# Investigation of the Molecular Basis of Primary Ciliary Dyskinesia and Situs Inversus amongst the Amish

Volume 1 of 1

Submitted by **Oluwaseun Temiladeoluwa Rachael Olubodun** to the

University of Exeter

as a thesis for the degree of

**Masters by Research** in Medical Studies

In July **2018**

This thesis is available for Library use on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

I certify that all material in this thesis which is not my own work, has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

# Acknowledgements

I would like to extend my utmost thanks to all the families that were participants in these studies and to the wider Amish community for their ongoing support of the work and broader agenda of the Windows of Hope project.

I am grateful to my supervisors, Dr Emma Baple and Professor Andrew Crosby, who have made this masters project such a worthwhile and enjoyable learning experience.

I would also like to extend my heartfelt thanks to Dr Kate Ellacott who has encouraged and supported me throughout my endeavours this year – regardless of the nature of the challenges that arose.

None of the work produced in the following thesis would have been possible without the Windows of Hope project team; in particular, Professor Harold Cross and Roselyn Coblentz.

I would also like to thank all members of the Crosby Baple Neurogenetics group for their help and support; with special thanks to Dr Barry Chioza and Dr Gaurav Harlalka, without whom this thesis would not have come to fruition.

Crosby Baple Neurogenetics group members: Fatema Al-Salmi, Ilaria D'Atri, James Fasham, Stanya Gardner, Beth Hilditch, Hannah Jones, Simon Lam, Serene Lin, Lettie Rawlins, Olivia Rickman, Shamim Saleha and Claire Salter.

## Abstract

Cilia are hair-like projections from eukaryotic cells: they are complex organelles that can be split into motile (normal/embryonic nodal) and non-motile types. Primary ciliary dyskinesia (PCD) is genetically heterogeneous, typically autosomal recessive disorder, characterised by oto-sino-pulmonary disease, male infertility and situs abnormalities, due to structural and motion defects of cilia. Early diagnosis and specialist intervention significantly improves clinical outcome. Situs inversus (SI) is a related condition involving, the mirror image rearrangement of the internal organs and has a more complex genetic basis; it can also be a component of a more complex syndrome. Pathogenic variants in >40 genes have been associated with these conditions, accounting for <70% of cases.

This project is part of a long-running clinical-genetic program based amongst the Amish communities in the USA which aims to define the frequencies and genetic causes of PCD/SI. A combination of clinical phenotyping, autozygosity mapping, and whole exome sequencing, were used to investigate Amish families affected by these disorders. This enabled the identification of the novel disease genes in each case. This study determined that both PCD and SI occur more commonly amongst the Amish than in non-Amish families. Genomic studies of affected individuals identified known causes of PCD, as well as two putative novel causes of PCD/SI. Notably, previous studies of knockout mice for one of the new candidate genes identified SI as a key phenotype, overlapping with the human condition. Preliminary genetic studies in other families with PCD and situs abnormalities failed to identify clear genetic causes, so studies are ongoing to

investigate more complex pathogenesis. While additional work is required to define the outcome of the sequence variants identified, these studies provide a greatly improved understanding of the frequencies and causes of PCD/SI in the Amish, and define two putative new candidate genetic causes.

# Contents

# List of Tables

# List of figures

# Abbreviations

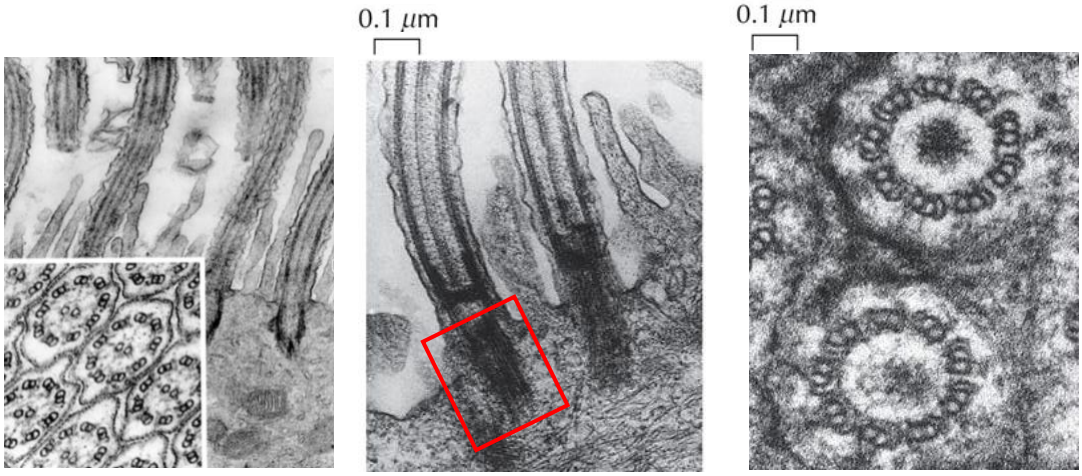
| | | | |
|---|---|---|---|
| 3’ | Three prime | m | Meter |
| 5’ | Five prime | M | Molar |
| APS | Ammonium persulphate | ml | Millilitre |
| AR | Autosomal recessive | $MgCl_2$ | Magnesium chloride |
| Arg | Arginine | MRI | Magnetic resonance imaging |
| ADP | Adenosine diphosphate | NaCl | Sodium chloride |
| ATP | Adenosine triphosphate | NaOH | Sodium hydroxide |
| BBS | Bardet-Biedl syndrome | NPHP | Nephronopthisis |
| Bp | Base pair | nNO | Nasal nitric oxide |
| °C | Degrees Celsius | PAGE | Polyacrylamide gel electrophoresis |
| cDNA | Coding deoxyribonucleic acid | PCR | Polymerase chain reaction |
| DNA | Deoxyribonucleic acid | PBS | Polyacrylamide buffered saline |
| dNTP | Deoxyribonucleotide triphosphate | RT-PCR | Reverse transcriptase polymerase chain reaction |
| Gln | Glycine | SAP | Shrimp alkaline phosphatase |
| Het | Heterozygous | SD | Standard deviation |
| kb | Kilobases | SDS | Sodium dodecyl sulfate |
| L | litre | SEM | Standard error of mean |
| μL | Microlitre | Ser | Serine |
| min | Minute | | |

| | |
|---|---|
| SNP | Single nucleotide polymorphism |
| TBE | Tris-Borate-EDTA |
| WHO | World Health Organisation |
| WOH | Window of Hope Amish project |

# 1 Introduction

## 1.1 Cilia

Cilia are complex evolutionarily conserved organelles that can be split into three types: normal motile, embryonic nodal motile and non-motile (primary or sensory cilia).[1] They are a distinct cellular compartment projecting from eukaryotic cells, structurally identical to eukaryotic flagella.[2] Distinctions between cilia and eukaryotic flagella are sometimes drawn due to differences in their function and length.[3] Cilia are hair-like attachments found on epithelial surfaces; each cell has around 200 cilia, anchored by basal bodies to the apical cytoplasm – extending from the cell surface into the extracellular surface.




**(a)** **(b)** **(c)**

***Figure 1-1 Apical section of ciliated epithelium shown under electron microscope[4,5]***

*(a) Bottom left corner show cilia in cross section, displaying peripheral pairs and central pairs of microtubules; the rest of the image shows a cross section of the cilia axoneme.*

*(b) Close-up image of basal bodies, where cilia growth is initiated and anchored (highlighted in red).*

*(c) Magnification showing cilia in cross-section – displaying peripheral pairs and central pairs of microtubules*

Cilia have a core axonemal cytoskeletal structure made up of approximately 250 proteins. A single cilium may be 1-10 micrometres long and less than a micrometre wide.[6] The axoneme is the central or inner strand of a cilium or flagellum, structured with longitudinal microtubules that make up the basic axonemal structure providing support, and facilitating bending and further motion of the cilia (***Figure 1-1***).[7]

## 1.2 Evolutionary conservation of cilia

This plain structure of cilia is highly conserved across species and facilitates many essential physiological functions.[2] The core set of proteins needed for cilia formation – intraflagellar transport proteins – are also highly conserved through evolution.[8] The same proteins are needed in *Chlamydomonas reinhardtii* as are needed in humans. Such conservation has been illustrated when studying the ciliopathy disorders: polycystic kidney disease, nephronopthisis (NPHP) and Bardet-Biedl syndrome (BBS). Such evolutionary conservation has allowed studies of ciliopathy disorders to benefit from the use of model organisms including *Chlamydomonas reinhardtii*, *Caenorrhabditis elegans*, zebrafish and mice.[8]

The cilia subtypes which include normal motile, embryonic nodal motile and non-motile cilia are determined by the arrangement of the microtubules of the axoneme and the presence, or absence, of inner and outer dynein arms.[2] The differences in ciliary structure and function are highlighted and further detailed in ***Figure 1-2***.

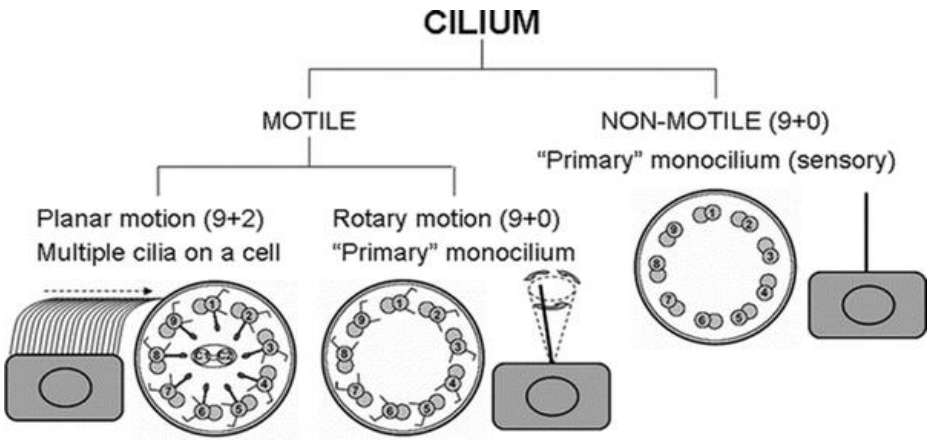



***Figure 1-2 Diversity of ciliary axoneme; adapted from[7]***

## 1.3 Ciliary subtypes

### 1.3.1 Normal motile cilium

Motile cilia have microtubules of **α**- and β-monomers of tubulin and associated accessory elements.[1] They have a 9+2 pattern, meaning that they have 9 peripheral pairs of microtubules, connected to each other by dynein arms (nexin linked protein complexes) –as below.

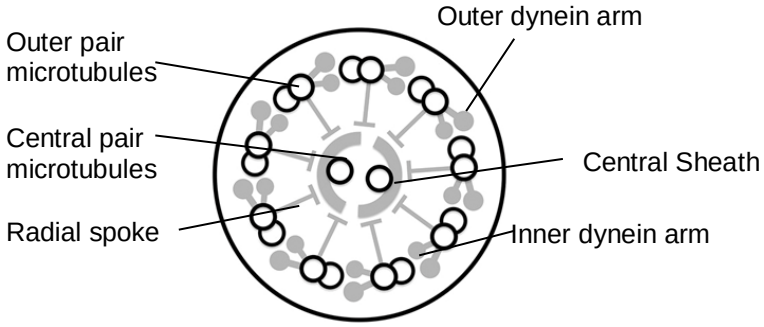



***Figure 1-3 Ultrastructure of normal motile cilium; adapted from[9,10]***

These microtubule doublets, or pairs, are then connected by radial spokes to the central pair of microtubules my radial spokes. The dynein arms contain ATP, whilst the microtubules within a pair can slide on one another to produce ciliary

motion. The extent to which cilia bend and microtubules slide are limited by the protein links formed between the peripheral doublets.

Peripheral microtubules – or doublets – contain outer and inner dynein arms that contain enzymes for ATP hydrolysis (defined in appendix 18.1). Doublets are connected within by nexin-dynein regulatory complexes and then joined by radial spokes to the central pair. This joining, to the central pair, provides structural support required during cilia bending.[1] This type of cilia can be found in large numbers on apical surfaces of upper and lower respiratory tracts, the brain ependyma (the lining the ventricles of the CNS), and oviducts of the female reproductive tract. Motile cilia and sperm flagella are often thought of as the same, except for subtle differences in dynein arm distribution along the axoneme.[2,11]

Normal motile cilia clear mucus, bacteria and toxic substances from conducting airways.[12-14] Mucociliary clearance is an important defence mechanism in the lungs, relying on the motive force of cilia to clear bacteria and toxic substances from airways. Cilia produce their motility by ATP hydrolysis in the dynein arms: this process produces a force that induces sliding of adjacent axonemal structures, therefore generating ciliary waveform in human airways. Cilia beat at around 200 beats per minute, coordinated by their planar orientation and cilia driven fluid flow, regulated by multiple signalling molecules, including cAMP, cGMP, and NO.[15] Cilia are coordinated in their action, bending in synchrony in wave-like movements at 16Hz.[12,15] Cilia's forward strokes are more rapid than

the recovery (or backward) stroke. The rapid beat extends into the mucus layer and is facilitated by the lack of friction between each cilia – this is because of their planar oridentation, negative charge of glycoproteins that coat them and the flow of overlying fluid.[15]

### 1.3.2 Embryonic nodal motile cilium

These cilia are also classed as motile, but create rotary motion as opposed to the flexural bending typically seen in epithelial cilia.[16,17] Each cilium has nine peripheral doublets and dynein arms but, unlike normal motile cilia, do not have a central pair of microtubules; they are therefore said to have a 9+0 pattern.[12] Their rotary motion facilitates the lateral movement of organs during embryogenesis by causing the leftward flow of extracellular fluid.[18] Embryonic nodal motile cilia are primary cilia in the ventral node. The ventral node is an enlarged group of cells and is the point at which nodal flow is generated. Nodal flow is the autonomous leftward movement of this fluid by the rotary motion of the cilia to create left-right asymmetry.[12,19] Abnormal function of this group of cilia can lead to problems with left-right body orientation leading to a range of clinical presentations, including dextrocardia, situs ambiguus or situs inversus totalis.[12]

### 1.3.3 Single non-motile (sensory or primary) cilium

These are the immotile cilia responsible for sensing changes in the environment. The have a 9+0 pattern, meaning that there are 9 peripheral pairs of microtubules, not joined by dynein arms and there are no inner or central pair of microtubules.[12] These non-motile cilia are important for homeostasis; they are

present in most cells in the body, sense changes to the local environment and have a role in planar cell polarity.[1]

### 1.3.4 Ciliary dysfunction

The different types of cilia play crucial roles during human development and are important for the normal function of several organs. It is perhaps not unexpected that abnormalities of ciliary function have wide ranging effects on many organ systems in the body including the brain, eye, kidney, liver, heart and skeleton.[6,20,21] Disorders of ciliary function can therefore present with a range of clinical symptoms, which include: intellectual disability, blindness, chronic oto-pulmonary infection, cystic kidneys, skeletal dysplasia and infertility. ***Figure 1-4*** illustrates some of the potential impacts of dysfunctional cilia, although this is by no means an exhaustive list.

## 1.4 Ciliopathies

### 1.4.1 Definition:

Ciliopathies are genetic conditions that arise because of ciliary dysfunction. Affected structures can include the cilia themselves and, also, the basal bodies that anchor cilia in place.[22,23] There are a whole spectrum of different disorders classified as ciliopathies; the table overleaf describes some of the different types of ciliopathy disorders (***Table 1-1***).

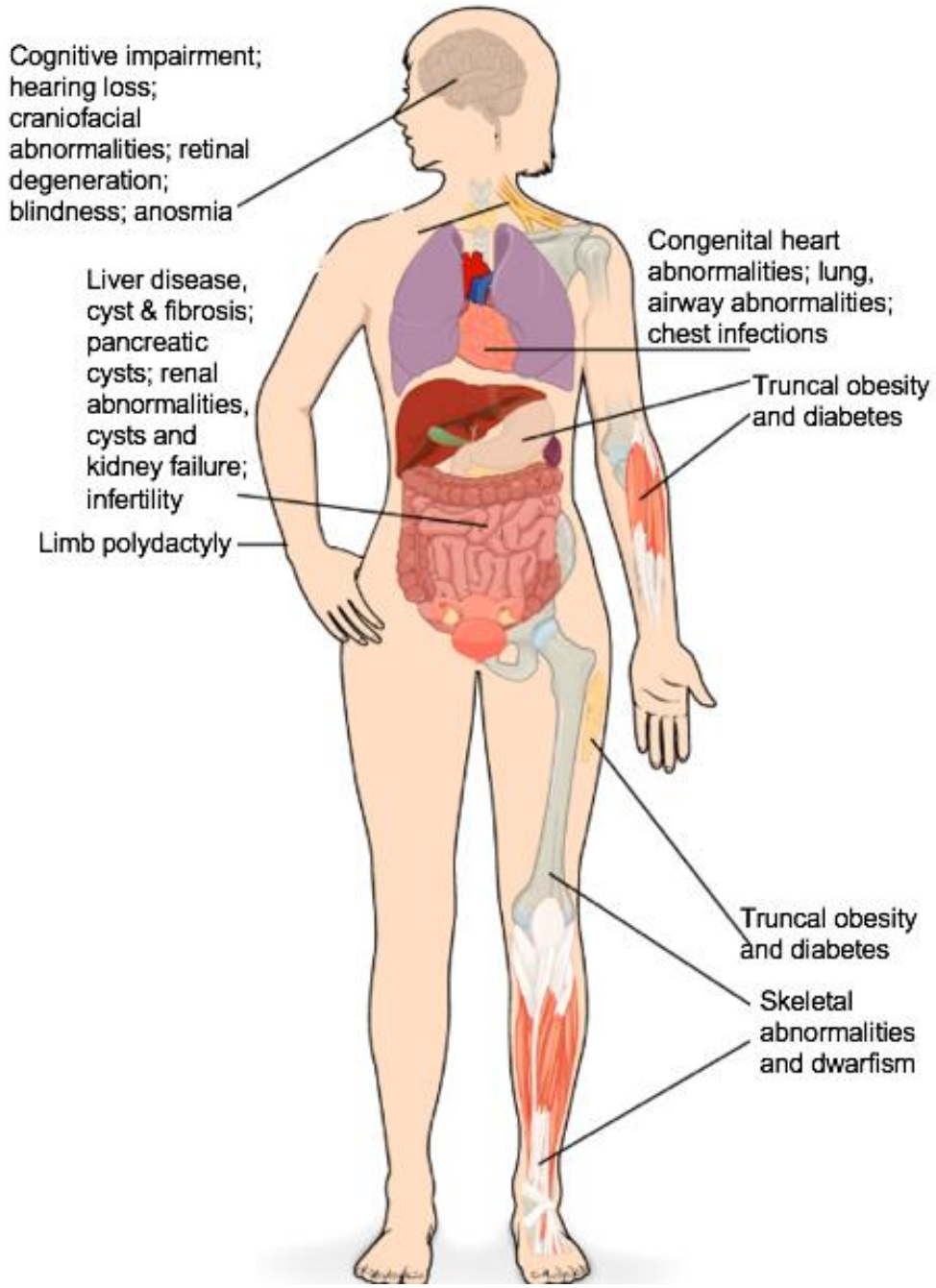



***Figure 1-4 Impacts of defective cilia. Adapted from [24,25]***

### 1.4.2 Ciliopathy summary

| Ciliopathy | Key Notes | Genes implicated |
|---|---|---|
| **Alström syndrome**[26-29] | • Inheritance: AR; prevalence: N/K<br>• Description: affects metabolism, with obesity, insulin resistance and type two diabetes mellitus, cardiomyopathy, renal and hepatic dysfunction, hyperlipidaemia, retinal degeneration, and sensorineural hearing loss. | *ALMS1* |
| **Bardet-Biedl syndrome**[30,31] | • Inheritance: AR; prevalence: 1 in every 100,000 births<br>• Description: kidney abnormalities/failure; global developmental delay, truncal obesity, poorly developed male genitalia, visual impairment due to red-cone dystrophy (confused with retinitis pigmentosa). | *BBS1, BBS2, ARL6/BBS3, BBS4, BBS5, BBS7, TTC8/BBS8, BBS10, TRIM32/BBS11 BBS12, CCDC28B, CEP290, TMEM67, MKS1, MKKS, BBS6* |
| **Ellis-van Creveld syndrome**[32,33] | • Inheritance: AR; prevalence: In the Lancaster Amish, 1:200; worldwide said to be ~1:6,000<br>• Description: ciliopathy syndrome, with short limb dwarfism with abnormal finger (appendix 7.2) development congenital heart defects and underdevelopment of male genitalia. | *EVC, EVC2* |
| **Asphyxiating thoracic dystrophy**[34] | • Inheritance: AR; prevalence: ~ 1 in 100,000 births<br>• Description: part of the short-rib polydactyly syndrome spectrum with narrowing of the thorax, subsequent poor lung development, presenting as respiratory distress in neonates (appendix 7.2), impaired growth of limbs. | *IFT80, DYNC2H1. CEP120, CSPP1, DYNC1H1, IFT140, IFT172, TTC21B, WDR19, WDR34, WDR35, WDR60.* |
| **Joubert syndrome**[8,12,35-38] | • Inheritance: variable from AR to X-linked chromosomal recessive. Prevalence N/K.<br>• Description: rare, incurable neurodevelopmental disorder: low muscle tone, developmental delay, breathing and movement difficulties; kidney disease (nephronophthisis and renal insufficiency) and retinal degeneration (appendix 7.2). | *CEP290, RPGRIP1L, TMEM67, INPP5E, TMEM216, AHII, ARLI3B, CC2D2A* |
| **Meckel-Gruber syndrome**[8,12,37-41] | • Inheritance: AR; prevalence worldwide: 1 in 13,250-14,000; in the Finnish and Belgian (1 in 9,000 and 3,000, respectively)<br>• Description: associated with short life expectancy/ incompatible with life. Occipital encephalocele, bilateral dysplastic and cystic kidneys, post axial polydactyly and liver fibrosis. +/-(neural tube defects, congenital heart disease, urinary tract and wider genitalia malformations). | *B9D1, B9D2, CC2D2A, CEP290. MKS1, RPGRIPL, TMEM67, TMEM216, AHII* |
| **Nephronophthisis**[8,12,37,42] | • Inheritance: AR; prevalence is 1-9 / 100 000 | *XPNPEP3* |

| | | |
|---|---|---|
| | • Description: disorder of the kidney; leading cause of non-hereditary end stage renal failure in children and young adults under the age of 30. Polyuria, polydipsia and anaemia; small kidneys, with sporadic cysts, cell infiltrates and interstitial fibrosis. | *NPHP1, INV, NPHP3, NPHP4, CEP290, GLIS2, RPGRIP1L, NEK8, TMEM67* |
| **Polycystic Kidney Disease**[43,44] | • Inheritance: AR or autosomal dominant inheritance; prevalence: ~ 1 in 500 – 1000; 12.5 million people affected worldwide<br>• Description: umbrella term for genetic disease causing kidney failure in adults and children. Fluid-filled cysts, enlarging, multiplying and bilaterally, leading to progressive renal failure. Poor kidney function and excretion of waste products +/- (liver, pancreas, spleen, brain and intestine dysfunction) | |
| **Retinitis Pigmentosa**[45,46] | • Inheritance: variable; prevalence: N/K<br>• Description: umbrella term for inherited disease of the retina manifesting as progressive vision loss, initial poor night vision and then loss of peripheral vision. Treatment is supportive; age at presentation varies; no difference in incidence between genders or races | Thought to be mutation in up to 200 genes, with several forms of retinitis pigmentosa recognised as ciliopathies. |
| **Senior-Løken syndrome**[8,12,37,38,47,48] | • Inheritance: AR; prevalence: 1 in a 1,000,000 people worldwide (very few recorded in medical literature)<br>• Decription: a combination of nephronophthisis and an eye condition – Leber congenital amaurosis. Retina: photophobia, nystagmus, hyperopia. Kidney: polyuria, polydipsia and anaemia; small kidneys, with sporadic cysts, cell infiltrates and interstitial fibrosis. Presents from first few years of life to late childhood. | *NPHP1, INV, NPHP3, NPHP4, IQCB1, CEP290, RPGRIP1L, NEK8, SDCCAG8* |
| **Usher syndrome**[46,49] | • Inheritance: AR; prevalence:<br>• Description affects hearing, vision and balance. Three subtypes;<br>Type I: balance problems, late sitting and walking milestones. Profound deafness; cochlear implants are used for speech and language development. Progressive vision problems present aged of 8-12 years old (poor night vision to complete blindness)<br>Type II: moderate/ severe hearing problems, good balance; older age of onset ($4^{th}$-$5^{th}$ decade of life)<br>Type III: normal hearing and vision progressing to profound hearing and vision loss, variably. Some will have balance problems but difficult to predict how severely an individual will be affected. | *Type I: MY07A, USH1C, CDH23, CPHD15, SANS*<br><br>*Type II: USH2A, VLGR1, WHRN*<br><br>*Type III: USH3A* |

***Table 1-1 Summary of different types of ciliopathy disorders***

**Table key:** AR – autosomal recessive; N/K – not known

## 1.5 Primary ciliary dyskinesia (PCD)

### 1.5.1 Definition

PCD is a genetically heterogeneous monogenic autosomal-recessive disorder of cilia structure and function and is considered to be one of the ciliopathy disorders.[1,12] PCD is also known as immotile cilia syndrome and Kartagener's syndrome.[12] PCD is characterised by three hallmark symptoms: oto-sino-pulmonary manifestations, organ laterality defects and infertility. The term 'primary' in the context of PCD makes the distinction from acquired ciliary defects which can occur as a result of environmental factors such as infection or inflammation. Interestingly, an estimated 30% of patients with PCD who are thought to have a normal ciliary structure.[1]

### 1.5.2 Epidemiology and clinical presentation of PCD

PCD has an estimated incidence of 1 per 10,000-20,000 – based on population surveys in Norway and Japan – but prevalence is difficult to ascertain due to inconsistent and sub-optimal diagnostic approaches.[1,50,51] The clinical features of PCD which are described above, are result from defective motile cilia in the conducting airways, paranasal sinuses, middle ear and reproductive tract. It is the defective embryonic nodal motile cilium during critical stages in embryogenesis that result in the situs abnormalities which are seen in 50-60% of PCD cases.[1]

The age at presentation and range of severity of PCD can be quite variable. Affected individuals most often present with nasal congestion and wet cough soon after birth, recurrent chronic middle ear and sinus infections are commonplace by 6-8 years of age.[1] Bronchiectasis is present in 50-70% of older paediatric patients and by adulthood is almost universal. In adulthood, individuals with PCD are almost universally are affected by recurrent and chronic bacterial infections in the lower airways.[1] It is important to be aware that be milder or atypical presentations of PCD do occur, with some patients suffering fewer recurrent infections or not exhibiting symptoms until their teenage years or adulthood.[52,53]

#### *1.5.2.1 Oto-sino-pulmonary manifestations*

**Pulmonary manifestations:** The presence of a wet cough all year round is one of the characteristic clinical signs associated with PCD in infants and children. The frequency of neonatal respiratory distress highlights the importance of normal ciliary function in clearing the foetal lung.[12] PCD patients suffer a vicious cycle of infection and inflammation that results in impaired airway clearance, airway damage and chronic colonisation and infection with a variety of microorganisms. This colonisation leads to further infection and inflammation and destruction of conducting airways and alveolar surfaces. Impaired mucociliary clearance can lead to recurrent bronchitis, pneumonia and eventually bronchiectasis in the middle and lower lobes. Bronchiectasis is the clinical consequence of chronic inflammation and destruction of bronchial walls,

particularly affecting medium sized airways in segmental or sub-segmental bronchi.[12] This bronchiectasis leads to ~25% patients developing respiratory failure in adulthood (***Figure 1-5***).[54,55]

**Oto-sinus symptoms:** Rhinitis, nasal congestion, recurrent sinus infections and Nasal polyps are all features of PCD. Chronic otitis media and subsequent conductive hearing loss – permanent or transient  are also common features of this condition – and re thought to be due to dysfunctional cilia along the Eustachian tube.[1,56] Interestingly, sinusitis becomes more of a problem as affected individuals age because sinuses develop as children grow.[57]

### *1.5.2.2 Organ laterality defects*

As described above, during embryogenesis, an individual's thoraco-abdominal orientation is determined by unidirectional rotation of the nodal cilia. When the cilia function is abnormal, organ orientation is random.[12] Subsequent organ laterality defects, or situs abnormalities, are said to be present in between 50 and 60% of adults found to have PCD.[1] PCD is associated with left-right asymmetric organ deflection, most commonly situs inversus totalis – mirror-image reversal of the organs in the chest and abdominal cavity.[58,59] This has implications for studies of both situs inversus and ciliary dyskinesia when studied in combination and isolation. The genetic underpinning of organ laterality defects is thought to be a complex causative relationship, as further discussed later in this thesis, in relation to situs inversus (1.6).

#### *1.5.2.3 Infertility*

The flagella that facilitate movement of the sperm are dysfunctional in males with PCD. This leads to infertility in around 98-99% of affected males.[1,12] Infertility and subfertility in females affected by PCD is less common and reported less clearly.[60,61] Fertility problems in women with PCD are thought to be due to abnormal cilia in the fallopian tubes, leading to impaired movement of ova and an increasing the risk of ectopic pregnancy.[62]

### 1.5.3 Diagnosis of PCD

There is no standardised protocol for diagnosis and subsequent treatment of PCD. Establishing a definitive diagnosis can be challenging because of the variable phenotype seen in affected patients. Patient clinical history often yields symptomatic clues suggestive of PCS and the diagnosis of PCD is often centred on the observation of these signs and symptoms. Respiratory distress in the form of tachypnea and increased work of breathing is reported in around 80% of full term neonates with PCD, sometimes with the need for supplemental oxygen for a period varying from a few hours, to weeks. This may lead to a diagnosis of transient tachypnea of neonate or neonatal pneumonia. Medical imaging, respiratory microbiology and pulmonary physiology may point to the early onset of lung disease and to abnormal airflow mechanics.[1]

Radiographic abnormalities are often present but frequently remain undetected, owing to our disinclination to expose children to radiation. This means diagnoses of sinusitis and bronchiectasis can be missed in infants and pre-schoolers without CT scanning for significant findings (atelectasis, peri-hilar, or interstitial, markings, air trapping, ground glass opacities, fibrosis of airway walls and mucus plugging).[1] 50-70% of paediatric patients with PCD will have bronchiectasis on imaging, which is almost universally noted in affected adults (***Figure 1-5***).[1]

Diagnosis can be made through cilia ultrastructure examination and analysis under electron microscope and digital video technology. This was formally the gold standard, identifying abnormalities like short or absent dynein arms, and dyskinetic ciliary motion or reduced ciliary beat frequency, can be identified and may support a diagnosis.[63] However, the method is not standardised or readily available and is dependent on technical and operator factors.[12] Furthermore, there are 30% individuals with a PCD phenotype have no apparently dysfunctional cilia. Apparently normal cilia function would, ordinarily, exclude PCD preclude diagnosis so this technique cannot be used to make a diagnosis of PCD in isolation.[1,12] As PCD is often associated with situs inversus, one strong indicator of the disorder would be detxrocardia evident on imaging (**Figure 1-7**).

There are a number of different approaches to diagnosis PCD. Diagnosis can be made through cilia ultrastructure examination and analysis under electron microscope and digital video technology. This was formally the gold standard,

identifying abnormalities like short or absent dynein arms, and dyskinetic ciliary motion or reduced ciliary beat frequency, can be identified and may support a diagnosis.[63] However, the method is not standardised or readily available and is dependent on technical and operator factors.[12] Furthermore, there are 30% individuals with a PCD phenotype have no apparently dysfunctional cilia. Apparently normal cilia function would, ordinarily, exclude PCD preclude diagnosis so this technique cannot be used to make a diagnosis of PCD in isolation.[1,12]

Nasal nitric oxide (nNO) is usually low in individuals affected by PCD and its measurement makes up part of a complex diagnostic work up.[64,65] The test measure the amount of nitric oxide in one's nostril, whilst they hold their breath. Such a reading can be a strong preliminary indicator of PCD and is a non-invasive method of gathering supporting evidence.[66,67]

Establishing PCD diagnoses has been helped by the discovery of mutations in multiple PCD associated genes: this project aims to uncover other genes associated with PCD in this way.[37] Comprehensive genetic testing can provide an accurate molecular diagnosis for individuals at a younger age, improving their long-term outlook. Within the community setting genetic testing for specific founder variants is often the most cost effective and efficient way to reach a diagnosis (***Table 1-6***).[68] Approximately two-thirds of those that undergo genetic testing will receive a definitive diagnosis. Importantly, a definitive molecular

diagnosis facilitates informed genetic counselling to be offered to related family members.[69]

Alternatively gene panelling may be used, to include all genes known to be associated with PCD, as well as other genes of interest: the genes included in such panels can vary, and with it the multi-gene panel's sensitivity.[68]

Some diagnostic centres in Europe are employing the use of immunofluorescence analysis of ciliary proteins, using antibodies directed against the main axonemal components to pinpoint structural and functional defects.[12] The hope is to develop screening panels of antibodies to enable enabling screening of cilia samples.

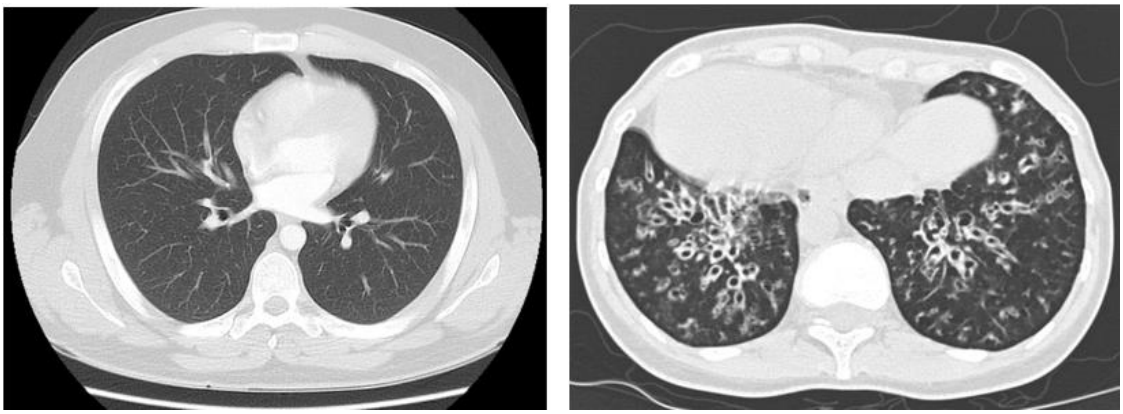


***Figure 1-5 (left to right) apparently normal chest CT scan versus a CT scan of a patient with PCD[70,71]***

### 1.5.4 Genetics of PCD

#### *1.5.4.1 Known genetic causes of PCD*

Biallelic mutations in 38 genes have been associated with PCD to date, accounting for around 70% of PCD cases.[72,73] A very small minority of cases have been reported to show X-linked recessive inheritance and a very few sporadic cases have been described, resulting from *de novo* mutations in the cytoplasmic axonemal dynein assembly factor PIH1D3, for example[74,75].

For the most part, PCD results from pathogenic sequence alterations in genes that encode proteins that make up the core axoneme structure of motile cilia – a component of the motile cilia structure, said to be essential for movement.[2] Other genes associated with PCD encode proteins integral to the ciliary dynein motor structures and other proteins essential for the active transportation of ciliary dynein motors from where they are assembled in the cytoplasm into the axoneme.[2] Dynein motors allow a single dynein molecule to transport its cargo by travelling along the distance of a microtubule, without becoming detached: it is a collection of complex proteins made up of smaller polypeptide units.[76]

**Table 1-2** describes the recognised genetic causes of PCD. The genes in which Amish founder mutations have been identified are highlighted in bold. **Figure 1-6** illustrates the probable localisation of the above listed genes' products within the human ciliary structure components.

| Gene | Chromosome | Component of axonemal ultra-structure affected |
|---|---|---|
| ***DNAI1***[12,77-81] | 3p21 | ODA & IC |
| *DNAI2*[12,78,83,84] | 17q25 | ODA & IC |
| ***DNAH5***[12,81,83,85,86] | 5q15 | ODA & HC |
| *DNAH11*[12,82,87-89] | 7q15 | ODA & HC |
| *NME8 (TXNDC3)*[12,86,87] | 7q14 | ODA, IC +/- LC |
| *KTU/PF13 (DNAAF2)*[12,82,88] | 14q21 | cytoplasmic protein (dynein arms assembly) |
| *LRRC50 or DNAAF1*[12,89-91] | 16q24 | cytoplasmic protein (dynein arms assembly) |
| *CCDC39*[12,82,92] | 3q26 | ciliary axonemal disorganisation; NL/DRC factor |
| *CCDC40*[12,82,93] | 17q25 | ciliary axonemal disorganisation; NL/DRC factor |
| RSPH4A[12,82,94] | | radial spokes; transposition defect |
| *RSPH9*[12,82,86,94] | | radial spokes; central pair defect |

| Gene | Chromosome | Component of axonemal ultra-structure affected |
|---|---|---|
| *RSPH3*[82] | 6p25 | Central pair or radial spokes |
| *ZMYND10*[82] | 3p21 | cytoplasmic protein (dynein arms assembly) +/- transport factor |
| *C21orf59*[82] | 21q22 | cytoplasmic protein (dynein arms assembly) +/- transport factor |
| *ODA8, ODA5, ODA10* | | |
| TTC25[82] | | ODA targeting/docking factor |
| *SPAG1*[82,86] | 8q22 | |
| CCDC164[82] | | NL subunit |
| *CCDC65*[82,86] | 12q13 | NL subunit |
| *HYDIN*[82,86] | 16q22 | CP subunit |
| *CCNO*[82,86] | 5q11 | CCNO: cytoplasmic centriole assembly and docking factor |
| *MCIDAS*[82] | | MCIDAS: nuclear regulator of CCNO and FOXJ1 |

| | | | | | |
|---|---|---|---|---|---|
| ***HEATR2 (DNAAF5)***[82,95,96] | 7p22 | cytoplasmic protein (dynein arms assembly) | *OFD1*[82] | | |
| *CFAP300*[2] | | ODA & IDA | *RPGR*[82] | | |
| *DNAAF3*[82,86,97] | 19q3 | cytoplasmic protein (dynein arms assembly) | *DNAL1*[86] | 14q24 | |
| *DNAAF4 (DYX1C1)*[86,98,99] | 15q21 | cytoplasmic protein (dynein arms assembly) | CCDC103[100,101]* | 17q21 | ODA & IDA |
| *LRRC6*[82,102-104] | 8q24 | ODA & IDA | *C11orf70* | | outer and inner dynein arms |
| *PIH1D3*[75,105] | Xq22 | ODA +/- IDA | *RSPH1*[82,86] | 21q22 | Central pair or radial spokes |
| *CCDC151*[86,106,107] | 19p13 | | *DRC1*[86] | 2p23 | |
| *ARMC4*[86,107-109] | 10p | cytoplasmic protein (outer dynein arms assembly) | *causative PCD loci*[38,86] | 15q13-q15 | |
| ***CCDC114***[82] | 19q13 | ODA targeting/ docking factor | *causative PCD loci*[38,86] | 15q24-q25 | |

***Table 1-2 Genes implicated in PCD (as of July 2018),***

*The genes identified in Amish PCD cases are highlighted in **bold**; ODA – outer dynein arm; IDA – inner dynein arm; IC – intermediate chain; LC – light chain; HC – heavy chain; NL – nexin link; DRC - dynein regulatory complex*

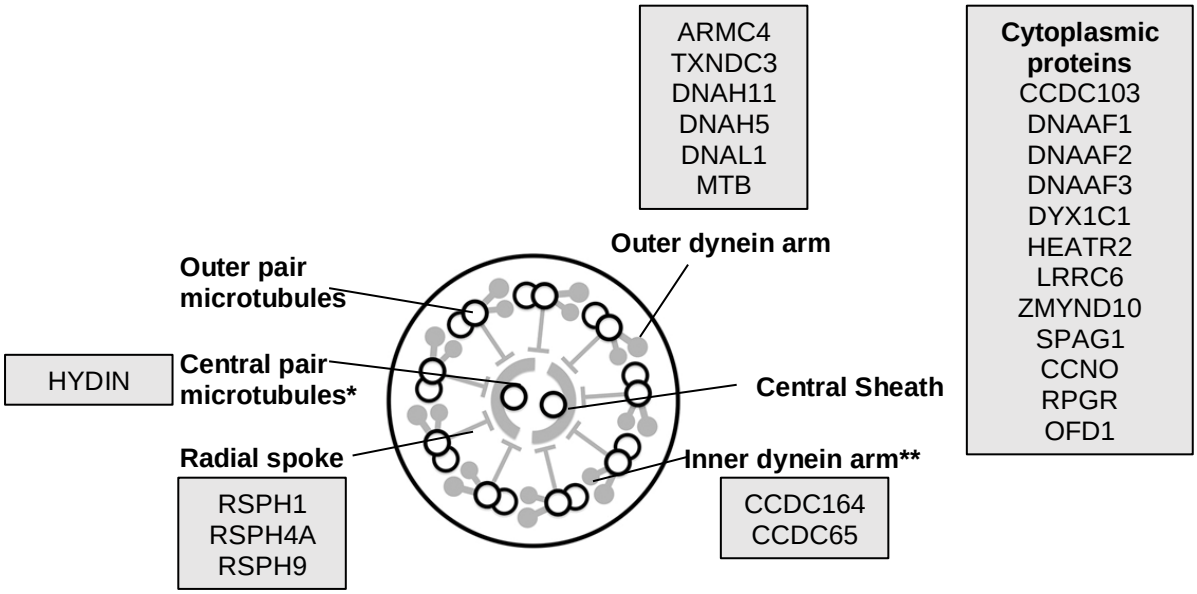



***Figure 1-6 Ultrastructure of normal motile cilium. Adapted from[9,10,110-112]***

*Illustration of the probable localisation or action of gene mutations (listed in* **Table 1-2***) within the human ciliary structure components.*

** 'central pair microtubules' also encompasses their projections and interconnections*

*** 'inner dynein arm' here encompasses the nexin–dynein regulatory complex[111,112]*

Mutations affecting different parts of the axoneme have different effects. Mutations coding for the components of the central apparatus affect the central pair of microtubules and the radial spokes and tend not to cause laterality defects.[1] Mutations

coding for the proteins that make up the dynein arms and the determine their assembly tend to have a more profound effect on cilia's structure and therefore function.[113]

### *1.5.4.2 Genetic modifiers of PCD*

Different degrees of PCD severity are seen depending on the underlying genetic cause of the disorder. Mutation of proteins that make up the components of the ciliary central apparatus, (the central pair of microtubules and the radial spokes), tend not to cause laterality defects. Mutations in the proteins that make up the dynein arms and the determine their assembly tend to have a more profound effect on cilia structure and function. [113,114]

Interestingly, phenotype also varies between affected family members with the same genotype. Several explanations have been proposed for this, including the possibility of genetic modifiers. These are genetic sequence variations that on their own are insufficient to cause disease but when present in conjunction with disease associated variants, may modify the phenotype so that a patient presents with more or less severe clinical signs and symptoms of PCD.[114] Mouse model studies aiming to identify genetic modifiers of PCD are ongoing.[115]

The phenotypic variability observed amongst PCD patients, highlights the importance of increasing our understanding of the mechanisms that underlie the diverse spectrum of PCD phenotypes. This in turn, may identify potential therapeutic targets which would have wider implications for the prevention and treatment of ciliopathy disorders in general.

#### *1.5.4.3 Additional features reported in PCD*

PCD has been reported in conjunction with X-linked retinitis pigmentosa and sensory hearing deficits resulting from mutations in the RPGR gene. Proteins encoded by this gene are required for photoreceptor maintenance and viability.[116,117] PCD has also been associated with a novel syndrome, reported in a single family and due to X-linked mutation in the oral-facial-digital type 1 gene (OFD1 gene).[117,118] The syndrome is characterised by X-linked recessive inheritance, learning difficulties, macrocephaly and PCD.[117]

#### *1.5.4.4 Management of PCD*

There is no standardised protocol for management: the mainstay of treatment is physiotherapy and early intervention with antibiotics creating a targeted approach to prioritise airway clearance, surveillance of pulmonary function and monitoring of respiratory microbiology. Further detail regarding their treatment protocol can be found in appendix 7.3. Post diagnoses, lung function tests are used to obtain a baseline from which to monitor an individual, over a period of time, and personalise their care.[12]

Air flow mechanics can be monitored using spirometry: in PCD, abnormal air flow mechanics are noted in infants and young children, some cases progressing so that, by the ages of 6-8, the forced expiratory volume in 1 second is significantly reduced (FEV1 – further defined in Appendix 18.3). Disease course is highly heterogeneous

but it is said that, with aggressive treatment, lung function can be stabilised, preventing decline.[119] Whilst diagnosis can be made with CT and or spirometry, they are more useful for monitoring as they are said to be more effective in assessing the extent and severity of lung abnormalities in early PCD disease.[1] CT, in particular can be used to map the patterns of distribution of lung disease, in support of a PCD diagnosis as bronchiectasis having a predilection for the middle and lower lobes.[12]

Respiratory microbiology can be used to tailor treatment to individuals with appropriate antibiotics, if necessary. Patients affected by PCD typically culture Haemophilus influenza, Staphylococcus aureus and Streptococcus pneumonia. Children affected by PCD can intermittently culture Pseudomonas aeruginosa, which can become a chronic airway infection in teenage and adulthood.[1]

## 1.6 Situs inversus totalis (SI)

### 1.6.1 Definition

Situs inversus totalis (commonly referred to as situs inversus) is characterised by abnormal position of the internal organs in the chest and abdominal cavity.[58,120-122] There is malpositioning of the heart on the right side of the body, known as dextrocardia and organs have a mirror image reversal. This can be observed in chest x-rays and computed tomography scans (**Figure 1-7**).[99,121]

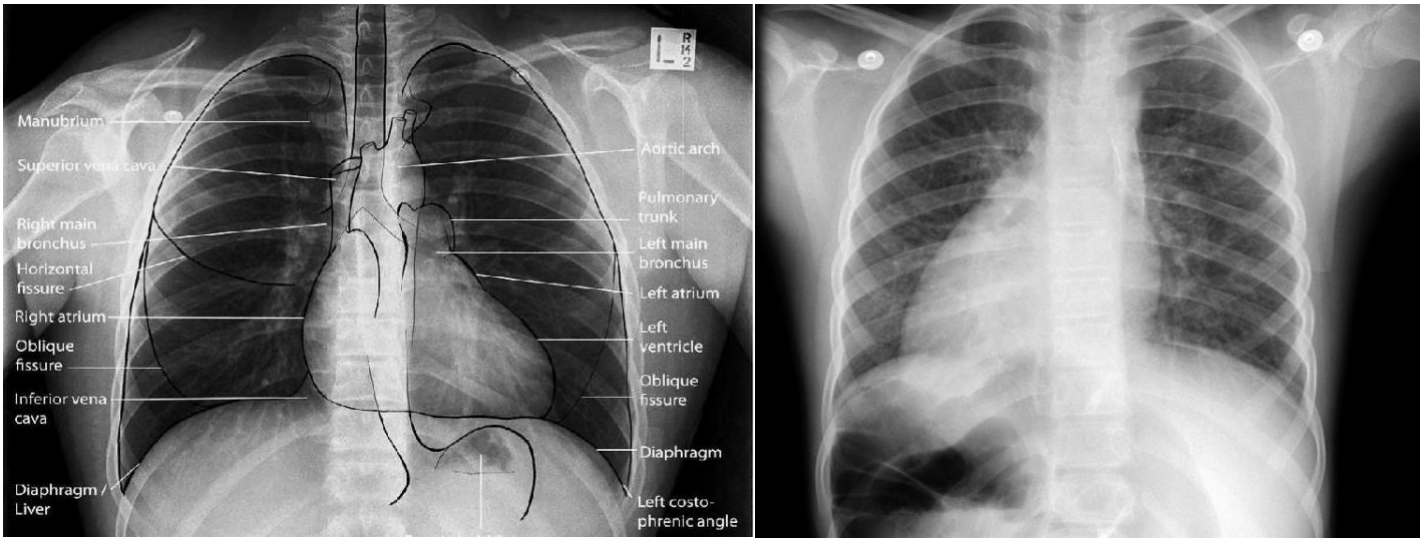



***Figure 1-7 Anterior-posterior chest x-rays showing situs solitus (left) and situs inversus (right), adapted from[99,121]*** Notable in the situs inversus x-ray demonstrating dextrocardia with the gastric bubble in the stomach underneath, whilst the opacity on the left hand side is the liver

### 1.6.2 Epidemiology and presentation of situs inversus totalis

In radiographs taken of patients with SI, the heart can be seen on the right side, as opposed to the usual positioning on the left. Rarely, situs inversus can coexist with congenital heart defects, the most common being transposition of greater vessels. Whilst up to 60% of individuals diagnosed with PCD have situs inversus totalis, a further 6% may have situs ambiguus, or heterotaxy – simply defined as incomplete situs inversus totalis.[58] Organs may be duplicated or absent, presented as an incomplete mirror image reversal of the organs in the thoracic and abdominal cavities; congenital heart defects are often also seen in these patients.[12]

### 1.6.3 Diagnosis of SI

Diagnosis of SI is usually made following through examination and radiographic imaging of the cardiothoracic and abdominal cavities.[123] However, diagnosis of SI is often incidental, the condition coming to light when the patient is receiving medical attention for another medical condition. Appendicitis often reveals an appendix on the left side of the abdomen when it should be on the right.[124] SI can also come to light during a cardiological examination or when performing a chest x-ray or CT scan in patients for another reason. Individuals with isolated SI are entirely healthy, so there are no risks related to delayed diagnosis.[123,125]

### 1.6.4 Genetics of SI

The genetics of SI is complex, several familial cases have been reported, in which inheritance has been described as either autosomal recessive (most commonly), autosomal dominant, or X-linked.[126] Where situs inversus is associated with another underlying syndrome or condition, the inheritance pattern will be that of the underlying condition.

SI appears to be genetically heterogeneous, meaning that different genetic factors or genes may cause the condition among different people or families. There are over 60 genes implicated in determining organ laterality.[123,127] A summary of the genes more commonly implicated in SI can be found in ***Table 1-3.***[123]

| Gene | Phenotype | Gene | Phenotype |
|---|---|---|---|
| *CFAP53*[127,128] | D, SI, SA | *ANKS3*[127,128] | dextrocardia; situs inversus totalis |
| *NODAL*[127,128] | D, SI | *LEFTY A* [129] | heterotaxy |
| *HES7*[127,130] | D, SI | *ACVR2B*[129] | heterotaxy |
| *NME7*[127,128] | D, SI | *CFC1*[129] | heterotaxy |
| *PKD1L1*[127,128,131] | D, SI, SA | *INVERSINE*[129] | heterotaxy |
| *MMP21*[127,128,131] | D, SI, SA | *ZIC3*[129] | heterotaxy |
| *CFAP52*[127,128] | D, SI | | |

***Table 1-3 Genes previously associated with situs abnormalities,*** *D: dextrocardia, SI situs inversus, SA: situs ambiguus; respective references listed alongside gene*

### *1.6.4.1 Genetic modifiers*

| Gene | Phenotype |
|---|---|
| CITED2[127] | D, SI |
| ACVR2B[131] | SA |
| LEFTY2[131] | SA |
| ZIC3[131] | SA |
| | |
| **Genes undergoing testing** | **Phenotype** |
| CFC1[131] | SA |
| *NODAL*[127,128,131] | SA |

***Table 1-4 Genes implicated in modifying susceptibility to situs abnormalities****, D: dextrocardia, SI situs inversus, SA: situs ambiguus; respective references listed alongside gene*

### *1.6.4.2 Additional clinical features associated with SI*

As mentioned previously, there is a recognised association between PCD and SI, with 25% of patients with SI also having PCD.[132] There is a 5-10% prevalence of congenital heart disease in individuals with situs inversus totalis, usually transposition of the greater vessels.[132,133]

### 1.6.5 Management of SI

No treatment is required for isolated SI, the issues tend to arise when treating an individual for something unrelated.[124] For example, most donor organs will come from individuals with situs solitus; therefore, those requiring transplantation of an asymmetrical organ, will have the added challenges of trying to harvest the organ from the donor cavity and overcome the surgical challenges that come with an ill-fitting recipient cavity and adjoining blood vessels.[124]

Treatment is not required as those with SI are otherwise healthy; if associated with another condition, then an individual may need treatment for the associated condition's signs and symptoms.[123] Diagnosis influences management of an affected individuals rarely, tending to be emergency situations and surgical complications caused by failure to recognised reversed anatomy. [123]

## 1.7 The Amish and Windows of Hope project

### 1.7.1 History and culture

The Amish are plain, rural-living traditional Christians, of Swiss-German descent that moved to the United States (mainly Pennsylvania, Ohio and Indiana) from as early as 1693, through into the 1800s, fleeing persecution, in amidst Europe's 16th Century Protestant Radical Reformation.[134-137] The "Amish" – previously called the "Amman-ish" – are named after the Anabaptist leader Bishop Jacob Ammann (a branch of Anabaptists) and are encouraged to abide by strict church discipline, employing the

use of “Meidung” or “Bann”, shunning or excommunication respectively. Members of the Amish church live their lives according to the New Testament, abiding by principles of non-resistance, tolerance and peace. This standpoint was considered a capital offence by the Catholic church in the 1500s, leading to the persecution of many.[134-137]

The Amish accept basic Christian beliefs, with specific emphases and interpretations that have evolved over time. Professor Donald Kraybill, widely considered to be the foremost expert on the Amish, defines the Amish as: “any group that affirms the basic tenants of Amish belief, including adult baptism, non-violence and separation from the outside world”.[135,137] Optional additions to this definition include their chosen mode of transport (horse and carriage), language (German-derived dialect) and their self-identification as Amish. The Amish employ the use of three languages in different cultural contexts: Pennsylvania Dutch – an evolution of a Swiss-German dialect that incorporates English – is used daily; high German is used in church for bible readings and hymns and English is taught at school to facilitate communication with the outside world.[137]

Amish groups may differ on religious and cultural grounds and or geographical location and origin, with affiliations graded ‘low’ to ‘high’ order, determined by the extent of their separation from modern society.[134]

### 1.7.2 Migration and population size

Fleeing religious persecution and compulsory military service, migration occurred in two waves. Around 3000 individuals are thought to have settled the original settlements in Pennsylvania (in the 1700s and 1800s), and Holmes County, Ohio (in the early 1800s), with subsequent migrations throughout the USA (***Figure 1-8; appendix 7.4*** ). Originally, these individuals came from families of similar descent with common ancestors so that the gene pool of the modern-day Amish is particularly modest, with each region displaying its own genetic architecture. The Amish are one of the fastest growing populations worldwide, with their population reportedly doubling every 15 to 20 years.[135,138-140] As of June 2017, there were Old Order communities in 31 US states.[141,142]

| State | Amish Population in | | | | Amish population increase 1992-2017 (%) |
|---|---|---|---|---|---|
| | **1992** | **2000** | **2010** | **2017** | |
| Pennsylvania | 35200 | 40100 | 59350 | 59350 | 169 |
| Ohio | 43200 | 49750 | 5859 | 58590 | 136 |
| Indiana | 25200 | 32650 | 43710 | 43710 | 173 |
| Wisconsin | 7800 | 10250 | 15360 | 15360 | 197 |
| New York | 4700 | 5000 | 12015 | 12015 | 256 |
| Michigan | 6500 | 9300 | 11350 | 11350 | 175 |
| Missouri | 5200 | 6100 | 9475 | 9475 | 182 |
| Kentucky | 1500 | 5150 | 7750 | 7750 | 517 |

***Table 1-5 Amish population change in USA's 8 most populated states***[135,138-140]


***Figure 1-8 Map of Canada and USA, denoting the region in which Amish families reside***[135,136,139-146] IN – Indiana; OH – Ohio; ON – Ontario; MO – Missouri; MS – Minnesota PA – Pennsylvania

### 1.7.3 Lifestyle

The Amish live in rural communities and Amish children tend to be educated separately from the general population, and only up to the age of 14.[141,143] Women tend to be home makers, whilst men largely work in farming or in factories, carpentry, construction and other smaller industries. Their use and interaction with technology is

limited and their preferred mode of travel is by horse and buggy, only using motorised transport for longer journeys and medical emergencies.[135,147]

***Figure 1-9 Amish horse and buggy transportation[148]***

With regards to healthcare, the Amish do not, in general, engage with health and social care insurance provisions, such as Medicare. The US federal government, via Medicare, ensures health insurance for vulnerable groups (65+ years old, disability, end stage renal disease or low income).[137,149] There are some Amish communities that have negotiated discounted insurance rates and cash payments with local healthcare; furthermore, families unable to pay may be aided by their church district.[141,146]

### 1.7.4 The Windows of Hope project

Professor Andrew Crosby and Professor Harold Cross established the Amish Windows of Hope (WoH) project in 2000. The project is now led by Prof Crosby and Dr Baple and is based at University of Exeter. The WoH project is a translational

research study, with the primary aim being to support local clinical teams who provide care for Amish families to diagnose and treat inherited conditions present amongst the Amish.

The study has its roots in the pioneering 1960s studies of McKusick and Cross, who identified, characterised and published several neurological disorders present amongst the Holmes County Amish community.[150-155] In 2000, the WoH team began genomic studies to define the genetic and molecular basis of the conditions described by McKusick and Cross.

In the last 18 years, WOH has been responsible for the identification and delineation of 18 novel inherited diseases in the Amish community. Project discoveries directly benefit the communities the project serves, delivering educational, diagnostic and therapeutic benefits. Additionally, almost all the conditions originally identified in the Amish have subsequently been documented in other families worldwide, highlighting the importance and impact of founder population studies.

### 1.7.5 The founder effect

The Amish's migration wave represents a 'bottleneck' event, reducing the genetic diversity of the group (**Figure 1-10**).[156] With subsequent population expansion, certain disease-associated alleles have accumulated to high prevalence in the community. This has led to a high incidence of certain (and particularly autosomal recessive) inherited disorders in the Amish, although other conditions which may be common

elsewhere may occur at low frequency or be absent in the community. Each Amish community displays its own genetic distinctiveness, with specific genetic disorders being present at a higher frequency in each group.

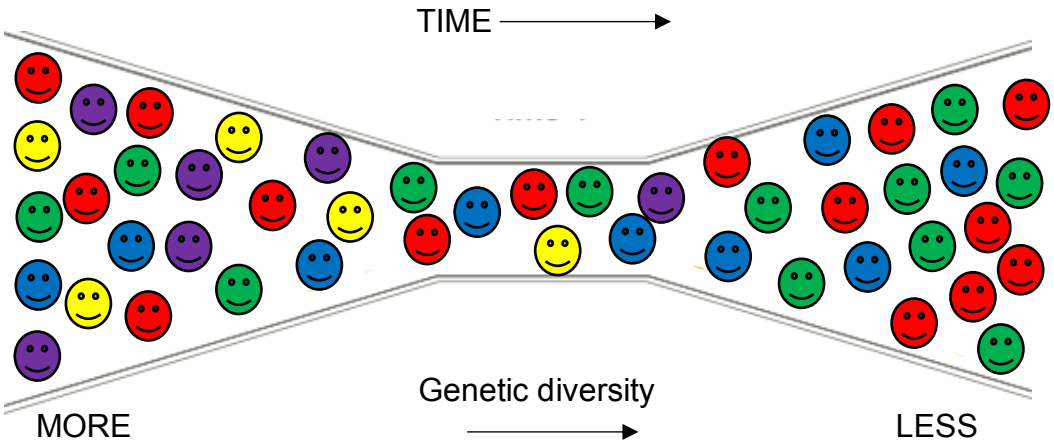



***Figure 1-10 Illustration of a bottleneck event, giving rise to the founder effect. Adapted from[157,158]***

A significant contribution to our knowledge and understanding of human development and disease processes have been made by the study of single gene disorders which are rare in the general population, but which occur with increased frequency in certain consanguineous populations or populations in which there is a high frequency of inter-community marriage.[159,160] When the ancestry of a community comprises only a limited number of founder individuals, the frequency of autosomal recessive alleles that were present in the founders may increase within that population such that if any two individuals from that community are selected at random, the chance that both will be carriers will be high. This is known as the founder effect.[156] The disease causing

gene variant is often surrounded by a common or shared region of the genome, otherwise known as a haplotype.[161] Thus, when looking to identify a causative gene variant responsible for an autosomal recessive condition, one must look to first identify common regions of homozygosity shared by multiple affected siblings or individuals.[12] These regions of homozygosity can then be cross-referenced with exome and genome sequencing data to identify potential disease causing variants (**Figure 1-11**). In so doing, a large number of rare autosomal recessive conditions[a] have been identified and described within the Amish.[162,163] These findings have increased scientific understanding of molecular pathways and disease pathology.

[a] ***Autosomal recessive conditions discovered amongst the Amish:*** *Ellis-Van Creveld syndrome, Primary Ciliary Dyskinesia, Alstroem syndrome, Dilated cardiomyopathy with ataxia syndrome, Renpenning syndrome, Sitosterolemia, Beta-glucuronidase deficiency mucopolysaccharidosis, McKusick-Kaufman syndrome and Limb Girdle Dystrophy type 21.*

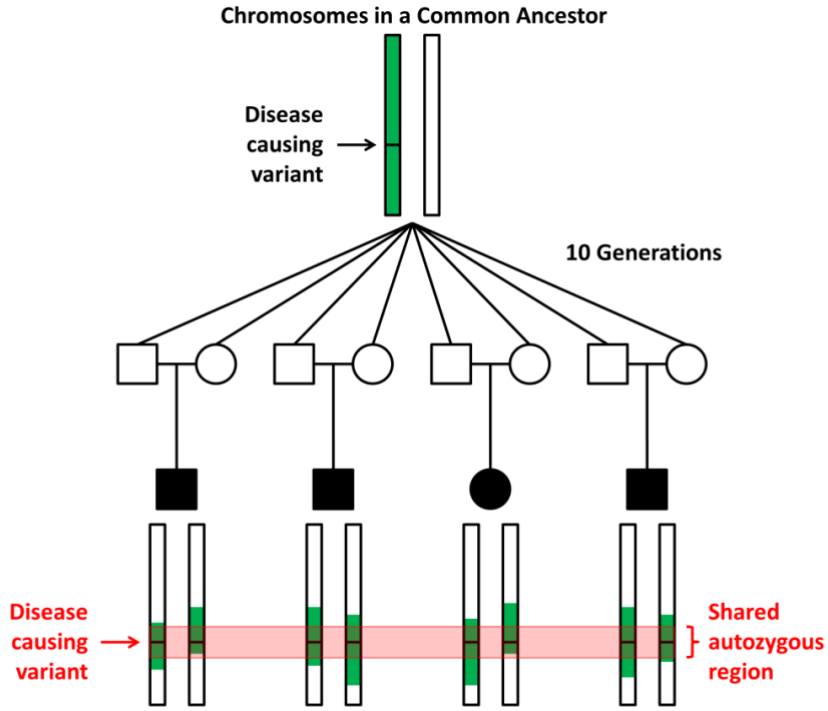



***Figure 1-11 Inheritance of autozygous regions. Taken from[134]***

### 1.7.6 PCD (and other ciliopathies) in the Amish

Studies of inherited disease amongst the Amish have made substantial contribution to the field of human genetics in the last half a century. The table below summarises the genetic causes of PCD that have been identified in the Amish, with the specific communities in which each is known to occur also described.

| Gene | Mutation | Community |
|---|---|---|
| **DNAH5**[164] | 35200<br>c.4348C>T (p.Q1450X) | Amish – Pennsylvania<br>Amish – Wisconsin<br>Mennonite – Arkansas |
| **DNAH5**[164] | c.10815delT (p.P3606HfsX23) | Amish - Wisconsin |
| **DNAI1**[164] | g.IVS1+2_3insT (c.48+2_3insT) | Amish –Mennonite – Midwestern USA |
| **HEATR2**[96,164] | (p.L795P) | Amish –Mennonite – Midwestern USA |

***Table 1-6 Genes and mutations known to cause PCD in the Amish,*** The study references are listed alongside the gene name

DNAH5 was one for the first genes to have been implicated in PCD pathogenesis. Since the original publication, more than 60 mutations in this gene have been identified in patients with PCD, making it one of the most common genes to be linked to the disease[164]. A number of other ciliopathy disorders have been identified in the Amish including: Alstrom syndrome, Bardet-Biedl syndrome, Ellis van Creveld syndrome and Mckusick-Kaufman syndrome.[165,166] The Amish represent an ideal population in which to look for genetic modifiers of PCD because of the reduced genetic variability within the community and the relatively uniform environment and lifestyle.[167]

# 2 Hypotheses and aims

The overarching objective of this project was to learn more about the spectrum of sequence alterations responsible for PCD and SI in Amish populations. This involved two main aims;

1. To genotype a large cohort of Amish families with individuals with PCD and SI for genetic mutations known to cause these conditions in the Amish
2. To evaluate whole genome SNP mapping and exome data, as part of an ongoing research study to identify putative new genetic causes of these conditions.

Together these studies will further our understanding of cilia function and disease pathogenesis, with the potential to identify new therapeutic targets, while furthering understanding of the spectrum of clinical features seen in patients with ciliopathies. Ultimately, this work will translate scientific findings to provide accurate molecular diagnostic testing and inform genetic counselling and targeted management protocols.

# 3 Material and methods

## 3.1 Buffers, reagents and stock solutions

| Buffer, reagent or stock solution | Constituents |
|---|---|
| **Agarose loading buffer** | 40% ficoll, 0.2% xylene cyanol, 0.2% bromophenol blue |
| **Polyacrylamide gel electrophoresis solution** | acrylamide and BIS, Amersco products |
| **Urea gel** | |
| **Ammonium persulphate (APS)** | 98-100% APS |
| **PAGE loading buffer** | formamide, bromphenol blue, xylenecynol |
| **Cell lysis buffer** | 31.84% guanidinium chloride, 25-50%urea, <2% sodium dodecyl sulphate (SDS) |
| **Proteinase K solution** | 25-50% glycerol, <2.5% proteinase, triirachium album serine |
| **Binding buffer** | polyethlyene Glycol 15-20% |
| **Column wash solution used for DNA extraction** | propane-1,2-diol 50-75% |
| **Dreamtaq polymerase** | 0.3% DreamTaq DNA Polymerase, 99.7% 10X DreamTaq Buffer (includes $MgCl_2$) |

***Table 3-1 Buffers, reagent and stock solutions used in the project,[168]***

## 3.2 Subjects and Samples

WOH project studies are reviewed and approved by the institutional Review Board of the Office for Responsible Conduct of Research, based at the University of Arizona, in Tuscon, Arizona, USA and the University of Exeter Research Ethics Committee. The study is in accordance with the principles of the Declaration of Helsinki and in accordance with government guidelines on research governance

in health and social care (found at https://www.gov.uk/government/publications/research-governance-framework-for-health-and-social-care-second-edition) and the Human Tissue Act 2004.

Participants were recruited through the Amish community PCD and SI family support meetings. A full medical and developmental history and systemic examination for the purposes of the research study was obtained for all individuals recruited with informed consent. Additional clinical records (including results of any investigations) were requested using the appropriate information release. Blood and/or buccal samples were collected for DNA extraction.

## 3.3 Data management

Upon sample collection, families are assigned a pedigree and individual identification number (ID). Clinical and molecular information was then recorded in a password protected database. Any further work undertaken was stored digitally, encrypted and password protected.

Family pedigrees are constructed using online a Swiss Anabaptist Genealogical database (www.saga-omii.org) and stored using Progeny, 800 Palm Trail, Suite 200, Delray Beach, FL 33483 USA – see http://www.progenygenetics.com.

## 3.4 Molecular methods

### 3.4.1 DNA extraction

The DNA of selected individuals was extracted from whole blood at room temperature using the Promega ReliaPrep™ Blood gDNA Miniprep system. Blood was thoroughly mixed before DNA extraction. Volume of 200μl of whole blood used according to the following procedure:

- Volume of 200μl of whole blood used. The blood was completely thawed and shaken at room temperature.
- Filter tip pipette tips used as a preventative measure against contamination of samples and equipment.
- 20μl of Proteinase K was dispensed into a 1.5ml microcentrifuge tube and 200μl of blood was briefly mixed with it, using repeat pipetting technique.
- 200μl of cell lysis buffer was added to the microcentrifuge tube, which was then vortexed for 10 to 20 seconds – allowing mixing of the contents in its entirety – to ensure good yield.
- The tube and its contents were then incubated at 56$^{o}$C for 10 minutes.
- During incubation, ReliaPrep™ binding columns were placed into empty collection tubes.
- The tube is removed from the heating block and 250μl of binding buffer is added before vortexing for 10-20 seconds.
- The contents of the tube are then added to the ReliaPrep™ binding column and centrifuged at maximum speed (13, 000 rpm for 1 minute) for a minute.

- The lysate colour should be dark green in colour, and that no residue remains visible on top of the membrane. If residue remains, then the column should be spun for a further minute.
- Binding columns placed in clean collection tubes (flow through liquid discarded as toxic waste in sharps bins)
- 500μl of Column Wash solution added to the column was spun at maximum speed for 3 minutes; further spinning for a minute was initiated if any solution still visible – again, flow through liquid being discarded as hazardous waste.
- The above step is then repeated three times and then the column is placed in a clean 1.5ml centrifuge tube. 200μl of nuclease-free water is added and spun at maximum speed for a minute, to extract the DNA via the action of a solvent. This can be done with a smaller volume – as little as 50μl: concentration is significantly increased but the yield can be reduced by up to 25-30%.
- ReliaPrep™ binding columns are then discarded and the eluate is then measured using nanodrop: DNA's concentration and purity can be assessed, diluting with molecular grade water (if necessary) to achieve a target concentration of 10-30ng/μl.
- Process adapted from the Promega ReliaPrep™ Blood gDNA Miniprep system technical manual.[169]

### 3.4.2 Single nucleotide polymorphism (SNP) genotyping

SNP genotyping was undertaken using Illumina CytoSNP-12v2.1 330K arrays when using the Infinium® HD Assay Ultra manual protocol. This was carried out by Source Bioscience Sequencing, Cambridge UK. Chips that can hold 12 samples each, is hybridised with DNA that is extracted from venous blood or buccal samples, using the process outlined in **DNA extraction**. 200ng of DNA per sample, at 50ng/μl concentration, is needed for the assay that is conducted over a 3-day period:

**Day One:** DNA samples are denatured using a buffer comprising 0.1N NaOH, and neutralised in the pre-PCR room, before isothermal whole genome amplification overnight.

**Day Two:** DNA is fragmented using Illumina fragmentation buffer (FBS), and precipitated using isopropanolol. DNA is then collected by centrifugation at 4$^{o}$C, before it re-suspends, and then denatures at 95$^{o}$C for 20 minutes. Samples applied to a BeadChip, and separated by an IntelliHyb® seal, are then placed in a humidified Hyb chamber and incubate for 16-24 hours at 48$^{o}$C in the Illumina Hybridisation Oven. Amplified ad fragmented DNA samples anneal to locus-specific50-mers.

**Day Three:** The following day, hybridised and non-specifically hybridised DNA are washed away using a buffer (PB1) that contained aliphatic amide. BeadChips

are then prepared for single base extension of the oligonucleotides and the captured DNA, on the BeadChips, are used as a template to incorporate detectable labels.

The BeadChips are stained using Illumina XStain HD BeadChip process before imaging, using an Illumina BeadArray reader. This reader used laser to excite the fluorophores of single base extension products on the beads, allowing the scanner to record images of the light displayed so the genotype can then be determined.

Illumina GenomeStudio Integrated Informatics Platform is used to extract, visualise and analyse the genotyping data collected. Data analysed is the imported into Microsoft Excel for initial analysis, using IBD finder software (http://dna.leeds.ac.uk/ibdfinder), to identify to identify regions of homozygosity are initially greater than 1Mb.

Assistance and support for the SNP genotyping was provided by Dr Gaurav Harlalka and Dr Barry Chioza, University of Exeter.

### 3.4.3 Primer design

Gene sequences were collected using the Ensembl Genome Browser – GRCh37, assembled in 2009.[170] Primer3 software and the ExonPrimer link on the UCSC

genome browser were then used to design primers for use in PCR amplification.[171,172] Primer sequences can be found in appendix 7.5.

The criteria for primer design is outlines in the textbox below:

- 18 to 27 bases long, with length influencing primer's specificity and annealing temperature.
- Melting temperature for either primer to be similar, thus maintaining PCR efficiency. This is usually between 57-63$^{o}$C
- Guanine-cytosine (GC) base content, where possible, will be below 40-60%
- If primers had inter or intra-primer efficiency then they were designed to be no more than 3 bases long to prevent formation of secondary structures and primer dimers.
- Sequences selected were specific and a 100% match to the region to be amplified. Further *in silico* PCR analysis and BLAST searches were facilitated by UCSC Genome bioinformatics website.

***Table 3-2 Criteria for primer design***

### 3.4.4 Polymerase chain reaction (PCR)

PCR is an in vitro method of cloning DNA, in molecular genetics, to facilitate the exponential amplification of selected lengths of DNA – starting with very little DNA. The "touchdown" PCR protocol was used to improve PCR's specificity, by helping to avoid amplification of non-specific products. This involves incrementally lowering (by 2$^{o}$C) of the temperature during the annealing process

every two cycles until the assigned melting temperature is reached. See table below for further details.

| Number of cycles | Temperature (°C) | Time (Seconds) |
|---|---|---|
| 1 | 98 | 120 |
| 2 | 95<br>Anneal +4<br>72 | 30<br>45<br>45 |
| 2 | 95<br>Anneal +2<br>72 | 30<br>45<br>45 |
| 35 | 95<br>Anneal<br>72 | 30<br>45<br>45 |
| 1 | 72: elongation | 300 |

***Table 3-3 Touchdown PCR program particulars***

Throughout, a standard protocol to amplify DNA was used, facilitated by a final volume mix of 10μl, as detailed in table below:

| Component | Volume (μl) |
|---|---|
| DNA (control or sample) or double distilled $H_2O$ | 0.8 |
| Forward primer (Sigma) | 0.4 |
| Reverse primer (Sigma) | 0.4 |
| dNTPs (10mM) | 0.4 |
| 10x buffer (15mM $MgCl_2$, Thermoscientific) | 1.0 |
| Double distilled $H_2O$ | 6.9 |
| Dream Taq polymerase (5U/μl) | 0.1 |
| **Total** | **10.0** |

***Table 3-4 10μl PCR reaction components***

PCR reactions were carried out in an Eppendorf 96 well thermocycler machine and primer pairs were first trialled on control DNA samples that were known to amplify well from previous reactions. PCR reactions producing weak or no product prompted a review of mixture contents and amplifications – as per notes on optimisation of amplification conditions. A negative control, using water was included in every reaction to confirm that the DNA amplified was not the result of contamination. Should results display negative control amplification then results were disregarded as unreliable and experiment was repeated.

### 3.4.5 Optimisation of amplification conditions

Optimal annealing temperatures for the primer pairs were determined by carrying out PCR reactions across a temperature gradient of 55-67$^{o}$C.When working with primers with a guanine-cytosine content greater than 60%, a second gradient with 10% dimethyl sulfoxide (DMA, Fisher Scientific) was performed with alterations to primer concentration and magnesium concentration as required.

### 3.4.6 Agarose gel electrophoresis

Agarose gel electrophoresis of PCR products was used to ensure DNA was amplified sufficiently, with electrophoresis on a 1.5% agarose gel made by adding 1.5g of agarose to 100ml of 1x lithium acetate borate (LAB) buffer and dissolving agarose by gentle heating using a microwave. As the solution cooled, but before gels set, a DNA intercalating agent – called ethidium bromide (EtBr) – was added. EtBr fluoresces brightly when exposed to UV light and so, if it successful,

intercalates into the amplified DNA. Under UV light, EtBr can confirm a successful PCR reaction by showing visualising the amplified DNA. After adding EtBr, gel is poured into casting tray, with a comb in place to create wells and allowed to set.

Gel is then placed in an electrophoresis tank filled with just enough LAB to cover the surface of the gel. Samples were loaded alongside a 1Kb DNA ladder (from Thermo Scientific) to size up the PCR products. 2μl of the gene ladder was loaded in to the well, diluted with agarose gel loading buffer so that it sits at a final concentration of 50 ng/μl. 5μl of each sample was loaded in subsequent wells and electrophoresed at 150V, for 30 minutes, or until the fragments were suitably resolved. A UV transilluminator was used to visualise the PCR product in the gel and photographs were subsequently taken with a digital orange filtered camera. If the DNA fragments were larger, i.e. greater than 500bp, then lower concentrations (0.8-1.0%) of agarose gels were uses.

### 3.4.7 Purification of PCR products

To successfully sequence the PCR products, primers and dNTPs, not been incorporated, had to be removed. Cleaning up PCR products was therefore undertaken with ExoSAP-IT: a substance made of exonuclease-1(EXO) and shrimp alkaline phosphatase (SAP). EXO degrades single stranded DNA in a 3 prime to 5 prime direction. Deoxyribonucleoside 5 prime-monophosphates is then released in a stepwise manner, releasing dNTPs from unincorporated primers. SAP catalyses the release of 5 prime- and 3 prime-phosphate groups from extra

nucleotides. 2μl of the ExoSAP-IT was added to each 5μl of PCR product and then incubated at 37ºC for 15 minutes, providing enzymes' optimal temperature, and then at 95ºC for 15 minutes to terminate enzyme activity.

### 3.4.8 Sequencing reaction

Purified PCR products were sequenced using BigDye Terminator Cycle sequencing kit with the forward and reverse primers in separate reactions. The kit contains: [change this - Taq DNA Polymerase, dNTPs, ddNTPs-dye terminators, $MgCl_2$ buffer. In dye-terminator sequencing, each of the four dideoxynucleotide chain terminators is labelled with fluorescent dyes, each of which emit light at different wavelengths. The protocol used for the sequencing reaction is detailed in the table below.

| Component | Volume (μl) |
|---|---|
| Big Dye | 0.5 |
| Big Dye Buffer | 1.75 |
| Molecular grade $H_2O$ | 4.25 |
| Forward or Reverse primer (10nM) | 0.5 |
| Cleaned PCR products | 3.0 |

***Table 3-5 Components of a 10μl reaction***

Sequencing Reaction Program:

1. Denaturation: 96ºC for 30 seconds
2. Annealing: 50ºC for 15 seconds
3. Elongation: 60 ºC for 4 minutes
4. Repeat above steps for 25 cycles

### 3.4.9 Sequencing reaction purification

Sequencing reaction purification outsourced to Source Bioscience Sequencing, Cambridge UK; purified products were run on an automated DNA sequencer, according to their own protocol. The resulting sequence was analysed using Chromas light (Technelysium Pty Ltd) and Finch TV and incorporated into further study.

Major assistance and support in analysis was provided by Dr Gaurav Harlalka and Dr Barry Chioza, University of Exeter, without who this project would not have come to fruition.

### 3.4.10 Restriction digest

A restriction fragment length polymorphism (RFLP) in a DNA polymorphism either creates or destroys a recognition site on which a restriction enzyme works. Most restriction enzymes recognise palindromic sequences of DNA and RFLP analysis detects variation in the DNA sequence – detectable when restriction enzymes are used to fragment the DNA. One can then ascertain the size of these subsequent fragments by performing electrophoreses on agarose gels. This technique allowed for the identification of previously unknown variants (identified during sequencing) in a control population.

NEBcutter v2.0 website (https://tools.neb.com/NEBcutter2/) was used to select the most appropriate restriction enzyme and then forward and reverse primers, along with their sandwiched, intermediate sequence were inputted into the NEBcutter programme – in both their wildtype and variant forms. Enzymes used were selected because they cut the variant and wild type (WT) sequences differently. In addition, the enzyme had to produce products that were sufficiently distinguishable, between the wildtype and variant sequences, when run on an agarose gel. If variant had no effect on recognised restriction site, or there were pseudogenes present, result interpretation would be complicated; however, controls were sequenced to identify variant.

20μl PCR reactions were set up to amplify a sequence that contains the variant in question: 5μl of this is then run on an agarose gel to ensure successful amplification. Restriction digests are then set up using 5μl of PCR product, in conditions specifically suited to the enzyme, made up to a final volume of 30μl and then incubated at the required temperature. Control DNA samples, negative controls (using water in place of DNA) and known homozygotes and heterozygotes for the variant were then used in this restriction digest.

Known heterozygotes and homozygotes provide the control bands in the restriction digest, against which you can confirm digestion: viewing fragments these on agarose gels would correspond with either wildtype or mutant variant sequences. Electrophoresis is the performed on these incubated samples at 3%

agarose gel for 60 minutes at 100V. The bands produced can be analysed to seek out the specific variant, and frequency of such variant will be estimated by performing restriction digestion multiple control DNA samples.

### 3.4.11 Denaturing polyacrylamide gel electrophoresis (PAGE)

PAGE was used to genotype individuals for the new candidate gene, which corresponded to a 4pb deletion. PCR product used was generated for affected and unaffected individuals, parents and siblings, using the above-mentioned protocol – confirming the presence of PCR product using agarose gel electrophoresis. Successfully amplified products for everyone were then run on polyacrylamide gels, via electrophoresis. The size of an individual's microsatellite DNA can be assessed and compared to parents and siblings, as a crude and visual way of genotyping.

Preparing page gels: Polyacrylamide gel was made by mixing 50ml of 8% PAGE solution (contained acrylamide and BIS, Amresco products), 80μl of TBE-urea gel and 80μl of 25% ammonium persulphate (APS). This solution was then mixed for 20 seconds, using a magnetic stirrer and then poured into the gel casting apparatus. 20 teeth comb was then inserted into the apparatus to create wells, before leaving the gel to set for 40-60 minutes. Some of the gel mixture is left in the beaker to ascertain whether the gel had polymerised or solidified.

Electrophoresis: In setting up the electrophoresis tank, it is possible to run two gels: tank was filled with a litre of warm Tris-borate-EDTA (TBE), and the two plates were loaded so that there were no air bubbles at the bottom side of the loaded gel and that the buffer covers the lower sides of both glass gels.

The PCR products were mixed with denaturing PAGE loading buffer (10ml foramide, 70μl bromophenolblue and 70μl xylene cyanol). The brightness of the PCR product band in amplification on the agarose gel, when viewed under UV transillumination, was considered when deciding the ratios or relative volumes of the PCR product and the PAGE loading buffer. 2μl of PCR: 2μl of loading dye was the typical ratio used. This mixture was denatured by heating to a temperature of 95$^{o}$C for 4 minutes in a thermocycler and then stored on ice.

Before loading the PCR product, and loading buffer solution, wells were washed out using TBE buffer via a syringe and needle to remove urea that remained in the wells. 1.1μl of the denatured product was loaded in each well, before running the gel at 100mAmp, for a length of time determined by the size of the microsatellite PCR product. Run time is usually 50 minute per 100 base pairs.

### 3.4.12 Silver staining

Upon completion, PAGE gels were detached from glass and a mark was made on the left of the first gel and the right-hand side of the second gel to distinguish them from each other, and to mark from which end the gel was loaded. The gel

was then washed with distilled water and then silver stained, according to the following recorded steps. Gels were immersed in the following solutions in turn, whilst on a desktop shaker – rinsing the gel in distilled water between each step.

Solution 1 for 5-10 minutes in 268.5ml dd$H_2O$, 30ml ethanol and 1.5ml acetic acid.

Solution 2 for 15 minutes in 300ml dd$H_2O$ and 0.3g $AgNO_3$ – ensuring to thoroughly rinse gels with dd$H_2O$ here.

Solution 3 for 10-15 minutes, or until bands become visible, in 300ml dd$H_2O$ and 4.5g NaOH and 450µl formaldehyde.

Solution 4 for 5-15 minutes, to prevent the development of other bands, 300ml dd$H_2O$ and 2.2g $Na_2CO_3$.

### 3.4.13 Storing and reading PAGE gels

Gels are wrapped in saran wrap, labelled and stored in the fridge. PAGE gels were then dried with blotting paper and drying machine at 75$^oC$ (time dependent on the size of gel) so that they can then be read.

### 3.4.14 Exome sequencing

Whole exome sequencing (WES) analysis was undertaken inhouse on a NextSeq500 (Illumina, San Diego, CA, USA). DNA samples were targeted using Agilent Sureselect Whole Exome v6. The reads were aligned using BWA-MEM (v0.7.12), with mate-pairs fixed and duplicates removed using Picard (v1.129).

InDel realignment and base quality recalibration were performed using GATK (v3.4-46). SNVs and InDels were detected using GATK HaplotypeCaller and annotated using Alamut batch (v1.4.4). Read depth was determined for the whole exome using GATK DepthOfCoverage. This conforms to GATK Best Practices. Whole exome sequencing data was cross referenced with genome wide SNP genotyping data and filtered for rare, non-synonymous exonic or splice variants, with a population frequency of <0.01 in control databases (including the Genome Aggregation Database; gnomAD, the Exome Aggregation Consortium; ExAC, and the 1000 Genomes Project). Variants were filtered out if there was a depth less than 4x or greater than 1200x, or a consensus quality of less than 20.

# 4 Results

This study stems from a clinical-diagnostic project to define the causes of PCD and SI in the Amish community. This involved genotyping of known genetic causes of these disorders within the community and whole exome sequencing (WES; undertaken in the University of Exeter diagnostic pipeline) alongside whole genome SNP mapping (kindly undertaken by Dr Barry Choiza, with data made available for analysis as part of this project) to investigate for known as well as candidate novel causes of the clinical phenotype. In total, 37 Amish individuals (aged 4y1m-71y3m) affected by PCD and or SI were recruited to the research programme.

## 4.1 Whole genome SNP mapping, dideoxy DNA sequencing and restriction digest identifies HYDIN mutation as a cause of PCD in the Indiana Amish community

Within the Indiana Amish community, a family (**Figure 4-1**) with 3 siblings affected with classical PCD (**Table 4-1**), were born to unaffected parents. A1422 had previously undergone diagnostic molecular genetic testing using a target next generation sequencing panel test for PCD through Invitae. The targeted panel included 32 genes previously associated with PCD, no pathogenic or likely pathogenic mutations were identified. The affected individuals were then screened for founder variants known to cause PCD in the Amish, using dideoxy

DNA sequencing (primer sequences can be found in appendix 7.5). No previously reported variant was identified.

Assuming that a founder gene mutation was responsible for the phenotype seen, autozygosity mapping was undertaken to identify candidate homozygous genomic regions that could harbour gene mutations. This involved genotyping using a genome-wide SNP microarray (Illumina Human CytoSNP-12c2.1 330K) with DNA from affected individuals, unaffected siblings and parents. This revealed a single notable region of homozygosity on chromosome 16 of ~15Mb, as well as two smaller regions of homozygosity on chromosome 7, all shared by the 3 affected individuals (**Table 4-2**). These are genomic regions where affected individuals share the same haplotype, and so may harbour candidate homozygous mutations. Within these genomic regions, no other variants were identified that could not be excluded in genome databases.

| | **A1423/ VI:3** | **A1422/ VI:6** | **A1421/ VI:1** |
|---|---|---|---|
| **Gender** | M | M | F |
| **Age** | 16y5m | 28y8m | 32y5m |
| **Recurrent sinopulmonary infections** | ✓ | ✓ | ✓ |
| **Situs Inversus Totalis** | - | - | - |
| **Infertility** | ✓ | ✓ | ✓ |

**Table 4-1 Clinical presentation of individuals in Indiana Amish family with PCD** ✓, indicates presence of the feature in affected individuals; - indicated features that are absent

| Chromosome | Size region | Genes in the region |
|---|---|---|
| chr16: 66,398,520-81,479,031 [HG38] | 15.08Mbp | 319 |
| chr7: 152,907,691-154,174,878 [HG38] | 1.3Mbp | 11 |
| chr7: 53,831,459-54,934,939 [HG38] | 1.1Mbp | 10 |

***Table 4-2 Notable regions of homozygosity***

In order to identify the causative mutation, whole exome sequence analysis was undertaken. A single affected individual (VI:3, Figure 4-1) was sequenced to generate a novel variant profile. After cross referencing with the SNP data, only a single candidate homozygous variant was identified that was of clear relevance to the phenotype, a novel nonsense mutation in *HYDIN* (NM_001198542; c.2128G>T; p.Glu710*), a gene previously known to be associated with PCD and located within the largest of the shared regions of homozygosity. This variant is in the exon 16 out of 19 exons and would be predicted to result in nonsense mediated mRNA decay. The variant was further validated by restriction digest and was found to co-segregate within the family as appropriate for an autosomal recessive disorder. (**Figure 4-1**, **Figure 4-2** and **Figure 4-3**).

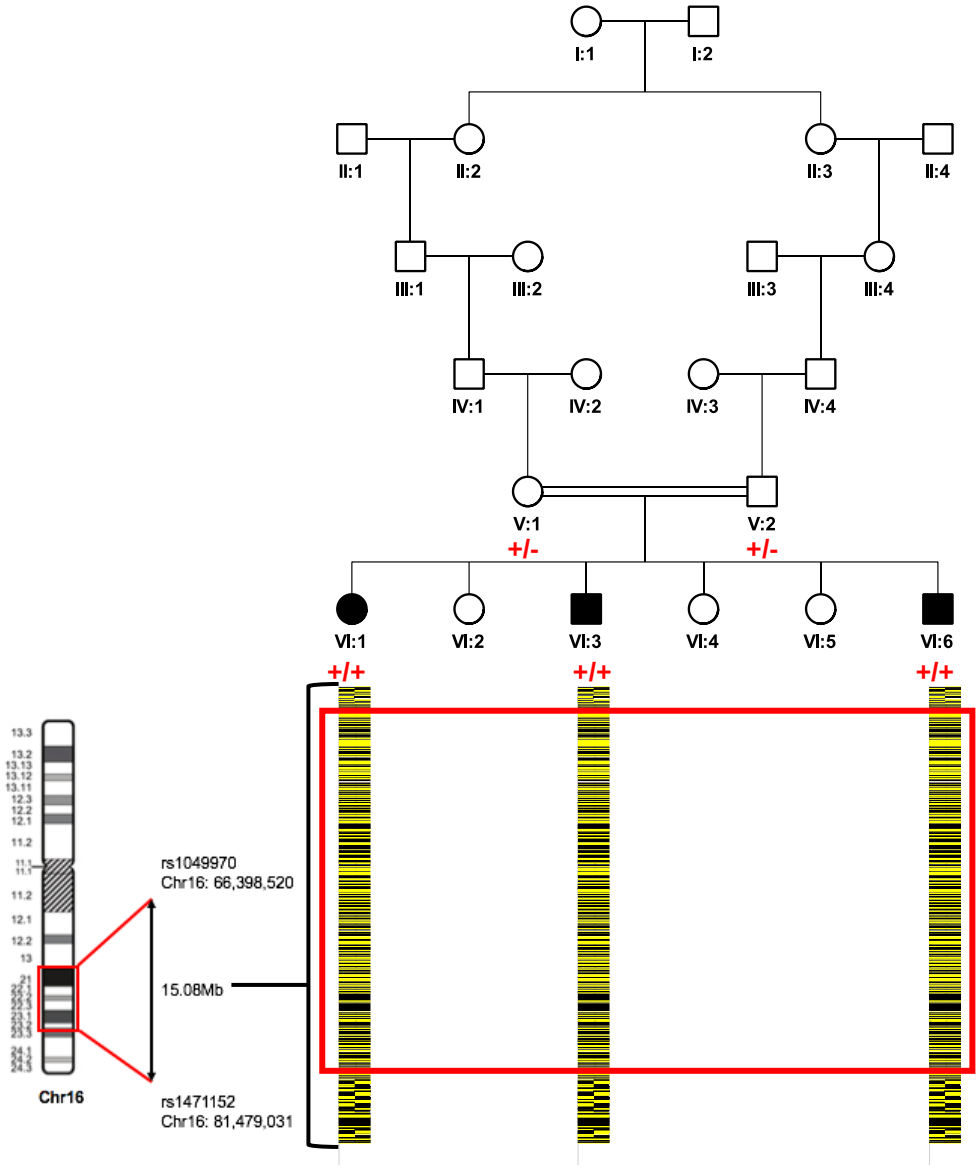



**Figure 4-1 A simplified pedigree of the extended Amish family investigated**, with pictorial representation of genotypes across ~15Mb of chromosome 16 encompassing the disease locus with (+) denoting c.2128G>T mutant and (-) denoting wild-type. Below each individual lies a yellow-black pictorial representation of the SNP genotype of affected individuals, indicating homozygosity for a ~15Mb region of chromosome 20 encompassing the *HYDIN* locus (red box) – also highlighted in red on chromosome 16 ideogram (left). All affected individuals were subsequently shown to be homozygous for the *HYDIN* (NM_001198542 c.2128G>T) variant. Parental samples were heterozygous and unaffected siblings were either wildtype or heterozygous carriers.

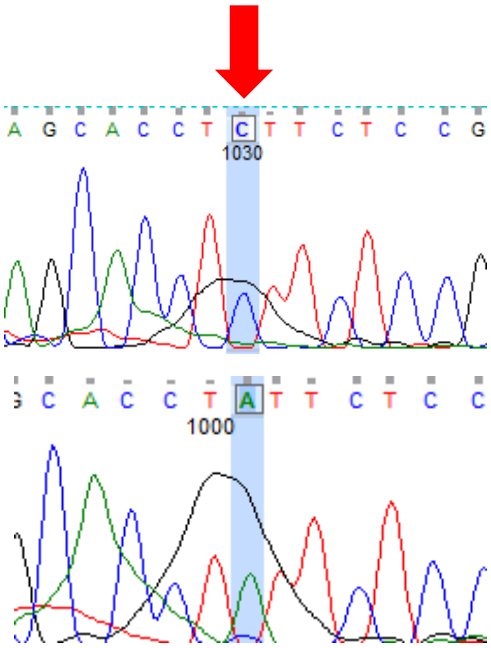



***Figure 4-2 Chromatogram showing the DNA sequence at the position of the HYDIN variant (NM_001198542 c.2128G>T) in a wild-type individual (top panel), and a homozygous affected individual (lower panel).***

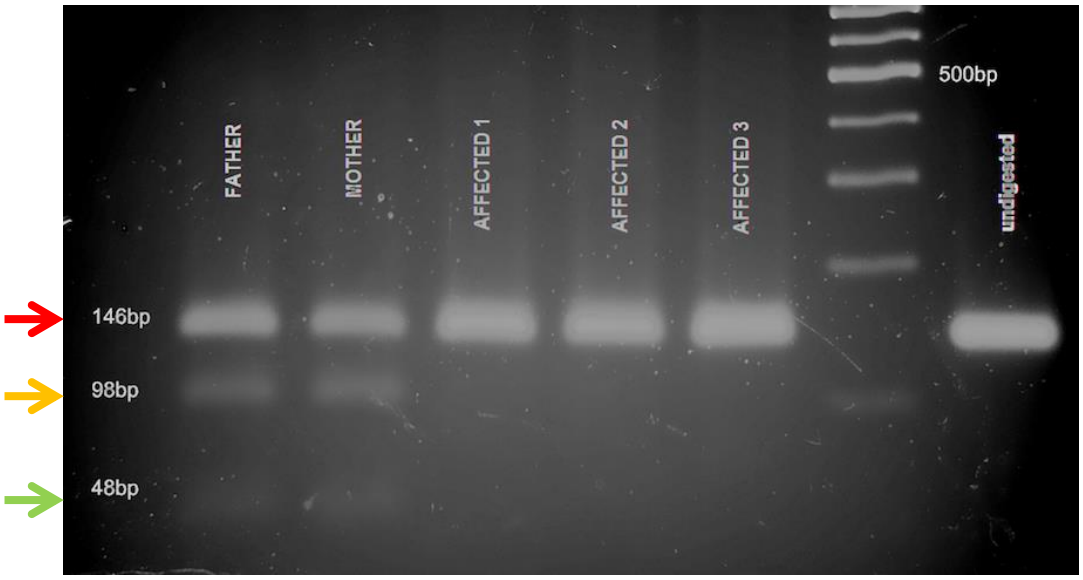



***Figure 4-3 HYDIN EarI restriction digest***

A1420 – father; A1419 – mother (both unaffected); A1421, A1422, A1423 – affected 1, 2, 3. The c.2128G>T alteration abolishes a EarI restriction site, and so the mutant allele remains undigested. The red, amber and green arrows highlight the digested (WT) 146bp (red) PCR product into 98bp (yellow) and 48bp (green) portions of DNA

The variant's frequency was investigated in the Amish patient cohort screening the DNA from all the other study participants with PCD or SI by dideoxy DNA sequencing (Primer sequences can be found in appendix 7.5). No other individuals were found to be homozygous or heterozygous for the variant. Inspection of the Amish exome database of 76 Amish controls identified no heterozygous carriers or homozygous individuals. Given the clinical picture seen in affected family members and *HYDIN's* well-defined role as a gene mutation in PCD, this variant was considered most likely to explain the clinical findings in this family. This appears to be a novel variant in a gene known to cause PCD as it is unreported on the gnomAD database. Given the conservation of the Glu710 residue (**Figure 4-4**) and the surrounding amino acids, it seems that this *HYDIN* variant introduces nonsense-mediated mRNA decay.[173,174] **b**

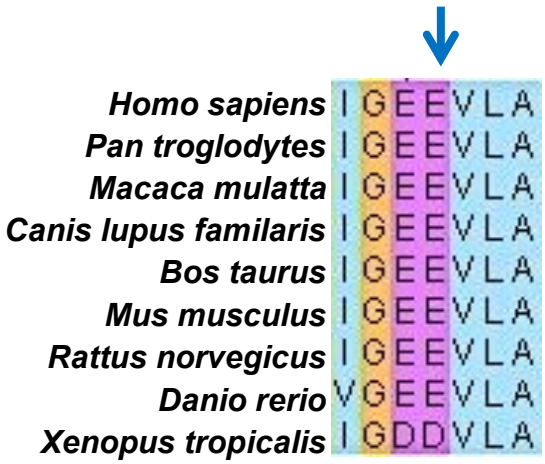



***Figure 4-4 Conservation of Glu710*** A clustal W2 alignment of amino acid sequences of *HYDIN* in Homo sapiens and 8 other species, illustrating stringent

[b] Nonsense-mediated mRNA decay is a surveillance pathway that reduces the errors in expression by eliminating mRNA transcripts that contain premature stop codons and subsequent proteins.[172,173]

conservation of Glu710 residue – indicated using blue arrow. (Created using Jalview software)

## 4.2 Whole genome SNP mapping and WES identifies *DNAH11* mutation as a possible cause of SI in the Missouri Amish community

Within the Missouri Amish community, a family that features 2 siblings affected by the same clinical phenotype, born to unaffected parents. Their clinical phenotype was as follows:

| | **A1425/ VII:3** | **A1426/ VIII:8** |
|---|---|---|
| **Gender** | M | M |
| **Age** | 70y8m | 63y11m |
| **Recurrent sinopulmonary infections** | ✓ (mild) | ✓ (mild) |
| **Situs inversus totalis** | ✓ | ✓ |
| **Infertility** | N/K | N/K |

***Table 4-3 Clinical presentation of individuals in Missouri Amish family with PCD/SI***

**NB:** ✓, indicates presence of said feature in affected individuals; - indicated features that are absent; N/K, indicates not known

Assuming that a founder mutation was responsible for the phenotype seen, autozygosity mapping was undertaken to identify candidate homozygous regions, that could harbour founder gene mutations. This involved genotyping using a genome-wide SNP microarray (Illumina Human CytoSNP-12c2.1 330K) with DNA from affected individuals. This revealed four small, but notable, regions

of homozygosity on chromosomes 7, 8, 13 and 20, detailed in **Table 4-4**. These are genomic regions where affected individuals share the same haplotype and so may harbour the same candidate homozygous mutations.

| Chromosome | Size region | Genes in the region |
|---|---|---|
| 8:19,596,237-22,531,463 [HG19] | 2.9Mbp | 27 |
| 13:47,662,178-49,308,307 [HG19] | 1.6Mbp | 8 |
| 20:30,567,677-31,619,297 [HG19] | 1.1Mbp | 15 |
| 7:157,714,638-159,119,486 [HG19] | 1.4Mbp | 5 |

***Table 4-4 Notable regions of homozygosity***

WES was also undertaken on a single affected family member (A1525) and the data was cross referenced with SNP mapping collected. This identified a single candidate variant in the homozygous regions, a deleterious missense mutation in *DNAH11* (NM_003777.3; c.10494G>A; p.Arg3498His), predicted to be deleterious by the *in-silico* prediction programmes SIFT and Polyphen. Biallelic variants in *DNAH11* have previously been described in association with PCD. This variant was validated by dideoxy DNA sequencing and was found to co-segregate as appropriate for an autosomal recessive disorder, as displayed in figures **Figure 4-5** and **Figure 4-6**. (Primer sequences can be found in appendix 7.5)

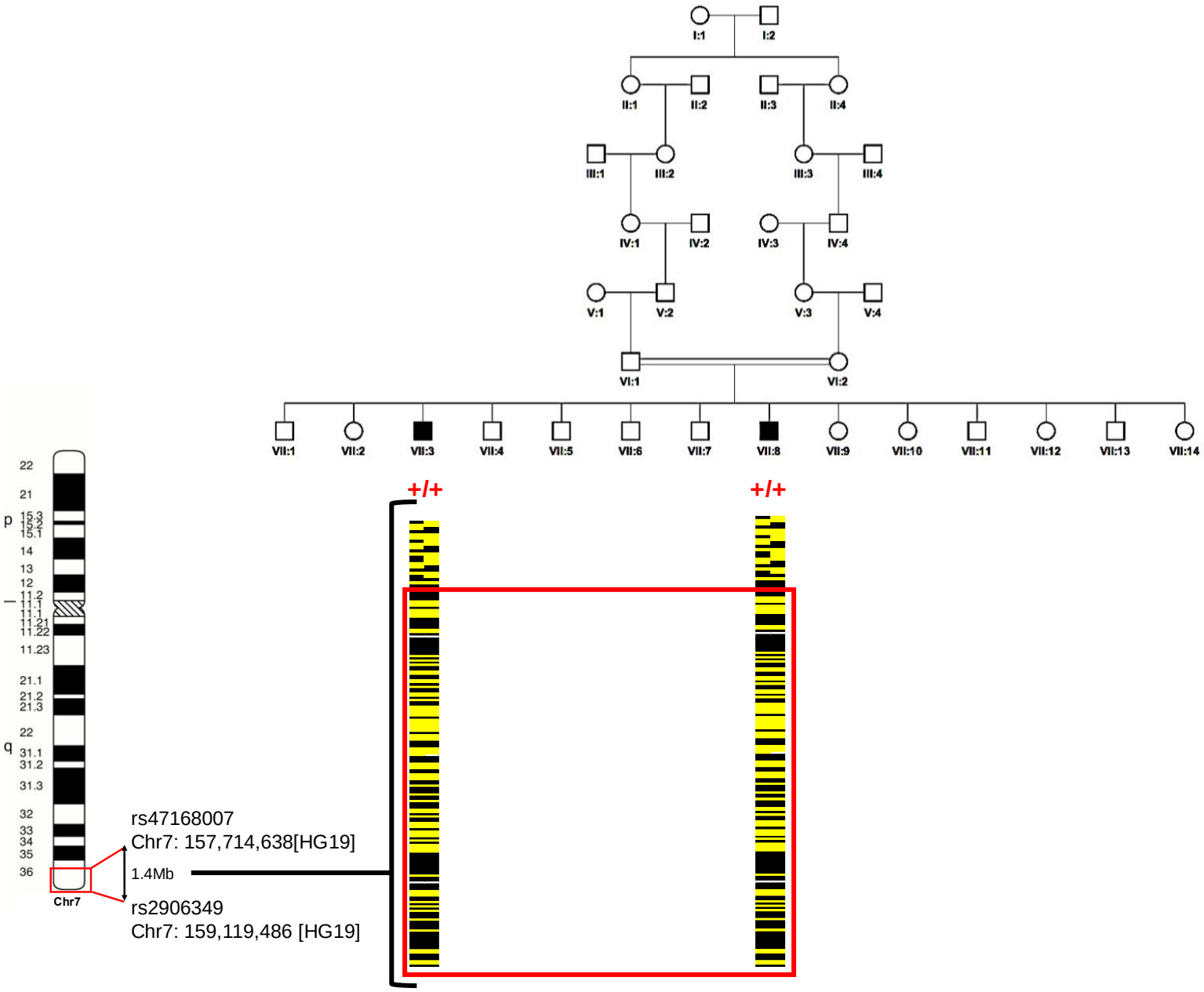



**Figure 4-5 A simplified pedigree of this Missouri Amish extended family investigated,** with a pictorial representation of the genotypes ~1.4Mb region of chromosome 7 encompassing the *DNAH11* locus with (+) denoting c.10494G>A mutant and (-) denoting wild-type. Below each individual lies a yellow-black pictorial representation of the SNP genotype of affected individuals, indicating homozygosity for a ~1.4Mb region of chromosome 7 encompassing the *DNAH11* locus (red box) – also highlighted in red on chromosome 7 ideogram (left). All affected individuals were subsequently shown to be homozygous for *DNAH11* (NM_003777.3 c.10494G>A) variant. Parental samples were heterozygous and unaffected siblings were either wildtype or heterozygous carriers.

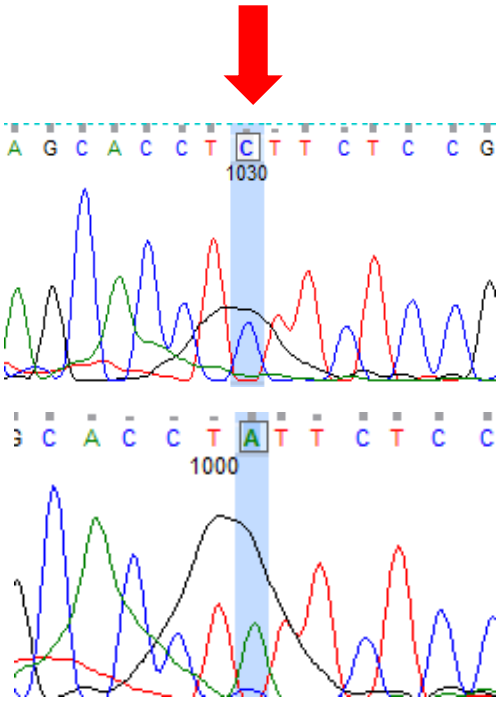



***Figure 4-6 Chromatogram showing the DNA sequence at the position of the DNAH11 variant (NM_003777.3 c.10494G>A) in a wild-type individual (top panel), and a homozygous affected individual (lower panel).***

To investigate the frequency of the variant in the Amish patient cohort, DNA from other participants was screened with dideoxy DNA sequencing. This identified 4 further individuals homozygous for the variant out of 24 individuals successfully genotyped. By comparison, inspection of the Amish exome database of 76 Amish controls, identified only 3 heterozygous carriers, with no homozygous individuals. Given the clinical picture seen in affected family members and the well-defined role of *DNAH11* gene mutation in PCD, this variant was considered as a possible case of SI in this family. The variant codes for the Arg3498 residue, which is stringently conserved (as seen in **Figure 4-7**).

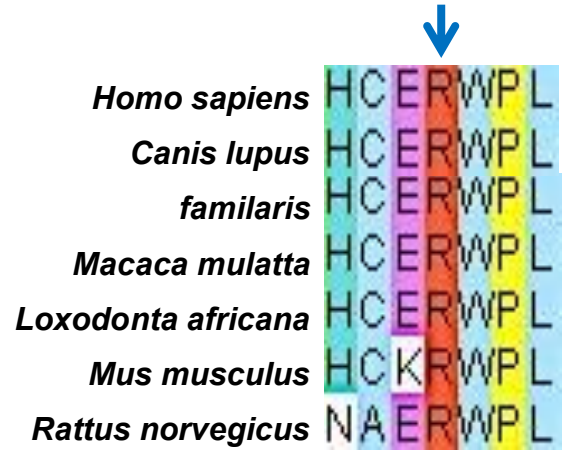



***Figure 4-7 Conservation of Arg3498***

A clustal W2 alignment of amino acid sequences of *DNAH11* in *Homo sapiens* and 6 other species, illustrating stringent conservation of Arg3498 residue – indicated using blue arrow. (Created using Jalview software)

## 4.3 Genetic studies identify a candidate new genetic causes of disease, in genes not previously associated with SI and PCD

WES was also undertaken in 11 additional individuals with SI, from the North American Amish community and 2 siblings with PCD from the Canadian Amish community to exclude mutations in other known genetic causes of the condition, as well as identify candidate new genetic causes of these conditions. Sequencing data was filtered to identify homozygous and potential compound heterozygous variants. Variants were filtered according to frequency (<0.01) with priority given initially to likely deleterious variants (frameshift, nonsense, missense and essential splice site). Further consideration was given for those that were predicted as deleterious with an unreported frequency. This identified two stand-out candidate genes that warranted initial further investigation.

### 4.3.1 A candidate gene for SI, encoding a meiosis specific protein

One of the most interesting findings from the WES analysis was a homozygous 4bp deletion, (c.407_410del) in exon 4 of a 10-exon gene. The gene encodes a meiosis specific protein (termed '*MSP*') and the variant is predicted to cause a reading frameshift resulting in a loss of function mutation. The variant is novel, (not reported on the gnomAD, ExAC or 1000 genomes databases). The *MSP* variant was identified initially in a single individual with isolated SI totalis. This finding was notable as the gene has recently been shown to be associated with a phenotype of SI and hydrocephalus in a mouse model, identifying the variant as a strong candidate for SI in humans. To investigate this, further genotyping of the rest of the cohort of Amish SI and PCD cases was undertaken using PAGE to determine the frequency of the variant (**Figure 4-8**).

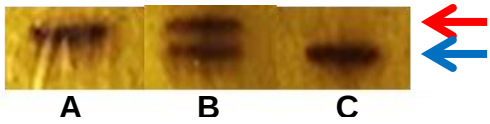



***Figure 4-8 An example of the PAGE genotyping defining the wild type allele (upper red arrow) and deleted allele (lower blue arrow). A: wild type, B: heterozygous carrier, C: homozygous for the candidate c.407_410del MSP pathogenic variant.***

This identified 4 further individuals with isolated SI totalis who were homozygous for the *MSP* variant out of 36 individuals successfully genotyped. By comparison, inspection of the Amish exome database of 76 Amish controls, identified only 4 heterozygous carriers for the *MSP* 4bp deletion variant, with no homozygous individuals.

### 4.3.2 A candidate gene for SI encoding a DIX domain containing protein

WES of the affected SI cohort also identified a second candidate genetic cause of the condition, a homozygous nonsense variant (c.1168c>T; p.Gln390*) in exon 12 of a 21-exon DIX (dishevelled-axin) domain-containing gene.[175,176] This DIX domain gene is a positive effector on the Wnt signalling pathway[c]. The variant is novel (not reported on the gnomAD, ExAC or 1000 genomes databases) and disrupts an amino acid that is stringently conserved (**Figure 4-9**), although is predicted to be a nonsense mutation and the transcript would be expected to undergo nonsense mediated mRNA decay.

To investigate the frequency of the candidate variant in the Amish patient cohort, DNA from other participants was screened with dideoxy DNA sequencing. (Primer sequences can be found in appendix 7.5) This identified 4 individuals homozygous for the variant out of 24 individuals successfully genotyped. By comparison, inspection of the Amish exome database of 76 Amish controls, identified only 3 heterozygous carriers, with 1 homozygous individuals; it is not known for certain whether this control individual has SI or not. It is possible that this individual may have SI, but given the lack of any clinical symptoms

[c] Wnt signalling pathways are a group of pathways, made from proteins that pass signals into a cell through cell surface receptors. There are two types: the canonical or β-catenin dependent pathway and the non-canonical or β-catenin-independent.[174-177]

associated with pure SI, it is entirely possible that the condition may not have been diagnosed.

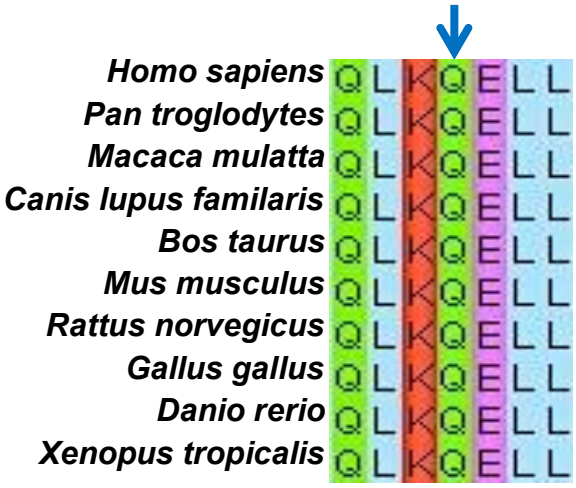



***Figure 4-9 Conservation of Gln390***

A clustal W2 alignment of amino acid sequences of this *DIX*-domain gene in *Homo sapiens* and 9 other species, illustrating stringent conservation of Gln390 residue – indicated using blue arrow. (Created using Jalview software)

## 4.4 Whole genome sequencing SNP mapping and WES fails to identify the cause of PCD in a Canadian Amish sibship

Within Canadian Amish community, 2 siblings affected by PCD were born to unaffected parents. The inheritance of the disorder in the family was most compatible with autosomal recessive inheritance (**Figure 4-10**). The clinical phenotype of the affected siblings is described in (**Table 4-5**):

| | A1510 | A1511 |
|---|---|---|
| Gender | M | F |
| Age | 20y0m | 30y3m |
| Recurrent sinopulmonary infections | ✓ | ✓ |
| Situs inversus totalis | ✓ | ✓ |
| Infertility | N/K | N/K |

***Table 4-5 A comparison of the clinical phenotype of individuals in Canadian Amish family***

**Table key:** N/K, not known; ✓, indicates presence of said feature in affected individuals; N/K, indicates not known

Assuming that a founder mutation was the most likely cause of PCD in this family, autozygosity mapping was undertaken to identify candidate homozygous genomic regions that could harbour gene mutations. Using a genome-wide SNP microarray (Illumina Human CytoSNP-12c2.1 330K), genotyping was carried out using the DNA from affected individuals, and unaffected parents. This identified three small regions of homozygosity on chromosomes 2, 7 and 11 (**Table 4-6**)

common to the affected individuals. These are genomic regions where affected individuals share the same haplotype, and so may harbour candidate homozygous mutations.

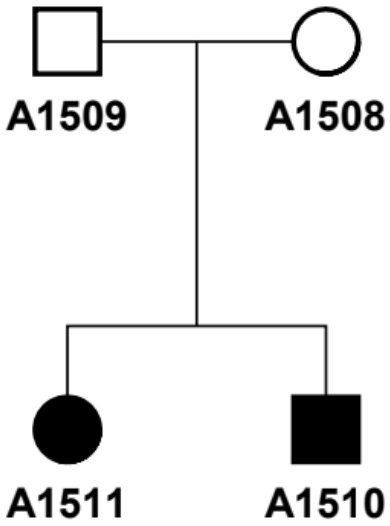



***Figure 4-10 Simplified pedigree of nuclear Canadian Amish family***

| Chromosome | Size region | Genes in the region |
|---|---|---|
| chr11: 100,246,575-104,857,221 [HG19] | 4.6Mbp | 27 |
| chr 2: 240,482,934-243,029,573 [HG19] | 2.5Mbp | 36 |
| chr7: 65,437,774-66,962,579 [HG19] | 1.5Mbp | 9 |

***Table 4-6 Notable regions of homozygosity***

After establishing the shared regions of homozygosity between the siblings, WES was carried out on one of the affected siblings (A1511), to identify candidate rare or novel homozygous variants in these regions as well potential compound heterozygous variant genome wide.

The data was filtered, looking for rare and novel deleterious variants (frameshift, nonsense, missense, and essential splice site variants initially), with an initial threshold of 0.01 for the allele frequency. If a variant did not have a SIFT or POLYPHEN prediction, or it did not have a recorded allele frequency then, they were still included in the shortlist. If a gene was found to have any individuals reported to be homozygous on gnomAD, then the variant was also excluded, leaving only a very short list of variants for consideration. Following this analysis, only a single gene variant was identified in the homozygous regions, in the MMP1 gene – a gene which codes for an interstitial collagenase protein (further discussed in chapter 5.1). Other genes identified genome wide, that may have fallen into a smaller homozygous region were also considered: variants in the RNF11 and ASB18 genes listed below (**Table 4-7**) were also considered as they met the above listed criteria.

| Chromosome | Gene | Gene Transcript | dbSNP | ExAC_NFE allele frequency | gnomAD Allele frequency | Homs | Hets | SIFT prediction | POLYPHEN prediction |
|---|---|---|---|---|---|---|---|---|---|
| ***Looking in homozygous regions – chromosome 2, 7 and 11**** | | | | | | | | | |
| 11 | MMP1 | NM_002421.3:p.Ala350Thr/c.1048G>A | rs149950482 | 0.003 | 0.0002138 | 0 | 151 | D | B |
| ***Looking genome wide*** | | | | | | | | | |
| 1 | RNF11 | NM_014372.4:p.Pro38Arg/c.113C>G | rs775267881 | 0.00001861 | 0.00001890 | 0 | 2 | T | P |
| 2 | ASB18 | NM_212556.2:p.Gly242Arg/c.724G>C | 0 | 0 | 0 | 0 | 0 | D | D |

***Table 4-7 Exome sequencing analysis for A1511***

**Table key:** * – Chromosome 2 and 7 had no variants that fulfilled the criteria for further consideration. B – benign, D – damaging, P – probable, T – tolerable

# 5 Discussion and future work

## 5.1 Discussion

This thesis describes the investigation of molecular causes of SI and PCD within Amish families, undertaken as part of an ongoing clinical-genetic program to improve knowledge of inherited disease in the community. As part of the Windows of Hope program, 37 individuals (aged 4y1m-71y3m) were recruited with SI or PCD. Mapping, traditional and next generation sequencing technologies were used to identify genes and mutations responsible for the clinical picture. Causative mutations were verified by co-segregation and regional control analysis.

In studying the causes of disease in families from the Indiana Amish community, there were three siblings with the same clear clinical phenotype of PCD, born to unaffected parents. Genomic studies identified a nonsense variant in the *HYDIN* gene (NM_001198542 c.2128G>T), cosegregating, as appropriate for an autosomal recessive condition, as the likely cause. All affected individuals in this family were found to be homozygous for the variant, which is predicted to result in nonsense mediated mRNA decay of the resulting transcript.

The *HYDIN* gene codes for the axonemal central pair apparatus protein, and so is involved in axonemal central apparatus assembly, cilia movement and brain development (characterised by accumulation of cerebrospinal fluid in the

ventricles of the brain).[177,178] It is therefore associated with PCD and hydrocephalus and is found in the axonemal central pair of microtubules, crucial for functional movement of cilia; sperm motility is also reduced.[179,180]

Studies show that *HYDIN* dysfunction impairs ciliary motility in mice; those affected develop hydrocephalus. Mutant *HYDIN* cilia are incapable of bending normally, with reduced ciliary beat frequency and often stall. Affected cilia are therefore incapable of generating fluid flow, with these defects apparent in the trachea and in the brain. Subsequent hydrocephalus is the result of impaired cilia-generated fluid flow in the brain because of dysfunctional cilia.[181]

*HYDIN* mutation directly affects the central pair of microtubules (CP), explaining why those that are homozygous for *HYDIN* mutations do not also present with SI. Nodal cilia (with a 9+0 arrangement) do not have a central pair of microtubules and are therefore unaffected by *HYDIN* mutations. It is the action of nodal cilia, during embryogenesis, that determines the randomisation of organs in left-right asymmetry. This is consistent with the presentation of affected individuals in this study. This would be further supported by a collection of patients affected by PCD, with central pair defects but no situs abnormalities.[182-186]

While not a common cause of the condition, *HYDIN* gene mutation is a well-established cause of PCD. Recessive *HYDIN* mutations have been shown to cause PCD, without situs abnormalities in the past; 6 individuals form a another

founder community, based in the Faroe Island, shared a haplotype, within which *HYDIN* gene lies.[187] In these individuals, *HYDIN* mutant respiratory cilia lack central pair of microtubules: similar finding have been reported in *HYDIN*-deficient *Chlamydomonas* and mice.[188,189]

In short, this *HYDIN* variant appears to be a clear loss of function variant that presents as a novel cause of PCD in the Amish. When looking at the frequency of the variant in Amish controls, there were no homozygotes, further strengthening our case. What's more, there were, unsurprisingly, no heterozygous carriers of the variant: the family in which this variant was identified were from an Amish community in Indiana, whilst the majority of Amish controls were collected from Amish settlements in Ohio. This may point to further variation of the Amish gene pool, distinguished by geographical location of an extended families within communities.

The identification of a PCD-causing variant in a gene previously linked with PCD has provided important information that can inform further gene panel and screening studies in Amish founder communities. With further investigation, the expectation is the mutation will continue to be found in Amish patients with diagnoses yet to be made: evidence will suggest that this variant is enriched in the community with the number of heterozygous carriers more commonplace than in other populations. This furthers the scientific community's understanding of autosomal recessive inheritance in founder communities and has broadened

our appreciation for the gene pool that exists in the Amish. The findings are highly suggestive that the *HYDIN* mutation is responsible for clinical picture seen, and providing a diagnosis for the Amish family. Without such studies, many of the affected individuals and their families would still be without a clinical diagnosis, follow up and the necessary education and counselling. In the future, it would be useful to consider functionality and consider how biological process may, or may not, be different because of harbouring the mutation. Drawing connections will allow us to outline a clear molecular link between the disease-causing gene alteration and cellular and clinical phenotypes observed.

In studying the Missouri Amish community, there were 2 siblings with a clinical phenotype of SI and PCD, born to unaffected parents. SNP mapping and exome data was collected from the siblings and further analysed. During subsequent analysis, *DNAH11* was the only standout potential causative gene and, by performing dedioxy DNA sequencing (**Figure 4-6**) on aforementioned individuals, the *DNAH11* variant (NM_003777.3 c.10494G>A), was found to cosegregate as appropriate for an autosomal recessive condition. All affected individuals were found to be homozygous for a c.10494G>A Amish variant in *DNAH11*, which results in the replacement of the stringently conserved arginine residue by histidine at position 3498 (p.Arg3498His) of the protein (**Figure 4-7**).

The identification of a known PCD-causing *DNAH11* has provided important information to further inform gene panel and screening studies in Amish

communities, developing our understanding of autosomal inheritance in founder communities. This work also broadened the depth of knowledge we possess about the Amish gene pool. The frequency of this variant was ascertained in amongst other Amish individuals with SI or PCD. Aside from the individuals already identified, no further homozygous individuals were discovered in those so far screened within the SI/PCD Amish cohort. However, within Amish controls there were 3 heterozygous carriers. This is an allele frequency of approximately 0.03947, compared to a published gnomAD allele frequency of 0.0006482, providing evidence that there is evidence of enrichment within the community. It is notable that there are 82 known heterozygous carriers of the variant on the gnomAD database, which may be higher than expected for such a condition. Whilst this isn't sufficient evidence to exclude *DNAH11* as the cause, it does cast some doubt. However, further consideration should be given to a milder more mixed phenotype seen in affected individuals homozygous for this variant as one explanation could be that the *DNAH11* variant is responsible for a milder phenotype, which would be consistent with the SI phenotype seen in this study and which may not necessarily be identified in affected individuals.

*DNAH11* is a gene responsible for generating respiratory cilia proteins – known as axonemal dynein heavy chain 11 proteins, responsible for the generation of movement in the cilia through ATP and nucleotide binding. In so doing, *DNAH11*'s dysfunction can affect a range of biological processes, including determination of left-right asymmetry, sperm flagella motility, regulation of cilia

beat frequency. Some evidence suggests *DNAH11* may also have a role in the development of the heart, as well as learning and memory.[190-192] Interestingly, only half of those with PCD caused by a mutation in this gene will have situs abnormalities due to dysfunctional embryonic nodal cilia: investigating this in depth may shed light on how and why situs abnormalities occur. In fact, a murine model demonstrated how a mutation in the homologous mouse gene *DNAH11*, was characterised by SI and immotile cilia in the embryonic node.[192]

Taken together, the evidence, so far collected, is supportive of the *DNAH11* (c.10494G>A) variant as the most likely causative variant in the Amish family investigated. Further investigations of additional families will ultimately help determine this.

As well as mutations in genes previously known to cause SI/PCD, our studies also identified candidate new genetic causes of the condition. Our studies identified a deleterious 4bp deletion (c.407_410del; p.Glu136Glyfs*16) in the coding region of a meiosis specific protein (termed '*MSP*'). This is novel mutation: unreported on the gnomAD database. Of the 35 individuals successfully genotyped in the SI/PCD Amish cohort, 4 were homozygous for the variant, with 4 heterozygous carriers. In the 76 Amish controls, a further 4 individuals were found only to be heterozygous carriers for the variant. The *MSP* gene is not tolerant of change: there are 63 loss of function variants published on gnomAD – all with a very low allele frequency, and importantly with no listed homozygotes.

The *MSP* gene is protein coding and thought to play a role in meiotic division and germ cell differentiation. On a molecular level the gene has been linked with protein binding, affecting cilia organisation and assembly, left-right asymmetry, mitosis and spermatogenesis. A mouse model has been generated in which mice have been found to share an overlapping phenotype of SI with the Amish individuals investigated here. While this is strongly supportive and evidence that this is a candidate gene that warrants further investigation, further work is required. The next steps may involve further investigation of mouse model alongside affected individuals to more precisely define overlapping phenotypical outcomes. Additionally, the identification of additional Amish as well as non-Amish individuals with the condition, will help confirm the gene as causation. Finally, further molecular studies to investigate the ciliary role of MSP will provide further insight into the pathogenesis of SI, and the basis of this condition.

Whole exome sequencing (WES) also identified a second candidate gene, a dishevelled-axin domain-containing (*DIX*) gene known to encode a molecule with aa positive effect on the Wnt signalling pathways[4]. Wnt signalling pathways are a group of pathways, made from proteins that pass signals into a cell through cell surface receptors.[176,193,194] The variant identified in the Amish – novel and

[4]Wnt, here, is a blend of old gene nomenculature: Wnt is the blending of int and Wg, meaning "Wingless-related integration site"[193]

unreported, was a homozygous nonsense mutation (c.1168c>T; p.Gln390*), predicted to lead to either early truncation of the protein product or nonsense mediated mRNA decay, making it a strong candidate for SI and or PCD. Whilst the mutation was unreported on gnomAD, sequencing identified 4 homozygous individuals, out of 24 with SI that were successfully genotyped. In the Amish control cohort, there were 3 heterozygous carriers and one homozygote: while not conclusive, this demonstrates enrichment of this mutation, in the affected group. When considering the *DIX* gene itself: there are 43 loss of function mutations and only one of these has any homozygotes in the gnomAD database, indicating that the *DIX* domain gene is intolerant of deleterious alterations. A Multiz alignment of amino acids coded for by this *DIX*-domain gene highlights its stringent conservation across the different species, alluding to its utility and potential necessity for correct cilia function (**Figure 4-9**)

Interestingly, the *DIX* mutation was originally identified an individual who was initially investigated as part of another study, exploring new causes of global developmental delay. This individual had a clinical picture of global developmental delay with unknown genetic cause. A further question could be raised as to why this individual did not, also, have a diagnosis of SI as well. This could be explained by the fact that it is not uncommon for members of a founder community to have more than one genetic disease. It is not implausible for a single individual to have inherited more than one disease causing variant. In addition, the individuals that were homozygous for this variant had variable

phenotypes: whilst some had a clinical picture of isolated SI, other's phenotype included signs and symptoms of mild PCD. One explanation could be that this mutation leads to a variable phenotype that may include PCD symptoms. Studies of DIX gene function do not relate to a specific area of biology from which we can draw a clear link to what is already known about the molecular mechanisms that underlie ciliopathies. Using mouse models, the gene has been linked to lung cancer – not overtly related to PCD. The *DIX*-domain gene is also thought to be a positive effector on Wnt signalling pathways, particularly the canonical or β-catenin dependent pathway. While increase canonical Wnt signalling has been implicated in ciliopathies affecting sensory and motile cilia, it is not obviously clear whether this is concomitant with the molecule being a cause of the condition.[195,196] This is a complex area of study and requires copious further investigation.

Finally, genetic studies also involved an Amish family from a settlement in Ontario, Canada. The family comprised two affected siblings, both presenting with classic signs of PCD and dextrocardia, born to unaffected parents. The condition in this family is thus indicative of a monogenic autosomal recessive cause of their disorder. Assuming this, SNP mapping was used to identify regions of the genome shared between the siblings, as a starting point to look for pathogenic variants. Through exome data analysis, considering deleterious homozygous variants, as well as compound heterozygous variants, identified only one candidate variant to investigate within the earmarked homozygous

regions in MMP1. The parameters used to filter candidate variants included ExAC and gnomAD allele frequencies of <0.01 and SIFT and POLYPHEN deleterious predictions. However, many questions remain regarding the relation between causative gene and the subsequent phenotype. It is possible that there is no clear monogenic cause and that there is a more complex genetic cause, perhaps a deep intronic cause of the disorder. The MMP1 gene variant (NM_002421.3:p.Ala350Thr/c.1048G>A) was the only deleterious change in the genome that was of potential interest. This variant has a high gnomAD carrier frequency, with 151 heterozygotes listed, decreasing the likelihood that this is disease-causing. However, there are no known homozygotes for the variant listed, and so this does not conclusively exclude the variant as causative, The MMP1 gene encodes for a interstitial collagenase protein, where dysfunction is linked with collagen degradation airway remodelling and chronic obstructive pulmonary disease (COPD).[197,198] On first review this makes the link to the PCD/SI phenotype unlikely; however, it's possible that the two may be linked: bronchiectasis and COPD can be seen secondary to cases of PCD – perhaps alluding to the genes relevance.[199] Another explanation is that the individual exomed in this instance has a clinical phenotype that features some mild collagen degradation, at a severity so mild so not disease classified.

It is possible that other potential disease loci were missed when searching with this criterion; the search criteria was then extended to look genome wide, considering homozygous and compound heterozygous variants. An additional

homozygous candidate was identified in RNF11 (NM_014372.4:p.Pro38Arg/c.113C>G): a missense variant, predicted as possibly damaging by POLYPHEN. There were no recorded homozygotes and only 2 heterozygous carriers published on gnomAD so initial thoughts were that this could be a promising candidate. However, cross-referencing of the two siblings' mapping data can exclude this as a potential cause; whilst the genotype of the proband was homozygous, the other sibling was heterozygous at the same loci, making it highly unlikely that this is the causative variant. For it to be a viable candidate, the haplotype seen in the mapping data would have to be identical at (and surrounding) the potential disease loci. In addition, published mutations in this gene suggest relevance to an alternative phenotype – namely breast and prostate carcinogenesis.[200]

The second candidate identified, looking genome wide, was the ASB18 variant (NM_212556.2:p.Gly242Arg/c.724G>C), novel and unreported on gnomAD. Cross-reference with the SNP mapping data confirmed that the siblings were homozygous and shared a haplotype in the surrounding genomic region (chr2 ~237Mb). Not much is known about this gene's structure or function but further investigation is warranted. ASB18 codes for proteins involved in ubiquination – part of the protein modification process but do not appear to be related to PCD/SI.[201,202] Using STRING analysis link can further be drawn from ASB18 to UBA52; UBA52 is a gene that has been noted as overexpressed during airway epithelium differentiation.[203] Now given that airways epithelium are often ciliated,

it would be unwise to discount ASB18 as a potential candidate gene – particularly as UBA52 has been linked to genes involved in ciliogenesis, highlighting this as an area of potential further research.[204]

## 5.2 Future work

The results regarding variants in genes known to cause PCD (NM_001198542 c.2128G>T) *HYDIN* & (NM_003777.3 c.10494G>A) *DNAH11*) present some interesting reading. If, or when, they are validated by further research, they will be incorporated in to future diagnostic testing panels to help guide molecular genetic testing/patient counselling. Additional work is required to establish exactly which molecular processes are perturbed by the mutations discussed and more individuals within the Amish community need to be screened in order establish revised control frequencies with a bigger sample size. Future work will focus on techniques that can be used to verify the highlighted variants as causative so that they can inform future diagnostic practices.

Regarding the two candidate new disease genes (*MSP* and *DIX*), the next step would then be to conduct functional studies with the aim of assessing the effect of mutations on proteins and their molecular interactions. In addition, it is important to consider tissue expression analysis during development (for example in mouse tissue block hybridisation studies), protein localisation studies to explore cellular roles, and molecular pathway analysis to further explore functional molecular complexes to improve our understanding of the biological

mechanisms. To assess the direct effect of mutation on protein interactions, we may need to consider western blotting to consider levels of proteins and their stability. Other techniques that could be explored include fluorescence activated cell sorting, surface plasmon resonance, immunoprecipitation and invitro assays. Finally, to make the dataset more robust, endogenous expression of native proteins in abnormal cell phenotypes, would strengthen the evidence surrounding supporting the labelling of a deleterious mutation as causative. As the variants in the two new candidate genes are nonsense mutations, it is unlikely that there will be any gene protein product from the mutant gene, and therefore we would expect complete loss of molecular function.

The work here is ongoing and further data analysis needs to be conducted before moving onto in-depth functional studies. This work highlights the power of new and emerging genomic technologies when combined with robust clinical phenotyping practices and a thorough understanding of the genetic make-up of a specific population.

# 6 Conclusion

## 6.1 Clinical benefits of community studies to the Amish community

This work within this study makes up a small subsection of a long-standing clinical genetic project, under the name of ‘Windows of Hope’ (WOH), based in amongst the Amish communities of Ohio. The project’s overarching aims are to determine the genetic basis of inherited disease occurring in the Amish community with the

hope of translating this information into a format that can directly benefit the study participants, their families and their wider communities, as well as the wider general population. WOH has had considerable success in achieving these aims, discovering 18 novel conditions that exist within the community, with several novel conditions undergoing further study. This has directly impacted the Amish community' with improvements and developments in molecular genetic testing and counselling, and patient management strategies.

Recently discovered disorders within the community are still sometimes poorly understood: with the carrier frequency being relatively high, many children and families in the community undergo rigorous, costly, and painful investigations, in pursuit of answers. The project aims to address this by improving awareness of the clinical phenotype and incidence of conditions, working with community and specialist healthcare providers to combat unnecessary medicalisation. The specialist knowledge acquired from genetic, molecular and further clinical study outlined in this thesis will empower patients and healthcare providers to engage with appropriate levels of medical care and follow up. Project findings are fed back promptly, at regular intervals, to the community: this ranges from family support group meetings, community liaison and patient focused disease-specific education. Meeting with local and specialist healthcare professionals are held to improve healthcare outcomes, discussing and evaluating diagnostic and management of newly identified conditions.

Technological advances in genetics have improved the efficiency of the field: the speed and breadth of diagnostic testing has facilitated the study of PCD and situs abnormalities where phenotype and genotype so often vary. Next generation sequencing has facilitated screening for multiple causative gene variants in parallel, permitting a more liberal delineation of an individual's precise phenotype. These practices, thus, have particularly utility in self-contained communities like the Amish, as causative mutations are so commonly founder mutations, facilitating the development of gene panels to be used for future screening. Further studies in this field would facilitate personalisation of these gene panels to suit church groups within the Amish, where some Mendelian conditions show varying prevalence in different communities.

Long term, the work of this project feeds into the wider aims of WOH wider aims to utilise these new technologies to develop a personalised new-born screening program, facilitating early diagnosis and tailored management protocols to improve health outcomes for the affected individuals and their families.

## 6.2 Benefits of community based studies to the wider society

The work of this project, and others like it, provide numerous and varying benefits to wider society. As discussed, Amish ancestry and endogamous marriages mean that mutation responsible for genetic conditions are enriched, making identifying such mutation more accessible. Importantly, the majority of genetic conditions identified in the Amish community have then been identified in the

wider population, underscoring the importance of Amish genetic study discoveries. Elis-van-Creveld syndrome (EVC) – a ciliopathy is an example of one such genetic disorder: mutations in EVC and EVC2 gene, result in small EVC and EVC2 proteins, respectively.[32,33]

## 6.3 Scientific benefits of community based studies

The study of rare disorders has recognised value, in terms of the insights such studies can provide into biological pathways, highlighting other potential candidate genes for study and furthering the understanding of related disorders. More recently, there are an increasing number of genes implicated in the causation of ciliopathies, specifically PCD and SI. These studies add to the body of evidence that develops our understanding of the molecules involved in cilia assembly, function and movement.

The next challenge is to use this knowledge to further investigate the complex relationship between causative genes and the resulting phenotype. This will help to establish, conclusively, new genetic causes of PCD and SI; from there, this information can be used to manage and treat the symptoms and circumvent the structural defects that lead to the poor health outcomes associated. Future genetic study holds the potential promise of introducing such therapeutic interventions and, in doing so could transform our future practice of medicine and how we manage Mendelian inherited disease.

# 7 Appendices

## 7.1 ATP Hydrolysis

**ATP hydrolysis**

ATP – or adenosine triphosphate – is the energy that cells uses to build molecules, transport tissues and facilitate enzyme activity, essential for metabolism.

ATP hydrolysis is the use of water to break down ATP into adenosine diphosphate (or ADP), a hydronium ion and phosphoryl group, in the process releasing energy. This process is usually in conjunction with another reaction it is fuelling.

Ciliary motility is dictated by the dynein's ability perform ATP hydrolysis: cilia exhibiting a high beat frequency have high dynein ATPase activity. Dynein arms providing the ATP-driven motility has implications for wider ciliary function if genetic mutations result in dysfunctional or absent dynein arms. This is particularly relevant as dynein arms activity can be modulated to change ciliary beat frequency in response to significant signalling molecules, when necessary.

## 7.2 Medical imagining in ciliopathies

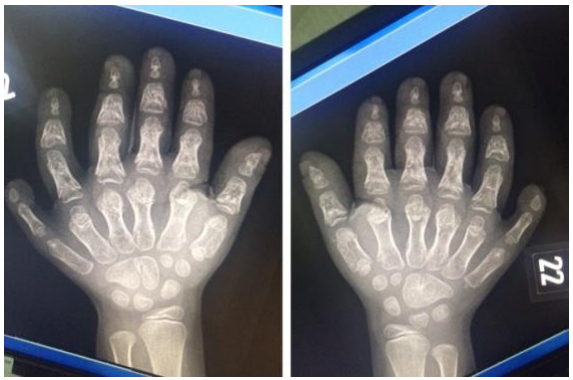


***Figure 7-1 Cone-shaped epiphyses seen on X-ray of patient with Ellis-van Creveld. Images taken from[32]***

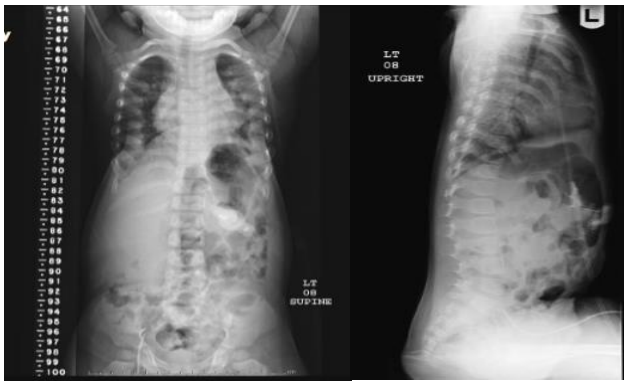



***Figure 7-2 Frontal and lateral x-ray of patient with asphyxiating thoracic dystrophy; characterised by the bell-shaped thorax seen on the left and horizontally oriented ribs. Images taken from***

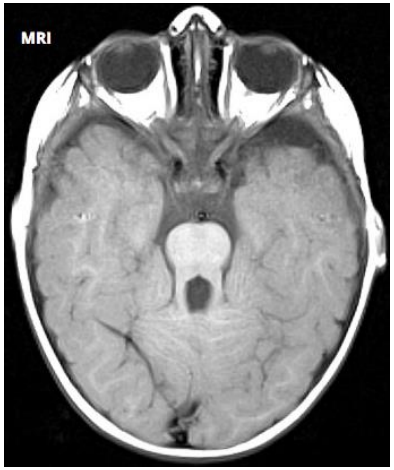



***Figure 7-3 MRI scan - Axial view - A brain showing the "molar tooth sign", pathognomonic for Joubert syndrome. Taken from[205]***

## 7.3 Components of PCD treatment

| |
|---|
| Airway clearance is achieved with daily prescribed cardiovascular exercise, percussion and use of positive expiratory pressure devices. Aims include mobilising and aiding expectoration of pulmonary secretions, improving ventilation efficiency and maintaining, or improving, exercise tolerance, with a view to reduce breathlessness and chest pain.[12] |
| Studying patient flora to identify problematic organisms that may need treatment is a method to reduce the occurrence and reoccurrence on infection: Haemophilus influenza, Streptococcus pneumoniae, Staphylococcus aureus and opportunistic pathogens, such as Pseudomonas aeruginosa, are a small selection of the infections clinicians aim to eradicated by performing airway cultures ever 3-6 months.[12] |
| Antibiotics are used to treat bacterial infections identified in airways, based on the above cultures and sensitivities. There is limited evidence on the efficacy of antibiotics in exacerbation but there are schools of thought to suggest that antibiotics regimens improve symptoms and hasten recovery, considering the use of long term us in chronic cases, rotating antibiotic cycles to avoid antimicrobial resistance.[12] |
| Extrapulmonary disease in adults, including chronic rhinitis and sinusitis, is responsible for significant morbidity but no one treatment has been deemed to be unequivocally effective. Clinicians opt for intranasal corticosteroids, sinus lavage and use of antibiotics sparingly. In extreme cases, endoscopic surgery can be used to promote drainage and better delivery of topical medications. Other extrapulmonary implications include infertility, for which assisted fertilisation of in-vitro fertilisation can be used. ~25% of those affected by PCD go on to develop respiratory failure and can be considered for a lung transplant. |

***Table 7-1 Components of PCD treatment[12]***

## 7.4 Amish migration and Settlement

| State | Amish Population in | | | | Amish population increase 1992-2017 (%) |
|---|---|---|---|---|---|
| | 1992 | 2000 | 2010 | 2017 | |
| **Pennsylvania** | 35200 | 40100 | 59350 | 59350 | 169 |
| **Ohio** | 43200 | 49750 | 5859 | 58590 | 136 |
| **Indiana** | 25200 | 32650 | 43710 | 43710 | 173 |
| **Wisconsin** | 7800 | 10250 | 15360 | 15360 | 197 |
| **New York** | 4700 | 5000 | 12015 | 12015 | 256 |
| **Michigan** | 6500 | 9300 | 11350 | 11350 | 175 |
| **Missouri** | 5200 | 6100 | 9475 | 9475 | 182 |
| **Kentucky** | 1500 | 5150 | 7750 | 7750 | 517 |

***Table 7-2 Amish population change in USA's 8 most populated states***[135,138-140]

Settlement sizes and changes in population sizes are stipulated above but it's worth noting the largest settlements within these states: Lancaster County (in south-eastern Pennsylvania), Holmes County (in north-eastern Ohio) and Elkhart and LaGrange (in north-eastern Indiana) remain the largest communities – with almost half of the population of Holmes County recorded as Amish.[141,142]

Aside from the above populous states, the Amish also reside in Oxford, Norfolk and Bruce counties in Ontario, Canada. The Amish reside in Europe, without a split between the Old Order Amish and Amish Memonites; in the 19$^{th}$ and 20$^{th}$

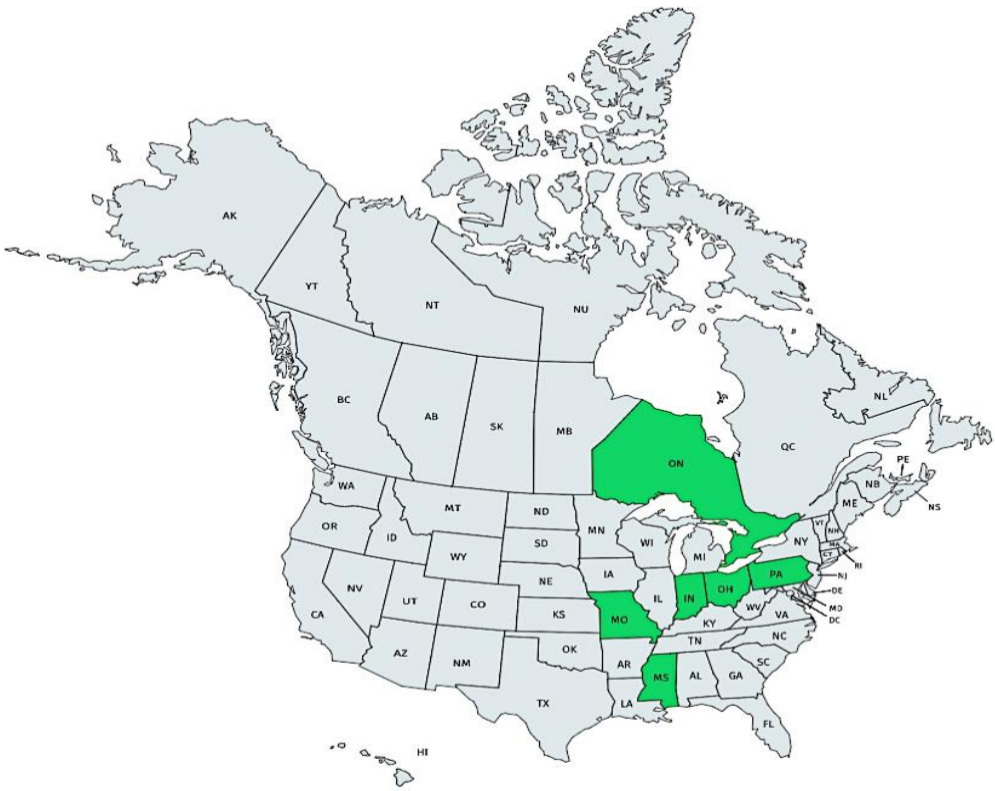



***Figure 7-4 Map of Canada and USA, denoting the region in which Amish families reside***[135,136,139-146] IN – Indiana; OH – Ohio; ON – Ontario; MO – Missouri; MS – Minnesota  PA – Pennsylvania

centuries, many communities integrated and, subsequently, many “Amish” communities have lost their separate identity and culture. The last of the European Amish churches joined the Mennonites in 1937, in Ixheim in southwest Germany.[145] The Amish have made attempts to settle in Latin America, albeit temporarily, settling in Mexico and Honduras; since then there have be resettlements in Catamarca, Argentina and Colombia Naranjita, Bolivia.[136,144,146]

## 7.5 Primer sequences

| Primer name | Forward primer | Reverse primer |
|---|---|---|
| *DNAH11* | CCGAAAATGCCGCTATCCTAA | TGGTGGAGAGGTCTAAGAAACA |
| *HYDIN* | GAGCAATACCTTGCTGTAAGTAA | CTTATGCAGGTATGTGGCTTC |
| *MSP* gene | GAAATCACGCAGGATTTTCTAC | AAACAGCATTGAGCTTAGAGAAT |
| *DIX-domain* gene | GCCTTGCAGCAGAGATTGAC | CAGACACTTAAGGCAGGCAT |

# 8 Glossary

- **ATP:** ATP – or adenosine triphosphate – is the energy that cells uses to build molecules, transport tissues and facilitate enzyme activity, essential for metabolism.[206]
- **ATP hydrolysis:** ATP hydrolysis is the use of water to break down ATP into adenosine diphosphate (or ADP), a hydronium ion and phosphoryl group, in the process releasing energy. This process is usually in conjunction with another reaction it is fuelling.[206]
- **Dynein motors:** a collection of complex proteins, made up of smaller polypeptide units, that allow a single dynein molecule to transport its cargo by travelling along the distance of a microtubule, without becoming detached.[76]
- **Nonsense mediated decay:** Nonsense-mediated mRNA decay is a surveillance pathway that reduces the errors in expression by eliminating mRNA transcripts that contain premature stop codons and subsequent proteins.[172-174]
- **Proband:** an individual that is the focus of a study/ or reported on; they are usually the first affected individual in a family being investigated for a genetic disorder.[74]
- **Wnt signalling pathway:** Wnt signalling pathways are a group of pathways, made from proteins that pass signals into a cell through cell surface receptors. There are two types: the canonical or β-catenin dependent pathway and the non-canonical or β-catenin-independent.[174-177]

# 9 Bibliography


1. Knowles M, Zariwala M, Leigh M. Primary Ciliary Dyskinesia. *Clin Chest Med.* 2016;37(3):449-461.
2. Fassad MRS, A. le Borgne, P. Koll, F. Patel, M. Dixon, M. Hayward, J. Richardson, C. Frost, E. Jenkins, L. Cullup, T. Chung, E.M.K. Lemullois, M. Aubusson-Fleury, A. Hogg, C. Mitchell, D.R., Tassin, A.M. Mitchinson, H.M. C11orf70 Mutations Disrupting the Intraflagellar Transport-Dependent Assembly of Multiple Axonemal Dyneins Cause Primary Ciliary Dyskinesia. *Am J Hum Genet.* 2018;102(5):956-972.
3. Wikipedia. Cilium. 2018; https://en.wikipedia.org/wiki/Cilium. Accessed 28/07/2018.
4. Junqueira L, Carneiro J. *Basic Histology.* 11th edition ed. New York: McGraw-Hill; 2005.
5. Lynch C. Electron Micrographs of Basal bodies. *Cytoskeleton and Cell Movement* 2016; http://slideplayer.com/slide/6319335/21/images/48/Fig+12.55+Electron+micrographs+of+basal+bodies.jpg. Accessed 27/07/2018.
6. Harris T. Structure and Function of Cilia. 2018; http://www.ciliopathyalliance.org/cilia/structure-and-function-of-cilia.html. Accessed 27/07/2018.
7. Porter ME, Sale WS. The 9 + 2 Axoneme Anchors Multiple Inner Arm Dyneins and a Network of Kinases and Phosphatases That Control Motility. *J Cell Biol.* 2000;151(5):37-42.
8. Hurd TW, Hildebrandt F. Mechanisms of Nephronophthisis and Related Ciliopathies. *Nephron Exp Nephrol.* 2011;118(1):e9-e14.
9. Keishi N, Sen T. Cilia in the choroid plexus: their roles in hydrocephalus and beyond. *Frontiers in Cellular Neuroscience.* 2015;9:1-5.
10. Lee JE, Gleeson JG. A systems-biology approach to understanding the ciliopathy disorders. *Genome Medicine.* 2011;3(9):59.
11. Fliegauf M, Olbrich H, Horvath J, et al. Mislocalization of DNAH5 and DNAH9 in Respiratory Cells from Patients with Primary Ciliary Dyskinesia. *Am J Respir Crit Care Med.* 2005;171(12):1343-1349.
12. Lobo LJ, Zariwala MA, Noone PG. *Bronchiectasis.* 2011.
13. Knowles MR, Boucher RC. Mucus clearance as a primary innate defense mechanism for mammalian airways. *J Clin Invest.* 2002;109(5):571-577.
14. Tilley AE, Walters MS, Shaykhiev R, Crystal RG. Cilia dysfunction in lung disease. *Annu Rev Physiol.* 2015;77:379-406.
15. Salathe M. Regulation of Mammalian Ciliary Beating. *Annu Rev Physiol.* 2007;69(1):401-422.
16. Satir P, Christensen ST. Structure and function of mammalian cilia. *Histochem Cell Biol.* 2008;129(6):687-693.
17. Satir P, Heuser T, Sale WS. A Structural Basis for How Motile Cilia Beat. *BioScience.* 2014;64(12):1073-1083.
18. Basudha BM, Brueckner. Cilia : Multifunctional Organelles at the Center of Vertebrate Left–Right Asymmetry. *Current Topics in Developmental Biology.* 2008;85:151-174.

19. Hirokawa N, Tanaka Y, Okada Y, Takeda S. Nodal flow and the generation of left-right asymmetry. *Cell.* 2006;125(1):33-45.
20. Lie H, Ferkol T. Primary ciliary dyskinesia: recent advances in pathogenesis, diagnosis and treatment. *Drugs.* 2007;67(13):1883-1892.
21. Damseh N, Quercia N, Rumman N, Dell SD, Kim RH. Primary ciliary dyskinesia: mechanisms and management. In: *Appl Clin Genet.* Vol 10.2017:67-74.
22. Lee JH, Gleeson JG. The role of primary cilia in neuronal function. *Neurobiology of Disease.* 2010;38(2):167-172.
23. Adams MS, U.M. Logan, C.V. Johnson, C.A. Recent advances in the molecular pathology, cell biology and genetics of ciliopathies. *J Med Genet.* 2008;45(1):257-267.
24. Harris T. Impact of Defective Cilia. 2018; http://www.ciliopathyalliance.org/cilia/impact-of-defective-cilia.html. Accessed 28/07/2018.
25. OpenStax. Anatomy & Physiology In: OpenStax; 2018: https://cnx.org/contents/FPtK1zmh@9.1:I6lVthMD@5/Introduction. Accessed 28/07/2018.
26. Harris T. Alström Syndrome. 2018; http://www.ciliopathyalliance.org/ciliopathies/alstrom-syndrome.html. Accessed 28/07/2018.
27. UK AS. What is Alström Syndrome? 2018; http://www.alstrom.org.uk/what-is/. Accessed 28/07/2018.
28. Reference GH. Alström syndrome. 2018; https://www.ncbi.nlm.nih.gov/pubmed/. Accessed 28/07/2018.
29. Marshall J, Hinman E, Collin G, et al. Spectrum of ALMS1 variants and evaluation of genotype-phenotype correlations in Alström syndrome - Marshall - 2007 - Human Mutation - Wiley Online Library. *Hum Mutat.* 2018;28(11).
30. Harris T. Bardet-Biedl Syndrome. 2018; http://www.ciliopathyalliance.org/ciliopathies/bardet-biedl-syndrome.html. Accessed 28/07/2018.
31. (OMIM) OMIiM. BARDET-BIEDL SYNDROME 1; BBS1 2018; http://omim.org/entry/209900. Accessed 28/07/2018.
32. Al-Fardan A, Al-Qattan MM. Wide-spread cone-shaped epiphyses in two Saudi siblings with Ellis-van Creveld syndrome. *International Journal of Surgery Case Reports.* 2017;39:212-217.
33. D'Asdia MCT, I. Consoli, F. Ferese, R. Magliozzi. M. Bernardini, L. Guida, V. Digilio, M.C Marino, B. Dallapiccola, B. De Luca, A. Novel and recurrent EVC and EVC2 mutations in Ellis-van Creveld syndrome and Weyers acrofacial dyostosis. *Eur J Med Genet.* 2013;56(2):80-87.
34. Center GaRDI. Jeune syndrome. 2018; https://rarediseases.info.nih.gov/diseases/3049/jeune-syndrome. Accessed 28/07/2018.
35. Harris T. Joubert Syndrome. 2018; http://www.ciliopathyalliance.org/ciliopathies/joubert-syndrome.html. Accessed 28/07/2018.
36. St GO. Joubert's syndrome. 2018; https://www.gosh.nhs.uk/conditions-and-treatments/conditions-we-treat/jouberts-syndrome. Accessed 28/07/2018.

37. Merveille A-C, Davis EE, Becker-Heck A, et al. CCDC39 is required for assembly of inner dynein arms and the dynein regulatory complex and for normal ciliary motility in humans and dogs. *Nat Genet.* 2011;43(1):72-78.
38. Castleman VH, Romio L, Chodhari R, et al. Mutations in radial spoke head protein genes RSPH9 and RSPH4A cause primary ciliary dyskinesia with central-microtubular-pair abnormalities. *Am J Hum Genet.* 2009;84(2):197-209.
39. Reference GH. Meckel syndrome. 2018; https://ghr.nlm.nih.gov/condition/meckel-syndrome. Accessed 28/07/2018.
40. Kheir AEM, Imam A, Omer IM, et al. Meckel-Gruber syndrome: A rare and lethal anomaly. *Sudan J Paediatr.* 2012;12(1):93-96.
41. Logan CV, Abdel-Hamed Z, Johnson CA. Molecular Genetics and Pathogenic Mechanisms for the Severe Ciliopathies: Insights into Neurodevelopment and Pathogenesis of Neural Tube Defects. *Molecular Neurobiology.* 2018;43(1):12-26.
42. Orphanet. Nephronophthisis. 2018; https://www.orpha.net/consor/cgi-bin/OC_Exp.php?Lng=EN&Expert=655.
43. Reference GH. Polycystic kidney disease. 2018; https://ghr.nlm.nih.gov/condition/polycystic-kidney-disease. Accessed 27/07/2018.
44. Harris T. Polycystic Kidney Disease. 2018; http://www.ciliopathyalliance.org/ciliopathies/polycystic-kidney-disease.html. Accessed 27/07/2018.
45. Borislav B. Retinitis Pigmentosa. 2018; http://www.ciliopathyalliance.org/ciliopathies/retinitis-pigmentosa.html. Accessed 28/07/2018.
46. Oishi M, Department of Ophthalmology and Visual Sciences KUGSoM, Kyoto, Japan, Oishi A, et al. Comprehensive Molecular Diagnosis of a Large Cohort of Japanese Retinitis Pigmentosa and Usher Syndrome Patients by Next-Generation Sequencing. *Investigative Ophthalmology & Visual Science.* 2017;55(11):7369-7375.
47. Senior Loken Syndrome | Genetic and Rare Diseases Information Center (GARD) â□“ an NCATS Program. 2018; https://www.ncbi.nlm.nih.gov/pubmed/.
48. Reference GH. Senior-Løken syndrome. 2018; https://www.ncbi.nlm.nih.gov/pubmed/.
49. Harris T. Usher Syndrome. 2018; http://www.ciliopathyalliance.org/ciliopathies/usher-syndrome.html. Accessed 27/07/2018.
50. Torgersen J. Transposition of viscera, bronchiectasis and nasal polyps; a genetical analysis and a contribution to the problem of constitution. *Acta radiol.* 1947;28(1):17-24.
51. Katsuhara K, Kawamoto S, Wakabayashi T, Belsky J. Situs Inversus Totalis and Kartagener's Syndrome in a Japanese Population. *Chest.* 2016;61(1).
52. Primary Ciliary Dyskinesia | National Heart, Lung, and Blood Institute. 2018; https://www.nhlbi.nih.gov/health-topics/primary-ciliary-dyskinesia. Accessed 27/07/2017.
53. Accpchest. Primary Ciliary Dyskinesia (PCD) | CHEST Foundation. 2018; https://foundation.chestnet.org/patient-education-resources/primary-ciliary-dyskinesia-pcd/. Accessed 27/07/2017.

54. Noone PG, Leigh MW, Sannuti A, et al. Primary ciliary dyskinesia: diagnostic and phenotypic features. *Am J Respir Crit Care Med.* 2004;169(4):459-467.
55. Lobo LJ, Zariwala MA, Noone PG. Primary ciliary dyskinesia. *QJM.* 2014;107(9):691-699.
56. ENT Issues: PCD Support. 2018; http://pcdsupport.org.uk/what-is-pcd/general-pcd-information/ent-issues/. Accessed 27/07/2018.
57. PCD and the Nose â□“ PCD Support. 2018; http://pcdsupport.org.uk/what-is-pcd/general-pcd-information/pcd-and-the-nose/.
58. Rungha R, Jha RK. Situs inversus incomplitus: A diagnostic dilemma. *JIACM.* 2016;17(1):64-65.
59. Pennekamp P, Menchen T, Dworniczak B, Hamada H. Situs inversus and ciliary abnormalities: 20 years later, what is the connection? *Cilia.* 2015;4:1.
60. Blyth MW, D. Ectopic pregnancy in primary ciliary dyskinesia. *Journal of Obstetrics and Gynecology.* 2008;28(3):358.
61. Leigh MW. Primary Ciliary Dyskinesia. 2012; https://www.sciencedirect.com/topics/medicine-and-dentistry/primary-ciliary-dyskinesia. Accessed 28/07/2018.
62. Reference GH. Primary ciliary dyskinesia. 2018; https://ghr.nlm.nih.gov/condition/primary-ciliary-dyskinesia. Accessed 27/07/2018.
63. Lucas JS, Burgess A, Mitchison HM, et al. Diagnosis and management of primary ciliary dyskinesia. *Arch Dis Child.* 2014;99(9):850-856.
64. Collins S, Behan L, Evans H, et al. The accuracy of nasal nitric oxide testing in PCD diagnostics is population specific. 2015.
65. Shapiro AJ, Josephson M, Rosenfeld M, et al. Accuracy of Nasal Nitric Oxide Measurement as a Diagnostic Test for Primary Ciliary Dyskinesia. A Systematic Review and Meta-analysis. *Ann Am Thorac Soc.* 2017;14(7):1184-1196.
66. PCD diagnostic service. 2018; http://www.uhs.nhs.uk/OurServices/Childhealth/Primary-ciliary-dyskinesia/PCD-diagnostic-service.aspx. Accessed 28/07/2018.
67. Lucas JS, Walker WT. Nasal Nitric Oxide Is an Important Test in the Diagnostic Pathway for Primary Ciliary Dyskinesia. *Ann Am Thorac Soc.* 2013;10(6):645-647.
68. Davis SD, Ferkol TW, Rosenfeld M, et al. Clinical Features of Childhood Primary Ciliary Dyskinesia by Genotype and Ultrastructural Phenotype. *Am J Respir Crit Care Med.* 2015;191(3):316-324.
69. Gene Dossiers. 2018; https://ukgtn.nhs.uk/find-a-test/gene-dossiers/. Accessed 27/07/2018.
70. Dixon A. Normal CT chest | Radiology Case | Radiopaedia.org. 2018; https://radiopaedia.org/cases/normal-ct-chest. Accessed 28/07/2018.
71. Radiopaedia. Radiology Quiz 12819 | Radiopaedia.org. 2018; https://radiopaedia.org/cases/12819/studies/12940?referrer=%2Farticles%2Fprimary-ciliary-dyskinesia%23image_list_item_631727. Accessed 28/07/2018.
72. Walker WT, Jackson CL, Lackie PM, Hogg C, Lucas JS. Nitric oxide in primary ciliary dyskinesia. *Eur Respir J.* 2012;40(4):1024-1032.
73. Leigh MW, Hazucha MJ, Chawla KK, et al. Standardizing Nasal Nitric Oxide Measurement as a Test for Primary Ciliary Dyskinesia. *Ann Am Thorac Soc.* 2013;10(6):574-581.

74. Institute NHGR. Proband. *Talking Glossary of Genetic Terms* 2018; https://www.genome.gov/glossary/index.cfm?id=164. Accessed 28/07/2018.
75. Olcese C, Patel MP, Shoemark A, et al. X-linked primary ciliary dyskinesia due to mutations in the cytoplasmic axonemal dynein assembly factor PIH1D3. *Nature Communications.* 2017;8:14279.
76. McKenney RJ, Huynh W, Tanenbaum ME, Bhabha G, Vale RD. Activation of cytoplasmic dynein motility by dynactin-cargo adapter complexes. *Science.* 2014;345(6194):337-341.
77. Zariwala M, Noone P, Sannuti A, et al. Germline Mutations in an Intermediate Chain Dynein Cause Primary
Ciliary Dyskinesia. *Am J Respir Cell Mol Biol.* 2001;25.
78. Pennarun G, Escudier E, Chapelin C, et al. Loss-of-Function Mutations in a Human Gene Related to Chlamydomonas reinhardtii Dynein IC78 Result in Primary Ciliary Dyskinesia. In: *Am J Hum Genet.* Vol 65.1999:1508-1519.
79. Guichard C, Harricane M-C, Lafitte J-J, et al. Axonemal Dynein Intermediate-Chain Gene (DNAI1) Mutations Result in Situs Inversus and Primary Ciliary Dyskinesia (Kartagener Syndrome). *American Journal of Human Genetics.* 2001;68(4):1030-1035.
80. Failly M, Saitta A, Munoz A, et al. DNAI1 mutations explain only 2% of primary ciliary dykinesia. *Respiration.* 2008;76(2):198-204.
81. Zariwala MA, Leigh MW, Ceppa F, et al. Mutations of DNAI1 in Primary Ciliary Dyskinesia: Evidence of Founder Effect in a Common Mutation. In: *Am J Respir Crit Care Med.* Vol 174.2006:858-866.
82. Mirra V, Werner C, Santamaria F. Primary Ciliary Dyskinesia: An Update on Clinical Aspects, Genetics, Diagnosis, and Future Treatment Strategies. *Front Pediatr.* 2017;5.
83. Pennarun GC, Catherine. Escudier, Estelle. Bridoux, Anne Marie. Dastot, Florence. Cacheux, V, alère. Goossens MA, Serge. Duriez, Bénédicte. The human dynein intermediate chain 2 gene ( DNAI2): cloning, mapping, expression pattern, and evaluation as a candidate for primary ciliary dyskinesia | SpringerLink. *Human Genetics.* 2000;107(6):642-649.
84. Loges NT, Olbrich H, Fenske L, et al. DNAI2 mutations cause primary ciliary dyskinesia with defects in the outer dynein arm. *American Journal of Human Genetics.* 2008;83(5):547-558.
85. Failly M, Bartoloni L, Letourneau A, et al. Mutations in DNAH5 account for only 15% of a non-preselected cohort of patients with primary ciliary dyskinesia. 2009.
86. Online Mendelian Inheritance in Man (OMIM). 2018; https://www.omim.org/entry/244400. Accessed 28/07/2018.
87. Duriez B, Duquesnoy P, Escudier E, et al. A common variant in combination with a nonsense mutation in a member of the thioredoxin family causes primary ciliary dyskinesia. *Proc Natl Acad Sci U S A.* 2007;104(9):3336-3341.
88. Omran H, Kobayashi D, Olbrich H, et al. Ktu/PF13 is required for cytoplasmic pre-assembly of axonemal dyneins. *Nature.* 2008;456(7222):611-616.
89. Loges NT, Olbrich H, Becker-Heck A, et al. Deletions and Point Mutations of LRRC50 Cause Primary Ciliary Dyskinesia Due to Dynein Arm Defects. *Am J Hum Genet.* 2009;85(6):883-889.
90. Duquesnoy P, Escudier E, Vincensini L, et al. Loss-of-Function Mutations in the Human Ortholog of Chlamydomonas reinhardtii ODA7 Disrupt Dynein Arm

Assembly and Cause Primary Ciliary Dyskinesia. *Am J Hum Genet.* 2009;85(6):890-896.
91. Sullivan-Brown J, Schottenfeld J, Okabe N, et al. Zebrafish mutations affecting cilia motility share similar cystic phenotypes and suggest a mechanism of cyst formation that differs from pkd2 morphants. *Dev Biol.* 2008;314(2):261-275.
92. Merveille AC, Davis EE, Becker-Heck A, et al. CCDC39 is required for assembly of inner dynein arms and the dynein regulatory complex and for normal ciliary motility in humans and dogs. *Nat Genet.* 2011;43(1):72-78.
93. Becker-Heck A, Zohn I, Okabe N, et al. The coiled-coil domain containing protein CCDC40 is essential for motile cilia function and left-right axis formation. *Nat Genet.* 2011;43(1):79-84.
94. Castleman VH, Romio L, Chodhari R, et al. Mutations in Radial Spoke Head Protein Genes RSPH9 and RSPH4A Cause Primary Ciliary Dyskinesia with Central-Microtubular-Pair Abnormalities. *Am J Hum Genet.* 2009;84(2):197-209.
95. Diggle CP, Moore DJ, Mali G, et al. HEATR2 Plays a Conserved Role in Assembly of the Ciliary Motile Apparatus. *PLoS Genet.* 2014;10(9).
96. Horani A, Druley Todd E, Zariwala Maimoona A, et al. Whole-Exome Capture and Sequencing Identifies HEATR2 Mutation as a Cause of Primary Ciliary Dyskinesia. *Am J Hum Genet.* 2012;91(4):685-693.
97. Mitchison HM, Schmidts M, Loges NT, et al. Mutations in axonemal dynein assembly factor DNAAF3 cause primary ciliary dyskinesia. *Nat Genet.* 2012;44(4):381-389, S381-382.
98. Tarkar A, Loges NT, Slagle CE, et al. DYX1C1 is required for axonemal dynein assembly and ciliary motility. *Nat Genet.* 2013;45(9):995-1003.
99. Chandrasekar GL, V. Hultenby, K. Tapia-Páez, I. Kere, J. . The Zebrafish Orthologue of the Dyslexia Candidate Gene DYX1C1 Is Essential for Cilia Growth and Function. *PLoS ONE.* 2013;8(5):e63123.
100. Panizzi JR, Becker-Heck A, Castleman VH, et al. CCDC103 mutations cause primary ciliary dyskinesia by disrupting assembly of ciliary dynein arms. *Nat Genet.* 2012;44(6):714-719.
101. King SM, Patel-King RS. The Oligomeric Outer Dynein Arm Assembly Factor CCDC103 Is Tightly Integrated within the Ciliary Axoneme and Exhibits Periodic Binding to Microtubules*. *J Biol Chem.* 2015;290(12):7388-7401.
102. Kott E, Duquesnoy P, Copin B, et al. Loss-of-Function Mutations in LRRC6, a Gene Essential for Proper Axonemal Assembly of Inner and Outer Dynein Arms, Cause Primary Ciliary Dyskinesia. *Am J Hum Genet.* 2012;91(5):958-964.
103. Horani A, Ferkol TW, Shoseyov D, et al. LRRC6 Mutation Causes Primary Ciliary Dyskinesia with Dynein Arm Defects. *PLoS ONE.* 2013;8(3).
104. Transport of the outer dynein arm complex to cilia requires a cytoplasmic protein Lrrc6 - Inaba - 2016 - Genes to Cells - Wiley Online Library. 2018.
105. Paff T, Loges NT, Aprea I, et al. Mutations in PIH1D3 Cause X-Linked Primary Ciliary Dyskinesia with Outer and Inner Dynein Arm Defects. *Am J Hum Genet.* 2017;100(1):160-168.
106. Horani A, Ferkol TW, Dutcher SK, Brody SL. Genetics and Biology of Primary Ciliary Dyskinesia. *Paediatr Respir Rev.* 2016;18:18-24.
107. Hjeij R, Onoufriadis A, Watson C, et al. CCDC151 Mutations Cause Primary Ciliary Dyskinesia by Disruption of the Outer Dynein Arm Docking Complex Formation. *Am J Hum Genet.* 2014;95(3):257-274.

108. Hjeij R, Lindstrand A, Francis R, et al. ARMC4 Mutations Cause Primary Ciliary Dyskinesia with Randomization of Left/Right Body Asymmetry. *Am J Hum Genet.* 2013;93(2):357-367.
109. Onoufriadis A, Shoemark A, Munye MM, et al. Combined exome and whole-genome sequencing identifies mutations in ARMC4 as a cause of primary ciliary dyskinesia with defects in the outer dynein arm. *J Med Genet.* 2014;51(1):61-67.
110. Kurkowiak M, Ziętkiewicz E, Witt M. Recent advances in primary ciliary dyskinesia genetics. *J Med Genet.* 2015;52(1):1-9.
111. Bower R, Tritschler D, VanderWaal K, et al. The N-DRC forms a conserved biochemical complex that maintains outer doublet alignment and limits microtubule sliding in motile axonemes. In: *Mol Biol Cell.* Vol 24.2013:1134-1152.
112. Lin J, Tritschler D, Song K, et al. Building Blocks of the Nexin-Dynein Regulatory Complex in Chlamydomonas Flagella. *Journal of Biological Chemistry.* 2011;286(33):29175-29191.
113. Onoufriadis A, Paff T, Antony D, et al. Splice-Site Mutations in the Axonemal Outer Dynein Arm Docking Complex Gene CCDC114 Cause Primary Ciliary Dyskinesia. *Am J Hum Genet.* 2013;92(1):88-98.
114. Slavotinek A, Biesecker LG. Genetic modifiers in human development and malformation syndromes, including chaperone proteins. *Human Molecular Genetics.* 2018;12(suppl_1).
115. Lee L. Riding the wave of ependymal cilia: Genetic susceptibility to hydrocephalus in primary ciliary dyskinesia. *Journal of Neuroscience Research.* 2013;91(9).
116. Moore A, Escudier E, Roger G, et al. RPGR is mutated in patients with a complex X linked phenotype combining primary ciliary dyskinesia and retinitis pigmentosa. *J Med Genet.* 2006;43(4):326-333.
117. Zariwala MA, Knowles MR, Omran H. Genetic Defects in Ciliary Structure and Function. *Annu Rev Physiol.* 2007;69(1):423-450.
118. Reference GH. OFD1 gene. 2018; https://www.ncbi.nlm.nih.gov/pubmed/. Accessed 28/07/2018.
119. Rubbo B, Lucas JS. Clinical care for primary ciliary dyskinesia: current challenges and future directions. *European Respiratory Review.* 2017;26(170023).
120. Nall R. Situs Inversus: Definition and Patient Education. 2018; https://www.healthline.com/health/situs-inversus. Accessed 28/07/2018.
121. Radeopaedia. Situs Inversus - Radiology Quiz 7446 2018; https://radiopaedia.org/cases/7446/studies/8270?referrer=%2Farticles%2Fsitus-inversus%23image_list_item_131159
https://radiopaedia.org/articles/situs-inversus. Accessed 28/07/2018.
122. Situs inversus. In. *Wikipedia*2017.
123. Situs inversus | Genetic and Rare Diseases Information Center (GARD) â□“ an NCATS Program. 2018; https://www.ncbi.nlm.nih.gov/pubmed/.
124. Bass KD, Presbyterian St Luke's Columbia Medical Center D, CO, USA., The Children's Hospital D, CO, USA., et al. Laparoscopic Ladd's procedure in infants with malrotation. *Journal of Pediatric Surgery.* 1998;33(2):279-281.
125. Fulcher AS, Turner MA. Abdominal Manifestations of Situs Anomalies in Adults.
126. Gedda L, Sciacca A, Brenci G, et al. Situs viscerum specularis in monozygotic twins. *Acta geneticae medicae et gemellologiae.* 1984;33(1):81-85.

127. Dextrocardia with Situs Inversus. 2018; http://www.malacards.org/card/dextrocardia_with_situs_inversus. Accessed 28/07/2018.
128. OrphaNet. Situs inversus totalis. 2018; https://www.orpha.net/consor/cgi-bin/Disease_Search.php?lng=EN&data_id=14774&Disease_Disease_Search_diseaseGroup=situs-inversus&Disease_Disease_Search_diseaseType=Pat&Disease(s)/group%20of%20diseases=Situs-inversus-totalis&title=Situs-inversus-totalis&search=Disease_Search_Simple. Accessed 28/07/2018.
129. OrphaNet. Heterotaxia. 2018; https://www.orpha.net/consor/cgi-bin/Disease_Search.php?lng=EN&data_id=8556&Disease_Disease_Search_diseaseGroup=Heterotaxia&Disease_Disease_Search_diseaseType=Pat&Disease(s)/group%20of%20diseases=Heterotaxia&title=Heterotaxia&search=Disease_Search_Simple. Accessed 28/07/2018.
130. Sparrow DB, Faqeih EA, Sallout B, et al. Mutation of HES7 in a large extended family with spondylocostal dysostosis and dextrocardia with situs inversus - Sparrow - 2013 - American Journal of Medical Genetics Part A - Wiley Online Library. *Am J Med Genet.* 2013;161(9).
131. OrphaNet. Situs ambiguus. 2018; https://www.orpha.net/consor/cgi-bin/Disease_Search.php?lng=EN&data_id=17142&Disease_Disease_Search_diseaseGroup=Situs-ambiguus&Disease_Disease_Search_diseaseType=Pat&Disease(s)/group%20of%20diseases=Situs-ambiguus&title=Situs-ambiguus&search=Disease_Search_Simple. Accessed 28/07/2018.
132. Rugina ALD, A.G. Nistor, N. Mihaila, D. Primary ciliary dyskinesia diagnosed by electron microscopy
in one case of Kartagener syndrome. *Romainian Journal of Morphology & Embryology.*
133. Leadbeater V. Reversed Heart, Situs Inversus and Dextrocardia. In. *The ID Band Company Blog*2015.
134. Baple E. *Investigation of the molecular basis of three new disorders of brain growth and development*
*identified amongst the Amish*. Exeter: Medical Studies, University of Exeter; 2014.
135. Kraybill D. *The Riddle of The Amish Culture.* 2001.
136. Nolt S. *A History of the AMish.* Good Books; 1992.
137. Kraybill D, Johnson-Weiner K, Nolt S. *The Amish.* John Hopkins University Press; 2013.
138. Hostetler JA. *Amish Society.* Baltimore1993.
139. Amish Population Change 2012-2017. In: College YCfAaPS-E, ed.
140. Bodies AoSAR. U.S. Religion Census | Religious Statistics & Demographics. 2018; http://www.usreligioncensus.org/. Accessed 26/07/2018, 2018.
141. Hurst C, McConnell D. *An Amish Paradox.* Johns Hopkins University Press; 2010.
142. Twelve largest settlements, 2017. 2013; http://groups.etown.edu/amishstudies/statistics/largest-settlements/.
143. Hostetler J. *Amish Society.* Baltimore1993.
144. Anderson C, Anderson J. The Amish Settlement in Honduras, 1968-1978 *Journal of Amish and Plain Anabaptist Studies.* 2016;4(1):1-50.

145. Online GAME. Ixheim (Rheinland-Pfalz, Germany) 1957; http://gameo.org/index.php?title=Ixheim_(Rheinland-Pfalz,_Germany)&oldid=145494. Accessed 2018 July 26, 2018.
146. America A. 2016 Amish Population: Two New Settlements In South America. 2018; http://amishamerica.com/2016-amish-population/. Accessed 26/07/2018, 2018.
147. Kraybill D, Olshan M. *Amish Struggle with Modernity.* University Press of New England; 1994.
148. Baple E. *Investigation of the molecular basis of three new disorders of brain growth and development identified amongst the Amish*. Exeter: Medical Studies, University of Exeter; 2014.
149. Baple E. *Investigation of the molecular basis of three new disorders of brain growth and development identified amongst the Amish*. Exeter, University of Exeter; 2014.
150. Bowman HS, Mc Kusick VA, Dronamraju KR. Pyruvate Kinase Deficient Hemolytic Anemia in an Amish Isolate. *Am J Hum Genet.* 1965;17(1):1-8.
151. Cross HE, McKusick VA. The Troyer Syndrome: A Recessive Form of Spastic Paraplegia With Distal Muscle Wasting. *Archives of Neurology.* 2018;16(5):473-485.
152. Cross HE, McKusick VA. A survey of neurological disorders in a genetic isolate. *Neurology.* 1967;17(8 Pt 1):743-751.
153. McKusick VA. Mendelian Inheritance in Man and Its Online Version, OMIM. *Am J Hum Genet.* 2007;80(4):588-604.
154. Cross HE, McKusick VA. The Mast Syndrome: A Recessively Inherited Form of Presenile Dementia With Motor Disturbances. *Archives of Neurology.* 2018;16(1):1-13.
155. McKusick VA, Egeland JA, Eldridge R, Krusen DE. Dwarfism in the Amish: The Ellis-van Creveld Syndrome. *Bulletin of the Johns Hopkins Hospital.* 1964;115:306-336.
156. Alkuraya F. The application of next-generation sequencing in the autozygosity mapping of human recessive diseases. *Hum Genetics.* 2013;132(11):1197–1211.
157. Bottlenecks and founder effects. 2018.
158. Ploshnitsa AI, Goltsman ME, Macdonald DW, Kennedy LJ, Sommer S. Impact of historical founder effects and a recent bottleneck on MHC variability in Commander Arctic foxes (Vulpes lagopus). *Ecol Evol.* 2012;2(1):165-180.
159. Ruiz-Perez VL, Ide SE, Strom TM, et al. Mutations in a new gene in Ellis-van Creveld syndrome and Weyers acrodental dysostosis. *Nature Genetics.* 2000;24(3):283.
160. Amselem S, Duquesonoy P, Attree O, et al. Laron dwarfism and mutations of growth hormone receptor gene. *New England Journal of Medicine.* 1989(321):989-995
161. Lander ES, Bostein D. Homozygosity mapping: a way to map human recessive trains with DNA of inbred children. *Science.* 1987(236):1567-1570.
162. Carr IB, S. O'Sullivan, J. Aldahmesh, M. Shamseldin, H. Markham, A. Bonthron, D. Black, G. Alkuraya, F. Autozygosity Mapping with Exome Sequence Data *Hum Mutat.* 2013;34(1).

163. DePristo M, Banks E, Poplin R, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet.* 2011;43(5):491-498.
164. Ferkol TW, Puffenberger EG, Lie H, et al. Primary Ciliary Dyskinesia-Causing Mutations in Amish and Mennonite Communities. *J Pediatr.* 2013;163(2):383-387.
165. Nakana TB, L.G. . No evidence for triallelic inheritance of MKKS/BBS loci in Amish Mckusick–Kaufman syndrome - Nakane - 2005 - American Journal of Medical Genetics Part A - Wiley Online Library. *Am J Med Genet.* 2018;138A(1).
166. Jinks RN, 1 Department of Biology and Biological Foundations of Behaviour Program FaMC, Lancaster, PA 17604, USA, Puffenberger EG, et al. Recessive nephrocerebellar syndrome on the Galloway-Mowat syndrome spectrum is caused by homozygous protein-truncating mutations of WDR73. *Brain.* 2018;138(8):2173-2190.
167. Schwartz PJ, Brink PA. Of Founder Populations, Long QT Syndrome, and Destiny. *Heart Rhythm.* 2009;6(11 Suppl):S25-33.
168. Promega. Kit Components: ReliaPrep™ Blood gDNA Miniprep System,
100 preps2017.
169. Promega. Technical manual: ReliaPrep™ Blood gDNA Miniprep System. In:2012.
170. Ensembl genome browser 92. 2018. https://www.ensembl.org/index.html. Accessed 28/07/2018.
171. Research WIfB. Primer3 Input (version 0.4.0). 2018; http://bioinfo.ut.ee/primer3-0.4.0/primer3/. Accessed 28/07/2018.
172. UCSC. UCSC Genome Browser Home. 2018; http://genome.ucsc.edu/index.html. Accessed 28/07/2018.
173. Baker KEP, R. Nonsense-mediated mRNA decay: terminating erroneous gene expression. *Current Opinion in Cell Biology.* 2004;16(3):293-299.
174. Chang Y-F, Imam JS, Wilkinson MF. The Nonsense-Mediated Decay RNA Surveillance Pathway. *Annual Review of Biochemistry.* 2007;76:51-74.
175. Shiomi K, Uchida H, Keino-Masu K, Masu M. Ccd1, a novel protein with a DIX domain, is a positive regulator in the Wnt signaling during zebrafish neural patterning. *Current biology : CB.* 2003;13(1):73-77.
176. Zhan T, Rindtorff N, Boutros M. Wnt signaling in cancer. *Oncogene.* 2016;36(11):1461.
177. Database GHG. HYDIN Gene - GeneCards | HYDIN Protein | HYDIN Antibody. 2018; https://www.genecards.org/cgi-bin/carddisp.pl?gene=HYDIN. Accessed 28/07/2018.
178. UniProtKB. Hydrocephalus-inducing protein homolog - Homo sapiens (Human) - HYDIN gene & protein. 2018; https://www.uniprot.org/uniprot/Q4G0P3. Accessed 28/07/2018.
179. RGD. Hydin (Hydin, axonemal central pair apparatus protein) - Rat Genome Database. 2018.
180. Proteopedia. Solution structure of the C-terminal PapD-like domain from human HYDIN protein. 2018; http://proteopedia.org/wiki/index.php/2e6j. Accessed 28/07/2018.
181. Lechtreck K-F, Delmotte P, Robinson ML, Sanderson MJ, Witman GB. Mutations in Hydin impair ciliary motility in mice. 2008.

182. Chilvers MA, Rutman A, O'Callaghan C. Functional analysis of cilia and ciliated epithelial ultrastructure in healthy children and young adults. *Thorax.* 2003;58(4):333-338.
183. Stannard W, Rutman A, Wallis C, O'Callaghan C. Central microtubular agenesis causing primary ciliary dyskinesia. *Am J Respir Crit Care Med.* 2004;169(5):634-637.
184. Carlén B, Stenram U. Primary Ciliary Dyskinesia: A Review. *Ultrastructural Pathology.* 2005;29(3-4):217-220.
185. Sturgess JMC, J. Wong, J. Aspin, N. Turner, P. Cilia with Defective Radial Spokes — A Cause of Human Respiratory Disease | NEJM. *N Engl J Med.* 1979(300):53-56.
186. Tamalet A, Clement A, Roudot-Thoraval F, et al. Abnormal Central Complex Is a Marker of Severity in the Presence of Partial Ciliary Defect. 2001.
187. Olbrich H, Schmidts M, Werner C, et al. Recessive HYDIN Mutations Cause Primary Ciliary Dyskinesia without Randomization of Left-Right Body Asymmetry. *Am J Hum Genet.* 2012;91(4):672-684.
188. Lechtreck KF, Witman GB. Chlamydomonas reinhardtii hydin is a central pair protein required for flagellar motility. *J Cell Biol.* 2007;176(4):473-482.
189. Lechtreck KF, Delmotte P, Robinson ML, Sanderson MJ, Witman GB. Mutations in Hydin impair ciliary motility in mice. *J Cell Biol.* 2008;180(3):633-643.
190. UniProtKB. Dynein heavy chain 11, axonemal - Homo sapiens (Human) - DNAH11 gene & protein. 2018; https://www.uniprot.org/uniprot/Q96DT5. Accessed 28/07/2018.
191. Gene: DNAH11 (ENSG00000105877) - Summary - Homo sapiens - Ensembl genome browser 92. 2018; https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000105877;r=7:21543215-21901839. Accessed 28/07/2018.
192. Bartoloni L, Blouin J-L, Pan Y, et al. Mutations in the DNAH11 (axonemal heavy chain dynein type 11) gene cause one form of situs inversus totalis and most likely primary ciliary dyskinesia. *Proceedings of the National Academy of Sciences of the United States of America.* 2002;99(16):10282-10286.
193. Komiya Y, Habas R. Wnt signal transduction pathways. *Organogenesis.* 2008;4(2):68-75.
194. NUSSE R. Wnt signaling in disease and in development. *Cell Research.* 2005;15(1):28.
195. Zariwala M, Gee H, Kurkowiak M, et al. ZMYND10 Is Mutated in Primary Ciliary Dyskinesia and Interacts with LRRC6. *Am J Hum Genet.* 2013;93(2):336-345.
196. Germino GG. Linking cilia to Wnts. *Nat Genet.* 2005;37(5):455.
197. UniProtKB. Interstitial collagenase precursor - Homo sapiens (Human) - MMP1 gene & protein. 2018; https://www.uniprot.org/uniprot/P03956. Accessed 28/07/2018.
198. Stankovic M, Nikolic A, Divac A, et al. Matrix Metalloproteinases Gene Variants in Idiopathic Disseminated Bronchiectasis. *Journal of Investigative Medicine.* 2009;57(3).
199. Tsang K, Bilton D. Clinical challenges in managing bronchiectasis. *Clinical Journal of the Asian Pacific Society of Respirology.* 2009(14):637-650.
200. OMIM Entry - * 612598 - RING FINGER PROTEIN 11; RNF11. 2018.

201. UniProtKB. Ankyrin repeat and SOCS box protein 18 - Homo sapiens (Human) - ASB18 gene & protein. 2018; https://www.uniprot.org/uniprot/Q6ZVZ8. Accessed 28/07/2018.
202. Kohroki JN, T. Nakamura, T. Masuho, Y. ASB proteins interact with Cullin5 and Rbx2 to form E3 ubiquitin ligase complexes - Kohroki - 2005 - FEBS Letters - Wiley Online Library. *FEBS Press.* 2018;579(30).
203. Identification of transcripts overexpressed during airway epithelium differentiation. *European Respiratory Journal.* 2008;32(1):121-128.
204. UBA52 26 items (Homo sapiens) *STRING network view* 2017; http://version10.string-db.org/cgi/network.pl?taskId=ugjgOxWhknjR. Accessed 28/07/2018.
205. Gaillard F. Molar tooth sign (CNS) - Radiology Reference Article. 2018; https://radiopaedia.org/articles/molar-tooth-sign-cns-1 -image_list_item_246454. Accessed 26/07/2018.
206. Academy K. Overview of ATP hydrolysis. 2018; https://www.khanacademy.org/test-prep/mcat/biomolecules/overview-metabolism/a/overview-of-atp-hydrolysis. Accessed 26/07/2018.