The importance of genomic profiling for differential diagnosis of pediatric lung disease patients with suspected ciliopathies

Marina Anđelković¹, Vesna Spasovski¹, Miša Vreća¹, Aleksandar Sovtić², Milan Rodić², Jovana Komazec¹, Sonja Pavlović¹†, Predrag Minić²,³

1Institute of Molecular Genetics and Genetic Engineering, University of Belgrade, Belgrade, Serbia;
2Dr Vukan Ćupić Mother and Child Health Care Institute of Serbia, Belgrade, Serbia;
3University of Belgrade, Faculty of Medicine, Belgrade, Serbia

Received: October 12, 2018
Revised: December 27, 2018
Accepted: December 29, 2018
Online First: March 4, 2019
DOI: https://doi.org/10.2298/SARH181012012A

Accepted papers are articles in press that have gone through due peer review process and have been accepted for publication by the Editorial Board of the Serbian Archives of Medicine. They have not yet been copy edited and/or formatted in the publication house style, and the text may be changed before the final publication. Although accepted papers do not yet have all the accompanying bibliographic details available, they can already be cited using the year of online publication and the DOI, as follows: the author’s last name and initial of the first name, article title, journal title, online first publication month and year, and the DOI; e.g.: Petrović P, Jovanović J. The title of the article. Srp Arh Celok Lek. Online First, February 2017.
When the final article is assigned to volumes/issues of the journal, the Article in Press version will be removed and the final version will appear in the associated published volumes/issues of the journal. The date the article was made available online first will be carried over.

Correspondence to:
Sonja PAVLOVIĆ
Institute of Molecular Genetics and Genetic Engineering, University of Belgrade
Vojvode Stepe 444a, 11010 Belgrade, Serbia
E-mail: sonja@sezampro.rs
The importance of genomic profiling for differential diagnosis of pediatric lung disease patients with suspected ciliopathies

Значај геномског профилисања за диференцијалну дијагнозу педијатријских пацијената са болестима плућа сусpectedних на цилиопатије

INTRODUCTION

Aberrant function of the axonemal structure has been related to the class of disorders collectively known as “ciliopathies” which includes primary ciliary dyskinesia (PCD), Alstrom...
syndrome, Bardet-Biedl syndrome, Joubert syndrome, nephronophthisis polycystic kidney disease, polycystic liver disease, Meckel-Gruber syndrome, hydrocephalus, Jeune syndrome, and laterality defects [1, 2, 3, 4]. PCD was the first disease associated with the aberrant function of motile cilia.

Primary ciliary dyskinesia (PCD (MIM 244400)) is a very clinically and genetically diverse group of disorders with aberrant ciliary motility resulting in respiratory tract disease. Iannaccone et al. [5] found that the disease is inherited in autosomal recessive or X-linked manner. Lucas et al. [6] found that aberrant ciliary function in infancy leads to neonatal respiratory distress in term infants, wet cough, chronic rhinosinusitis, hearing impairment, bronchiectasis, and about 50% of patients have situs inversus (SI). Noone et al. [7] proposed an explanation for the association between PCD and SI, that the normal function of cilia plays a role in regular organ orientation, whereas orientation of the organs in PCD is a random event due to abnormal function of cilia in early embryonic development. Although the diagnosis of PCD is much easier when abnormal placement of the internal organs is present, SI also can mislead physicians to suppose that patient has PCD, which does not have to be correct since SI is common within ciliopathies, can be isolated (without ciliary dysfunction), or can be an unusual state of cystic fibrosis (CF) [7, 8]. In the recent study, Driscoll et al. [9] described that bronchiectasis was reported in 37% of patients with autosomal-dominant polycystic kidney disease as a consequence of sensory cilia dysfunction. Morini et al. [10] said that, similar to cystic fibrosis, in the early years of life, an infant with ciliopathy may be undernourished, which makes the diagnosis a challenge.

The aim of our study was to point out the significance of genomic profiling for differential diagnosis of patients with comprehensive or partial clinical presentation of PCD, using next-generation sequencing (NGS) approach. This approach allows us to analyze causative and candidate PCD genes, but also genes involved in other pediatric lung diseases with clinical presentation similar to PCD, and to establish the correct diagnosis. We also wanted to design a strategy for genomic testing relevant for differential diagnosis of pediatric lung disease patients with suspected ciliopathies.

METHODS

Written informed consent was obtained from all participants and the study protocol was approved by the Ethics Committee of the Dr Vukan Čupić Mother and Child Health Care Institute of
Serbia in Belgrade, Serbia, and has, therefore, been performed in accordance with the standards laid down in the 1964 Helsinki Declaration and its later amendments. All patients were physically examined and in all of them ciliopathy was detected using optical microscopy. The samples were referred to the Institute of Molecular Genetics and Genetic Engineering for genetic profiling of clinically suspected ciliopathy.

**Diagnostic work up**

We conducted genetic analysis of 29 (21 previously reported) subjects from 18 unrelated families of Serbian descent. Subjects with two or more clinical symptoms of PCD, as well as their relatives were included in this study. PCD diagnosis was made based on the clinical symptoms such as: chronic sinusitis, neonatal respiratory distress, bronchiectasis, recurrent pneumonia, SI, cilia motility, infertility, and all patients were treated at the Mother and Child Health Care Institute. Nasal cells sampled by nasal brushing procedure were used to examine cilia motility using optical microscope. Absence of CF was verified by analysis of the presence for *Pseudomonas aeruginosa* and sweat chloride test. The control group for candidate disease-causing variants detected using NGS approach, consisted of in-house collection of 69 subjects from general population of Serbia sequenced by NGS methodology.

**Genomic analysis**

Genomic DNA obtained from 29 subjects (21 patients, two siblings and six parents) was isolated from peripheral blood by QIAampDNA Mini Kit (Qiagen GmbH, Hilden, Germany). Qubit 3.0 fluorimeter (Invitrogen, Carlsbad, CA, USA) was used for measuring the quality and quantity of isolated DNA. Qubit dsDNA HS (Invitrogen) Assay Kit with range of 0.2–100 ng was used. Patients were analyzed by the NGS approach using Clinical-Exome Sequencing TruSight One Gene Panel (Illumina, San Diego, CA, USA). The relatives were analyzed only for detected gene variants in probands by direct Sanger sequencing.

**A gene selection strategy and variant filtering**

To identify the mutational profile involved in the pathogenesis of PCD, we analyzed disease-associated genes according to the Online Mendelian Inheritance in Man (OMIM). We also searched PubMed for studies describing clinical symptoms and genetic background of PCD patients. The first
search implied the following terms: “primary ciliary dyskinesia,” “immotile cilia syndrome,” “ciliary motility disorders,” “Kartagener syndrome,” abnormalities in the ciliary ultrastructure”, “abnormalities in the beating pattern,” “underlying genetic causes of the ciliary dysfunction,” “genes encoding ODA components,” “genes related to assembling of dynein arms,” “genes encoding radial spokes or CP component,” “genes encoding proteins of the N-DRC,” “genetic testing in PCD,” “novel clinical phenotypes for PCD,” “male infertility in PCD,” etc. Then, besides the genes described in the literature, we added genes that are coding for proteins that participate in protein-protein interactions with PCD disease-associated proteins in accordance to String Interaction Network (https://string-db.org/), and genes from the same gene family as PCD disease-causing genes. So, the first step of the analysis comprised 29 genes: (CCDC8, CCDC39, CCDC40, CCDC50, CCDC88, CCDC103, DNAAF1 (LRRC50), DNAAF2 (KTU), DNAAF3, DNAI1, DNAI2, DNAL1, DNAH1, DNAH5, DNAH9, DNAH11, DYNC1H1, HEATR2, HYDIN, LRRC6, NME1, NME8 (TXNDC3), OFD1, RPGR, RSPH4A, RSPH9, SPAG16, SPAG17, and TCTE1).

To identify mutational profile of patients in which we didn’t find variants in above listed genes, we analyzed genes related to individual symptoms of the disease. We searched on PubMed for the following terms: “bronchiectasis”, “bronchiectasis causes,” “chronic cough,” “mucus production,” “increased mucus secretion,” “CF bronchiectasis,” “bronchiectasis without CF,” “respiratory distress in term neonates,” “surfactant metabolism,” “highly expressed lung proteins”, etc. We selected 45 genes for these symptoms that were present in TruSight One Gene Panel: ABCA3, ABCA1, ABCB1, ABCC3, ABCA10, ABCB11, ABCC4, ABCA12, ABCB4, ABCC6, ABCA13, ABCB6, ABCC8, ABCA2, ABCB7, ABCC9, ABCC1, ABCD1, ABCA4, ABCC11, ABCD4, ABCA7, ABCC2, ABCG1, ABCG2, ABCG5, ABCG8, CFTR, MUC1, MUC13, MUC2, MUC3A, MUC4, MUC5B, MUC6, MUC7, SFTPA1, SFTPA2, SFTPB, SFTPC, SFTP, SCNN1A, SCNN1B, SCNN1G, and SLC26A9.

The last group of analyzed genes in our genomic analysis was related to sensory ciliopathies, present in TruSight One Gene Panel, 19 genes in total: (BBS1, BBS2, BBS3, BBS4, BBS5, BBS6, BBS7, BBS8, BBS9, BBIP10, KIF3A, KIF3B, KAP, IFTA, IFTB, IFT43, IFT80, IFT122, and TTC21B). The searching terms were as follows: “diseases associated with ciliary dysfunction,” “ciliopathies,” “sensory ciliopathies.”
Also, all patients were analyzed for presence of homozygous or compound heterozygous variants in CFTR gene to exclude the possibility of CF as final diagnosis.

Overall, we have performed bioinformatic analysis of 93 genes (Table 1).

**Bioinformatic analysis**

Illumina Clinical-Exome Sequencing TruSight One Gene Panel includes all known disease-associated genes described in the OMIM database until 2013, designed to cover all exons and flanking intronic regions of 4,813 genes (~ 62,000 exons). Illumina MiSeq machine was used to generate comprehensive genetic libraries. The software for data analyzing gained using this approach was Illumina Variant Studio 2.0 and 3.0 (Illumina). After automatically annotation of all samples, we have accessed to variant filtering. Within this Illumina software there are numerous information about the analyzed variant, such as genomic coordinates, classifications, conservation scores, cDNA, CDS, and protein positions, transcripts and consequences, but also information about allele frequencies (1,000 Genomes and ExaC) and functional impact of variants (SIFT and PolyPhen2). Only variants that had allele frequencies less than 5%, predicted as damaging, and absent in 69 control samples (TruSight One base), were further analyzed and considered as candidate variants. If the analyzed genetic variant had frequency > 5% in European population, but wasn’t detected in our control samples, and was predicted as pathogenic, we took it into account. The databases used to additionally determine candidate variants: VarSome (which includes the databases: ClinVar, classification according to American College of Medical Genetics and Genomics (ACMG)), gnomAD genomes, gnomAD exomes, The Human Gene Mutation Database (HGMD), 1000 Genomes Project, dbSNP, Exome Aggregation Consortium (ExaC), and Ensembl. Tools used for pathogenicity scoring were: MutationTaster, Deleterious Annotation of Genetic Variants Using Neural Networks (DANN), and Functional Analysis through Hidden Markov Models (FATHMM-MKL). The effects on protein level were examined with *in silico* tools Provean, SIFT, and PolyPhen2.

**RESULTS**

**Strategy for differential diagnosis of pediatric lung disease patients with suspected ciliopathies**
We have proposed an algorithm for differential diagnosis between PCD and other pediatric lung diseases with similar clinical presentation (Figure 1). Clinical diagnosis of PCD was the criterion for genomic profiling of the patients, and it was established when a patient exhibited two or more symptoms of the clinical presentation of PCD.

Using the NGS for analysis of 4813 genes, we generated large libraries with approximately nine thousand variants per patient. Overall, we have done bioinformatic analysis of 93 genes. First, we prioritized 29 PCD causative genes and candidate genes and established a diagnosis for 11/21 (52.38%) patients from nine unrelated families. Than we expanded our genomic analysis to 45 genes related to individual symptoms such as bronchiectasis, atopic asthma, sinusitis and neonatal respiratory distress, and established the diagnosis for 6/21 (28.57%) patients. Thus, total diagnostic rate in our study reached 80.95 %. The last group of analyzed genes were 19 genes related to sensory ciliopathies, but we weren’t able to detect homozygous, compound heterozygous or trans allelic pathogenic variants and establish the diagnosis in the remaining 4/21 (19.04%) patients. However, in 3/21 patients we have found mono allelic pathogenic variants within PCD disease-causing genes (14.28%). That resulted in only 1/21 (4.76%) patient, in whom the genetic background of the disease remained unidentified.

Andelković et al. [11] claimed that genes and variants detected in PCD related genes in 11/21 patients, and mono allelic variants detected in 3/21 patients were previously reported. Among six patients with variants in genes related to individual symptoms of the disease, three (50%) had the same mutation in SCN1A gene (NM_001159576.1, c. 1654T>C, p. Trp552Arg), in two out of three patients, this gene variant was associated with gene variant in CFTR gene (NM_000492.3, c. 3485G>A, p.Arg1162Leu), while in the third patient it was present in homozygous state. In the remaining three patients, gene variants were detected in heterozygous state in ABCA3 (NM_001089.2, c.2125C>T, p.Arg709Trp), in homozygous state in SLC26A9 (NM_134325.2, c.514G>A, p.Val172Met), and in compound heterozygous state in MUC2 (NM_002457.2, c.5735G>T, p.Gly1918Val; c.8084T>G, p.Ile2683Ser) genes. All variants were characterized as pathogenic, damaging or deleterious according to various databases and in silico prediction scores and tools, and
had allele frequency less than 5% in European population. Detected variants in genes related to bronchiectasis, CF, neonatal respiratory distress and asthma are listed in Table 2.

The algorithm we designed provided the mutation detection rate of more than 95%. It also contributed to the improvement of diagnosis rate from 52% (PCD patients) to 81% (all analyzed patients).

**DISCUSSION**

In order to reach an exact diagnosis in patients that exhibited the clinical presentation of PCD, which is similar to clinical presentations of other pediatric pulmonary disorders associated with ciliopathy, we have established the strategy for differential diagnosis based on genomic profiling using NGS methodology. In this study, genomic analysis revealed that 52% of patients had PCD. Additionally, we established the diagnosis of other pediatric lung diseases for 29% of patients.

**Patients with lung diseases presenting overlapping clinical symptoms with PCD patients**

In 6/21 patients without variants in PCD related genes, but with clinical presentation of PCD, extended genomic analysis revealed that this heterogeneous group have genetic background indicating other lung diseases, such as ectopic asthma, NDRS and bronchiectasis without CF. Pathogenic variants in homozygous, compound heterozygous, and trans allelic state were detected in *ABCA3*, *CFTR*, *MUC2*, *SCNN1A*, and *SLC26A9* genes.

Hanukoglu and Hanikoglu [12] reported that the *SCNN1A* gene encodes the alpha subunit of the ENaC (amiloride-sensitive epithelial sodium channel), a channel that allows the movement of sodium ions from the lumen into epithelial cells through the apical cell membrane. Fajac and Viel [13] reported in their previous study that mutations in *SCNN1* gene family might be deleterious for ENaC channel and lead to bronchiectasis, especially in patients that are trans-heterozygotes for ENaC/CFTR mutations. *SCNN1A* gene might be included on the list of potential gene candidates for PCD, since bronchiectasis is one of the well-known PCD clinical manifestations.

Our patient that harbored heterozygous genetic variant in *ABCA3* gene exhibited recurrent bronchiolitis (with hospitalization) and SI. Bronchiolitis affects the tiny airways, and as these airways become inflamed, they swell and fill with mucus, which make breathing difficult. Yamano et al. [14]
stated that the \textit{ABCA3} gene encodes a protein that is localized to the limiting membrane of lamellar bodies in alveolar type II cells. Matsumura et al. \cite{15} indicated in their studies in surrogate cell model systems that \textit{ABCA3} pathogenic gene variants impair the metabolism of surfactant by altering intracellular trafficking, the way of folding the \textit{ABCA3} protein or impairing hydrolysis of ATP. Newborns with homozygous or compound heterozygous mutations in \textit{ABCA3} gene develop progressive, lethal, neonatal-onset respiratory distress due to surface tension–lowering function of lung surfactant \cite{16, 17}. Stahlman et al. \cite{18} reported that in newborns with a single gene variant, if combined with developmental immaturity, it could reduce \textit{ABCA3} gene expression below a functional threshold and the consequence is NRDS.

Patient with homozygous pathogenic variants in \textit{MUC2} gene exhibits symptoms of asthma, bronchitis, and sinusitis. The protein encoded by \textit{MUC2} gene is secreted and forms an insoluble barrier composed of mucous (https://www.ncbi.nlm.nih.gov/gene/4583). \textit{MUC2} gene variants may lead to an altered glycoprotein folding or variable glycosylation levels. Clinically, these changes may modify the function of mucin \cite{19, 20, 21}. Kaliner et al. \cite{22} reported that the mucus is overproduced in the respiratory tract during acute conditions of the disease, and in chronic conditions (bronchitis, asthma, CF, and sinusitis), thereby contributing to mucus blockade of the airways.

In one patient, we have found pathogenic genetic variant in \textit{SLC26A9} gene. This genetic variant was not found in our in-house TruSight One base. \textit{SLC26A9} functions as a Cl––HCO3– exchanger, a Cl– channel and a Na+ coupled transporter \cite{23, 24}. Chang et al. \cite{23} reported that the \textit{SLC26A9} protein is localized to epithelia of the lung and stomach. CFTR and \textit{SLC26A9} have overlapping expression in lung epithelia and gut, making essential protein-protein interactions. Moreover, in a recent GWAS analysis of patients with CF, Sun et al. \cite{25} pointed out that expression of human \textit{SLC26A9} gene is associated with various CF phenotypes. Several studies have indicated that \textit{SLC26A9} mutations may impart large (monogenic) or small (polygenic) phenotypic effects in stomach-related diseases, CF, and/or other disorders in humans. Thus, Chen et al. \cite{26} concluded that the functional characterization of \textit{SLC26A9} variants are clinically very important.

Despite a similar clinical presentation in PCD patients and patients affected with other lung diseases, when the diagnosis of PCD was excluded, and the correct diagnosis of bronchiectasis,
NRDS and asthma was established, more adequate therapy was used for treatment of the diseases, and the quality of life for those pediatric patients was improved.

In the remaining three patients, monoallelic pathogenic variants were detected in above listed 93 genes. For one patient, we could not associate clinical presentation with alterations in any of the genes included in our algorithm. This suggests that our in-house made list of 93 genes should be increased. New studies and new knowledge will enable the establishment of the genetic background in those patients as well. It is also necessary to emphasize that large chromosomal rearrangements encompassing these 93 genes that could be present in our patients, cannot be detected using NGS methodology.

**CONCLUSION**

Our study indicates the significance of genomic profiling for differential diagnosis between patients with comprehensive or partial clinical presentation of PCD and patients with other pediatric lung diseases. The strategy we designed includes NGS analysis of 93 genes relevant for symptoms present in ciliopathies. The proposed diagnostic algorithm contributed to the improvement of diagnosis rate from 52% (PCD patients) to 81% (all analyzed patients). Given the heterogeneity of possible symptoms associated with PCD and ciliopathies in general, there is no uniform approach to establish a precise diagnose. Modern high-throughput genetic approaches allow the causative biallelic mutations identification in about 60% of patients. Although not yet applied for routine diagnostics, NGS turned out to be more efficient and cost-effective in diagnosing PCD and other pediatric lung diseases associated with ciliopathies, compared to traditional sequencing of single genes. An improved strategy for easier and faster establishment of final diagnosis of ciliopathies is mandatory and includes both, clinical and genetic confirmation of the disease. Our study is in accordance with this standpoint.

**ACKNOWLEDGMENT**

This research has been supported by grant III41004 from the Ministry of Education, Science and Technological Development of the Republic of Serbia.
CONFLICT OF INTEREST: None declared
REFERENCES

### Table 1. In-house designed list of genes related to pediatric lung diseases associated with ciliopathies; list of 93 genes identified on Illumina Clinical-Exome Sequencing TruSight One Gene Panel

<table>
<thead>
<tr>
<th>ABCA1</th>
<th>ABCB1</th>
<th>ABCC3</th>
<th>ABCG2</th>
<th>BBS5</th>
<th>CCDC50</th>
<th>DNAH5</th>
<th>HYDIN</th>
<th>KIF3B</th>
<th>MUC6</th>
<th>SCNN1</th>
<th>SFTPBP</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABCA10</td>
<td>ABCB11</td>
<td>ABCC4</td>
<td>ABCG5</td>
<td>BBS6</td>
<td>CCDC8</td>
<td>DNAH9</td>
<td>IFT122</td>
<td>LRRC6</td>
<td>MUC7</td>
<td>SCNN1B</td>
<td>SFTPC</td>
</tr>
<tr>
<td>ABCA12</td>
<td>ABCB4</td>
<td>ABCC6</td>
<td>ABCG8</td>
<td>BBS7</td>
<td>CCDC88</td>
<td>DNAH11</td>
<td>IFT43</td>
<td>MUC1</td>
<td>NME1</td>
<td>SCNN1G</td>
<td>SFTP</td>
</tr>
<tr>
<td>ABCA13</td>
<td>ABCB6</td>
<td>ABCC8</td>
<td>BBIP10</td>
<td>BBS8</td>
<td>CFTR</td>
<td>DNAI1</td>
<td>IFT80</td>
<td>MUC13</td>
<td>NME8</td>
<td>SLC26A9</td>
<td>TTC1</td>
</tr>
<tr>
<td>ABCA2</td>
<td>ABCB7</td>
<td>ABCC9</td>
<td>BBS1</td>
<td>BBS9</td>
<td>DNAAF1</td>
<td>DNAI2</td>
<td>IFTA</td>
<td>MUC2</td>
<td>OFD1</td>
<td>SPAG16</td>
<td>TTC21</td>
</tr>
<tr>
<td>ABCA3</td>
<td>ABCC1</td>
<td>ABCD1</td>
<td>BBS2</td>
<td>CCDC103</td>
<td>DNAAF2</td>
<td>DNAL1</td>
<td>IFTB</td>
<td>MUC5A</td>
<td>RPGR</td>
<td>SPAG17</td>
<td></td>
</tr>
<tr>
<td>ABCA4</td>
<td>ABCC11</td>
<td>ABCD4</td>
<td>BBS3</td>
<td>CCDC39</td>
<td>DNAAF3</td>
<td>DYN1H1</td>
<td>KAP</td>
<td>MUC4</td>
<td>RSPH4A</td>
<td>SFTPA1</td>
<td></td>
</tr>
<tr>
<td>ABCA7</td>
<td>ABCC2</td>
<td>ABCG1</td>
<td>BBS4</td>
<td>CCDC40</td>
<td>DNAH1</td>
<td>HEATR2</td>
<td>KIF3A</td>
<td>MUC5B</td>
<td>RSPH9</td>
<td>SFTPA2</td>
<td></td>
</tr>
</tbody>
</table>
Figure 1. Proposed algorithm for differential diagnosis between PCD and other lung diseases with suspected ciliopathies: candidates for a genetic testing were patients with situs inversus totalis or other situs abnormalities, with both upper and lower respiratory tract disease, history of unexplained neonatal respiratory distress, infertility in reproductive period, and immotility or abnormal cilia motility; 21 patients had more than two above mentioned symptoms and they were analyzed using TruSightOne panel; we have conducted bioinformatics analysis of next-generation sequencing data of all patients for presence of homozygous, compound heterozygous or trans allelic variants in 29 known primary ciliary dyskinesia (PCD) disease-causing and candidate genes; after we had established the diagnosis of PCD for 11/21 52.38% of patients, we extended our analysis on genes related to individual symptoms of the disease (45 genes), and genes related to sensory ciliopathies (19 genes); we reached the mutation detection rate of 95% (20/21), and diagnosis rate of 81% (17/21).