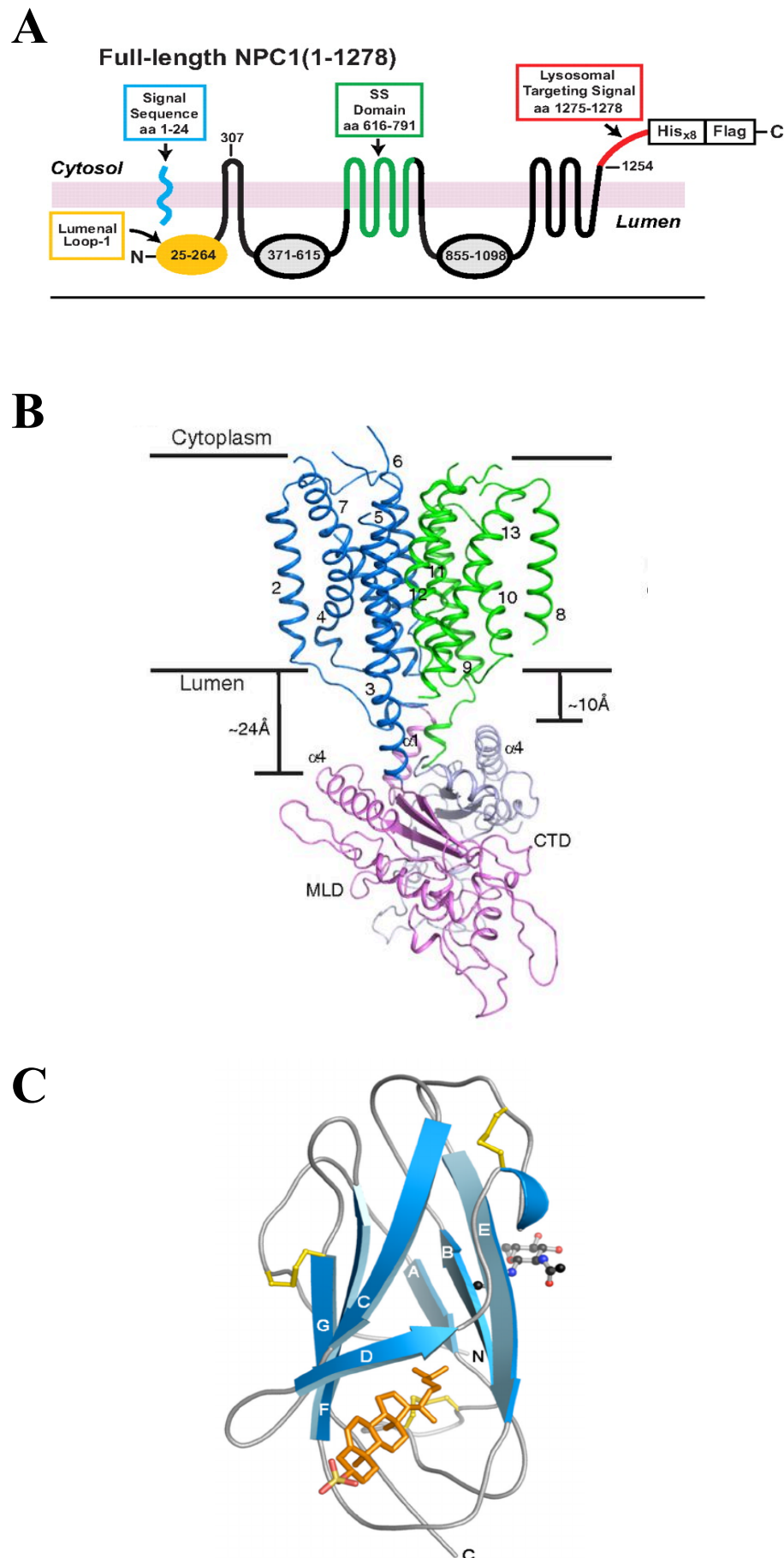


Figure 1.11: Schematic of NPC1 and NPC2 protein structures

Figure 1.11 (previous page): Schematic of NPC1 and NPC2 protein structures**NPC1**

A) Predicted topology of the full-length NPC1 protein [modified from 86]

B) Overall structure of NPC1 showing the N-terminal transmembrane domains (2-7, blue), the C-terminal transmembrane domains (8-13, green), the middle luminal domain (MLD, magenta) and C-terminal domain (CTD, light blue). Numbered helices of the luminal domains are involved in interactions and the distances of MLD (luminal loop-2) and CTD from the lysosomal membrane are marked. [modified from 87]

NPC2 [modified from 89]

C) A ribbon representation (blue) of the 7 β -strand fold NPC2 structure, stabilised by three disulphide bonds (yellow). Cholesterol-3-O-sulphate is bound (sulphur: yellow, oxygen: red, carbon: gold) between strands D and E. The bovine form of NPC2 is glycosylated, which were cleaved by EndoH_f for crystallisation. This leaves the innermost *N*-acetylglucosamine (GlcNAc) sugar attached, shown by the stick and ball structure at Asn-39 (carbon: black, oxygen: red, nitrogen: blue).

1.3.iii: Niemann-Pick type C disease phenotype and diagnosis

Onset of NPC can range from the perinatal period to well into adulthood. This is a continuum but is often subdivided into 5 stages; perinatal, early infantile (2 months – 2 years), late infantile (2 – 6 years), juvenile (6 – 15 years) and adult (>15 years)

(Figure 1.12). Correspondingly, patient lifespan varies from a few days to over 60 years old, although the majority die between 10 and 25 years of age [84].

Each stage is split into two forms; systemic onset and neurological onset. The exception to this is during the perinatal period where only the systemic symptoms are presented. However it has been reported that patients who die from a severe perinatal form have siblings with an early onset (infantile or juvenile) neurological form [84]. Overall the most common symptoms are hepatosplenomegaly, vertical supranuclear gaze palsy (VSGP), cataplexy and neurological degeneration, which includes cerebellar ataxia, dysphagia and dysarthria [84]. Systemic disease generally precedes onset of neurological symptoms but it is the progressive degeneration that causes fatality in most cases.

The varied age of onset and clinical phenotype of NPC can complicate diagnosis [98]. When classical symptoms, such as hepatosplenomegaly, ataxia and supranuclear gaze palsy, are present, suspecting a diagnosis of NPC is relatively easy. However, although peripheral symptoms are generally presented before the neurological symptoms, where this is absent a delay in diagnosis can occur [84].

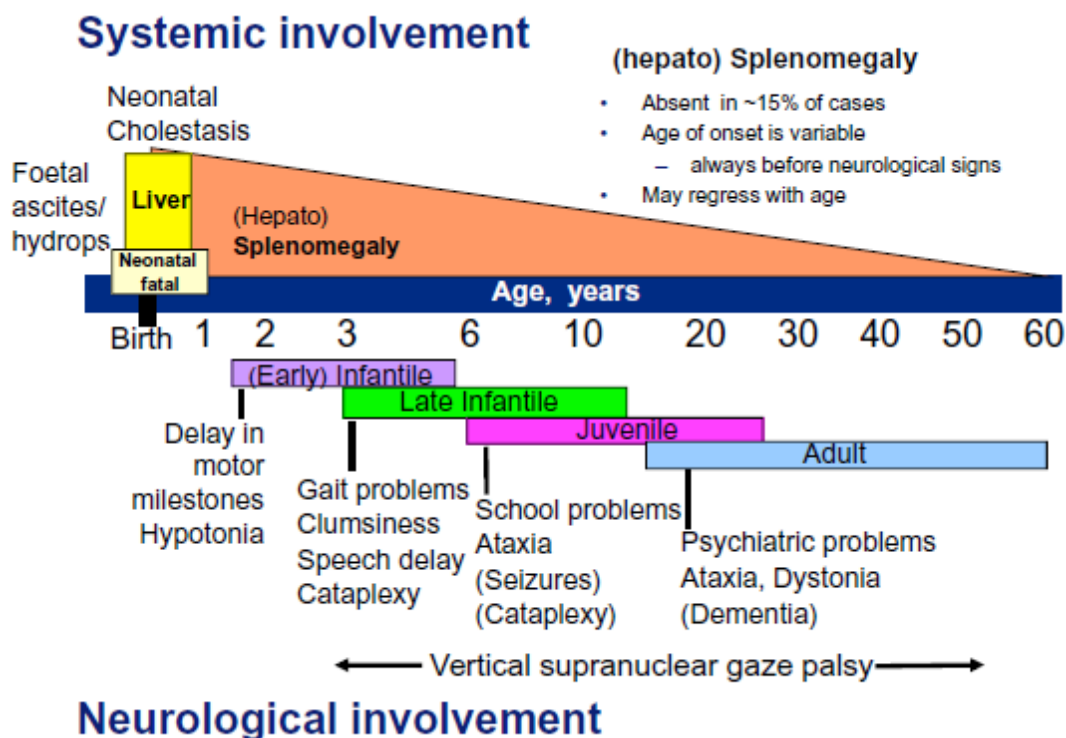


Figure 1.12: Age of onset of symptoms: [84]

Diagram showing the age of onset of NPC1, along with the age of presentation of first neurological symptoms. NPC1 has a complex pathological cascade including neuroinflammation, oxidative stress, and neuronal apoptosis, all of which contribute to the varied phenotype presented. Symptoms include progressive neurodegeneration, hepatosplenomegaly, ataxia, dysphagia, dysarthria, tremor and Vertical Supranuclear Gaze Palsy.

Diagnosis can be through behaviour or biochemical identification. Behavioural diagnosis relies on assessing muscle tone, movement, reflexes, hearing and swallowing [98]. Biochemical identification of the cholesterol storage and impaired trafficking is currently the most reliable form of diagnosis. Skin fibroblast samples are biopsied from suspected patients and are stained using filipin, an autofluorescing antibiotic that binds to unesterified cholesterol [100]. Examination of the cells using fluorescence microscopy reveals that NPC patients typically show strongly fluorescent vesicles in approximately 80% of cases. In the remaining cases, less staining is present and indicates a variant phenotype [101] (**Figure 1.13**). For diagnosis of these cases mutation analysis is often implemented [84]. Currently, there is an average diagnostic delay of around 5 years [102] but the development of biomarkers specific for NPC1 that are less invasive, such as oxysterols [5] and bile acids [103], will facilitate faster diagnosis in the future.

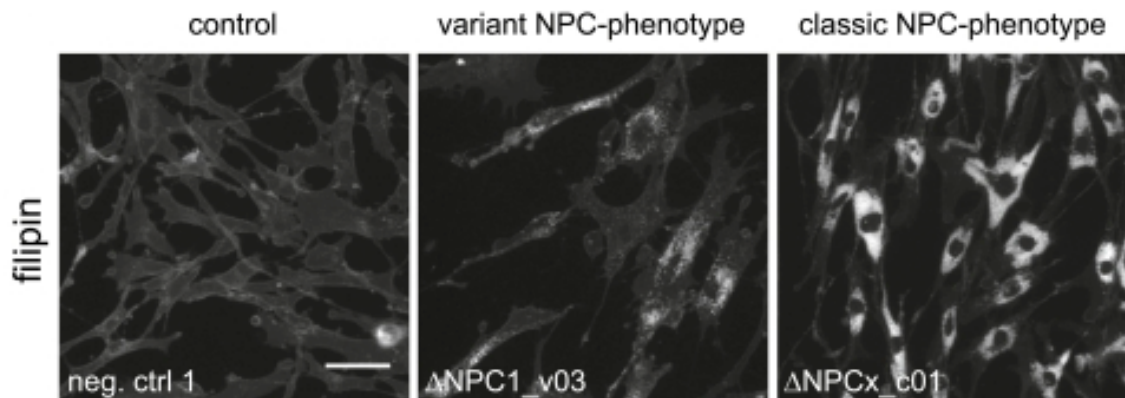


Figure 1.13: Filipin staining of NPC fibroblasts [99].

The left image shows control fibroblasts with very little staining. The right image shows the classical NPC phenotype with strong peri-nuclear staining indicating accumulation of unesterified cholesterol. The middle image shows the variant phenotype with less staining.

1.3.iv: Altered metabolism in Niemann-Pick type C disease

It has been shown that, although an important biomarker, free cholesterol accumulation may not be the primary offending metabolite in NPC. Glycosphingolipids (GSLs), sphingosine, sphingomyelin and phospholipids are also among the accumulating metabolites but the levels of each metabolite vary in individual organs [84]. High levels of glycosphingolipids accumulate in the brain, while sphingomyelin and cholesterol accumulate to a lesser extent, and there is more pronounced metabolite storage in the spleen compared to the liver [84].

A feature of other primary lysosomal sphingolipid storage disorders is abnormal endoplasmic reticulum (ER) calcium homeostasis [104]. Studies in NPC1 mutant cells show abnormal calcium homeostasis, with lower levels of calcium in the acidic compartment. In contrast, ER and mitochondrial calcium levels were normal, suggesting that the trafficking defect occurs before lipid storage in NPC1 pathogenesis [104]. This proposed the hypothesis that the defective trafficking and fusion in the endocytic pathway is caused by insufficient calcium release. Additionally, studies have shown that sphingosine rapidly accumulates when the NPC1 protein is inactivated, and that the low calcium concentrations in the acidic compartment can be caused by sphingosine [104]. Sphingosine can therefore be considered a candidate for being the primary offending metabolite in NPC and indeed NPC1 may function as a sphingosine transporter.

1.3.iv: Neurological defects in Niemann-Pick type C disease

Some areas of the brain are affected more than others in NPC disease, for example the cerebral neocortex and hippocampus are the least severely affected whereas the corpus callosum and cerebellum are the most affected [105]. The major neuronal cells in the cerebellum are Purkinje cells (PCs) and it was observed that these cells have a distinctive degeneration pattern in NPC1 murine models [106]. Up to P40 there was no difference between the numbers of PCs in NPC1 versus wild-type mice. However, at P60 there was a dramatic decrease in the number of PCs in the NPC1 brains [107]. The cerebellum is organised into 10 lobes and degeneration starts at lobe I, continuing through to lobe IX, although limited loss occurs in lobe IX [106,107] (**Figure 1.14**). This suggests that the PCs in lobes IX and X appear to have some protection and do not degenerate to the same extent, though the mechanism of why this occurs is not known.

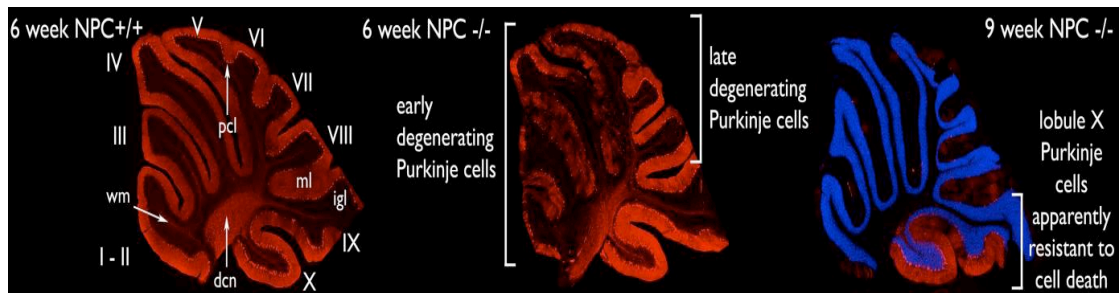


Figure 14: Degeneration of Purkinje Cells in the cerebellum (unpublished data; Dr. I. Williams, Platt Laboratory, 2010).

These immunofluorescence images depict the sagittal view of the cerebellum, staining Calbindin (CaBP) (red). Blue DAPI staining on the 9-week mutant shows the lobular structure. The anterior lobes degenerated first, with subsequent loss in the central and posterior lobes. Lobe IX showed little PC degeneration and lobe X showed no degeneration at all. Therefore the extent of PC loss correlates with disease progression.

1.4 Animal Models

Although patient case studies and cell culture are valuable methods of gaining information about the disorders, animal models allow the biochemical aspects to be further investigated. There are multiple animal models for both SLOS (**Table 1.1**) and NPC1 (**Table 1.2**).

1.4.i Smith-Lemli-Opitz Syndrome animal models

Animal	Model	Phenotype
Rat	AY9944, teratogenic model	Low plasma cholesterol, high plasma 7DHC, holoprosencephaly
Rat	BM15766, teratogenic model	Low plasma cholesterol, high plasma 7DHC, holoprosencephaly
Mouse	<i>Dhcr7</i> ^{Δ3-5/Δ3-5} genetic model	Low tissue/serum cholesterol, high tissue/serum 7DHC, craniofacial abnormalities, death by P1
Mouse	<i>Dhcr7</i> ^{T93M/Δ3-5} genetic model	Low tissue cholesterol, high tissue 7DHC, sterol levels normalise over time but remain different to normal, 2,3-toe syndactyly, neurological structure abnormalities

Table 1.1: Summary table of SLOS animal models

Rat models of SLOS have been created using chemical inhibitors of DHCR7, AY9944 (trans-1,4-bis(2-chlorobenzyl-amino-methyl)cyclohexane dihydrochloride) and BM15766. When these compounds are fed to pregnant rats at the beginning of gestation, the pups present an SLOS phenotype, with lower plasma cholesterol, increased plasma 7DHC and holoprosencephaly [59]. Although these compounds can be used in mice, a 10-20 times higher dose is required [29]. Therefore genetic-knockout murine models are preferable.

Two murine models exist for SLOS. One model is homozygous for a null mutation (*Dhcr7*^{Δ3-5/Δ3-5}). This model was developed by targeted homologous recombination in mouse embryonic stem cells to disrupt the *Dhcr7* gene [108]. This disruption involved insertion of a neomycin phosphotransferase gene and deletion of coding exons III, IV and part of V (Δ3-5). The heterozygous mice (*Dhcr7*^{+/Δ3-5}) appear phenotypically normal and so crossing heterozygous mice produces the null homozygous mice [108]. Biochemically, *Dhcr7*^{Δ3-5/Δ3-5} newborn pups have decreased cholesterol and increased 7DHC in tissues and serum compared to age-matched heterozygous or control pups. 7-dehydrodesmosterol was also reduced in the cortex and mid-brain of mutant pups. Phenotypically, the mutant pups were smaller in size (**Figure 1.15**), had craniofacial malformations, decreased movement and failed to feed, resulting in death soon after birth (P1). Cleft palate was also a rare finding [108].

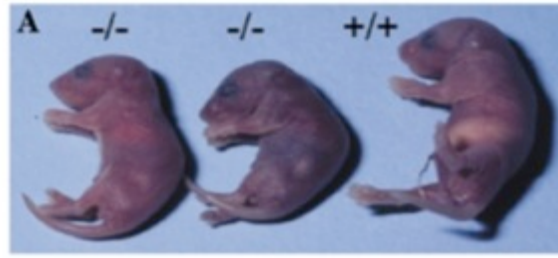


Figure 1.15: Characterisation of $Dhcr7^{\Delta3-5/\Delta3-5}$ mice [108].

New-born $Dhcr7^{\Delta3-5/\Delta3-5}$ mice (two on the left) were significantly smaller than age-matched wild type mice (right). The middle pup lacked nasal openings and cyanotic episodes.

Although tissues from these mice can be used for analysis, they are not suited for extended studies and so a milder model was developed. A commonly seen mutation in SLOS patients is the T93M point mutation, which generally relates to a classical clinical presentation [109]. Site-directed mutagenesis was implemented to create a dinucleotide change at codon 89 (T $\rightarrow$ M) in the mouse, which relates to codon 93 in humans. These $Dhcr7^{T93M/+}$ mice are phenotypically normal, viable and fertile [110]. Therefore these mice were inter-crossed (to establish $Dhcr7^{T93M/T93M}$ mice) and mated with $Dhcr7^{+/\Delta3-5}$ mice (to establish $Dhcr7^{T93M/\Delta3-5}$ mice). As $Dhcr7^{T93M/T93M}$ are also phenotypically normal, viable and fertile, these mice were also mated with $Dhcr7^{+/\Delta3-5}$ to achieve the $Dhcr7^{T93M/\Delta3-5}$ genotype. However, approximately a quarter of the expected $Dhcr7^{T93M/\Delta3-5}$ mice appeared to be lost early in embryonic development, possibly due to maternal placental abnormalities caused by the T93M homozygous mutation [110].

Although smaller in size (**Figure 1.16A**), survival of $Dhcr7^{T93M/\Delta3-5}$ mice was comparable to control mice [110]. Approximately two thirds of $Dhcr7^{T93M/\Delta3-5}$ mice are

born with 2-3 toe syndactyly (**Figure 1.16B**) and ventricular dilatation has been noted in one third of mice as early as three months of age (**Figure 1.16C**). Partial agenesis of the corpus callosum was also a rare finding. Biochemically, CNS and peripheral DHC levels in one day-old *Dhcr7*^{T93M/ Δ 3-5} are higher than both *Dhcr7*^{+/+} and *Dhcr7*^{T93M/T93M} and are found at similar levels to *Dhcr7* ^{Δ 3-5/ Δ 3-5} mice [110]. Cholesterol levels are also decreased. However, by 6 weeks of age kidney and liver DHC levels in *Dhcr7*^{T93M/ Δ 3-5} mice decreased, although CNS levels remained elevated. Cholesterol levels had also normalised in all tissues by this time [110]. As residual cholesterol synthesis in the mice was similar to that seen in genetically equivalent human fibroblasts, it was hypothesised that this could be explained in part by the mice having a higher *Dhcr7* expression and a higher level of cholesterol synthesis (50 mg/kg/day compared to 10 mg/kg/day) [110].

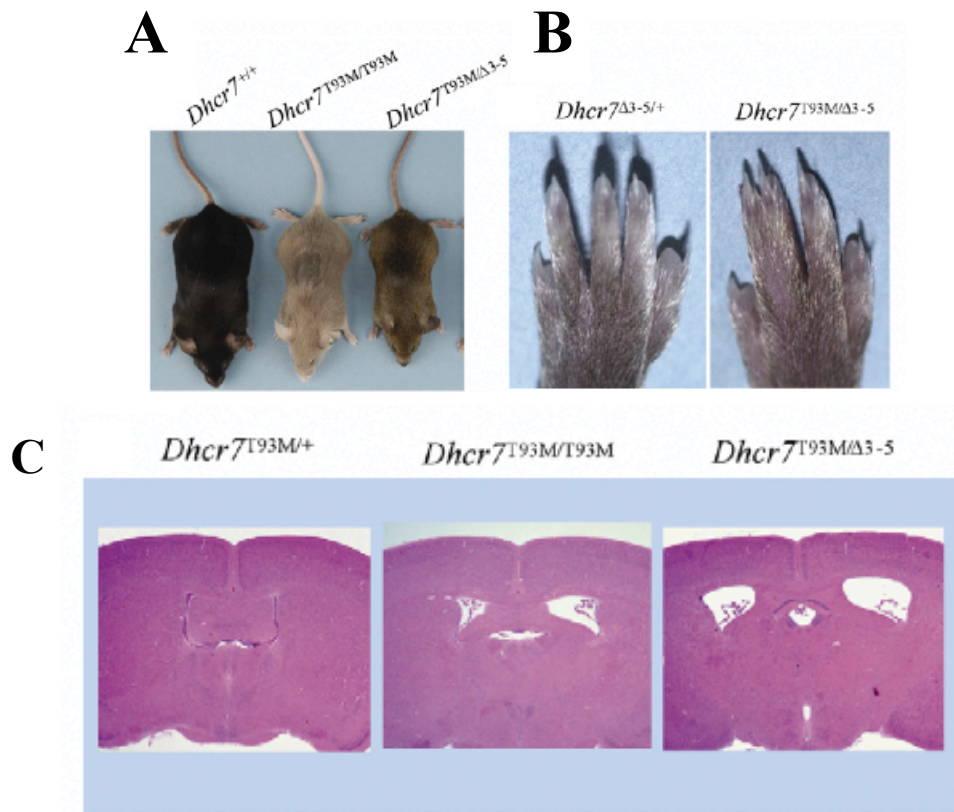


Figure 1.16: Characterisation of the *Dhcr7*^{T93M/Δ3-5} mice (modified from [110]).

A) *Dhcr7*^{T93M/T93M} mice (middle) are smaller than *Dhcr7*^{+/+} mice (left). *Dhcr7*^{T93M/Δ3-5} mice (right) are smaller than both other genotypes. **B)** Image showing 2,3-toe syndactyly in *Dhcr7*^{T93M/Δ3-5} mice (right) compared to a normal paw in *Dhcr7*^{+/Δ3-5} mice. **C)** Histochemical analysis of the cortex of *Dhcr7*^{T93M/+} (left), *Dhcr7*^{T93M/T93M} (middle) and *Dhcr7*^{T93M/Δ3-5} (right) mice. Ventricular dilatation can be seen to increase from the left to right images.

1.4.ii Niemann-Pick type C disease animal models

Animal	Model	NPC1 or NPC2?	Phenotype
Mouse	BALB/c- <i>npc1</i> ^{nh} Spontaneous mutation	NPC1	Severe infantile form, develops neurological symptoms at 6 weeks, death occurs 10-12 weeks
Mouse	C57BL/KsJ- <i>npc1</i> ^{spm} Spontaneous mutation	NPC1	Severe infantile form, develops neurological symptoms at 7 weeks, death occurs 12-14 weeks
Mouse	C57BL/6J <i>npc</i> ^{nmf164} Ethyl- nitrosourea mutation	NPC1	Juvenile form, develops neurological symptoms at ~ 11 weeks, death occurs 16-17 weeks
Mouse	Genetic mutation	NPC2	Develops neurological symptoms at 7-8 weeks, death occurs 13-18 weeks
Feline	Spontaneous mutation	NPC1	Juvenile form, develops neurological symptoms at 8-12 weeks, death occurs 8-12 months
Canine	Spontaneous mutation	NPC1	Develops neurological symptoms at 2 months, put down at 9 months

Table 1.2: Summary table of NPC animal models

Feline [111,112], canine [113] and three murine (BALB/c-*npc1*^{nih} [114], *npc*^{nmf164} [115] and C57BL/KsJ-*npc1*^{spm} [116]) models of NPC1 exist, with the BALB/c-*npc1*^{nih} model being the most commonly used (**Figure 1.17**). All except the *npc*^{nmf164} murine model are caused by spontaneous mutations.

The spontaneous mutation in the *npc1*^{nih} murine model results in truncation of the Npc1 protein before the SSD, eliminating 11 of the 13 transmembrane domains. This model corresponds to the severe infantile form of the human disease [117] and the mice are born phenotypically normal, developing neurological symptoms approximately by 6 weeks of age. Ataxia and tremor are the common hallmarks of disease progression, likely caused by the characteristic neurodegeneration of the cerebellum. Mice reach a terminal end stage of disease by 10 – 12 weeks without therapeutic intervention [118].

The second spontaneous murine model, C57BL/KsJ-*npc1*^{spm}, was observed by Miyawaki and colleagues in 1975 and also corresponds to a severe infantile disease form [119]. This model presents with hepatosplenomegaly, foamy cells, degeneration of Purkinje cells and storage of free cholesterol and sphingomyelin. As with the *npc*^{nih} model, the characteristic neurological symptoms of ataxia and tremor are present, but onset is about a week later, around 7 weeks of age [119]. Growth and weight also decrease and death follows at 12 – 14 weeks of age [119].

Unlike the other two mouse models, the *npc*^{nmf164} mouse model was created via ethylnitrosourea mutagenesis [115]. This resulted in a point mutation in the *Npc1* gene, causing an amino acid change in the large cysteine-rich luminal loop of Npc1 between the 8th and 9th transmembrane domains, p.D1005G. This mutation confers only a partial

loss of function of *Npc1* and therefore corresponds to a late-onset milder form of NPC than the other mouse models. The mice still present with the classical pathology associated with NPC1; increased storage of unesterified cholesterol, glycolipids and sphingomyelin, foamy macrophages, progressive neurodegeneration of the cerebellum and increased neuroinflammation [115]. However, the onset of ataxia and tremor does not occur until around 11 weeks of age [115]. Correspondingly, there was increased survival and onset and progression of weight loss was delayed. The increased survival of *npc^{nmf164}* homozygotes allows mating of homozygous pairs, which is a huge advantage as this enables generation of litters comprised of only homozygotes [115].

A knock-out NPC2 mouse model is also available, which presents with a phenotype similar to NPC1. This corresponds to the human form of the disease, where mutations in NPC1 or NPC2 result in similar phenotypes [120].

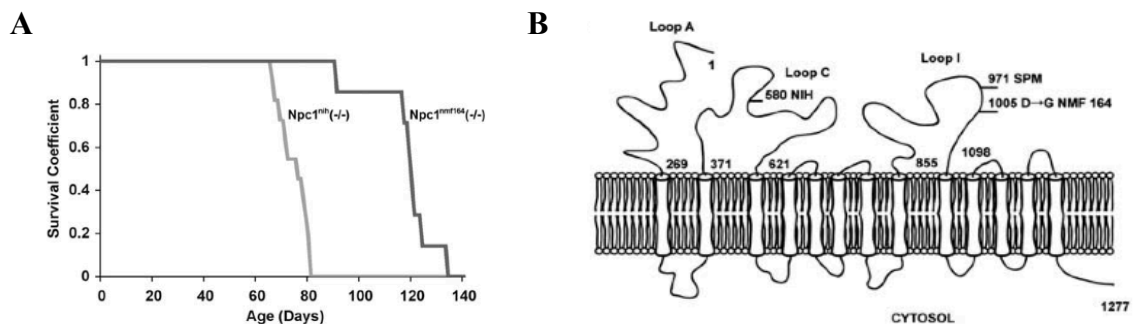


Figure 1.17: NPC1 mice [115].

A) Survival plot of *npc^{nih}* and *npc^{spm}* mice. It can be easily observed that the *npc^{spm}* mice live substantially longer. **B)** Diagram of the NPC1 protein with locations of the mutations of the three mouse models indicated. *Npc^{nih}*; amino acid 580 in loop C. *npc^{spm}*; amino acid 971 in loop 1. *npc^{nmf}*; amino acid 1005 in loop 1

The feline model of NPC originated from a spontaneous mutation in the affected domestic kitten, which presented with the neurological symptoms indicative of NPC, namely ataxia and tremor [112]. Further study showed that, similar to patients, cats initially presented with low birth weight, hepatomegaly, and subsequently neurological symptoms. Tremor appeared at 8 – 12 weeks of age, which progressed to ataxia and death by 8 – 10 months. Biochemically, the mutant cats show storage of glycolipids, unesterified cholesterol and sphingomyelin [112]. Fibroblasts cultured from NPC cats showed increased perinuclear filipin staining, confirming storage of unesterified cholesterol and mimicking the phenotype seen in patients. Histology of the brain showed the presence of axonal spheroids, ectopic dendritic growth and the characteristic degeneration of PCs in the cerebellum [112]. Similar to the npc^{nmf164} murine model, the feline model presents with a less-severe form of NPC1, representative of the juvenile neurological human form [121].

A spontaneous canine model has also been seen, although this model is not easily available [113]. At 2 months of age, a boxer dog showed signs of decreased motor ability. Further neurological symptoms progressed and the dog was euthanized at 8 months of age due to inability to stand. Post mortem analysis showed that although the majority of organs were normal, there was mild atrophy of the cerebellum caused by degeneration of PCs and granular cells. Elevated levels of unesterified cholesterol and GSLs were also present [113].

1.5 Therapeutic Intervention

1.5.i: Smith-Lemli-Opitz Syndrome

SLOS is likely caused by a combination of fixed developmental defects and functional defects due to the abnormal sterol levels [122]. Drug treatments mainly focus on the functional defects, although it is currently unclear whether the defects result from the increase in 7DHC or decrease in cholesterol.

To correct the cholesterol deficit, it has become standard of care to provide SLOS patients with cholesterol supplementation [56]. During infancy this is 30-40 mg/kg/day but this can be increased up to 300 mg/kg/day [59]. Cholesterol can be delivered in multiple ways; in a natural form, for example via eggs or cream, or as purified food grade cholesterol sprinkled on food or diluted in oil [56]. Cholesterol supplementation can provide the tissues with cholesterol but can also down-regulate cholesterol biosynthesis via HMGR, therefore reducing 7DHC. Dietary cholesterol can also down-regulate the LDL receptor, therefore increasing plasma LDL and total cholesterol concentrations. However, the efficacy is not consistent between trials [59]. Some studies reported a reduction in plasma 7DHC and increased plasma cholesterol [76], and although Nwokoro and Mulvihill reported elevated plasma cholesterol in children, this did not follow with the older patients [123], potentially due to the source of cholesterol and trial duration. Additionally, cholesterol supplementation has also been reported to improve growth, decrease photosensitivity and possibly improve behaviour, but these results are subjective [124].

Although cholesterol supplementation has the ability to increase the plasma cholesterol, dietary cholesterol does not pass through the BBB and therefore will have limited, if any, effect on neurological manifestations [59]. Simvastatin is an inhibitor of HMGR [125] that is able to cross the BBB [126] and therefore has the potential to improve cognition. On the other hand, decreasing cholesterol synthesis has the possibility of reducing the already low cholesterol levels in SLOS. An option is to give statins in combination with cholesterol supplementation [127].

One study found that in mildly affected patients, the combination of simvastatin and cholesterol supplementation had a similar positive effect to cholesterol supplementation alone. In both cases, the plasma 7+8DHC/cholesterol ratio decreased but via different mechanisms; the 7+8DHC levels decreased in the combination therapy, whereas the cholesterol levels increased with cholesterol supplementation alone [127]. However, simvastatin provided no significant benefit in more severe patients. In agreement with this, simvastatin proved toxic in two clinically severe cell lines homozygous for a null mutation [128]. This study also showed that simvastatin increased DHCR7 expression in both control and SLOS fibroblasts. Other components of the cholesterol synthesis pathway, such as HMGR, LDLR and SQS, were also up-regulated. It is hypothesised that Simvastatin causes increased expression of the allele with residual DHCR7 activity [128].

Many of the phenotypic features of SLOS are fixed developmental problems. Therefore in order to correct these, *in utero* therapy would likely have to be used. Lindegaard *et al* provided a proof-of-principle study focusing on up-regulating Abca1 transcription [129]. Abca1 is involved with the transport of cholesterol across the placenta from the

maternal to foetal circulation. This study showed that liver X receptor (LXR)-agonists, such as T0901317, increase *Abca1* transcription in placental tissue and treating *Dhcr7*^{-/-} mouse embryos with this compound increased cholesterol [129]. Although this is a promising therapeutic strategy, there are differences in mouse and human placental function and structure. The yolk sac in mice persists throughout gestation but is only important in the first 2-3 weeks of human gestation. Secondly, the anatomical structures of human and murine placentas differ, potentially affecting the transfer of nutrients [129].

1.5.ii: Niemann-Pick type C disease

Enzyme replacement therapy (ERT) is a common therapeutic approach for disorders caused by defective enzymes. Although lysosomal storage disorders, such as type 1 Gaucher disease and Fabry disease, have benefited from ERT, many lysosomal enzymes do not cross the BBB and consequently, ERT cannot be used to treat neurological manifestations [130]. A second approved method for type 1 Gaucher disease is oral administration of a small molecule, called miglustat. This approach is known as substrate reduction therapy (SRT). Gaucher disease is caused by defects in glucosylceramide synthase, which results in storage of sphingolipids [130]. Miglustat, an inhibitor of glucosylceramide synthase, affects glycosphingolipid (GSL) biosynthesis and thus reduces storage of GSLs in the lysosomes [131].

One of the hallmarks of NPC is the accumulation of GSLs, including GM2 and GM3 [132]. Therefore it was hypothesised that miglustat may also be beneficial in NPC. Miglustat has the advantage of being able to cross the BBB and therefore is likely to be

able to affect the cognitive aspects of NPC [133]. Zervas *et al* tested miglustat in both the murine *npc^{nih}* and cat models of NPC1 and both models showed a delayed onset and progression of the clinical symptoms, such as ataxia and tremor [134]. Survival was also improved and, biochemically, high performance lipid chromatography (HPLC) showed that miglustat lowered the GSL storage in the cerebral cortex of both mice and cats. After success in animal models, miglustat was then trialled in patients. It was shown that miglustat could improve symptoms, such as the ability to swallow, reduce lipid storage and normalise trafficking [133]. To date, miglustat is the only approved therapy for NPC [130].

Other monotherapies have also been tested in NPC, including ibuprofen and curcumin. Curcumin is a weak sarcoendoplasmic reticulum calcium transport ATPase (SERCA) antagonist and elevates cytosolic calcium. Curcumin treated mice had improved coat condition, weight gain, increased activity and survival and decreased tremor [104]. Williams *et al* also investigated monotherapy treatment of curcumin, ibuprofen and miglustat in NPC1 mice, but furthered the study by looking at combinations of these drugs [135]. The untreated mice die at a mean age of 10.5 weeks but survival was increased by all the monotherapies. The combination of curcumin and miglustat had the greatest survival, 83% above untreated. Onset of weight loss was delayed and motor function improved in all combinations containing miglustat. However, the triple therapy combination was the most effective [135]. This was also shown in histological analysis of the NPC1 mouse brain where the triple combination gave the best protection against microglial recruitment and activation and for PC survival [135] (**Figure 1.18**).

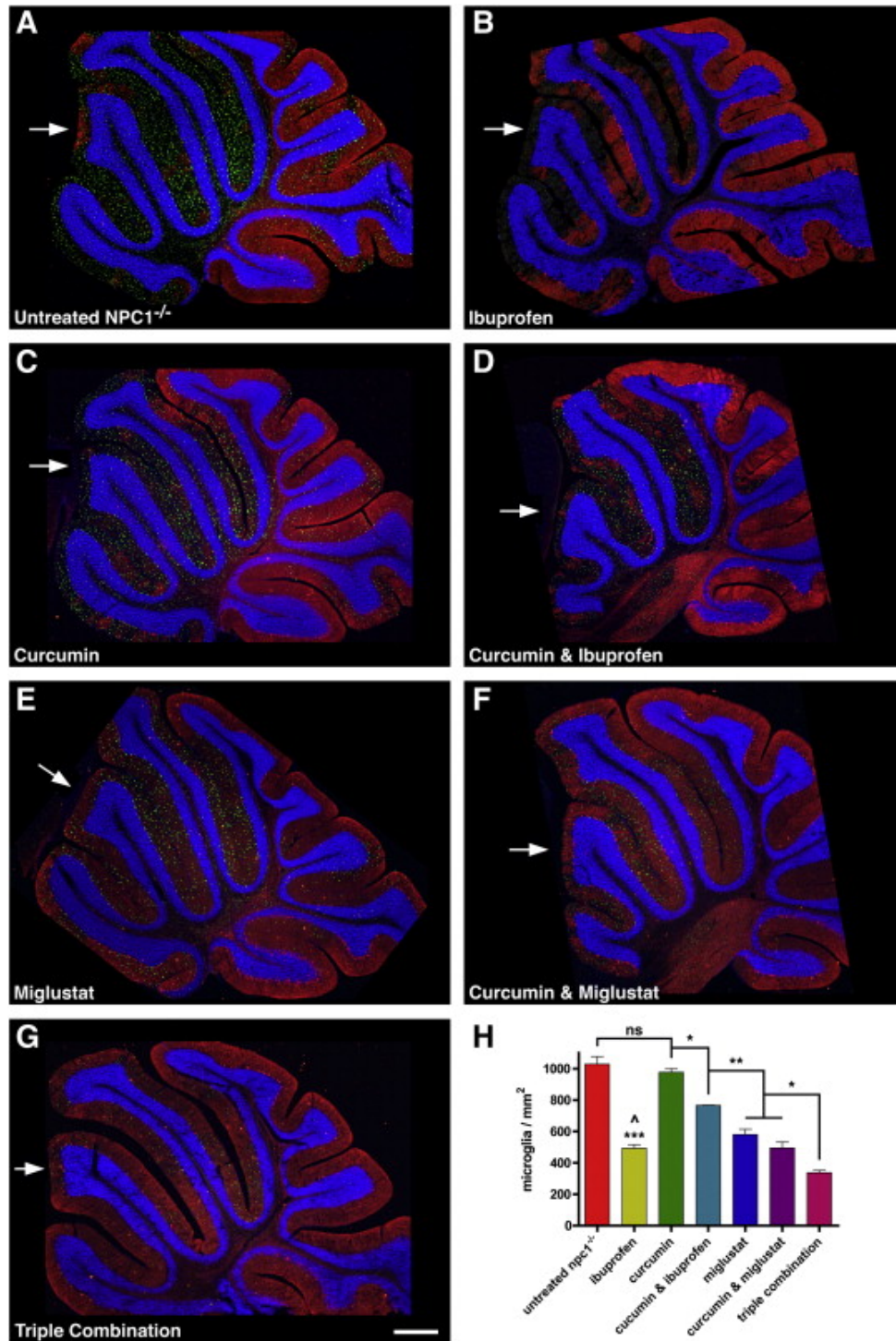


Figure 1.18: Therapeutic effect of monotherapies and combinations [135].

These images show the cerebellum of 7.5 week old *npc^{nh/-}* mice. Purkinje cells are in red, microglia in green and a nuclear stain in blue.

A) untreated; B) Ibuprofen monotherapy; C) Curcumin monotherapy; E) miglustat monotherapy; D) Curcumin + Ibuprofen double combination F) Curcumin + miglustat double combination; G): triple combination. H) Quantification of results

Cholesterol is the precursor for steroids and as NPC mice are hyperandrogenic and have underdeveloped reproductive organs, this is indicative of defective steroid biosynthesis [136]. Neurosteroids are formed in the brain from cholesterol and it has been shown that both pregnenolone, the first steroid produced from cholesterol, and allopregnenolone are present at decreased levels in NPC mice [136]. Studies showed that neurosteroidogenesis was not decreased in foetal mouse development but steadily diminished after birth, particularly the activity of 3 α -hydroxysteroid dehydrogenase (3 α HSD), which is an allopregnenolone-synthetic enzyme [136]. Therefore, allopregnenolone was administered to NPC1 mice and it was found that there were improvements in lifespan, increased PC survival and a decrease in the rate of decline of locomotor function and motor co-ordination. Drug efficacy was greater when allopregnenolone was administered earlier [136].

The vehicle for administering allopregnenolone was cyclodextrin (CD). Combination studies suggested that it might have been CD and not allopregnenolone that was causing the beneficial effects [137] and it was found that mice treated with CD lived significantly longer, had delayed weight loss, decreased cholesterol and ganglioside storage and rescue of PCs [137]. Treated mice also had improved behavioural characteristics, such as delayed ataxic gait and tremor [137]. However, some *in vivo* studies have reported that CD cannot cross the blood brain barrier and therefore it is unclear how CD is improving the neurological deficits in NPC. There is a suggestion that CD interacts with 24-hydroxycholesterol facilitating its movement across the blood brain barrier [138]. CD is also beneficial in NPC2 models [139].

Although these novel therapies are beneficial in animal models, translating these therapies to human trials is an on-going and challenging task. Bone marrow transplantation has been successful in other lysosomal storage diseases but had no efficacy in alleviating the neurological symptoms in either animal models or in a human patient [140]. Similarly, liver transplantation caused no improvement in neurological symptoms [141]

1.6 Aims of Thesis

The clinical sides of both SLOS and NPC have been relatively well studied. However, therapeutic options for both disorders are limited, partly due to uncertainty in what metabolite or pathway are causal. Throughout this thesis, I aim to help elucidate the mechanisms behind these disorders.

As described in this introduction, there are multiple animal models of both SLOS and NPC. Animal models generally have a dual purpose in clinical research; to gain insight into the biochemical and physical aspects of the disorder and to test therapies for safety and efficacy before proceeding to clinical trials. Efficacy can be tested via invasive methods, such as dissection and histological analysis, or non-invasive methods, for example, normalisation of behaviour.

The model I will use in this thesis is the *Dhcr*^{T93M/ Δ 3-5} murine model of SLOS.

Although the biochemistry of this model has been well characterised, the behavioural phenotype has not. Therefore I aim to begin to characterise the behavioural profile of the *Dhcr*^{T93M/ Δ 3-5} mouse model, with the future potential for testing novel therapies.

Limited neurological structural analysis has been carried out in this model and therefore I will study this in the mice that have undergone behavioural analysis. In this way, I aim to discover if there are any links between behaviour and neurological structure, which could provide potential therapeutic targets.

Miglustat is the only approved therapy for NPC and the NPC-like phenotype in SLOS suggests that miglustat may also be a potential therapy. The therapeutic mechanism of

miglustat is unknown and I aim to investigate this via RNAseq in control, SLOS and NPC patient fibroblast lines. This will reveal gene expression changes that could be used to discover through which pathway miglustat acts, if the pathways are the same in SLOS and NPC and provide potential new drug targets.

Once therapies have been developed they need to be administered within an optimum time window. As SLOS and NPC have broad severity spectrums or can present with atypical phenotypes, there may be discrepancies between the observed clinical incidence and the actual disease incidence. I aim to estimate the actual incidence via a bioinformatics approach using massively parallel exome sequencing data sets. This has the potential to identify groups of patients that may otherwise be missed and therefore enable the best possible treatment.

Chapter 2: Determining allelic frequency of Smith-Lemli-Opitz Syndrome and Niemann-Pick disease type C via use of massively parallel sequencing data sets

2.1 Introduction

Disease incidence is designated as the number of new cases within a specified time window in a given population [142]. For monogenic disorders, incidence can be estimated using allelic frequency (the incidence of a particular gene variant in a population) via the Hardy-Weinberg equation. Although deep sequencing can be expensive, multiple massively parallel sequencing data sets already exist and these can be used to identify allelic frequencies. Subsequently, an estimated incidence of monogenic disorders can be determined.

2.1.i: Smith-Lemli-Opitz Syndrome

The current incidence of SLOS is reported to be between 1:10,000 and 1:70,000 depending on the population studied [2]. The higher incidence is generally thought to be caused by founder effects in Eastern Europe [143]. *DHCR7* has a length of around 14 kb, constituting 9 exons, and is located in chromosome 11q12-13 [144,145]. The translated protein has a molecular weight of 54.4 kDa and comprises 475 amino acids

[146]. Over 164 mutations in *DHCR7* have been reported [143,147–151] but not all are present at the same frequency in every population.

Although c.964-1G>C is considered the most common mutation, comprising approximately one third of all SLOS alleles, this is not true for every population and this variant has decreasing prevalence moving eastward across Northern Europe [152].

There is currently a reported carrier frequency of 1% in North American populations [5] but this may vary in different geographical locations. For example, this frequency is increased (1.4% - 2%) in European Caucasians, which results in an estimated incidence of 1:20,000 – 1:40,000 [57,153]. The variants p.W151X (c.452G>A) and p.V326L (c.976G>T) are most prevalent in North East Europe and p.T93M (c.278C>T) in Mediterranean populations [154]. Interestingly, SLOS appears to be less common in individuals of African and Asian heritage [66].

The abnormal sterol levels present in SLOS, reported in 1993 [155], allowed both confirmation of clinically suspected cases and exclusion of misdiagnosis. This enabled incidence to be more reliably estimated and since the availability of the test the incidence of SLOS has been predicted to be 1:50,000 [156], 1:60,000 [157] and 1:70,000 live births [158]. However, in agreement with the distribution of SLOS alleles, these estimates are much higher in some European populations; for example 1:10,000 in the Czech Republic [156].

2.1.ii: Niemann-Pick type C disease

NPC is caused by mutations in *NPC1* and *NPC2* and the combined minimal incidence has previously been reported as 1:104,000 [84]. It is possible that this is an under ascertainment due to variation in presentation and limitations of diagnostic testing. Initially NPC was diagnosed via histopathological analysis of organs, such as the spleen, liver and brain [159] but, after the identification of NPC as a lipid-storage disorder in the 1980s [160], the filipin test started to be utilised [161,162]. Filipin is an auto-fluorescing polyene antibiotic that binds to unesterified cholesterol and has an excitation wavelength of 360nm and emission wavelength of 480nm [100]. As such, filipin staining of > 80% of patient fibroblasts presents with bright punctate staining, representing the storage of unesterified cholesterol in the lysosomes. Despite this, diagnosis via this method is not always straightforward. A subset of patients present with a classical NPC clinical phenotype but have a variant biochemical phenotype with intermediate to normal cholesterol esterification rates [101,163]. The *NPC1* gene was characterised in 1997 [164] allowing genetic analysis.

As technology has improved genetic testing has become more prevalent and has allowed pre-natal testing and genetic counselling [98]. Additionally, the incidence of NPC may be an underestimate due to the variability in clinical symptoms and biochemical phenotypes. Genetic profiling has enabled earlier detection of NPC but, as in SLOS, distribution of mutations is not equal across populations. As stated in the **Introduction**, NPC1 is a 1252 amino acid membrane protein with 13 transmembrane domains, 3 luminal loops, a SSD, a cleaved N-terminal signal sequence and a C-terminal sequence [92,93]. In contrast, NPC2 is a small soluble protein of 132 amino

acids [84]. The most common mutation in NPC1 patients of Western European descent is p.I1061T (c.3182T>C), accounting for approximately 20% of all recorded NPC1 alleles in the UK and France [165]. It is usually associated with classical low cholesterol-esterification rates and a juvenile age of onset [166]. On the other hand, p.P1007A (c.3019C>G) is the second most common mutation in Europeans but is associated with a 'variant' phenotype [167]. The most common mutation reported in NPC2 is p.E20X [168].

Despite improvements in diagnosis methodology, it often is necessary to use a variety of techniques (histology, clinical symptomatology, biochemical and molecular analysis) in order to achieve a correct diagnosis. This can lead to a delay of greater than 5 years between symptom onset and diagnosis [159].

As increasing numbers of potential therapeutic interventions are being developed for both SLOS and NPC, it is becoming more critical for accurate incidence to be determined in order to most effectively treat all patients. Therapeutic intervention has the greatest effect if administered within an optimum time window, which requires early identification of a disorder, usually before symptoms have started to appear. SLOS and NPC have broad severity spectrums and, sometimes, atypical presentations. Consequently, there may be under ascertainment of the actual incidence and determination of this incidence will enable "at risk" populations to be identified. This will condense the number of individuals to genetically counsel and screen, therefore reducing costs, and will potentially enable misdiagnosed or missed individuals to attain the best possible treatment.

2.1.iii: Aims of Chapter 2

- To establish an estimated incidence of SLOS and NPC using available massively parallel sequencing data sets
- Ascertain if there is potential under ascertainment of patients

2.2 Materials and Methods

2.2.i Data sets

Four large massively parallel sequencing data sets were used; version 3 release of the 1000 genome project [169], National Heart, Lung and Blood Institute GO Exome Sequencing Project (ESP) [170], ClinSeq [171] and a database from an National Institutes of Health (NIH) inter-institute collaboration on autism (primary investigators: F.D.P., J.E.B.-W., E Tierney, and A Thurm). Patients of NPC and SLOS were excluded, leading to different numbers of total alleles analysed.

For SLOS, the 1000 genome project contributed a maximum of 2,184 alleles, ESP contributed 12,988 alleles, ClinSeq contributed 1902 alleles and the inter-institute Autism project contributed 762 alleles. A maximum total of 17,836 alleles were analysed.

For NPC1 and NPC2, the 1000 genome project contributed a maximum of 2,184 alleles, ESP contributed 13,006 alleles, ClinSeq contributed 1902 alleles and the inter-institute Autism project contributed 662 alleles. A maximum total of 17,754 alleles were analysed.

2.2.ii Variant calls and annotation

Variant call data was generated by Perl script for areas overlapping the defined gene; *DHCR7*, *NPC1* or *NPC2*. Each call was evaluated for each allele at every base of the coding exons or occurring within 5 base pairs of an exon. Annotation of mutations was performed using SNPnexus [172] with Refseq annotations [127].

2.2.iii Determination of pathogenicity

Multiple processes were used to determine pathogenicity of each variant call, using both manual and bioinformatics approaches. Using the Human Gene Mutation Database (HGMD) [151] and the NIH NPC patient cohort database, variants that had previously been identified were determined. However, inclusion in HGMD does not always indicate pathogenicity in a patient, therefore primary literature was also consulted. Tertiary structures of NPC1 and NPC2 were modelled using I-TASSER [173] and variants were mapped onto these structures in order to identify conserved residues and possible interactions.

Multiple types of variant were discovered when probing the datasets; coding single nucleotide base variants (CSNBs), splice variants, insertions, deletions and indels. Three *in silico* programmes were used to predict the pathogenicity of CSNBs – PolyPhen-2 [174], SIFT [175] and Mutation assessor [176]. SIFT and Polyphen-2 designates each variant as “Benign”, “Possibly Damaging” or “Probably Damaging”. Variants that were predicted “Possibly Damaging” or “Probably Damaging” were considered pathogenic. Mutation assessor predicts the variants to be “Non-functional

low”, “Non-functional neutral”, “Functional (medium)” or “Functional (high)”.

Variants that were predicted “Functional (medium)” or “Functional (high)” were considered pathogenic. In cases where the software did not agree on predicted pathogenicity, the prediction of two out of the three was accepted. Pathogenicity of splice variants was predicted by MaxEntScan [177]. This software predicted variants as “Neutral”, “Negative” or “Strongly Negative”, with pathogenicity accepted when the prediction was “Negative” or “Strongly Negative”.

2.2.iv Molecular analysis

Some determinations of pathogenicity required identification of cell lines heterozygous for the required variant. Lymphoblast and fibroblasts were grown in T75 flasks at an approximate density of 5 million and 3 million cells respectively. Lymphoblasts were grown in Roswell Park Memorial Media (RPMI, cat.# 11875-093, ThermoFisher Scientific) with 15% fetal bovine serum (cat.# FB-12, Omega Scientific Inc.), 1X (2mM) L-glutamine (cat.# 25030-081, Life Technologies), 1X Penicillin (100 Units/ml) /Streptomycin (100µg/ml) (cat# 15140-122, Life Technologies) in a 5% CO₂ environment at 37°C. Fibroblasts were grown in Dulbecco’s Modified Eagle Medium (DMEM, cat.# 11995-065, ThermoFisher Scientific) with 10% fetal bovine serum (cat.# FB-12, Omega Scientific Inc.), 1X (2mM) L-glutamine (cat.# 25030-081, Life Technologies), 1X Penicillin (100 Units/ml) /Streptomycin (100µg/ml) (cat.# 15140-122, Life Technologies) in a 5% CO₂ environment at 37°C.

Cell pellets were collected from the cell lines and mRNA was isolated (RNeasy mini kit, cat.# 74106, Qiagen) and reverse transcribed (SuperScript III One-step RT-PCR

System, cat.# 10928-034, Life Technologies). The primers for the reverse transcription are listed below.

Primers for *DHCR7*:

- Forward (HSP1D) – 5' AGGTGTGCGCAGGACTTTA 3'.
- Reverse (HSP13) – 5' GCTGGGCTCTCTCCAGTTTA 3'

Primers for *NPC2*:

- Forward (NPC2 F3) – 5' GGTGGAGTGGCAACTTCAGG 3'
- Reverse (NPC2 R2) – 5' CACTGGATACCATTGGAGAGC 3'

A 1.5% agarose gel (Ultra-Pure Agarose, cat.# 15510-027, Life Technologies) was used to visualise the cDNA. Required bands were cut from the gel, purified and cloned (TOPO TA Cloning Kit for Sequencing, cat.# 45-0030, Life Technologies). Colonies were grown overnight in LB-ampicillin (cat.# BPL-2580, KD Medical) and plasmid DNA isolated (QiaSpin MiniPrep Kit, cat# 27106, Qiagen). Sequencing was then performed using the BigDye Terminator v3.1 Cycle sequencing Kit (cat.# 4336917, Life Technologies) and a 3500xL Genetic Analyser (Applied Biosystems).

2.2.v Determination of predicted disease incidence

To estimate a carrier frequency, the total number of variants that were predicted or shown to be pathogenic was divided by the total number of variants. The predicted incidence was then calculated using the Hardy-Weinberg equation. The Hardy-Weinberg equation relies on five assumptions [178]:

- 1) No genetic drift
- 2) A closed population (there is no genetic transfer between populations, from migration or transfer of gametes)
- 3) Mutations are negligible
- 4) Individuals are mating randomly
- 5) Natural selection is not operating on the population

The population sampled in this study is relatively large and therefore it can be assumed that there is limited genetic drift. We made the assumption that allelic frequencies were consistent across the different ethnic groups represented in the data sets. Two of the data sets incorporate two or more ethnic groups: the 1000 genomes project and ESP. The number of alleles represented by each group in the 1000 genomes project is negligible in comparison to the number of alleles analysed in total. However, the greatest potential error in this assumption would arise from ESP where there are a large number of individuals from either European or African descent.

NPC1 could potentially have a selective pressure, as it has association with filovirus' such as the Ebola virus [179]. However, Al-Daghri *et al.* determined that selective pressure on NPC1 was weak to neutral [180].

2.3 Results

2.3.i Smith-Lemli-Opitz Syndrome

A total of 17,836 chromosomes were analysed for *DHCR7* and from the four datasets, 515 nonsynonymous sequence variants were present (**Summary Table of *DHCR7* in Appendices**). As none of the chromosomes analysed were from patients, this resulted in an initial carrier frequency of 2.89%. These 515 nonsynonymous sequence variants comprised 74 distinct changes; 71 coding single nucleotide base variants (CSNBs), 2 splice mutations and 1 insertion. 43 of these distinct changes were unique, appearing only once in all four datasets. However, 9 of the 74 variants appeared 10 or more times, accounting for 408 of the total 515 alleles (79%).

The potential pathogenicity of the variants was established using a combination of literature searching and the HGMD®. From this, the variants that had been previously identified in patients were found. 25 variants had been previously identified in patients; 23 CSNBs and 2 splice mutations. PolyPhen-2 predicted 18 of the CSNBs to be Probably Damaging, 4 to be Possibly Damaging and 1 to be Benign. However, as the predicted benign variant (c.400G>T, p.V134L) was reported by Waterham [143] in an SLOS patient, we considered it to be pathogenic. The remaining unpublished 48 CSNBs had predictions of 17 Probably Damaging, 7 Possibly Damaging and 24 Benign. This gave a total of 47 out of 71 CSNBs that were considered pathogenic (66.2%).

Of the 2 splice variants MaxEntScan predicted c.99-4G>A as Neutral (non-pathogenic) and c.964-1G>C as Negative (pathogenic). c.964-1G>C is the most common mutation seen in SLOS patients with a previously reported carrier frequency of around 1% in Northern European populations. However, the carrier frequencies of this variant in each of the four data sets are as follows: 0.46% in 1000 Genome, 0.54% in ESP, 0.97% in ClinSeq and 0.66% in the Autism collaboration.

Although the ClinSeq frequency is in good concordance with the reported carrier frequency, the other results are approximately half what would be expected. The 1000 Genome project separates out the variants by population and therefore it is possible to restrict analysis to only Northern European populations. If this restriction is taken into account, the frequency increases to 0.81%, significantly closer to the reported frequency (**Table 2.2**). ClinSeq mainly comprises a North American cohort, also potentially explaining the higher frequency. ESP is comprised of two groups; European Americans (EA) and African Americans (AA). Although the number of alleles contributed by each population is not defined for this particular variant, the total alleles contributed are 8588 for EA and 4400 for AA. Using these numbers results in carrier frequencies of 0.63% for EA and 0.36% for AA. The EA frequency is still below expected, but as the frequency of c.964-1G>C varies across Europe and there is no further population segregation, this is not surprising. The interesting point to note is that the frequency is higher in EA than AA.

	1000 Genome	NE populations*	ESP	ClinSeq	Autism
c.964-1G>C variant	10	8	70	9	5
Total alleles analyzed	2184	988	12864	926	762
Rate	0.46%	0.81%	0.54%	0.97%	0.66%

Table 2.2: Table of c.964-1G>C variant for SLOS.

The first row designates the number of variants seen in each dataset. The second row shows the total number of alleles analyzed and the third row shows the carrier rate. It can be clearly seen when only the North Eastern populations are analysed in the 1000 genome project, the carrier frequency increases from 0.46% to 0.81%.

*Denoted Northern European populations or those with Northern European Ancestry as follows:

- CEU: Utah Residents (CEPH) with Northern and Western European ancestry
- CLM: Colombians from Medellin, Colombia
- FIN: Finnish in Finland
- GBR: British in England and Scotland
- IBS: Iberian population in Spain
- PUR: Puerto Ricans from Puerto Rico
- TSI: Toscani in Italia

The second splice variant, c.99-4G>A, is present at a frequency of 0.17%. Although MaxEntScan predicted this variant to be Neutral, as it was previously published as a disease causing mutation [143] it was initially considered to be pathogenic. However, the relatively high frequency rate (second only to c.964-1G>C) combined with only being reported in 1 patient, questioned this conclusion. mRNA was isolated from a cell line heterozygous for c.99-4G>A (NIMH 32), a control cell line and CWSLO (a patient

cell line containing the c.964-1G>C and p.T93M variants). The mRNA was then reverse transcribed and sequenced. No alternative splicing and two distinct normal alleles were observed (**Figure 2.1**) and therefore it was concluded that the c.99-4G>A variant was non-pathogenic. The insertion, c.849_856CTTCT, was considered to be pathogenic as it causes a frame shift (p.F284S*10) and early termination of the protein.

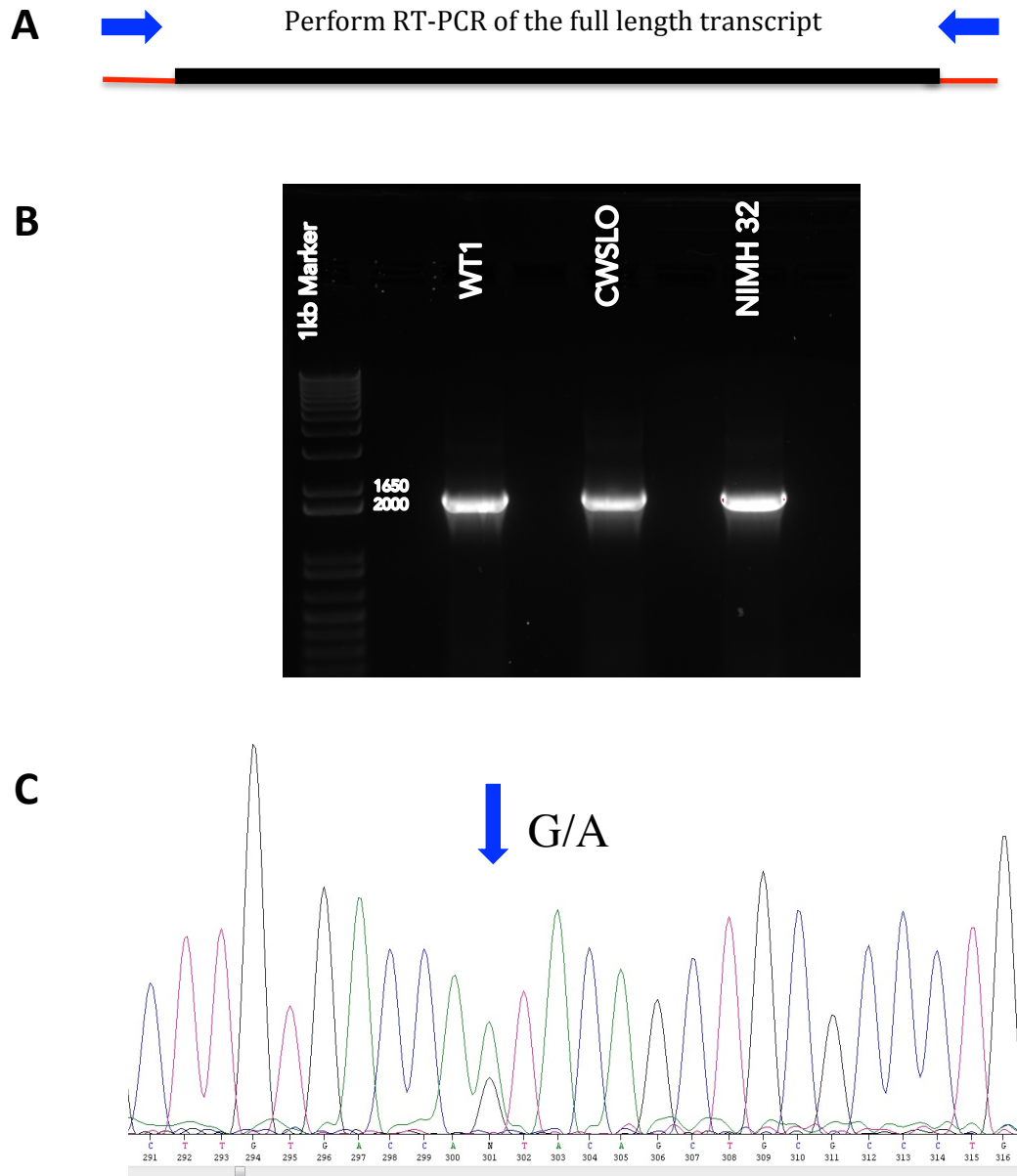


Figure 2.1: Analysis of the c.99-4G>A variant

- A) Reverse transcription was performed on the full length transcript
- B) ~1.7kb band is expected for the control cell line (WT1). CWSLO has an insert of ~130bp and as such, a doublet is expected. However, from this gel, it is difficult to determine small changes in molecular weight.
- C) The full length sequencing of NIMH 32 showed two distinct normal alleles with naturally occurring polymorphisms, G/A.

Using an *in silico* modelling program, Protter, CSNBs can be visualised on the structure of DHCR7 (**Figure 2.2**). This can show if the variants are clustered or spread throughout the protein. For DHCR7, the variants are relatively evenly dispersed, suggesting the absence of a particular location that causes the enzymatic defect.

Taking into consideration all the analyses, this resulted in a total of 191 pathogenic alleles and a carrier frequency of 1.07%. PolyPhen-2 includes a False Discovery Rate (FDR) for all CSNBs. For all three prediction categories, the FDR was 12% but if only Probably Damaging and Possibly Damaging were included the FDR was 5%. After including the FDR, this resulted in a final total of 181 pathogenic alleles and a carrier frequency of 1.01%. Using the Hardy-Weinberg equation a disease incidence of between 1/34,937 and 1/39,215 was calculated. This is in good concordance with the previously reported estimate of 1/40,000 [153,158], especially if *in utero* fatalities are taken into consideration.

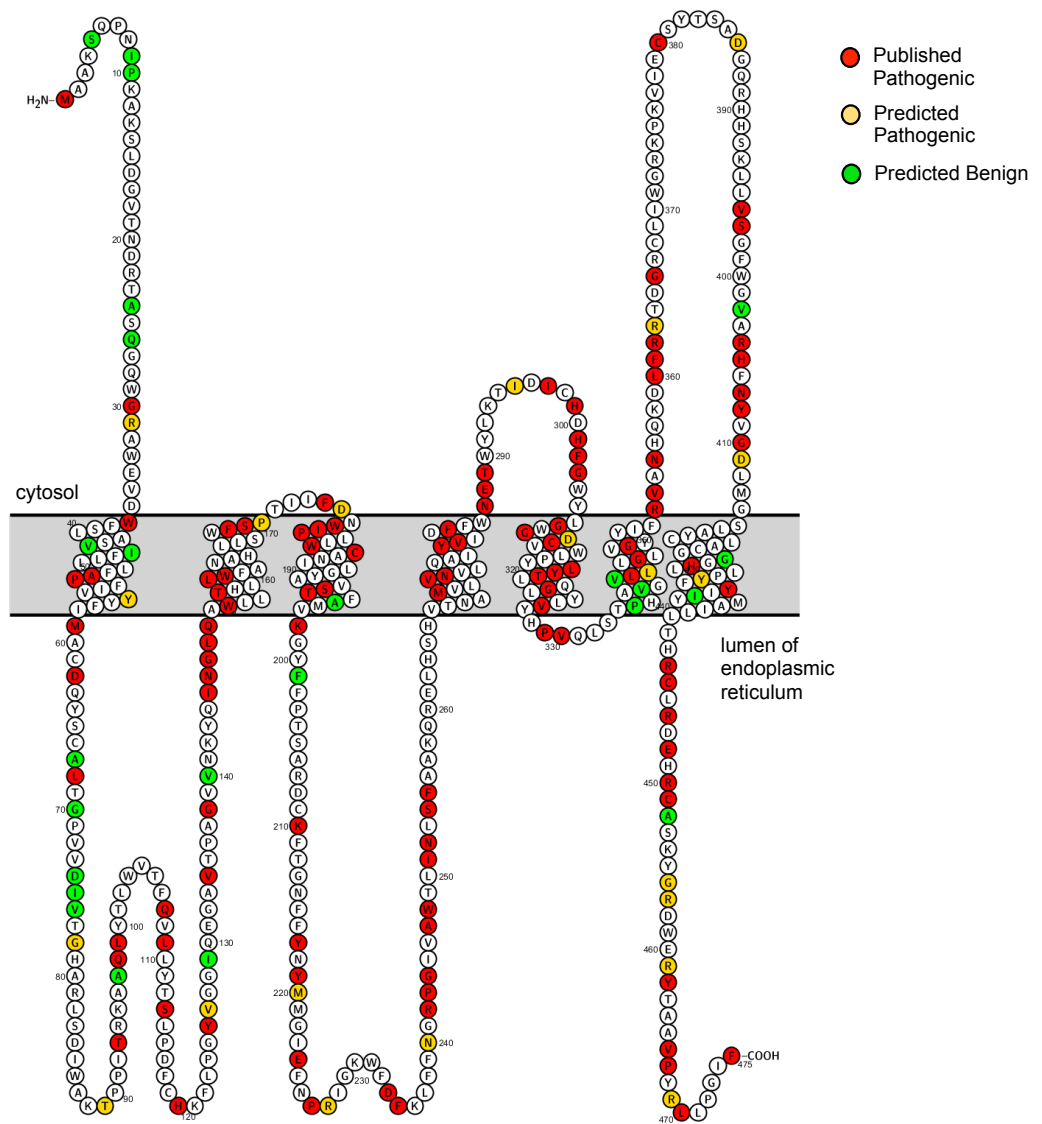


Figure 2.2: Model of DHCR7 protein using the program “Protter” [248]

This figure shows the position of published variants (red), variants that are predicted pathogenic (yellow) and those that are predicted to be benign (green). This demonstrates that the pathogenic mutations are spread relatively evenly across the entire length of the protein.

2.3.ii Niemann-Pick type C1 disease and Niemann-Pick type C2 disease

A total of 17,754 chromosomes were analysed for *NPC1* and *NPC2* sequence variants, with 16,455 and 271 variants found respectively. Of the variants found in *NPC1*, this was split into 147 distinct changes; 129 CSNBs, 9 splice variants and 9 insertions/deletions. The 271 variants found in *NPC2* were split into 14 distinct changes; 12 CSNB and 2 splice variants (**Summary Tables of *NPC1* and *NPC2* in Appendices**).

Due to the increased number of variants and the added complexity compared to SLOS, two additional *in silico* prediction software have been used to ascertain pathogenicity of CSNBs; SIFT and Mutation assessor. Where the software differed in prediction of pathogenicity, the prediction of two out of the three was used. With respect to *NPC1*, 30 distinct CSNB changes and 2 out of 9 splice variants were predicted to be pathogenic. Out of the total 129 CSNBs found in *NPC1*, 58 were initially ascertained to be pathogenic; 28 previously published, 30 predicted via *in silico* analysis. 4 CSNBs were known polymorphisms and 1 CSNB had been published as a Benign variant. 4 insertions/deletions had been previously published. *NPC2* had 8 CSNBs (6 not reported in HGMD) and 1 out of 2 splice variants that were predicted pathogenic. The FDR for *NPC1* and *NPC2* respectively were 0.04% and 0.03%. This effect is negligible and therefore was not included in the calculations for incidence. Additionally, one novel *NPC1* variant was found in the NIH cohort (c.2524T>C, p.F842L).

In *NPC1*, one indel (c.2020_2021del) was present 319 times but only in the ESP dataset. As it was unique to one dataset and present in such high numbers, this variant

was considered a technical artefact and removed from the calculations. 8 other indels were found in *NPC1*. These caused frame-shifts and were considered pathogenic. No indels were identified in *NPC2*.

NPC1

Based on these predictions, *NPC1* was initially designated to have 68 pathogenic changes; 58 CSNBs, 2 splice variants and 8 indels. Overall this totalled 371 pathogenic alleles, resulting in an estimated carrier rate of 2.09% and an estimated incidence of 1/9160. If these findings are superimposed onto a model of NPC1 along with all previously published variants (**Figure 2.3**) it can be seen that, although variants are found throughout the protein, the majority of pathogenic variants are found in transmembrane domains 3-13 and the third luminal loop. These regions contain the SSD and is in agreement with the previous report, which stated that a third of mutations are found between residues 855 and 1098 [84].

NPC1 has a previously reported incidence of 1/104,000 and there is a large discrepancy between this value and the incidence estimated from the datasets. Therefore manual curation was applied in order to discover the cause of the difference.

Of the 371 predicted pathogenic alleles, 254 (68%) resulted from 4 distinct variants.

- c.665A>G, p.N222S (0.400%)
- c.1532C>T, p.T511M (0.287%)
- c.2882A>G, p.N961S (0.389%)
- c.3598A>G, p.S1200G (0.355%)

The most commonly seen mutation in *NPCI* is p.I1061T (c.3182T>C), which has a frequency of 0.028% in the four datasets. The four high frequency variants are present at over 10 times this value and therefore it is not likely that these variants are associated with the classical form of NPC1. If these variants are excluded an estimated incidence of 1/92,104 (estimated carrier rate: 0.659%) is found, which is significantly closer to the previously reported incidence. Therefore the pathogenicity of these variants was investigated further.

Both p.N222S (c.665A>G) and p.N961S (c.2882A>G) have been seen in patients.

p.N222S (c.665A>G) was found in two cases:

- A 35 year old patient with adult onset NPC1 in combination with the common p.I1061T (c.3182T>C) variant. The patient presented with variant filipin staining, hepatosplenomegaly and later manifested ataxia [99].
- In two teenage sisters, from the NIH cohort, who were diagnosed based on splenomegaly, and were noted as visceral-only cases. The variant was found in combination with p.C468G (c.1402T>G), which is predicted by PolyPhen-2 to be Probably Damaging.

p.N961S (c.2882A>G) has been reported in an adult-onset case in a compound heterozygous state with p.S666N (c.1997G>A) [181]. p.S666N (c.1997G>A) has a predicted pathogenicity of Probably Damaging. The patient had subclinical hepatosplenomegaly and lymphadenopathy. Death was caused by acute pulmonary embolism and myocardial infarction.

From these case studies, it is possible that p.N222S (c.665A>G) and p.N961S (c.2882A>G) are associated with a late-onset or atypical form of NPC1. If they are pathogenic and fully penetrant, the estimated incidence would range from 1/19,077 to 1/36,420.

The remaining two high frequency variants, p.T511M (c.1532C>T) and p.S1200G (c.3598A>G) are predicted by *in silico* analysis to be pathogenic; p.T511M (c.1532C>T) is predicted to be Probably Damaging by PolyPhen-2 and p.S1200G (c.3598A>G) is predicted to be Probably Damaging by PolyPhen-2 and SIFT and Functional (medium) by Mutation Assessor. However, although previously published, neither variant has been reported in a classical NPC1 patient:

- p.T511M (c.1532C>T) was reported by Millat et al. [166] as a novel non-pathological CSNB.
- p.S1200G (c.3598A>G) was reported in the ZOOM study in an “NPC uncertain” patient. It was found in a compound heterozygous state with p.V664M (c.1990G>A), a known *NPC1* mutation [182]. However, plasma cholestane-3 β , 5 α , 6 β -triol, a biomarker for NPC [183], testing was negative.

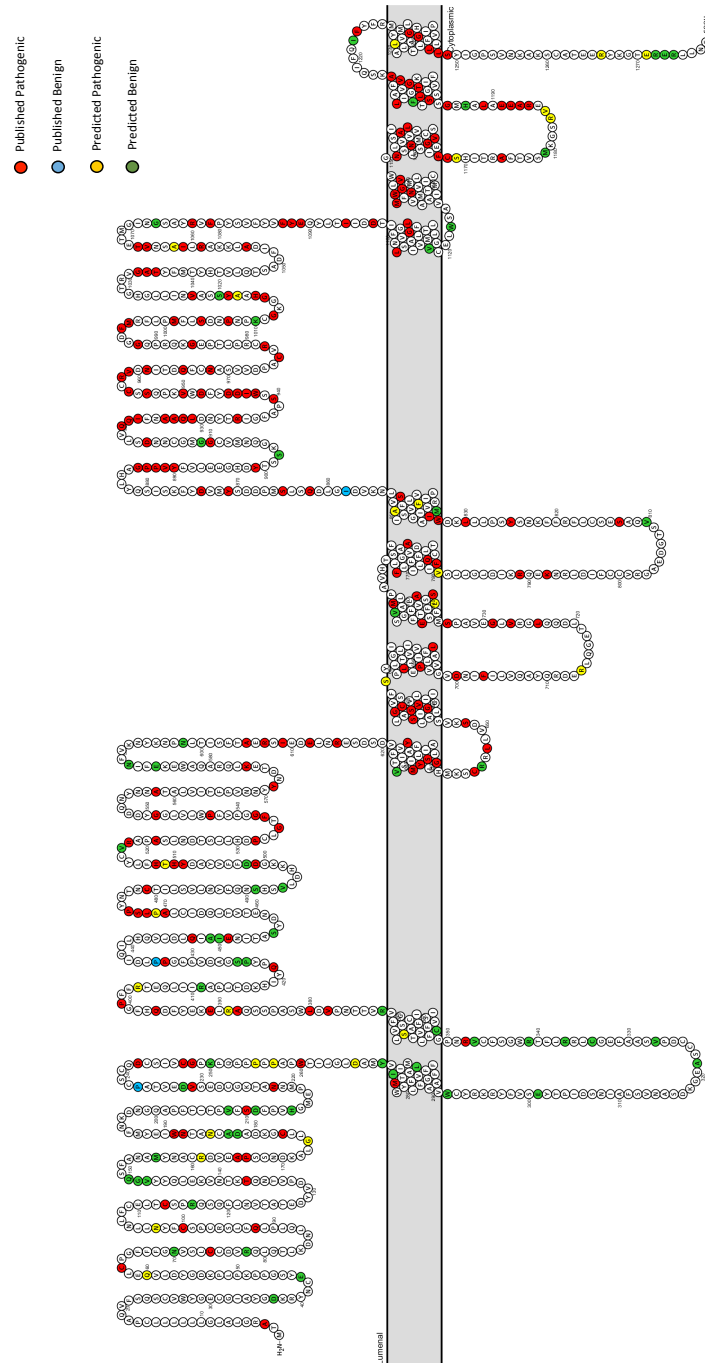


Figure 2.3: Model of the NPC1 protein using the program “Protter”

This figure shows the position of published pathogenic variants (red), published benign variants (blue), variants that are predicted pathogenic (yellow) and those that are predicted to be benign (green). The majority of the pathogenic variants are located within transmembrane domains 3-13 and the 3rd luminal loop.

NPC2

In *NPC2*, 151 pathogenic alleles were found, out of a total of 17,754 alleles. This resulted in a carrier rate of 0.85% and an estimated incidence of 1/55,297. As the fraction of *NPC2* cases has been reported to be between 3-5% of total cases (resulting in an incidence of approximately 1/2,000,000), this estimated incidence is too high and therefore manual curation was applied to discover why this was the case.

As with *NPCI*, particular variants stood out with significantly higher frequencies:

- c.441+1G>A
- c.88G>A, p.V30M

Both of these variants were reported in HGMD.

c.441+1G>A

c.441+1G>A was predicted to be Strongly Negative (strong pathogenic potential) by MaxEntScan. Wild-type cell lines along with cell lines heterozygous for this variant (NIMH 42 and NIMH 77) were identified. RNA was extracted and reverse transcribed to cDNA. When the cDNA was run out on a gel, only one band was seen for the wild-type cell line. In contrast, several bands were seen in NIMH 42 and NIMH 77 (**Figure 2.4A**). It was predicted that this was the result of multiple splicing events, which caused the insertion of additional amino acids. The most prominent of these involved an insertion of 16 bases, resulting in a frame shift that caused changes in the 4 terminal amino acids and addition of 86 amino acids onto the C-terminal (**Figures 2.4B and 4C**). However, there were several lines of evidence that suggested that this would result in a non-pathogenic protein:

- Modelling with I-TASSER shows that this variant does not affect the cholesterol binding pocket or protein stability
- NPC2 has been fused with the mCherry protein [184]. This results in a fully functioning protein that has correct trafficking and is able to rescue the cholesterol storage phenotype in a Balb/c *Npc2*^{-/-} MEF line.
- The variant has not been reported in an NPC2 patient

Taking into account these results, it was concluded that this c.441+1G>A was non-pathogenic.

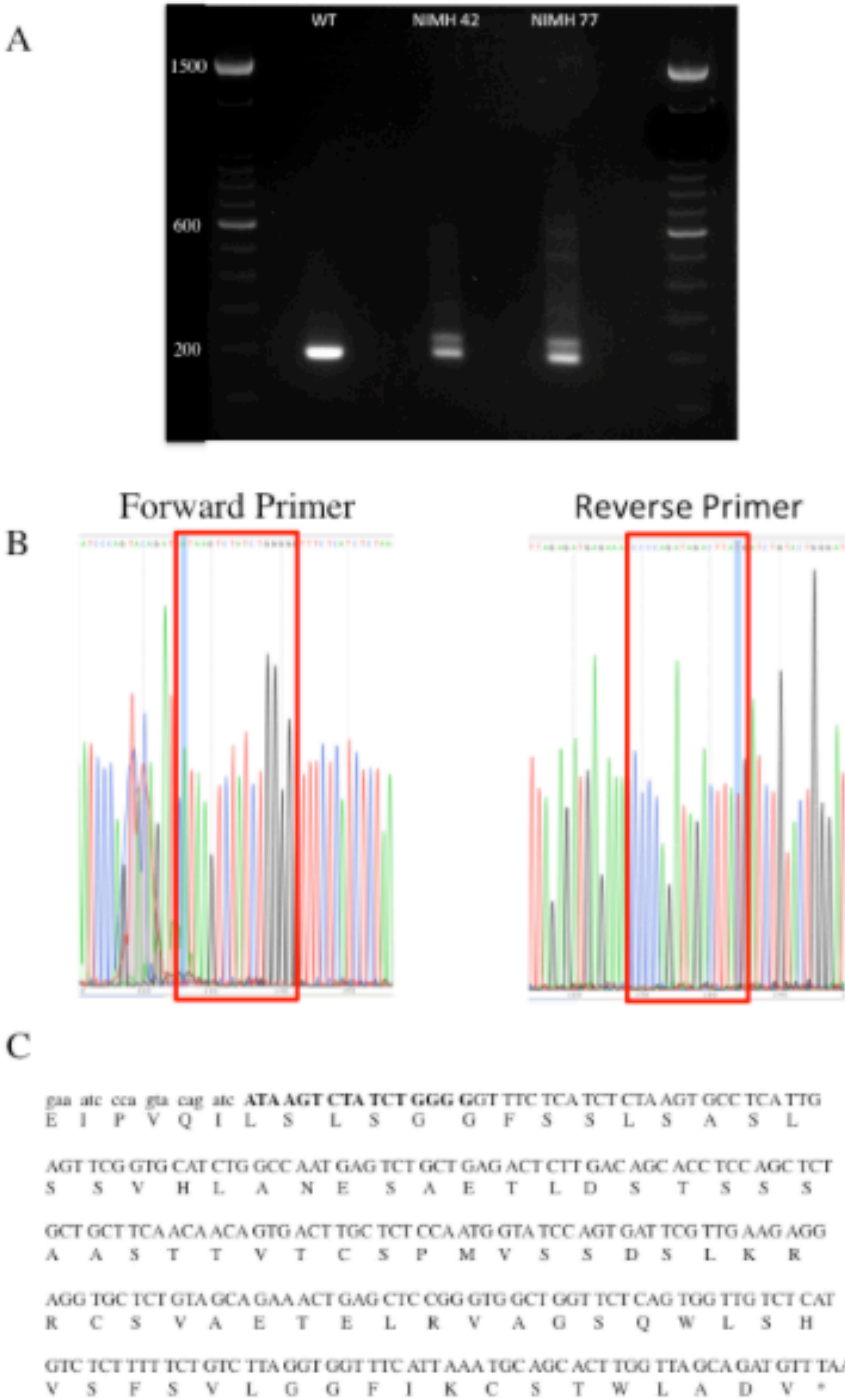


Figure 2.4: Analysis of the *NPC2* c.441+1G>A variant

- (A) Gel showing one band for the control (WT) cell line and multiple bands for the two heterozygous cell lines, NIMH 42 and NIMH 77.
- (B) Sequencing of the most prominent alternative splicing event. The red box designates the 16 additional nucleotides.
- (C) Sequence resulting from 16 base insertion (nucleotides in bold)

p.V30M (c.88G>A)

p.V30M (c.88G>A) was predicted to be Possibly Damaging by PolyPhen-2 and SIFT, but non-pathogenic by Mutation Assessor. Further analysis was carried out to confirm the pathogenic state of this variant.

- One patient with a variant NPC2 phenotype had been reported with the p.V30M (c.88G>A) variant, but a second mutation was not observed [185]. The patient also had near-normal levels of cholesterol esterification in fibroblasts.
- Sequence alignment and structure analysis suggested that p.V30M (c.88G>A) is not evolutionarily conserved and is in a structurally variable place away from the cholesterol binding pocket.

In this study, p.V30M (c.88G>A) is present at a higher frequency than any other known *NPC2* allele and if it is included in the calculation, the estimated incidence is 1/402,400 (18.6% of total NPC patients). Taking these factors into account, it was concluded that p.V30M (c.88G>A) was non-pathogenic.

Overall for *NPC2*, we were left with 21 pathogenic alleles (**Figure 2.5**) (0.118%), which resulted in an estimated incidence of 1/2,858,998.

The combined predicted incidence of NPC1 and NPC2 is 1/89,229, with NPC2 comprising 3.1%. This is slightly higher than the 1/104,000 cases described by Vanier, but the fraction of NPC2 cases is in agreement with previous reports.

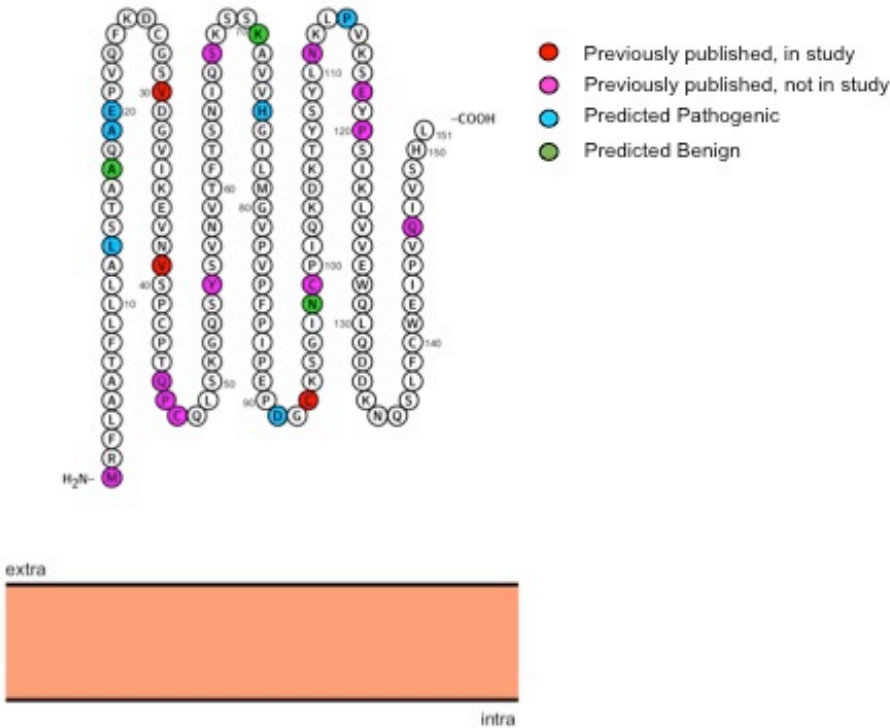


Figure 2.5: Model of the NPC2 protein using the program “Protter”

This figure shows the position of published pathogenic variants (red), published benign variants (blue), variants that are predicted pathogenic (yellow) and those that are predicted to be benign (green).

2.4 Discussion

The current reported disease incidence is based on clinical ascertainment but if atypical or pre-natal lethality is taken into account, this could be an under-ascertainment of true incidence. By utilizing massively parallel sequencing data sets, the incidence of SLOS and NPC was estimated using both bioinformatic tools and manual curation. Although, the apparent absence of statistics is a valid concern, there were multiple internal statistical approaches that were used in the generation of the final data counts. The GATK best Practices pipeline [186] was used in the mutation calling all of the exome datasets, which involved imposing filtering limits on the QUAL scores and other statistical tests, such as Fisher's Strandedness testing. This enabled false positives to be eliminated and therefore future analysis only dealt with reliable calls. Furthermore, the impact scoring software used for annotation, such as SIFT and Polyphen-2, also contained internal statistical analyses.

With respect to both SLOS and classical NPC1, the results do not suggest significant under-ascertainment of patients. The previously reported clinical incidence for SLOS is 1/40,000 and the incidence estimated from our study is between 1/34,937 and 1/39,215. With respect to classical NPC1, the estimated incidence was 1/92,000 for NPC1 and 1/2,900,000 for NPC2, resulting in a combined incidence of 1/89,000. This is not significantly under the clinical estimate of 1/104,000.

Based on the data, we hypothesised that there may be significant under-ascertainment of a late-onset form of NPC1. The clinical significance and phenotype are not fully delineated but previous case studies suggest a visceral-only or neurologically mild

phenotype. The two variants responsible for this atypical NPC1 phenotype (p.N222S (c.665A>G) and p.N961S (c.2882A>G)) have a combined carrier frequency of 0.8%. The clinical relevance is strongest for p.N222S (c.665A>G) as two independent cases with similar visceral and delayed neurological manifestations with variant filipin staining have been observed. If this variant is fully-penetrant and present in a compound heterozygous state with another pathogenic NPC1 variant, the incidence of NPC1 is estimated to be 1/35,667. If p.N961S (c.2882A>G) is included, the incidence increases to 1/19,077. This is drastically higher than the clinically reported incidence. However, only one case study is available for p.N961S (c.2882A>G) and there is no supporting diagnostic testing.

On the other end of the scale, the high-frequency variants that were regarded as non-pathogenic (p.T511M (c.1532C>T) and p.S1200G in *NPC1* and p.V30M (c.88G>A) in *NPC2*) may not have been clinically reported due to prenatal lethality. However, as NPC is an autosomal recessive disease, it is unlikely that there will be a situation that this variant results in solely prenatal lethality, such as dominant inhibitory function.

Although extremely powerful, estimating incidence using massively parallel-sequencing data sets is not without limitations. The number of alleles detected relies on the quality of sequencing and depth of coverage and there are multiple alignment and calling algorithms that can impact the number of variants called per sample. Additionally, exome sequencing does not always capture the whole exome and, correspondingly, we did not detect every variant in every allele. We also only detected 25 out of the 164 known SLOS variants (15.2%), although many of the common variants were detected. This could be due to the depth of coverage, which varies between sequencers and data

sets. The 1000 genomes project sequences to a depth of 4X resulting in detecting only variants that are present at a frequency of 1% or higher. This means exceptionally rare variants may be missed. However, the other data sets in this study sequenced to a greater depth of coverage, increasing the detection of rare variants.

While *in silico* analysis was mainly in agreement with published pathogenic alleles, additional manual curation was required to determine pathogenicity. For example, the variant p.V134L (c.400G>T) in SLOS was predicted by PolyPhen-2 as “Benign” but was previously reported in a patient and therefore was included as pathogenic. However, no functional analysis was included so it is possible that the variant is benign. Conversely, it is possible that variants predicted to be benign via *in silico* analysis might be subsequently found to be causative.

The 1000 genomes project, and to a lesser extent the ESP dataset, included ethnicity data, which enhanced the ability to assess the frequencies of variants within ethnic groups. The advantage of this was most evident in the analysis of the prevalent SLOS variant, c.964-1G>C. The previous carrier rate of this variant was based on populations of Northern European descent and it was only when analysis was restricted to this population that a similar rate was achieved. The lack of ethnicity data does not diminish the power of the data sets but exemplifies that inclusion of this data would increase versatility. As the cost of sequencing decreases, it is likely that this data will be more accessible and that the data sets will become even larger and more powerful tools.

Aside from estimating incidence, other important clinical data may be collected from observing the frequency of variants. For example, p.H215R (c.644A>G) is found to be

present in almost a third of NPC1 alleles and has been shown by multiple groups to be associated with obesity.

SLOS and NPC were used in this study to exemplify how massively parallel data sets can be used in general health for rare disorders and the main focus of this study was to predict disease incidence to determine if there is under-ascertainment of cases. Although a novel and interesting study, the volume of manual curation and molecular testing required limits the practicality for general health assessment. The potential discovery of a higher prevalence of late-onset NPC1 does highlight the importance of using genetics to aid the accurate assessment of incidence of disorders. If clinical reports are used alone, many patients could be misdiagnosed or missed due to prenatal lethality or atypical presentation, resulting in a lower incidence. In order to most effectively treat and improve lives for these patients, correct diagnosis is critical.

The availability of sequencing has remarkably improved over recent years and future development of these bioinformatic approaches could allow more accurate genetic risk assessments and aid diagnosis.

Chapter 3: Analysis of the Behaviour and Neurological

Structure of *Dhcr7*^{T93M/Δ3-5} mice

3.1: Introduction

The path to therapeutic intervention is a multi-step process that usually involves testing a drug in animal models before proceeding to clinical trials. Drug efficacy in an animal model can be measured using both invasive and non-invasive techniques. One of the most common non-invasive techniques is to study behaviour. As stated in the **Introduction**, dietary cholesterol supplementation is a standard of care for SLOS patients but the effects on behaviour are controversial, with reports of improvements in most aspects of the behavioural profile in some cases but not all [124]. Additionally, cholesterol does not cross the BBB and, as such, it is uncertain how cholesterol supplementation would improve the cognitive and neurological aspects of the disorder [124]. Consequently, a therapy that can cross the BBB, thus impacting the CNS, needs to be developed. This would then have the potential to improve the cognitive and behavioural aspects of the disorder.

SLOS has a very specific biochemical hallmark, namely increased levels of 7DHC and decreased levels of cholesterol, but the clinical phenotype is variable. As Dykens [187] noted, “a behavioural phenotype is not a set behavioural pattern but more the heightened probability or likelihood that people with a given syndrome will exhibit certain

behavioural and developmental sequelae relative to those without the syndrome”.

Despite this variation, assessing the behavioural profile can aid diagnosis in mild SLOS cases, which present with a subtle physical phenotype. Typical behaviours observed in SLOS patients are [69]:

- Hyperactivity
- Sleep-cycle disturbance
- Hypersensitivity and hyperacusis
- Irritability
- Increased anxiety
- Increased aggression and self-injurious behaviour.
- Autistic traits, including ritualistic behaviour
- Obsessive compulsive behaviours

By observing if the behavioural phenotype is replicated in mouse models, this would allow more effective treatments to be developed and tested. Two mouse models of SLOS (*Dhcr7*^{Δ3-5/Δ3-5} and *Dhcr7*^{tmlGst/J}) present with neonatal lethality [108,188,189] and, therefore, are unsuitable for behavioural study. Heterozygous mice (*Dhcr7*^{+/-}) of the latter model (*Dhcr7*^{tmlGst/J}) are viable and one study used them to assess the behavioural phenotype [188]. This study used a battery of behavioural tests, including the elevated zero maze, social dominance testing, marble burying, Y maze and rotor rod testing. Overall there were not many differences between controls (*Dhcr7*^{+/+}) and heterozygous (*Dhcr7*^{+/-}) mice but there are a few points to note:

- Both genotypes displayed greater anxiety in bright compared to dim light on the elevated zero maze (a common test for anxiety).

- 6-month old *Dhcr7*^{+/-} females displayed greater social dominance than controls but overall this difference was small.
- *Dhcr7*^{+/-} had normal locomotor ability but no difference compared to controls in entries into Y maze, suggesting intact spatial working memory.
- On the rotarod, *Dhcr7*^{+/-} females had superior performance compared to controls.
- Generally there was no difference in number of marbles buried, although males tended to partially bury more marbles than females, irrespective of genotype.

The limited differences in behaviour are not surprising, as the study used SLOS *Dhcr7*^{+/-} mice. Parents of SLOS patients have minimal or no clinical abnormalities, although they have slightly elevated plasma 7DHC levels [190]. Similarly, Wassif *et al* found that *Dhcr7*^{+/ Δ 3-5} mice had slightly elevated 7DHC levels in tissue [108] and Correa-Cerro *et. al* reported that *Dhcr7*^{T93M/+} mice are phenotypically normal [110]. As such, it is likely that there is minimal heterozygotic effect and suggests that in order for the SLOS behavioural phenotype to be observed in the mouse, a homozygous model is required. The severe homozygous *Dhcr7* ^{Δ 3-5/ Δ 3-5} model dies soon after birth (P1) due to failure to thrive caused by lack of feeding [108]. Consequently, this model is not suitable for behavioural studies. However, similar to the *Dhcr7*^{+/-} heterozygous mice, *Dhcr7*^{T93M/ Δ 3-5} mice live into adulthood and are suitable for behavioural studies.

Although biochemical and limited histology studies have been performed on *Dhcr7*^{T93M/ Δ 3-5} mice [110], their behavioural phenotype has not been investigated. As anxiety is a trait seen in SLOS patients [69] and was investigated in the previous study

using heterozygous mice, the *Dhcr7*^{T93M/Δ3-5} mice were tested using the elevated plus maze. The elevated plus maze originated from the observation in 1955 that rats preferentially explored closed areas of an elevated Y maze compared to open areas [191]. Montgomery proposed that the novel environment of the closed and open arms evoke both the fear and exploratory drives but the open arms cause a stronger fear response due to risk of predation. This results in greater exploration of the closed arms by healthy rodents [191]. A further study also supported this hypothesis as control rats avoided open arms of an elevated plus maze, but additionally found that rats treated with anxiolytic drugs had a significantly increased proportion of entries into the open arms [192]. This was shown to also be the case in mice [193].

Other techniques used to measure anxiety include assessing burrowing and Nestlet shredding [194]. Laboratory mice have been shown to remove food pellets placed in a 'burrow' in the cage and these behaviours have been suggested to represent several factors [195]. Defensive burying is a known paradigm of anxiety and the burrowing could represent this need for shelter and safety by providing the mice with a place to hide [195]. Burrowing could also be a form of tunnel maintenance with the food pellets representing a blockage in the tunnel [195]. Decreased burrowing has been shown to be associated with Scrapie disease, a neurodegenerative prion disorder [196].

As with burrowing, nest construction is also a natural behaviour of mice. Nesting is thought to represent shelter, aid reproduction and survival of mouse pups and help retain body heat [197]. Despite the reproductive nature of nesting, there is not a discernable difference between nest building of male and female mice [198]. Previous studies have used a variety of materials for nest construction, including salt marsh hay [199], paper

strips [200] and cotton dental balls, before using Nestlets (2" square nesting material composed of pulped cotton fiber, sold by Ancare ©) became the standard procedure. However, even then there was some debate about the optimum scoring method and scoring systems, ranging from a 2-point scale to an 8-point scale, have been used [198,200–202]. Although useful, this scoring system is subjective and therefore a quantifiable measure is often conjointly used, such as weighing the un-shredded Nestlet after the completion of the test [197]. Nest construction is decreased in mice with lesions of the hippocampus [203], medial preoptic area [200] and septum [201].

Dhcr7^{T93M/Δ3-5} mice have been observed to have functional locomotor abilities (J.Cross, unpublished data). Consequently the open field test is a potentially useful task [204] and can assess if the hyperactivity phenotype observed in patients is also present in the *Dhcr7*^{T93M/Δ3-5} model. The open field is also a further measure of anxiety due to the fact the mice tend to stay towards the edge of the field and do not venture into the open centre [205].

3.1.i: Aims of Chapter 3

The aims of this chapter are to:

- Assess the anxiety and hyperactivity behavioural phenotypes in the *Dhcr7*^{T93M/Δ3-5} mouse model and the extent to which they correspond to the patient phenotypes.
- Further investigate the neuroanatomical changes using histology and discover if there are links to behaviour

3.2: Materials and Methods

3.2.i: Mice

Control (*Dhcr7*^{+/+}) and age-matched *Dhcr7*^{T93M/ Δ 3-5} mice on a mixed BL6/129 background, were analysed at 3, 4 and 5-month time points (+/- 3 days). At each time point, the mice were analysed on five behavioural tests; burrowing, Nestlet (Ancare ©) shredding, marble burying, elevated plus maze and open field.

3.2.ii: Burrowing test

“Burrows” were constructed out of plastic laboratory wash bottles (250ml polyethylene wash bottle, cat.# 6640002998493, NIH supply centre). Clips were used as stands at the burrow entrance to prevent rolling and to give the burrow a downward slope. 200g of food pellets were placed into the burrow, and the burrow was positioned at the front right corner of each clean cage (**Figure 3.1**). At approximately 3.30pm (2.5 hours before the dark housing cycle), a single mouse was placed into each cage and observed for 2 hours. Behavioural notes were made and the pellets remaining in the burrow were weighed at the conclusion of the 2 hours. This is to observe if the mice used are vigorous burrowers that empty the entire 200g in 2 hours and to achieve familiarity with the burrow before the overnight step [195]. The pellets were replaced into the burrow and the burrow was returned to the same position in the cage. The cages were put back onto the rack and left overnight. The following morning (between 9am and 11am, 3-5 hours after the start of the light cycle), the food pellets remaining in the burrow were

weighed to assess extent of burrowing. The exact time of measurement is not critical as the mice are usually asleep during the light cycle and there is limited burrowing activity [195]. It was observed that the mice do not eat a significant amount of pellets (J. Cross, unpublished data). After each test, the burrow was cleaned with 70% ethanol to remove any odours that may affect subsequent tests.



Figure 3.1: Burrowing test setup

This image shows an example cage depicting how the burrow was set up. The burrow was filled with 200g of food pellets and placed in a clean cage against the wall. This was the same position for every cage.

3.2.iii: Nestlet test

Mice were separated into individual cages, containing food, water and sawdust bedding but no Nestlet, and were left to acclimatise for at least 3 hours. At approximately 4.30pm (1.5 hours before the lights were turned off), a Nestlet (Ancare ©) was weighed and put into the cage. The cages were placed back onto the rack and left undisturbed for 30 minutes. The Nestlet was then scored for the amount shredded and the remaining un-shredded Nestlet was weighed (**Table 3.1**). If the Nestlet was partly shredded but a nest was made, a score of 5 was given.

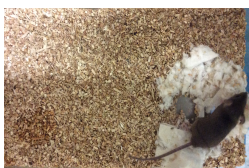
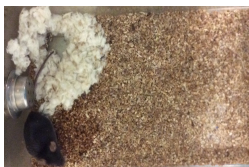
Score	Description	Image
1	No Shredding 0 - 1% shredded	
2	Slight Shredding 1 - 45% shredded	
3	Partial Shredding 45 - 90%	
4	Fully Shredded no nest (flat) 90 - 100% shredded	
5	Fully Shredded nest made (curved around mouse) 90 - 100% shredded	

Table 3.1 Nestlet shredding test setup and scoring**Table 3.1 (previous page): Nestlet shredding test setup and scoring**

The left-hand column designates the scoring system, ranging from 1-5 and a description of the score is denoted in the middle column. An example image is shown in the right-hand column. The Nestlet was placed in the centre of a clean cage, as shown in the Score 1 image. The cage lid was replaced and the cage put back onto the rack. The cage was then left for 30 minutes and scored. The Nestlet was replaced into the cage and the cage was left on the rack overnight before scoring a second time.

3.2.iv: Marble burying test

A clean cage was filled with approximately 10 cm depth of sawdust bedding and 12 green glass marbles (~1 cm diameter) were evenly spaced on top of the sawdust (**Figure 3.2**). Between 9am and 11am (3-5 hours after the start of the light cycle), a single mouse was placed into the cage. The cage was covered with an opaque cloth and left undisturbed for 30 minutes. The mouse was then removed from the cage and the number of marbles buried were scored. A buried marble was counted as a marble that was $\frac{3}{4}$ or more covered in sawdust. After each test, the marbles were cleaned with 70% ethanol to remove any odours that may affect subsequent tests.

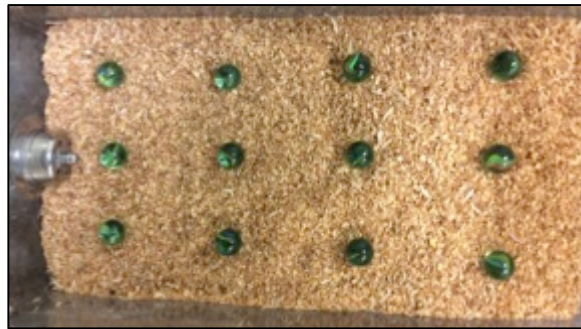


Figure 3.2: Marble burying test setup

An example cage showing the setup of the marble burying test. Twelve marbles were evenly spaced in a clean cage containing ~10cm of sawdust. A single mouse was placed in the cage and the cage was covered and left for 30 minutes before scoring.

3.2.v: Elevated plus maze

The maze was elevated approximately 1.25 meters off the ground (**Figure 3.3**). A camera (Microsoft HD) was positioned on the ceiling above the maze and connected to a laptop computer. The software used was AnyMaze[®], and set up for the elevated plus maze test as per manufacturers instructions. The mice were brought into the room, cages covered and the mice were allowed to acclimatise for 30 minutes. The mouse to be evaluated was placed in the middle of the maze and left to explore the maze for 5 minutes. The movements of the mouse were tracked and recorded by the camera and software. If the mouse fell off, it was replaced quickly and a note was made. At the end of the test, the mouse was returned to the cage. After each test, the elevated plus maze was cleaned with 70% ethanol to remove any odours that may affect subsequent tests. Measurements were taken within the first half of the light cycle.

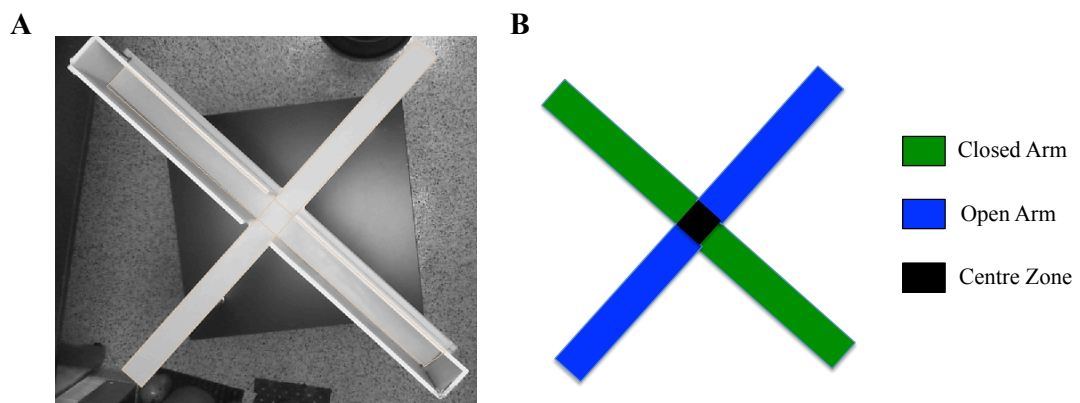


Figure 3.3: Elevated Plus Maze setup

- A) Image showing the elevated plus maze camera view.
- B) Schematic showing the location of the closed (green) and open (blue) arms of the elevated plus maze. The central zone is shown in black. The orientation of the arms relates to the maze shown in image A.

3.2.vi: Open field

A camera (Microsoft HD) was positioned on the ceiling above the open field and connected to a laptop computer. The open field was a 60 cm x 60 cm homemade plywood square arena (**Figure 3.4**). The edge zone was 10 cm diameter and the centre square was 20 cm diameter. The edge zone represents a protected area and the centre zone represents the most exposed area of the field. The software used was AnyMaze[®], and set up for the open field test as per manufacturers instructions. A single mouse was placed into the cage at the same position each time and allowed to explore for 15 minutes. The movements of the mouse were tracked and recorded by the camera and software. At the end of the test, the mouse was returned to the home cage and the field was cleaned with 70% ethanol to remove any odours that would affect subsequent tests. Measurements were taken within the first half of the light cycle.

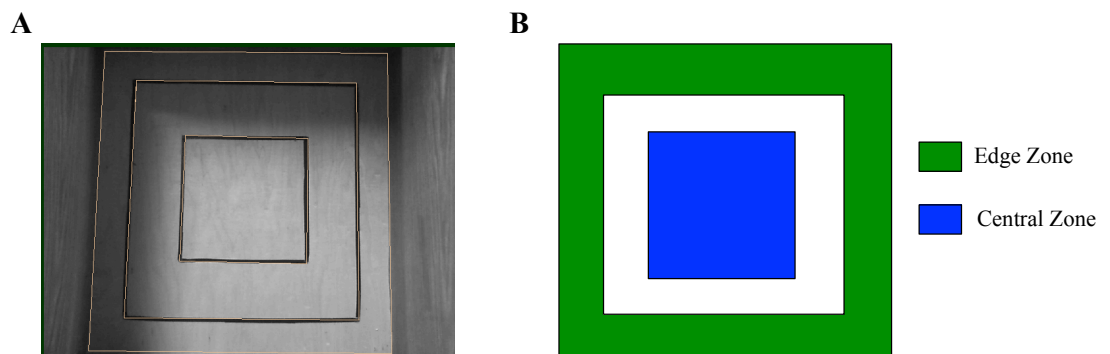


Figure 3.4: Open Field setup

- A) Image showing the open field camera view.
- B) Schematic showing the location of the edge (green) and central (blue) zones of the open field. The mouse is tracked across the whole field but only entries into the edge (green) and central (blue) zones were measured. The orientation of the zones relates to the open field shown in image A.

3.2.vii: Histology

Cerebral cortex from control and *Dhcr7*^{T93M/Δ3-5} mice were PFA perfused and separated from the cerebellum and mounted onto a metal, circular block. Prior to mounting, the block was covered with water, left to freeze and cut on the cryostat to gain a flat surface. The cortex was placed onto the block (coronal side separated from the cerebellum face down) and left at ~ -40°C for 5 minutes to adjust to the temperature. The cortex was then cryo-protected with ddH₂O and left to freeze at ~ -40°C. The block was then placed into the cutting space on the cryostat and left to adjust (~ -10°C) for 10 minutes. Brains were cut into 20μm coronal sections (Cryostat, Bright Model OTF) and serially divided into 5 wells filled with 1ml 1X PBS with 0.02% Triton-X.

Sections were mounted onto Superfrost Plus (cat.# J1800AMNZ ,ThermoScientific,) slides, covered and left to dry overnight. The slides were then stained using hematoxylin (Harris Hematoxylin Solution, cat.# HHS16, Sigma Aldrich) and eosin (Eosin Y solution, cat.# HT110116, Sigma Aldrich) using the following protocol:

- 3 minutes in hematoxylin (filtered (filter paper) Harris Hematoxylin Solution, cat.# HHS16, Sigma Aldrich)
- Rinsed in tap water until water ran clear
- 3 minutes in acid alcohol (1% concentrated hydrochloric acid in 95% ethanol)
- 1 minute soak in tap water
- 1 minute in Bluing Agent (1% magnesium sulphate, 0.07% sodium bicarbonate in 300ml tap water)
- Quick rinse in tap water and 95% ethanol

- 30 seconds in eosin (Eosin Y solution, cat.# HT110116, Sigma Aldrich)
- 30 seconds in 95% ethanol
- 3 minutes in 100% ethanol
- 10 seconds in HistoClear (cat.# HS-200, National Diagnostics)
- Coverslip with Histomount (cat.# HS-103, National Diagnostics)

Slides were imaged using Zeiss Axio Imager A.1 microscope and AxioVision Release 4.8.2 SP2 software. Measurements were taken using ImageJ (version 2.0.0, 2010).

3.2.viii: Statistical analysis

GraphPad Prism (Prism 6 Mac OS X, version 6.0b) was used for statistical analysis. A two-way ANOVA was implemented to ascertain interaction between gender and genotype. As there was no interaction, a multiple student t-test was used for the histological analysis, elevated plus maze and open field tests. A Mann-Whitney U test was used for the Nestlet shredding, burrowing and marble burying tests, as stated by Deacon [195,197]. Graphs show the mean and standard error of mean (SEM).

3.3: Results

3.3.i: Burrowing test

Rodents naturally have burrowing habits and laboratory mice were shown to readily burrow food pellets from a tube placed in the cage [206]. Using a previously outlined protocol as a basis [195], cages were set up with sawdust, water and a burrow containing 200g of food pellets. At approximately 3.30pm (2.5 hours before the start of the dark phase of the light cycle), mice were individually placed into the cages and observed for 2 hours.

The data was subjected to a two-way ANOVA that examined the effect of genotype and gender on burrowing. At both 2 hours and overnight, there was no significant interaction between gender and genotype (**Table 3.2**). As such, the data was further analysed via a Mann-Whitney U test.

	Time Point	Interaction	Significance
2 hours	3 month	$F(1,35) = 0.12$	No, $p = 0.7365$
	4 month	$F(1,51) = 0.19$	No, $P = 0.6641$
	5 month	$F(1,32) = 1.45$	No, $p = 0.2375$
Overnight - 2 hours	3 month	$F(1,35) = 1.29$	No, $p = 0.2645$
	4 month	$F(1,48) = 0.20$	No, $p = 0.6567$
	5 month	$F(1,32) = 0.28$	No, $p = 0.6025$

Table 3.2: 2-way ANOVA interaction scores for burrowing

A two-way ANOVA was implemented to discover the effect of genotype and gender on burrowing. At all time points there was no significant interaction between genotype and gender. Therefore, further analysis could be implemented via a Mann-Whitney U test.

Overall, at 3 months of age mice did not burrow a significant amount in 2 hours for either gender or genotype (**Figure 3.5A**). This was also true at 4 months, although the amount burrowed was slightly higher, especially for female *Dhcr7*^{T93M/ Δ 3-5} mice. One or two mice of each gender and genotype burrowed an exceptionally high amount in 2 hours. These mice tended to consistently burrow at unusually high levels over the time course compared to the other mice in the study. At 5 months, more variation occurred with 3 - 5 mice of each genotype and gender burrowing up to 200g but the remaining mice burrowing little. The mice responsible for the higher values tended to have been exposed to the burrow before, suggesting a learning phenotype or decreased anxiety of the novel object.

A ceiling effect was often reached for the overnight burrowing, with mice burrowing the entire contents of the burrow (**Figure 1B**). This was observed most at 5 months. At 3 months, there was no discernable difference between female mice, although *Dhcr7*^{T93M/ Δ 3-5} male mice trended to burrow more than their controls. However, at 4 months, the *Dhcr7*^{T93M/ Δ 3-5} mice burrowed significantly more in both genders. By 5 months, the majority of mice burrowed 60% or more of the food pellets. As with the 2-hour study, the mice that were exposed to the test for each time point tended to be responsible for the higher percentages of pellets removed from the burrow.

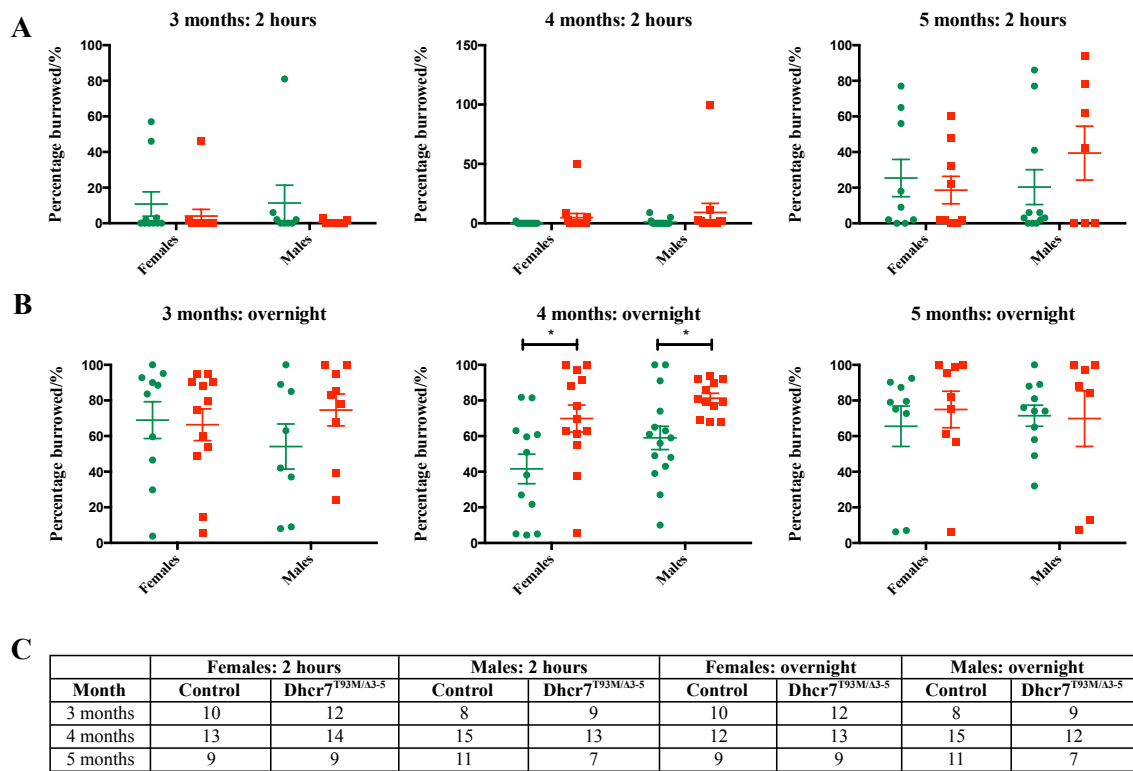


Figure 3.5: Burrowing test

Figure 3.5 (previous page): Burrowing test**A) Graphs showing the amount of food pellets removed from the burrow in 2 hours as a percentage of the total 200g.**

There was minimal burrowing at 3 (females: $p = 0.2486$, males: $p = 0.2177$) and 4 months (females: $p = 0.1474$, males: $p = 0.1777$) in control (green) or *Dhcr7*^{T93M/Δ3-5} (red) mice of either gender. Burrowing increased in all genotypes and genders at 5 months of age (females: $p = 0.7123$, males: $p = 0.6672$).

B) Graphs showing the amount of food pellets removed from the burrow overnight as a percentage of the amount remaining after 2 hours.

At 3 months there was a large amount of variation in control (green) and *Dhcr7*^{T93M/Δ3-5} (red) mice of both genders (females: $p = 0.7604$, males: $p = 0.2864$). This variation continued over the time course, although there was a significant increase in the amount burrowed by the *Dhcr7*^{T93M/Δ3-5} mice compared to controls at 4 months (females: $p = 0.0133$, males: $p = 0.0071$). This significance disappeared by 5 months (females: $p = 0.4220$, males: $p = 0.4368$).

C) Number of mice analysed at each time point.

Across the time course, there are occasional differences in the numbers analysed for 2 hours and overnight. This is due to some mice being excluded due to a water bottle leak, which caused flooding of the cage

3.3.ii: Nestlet shredding test

The data was subjected to a two-way ANOVA that examined the effect of genotype and gender on Nestlet shredding. At both 30 minutes and overnight, there was no significant interaction between gender and genotype (**Table 3.3**). As such, the data was further analysed via a Mann-Whitney U test, as described by Deacon in 2006 [197].

	Time Point	Interaction	Significance
30 min	3 month	F(1,38) = 0.07	No, p = 0.7960
	4 month	F(1,58) = 1.37	No, p = 0.2472
	5 month	F(1,51) = 1.33	No, p = 0.2537
Overnight	3 month	F(1,36) = 0.46	No, p = 0.5000
	4 month	F(1,43) = 0.03	No, p = 0.8645
	5 month	F(1,42) = 1.13	No, p = 0.2933

Table 3.3: 2-way ANOVA interaction scores for Nestlet shredding

A two-way ANOVA was implemented to discover the effect of genotype and gender on Nestlet shredding. At all time points there was no significant interaction between genotype and gender. Therefore, further analysis could be implemented via a Mann-Whitney U test.

The first study was to measure the amount of Nestlet shredded within 30 minutes (**Figure 3.6A**). At 3 months there was a trend for *Dhcr*^{T93M/Δ3-5} mice to be less likely to build full nests than control male mice ($p = 0.1000$). The scoring system was a more accurate measure of nest building as it was observed that although the *Dhcr*^{T93M/Δ3-5} mice did shred the entire Nestlet in some instances, they tended to not build a full nest, the difference of a score of 4 or 5 (**Figure 3.6C**). Female *Dhcr*^{T93M/Δ3-5} mice scored significantly less than the female control mice at 4 months of age (females: $p = 0.0044$) and male *Dhcr*^{T93M/Δ3-5} mice scored significantly less than male control mice at 5

months of age (males: $p = 0.0322$). This significance disappeared at 5 months of age, although the trend for $Dhcr7^{T93M/\Delta3-5}$ mice to score lower than control mice continued for both genders over the time course. There was variation in all genders and genotypes but this was greater in the $Dhcr7^{T93M/\Delta3-5}$ mice. A few control mice scored at a consistently low rate across the time course and therefore it is possible that these mice were mis-genotyped. However, the tissue was not available to check this.

Both genotypes overall shredded the majority of the Nestlet overnight (**Figure 3.6B**). As with the 30-minute study, $Dhcr7^{T93M/\Delta3-5}$ mice scored lower than control mice at all time points. Using the Mann-Whitney U test, the decrease was significant in female $Dhcr7^{T93M/\Delta3-5}$ mice at 3 and 4 months (females: 3 months: $p = 0.0201$, 4 months: $p = 0.0054$) and significant in male $Dhcr7^{T93M/\Delta3-5}$ mice at 4 and 5 months (males: 4 months: $p = 0.048$, 5 months: $p = 0.0124$). There was no reduction in locomotor activity in $Dhcr7^{T93M/\Delta3-5}$ mice compared to controls, as indicated in the open field test later in the chapter.

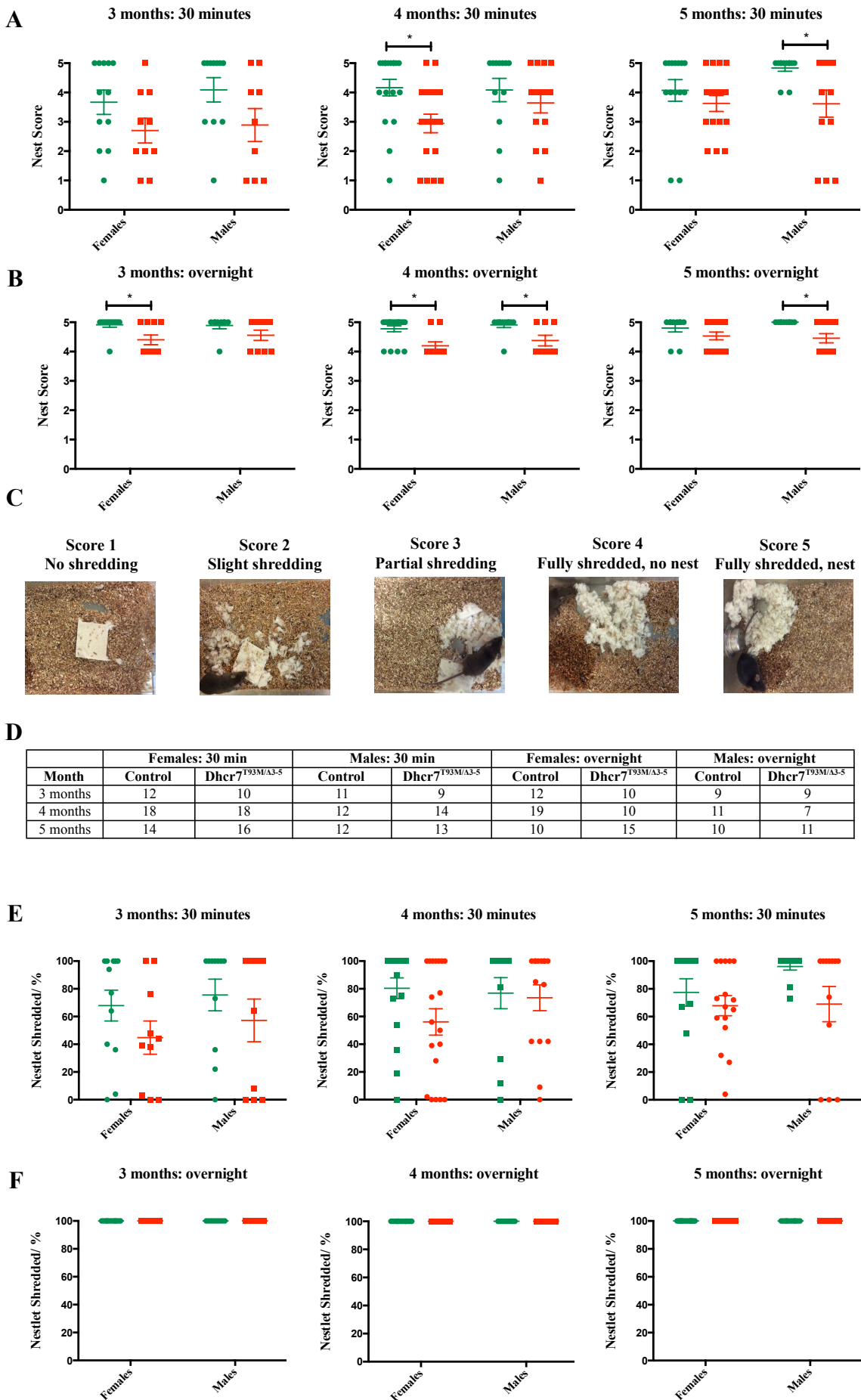


Figure 3.6: Nestlet shredding

Figure 3.6 (previous page): Nestlet shredding**A) Graphs showing Nestlet scores after 30 minutes.**

There was variation in control (green) and *Dhcr7*^{T93M/Δ3-5} (red) in both genders across the time course. The difference was not significant at 3 months (females: $p = 0.1208$, males: $p = 0.1000$) but was significant in female *Dhcr7*^{T93M/Δ3-5} mice at 4 months (females: $p = 0.0044$, males: $p = 0.2568$) and male *Dhcr7*^{T93M/Δ3-5} mice at 5 months (females: $p = 0.1427$, males: $p = 0.0322$).

B) Graphs showing Nestlet scores overnight

Across the time course, all mice shredded the entire Nestlet. However, *Dhcr7*^{T93M/Δ3-5} mice generally did not make a full nest from the Nestlet, causing a score of 4 instead of 5. This was significant in females at 3 months (females: $p = 0.0201$, males: $p = 0.1000$), both males and females at 4 months (females: $p = 0.0054$, males: $p = 0.0408$) and males at 5 months (females: $p = 0.2290$, males: $p = 0.0124$).

C) Scoring System

Representative images of the scoring system used to assess Nestlet shredding, ranging from 1 (left image) to 5 (right image).

D) Number of mice analysed at each time point.

Across the time course, there are occasional differences in the numbers analysed for 30 minutes and overnight. This is due to some mice being excluded due to a water bottle leak, which caused flooding of the cage.

E) Graphs showing amount of Nestlet shredded in 30 minutes.

There was variation in control (green) and *Dhcr7*^{T93M/Δ3-5} (red) in both genders across the time course and no significant differences in the amount of Nestlet shredded. (Females: $p = 0.1946$, males: $p = 0.8533$. 4 months: females: $p = 0.0793$, males: $p = 0.6953$. 5 months: females: $p = 0.2266$, males: $p = 0.0974$).

F) Graphs showing amount of Nestlet shredded overnight.

Mice of all genders and genotypes shredded 100% of the Nestlet overnight. From the weight of shredded Nestlet, there was no way to distinguish between mice that constructed a nest and those that did not. p -values for all mice were >0.9999 .

3.3.iii: Marble Burying

The data was subjected to a two-way ANOVA that examined the effect of genotype and gender on marble burying. At all time points, there was no significant interaction between gender and genotype (**Table 3.4**). As such, the data was further analysed via a Mann-Whitney U statistical test.

	Time Point	Interaction	Significance
Marble Burying	3 month	F(1,38) = 0.01	No, p = 0.9272
	4 month	F(1,27) = 1.53	No, p = 0.2267
	5 month	F(1,19) = 0.78	No, p = 0.3883

Table 3.4: 2-way ANOVA interaction scores for marble burying

A two-way ANOVA was implemented to discover the effect of genotype and gender on marble burying. At all time points there was no significant interaction between genotype and gender. Therefore, further analysis could be implemented via a Mann-Whitney U test.

Twelve marbles were evenly distributed in a clean cage and a single mouse was placed in the cage for 30 minutes. *Dhcr7*^{T93M/Δ3-5} mice appeared to bury slightly less marbles in 30 minutes, but this was only significant at 4 months in the male mice (males: $p = 0.0364$) (**Figure 3.7**). Only 3 control male mice were scored at this time point and could potentially explain the significant result. Therefore, more mice will need to be tested in order to establish if there was a significant difference.

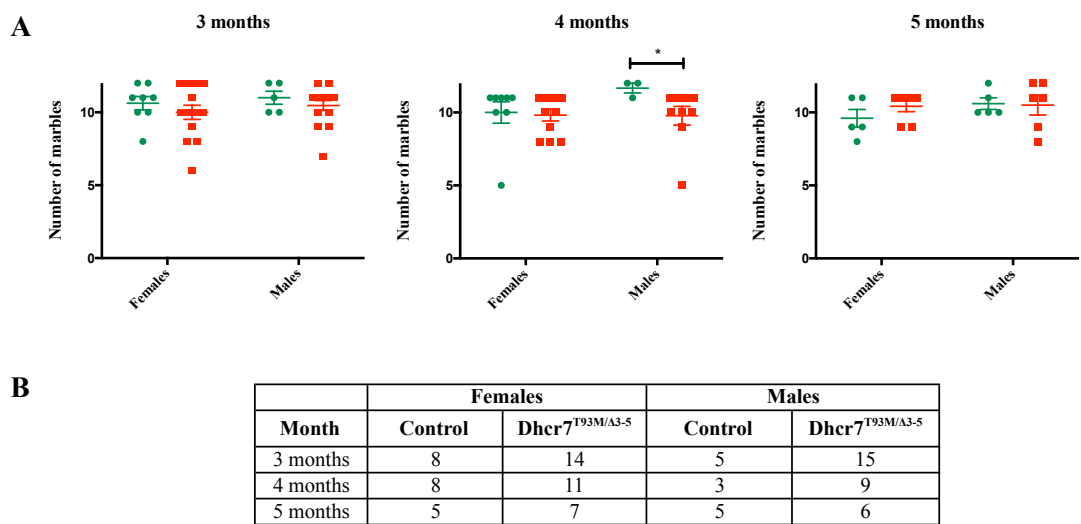


Figure 3.7: Marble burying

A) Graphs showing marble burying

Results were scored after 30 minutes. There was a slight tendency for *Dhcr7*^{T93M/A3-5} mice (red) to bury less than control mice (green). (3 months: females: $p = 0.4363$, males: $p = 0.5753$. 4 months: females: $p = 0.4419$, males: $p = 0.0364$. 5 months: females: $p = 0.4078$, males: $p = 0.9026$).

B) Number of mice analysed at each time point

3.3.iv: Elevated plus maze

The elevated plus maze is generally considered to be a measure of anxiety [193]. The maze consists of an elevated platform comprised of two open arms and two closed arms. Mice are anxious of open, elevated spaces and therefore have a tendency to stay within the closed zones. The time spent in the central zone between the arms is thought to represent risk-assessment or decision-making [207].

The data was subjected to a two-way ANOVA that examined the effect of genotype and gender on time spent on the open and central zones of the elevated plus maze. At all time points, there was no significant interaction between gender and genotype (**Table 3.5**). As such, the data was further analysed via a multiple t-test statistical test.

	Time Point	Interaction	Significance
Open Zone	3 month	$F(1,34) = 2.13$	No, $p = 0.1536$
	4 month	$F(1,38) = 0.01$	No, $p = 0.9317$
	5 month	$F(1,26) = 0.59$	No, $p = 0.4509$
Central Zone	3 month	$F(1,34) = 0.03$	No, $p = 0.8732$
	4 month	$F(1,38) = 0.49$	No, $p = 0.4882$
	5 month	$F(1,25) = 1.15$	No, $p = 0.2932$

Table 3.5: 2-way ANOVA interaction scores for the elevated plus maze

A two-way ANOVA was implemented to discover the effect of genotype and gender on time spent on the open and central zones of the elevated plus maze. At all time points there was no significant interaction between genotype and gender. Therefore, further analysis could be implemented via multiple t-test. However, although not significant, the interaction for the open zone at 3 months was under 0.2 (blue text).

There was a relatively large amount of variation between both genotypes, although female *Dhcr7*^{T93M/Δ3-5} mice spent significantly more time in the central zone than controls ($p = 0.0207$). Decision-making may be associated with the frontal cortex [208] and lesions in the frontal lobe have been seen in SLOS patients [68].

Overall, there was more variation in the time spent in the open arms in the *Dhcr7*^{T93M/Δ3-5} mice compared to control mice. At 3 months, female and male controls and male *Dhcr7*^{T93M/Δ3-5} mice spent less than 100 seconds (< 33% total test time) in the open zones (**Figure 3.8A**), whereas the female *Dhcr7*^{T93M/Δ3-5} mice spent significantly more time in the open zones (greater than 100 seconds of the total test time) (females: $p = 0.0151$). This continued to 4 months of age (females: $p = 0.0130$) but fewer female spent over 100 seconds in the open zones. At 4 months of age, all *Dhcr7*^{T93M/Δ3-5} male mice spent more time in the open zones than the male control mice. However, two male *Dhcr7*^{T93M/Δ3-5} mice spent a significantly longer time in the open zones than the other male *Dhcr7*^{T93M/Δ3-5} mice (4 months males: $p = 0.0120$). By 5 months of age, the time spent in the open zones was minimal for all genders and genotypes. This was also confirmed in the track plots (**Figure 3.8B**).

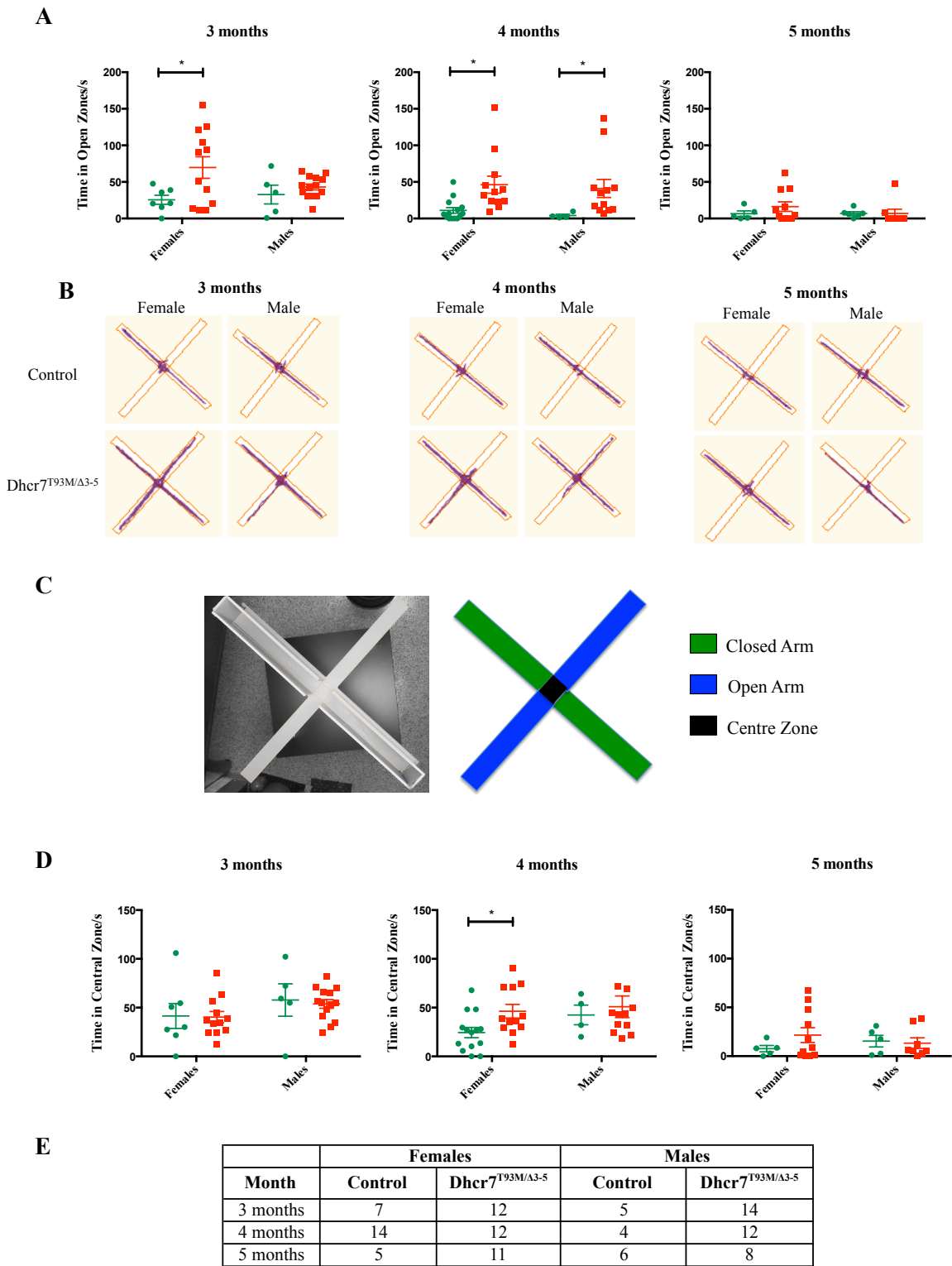


Figure 3.8: Elevated Plus Maze

Figure 3.8 (previous page): Elevated Plus Maze**A) Graphs showing the time spent in the open zones of the elevated plus maze.**

Dhcr7^{T93M/Δ3-5} (red) mice spent significantly more time in the open zones at 3 (females: $p = 0.0151$, males: $p = 0.4750$) and 4 months (females: $p = 0.0130$, males: $p = 0.0120$) of age than control (green) mice. This difference normalised by 5 months (females: $p = 0.2199$, males: $p = 0.9959$).

B) Representative track plots confirming the time spent in the open zones.

The left-right diagonal (\) are the closed zones. The right-left diagonal (/) are the open zones

C) Visual representation of the setup of the elevated plus maze.

The orientation of the setup relates to the track plots

D) Graphs showing the time spent in the central zone of the elevated plus maze.

The central zone is thought to represent decision-making and these graphs show the amount of the control (green) and *Dhcr7*^{T93M/Δ3-5} (red) spent in the central zone. There were no differences between the times spent in the central zones at 3 months between either genotype (females: $p = 0.9332$, males: $p = 0.8260$). At 4 months, female *Dhcr7*^{T93M/Δ3-5} spent significantly more time in the central zone than controls ($p = 0.0207$) but this was not the case for males ($p = 0.5848$). There were no significant differences at 5 months of age, although there tended to be more variation in the *Dhcr7*^{T93M/Δ3-5} mice (females: $p = 0.1179$, males: $p = 0.8026$)

E) Number of mice analysed at each time point

3.3.v: Open field

The data was subjected to a two-way ANOVA that examined the effect of genotype and gender on the mean gait speed. At all time points, there was no significant interaction between gender and genotype (**Table 3.6**). As such, the data was further analysed via a multiple t-test statistical test.

	Time Point	Interaction	Significance
Mean Speed	3 month	$F(1,30) = 3.79$	No, $p = 0.0609$
	4 month	$F(1,19) = 0.12$	No, $p = 0.7283$
	5 month	$F(1,22) = 3.08$	No, $p = 0.0930$

Table 3.6: 2-way ANOVA interaction scores for the open field test.

A two-way ANOVA was implemented to discover the effect of genotype and gender on mean gait speed. At all time points there was no significant interaction between genotype and gender. Therefore, further analysis could be implemented via multiple t-test. However, although not significant, the interaction for the open zone at 3 and 5 months was under 0.1 (blue text).

At 3 months, there was no difference between the mean gait speeds of the male control and *Dhcr*^{T93M/Δ3-5} mice (males: $p = 0.6793$). The *Dhcr*^{T93M/Δ3-5} female mice had a mean gait speed significantly higher than the control female mice (females: $p = 0.0049$) (**Figure 3.9A**). This difference disappeared over the time course. At 4 and 5 months there was little difference in the mean speeds of the mice. There were two female control mice that had a much higher mean gait speed across all time points than the other female control mice. Interestingly, the *Dhcr*^{T93M/Δ3-5} mice that had a higher mean speed also spent more time on the open zones of the elevated plus maze. As can be seen

from the track plots (**Figure 3.9B**), female *Dhcr7*^{T93M/Δ3-5} mice cover more ground at 3 months, including the central zones. This decreases over the time course and at 5 months, the female *Dhcr7*^{T93M/Δ3-5} mice overall stay to the edges, similar to the control female mice.

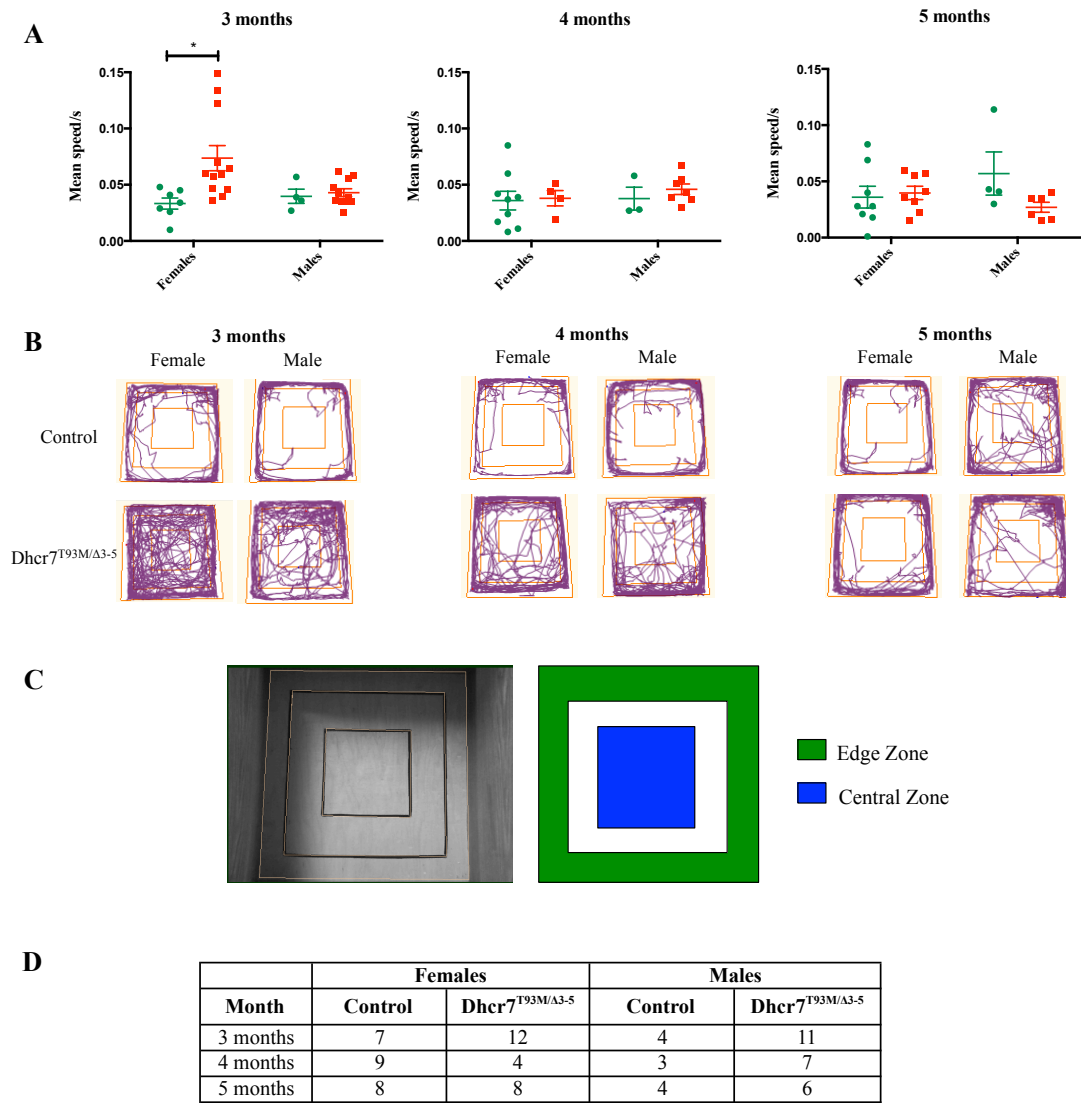


Figure 3.9: Open Field

Figure 3.9 (previous page): Open Field

A) **Graphs showing the mean travelling speed of the mice during the 15 minute test.** Female *Dhcr7*^{T93M/Δ3-5} (red) mice had a significantly higher mean speed than control (green) female mice at 3 months (female: $p = 0.0049$, males: $p = 0.6793$). This normalised up to 5 months of age. Overall, there was little difference in the mean speed of male mice (4 months: female: $p = 0.8483$, males: $p = 0.5129$. 5 months: females: $p = 0.7461$, males: $p = 0.2168$).

B) **Representative track plots of the open field.**

It can be seen that female *Dhcr7*^{T93M/Δ3-5} mice cross the centre multiple times at 3 months of age whereas the control female mice stay to the edge of the field. *Dhcr7*^{T93M/Δ3-5} male mice also cross the centre more often than control male mice. This difference normalised over time.

C) **Visual representation of the setup of the open field.**

The orientation of the setup relates to the track plots

D) **Number of mice analysed at each time point**

3.3.vi: Histological analysis

After the behavioural tests were completed, at 5 months of age both *Dhcr7*^{T93M/Δ3-5} and control mice were culled and the brains were dissected. Coronal sections (20μm) were cut and stained using hematoxylin and eosin (H&E) (**Figure 3.10A**). Hematoxylin stains the nuclei and eosin stains acidophilic substances, mainly present in the cytoplasm. *Dhcr7*^{T93M/Δ3-5} brains are smaller than control brains and, therefore, measurements were normalised to the whole section size and presented as a percentage of total section size (**Figure 3.10B**). It was found that a cross-section of the anterior third of the hippocampus in *Dhcr7*^{T93M/Δ3-5} mice, irrespective of gender, was smaller and misshapen compared to control (females: $p = 0.0033$, males: $p = 0.0193$). This was

particularly evident in female mice, where ventricular dilatation was also more prominent.

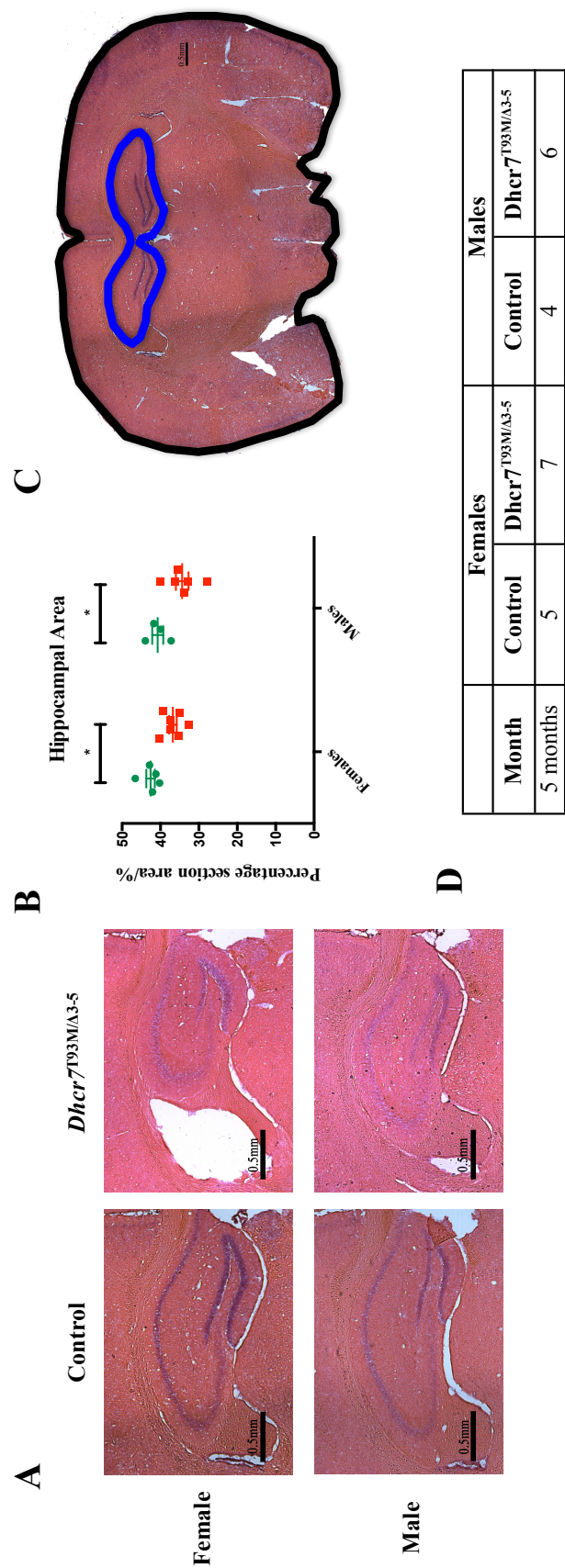


Figure 3.10: Hippocampal Analysis

Figure 3.10 (previous page): Hippocampal Analysis**A) Representative H&E images of the hippocampus in 5-month old mice.**

It can be seen that the hippocampus is smaller in the *Dhcr7*^{T93M/Δ3-5} mice, particular in the females (females: $p = 0.0033$, males: $p = 0.0193$).

Additionally, the fainter hematoxylin staining of the *Dhcr7*^{T93M/Δ3-5} hippocampus is suggestive of decreased neuronal content.

B) Graph quantifying the hippocampal area as a percentage of total section area.

The area is significantly smaller in female and male *Dhcr7*^{T93M/Δ3-5} (red) mice than control (green) mice.

C) Trace showing how measurements were obtained using ImageJ.

The area of the entire section (black line) and the area of the hippocampus (blue line) were measured. The following calculation was performed to gain the hippocampus area as a percentage of section area: (Hippocampal area/Section area) X 100

D) Number of 5 month old mice analysed

The anterior commissure was also noted to be smaller or absent in approximately 20% of *Dhcr7*^{T93M/Δ3-5} mice compared to control mice (females: $p = 0.0502$, males: $p = 0.1366$) (**Figure 3.11A and B**). A preliminary study in 10 month old female mice was also undertaken and it was observed that the anterior commissure was also smaller or absent in the *Dhcr7*^{T93M/Δ3-5} mice (females: $p = 0.1030$) (**Figure 3.11C**). This is only evident in a small number of mice, explaining the non-significant p -value.

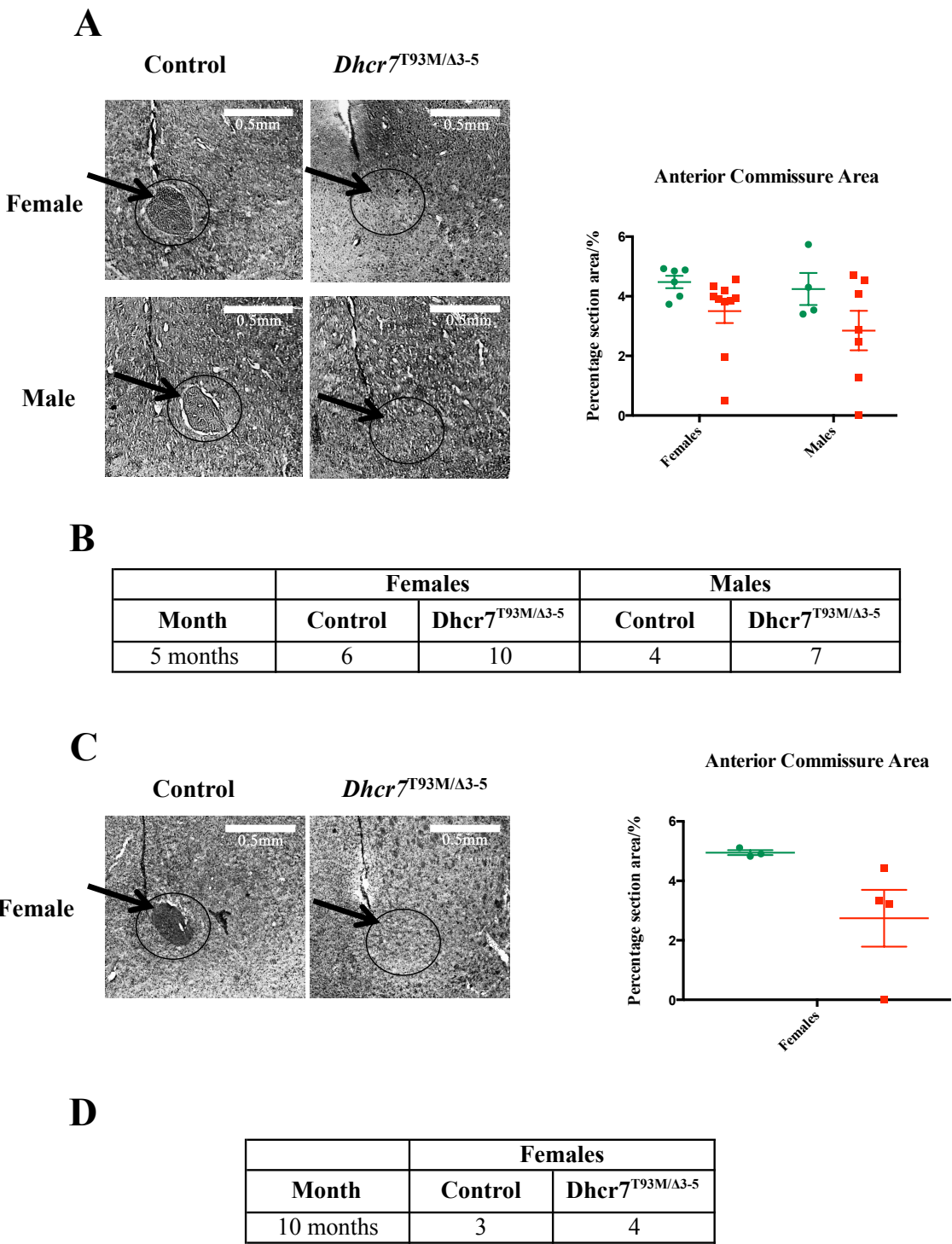


Figure 3.11: Anterior Commissure Analysis

Figure 3.11 (previous page): Anterior Commissure Analysis**A) Representative H&E images of the anterior commissure**

The black arrow and circle depict where the anterior is or should be. The images have been converted to grey-scale so the anterior commissure can be viewed more easily. It can be seen that the anterior commissure is smaller or absent (represented in the images in this figure) in approximately 20% of the *Dhcr7*^{T93M/Δ3-5} mice (females: $p = 0.0502$, males: $p = 0.1366$).

B) Graph quantifying the anterior commissure area as a percentage of total section area.

The area is smaller in about 20% of female and male *Dhcr7*^{T93M/Δ3-5} (red) mice than control (green) mice.

C) Preliminary study of 10 month female mice.

The black arrow shows the anterior commissure is present in the left control image but absent in the right *Dhcr7*^{T93M/Δ3-5} image. This is quantified in the graph (control in green, *Dhcr7*^{T93M/Δ3-5} in red) (females: $p = 0.1030$).

3.4: Discussion

It has been reported that laboratory mice usually remove ~ 70g in 2 hours and ~ 200g of food pellets overnight [194]. However, the results seen in this study showed very little burrowing in the 2-hour test. This could be explained by several factors. It has been previously noted that there is variation in how different strains perform in the test [209]. Deacon stated that C57BL/6 mice readily burrow but 129S2/SvHsd mice do not [206]. The mice used in this study were on a mixed BL6/129 background, which could explain the variation. Although using an isogenic background would likely obtain tighter results, general conclusions can be drawn from this data in the *Dhcr7*^{T93M/Δ3-5} mouse model. Additionally, although Deacon used adult mice, the study does not state the exact age of the mice used [206]. My results show that the amount burrowed in 2 hours increases up to 5 months. This could suggest that as mice age, they burrow more in the 2-hour time limit and if I had continued the time course beyond 5 months, my results may have agreed with the previous report. However, mice typically reach adulthood ~ 2 months of age and so this is unlikely. A more likely suggestion for why the amount burrowed increases over the time course is that the novelty of the burrow decreases and the mice become less anxious to approach it. Although there was variation in all genders and genotypes, overall, the *Dhcr7*^{T93M/Δ3-5} mice were marginally more likely to have increased burrowing compared to controls. However, this was only significant at 4 months in the overnight study. This could be showing the increased anxiety phenotype or hyperactivity, which is also noted in the increased gait speed of the *Dhcr7*^{T93M/Δ3-5} mice in the open field test.

In contrast, the Nestlet test and elevated plus maze are both suggestive of decreased anxiety in the *Dhcr7*^{T93M/Δ3-5} mice. As this is opposite to what is typically observed in patients, these findings are either the result of a secondary defect or suggest that the *Dhcr7*^{T93M/Δ3-5} model does not replicate the behavioural traits of patients.

Decreased shredding in the Nestlet test has been observed to be caused by serotonin uptake inhibitors [210] and also has been associated with lesions in the hippocampus [197]. The marble burying task is also reported to be affected by serotonin uptake inhibitors [210], and although no previous discernable differences were reported in the marble burying test, the *Dhcr7*^{T93M/Δ3-5} had a slight tendency to bury less marbles than control mice. However, this was only significant in male mice at 4 months, which could be due to small sample number. If found to hold with larger sample size, this, along with the Nestlet result, could suggest some very subtle alteration of serotonin production or signalling and is consistent with a previous study using a more severe mouse model of SLOS [211].

The more severe model also had hippocampal abnormalities and histological analysis of *Dhcr7*^{T93M/Δ3-5} cortical sections agreed with this, showing alterations in the hippocampus compared to controls. SHH signalling is an important factor in correct neural patterning and has been suggested to have a role in the development of dorsal midline structures [212], causing enlarged ventricles and affected development of the cortical hem. The cortical hem is a structure that has marked WNT expression [213] and is involved in hippocampal organization [214]. WNT signalling has been shown to be involved in SLOS pathogenesis [215] and suggests that the correction of this pathway could be a potential therapeutic target (discussed further in **Chapter 4**).

The hippocampus is the part of the brain responsible for learning and memory and spatial awareness. As such, the changes seen in the hippocampus could explain the decreased shredding observed in the Nestlet test and the potential delayed learning of the *Dhcr7*^{T93M/Δ3-5} mice, observed by the normalisation of the behaviours in the elevated plus maze and open field tasks. This could also provide an explanation for the decreased anxiety phenotype observed in the elevated plus maze test as the *Dhcr7*^{T93M/Δ3-5} mice may have altered special awareness and may not be fully aware of the elevation or lack of walls on the open arms. *Dhcr7*^{T93M/Δ3-5} mice were additionally observed to fall off the maze more often than controls, which could be a sign of altered special awareness. However, this observation could also be due to the *Dhcr7*^{T93M/Δ3-5} mice spending more time on the open arms, resulting in more opportunities to fall off the maze.

In all behavioural tasks, the *Dhcr7*^{T93M/Δ3-5} mice had most difference from controls at 3 months of age and this difference tended to normalise up to 5 months of age. The defects seen in the hippocampus may be causing a delay in learning, which would correspond to the cognitive defects seen in SLOS patients. Further behavioural tests, such as the Morris water maze to assess spatial awareness [216] and the object discrimination tests [217] to measure cognitive ability, would need to be carried out in the *Dhcr7*^{T93M/Δ3-5} mice to analyse this phenotype further.

Histology also revealed defects in the anterior commissure. The anterior commissure is a white matter tract that connects the two hemispheres of the brain. Although the function is not fully elucidated, it is suggested to have multiple functions, including

recognising pain sensation and chemoreception. It also may have a role in the connection between the amygdalas (emotional responses) and temporal lobes (speech, memory and hearing). As well as cognitive defects, SLOS patients present with hypersensitivity and autistic behaviours, particularly with regards to communication [218]. The smaller anterior commissure could be partly responsible for these behaviours but further study is needed to determine if this is the case.

Although MRI scans of SLOS patient brains found defects in the corpus callosum, cerebellum and septum pellucidum [67,68], no obvious defects in these structures were observed in the *Dhcr7*^{T93M/Δ3-5} model. This does not automatically discount the mouse model as a reliable model as even within patients there was variability. This partly depended on the severity of the patient's disease and, since the *Dhcr7*^{T93M/Δ3-5} mutation is mild, the defects may be subtle or only appear in a limited subpopulation. This is exemplified by the smaller anterior commissure only being observed in around 20% of cases in the *Dhcr7*^{T93M/Δ3-5} mouse model, generally corresponding to higher degrees of ventricular dilatation.

Ventricular dilatation was previously observed in approximately a third of patients [68] and *Dhcr7*^{T93M/Δ3-5} mice [110] and therefore if the numbers of mice studied was increased, the number of mice with smaller anterior commissures may reach a similar value. White matter is largely comprised of myelin and cholesterol is a large component of myelin. Unexpectedly, the MRI studies [68] did not find evidence of altered myelination, which could be due to the sensitivity of the MRI. Therefore it should be examined if myelin is altered in this SLOS mouse model using

immunohistochemistry techniques, for example, by staining for myelin-basic protein 2 (MBP-2).

In conclusion, *Dhcr7*^{T93M/ Δ 3-5} mice were observed to have some differences in behaviour and brain structure compared to control mice, with more variation and differences observed in the female mice. Although the 2-way ANOVA test did not report any interaction between gender and genotype, the *p*-value was below 0.2 at one or two time points for both the elevated plus maze and open field. This may explain the slight gender difference in the results of the behavioural tests, particularly the elevated plus maze and open field. A potential cause of gender differences could be due to hormone changes. As cholesterol is an important precursor to many hormones and as SLOS is associated with altered sterol levels, this could have more of an impact in the female hormonal cycle. The effect of the oestrous cycle on behavioural tests has been investigated in C57BL/6J mice [219]. This paper reported that C57BL/6J mice did not show differences at the points of the oestrous cycle but Balb/c mice were affected. As the *Dhcr7*^{T93M/ Δ 3-5} mice are on a mixed background, it is uncertain how much of an effect the oestrous cycle had on these results. Additionally, many behavioural tests are sensitive to strain differences and therefore an SLOS mouse model with an isogenic background should be tested to determine if these results are accurate.

Interestingly, preliminary results show that *Dhcr7*^{T93M/ Δ 3-5} mice bred on an isogenic C57BL/6J background do not present with the survival associated with the *Dhcr7*^{T93M/ Δ 3-5} mice bred on a BL6/129 mixed background. Out of 13 *Dhcr7*^{T93M/ Δ 3-5} isogenic C57BL/6J mice, 12 died at P1 and 1 survived. From 19 mice genotyped as *Dhcr7*^{T93M/+} (isogenic C57BL/6J), 17 survived and 2 died at P1 (Dr. Niamh Cawley,

unpublished data). This could suggest that there are modifier genes which affect the severity of this genotype [220] and could potentially explain the amount of variation observed in this study.

In my study, structural neurological differences were also seen but as the number of mice analysed was limited, only potential conclusions can be made. However, they are suggestive of subtle structural defects and therefore justify further analysis, particularly using mice with an isogenic background where the differences may be more distinct.

As SLOS does not have an optimal treatment pathway, if the results suggested by these tests are accurate, it could help aid the development of therapeutic interventions.

Chapter 4: Analysis of RNA expression in Smith-Lemli-Opitz Syndrome and Niemann-Pick type C Disease

4.1. Introduction

The molecular biology of a cell is complex and is still not completely understood, involving multiple interactions that are heavily regulated by feedback mechanisms and signalling pathways. The studies of these cellular processes are critical to discovering the cause of many debilitating disorders, therapeutic targets and the molecular mechanisms of drug treatments. These studies are split into different classes of ‘omics’, including:

- **Genomics**; the expression of all the genes in an organism
- **Transcriptomics**; the expression of the mRNA of an organism
- **Proteomics**; the study of the proteins of an organism
- **Metabolomics**; the study of the global metabolites in an organism under a given set of conditions

As the cellular environment changes, cells respond by altering regulation of gene transcription in order to maintain homeostasis. The process of transcription involves the formation of mRNA from DNA, including the splicing out of non-coding DNA.

Thus, the transcriptome reflects the genes that are being actively expressed at a given moment. In order to determine gene expression experimentally, the mRNA is reverse-transcribed to cDNA and then analysed via microarray, RNAseq or individual genetic assays, such as real time PCR (**Figure 4.1**).

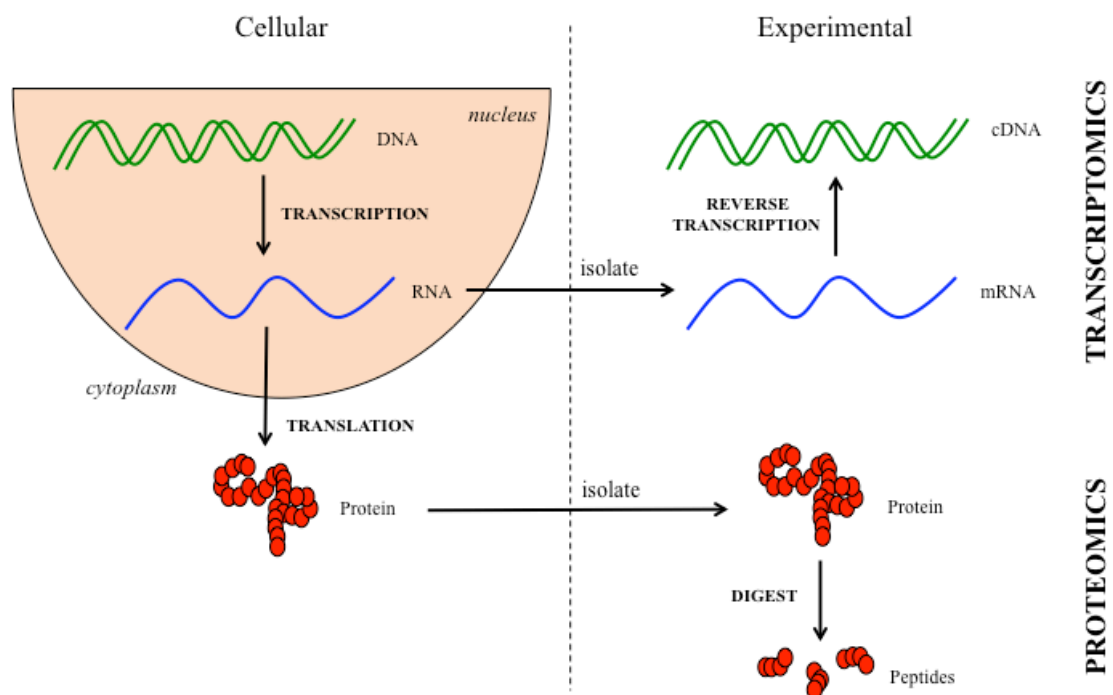


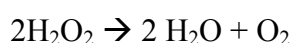
Figure 4.1: Transcription, Translation and their corresponding ‘omics’

The left-hand side of the dotted line represents the biological processes of transcription and translation and the right-hand side of the dotted line shows the beginning experimental processes of transcriptomics and proteomics.

Transcriptomics involves the reverse-transcription of mRNA to cDNA, which then goes onto further genetic analysis, e.g. TaqMan or microarrays. For proteomic analysis, the protein is digested and then further analysed using mass spectrometry.

The metabolome of SLOS has been relatively well studied, including the known hallmark of altered sterol levels, but it is uncertain if the pathogenesis of SLOS is caused by the decrease in cholesterol or the increase in 7DHC. This may be elucidated by the discovery of reactions specific to 7DHC.

Oxysterols are oxygenated derivatives of sterols, which can be formed via free radical or enzymatic oxidation and can exert multiple biological effects including cytotoxicity, suppression of immune response and regulation of cholesterol homeostasis [221]. This is usually referred to as oxidative stress. Oxidative stress is a broad term that describes an imbalance in redox couples, often via the production of reactive oxygen species (ROS). The concept of oxidative stress first began to be considered in 1954 [222] when researchers started to investigate the effects of free radicals and oxygen poisoning. Among some of the earliest discovered enzymes were superoxide dismutase and catalase, which catalyse the respective reactions:



McCord et al. suggested that superoxide dismutase is critical to the survival of oxygen-metabolising organisms [223], but, although superoxide, O_2^- , is the most predominant radical produced in free radical biology [224], not all free radicals are superoxide.

Lipid peroxidation is the process of forming lipid peroxide via the interaction of ROS and lipids. Lipid peroxidation has been suggested to be involved in the pathophysiology of many human disorders and is measured by the propagation rate

constant for free radical oxidation. 7DHC has an intensely increased propensity for lipid peroxidation compared to cholesterol and has a propagation rate of $2260 \text{ M}^{-1} \text{ s}^{-1}$, which is almost 200 times that of cholesterol, $11 \text{ M}^{-1} \text{ s}^{-1}$ [225]. This may result in increased levels of 7DHC-derived oxysterols, particularly in tissues where 7DHC levels are high. Oxygen and lipids are present at high concentrations in the brain and there are also fewer neurological defences against oxidative stress [226]. As such, the brain is thought to be particularly vulnerable to oxidative stress [227].

Oxidative stress has been previously suggested to be present in SLOS. Elevated levels of lipid biomarkers associated with oxidative stress have been seen in murine models [228] and DNA microarray analysis has shown increased expression of genes associated with oxidative stress and regulation of cell death [229]. In 2010, Korade *et al.* studied how 7DHC derived oxysterols can alter the gene expression in Neuro2a cells and they found that cellular proliferation is down regulated. Genes of cholesterol biosynthesis are also regulated, including *DHCR7*, *SQS*, *HMGCR*, *PRR13*, *ADAM19* and *SNAG1* [230], but the extent to which 7DHC-derived oxysterols affect the pathophysiology of SLOS is still uncertain.

Although the presence of oxidative stress is uncertain in SLOS, oxidative stress is known to be prevalent in NPC [231] and NPC fibroblasts have been shown to have changes, such as reduced catalase activity, consistent with oxidative stress [232]. A common treatment for NPC, miglustat, has been suggested to improve oxidative stress defences [233]. Miglustat is a small iminosugar molecule that inhibits glucosylceramide synthase [131] and subsequently causes reduced production of glycosphingolipids. Glycosphingolipids (GSLs) and cholesterol have been shown

to be important in membrane microdomains, also known as lipid rafts [234], and a murine model of NPC was found to have increased levels of GSLs in the endocytic pathway, including LacCer, GA2, GM3 and GM2 [132]. Increased levels of GSLs were also seen in the membrane microdomains of liver, spleen and cerebellum of this NPC model and treatment with miglustat, restored trafficking and localisation of GSLs. GSLs are also suggested to be involved in cellular proliferation and differentiation [235].

4.1.i: Aims of Chapter 4

- To further examine the extent of oxidative stress in SLOS via expression of genes involved in oxidative stress.
- Compare the transcriptome of wild type (WT), SLOS and NPC fibroblasts and the effect miglustat has on them.

4.2: Materials and Methods

4.2.i: RNA isolation

Half a cerebral cortex from female 3-month *Dhcr7*^{T93M/ Δ 3-5} and control (*Dhcr7*^{+/+}) mice were PBS perfused and homogenized (OMNI TH homogenizer) in 0.4ml Trizol (cat.# 15596018, Life Technologies). Once fully homogenized, 0.35ml Trizol was added. The RNA-extraction columns have a maximum capacity of 100 μ g RNA, therefore the homogenate was separated into 200 μ l samples and made up to 0.75ml with Trizol. The homogenate was mixed via pipetting and left to incubate at room temperature for 5 minutes. 0.2ml chloroform was added, vortexed and incubated for a further 2-3 minutes at room temperature. The solution was then centrifuged at 12,000g (10,600 RPM) for 15 minutes at 4°C (Centrifuge 5417R, Eppendorf). The upper aqueous phase was transferred to a new tube and vortexed gently (Thermomixer R, Eppendorf) while 200 μ l 100% RNase-free ethanol was added drop-wise. The solution was added to the RNeasy column (RNeasy Mini Kit, cat.# 74106, Qiagen) and RNA extracted using the manufacturer's protocol. The RNA concentration was then determined using a Nanodrop (Nanodrop 2000 Spectrophotometer, Thermo Scientific).

4.2.ii: cDNA synthesis and profiler array

cDNA was synthesized using 400ng of total RNA and the reagents and protocol of the RT² First Strand Kit (cat.# 330401, Qiagen). The resulting cDNA was mixed

with 2x RT² SYBR Green Mastermix and RNase-free water in the amounts specified by the protocol for the 384 (4 x 96) profiler array option (RT² Profiler PCR Mouse Oxidative Stress Array, cat.# PAMM-065ZE-4, Qiagen). 10µl of this mixture to the array plate was added and continued as per manufacturers instructions using a 7900 HT Fast Real Time PCR System (Applied Biosystems).

4.2.iii: Taqman validation

RNA was extracted from a heterozygous mouse (*Dhcr7*^{T93M/+}), as described above, and cDNA was made (High Capacity cDNA Archive kit, cat.# 4368813, ThermoFisher Scientific). The following TaqMan probes were used:

- *Epx* (Mm00514768_m1, Applied Biosystems)
- *Sod3* (Mm00448831_s1, Applied Biosystems)
- *Ccl5* (Mm01302427_m1, Applied Biosystems)

Each probe was validated against *Gapdh* using *Dhcr7*^{T93M/+} cDNA and concentrations ranging between 0 – 100ng to create a standard curve. 96-well plates (MicroAmp Optical 96-well Reaction Plate, cat.# 4306737, Applied Biosystems) were used and the assays run on an 7300 Real Time PCR System (Applied Biosystems). To validate, the slopes of the probe have to be within 0.1 of the slope of *Gapdh* (*Gapdh* Control Reagents, cat.# 1601134, Life Technologies). The Ct threshold was set at 0.15, which was within the linear range of the amplification curves.

RNA was isolated from eight female 3-month half cortexes (four control, four *Dhcr7*^{T93M/ Δ 3-5}) as previously described. These were different mice to those run on the profiler array. cDNA was made using the High Capacity cDNA Archive kit (cat.# 43-688-14, ThermoFisher Scientific). Due to competition between GAPDH and the probe, the probes were run in separate wells on the plate. cDNA had previously been extracted from E18.5 control and *Dhcr7* ^{Δ 3-5/ Δ 3-5} mice. As with the *Dhcr7*^{T93M/ Δ 3-5} 3 month mice, the RT-PCR was performed on the 7300 Real Time PCR System (Applied Biosystems) with *Gapdh*.

4.2.iv: Western blot

300 μ l RIPA lysis buffer (150mM sodium chloride, 1% Triton X-100, 0.5% sodium deoxycholate, 0.1% sodium dodecyl sulphate, 50mM Tris, pH 8.0) with protease inhibitors (5mM sodium fluoride, 1mM sodium orthovanadate) was added to ~5mg tissue from three female 3 month mouse half cerebral cortexes (one control, one *Dhcr7*^{T93M/+} and one *Dhcr7*^{T93M/ Δ 3-5}). The samples were homogenised using a Omni Bead Ruptor 24 (Omni International Inc.) and subsequently centrifuged at 12,000 RPM at 4°C. The supernatant was aspirated and protein concentration determined using a Bradford assay (DC Protein Assay Kit, cat.# 500-0112, Bio Rad).

7 μ l SeeBlue Plus2 ladder (cat.# LC5925, ThermoFisher Scientific) was loaded into a Bis-Tris 4-12% NuPAGE gel (1mm x 12 well, cat.# NP0342BOX, ThermoFisher Scientific). 13 μ l of lysate + loading buffer was added and the gel was run and transferred using an iBlot 1 (ThermoFisher Scientific). The resulting membrane was put in 20ml blocking buffer (Western Breeze kit, cat.# WB7106, ThermoFisher

Scientific) overnight and then probed using 1:1000 dilution Epx (cat.# sc-19147, Santa Cruz Biotechnology) and 1:10,000 dilution β -actin (cat.:Ab49900, Abcam) as the loading control. The anti-goat Western Breeze kit (cat.# WB7106, ThermoFisher Scientific) was used and bands imaged using a BioRad ChemiDoc XRS+ imager.

4.2.v: Cell culture for RNAseq

Cells from patient cell lines Control, A2SLOS (homozygous c.964-1G>C) and NPC25 (c.2979dupA, c.2103C>G; fs(exon 20), p.N701K) were plated in T75 flasks at a density of 200,000 cells and grown in Dulbecco's Modified Eagle Medium (DMEM, cat.# 11995-065, ThermoFisher Scientific) with 10% fetal bovine serum (cat.# FB-12, Omega Scientific Inc.), 1X (2mM) L-glutamine (cat.# 25030-081, Life Technologies), 1X Penicillin (100 Units/ml)/Streptomycin (100 μ g/ml) (cat.# 15140-122, Life Technologies) in a 5% CO₂ environment at 37°C overnight. Each flask contained two 18mm square coverslips (cat.# 12-541A, Fisher Scientific). The next day each flask was aspirated and washed twice with 1X PBS (cat.# 10010-023, Life Technologies). The cells were then grown in 20ml of the desired media for 5 days.

- **10% FBS:** Dulbecco's Modified Eagle Medium (DMEM) with 10% fetal bovine serum, 1X L-glutamate, 1X Penicillin/Streptomycin
- **7.5% LPDS:** McCoy's 5A media with 7.5% lipo-deficient serum, Insulin Growth Factors 1X L-glutamate, 1X Penicillin/Streptomycin

- **7.5% LPDS + 100 μ M miglustat:** McCoy's 5A media with 7.5% lipo-deficient serum, Insulin Growth Factors 1X L-glutamate, 1X Penicillin/Streptomycin, 100 μ M miglustat

The LPDS was made by Dr. Christopher Wassif and Dr. Stephanie Cologna as follows:

- 1) Dissolve 0.1 mg of EDTA in to 50 ml FBS Serum
- 2) Mix 128 ml n-butanol with 192 ml iso-propyl ether
- 3) Add 80 ml of the butanol/iso-propyl ether solution to 40 ml of the FBS/EDTA.
- 4) Shake for 1 hour in the dark.
- 5) Centrifuge at 5°C for 15 min at 5000 RPM.
- 6) Remove the bottom phase and re-centrifuge at 5°C for 15 min at 5000 RPM.
- 7) Remove the top phase and filter through a 0.8 μ m filter.
- 8) Lyophilize the solution.
- 9) Re-suspend in the same volume that was lyophilized and filter sterilize through 0.2 μ m filter. Freeze at -20°C.

4.2.vi: Collecting cell pellets

Media was aspirated and cells washed with 1X PBS. The coverslips were collected and fixed with 4% paraformaldehyde for 20 minutes (cat.# P6148, Sigma). 3ml

trypsin (0.05% Trypsin-EDTA, cat# 25300-054, Life Technologies) was added to the flasks, left to incubate at 37°C for 5 minutes. Trypsin was quenched with 3ml 1X PBS. There were two T75 flasks for each repeat. Half the cells from one flask were separated for FACS analysis and spun down (3000 RPM for 3.5 minutes, IEC Centra CL2). The remaining cells were spun down, frozen and saved for RNAseq analysis.

4.2.vii: Immunofluorescence

The coverslips were placed in petri dishes and rinsed in 1X PBS. 2 ml permeabilisation buffer was added and left for 30 minutes in the dark at 4°C. The coverslips were rinsed in 1X PBS three times. 2ml Incubation buffer was added, left for 48 hours in the dark at 4°C and then the coverslips were washed a further three times in 1X PBS. A nuclear stain was added for 30 minutes in the dark at room temperature, the coverslips were rinsed with 1X PBS and mounted onto microscope slides.

- **Permeabilisation buffer:** 1X PBS, 10% non-steroidal goat serum (cat.# G9023, Sigma), 2% 10% saponin (cat.# S4521-10G, Sigma) stock
- **Incubation buffer:** 1X PBS, 10% non-steroidal goat serum (cat.# G9023, Sigma), 0.5% 10% saponin stock, 0.25% 0.002% filipin (cat.# 08707, PolySciences Inc.) stock, 0.1% cholera toxin (cat.# C-22841, Molecular Probes)

- **Nuclear stain:** 1x PBS, 10% non-steroidal goat serum (cat.# G9023, Sigma), 0.5% 10% saponin stock, 0.05% YO-PRO (cat.# Y3603, Molecular Probes), 0.05% TO-PRO (cat.# T3605, Molecular Probes)

4.2.viii: FACS staining protocol

The cells for FACS analysis were re-suspended in 1X PBS/BSA (833µl 30% BSA (cat.# A9576, Sigma) in 50ml 1X PBS) solution and spun down (1400 RPM for 5 minutes at 4°C). The cells were washed twice in the PBS/BSA solution and then re-suspended in 200µl 1X PBS. The cells were filtered through a 35µm mesh (Falcon 352235) and split into two 100µl samples. 1µl live-dead stain (LIVE/DEAD Fixable Violet Dead Cell Stain Kit, cat.# L34955, ThermoFisher Scientific) was added to each 100µl sample, covered and left for 10 minutes. The samples were washed with 1X PBS and re-suspended in cytofix/cytoperm buffer (BD Biosciences kit, cat.# 512090KZ, BD Biosciences) for 10-15 minutes. Incubation buffer was added and left for 2 hours in the dark at 4°C. Samples were washed twice with 1X cytowash (cat# 554723, BD Biosciences) and 100µl cytofix/cytoperm buffer was added. Samples were kept in the dark until run on the FACS machine (BD LSR Fortessa). BD compbeads were used (cat.# 51-90-9001291 and 51-90-9000949)

- **Incubation buffer:** 1X cytowash, 0.001% filipin, 0.33% cholera toxin

4.2.ix: RNAseq

Cell pellets were obtained and RNA extracted using the RNeasy kit (RNeasy Mini Kit, cat.# 74106, Qiagen). RNA concentration was obtained using a Nanodrop 2000 (ThermoScientific). The RNAseq protocol was carried out by the Molecular Genomics Core (NIH) as follows:

The libraries were constructed with an Illumina TruSeq Stranded mRNA Library Preparation Kit (cat# RS-122-2101) to gain polyA selected RNA and the library was stranded. The samples were sequenced on a HiSeq2500 in RapidRun mode using Version 3 Chemistry and the 35 libraries were barcoded and pooled. The entire pool was run on 5 FlowCells and the sequencing was paired-end with 100 bp from each end. Alignment of the data was performed using RNA-STAR [236] and was performed against the Human genome (hg19 build, UCSC gene annotations). Reads at the gene level were counted using subhead featureCounts [237] and differential expression was determined using DESeq2 [238].

A ranking system was employed to the DESeq2 data using the following formula:

$$\frac{1}{p_{adj}} \times \frac{ABS(\log_2 \text{ fold change})}{\log_2 \text{ fold change}}$$

“p_{adj}” refers to the *p*-value which is normalised using Benjamini-Hochberg method.

Subsequent analysis was performed using Gene set enrichment analysis (GSEA)

software using the pre-ranked tool. Graphs were plotted using GraphPad Prism

(Prism 6 Mac OS X, version 6.0b).

4.3: Results

RNA is a useful tool to investigate disorders and can be used as a basis for multiple methods. In this chapter I use different techniques to analyse varying aspects of the mechanisms involved in SLOS and NPC.

Oxidative stress has been suggested to be present in SLOS so the first part of the results in this chapter focuses on isolating RNA from control and SLOS mouse brains for use in an oxidative stress array. This is combined with previous proteomics work to further interpret the results of the array.

The latter half of the chapter focuses on RNA isolated from fibroblasts from SLOS and NPC patients for RNAseq analysis. As there can be differences between disease models, I first compared relevant results from the array with those seen in the SLOS RNAseq dataset. However, the main focus of the RNAseq datasets was to compare the effects of miglustat on gene expression between SLOS and NPC. Therefore, the remainder of the chapter looks in depth at the RNAseq datasets in order to help elucidate its method of action in both disorders.

4.3.i: RT² Oxidative Stress Array

Oxidative stress has previously been suggested to be a part of SLOS pathogenesis. However, the majority of these studies have been in a rat model using a synthetic

inhibitor of Dhcr7. Therefore, I aimed to analyse expression of genes associated with oxidative stress in the *Dhcr7*^{T93M/Δ3-5} mouse model.

The Oxidative Stress RT² Profiler Array tested 84 genes associated with oxidative stress. RNA isolated from half a cerebral cortex from four 3-month old control and four *Dhcr7*^{T93M/Δ3-5} female mice were analysed in four replicates each. Cycles above 35 were considered negligible, although where only one out of the four replicates was above 35, this was discounted from the calculations. An average cycle number for each gene was calculated for each cortex. An average ΔCt was calculated for control and *Dhcr7*^{T93M/Δ3-5} phenotypes and from this a fold change was calculated. A p-value cut-off of less than 0.05 was put in place and fold change thresholds of ± 2, which corresponds to ± 1 log₂ fold change. Only three genes had p-values less than 0.05:

- **Eosinophil peroxidase (*Epx*);** p-value: 0.0056
- **Chemokine (C-C motif) ligand 5 (*Ccl5*);** p-value: 0.0069
- **Superoxide dismutase 3 (*Sod3*);** p-value: 0.0144

Only *Epx* had a significant p-value and was over the fold change threshold; log₂ fold change of -2.54. *Ccl5* had a log₂ fold change of -0.82 and *Sod3* a log₂ fold change of 0.48 (**Figure 4.2**).

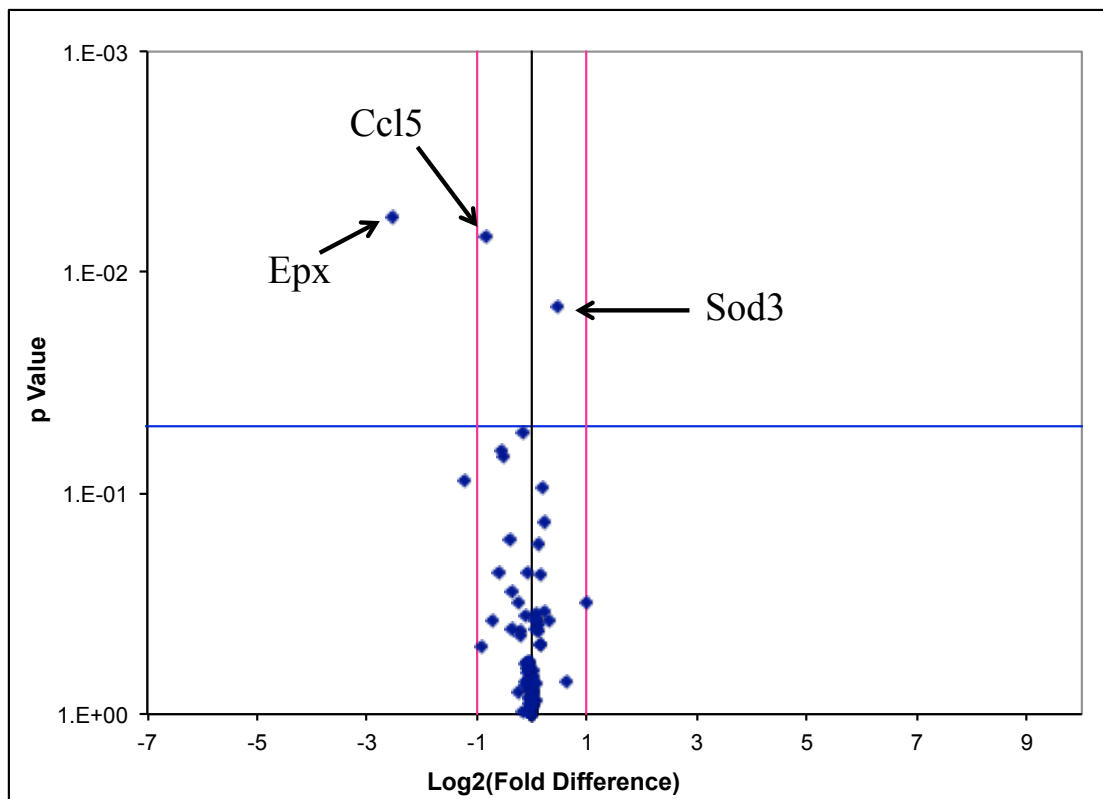


Figure 4.2: Volcano plot of the Oxidative Stress RT² Profiler Array

Three genes had a *p*-value of lower than 0.05; *Epx* (0.0056), *Ccl5* (0.0069) and *Sod3* (0.0144). Of these genes, only *Epx* has a log₂ fold change over the ± 1.0 threshold (-2.54). The remaining 81 genes were below the *p*-value threshold.

4.3.ii: Validation of significant genes

In order to validate these results, four other control and *Dhcr7*^{T93M/ Δ 3-5} half cerebral cortex from 3-month old female mice were used for TaqMan analysis. The average delta Ct for controls was used to assess variation between samples and the variation between control samples was minimal. Three repeats were used for each cortex.

In agreement with the Oxidative Stress Array, *Epx* had significant down-regulation in all four additional *Dhcr7*^{T93M/Δ3-5} cortexes (**Table 4.1**). For *Ccl5* and *Sod3*, there was not significant difference between control and *Dhcr7*^{T93M/Δ3-5} mice. However, “*Dhcr7*^{T93M/Δ3-5} C” had a higher log2 fold change for *Ccl5* and *Sod3*, suggesting there may be variation between samples.

	Log2 Fold Change		
	Epx	Ccl5	Sod3
Control A	-0.10	-0.13	-0.49
Control B	-0.09	-0.30	0.70
Control C	-0.30	0.58	-0.37
Control D	0.49	-0.16	0.16
<i>Dhcr7</i>^{T93M/Δ3-5} A	-3.34	-0.27	-0.17
<i>Dhcr7</i>^{T93M/Δ3-5} B	-5.66	0.72	-0.02
<i>Dhcr7</i>^{T93M/Δ3-5} C	-4.66	1.26	1.23
<i>Dhcr7</i>^{T93M/Δ3-5} D	-4.64	-0.17	0.34

Table 4.1: Validation of Epx, Ccl5 and Sod3 using TaqMan probes

Eight half cortexes from 3 month old female mice (four control (Control A-D), four *Dhcr7*^{T93M/Δ3-5} (*Dhcr7*^{T93M/Δ3-5} A-D), were analysed via RT-PCR using TaqMan probes and *Gapdh* as a normalization control. The average delta Ct of all control samples was used to assess variation between samples. The expression of all three probes had negligible expression in the control cortexes. In agreement with the profiler array, *Epx* had decreased expression in the *Dhcr7*^{T93M/Δ3-5} mice. For *Ccl5* and *Sod3*, “*Dhcr7*^{T93M/Δ3-5} C” had increased expression.

4.3.iii: Protein analysis of Epx

The RT² Oxidative Stress array analyses the expression of RNA but this does not always correspond to the levels of protein produced. Therefore, as Epx was consistently significantly down regulated, a preliminary investigation was implemented on the protein level. Epx comprises two chains – a heavy chain of 55 kDa and a light chain of 15 kDa. Correspondingly, three bands were seen on the Western Blot – 15 kDa, 55 kDa and 70 kDa. β -actin was used as a loading control. In contrast to what was seen in the RNA expression, there was no significant difference in protein level (**Figure 4.3**). However, as this was only a preliminary test with an N of 1, more testing is needed to establish if this is a true observation.

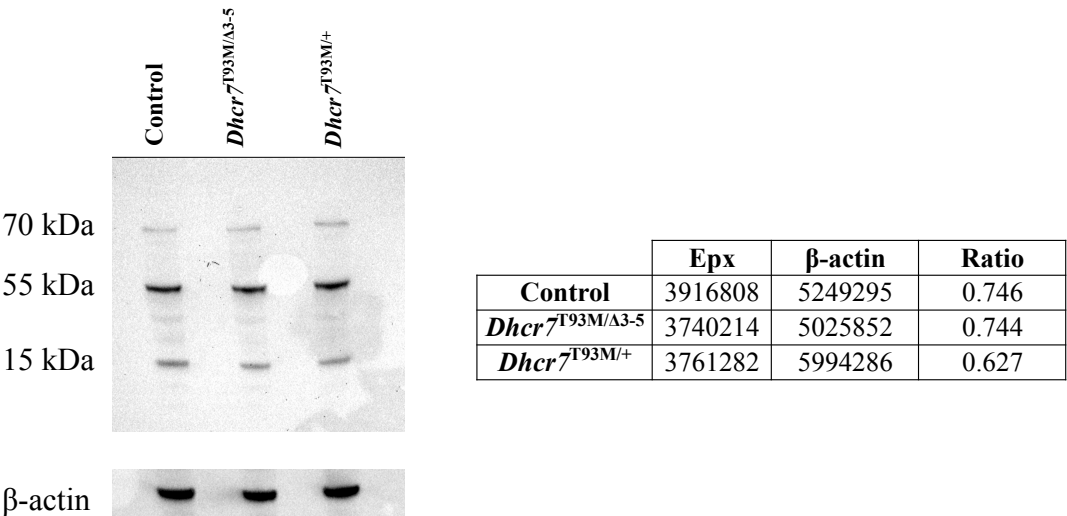


Figure 4.3: Western blot of Eosinophil peroxidase (Epx)

- A) Western blot of lysate from a control, *Dhcr7*^{T93M/Δ3-5} and *Dhcr7*^{T93M/+} half cortex, using β -actin as a loading control. Epx is comprises a heavy chain (55 kDa) and a light chain (15kDa). The entire protein is 70 kDa. The three major bands show these molecular weights.
- B) Using the density of the main 55kDa band, a ratio can be discovered to normalize to β -actin. This showed that there was no difference between any of the phenotypes.

4.3.iv: Superoxide dismutase: further investigation

Superoxide is a common component of oxidative stress and therefore up-regulation of genes that eradicate superoxide, such as superoxide dismutase, are an indicator that oxidative stress is a pathogenic factor. Superoxide dismutase is split into three types; superoxide dismutase 1 (Sod1, located in cytoplasm), superoxide dismutase 2 (Sod2, located in mitochondria) and superoxide dismutase 3 (Sod3, the extracellular form).

Although Sod3 was not above the log2 fold change threshold in the profiler array, a previous proteomics study showed that Sod1 was up regulated in both E18.5 *Dhcr7*^{Δ3-5/Δ3-5} mice and CSF from SLOS patients (**Figure 4.4A and B**). In contrast, the profiler array showed that both *Sod1* and *Sod2* were down regulated, although these values were not significant (**Figure 4.4C**).

From the results of the proteomics study and the profiler array, there seemed to be a potential change in regulation of the superoxide dismutases in SLOS. Therefore, to further test the RNA expression of *Sod3*, E18.5 *Dhcr7*^{Δ3-5/Δ3-5} mice were additionally used. To ensure a fair test, the RNA extracted from 3-month *Dhcr7*^{T93M/Δ3-5} mice used in the previous validation was also analysed. As before, the control mice had no significant changes in *Sod3* levels and “*Dhcr7*^{T93M/Δ3-5} C” had the highest increase in *Sod3*. At E18.5, there was very little change between the control and *Dhcr7*^{Δ3-5/Δ3-5} mice (**Figure 4.4D and E**).

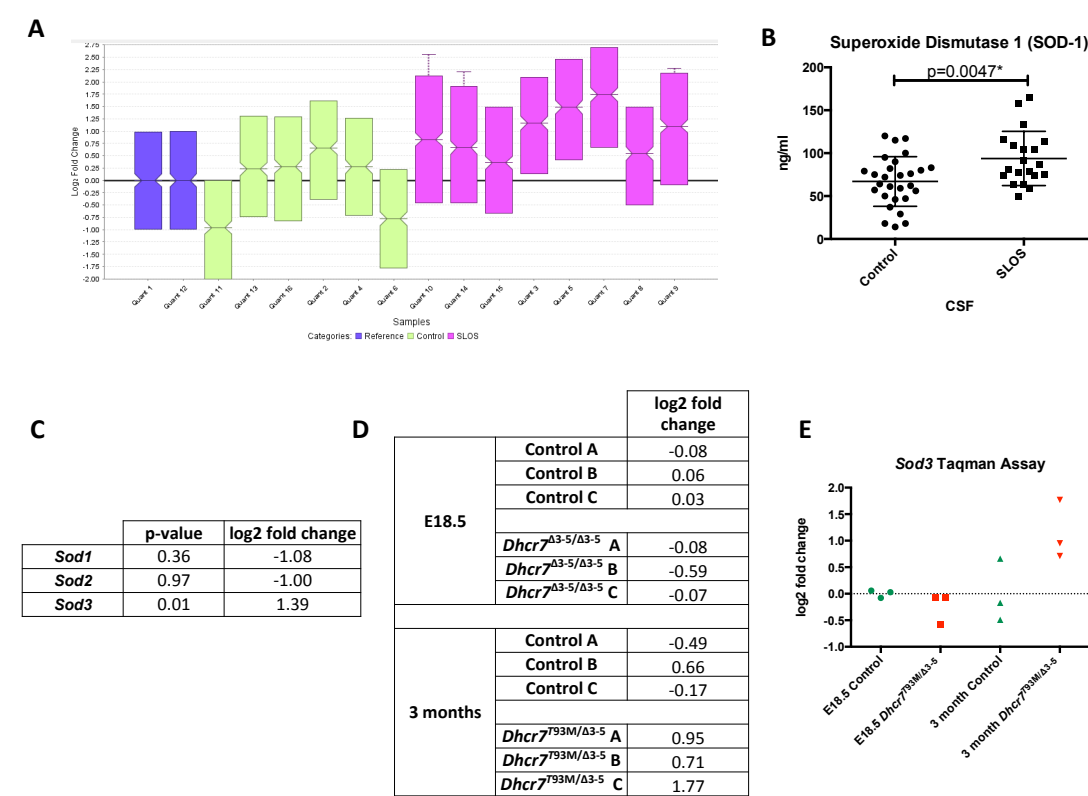


Figure 4.4: Superoxide dismutase analysis

- A) Proteomics analysis of Sod1 from control and *Dhcr*^{Δ3-5/Δ3-5} E18.5 mice
- B) Proteomics analysis of SOD1 from SLOS patient CSF.
- A + B*: Data from Dr. Stephanie Cologna and Christine Sheih
- C) Data for *Sod1* – *Sod3* from the RT² Profiler Array
- D) *Sod3* data comparing E18.5 control and *Dhcr*^{Δ3-5/Δ3-5} mice and 3 month control and *Dhcr*^{T93M/Δ3-5} mice.
- E) Graph plotting expression of *Sod3* E18.5 and 3 month data

4.3.v: RNAseq

Although there is a potentially interesting result in the profiler array for *Sod3*, this is not as significant as that seen from the human proteomics data. Therefore, I decided to analyse my RNAseq data, which shows expression from fibroblasts from a severe SLOS patient to elucidate if there is a difference between the mouse model and human data or if it is difference between RNA and protein expression. The cells were grown in 10% media and lipo-deficient serum (LPDS). As growing the cells in LPDS forces the SLOS phenotype to occur, the difference between cells in 10% and LPDS media represent the expression change in the SLOS phenotype.

In agreement with what is seen in both the array and validation in the *Dhcr7*^{T93M/Δ3-5} mouse, *SOD3* is up regulated. *SOD1* and *SOD2* are also both up regulated in this dataset, although the *SOD1* change does not have a significant *p*-value (**Table 4.2**). This agrees with the proteomics data but is in contrast with the profiler array. Similarly, *EPX* is slightly down regulated but this is not significant or dramatic as in the *Dhcr7*^{T93M/Δ3-5} mouse.

	p-value	log2 fold change
SOD1	0.057	0.45
SOD2	4.05 E-14	1.11
SOD3	1.44 E-10	1.73

	p-value	log2 fold change
EPX	0.86	-0.19

Table 4.2: RNAseq data of *SOD1* – 3 and *EPX*

This data is from the A2SLOS patient cell line RNAseq data set. *SOD1* had a non-significant negligible change. *SOD2* has a significant increase and a log2 fold change of 1.11 and *SOD3* has a significant increase with a log2 fold change of 1.73. *EPX* did not have either a significant *p*-value or log2 fold change.

4.3.vi: RNAseq – miglustat study

The main focus of the RNAseq data was to test the effect of the drug miglustat in SLOS and NPC patient fibroblasts. NPC is a lysosomal storage disorder and the accumulation of cholesterol in the lysosomes can be viewed using filipin staining. SLOS has been observed to have an NPC1-like phenotype with similar punctate staining but miglustat affects the filipin staining of SLOS and NPC fibroblasts in different ways, reducing filipin staining in SLOS but not NPC fibroblasts (**Figure 4.5**).

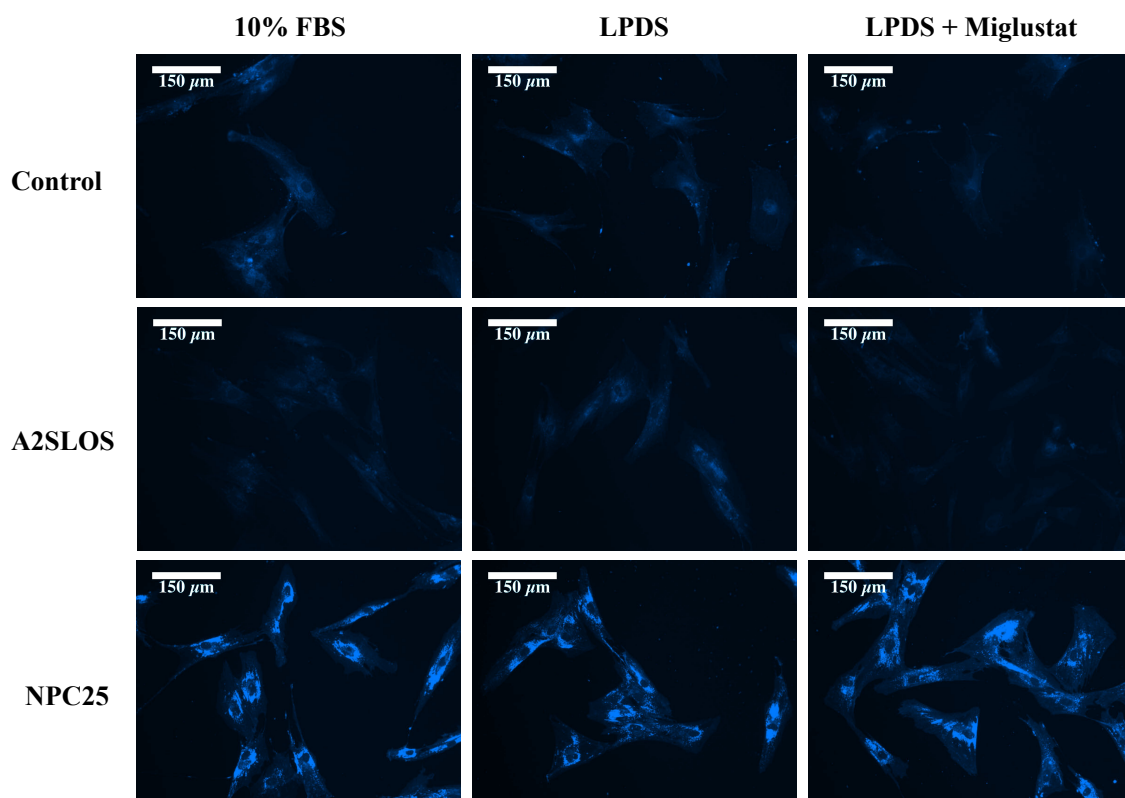


Figure 4.5: Filipin staining of control, A2SLOS and NPC25 fibroblasts

Figure 4.5 (previous page): Filipin staining of control, A2SLOS and NPC25 fibroblasts

Filipin staining (blue) for unesterified cholesterol in control (top), A2SLOS (middle) and NPC25 (bottom) fibroblasts grown in specified media for 5 days. In 10% media, only NPC25 has noticeable filipin staining. Filipin staining may increase slightly in A2SLOS and WT cells when they are grown in LDPS but the main point of interest is that miglustat does not decrease staining in NPC25 or WT cells but does appear to decrease staining in A2SLOS cells.

This staining was inconsistent and therefore FACS was used as a second method to confirm that miglustat was having an effect on the cells. The cells were stained with cholera toxin and measured via FACS (**Figure 4.6**). It was seen that when the cells were grown in 10% FBS, both control and A2SLOS had similar cholera toxin stains whereas NPC25 had a much higher intensity stain. When the cells were placed into LPDS, the staining for NPC25 stayed approximately the same but control and A2SLOS staining increased to a similar level. After miglustat was added, the staining decreased to the same level in all cell lines. Cholera toxin stains GM1 and, as miglustat disrupts the ganglioside synthesis pathway, the reduction in staining shows that the drug is affecting the cells.

FSC-A measures the forward scatter and corresponds to the size of the cells. In 10% FBS, all the cells are different sizes, ranging from control as the smallest to NPC25 cells, which are the largest. In LPDS, control cells remain the same size, but NPC25 and A2SLOS decrease in size, although NPC25 cells are still larger. With the addition of miglustat, NPC25 cells appear to stay the same size whereas control and A2SLOS increase slightly in size.

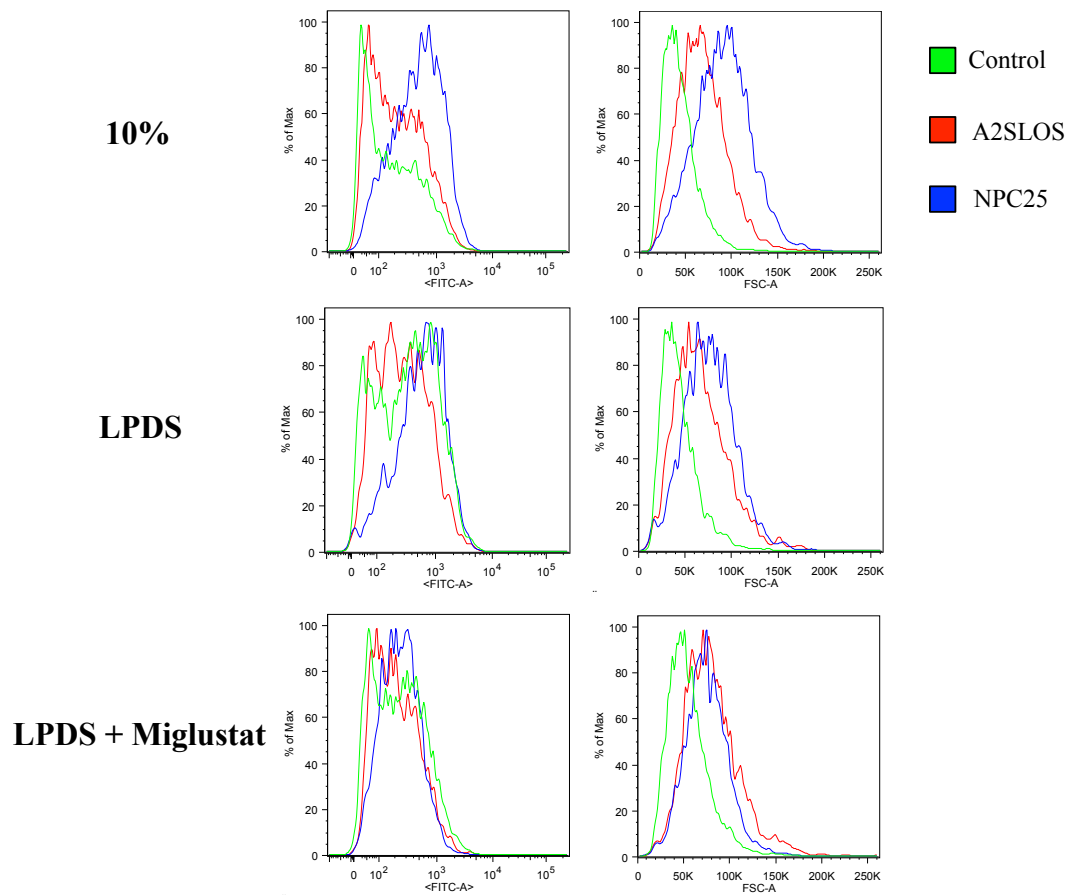


Figure 4.6: FACS plots from control, A2SLOS and NPC25 fibroblasts

FACS analysis of cholera toxin staining (left column, FITC-A) and cell size (right column, FSC-A). Control cells (green) have little change in either cholera toxin staining or cell size. A2SLOS cells (red) have increased cholera toxin staining in LPDS, which then decreases slightly in miglustat. Cell size of A2SLOS has a marginal decrease in LPDS versus 10% FBS, but then increases with the addition of miglustat. NPC25 cells (blue) have high cholera toxin staining in both 10% FBS and LPDS but this decreases to control levels in LPDS + miglustat. Cell size remains relatively constant, with only a slight decrease in the LPDS condition.

4.4.vii: RNAseq - results

The FACS analysis showed that miglustat was having an effect on the cells.

Therefore, the RNA could then be analysed via RNAseq to discover any changes in gene expression levels.

Two main values were gained from the RNAseq data; log₂ fold change and a normalised *p*-value (*padj*). This value is the *p*-value adjusted for the false discovery rate using the Benjamini-Hochberg procedure. Control cells grown in cholesterol-deficient media (LPDS) show several gene expression changes (**Figure 4.7**).

However, the number of genes with altered expression is greatly increased in A2SLOS and NPC25 compared to control. As A2SLOS cells require LPDS in order to display the disease phenotype, all cell lines were grown in LPDS with the addition of miglustat. They were then compared to their counterparts grown in LPDS to discover the effect of miglustat on gene expression. As can be seen in **Figure 4.7**, miglustat had very limited effect on control cells, with only two genes having a significant expression change above the ± 1 log₂ fold change threshold (*HEYL* and *OAS3*). Both A2SLOS and NPC25 had significantly more genes with altered expression.

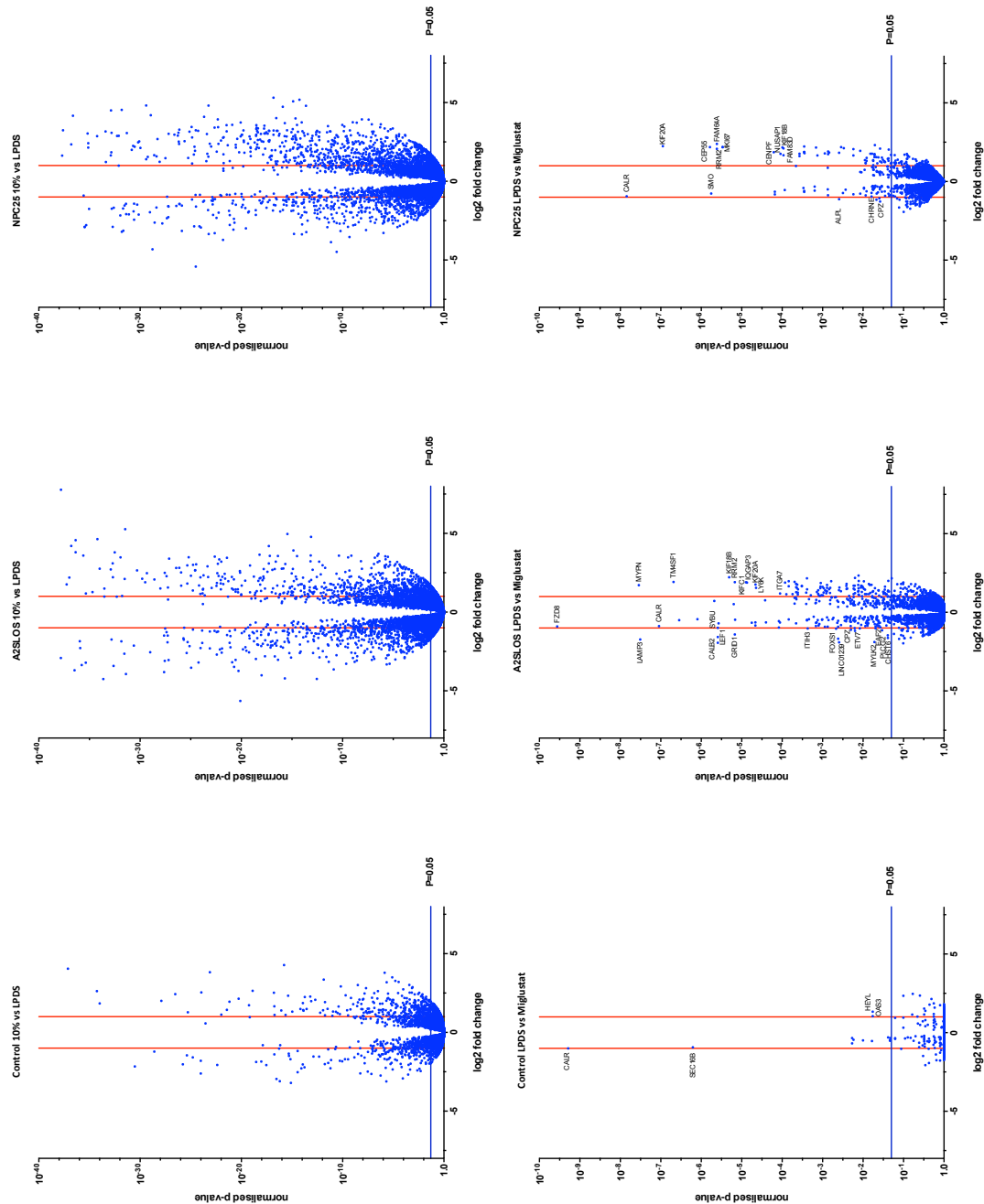


Figure 4.7: Volcano plots of RNAseq data

Figure 4.7 (previous page): Volcano plots of RNAseq data

The top row of images show the gene expression changes when control (left), A2SLOS (middle) and NPC25 (right) cells are grown in LPDS. The blue horizontal line represent the p -value of 0.05 and the red vertical lines show the $\pm 1.0 \log_2$ fold change. It can be seen that there are multiple significant changes in all three cell lines, but to a greater extent in A2SLOS and NPC25 compared to control.

The bottom row of images show the genes expression change as a result of the addition of miglustat. Control cells (left) have only two genes that pass both the p -value and fold change thresholds. A2SLOS (middle) and NPC25 (right) have multiple genes with significantly altered expression.

Out of the 21,836 genes analysed, A2SLOS cells were found to have 342 genes with a significant normalised p -value (>0.05) but only 131 of these had a $\log_2$ fold change of ± 1 (17 genes down-regulated, 114 genes up-regulated). If this threshold is tightened further to ± 1.5 , 58 genes have altered expression (8 genes down regulated, 50 genes up regulated).

NPC25 also had 21,836 genes analysed but had 145 with significant normalised p -values (>0.05) but only 65 of these had a $\log_2$ fold change of ± 1 (4 genes down-regulated, 61 genes up-regulated). If this threshold is tightened further to ± 1.5 , 43 genes have altered expression (all genes up-regulated).

Running the entire data set through GSEA using a ranked system, the changes in gene expression caused by the addition of miglustat can be split into hallmark pathways. If all 21,836 genes for both A2SLOS and NPC25 are run through the

software, 27 hallmark pathways are found; 9 for A2SLOS, 9 for NPC25 and 9 for both A2SLOS and NPC25. Interestingly, for the pathways denoted for both A2SLOS and NPC25, 4 pathways are up regulated (G2M Checkpoint, E2F Targets, Mitotic Spindle and Inflammatory Response), 2 are down regulated (Unfolded Protein Response, P52 Pathway) but the remaining 3 have opposite expression changes; Protein Secretion and MTORC1 Signalling (down regulated in A2SLOS and up regulated in NPC25) and Myogenesis (up regulated in A2SLOS and down regulated in NPC25). This ranking system takes into account genes that do not have significant change in expression. If the analysis is limited to only genes that have a significant p-value and a log2 fold change of ± 1.0 , then the results are reduced to 6 pathways; 5 for A2SLOS and 1 for both A2SLOS and NPC25 (**Figure 4.8**).

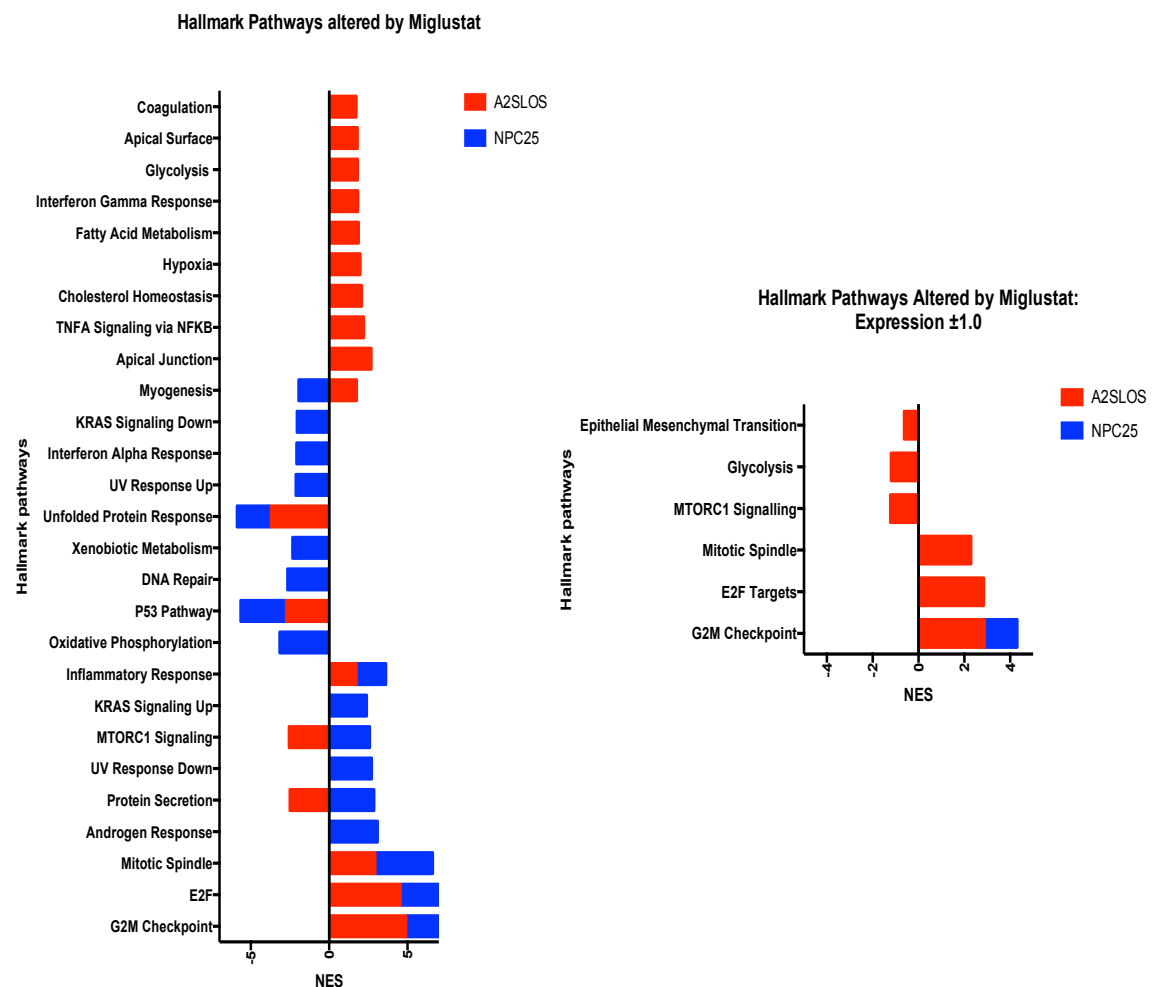


Figure 4.8: GSEA analysis of A2SLOS and NPC25 fibroblasts

Both graphs show the hallmark pathways against the normalized enrichment score (NES). NES is a value given by the GSEA software.

- A) This graph shows the pathways are found by analyzing all genes in the RNAseq data set for both A2SLOS (red) and NPC25 (blue).
- B) When the dataset is restricted to only significant changes, only 6 hallmark pathways are found.

To assess these results more closely, the functions of the most prominent genes were examined. From the volcano plots, there are certain genes that stand out (**Table 4.3A**). Although many of the expression changes are exacerbations of the genes altered in only LPDS, some genes have opposite expression (**Table 4.3B**).

A

A2SLOS			NPC25	
Up	Down		Up	Down
MYPN	LAMP3		KIF20A	
TM4SF1	CALB2		CEP55	
KIF18B	LEF1		RRM2	
RRM2	GRID1		FAM64A	
KIFC1			MKI67	
IQGAP3				
KIF20A				
LYK6				

B

A2SLOS	Function	NPC25	Function
ALDH8A1	Involved in retinoic acid biosynthesis	ALPL	May play a role in bone mineralisation
CPZ	Involved in the Wnt signaling pathway	CPZ	Involved in the Wnt signaling pathway
CRABP2		RPS4Y1	Component of the 40S ribosomal subunit
DES	Binds to MYPN, a muscle protein		
ETV7	Transcriptional repressor of genes involved in cell proliferation		
ITIH3	Can bind hyaluronic acid		
KRT15	Part of the keratin gene family		
KRT19			
KRT34			
LEF1			
MC4R	May increase neurogenesis		
NGFR	Necessary for the circadian oscillation of clock genes and glucose and lipid metabolism in liver		
SLC14A1	Mediates urea transport		

down to up regulation

up to down regulation

Table 4.3: Further in depth gene analysis

- A) This table shows the most significant genes from the dataset and pass the threshold of $\pm 1.0 \log_2$ fold change. The left hand side of the table shows the genes from A2SLOS and the right-hand side shows the genes from NPC25.
- B) This table shows differentially expressed genes in A2SLOS and NPC25 when compared between cells grown in LPDS and LPDS + miglustat. The left-hand side of the table shows the genes changed in A2SLOS and the right-hand side shows the genes changed in NPC25. The genes that were down regulated in LPDS but up regulated with the addition of miglustat are shaded green. The genes that were up regulated in LPDS but down regulated with the addition of miglustat are shaded red.

4.4: Discussion

4.4.i: Oxidative Stress Array

Oxidative stress has previously been implicated in SLOS but the data in this study does not support this suggestion. From the RT² PCR Profiler Array for Oxidative Stress, three genes were significant and only one of these was additionally above the log2 fold change threshold. However, previous studies have used different models of SLOS, which may cause a more severe and less varied phenotype than that seen in *Dhcr7*^{T93M/Δ3-5} model. This could cause oxidative stress markers to be present at more elevated levels than are seen in this study. Also, as there is variability within the *Dhcr7*^{T93M/Δ3-5} model, it is possible that differences in expression between control and *Dhcr7*^{T93M/Δ3-5} may become apparent if testing numbers are increased. However, there was not large variability in expression between the four controls and *Dhcr7*^{T93M/Δ3-5} mice to suggest that would be the case for the majority of genes in the assay.

The one gene that showed some variation between the samples is *Sod3*. *Sod3* had a log2 fold change of 0.48 in the profiler array and, although not above the threshold, it was one of three genes with significant *p*-values and therefore was validated using six additional cerebral cortexes. There were no significant change in E18.5 severe mice but in the 3-month *Dhcr7*^{T93M/Δ3-5} model there was a significant increase one mouse with a higher log2 fold change. This variation in the *Dhcr7*^{T93M/Δ3-5} model is not unexpected and agrees with what was seen in the histology work (**Chapter 3**).

Proteomics studies on SOD1 had previously shown the protein to be up regulated in both *Dhcr7*^{Δ3-5/Δ3-5} E18.5 mice and CSF from patients. There was no significant difference in my array in *Sod1* or *Sod2* but there was in *Sod3*. I therefore investigated *Sod3* in E18.5 *Dhcr7*^{Δ3-5/Δ3-5} mice, but there was no significant difference in the expression of *Sod3*. There are several suggestions as to why this is the case. There are multiple pathways that can cause protein levels to change but leave RNA expression unaltered. As noted by Maier *et al.* [239], transcription and translation do not always have a simple correlation and are affected by multiple processes, including the half-life of the protein, regulatory proteins and experimental error and noise. Also, as there is variation in the SLOS mouse models and only three mice were analysed, more mice are required to draw firm conclusions.

The gene that was significant in both *p*-value and log2 fold change from the RT² Profiler Array was *Epx*, eosinophil peroxidase. Expression was significantly decreased, with a log2 fold change of -2.59, in the *Dhcr7*^{T93M/Δ3-5} cortex compared to controls. However, when the protein level was investigated via Western Blot, there was no difference between control and *Dhcr7*^{T93M/Δ3-5} mice. As noted above, there are multiple pathways that can cause protein levels to change but leave RNA expression unaltered. Additionally, the western blot is based off an N of 1 and, therefore, only limited conclusions can be drawn. The difference between RNA and protein expression is intriguing and would benefit from further investigation, both by expanding the number of samples studied and by additional techniques, such as mass spectrometry. In contrast to the murine data, the RNAseq human data did not

support the change in expression of *EPX*. The Allen Mouse Brain Atlas [240] does not show a large amount of *EPX* expression in the brain. If *EPX* is only present in low concentrations, even a small change would result in a large fold change, but this may not be biologically significant. Additionally, the experiments were also performed on different tissues and it is possible that there are differences between the mouse model and patients.

4.4.ii: RNAseq – Miglustat study

Miglustat is an approved therapy for NPC in Europe and is noted to reduce GSL accumulation in lysosomes. SLOS is suggested to have an NPC1-like phenotype and therefore miglustat may have therapeutic benefits in SLOS [241]. If patient fibroblasts are treated with miglustat, it has been reported that the drug decreases filipin staining in SLOS but not NPC. This suggests that the drug may be acting via different pathways in the two disorders and this was investigated using RNAseq.

Sequencing was performed at 100x coverage and was a 2 x 100bp paired-end read. Paired-end sequencing allows additional positioning information and enables improved identification of splice variants and CSNBs. However, it is a more costly and time-consuming process as paired-end runs read from one end to the other and then read another round starting from the opposite end. The resulting sequences were then aligned using RNA-star [236] against the Human genome and gene expression levels were counted using subhead featureCounts [237]. As artefacts and biases can alter the results, normalisation has to occur.

Only one sample, with three replicates, was analysed per condition, therefore, as there is high uncertainty of within-group variance, inferential statistical methods cannot not be used. There are many bioinformatics methods available for normalisation but we used DESeq2.

DESeq2 corrects for library size and RNA composition bias, which can arise when a small number of genes are very highly expressed in one experiment condition but not in another. It additionally fits negative binomial generalised linear models for each gene and uses the Wald test for significance testing. It automatically counts outliers using Cooks' distance and removes these genes from analysis, as well as removing genes whose mean of normalized counts is below a certain threshold. Therefore, the first step was to normalise each replicate as a percentage of total mRNA and these values were then averaged across replicates within each genotype. DESeq2 uses these values and the resulting errors to calculate the significance of differences between genotypes.

GSEA software was used to compute the biological effect of the expression changes calculated by DESeq2. The GSEA software uses four main statistical approaches to create an analysis report. These are Enrichment Score, Normalised Enrichment Score (NES), Nominal P Value and False Discovery Rate. The FDR is the estimated probability that a gene set with a given NES represents a false positive finding. For example, if the FDR is 25%, the result is valid 3 out of 4 times. I used an FDR cut-off of 5% as I have a small number of samples and therefore needed a more stringent significance cut-off.

Control cells showed very little gene expression change when miglustat was added to the cells, showing that the drug is affecting the specific disease. More genes were affected in A2SLOS compared to NPC25, potentially confirming the suggestion that the drug acts through different pathways in each disorder. However, in order to determine if this was the case, detailed analysis was carried out on the data.

In both disorders, the majority of genes with aberrant expression were involved in the cell cycle, particularly those associated with mitotic spindles and the G2M checkpoint. The function of the up regulated genes associated with these pathways was to increase cell growth and proliferation. In agreement with this, changes in gene expression to repress apoptosis were also observed, for example, *FOXSI* is an inhibitor for *FOXO3* and *FOXO4* expression. These latter proteins are involved in the up regulation of genes required for apoptosis therefore *FOXSI* indirectly inhibits apoptosis. This correlates with the up regulation of *BIRC5*, which is an inhibitor of apoptosis.

From the volcano plots, certain genes with the lowest p-values, i.e. most significance can be easily picked out. In A2SLOS, there were 8 of these that were up-regulated and 4 that were down-regulated. In NPC25, 5 were up-regulated. In both disorders, the majority of up-regulated genes were involved in cell proliferation (*TM4SF*, *KIF18B*, *KIFC1*, *KIF20A*, *FAM64A*, *CEP55*, *MKI67*, *IQGAP3* and *RRM2*). This agrees with the observations, reviewed by Hakomori [235], that glycolipids are involved in cell growth regulation. Retinoid compounds

induce differentiation and genes involved in retinoic acid synthesis, such as *ALDH8A1*, are up-regulated in A2SLOS and NPC25 cells treated with miglustat.

Interestingly, *RRM2* also has a role in the inhibition of WNT signalling and the WNT signalling pathway has been shown to have differential expression in both SLOS and NPC [215,242]. In control cells grown in LPDS, *WNT9A* is the only WNT gene affected, with a down-regulation of -1.19. On the other hand, A2SLOS cells grown in LPDS have highly significant differential expression of *WNT10B* (-2.23), *WNT4* (-2.31) and *WNT2* (1.18), in addition to *WNT9A* (-2.25). Similarly, NPC25 fibroblasts grown in LPDS have highly significant differential expression of *WNT16* (2.15), *WNT2* (1.37), *WNT4* (-2.44) and *WNT9a* (-1.74).

Multiple other genes involved in the WNT pathway are additionally up- or down-regulated in A2SLOS and NPC25 but not control cells. For example, *FZD8* has the most significant gene expression change in A2SLOS and, although not below the -1.0 threshold, *FZD8* has a log2 fold change of -0.90. *CRABP2* and *LEF1* (one of the most significantly down-regulated genes in A2SLOS) are also key members of the WNT pathway. Interestingly, miglustat reverses expression of these genes; *CPZ* and *RRM2* (A2SLOS and NPC25) and *LEF1* (A2SLOS) are up-regulated in LPDS but down-regulated with the addition of miglustat. Therefore miglustat may help correct the aberration in this pathway but more research will be needed to discover if this is the case.

With regards to the second most significant gene expression change in A2SLOS, *MYPN* is a muscle protein with the ability to bind other proteins, such as DES.

Correspondingly in A2SLOS, *MYPN* is the most significantly up-regulated gene, with a log2 fold change of 1.73, and up-regulation of *DES* is also observed (log2 fold change of 1.29) although the significance of this is not certain. Other highlighted genes are involved in keratin synthesis (*KRT15*, *KRT19* and *KRT34*) but previous studies have identified false positivity associated with these genes, including the presence of pseudogenes with sequence similarity [243]. As such, the up-regulation of these genes in this study may be considered artefacts.

MC4R and NGFR are receptors of the nervous system. Antagonists of MC4R have been found to increase neurogenesis and cognitive recovery in Alzheimer's disease [244,245]. NGFR is suggested to be involved in the circadian oscillations of clock genes and this could correspond to the sleep-cycle disturbance seen in SLOS patients.

The most significantly down regulated genes in A2SLOS were *LAMP3*, *CALB2* and *GRID1*. These all have a role in signal transmission or function of neural cells. More study will need to be carried out to enable the result of the down regulation of these genes. No genes had highly significant down regulation in NPC25.

Although not among the highly significant genes, *RPS4Y1* is significantly down-regulated in NPC25 fibroblasts grown in LPDS but up-regulated with the addition of miglustat. This gene encodes for a component of the 40S ribosomal subunit. An interesting point to note is that in an NPC1 yeast synthetic lethal screen (work carried out by Alexandria Colaco, DPhil candidate, University of Oxford), ribosomal proteins were also highlighted.

Although there are some potentially interesting findings in the RNAseq study, the significantly altered genes, such as those involved in the cell cycle and WNT pathway, need to be validated via other methods, for example, using TaqMan probes. If the changes in expression are confirmed, further experiments would be required to test the resulting protein levels of these genes to ascertain if there are any correlations. Additionally, other cell lines of varying severity levels would need to be investigated to help gain the extent of biological significance. If these results are shown to be significant, they could provide novel insights into the pathogenic mechanisms of SLOS and NPC.

Chapter 5: Concluding Remarks

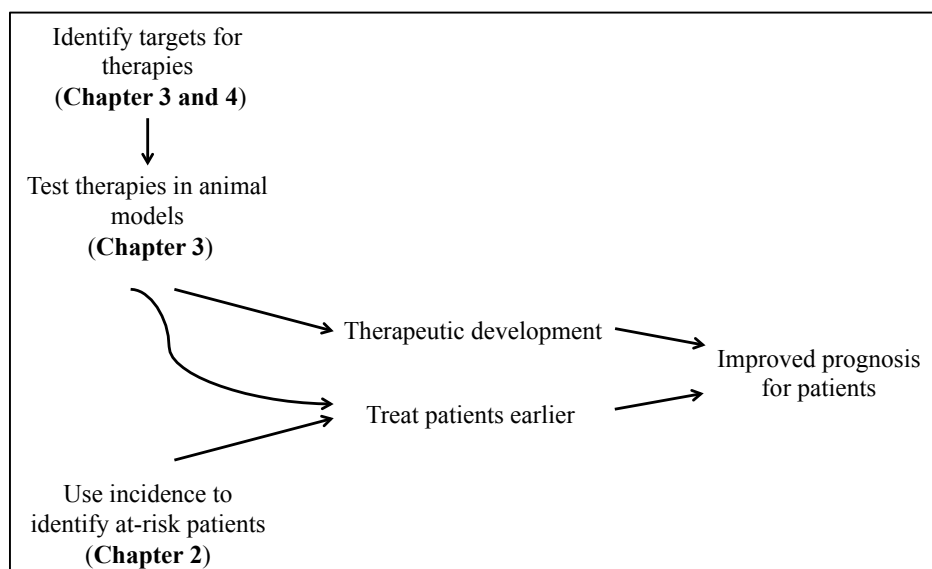


Figure 1: Summary figure showing how the work in this thesis can impact future study.

Research into the properties and biological functions of cholesterol has been a continuous challenge for over 250 years. As with many things in life, cholesterol levels in the body need to be balanced, as exemplified in SLOS and NPC. Although both are monogenic disorders caused by defective cholesterol homeostasis, SLOS has decreased levels of cholesterol and increased levels of the cholesterol precursor 7DHC, while NPC has elevated storage of un-esterified cholesterol in late endosomes/lysosomes. Despite resulting from very different defects in cholesterol homeostasis, both disorders are devastating and have distinctive phenotypes.

Not only is cholesterol critical for correct development, it is also the precursor to multiple important molecules that play a role throughout life. **Chapter 1** reviewed the

current literature regarding cholesterol metabolic pathways, its regulation and defective homeostasis in SLOS and NPC. Although research into both disorders has been extensive, limited therapeutic options are available for both monogenic diseases.

In order for a treatment to have the greatest effect, it requires administration within an optimum time window. Therefore, the earlier the treatment is administered, the more likely it is to maximise clinical benefit. Early treatment is also important in degenerative disorders, such as NPC, because if treatment is administered before the majority of degeneration has occurred, the disease course has the potential to be stabilised.

Monogenic disorders follow the principle of Mendelian genetics and can usually be followed through family trees. As SLOS and NPC are both recessive disorders, if these families are not closely followed, many people could be unknown carriers and at risk of passing the disorder to their children. Additionally, both disorders also have broad severity spectrums, ranging from *in utero* lethality to late-onset mild cases. With so much variation, how can “at risk” populations be diagnosed early enough to be treated before the onset of symptoms?

By estimating the incidence of monogenic disorders, the potentially “at risk” populations can be identified. This constricts the number of individuals to genetically counsel and screen, therefore reducing costs, preventing affected cases and increasing feasibility. **Chapter 2** investigated the novel use of massively parallel sequencing data sets to estimate incidence of SLOS, NPC1 and NPC2. Multiple causative mutations have been discovered for these disorders but carrier frequencies of individual mutations

vary across geographical populations. This chapter used existing population data from the 1000 genomes project to exemplify how the carrier frequency of the most frequent SLOS mutation, c.964-1G>C, had carrier frequencies ranging from 0.46% - 0.81% depending on the population studied. Accurate population genetics requires data from a large cohort and, although the time taken and cost of sequencing has dramatically reduced over recent years, it is still a significant investment. This exemplifies the scientific and social benefit that can be derived from making multiple experiments available to other researchers.

From the four data sets used in **Chapter 2**, a maximum of 17,836 alleles were analysed. The cost of sequencing this number of alleles for an individual study would be significant but, as this data was already available, it added great depth to the study at minimal cost. This chapter discovered that the estimated incidences of SLOS and NPC were approximately the same as that reported from previous clinical studies, 1/40,000 and 1/89,000 respectively. However, a late-onset form of NPC1 may be more prevalent than previously thought, with an estimated incidence as high as 1/19,077. In this case, many patients may be being misdiagnosed or missed due to atypical presentation or mild symptoms. If these patients can be identified, this would aid the administration of the most effective treatments.

Chapters 3 and 4 investigated aspects of the methodology involved in development of therapies. After the development of novel therapeutic intervention, before progressing to clinical trials, drugs must be tested in animal models to test efficacy and toxicity. One useful, non-invasive technique is to observe alterations in behaviour. As the behaviour of the *Dhcr7*^{T93M/Δ3-5} SLOS mouse model has not previously been

characterised, **Chapter 3** began to identify if the behavioural characteristics replicate those seen in patients. As the mice are bred on a mixed background, tentative conclusions can be drawn, but a homogenous background model is currently being developed that may tighten the data. Interestingly, the isogenic C57BL/6 *Dhcr7*^{T93M/Δ3-5} model does not have the same survival rate seen when the same genotype is present on a mixed background. This led to the suggestion that genetic modifiers may be involved and this could potentially explain some of the variation seen in this study.

Corresponding to patient behaviour, hyperactivity was observed in *Dhcr7*^{T93M/Δ3-5} mice, though to a greater extent in females. In contrast, *Dhcr7*^{T93M/Δ3-5} mice exhibited decreased anxiety, which is contradictory to patient studies. These data may be explained by the neurological studies of the *Dhcr7*^{T93M/Δ3-5} brains, which revealed abnormalities of the anterior commissure and hippocampus.

Interestingly, the hippocampus contains multiple serotonin receptors and inhibitors of serotonin uptake have been shown to decrease Nestlet shredding in mice [210]. If the RNAseq data (**Chapter 4**) is also taken into consideration, 5-hydroxytryptamine (serotonin) receptor 2A (*HTR2A*) has a log2 fold change of -1.64 when A2SLOS cells are grown in LPDS compared to 10% FBS (i.e. when A2SLOS cells present with the SLOS phenotype). In contrast, 5-hydroxytryptamine receptor 1B (*HTR1B*) has a log2 fold change of 1.97. Although it is currently not certain how serotonin may contribute to the phenotype of SLOS, if this is established, it could provide a potential therapeutic target.

Prior neurological studies of SLOS patients have also exposed structural defects, although these exact changes were not found in the *Dhcr7*^{T93M/Δ3-5} model. Nevertheless, the structural differences that were found may be used as a guide for areas that justify further research. For example, in both patients and the *Dhcr7*^{T93M/Δ3-5} model there was atrophy of white matter; the corpus callosum in patients and anterior commissure in the mouse model. Myelin is the main component of white matter, comprising 70% lipids, 25% of which is cholesterol [45]. It is not known if myelin is altered in SLOS, but the data from both this and previous studies suggests that it is an area worthy of further investigation.

The main focus of the RNAseq data in **Chapter 4** was to assess the pathways altered by miglustat in SLOS and NPC patient fibroblasts. This study highlighted that the two main pathways altered by miglustat were cell growth and WNT signalling. Miglustat alters the glycolipid pathway and, as studies have shown that this pathway is involved in cell proliferation, this could be the cause of the altered genes involved in cell growth. The altered genes in the WNT signalling pathway found from the study is an interesting result as previous studies in iPS and neuronal cells have found alterations in WNT signalling in SLOS [215] and NPC [242]. As WNT signalling is crucial to cell development [246], treatment would likely have most effect if delivered early in life, emphasising the importance of identifying patients (**Chapter 2**). However, as treatment of the WNT signalling pathway is complex [247], and the RNAseq data was only based on three cell lines, a thorough investigation into the alterations of this pathway in SLOS and NPC would be required.

The data in this thesis was specified towards SLOS and NPC, yet many of the techniques employed could be translated to other recessive monogenic disorders, for example, the estimation of incidence studied in **Chapter 2** and the investigation into drug mechanisms using RNAseq (**Chapter 4**). The RNAseq methods could also be used to determine pathways of additional drug treatments, such as curcumin, a drug that has been shown to be effective in NPC [135], but also has an affect on the WNT pathway [247]. Once the behavioural phenotype of the *Dhcr*^{T93M/Δ3-5} mouse model (on the homozygous or heterozygous background) has been fully elucidated, these treatments could be administered to the mice and alterations in the behaviour and neurological structure observed.

Ultimately, the goal for all clinical research is to improve the prognosis, if not cure, rare and common disorders. The research contained in this thesis is only the tip of the iceberg in the vast ocean that is the development of therapies for human diseases. Despite this, each development, however small, helps understand the world around us and to create a better quality of life for patients suffering from devastating disorders.

“Research means to investigate something you do not know or understand”

“Research is creating new knowledge.”

————— Neil Armstrong —————

Future Direction

This thesis showed how existing datasets could be utilised in order to estimate incidence of rare, monogenic disorders. However, this approach was not ideal as multiple literature searches and molecular analyses were required, although it was useful for analysing population genetics. If future studies include greater population data, these datasets could have a use for analysis of population genetics of Mendelian disorders. This would identify populations that had the greatest risk and would potentially enable undiagnosed patients to be discovered. Using existing large datasets also could be used to identify secondary risk factors, for example, the SLOS study highlighted a gene associated with obesity. This could help guide health decisions.

This thesis began to explore other avenues to discover the mechanisms behind the pathogenesis of SLOS and NPC. RNAseq is a powerful tool that can analyse the transcriptome effectively and at a relatively low cost. This study used one drug treatment, miglustat, on only one sample cell line from control, SLOS and NPC populations. However, the results from this study require validation using other real-time PCR methods and also proteomics analysis to discover how the gene expression relates to protein levels. Future directions would include analysing multiple cell lines across the severity spectrums of SLOS and NPC to establish if there are differences in the transcriptomes. Additionally, it would be useful to test other drug treatments, for example the antioxidant δ -tocopherol or cyclodextrin. As was seen in the miglustat study, this could highlight further avenues in which to explore novel therapeutic interventions.

Currently, there is not an ideal SLOS mouse model in which to test drug treatments. However, murine models using the *Dhct*^{T93M/Δ3-5} genotype are being developed on two isogenic backgrounds. Initial observations of these mice have produced some surprising results, with the C57BL/6 isogenic model dying early in life, but the FVB/N isogenic background model produces survival similar to the mixed background model. This suggests that there may be one or more genetic modifiers and therefore future study should aim to identify them by analysing the genetic profiles of these models.

References

1. McMichael J, Brigden WW. Diseases of the Cardiovascular System (Medical). *Annu Rev Med* 1950;1:61–78.
2. Scirica BM, Cannon CP. Treatment of Elevated Cholesterol. *Circulation* 2005;111.
3. Orth M, Bellosta S. Cholesterol: Its Regulation and Role in Central Nervous System Disorders. *Cholesterol* 2012;2012:1–19.
4. Cooper GM, Hausman RE. *The Cell: A Molecular Approach* 2nd Edition. 2007.
5. Porter FD, Herman GE. Malformation syndromes caused by disorders of cholesterol synthesis. *J Lipid Res* 2010;52:6–34.
6. Farese, Jr R V, Herz J. Cholesterol metabolism and embryogenesis. *Trends Genet* 1998;14:115–120.
7. King M. Cholesterol: Synthesis, Metabolism, Regulation [Internet]. 2016; Available from: <http://themedicalbiochemistrypage.org/cholesterol.php>
8. Goldstein JL, Brown MS. Regulation of the mevalonate pathway. *Nature* 1990;343:425–430.
9. Dietschy JM. Regulation of cholesterol metabolism in man and in other species. *Klin Wochenschr* 1984;62:338–345.
10. Rashid R, Ameer F, Kalbacher H, Scandiuizzi L, Zaidi N. Aberrant de novo cholesterologenesis: Clinical significance and implications. *Clin. Chim. Acta*. 2015;450:356–361.
11. Liscum L. Chapter 14 - Cholesterol biosynthesis. 2008.

12. Friesen JA, Rodwell VW. The 3-hydroxy-3-methylglutaryl coenzyme-A (HMG-CoA) reductases. *Genome Biol* 2004;5:248.
13. Liscum L, Finer-Moore J, Stroud RM, Luskey KL, Brown MS, Goldstein JL. Domain structure of 3-hydroxy-3-methylglutaryl coenzyme A reductase, a glycoprotein of the endoplasmic reticulum. *J Biol Chem* 1985;260:522–30.
14. Olender EH, Simon RD. The intracellular targeting and membrane topology of 3-hydroxy-3-methylglutaryl-CoA reductase. *J Biol Chem* 1992;267:4223–4235.
15. Liscum L, Cummings RD, Anderson RG, DeMartino GN, Goldstein JL, Brown MS. 3-Hydroxy-3-methylglutaryl-CoA reductase: a transmembrane glycoprotein of the endoplasmic reticulum with N-linked “high-mannose” oligosaccharides. *Proc Natl Acad Sci USA* 1983;80:7165–7169.
16. Bloch K, Rittenberg D. On the Utilization of Acetic Acid for Cholesterol Formation. *J Biol Chem* 1942;145:625–636.
17. Kandutsch AA, Russell AE. Preputial gland tumor sterols. 3. A metabolic pathway from lanosterol to cholesterol. *J Biol Chem* 1960;235.
18. Mitsche MA, McDonald JG, Hobbs HH, Cohen JC. Flux analysis of cholesterol biosynthesis in vivo reveals multiple tissue and cell-type specific pathways. *Elife* 2015;4:1–21.
19. Goldstein JL, Brown MS. The Low-Density Lipoprotein and it’s relation to Atherosclerosis. *Ann Rev Biochem* 1977;46:897–930.
20. Gil G, Faust JR, Chin DJ, Goldstein JL, Brown MS. Membrane-bound domain of HMG CoA reductase is required for sterol-enhanced degradation of the enzyme. *Cell* 1985;41:249–258.
21. Chun KT, Simoni RD. The role of the membrane domain in the regulated degradation of 3-hydroxy-3-methylglutaryl coenzyme A reductase. *J Biol Chem*

- 1992;267:4236–4246.
22. Jingami H, Brown MS, Goldstein JL, Anderson RGW, Luskey KL. Partial Deletion of Membrane-bound Domain of 3-Hydroxy-3-methyl- glutaryl Coenzyme A Reductase Eliminates Sterol-enhanced Degradation and Prevents Formation of Crystalloid Endoplasmic Reticulum. 1987;104:1693–1704.
 23. Xu L, Simoni RD. The inhibition of degradation of 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) reductase by sterol regulatory element binding protein cleavage-activating protein requires four phenylalanine residues in span 6 of HMG-CoA reductase transmembrane domain. Arch Biochem Biophys 2003;414:232–43.
 24. Hua X, Nohturfft A, Goldstein JL, Brown MS. Sterol resistance in CHO cells traced to point mutation in SREBP cleavage-activating protein. Cell 1996;87:415–426.
 25. Liscum L. Cholesterol biosynthesis. In: Biochem. Lipo. Mem. 2002.
 26. Nohturfft A, Yabe D, Goldstein JL, Brown MS, Espenshade PJ. Regulated step in cholesterol feedback localized to budding of SCAP from ER membranes. Cell 2000;102:315–323.
 27. Shimano H. Sterol regulatory element-binding proteins (SREBPs): transcriptional regulators of lipid synthetic genes. Prog Lipid Res 2001;40:439–52.
 28. Brown MS, Goldstein JL. Cholesterol feedback: from Schoenheimer’s bottle to Scap’s MELADL. J Lipid Res 2009;50 Suppl:S15–S27.
 29. Roux C, Wolf C, Mulliez N, Gaoua W, Cormier V, Chevy F, Citadelle D. Role of cholesterol in embryonic development. Am J Clin Nutr 2000;71:1270S–9S.
 30. Gurdon JB, Bourillot PY. Morphogen gradient interpretation. Nature 2001;413:797–803.

31. Jeong J, McMahon AP. Cholesterol modification of Hedgehog family proteins. *J. Clin. Invest.* 2002;110:591–596.
32. Echelard Y, Epstein DJ, St-Jacques B, Shen L, Mohler J, McMahon JA, McMahon AP. Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the regulation of CNS polarity. *Cell* 1993;75:1417–1430.
33. Porter JA, Young KE, Beachy PA. Cholesterol modification of hedgehog signaling proteins in animal development. *Science* 1996;274:255–259.
34. Porter JA, Ekker SC, Park WJ, Von Kessler DP, Young KE, Chen CH, Ma Y, Woods AS, Cotter RJ, Koonin E V., Beachy PA. Hedgehog patterning activity: Role of a lipophilic modification mediated by the carboxy-terminal autoprocessing domain. *Cell* 1996;86:21–34.
35. Roelink H, Porter JA, Chiang C, Tanabe Y, Chang DT, Beachy PA, Jessell TM. Floor plate and motor neuron induction by different concentrations of the amino-terminal cleavage product of sonic hedgehog autoproteolysis. *Cell.* 1995;81:445–455.
36. Gilbert SF. *Differentiation of the Neural Tube* (6th Edition). Sinauer Associates; 2000.
37. Riddle RD, Johnson RL, Laufer E, Tabin C. Sonic hedgehog mediates the polarizing activity of the ZPA. *Cell* 1993;75:1401–1416.
38. Tickle C, Barker H. *The Sonic hedgehog gradient in the developing limb*. Wiley Interdiscip Rev Dev Biol 2013;2:275–290.
39. R&D Systems Inc. Sonic Hedgehog: a morphogen involved in axon guidance [Internet]. Available from: <https://www.rndsystems.com/resources/articles/sonic-hedgehog-morphogen-involved-axon-guidance>

40. Lum L, Beachy PA. The Hedgehog response network: sensors, switches, and routers. *Science* 2004;304:1755–9.
41. Burke R, Nellen D, Bellotto M, Hafen E, Senti K a, Dickson BJ, Basler K. Dispatched, a novel sterol-sensing domain protein dedicated to the release of cholesterol-modified hedgehog from signaling cells. *Cell* 1999;99:803–815.
42. Murone M, Rosenthal A, de Sauvage FJ. Sonic hedgehog signaling by the patched-smoothened receptor complex. *Curr Biol* 1999;9:76–84.
43. Rimkus TK, Carpenter RL, Qasem S, Chan M, Lo H-W. Targeting the Sonic Hedgehog Signaling Pathway: Review of Smoothened and GLI Inhibitors. *Cancers (Basel)* 2016;8:22.
44. Ohvo-Rekila H, Ramstedt B, Leppimäki P, Peter Slotte J. Cholesterol interactions with phospholipids in membranes. *Prog. Lipid Res.* 2002;41:66–97.
45. Saher G, Brügger B, Lappe-Siefke C, Möbius W, Tozawa R, Wehr MC, Wieland F, Ishibashi S, Nave K-A. High cholesterol level is essential for myelin membrane growth. *Nat Neurosci* 2005;8:468–475.
46. Rezen T, Rozman D, Pascucci JM, Monostory K. Interplay between cholesterol and drug metabolism. *Biochim Biophys Acta - Proteins Proteomics* 2011;1814:146–160.
47. Griffiths WJ, Abdel-Khalik J, Hearn T, Yutuc E, Morgan AH, Wang Y, Javitt NB, Schroepfer GJ, Russell DW, Iuliano L, Murphy RC, Johnson KM, Leonarduzzi G, Gargiulo S, Gamba P, Testa G, Sottero B, Rossin D, Staurengli E, Poli G, Dietschy JM, Turley SD, Lütjohann D, Breuer O, Ahlborg G, Nennesmo I, Siden A, Diczfalusy U, Björkhem I, Janowski BA, Grogan MJ, Jones SA, Wisely GB, Kliewer SA, Corey EJ, Mangelsdorf DJ, Radhakrishnan A, Ikeda Y, Kwon HJ, Brown MS, Goldstein JL, Horton JD, Goldstein JL,

- Brown MS, Gill S, Chow R, Brown AJ, Wang Y, Muneton S, Sjövall J, Jovanovic JN, Griffiths WJ, Tint GS, Yu H, Shang Q, Xu G, Patel SB, Wang Y, Karu K, Meljon A, Turton J, Yau JL, Seckl JR, Wang Y, Griffiths WJ, Goyal S, Xiao Y, Porter NA, Xu L, Guengerich FP, Meljon A, Wang Y, Griffiths WJ, Jansen M, Wang W, Greco D, Bellenchi GC, Di PU, Brown AJ, Ikonen E, Theofilopoulos S, Wang Y, Kitambi SS, Sacchetti P, Sousa KM, Bodin K, Kirk J, Salto C, Gustafsson M, Toledo EM, Sacchetti P, Sousa KM, Hall AC, Liste I, Steffensen KR, Theofilopoulos S, Parish CL, Hazenberg C, et al. Current trends in oxysterol research. *Biochem Soc Trans* 2016;44:652–8.
48. Bielska AA, Schlesinger P, Covey DF, Ory DS. Oxysterols as non-genomic regulators of cholesterol homeostasis. *Trends Endocrinol Metab* 2012;23:99–106.
49. McLean KJ, Hans M, Munro AW. Cholesterol, an essential molecule: diverse roles involving cytochrome P450 enzymes. *Biochem Soc Trans* 2012;40:587–593.
50. Staels B, Fonseca VA. Bile acids and metabolic regulation: mechanisms and clinical responses to bile acid sequestration. *Diabetes Care* 2009;32:S237–S245.
51. Pikuleva IA. Cholesterol-metabolizing cytochromes P450. *Drug Metab Dispos* 2006;34:513–520.
52. Lipesheva GI, Waterman MR. CYP51 - The omnipotent P450. *Mol Cell Endocrinol* 2004;215:165–170.
53. Tint G. Cholesterol defect in Smith-Lemli-Opitz syndrome. *Am J Med Genet* 1993;15:573–4.
54. Kelley RI, Herman GE. Inborn errors of sterol biosynthesis. *Annu Rev Genomics Hum Genet* 2001;2:299–341.
55. Herman GE, Kratz L. Disorders of sterol synthesis: Beyond Smith-Lemli-Opitz

- syndrome. *Am J Med Genet Part C Semin Med Genet* 2012;160 C:301–321.
56. Bianconi SE, Cross JL, Wassif CA, Porter FD. Pathogenesis, Epidemiology, Diagnosis and Clinical Aspects of Smith-Lemli-Opitz Syndrome. *Expert Opin orphan drugs* 2015;3:267–280.
57. Tint GS, Irons M, Elias ER, Batta AK, Frieden R, Chen TS, Salen G. Defective cholesterol biosynthesis associated with the Smith-Lemli-Opitz syndrome. *N Engl J Med* 1994;330:107–13.
58. Yu H, Patel SB. Recent insights into the Smith-Lemli-Opitz syndrome. *Clin Genet [Internet]* 2005;68:383–91. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1350989&tool=pmcentrez&rendertype=abstract>
59. DeBarber AE, Eroglu Y, Merkens LS, Pappu AS, Steiner RD. Smith-Lemli-Opitz syndrome. *Expert Rev Mol Med* 2011;13:e24.
60. Nowaczyk MJM, Irons MB. Smith-Lemli-Opitz syndrome: Phenotype, natural history, and epidemiology. *Am J Med Genet Part C Semin Med Genet* 2012;160 C:250–262.
61. Keller RK, Arnold TP, Fliesler SJ. Formation of 7-dehydrocholesterol-containing membrane rafts in vitro and in vivo, with relevance to the Smith-Lemli-Opitz syndrome. *J Lipid Res* 2004;45:347–355.
62. Kovarova M, Wassif CA, Odom S, Liao K, Porter FD, Rivera J. Cholesterol deficiency in a mouse model of Smith-Lemli-Opitz syndrome reveals increased mast cell responsiveness. *J Exp Med* 2006;203:1161–71.
63. Shrivastava S, Paila YD, Dutta A, Chattopadhyay A. Differential effects of cholesterol and its immediate biosynthetic precursors on membrane organization. *Biochemistry* 2008;47:5668–5677.

64. Rossi M, Federico G, Corso G, Parenti G, Battagliese A, Frascogna AR, della Casa R, dello Russo A, Strisciuglio P, Saggese G, Andria G. Vitamin D status in patients affected by Smith-Lemli-Opitz syndrome. *J Inherit Metab Dis* 2005;28:69–80.
65. Bialer MG, Penchaszadeh VB, Kahn E, Libes R, Krigsman G, Lesser ML. Female external genitalia and müllerian duct derivatives in a 46,XY infant with the Smith-Lemli-Opitz syndrome. *Am J Med Genet* 1987;28:723–731.
66. Kelley RI, Hennekam RC. The Smith-Lemli-Opitz syndrome. *J Med Genet* 2000;37:321–35.
67. Lee RWY, Yoshida S, Jung ES, Mori S, Baker EH, Porter FD. Corpus callosum measurements correlate with developmental delay in Smith-Lemli-Opitz syndrome. *Pediatr Neurol* 2013;49:107–112.
68. Lee RWY, Conley SK, Gropman A, Porter FD, Baker EH. Brain magnetic resonance imaging findings in Smith-Lemli-Opitz syndrome. *Am J Med Genet Part A* 2013;161:2407–2419.
69. Tierney E, Nwokoro NA, Kelley RI. Behavioral phenotype of RSH/Smith-Lemli-Opitz syndrome. *Ment Retard Dev Disabil Res Rev* 2000;6:131–134.
70. Woollett LA. Maternal cholesterol in fetal development: Transport of cholesterol from the maternal to the fetal circulation. *Am. J. Clin. Nutr.* 2005;82:1155–1161.
71. Baardman ME, Kerstjens-Frederikse WS, Berger RMF, Bakker MK, Hofstra RMW, Plösch T. The role of maternal-fetal cholesterol transport in early fetal life: current insights. *Biol Reprod* 2013;88:24.
72. Gofflot F, Gaoua W, Bourguignon L, Roux C, Picard JJ. Expression of sonic hedgehog downstream genes is modified in rat embryos exposed in utero to a distal inhibitor of cholesterol biosynthesis. *Dev Dyn* 2001;220:99–111.

73. Lanoue L, Dehart DB, Hinsdale ME, Maeda N, Tint GS, Sulik KK. Limb, genital, CNS, and facial malformations result from gene/environment-induced cholesterol deficiency: further evidence for a link to sonic hedgehog. *Am J Med Genet* 1997;73:24–31.
74. Blassberg R, Macrae JI, Briscoe J, Jacob J. Reduced cholesterol levels impair Smoothed activation in Smith-Lemli-Opitz syndrome. *Hum Mol Genet* 2016;25:693–705.
75. Cooper MK, Wassif CA, Krakowiak PA, Taipale J, Gong R, Kelley RI, Porter FD, Beachy PA. A defective response to Hedgehog signaling in disorders of cholesterol biosynthesis. *Nat Genet* 2003;33:508–13.
76. Linck LM, Lin DS, Flavell D, Connor WE, Steiner RD. Cholesterol supplementation with egg yolk increases plasma cholesterol and decreases plasma 7-dehydrocholesterol in Smith-Lemli-Opitz syndrome. *Am J Med Genet* 2000;93:360–5.
77. Andersson HC, Frentz J, Martinez JE, Tuck-Muller CM, Bellizaire J. Adrenal insufficiency in Smith-Lemli-Opitz syndrome. *Am J Hum Genet* 1999;82:382–384.
78. Bianconi SE, Conley SK, Keil MF, Sinaii N, Rother KI, Porter FD, Stratakis CA. Adrenal function in Smith-Lemli-Opitz syndrome. *Am J Med Genet Part A* 2011;155:2732–2738.
79. Marcos J, Guo LW, Wilson WK, Porter FD, Shackleton C. The implications of 7-dehydrosterol-7-reductase deficiency (Smith-Lemli-Opitz syndrome) to neurosteroid production. *Steroids* 2004;69:51–60.
80. Griffiths WJ, Abdel-Khalik J, Crick PJ, Ogundare M, Shackleton CH, Tuschl K, Kwok MK, Bigger BW, Morris AA, Honda A, Xu L, Porter NA, Björkhem I,

- Clayton PT, Wang Y. Sterols and oxysterols in plasma from Smith-Lemli-Opitz syndrome patients. *J Steroid Biochem Mol Biol* 2016;
81. Honda A, Salen G, Shefer S, Batta AK, Honda M, Xu G, Stephen Tint G, Matsuzaki Y, Shoda J, Tanaka N. Bile acid synthesis in the Smith-Lemli-Opitz syndrome: effects of dehydrocholesterols on cholesterol 7 α -hydroxylase and 27-hydroxylase activities in rat liver. *J Lipid Res* 1999;40:1520–1528.
82. Wassif CA, Vied D, Tsokos M, Connor WE, Steiner RD, Porter FD. Cholesterol storage defect in RSH/Smith–Lemli–Opitz syndrome fibroblasts. *Mol Genet Metab* 2002;75:325–334.
83. Niemann A. Ein unbekanntes Krankheitsbild. *Jahrb für Kinderheilkd* 1914;79:1–10.
84. Vanier MT. Niemann-Pick disease type C. *Orphanet J Rare Dis* 2010;5:16.
85. Crocker AC. The Cerebral Defect in Tay-Sachs Disease and Niemann-Pick Disease. *J Neurochem* 1961;7:69–80.
86. Brady R O, Kanfer JN, Mock MB, Fredrickson DS. The metabolism of sphingomyelin. II. Evidence of an enzymatic deficiency in Niemann-Pick disease. *Proc Natl Acad Sci USA* 1966;55:366–369.
87. Winsor EJ, Welch JP. Genetic and demographic aspects of Nova Scotia Niemann-Pick disease (type D). *Am J Hum Genet* 1978;30:530–8.
88. Zhang J, Liu Q. Cholesterol metabolism and homeostasis in the brain. *Protein Cell* 2015;6:254–264.
89. Naureckiene S, Sleat DE, Lackland H, Fensom A, Vanier MT, Wattiaux R, Jadot M, Lobel P. Identification of HE1 as the Second Gene of Niemann-Pick C Disease. *Science* 2000;290:2298–301.
90. Goldman SDB, Krise JP. Niemann-Pick C1 functions independently of Niemann-

- Pick C2 in the initial stage of retrograde transport of membrane-impermeable lysosomal cargo. *J Biol Chem* 2010;285:4983–4994.
91. Scott C, Ioannou YA. The NPC1 protein: Structure implies function. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids*. 2004;1685:8–13.
92. Infante RE, Radhakrishnan A, Abi-Mosleh L, Kinch LN, Wang ML, Grishin N V., Goldstein JL, Brown MS. Purified NPC1 protein II. Localization of sterol binding to a 240-amino acid soluble luminal loop. *J Biol Chem* 2008;283:1064–1075.
93. Li X, Wang J, Coutavas E, Shi H, Hao Q, Blobel G, Maxfield FR, Pfeiffer SR, Rothman JE. Structure of human Niemann–Pick C1 protein. *Proc Natl Acad Sci* 2016;113:8212–17.
94. Kwon HJ, Abi-Mosleh L, Wang ML, Deisenhofer J, Goldstein JL, Brown MS, Infante RE. Structure of N-Terminal Domain of NPC1 Reveals Distinct Subdomains for Binding and Transfer of Cholesterol. *Cell* 2009;137:1213–1224.
95. Abdul-Hammed M, Breiden B, Adebayo M a, Babalola JO, Schwarzmunn G, Sandhoff K. Role of endosomal membrane lipids and NPC2 in cholesterol transfer and membrane fusion. *J Lipid Res* 2010;51:1747–1760.
96. Xu S, Benoff B, Liou HL, Lobel P, Stock AM. Structural basis of sterol binding by NPC2, a lysosomal protein deficient in Niemann-Pick type C2 disease. *J Biol Chem* 2007;282:23525–23531.
97. Friedland N, Liou H-L, Lobel P, Stock AM. Structure of a cholesterol-binding protein deficient in Niemann-Pick type C2 disease. *Proc Natl Acad Sci U S A* 2003;100:2512–7.
98. Wraith JE, Baumgartner MR, Bembi B, Covanis A, Levade T, Mengel E, Pineda M, Sedel F, Top??u M, Vanier MT, Widner H, Wijburg FA, Patterson MC.

- Recommendations on the diagnosis and management of Niemann-Pick disease type C. *Mol Genet Metab* 2009;98:152–165.
99. Tängemo C, Weber D, Theiss S, Mengel E, Runz H. Niemann-Pick Type C disease: characterizing lipid levels in patients with variant lysosomal cholesterol storage. *J Lipid Res* 2011;52:813–825.
100. Vanier MT, Latour P. Laboratory diagnosis of Niemann-Pick disease type C: The filipin staining test. *Methods Cell Biol* 2015;126:357–375.
101. Vanier MT, Rodriguez-Lafrasse C, Rousson R, Gazzah N, Juge MC, Pentchev PG, Revol A, Louisot P. Type C Niemann-Pick disease: spectrum of phenotypic variation in disruption of intracellular LDL-derived cholesterol processing. *BBA - Mol Basis Dis* 1991;1096:328–337.
102. Yanjanin NM, Vélez JI, Gropman A, King K, Bianconi SE, Conley SK, Brewer CC, Solomon B, Pavan WJ, Arcos-Burgos M, Patterson MC, Porter FD. Linear clinical progression, independent of age of onset, in Niemann-Pick disease, type C. *Am J Med Genet B Neuropsychiatr Genet* 2010;153B:132–40.
103. Mazzacuva F, Mills P, Mills K, Camuzeaux S, Gissen P, Nicoli E-R, Wassif C, te Vruchte D, Porter FD, Maekawa M, Mano N, Iida T, Platt F, Clayton PT. Identification of novel bile acids as biomarkers for the early diagnosis of Niemann-Pick C disease. *FEBS Lett* 2016;590:1651–1662.
104. Lloyd-Evans E, Morgan AJ, He X, Smith DA, Elliot-Smith E, Sillence DJ, Churchill GC, Schuchman EH, Galione A, Platt FM. Niemann-Pick disease type C1 is a sphingosine storage disease that causes deregulation of lysosomal calcium. *Nat Med* 2008;14:1247–55.
105. Ong WY, Kumar U, Switzer RC, Sidhu A, Suresh G, Hu CY, Patel SC. Neurodegeneration in Niemann-Pick type C disease mice. *Exp Brain Res*

- 2001;141:218–31.
106. Sarna JR, Larouche M, Marzban H, Sillitoe R V, Rancourt DE, Hawkes R. Patterned Purkinje Cell Degeneration in Mouse Models of Niemann-Pick Type C Disease. *J Comp Neurol* 2003;456:279–291.
107. Higashi Y, Murayama S, Pentchev PG, Suzuki K. Cerebellar degeneration in the Niemann-Pick type C mouse. *Acta Neuropathol* 1993;85:175–84.
108. Wassif CA, Zhu P, Kratz L, Krakowiak PA, Battaile KP, Weight FF, Grinberg A, Steiner RD, Nwokoro NA, Kelley RI, Stewart RR, Porter FD. Biochemical, phenotypic and neurophysiological characterization of a genetic mouse model of RSH/Smith–Lemli–Opitz syndrome. *Hum Mol Genet* 2001;10:555–564.
109. Correa-Cerro LS, Wassif CA, Waye JS, Krakowiak PA, Cozma D, Dobson NR, Levin SW, Anadiotis G, Steiner RD, Krajewska-Walasek M, Nowaczyk MJM, Porter FD. DHCR7 nonsense mutations and characterisation of mRNA nonsense mediated decay in Smith-Lemli-Opitz syndrome. *J Med Genet* 2005;42:350–7.
110. Correa-Cerro LS, Wassif CA, Kratz L, Miller GF, Munasinghe JP, Grinberg A, Fliesler SJ, Porter FD. Development and characterization of a hypomorphic Smith-Lemli-Opitz syndrome mouse model and efficacy of simvastatin therapy. *Hum Mol Genet* 2006;15:839–851.
111. Lowenthal AC, Cummings JF, Wenger DA, Thrall MA, Wood PA, De Lahunta A. Feline sphingolipidosis resembling Niemann-Pick disease type C. *Acta Neuropathol* 1990;81:189–197.
112. Brown DE, Thrall MA, Walkley SU, Wenger DA, Mitchell TW, Smith MO, Royals KL, March PA, Allison RW. Feline Niemann-Pick disease type C. *Am J Pathol* 1994;144:1412–1415.
113. Kuwamura M, Awakura T, Shimada A, Umemura I, Kagota K, Kawamura N,

- Naikp M. Type C Niemann-Pick disease in a boxer dog. *Acta Neuropathol* 1993;85:345–348.
114. Pentchev PG, Gal AE, Booth AD, Omodeo-Sale F, Fouks J, Neumeyer BA, Quirk JM, Dawson G, Brady RO. A lysosomal storage disorder in mice characterized by a dual deficiency of sphingomyelinase and glucocerebrosidase. 1980;619:669–79.
115. Maue RA, Burgess RW, Wang B, Wooley CM, Seburn KL, Vanier MT, Rogers MA, Chang CC, Chang T-Y, Harris BT, Graber DJ, Penatti CAA, Porter DM, Szwergold BS, Henderson LP, Totenhagen JW, Trouard TP, Borbon IA, Erickson RP. A novel mouse model of Niemann–Pick type C disease carrying a D1005G-Npc1 mutation comparable to commonly observed human mutations. *Hum Mol Genet* 2012;21:730–750.
116. Ohno K, Nanba E, Miyawaki S, Sakiyama T, Kitagawa T, Takeshita K. A cell line derived from sphingomyelinosis mouse shows alterations in intracellular cholesterol metabolism similar to those in type C Niemann-Pick disease. *Cell Struct Funct* 1992;17:229–235.
117. Loftus SK, Morris JA, Carstea ED, Gu JZ, Cummings C, Brown A, Ellison J, Ohno K, Rosenfeld MA, Tagle DA, Pentchev PG, Pavan WJ. Murine Model of Niemann-Pick C Disease : Mutation in a Cholesterol Homeostasis Gene. *Science* (80-) 1997;277:232–235.
118. Voikar V, Rauvala H, Ikonen E. Cognitive deficit and development of motor impairment in a mouse model of Niemann-Pick type C disease. *Behav Brain Res* 2002;132:1–10.
119. Miyawaki S, Mitsuoka S, Sakiyama T, Kitagawa T. Sphingomyelinosis, a new mutation in the mouse: a model of Niemann-Pick disease in humans. *J Hered*

- 1982;73:257–263.
120. Sleat DE, Wiseman JA, El-Banna M, Price SM, Verot L, Shen MM, Stephen Tint G, Vanier MT, Walkley SU, Lobel P, Scott MP. Genetic evidence for nonredundant functional cooperativity between NPC1 and NPC2 in lipid transport. *Proc Natl Acad Sci* 2004;101:5886–91.
121. Vite CH, Bagel JH, Swain GP, Prociuk M, Sikora TU, Stein VM, O'Donnell P, Ruane T, Ward S, Crooks A, Li S, Mauldin E, Stellar S, De Meulder M, Kao ML, Ory DS, Davidson C, Vanier MT, Walkley SU. Intracisternal cyclodextrin prevents cerebellar dysfunction and Purkinje cell death in feline Niemann-Pick type C1 disease. *Sci Transl Med* 2015;7:276ra26.
122. Koo G, Conley SK, Wassif CA, Porter FD. Discordant phenotype and sterol biochemistry in Smith-Lemli-Opitz syndrome. *Am J Med Genet Part A* 2010;152:2094–2098.
123. Nwokoro N, Mulvihill J. Cholesterol and bile acid replacement therapy in children and adults with Smith-Lemli-Opitz (SLO/RSH) syndrome. *Am J Med Genet* 1997;68:315–21.
124. Tierney E, Conley SK, Goodwin H, Porter FD. Analysis of short-term behavioral effects of dietary cholesterol supplementation in Smith-Lemli-Opitz syndrome. *Am J Med Genet A* 2010;152A:91–95.
125. Jira PE, Wevers RA, de Jong J, Rubio-Gozalbo E, Janssen-Zijlstra FS, van Heyst AF, Sengers RC, Smeitink JA. Simvastatin. A new therapeutic approach for Smith-Lemli-Opitz syndrome. *J Lipid Res* 2000;41:1339–46.
126. Saheki A, Terasaki T, Tamai I, Tsuji A. In vivo and in vitro blood-brain barrier transport of 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) reductase inhibitors. *Pharm Res* 1994;11:305–311.

127. Haas D, Garbade SF, Vohwinkel C, Muschol N, Trefz FK, Penzien JM, Zschocke J, Hoffmann GF, Burgard P. Effects of cholesterol and simvastatin treatment in patients with Smith-Lemli-Opitz syndrome (SLOS). *J Inher Metab Dis* 2007;30:375–387.
128. Wassif CA, Krakowiak PA, Wright BS, Gewandter JS, Sterner AL, Javitt N, Yergey AL, Porter FD. Residual cholesterol synthesis and simvastatin induction of cholesterol synthesis in Smith-Lemli-Opitz syndrome fibroblasts. *Mol Genet Metab* 2005;85:96–107.
129. Lindegaard ML, Wassif CA, Vaisman B, Amar M, Wasmuth E V., Shamburek R, Nielsen LB, Remaley AT, Porter FD. Characterization of placental cholesterol transport: ABCA1 is a potential target for in utero therapy of Smith-Lemli-Opitz syndrome. *Hum Mol Genet* 2008;17:3806–3813.
130. Platt FM. Sphingolipid lysosomal storage disorders. *Nature* 2014;510:68–75.
131. Platt FM, Neises GR, Dwek RA, Butters TD. N-butyldeoxynojirimycin is a novel inhibitor of glycolipid biosynthesis. *J Biol Chem* 1994;269:8362–8365.
132. Te Vruchte D, Lloyd-Evans E, Veldman RJ, Neville DCA, Dwek RA, Platt FM, Van Blitterswijk WJ, Sillences DJ. Accumulation of glycosphingolipids in Niemann-Pick C disease disrupts endosomal transport. *J Biol Chem* 2004;279:26167–26175.
133. Patterson MC, Vecchio D, Prady H, Abel L, Wraith JE. Miglustat for treatment of Niemann-Pick C disease: a randomised controlled study. *Lancet Neurol* 2007;6:765–772.
134. Zervas M, Somers KL, Thrall MA, Walkley SU. Critical role for glycosphingolipids in Niemann-Pick disease type C. *Curr Biol* 2001;11:1283–1287.

135. Williams IM, Wallom KL, Smith DA, Al Eisa N, Smith C, Platt FM. Improved neuroprotection using miglustat, curcumin and ibuprofen as a triple combination therapy in Niemann-Pick disease type C1 mice. *Neurobiol Dis* 2014;67:9–17.
136. Griffin LD, Gong W, Verot L, Mellon SH. Niemann-Pick type C disease involves disrupted neurosteroidogenesis and responds to allopregnanolone. *Nat Med* 2004;10:704–11.
137. Davidson CD, Ali NF, Micsenyi MC, Stephney G, Renault S, Dobrenis K, Ory DS, Vanier MT, Walkley SU. Chronic cyclodextrin treatment of murine Niemann-Pick C disease ameliorates neuronal cholesterol and glycosphingolipid storage and disease progression. *PLoS One* 2009;4:e6951.
138. Pontikis CC, Davidson CD, Walkley SU, Platt FM, Begley DJ. Cyclodextrin alleviates neuronal storage of cholesterol in Niemann-Pick C disease without evidence of detectable blood-brain barrier permeability. *J Inherit Metab Dis* 2013;36:491–498.
139. Rosenbaum AI, Zhang G, Warren JD, Maxfield FR. Endocytosis of beta-cyclodextrins is responsible for cholesterol reduction in Niemann-Pick type C mutant cells. *Proc Natl Acad Sci* 2010;107:5477–5482.
140. Hsu YS, Hwu WL, Huang SF, Lu MY, Chen RL, Lin DT, Peng SS, Lin KH. Niemann-Pick disease type C (a cellular cholesterol lipidosis) treated by bone marrow transplantation. *Bone Marrow Transplant* 1999;24:103–107.
141. Gartner JC, Bergman I, Malatack JJ, Zitelli BJ, Jaffe R, Watkins JB, Shaw BW, Iwatsuki S, Starzl TE. Progression of neurovisceral storage disease with supranuclear ophthalmoplegia following orthotopic liver transplantation. *Pediatrics* 1986;77:104–106.
142. Last JM. A dictionary of epidemiology. *Int. J. Epidemiol.* 1986;15:277.

143. Waterham HR, Hennekam RCM. Mutational spectrum of Smith-Lemli-Opitz syndrome. *Am J Med Genet Part C Semin Med Genet* 2012;160 C:263–284.
144. Wassif CA, Maslen C, Kachilele-Linjewile S, Lin D, Linck LM, Connor WE, Steiner RD, Porter FD. Mutations in the Human Sterol Δ 7-Reductase Gene at 11q12-13 Cause Smith-Lemli-Opitz Syndrome. *Am J Hum Genet* 1998;63:55–62.
145. Fitzky BU, Witsch-Baumgartner M, Erdel M, Lee JN, Paik YK, Glossmann H, Utermann G, Moebius FF. Mutations in the Delta7-sterol reductase gene in patients with the Smith-Lemli-Opitz syndrome. *Proc Natl Acad Sci U S A* 1998;95:8181–6.
146. Waterham HR, Wijburg FA, Hennekam RC, Vreken P, Poll-The BT, Dorland L, Duran M, Jira PE, Smeitink JA, Wevers RA, Wanders RJ. Smith-Lemli-Opitz syndrome is caused by mutations in the 7-dehydrocholesterol reductase gene. *Am J Hum Genet* 1998;63:329–38.
147. Balogh I, Koczok K, Szabo GP, Torok O, Hadzsiev K, Csabi G, Balogh L, Dzsudzsak E, Ajzner E, Szabo L, Csakvary V, Olah A V. Mutational spectrum of Smith-Lemli-Opitz syndrome patients in hungary. In: *Molecular Syndromology*. 2012. p. 215–222.
148. Ellingson M., Wick M., White W., Raymond K., Saenger A., Pichurin P., Wassif C., Porter F., Babovic-Vuksanovic D. Pregnancy in an individual with mild Smith-Lemli-Opitz syndrome. 2014;85:495–497.
149. Sparks SE, Wassif CA, Goodwin H, Conley SK, Lanham DC, Kratz LE, Hyland K, Gropman A, Tierney E, Porter FD. Decreased cerebral spinal fluid neurotransmitter levels in Smith-Lemli-Opitz syndrome. *J Inherit Metab Dis* 2014;37:415–420.

150. Olah A V., Szabo GP, Varga J, Balogh L, Csabi G, Csakvary V, Erwa W, Balogh I. Relation between biomarkers and clinical severity in patients with Smith-Lemli-Opitz syndrome. *Eur J Pediatr* 2013;172:623–630.
151. Stenson PD, Mort M, Ball E V, Howells K, Phillips AD, Thomas NS, Cooper DN. The Human Gene Mutation Database: 2008 update. *Genome Med* 2009;1:13.
152. Witsch-Baumgartner M, Ciara E, Löffler J, Menzel HJ, Seedorf U, Burn J, Gillessen-Kaesbach G, Hoffmann GF, Fitzky BU, Mundy H, Clayton P, Kelley RI, Krajewska-Walasek M, Utermann G. Frequency gradients of DHCR7 mutations in patients with Smith-Lemli-Opitz syndrome in Europe: evidence for different origins of common mutations. *Eur J Hum Genet EJHG* [Internet] 2001;9:45–50. Available from:
https://drive.google.com/file/d/0Byg29d4uCW_zUjBrQnl2TU1uUUE/view%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/11175299
153. Lowry RB, Yong SL. Borderline normal intelligence in the Smith-Lemli-Opitz (RSH) syndrome. *Am J Med Genet* 1980;5:137–143.
154. Witsch-Baumgartner M, Schwentner I, Gruber M, Benlian P, Bertranpetit J, Bieth E, Chevy F, Clusellas N, Estivill X, Gasparini G, Giros M, Kelley RI, Krajewska-Walasek M, Menzel J, Miettinen T, Ogorelkova M, Rossi M, Scala I, Schinzel A, Schmidt K, Schöntitzer D, Seemanova E, Sperling K, Syrrou M, Talmud PJ, Wollnik B, Krawczak M, Labuda D, Utermann G. Age and origin of major Smith-Lemli-Opitz syndrome (SLOS) mutations in European populations. *J Med Genet* 2008;45:200–9.
155. Irons, M, Elias, ER, Salen, G, Tint, GS, Batta A. Defective cholesterol biosynthesis in Smith-Lemli-Opitz syndrome. *Lancet* 1993;3857:1414.

156. Kelley RI. Editorial : A New Face for an Old Syndrome. *Am. J. Med. Genet.* 1997;256:251–256.
157. Ryan A, Bartlett K, Clayton P, Eaton S, Mills L, Donnai D, Winter RM, Burn J. Smith-Lemli-Opitz syndrome: a variable clinical and biochemical phenotype. *J Med Genet* 1998;35:558–565.
158. Nowaczyk MJM, Zeesman S, Wayne JS, Douketis JD. Incidence of Smith-Lemli-Opitz syndrome in Canada: Results of three-year population surveillance. *J Pediatr* 2004;145:530–535.
159. Jahnova H, Dvorakova L, Vlaskova H, Hulkova H, Poupetova H, Hrebicek M, Jesina P. Observational, retrospective study of a large cohort of patients with Niemann-Pick disease type C in the Czech Republic: a surprisingly stable diagnostic rate spanning almost 40 years. *Orphanet J Rare Dis* 2014;9:140.
160. Pentchev PG, Comly ME, Kruth HS, Vanier MT, Wenger DA, Patel S, Brady RO. A defect in cholesterol esterification in Niemann-Pick disease (type C) patients. *Proc. Natl. Acad. Sci.* 1985;82:8247–8251.
161. Kruth HS, Comly ME, Butler JD, Vanier MT, Fink JK, Wenger DA, Patel S, Pentchev PG. Type C Niemann-Pick disease. Abnormal metabolism of low density lipoprotein in homozygous and heterozygous fibroblasts. *J Biol Chem* 1986;261:16769–16774.
162. Blanchette-Mackie EJ, Dwyer NK, Amende LM, Kruth HS, Butler JD, Sokol J, Comly ME, Vanier MT, August JT, Brady RO. Type-C Niemann-Pick disease: low density lipoprotein uptake is associated with premature cholesterol accumulation in the Golgi complex and excessive cholesterol storage in lysosomes. *Proc Natl Acad Sci U S A* 1988;85:8022–6.
163. Vanier MT, Duthel S, Rodriguez-Lafrasse C, Pentchev P, Carstea ED. Genetic

- heterogeneity in Niemann-Pick C disease: a study using somatic cell hybridization and linkage analysis. *Am J Hum Genet* 1996;58:118–25.
164. Carstea ED, Morris JA, Coleman KG, Loftus SK, Zhang D, Cummings C, Gu J, Rosenfeld MA, Pavan WJ, Krizman DB, Nagle J, Polymeropoulos MH, Sturley SL, Ioannou YA, Higgins ME, Comly M, Cooney A, Brown A, Kaneski CR, Blanchette-Mackie EJ, Dwyer NK, Neufeld EB, Chang T-Y, Liscum L, Strauss JF, Ohno K, Zeigler M, Carmi R, Sokol J, Markie D, O'Neill RR, van Diggelen OP, Elleder M, Patterson MC, Brady RO, Vanier MT, Pentchev PG, Tagle DA. Niemann-Pick C1 disease gene: Homology to mediators of cholesterol homeostasis. *Science* 1997;277:228–231.
165. Imrie J, Dasgupta S, Besley GTN, Harris C, Heptinstall L, Knight S, Vanier MT, Fensom a H, Ward C, Jacklin E, Whitehouse C, Wraith JE. The natural history of Niemann-Pick disease type C in the UK. *J Inherit Metab Dis* 2007;30:51–9.
166. Millat G, Marçais C, Rafi MA, Yamamoto T, Morris JA, Pentchev PG, Ohno K, Wenger DA, Vanier MT. Niemann-Pick C1 disease: the I1061T substitution is a frequent mutant allele in patients of Western European descent and correlates with a classic juvenile phenotype. *Am J Hum Genet* 1999;65:1321–9.
167. Vanier MT, Millat G. Niemann-Pick disease type C. *Clin Genet* 2003;64:269–281.
168. Millat G, Chikh K, Naureckiene S, Sleat DE, Fensom a H, Higaki K, Elleder M, Lobel P, Vanier MT. Niemann-Pick disease type C: spectrum of HE1 mutations and genotype/phenotype correlations in the NPC2 group. *Am J Hum Genet* 2001;69:1013–1021.
169. Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, Kang HM, Marth GT, McVean GA. An integrated map of genetic variation from

- 1,092 human genomes. *Nature* 2012;491:56–65.
170. EVS. Exome Variant Server. NHLBI GO Exome Seq Proj [Internet] 2014; Available from: <http://evs.gs.washington.edu/EVS/>
171. Biesecker LG, Mullikin JC, Facio FM, Turner C, Cherukuri PF, Blakesley RW, Bouffard GG, Chines PS, Cruz P, Hansen NF, Teer JK, Maskeri B, Young AC, Manolio TA, Wilson AF, Finkel T, Hwang P, Arai A, Remaley AT, Sachdev V, Shamburek R, Cannon RO, Green ED. The ClinSeq Project: Piloting large-scale genome sequencing for research in genomic medicine. *Genome Res* 2009;19:1665–1674.
172. Chelala C, Khan A, Lemoine NR. SNPnexus: A web database for functional annotation of newly discovered and public domain single nucleotide polymorphisms. *Bioinformatics* 2009;25:655–661.
173. Roy A, Kucukural A, Zhang Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc* 2010;5:725–738.
174. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. A method and server for predicting damaging missense mutations. *Nat Methods* 2010;7:248–249.
175. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* 2009;4:1073–1081.
176. Reva B, Antipin Y, Sander C. Predicting the functional impact of protein mutations: Application to cancer genomics. *Nucleic Acids Res* 2011;39:e118.
177. Yeo G, Burge CB. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J Comput Biol* 2004;11:377–394.
178. Ryckman K, Williams SM, Ryckman K, Williams SM. Calculation and Use of

- the Hardy-Weinberg Model in Association Studies. Hoboken, NJ, USA: John Wiley & Sons, Inc.; 2008.
179. Carette JE, Raaben M, Wong AC, Herbert AS, Obernosterer G, Mulherkar N, Kuehne AI, Kranzusch PJ, Griffin AM, Ruthel G, Dal Cin P, Dye JM, Whelan SP, Chandran K, Brummelkamp TR. Ebola virus entry requires the cholesterol transporter Niemann-Pick C1. *Nature* 2011;477:340–3.
180. Al-Daghri NM, Cagliani R, Forni D, Alokail MS, Pozzoli U, Alkharfy KM, Sabico S, Clerici M, Sironi M. Mammalian NPC1 genes may undergo positive selection and human polymorphisms associate with type 2 diabetes. *BMC Med* 2012;10:140.
181. Dvorakova L, Sikora J, Hrebicek M, Hulkova H, Bouckova M, Stolnaja L, Elleder M. Subclinical course of adult visceral Niemann-Pick type C1 disease. A rare or underdiagnosed disorder? *J Inherit Metab Dis* 2006;29:591.
182. Bauer P, Balding DJ, Klünemann HH, Linden DEJ, Ory DS, Pineda M, Priller J, Sedel F, Muller A, Chadha-Boreham H, Welford RWD, Strasser DS, Patterson MC. Genetic screening for Niemann-Pick disease type C in adults with neurological and psychiatric symptoms: Findings from the ZOOM study. *Hum Mol Genet* 2013;22:4349–4356.
183. Boenzi S, Deodato F, Taurisano R, Goffredo BM, Rizzo C, Dionisi-Vici C. Evaluation of plasma cholestane-3 β ,5 α ,6 β -triol and 7-ketocholesterol in inherited disorders related to cholesterol metabolism. *J Lipid Res* 2016;57:361–367.
184. Huang L, Pike D, Sleat DE, Nanda V, Lobel P. Potential pitfalls and solutions for use of fluorescent fusion proteins to study the lysosome. *PLoS One* 2014;9.
185. Park WD, O'Brien JF, Lundquist PA, Kraft DL, Vockley CW, Karnes PS,

- Patterson MC, Snow K. Identification of 58 novel mutations in Niemann-Pick disease type C: Correlation with biochemical phenotype and importance of PTC1-like domains in NPC1. *Hum Mutat* 2003;22:313–325.
186. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 2010;20:1297–303.
 187. Dykens EM. Measuring behavioral phenotypes: Provocations from the “new genetics.” *Am J Ment Retard* 1995;99:522–532.
 188. Korade Z, Folkes OM, Harrison FE. Behavioral and serotonergic response changes in the Dhcr7-HET mouse model of Smith-Lemli-Opitz syndrome. *Pharmacol Biochem Behav* 2013;106:101–108.
 189. Fitzky BU, Moebius FF, Asaoka H, Waage-Baudet H, Xu L, Xu G, Maeda N, Kluckman K, Hiller S, Yu H, Batta AK, Shefer S, Chen T, Salen G, Sulik K, Simoni RD, Ness GC, Glossmann H, Patel SB, Tint GS, Hoffmann G, Clayton P, Mills K, Keeling J, FitzPatrick D, Kelley R, Braverman N, König A, Happel R, Bornholdt D, Engel H, Grzeschik K-H, Grange D, Kratz L, Braverman N, Kelley R, Lowry R, Yong S-L, Cuniff C, Kratz L, Moser A, Natowicz M, Kelley R, Witsch-Raumgartner M, Smith D, Lemli L, Opitz J, Curry C, Tint G, Shefer S, Moebius F, Fitzky B, Lee J, Paik Y-K, Glossmann H, Fitzky B, Tint G, Kelley R, Witsch-Baumgartner M, Porter J, Young K, Beachy P, Tint G, Incardona J, Gaffield W, Kapur R, Roelink H, Honda M, Honda M, Steiner R, Linck L, Flavell D, Lin D, Connor W, Shehee W, Oliver P, Smithies O, Batta A, Tint G, Abuelo D, Shefer S, Salen G, Honda M, Nguyen L, Wu J, Zhu Y, Patel S, Chambers C, Ness G, Ness G, Sample C, Smith M, Pendleton L, Eichler D, Ness

- G, Zhao Z, Skalnik D, Narita H, et al. 7-Dehydrocholesterol-dependent proteolysis of HMG-CoA reductase suppresses sterol biosynthesis in a mouse model of Smith-Lemli-Opitz/RSH syndrome. *J Clin Invest* 2001;108:905–15.
190. Kelley RI, Hennekam RC. The Smith-Lemli-Opitz syndrome. *J Med Genet* [Internet] 2000;37:321–35. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1734573&tool=pmcentrez&rendertype=abstract>
191. Montgomery KC. The relation between fear induced by novel stimulation and exploratory drive. *J Comp Physiol Psychol* 1955;48:254–260.
192. Handley SL, Mithani S. Effects of alpha-adrenoceptor agonists and antagonists in a maze-exploration model of “fear”-motivated behaviour. *Naunyn Schmiedebergs Arch Pharmacol* 1984;327:1–5.
193. Lister RG. The use of a plus-maze to measure anxiety in the mouse. *Psychopharmacology (Berl)* 1987;92:180–185.
194. Deacon R. Assessing burrowing, nest construction, and hoarding in mice. *J Vis Exp* 2012;e2607.
195. Deacon RMJ. Burrowing in rodents: a sensitive method for detecting behavioral dysfunction. *Nat Protoc* 2006;1:118–121.
196. Deacon RM, Raley JM, Perry VH, Rawlins JN. Burrowing into prion disease. *Neuroreport* 2001;12:2053–2057.
197. Deacon RMJ. Assessing nest building in mice. *Nat Protoc* 2006;1:1117–9.
198. Lisk RD, Pretlow R a, Friedman SM. Hormonal stimulation necessary for elicitation of maternal nest-building in the mouse (*Mus musculus*). *Anim Behav* 1969;17:730–737.
199. Lisk RD. Oestrogen and progesterone synergism and elicitation of maternal nest-

- building in the mouse (*Mus musculus*). *Anim Behav* 1971;19:606–610.
200. Numan M, Rosenblatt JS, Komisaruk BR. Medial preoptic area and onset of maternal behavior in the rat. *J Comp Physiol Psychol* 1977;91:146–64.
201. Slotnick BM, Nigrosh BJ. Maternal behavior of mice with cingulate cortical, amygdala, or septal lesions. *J Comp Physiol Psychol* 1975;88:118–127.
202. Kim JJ, Shih JC, Chen K, Chen L, Bao S, Maren S, Anagnostaras SG, Fanselow MS, De Maeyer E, Seif I, Thompson RF. Selective enhancement of emotional, but not motor, learning in monoamine oxidase A-deficient mice. *Proc Natl Acad Sci* 1997;94:5929–5933.
203. Deacon RMJ, Croucher A, Rawlins JN. Hippocampal cytotoxic lesion effects on species-typical behaviours in mice. *Behav Brain Res* 2002;132:203–213.
204. Gould TD, Dao DT, Kovacsics CE. The Open Field Test. In: *Mood and Anxiety Related Phenotypes in Mice*. 2009. p. 1–20.
205. Prut L, Belzung C. The open field as a paradigm to measure the effects of drugs on anxiety-like behaviors: A review. *Eur. J. Pharmacol.* 2003;463:3–33.
206. Deacon RMJ. Burrowing: A sensitive behavioural assay, tested in five species of laboratory rodents. *Behav Brain Res* 2009;200:128–133.
207. Braun AA, Skelton MR, Vorhees C V, Williams MT. Comparison of the elevated plus and elevated zero mazes in treated and untreated male Sprague-Dawley rats: effects of anxiolytic and anxiogenic agents. *Pharmacol Biochem Behav* 2011;97:406–15.
208. Rahnev D, Nee DE, Riddle J, Larson AS, D'Esposito M. Causal evidence for frontal cortex organization for perceptual decision making. *Proc Natl Acad Sci* [Internet] 2016;201522551. Available from: <http://www.pnas.org/lookup/doi/10.1073/pnas.1522551113>

209. Volkar, V, Kiks, S, Vasar E, Rauvala, H. Strain and gender differences in the behavior of mouse lines commonly used in transgenic studies. *Physiol Behav* 2001;72:271–281.
210. Li X, Morrow D, Witkin JM. Decreases in nestlet shredding of mice by serotonin uptake inhibitors: Comparison with marble burying. *Life Sci* 2006;78:1933–1939.
211. Waage-Baudet H, Lauder JM, Dehart DB, Kluckman K, Hiller S, Tint GS, Sulik KK. Abnormal serotonergic development in a mouse model for the Smith-Lemli-Opitz syndrome: Implications for autism. *Int J Dev Neurosci* 2003;21:451–459.
212. Himmelstein DS, Bi C, Clark BS, Bai B, Kohtz JD. Balanced Shh signaling is required for proper formation and maintenance of dorsal telencephalic midline structures. *BMC Dev Biol* 2010;10:118.
213. Grove E a, Tole S, Limon J, Yip L, Ragsdale CW. The hem of the embryonic cerebral cortex is defined by the expression of multiple Wnt genes and is compromised in Gli3-deficient mice. *Development* 1998;125:2315–2325.
214. Mangale VS, Hirokawa KE, Satyaki PR V., Gokulchandran N, Chikbire S, Subramanian L, Shetty AS, Martynoga B, Paul J, Mai M V., Li Y, Flanagan LA, Tole S, Monuki ES. Lhx2 Selector Activity Specifies Cortical Identity and Suppresses Hippocampal Organizer Fate. *Science* 2008;319:304–309.
215. Francis KR, Ton AN, Xin Y, O'Halloran PE, Wassif CA, Malik N, Williams IM, Cluzeau C V, Trivedi NS, Pavan WJ, Cho W, Westphal H, Porter FD. Modeling Smith-Lemli-Opitz syndrome with induced pluripotent stem cells reveals a causal role for Wnt/ β -catenin defects in neuronal cholesterol synthesis phenotypes. *Nat Med* 2016;22:388–96.
216. Nunez J. Morris Water Maze Experiment. *J Vis Exp* 2008;12–13.

217. Leger M, Quiedeville A, Bouet V, Haelewyn B, Boulouard M, Schumann-Bard P, Freret T. Object recognition test in mice. *Nat Protoc* 2013;8:2531–2537.
218. Sikora DM, Pettit-Kekel K, Penfield J, Merkens LS, Steiner RD. The near universal presence of autism spectrum disorders in children with Smith-Lemli-Opitz syndrome. *Am J Med Genet Part A* 2006;140:1511–1518.
219. Meziane H, Ouagazzal AM, Aubert L, Wietrzyk M, Krezel W. Estrous cycle effects on behavior of C57BL/6J and BALB/cByJ female mice: Implications for phenotyping strategies. *Genes, Brain Behav* 2007;6:192–200.
220. Nadeau JH. Modifier genes in mice and humans. *Nat Rev Genet* 2001;2:165–174.
221. Schroepfer GJ. Oxysterols: modulators of cholesterol metabolism and other processes. *Physiol Rev* 2000;80:361–554.
222. Gerschman R, Gilbert DL, Nye SW, Dwyer P, Fenn WO. Oxygen poisoning and x-irradiation: a mechanism in common. *Science* 1954;119:623–626.
223. McCord JM, Keele BB, Fridovich I. An enzyme-based theory of obligate anaerobiosis: the physiological function of superoxide dismutase. *Proc Natl Acad Sci U S A* 1971;68:1024–7.
224. Hybertson BM, Gao B, Bose SK, McCord JM. Oxidative stress in health and disease: The therapeutic potential of Nrf2 activation. *Mol Aspects Med* 2011;32:234–246.
225. Xu L, Davis TA, Porter NA. Rate constants for peroxidation of polyunsaturated fatty acids and sterols in solution and in liposomes. *J Am Chem Soc* 2009;131:13037–13044.
226. Vázquez MC, Balboa E, Alvarez AR, Zanolungo S. Oxidative stress: A pathogenic mechanism for Niemann-Pick type C disease. *Oxid Med Cell Longev* 2012;2012:205713.

227. Adibhatla RM, Hatcher JF. Altered lipid metabolism in brain injury and disorders. *Subcell Biochem* 2008;49:241–68.
228. Korade Z, Xu L, Mirnics K, Porter N a. Lipid biomarkers of oxidative stress in a genetic mouse model of Smith-Lemli-Opitz syndrome. *J Inherit Metab Dis* 2012;36:113–22.
229. Siddiqui, A.M.;Wassif, C.A.;Porter, F.D;Richards, M.J.;Fliesler SJ. A Systems-Level Approach to Temporal Transcriptional Profiling of Retinas in a Rat Model of Smith-Lemli-Opitz Syndrome. *ARVO Annu Meet Abstr* 2008;
230. Korade Z, Xu L, Shelton R, Porter N a. Biological activities of 7-dehydrocholesterol-derived oxysterols: implications for Smith-Lemli-Opitz syndrome. *J Lipid Res* 2010;51:3259–3269.
231. Fu R, Yanjanin NM, Bianconi S, Pavan WJ, Porter FD. Oxidative stress in Niemann-Pick disease, type C. *Mol Genet Metab* 2010;101:214–218.
232. Zampieri S, Mellon SH, Butters TD, Nevyjel M, Covey DF, Bembi B, Dardis A. Oxidative stress in NPC1 deficient cells: Protective effect of allopregnanolone. *J Cell Mol Med* 2009;13:3786–3796.
233. Ribas GS, Pires R, Coelho JC, Rodrigues D, Mescka CP, Vanzin CS, Biancini GB, Negretto G, Wayhs CAY, Wajner M, Vargas CR. Oxidative stress in Niemann-Pick type C patients: a protective role of N-butyl-deoxynojirimycin therapy. *Int J Dev Neurosci* 2012;30:439–44.
234. Calder PC, Yaqoob P. Lipid rafts - composition, characterization, and controversies. *J Nutr* 2007;137:545–547.
235. Hakomori S. Glycosphingolipids in Cellular Interaction, Differentiation and Oncogenesis. *Ann Rev Biochem* 1981;50:733–64.
236. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P,

- Chaisson M, Gingeras TR. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* 2013;29:15–21.
237. Liao Y, Smyth GK, Shi W. FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 2014;30:923–930.
238. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;15:550.
239. Maier T, Güell M, Serrano L. Correlation of mRNA and protein in complex biological samples. *FEBS Lett*. 2009;583:3966–3973.
240. Allen Institute for Brain Science. Allen Mouse Brain Atlas [Internet]. Allen Mouse Brain Atlas. 2009;Available from: <http://mouse.brain-map.org>
241. Wassif, C.A; Lloyd-Evans, E.;Haslett, L.J.;Williams, I.M.; Toth, C.L.;Porter, F.D.;Platt FM. Sterol precursors induce Niemann-Pick C disease phenotypes in Smith-Lemli-Opitz Syndrome causing defective LDL-cholesterol utilization that is corrected by imino-sugars. 2012.
242. Efthymiou AG, Steiner J, Pavan WJ, Wincovitch S, Larson DM, Porter FD, Rao MS, Malik N. Rescue of an in vitro neuron phenotype identified in Niemann-Pick disease, type C1 induced pluripotent stem cell-derived neurons by modulating the WNT pathway and calcium signaling. *Stem Cells Transl Med* 2015;4:230–238.
243. Lacroix M. Significance, detection and markers of disseminated breast cancer cells. *Endocr. Relat. Cancer*. 2006;13:1033–1067.
244. Giuliani D, Neri L, Canalini F, Calevro A, Ottani A, Vandini E, Sena P, Zaffe D, Guarini S. NDP- α -MSH induces intense neurogenesis and cognitive recovery in Alzheimer transgenic mice through activation of melanocortin MC₄ receptors. *Mol Cell Neurosci* 2015;67:13–21.

245. Ramirez D, Saba J, Carniglia L, Durand D, Lasaga M, Caruso C. Melanocortin 4 receptor activates ERK-cFos pathway to increase brain-derived neurotrophic factor expression in rat astrocytes and hypothalamus. *Mol Cell Endocrinol* 2014;411:28–37.
246. Logan CY, Nusse R. The Wnt signaling pathway in development and disease. *Annu Rev Cell Dev Biol* 2004;20:781–810.
247. Blagodatski A, Poteryaev D, Katanaev VL. Targeting the Wnt pathways for therapies. *Mol Cell Ther* 2014;2:28.
248. Omasits U, Ahrens CH, Müller S, Wollscheid B. Protter: Interactive protein feature visualization and integration with experimental proteomic data. *Bioinformatics* 2014;30:884–886.

Appendices

***DHCR7* Summary Table**

Summary table of variants in *DHCR7*

The 'cDNA' column designates the variant change seen in the cDNA from the data sets. The 'Protein' column shows the amino acid change in the protein. 'rs#' is the reference number for the CSNB. The 'Polyphen/Published/MaxEntScan' column refers to the prediction of the variant by Polyphen-2 or MaxEntScan and if the variant has been previously published. As all variants were not discovered in every allele, the 'total alleles' column details the total number of alleles analysed for that particular variant. The subsequent four columns show the number of times that variant occurs within each data set. The final column, 'rate' is the carrier frequency for that variant.

cDNA	Protein	rs#	Polyphen/Published/MaxEntScan	Total Alleles	Total Variants	NHLBI	1000 Genome	Clinseq	Autism	Rate
c.3G>A	p.M1I	rs121909767	POSSIBLY DAMAGING/Published	17836	1	1	0	0	0	0.006%
c.14C>T	p.S5L	rs1127869	BENIGN	17836	2	0	2	0	0	0.011%
c.25A>T	p.I9F	rs115595829	BENIGN	17502	171	153	13	3	2	0.977%
c.28C>G	p.P10A	rs139166382	BENIGN	17836	3	3	0	0	0	0.017%
c.70G>T	p.A24S	rs146867923	BENIGN	17736	43	37	4	0	2	0.242%
c.77A>C	p.Q26P	rs199815542	BENIGN	17836	1	0	0	1	0	0.006%
c.89G>C	p.G30A	rs200334114	PROBABLY DAMAGING/Published	17836	1	0	1	0	0	0.006%
c.91C>T	p.R31C	rs367585401	PROBABLY DAMAGING	17836	1	1	0	0	0	0.006%
c.92G>A	p.R31H	rs370307688	PROBABLY DAMAGING	17836	2	2	0	0	0	0.011%
c.99(-4)G>A	intronic	rs140748737	Neutral/Published	17736	27	0	19	0	8	0.152%
c.127G>A	p.V43I	rs200984695	BENIGN	17836	1	1	0	0	0	0.006%
c.131T>C	p.I44T	rs142897396	BENIGN	17836	5	5	0	0	0	0.028%
c.165C>G	p.Y55X	rs374941029	PROBABLY DAMAGING	17836	1	1	0	0	0	0.006%
c.199G>A	p.A67T	rs143999854	BENIGN	17746	12	10	1	1	0	0.068%
c.208G>A	p.G70S	rs144512551	BENIGN	17772	22	19	2	1	0	0.124%
c.221A>C	p.D74A	rs374874199	BENIGN	17836	1	1	0	0	0	0.006%
c.223A>T	p.I75F	rs370748173	BENIGN	17836	2	2	0	0	0	0.011%
c.226G>A	p.V76I		BENIGN	17502	1	0	0	0	1	0.006%
c.232G>A	p.G78R	rs373352413	PROBABLY DAMAGING	17836	1	1	0	0	0	0.006%
c.265A>T	p.T89S	rs375997113	POSSIBLY DAMAGING	17836	1	1	0	0	0	0.006%
c.278C>T	p.T93M	rs80338853	PROBABLY DAMAGING/Published	17836	1	1	0	0	0	0.006%
c.289G>A	p.A97T	rs150563256	BENIGN	17836	5	1	4	0	0	0.028%
c.292C>T	p.Q98X	rs104886039	PROBABLY DAMAGING/Published	17822	2	1	0	1	0	0.011%
c.321G>C	p.Q107H	rs104886040	PROBABLY DAMAGING/Published	17836	1	1	0	0	0	0.006%
c.376G>A	p.V126I	rs143587828	PROBABLY DAMAGING	17836	2	2	0	0	0	0.011%
c.385A>G	p.I129V	rs138215017	BENIGN	17836	1	1	0	0	0	0.006%
c.400G>T	p.V134L	rs201466849	BENIGN/Published	17308	3	1	1	0	1	0.017%
c.418G>A	p.V140M	rs373908315	BENIGN	17836	1	1	0	0	0	0.006%
c.452G>A	p.W151X	rs11555217	PROBABLY DAMAGING/Published	17422	15	11	0	4	0	0.086%
c.461C>G	p.T154R	rs143312232	PROBABLY DAMAGING/Published	17494	10	5	0	0	5	0.057%
c.506C>T	p.S169L	rs80338855	POSSIBLY DAMAGING/Published	17836	1	1	0	0	0	0.006%
c.509C>G	p.P170R	rs150459687	PROBABLY DAMAGING	17836	1	1	0	0	0	0.006%
c.523G>A	p.D175N	rs368269558	POSSIBLY DAMAGING	17836	1	1	0	0	0	0.006%
c.583G>A	p.A195T	rs151170252	BENIGN	17836	3	3	0	0	0	0.017%
c.603C>A	p.F201L	rs377727130	BENIGN	17836	1	1	0	0	0	0.006%

c.603C>A	p.F201L	rs377727130	BENIGN	17836	1	1	1	0	0	0	0.006%
c.658A>C	p.M220L	rs200659923	POSSIBLY DAMAGING	17836	1	0	0	1	0	0	0.006%
c.670G>A	p.E224K	rs373121544	PROBABLY DAMAGING/Published	17836	1	1	1	0	0	0	0.006%
c.683G>A	p.R228Q	rs201556114	POSSIBLY DAMAGING	17836	1	0	0	1	0	0	0.006%
c.719A>G	p.N240S	rs148609143	PROBABLY DAMAGING	17836	5	5	5	0	0	0	0.028%
c.724C>T	p.R242C	rs80338856	PROBABLY DAMAGING/Published	17836	2	2	2	0	0	0	0.011%
c.822C>A	p.N274K	rs139787408	POSSIBLY DAMAGING/Published	17836	1	1	1	0	0	0	0.006%
c.849_856CTTCT				17824	1	0	0	0	1	0	0.006%
c.852C>A	p.F284L	rs184297154	PROBABLY DAMAGING/Published	17836	1	0	0	1	0	0	0.006%
c.883A>G	p.I295V	rs201574502	PROBABLY DAMAGING	17836	1	0	0	1	0	0	0.006%
c.907G>A	p.G303R	rs142808899	PROBABLY DAMAGING/Published	17836	4	4	4	0	0	0	0.022%
c.928G>A	p.D310N	rs370955781	PROBABLY DAMAGING	17836	1	1	1	0	0	0	0.006%
c.964(-1)G>C	intrinsic	rs138659167	Negative/Published	16736	94	70	70	10	9	5	0.562%
c.988G>A	p.V330M	rs139724817	POSSIBLY DAMAGING	17812	8	5	5	2	0	1	0.045%
c.1004C>G	p.P335R	rs77762671	BENIGN	17836	2	0	0	2	0	0	0.011%
c.1012G>A	p.V338M	rs72954276	BENIGN	17568	14	9	9	4	1	0	0.080%
c.1018G>A	p.V340I	rs148081697	BENIGN	17830	1	1	1	0	0	0	0.006%
c.1025T>C	p.L342P	rs199957106	POSSIBLY DAMAGING	17836	1	0	0	1	0	0	0.006%
c.1084C>T	p.R362C	rs371302153	PROBABLY DAMAGING/Published	17832	1	1	1	0	0	0	0.006%
c.1085G>A	p.R362H	rs142213147	PROBABLY DAMAGING	17836	2	2	2	0	0	0	0.011%
c.1088G>A	p.R363H	rs200539324	POSSIBLY DAMAGING	17836	1	0	0	1	0	0	0.006%
c.1138T>C	p.C380R	rs373306653	PROBABLY DAMAGING/Published	17836	1	1	1	0	0	0	0.006%
c.1156G>A	p.D386N	rs369837196	PROBABLY DAMAGING	17836	1	1	1	0	0	0	0.006%
c.1204G>A	p.V402M	rs200099137	BENIGN	17822	1	1	1	0	0	0	0.006%
c.1210C>T	p.R404C	rs61757582	PROBABLY DAMAGING/Published	17574	1	0	0	0	1	0	0.006%
c.1228G>C	p.G410R	rs372055524	PROBABLY DAMAGING/Published	17502	1	0	0	0	0	1	0.006%
c.1231G>A	p.D411N	rs372055524	PROBABLY DAMAGING	17828	2	2	2	0	0	0	0.011%
c.1270G>T	p.G424C	rs368150818	PROBABLY DAMAGING	17810	1	1	1	0	0	0	0.006%
c.1270G>A	p.G424S		BENIGN	17442	1	0	0	0	1	0	0.006%
c.1290C>G	p.Y430X	rs140791666	PROBABLY DAMAGING	17828	1	1	1	0	0	0	0.006%
c.1300A>G	p.I434V	rs375187933	BENIGN	17832	1	1	1	0	0	0	0.006%
c.1336C>T	p.R446W	rs145043679	PROBABLY DAMAGING/Published	17834	2	2	2	0	0	0	0.011%
c.1342G>A	p.E448K	rs80338864	PROBABLY DAMAGING/Published	17572	1	0	0	0	1	0	0.006%
c.1354G>A	p.A452T	rs140400648	BENIGN	17836	2	0	0	2	0	0	0.011%
c.1366G>A	p.G456S	rs201847193	PROBABLY DAMAGING	17836	1	0	0	1	0	0	0.006%
c.1369C>T	p.R457W	rs371873032	PROBABLY DAMAGING	17832	2	2	2	0	0	0	0.011%
c.1381C>T	p.R461C	rs199506852	POSSIBLY DAMAGING	17836	2	0	0	2	0	0	0.011%
c.1384T>C	p.Y462H	rs201270451	PROBABLY DAMAGING/Published	17630	1	0	0	0	1	0	0.006%
c.1405C>T	p.R469C	rs148660993	PROBABLY DAMAGING	17818	2	2	2	0	0	0	0.011%
c.1406G>A	p.R469H	rs201150384	PROBABLY DAMAGING	17836	1	0	0	1	0	0	0.006%

NPC1 Summary Table

Summary tables of *NPC1* and *NPC2*

The ‘cDNA’ column designates the variant change seen in the cDNA from the data sets. The ‘Protein’ column shows the amino acid change in the protein. ‘rs#’ is the reference number for the CSNB. The ‘Polyphen/Published/MaxEntScan’ column refers to the prediction of the variant by Polyphen-2 or MaxEntScan and if the variant has been previously published. The next two columns are the pathogenicity predictions of each variant as designated by SIFT and Mutation Assessor respectively. Not all variants locations were sequenced in all alleles and were, therefore, not ascertained in every sample. Correspondingly, the ‘total alleles’ column details the total number of alleles analysed for that particular variant. The subsequent four columns show the number of times that variant occurs within each data set. The final column, ‘rate’ is the carrier frequency for that variant.

cDNA	Protein	rs#	Polyphen-2/MaxEntScan/Published	SIFT	MutationAssessor Pred	Total Alleles	Total Variants	NHLBI	1000 Genome	Clinseq	Autism	Rate
c.110A>G	p.D37G		BENIGN	BENIGN	Predicted non-functional (low)	17754	1	0	0	1	0	0.006%
c.127G>A	p.E43K	rs138277307	BENIGN	BENIGN	Predicted non-functional (neutral)	17754	1	1	0	0	0	0.006%
c.180G>T	p.Q60H	rs145666943	BENIGN	POSSIBLY DAMAGING	Predicted functional (medium)	17754	5	4	1	0	0	0.028%
c.181-4A>C	intronic	rs374571310	Neutral			17754	2	2	0	0	0	0.011%
c.181-3A>G	intronic	rs371126954	Strong Negative Effect			17754	1	1	0	0	0	0.006%
c.209A>G	p.N70S	rs200291759	BENIGN	BENIGN	Predicted non-functional (low)	17754	1	0	1	0	0	0.006%
c.233G>A	p.R78Q	rs373274825	BENIGN	BENIGN	Predicted non-functional (neutral)	17754	1	1	0	0	0	0.006%
c.346C>T	p.R116X	rs144973225	PROBABLY DAMAGING/Published			17754	1	1	0	0	0	0.006%
c.347G>A	p.R116Q	rs140952850	BENIGN	BENIGN	Predicted non-functional (neutral)	17754	2	2	0	0	0	0.011%
c.410C>T	p.T137M	rs372947142	BENIGN/Published	BENIGN	Predicted non-functional (low)	17754	1	1	0	0	0	0.006%
c.424_425insGA	p.K142Rfs					17266	1	1	0	0	0	0.006%
c.442G>C	p.V148L	rs200323346	BENIGN	BENIGN	Predicted non-functional (neutral)	17754	2	0	2	0	0	0.011%
c.445G>A	p.G149R	rs143205855	BENIGN	BENIGN	Predicted non-functional (low)	17754	2	1	1	0	0	0.011%
c.449A>G	p.Q150R	rs375940577	BENIGN	BENIGN	Predicted non-functional (neutral)	17754	1	1	0	0	0	0.006%
c.466A>G	p.M156V	rs149074243	BENIGN	BENIGN	Predicted non-functional (low)	17754	6	6	0	0	0	0.034%
c.467T>C	p.M156T	rs147615070	BENIGN	BENIGN	Predicted non-functional (low)	17754	2	1	0	1	0	0.011%
c.481C>T	p.R161W	rs141243713	PROBABLY DAMAGING	PROBABLY DAMAGING	Predicted functional (medium)	17754	2	2	0	0	0	0.011%
c.520G>C	p.G174R	rs370098528	PROBABLY DAMAGING	POSSIBLY DAMAGING	Predicted functional (medium)	17754	1	1	0	0	0	0.006%
c.544G>A	p.D182N	rs201021988	BENIGN	BENIGN	Predicted non-functional (neutral)	17754	2	1	1	0	0	0.011%
c.547G>A	p.A183T	rs111256741	BENIGN	BENIGN	Predicted non-functional (neutral)	17754	7	4	3	0	0	0.039%
c.548C>T	p.A183V	rs192963719	BENIGN	BENIGN	Predicted non-functional (neutral)	17754	1	0	1	0	0	0.006%
c.553A>G	p.N185D	rs139485263	BENIGN	POSSIBLY DAMAGING	Predicted functional (medium)	17752	4	3	0	1	0	0.023%
c.622G>C	p.V208L	rs372416248	BENIGN	BENIGN	Predicted non-functional (low)	17754	1	1	0	0	0	0.006%
c.631G>T	p.D211V	rs367851289	PROBABLY DAMAGING	BENIGN	Predicted non-functional (low)	17754	1	1	0	0	0	0.006%
c.644A>G	p.H215R	rs1805081	BENIGN/known polymorphism	BENIGN	Predicted non-functional (low)	17696	5250	3849	535	646	220	29.668%
c.665A>G	p.N222S	rs55680026	BENIGN/Published	BENIGN	Predicted non-functional (low)	17748	71	59	2	9	1	0.400%
c.688_693del	p.S230_V231del		Published			17226	3	3	0	0	0	0.017%
c.695A>G	p.D232G	rs201956601	BENIGN	BENIGN	Predicted functional (medium)	17754	1	0	1	0	0	0.006%
c.709C>T	p.P237S	rs80358251	BENIGN/Published benign	BENIGN	Predicted non-functional (low)	17730	183	150	14	15	4	1.032%
c.749A>C	p.K250T		BENIGN	BENIGN	Predicted non-functional (low)	17754	1	0	0	0	1	0.006%
c.763C>T	p.P255S	rs373815982	POSSIBLY DAMAGING	BENIGN	Predicted non-functional (low)	17754	1	1	0	0	0	0.006%
c.764C>G	p.P255R	rs371023983	PROBABLY DAMAGING	BENIGN	Predicted functional (medium)	17754	1	1	0	0	0	0.006%
c.769C>T	p.P257S	rs368776731	BENIGN	POSSIBLY DAMAGING	Predicted functional (medium)	17754	1	1	0	0	0	0.006%
c.782C>T	p.T261M	rs374169117	BENIGN	BENIGN	Predicted non-functional (neutral)	17754	1	1	0	0	0	0.006%
c.797A>G	p.D266G	rs370188327	POSSIBLY DAMAGING	POSSIBLY DAMAGING	Predicted functional (medium)	17754	1	1	0	0	0	0.006%
c.806A>G	p.Y269C		POSSIBLY DAMAGING	BENIGN	Predicted non-functional (neutral)	17746	1	0	0	1	0	0.006%
c.811A>G	p.I271V	rs370810779	BENIGN	BENIGN	Predicted non-functional (low)	17754	1	1	0	0	0	0.006%
c.841C>T	p.L281F	rs377132020	BENIGN	BENIGN	Predicted non-functional (low)	17754	1	1	0	0	0	0.006%
c.873G>T	p.W291C	rs138151007	BENIGN			17754	10	7	2	1	0	0.056%
c.901G>A	p.E301K	rs150154006	POSSIBLY DAMAGING	BENIGN	Predicted non-functional (low)	17754	1	1	0	0	0	0.006%
c.962C>T	p.A321V	rs138079168	BENIGN	BENIGN	Predicted non-functional (low)	17750	2	1	1	0	0	0.011%
c.979G>A	p.V327I	rs141361998	BENIGN	BENIGN	Predicted non-functional (neutral)	17754	1	1	0	0	0	0.006%
c.1001G>C	p.C334S	rs199693280	BENIGN	BENIGN	Predicted non-functional (neutral)	17754	1	0	0	1	0	0.006%
c.1010G>A	p.R337Q	rs373390781	BENIGN	BENIGN	Predicted non-functional (low)	17754	1	1	0	0	0	0.006%
c.1022G>C	p.R341P	rs370181667	BENIGN	BENIGN	Predicted non-functional (low)	17754	1	1	0	0	0	0.006%
c.1039G>A	p.V347I	rs376741451	BENIGN	BENIGN	Predicted non-functional (low)	17754	2	2	0	0	0	0.011%
c.1055G>T	p.C352F	rs149020783	BENIGN	BENIGN	Predicted non-functional (low)	17754	2	2	0	0	0	0.011%

c.1094C>T	p.S365L	rs200243024	POSSIBLY DAMAGING			17754	2	0	0	2	0	0.011%
c.1115G>A	p.R372Q	rs150053420	BENIGN	BENIGN	Predicted non-functional (neutral)	17754	1	1	0	0	0	0.006%
c.1166G>A	p.R389H	rs373751051	POSSIBLY DAMAGING	POSSIBLY DAMAGING	Predicted functional (medium)	17754	1	1	0	0	0	0.006%
c.1208T>C	p.F403S	rs371234970	PROBABLY DAMAGING	PROBABLY DAMAGING	Predicted functional (medium)	17754	1	1	0	0	0	0.006%
c.1211G>A	p.R404Q	rs139751448	PROBABLY DAMAGING/Published	PROBABLY DAMAGING	Predicted functional (high)	17754	1	1	0	0	0	0.006%
c.1232G>A	p.R411Q	rs77080672	BENIGN	BENIGN	Predicted non-functional (neutral)	17754	34	26	5	1	2	0.192%
c.1270C>G	p.P424A	rs143797098	BENIGN	BENIGN	Predicted non-functional (neutral)	17754	1	1	0	0	0	0.006%
c.1274C>T	p.S425L	rs140149624	BENIGN	BENIGN	Predicted non-functional (neutral)	17754	1	1	0	0	0	0.006%
c.1300C>T	p.P434S	rs61731962	BENIGN/known polymorphism	BENIGN	Predicted non-functional (low)	17752	212	189	18	3	2	1.194%
c.1346C>T	p.A449V	rs372289265	BENIGN	BENIGN	Predicted non-functional (low)	17754	1	1	0	0	0	0.006%
c.1348A>G	p.I450V	rs141892620	BENIGN	BENIGN	Predicted non-functional (low)	17752	5	3	0	2	0	0.028%
c.1367C>T	p.S456F	rs374159264	BENIGN	BENIGN	Predicted non-functional (low)	17754	1	1	0	0	0	0.006%
c.1412C>T	p.P471L	rs201226297	PROBABLY DAMAGING	PROBABLY DAMAGING	Predicted functional (medium)	17754	1	0	1	0	0	0.006%
c.1421C>T	p.P474L	rs372445155	PROBABLY DAMAGING/Published	PROBABLY DAMAGING	Predicted functional (medium)	17754	1	1	0	0	0	0.006%
c.1472G>A	p.S491N	rs370758521	BENIGN	BENIGN	Predicted non-functional (low)	17754	1	1	0	0	0	0.006%
c.1480G>A	p.V494M	rs199812609	BENIGN	BENIGN	Predicted non-functional (low)	17754	1	1	0	0	0	0.006%
c.1506C>G	p.D502E	rs191537721	BENIGN	BENIGN	Predicted non-functional (low)	17754	1	0	1	0	0	0.006%
c.1532C>T	p.T511M	rs13381670	PROBABLY DAMAGING			17754	51	42	8	0	1	0.287%
c.1549G>A	p.V517I	rs201791992	BENIGN	BENIGN	Predicted non-functional (low)	17754	1	0	1	0	0	0.006%
c.1552C>T	p.R518W	rs377515417	PROBABLY DAMAGING/Published	PROBABLY DAMAGING	Predicted functional (medium)	17754	2	2	0	0	0	0.011%
c.1561G>T	p.A521S	rs138184115	BENIGN/Published	BENIGN	Predicted non-functional (neutral)	17754	9	9	0	0	0	0.051%
c.1628C>T	p.P543L	rs369368181	PROBABLY DAMAGING/Published	PROBABLY DAMAGING	Predicted functional (medium)	17754	1	1	0	0	0	0.006%
c.1672G>T	p.A558S	rs201156397	POSSIBLY DAMAGING/Published	PROBABLY DAMAGING	Predicted functional (medium)	17704	1	0	0	1	0	0.006%
c.1756G>A	p.E586K	rs369753548	BENIGN	BENIGN	Predicted non-functional (low)	17754	2	2	0	0	0	0.011%
c.1766A>G	p.N589S	rs147021046	BENIGN	BENIGN	Predicted non-functional (low)	17732	8	7	0	1	0	0.045%
c.1780_1781insT	p.Y594Lfs					17266	1	1	0	0	0	0.006%
c.1793A>G	p.N598S	rs201236716	BENIGN	PROBABLY DAMAGING	Predicted non-functional (low)	17754	1	0	1	0	0	0.006%
c.1870G>A	p.V624I	rs76615690	BENIGN	BENIGN	Predicted non-functional (neutral)	17754	2	0	2	0	0	0.011%
c.1901A>G	p.V634C	rs202140203	PROBABLY DAMAGING/Published	PROBABLY DAMAGING	Predicted functional (medium)	17752	1	0	0	1	0	0.006%
c.1936C>T	p.R646C	rs368129141	POSSIBLY DAMAGING	BENIGN	Predicted non-functional (neutral)	17754	1	1	0	0	0	0.006%
c.1937G>A	p.R646H	rs112387560	BENIGN	BENIGN	Predicted non-functional (low)	17754	7	4	3	0	0	0.039%
c.1976C>T	p.A659V	rs140786703	POSSIBLY DAMAGING	PROBABLY DAMAGING	Predicted functional (medium)	17754	1	1	0	0	0	0.006%
c.1990G>A	p.V664M	rs376213990	PROBABLY DAMAGING/Published	PROBABLY DAMAGING	Predicted functional (medium)	17754	1	1	0	0	0	0.006%
c.2020_2021del	p.V674Lfs					17266	319	319	0	0	0	1.848%
c.2027G>C	p.S676T		PROBABLY DAMAGING	PROBABLY DAMAGING	Predicted functional (medium)	17754	1	0	0	0	1	0.006%
c.2083C>G	p.L695V	rs370323921	PROBABLY DAMAGING/Published	PROBABLY DAMAGING	Predicted functional (high)	17754	1	1	0	0	0	0.006%
c.2141G>A	p.R714H	rs375047023	PROBABLY DAMAGING	PROBABLY DAMAGING	Predicted non-functional (low)	17754	2	2	0	0	0	0.011%
c.2209C>G	p.L737V	rs201100763	PROBABLY DAMAGING	PROBABLY DAMAGING	Predicted functional (medium)	17754	1	0	0	1	0	0.006%
c.2257G>A	p.V753M	rs146874573	BENIGN	PROBABLY DAMAGING	Predicted non-functional (low)	17754	2	2	0	0	0	0.011%
c.2338G>A	p.V780M	rs193182840	PROBABLY DAMAGING	PROBABLY DAMAGING	Predicted functional (medium)	17754	1	0	1	0	0	0.006%
c.2428G>T	p.V810L	rs145362908	BENIGN	BENIGN	Predicted non-functional (neutral)	17754	8	5	3	0	0	0.045%
c.2428G>C	p.V810L	rs145362908	BENIGN	BENIGN	Predicted non-functional (neutral)	17752	1	0	0	1	0	0.006%
c.2501T>C	p.M834T	rs373435883	BENIGN	BENIGN	Predicted non-functional (neutral)	17754	1	1	0	0	0	0.006%
* c.2524T>C	p.F842L	rs190298665	PROBABLY DAMAGING			17754	1	0	1	0	0	0.006%
c.2525T>C	p.F842S	rs374068891	PROBABLY DAMAGING	PROBABLY DAMAGING	Predicted functional (medium)	17754	1	1	0	0	0	0.006%
c.2551G>A	p.A851T	rs139297968	POSSIBLY DAMAGING	PROBABLY DAMAGING	Predicted functional (medium)	17754	2	2	0	0	0	0.011%
c.2572A>G	p.I858V	rs1805082	BENIGN/known polymorphism	BENIGN	Predicted non-functional (neutral)	17752	8005	5758	1100	864	283	45.094%
c.2605-6_2605-3del	intronic		Neutral			17266	1	1	0	0	0	0.006%
c.2621A>T	p.D874V	rs372030650	POSSIBLY DAMAGING/Published	BENIGN	Predicted functional (medium)	17754	1	1	0	0	0	0.006%

c.2705C>G	p.S902C	rs374656358	BENIGN	BENIGN	Predicted non-functional (low)	17754	1	1	0	0	0	0	0.006%
c.2731G>A	p.G911S	rs34302553	BENIGN	POSSIBLY DAMAGING	Predicted non-functional (low)	17754	69	63	3	0	0	3	0.389%
c.2796-4C>T	intronic	rs374406578	Neutral			17754	1	1	0	0	0	0	0.006%
c.2800C>T	p.R934X	rs370721218	PROBABLY DAMAGING/Published			17754	1	1	0	0	0	0	0.006%
c.2819C>T	p.S940L	rs143124972	PROBABLY DAMAGING/Published			17754	1	1	0	0	0	0	0.006%
c.2873G>A	p.R958Q	rs120074132	PROBABLY DAMAGING/Published	PROBABLY DAMAGING	Predicted functional (medium)	17754	1	1	0	0	0	0	0.006%
c.2882A>G	p.N961S	rs34084984	BENIGN/Published	BENIGN	Predicted non-functional (low)	17754	69	56	11	1	1	1	0.389%
c.2908_2909insTT	p.S970Ffs					17266	1	1	0	0	0	0	0.006%
c.2911+4C>T	intronic	rs186588103	Neutral			17754	3	3	0	0	0	0	0.017%
c.2929+4C>T	intronic	rs186588103	Neutral			17754	1	0	1	0	0	0	0.006%
c.2972_2973del	p.991_fs		Published	POSSIBLY DAMAGING	Predicted functional (medium)	17724	2	0	0	0	1	1	0.111%
c.2974G>T	p.G929W	rs80358254	PROBABLY DAMAGING/Published	POSSIBLY DAMAGING	Predicted functional (medium)	17754	1	1	0	0	0	0	0.006%
c.3011C>T	p.S1004L	rs150334966	PROBABLY DAMAGING/Published			17750	11	8	0	2	1	1	0.062%
c.3019C>G	p.P1007A	rs80358257	PROBABLY DAMAGING/Published	PROBABLY DAMAGING	Predicted functional (medium)	17754	3	2	1	1	0	0	0.017%
c.3028A>C	p.K1010Q	rs191876836	BENIGN	BENIGN	Predicted functional (medium)	17754	1	0	1	0	0	0	0.006%
c.3047A>G	p.H1016R	rs140211089	POSSIBLY DAMAGING/Published	PROBABLY DAMAGING	Predicted functional (medium)	17754	1	1	0	0	0	0	0.006%
c.3057G>A	p.A1018T	rs146666146	PROBABLY DAMAGING	POSSIBLY DAMAGING	Predicted functional (high)	17754	2	2	0	0	0	0	0.111%
c.3059G>C	p.S1020T	rs374719153	BENIGN	BENIGN	Predicted functional (medium)	17754	1	1	0	0	0	0	0.006%
c.3182T>C	p.I1061T	rs80358259	BENIGN/Published	POSSIBLY DAMAGING	Predicted functional (medium)	17754	5	5	0	0	0	0	0.028%
c.3184G>A	p.A1062T	rs369960141	POSSIBLY DAMAGING	POSSIBLY DAMAGING	Predicted functional (medium)	17754	1	1	0	0	0	0	0.006%
c.3217G>A	p.G1073S	rs141440861	BENIGN	BENIGN	Predicted non-functional (neutral)	17750	23	18	4	1	0	0	0.130%
c.3265G>A	p.E1089K	rs374526072	PROBABLY DAMAGING/Published	POSSIBLY DAMAGING	Predicted functional (high)	17754	1	1	0	0	0	0	0.006%
c.3343G>T	p.V1115F	rs34226296	BENIGN	BENIGN	Predicted non-functional (neutral)	17754	43	34	7	0	0	2	0.242%
c.3364T>C	p.W1122R	rs148571882	BENIGN	BENIGN	Predicted non-functional (low)	17754	2	1	0	0	0	1	0.011%
c.3422T>G	p.V1141G	rs144725473	PROBABLY DAMAGING/Published	POSSIBLY DAMAGING	Predicted non-functional (low)	17754	1	1	0	0	0	0	0.006%
c.3477+4A>G	intronic	rs114073738	Neutral			17754	28	28	0	0	0	0	0.158%
c.3498+4A>G	intronic		Negative			17754	4	0	4	0	0	0	0.023%
c.3506G>T	p.S1169I	rs139612110	PROBABLY DAMAGING	PROBABLY DAMAGING	Predicted functional (medium)	17754	2	2	0	0	0	0	0.011%
c.3535A>G	p.M1179V	rs61731969	BENIGN	BENIGN	Predicted non-functional (neutral)	17694	61	54	5	2	0	0	0.345%
c.3548G>A	p.R1183H	rs148035987	PROBABLY DAMAGING	PROBABLY DAMAGING	Predicted functional (medium)	17754	8	7	0	0	0	1	0.045%
c.3550G>A	p.V1184M		POSSIBLY DAMAGING	PROBABLY DAMAGING	Predicted functional (medium)	17676	1	0	0	0	1	0	0.006%
c.3556C>T	p.R1186C	rs145297180	PROBABLY DAMAGING	PROBABLY DAMAGING	Predicted functional (medium)	17754	2	2	0	0	0	0	0.011%
c.3557G>A	p.R1186H	rs200444084	PROBABLY DAMAGING/Published	PROBABLY DAMAGING	Predicted functional (medium)	17754	2	1	1	0	0	0	0.011%
c.3560C>T	p.A1187V	rs113371321	POSSIBLY DAMAGING/Published	PROBABLY DAMAGING	Predicted functional (high)	17754	1	0	1	0	0	0	0.006%
c.3566A>G	p.E1189G	rs369098773	POSSIBLY DAMAGING/Published	BENIGN	Predicted non-functional (low)	17754	1	1	0	0	0	0	0.006%
c.3577C>T	p.H1193Y	rs375309094	BENIGN	BENIGN	Predicted non-functional (neutral)	17754	1	1	0	0	0	0	0.006%
c.3598A>G	p.S1200G	rs35248744	PROBABLY DAMAGING/Published	PROBABLY DAMAGING	Predicted functional (medium)	17754	63	61	1	0	0	1	0.355%
c.3611_3614del	p.L1204Qfs		Published			17264	1	1	0	0	0	0	0.006%
c.3619T>C	p.F1207L	rs140827681	BENIGN	POSSIBLY DAMAGING	Predicted non-functional (low)	17754	1	1	0	0	0	0	0.006%
c.3667A>G	p.I1223V	rs368658600	BENIGN	POSSIBLY DAMAGING	Predicted functional (medium)	17754	1	1	0	0	0	0	0.006%
c.3689T>C	p.L1230S	rs374150662	PROBABLY DAMAGING	PROBABLY DAMAGING	Predicted functional (medium)	17754	1	1	0	0	0	0	0.006%
c.3741_3745A			Published			17748	1	0	0	0	1	0	0.006%
c.3742_3745del	p.L1248Vfs		Published			17266	1	1	0	0	0	0	0.006%
c.3755-5_3755-4insTC	intronic		Neutral			17266	4	4	0	0	0	0	0.023%
c.3796C>T	p.R1266X	rs376164368	PROBABLY DAMAGING			17754	1	1	0	0	0	0	0.006%
c.3797G>A	p.R1266Q	rs1805084	BENIGN/Known polymorphism	BENIGN	Predicted non-functional (low)	17754	1724	1275	307	101	41	9.710%	
c.3799T>G	p.Y1267D	rs373435628	BENIGN	BENIGN	Predicted non-functional (low)	17754	1	1	0	0	0	0	0.006%
c.3811G>C	p.E1271Q	rs140527006	POSSIBLY DAMAGING	BENIGN	Predicted functional (medium)	17754	1	1	0	0	0	0	0.006%
c.3814C>T	p.R1272C	rs200264267	PROBABLY DAMAGING	BENIGN	Predicted non-functional (neutral)	17754	1	0	1	0	0	0	0.006%
c.3818A>G	p.E1273G	rs374032318	BENIGN	BENIGN	Predicted non-functional (low)	17754	1	1	0	0	0	0	0.006%
c.3821G>A	p.R1274Q	rs151305963	BENIGN	BENIGN	Predicted non-functional (neutral)	17754	5	4	0	0	0	1	0.028%

NPC2 Summary Table

cdNA	Protein	rs#	Polyphen-2/MaxEntScan/Published	SIFT	MutationAssessor Pred	Total Alleles	Total Variants	NHLBI	1000 Genome	Clinseq	Autism	Rate
c.38T>C	p.L13P	rs147602717	PROBABLY DAMAGING	POSSIBLY DAMAGING	Predicted functional (medium)	17360	5	4	0	1	0	0.029%
c.49G>A	p.A17T	rs145302203	BENIGN	BENIGN	Predicted non-functional (low)	17748	1	1	0	0	0	0.006%
c.56C>A	p.A19D	rs369392502	PROBABLY DAMAGING	PROBABLY DAMAGING	Predicted functional (medium)	17748	1	1	0	0	0	0.006%
c.58G>T	p.E20X	rs80358260	Published	PROBABLY DAMAGING	Predicted functional (medium)	17750	1	1	0	0	0	0.006%
c.88G>A	p.V30M	rs151220873	PROBABLY DAMAGING/Published	POSSIBLY DAMAGING	Predicted non-functional (low)	17252	34	25	1	4	4	0.197%
c.115G>A	p.V39M	rs80358261	PROBABLY DAMAGING/Published	PROBABLY DAMAGING	Predicted functional (medium)	17754	1	1	0	0	0	0.006%
c.212A>G	p.K71R	rs142075589	POSSIBLY DAMAGING	BENIGN	Predicted non-functional (neutral)	17754	3	2	1	0	0	0.017%
c.224A>T	p.H75L	rs369221608	PROBABLY DAMAGING		Predicted functional (medium)	17754	1	1	0	0	0	0.006%
c.271G>A	p.D91N	rs148607507	POSSIBLY DAMAGING	POSSIBLY DAMAGING	Predicted non-functional (low)	17754	10	8	0	2	0	0.056%
c.278G>T	p.C93F	rs143960270	PROBABLY DAMAGING/Published		Predicted functional (high)	17754	1	1	0	0	0	0.006%
c.292A>C	p.N98H	rs142858704	BENIGN	BENIGN	Predicted non-functional (low)	17754	12	11	1	0	0	0.068%
c.340C>G	p.P114A	rs371363324	PROBABLY DAMAGING		Predicted functional (high)	17754	1	1	0	0	0	0.006%
c.441+1G>A	intronic	rs140130028	Strong Negative/Published			17752	96	83	0	6	7	0.541%
c.442-4A>C	intronic	rs114950106	Neutral			17754	104	104	0	0	0	0.586%

Research Publications

- Bianconi SE, **Cross JL**, Wassif CA, Porter FD (2015); “*Pathogenesis, Epidemiology, Diagnosis and Clinical Aspects of Smith-Lemli-Opitz Syndrome*”; Expert Opinion in Orphan Drugs
- Wassif CA*, **Cross JL*** et. al. (2015); “*High Incidence of unrecognised visceral/neurological late-onset Niemann-Pick disease, type C1, predicted by analysis of massively parallel sequencing data sets*”; Genetics in Medicine
- **Cross JL***, Iben J* et. al. (2015); “*Determination of the allelic frequency in Smith-Lemli-Opitz Syndrome by analysis of massively parallel sequencing data sets*”; Clinical Genetics