Genetic Diagnosis and Respiratory Management of Primary Ciliary Dyskinesia

Paff, T.

2017

document version
Publisher's PDF, also known as Version of record

4

DIAGNOSTIC YIELD OF A TARGETED GENE PANEL IN DUTCH PRIMARY CILIARY DYSKINESIA PATIENTS

Tamara Paff; Irsan E. Kooi; Youssef Moutaouakil; Elise Riesebos; Erik A. Sistermans; Johannes M.A. Daniels; Janneke M. Weiss; Hans W.M. Niessen; Eric G. Haarman; Gerard Pals; Dimitra Micha

Peer review by Human Mutation (August 2017)

ABSTRACT

We evaluated the expression profile of genes related to primary ciliary dyskinesia (PCD) during *in vitro* ciliogenesis. We selected candidate genes with a similar expression profile and tested the diagnostic yield of targeted-exome sequencing in a Dutch cohort of PCD patients. Twelve healthy non-smoking adults underwent nasal epithelial curette biopsies. We determined differential gene expression in ciliated cells compared to epithelial precursor cells by RNA sequencing. A targeted panel of 26 PCD-related genes and 284 candidate genes with significant upregulation in ciliated cells was sequenced in 74 Dutch PCD patients. All genes reported as PCD causative, except *NME8*, showed significant upregulation during *in vitro* ciliogenesis of human airway cells.

We observed a 67.6% diagnostic yield when testing the targeted-exome panel in a Dutch cohort of 74 patients. The candidate genes included two recently published PCD-related genes *DNAJB13* and *PIH1D3*; identification of the latter was a direct result of this study.We demonstrate a highly sensitive and moderately specific approach for identification of PCD-related genes, based on significant gene upregulation during *in vitro* ciliogenesis. Targeted-exome sequencing had a diagnostic yield of 67.6% in a Dutch PCD population with relatively high percentage of *DNAI and HYDIN* mutations.

INTRODUCTION

Primary ciliary dyskinesia (PCD) is a rare inherited disorder, characterized by dysfunction of motile cilia that line the respiratory tract, fallopian tubes and embryonic node [1]. PCD patients subsequently suffer from recurrent airway infections, situs abnormalities and sub– or infertility [2]. Early diagnosis is important to improve respiratory outcome [3–6]. However, PCD is underdiagnosed or diagnosed late due to unawareness of physicians and complex diagnostic investigations, requiring expensive infrastructure and an experienced team [7, 8]. Diagnosis is primarily based on a combination of evaluation of ciliary ultrastructure by transmission electron microscopy (TEM) and motility by high-speed videomicroscopy (HSVM) [9, 10]. As the number of identified PCD-related genes increases, the North American Genetic Disorders of Mucociliary Clearance Consortium recently proposed a more central role for genetic testing in the diagnostic pathway [11].

Motile cilia have a complex ultrastructure composed of microtubule doublets connecting to multi-protein complexes, such as dynein arms, radial spokes and nexin-dynein regulatory complexes [1]. The axoneme comprises of > 200 distinct proteins [12]. Theoretically, defects in any one of the proteins required to build or regulate the cilium may cause PCD [13]. Therefore, many efforts currently aim to unravel the entire ciliome. Several databases have incorporated results of proteomic, transcriptomic and comparative genomic studies, of which the Cildb is the most comprehensive, covering 55 high-throughput studies in 32 species with cilia or flagella [14]. Regarding these incorporated human ciliary studies, most of the currently known PCD-related genes have been captured by investigating differential gene expression during *in vitro* ciliogenesis of human epithelium using an air-liquid interface (ALI) culture system [15]. This demonstrates that a cell culture system can be a unique model to examine the factors and pathways that regulate mucociliary differentiation, to potentially identify novel PCD genes. As the proportion of ciliated cells that grow in ALI cultures may be less than in human bronchial epithelium, the motile cilia signature may be underestimated with this approach, providing a possible explanation for the lack of detection of several PCD-related genes [16]. In addition to using a candidate gene approach that has led to identification of various PCD-related genes, the widespread use of next generation sequencing (NGS) allowed a large part of recent gene discovery. The extensive genetic heterogeneity in PCD makes NGS approaches also attractive for use in the diagnostic setting, as it offers parallel sequencing of multiple genes or of the entire exome. The diagnostic yield of targeted NGS panels used in Europe and the US currently varies between 43-76% [17–19]. This means that genetic testing cannot yet provide enough certainty to be used as the sole test for PCD. However, considering the pace of new discoveries and advances in sequencing technologies and data analysis, it is reasonable to believe that genetic analysis will be established as the

preferred diagnostic tool in the future. Genetic characterization of international PCD cohorts is pivotal to achieve this. Genetic testing may contribute to the confirmation of PCD diagnosis in patients that have been diagnosed solely based on clinical symptoms or in which results remain inconclusive. Further, genotyping facilitates evaluation and follow-up of genotype-phenotype relationships and adequate genetic counseling of families [20–22]. In this study we therefore aimed to determine the genetic defects in Dutch PCD patients with a targeted gene panel. The gene panel included 26 PCD-related genes known at the beginning of this study and a set of candidate genes. To prioritize the PCD candidate genes we investigated the transcriptome of human airway cells in a monolayer-suspension cell culture, which is aimed at culturing ciliated cells. We hypothesized that PCD-related genes show a distinct expression profile with significant upregulation during ciliogenesis. We tested this proof of principle by using RNA sequencing to analyse global mRNA expression in human airway cells of healthy individuals in a sequential monolayer-suspension cell culture system in which the mRNA expression profile of ciliary cells, grown in suspension, was compared with unciliated precursor cells, grown as monolayer on collagen.

METHODS AND MATERIALS

Subjects

Twelve healthy non-smoking adults were included to undergo nasal curette biopsies for studying differential gene expression during *in vitro* ciliogenesis. All Dutch university hospitals and the PCD patient organisation were requested to inform patients about the study. Participants received detailed information from the researcher upon request. PCD patients from the town of Volendam were excluded as a founder mutation has been identified recently [23]. Informed consent was obtained from all participants prior to study procedures. Studies were approved by the medical ethical committee of the VU University Medical Center.

PCD diagnosis

PCD diagnosis has evolved quickly in the past years, so not all patients have been diagnosed in a similar way [24]. Therefore we classified patients according to European Respiratory Society guidelines as "definite PCD" when they had clinical symptoms and an ultrastructural defect, observed by TEM [10]. Patients with clinical symptoms, low nasal nitric oxide and/or with a motility defect observed by HSVM were classified as "probable PCD" and patients with a previous diagnosis based on only clinical symptoms classified as "PCD clinical diagnosis" [24]. We compared diagnostic status of patients before and after genetic testing.

In vitro ciliogenesis

An experienced pediatric pulmonologist collected nasal epithelial cells from 12 healthy, non-smoking adult subjects [25]. Cells were cultured using a sequential monolayer-suspension culture technique, adapted from Jorissen et al. [26]. For this, cells were seeded on 0.2% rat tail collagen matrix followed by culturing in monolayer medium consisting of Dulbecco's Modified Eagle Medium F-12 Nutrient Mixture media supplemented with 2% Ultroser G, 1% Glutamax and 1% penicillin/streptomycin at 37ºC and 5% CO_2. The monolayer medium promoted the growth of precursor cells without cilia (monolayer phase). After two weeks of culture, cells reached confluency. Subsequently, the collagen matrix was digested with Clostridium histolyticum collagenase to detach the cell layer. The cells were then cultured in suspension medium consisting of Dulbecco's Modified Eagle Medium F-12 Nutrient Mixture media supplemented with 10% NU serum, 1% Glutamax and 1% penicillin/streptomycin on a shaking platform at 55tpm for 1 week. The cells were allowed to grow for 2 more weeks in suspension media during which they gradually became fully ciliated (suspension phase). Ciliogenesis was confirmed by microscopic examination. HSVM was used to examine ciliary function with regard to frequency, amplitude as well as intracellular and intercellular coordination.

Cell harvest and RNA isolation

Cells from 12 subjects were harvested at 2 different time points during the monolayer-suspension cell culture to isolate RNA for RNA sequencing analysis: (T1) on day 14 (unciliated cells in monolayer phase) and (T2) on day 42 (28 days after re-differentiation switch (suspension phase); visible ciliated cells) (figure 1). Normal ciliary beat frequency and beat pattern at T2 were confirmed by HSVM prior to RNA isolation. Cells were centrifuged at 300g for 10min after which the cell pellet was resuspended in 1ml Earle's Balanced Salt Solution. After centrifugation at 300g for 2min the cell pellet was used to isolate RNA with the Quick-RNA Miniprep kit (ZYMO Research, Irvine, CA, USA), according to the manufacturer's protocol. RNA concentration and quality of the RNA samples was determined using the NanoDrop 1000 (Thermo Fisher Scientific, Waltham, MA, USA) and/or TapeStation 2200 (Agilent technologies, Santa Clara, CA, USA). RNA was stored at -80°C until further processing.

Gene expression analysis by RNA sequencing

RNA samples were prepared for sequencing using the TruSeq Stranded Total mRNA Library Prep Kit (Illumina Inc. San Diego, USA) according to the manufacturer's instruction. Quality of the cDNA libraries was accessed using TapeStation 2200 (Agilent technologies, Santa Clara, CA, USA). Bar-coded libraries were pooled and run on an Illumina HiSeq 2500 (San Diego, CA, USA) for sequencing. Sequencing reads were cleaned by 5'-end quality trimming and

Illumina-adapter clipping by Trimmomatic [27]. Pre-alignment quality control of the cleaned sequencing reads was done with FastQC [28]. Spliced alignment to reference genome hg19/GRCh37 of cleaned sequencing reads was done with STAR, guided by gene annotation in the refGene UCSC table [29]. Post-alignment quality control including genebody coverage analysis was done with RSeQC [30]. Read summarization was performed with HTseq using strandness-aware union intersection of uniquely aligned reads (mapping quality threshold 35) with refGene gene annotation [31]. A pre-normalization filter was used excluding genes that had less than 5 reads in more than 80% of the samples. Using edgeR, trimmed mean of M-values (TMM) normalization was performed followed by differential expression between monolayer (cilium deficient) and suspension (cilium proficient), paired by patient using a generalized linear model [32]. Genes with false discovery rate (FDR) adjusted p-values <0.05 and logCPM >0 were considered significant. Downstream analyses and visualizations of differential gene expression including heat maps and boxplots were performed using R version 3.2.2.

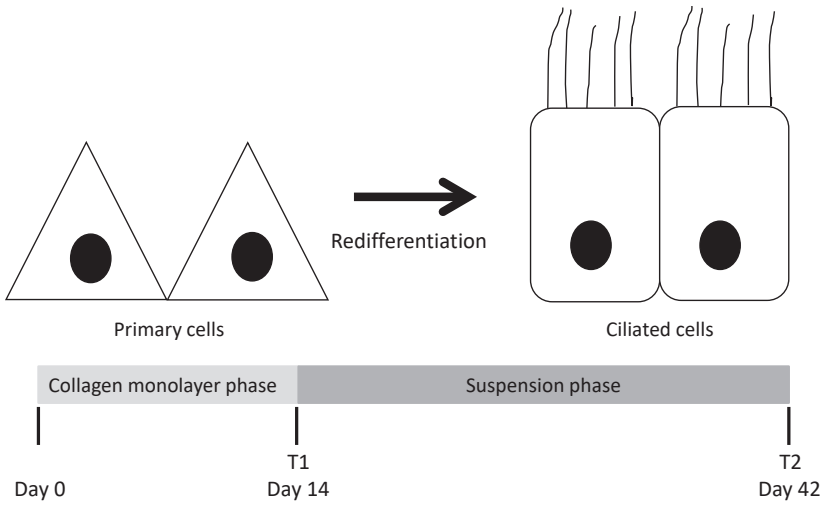

Figure 1. Schematic overview of in vitro ciliogenesis

Schematic overview of in in vitro ciliogenesis using a collagen monolayer-suspension cell culture system. RNA was isolated at T1 (after 14 days) and at T2 (after 42 days).

DNA extraction

Saliva from PCD patients was collected using Oragene-250 self-collection saliva kits (DNA Genotek Inc., Canada) that were delivered to patients' homes. DNA extraction was performed according to the manufacturer's protocol. Briefly, the collected cells in the saliva were subjected to lysis using a purifying buffer provided in the kit to remove proteins. This was followed by by incubation on ice and DNA precipitation with 100% ethanol. The DNA

was rehydrated in TE (Tris–HCl 1 mM pH 8.0 and 0.5 mM EDTA pH 8.0). DNA quantity and quality were measured by ultraviolet spectrometry with the NanoDrop 1000 (Thermo Fisher Scientific, USA). DNA samples were stored at -20°C.

Targeted gene panel

Twenty-six PCD-related genes that were known at the time of composing the gene panel were included (table S4.1). Additional candidate genes were selected from the DEG list based on showing a clear link to ciliary or flagellar proteins or processes in published candidate gene lists and the Gene Network database (table S4.1) [14].

DNA Sequencing

DNA was sheared by sonification to an average size of 200 bp with the Covaris S220 (Covaris, Woburn, MA, USA). Sheared DNA was subjected to Illumina paired-end DNA library preparation and enriched for target sequences (MYbaits Target Enrichment kit, MYcroarray, USA). Target sequences included the 310 genes listed in table S4.1 or the entire exome combined with in silico analysis of the genes listed. Enriched libraries were sequenced with the HiSeq 2500 platform (Illumina, USA) as paired-end 100 bp reads. Sequencing reads were cleaned by 5'-end quality trimming and Illumina-adapter clipping by Trimmomatic [27]. Pre-alignment quality control of the cleaned sequencing reads was done with FastQC [28]. 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 were annotated with ANNOVAR.

Analysis of variant pathogenicity after sequencing

Filtering of variants for novelty was performed by exclusion of variants with frequency >0.01 in the 1000 Genomes database or the Exome Sequencing Project (ESP). We focused on nonsynonymous mutations, frame-shift mutations, splice-site mutations and indels following an autosomal-recessive or X-linked inheritance pattern. Variants in one of the PCD-related genes were considered disease-causing when they correlated with ultrastructural and/or ciliary motility defect observed by electron microscopy and HSVM, respectively. Expected pathogenicity of nonsynonymous mutations was evaluated by a combination of prediction scores such as PolyPhen, SIFT and Mutation Taster. The predicted effect of splice-site mutations (Alamut Visual, Interactive Biosoftware, USA) were confirmed in cDNA from respiratory epithelial cells after sequential monolayer-suspension culture whenever possible. For this RNA was isolated as previously described. cDNA syhthesis was performed with the SuperScript® VILO cDNA Synthesis Kit (Thermo fisher Scientifi Inc., Waltham, USA) according to manufacturer's instructions.

Analysis of variants in *HYDIN*

As *HYDIN* has a pseudogene on chromosome 1 (*HYDIN2*) with very high homology to exon 6-83, we confirmed likely pathogenic *HYDIN* variants, occuring in around 25% or 50% of sequence data, with allele-specific PCRs. We designed the primers by aligning *HYDIN* and *HYDIN2* genomic DNA sequences in BLAST to have a single base pair difference at the 3' ends [33]. Primer sequences are presented in table S4.2. We expected that sequencing reads that could be mapped to both the *HYDIN gene* and the *HYDIN2* gene would automatically be discarded. We re-mapped the sequencing data to the reference genome after removing the *HYDIN2* gene from this genome on chromosome 1 to investigate whether this method allows identification of additional variants that would have been missed by standard analysis. Suspected variants were resequenced with Sanger sequencing, using PCR primers specific for the active gene on chromosome 16.

RESULTS

Differential gene expression during in vitro ciliogenesis using RNA sequencing

RNA samples from one healthy control did not pass quality control steps and were discarded from the analysis. Two differentially expressed gene clusters ($p<0.05$, FDR<0.05) were identified (figure 2). One cluster (5198/25963, 20.0 %) included genes with increased expression during the monolayer phase of cell culture (T1) and the other cluster included genes with increased expression during the suspension phase (T2) (5499/25963, 21.2%).The latter cluster included 35 of 36 (all but *NME8*) currently known PCD-related genes (figure 3, table S4.3). In contrast to other PCD-related genes, *NME8* showed no expression using RNA sequencing of human airway cells in 12 healthy controls. All other PCD-related genes either showed relatively low expression before ciliogenesis and high expression during ciliogenesis or a relatively high expression throughout the entire experiment, whereas *NME8* showed no expression throughout the entire experiment. We propose this gene cluster can be best used in combination with exome data to prioritize analysis of gene variants that may play an important role in ciliogenesis, thereby aiding PCD gene discovery.

Diagnostic yield of the targeted exome panel

A total of 83 PCD patients applied to participate in this study, of which 74 were included (figure 4). In 50/74 patients (67.6%) we found biallelic mutations that are expected to be pathogenic in a total of 12 PCD-related genes, including in the recently identified *PIH1D3* gene (table 1, figure 5). Most gene defects were found in *DNAH5* (13/74 patients, 17.6%) and *DNAI1* (12/74 patients, 16.2%). The relatively high proportion of *DNAI1* mutations in our cohort appears largely due to the common c.48+2dup (also known as IVS1+2insT) founder

mutation that was present in 15/24 (62.5%) alleles in patients with biallelic mutations in *DNAI1*. Further, we observed a very high frequency (6/74, 8.1%) of *HYDIN* mutations in our population. We used *HYDIN*-specific primers to confirm likely pathogenic mutations. Re-mapping sequencing data, after removing the pseudogene (*HYDIN2)* from the reference genome resulted in two additional variants that would otherwise have been missed. Further, we identified a hemizygous mutation in the novel X-linked gene *PIH1D3* in a male with PCD without syndromic coseggregation, which has recently been published [34]. Among the 284 selected candidate genes in our gene panel was also *DNAJB13*, which has only recently been identified as a novel PCD-related gene [35]. In 7/74 patients (9.5%) we only observed monoallelic mutations in currently known PCD-related genes (table 2). In these cases we may have missed deep intronic mutations that may lead to splice defects. In 17/74 patients (23.0%) the analysis revealed no disease-causing mutations. These patients all had multiple nonsynonymous variants of unknown significance (VUS) in multiple PCD-related genes and/or candidate genes. Figure 4 gives an overview of the diagnostic classification of PCD patients before and after genetic testing.

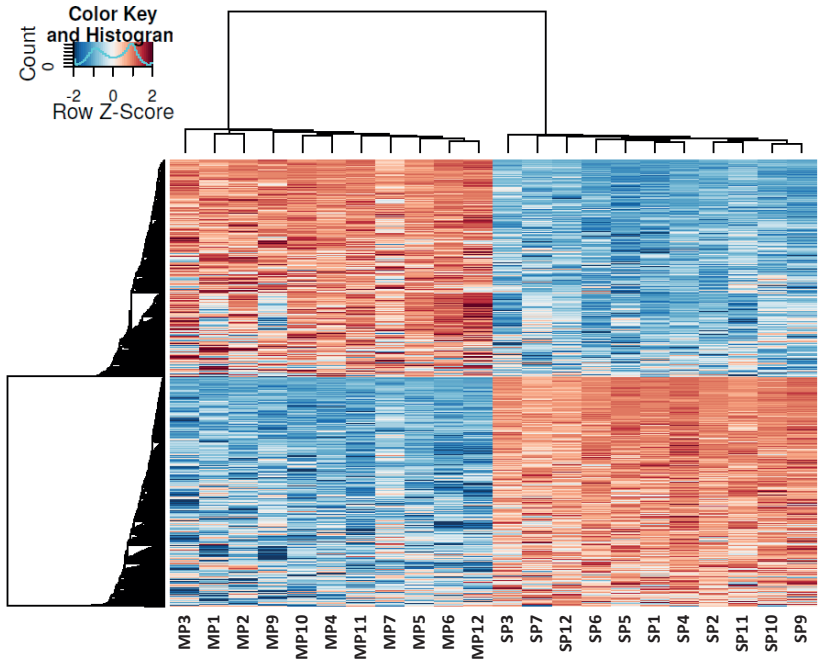

Figure 2. Heatmap of gene expression during in vitro ciliogenesis

Heat map showing gene expression in human respiratory epithelial cells during collagen monolayer-suspension cell culture. Changes at two different timepoints (T1: monolayer phase and T2: suspension phase) of 11 samples are depicted. Red colour indicates high expression and blue indicates low expression. MP: monolayer phase; SP: suspension phase.

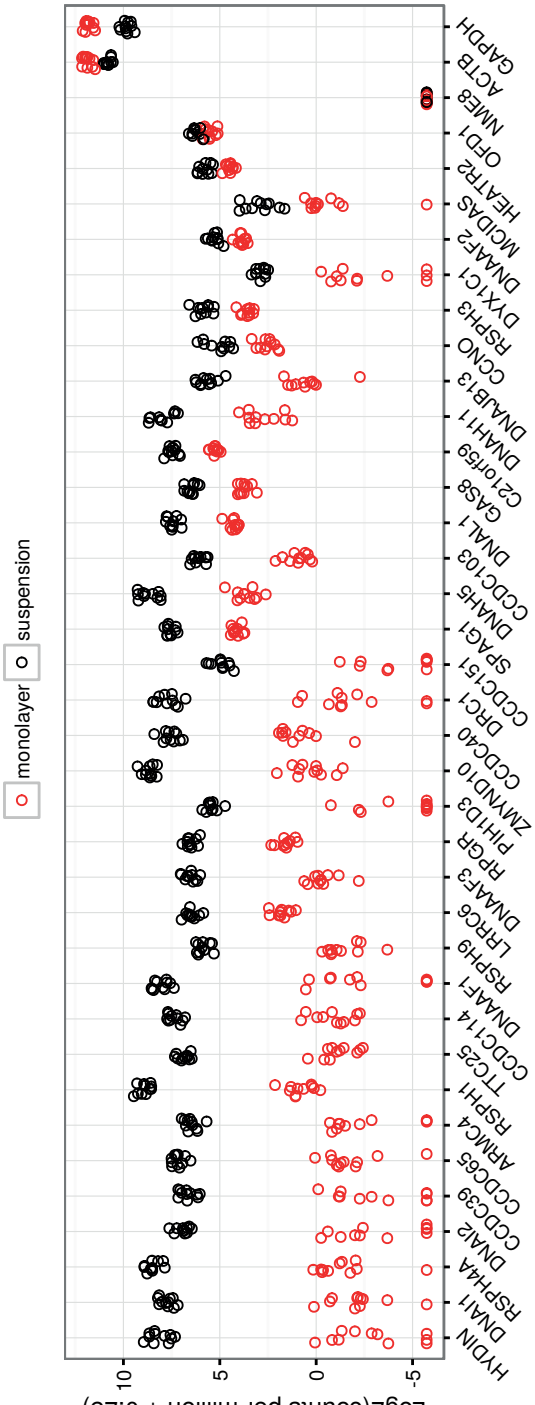

Figure 3. Grid of gene expression of PCD-related genes during in vitro ciliogenesis

Graphical representation of gene expression for all PCD-related genes. Levels of gene expression in 11 samples of human respiratory epithelial cells are shown in red at T1 (monolayer phase) and in black at T2 (suspension phase) on a log2 (counts per million) scale. ACTB and GAPDH are added to compare gene expression of household genes.

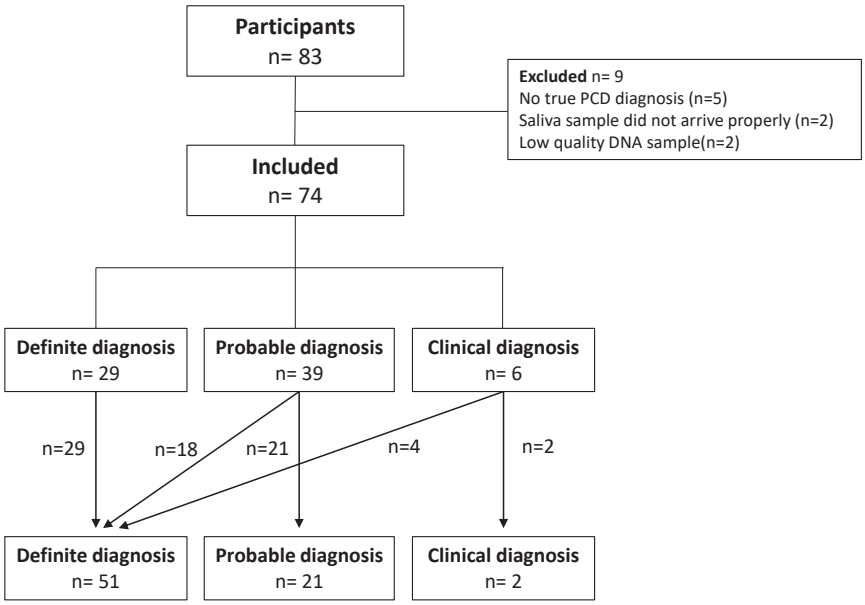

Figure 4. Overview of diagnostic certainty in PCD patients before and after genetic testing.

Overview of diagnostic certainty in PCD patients included in the study, before and after genetic testing. Definite diagnosis: hallmark ultrastructural defect and/or biallelic pathogenic mutation; Probable diagnosis: hallmark ciliary motion defect and/or low nasal nitric oxide; Clinical diagnosis: only clinical PCD symptoms.

Table 1. Overview of biallelic mutations related to PCD diagnosis.

Patient ID	Gene	Variants and protein change	Effect	Zygosity	Motion defect	Ultrastructural defect	Ref.
PCD-2035	DNAH5	c.10384C>T:p.Q3462X; c.10815delT:p.D3606Hfs*23	Stopgain; Frameshift deletion	HET	Unknown	ODA + IDA	[17, 36]
PCD-0106	DNAH5	c.4664G>A:pW1555X; c.376delG:p.V126Yfs*23	Stopgain; Frameshift deletion	HET	Immotile	ODA + IDA	Novel
PCD-0032	DNAH5	c.10384C>T:p.Q3462X; c.13338+5G>A	Stopgain; Splicing	HET	Immotile	ODA + IDA	[17, 36]
PCD-0031	DNAH5	c.13729C>T:p.R4577X; c.3989_3990del:p.E1330Afs*3	Stopgain; Frameshift deletion	HET	Immotile	ODA + IDA	Novel
PCD-0103	DNAH5	c.5495T>A:p.L1832X; c.10815delT:p.D3606Hfs*23	Stopgain; Frameshift deletion	HET	Immotile	ODA	Novel; [36]
PCD-0118	DNAH5	c.10384C>T:p.Q3462X; c.8368delT:p.Y2790Mfs*16	Stopgain; Frameshift deletion	HET	Immotile	ODA + IDA	[17]; Novel
PCD-2040	DNAH5	c.2578-1G>A; c.10815delT:p.D3606Hfs*23	Splicing; Frameshift deletion	HET	Immotile	ODA	Novel; [36]
PCD-2057	DNAH5	c.10384C>T:p.Q3462X; c.4360C>T:p.R1454X	Stopgain; Stopgain	HET	Unknown	ODA + IDA	[17, 37]
PCD-3058 [Sib. 3059]	DNAH5	c.1089+1G>A; c.5177T>C:p.L1726P	Splicing; Nonsynonymous	HET	Immotile	ODA+IDA	[17, 38]
PCD-3059	DNAH5	c.1089+1G>A; c.5177T>C:p.L1726P	Splicing; Nonsynonymous	HET	Immotile	ODA + IDA	[17, 38]
PCD-1162	DNAH5	c.5647C>T:p.R1883*; c.1852C>T:p.R618*	Stopgain; Stopgain	HET	Immotile	ODA	[38]; Novel
PCD-3152 [Sib. 3153]	DNAH5	c.1730+2T>C; c.1089+1G>A	Splicing; Splicing	HET	Immotile	Unknown	Novel; [38]
PCD-3153	DNAH5	c.1730+2T>C; c.1089+1G>A	Splicing; Splicing	HET	Immotile	ODA	Novel; [38]
PCD-0051	DNAI1	c.48+2dup; c.48+2dup	Splicing; Splicing	HOM	Immotile	ODA + IDA	[39]

Table 1. Overview of biallelic mutations related to PCD diagnosis. *(Continued)*

Patient ID	Gene	Variants and protein change	Effect	Zygosity	Motion defect	Ultrastructural defect	Ref.
PCD-0108	DNAI1	c.48+2dup; c.48+2dup	Splicing; Splicing	HOM	Low frequency, stiff	ODA	[39]
PCD-0154	DNAI1	c.48+2dup; c.48+2dup	Splicing; Splicing	HOM	Immotile	ODA	[39]
PCD-3146	DNAI1	c.48+2dup; c.48+2dup	Splicing; Splicing	HOM	Low frequency, stiff	ODA	[39]
PCD-8073	DNAI1	c.48+2dup; c.48+2dup	Splicing; Splicing	HOM	Unknown	Unknown	[39]
PCD-0102	DNAI1	c.48+2dup; c.1229delT:p.I410Afs*61	Splicing; Frameshift deletion	HET	Immotile	ODA	[39]; Novel
PCD-5164	DNAI1	c.48+2dup; c.1031+5G>T	Splicing; Splicing	HET	Immotile	Unknown	[39]; Novel
PCD-5165 [Sib.5164]	DNAI1	c.48+2dup; c.1031+5G>T	Splicing; Splicing	HET	Low frequency, stiff	ODA	[39]; Novel
PCD-1136	DNAI1	c.814C>T:p.Q272X; c.814C>T:p.Q272X	Stopgain; Stopgain	HOM	Unknown	Unknown	Novel
PCD-1137 [Sib. 1136]	DNAI1	c.814C>T:p.Q272X; c.814C>T:p.Q272X	Stopgain; Stopgain	HOM	Almost no cilia, immotile	Unknown	Novel
PCD-8170	DNAI1	c.48+2dup; c.577_578delAC:p.T193*	Splicing; Frameshift deletion	HET	Immotile	ODA + IDA	[39]; Novel
PCD-8076	DNAI1	c.48+2dup; c.1243delT:p.Y415Tfs*56	Splicing; Frameshift deletion	HET	Immotile	Unknown	[39]; Novel
PCD-0029	HYDIN	c.6669+1G>A; c.6669+1G>A	Splicing; Splicing	HOM	Low frequency, stiff, lack of coordination	Normal	Novel
PCD-0043	HYDIN	c.8356C>T:p.R2786*; c.13867G>T:p.G4623*	Stopgain; Stopgain	HET	Low frequency, stiff, circular movement	Normal, some single tubuli	Novel
PCD-1124	HYDIN	c.5388C>A:p.Y1796*; c.12444-3C>G	Stopgain; Splicing	HET	Low frequency, stiff, lack of coordination	Normal, some single tubuli	Novel

Table 1. Overview of biallelic mutations related to PCD diagnosis. *(Continued)*

Patient ID	Gene	Variants and protein change	Effect	Zygosity	Motion defect	Ultrastructural defect	Ref.
PCD-8179	HYDIN	c.8674_8675delinsG:p.Q2892Gfs*3; c.8674_8675delinsG:p.Q2892Gfs*3	Frameshift delins; Frameshift delins	HOM	Low frequency, stiff, lack of coordination	Normal	Novel
PCD-8078 [Sib.8179]	HYDIN	c.8674_8675delinsG:p.Q2892Gfs*3; c.8674_8675delinsG: p.Q2892Gfs*3	Frameshift delins; Frameshift delins	HOM	Low frequency, stiff, lack of coordination	Unknown	Novel
PCD-8184	HYDIN	c.8674_8675delinsG: p.Q2892Gfs*3; c.8674_8675delinsG: p.Q2892Gfs*3	Frameshift delins; Frameshift delins	HOM	Low frequency, stiff, lack of coordination	Normal	Novel
PCD-2060	DNAAF1	c.811C>T:p.R271*; c.811C>T:p.R271*	Stopgain; Stopgain	HOM	Immotile	ODA + IDA	[40]
PCD-3044	DNAAF1	c.1528+2T>C; c.1528+2T>C	Splicing; Splicing	HOM	Low frequency, stiff, lack of coordination	ODA + IDA	Novel
PCD-3045 [Sib. 3044]	DNAAF1	c.1528+2T>C; c.1528+2T>C	Splicing; Splicing	HOM	Low frequency, stiff, lack of coordination	ODA + IDA	Novel
PCD-4067	DNAAF1	c.1528+2T>C; c.1528+2T>C	Splicing; Splicing	HOM	Low frequency, stiff, lack of coordination	ODA + IDA	Novel
PCD-0142	DNAH11	c.10174C>T:p.R3392C; c.12889C>T:p.R4297W	Nonsynonymous; Nonsynonymous	HET	High frequency, stiff	Normal	Novel; Novel
PCD-0038	DNAH11	c.9824A>C:p.Y3275S; c.13304-1G>A	Nonsynonymous; Splicing	HET	High frequency, stiff	Normal	Novel; Novel
PCD-8181	DNAH11	c.793_794insCAGCT:p.R265Pfs*5; c.10568+1G>A	Frameshift insertion; Splicing	HET	High frequency, stiff	Normal	Novel; Novel
PCD-2061	DNAH11	c.7913A>G: p.Q2638R; c.7913A>G: p.Q2638R	Nonsynonymous; Nonsynonymous	HOM	High frequency, stiff	Normal	Novel
PCD-0101	CCDC103	c.461A>C:p.H154P; c.461A>C:p.H154P	Nonsynonymous; Nonsynonymous	HOM	Low frequency, stiff, lack of coordination	ODA	[41]
PCD-0115	CCDC103	c.461A>C:p.H154P; c.461A>C:p.H154P	Nonsynonymous; Nonsynonymous	HOM	Low frequency, stiff, lack of coordination	ODA + IDA	[41]
PCD-0009	CCDC40	c.1677_1678insAC:p.E559Rfs*11; c.1677_1678insAC:p.E559Rfs*11	Frameshift insertion; Frameshift insertion	HOM	Low frequency, stiff, lack of coordination	IDA + oriëntation	Novel

Table 1. Overview of biallelic mutations related to PCD diagnosis. *(Continued)*

Patient ID	Gene	Variants and protein change	Effect	Zygosity	Motion defect	Ultrastructural defect	Ref.
PCD-0110	CCDC40	c.248delC:p.A83Vfs*84; c.C1855T:p.Q619*	Frameshift deletion; Stopgain	HET	Low frequency, stiff, lack of coordination	Oriëntation	[42]; Novel
PCD-0128	CCDC39	c.2347_2351del:p.F783Yfs*3; c.2347_2351del:p.F783Yfs*3	Frameshift deletion; Frameshift deletion	HOM	Low frequency, stiff,	IDA	Novel
PCD-3163	CCDC39	c.610-2A>G; c. -13 to c.7del	Splicing; Splicing	HET	Low frequency, stiff, lack of coordination	Oriëntation	[43]; Novel
PCD-0050	DNAAF2	c.1901T>C:p.F634S; c.998 C>T:p.A333V	Nonsynonymous; Splicing	HET	Immotile	ODA + partial IDA	Novel; Novel
PCD-3147	HEATR2	c.50_51insG:p.E19Gfs*5; c.1499G>T:p.C500F	Frameshift insertion; Nonsynonymous	HET	Low frequency, lack of coordination	ODA + IDA	Novel; Novel
PCD-3148 [Sib. 3147]	HEATR2	c.50_51insG:p.E19Gfs*5; c.1499G>T:p.C500F	Frameshift insertion; Nonsynonymous	HET	Low frequency, lack of coordination	Unknown	Novel; Novel
PCD-0030	CCNO	c.787insG:pR263Afs; c.787insG:pR263Afs	Frameshift insertion; Frameshift insertion	HOM	No cilia	No cilia	Novel
PCD-8083	PIH1D3	c.357_363delGGTGGGA:p.V120Lfs*6	Frameshift deletion	HEMI	Immotile	ODA + IDA	[34]

Overview of biallelic mutations and observed defects in ciliary motion and ultrastructure in PCD patients. HOM: homozygous; HET: heterozygous; HEMI: hemizygous; ODA: outer dynein arm; IDA: inner dynein arm; Ref.: reference. DNAH5 (NM_001369.2), DNAI1 (NM_001281428.1), HYDIN (NM_001270974.2), DNAAF1 (NM_178452.5), DNAH11 (NM_001277115.1), CCDC103 (NM_213607.2), CCDC40 (NM_017950.3), CCDC39 (NM_181426.1), HEATR2 (NM_017802.3), DNAAF2 (NM_018139.2), CCNO (NM_021147.3), PIH1D3 (NM_001169154.1).

Table 2. Overview of monoallelic mutations potentially related to PCD diagnosis

Patient ID	Gene	Variant and protein change	Effect	Zygosity	HSVM defect	EM defect	Ref.
PCD-8072	DNAI1	c.48+2dup; [unknown]	Splicing	HET	Immotile	CP defect, partial IDA	[39]
PCD-8077	DNAI1	c.48+2dup; [unknown]	Splicing	HET	Unknown	Unknown	[39]
PCD-1123	DNAH11	c.7279C>G:p.Q2426E; [unknown]	Nonsynonymous	HET	Lack of coordination	Normal	Novel
PCD-1127	DNAH11	c.3291C>A:p.S1097R; [unknown]	Nonsynonymous	HET	Lack of coordination	Normal	Novel
PCD-0041	CCDC39	c.1960G>A:pG654R; [unknown]	Nonsynonymous	HET	Low frequency, stiff, lack of coordination	Partial IDA	Novel
PCD-7171	HYDIN	c.14625_14626delGC:pP4876Lfs*16; [unknown]	Frameshift deletion	HET	Low frequency, stiff, lack of coordination	Normal	Novel
PCD-0011	HYDIN	c.14625_14626delGC:pP4876Lfs*16; [unknown]	Frameshift deletion	HET	Stiff	Normal	Novel

Overview of monoallelic mutations and observed defects in ciliary motion and ultrastructure in PCD patients. HOM: homozygous; HET: heterozygous; HEMI: hemizygous; ODA: outer dynein arm; IDA: inner dynein arm; Ref.: reference. DNAI1 (NM_001281428.1), DNAH11 (NM_001277115.1), CCDC39 (NM_181426.1), HYDIN (NM_001270974.2).

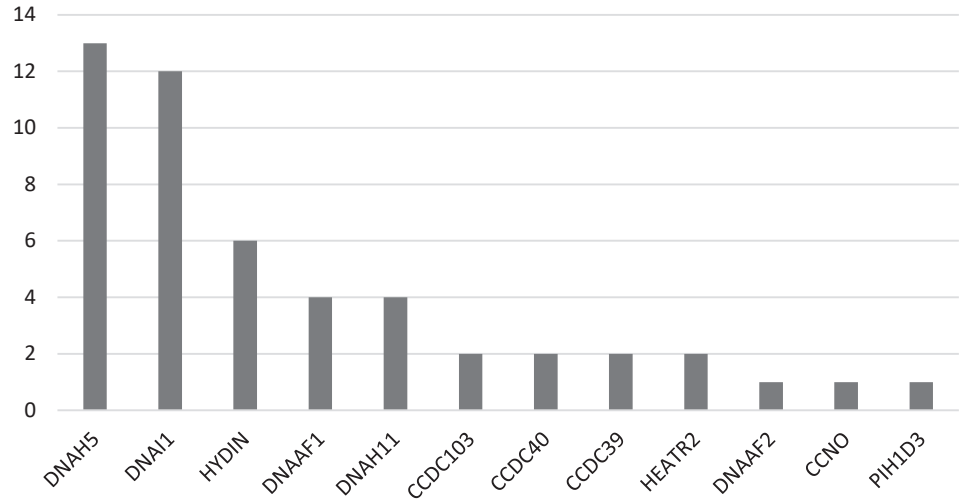

Figure 5. Overview of PCD-related genes in which bi-allelic pathogenic mutations were found in 74 Dutch PCD patients.

Data are depicted as numbers. DNAH5 (NM_001369.2), DNAI1 (NM_001281428.1), HYDIN (NM_001270974.2), DNAAF1 (NM_178452.5), DNAH11 (NM_001277115.1), CCDC103 (NM_213607.2), CCDC40 (NM_017950.3), CCDC39 (NM_181426.1), HEATR2 (NM_017802.3), DNAAF2 (NM_018139.2), CCNO (NM_021147.3), PIH1D3 (NM_001169154.1).

DISCUSSION

In the present study, we observed that all but one PCD-related gene (*NME8*) showed significant upregulation during *in vitro* ciliogenesis of human respiratory cells. We demonstrate that the identified differentially expressed gene cluster including ~ 5500 genes is highly sensitive for detection of currently known PCD-related genes, but also involves many unrelated genes. When using a targeted exome panel including a set of 26 PCD-related genes and 284 selected candidate genes, we observed a diagnostic yield of 67.6% in 74 Dutch PCD patients. The candidate genes included two recently identified PCD-related genes: *DNAJB13* and *PIH1D3*, of which the latter was a direct discovery of the current study [34, 35]. This underlines the value of using differentially expressed genes during *in vitro* ciliogenesis of human airway cells, to prioritize PCD candidate genes.

Differentially expressed gene cluster

In comparison to other human ciliary studies included in the Cildb, the DEG cluster analysis of ciliogenesis gene expression presented here is the only approach that detected nearly all currently known PCD-related genes. The unique aspect of our approach is the comparison of gene expression during in vitro ciliogenesis with gene expression in the precursor cells,

allowing us to identify upregulated genes. A cluster of ~5500 genes cannot be easily incorporated in a targeted-exome sequencing gene panel due to its large size. However, this cluster analysis as an *in silico* gene panel after whole exome sequencing, is a potentially powerful tool to prioritize variants in novel candidate PCD-related genes in patients that do not harbour pathogenic variants in any of the currently known PCD-related genes. This would also reduce the number of variants to investigate to approximately 21% of complete whole-exome data.

NME8

Interestingly, *NME8* gene expression during *in vitro* ciliogenesis was different from other PCD-related genes, with a lack of expression throughout the entire culture. Other PCD-related genes showed either 1) a relatively high expression throughout the two phases of the cell culture with a modest increase in the suspension phase in which ciliogenesis takes place or 2) a relatively low expression in the unciliated monolayer phase of cell culture followed by a major increase in expression during the suspension phase. Duriez and co-workers identified *NME8* (also known as *TXNDC3*) as a candidate PCD gene in 2007 as the human ortholog of the sea urchin gene encodes a component of sperm outer dynein arms [44]. They have described two compound heterozygous mutations in a girl with PCD from consanguineous parents. These include a nonsynonymous variant predicted to be pathogenic and a common variant predicted to lower expression of the *TXNDC3d7* isoform, which is suggested to be detrimental. The distinct behaviour of *NME8* in comparison to other known PCD-related genes could indicate that our approach is not 100% sensitive in identifying genes that are vital for proper ciliary function. The *in vitro* monolayer-suspension cell culture system may not reflect all processes that occur during *in vivo* ciliogenesis. On the other hand one might argue that the mutations described in *NME8* may not be causative of PCD in the described patient. It is unknown whether homozygous mutations, which are highly likely in a consanguiseous family, in other genes with ODA defects have been ruled out in this patient. To our knowledge, this is the only patient reported with possible disease-causing mutations in *NME8*. It is, however, not uncommon given that many private mutations have been reported in PCD.

Diagnostic yield of gene panel

The obtained diagnostic yield of 67.6% in a Dutch PCD cohort is comparable to the yield observed in three other studies in which genetic panels were used in combination with Sanger pre-sequencing or targeted copy number variation (CNV) analysis, aimed at published exon deletions in known PCD-related genes [17–19]. The latter increased the diagnostic yield by 8.9% to 76% in a study by Marshall and co-workers [18]. The current study did not include PCD patients from the town of Volendam. If we take into account that

this town currently has 56 identified PCD patients all harbouring the same homozygous founder mutation, the diagnostic yield is potentially much higher. Our results confirm that the majority of unrelated PCD patients have defects in a core of 5-10 PCD genes, whereas most other mutations are private. Despite this similarity, the distribution of genetic defects is quite different from other countries. *DNAI1*, *CCDC114* and *HYDIN* mutations are much more prevalent in Dutch PCD patients than in others. In case of *DNAI1* and *CCDC114* this is due to founder mutations. To our knowledge, this is the first time that such a high frequency of *HYDIN* mutations is demonstrated. It is unknown whether this is specific to the Dutch population or whether this is a result of the improved extensive analysis of *HYDIN* variants and re-mapping of sequencing data, which prevents automatic discarding of reads which also map to the pseudogene. This proof of concept study remains to be validated in other PCD patient cohorts but definitely offers a novel method to deal with pseudogenes in NGS data analysis.

Variation in ultrastructure and motility defects

There was a slight difference in TEM defects observed in some of the patients with biallelic mutations in the same genes, such as in *DNAH5*. This has been observed in other studies as well and is expected to reflect the difficult evaluation of the IDAs [17]. We also see slight variation in motility defects as a consequence of *DNAAF1* mutations, from complete immotility to severly reduced ciliary beat frequency and amplitude. This has been previously observed in other dynein assembly factors *DNAAF2* and *4*, but it is uncommon in *DNAAF1*.

Unsolved cases

90% of patients in which we did not find biallelic mutations or monoallelic mutations in PCD-related genes had a "likely PCD diagnosis", 5% had "clinical PCD diagnosis" and 5% had a "definite PCD diagnosis". Gene defects in these patients are expected to be identified in one of the 7 currently known PCD-related genes that were not included in the gene panel or in yet to discover PCD-related genes. Another possibility is that some of these individuals may not be true PCD patients. However, this is unlikely as 95% of them repeatedly showed abnormal ciliary beat pattern or ciliary immotility. Finally, many patients had compound heterozygous non-synonymous variants in PCD-related genes, with very low or unknown population frequencies, which may be causative of PCD. The possible causative effect of these variants remains to be studied. It could be reasoned that non-synonymous variants may have relatively small effects in large genes, such as *DNAH5*. On the contrary, however, in other diseases where large genes are affected, such as *FBN1* in Marfan syndrome, the majority of causative mutations is non-synonymous [45].

Genetic testing approach

We cannot yet rely on genetic testing as a sole diagnostic test, as approximately 30% of patient cases remain unsolved [17–19]. Most likely there are only private mutations left to identify in PCD. Therefore, two possible strategies can be proposed to aid genetic testing in diagnostics; 1) a two-tiered approach with a (country specific) targeted gene panel with subsequent exome sequencing in patients in which no mutations are found, or 2) directly applying exome sequencing in all patients. The latter approach is becoming more frequently used as costs lower and coverage homogeneity is improved. When choosing exome sequencing we propose to use *in silico* panels of 1) the 36 currently known PCD-related genes and/or 2) the cluster differentially expressed genes during *in vitro* ciliogenesis of human airway cells, to prioritize variants in possibly novel PCD-related genes. This approach enables pre-selection of variants in genes with a similar expression to currently known PCD-related genes and possible reduction of the total exome data analysis to ~21%.

To summarize, we present a differentially expressed gene cluster with significant upregulation during *in vitro* ciliogenesis of human airway cells. This cluster identified all but one currently known PCD-related genes. From these, only *NME8* exhibited very low expression throughout the entire experiment. Using this gene cluster in the analysis of exome sequencing data may aid in the PCD gene discovery by prioritizing genes that could play an important role in ciliogenesis. Further, we observed a 67.6% diagnostic yield when testing a gene panel of 26 known PCD-related genes and 284 selected candidate genes in a Dutch cohort, not including PCD patients from the Volendam population harbouring a *CCDC114* founder mutation. Dutch PCD patients appear to have a relatively high percentage of *DNAI, CCDC114* and *HYDIN* mutations compared to data from other countries. The candidate genes that were selected from the differentially expressed gene cluster also included two novel PCD-related genes *DNAJB13* and *PIH1D3*, of which identification of the latter was a direct result of this study. This illustrates the potential of the differentially expressed gene cluster approach in the discovery of novel PCD-related genes.

SUPPLEMENTAL DATA

Supplemental data include three tables and can be found with this article online.

ACKNOWLEDGEMENTS

The authors thank the healthy volunteers and the patients for their participation in this study. We also acknowledge F. Pronk, W. Van Leeuwen and G. Koney, for their assistance

in sequencing; M. Massink for his assistance in data analysis, J. Fritz for his assistance with electron microscopy. We thank prof. G.H. Koppelman for his expert scientific support. We are grateful for receiving grants from E. Mulder and his "Dutch tough mudder team" and from Fonds NutsOhra to fund part of this study.
Disclosure statement: the authors declare no conflict of interest.

REFERENCES

1. Fliegauf M, Benzing T, Omran H. When cilia go bad: cilia defects and ciliopathies. *Nat. Rev. Mol. Cell Biol.* 2007; 8: 880–893.
2. Goutaki M, Meier A, Halbeisen F, Lucas J, Dell S, Maurer E, Casaulta C, Jurca M, Spycher B, Kuehni C. Clinical manifestations in Primary Ciliary Dyskinesia: systematic review and meta-analysis. *Eur. Respir. J.* 2016; 48: 1081–1095.
3. Ellerman A, Bisgaard H. Longitudinal study of lung function in a cohort of primary ciliary dyskinesia. *Eur. Respir. J.* 1997; 10: 2376–2379.
4. Hellinckx J, Demedts M, De Boeck K. Primary ciliary dyskinesia: evolution of pulmonary function. *Eur. J. Pediatr.* 1998; 157: 422–426.
5. Marthin JK, Petersen N, Skovgaard LT, Nielsen KG. Lung function in patients with primary ciliary dyskinesia: a cross-sectional and 3-decade longitudinal study. *Am. J. Respir. Crit. Care Med.* 2010; 181: 1262–1268.
6. Shah A, Shoemark A, MacNeill SJ, Bhaludin B, Rogers A, Bilton D, Hansell DM, Wilson R, Loebinger MR, Noone P, Leigh M, Sannuti A, Magnin M, Cros P, Beydon N, Ellerman A, Bisgaard H, Vallet C, Escudier E, Roudot-Thoraval F, Boon M, Smits A, Cuppens H, Davis S, Ferkol T, Rosenfeld M, Maglione M, Bush A, Montella S, Irving S, et al. A longitudinal study characterising a large adult primary ciliary dyskinesia population. *Eur. Respir. J.* 2016; 169: 816–825.
7. Behan L, Dunn Galvin A, Rubbo B, Masefield S, Copeland F, Manion M, Rindlisbacher B, Redfern B, Lucas JS. Diagnosing primary ciliary dyskinesia: an international patient perspective. *Eur. Respir. J.* 2016; 48: 1096–1107.
8. Kuehni CE, Frischer T, Strippoli M-PF, Maurer E, Bush A, Nielsen KG, Escribano A, Lucas JSA, Yiallouros P, Omran H, Eber E, O'Callaghan C, Snijders D, Barbato A. Factors influencing age at diagnosis of primary ciliary dyskinesia in European children. *Eur. Respir. J.* 2010; 36: 1248–1258.
9. Barbato A, Frischer T, Kuehni C, Snijders D, Azevedo I, Baktai G, Bartoloni L, Eber E, Escribano A, Haarman E, Hesselmar B, Hogg C, Jorissen M, Lucas J, Nielsen K, O'Callaghan C, Omran H, Pohunek P, Strippoli M, Bush A. Primary ciliary dyskinesia: a consensus statement on diagnostic and treatment approaches in children. *Eur. Respir. J.* 2009; 34: 1264–1276.
10. 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, et al. European Respiratory Society guidelines for the diagnosis of primary ciliary dyskinesia. *Eur. Respir. J.* 2017; 49: pii. 1601090.
11. Shapiro AJ, Zariwala MA., Ferkol T, Davis SD, Sagel SD, Dell SD, Rosenfeld M, Olivier KN, Milla C, Daniel SJ, Kimple AJ, Manion M, Knowles MR, Leigh MW. Diagnosis, monitoring, and treatment of primary ciliary dyskinesia: PCD foundation consensus recommendations based on state of the art review. *Pediatr. Pulmonol.* 2016; 51: 115–132.
12. Ostrowski LE, Blackburn K, Radde KM, Moyer MB, Schlatzer DM, Moseley A, Boucher RC. A proteomic analysis of human cilia: identification of novel components. *Mol. Cell. Proteomics* 2002; 1: 451–465.
13. Horani A, Ferkol T, Dutcher S, Brody S. Genetics and biology of primary ciliary dyskinesia. *Paediatr. Respir. Rev.* 2016; Mar: 18–24.
14. Arnaiz O, Cohen J, Tassin A-M, Koll F. Remodeling Cildb, a popular database for cilia and links for ciliopathies. *Cilia* 2014; 3: 3–9.

15. Ross AJ, Dailey L a., Brighton LE, Devlin RB. Transcriptional profiling of mucociliary differentiation in human airway epithelial cells. *Am. J. Respir. Cell Mol. Biol.* 2007; 37: 169–185.

16. Dvorak A, Tilley AE, Shaykhiev R, Wang R, Crystal RG. Do airway epithelium air-liquid cultures represent the in vivo airway epithelium transcriptome? *Am. J. Respir. Cell Mol. Biol.* 2011; 44: 465–473.

17. 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. *Pediatr. Pulmonol.* 2016; 51: 498–509.

18. Marshall CR, Scherer SW, Zariwala M, Lau L, Paton T, Stockley T, Jobling RK, Ray PN, Knowles MR, Hall D, Dell SD, Kim RH. Whole Exome Sequencing and Targeted Copy Number Analysis in Primary Ciliary Dyskinesia. *G3* 2015; 5: 1775–1781.

19. Boaretto F, Snijders D, Salvoro C, Spalletta A, Mostacciuolo M, Collura M, Cazzato S, Girosi D, Silvestri M, Rossi G, Barbato A, Vazza G. Diagnosis of Primary Ciliary Dyskinesia by a Targeted Next-Generation Sequencing Panel: Molecular and Clinical Findings in Italian Patients. *J. Mol. Diagnostics* 2016; 18: 912–922.

20. Knowles MR, Ostrowski LE, Leigh MW, Sears PR, Davis SD, Wolf WE, Hazucha MJ, Carson JL, Olivier KN, Sagel SD, Rosenfeld M, Ferkol TW, Dell SD, Milla CE, Randell SH, Yin W, Sannuti A, Metjian HM, Noone PG, Noone PJ, Olson CA, Patrone M V, Dang H, Lee H-S, Hurd TW, Gee HY, Otto E, Halbritter J, Kohl S, Kircher M, et al. Mutations in RSPH1 cause primary ciliary dyskinesia with a unique clinical and ciliary phenotype. *Am. J. Respir. Crit. Care Med.* 2014; 189: 707–717.

21. Amirav I, Wallmeier J, Loges NT, Menchen T, Pennekamp P, Mussaffi H, Abitbul R, Avital A, Bentur L, Dougherty GW, Nael E, Lavie M, Olbrich H, Werner C, Kintner C, Omran H, Alkrinawi S, Aviram M, Rotschild M, Blau H, Kerem E, Cohen-Cymberknoh M, Shoseyov D, Springer C, Hevroni A, Dabbah H, Elizur A, Picard E, Goldberg S, Efrati O, et al. Systematic Analysis of CCNO Variants in a Defined Population: Implications for Clinical Phenotype and Differential Diagnosis. *Hum. Mutat.* 2016; 37: 396–405.

22. Davis SD, Ferkol TW, Rosenfeld M, Lee HS, Dell SD, Sagel SD, Milla C, Zariwala MA., Pittman JE, Shapiro AJ, Carson JL, Krischer JP, Hazucha MJ, Cooper ML, Knowles MR, Leigh MW. Clinical features of childhood primary ciliary dyskinesia by genotype and ultrastructural phenotype. *Am. J. Respir. Crit. Care Med.* 2015; 191: 316–324.

23. Onoufriadis A, Paff T, Antony D, Shoemark A, Micha D, Kuyt B, Schmidts M, Petridi S, Dankert-Roelse JE, Haarman EG, Daniels JMA, Emes RD, Wilson R, Hogg C, Scambler PJ, Chung EMK, 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: 88–98.

24. Goutaki M, Maurer E, Halbeisen FS, Amirav I, Barbato A, Behan L, Boon M, Casaulta C, Clement A, Crowley S, Haarman E, Hogg C, Karadag B, Koerner-Rettberg C, Leigh MW, Loebinger MR, Mazurek H, Morgan L, Nielsen KG, Omran H, Schwerk N, Scigliano S, Werner C, Yiallouros P, Zivkovic Z, Lucas JS, Kuehni CE. The international primary ciliary dyskinesia cohort (iPCD Cohort): methods and first results. *Eur. Respir. J.* 2017; 49: 1601181.

25. Lucas J, Paff T, Goggin P, Haarman E. Diagnostic Methods in Primary Ciliary Dyskinesia. *Paediatr. res* 2016; 18: 8–17.

26. Willems T, Jorissen M. Sequential monolayer-suspension culture of human airway epithelial cells. *J. Cyst. Fibros.* 2004; : Supplement 2:53-4.

27. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 2014; 30: 2114–2120.

28. Andrews S. FastQC: a quality control tool for high throughput sequence data. 2010. p. www.bioinformatics.babraham.ac.uk/projects/fastqc.

29. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* 2013; 29: 15–21.

30. Wang L, Wang S, Li W. RSeQC: quality control of RNA-seq experiments. *Bioinformatics* 2012; 28: 2184–2185.

31. Anders S, Pyl PT, Huber W. HTSeq-A Python framework to work with high-throughput sequencing data. *Bioinformatics* 2015; 31: 166–169.

32. Robinson MD, McCarthy DJ, Smyth GK. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2009; 26: 139–140.

33. Olbrich H, Schmidts M, Werner C, Onoufriadis A, Loges NT, Raidt J, Banki NF, Shoemark A, Burgoyne T, Al Turki S, Hurles ME, Köhler G, Schroeder J, Nürnberg G, Nürnberg P, Chung EMK, Reinhardt R, Marthin JK, Nielsen KG, Mitchison HM, Omran H. Recessive HYDIN Mutations Cause Primary Ciliary Dyskinesia without Randomization of Left-Right Body Asymmetry. *Am. J. Hum. Genet.* 2012; 91: 672–684.

34. Paff T, Loges NT, Aprea I, Wu K, Bakey Z, Haarman EG, Daniels JMA, Sistermans EA, Bogunovic N, Dougherty GW, Höben IM, Große-Onnebrink J, Matter A, Olbrich H, Werner C, Pals G, Schmidts M, Omran H, Micha D. Mutations in PIH1D3 Cause X-Linked Primary Ciliary Dyskinesia with Outer and Inner Dynein Arm Defects. *Am. J. Hum. Genet.* 2017; 100: 160–168.

35. El Khouri E, Thomas L, Jeanson L, Bequignon E, Vallette B, Duquesnoy P, Montantin G, Copin B, Dastot-Le Moal F, Blanchon S, Papon JF, Lorès P, Yuan L, Collot N, Tissier S, Faucon C, Gacon G, Patrat C, Wolf JP, Dulioust E, Crestani B, Escudier E, Coste A, Legendre M, Touré A, Amselem S. Mutations in DNAJB13, Encoding an HSP40 Family Member, Cause Primary Ciliary Dyskinesia and Male Infertility. *Am. J. Hum. Genet.* 2016; 99: 489–500.

36. Hornef N, Olbrich H, Horvath J, Zariwala MA, Fliegauf M, Loges NT, Wildhaber J, Noone PG, Kennedy M, Antonarakis SE, Blouin J-L, Bartoloni L, Nüsslein 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. *Am. J. Respir. Crit. Care Med.* 2006; 174: 120–126.

37. Olbrich H, Häffner K, Kispert A, Völkel A, Volz A, Sasmaz G, Reinhardt R, Hennig S, Lehrach H, Konietzko N, Zariwala M, Noone PG, Knowles M, Mitchison HM, Meeks M, Chung EMK, Hildebrandt F, Sudbrak R, Omran H. Mutations in DNAH5 cause primary ciliary dyskinesia and randomization of left-right asymmetry. *Nat. Genet.* 2002; 30: 143–144.

38. 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 J-L. Mutations in DNAH5 account for only 15% of a non-preselected cohort of patients with primary ciliary dyskinesia. *J. Med. Genet.* 2009; 46: 281–286.

39. Pennarun G, Escudier E, Chapelin C, Bridoux M, Cacheux V, Roger G, Clément A, Goossens M, Amselem S, Duriez B. Loss-of-function mutations in a human gene related to Chlamydomonas reinhardtii dynein IC78 result in primary ciliary dyskinesia. *Am. J. Hum. Genet.* 1999; 65: 1508–1519.

40. Loges NT, Olbrich H, Becker-Heck A, Häffner K, Heer A, Reinhard C, Schmidts M, Kispert A, Zariwala MA, Leigh MW, Knowles MR, Zentgraf H, Seithe H, Nürnberg G, Nürnberg P, Reinhardt R, Omran H. Deletions and point mutations of LRRC50 cause primary ciliary dyskinesia due to dynein arm defects. *Am. J. Hum. Genet.* 2009; 85: 883–889.

41. 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, Häffner K, Hellman N, Chodhari R, Gupta A, Kramer-Zucker A, Olale F, Burdine RD, Schier AF, O'Callaghan C, Chung EMK, Reinhardt R, Mitchison HM, King SM, Omran H, Drummond I a. CCDC103 mutations cause primary ciliary dyskinesia by disrupting assembly of ciliary dynein arms. *Nat. Genet.* 2012; 44: 714–719.

42. 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 K V, 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.* 2010; 43: 79–84.

43. Blanchon S, Legendre M, Copin B, Duquesnoy P, Montantin G, Kott E, Dastot F, Jeanson L, Cachanado M, Rousseau, Papon JF, Beydon N, Brouard J, Crestani B, Deschildre a., Desir J, Dollfus H, Leheup B, Tamalet A, Thumerelle C, Vojtek a.-M, 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: 410–416.

44. Duriez B, Duquesnoy P, Escudier E, Bridoux A-M, Escalier D, Rayet I, Marcos E, Vojtek A-M, Bercher J-F, Amselem S. 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: 3336–3341.

45. Franken R, Den Hartog AW, Radonic T, Micha D, Maugeri A, Van Dijk FS, Meijers-Heijboer HE, Timmermans J, Scholte AJ, Van Den Berg MP, Groenink M, Mulder BJM, Zwinderman AH, De Waard V, Pals G. Beneficial Outcome of Losartan Therapy Depends on Type of FBN1 Mutation in Marfan Syndrome. *Circ. Cardiovasc. Genet.* 2015; 8: 383–388.