PLOS ONE

RESEARCH ARTICLE

Genomic profiling supports the diagnosis of primary ciliary dyskinesia and reveals novel candidate genes and genetic variants

Marina Andjelkovic[1], Predrag Minic[2,3], Misa Vreca[1], Maja Stojiljkovic[1], Anita Skakic[1], Aleksandar Sovtic[2], Milan Rodic[2], Vesna Skodric-Trifunovic[3,4], Nina Maric[5], Jelena Visekruna[2], Vesna Spasovski[1], Sonja Pavlovic[1]*

1 Laboratory for Molecular Biomedicine, Institute of Molecular Genetics and Genetic Engineering, University of Belgrade, Belgrade, Serbia, **2** Mother and Child Health Care Institute of Serbia „Dr Vukan Cupic", Belgrade, Serbia, **3** School of Medicine, University of Belgrade, Belgrade, Serbia, **4** Clinic for Pulmonology, Clinical Center of Serbia, Belgrad, Serbia, **5** Clinic for children diseases, University Clinical Center of the Republic of Srpska, Banja Luka, Bosnia and Herzegovina

* sonya@sezampro.rs

Check for updates

OPEN ACCESS

Citation: Andjelkovic M, Minic P, Vreca M, Stojiljkovic M, Skakic A, Sovtic A, et al. (2018) Genomic profiling supports the diagnosis of primary ciliary dyskinesia and reveals novel candidate genes and genetic variants. PLoS ONE 13(10): e0205422. https://doi.org/10.1371/journal.pone.0205422

Editor: Mohammad R. Akbari, University of Toronto, CANADA

Received: April 13, 2018

Accepted: September 25, 2018

Published: October 9, 2018

Copyright: © 2018 Andjelkovic et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files. GenBank accession numbers for the sequences used in the analyses were as follows: NM_017950.3 (CCDC40), NM_012144.2 (DNAI1), NM_001369.2 (DNAH5), NM_031427.3 (DNAL1), NM_001277115.1(DNAH11), NM_012472.3 (LRRC6), NM_024532.4 (SPAG16), and NM_206996.2 (SPAG17).

Abstract

Primary ciliary dyskinesia (PCD) is a rare inherited autosomal recessive or X-linked disorder that mainly affects lungs. Dysfunction of respiratory cilia causes symptoms such as chronic rhinosinusitis, coughing, rhinitis, conductive hearing loss and recurrent lung infections with bronchiectasis. It is now well known that pathogenic genetic changes lead to ciliary dysfunction. Here we report usage of clinical-exome based NGS approach in order to reveal underlying genetic causes in cohort of 21 patient with diagnosis of PCD. By detecting 18 (12 novel) potentially pathogenic genetic variants, we established the genetic cause of 11 (9 unrelated) patients. Genetic variants were detected in six PCD disease-causing genes, as well as in *SPAG16* and *SPAG17* genes, that were not detected in PCD patients so far, but were related to some symptoms of PCD. The most frequently mutated gene in our cohort was *DNAH5* (27.77%). Identified variants were in homozygous, compound heterozygous and trans-heterozygous state. For detailed characterization of one novel homozygous genetic variant in *DNAI1* gene (c. 947_948insG, p. Thr318TyrfsTer11), RT-qPCR and Western Blot analysis were performed. Molecular diagnostic approach applied in this study enables analysis of 29 PCD disease-causing and related genes. It resulted in mutation detection rate of 50% and enabled discovery of twelve novel mutations and pointed two possible novel PCD candidate genes.

Introduction

Primary ciliary dyskinesia (PCD (OMIM #244400)) is a rare inherited autosomal recessive or X-linked disorder, which affects lungs, reproductive organs, and organ laterality. Major characteristics of PCD are ultrastructural defects of cilia leading to ciliary immotility or abnormal ciliary motility [1, 2]. PCD presents with neonatal respiratory distress in 80% of cases,

Funding: This work has been funded by grant from the Ministry of Education, Science and Technological Development, Republic of Serbia (III 41004) (http://www.mpn.gov.rs/), SP. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

recurring acute rhinosinusitis [3], rhinitis [4], sinusitis, conductive hearing loss due to otitis media, recurrent or chronic lung infections with bronchiectasis and progressively declining lung function[5]. Since motile cilia are present throughout the respiratory tract, every loss of structural/functional integrity of cilia leads to disorder in the primary innate defense mechanism of mucociliary clearance [6]. Around 50% of PCD patients have *situs inversus* (SI) [3], and reduced fertility in males and females is observed [7, 8]. The estimated prevalence of PCD is 1 in 15.000–30.000 live births, but recent studies propose that prevalence is higher in consanguineous populations[9]. Early recognition of PCD and the correct diagnosis are often delayed due to the clinical symptoms overlapping with other chronic airway disorders.

The clinical features of PCD are nowadays well recognized, but the diagnosis is still challenging, especially when patients represent non-specific signs and symptoms, and when necessary equipment and screening tests including nasal nitric oxide (nNO) measurements, nasal mucociliary transport tests, and saccharine test, are not available to physicians. Even transmission electron microscopy, which is considered as "gold standard" for PCD, cannot resolve 30% of PCD patients with normal cilia ultrastructure [10–15]. Therefore, implementation of genetic and molecular diagnosis of PCD is rather a necessity, regardless of whether it is a confirmation of a clinical diagnosis or suspicion of PCD. The identification of PCD genes has been based on linkage studies, candidate gene approaches and proteomic analyses, combined with sequencing of potentially causative genes [16–19]. Recently, the availability of high-throughput sequencing techniques contributed to the swift identification of new PCD-causative genes, resulting in more than 40 genes to be identified so far, which allowed 65% of cases to be genetically described [15, 20–23].

Herein, we analyzed the cohort of clinically suspected PCD patients from Serbia using a Clinical-Exome Next-Generation Sequencing (NGS). Using this approach we confirmed some genetic variants already classified as disease-causing mutations, and discovered 12 novel genetic variants that had never been reported to be associated with PCD. In this study we also describe functional characterization of the novel, potentially pathogenic variant in the *DNAI1 (CILD1)* gene found in two siblings.

Materials and methods

Subjects

This study has been approved by the Ethics Committee of the Mother and Child Health Care Institute of Serbia Dr Vukan Cupic, Belgrade, Serbia. The study has been performed in accordance with the ethical standards lain down in the Helsinki Declaration and its later amendments. Informed consent was obtained from all participants included in the study.

Inclusion criteria for PCD diagnosis included symptoms such as neonatal respiratory distress, chronic sinusitis, bronchiectasis, recurrent pneumonia and SI. For identification of cilia motility, respiratory epithelial cells were obtained by nasal brushing, and cilia movements were detected using optical microscope. Exclusion criteria implied the absence of cystic fibrosis (CF), for which purpose patients were tested for presence of *Pseudomonas aeruginosa* and sweat chloride test.

Extensive genetic analysis on 21 subjects from 18 unrelated families was conducted. Control group was consisted of in-house TruSight One collection of 69 subjects from general population of Serbia. Co-segregation analysis was performed for three families. For that purpose parents from these families as well as a child suspected to have PCD from third family, were recruited and tested by Sanger sequencing.

Genomic DNA was isolated from peripheral blood by a salting-out method. The samples were then referred to the Institute of Molecular Genetics and Genetic Engineering (IMGGE)

for genetic profiling. The quality and quantity of isolated DNA was measured using Qubit 3.0 fluorimeter (Invitrogen, USA).

Genetic and Bioinformatic analysis

Twenty one patients were analyzed by NGS approach using the Illumina Clinical-Exome Sequencing TruSight One Gene Panel. This panel includes all the known disease-associated genes described in the OMIM database until 2013, designed to cover all exons and flanking intronic regions of 4813 genes (~ 62.000 exons). 5ng/ul (10ul) of genomic DNA was used in all reactions, and quantification of DNA in all steps of NGS sequencing was determined using Qubit 3.0 fluorimeter (Invitrogen, USA). Qubit dsDNA HS Assay Kit with range of 0.2-100ng was used (Invitrogen, USA). Illumina MiSeq platform was used to generate libraries and to produce the VCF files. The software for analyzing data gained using NGS approach was IlluminaVariant Studio 2.0 and 3.0 (Illumina, San Diego, USA). VCF file contained approximately 8–10.000 variants per sample. After annotation of all samples we have accessed to variant filtering. The Filters section within IlluminaVariant Studio contains multiple parameters from which we have used the following: 1. General section: genotype-all, variant type-all, chromosomes-all; 2. Variant section: Pass Filter, 3. Consequence section: ClinVar pathogenic, Show only variants that are-all (missense, frameshift, stop gained, stop lost, initiator codon, in-frame insertion, in-frame deletion, splice); 4. Population frequency section: Thousand Genomes: set all frequencies on less than 5%, Exac: set all frequencies on less than 5%; 5. Classification section: Filter by classification-presumed pathogenic, pathogenic, unknown significance. After applying these filters, we got around 100 variants per sample. Than we prioritized 29 genes of interest for PCD (disease-causing genes according to OMIM, candidate genes coding for proteins which directly interacts with PCD disease-associated proteins according to String Interaction Network (https://string-db.org/), and genes belonging to the same gene family as disease-causing genes) (Table 1), and got three to five variants per sample. Variants that had allele frequencies less than 5% (according to 1000 Genomes and ExaC), predicted as damaging (according to SIFT and PolyPhen2 which are implemented into VariantStudio), and absent in our in-house TruSight One collection of 69 samples sequenced with the same methodology, were further analyzed and considered as candidate variants (one or two per sample). In-house TruSight One base served for exclusion of population specific variants. If some variant had higher frequency than 5% in European population and predicted as pathogenic, but wasn't detected in our patients' cohort, we took it into consideration.

Mutation validation and co-segregation analysis

The following databases were used to additionally determine candidate variants: VarSome (which includes the following databases: gnomAD genomes, gnomAD exomes, ClinVar, 1000 Genomes Project, and classification according to American College of Medical Genetics and Genomics (ACMG)), dbSNP, Exome Aggregation Consortium (ExaC), Ensembl, and The Human Gene Mutation Database (HGMD). Prediction tools used for pathogenicity scoring were as follows: Deleterious Annotation of Genetic Variants Using Neural Networks (DANN), MutationTaster, and Functional Analysis through Hidden Markov Models (FATHMM-MKL). Effects on protein level were investigated with SIFT, Provean, and PolyPhen2 in silico tools.

We used I-TASSER server (http://zhanglab.ccmb.med.umich.edu/I-TASSER/) to generate a PDB files and for identification of amino acids required for protein-protein interactions (PPIs). For analysis of protein structure, we used UCFS Chimera tool (http://www.rbvi.ucsf.edu/chimera).

Table 1. A list of analyzed PCD-related genes.

PCD causative genes	Protein localization/function	Variants detected in our study	The incidence of variants in our study
CCDC39	N-DRC	-	0%
CCDC40	N-DRC	+	11.11%
CCDC103	Cytoplasmatic, ODA assembling	-	0%
DNAAF1 (LRRC50)	Cytoplasmatic, DA assembling	-	0%
DNAAF2 (KTU)	Cytoplasmatic, DA assembling	-	0%
DNAAF3	Cytoplasmatic, DA assembling	-	0%
DNAI1	ODA	+	16.66%
DNAI2	ODA	-	0%
DNAL1	ODA	+	16.66%
DNAH1	ODA	-	0%
DNAH5	ODA	+	27.77%
DNAH9	ODA	-	0%
DNAH11	ODA	+	5.55%
HEATR2	Cytoplasmatic, DA assembling	-	0%
HYDIN	CP	-	0%
LRRC6	Cytoplasmatic, DA assembling	+	5.55%
NME8 (TXNDC3)	ODA	+	0%
OFD1	Cytoplasmatic	-	0%
RPGR	Cytoplasmatic	-	0%
RSPH4A	RS	-	0%
RSPH9	RS	-	0%
PCD candidate genes			
CCDC8	Regulate microtubule dynamics	-	0%
CCDC50	Associates with microtubule-based structures	-	0%
CCDC88	Role in ciliogenesis and cilium morphology	-	0%
DYNC1H1	Microtubule motor activity	-	0%
NME1	Centrosome	-	0%
SPAG16	Role in motile ciliogenesis	+	11.11%
SPAG17	Proper function of the axoneme	+	5.55%
TCTE1	Component of axonemal dynein and cytoplasmic dynein 1	-	0%

CP, central pair; DA, dynein arms; ENaC, amiloride-sensitive epithelial sodium channel; N-DRC, nexin-dynein regulatory complex; ODA, outer dynein arm; RS, radial spoke.

https://doi.org/10.1371/journal.pone.0205422.t001

The results obtained by NGS method were further verified by direct sequencing using the Big Dye terminator cycle sequencing kit and the ABI PRISM 310 automated sequencer (Applied Biosystems, USA). Primer sequences are listed in S1 Table. Co-segregation analysis was performed for three families which were available for the analysis.

RNA extraction and synthesis of cDNA

For extraction of total RNA peripheral blood mononuclear cells from two probands, their parents, and 11 healthy controls were purified on Ficoll-PaqueTM Plus (GE Healthcare, USA) density gradient. TRI Reagent Solution (Ambion, USA) was used for isolation of total RNA, according to the manufacturer's protocol. Complementary DNA (cDNA) was synthesized from 1μg of total RNA using RevertAid M-MuLV Reverse Transcriptase (ThermoFisher Scientific, USA).

Relative quantification (RT-qPCR)

Expression of mRNA was determined using KAPA PROBE and SYBR Green Universal qPCR kit (KAPA Biosystems, USA). Real-time PCR was performed using ABI 7500 Real-Time PCR System (Applied Biosystems, USA). The Glyceraldehyde 3-phosphate dehydrogenase gene (*GAPDH*) was used as an endogenous control and median expression level of the healthy control group was used as calibrator. Relative quantification analysis was performed by a comparative ddCT method. All experiments were performed in duplicates. Primers for DNAI1 transcript designed according to instructions listed in [24], and their sequences are listed in S2 Table.

Western blot analysis

Peripheral blood from two patients with the novel genetic variant, their parents, and healthy control was used for isolation of total cellular proteins for Western blot analysis. Peripheral blood cells were lysed with Lysis Buffer, treated with TEN (Tris-EDTA-NaCl) containing protease inhibitor (Roche, Switzerland) in 0.25M TrisHCl using three freeze- and- thaw cycles. For antibody validation, proteins from Hek293 cell line were used as negative control. DNAI1 protein was detected by Western blot using rabbit Anti-Dynein intermediate chain 1 monoclonal antibody (Abcam, United Kingdom), anti-rabbit IgG horseradish peroxidase conjugate (GE Healthcare, USA) and the enhanced chemiluminescence detection system (GE Healthcare, USA).

Results

Clinical phenotype and mutation profiling

Study included 21 subjects, 10 males (47.60%) and 11 females (52.40%) with an age range of 1 to 35 years, and two family members (parents). *Situs solitus* was observed in 11/21 (52.4%) patients, *situs inversus* was present in 10/21 (47.60%) patients, whereas *situs ambiguous* was not detected. Neonatal respiratory distress was present in 17/21 (80.95%), 19/21 had chronic sinusitis (90.47%), 17/21 had bronchiectasis (80.95%), and 9/21 had recurrent pneumonia (42.86%). Microscopically observed immotile cilia was present in 17/21 (80.95%), 4/21 patients had motile cilia (19.05%). In the group of patients with motile cilia 1/4 (25.00%) had normal ciliary beat pattern, whereas 3/4 had pathological ciliary beat pattern (Table 2). All patients were negative for presence of *Pseudomonas aeruginosa*, and concentration of chloride in sweat was not elevated. Additionally, detected symptoms were chronic secretory otitis media and haemoptysis. Parents of all analyzed patients had confirmed that consanguinity is not present in their families.

In order to determine the genetic basis of 21 patients suspected to have PCD, we used TruSight One Gene Panel and generated comprehensive libraries. We prioritized 29 genes listed in Table 1, and used the recessive disease model which included homozygotes, compound heterozygotes and trans-heterozygotes. Analysis of the 29 genes revealed a total of 2210 variants. Then we searched for this variants within in-house TruSight One base and discarded the variants that also exist in general population of Serbia for further analysis. Overall, we have detected 18 different pathogenic genetic variants in PCD disease-causing and candidate genes in 11 patients (9 unrelated patients), which may be considered as potential genetic causes of the disease, since various prediction tools suggested their pathogenicity. Among 9 PCD patients, homozygous variants were detected in 3 patients, 4 carried compound heterozygous variants, and 2 had trans-heterozygous variants. In 3 of 4 patients with compound heterozygous variants, the inheritance was not confirmed due to the unavailability of the parental

Table 2. Clinical and genetic characteristics of analysed patients in population of Serbia.

	Patients
	21 (100.00%)
Males	11 (52.40%)
Females	10 (47.60%)
Diagnostic procedures	
Symptoms	
Bronchiectasis	17 (80.95%)
Chronic sinusitis	19 (90.47%)
Neonatal respiratory distress	17 (80.95%)
Recurrent pneumonia	9 (42.86%)
Organ laterality	
Situs inversus	10 (47.60%)
Situs solitus	11 (52.40%)
Situs ambiguous	0
Cilia motility	
Motile:	4 (19.05%)
Normally ciliary beat pattern	1/4 (25.00%)
Pathological ciliary beat pattern	3/4 (75.00%)
Non-motile	17 (80.95%)
Genetics	
Mutations detected in PCD causative and candidate genes	9 (50%)
Homozygous mutations in *CFTR* gene	0

https://doi.org/10.1371/journal.pone.0205422.t002

samples. Disease-causing variants were detected in *CCDC40*, *DNAI1*, *DNAL1*, *DNAH5*, *DNAH11*, *LRRC6*, *SPAG16* and *SPAG17* genes (Table 3). The most frequently mutated alleles were within *DNAH5* gene (27.77%) and none of 21 analysed patient had homozygous mutation in *CFTR* gene. We also included three monoallelic variants in Table 3, (variant c.8999G>A in *DNAH5* gene; variant c.27T>G, in *LRCC6* gene; and variant c.8555C>G, in *DNAH11* gene,) which were detected in our patients, and shown to be potentially pathogenic, but these patients were not included in the above-mentioned 9 PCD patients.

RT-qPCR and Western blot analysis

Among patients with detected variants in PCD causative genes, we found a novel homozygous variant in *DNAI1* gene (c. 947_948insG, p. Thr318TyrfsTer11), (GenBank: NM_012144.2), in two affected siblings (designated as P9 and P10). Their parents were heterozygous carriers of detected variant (Fig 1A and 1B).We analysed DNAI1 mRNA from peripheral blood mononuclear cells using RT-qPCR to investigate the potential effect of the nucleotide insertion on transcript stability. We found that the DNAI1 transcript was expressed 30–45% lower in two affected patients and their parents compared to controls (Fig 1C). DNAI1 protein was analysed using Western Blot method to examine the impact of frameshift stop mutation on the protein level, and full-length protein in patients was detected (Fig 1D, lanes 3 and 4). The amount of detected protein in patients was lower in comparison to parents (Fig 1D, lanes 5 and 6) and positive control (Fig 1D, lane 1).

Computational analyses

In silico modelling implied that amino acid change and consequential downstream introduction of multiple stop codons led to protein truncation (Fig 2A). We identified positions and

Table 3. Spectrum of detected genetic variants in Serbian patients with PCD.

GENE	dbSNP ID	Genomic coordinates	Genetic variants		Allele freq Eur (%)[a]	ACMG classification[b]	DANN[c]	SIFT/ Provean[d]	Patient ID	Reference from HGMD/ ClinVar
			Nucleotide change	Amino acid change						
CCDC40	rs397515393	chr17-78013765	c. 248delC	p.Ala83ValfsTer84	0	Pathogenic (PVS1, PS3)	NA	Not applicable	P5	[25]
	rs747233125	chr17-78060006	c. 2440C>T	p.Arg814Ter	0	Pathogenic (PVS1, PM2, PM4, PP2)	0.9899	Damaging	P13, P14	[26]
DNAI1	-	chr9-34500765	c. 947_948insG	p. Thr318TyrfsTer11	unknown	NA[e]	NA	Not applicable	P9, P10	novel
	rs867262419	chr9-34512140	c.1345_1349delCTTAA	p. Asn450LeufsTer6	0	NA[e]	NA	Not applicable	P21	novel
	-	chr9-34514506	c.1684G>A	p.Asp562Asn	0	Uncertain Significance (PM2, PP3)	0.9993	Damaging	P21	novel
DNAH5	-	chr5-13864748	c. 4356-2A>G	-	0	Pathogenic (PVS1, PM2, PP3)	0.9949	Not applicable	P4	novel
	-	chr5-13809281	c.7624T>C	p.Trp2542Arg	0	Uncertain Significance (PM2, PP3)	0.9919	Damaging	P4	novel
	rs137949961	chr5-13777417	c.8999G>A	p.Arg3000Gln	0.26	Uncertain Significance (PP3)	0.9989	Damaging	P6	ClinVar 219734
	rs140782270	chr5-13914743	c.1206T>A	p.Asn402Lys	0.13	Uncertain Significance (PM2, PP3)	0.9979	Damaging	P20	ClinVar 188080
	-	chr 5–13718980	c.8012A>G	p.Gln2701Arg	0	NA[e]	NA	Damaging	P20	novel
DNAL1	rs751754576	chr14-74154044	c.347A>T	p.Lys116IIe	0	Uncertain Significance (PM2, PP3)	0.9926	Damaging	P1	novel
	rs535885451	chr14-74154047	c.350T>G	p.Leu117Trp	0	Uncertain Significance (PM1, PM2, PP3)	NA	Damaging	P1	novel
	-	chr14-74156171	c.485G>A	p.Trp162Ter	0	Uncertain Significance (PM2, PM4, PP3)	0.9946	Damaging	P1	novel
LRRC6	rs200906172	chr8-133673857	c.27T>G	p.Ile9Met	0	Uncertain Significance (PM2, BS4)	0.9907	Damaging	P2	ClinVar 473106
DNAH11	-	chr7-21788220	c.8533C>G	p. Arg2845Gly	0.13	Likely benign (PM1, PM2, PP5, BP1, BP4, BP6)	0.9813	Damaging	P3	ClinVar 359658
SPAG16	rs61752199	chr2-214354811	c.1067G>A	p.Ser356Asn	2	Uncertain Significance (PM1, BS1)	0.9964	Damaging	P8	novel
	-	chr2-214174834	c. 331G>A	p.Asp111Asn	0	Uncertain Significance (PP3)	0.9993	Damaging	P11	novel

(*Continued*)

Table 3. (Continued)

GENE	dbSNP ID	Genomic coordinates	Genetic variants		Allele freq Eur (%)[a]	ACMG classification[b]	DANN[c]	SIFT/ Provean[d]	Patient ID	Reference from HGMD/ ClinVar
			Nucleotide change	Amino acid change						
SPAG17	rs17185492	chr1-118644524	c. 473A>T	p.Glu158Val	17	Benign (PP2, BS1, BA1,)	0.9927	Damaging	P8, P11	novel

All identified genetic variants were numbered based on cDNA reference sequences and as recommended by the Human Genome Variation Society (http://www.hgvs.org/mutnomen). RefSeq accession numbers for the sequences used in the analyses were as follows: NM_017950.3 (*CCDC40*), NM_012144.2 (*DNAI1*), NM_001369.2 (*DNAH5*), NM_031427.3 (*DNAL1*), NM_001277115.1 (*DNAH11*), NM_012472.3 (*LRRC6*), NM_024532.4 (*SPAG16*), and NM_206996.2 (*SPAG17*).

[a] According to VarSome (1000 Genome Project, ExaC)

[b] ACMG: American College of Medical Genetics and Genomics.

[c] Deleterious annotation of genetic variants using neural networks (DANN). The value range is from 0 to 1, with 1 given to the variants predicted to be the most damaging.

[d] SIFT (sorts intolerant from tolerant) is an in silico prediction tool based on sequence homology derived from closely related sequences collected through PSI-BLAST. Provean (Protein Variation Effect Analyzer) is an in silico tool that predicts how nonsynonymous or in-frame indel variant will affect a protein's biological function.

[e] This variant does not have automated ACMG/AMP 2015 interpretation

https://doi.org/10.1371/journal.pone.0205422.t003

amino acids in polypeptide chain participating in PPIs and established that the region in DNAI1 protein for PPIs is located downstream of premature stop codon (S3 Table). If for some reason, this mutated protein still exists in cells, its function is disrupted due to the absence of accurate amino acids in targeted position for PPIs (Fig 2A). Amino acid sequence alignment indicated high evolutionary conservation of residues affected by this variant among the DNAI1 orthologs in all analyzed vertebrate species (Fig 2B).

Discussion

In the present study, we analysed the largest PCD cohort from Southeastern Europe ever reported. This is the first study which includes the genetic and molecular profiling, as well as phenotypic characterisation of PCD patients of Serbian ethnicity. Although large cohorts of PCD patients have been reported in Europe, only a small number of studies included molecular analysis of PCD patients [17, 20, 21, 27]. More often they included individual families or analysis of specific genes [22, 25, 28, 29]. Usage of custom-made population-specific platforms is convenient for detailed analysis of known PCD causative genes, but omits to detect novel candidate genes. Such limitations are circumvent by usage of whole-exome or whole-genome approaches. The approach presented in this study enables analysis of all exons of genes reported to be related to some human disorder, and thus enables discovery of potential candidate genes for PCD. The main limitation of this panel is the lack of all known PCD related genes, which can result in unresolved cases and lack of final genetic diagnosis.

We established the genetic cause of 9 (50%) unrelated patients involved in this study using NGS approach, which is consistent with the results published earlier [18, 21]. The most frequently mutated gene in our population, so far, is *DNAH5* (27.77%), which is consistent with results of previous studies on European population [21, 30]. Homozygous genetic variants were not detected in *CFTR* gene, as well as, presence of *Pseudomonas aeruginosa* and elevated chloride in sweat, which enabled us to exclude CF as potential diagnosis of analysed patients (Table 1).

DNAI1, *DNAH5*, *DNAH11*, and *DNAL1* genes are part of outer dynein arms (ODA)[31], and mutations in these genes were detected in 55.55% of our patients. Mutations in *LRCC6*, *SPAG16* and *SPAG17* genes found in 3 patients are responsible for combined defects in outer

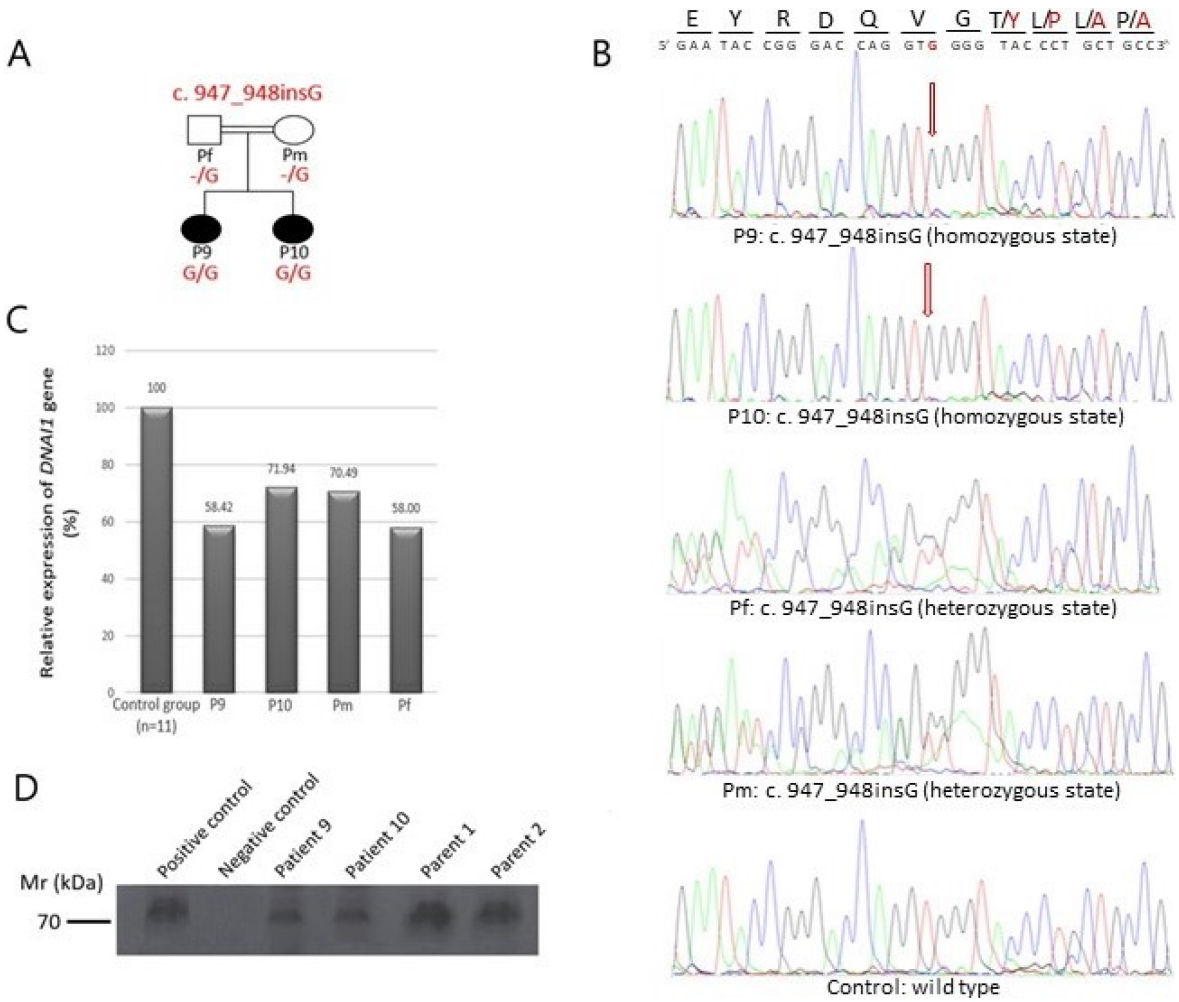

Fig 1. Confirmation by direct sequencing, expressional analysis and protein analysis of novel variant, c.947_948insG, found in *DNAI1* gene. (A) Pedigree structure of one PCD family with two affected sisters. (B) Sanger sequence confirmation of the 1 bp insertion (NM_012144.2: c.947_948insG) in exon 9 of *DNAI1* gene. First two sequences designated as P9 and P10 originated from the affected patients and displayed homozygous insertion of G nucleotide. Second two sequences (Pf and Pm) belongs to their parents and displayed insertion of G nucleotide on the one allele, and the last sequence originated from healthy control. For direct sequencing reverse primer was used, and therefore the sequences are aligned from the reverse side. (C) Comparison of the RT-qPCR profile of control group (mean of healthy control samples represents 100% expression), affected siblings (P9 and P10) and their parents (Pm and Pf). Relative expression of the DNAI1 transcript was 30–45% lower in two affected patients and their parents compared to healthy controls. (D) Immunoblotting of proteins from two patients with novel genetic variant, their parents and healthy control with rabbit Anti-Dynein intermediate chain 1 antibody, which reacts with human DNAI1 protein. Using Western Blot method we detected full length protein in patients (lanes 3 and 4), but the amount was lower in comparison to parents (lanes 5 and 6) and positive control (lane 1). Lane 2 represents negative control. The blot shown is representative of three independent experiments.

https://doi.org/10.1371/journal.pone.0205422.g001

and inner dynein arms (ODA/IDA). Mutations in *CCDC40* gene lead to defects in IDA and nexin-dynein regulatory complexes (N-DRCs)[32]. To the best of our knowledge, this is the first report of genetic variants in *SPAG16* and *SPAG17* genes detected in PCD patients, although these genes are correlated with some symptoms in the literature [33]. *SPAG 17* encodes for central pair protein present in the "9 + 2" axonemes. The encoded protein is required for the proper function of the axoneme. Mutations in the orthologous gene in mice lead to PCD characterized by immotile nasal and tracheal cilia, reduced clearance of nasal mucus, respiratory distress, and neonatal lethality due to impaired airway mucociliary clearance [33]. The *SPAG16* gene encodes two major transcripts, SPAG16L and SPAG16S. This two

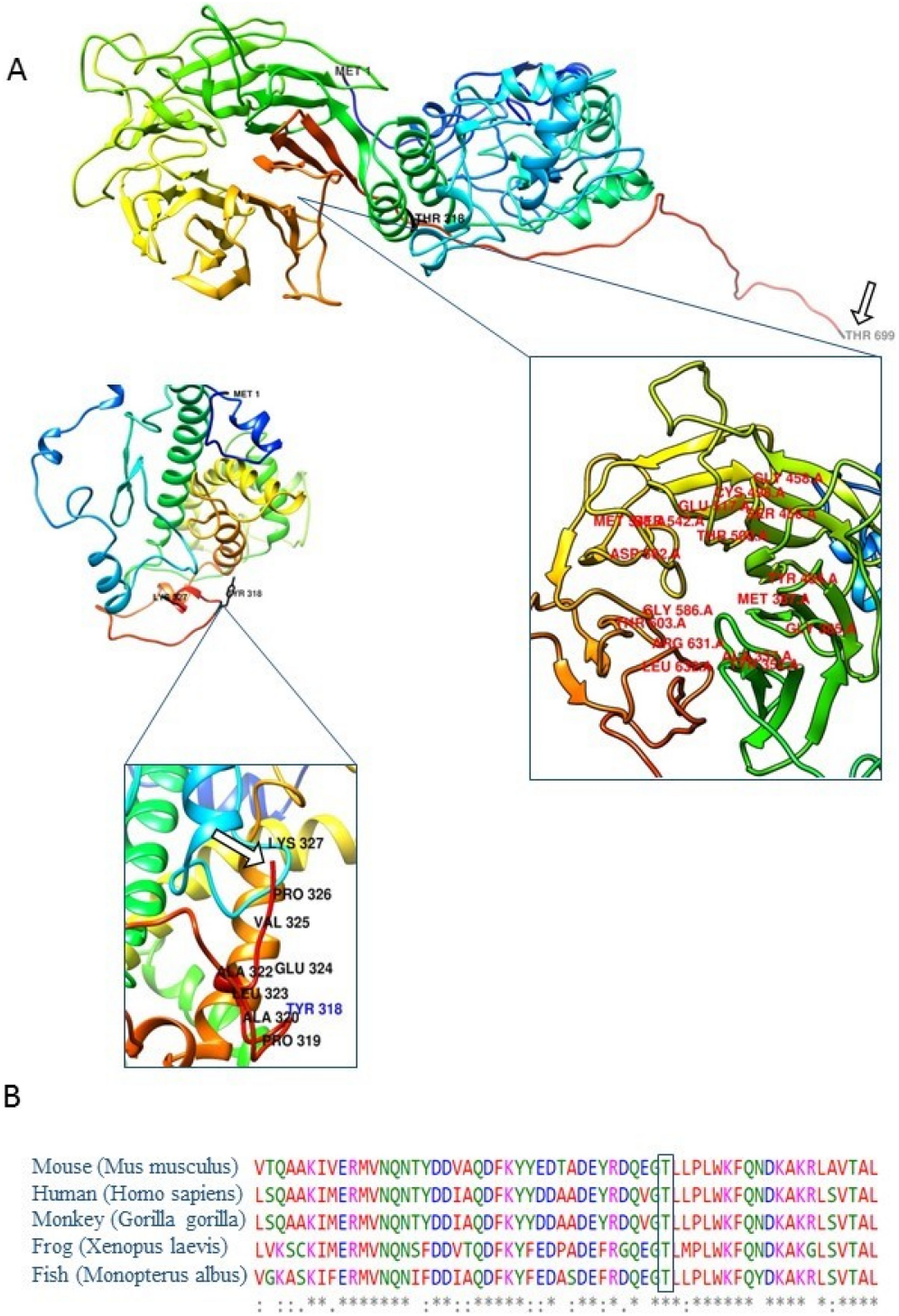

Fig 2. Three-dimensional molecular models of DNAI1 protein with close-up views of the regions harbouring novel frameshift genetic variant, and regions for PPIs. (A) View of the wild type (top picture) and mutated (bottom picture) DNAI1 protein. Wild type DNAI1 protein consists of 699 amino acids (GenBank accession number: NP_036276.1). Mutated is shorter than wild type and consists of 327 amino acids. Proteins begin to differ after position 318 in polypeptide chains. White arrows point to the last amino acid in polypeptide chains (Thr 699 in wild type and Lys 327 in mutated protein). Labelled amino acids in close-up view of wild type protein are representing ligand binding sites. Close-up view of mutated protein is displaying the altered region and truncation of protein due to frameshift stop mutation. Insertion has led to changes in open reading frame and consequently incorporation of

incorrect amino acids downstream of insertion and premature stop codon. Images were prepared using UCFS Chimera software (http://www.rbvi.ucsf.edu/chimera). (B) Amino acid sequence alignment with ClustalW2. Five sequences were aligned and results indicate to high evolutionary conservation of residues affected by p. Thr318TyrfsTer11 variant among the DNAI1 orthologs in all analyzed vertebrate species (mouse, monkey, frog, fish and human). The amino acid marked with a rectangle displayed absolute evolutionary conservation among analysed species.

https://doi.org/10.1371/journal.pone.0205422.g002

proteins associate with the axoneme of sperm tail and the nucleus of postmeiotic germ cells, respectively [34]. Previous studies suggested that SPAG17, SPAG6, and SPAG16L form an interactome in the mammalian central apparatus. The significance of these interactions is inferred from the fact that deficiency in two of these proteins result in defects in spermatogenesis, sperm motility, and ciliary dysfunction [35].

Characterization of homozygous genetic variant

Among patients with detected genetic variants in PCD causative genes, we found new potentially pathogenic variant (c.947_948insG located at exon 9 in *DNAI1* gene), presented in Axonemal dynein intermediate-chain gene (*DNAI1* (OMIM #604366)) was the first reported gene in which mutations were found to be associated with PCD, and so far 73 genetic variants were detected, according to ClinVar. Segregation analysis confirmed that homozygous variant in *DNAI1* gene that has been found in two siblings, was consistent with recessive inheritance of the mutation (Fig 1). This c.947_948insG genetic variant leads to a frameshift and an introduction of a UGA stop codon after 33 nucleotides, which would lead to synthesis of truncated protein of 327 amino acids (Fig 2). On the transcription level, we showed that DNAI1 mRNA transcript containing a premature stop codon is less expressed than wild type transcript (Fig 1C). This c.947_948insG transcript is most likely marked for degradation by nonsense-mediated mRNA decay, but the efficiency of this process is different in different cell types, and therefore some amount of this transcript remains in cells (Fig 1C, lanes: P9, P10, Pm, and Pf). Analysis of protein containing p.Thr318TyrfsTer11 amino acid change using Western Blot method surprisingly showed full-length protein in patients (Fig 1D, lanes P9, and P10) regardless of the presence of premature stop codon. In order to explain such result, we translated wild type and mutated mRNA sequence using Translate Tool (http://web.expasy.org/translate/) and discovered that this frameshift mutation leads to insertion of sixteen UGA stop codons downstream of premature UGA stop codon (S1 and S2 Figs). While the three stop codons typically lead to termination of translation, the process of protein synthesis termination, although effective, is not 100% efficient. Several natural mechanisms of termination suppression exist, including ribosomal frameshifting [36] suppressor tRNAs [37] and high-translational UGA readthrough, which we assume to be responsible for the presence of full length protein in our patients. Additionally, a selenocysteine tRNA can insert a selenocysteine (Sec) instead UGA stop codons, if a selenocysteine insertion sequence (SECIS element) is present in the untranslated regions [38, 39]. Although the occurrence of selenoprotein genes is limited, the Sec UGA codon has become the first addition to the universal genetic code.

We analysed the possibility of presence of the SECIS element in 3'UTR region of DNAI1 mRNA. Analysis of 60 nucleotides downstream of the originally stop codon for potential RNA secondary structures using mfold online program (http://unafold.rna.albany.edu/) [40] showed that dG of DNAI1 downstream sequence is in the range of the dG of downstream sequences of human selenoprotein P (SelP) gene (dG DNAI1: -14kcal/mol; dG SelP: -2.2kcal/mol to-16.6kcal/mol) [41]. This is an important finding since SelP contains multiple UGA codons [42], as well as DNAI1. Relying on previous studies on selenocysteine, our *in silico* findings, as well as high evolutionary conservation of this part of DNAI1 protein among

ortholog species, we postulate a hypothesis that, for some reason, c.947_948insG mutation prefers the synthesis beyond the stop codon and insertion of selenocysteins for UGA stop codons rather than degradation of protein but biological significance of UGA recoding in DNAI1 protein still remains unclear.

Genetic diagnosis of PCD, concluding remarks and future perspectives

There is currently no "gold standard" test to diagnose PCD [43]. PCD diagnostic procedures are complex, require expensive infrastructure, and an experienced team of clinicians, geneticists and microscopists [43–45]. Molecular diagnostics is a useful tool, especially in diseases with overlapping clinical presentation. Nevertheless, due to genetic heterogeneity of PCD, approximately 35% of PCD cases still lack confirmation of genetic basis of the disease [46]. NGS approach allows us to generate an extensive data libraries in short period of time and to analyse many genetic variants and potential candidate genes for PCD. Analysis of genetic background of PCD will lead to better understanding of the diseases and, eventually, to the design of new, molecular-targeted therapy.

Supporting information

S1 Fig. Translation of mutated mRNA sequence. Frameshift mutation led to premature stop codon (marked with red rectangle) and a consequentially insertion of sixteen stop codons. Codon for all seventeen stops was UGA (DNAI1 mRNA: NM_12144.3).
(PDF)

S2 Fig. Nucleotide sequence alignment with ClustalW2. Two sequences were aligned from TGA stop codon and results indicate to evolutionary conservation of 53.3% among the mutated DNAI1 mRNA and SelP mRNA transcripts.
(PDF)

S1 Table. Primers designed for validation of detected disease-causing mutations found in our cohort of PCD patients.
(PDF)

S2 Table. Primers designed for RT-qPCR analysis of *DNAI1* gene.
(PDF)

S3 Table. Positions of amino acids in DNAI1 protein essential for protein-protein interactions.
(PDF)

Acknowledgments

This work has been funded by grant from the Ministry of Education, Science and Technological Development, Republic of Serbia (III 41004).

Author Contributions

Conceptualization: Marina Andjelkovic, Predrag Minic, Vesna Spasovski, Sonja Pavlovic.

Data curation: Predrag Minic, Aleksandar Sovtic, Milan Rodic, Vesna Skodric-Trifunovic, Nina Maric, Jelena Visekruna.

Formal analysis: Marina Andjelkovic, Anita Skakic, Aleksandar Sovtic, Milan Rodic, Vesna Skodric-Trifunovic, Nina Maric, Jelena Visekruna.

Investigation: Marina Andjelkovic, Misa Vreca, Anita Skakic, Nina Maric, Vesna Spasovski.

Methodology: Marina Andjelkovic, Misa Vreca, Anita Skakic.

Project administration: Marina Andjelkovic, Maja Stojiljkovic, Sonja Pavlovic.

Supervision: Maja Stojiljkovic, Vesna Spasovski, Sonja Pavlovic.

Validation: Misa Vreca.

Writing – original draft: Marina Andjelkovic.

Writing – review & editing: Predrag Minic, Maja Stojiljkovic, Vesna Spasovski, Sonja Pavlovic.

References

1. Afzelius BA. A human syndrome caused by immotile cilia. Science. 1976; 193(4250):317–9. PMID: 1084576
2. Knowles MR, Daniels LA, Davis SD, al. e. Primary ciliary dyskinesia. Recent advances in diagnostics, genetics, and characterization of clinical disease. American journal of respiratory and critical care medicine. 2013; 188(8):913–22. https://doi.org/10.1164/rccm.201301-0059CI PMID: 23796196
3. Boon M, Smits A, Cuppens H, Jaspers M, Proesmans M, Dupont LJ, et al. Primary ciliary dyskinesia: critical evaluation of clinical symptoms and diagnosis in patients with normal and abnormal ultrastructure. Orphanet J Rare Dis. 2014; 9:11. https://doi.org/10.1186/1750-1172-9-11 PMID: 24450482
4. Coren ME, Meeks M, Morrison I, Buchdahl RM, Bush A. Primary ciliary dyskinesia: age at diagnosis and symptom history. Acta Paediatr. 2002; 91(6):667–9. PMID: 12162599
5. Frija-Masson J, Bassinet L, Honore I, al. e. Clinical characteristics, functional respiratory decline and follow-up in adult patients with primary ciliary dyskinesia. Thorax. 2017; 72(2):154–60. https://doi.org/10.1136/thoraxjnl-2015-207891 PMID: 27382041
6. Daniels ML, Noone PG. Genetics, diagnosis, and future treatment strategies for primary ciliary dyskinesia. Expert Opin Orphan Drugs. 2015; 3(1):31–44. https://doi.org/10.1517/21678707.2015.989212 PMID: 26998415
7. Goutaki M, Meier AB, Halbeisen FS, Lucas JS, Dell SD, Maurer E, et al. Clinical manifestations in primary ciliary dyskinesia: systematic review and meta-analysis. The European respiratory journal. 2016; 48(4):1081–95. https://doi.org/10.1183/13993003.00736-2016 PMID: 27492829
8. Halbert SA, Patton DL, Zarutskie PW, Soules MR. Function and structure of cilia in the fallopian tube of an infertile woman with Kartagener's syndrome. Hum Reprod. 1997; 12(1):55–8. PMID: 9043902
9. O'Callaghan C, Chetcuti P, Moya E. High prevalence of primary ciliary dyskinesia in a British Asian population. Arch Dis Child. 2010; 95(1):51–2. https://doi.org/10.1136/adc.2009.158493 PMID: 19720631
10. Noone PG, Leigh MW, Sannuti A, Minnix SL, Carson JL, Hazucha M, et al. Primary ciliary dyskinesia: diagnostic and phenotypic features. American journal of respiratory and critical care medicine. 2004; 169(4):459–67. https://doi.org/10.1164/rccm.200303-365OC PMID: 14656747
11. Knowles MR, Leigh MW, Carson JL, Davis SD, Dell SD, Ferkol TW, et al. Mutations of DNAH11 in patients with primary ciliary dyskinesia with normal ciliary ultrastructure. Thorax. 2012; 67(5):433–41. https://doi.org/10.1136/thoraxjnl-2011-200301 PMID: 22184204
12. Morillas HN, Zariwala M, Knowles MR. Genetic causes of bronchiectasis: primary ciliary dyskinesia. Respiration; international review of thoracic diseases. 2007; 74(3):252–63. https://doi.org/10.1159/000101783 PMID: 17534128
13. O'Callaghan C, Rutman A, Williams GM, Hirst RA. Inner dynein arm defects causing primary ciliary dyskinesia: repeat testing required. The European respiratory journal. 2011; 38(3):603–7. https://doi.org/10.1183/09031936.00108410 PMID: 21406509
14. Olin JT, Burns K, Carson JL, Metjian H, Atkinson JJ, Davis SD, et al. Diagnostic yield of nasal scrape biopsies in primary ciliary dyskinesia: a multicenter experience. Pediatric pulmonology. 2011; 46(5):483–8. https://doi.org/10.1002/ppul.21402 PMID: 21284095
15. Guo T, Tan ZP, Chen HM, Zheng DY, Liu L, Huang XG, et al. An effective combination of whole-exome sequencing and runs of homozygosity for the diagnosis of primary ciliary dyskinesia in consanguineous families. Scientific reports. 2017; 7(1):7905. https://doi.org/10.1038/s41598-017-08510-z PMID: 28801648
16. Pennarun G, Escudier E, Chapelin C, Bridoux AM, Cacheux V, Roger G, et al. 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(6):1508–19. https://doi.org/10.1086/302683 PMID: 10577904

17. Geremek M, Schoenmaker F, Zietkiewicz E, Pogorzelski A, Diehl S, Wijmenga C, et al. Sequence analysis of 21 genes located in the Kartagener syndrome linkage region on chromosome 15q. Eur J Hum Genet. 2008; 16(6):688–95. https://doi.org/10.1038/ejhg.2008.5 PMID: 18270537

18. Horani A, Brody SL, Ferkol TW. Picking up speed: advances in the genetics of primary ciliary dyskinesia. Pediatr Res. 2014; 75(1–2):158–64. https://doi.org/10.1038/pr.2013.200 PMID: 24192704

19. Djakow J, Svobodova T, Hrach K, Uhlik J, Cinek O, Pohunek P. Effectiveness of sequencing selected exons of DNAH5 and DNAI1 in diagnosis of primary ciliary dyskinesia. Pediatric pulmonology. 2012; 47 (9):864–75. https://doi.org/10.1002/ppul.22520 PMID: 22416021

20. Boaretto F, Snijders D, Salvoro C, Spalletta A, Mostacciuolo ML, Collura M, et al. Diagnosis of Primary Ciliary Dyskinesia by a Targeted Next-Generation Sequencing Panel: Molecular and Clinical Findings in Italian Patients. J Mol Diagn. 2016; 18(6):912–22. https://doi.org/10.1016/j.jmoldx.2016.07.002 PMID: 27637300

21. Djakow J, Kramna L, Dusatkova L, Uhlik J, Pursiheimo JP, Svobodova T, et al. An effective combination of sanger and next generation sequencing in diagnostics of primary ciliary dyskinesia. Pediatric pulmonology. 2016; 51(5):498–509. https://doi.org/10.1002/ppul.23261 PMID: 26228299

22. Onoufriadis A, Shoemark A, Munye MM, James CT, Schmidts M, Patel M, et al. Combined exome and whole-genome sequencing identifies mutations in ARMC4 as a cause of primary ciliary dyskinesia with defects in the outer dynein arm. Journal of medical genetics. 2014; 51(1):61–7. https://doi.org/10.1136/jmedgenet-2013-101938 PMID: 24203976

23. Takeuchi K, Kitano M, Kiyotoshi H, Ikegami K, Ogawa S, Ikejiri M, et al. A targeted next-generation sequencing panel reveals novel mutations in Japanese patients with primary ciliary dyskinesia. Auris, nasus, larynx. 2018; 45(3):585–91. https://doi.org/10.1016/j.anl.2017.09.007 PMID: 28939216

24. Thornton B, Basu C. Real-time PCR (qPCR) primer design using free online software. Biochem Mol Biol Educ. 2011; 39(2):145–54. https://doi.org/10.1002/bmb.20461 PMID: 21445907

25. Antony D, Becker-Heck A, Zariwala MA, Schmidts M, Onoufriadis A, Forouhan M, et al. Mutations in CCDC39 and CCDC40 are the major cause of primary ciliary dyskinesia with axonemal disorganization and absent inner dynein arms. Human mutation. 2013; 34(3):462–72. https://doi.org/10.1002/humu.22261 PMID: 23255504

26. Becker-Heck A, Zohn IE, Okabe N, Pollock A, Lenhart KB, Sullivan-Brown J, et al. The coiled-coil domain containing protein CCDC40 is essential for motile cilia function and left-right axis formation. Nature genetics. 2011; 43(1):79–84. https://doi.org/10.1038/ng.727 PMID: 21131974

27. Casey JP, McGettigan PA, Healy F, Hogg C, Reynolds A, Kennedy BN, et al. Unexpected genetic heterogeneity for primary ciliary dyskinesia in the Irish Traveller population. Eur J Hum Genet. 2015; 23 (2):210–7. https://doi.org/10.1038/ejhg.2014.79 PMID: 24824133

28. Onoufriadis A, Shoemark A, Schmidts M, Patel M, Jimenez G, Liu H, et al. Targeted NGS gene panel identifies mutations in RSPH1 causing primary ciliary dyskinesia and a common mechanism for ciliary central pair agenesis due to radial spoke defects. Hum Mol Genet. 2014; 23(13):3362–74. https://doi.org/10.1093/hmg/ddu046 PMID: 24518672

29. Loges NT, Olbrich H, Fenske L, Mussaffi H, Horvath J, Fliegauf M, et al. DNAI2 mutations cause primary ciliary dyskinesia with defects in the outer dynein arm. Am J Hum Genet. 2008; 83(5):547–58. https://doi.org/10.1016/j.ajhg.2008.10.001 PMID: 18950741

30. Olbrich H, Haffner K, Kispert A, Volkel A, Volz A, Sasmaz G, et al. Mutations in DNAH5 cause primary ciliary dyskinesia and randomization of left-right asymmetry. Nature genetics. 2002; 30(2):143–4. https://doi.org/10.1038/ng817 PMID: 11788826

31. Raidt J, Wallmeier J, Hjeij R, Onnebrink JG, Pennekamp P, Loges NT, et al. Ciliary beat pattern and frequency in genetic variants of primary ciliary dyskinesia. The European respiratory journal. 2014; 44 (6):1579–88. https://doi.org/10.1183/09031936.00052014 PMID: 25186273

32. Merveille AC, Davis EE, Becker-Heck A, Legendre M, Amirav I, Bataille G, et al. CCDC39 is required for assembly of inner dynein arms and the dynein regulatory complex and for normal ciliary motility in humans and dogs. Nature genetics. 2011; 43(1):72–8. https://doi.org/10.1038/ng.726 PMID: 21131972

33. Teves ME, Zhang Z, Costanzo RM, Henderson SC, Corwin FD, Zweit J, et al. Sperm-associated antigen-17 gene is essential for motile cilia function and neonatal survival. Am J Respir Cell Mol Biol. 2013; 48(6):765–72. https://doi.org/10.1165/rcmb.2012-0362OC PMID: 23418344

34. Zhang Z, Zariwala MA, Mahadevan MM, Caballero-Campo P, Shen X, Escudier E, et al. A heterozygous mutation disrupting the SPAG16 gene results in biochemical instability of central apparatus components of the human sperm axoneme. Biol Reprod. 2007; 77(5):864–71. https://doi.org/10.1095/biolreprod.107.063206 PMID: 17699735

35. Zhang Z, Jones BH, Tang W, Moss SB, Wei Z, Ho C, et al. Dissecting the axoneme interactome: the mammalian orthologue of Chlamydomonas PF6 interacts with sperm-associated antigen 6, the

mammalian orthologue of Chlamydomonas PF16. Mol Cell Proteomics. 2005; 4(7):914–23. https://doi.org/10.1074/mcp.M400177-MCP200 PMID: 15827353

36. Weiss RB, Dunn DM, Atkins JF, Gesteland RF. Ribosomal frameshifting from -2 to +50 nucleotides. Prog Nucleic Acid Res Mol Biol. 1990; 39:159–83. PMID: 2247607

37. Beier H, Grimm M. Misreading of termination codons in eukaryotes by natural nonsense suppressor tRNAs. Nucleic Acids Res. 2001; 29(23):4767–82. PMID: 11726686

38. Low SC, Berry MJ. Knowing when not to stop: selenocysteine incorporation in eukaryotes. Trends Biochem Sci. 1996; 21(6):203–8. PMID: 8744353

39. Copeland PR, Stepanik VA, Driscoll DM. Insight into mammalian selenocysteine insertion: domain structure and ribosome binding properties of Sec insertion sequence binding protein 2. Mol Cell Biol. 2001; 21(5):1491–8. https://doi.org/10.1128/MCB.21.5.1491-1498.2001 PMID: 11238886

40. Mathews DH, Sabina J, Zuker M, Turner DH. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. J Mol Biol. 1999; 288(5):911–40. https://doi.org/10.1006/jmbi.1999.2700 PMID: 10329189

41. Howard MT, Aggarwal G, Anderson CB, Khatri S, Flanigan KM, Atkins JF. Recoding elements located adjacent to a subset of eukaryal selenocysteine-specifying UGA codons. Embo j. 2005; 24(8):1596–607. https://doi.org/10.1038/sj.emboj.7600642 PMID: 15791204

42. Tujebajeva RM, Copeland PR, Xu XM, Carlson BA, Harney JW, Driscoll DM, et al. Decoding apparatus for eukaryotic selenocysteine insertion. EMBO Rep. 2000; 1(2):158–63. https://doi.org/10.1093/embo-reports/kvd033 PMID: 11265756

43. Lucas JS, Leigh MW. Diagnosis of primary ciliary dyskinesia: searching for a gold standard. The European respiratory journal. 2014; 44(6):1418–22. https://doi.org/10.1183/09031936.00175614 PMID: 25435529

44. O'Callaghan C, Chilvers M, Hogg C, Bush A, Lucas J. Diagnosing primary ciliary dyskinesia. Thorax. 2007; 62(8):656–7. https://doi.org/10.1136/thx.2007.083147 PMID: 17687094

45. Knowles MR, Daniels LA, Davis SD, Zariwala MA, Leigh MW. Primary ciliary dyskinesia. Recent advances in diagnostics, genetics, and characterization of clinical disease. American journal of respiratory and critical care medicine. 2013; 188(8):913–22. https://doi.org/10.1164/rccm.201301-0059CI PMID: 23796196

46. Kurkowiak M, Zietkiewicz E, Witt M. Recent advances in primary ciliary dyskinesia genetics. Journal of medical genetics. 2015; 52(1):1–9. https://doi.org/10.1136/jmedgenet-2014-102755 PMID: 25351953

Met I P A S A K A P H K Q P H K Q S I S I G R G T R K R D E D S G T E V G
E G T D E W A Q S K A T V R P P D Q L E L T D A E L K E E F T R I L T A N
N P H A P Q N I V R Y S F K E G T Y K P I G F V N Q L A V H Y T Q V G N L
I P K D S D E G R R Q H Y R D E L V A G S Q E S V K V I S E T G N L E E D
E E P K E L E T E P G S Q T D V P A A G A A E K V T E E E L **Met** T P K Q P
K E R K L T N Q F N F S E R A S Q T Y N N P V R D R E C Q T E P P P R T
N F S A T A N Q W E I Y D A Y V E E L E K Q E K T K E K E K A K T P V A K
K S G K **Met** A **Met** R K L T S **Met** E S Q T D D L I K L S Q A A K I **Met** E
R **Met** V N Q N T Y D D I A Q D F K Y Y D D A A D E Y R D Q V G Y P A A A
L E V P K **Stop** Q S Q A P V R H C P L L E S K V Q G S V C S G I W L L **Stop** L
H E A E P G H A A A L Q P E E P Q L P **Stop** V H V Q Q Q Q R R H V S R H
P R G P P L P G G S R P L **Stop** R Q R G H L Q P Q E A P L P A L L L Q L S
Q V W Q A L R P C V A G Q V A E G **Stop** H G P K P **Stop** L L L C V I **Stop** R
Q D C V L D S R E E K A G S H R C H Q A E G G R Q H H G S S **Stop** G V A
A A P S G L W H C L **Stop** L P Q R D **Stop** L H V P S G H R G G K N L Q V
L **Stop** I L L Q P I P R H L **Stop** R P Q H V S G H C V L E P I P H Q G L H V L
Q L R L D S E D L G P H H Q D P D V H L **Stop** P E L S R G **Stop** C G L G A
I L F Y C V R S S H H R W E G P H I **Stop** L S H Q Q V **Stop** G H L Q P A C
G G Q K E Q A H P R A V Q S H P P H H H C G R **Stop**

```
3'UTR of SelP mRNA     ----ATAATTAGCAGTTTAGAATGGAGGAAG--AACAACAAAGACA-----TGCTTTCCA
3'UTR of DNAI1 mRNA    CTGGACAAACTGCTGAACCTGGTGAGGGAAGTGAAAATCA-AGACCTGAGGGGCTGGCCT
                           * **: :**:*:: . ..**..*****  **.*:** ****.      ***  **:

3'UTR of SelP mRNA     TTTTTTTCTTTACTTATCTCTC-AAAACAATATTACTTTGTCTTTTCAATCTTCTACTTT
3'UTR of DNAI1 mRNA    CAGT-CTCTGT-CCCATCGCTTGAATACAGTACTCCTAGGGCTTGACCCT---------
                        : *  *** * *  *** **  **:***.** *.**: * *** :*..*
```

Title of primers	Orientation	Sequence of primers	Length in bp
DNAL1_116/117F	Forward	5'-CTCCCACACAACTACTAGCTC-3'	21
DNAL1_116/117R	Reverse	5'-CTGTTTCTCTCTGAGCTTAGCC-3'	22
DNAL1_162F	Forward	5'-ATCTCCTGATGTCGTGATCCGC-3'	22
DNAL1_162R	Reverse	5'-CACCTTGGCTCAAACCTAGACC-3'	22
LRRC6_9F	Forward	5'-CAAAAACTCCAGGAGGAGGTACC-3'	23
LRRC6_9R	Reverse	5'-CACAGCACTTTCCAAAGTTCTCTC-3'	24
LRRC6_466F	Forward	5'-GGATGGGTGAGTTGAATCTG-3'	20
LRRC6_466R	Reverse	5'-GTTAAAACACTGGAGCAGCG-3'	20
DNAH5_379F	Forward	5'-GAGCTTCAAAGCAGCCATGTAG-3'	22
DNAH5_379R	Reverse	5'-CTCCCTCACATCCTTCCATTTC-3'	22
DNAH5_2542F	Forward	5'-GGGGAGAATTTGGAAGGTGAC-3'	21
DNAH5_2542R	Reverse	5'-GGATGCTTCATGTCTCCCC-3'	19
DNAH5_INTF	Forward	5'-GGAGGAGAGGAGCTTTTTGG-3'	20
DNAH5_INTR	Reverse	5'-GGGGTGAAAAGAGAACTTGGC-3'	21
CCDC40_83F	Forward	5'-CAGGTGTCACCACCAGAG-3'	18
CCDC40_83R	Reverse	5'-CTCTTGCTGGAAGCCTTGG-3'	19
SCNN1A_522F	Forward	5'-GCTGTTTCACCAAGTGCC-3'	18
SCNN1A_522R	Reverse	5'-CCTAGGAAAGAATGGGGGTG-3'	20
DNAH11_2600F	Forward	5'-CCACAGGGTGGATGAAATCG-3'	20
DNAH11_2600R	Reverse	5'-CCAACTTTCCAAAGATACAAAGGG-3'	24
DNAH11_2852F	Forward	5'-CCTGGGCAACAGAGCAAGTC-3'	20
DNAH11_2852R	Reverse	5'-GCCTTGGCCAAATCTGGGAG-3'	20
DNAI1_947_948F	Forward	5'-CCTGGGTTTGCCATAAAGCG-3'	20
DNAI1_947_948R	Reverse	5'-GTACTGGGGGTTTGGGATGTAGAG-3'	24
CCDC40_2440F	Forward	5'-CGGTCTAGGGGGAAGAAGTG-3'	20

CCDC40_2440R	Reverse	5'-GCAGGCAGTGGTTCACATTTAC-3'	22
DNAH5e13_F	Forward	5'-CAGTACCCGTGGAACCTGAC-3'	20
DNAH5e13_R	Reverse	5'-CTATGACCTTGTGACCCAAATTGCC-3'	25
DNAH5e54_F	Forward	5'-CATACGTATACGATAGAACAAAATCAGC-3'	28
DNAH5e54_R	Reverse	5'-CCCAATAGCACTTATTCACACCAC-3'	24
SPAG17e5_F	Forward	5'-CCTCATGCATAAATGTGCCTTGTCAC-3'	26
SPAG17e5_R	Reverse	5'-GCCACTGTGTTTCACTCGTGAAAG-3'	24
SPAG16e4_F	Forward	5'-GATAGGAGGACAGAGATTTGTTTGC-3'	26
SPAG16e4_R	Reverse	5'-GGTTTCAGGGCTGGAGATG-3'	19
CFTR22e_F	Forward	CCTGTTAGTTCATTGAAAAGCCCG	25
CFTR22e_R	Reverse	CTGCTTGCTTTGTTAGACTGTGTTC	24
SPAG16e10_F	Forward	5'-CAGAGTATCAACCTGTGCAGTG-3'	22
SPAG16e10_R	Reverse	5'-CTTAGTCTGTACTGCCACCAAG-3'	22
DNAH5e10_F	Forward	5'-CCCAAGCTCAAGTTAGAAGATTGTG-3'	25
DNAH5e10_R	Reverse	5'-GACATAGTAGGCTCCCAAGGAG-3'	22
DNAH5e49_F	Forward	5'-GGAAACACTCAAGGAGTTAGGCTC-3'	24
DNAH5e49_R	Reverse	5'-CAAGCAGGTGTTCTTCCTCACC-3'	22
DNAI1eg14_F	Forward	5'-GACAGAGATGGATGGAAGGAGGGAG-3'	25
DNAI1eg14_R	Reverse	5'-CTTTGCCATTTAGCTCCCTGCCTC-3'	24
DNAI1eg17_F	Forward	5'-CCAGGGAAGTGGTGGCTGC-3'	19
DNAI1eg17_R	Reverse	5'-CTGCAGAGGTGGAGGGTGGAAAG-3'	23

Title of primers	Orientation	Sequence of primers	Length in bp
DNAI1_F_Exp	Forward	5'-GTCCCAAGCTGCTAAGATCATGGAGCG-3'	27
DNAI1_R_Exp	Reverse	5'-CAGGGTACCCACCTGGTCCC-3	20

Position	337	352	385	387	404	436	456	458	498
Residue	A	Y	G	M	V	W	S	G	C
Position	500	517	541	542	562	586	603	631	633
Residue	T	E	M	S	D	G	T	R	T