Supplementary methods

Ethics

Local and national research and ethical approvals were obtained and adhered to (NRES Committee South Central Hampshire Ethics 06/Q1702/109, London Bloomsbury Research Ethics Committee 08/H0713/82 and Ile-de-France Ethics Committee CPP07729).

Genetics

Patients were screened by the next generation and Sanger sequencing methods summarised. Genetic analysis was evaluated by geneticists and clinicians specialised in PCD, with a confirmed genetic diagnosis defined as the presence of autosomal bi-allelic or single X-linked hemizygous variants classified as pathogenic according to international guidelines [1, 2]. Of 292 genetically screened patients in the discovery cohort, using these criteria we confirmed a genetic diagnosis in 199 patients. We excluded 93 patients carrying variants judged to be of uncertain significance, which included single variants in PCD genes predicted pathogenic/likely pathogenic but without a second variant identified; variants identified in candidate rather than known PCD genes; and variants of uncertain pathogenic effect for example if TEM data inconsistent.

Discovery and validation cohorts

The discovery group consisted of PCD patients from University Hospital Southampton (UHS) and the Royal Brompton Hospital (RBH), London, genotyped at University College London (UCL). Clinical and diagnostic data were collected retrospectively from electronic and paper-based medical records for all patients with a conclusive genetic result available up to July 2017. The

validation group consisted of patients genotyped from UHS and RBH between July 2017 and May 2019, and at Trousseau, Cochin and Creteil hospitals in France and Emma Children's Hospital in the Netherlands up to May 2019. Study data were collected applying the definitions according to the study coding protocol. Ciliary beat pattern and TEM were reviewed by specialists, blinded to all genetic data.

The phenotypic data collected from both validation and discovery cohorts and used for clustering were based upon 12 clinical and diagnostic variables: BMI, FEV_1 z-score, FVC z-score, NRDS, wet cough, rhinitis, glue ear, cardiac situs, CHD, nNO, CBP and TEM, as described in the main manuscript. Data found not to shape the model during development were excluded; this include age at diagnosis, height, weight, mutation type and ethnicity. Additional data were collected on clinical and diagnostic characteristics (see Table E1) but were not included in the modelling; these were used to explore the model. Each variable was used to colour the nodes by the categories detailed in Table E1 to further explore potential clusters of phenotypic data.

Ciliary beat pattern was described and categorised according to the predominant finding from the following terms: normal, completely immotile, weak residual movement, stiff, rotating, staggered beat, lack of cilia. Transmission electron microscopy was categorised as one of the following terms: non diagnostic, isolated ODA defect, ODA & IDA defect, MTD & IDA defect or isolated IDA defect, CC defect or lack of cilia.

Topological data analysis (TDA)

Topology is a branch of applied mathematics that is primarily concerned with the study of shape of data and is specifically designed to identify structural characteristics of high-dimensional

datasets. TDA [3] consists of a set of techniques for data analyses based on the reproduction of the structure of complex datasets into a geometric shape, that captures the essential features similarly to how a topographical map captures features of a landscape. It does so by dividing (or binning) the dataset through the application of a distance metric (e.g. a similarity measure) and then performing clustering within each of those separate segments. These are then visually represented as nodes of a network, each of which correspond to a collection of datapoints. TDA does not produce distinct clusters as traditional clustering techniques do but rather a network where points are connected depending on (dis)similarity between combination of the features of variables included in the model. In Symphony AyasdiAI, a user-friendly software that combines TDA with machine learning, different colours can be applied to the nodes of the network using any of the metrics or variables in the dataset, in order to inspect the data for patterns and hotspots.

TDA is an unsupervised data-driven technique, with no prior hypothesis needed. The outcome of interest should not be included in the clustering, which in this study were the genetic data. After the models were developed, we inspected the data by colouring the nodes by the different genes in order to identify any clusters or hotspots that would require further interrogation.

Machine learning was used in the lenses that were applied to our model. These lenses only provide the visualisation of the network through the application of a layout algorithm and therefore do not influence the clustering itself. In order to construct the topological models, we applied a variety of lenses. Lenses can be derived from statistical measures such as mean, from geometry such as centrality, from dimensionality-reduction techniques such as principal component analysis (PCA), or even from a variable in the dataset. After exploring several different lenses, we selected locally linear embedding (LLE) lenses as the most relevant to our dataset because they showed distinct clusters for further exploration. Similarly, we selected correlation as a metric after evaluating other

metrics. Correlation seemed an appropriate choice due to the differences of variance between variables, and the various categorical variables included in our dataset.

TDA deals with missing values individually; where they are missing, the TDA network will be mapped without that data point for that individual for that specific variable. The individual will still be plotted into the network according to similarities in variables for which there are data. For example, if an individual has values for BMI, NRDS, wet cough, rhinitis, glue ear, cardiac situs, CHD, nNO, CBP and TEM but data were missing for FEV_1 and FVC due to the age of the patient, this patient would be clustered in the TDA network according to similarities in the other ten variables for which we had measurements. This will have no effect on whether a patient clusters with patients that have similar values for these ten variables, they simply will not be clustered according to FEV_1 and FVC.

Additionally, TDA is highly robust in handling missing data, as has been shown in the literature [see references quoted in main paper] and also in a white paper by Glushakov et al [4]. In their study, the authors intentionally deleted values from their dataset in order to test the robustness of the TDA approach and found that the topological models were geometrically stable even when 90% of data were missing. We are therefore confident that missing data in our datasets did not affect the shape or clustering in the topological models.

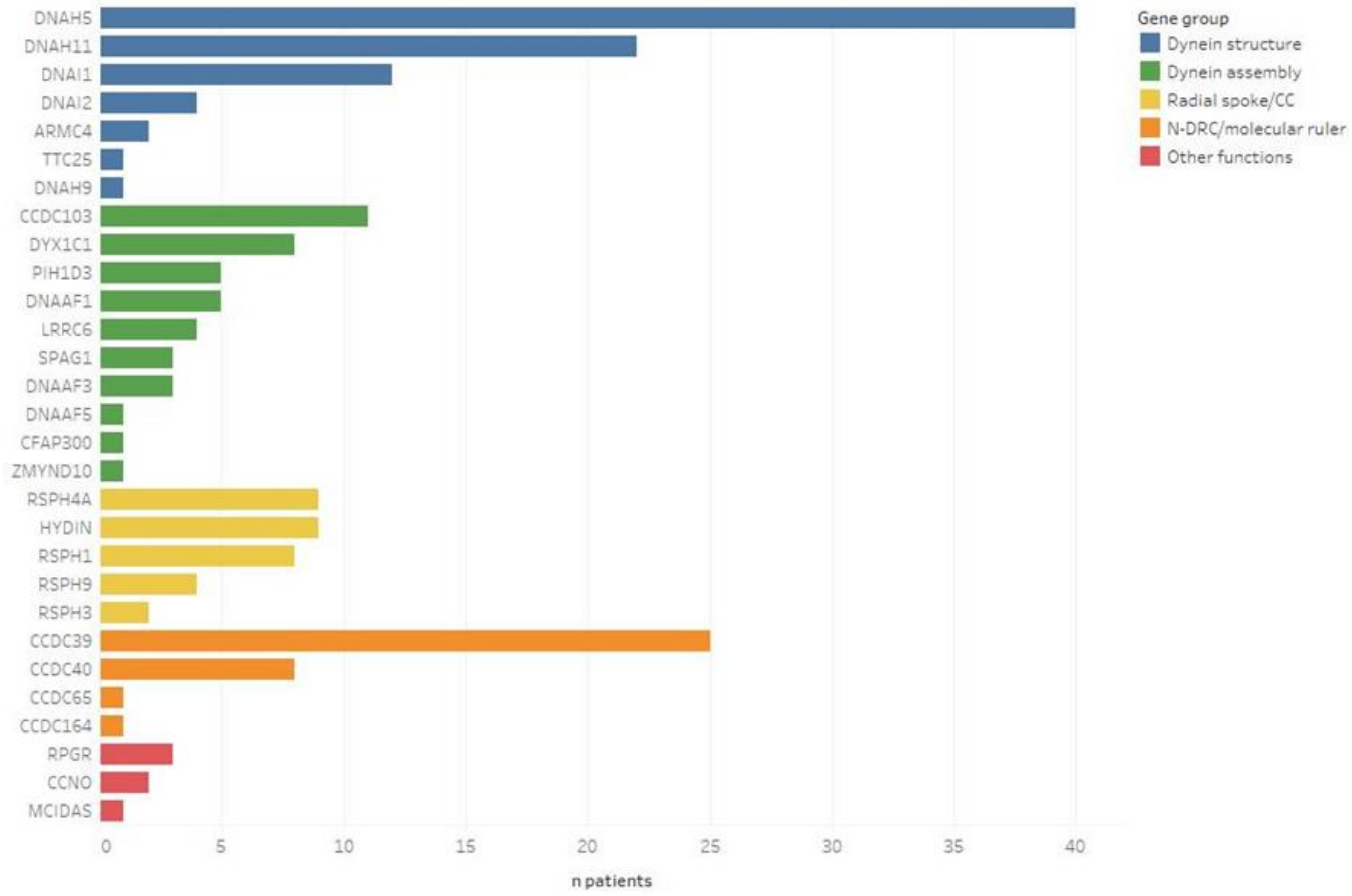
Summary of clinical diagnostic methods by group

Method	University Hospital Southampton	Royal Brompton London	Hôpital Trousseau, Paris	Hôpital Intercommunal Créteil	Hôpital Kremlin-Bicêtre, Le Kremlin-Bicêtre	Hôpital Cochin, Paris	Amsterdam UMC
Genotyping	**Discovery group:** DNA extracted from blood using salting out technique and stored in -20 until further use. Next-generation sequencing performed as previously described [5], either by whole exome sequencing (WES) or targeted gene panel sequencing (Illumina TruSeq Custom Amplicon, Agilent SureSelect Focused Exome or SureSelectXT custom panel), including all known PCD genes and other candidate genes, on an Illumina platform. Variant analysis used an in-house bioinformatics pipeline similar to [6] with variant confirmation by Sanger sequencing with parental segregation. **Validation Group:** Wessex Clinical Exome analysis using the llumina TruSight One Sequencing Panel; 29 PCD gene panel applied to NGS sequence data. Confirmation by Sanger sequencing with parental segregation.	DNA extracted from blood using salting out technique and stored in -20 until further use. Next-generation sequencing performed as previously described [5], either by whole exome sequencing (WES) or targeted gene panel sequencing (Illumina TruSeq Custom Amplicon, Agilent SureSelect Focused Exome or SureSelectXT custom panel), including all known PCD genes and other candidate genes, on an Illumina platform. Variant analysis used an in-house bioinformatics pipeline similar to [6] with variant confirmation by Sanger sequencing with parental segregation. For a number of patients, variants were identified by candidate gene Sanger sequencing.	Genomic DNA was extracted from whole blood (EDTA sampling) either with the Maxwell 16 IVD device (Promega) or with a FlexiGene kit (Qiagen). DNA was analysed by a targeted capture panel (SeqCap EZ Choice, Roche Diagnostics) including all the known PCD genes and candidate genes. Librairies were sequenced on a MiSeq sequencer (Illumina). Data was analysed with a in-house double pipeline base on Bwa and Bowtie. Sequencing depth of the regions of interest was over 50X. DNA from relatives and control samples from the probands were analysed by Sanger sequencing (BigDye v3.1, Life Technologies) on a 3130XL sequencer (Life Technologies).	Performed in Hôpital Trousseau, Paris	Performed in Hôpital Trousseau, Paris	Performed in Hôpital Trousseau, Paris	DNA extracted from blood using a Chemogen robot and stored in -20 until further use. DNA sequencing was done using whole exome sequencing with targeted analysis, including all known PCD genes. Enriched libraries were sequenced with the HiSeq or Nextseq platforms (Illumina, San Diego, CA) as paired-end 100 bp reads. Sequencing reads were cleaned by 5′-end quality trimming and Illumina-adapter clipping by Trimmomatic. Prealignment quality control of the cleaned sequencing reads was done with FastQC. Clean reads were mapped to reference genome hg19 (GRCh37) using BWA-MEM. The genome analysis toolkit was used for recalibrating quality scores, realignment around indels, marking PCR duplicates, and variant calling and variants were annotated with ANNOVAR. All

							mutations were confirmed by Sanger sequencing. The analysis of the sequencing data was done using in-house bioinformatics pipeline. Sequencing and data analysis done at the Department of Clinical Genetics, Amsterdam UMC, Vrije Universiteit Amsterdam.
Nasal nitric oxide analysis	Ecomedics CLD 88 Exhalyzer; exhalation against resistance; sampling 0.33 l/min	Logan LR5000 Chemiluminscence Analyser (Rochester Kent); breath hold sampling 0.25 l/min	NIOX Flex up to 2014 ; From 2014 up to now CLD 88 sp Ecophysics chemiluminescence NO analyser ; sampling flow rate of 0.3 $L.min^{-1}$; measurements during breathhold, expiration against resistance and tidal breathing	EVA4000 chemiluminescent analyzer (Seres, France) ; breath hold sampling 1.3 l/min followed ATS/ERS standards	FeNO+ medisoft biochemical analyser (Sorinnes, Belgium), NO nasal at a sample flow rate of 100ml/s through a nasal catheter, breathing through resistance for velum closing	Chemiluminescence Analyser (EndoNO 8000®, SERES, Aix-en-Provence, France), breath hold analysis 1.3 l/min	Niox vero, exhalation against resistance, sampling 0.33l/min
Electron microscope	60,000x magnification (minimum) by Hitachi H7000; 100-300 cilia were imaged in transverse section for assessment of axonemal structure. Quantitative analysis determined ciliary ultrastructure.	60,000x magnification (minimum) by Hitachi H7000; 100-300 cilia were imaged in transverse section for assessment of axonemal structure. Quantitative analysis determined ciliary ultrastructure.	Performed in Hôpital Intercommunal Créteil	Analyses were carried out in the Pathology Department, in collaboration with the ICM-QUANT platform (Institut du Cerveau et de la Moelle Epinière, Paris). 80,000x magnification (minimum) by Hitachi HT7700; at least 100 cilia were imaged in transverse section for assessment of axonemal structure. Quantitative analysis determined ciliary ultrastructure.	Performed in Hôpital Intercommunal Créteil	Performed in Hôpital Intercommunal Créteil	The samples were screened by using a transmission electron microscope (Tecnai $12G^2$, FEI Company) and minimal 20 images of representative cross-sections were taken with a VELETA side-entry camera at a magnification of at least x60.000.
High-speed video microscopy equipment	0.5 mm coverwell imaging chamber (Sigma-Aldrich, Poole, UK) mounted onto a glass slide; Olympus IX71 inverted microscope and	0.5 mm coverwell imaging chamber (Sigma-Aldrich, Poole, UK) mounted onto a glass slide; Leica DM-LB upright microscope with x100 oil plan	Glass slide with coverslip; Nikon Eclipse Ci upright microscope with x100 oil plan objective lens; room temperature; PL-A741 high-speed video	Performed in Hôpital Trousseau, Paris	Performed in Hôpital Trousseau, Paris	Performed in Hôpital Trousseau, Paris	0.5 mm coverwell imaging chamber (Sigma-Aldrich, Poole, UK) mounted onto a glass slide; Zeiss AX10 Observer.A1 inverted microscope and

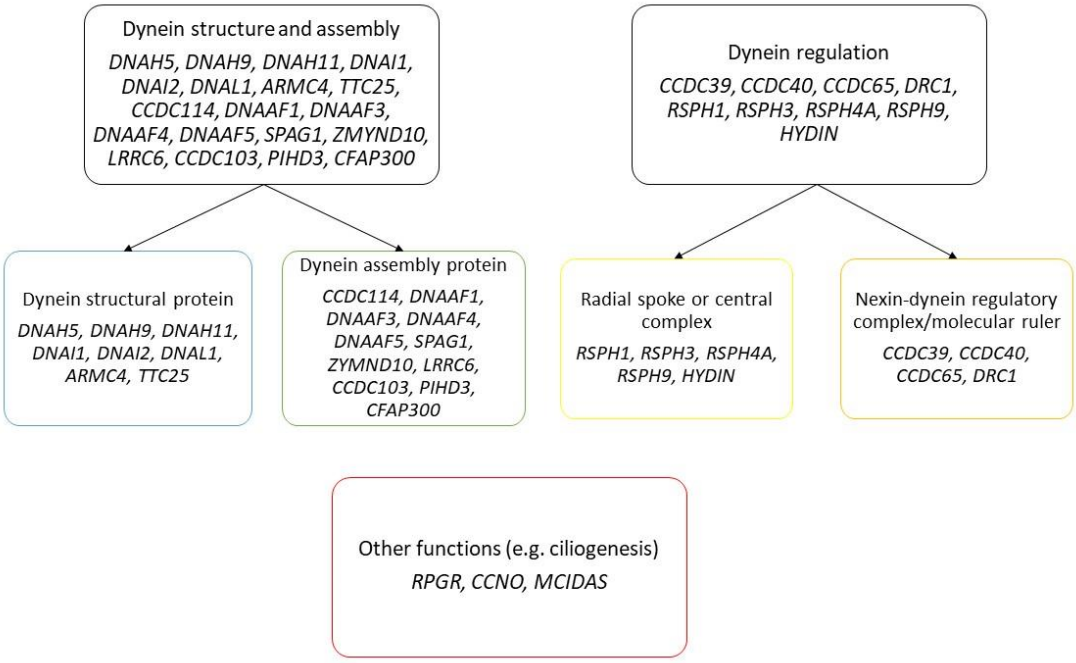
	condenser;x100 UPlan wide aperture oil objective; 37ºC heated environmental chamber (Solent Scientific, Southampton, UK); Photron FASTCAM MC2 high-speed video digital camera and Photron software.	objective lens; 37ºC heated stage; anti-vibration table (Wentworth Laboratories Ltd, Sandy, UK); Troubleshooter TS-5 Fastec imaging.	digital camera (PixeLINK, Ottawa, Canada).				condenser;Basler aVA1000 High-speed video digital camera and Strempix software.
High-speed video microscopy analyses	Images were digitally recorded using a high-speed camera at a rate of 500 frames per second (fps) and reviewed at reduced frame rates (30-60 fps) for analysis of ciliary beat pattern (CBP) and ciliary beat frequency (CBF).	Images were digitally recorded using a high-speed camera at a rate of 500 frames per second (fps) and reviewed at reduced frame rates (30-60 fps) for analysis of ciliary beat pattern (CBP) and ciliary beat frequency (CBF).	Images were digitally recorded using a high-speed camera at a rate of 355 frames per second (fps). Each movie was composed of 1,800 frames with a definition of 256 x 192 pixels (pixel size: 0.13 x 0.13 μm^2); twenty distinct areas containing intact undisrupted ciliated epithelial edges greater than 50 µm were recorded for analysis of ciliary beat pattern (CBP) and ciliary beat frequency (CBF).	Performed in Hôpital Trousseau, Paris	Performed in Hôpital Trousseau, Paris	Performed in Hôpital Trousseau, Paris	Images were digitally recorded using a high-speed camera at a rate of 120 frames per second (fps) and reviewed at reduced frame rates (10-20 fps) for analysis of ciliary beat pattern (CBP) and ciliary beat frequency (CBF).
Spirometry	FEV_1 on day of diagnostic testing, or first available result. Followed ATS/ERS Standards.	FEV_1 on day of diagnostic testing, or first available result. Followed ATS/ERS Standards.	FEV_1 on day of diagnostic testing, or first available result. Followed ATS/ERS Standards.	FEV_1 on day of diagnostic testing, or first available result. Followed ATS/ERS Standards.	FEV_1 on day of diagnostic testing, or first available result. Followed ATS/ERS Standards.	FEV_1 on day of diagnostic testing, or first available result. Followed ATS/ERS Standards. Jaeger MasterScreen Body (CAREFUSION, Hoechberg, Germany).	FEV_1 on day of diagnostic testing, or first available result. Followed ATS/ERS Standards.

Figure E1. Genetic results in 197 PCD patients from the *validation* group.



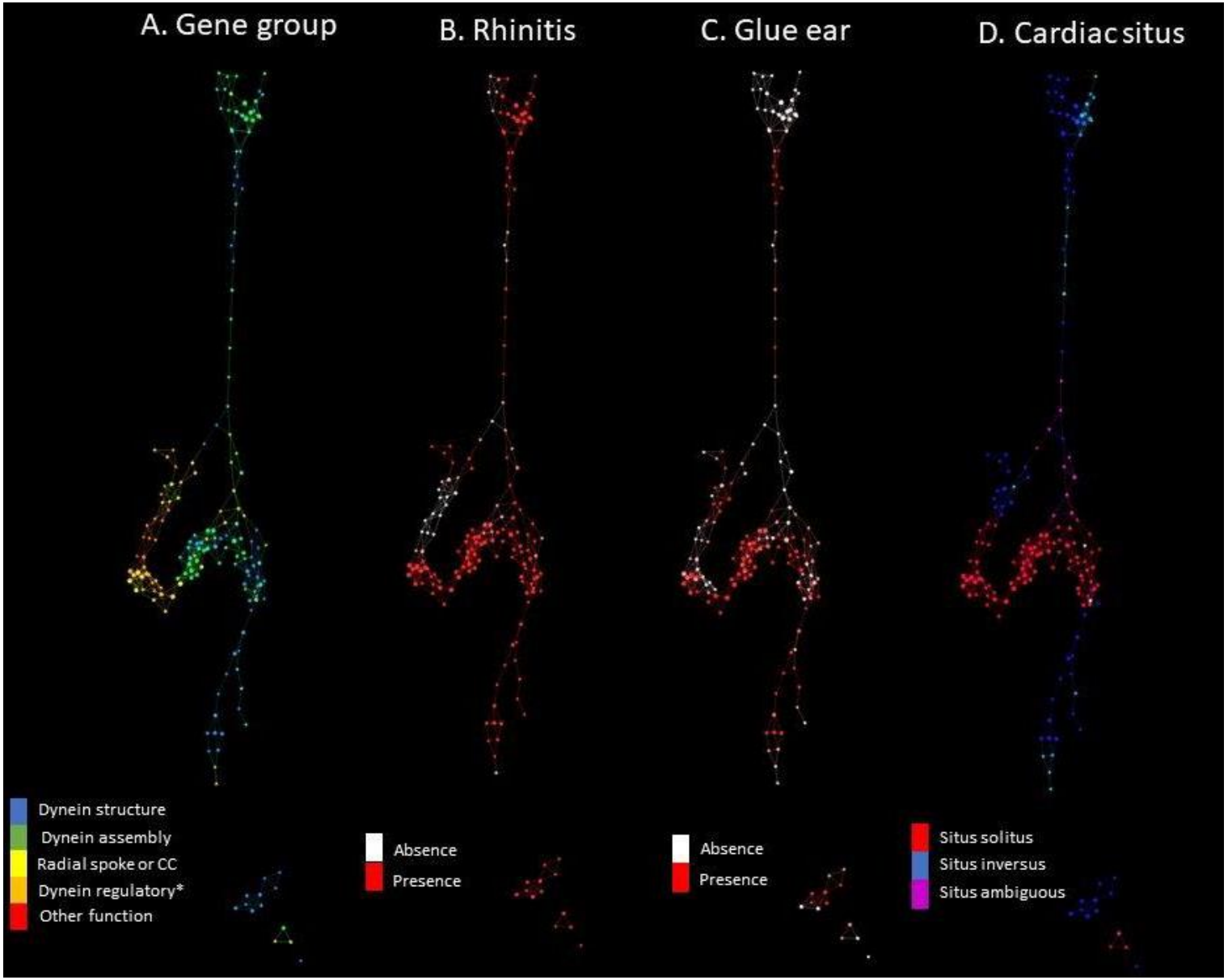
Patients in the validation cohort all had a confirmed clinical genetic diagnosis based upon PCD clinical experts identifying pathogenic or likely pathogenic variants, using identical diagnostic criteria for variant classification to that used for the discovery cohort (data not shown).

Figure E2. Stratification of all 31 PCD-causative genes in the overall study cohorts, placed into functional gene groups according to the ciliary components they encode.



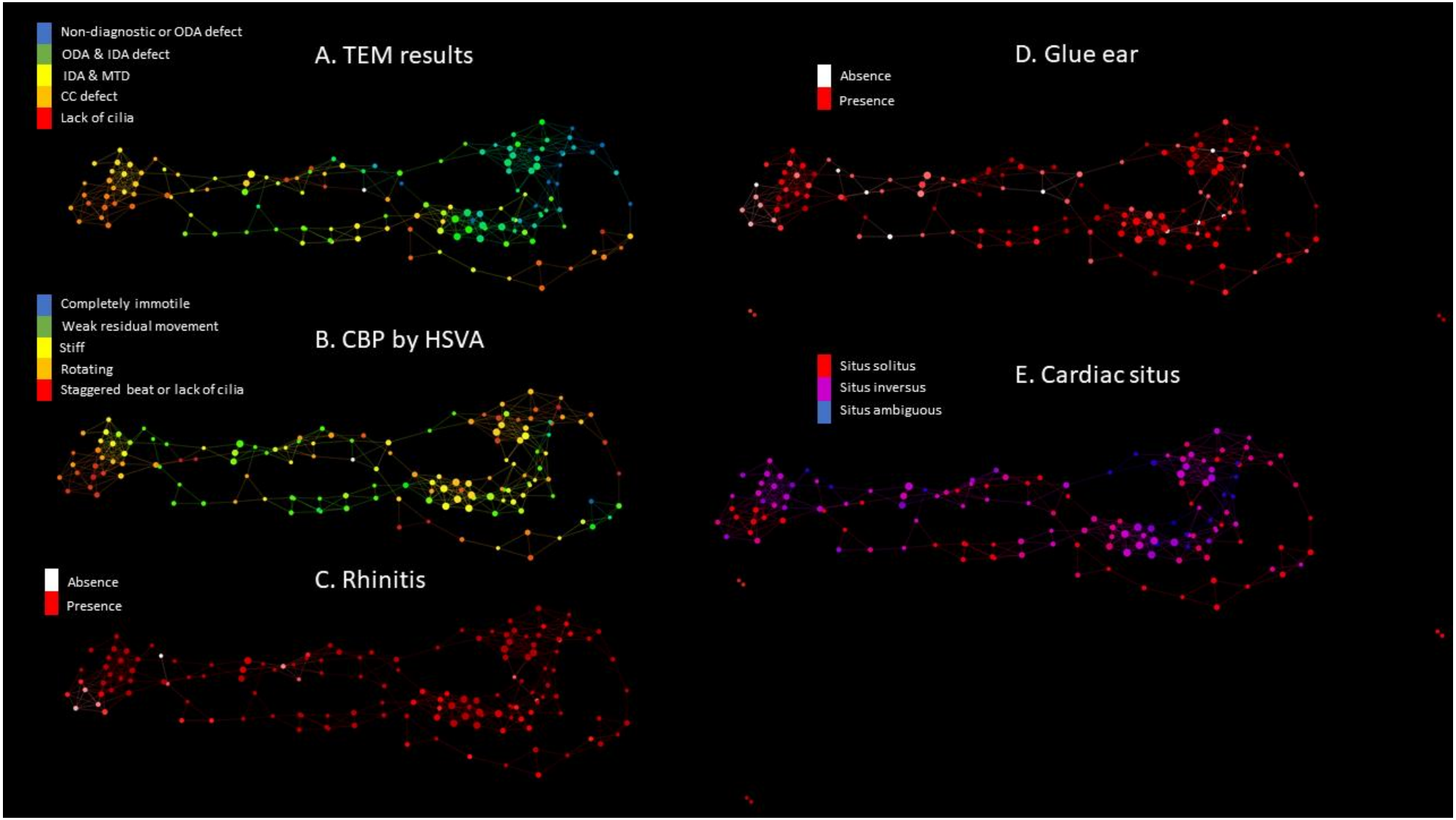
Each box defines a group: dynein structural protein, dynein assembly protein, radial spoke or central complex, dynein regulatory/molecular ruler, and other functions. Colours represent the gene groups: blue for genes involved in dynein structure, green in dynein assembly, yellow in radial spoke and central complex, orange in nexin-dynein regulatory complex/molecular ruler, and red in other functions.

Figure E3. Topological discovery model.



Models A-D are coloured by the following features: A. Gene group; B. History of rhinitis; C. History of glue ear; D. Cardiac situs. Node size represents the number of subjects. Each node represents combinations of features, connections represent that there are patients shared between the two nodes.

Figure E4. Topological validation model.



Models A-E are coloured by the following features: A. Transmission electron microscopy (TEM) findings; B. Ciliary beat pattern (CBP) by high-speed video analysis (HSVA); C. Rhinitis; D. Glue ear; E. Cardiac situs. Node size represents the number of subjects. Each node represents combinations of features, connections represent that there are patients shared between the two nodes.

Table E1. Description of data coding for clinical characteristics included in the study

Clinical characteristic	Description
Study ID	Unique ID (e.g. 0X-XXX)
DOB	
Date of LF (lung function) test	Closest to age at diagnosis Date format (e.g. DD-MM-YYYY)
Gender	Male = 1 Female = 2
Consanguinity	Up to 3rd degree cousins No = 0 Yes = 1
Number of siblings with PCD	Siblings with confirmed PCD
Ethnicity	Global Lung Function Initiative categories [7]
Weight	in kg
Height	in cm
BMI	Calculate BMI z-scores [8]
FEV_1	in litres. Calculate FEV_1 z-scores [7]
FVC	in litres
Date of diagnosis	Date format (e.g. DD-MM-YYYY)
Age at diagnosis	in years (1 decimal point)
Neonatal respiratory distress syndrome	Present Absent Unknown
History of wet cough	Present Absent Unknown
History of rhinitis	Present Absent Unknown
History of glue ear	Present Absent Unknown
Cardiac situs	Levocardia Dextrocardia Not applicable
Situs inversus totalis	Yes No Unknown
Echo done?	Yes No
Echo normal?	Yes No
Cardiac anatomy normal according to investigations?	Yes No Not applicable
Echo details, if abnormal	Free text
Abd USG	Performed Not performed
Abd USG normal?	Yes No Not applicable
nNO	in nL/min

Gene	Free text
Mutation	Free text
Transmission electron microscopy	
n cilia counted for arms	Calculate % of cilia with dynein arms
n cilia counted for microtubules	Calculate % of cilia with microtubules present
Both arms present	Calculate %
Inner arms missing	Calculate %
Outer arms missing	Calculate %
Both arms missing	Calculate %
Microtubular arrangement normal 9+2	Calculate %
Microtubules dis-arranged	Calculate %
Extra tubule	Calculate %
Single tubule	Calculate %
Central pair transposition	Calculate %
One of the central pair missing	Calculate %
Both central pair missing	Calculate %
Compound	Calculate %
TEM defect	Normal ODA IDA I&ODA IDA&MTD MTD Central complex defect Lack of cilia Inconclusive Note done
If ODA or I&ODA only please select if ODA is predominantly	Present Truncated Absent
If ODA or I&ODA only please select if ODA is present but not predominant	Present Truncated Absent
If ODA or I&ODA only please select if ODA is present but not predominant	Present Truncated Absent
High-speed video analysis	
CBP side view predominant finding	Normal Completely immotile Weak residual movement Stiff Rotating Staggered beat Long with bulbous tips Lack of cilia Not done
CBP present but not predominant 1	Normal Completely immotile Weak residual movement Stiff Rotating Staggered beat Long with bulbous tips Lack of cilia Not done
CBP present but not predominant 2	Normal Completely immotile

	Weak residual movement Stiff Rotating Staggered beat Long with bulbous tips Lack of cilia Not done
CBP present but not predominant 3	Normal Completely immotile Weak residual movement Stiff Rotating Staggered beat Long with bulbous tips Lack of cilia Not done
CBP present but not predominant 4	Normal Completely immotile Weak residual movement Stiff Rotating Staggered beat Long with bulbous tips Lack of cilia Not done
CBP present but not predominant 5	Normal Completely immotile Weak residual movement Stiff Rotating Staggered beat Long with bulbous tips Lack of cilia Not done
CBP present but not predominant 6	Normal Completely immotile Weak residual movement Stiff Rotating Staggered beat Long with bulbous tips Lack of cilia Not done
CBP top view predominant finding	Normal Completely immotile Weak residual movement Stiff Rotating Staggered beat Long with bulbous tips Lack of cilia Not done
CBP top present but not predominant 1	Normal Completely immotile Weak residual movement Stiff Rotating

	Staggered beat Long with bulbous tips Lack of cilia Not done
If stiff report location	Apical Basal Global
Syncronisation of CBP present	Yes No Not applicable
CBF	in Hz/min
Comments	Free text

Table E2. Diagnostic characteristics of patients in the *discovery* group, stratified by predefined gene groups. Genes are ordered according to gene distribution in the study population.

Diagnostic characteristic	**Dynein structure** (*DNAH5, DNAH11, DNAI1, ARMC4, DNAI2, DNAL1*) (n=89)	**Dynein assembly** (*CCDC103, DNAAF3, LRRC6, DNAAF4, SPAG1, ZYMND10, DNAAF1, CCDC114, PIHD3)* (n=52)	**Radial spoke or central complex** *(RSPH4A, RSPH9, RSPH1, HYDIN)* (n=18)	**N-RC/molecular ruler** *(CCDC40, CCDC39, DRC1, CCDC164)* (n=33)	**Other functions** *(RPGR, CCNO, MCIDAS)* (n=7)	**All**	**p-value**
Median nNO level in nL/min (IQR); n=149	11.0 (6.8 to 18.8)*	17.8 (7.8 to 33.6)	23.0 (11.0 to 34.2)	12.6 (5.4 to 18.8)	39.9 (15.3 to 96.9)*	13.0 (7.4 to 24.0)	**0.0071**
TEM findings, n=187							
Non-diagnostic TEM (%)	25 (30.1)	4 (7.7)	4 (23.5)	2 (6.9)	2 (33.3)	37 (19.8)	
Isolated ODA defect (%)	51 (61.5)	9 (17.3)	0	0	0	60 (32.1)	
ODA & IDA defect (%)	6 (7.2)	34 (65.4)	0	0	0	40 (21.4)	
MTD & IDA defect or isolated IDA defect (%)	0	4 (7.7)	0	27 (93.1)	0	31 (16.6)	
CC defect (%)	0	0	13 (76.5)	0	0	13 (7.0)	
Lack of cilia (%)	1 (1.2)	1 (1.9)	0	0	4 (66.7)	6 (3.2)	
CBP predominant side view, n=133							
Normal (%)	0	3 (8.8)	0	0	0	3 (2.7)	
Completely immotile (%)	35 (58.3)	25 (73.5)	0	5 (22.7)	3 (42.9)	68 (51.1)	
Weak residual movement (%)	8 (13.3)	0	0	1 (4.6)	0	9 (6.8)	
Stiff (%)	16 (26.7)	6 (17.7)	3 (30.0)	11 (50.0)	2 (28.6)	38 (28.6%	
Rotating (%)	0	0	7 (70.0)	0	0	7 (5.3)	
Staggered beat (%)	0	0	0	5 (22.7)	0	5 (3.8)	
Lack of cilia (%)	1 (1.7)	0	0	0	2 (28.6)	3 (2.3)	

*nNO= nasal nitric oxide (normal levels <77nl/min), TEM = Transmission electron microscopy, ODA= outer dynein arm, IDA = inner dynein arm, CC = central complex, CBP= ciliary beat pattern, ODA= outer dynein arm, IDA= inner dynein arm, MTD= microtubular disorganisation; * = significant difference between the pairs, Dunn's pairwise comparison with Holm-Sidak adjustment. P values $\leq$0.05 highlighted.

Table E3. Clinical characteristics of patients in the *discovery* group, stratified by predefined gene groups. Genes are ordered according to gene distribution in the study population.

Clinical characteristic	**Dynein structure** *(DNAH5, DNAH11, DNAI1, ARMC4, DNAI2, DNAL1)* (n=89)	**Dynein assembly** *(CCDC103, DNAAF3, LRRC6, DNAAF4, SPAG1, ZYMND10, DNAAF1, CCDC114 PIHD3)* (n=52)	**Radial spoke/ central complex** *(RSPH4A, RSPH9, RSPH1, HYDIN)* (n=18)	**N-DRC/molecular ruler** *(CCDC40, CCDC39, CCDC65, DRC1)* (n=33)	**Other functions** *(RPGR, CCNO, MCIDAS)* (n=7)	**All**	**p-value**
Male (%)	34 (38.2)	27 (51.9)	9 (50.0)	12 (36.4)	5 (71.4)	87 (43.7)	0.226
Ethnicity (n=191)							
White-British (%)	50 (58.1)	7 (13.5)	3 (16.7)	14 (50.0)	1 (14.3)	75 (39.3)	
White Irish (%)	0	5 (9.6)	3 (16.7)	1 (3.6)	4 (57.1)	13 (6.8)	
White-other (%)	10 (11.6)	4 (7.7)	1 (1.6)	5 (17.9)	1 (14.3)	21 (11.0)	
Indian (%)	4 (4.7)	5 (9.6)	0	1 (3.6)	0	10 (5.3)	
Pakistani (%)	6 (7.0)	18 (34.6)	3 (16.7)	2 (7.1)	0	29 (15.2)	
Bangladeshi (%)	0	2 (3.9)	1 (5.6)	0	0	3 (1.6)	
Sri Lankan (%)	3 (3.5)	2 (3.9)	0	0	0	5 (2.6)	
Middle East (%)	1 (1.2)	1 (1.9)	5 (27.8)	1 (3.6)	0	8 (4.2)	
Black (%)	7 (8.1)	0	1 (5.6)	1 (3.6)	0	9 (4.7)	
Chinese (%)	0	3 (5.8)	0	0	0	3 (1.6)	
Mixed (%)	1 (1.2)	1 (1.9)	0	1 (3.6)	0	3 (1.6)	
Other (%)	4 (4.7)	4 (7.9)	1 (5.6)	2 (7.1)	1 (14.3)	12 (6.3)	
Mean FEV_1 z-scores (SD), n=138	-1.4 (1.4) $^{+}$	-1.9 (1.4)	-1.7 (2.1)	-2.7 (1.6)$^{+}$	-2.7 (2.7)	-1.8 (1.6)	**0.0069**
Median age at diagnosis (IQR) n=184	9.1 (2.0 to 23.2)	7.3 (2.3 to 12.5)	9.5 (8.4 to 15.4)	7.5 2.0 to 13.8)	10.2 (5.8 to 12.7)	9.0 (2.9 to 15.4)	0.667
Neonatal respiratory distress (%)	31 (54.4)	31 (88.6)	7 (63.6)	15 (65.2)	3 (42.9)	87 (65.4)	**0.006**
Wet cough (%)	66 (94.3)	40 (100)	14 (100)	25 (96.2)	5 (71.4)	150 (95.5)	**0.042**
Rhinitis (%)	65 (91.6)	38 (95.0)	11 (91.7)	18 (72.0)	5 (71.4)	137 (88.4)	**0.027**
Glue ear (%)	38 (57.6)	19 (51.4)	9 (81.8)	9 (39.1)	4 (57.1)	79 (54.9)	0.206
Situs solitus (%)	31 (37.8)	19 (37.3)	18 (100)	18 (58.1)	7 (100%)	93 (49.2)	**<0.001**

$^{+}$ difference between groups was statistically significant (ANOVA followed by Tukey for pairwise comparisons). P values $\leq$0.05 highlighted.

Table E4. Variants defined in 199 PCD patients from the *discovery* cohort.

Pt ID	Gene	Allele 1	Allele 2	Variant classification		Reference	
				Allele 1	Allele 2	A1	A2
01-205	*ARMC4 (NM_001290020.1)*	c.1233_1234delinsT, p.Leu411Phefs*48	c.1969C>T, p.Gln657*	Frameshift (5)	Nonsense (5)	NA	[9]
01-214	*ARMC4 (NM_001290020.1)*	c.1283C>G, p.Ser428*	c.1283C>G, p.Ser428*	Nonsense (5)	Nonsense (5)	NA	NA
01-072	*ARMC4 (NM_001290020.1)*	c.2675C>A, p.Ser892*	c.2675C>A, p.Ser892*	Nonsense (5)	Nonsense (5)	[10]	[10]
01-073	*ARMC4 (NM_001290020.1)*	c.2675C>A, p.Ser892*	c.2675C>A, p.Ser892*	Nonsense (5)	Nonsense (5)	[10]	[10]
01-098	*CCDC103 (NM_001258395.1)*	c.461A>C, p.His154Pro	c.461A>C, p.His154Pro	Missense (5)	Missense (5)	[11-13]	[11-13]
01-099	*CCDC103 (NM_001258395.1)*	c.461A>C, p.His154Pro	c.461A>C, p.His154Pro	Missense (5)	Missense (5)	[11-13]	[11-13]
01-100	*CCDC103 (NM_001258395.1)*	c.461A>C, p.His154Pro	c.461A>C, p.His154Pro	Missense (5)	Missense (5)	[11-13]	[11-13]
01-103	*CCDC103 (NM_001258395.1)*	c.461A>C, p.His154Pro	c.461A>C, p.His154Pro	Missense (5)	Missense (5)	[11-13]	[11-13]
01-123	*CCDC103 (NM_001258395.1)*	c.461A>C, p.His154Pro	c.461A>C, p.His154Pro	Missense (5)	Missense (5)	[11-13]	[11-13]
01-124	*CCDC103 (NM_001258395.1)*	c.461A>C, p.His154Pro	c.461A>C, p.His154Pro	Missense (5)	Missense (5)	[11-13]	[11-13]
01-156	*CCDC103 (NM_001258395.1)*	c.461A>C, p.His154Pro	c.461A>C, p.His154Pro	Missense (5)	Missense (5)	[11-13]	[11-13]
01-170	*CCDC103 (NM_001258395.1)*	c.461A>C, p.His154Pro	c.461A>C, p.His154Pro	Missense (5)	Missense (5)	[11-13]	[11-13]
01-201	*CCDC103 (NM_001258395.1)*	c.461A>C, p.His154Pro	c.461A>C, p.His154Pro	Missense (5)	Missense (5)	[11-13]	[11-13]
02-051	*CCDC103 (NM_001258395.1)*	c.461A>C, p.His154Pro	c.461A>C, p.His154Pro	Missense (5)	Missense (5)	[11-13]	[11-13]
02-018	*CCDC103 (NM_001258395.1)*	c.461A>C, p.His154Pro	c.461A>C, p.His154Pro	Missense (5)	Missense (5)	[11-13]	[11-13]
02-033	*CCDC103 (NM_001258395.1)*	c.461A>C, p.His154Pro	c.461A>C, p.His154Pro	Missense (5)	Missense (5)	[11-13]	[11-13]
02-034	*CCDC103 (NM_001258395.1)*	c.461A>C, p.His154Pro	c.461A>C, p.His154Pro	Missense (5)	Missense (5)	[11-13]	[11-13]
01-079	*CCDC114 (NM_144577.3)*	c.287del, p.Lys96Argfs*23	c.287del, p.Lys96Argfs*23	Frameshift (5)	Frameshift (5)	NA	NA
01-029	*CCDC114 (NM_144577.3)*	c.486+1G>A	c.486+1G>A	Essential splice (5)	Essential splice (5)	[14]	[14]
01-074	*CCDC39 (NM_181426.1)*	c.1315A>T, p.Lys439*	c.1315A>T, p.Lys439*	Nonsense (5)	Nonsense (5)	NA	NA
01-064	*CCDC39 (NM_181426.1)*	c.1450del, p.Ile484Leufs*47	c.357+1G>C	Frameshift (5)	Essential splice (5)	[15]	[16]
01-030	*CCDC39 (NM_181426.1)*	c.1795C>T, p.Arg599*	c.1795C>T, p.Arg599*	Nonsense (5)	Nonsense (5)	[15]	[16]
01-045	*CCDC39 (NM_181426.1)*	c.2039_2040del, p.Cys680Phefs*9	c.526_527del, p.Leu176Alafs10*	Frameshift (5)	Frameshift (5)	[17]	NA
01-093	*CCDC39 (NM_181426.1)*	c.2040_2043del, p.Cys680Trpfs*15	c.440T>G, p.Leu147*	Frameshift (5)	Nonsense (5)	[17]	NA
01-063	*CCDC39 (NM_181426.1)*	c.2245G>T, p.Glu749*	c.2245G>T, p.Glu749*	Nonsense (5)	Nonsense (5)	[15]	[15]
01-016	*CCDC39 (NM_181426.1)*	c.2596G>T, p.Glu866*	c.2596G>T, p.Glu866*	Nonsense (5)	Nonsense (5)	[15]	[15]
02-028	*CCDC39 (NM_181426.1)*	c.664G>T, p.Glu222*	c.526_527del, p.Leu176Alafs10*	Nonsense (5)	Frameshift (5)	[15]	[15]
01-086	*CCDC39 (NM_181426.1)*	c.669_670insTA	c.610-2A>G	Frameshift (5)	Essential splice (5)	NA	[16]
01-200	*CCDC39 (NM_181426.1)*	c.830_831delCA, p.Asn276Lysfs*4	c.830_831delCA, p.Asn276Lysfs*4	Frameshift (5)	Frameshift (5)	[15]	[15]
01-102	*CCDC40 (NM_017950.3)*	c.1414del, p.Arg472Glyfs*3	c.3097A>T, p.Lys1033*	Frameshift (5)	Nonsense (5)	NA	[18]
01-179	*CCDC40 (NM_017950.3)*	c.1415delC, p.Arg472fs3*	c.1415delC, p.Arg472fs3*	Frameshift (5)	Frameshift (5)	[15]	[15]
02-049	*CCDC40 (NM_017950.3)*	c.1819_1823delinsT, p.Leu607Trpfs*33	c.1819_1823delinsT, p.Leu607Trpfs*33	Frameshift (5)	Frameshift (5)	NA	NA
01-138	*CCDC40 (NM_017950.3)*	c.248del, p.Ala83Valfs*84	c.552+6T>A	Frameshift (5)	Splice site (4)	[15, 19, 20]	[15, 19, 20]
01-054	*CCDC40 (NM_017950.3)*	c.248del, p.Ala83Valfs*84	c.248del, p.Ala83Valfs*84	Frameshift (5)	Frameshift (5)	[15, 19, 20]	[15, 19, 20]
01-068	*CCDC40 (NM_017950.3)*	c.248del, p.Ala83Valfs*84	c.248del, p.Ala83Valfs*84	Frameshift (5)	Frameshift (5)	[15, 19, 20]	[15, 19, 20]
02-045	*CCDC40 (NM_017950.3)*	c.248del, p.Ala83Valfs*84	c.248del, p.Ala83Valfs*84	Frameshift (5)	Frameshift (5)	[15, 19, 20]	[15, 19, 20]
02-067	*CCDC40 (NM_017950.3)*	c.248del, p.Ala83Valfs*84	c.748C>T, p.Glu250*	Frameshift (5)	Nonsense (5)	[15, 19, 20]	[15, 19, 20]
01-216	*CCDC40 (NM_017950.3)*	c.248del, p.Ala83Valfs*84	c.2450-2A>G	Frameshift (5)	Essential splice (5)	[15, 19, 20]	[15, 19, 20]
01-187	*CCDC40 (NM_017950.3)*	c.248del, p.Ala83Valfs*84	c.248del, p.Ala83Valfs*84	Frameshift (5)	Frameshift (5)	[15, 19, 20]	[15]
01-005	*CCDC40 (NM_017950.3)*	c.2712-1G>T	c.2712-1G>T	Essential splice (5)	Essential splice (5)	[15]	[15]
01-031	*CCDC40 (NM_017950.3)*	c.2712-1G>T	c.2712-1G>T	Essential splice (5)	Essential splice (5)	[15]	[15]
02-021	*CCDC40 (NM_017950.3)*	c.2712-1G>T	c.2712-1G>T	Essential splice (5)	Essential splice (5)	[15]	[15]
01-111	*CCDC40 (NM_017950.3)*	c.2712-1G>T	c.248del, p.Ala83Valfs*84	Essential splice (5)	Frameshift (5)	[15]	[15]
01-092	*CCDC40 (NM_017950.3)*	c.3181-3C>G	c.3181-3C>G	Splice site (3)	Splice site (3)	NA	NA
01-215	*CCDC40 (NM_017950.3)*	c.712G>T, p.Glu238*	c.940-2A>G	Nonsense (5)	Essential splice (5)	NA	[15]
01-137	*CCDC40 (NM_017950.3)*	c.940-2A>G	c.248del, p.Ala83Valfs*84	Essential splice (5)	Frameshift (5)	[15]	[15, 19, 20]
01-160	*CCDC65 (NM_033124.4)*	c.658G>T, p.Glu220*	c.658G>T, p.Glu220*	Nonsense (5)	Nonsense (5)	NA	NA
01-161	*CCDC65 (NM_033124.4)*	c.658G>T, p.Glu220*	c.658G>T, p.Glu220*	Nonsense (5)	Nonsense (5)	NA	NA

01-122	*CCDC65 (NM_033124.4)*	c.877_878del, p.Ile293Profs*2	c.877_878del, p.Ile293Profs*2	Frameshift (5)	Frameshift (5)	[21]	[21]
01-109	*CCDC65 (NM_033124.4)*	c.913C>T, p.Arg305*	c.913C>T, p.Arg305*	Nonsense (5)	Nonsense (5)	NA	NA
01-139	*CCNO (NM_021147.3)*	c.258_262dup, p.Gln88Argfs*8	c.258_262dup, p.Gln88Argfs*8	Frameshift (5)	Frameshift (5)	[22]	[22]
02-017	*CCNO (NM_021147.3)*	c.538dupC, p.Val180Glyfs*55	c.538dupC, p.Val180Glyfs*55	Frameshift (5)	Frameshift (5)	NA	NA
01-126	*DNAAF1 (NM_178452.5)*	c.285del, p.Lys95Asnfs*14	c.1484del, p.Pro495Glnfs*40	Frameshift (5)	Frameshift (5)	NA	[23]
01-127	*DNAAF1 (NM_178452.5)*	c.285del, p.Lys95Asnfs*14	c.1484del, p.Pro495Glnfs*40	Frameshift (5)	Frameshift (5)	NA	[23]
01-128	*DNAAF1 (NM_178452.5)*	c.285del, p.Lys95Asnfs*14	c.1484del, p.Pro495Glnfs*40	Frameshift (5)	Frameshift (5)	NA	[23]
01-223	*DNAAF1 (NM_178452.6)*	Deletion of exons 1-3	Deletion of exons 1-3	CNV (5)	CNV (5)	NA	NA
01-186	*DNAAF3 (NM_001256715.1)*	c.1030_1031delinsG, p.Pro344Glyfs*64	c.1273G>T, p.Gly425*	Frameshift (5)	Nonsense (5)	NA	NA
01-113	*DNAAF3 (NM_001256715.1)*	c.162_164delinsG, p.Val55Glyfs*28	c.162_164delinsG, p.Val55Glyfs*28	Frameshift (5)	Frameshift (5)	NA	NA
01-185	*DNAAF3 (NM_001256715.1)*	c.228+5G>C	c.228+5G>C	Splice site (3)	Splice site (3)	NA	NA
01-112	*DNAAF3 (NM_001256715.1)*	c.481-1G>A	c.481-1G>A	Essential splice (5)	Essential splice (5)	NA	NA
01-047	*DNAAF3 (NM_001256715.1)*	c.609_610delinsTGGGA, p.Ala272delinsGlyThr	c.296del, p.Glu167Glyfs*88	Inframe delins (5)	Frameshift (5)	NA	NA
01-089	*DNAAF3 (NM_001256715.1)*	c.621dupT, p.Val208Cysfs*12	c.621dupT, p.Val208Cysfs*12	Frameshift (5)	Frameshift (5)	[24]	[24]
01-090	*DNAAF3 (NM_001256715.1)*	c.621dupT, p.Val208Cysfs*12	c.621dupT, p.Val208Cysfs*12	Frameshift (5)	Frameshift (5)	[24]	[24]
01-174	*DNAAF3 (NM_001256715.1)*	c.621dupT, p.Val208Cysfs*12	c.621dupT, p.Val208Cysfs*12	Frameshift (5)	Frameshift (5)	[24]	[24]
01-131	*DNAAF3 (NM_001256715.1)*	c.901C>T, p.Gln301*	c.901C>T, p.Gln301*	Nonsense (5)	Nonsense (5)	NA	NA
01-070	*DNAAF3 (NM_001256715.1)*	c.997dup, p.Asp333Glyfs*64	c.570G>A, p.Trp190*	Frameshift (5)	Nonsense (5)	NA	NA
01-088	*DNAAF4 (NM_130810.3)*	3.5 kb deletion of exon 7	3.5 kb deletion of exon 7	CNV (5)	CNV (5)	[25]	[25]
01-232	*DNAAF4 (NM_130810.3)*	3.5 kb deletion of exon 7	3.5 kb deletion of exon 7	CNV (5)	CNV (5)	[25]	[25]
02-022	*DNAAF4 (NM_130810.3)*	3.5 kb deletion of exon 7	3.5 kb deletion of exon 7	CNV (5)	CNV (5)	[25]	[25]
02-010	*DNAAF4 (NM_130810.3)*	3.5 kb deletion of exon 7	3.5 kb deletion of exon 7	CNV (5)	CNV (5)	[25]	[25]
02-019	*DNAAF4 (NM_130810.3)*	3.5 kb deletion of exon 7	3.5 kb deletion of exon 7	CNV (5)	CNV (5)	[25]	[25]
01-085	*DNAAF4 (NM_130810.3)*	c.390_393del, p.Val132*	c.390_393del, p.Val132*	Nonsense (5)	Nonsense (5)	[25]	[25]
01-136	*DNAAF4 (NM_130810.3)*	c.808C>T, p.Arg270*	c.808C>T, p.Arg270*	Nonsense (5)	Nonsense (5)	[25]	[25]
01-176	*DNAH11 (NM_001277115.1)*	c.13040T>C, p.Leu4347Pro	Deletion of exons 68-75	Missense (3)	CNV (5)	NA	NA
02-073	*DNAH11 (NM_001277115.1)*	c.13270G>T, p.Glu4424*	c.13373C>T, p.Pro4458Leu	Nonsense (5)	Missense (5)	NA	[26]
01-040	*DNAH11 (NM_001277115.1)*	c.13531_13532ins13, p.Ala4511Valfs*13	c.3727G>T, p.Glu1243*	Frameshift (5)	Nonsense (5)	[27]	[27]
01-041	*DNAH11 (NM_001277115.1)*	c.13531_13532ins13, p.Ala4511Valfs*13	c.3727G>T, p.Glu1243*	Frameshift (5)	Nonsense (5)	[27]	[27]
01-042	*DNAH11 (NM_001277115.1)*	c.13531_13532ins13, p.Ala4511Valfs*13	c.3727G>T, p.Glu1243*	Frameshift (5)	Nonsense (5)	[27]	[27]
01-043	*DNAH11 (NM_001277115.1)*	c.13531_13532ins13, p.Ala4511Valfs*13	c.3727G>T, p.Glu1243*	Frameshift (5)	Nonsense (5)	[27]	[27]
01-095	*DNAH11 (NM_001277115.1)*	c.2832dup, p.Gln945Serfs*10	c.13240dup, p.Thr4414Asnfs*34	Frameshift (5)	Frameshift (5)	[27]	[27]
01-133	*DNAH11 (NM_001277115.1)*	c.3220G>T, p.Glu1074*	c.13069C>T, p.Arg4357*	Nonsense (5)	Nonsense (5)	[27]	[27]
01-147	*DNAH11 (NM_001277115.1)*	c.3380G>A, p.Trp1127*	c.3380G>A, p.Trp1127*	Nonsense (5)	Nonsense (5)	NA	NA
01-157	*DNAH11 (NM_001277115.1)*	c.3544C>T, p.Arg1182*	c.8798-5G>A	Nonsense (5)	Splice site (3)	[27]	[27]
02-063	*DNAH11 (NM_001277115.1)*	c.4333C>T, p.Arg1445*	c.9783G>C, p.Glu3261Asp	Nonsense (5)	Missense (3)	[26]	NA
02-062	*DNAH11 (NM_001277115.1)*	c.4333C>T, p.Arg1445*	c.4333C>T, p.Arg1445*	Nonsense (5)	Nonsense (5)	[26]	[26]
02-068	*DNAH11 (NM_001277115.1)*	c.4333C>T, p.Arg1445*	c.8698C>T, p.Arg2900*	Nonsense (5)	Nonsense (5)	[26]	[28]
02-038	*DNAH11 (NM_001277115.1)*	c.4333C>T, p.Arg1445*	c.13171C>T, p.Gln4391*	Nonsense (5)	Nonsense (5)	[26]	NA
01-065	*DNAH11 (NM_001277115.1)*	c.4410_4413del	c.7663C>T, p.Gln2555*	Frameshift (5)	Nonsense (5)	[27]	[27]
01-163	*DNAH11 (NM_001277115.1)*	c.4552C>T, p.Gln1518*	c.5778+1G>A, p.Val1821Thrfs*7	Nonsense (5)	Essential splice (5)	NA	[26]
01-084	*DNAH11 (NM_001277115.1)*	c.5506C>T, p.Arg1836*	c.5636T>A, p.Leu1879*	Nonsense (5)	Nonsense (5)	[27]	[27]
02-079	*DNAH11 (NM_001277115.1)*	c.5593C>T, p.Arg1865*	c.5593C>T, p.Arg1865*	Nonsense (5)	Nonsense (5)	NA	NA
01-158	*DNAH11 (NM_001277115.1)*	c.5924+1G>C	c.5924+1G>C	Essential splice (5)	Essential splice (5)	NA	NA
01-082	*DNAH11 (NM_001277115.1)*	c.6506C>T, p.Ser2169Leu	c.6506C>T, p.Ser2169Leu	Missense (3)	Missense (3)	[28]	[28]
01-083	*DNAH11 (NM_001277115.1)*	c.6506C>T, p.Ser2169Leu	c.6506C>T, p.Ser2169Leu	Missense (3)	Missense (3)	[28]	[28]
02-016	*DNAH11 (NM_001277115.1)*	c.6664C>T, p.Arg2222*	c.6682A>T, p.Lys2228*	Nonsense (5)	Nonsense (5)	NA	NA
01-221	*DNAH11 (NM_001277115.1)*	c.7472G>C, p.Arg2491Pro	c.6565C>T, p.Arg2189*	Missense (5)	Nonsense (5)	NA	NA
02-050	*DNAH11 (NM_001277115.1)*	c.8719C>T, p.Arg2907*	c.8719C>T, p.Arg2907*	Nonsense (5)	Nonsense (5)	[28]	[28]
01-169	*DNAH11 (NM_001277115.1)*	c.8932C>T, p.Gln2978*	c.853_857delinsG, p.Arg285Glufs*22	Nonsense (5)	Frameshift (5)	NA	NA
01-220	*DNAH11 (NM_001277115.1)*	c.9581_9582del, p.Leu3194Glnfs*10	c.4333C>T, p.Arg1445*	Frameshift (5)	Nonsense (5)	NA	[26]
01-178	*DNAH5 (NM_001369.2)*	c.10601T>C, p.Phe3534Ser	c.13458_13459insT, p.Asn4487fs*1	Missense (4)	Frameshift (5)	NA	[29, 30]
01-062	*DNAH5 (NM_001369.2)*	c.10616G>C, p.Arg3539Pro	c.7915C>T, p.Arg2639*	Missense (5)	Nonsense (5)	[31]	[32]

02-053	*DNAH5 (NM_001369.2)*	c.10815del, p.Pro3606Hisfs*22	c.6070-6071delAC, p.Gln2024Valfs*8	Frameshift (5)	Frameshift (5)	[20, 33, 34]	NA
02-054	*DNAH5 (NM_001369.2)*	c.10815del, p.Pro3606Hisfs*22	c.6070-6071delAC, p.Gln2024Valfs*8	Frameshift (5)	Frameshift (5)	[20, 33, 34]	NA
02-072	*DNAH5 (NM_001369.2)*	c.10815del, p.Pro3606Hisfs*22	c.5537T>C, p.Leu1846Pro	Frameshift (5)	Missense (5)	[20, 33, 34]	NA
02-048	*DNAH5 (NM_001369.2)*	c.10815del, p.Pro3606Hisfs*22	c.10616C>T, p.Arg3539Cys	Frameshift (5)	Missense (5)	[20, 33, 34]	[33]
02-023	*DNAH5 (NM_001369.2)*	c.10815del, p.Pro3606Hisfs*22	c.9720+5G>A	Frameshift (5)	Splice site (4)	[34]	NA
01-175	*DNAH5 (NM_001369.2)*	c.10815del, p.Pro3606Hisfs*22	c.13458_13459insT, p.Asn4487fs*1	Frameshift (5)	Frameshift (5)	[34]	[29, 30]
01-211	*DNAH5 (NM_001369.2)*	c.10815del, p.Pro3606Hisfs*22	c.10815del, p.Pro3606Hisfs*22	Frameshift (5)	Frameshift (5)	[33]	[33]
01-230	*DNAH5 (NM_001369.2)*	c.10815del, p.Pro3606Hisfs*22	c.2410G>T, p.Glu804*	Frameshift (5)	Nonsense (5)	[34]	NA
01-145	*DNAH5 (NM_001369.2)*	c.10825C>T, p.Gln3609*	c.3466del, p.Ile1156Leufs*24	Nonsense (5)	Frameshift (5)	NA	NA
01-146	*DNAH5 (NM_001369.2)*	c.10825C>T, p.Gln3609*	c.3466del, p.Ile1156Leufs*24	Nonsense (5)	Frameshift (5)	NA	NA
01-120	*DNAH5 (NM_001369.2)*	c.12705+1del	c.6249G>A, p.Met2083Ile	Essential splice (5)	Missense (5)	NA	[35]
01-191	*DNAH5 (NM_001369.2)*	c.13285C>T, p.Arg4429*	c.8642C>G, p.Ala2881Gly	Nonsense (5)	Missense (5)	NA	[34]
01-134	*DNAH5 (NM_001369.2)*	c.13285C>T, p.Arg4429*	c.13285C>T, p.Arg4429*	Nonsense (5)	Nonsense (5)	NA	NA
01-135	*DNAH5 (NM_001369.2)*	c.13285C>T, p.Arg4429*	c.13285C>T, p.Arg4429*	Nonsense (5)	Nonsense (5)	NA	NA
01-143	*DNAH5 (NM_001369.2)*	c.13338+1G>C	c.11437C>T, p.Arg3813Trp	Essential splice (5)	Missense (5)	NA	[18]
01-116	*DNAH5 (NM_001369.2)*	c.13399C>T, p.Gln4467*	c.13399C>T, p.Gln4467*	Nonsense (5)	Nonsense (5)	NA	NA
02-039	*DNAH5 (NM_001369.2)*	c.13458_13459insT, p.Asn4487fs*1	c.13338+1G>C	Frameshift (5)	Essential splice (5)	[29, 30]	NA
02-009	*DNAH5 (NM_001369.2)*	c.13458_13459insT, p.Asn4487fs*1	c.6930_6934delinsG, p.Asn2310Lysfs*15	Frameshift (5)	Frameshift (5)	[29, 30]	NA
01-048	*DNAH5 (NM_001369.2)*	c.13486C>T, p.Arg4496*	c.13458_13459insT, p.Asn4487fs*1	Nonsense (5)	Frameshift (5)	[29, 36, 37]	[29, 30]
02-024	*DNAH5 (NM_001369.2)*	c.13836G>A, p.Trp4612*	c.5710-2A>G, p.Cys1904-Lys1909del	Nonsense (5)	Essential splice (5)	NA	[33]
01-144	*DNAH5 (NM_001369.2)*	c.1828C>T, p.Gln610*	c.5563dup, p.Ile1855Asnfs*6	Nonsense (5)	Frameshift (5)	[32]	NA
01-189	*DNAH5 (NM_001369.2)*	c.232C>T, p.Arg78*	c.10815del, p.Pro3606Hisfs*22	Nonsense (5)	Frameshift (5)	[29, 36]	[34]
01-181	*DNAH5 (NM_001369.2)*	c.2710G>T, p.Glu904*	c.2710G>T, p.Glu904*	Nonsense (4)	Nonsense (4)	NA	NA
01-051	*DNAH5 (NM_001369.2)*	c.2893C>T, p.Gln965*	c.975-2A>G	Nonsense (5)	Essential splice (5)	NA	NA
01-206	*DNAH5 (NM_001369.2)*	c.5177T>C, p.Leu1726Pro	c.1730G>C, p.Arg577Thr	Missense (5)	Missense (5)	[31]	[29]
01-207	*DNAH5 (NM_001369.2)*	c.5177T>C, p.Leu1726Pro	c.1730G>C, p.Arg577Thr	Missense (5)	Missense (5)	[31]	[29]
02-029	*DNAH5 (NM_001369.2)*	c.5710-2A>G, p.Cys1904-Lys1909del	c.5710-2A>G, p.Cys1904-Lys1909del	Essential splice (5)	Essential splice (5)	[33]	[33]
01-190	*DNAH5 (NM_001369.2)*	c.5890_5894dup, p.Leu1966Serfs*9	c.6791G>A, p.Ser2264Asn	Frameshift (5)	Missense (5)	NA	[29]
02-011	*DNAH5 (NM_001369.2)*	c.6261T>G, p.Tyr2087*	c.6261T>G, p.Tyr2087*	Nonsense (5)	Nonsense (5)	NA	NA
02-026	*DNAH5 (NM_001369.2)*	c.6304C>T, p.Arg2102Cys	c.2052+1G>T	Missense (3)	Essential splice (5)	NA	NA
01-196	*DNAH5 (NM_001369.2)*	c.6763C>T, p.Arg2255*	c.9480T>A, p.Cys3160*	Nonsense (5)	Nonsense (5)	NA	NA
01-132	*DNAH5 (NM_001369.2)*	c.8383C>T, p.Arg2795*	c.5484+1G>A	Nonsense (5)	Essential splice (5)	NA	NA
01-209	*DNAH5 (NM_001369.2)*	c.8404C>T, p.Gln2802*	c.6249G>A, p.Met2083Ile	Nonsense (5)	Missense (5)	[29, 36]	NA
01-015	*DNAH5 (NM_001369.2)*	c.9516dup, p.Val3173Argfs*14	c.9516dup, p.Val3173Argfs*14	Frameshift (5)	Frameshift (5)	NA	NA
01-115	*DNAH5 (NM_001369.2)*	c.9694C>T, p.Gln3232*	c.9694C>T, p.Gln3232*	Nonsense (5)	Nonsense (5)	NA	NA
02-046	*DNAI1 (NM_012144.3)*	c.1490G>A, p.Gly497Asp	c.48+2dup, p.Ser17Valfs*12	Missense (5)	Essential splice (5)	[38]	[38]
01-044	*DNAI1 (NM_012144.3)*	c.1603del, p.Thr535Profs*31	c.1603del, p.Thr535Profs*31	Frameshift (5)	Frameshift (5)	NA	NA
01-069	*DNAI1 (NM_012144.3)*	c.1603del, p.Thr535Profs*31	c.1603del, p.Thr535Profs*31	Frameshift (5)	Frameshift (5)	NA	NA
01-087	*DNAI1 (NM_012144.3)*	c.1603del, p.Thr535Profs*31	c.1603del, p.Thr535Profs*31	Frameshift (5)	Frameshift (5)	NA	NA
02-013	*DNAI1 (NM_012144.3)*	c.1612G>A, p.Ala538Thr	c.1612G>A, p.Ala538Thr	Missense (5)	Missense (5)	[38]	[38]
02-031	*DNAI1 (NM_012144.3)*	c.1612G>A, p.Ala538Thr	c.1612G>A, p.Ala538Thr	Missense (5)	Missense (5)	[38]	[38]
01-021	*DNAI1 (NM_012144.3)*	c.48+2dup, p.Ser17Valfs*12	c.48+2dup, p.Ser17Valfs*12	Essential splice (5)	Essential splice (5)	[38]	[38]
01-022	*DNAI1 (NM_012144.3)*	c.48+2dup, p.Ser17Valfs*12	c.48+2dup, p.Ser17Valfs*12	Essential splice (5)	Essential splice (5)	[38]	[38]
01-140	*DNAI1 (NM_012144.3)*	c.48+2dup, p.Ser17Valfs*12	c.48+2dup, p.Ser17Valfs*12	Essential splice (5)	Essential splice (5)	[38]	[38]
02-006	*DNAI1 (NM_012144.3)*	c.48+2dup, p.Ser17Valfs*12	c.48+2dup, p.Ser17Valfs*12	Essential splice (5)	Essential splice (5)	[38]	[38]
02-058	*DNAI1 (NM_012144.3)*	c.48+2dup, p.Ser17Valfs*12	c.48+2dup, p.Ser17Valfs*12	Essential splice (5)	Essential splice (5)	[38]	[38]
02-059	*DNAI1 (NM_012144.3)*	c.48+2dup, p.Ser17Valfs*12	c.48+2dup, p.Ser17Valfs*12	Essential splice (5)	Essential splice (5)	[38]	[38]
02-069	*DNAI1 (NM_012144.3)*	c.48+2dup, p.Ser17Valfs*12	c.48+2dup, p.Ser17Valfs*12	Essential splice (5)	Essential splice (5)	[38]	[38]
01-001	*DNAI1 (NM_012144.3)*	c.48+2dup, p.Ser17Valfs*12	c.1612G>A, p.Ala538Thr	Essential splice (5)	Missense (5)	[38]	[38]
02-061	*DNAI1 (NM_012144.3)*	c.48+2dup, p.Ser17Valfs*12	c.1612G>A, p.Ala538Thr	Essential splice (5)	Missense (5)	[38]	[38]
01-028	*DNAI2 (NM_023036.4)*	c.1304G>A, p.Trp435*	c.1304G>A, p.Trp435*	Nonsense (5)	Nonsense (5)	[35]	[35]
01-101	*DNAI2 (NM_023036.4)*	c.1304G>A, p.Trp435*	c.1304G>A, p.Trp435*	Nonsense (5)	Nonsense (5)	[35]	[35]
01-229	*DNAI2 (NM_023036.4)*	c.883C>T, p.Arg295*	c.883C>T, p.Arg295*	Nonsense (5)	Nonsense (5)	NA	NA

01-097	*DNAL1 (NM_031427.3)*	c.225_229del, p.Leu75Phefs*30	c.225_229del, p.Leu75Phefs*30	Frameshift (5)	Frameshift (5)	NA	NA
01-055	*DRC1 (NM_145038.2)*	c.352C>T, p.Gln118*	c.2020C>T, p.Gln674*	Nonsense (5)	Nonsense (5)	[39]	NA
01-056	*DRC1 (NM_145038.2)*	c.352C>T, p.Gln118*	c.2020C>T, p.Gln674*	Nonsense (5)	Nonsense (5)	[39]	NA
01-119	*HYDIN (NM_001270974.2)*	c.13709del, p.Pro4570Leufs*22	c.13709del, p.Pro4570Leufs*22	Frameshift (5)	Frameshift (5)	NA	NA
01-121	*HYDIN (NM_001270974.2)*	c.2194dup, p.Tyr732Leufs*2	c.2194dup, p.Tyr732Leufs*2	Frameshift (5)	Frameshift (5)	NA	NA
02-027	*LRRC6 (NM_012472.4)*	c.299T>C, p.Ile100Thr	c.630del, p.Trp210Cysfs*12	Missense (5)	Frameshift (5)	[37]	[20]
01-057	*LRRC6 (NM_012472.4)*	c.630del, p.Trp210Cysfs*12	c.630del, p.Trp210Cysfs*12	Frameshift (5)	Frameshift (5)	[20]	[20]
01-094	*LRRC6 (NM_012472.4)*	c.630del, p.Trp210Cysfs*12	c.630del, p.Trp210Cysfs*12	Frameshift (5)	Frameshift (5)	[20]	[20]
01-129	*LRRC6 (NM_012472.4)*	c.630del, p.Trp210Cysfs*12	c.630del, p.Trp210Cysfs*12	Frameshift (5)	Frameshift (5)	[20]	[20]
01-130	*LRRC6 (NM_012472.4)*	c.630del, p.Trp210Cysfs*12	c.630del, p.Trp210Cysfs*12	Frameshift (5)	Frameshift (5)	[20]	[20]
01-184	*LRRC6 (NM_012472.4)*	c.630del, p.Trp210Cysfs*12	c.630del, p.Trp210Cysfs*12	Frameshift (5)	Frameshift (5)	[20]	[20]
01-204	*LRRC6 (NM_012472.4)*	c.630del, p.Trp210Cysfs*12	c.630del, p.Trp210Cysfs*12	Frameshift (5)	Frameshift (5)	[20]	[20]
01-218	*LRRC6 (NM_012472.4)*	c.630del, p.Trp210Cysfs*12	c.630del, p.Trp210Cysfs*12	Frameshift (5)	Frameshift (5)	[20]	[20]
01-010	*LRRC6 (NM_012472.4)*	c.183T>G, p.Asn61Lys	c.179-1G>A	Missense (4)	Essential splice (5)	NA	NA
01-011	*LRRC6 (NM_012472.4)*	c.183T>G, p.Asn61Lys	c.179-1G>A	Missense (4)	Essential splice (5)	NA	NA
01-142	*LRRC6 (NM_012472.4)*	c.793del, p.Arg266Aspfs*13	c.239_243del, p.Lys80Argfs*7	Frameshift (5)	Frameshift (5)	NA	NA
01-203	*MCIDAS (NM_001190787.1)*	c.332_333delinsG, p.Ala111Glyfs*22	c.332_333delinsG, p.Ala111Glyfs*22	Frameshift (5)	Frameshift (5)	NA	NA
01-007	*PIH1D3 (NM_001169154.1)*	c.127G>T, p.Glu43*	X-linked hemizygous	Nonsense (5)	-	[40]	-
01-164	*PIH1D3 (NM_001169154.1)*	c.266G>A, p.Trp89*	X-linked hemizygous	Nonsense (5)	-	[40]	-
01-075	*RPGR (NM_001034853.1)*	c.633del, p.Tyr212Metfs*11	X-linked hemizygous	Frameshift (5)	-	NA	-
02-007	*RPGR (NM_001034853.1)*	c.646G>T, p.Glu216*	X-linked hemizygous	Nonsense (5)	-	NA	-
02-012	*RPGR (NM_001034853.1)*	c.646G>T, p.Glu216*	X-linked hemizygous	Nonsense (5)	-	NA	-
02-037	*RPGR (NM_001034853.1)*	c.706C>T, p.Gln236*	X-linked hemizygous	Nonsense (5)	-	NA	-
01-208	*RSPH1 (NM_080860.3)*	c.275-2A>C, p.Gly92Alafs*10	c.275-2A>C, p.Gly92Alafs*10	Essential splice (5)	Essential splice (5)	[41]	[41]
02-005	*RSPH1 (NM_080860.3)*	c.275-2A>C, p.Gly92Alafs*10	c.275-2A>C, p.Gly92Alafs*10	Essential splice (5)	Essential splice (5)	[41]	[41]
02-008	*RSPH1 (NM_080860.3)*	c.275-2A>C, p.Gly92Alafs*10	c.275-2A>C, p.Gly92Alafs*10	Essential splice (5)	Essential splice (5)	[41]	[41]
01-199	*RSPH4A (NM_001010892.2)*	c.1351C>T, p.Gln451*	c.116C>A, p.Ser39*	Nonsense (5)	Nonsense (5)	NA	[42]
01-173	*RSPH4A (NM_001010892.2)*	c.1962_1966delinsC, p.Asp655Ilefs*83	c.1962_1966delinsC, p.Asp655Ilefs*83	Frameshift (5)	Frameshift (5)	NA	NA
01-026	*RSPH4A (NM_001010892.2)*	c.325C>T, p.Gln109*	c.1468C>T, p.Arg490*	Nonsense (5)	Nonsense (5)	[43]	[43]
01-037	*RSPH4A (NM_001010892.2)*	c.460C>T, p.Gln154*	c.460C>T, p.Gln154*	Nonsense (5)	Nonsense (5)	[43]	[43]
01-038	*RSPH4A (NM_001010892.2)*	c.460C>T, p.Gln154*	c.460C>T, p.Gln154*	Nonsense (5)	Nonsense (5)	[43]	[43]
01-039	*RSPH4A (NM_001010892.2)*	c.460C>T, p.Gln154*	c.460C>T, p.Gln154*	Nonsense (5)	Nonsense (5)	[43]	[43]
01-081	*RSPH4A (NM_001010892.2)*	c.460C>T, p.Gln154*	c.460C>T, p.Gln154*	Nonsense (5)	Nonsense (5)	[43]	[43]
02-057	*RSPH4A (NM_001010892.2)*	c.166dup, p.Arg56Profs*11	c.166dup, p.Arg56Profs*11	Frameshift (5)	Frameshift (5)	[12]	[12]
01-033	*RSPH9 (NM_001193341.1)*	c.801_803delGAA, p.Lys268del	c.801_803delGAA, p.Lys268del	Inframe AA del (5)	Inframe AA del (5)	[43]	[43]
01-034	*RSPH9 (NM_001193341.1)*	c.801_803delGAA, p.Lys268del	c.801_803delGAA, p.Lys268del	Inframe AA del (5)	Inframe AA del (5)	[43]	[43]
01-035	*RSPH9 (NM_001193341.1)*	c.801_803delGAA, p.Lys268del	c.801_803delGAA, p.Lys268del	Inframe AA del (5)	Inframe AA del (5)	[43]	[43]
01-036	*RSPH9 (NM_001193341.1)*	c.801_803delGAA, p.Lys268del	c.801_803delGAA, p.Lys268del	Inframe AA del (5)	Inframe AA del (5)	[43]	[43]
01-071	*RSPH9 (NM_001193341.1)*	c.801_803delGAA, p.Lys268del	c.801_803delGAA, p.Lys268del	Inframe AA del (5)	Inframe AA del (5)	[43]	[43]
01-194	*SPAG1 (NM_003114.4)*	c.1519dupA, p.Ile507Asnfs*5	c.1519dupA, p.Ile507Asnfs*5	Frameshift (5)	Frameshift (5)	[44]	[44]
01-195	*SPAG1 (NM_003114.4)*	c.1519dupA, p.Ile507Asnfs*5	c.1519dupA, p.Ile507Asnfs*5	Frameshift (5)	Frameshift (5)	[44]	[44]
01-025	*ZMYND10 (NM_015896.2)*	c.47T>G, p.Val16Gly	c.593_594del, p.Val198Glyfs*13	Missense (5)	Frameshift (5)	[45]	[45]
01-077	*ZMYND10 (NM_015896.2)*	c.65del, p.Phe22Serfs*21	c.65del, p.Phe22Serfs*21	Frameshift (5)	Frameshift (5)	[45]	[45]
01-078	*ZMYND10 (NM_015896.2)*	c.65del, p.Phe22Serfs*21	c.65del, p.Phe22Serfs*21	Frameshift (5)	Frameshift (5)	[45]	[45]
02-060	*ZMYND10 (NM_015896.2)*	c.47T>G, p.Val16Gly	c.47T>G, p.Val16Gly	Missense (5)	Missense (5)	[20, 45]	[20, 45]

Variants pathogenicity classified according to ACMG guidelines as Class 5 (pathogenic), Class 4 (likely pathogenic) or Class 3 (variant of uncertain significance, VUS) [2]. Class 3 variants (n=8) were included if variant present in combination with a Class 5 variant in the patient, or additional phenotypes suggested the Class 3 variant was highly likely causal although unpublished.

Table E5. Diagnostic characteristics of patients in the *validation* group, stratified by predefined gene groups. Genes are ordered according to gene distribution in the study population.

Diagnostic characteristic	**Dynein structure** *(DNAH5, DNAH11, DNAI1, DNAI2, ARMC4, DNAH9, TTC25)* (n=82)	**Dynein assembly** *(CCDC103, DNAAF4, PIHD3, DNAAF1, LRRC6, DNAAF3, SPAG1, DNAAF5, ZYMND10, CFAP300)* (n=42)	**Radial spoke/ central complex** *(RSPH4A, HYDIN, RSPH1, RSPH9, RSPH3)* (n=32)	**N-DRC/molecular ruler** *(CCDC39, CCDC40, CCDC65, DRC1)* (n=35)	**Other functions** *(RPGR, CCNO, MCIDAS)* (n=6)	**All**	**p-value**
Median nNO level in nL/min (IQR); n=138	16 (8.1 to 23.6)	14.4 (8 to 25)	22.9 (7.6 to 40.5)	13 (9.9 to 23)	35 (15.9 to 54)	16.3 (8.4 to 28)	0.7038
TEM findings, n=178							
Non-diagnostic TEM (%)	21 (28.4)	3 (8.3)	7 (22.6)	1 (2.9)	0	32 (18)	
Isolated ODA defect (%)	38 (51.4)	1 (2.8)	0	0	0	39 (21.9)	
ODA & IDA defect (%)	14 (18.9)	31 (86.1)	0	1 (2.9)	2 (66.7)	48 (27)	
MTD & IDA defect or isolated IDA defect (%)	0	1 (2.8)	1 (3.2)	32 (94.1)	0	34 (19.1)	
CC defect (%)	0	0	22 (71)	0	0	22 (12.4)	
Lack of cilia (%)	1 (1.4)	0	1 (3.2)	0	1 (33.3)	3 (1.7)	
CBP predominant side view, n=133							
Normal (%)	2 (2.6)	3 (9.1)	6 (20.7)	0	2 (40)	13 (7.4)	
Completely immotile (%)	34 (44.7)	27 (81.8)	1 (3.5)	14 (42.4)	1 (20)	77 (43.8)	
Weak residual movement (%)	29 (38.2)	3 (9.1)	6 (20.7)	12 (36.4)	0	50 (28.4)	
Stiff (%)	11 (14.5)	0	6 (20.7)	7 (21.2)	0	24 (13.6)	
Rotating (%)	0	0	10 (34.5)	0	0	10 (5.7)	
Staggered beat (%)	0	0	0	0	2 (40)	2 (1.1)	
Lack of cilia (%)	0	0	0	0	0	0	

*nNO= nasal nitric oxide (normal levels <77nl/min), TEM = Transmission electron microscopy, ODA= outer dynein arm, IDA = inner dynein arm, CC = central complex, CBP= ciliary beat pattern, ODA= outer dynein arm, IDA= inner dynein arm, MTD= microtubular disorganisation.

Table E6. Clinical characteristics of patients in the *validation* group, stratified by predefined gene groups. Genes are ordered according to gene distribution in the study population.

Clinical characteristic	**Dynein structure** *(DNAH5, DNAH11, DNAI1, DNAI2, ARMC4, DNAH9, TTC25)* (n=82)	**Dynein assembly** *(CCDC103, DNAAF4, PIHD3, DNAAF1, LRRC6, DNAAF3, SPAG1, DNAAF5, ZYMND10, CFAP300)* (n=42)	**Radial spoke/ central complex** *(RSPH4A, HYDIN, RSPH1, RSPH9, RSPH3)* (n=32)	**N-DRC/molecular ruler** *(CCDC39, CCDC40, CCDC65, DRC1)* (n=35)	**Other functions** *(RPGR, CCNO, MCIDAS)* (n=6)	**All**	**p-value**
Male (%)	41 (50)	22 (52.4)	14 (43.8)	23 (65.7)	4 (66.7)	104 (52.8)	0.393
Ethnicity, n=185							
White-British (%)	15 (20.0)	4 (11.1)	3 (9.4)	3 (9.4)	0	25 (13.5)	
White-Irish (%)	0	2 (5.6)	4 (12.5)	0	0	6 (3.2)	
White-other (%)	33 (41.8)	10 (27.8)	13 (40.6)	10 (31.3)	4 (66.7)	70 (37.8)	
Indian (%)	1 (1.3)	0	1 (3.1)	1 (3.1)	0	3 (1.6)	
Pakistani (%)	1 (1.3)	5 (13.9)	1 (3.1)	2 (6.3)	1 (16.7)	10 (5.4)	
Bangladeshi (%)	1 (1.3)	0	0	0	0	1 (0.5)	
Black (%)	2 (2.5)	3 (8.3)	1 (3.1)	1 (3.1)	0	7 (3.8)	
Chinese (%)	1 (1.3)	0	0	0	0	1 (0.5)	
Mixed (%)	5 (6.3)	0	0	1 (3.1)	0	6 (3.2)	
Other (%)	20 (25.3)	12 (33.3)	9 (28.1)	14 (43.8)	1 (16.7)	56 (30.3)	
Median FEV_1 z-scores (IQR), n=169	-1.3 (1.5)$^+$	-1.5 (1.6)	-2.1 (1.8)	-2.6 (1.5)$^+$	-2.6 (1.7)	-1.8 (1.6)	**0.0008**
Median age at diagnosis (IQR) n=184	14 (4.9 to 17.8)	14.3 (5.5 to 19.1)	15.9 (7.2 to 21.9)	13.9 (3.5 to 21.5)	20.4 (6.1 to 36)	14.5 (6 to 19.5)	0.435
Neonatal respiratory distress (%)	41 (56.9)	21 (60)	14 (50)	20 (69)	3 (50)	99 (58.2)	0.650
Wet cough (%)	78 (96.3)	38 (95)	29 (93.6)	31 (91.2)	5 (83.3)	181 (94.3)	0.431
Rhinitis (%)	77 (96.3)	37 (90.2)	26 (83.9)	31 (91.2)	5 (83.3)	176 (91.7)	0.150
Glue ear (%)	55 (69.6)	26 (66.7)	25 (83.3)	23 (69.7)	4 (66.7)	133 (71.1)	0.574
Situs solitus (%)	38 (48.1)	17 (41.5)	30 (100)	22 (62.9)	6 (100)	113 (59.2)	**<0.001**

$^+$ difference between groups was statistically significant (ANOVA followed by Tukey for pairwise comparisons). P values $\leq$0.05 highlighted. IQR: interquartile range, NRDS: neonatal respiratory distress syndrome, CHD: congenital heart defect.

Table E7. Summary of diagnostic test results for all patients included in the study, stratified by gene group. Genes are ordered according to gene distribution in the study population.

Diagnostic test	**Dynein structure** *(DNAH5, DNAH11, DNAI1, DNAI2, ARMC4, DNAL1, DNAH9, TTC25)*, (n=171)	**Dynein assembly** *(CCDC103, DNAAF4, LRRC6, DNAAF3, DNAAF1, PIHD3, SPAG1, ZYMND10, CCDC114, DNAAF5, CFAP300)*, (n=94)	**Radial spoke/ central complex** *(RSPH4A, RSPH1, HYDIN, RSPH9, RSPH3)*, (n=50)	**N-DRC/molecular ruler** *(CCDC39, CCDC40, CCDC65, DRC1)*, (n=68)	**Other function** *(RPGR, CCNO, MCIDAS)*, (n=13)	**All**
nNO findings (%), n=287						
Median nNO level in nL/min [IQR]; n=287	12.1 [7.2 to 21.3]	15.3 [7.9 to 30.4]	23 [9.8 to 36]	12.8 [7.5 to 20]	39.9 [15.6 to 75.5]	14.4 [8.0 to 26.0]
n patients with nNO<77 nL/min (%)	120 (95.2)	53 (88.3)	34 (89.5)	54 (98.2)	6 (75)	267 (93.0)
TEM findings (%), n=365						
Non-diagnostic TEM	46 (29.3)	7 (8.0)	10 (20.8)	3 (4.8)	2 (22.2)	68 (18.6)
Isolated ODA defect	89 (56.7)	10 (11.4)	0	0	0	99 (27.1)
ODA & IDA defect	20 (12.7)	65 (73.9)	0	1 (1.6)	(22.2)	88 (24.1)
MTD & IDA defect or isolated IDA defect	0	5 (5.7)	1 (2.1)	59 (93.7)	0	65 (17.8)
CC defect	0	0	35 (72.9)	0	0	35 (9.6)
Lack of cilia	2 (1.3)	1 (1.1)	2 (4.2)	0	5 (55.6)	10 (2.7)
CBP predominant side view (%), n=309						
Normal	2 (1.5)	6 (9.0)	6 (15.4)	0	2 (16.7)	16 (5.2)
Completely immotile	69 (50.7)	52 (77.6)	1 (2.6)	19 (34.6)	4 (33.3)	145 (46.9)
Weak residual movement	37 (27.2)	3 (4.5)	6 (15.4)	13 (23.6)	0	59 (19.1)
Stiff	27 (19.9)	6 (9.0)	9 (23.1)	18 (32.7)	2 (16.7)	62 (20.1)
Rotating	0	0	16 (41.0)	0	0	16 (5.2)
Staggered beat	0	0	1 (2.6)	5 (9.1)	2 (16.7)	8 (2.6)
Lack of cilia	1 (0.7)	0	0	0	2 (16.7)	3 (1.0)

N-DRC = nexin-dynein regulatory complex, nNO = nasal nitric oxide (normal levels <77nl/min), IQR = interquartile range, TEM = Transmission electron microscopy, ODA= outer dynein arm, IDA = inner dynein arm, CC = central complex, MTD = microtubular disorganisation, CBP= ciliary beat pattern.

Table E8. Summary of clinical characteristics for all patients included in the study, stratified by gene group. Genes are ordered according to gene distribution in the study population.

Clinical characteristic	**Dynein structure** *(DNAH5, DNAH11, DNAI1, DNAI2, ARMC4, DNAL1, DNAH9, TTC25)*, (n=171)	**Dynein assembly** *(CCDC103, DNAAF4, LRRC6, DNAAF3, DNAAF1, PIHD3, SPAG1, ZYMND10, CCDC114, DNAAF5, CFAP300)*, (n=94)	**Radial spoke/ central complex** *(RSPH4A, RSPH1, HYDIN, RSPH9, RSPH3)*, (n=50)	**N-DRC/ molecular ruler** *(CCDC39, CCDC40, CCDC65, DRC1)*, (n=68)	**Other function** *(RPGR, CCNO, MCIDAS)*, (n=13)	**All**	**p-value**
Male (%), n=396	75 (43.9)	49 (52.1)	23 (46.0)	35 (51.5)	9 (69.2)	191 (48.2)	0.226
Mean FEV_1 z-scores (SD), n=275	-1.3 (1.4)$^{+}$	-1.7 (1.5)$^{\#}$	-1.7 (1.9)	-2.5 (1.5)$^{+\#}$	-2.8 (2.2)	-1.7 (1.6)	**<0.001**
Median age at diagnosis (IQR) n=353	12 (3 to 20)	9.6 (3.1 to 16.3)	14.6 (7.8 to 18.7)	10.9 (2.2 to 15.4)	12.1 (6 to 20.4)	11.1 (4.2 to 17.8)	0.235
Neonatal respiratory distress syndrome (%), n=305	72 (55.8)	52 (73.2)	21 (52.5)	35 (67.3)	6 (46.2)	186 (61.0)	0.056
Wet cough (%), n=351	144 (95.4)	78 (95.3)	43 (93.5)	56 (93.3)	10 (76.9)	331 (94.3)	0.165
Rhinitis (%), n=349	142 (94.0)	75 (91.5)	37 (84.1)	49 (83.1)	10 (76.9)	313 (89.7)	**0.028**
Glue ear (%), n=333	93 (64.1)	45 (58.4)	34 (81.0)	32 (57.1)	8 (61.5)	214 (63.7)	0.099
Situs solitus (%), n=380	69 (42.9)	36 (39.1)	48 (100)	40 (60.6)	13 (100)	206 (54.2)	**<0.001**

N-DRC = nexin-dynein regulatory complex; SD = standard deviation; IQR = interquartile range; + # difference between groups was statistically significant (ANOVA followed by Tukey for pairwise comparisons).

References

1. Lucas JS, Barbato A, Collins SA, Goutaki M, Behan L, Caudri D, Dell S, Eber E, Escudier E, Hirst RA, Hogg C, Jorissen M, Latzin P, Legendre M, Leigh MW, Midulla F, Nielsen KG, Omran H, Papon J-F, Pohunek P, Redfern B, Rigau D, Rindlisbacher B, Santamaria F, Shoemark A, Snijders D, Tonia T, Titieni A, Walker WT, Werner C, Bush A, Kuehni CE. European Respiratory Society guidelines for the diagnosis of primary ciliary dyskinesia. *European Respiratory Journal* 2017: 49(1): 1601090.
2. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* 2015: 17(5): 405-424.
3. Carlsson G. Topology and data. *Bull Amer Math Soc* 2009: 46(2): 255-308.
4. Glushakov S, Kotenko, I., Rekalo, A. Handling Missing Data in Clinical Trials Using Topological Data Analysis. PhUSE EU Connect 2018 (https://www.lexjansen.com/phuse/2018/ml/ML07.pdf).
5. Best S, Shoemark A, Rubbo B, Patel MP, Fassad MR, Dixon M, Rogers AV, Hirst RA, Rutman A, Ollosson S, Jackson CL, Goggin P, Thomas S, Pengelly R, Cullup T, Pissaridou E, Hayward J, Onoufriadis A, O'Callaghan C, Loebinger MR, Wilson R, Chung EM, Kenia P, Doughty VL, Carvalho JS, Lucas JS, Mitchison HM, Hogg C. Risk factors for situs defects and congenital heart disease in primary ciliary dyskinesia. *Thorax* 2019: 74(2): 203-205.
6. Omoyinmi E, Standing A, Keylock A, Price-Kuehne F, Melo Gomes S, Rowczenio D, Nanthapisal S, Cullup T, Nyanhete R, Ashton E, Murphy C, Clarke M, Ahlfors H, Jenkins L, Gilmour K, Eleftheriou D, Lachmann HJ, Hawkins PN, Klein N, Brogan PA. Clinical impact of a targeted next-generation sequencing gene panel for autoinflammation and vasculitis. *PloS one* 2017: 12(7): e0181874.
7. Quanjer PH, Stanojevic S, Cole TJ, Baur X, Hall GL, Culver BH, Enright PL, Hankinson JL, Ip MS, Zheng J, Stocks J, Initiative ERSGLF. Multi-ethnic reference values for spirometry for the 3-95-yr age range: the global lung function 2012 equations. *Eur Respir J* 2012: 40(6): 1324-1343.
8. Group WHOMGRS. WHO Child Growth Standards based on length/height, weight and age. *Acta paediatrica* 2006: 450: 76-85.
9. Hjeij R, Lindstrand A, Francis R, Zariwala MA, Liu X, Li Y, Damerla R, Dougherty GW, Abouhamed M, Olbrich H, Loges NT, Pennekamp P, Davis EE, Carvalho CM, Pehlivan D, Werner C, Raidt J, Kohler G, Haffner K, Reyes-Mugica M, Lupski JR, Leigh MW, Rosenfeld M, Morgan LC, Knowles MR, Lo CW, Katsanis N, Omran H. ARMC4 mutations cause primary ciliary dyskinesia with randomization of left/right body asymmetry. *Am J Hum Genet* 2013: 93(2): 357-367.
10. Onoufriadis A, Shoemark A, Munye MM, James CT, Schmidts M, Patel M, Rosser EM, Bacchelli C, Beales PL, Scambler PJ, Hart SL, Danke-Roelse JE, Sloper JJ, Hull S, Hogg C, Emes RD, Pals G, Moore AT, Chung EM, Uk10K, Mitchison HM. 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.
11. Panizzi JR, Becker-Heck A, Castleman VH, Al-Mutairi DA, Liu Y, Loges NT, Pathak N, Austin-Tse C, Sheridan E, Schmidts M, Olbrich H, Werner C, Haffner K, Hellman N, Chodhari R, Gupta A, Kramer-Zucker A, Olale F, Burdine RD, Schier AF, O'Callaghan C, Chung EM, Reinhardt R, Mitchison HM, King SM, Omran H, Drummond IA. CCDC103 mutations cause primary ciliary dyskinesia by disrupting assembly of ciliary dynein arms. *Nat Genet* 2012: 44(6): 714-719.
12. Casey JP, McGettigan PA, Healy F, Hogg C, Reynolds A, Kennedy BN, Ennis S, Slattery D, Lynch SA. Unexpected genetic heterogeneity for primary ciliary dyskinesia in the Irish Traveller population. *Eur J Hum Genet* 2015: 23(2): 210-217.
13. D'Andrea G, Schiavulli M, Dimatteo C, Santacroce R, Guerra E, Longo VA, Grandone E, Margaglione M. Homozygosity by descent of a 3Mb chromosome 17 haplotype causes coinheritance of Glanzmann thrombasthenia and primary ciliary dyskinesia. *Blood* 2013: 122(26): 4289-4291.
14. Onoufriadis A, Paff T, Antony D, Shoemark A, Micha D, Kuyt B, Schmidts M, Petridi S, Dankert-Roelse JE, Haarman EG, Daniels JM, Emes RD, Wilson R, Hogg C, Scambler PJ, Chung EM, Uk10K, Pals G, Mitchison HM. 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.
15. Antony D, Becker-Heck A, Zariwala MA, Schmidts M, Onoufriadis A, Forouhan M, Wilson R, Taylor-Cox T, Dewar A, Jackson C, Goggin P, Loges NT, Olbrich H, Jaspers M, Jorissen M, Leigh MW, Wolf WE, Daniels ML, Noone PG, Ferkol TW, Sagel SD, Rosenfeld M, Rutman A, Dixit A, O'Callaghan C, Lucas JS, Hogg C, Scambler PJ, Emes RD, Uk10k, Chung EM, Shoemark A, Knowles MR, Omran H, Mitchison HM. Mutations in CCDC39 and CCDC40 are the major cause of primary ciliary dyskinesia with axonemal disorganization and absent inner dynein arms. *Hum Mutat* 2013: 34(3): 462-472.

16. Merveille AC, Davis EE, Becker-Heck A, Legendre M, Amirav I, Bataille G, Belmont J, Beydon N, Billen F, Clement A, Clercx C, Coste A, Crosbie R, de Blic J, Deleuze S, Duquesnoy P, Escalier D, Escudier E, Fliegauf M, Horvath J, Hill K, Jorissen M, Just J, Kispert A, Lathrop M, Loges NT, Marthin JK, Momozawa Y, Montantin G, Nielsen KG, Olbrich H, Papon JF, Rayet I, Roger G, Schmidts M, Tenreiro H, Towbin JA, Zelenika D, Zentgraf H, Georges M, Lequarre AS, Katsanis N, Omran H, Amselem S. 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.
17. Blanchon S, Legendre M, Copin B, Duquesnoy P, Montantin G, Kott E, Dastot F, Jeanson L, Cachanado M, Rousseau A, Papon JF, Beydon N, Brouard J, Crestani B, Deschildre A, Desir J, Dollfus H, Leheup B, Tamalet A, Thumerelle C, Vojtek AM, Escalier D, Coste A, de Blic J, Clement A, Escudier E, Amselem S. Delineation of CCDC39/CCDC40 mutation spectrum and associated phenotypes in primary ciliary dyskinesia. *J Med Genet* 2012: 49(6): 410-416.
18. Kim RH, D AH, Cutz E, Knowles MR, Nelligan KA, Nykamp K, Zariwala MA, Dell SD. The role of molecular genetic analysis in the diagnosis of primary ciliary dyskinesia. *Annals of the American Thoracic Society* 2014: 11(3): 351-359.
19. Becker-Heck A, Zohn IE, Okabe N, Pollock A, Lenhart KB, Sullivan-Brown J, McSheene J, Loges NT, Olbrich H, Haeffner K, Fliegauf M, Horvath J, Reinhardt R, Nielsen KG, Marthin JK, Baktai G, Anderson KV, Geisler R, Niswander L, Omran H, Burdine RD. 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.
20. Zariwala MA, Gee HY, Kurkowiak M, Al-Mutairi DA, Leigh MW, Hurd TW, Hjeij R, Dell SD, Chaki M, Dougherty GW, Adan M, Spear PC, Esteve-Rudd J, Loges NT, Rosenfeld M, Diaz KA, Olbrich H, Wolf WE, Sheridan E, Batten TF, Halbritter J, Porath JD, Kohl S, Lovric S, Hwang DY, Pittman JE, Burns KA, Ferkol TW, Sagel SD, Olivier KN, Morgan LC, Werner C, Raidt J, Pennekamp P, Sun Z, Zhou W, Airik R, Natarajan S, Allen SJ, Amirav I, Wieczorek D, Landwehr K, Nielsen K, Schwerk N, Sertic J, Kohler G, Washburn J, Levy S, Fan S, Koerner-Rettberg C, Amselem S, Williams DS, Mitchell BJ, Drummond IA, Otto EA, Omran H, Knowles MR, Hildebrandt F. ZMYND10 is mutated in primary ciliary dyskinesia and interacts with LRRC6. *Am J Hum Genet* 2013: 93(2): 336-345.
21. Austin-Tse C, Halbritter J, Zariwala MA, Gilberti RM, Gee HY, Hellman N, Pathak N, Liu Y, Panizzi JR, Patel-King RS, Tritschler D, Bower R, O'Toole E, Porath JD, Hurd TW, Chaki M, Diaz KA, Kohl S, Lovric S, Hwang DY, Braun DA, Schueler M, Airik R, Otto EA, Leigh MW, Noone PG, Carson JL, Davis SD, Pittman JE, Ferkol TW, Atkinson JJ, Olivier KN, Sagel SD, Dell SD, Rosenfeld M, Milla CE, Loges NT, Omran H, Porter ME, King SM, Knowles MR, Drummond IA, Hildebrandt F. Zebrafish Ciliopathy Screen Plus Human Mutational Analysis Identifies C21orf59 and CCDC65 Defects as Causing Primary Ciliary Dyskinesia. *Am J Hum Genet* 2013: 93(4): 672-686.
22. Wallmeier J, Al-Mutairi DA, Chen CT, Loges NT, Pennekamp P, Menchen T, Ma L, Shamseldin HE, Olbrich H, Dougherty GW, Werner C, Alsabah BH, Kohler G, Jaspers M, Boon M, Griese M, Schmitt-Grohe S, Zimmermann T, Koerner-Rettberg C, Horak E, Kintner C, Alkuraya FS, Omran H. Mutations in CCNO result in congenital mucociliary clearance disorder with reduced generation of multiple motile cilia. *Nat Genet* 2014: 46(6): 646-651.
23. Watson CM, Crinnion LA, Morgan JE, Harrison SM, Diggle CP, Adlard J, Lindsay HA, Camm N, Charlton R, Sheridan E, Bonthron DT, Taylor GR, Carr IM. Robust diagnostic genetic testing using solution capture enrichment and a novel variant-filtering interface. *Hum Mutat* 2014: 35(4): 434-441.
24. Mitchison HM, Schmidts M, Loges NT, Freshour J, Dritsoula A, Hirst RA, O'Callaghan C, Blau H, Al Dabbagh M, Olbrich H, Beales PL, Yagi T, Mussaffi H, Chung EM, Omran H, Mitchell DR. Mutations in axonemal dynein assembly factor DNAAF3 cause primary ciliary dyskinesia. *Nat Genet* 2012: 44(4): 381-389, S381-382.
25. Tarkar A, Loges NT, Slagle CE, Francis R, Dougherty GW, Tamayo JV, Shook B, Cantino M, Schwartz D, Jahnke C, Olbrich H, Werner C, Raidt J, Pennekamp P, Abouhamed M, Hjeij R, Kohler G, Griese M, Li Y, Lemke K, Klena N, Liu X, Gabriel G, Tobita K, Jaspers M, Morgan LC, Shapiro AJ, Letteboer SJ, Mans DA, Carson JL, Leigh MW, Wolf WE, Chen S, Lucas JS, Onoufriadis A, Plagnol V, Schmidts M, Boldt K, Uk10K, Roepman R, Zariwala MA, Lo CW, Mitchison HM, Knowles MR, Burdine RD, Loturco JJ, Omran H. DYX1C1 is required for axonemal dynein assembly and ciliary motility. *Nat Genet* 2013: 45(9): 995-1003.
26. Knowles MR, Leigh MW, Carson JL, Davis SD, Dell SD, Ferkol TW, Olivier KN, Sagel SD, Rosenfeld M, Burns KA, Minnix SL, Armstrong MC, Lori A, Hazucha MJ, Loges NT, Olbrich H, Becker-Heck A, Schmidts M, Werner C, Omran H, Zariwala MA, Genetic Disorders of Mucociliary Clearance C. Mutations of DNAH11 in patients with primary ciliary dyskinesia with normal ciliary ultrastructure. *Thorax* 2012: 67(5): 433-441.
27. Shoemark A, Burgoyne T, Kwan R, Dixon M, Patel MP, Rogers AV, Onoufriadis A, Scully J, Daudvohra F, Cullup T, Loebinger MR, Wilson R, Chung EMK, Bush A, Mitchison HM, Hogg C. Primary ciliary dyskinesia with normal ultrastructure: three-dimensional tomography detects absence of DNAH11. *Eur Respir J* 2018: 51(2).
28. Lucas JS, Adam EC, Goggin PM, Jackson CL, Powles-Glover N, Patel SH, Humphreys J, Fray MD, Falconnet E, Blouin JL, Cheeseman MT, Bartoloni L, Norris DP, Lackie PM. Static respiratory cilia associated with mutations in Dnahc11/DNAH11: a mouse model of PCD. *Hum Mutat* 2012: 33(3): 495-503.
29. Hornef N, Olbrich H, Horvath J, Zariwala MA, Fliegauf M, Loges NT, Wildhaber J, Noone PG, Kennedy M, Antonarakis SE, Blouin JL, Bartoloni L, Nusslein T, Ahrens P, Griese M, Kuhl H, Sudbrak R, Knowles MR, Reinhardt R, Omran H.

DNAH5 mutations are a common cause of primary ciliary dyskinesia with outer dynein arm defects. *American journal of respiratory and critical care medicine* 2006: 174(2): 120-126.
30. Berg JS, Evans JP, Leigh MW, Omran H, Bizon C, Mane K, Knowles MR, Weck KE, Zariwala MA. Next generation massively parallel sequencing of targeted exomes to identify genetic mutations in primary ciliary dyskinesia: implications for application to clinical testing. *Genet Med* 2011: 13(3): 218-229.
31. Djakow J, Kramna L, Dusatkova L, Uhlik J, Pursiheimo JP, Svobodova T, Pohunek P, Cinek O. An effective combination of sanger and next generation sequencing in diagnostics of primary ciliary dyskinesia. *Pediatric pulmonology* 2016: 51(5): 498-509.
32. Olbrich H, Haffner K, Kispert A, Volkel A, Volz A, Sasmaz G, Reinhardt R, Hennig S, Lehrach H, Konietzko N, Zariwala M, Noone PG, Knowles M, Mitchison HM, Meeks M, Chung EM, Hildebrandt F, Sudbrak R, Omran H. Mutations in DNAH5 cause primary ciliary dyskinesia and randomization of left-right asymmetry. *Nat Genet* 2002: 30(2): 143-144.
33. Failly M, Bartoloni L, Letourneau A, Munoz A, Falconnet E, Rossier C, de Santi MM, Santamaria F, Sacco O, DeLozier-Blanchet CD, Lazor R, Blouin JL. Mutations in DNAH5 account for only 15% of a non-preselected cohort of patients with primary ciliary dyskinesia. *J Med Genet* 2009: 46(4): 281-286.
34. Raidt J, Wallmeier J, Hjeij R, Onnebrink JG, Pennekamp P, Loges NT, Olbrich H, Haffner K, Dougherty GW, Omran H, Werner C. Ciliary beat pattern and frequency in genetic variants of primary ciliary dyskinesia. *Eur Respir J* 2014: 44(6): 1579-1588.
35. Knowles MR, Leigh MW, Ostrowski LE, Huang L, Carson JL, Hazucha MJ, Yin W, Berg JS, Davis SD, Dell SD, Ferkol TW, Rosenfeld M, Sagel SD, Milla CE, Olivier KN, Turner EH, Lewis AP, Bamshad MJ, Nickerson DA, Shendure J, Zariwala MA, Genetic Disorders of Mucociliary Clearance C. Exome sequencing identifies mutations in CCDC114 as a cause of primary ciliary dyskinesia. *Am J Hum Genet* 2013: 92(1): 99-106.
36. Xiong HY, Alipanahi B, Lee LJ, Bretschneider H, Merico D, Yuen RK, Hua Y, Gueroussov S, Najafabadi HS, Hughes TR, Morris Q, Barash Y, Krainer AR, Jojic N, Scherer SW, Blencowe BJ, Frey BJ. RNA splicing. The human splicing code reveals new insights into the genetic determinants of disease. *Science* 2015: 347(6218): 1254806.
37. Boaretto F, Snijders D, Salvoro C, Spalletta A, Mostacciuolo ML, Collura M, Cazzato S, Girosi D, Silvestri M, Rossi GA, Barbato A, Vazza G. Diagnosis of Primary Ciliary Dyskinesia by a Targeted Next-Generation Sequencing Panel: Molecular and Clinical Findings in Italian Patients. *The Journal of molecular diagnostics : JMD* 2016: 18(6): 912-922.
38. Zariwala MA, Leigh MW, Ceppa F, Kennedy MP, Noone PG, Carson JL, Hazucha MJ, Lori A, Horvath J, Olbrich H, Loges NT, Bridoux AM, Pennarun G, Duriez B, Escudier E, Mitchison HM, Chodhari R, Chung EM, Morgan LC, de Iongh RU, Rutland J, Pradal U, Omran H, Amselem S, Knowles MR. Mutations of DNAI1 in primary ciliary dyskinesia: evidence of founder effect in a common mutation. *American journal of respiratory and critical care medicine* 2006: 174(8): 858-866.
39. Wirschell M, Olbrich H, Werner C, Tritschler D, Bower R, Sale WS, Loges NT, Pennekamp P, Lindberg S, Stenram U, Carlen B, Horak E, Kohler G, Nurnberg P, Nurnberg G, Porter ME, Omran H. The nexin-dynein regulatory complex subunit DRC1 is essential for motile cilia function in algae and humans. *Nat Genet* 2013: 45(3): 262-268.
40. Olcese C, Patel MP, Shoemark A, Kiviluoto S, Legendre M, Williams HJ, Vaughan CK, Hayward J, Goldenberg A, Emes RD, Munye MM, Dyer L, Cahill T, Bevillard J, Gehrig C, Guipponi M, Chantot S, Duquesnoy P, Thomas L, Jeanson L, Copin B, Tamalet A, Thauvin-Robinet C, Papon JF, Garin A, Pin I, Vera G, Aurora P, Fassad MR, Jenkins L, Boustred C, Cullup T, Dixon M, Onoufriadis A, Bush A, Chung EM, Antonarakis SE, Loebinger MR, Wilson R, Armengot M, Escudier E, Hogg C, Group UKR, Amselem S, Sun Z, Bartoloni L, Blouin JL, Mitchison HM. X-linked primary ciliary dyskinesia due to mutations in the cytoplasmic axonemal dynein assembly factor PIH1D3. *Nature communications* 2017: 8: 14279.
41. Kott E, Legendre M, Copin B, Papon JF, Dastot-Le Moal F, Montantin G, Duquesnoy P, Piterboth W, Amram D, Bassinet L, Beucher J, Beydon N, Deneuville E, Houdouin V, Journel H, Just J, Nathan N, Tamalet A, Collot N, Jeanson L, Le Gouez M, Vallette B, Vojtek AM, Epaud R, Coste A, Clement A, Housset B, Louis B, Escudier E, Amselem S. Loss-of-function mutations in RSPH1 cause primary ciliary dyskinesia with central-complex and radial-spoke defects. *Am J Hum Genet* 2013: 93(3): 561-570.
42. Daniels ML, Leigh MW, Davis SD, Armstrong MC, Carson JL, Hazucha M, Dell SD, Eriksson M, Collins FS, Knowles MR, Zariwala MA. Founder mutation in RSPH4A identified in patients of Hispanic descent with primary ciliary dyskinesia. *Hum Mutat* 2013: 34(10): 1352-1356.
43. Castleman VH, Romio L, Chodhari R, Hirst RA, de Castro SC, Parker KA, Ybot-Gonzalez P, Emes RD, Wilson SW, Wallis C, Johnson CA, Herrera RJ, Rutman A, Dixon M, Shoemark A, Bush A, Hogg C, Gardiner RM, Reish O, Greene ND, O'Callaghan C, Purton S, Chung EM, Mitchison HM. 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.

44. Marshall CR, Scherer SW, Zariwala MA, Lau L, Paton TA, Stockley T, Jobling RK, Ray PN, Knowles MR, Hall DA, Dell SD, Kim RH. Whole-Exome Sequencing and Targeted Copy Number Analysis in Primary Ciliary Dyskinesia. *G3 (Bethesda, Md)* 2015: 5(8): 1775-1781.
45. Moore DJ, Onoufriadis A, Shoemark A, Simpson MA, zur Lage PI, de Castro SC, Bartoloni L, Gallone G, Petridi S, Woollard WJ, Antony D, Schmidts M, Didonna T, Makrythanasis P, Bevillard J, Mongan NP, Djakow J, Pals G, Lucas JS, Marthin JK, Nielsen KG, Santoni F, Guipponi M, Hogg C, Antonarakis SE, Emes RD, Chung EM, Greene ND, Blouin JL, Jarman AP, Mitchison HM. Mutations in ZMYND10, a gene essential for proper axonemal assembly of inner and outer dynein arms in humans and flies, cause primary ciliary dyskinesia. *Am J Hum Genet* 2013: 93(2): 346-356.