SHORT THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY (PHD)

Challenges in the prenatal and postnatal genetic diagnostics of severe monogenic disorders

by Katalin Koczok, MD

UNIVERSITY OF DEBRECEN
DOCTORAL SCHOOL OF MOLECULAR CELL AND IMMUNE BIOLOGY

Debrecen, 2019

SHORT THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY (PHD)

Challenges in the prenatal and postnatal genetic diagnostics of severe monogenic disorders

by Katalin Koczok, MD

Supervisor:
István Balogh, PhD

UNIVERSITY OF DEBRECEN
DOCTORAL SCHOOL OF MOLECULAR CELL AND IMMUNE BIOLOGY

DEBRECEN, 2019

CHALLENGES IN THE PRENATAL AND POSTNATAL GENETIC DIAGNOSTICS OF SEVERE MONOGENIC DISORDERS

By Katalin Koczok, MD

Supervisor: Dr. István Balogh, PhD

Doctoral School of Molecular Cell and Immune Biology, University of Debrecen

Head of the **Examination Committee**:	Prof. Dr. Gábor Szabó, DSc
Members of the Examination Committee:	Prof. Dr. Margit Balázs, DSc
	Dr. Attila Patócs, DSc

The Examination takes place at the In Vitro Diagnostic Building, Library of the Department of Laboratory Medicine, Faculty of Medicine, University of Debrecen,
22 March 2019, 11:00 a.m.

Head of the **Defense Committee**:	Prof. Dr. Gábor Szabó, DSc
Reviewers:	Prof. Dr. György Panyi, DSc
	Dr. Kálmán Tory, PhD
Members of the Defense Committee:	Prof. Dr. Margit Balázs, DSc
	Dr. Attila Patócs, DSc

The PhD Defense takes place at the Lecture Hall of Bldg. A, Department of Internal Medicine, Faculty of Medicine, University of Debrecen
22 March 2019, 2:00 p.m.

1. Introduction

Molecular genetic diagnostics like other laboratory processes can be divided into preanalytical, analytical and postanalytical phases. In the thesis some special, challenging aspects of these three phases are presented exemplifying how complex and intriguing genetic testing can be. Molecular genetic testing of severe monogenic disorders plays an important role in establishing the diagnosis, possibly giving the prognosis based on genotype-phenotype correlations, aiding in reproductive decision-making, and enabling prenatal diagnostics.

In prenatal molecular genetic diagnostics invasive sampling procedures (e.g. chorionic villus sampling or amniocentesis) may lead to maternal cell contamination (MCC), which is considered as a preanalytical risk for each fetal analysis. In the ***first study*** (***Interfering effect of maternal cell contamination on invasive prenatal molecular genetic testing***) we tested the analytical sensitivity of three molecular genetic methods to MCC: 1) Sanger DNA sequencing, (2) multiplex ligation-dependent probe amplification (MLPA), and a next-generation sequencing (NGS) method, 3) pyrosequencing.

Our ***second study*** (***A novel missense mutation in the DMD gene***) gives an example on a special aspect of the analytical phase, namely the analytical specificity of MLPA method and highlights the importance of confirmatory testing of single-exon deletions detected by MLPA. In addition, we show the difficulty of assessing a novel missense mutation's pathogenicity in the case of a large protein like dystrophin. Necessity of muscle biopsy and dystrophin immunohistochemistry studies underlines the complexity of this case.

In the ***third study*** (***Mutational spectrum of Hungarian Smith-Lemli-Opitz syndrome patients***) we described the mutational spectrum of 13 Smith-Lemli-Opitz syndrome (SLOS) patients diagnosed between 2008-2012 in our laboratory and compared our results to available data in the literature about *DHCR7* gene mutation types, mutation localization, genotypes, and mutational spectrum of other European populations. In addition we summarized biochemical and clinical features of the patients. A novel missense mutation was also detected, pathogenicity of which was assessed using *in silico* prediction tools and a recently published variant classification system.

1.1 Review of the literature

1.1.1 Interfering effect of maternal cell contamination on invasive prenatal molecular genetic testing

Prenatal molecular genetic testing is now widely available for a number of severe Mendelian disorders. Although noninvasive testing for single-gene disorders using cell-free fetal DNA in the maternal plasma is emerging, invasive prenatal diagnostics still accounts for the majority of prenatal testing in many genetic laboratories. Invasive sampling, however, raises the risk of maternal cell contamination. In chorionic villus (CVS) and amniotic fluid (AF) samples, MCC is a preanalytical risk factor for each prenatal molecular analysis. CVS can be contaminated by maternal decidua, while the presence of maternal blood cells can complicate AF sample analysis. Maternal cell contamination occurs less frequently in cultured AF samples because culturing favours the growth of amniocytes over maternal peripheral blood cells. Cultured CVS is more prone to contamination than direct ones, the degree of MCC can be reduced by careful dissection of maternal decidua from chorionic villi before culture set-up. Since PCR-based methods are highly sensitive to DNA contamination, genotype of the fetus may be falsely determined in the presence of MCC.

It is recommended for the genetic laboratory to determine the routinely used diagnostic tests' sensitivity to MCC. It has been shown that significant MCC level affecting result interpretation of prenatal tests is dependent on the applied method, and the presence of MCC does not necessarily invalidate prenatal diagnostic test results. Chamberlain et al investigated the sensitivity of multiplex PCR used for the detection of deletions in the dystrophin gene and showed that 3-5% MCC could be tolerated if the cycle number was limited to 25. In contrast with, Lamb et al reported that only high levels of MCC (>20%) proved to be significant when using a 135K oligonucleotide-based microarray. According to Hessner et al, allele-specific PCR used to detect paternally inherited alloalleles in fetuses at risk for immune cytopenic disorders was still informative in the presence of 94-99.8% MCC.

The importance of determining the routinely used molecular methods' sensitivity to MCC is underlined by the fact that MCC was shown to be present in up to 7.1% of different fetal samples.

1.1.2 Molecular genetic diagnostics of dystrophinopathies, dystrophin protein

Dystrophinopathies are the most common neuromuscular disorders. Mutations in the *DMD* gene (Xp21.2-p21.1, MIM 300377) encoding dystrophin protein lead to the X-linked allelic disorders Duchenne (DMD, MIM 310200) and Becker muscular dystrophy (BMD, MIM

300376), both diseases are characterized by progressive muscle wasting. DMD is the more severe form caused by mutations that lead to nearly complete absence of dystrophin protein. BMD, the milder variant, is characterized by remarkable variability in presentation and progression of symptoms and associated with abnormal but partially functional dystrophin protein.

Dystrophin is a sarcolemma-associated protein interacting with integral membrane and cytoplasmic proteins (sarcoglycans, dystroglycans, syntrophins, and dystrobrevins) assembled in the dystrophin-glycoprotein complex. These interactions enable a link formed between the extracellular matrix and the cytoskeletal actin. Dystrophin's main role is the stabilization of sarcolemma and protection of muscle fibers from contraction-induced damage. Three full-length isoforms of the protein are generated through different promoters (brain, Purkinje and muscle). The full-length human dystrophin protein transcribed from the muscle promoter consists of 3685 amino acids with a molecular weight of 427 kDa and is predicted to fold into four structural domains: 1) the N-terminal actin-binding domain (N-ABD), 2) a central rod domain (containing a second ABD), 3) a cysteine-rich domain binding to the integral membrane protein beta-dystroglycan, and 4) a C-terminal domain, which binds to dystrobrevin and syntrophins.

Clinical suspicion of DMD/BMD based on elevated serum creatine kinase (CK) and/or muscle biopsy results can be confirmed by molecular genetic testing of the *DMD* gene. The first tier genetic analysis is screening for deletions/duplications, which account for about 65%–80% of genetic alterations in DMD/BMD patients. If no deletion or duplication is found, sequencing of the *DMD* gene as a second tier test can be offered. The great majority of point mutations (20-35% of causative genetic alterations in the *DMD* gene) are nonsense mutations, small frameshift deletions/insertions and splice-site mutations.

The *DMD* gene is the largest human gene spanning more than 2.5 million base pairs and consisting of 79 exons. The high mutation rate of the gene has been claimed to be the consequence of very large introns. In one of three patients the disease is the result of a *de novo* mutation. An apparently sporadic case can be the result of germline mosaicism of the mother, which has to be taken into account when considering the recurrence risk.

Structural elements of dystrophin protein (*i.e.* maintaining the link between actin and the extracellular matrix) are well characterized. Based on the crystal structure an antiparallel dimer is formed by the N-ABDs of two dystrophin molecules. The N-ABD of human dystrophin is mainly helical and contains two calponin homology (CH) domains. The CH1 and CH2 subdomains are connected by a helical linker region (helix I). CH2 subdomain contains major

(A, C, E, and G) and short (B, D, and F) helices. Overall fold of CH1 closely resembles that of CH2 but lacks the short B and D helices, the latter being substituted by loop regions in this subdomain. Three actin-binding sites are responsible for interactions with actin molecules.

Missense mutations are rare cause of DMD/BMD, but, interestingly, about half of them affect N-ABD of dystrophin protein.

1.1.3 Smith-Lemli-Opitz syndrome

Smith-Lemli-Opitz syndrome (SLOS, MIM 270400) is the first described and most common cholesterol biosynthesis disorder caused by deficiency of 7-dehydrocholesterol reductase enzyme (DHCR7).

The human DHCR7 is a NADPH-dependent transmembrane protein localized to the endoplasmic reticulum, its predicted molecular weight is 54.5 kDa. Due to deficient DHCR7 enzyme activity production of cholesterol and desmosterol is decreased in SLOS, while 7-dehydrocholesterol (7-DHC) and its isomer 8-dehydrocholesterol (8-DHC) is elevated in the serum and tissues of patients. Clinical diagnosis of SLOS can be confirmed by measurement of increased serum/plasma 7-DHC concentration. Serum cholesterol concentration can be in the normal range in about 10% of patients, therefore this parameter cannot be used for exclusion of the disease.

Clinically SLOS is characterized by failure to thrive, developmental delay, mental retardation and behavioral disturbances. Multiple congenital malformations are common and the “SLOS face” is very distinctive. Syndactyly of the 2nd and 3rd toes is present in more than 95% of patients.

Cholesterol has several important biological functions. It is an essential lipid component of the mammalian cell membrane, main constituent of myelin, precursor of steroid hormones, neurosteroids, oxysterols, and bile acids. Its role in embryogenesis and morphogenesis has been described. Cholesterol is a structural component of lipid rafts, altered sterol composition of the cell membrane influences its physico-chemical and functional properties. The effect of substitution of cholesterol with 7-DHC on several metabolic pathways has not been defined yet. Cholesterol deficiency and increased concentration of 7-/8-DHC and their metabolites are all likely to contribute to the broad phenotypic spectrum of the disease.

SLOS is a monogenic disorder with autosomal recessive inheritance pattern, 217 mutations have been described in the *DHCR7* gene (11q13.4, MIM 602858) according to the Human Gene Mutation Database (Professional Human Gene Mutation Database, 2018.2

release). DHCR7 is ubiquitously expressed with the highest mRNA levels in the adrenal gland, liver, testis, and the brain.

SLOS is relatively common in the Caucasians with 1:10 000-1:70 000 incidence in populations of Northern and Central European heritage. In spite of high carrier frequency of null mutations, incidence of 0/0 genotype is low, which is probably the consequence of prenatal loss and neonatal mortality.

DHCR7 gene mutations show an intersting geographical distribution in Europe, which is most likely the consequence of founder effect, genetic drift and recurrent mutations.

2. Objectives

1) Our first aim was to test the analytical sensitivity of three molecular genetic methods to MCC:

i) Sanger DNA sequencing
ii) multiplex ligation-dependent probe amplification
iii) pyrosequencing.

2) As part of our molecular genetic diagnostic work we detected a novel missense mutation in the 4th exon of *DMD* gene affecting the 76th residue of N-ABD of dystrophin protein in a 14-month-old male patient. In our second study we aimed to investigate phenotypic effect of this novel mutation based on the clinical picture, muscle biopsy and dystrophin immunohistochemistry studies, and *in silico* analysis of the protein structure.

3) In the third study we aimed to summarize molecular genetic, clinical and biochemical (cholesterol and 7-DHC concentration) data of 13 Hungarian SLOS patients.

3. Materials and methods

3.1 Interfering effect of maternal cell contamination on invasive prenatal molecular genetic testing

3.1.1 MCC simulation

In all experiments, genomic DNA mixtures were created by contaminating a wild type (used as "fetal") DNA sample with a heterozygous (used as "maternal") DNA sample simulating 1%, 5%, 10%, 20%, 30%, and 40% MCC levels corresponding to 0.5%, 2.5%, 5%, 10%, 15%, and 20% mutant allele per DNA mixture, respectively.

Genomic DNA was extracted from peripheral blood leukocytes using QIAamp DNA Blood Mini Kit (Qiagen, Hilden, Germany). DNA concentration was measured on NanoDrop 2000 Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) according to the manufacturer's instructions.

3.1.2 Determination of Sanger DNA sequencing's sensitivity to MCC

The sensitivity to MCC was tested for 6 small-scale mutations: (1) 5 single-base substitutions and (2) a 3-bp deletion.

1. A wild type DNA sample ("fetal") was mixed with a DNA sample ("maternal") heterozygous for the following base substitutions (NM_138694.3: c.8870T>C, c.3407A>G, c.7916C>A, c.6992T>A, and c.107C>T) in the *PKHD1* gene (fibrocystin/polyductin, 6p12.3-p12.2, MIM 606702). Mutations in the *PKHD1* gene lead to autosomal recessive polycystic kidney disease (ARPKD, MIM 263200).
2. A wild type DNA sample ("fetal") was mixed with a DNA sample ("maternal") heterozygous for deletion of 3 nucleotides (NM_000492.3: c.1521_1523delCTT, legacy nomenclature: F508del) in the *CFTR* gene (cystic fibrosis transmembrane conductance regulator, 7q31.2, MIM 602421). Mutations in the *CFTR* gene lead to cystic fibrosis (CF, MIM 219700), a disease inherited in autosomal recessive manner.

Amplicon sequencing was performed using the BigDyeTerminator v3.1 Cycle Sequencing kit (Applied Biosystems, Foster City, CA, USA) according to the manufacturer's instructions. Samples were run on the ABI PRISM 310 Genetic Analyzer, and data were analysed using the Sequencing Analysis Software (Applied Biosystems).

3.1.3 Determination of MLPA's sensitivity to MCC

The sensitivity of MLPA to MCC was tested for a 1) multiexon deletion and a 2) multiexon duplication. All MLPA experiments were performed according to the manufacturer's instructions. Data were analysed using block normalization.

1) A wild type female DNA sample ("fetal") was mixed with a female DNA sample ("maternal") heterozygous for deletion of exons 1 to 3 (NG_008197.1(NM_000381.3):c.(?_-365)_(864+1_865-1)del) in the *MID1* gene (midline 1, Xp22.2, MIM 300552). Mutations in the *MID1* gene lead to X-linked Opitz G/BBB syndrome (XLOS, MIM 300000). The MLPA analysis was performed using the SALSA MLPA probemix P233-B2 MID1 (MRC-Holland, Amsterdam, the Netherlands). Because of X-linked inheritance pattern, deletion of the probe's recognition sequence leads to complete absence of the corresponding probe amplification product in males, while a heterozygous deletion in females is recognizable by 35% to 50% reduction of dosage quotient.

2) A wild type DNA sample ("fetal") was mixed with a DNA sample ("maternal") heterozygous for duplication of exons 33 to 35 (NG_008753.1(NM_138694.3):c.(5236+ 1_5237-1)_(5751+1_5752-1)dup) in the *PKHD1* gene. The MLPA analysis was performed using the SALSA MLPA probemix P341-B2/P342-B2 PKHD1 (MRC-Holland). Heterozygous duplication results in a typical increase of dosage quotient (range: 1.30-1.65).

3.1.4 Determination of the NGS method's sensitivity to MCC

A wild type DNA sample ("fetal") was mixed with a DNA sample ("maternal") heterozygous for a single-base substitution (NM_000138.4:c.4727T>C) in the *FBN1* gene (fibrillin 1, 15q21.1, MIM 134797). Mutations in the *FBN1* gene lead to Marfan syndrome (MFS, MIM 154700), a disease inherited in autosomal dominant manner. Bidirectional pyrosequencing was performed on GS Junior System, and data were analysed using the GS Amplicon Variant Analyzer software (Roche 454 Life Sciences, Branford, CT, USA). Samples were run in duplicates.

3.2 A novel missense mutation in the *DMD* gene

3.2.1 DNA isolation

Genomic DNA was extracted from peripheral blood leukocytes as described in **3.1.1**.

3.2.2 Deletion/duplication analysis of *DMD* gene

MLPA analysis for the detection of deletions/duplications in the *DMD* gene was performed using the SALSA MLPA probe mixes P034-B2 DMD and P035-B1 DMD (MRC-Holland). Data were analyzed according to the manufacturer's instructions. Briefly, relative peak heights were generated after intra- and inter-sample normalization using sex-matched control samples. Deletion of the probe's recognition sequence leads to complete absence of the corresponding probe amplification product in males.

3.2.3 Sanger DNA sequencing of exon 4 of *DMD* gene

After PCR amplification of the coding sequence and adjacent intronic regions of exon 4 of *DMD* gene bidirectional Sanger DNA sequencing as described in **3.1.2**.

3.2.4 Sequencing of the entire coding region of *DMD* gene

Sequencing of *DMD* gene was performed by bidirectional pyrosequencing on GS Junior system using the DMD MASTR assay (Multiplicom, Niel, Belgium). Data were analyzed with the GS Amplicon Variant Analyzer software (Roche 454 Life Sciences).

3.2.5 Muscle histology, dystrophin immunohistochemistry

Frozen sections of 7 micrometer thickness from muscle biopsy specimen were processed for histology, histochemistry and immunohistochemistry according to standard protocols. Regarding detection of sarcolemmal and sarcolemma associated proteins, immunohistochemistry was performed by using antibodies targeting dystrophin epitopes 1, 2, and 3, corresponding to central/core, C-terminal and N-terminal regions, respectively, (Novocastra, Newcastle, UK, distributed by Biomarker, Gödöllő, Hungary; primary antibody dilution was 1:20 in all cases), sarcoglycan alpha, beta, gamma, and delta (Novocastra, 1:50, 1:100, 1:100, and 1:50, respectively); merosin (Novocastra, 1:100) and spectrin (Novocastra, 1:100). Primary antibodies were incubated on slides for 1 hour at room temperature. Visualization of antigen-primary antibody binding was performed by using MACH 4 Universal HRP-Polymer Detection System (Biocare Medical, Pacheco, California, Amerikai Egyesült Államok) according tot he manufacturer's instructions. (*Dr. Tibor Hortobágyi, Division of Neuropathology, Department of Pathology, University of Debrecen)*

3.2.6 *In silico* analyses

Crystal structures of the ABDs of dystrophin (PDB ID:1DXX), utrophin (PDB ID: 1QAG), α-actinin 3 (PDB ID: 1WKU), and fimbrin (PDB ID: 1AOA) proteins available in Protein Data Bank database were used, while other protein information were downloaded from UniProt database for dystrophin (UniProt ID: P11532), utrophin (UniProt ID: P46939), α-actinin 3 (UniProt ID: Q08043), and fimbrin (UniProt ID: P13797). Structural figures were prepared by PyMOL Molecular Graphics System (Version 1.3 Schrödinger, LLC). Structure-based sequence alignment was performed by using SALIGN web server. Secondary structure predictions were performed by JPred4 server. FoldX algorithm and SDM web server were used to calculate protein stability changes (ΔΔG, kcal/mol) upon point mutations using crystal structure of dystrophin ABD. Aggregation properties were predicted based on the structure of chain A of N-ADB by Aggrescan3D server using dynamic mode of prediction.

(*Dr. János András Mótyán, Department of Biochemistry and Molecular Biology, University of Debrecen*)

3.3 Mutational spectrum of Hungarian Smith-Lemli-Opitz syndrome patients

3.3.1 Patients

Between 2008-2012 we investigated 13 (12 unrelated) patients from different parts of Hungary with typical clinical features of SLOS. The clinical diagnosis of SLOS was confirmed by measurement of 7-DHC and further molecular genetic analysis of the *DHCR7* gene was performed. Clinical information was available in each case and the severity score was calculated according to Kelley and Hennekam.

3.3.2 Methods

3.3.2.1 Biochemical measurements

Serum samples were used for 7-DHC measurements carried out by the modified method of Honda et al. Total cholesterol was measured by an enzymatic colometric method.

3.3.2.2 Analysis of the coding region of the DHCR7 gene

DNA isolation from blood leukocytes was performed as described in **3.1.1** and molecular genetic analysis of *DHCR7* was performed by amplifying exons 3–9 and the exon flanking intronic regions. PCR primers used were as described by Fitzky et al except for exon 7 (e7F: 5'-GCT GGG CTC TCG CTA AGT AA-3', e7R: 5'-GCA GTA GAT TAA GGT CAT GGG AAT-3') and exon9a (e9aR: 5'-ATG TAG AAG TAG GGC AGC AGG TGG C-3'). Sanger

DNA sequencing of PCR products was performed as described in **3.1.2**. Electropherograms of the sequenced PCR products were compared to the reference sequence of the *DHCR7* gene: NG_012655.2, NM_001360.2. In each case DNA samples of the parents were analyzed to prove that the detected mutations are found in *trans* in the patient. In the case of patient 11, only the parents' DNA samples were available.

3.3.2.3 Analysis of the non-coding exons and regulatory sequences of the DHCR7 gene

Analysis of the putative promoter region and the 3' polyadenylation site was performed using the following primers: PM1F (5'-CCT GAC ACC CATGTG TTT ACC-3'), PM1R (5'-CCG GAC TCG AGA TTG ACG-3'), PM2F (5'-AGC TGG GAT CCC GAA GAA G-3'), PM2R (5'-CAC GCC GCC TAC CCT CTA-3'), PM3-e1F (5'-GCA ATC GCTGAC ATC ATC C-3'), PM3-e1R (5'-CCT GTG AGT GGG CACCTG-3'), e2F (5'-GAG CTT CTG CCC TCT CCT G-3'), e2R (5'-TCA ACG CTG TGA AGC CAT AG-3'), 3UTR1F (5'-CTA CATGGC CAT CCT GCT G-3'), 3UTR1R (5'-CCC CAG GGA CACTGA TTA GA-3'), 3UTR2F (5'-GTT CCT TGC TTT TGC CTTCA-3') és 3UTR2R (5'-GGA GGG GGA TCT AGA GCT GA-3'). Sanger DNA sequencing of PCR products was performed as described in **3.1.2**.

3.3.2.4 Sequencing of the DHCR7 mRNA

Total RNA was extracted from EDTA-anticoagulated blood samples using Trizol reagent (Invitrogen, Carlsbad, CA, USA) and reverse transcription was carried out using High Capacity cDNA kit (Thermo Fisher Scientific). *DHCR7* cDNA was amplified in 4 overlapping fragments using *Pfu* polymerase (Agilent Technologies, Santa Clara, CA, USA). Following primers were used for the amplification: DHCR7mR1F (5'-TGA CAG AACCGC ATC TCA A-3'), DHCR7mR1R (5'-CTC CTA CGT AGCCGG GTA GA-3'), DHCR7mR2F (5'-TAC CTT GTG GGT CACCTT CC-3'), DHCR7mR2R (5'-ATT GGT CAC ATG GCT GTGG-3'), DHCR7mR3F (5'-GAT CGG GAA GTG GTT TGA CT-3'), DHCR7mR3R (5'-GCC ACC CGG AAG ATG TAG TA-3'), DHCR7mR4F (5'-GAC TGT GTC TGG CTG CCT TA-3') és DHCR7mR4R (5'-AGG ATG GCC ATG TAG ATG ATG-3'). Sanger DNA sequencing of PCR products was performed as described in **3.1.2**.

3.3.2.5 Assessing the pathogenicity of the novel variant in the DHCR7 gene

In order to assess possible deleterious effect of p.Tyr125Cys mutation, *in silico* analysis was performed using Sorting Intolerant From Tolerant (SIFT) prediction software. Sequence alignment of DHCR7 proteins was carried out using ClustalW. The presence of the c.374A>G

(p.Tyr125Cys) mutation in 50 healthy individuals was tested by restriction fragment length polymorphism (RFLP) analysis using *Sna*BI enzyme (New England Biolabs, Ipswich, MA, USA) digestion of PCR-amplified products.

3.4 Classification of sequence variants

Novel sequence variants were classified according to the variant classification recommendation (2015) of the American College of Medical Genetics and Genomics. This classification system applies two sets of criteria for classifying detected variants into either „pathogenic" or „likely pathogenic", or „benign" or „likely benign" categories.

4. Results

4.1 Interfering effect of maternal cell contamination on invasive prenatal molecular genetic testing

4.1.1 Determination of Sanger DNA sequencing's sensitivity to MCC

We performed mixing experiments for 6 different small-scale mutations in order to examine the effect of MCC on Sanger DNA sequencing.

When a wild type DNA sample ("fetal") was mixed with a DNA sample ("maternal") heterozygous for the following base-substitutions, limit of detection (LOD) of the mutant allele was as follows: 1) NM_138694.3:c.8870T>C, 2.5% mutant allele (5% MCC); 2) and 3) NM_138694.3:c.3407A>G and c.7916C>A, 5% mutant allele (10% MCC); 4) NM_138694.3:c.6992T>A 10% mutant allele (20% MCC); 5) NM_138694.3:c.107C>T, 15% mutant allele (30% MCC).

In the mixing experiment when the wild type DNA sample ("fetal") was mixed with a DNA sample ("maternal") heterozygous for a 3-bp deletion (NM_000492.3: c.1521_1523delCTT), the LOD was 10% mutant allele (20% MCC).

4.1.2 Determination of MLPA's sensitivity to MCC

We performed mixing experiments for 2 different mutation types: 1) a multiexon deletion and 2) a multiexon duplication.

1) In the first experiment, a wild type female DNA sample ("fetal") was mixed with a female DNA sample ("maternal") heterozygous for deletion of exons 1 to 3 in the *MID1* gene. The mixing experiment showed that up to 30% simulated MCC, there was no impact on genotyping, dosage quotients of probes hybridizing to the deleted region were in the normal range. Fourty

percent MCC corresponding to 20% mutant allele resulted in dosage quotients of 0.70 to 0.89 of relevant probes, which result may already lead to diagnostic uncertainty.
2) In the second experiment, a wild type DNA sample ("fetal") was mixed with a DNA sample ("maternal") heterozygous for duplication of exons 33 to 35 in the *PKHD1* gene. Data from the mixing experiments showed that even 40% simulated MCC (20% mutant allele) had no impact on genotyping as dosage quotients of probes recognizing the duplicated region were in the normal range.

4.1.3 Determination of pyrosequencing-based NGS's sensitivity to MCC

A wild type DNA sample ("fetal") was mixed with a DNA sample ("maternal") heterozygous for a single-base substitution (NM_000138.4: c.4727T>C) in the *FBN1* gene. As low as 0.5% mutant allele could be detected by pyrosequencing at high coverage. In the range of 2.5%, 5%, 10%, 15%, and 20% mutant allele fractions (mean coverage 752x, range: 437-1165x), quantification was accurate as the difference between expected and observed mutant allele fractions was acceptably low.

4.2 A novel missense mutation in the *DMD* gene

4.2.1 Case history

A six-month-old male patient was sent for neurologic examination because of mild generalized muscle weakness, hypotonia, and delayed motor development. Perinatal history, mental development, and sensation were found to be normal. As in spite of sensory-motor training no improvement of symptoms was observed, a neuromuscular disease was suspected. Electromyography showed myopathic features. Laboratory investigation at 1-year-old age showed markedly elevated serum creatine kinase (1497 U/L, reference range: 24–195 U/L) prompting the consideration of dystrophinopathy and *DMD* gene analysis was requested at the age of 14 months. There was no family history of DMD/BMD. Follow-up physical examination revealed difficulties in sitting up from lying, scapula alata, decreased muscle bulk on upper extremities and chest in addition to mild calf hypertrophy at the age of 25 months when muscle biopsy was performed from the quadriceps muscle. The serum creatine kinase activity showed an increasing tendency during the follow-up period (12-33 months).

4.2.2 Molecular genetic testing of *DMD* gene

Deletion/duplication testing of *DMD* gene by MLPA showed complete absence of the amplification product of probes hybridizing to exon 4. PCR amplification of exon 4, however,

did not confirm the deletion. Sanger sequencing of exon 4 revealed a novel point mutation (LRG_199t1: c.227A>T, LRG_199p1: p.Asn76Ile) leading to the change of Asn residue to Ile at the 76th position. The mutation proved to be *de novo* being not detectable in the mother's genomic DNA. In order to exclude the possibility of another potentially pathogenic mutation in the *DMD* gene, the whole coding region of the gene was sequenced. Data analysis revealed three hemizygous variants. Besides the unreported missense mutation c.227A>T, previously detected by Sanger sequencing, two other variants were found (both present in the Single Nucleotide Polymorphism Database (dbSNP) with >5% minor allele frequency (MAF)).

.

4.2.3 Muscle histology, dystrophin immunohistochemistry

Biopsy showed mild dystrophic features with mild variation in fiber size, occasional hypercontracted fibers, mild patchy increase in internal nucleation, and clusters of basophilic regenerating fibers. Dystrophin expression with antibody 1 showed variable intensity of sarcolemmal staining, segmentally severely reduced or lost. With antibody 2 there was intense, linear sarcolemmal expression consistent with normal staining pattern. With antibody 3 the majority of the fibers was negative, with revertant fibers with complete or segmental sarcolemmal expression in less than 5% of fibers. The spectrin immunohistochemistry revealed normal, linear sarcolemmal expression. The α, β, γ, and δ sarcoglycan staining showed linear sarcolemmal expression of moderate to high intensity, in occasional fibers with segmental decrease of expression. Merosin showed strong, linear sarcolemmal expression. (*Dr. Tibor Hortobágyi, Division of Neuropathology, Department of Pathology, University of Debrecen)*

4.2.4 Predicting the effects of Asn76Ile mutation

4.2.4.1 Asn76 residue of human dystrophin

Structure of the N-ABD of human dystrophin was analyzed to explore the possible role of Asn76 residue in domain stabilization. Based on the crystal structure, the Asn76 residue is located within the N-ABD, in helix E of CH1 subdomain. According to published data, Asn76 does not constitute a part of any actin-binding site and is not involved directly in actin binding. Side chain of Asn76 is not exposed to the surface of CH1 subdomain, rather it is buried between C and E helices and loop D. Side chain atoms of Asn76 are within hydrogen bond distance with main chain atoms of Asp46 and Glu65 residues in the crystal structure of N-ADB (Asn76.ND2-Asp46.O and Asn76.OD1-Glu65.N, respectively). These H-bonds may provide stabilizing interactions within the CH1 subdomain, making contact between helix E (Asn76), loop B (next to the N-terminal end of helix C) (Asp46), and loop D (Glu65).

4.2.4.2 Structure-based alignment of CH1 subdomain-containing proteins

Sequence alignment of human dystrophin, utrophin, α-actinin 3, and fimbrin proteins' N-ABDs showed high conservation of Asn76 residue (dystrophin numbering). Conservation of Asp46 and Glu65 residues interacting with Asn76 was also investigated. The Asp residue (Asp46 according to dystrophin numbering) was found to be not fully conserved in the case of aligned CH1 subdomain-containing human protein sequences, while 1 and 4 isoforms of α-actinin contain Asp residue, Asn can be found in this position in isoforms 2 and 3. Residues corresponding to Glu65 of human dystrophin show the conservation of side chains having strongly similar properties (Asn, Asp, or Glu) in this position.

4.2.4.3 Predicted effects of Asn76Ile mutation

Secondary structure prediction was performed using the sequence of p.Asn76Ile mutant N-ABD and showed no disruption of helix E, the mutant protein was predicted to maintain the overall secondary structural organization of the wild type N-ABD. Structure of the mutant human dystrophin protein showed the loss of the conserved H-bonds of Asn76 with the neighboring Asp46 and Glu65 residues, due to the substitution of a polar Asn residue to a hydrophobic Ile. The mutation was predicted to be highly destabilizing ($\Delta\Delta G > 3$ kcal/mol) implying its possible association with protein malfunction. The impact of p.Asn76Ile mutation on dystrophin aggregation propensity/solubility was also predicted. The analysis showed that both wild type (Asn76) and mutated (Ile76) residues are buried and are not part of aggregation prone regions in the CH1 subdomain. (*Dr. János András Mótyán, Department of Biochemistry and Molecular Biology, University of Debrecen*)

4.3 Mutational spectrum of Hungarian Smith-Lemli-Opitz syndrome patients

4.3.1 Serum cholesterol and 7-DHC measurements

Both cholesterol and 7-DHC levels were measured in the same serum samples of the patients. Extremely elevated serum 7-DHC levels were measured in all patients, ranging from 87 mg/L to 302 mg/L, which is 580-2013-fold higher than the reference value (<0.15 mg/L). Cholesterol levels were low 8/13 patients compared to the lower limit of their age-specific reference range. 7-DHC/cholesterol ratio was abnormal in all patients.

4.3.2 Molecular genetic analysis of *DHCR7* gene

4.3.2.1 DHCR7 gene mutation types in Hungarian SLOS patients

In 12/13 patients both causative mutations were identified, in one case only one *DHCR7* mutation was detected in heterozygous form. In 12 unrelated individuals 23 SLOS alleles were identified: 56,5% (13/23) missense, 26,1% (6/23) nonsense and 17,4% (4/23) splicing mutations. In 11 unrelated patients both disease causing mutations were determined. In all individuals but one compound heterozygosity was detected. In one patient homozygosity for the most common null allele (c.964–1G>C) was established (9.1%, 1/11). In the remaining 10 cases *DHCR7* genotypes were as follows: in 7/11 patients (63.6%) null allele (nonsense or splicing mutation)/missense mutation while in 3/11 patients (27.3%) two different missene mutations were detected.

4.3.2.2 Mutational spectrum of Hungarian SLOS patients

We compared the mutational spectrum of Hungarian SLOS patients to mutations previously described in different European populations. The panethnic c.964–1G>C *DHCR7* mutation, which is most common in Great Britain and also frequent in Spain, Italy, Germaly and Austria, but rare in Poland, the Czech Republic, Slovakia and Southern Italy, was the second most common *DHCR7* mutation among Hungarian SLOS patients (4 alleles). p.Trp151* mutation was the most common (6 alleles) in our patient group. p.Trp151* is frequent in Poland, Germany, Austria, Czech Republic, and Slovakia, but it is very rarely detected in Great Britain, Spain and Italy. The other typical Eastern European mutation p.Val326Leu shows a very similar geographical distribution as the p.Trp151* mutation, it was the third most common SLOS-causing variant in the Hungarian patients (3 alleles). p.Ser397Leu and p.Tyr432Cys mutations were both detected on 2-2 alleles (8.3-8.3%). All other (previously described) *DHCR7* mutations were detected on one allele: p.Ala247Val, p.Gly244Arg, p.Leu109Pro, p.Arg443His, and p.Arg242His.

4.3.2.3 A novel missense mutation in the DHCR7 gene

In addition to the known nonsense, missense and splicing mutations, 1 previously unidentified mutation, c.374A>G (p.Tyr125Cys) was also detected. The pathogenicity of this mutation is supported by the following facts: 1) the affected residue is phylogenetically highly conserved; 2) the Grantham difference is 194 and 3) the affected residue does not tolerate the tyrosine to cysteine missense mutation according to the SIFT prediction. In addition, this was the only candidate mutation detected in *trans* with another known SLOS-causing mutation, c.976G>T

(p.Val326Leu) in the affected patient. The closest amino acid positions that have been subject to pathogenic mutations leading to SLOS are 119 (p.His119Leu) and 138 (p.Gly138Val). RFLP analysis was performed to exclude the possibility of c.374A>G (p.Tyr125Cys) being a benign variant in the Hungarian population. The mutation was not present in 50 healthy individuals representing 100 alleles.

4.3.2.4 Tiered molecular genetic analysis of DHCR7 gene

In patient 13 a 3-tier molecular genetic analysis was performed. The first tier analysis that involved sequencing of the coding exons revealed 1 previously described pathogenic mutation in a heterozygous form, namely c.452G>A (p.Trp151*). In the second tier molecular testing sequence analysis of the non-coding exons, the putative promoter and part of the 3' untranslated region of the gene was performed. This analysis did not reveal any other potentially causative genetic alteration. Next, cDNA sequencing was performed leading to the following conclusions: first, no other mutation that affects the coding sequence could be detected. Second, analysis at the RNA level confirmed the presence of nonsense-mediated mRNA decay (NMD) as the mRNA derived from blood leukocytes lacked the p.Trp151* allele (NMD for this nonsense mutation has been shown before analyzing skin fibroblasts- derived mRNA).

5. Discussion

5.1 Interfering effect of maternal cell contamination on invasive prenatal molecular genetic testing

In this study, the sensitivity of 3 molecular genetic methods to MCC was determined: Sanger DNA sequencing, MLPA, and pyrosequencing.

For Sanger DNA sequencing, mixing experiments of 6 different small-scale mutations showed variable LOD of the mutant allele in the range of 2.5% to 15% corresponding to 5% to 30% simulated MCC. Based on our data at and above this MCC level, the "fetal" genotype may become equivocal and lead to diagnostic uncertainty or even a false fetal genotype when in practice.

Our results are in accordance with previous publications evaluating analytical sensitivity of Sanger sequencing. For most of the mutations tested, the LOD proved to be in the range of 15% to 20% mutant allele showing insensitivity of Sanger sequencing to detect low-grade mosaicism.

Depending on the amount of maternal DNA present in the fetal sample, lack of quantitation and the abovementioned variability of sensitivity may render interpretation of prenatal results difficult.

In summary, although our experiments and available data in the literature suggest low sensitivity of Sanger sequencing to MCC for most of the mutations, it has to be emphasized that significancy level may depend on many factors such as type of the mutation, sequence enviroment, baseline variability, mode of inheritance, and quality of the DNA sample as well, not to mention PCR bias; thus, judgement on a case to case basis and bidirectional sequencing are of great importance.

In the first experiment determining the sensitivity of MLPA method to MCC, a wild type female DNA was contaminated with a female DNA heterozygous for deletion of exons 1 to 3 in the *MID1* gene. Up to 30% simulated MCC (corresponding to 15% mutant allele), there was no effect on genotyping. Fourty percent MCC (corresponding to 20% mutant allele) resulted in 30% maximum reduction of dosage quotient, which may already cause diagnostic uncertainty.

In the second MLPA experiment, a wild type DNA sample was mixed with a DNA sample heterozygous for duplication of exons 33 to 35 in the *PKHD1* gene. Even 40% simulated MCC (corresponding to 20% mutant allele) had no impact on genotyping, dosage quotients of probes hybridizing to the duplicated region were still in the normal range.

Van Veghel-Plandsoen et al investigated MLPA's sensitivity to detect low-grade mosaicism; according to the authors, a multiexon duplication and a whole gene deletion was detectable at about 20% and 15% mutant allele fractions, respectively. Similarly to our results, MLPA was less sensitive to detect duplications than deletions. Although these data demonstrate insensitivity of MLPA method to maternal DNA contamination, accurate fetal genotyping necessitates simultaneous MCC level determination, and the maternal genotype has to be taken into consideration as well if MCC is present.

In the third group of experiments, we investigated the sensitivity of pyrosequencing to MCC. Our experiments demonstrated that as low as 0.5% mutant allele was detectable at high coverage; however, at this very low level, the mutant allele was underestimated. Mutant allele fractions were accurately (for the purpose of invasive prenatal molecular diagnosis) quantified by pyrosequencing in the range of 2.5% to 20%; therefore, if MCC analysis is performed on the same fetal sample and the maternal genotype is known, correct interpretation of diagnostic testing data may still be possible in the presence of 1% to 40% MCC.

Sensitivity of different pyrosequencing assays developed for *KRAS* mutation testing has been shown before to be about 5%. According to Rohlin et al, pyrosequencing can achieve even higher analytical sensitivity, 1% mutant allele fraction is detectable at very high sequence coverage. Jamuar et al investigated the capability of another massively parallel sequencing method for the detection of mosaic mutations and found that 5% mutant allele fraction is readily detectable at a mean read depth of 200x.

Massively parallel sequencing technology has the ability to quantitate mutant and wild type allele fractions within a sample. Therefore, although these techniques are sensitive to MCC at high coverage if MCC analysis is performed on the same fetal sample and the maternal genotype is known, precise determination of allelic ratios might enable correct interpretation of diagnostic testing data even in the presence of maternal DNA.

Based on recent knowledge, even if prenatal samples are not purely of fetal origin, correct prenatal molecular diagnosis and reporting are possible and repeated sampling can be avoided in many of the cases.

5.2 A novel missense mutation in the *DMD* gene

BMD/DMD was suspected in the case of a 14-month-old male patient because of markedly elevated creatine kinase level, mild hypotonia, and gross motor delay. For the molecular genetic confirmation of dystrophinopathy a tiered genetic testing of the *DMD* gene was performed starting with MLPA analysis showing deletion of exon 4. Single-exon deletions detected by MLPA always need to be confirmed by another method. Successful PCR amplification of exon 4 proved that our MLPA result must have been false positive. Sanger sequencing of exon 4 showed A-to-T transversion at the 227th nucleotide position lying directly at the ligation site of the 5' MLPA probe, which most likely led to unsuccessful ligation and/or hybridization. Sequencing of the entire coding region of *DMD* gene revealed no other potentially pathogenic variant except for the novel point mutation c.227A>T, previously detected by Sanger sequencing.

The c.227A>T variant was classified as likely pathogenic according to recently published ACMG variant classification criteria upon the following: i) the detected variant is *de novo* (the mutation was absent in the mother of the patient and there was no family history of BMD/DMD), ii) the detected variant is absent from individuals in the 1000 Genomes Project and Exome Aggregation Consortium (ExAC), iii) the affected amino acid residue is evolutionary conserved between species, in addition to it sequence and secondary structure comparison of other N-ABD containing human proteins also showed conservation of Asn76;

and iv) the patient's phenotype is highly specific for BMD/DMD (both markedly elevated CK levels and the clinical picture suggested dystrophinopathy). In order to further (indirectly) support pathogenicity of the detected variant and discriminate between BMD and DMD, muscle biopsy was performed.

Light microscopic examination revealed mild myopathic changes with regenerative activity. The expression pattern of dystrophin was pathological, with severely reduced and segmentally absent staining with one of the anti-dystrophin antibodies (dystrophin-1, targeting the core region of the protein), and with absent or severally reduced staining with dystrophin-3 antibody (targeting the N-terminal end of the protein). These findings are consistent with the *DMD* gene mutation detected by the mutation analysis and are suggestive of BMD.

Although missense mutations are rarely detected in DMD/BMD patients, about 50% of them occurs in the N-ABD of dystrophin protein. We studied the dystrophin protein at the level of the N-ABD and used *in silico* methods to predict the effects of the detected point mutation on protein structure and stability, using the sequence and crystal structure of human dystrophin N-ABD.

Not only sequence alignment but structural comparison of some CH1 subdomain-containing proteins also showed the conservation of Asn residue in this position (Asn76 according to dystrophin numbering).

Substitution of Asn76 residue to Ile showed no impact on the secondary structure of helix E. Stability analysis showed highly destabilizing nature of p.Asn76Ile non-synonymous mutation. As p.Asn76Ile mutation has no effect on the secondary structure of ABD based on our predictions, it probably causes local conformational changes due to the disruption of stabilizing interactions within the CH1 subdomain. The mutation may alter the flexibility of loop B (due to the loss of Asn76-Asp46 H-bond between helix E and loop B) and spatial positions of loop D residues (due to the loss of Asn76-Glu65 H-bond between helix E and loop D), as well. Conformation of helix C could be also affected by the p.Asn76Ile mutation, because there is no direct hydrogen bond between helix E and helix C, only Asp46 located next to the N-terminal end of helix C interacts with Asn76 residue.

Neither the wild type nor the p.Asn76Ile mutant ABD showed increased aggregation propensity of CH1 subdomain.

In silico analyses of the protein structure support the mutation pathogenicity but cannot provide information about disease severity. In this case, based on dystrophin immunohistochemistry results, we think that p.Asn76Ile mutation most likely leads to BMD, but the phenotype clearly needs to be reevaluated over time on a clinical basis.

As the mutation is a *de novo* point mutation for counseling purposes 2% recurrence risk has to be taken into account.

From the molecular genetic analysis point of view, our results highlight the importance of confirmatory testing of single-exon deletion MLPA results.

5.3 Mutational spectrum of Hungarian Smith-Lemli-Opitz syndrome patients

Hereditary disorders of cholesterol biosynthesis are characterized by cholesterol deficiency and presumably it is the accumulation of different sterol precursors that lead to the distinct phenotypic presentation of different syndromes. The first tier laboratory test for the confirmation of clinical diagnosis/suspicion of SLOS is measurement of serum/plasma 7-DHC concentration.

Extremely elevated serum 7-DHC levels were measured in all patients. Cholesterol concentration was low in about 2/3 of patients when compared to their age-specific reference range.

Molecular genetic analysis of *DHCR7* gene led to the idetification of both mutations in 11 unrelated patients, in one case only one mutation was detected in heterozygous form (23/24 alleles, 95.8%). This is in good agreement with available data in the literature about 95-96% mutation detection rate when sequencing the coding region and exon-intron boundaries of *DHCR7* gene.

In spite of a tiered genetic approach involving analysis of the non-coding exons, the putative promoter and 3' UTR in the case of patient 13 only one previously published SLOS-causing mutation was identified in heterozygous form (c.452G>A, p.Trp151*). mRNA (isolated from blood leukocytes) sequencing revealed NMD for p.Trp151* mutation, which phenomenon has been shown before analyzing skin fibroblasts-derived mRNA. mRNA sequencing did not reveal any other potentially causative genetic alteration. In this case diagnosis of SLOS could still be unequivocally established based on the clinical symptoms and elevated serum 7-DHC concentration.

DHCR7 mutations detected in our patients were clustered in exons 5-7 and intron 8-exon 9, which distribution is in good agreement with literature data. About half of the mutations (56.5%) were missense, which is lower than the published frequency of missense mutations in SLOS patients (~90%). The difference between our and previously reported data might be explained by the low number of patients investigated in our study leading to an estimation bias. About two third of missense mutations (69%) localized to transmembrane (TM) domains (p.Leu109Pro, p.Arg242His, p.Gly244Arg, p.Ala247Val, p.Val326Leu, and p.Tyr432Cys)

similarly to previously published results. Interestingly, 2/3 of TM mutations (p.Arg242His, p.Gly244Arg, p.Ala247Val, p.Val326Leu) affect the sterol sensing domain of the protein. p.Ser397Leu detected on 2 alleles affects the highly conserved sterol reductase „signature" motif (4th cytoplasmic loop), while p.Arg443His is localized to the C-terminus and affects sterol reductase domain 2 and presumably leads to decreased protein stability. In addition to the missense mutations, nonsense and splicing mutations were detected in 26.1% and 17.4% respectively, all of the them resulting in null alleles.

Genotype distrubution was in good agreement with available literature data. In all patients but one (9.1%, homozygous for c.964–1G>C) compound heterozygosity was detected: either in the form of null allele/missense mutation in 63.6% or two diferent missense mutations in 27.3% of the patients.

DHCR7 gene mutations show an interesting geographical distribution across Europe, which phenomenon has been observed for other monogenic disorders as well. A null mutation, p.Trp151* was the most common in the Hungarian SLOS patients. p.Trp151* mutation probably arose in Southern Poland and shows an East to West gradient across Northern Europe. Another null mutation, c.964–1G>C disrupts the splice acceptor site in intron 8, was the second most common in our patient cohort. c.964–1G>C is the most frequently reported SLOS-causing mutation, which presumably arose in the British Isles and shows a West to East gradient across Northern Europe. Another typical „East European mutation", p.Val326Leu, was found to be the third most frequent in the Hungarian patients and together with p.Trp151* and c.964–1G>C mutations accounted for 54% of SLOS alleles in Hungary. Haplotype analysis showed that probably all three mutations are founder mutations. SLOS alleles with a frequency of ≤8,3% in the Hungarian patients are rare in other European populations too.

A novel missense mutation was also detected: c.374A>G (p.Tyr125Cys). According to SIFT's prediction a tyrosine to cysteine substitution is not tolerated in this position, sequence alignment of DHCR7 orthologs showed conservation of Tyr125 residue, the Grantham score was 194, i.e. physico-chemical characteristics of the two amino acids are greatly different. We could not detect c.374A>G in 50 healthy control samples (100 alleles), which means, that this genetic alteration is most likely not a Hungarian benign variant, in addition, it is not present in population databases such as ExAC, or 1000 Genomes Project. Pathogenicity of the novel variant is further supported by the fact that this was the only candidate causative mutation detected in *trans* with a previously reported pathogenic *DHCR7* mutation. If missense mutations are common mechanism of the disease (which is true for SLOS) this is a supporting evidence too. The patient's phenotype, being highly specific for SLOS, is another supporting

evidence. Taken together c.374A>G (p.Tyr125Cys) variant can be classified as „likely pathogenic".

Between 2008-2018 prenatal molecular analysis of CVS or AF samples was performed in 10 families affected with SLOS. In two cases the fetuses were affected with homozygous or compound heterozygous genotypes, in one case carrier status (the child has not been born yet) was confirmed and 7 healthy children were born (in 4 cases carrier state, while in 3 others homozygous normal genotype was detected).

These three studies demonstrate that molecular genetic diagnostics and research are closely connected and inseparable, and that collaboration and consultation with clinicians and other diagnostic specialists is very important, as well as carefully taking into account each element of the diagnostic process when interpreting molecular genetic results.

6. Summary

In our work we studied some of the special and challenging aspects of prenatal and postnatal molecular genetic diagnostics of severe monogenic disorders.

In our first study we investigated the sensitivity of three molecular genetic methods to MCC. Sensitivity of Sanger DNA sequencing proved to be variable, significant MCC was in the range of 5-30% depending on the type of mutation. In the case of MLPA method only high level (≥40%) MCC interfered with 'fetal' genotyping. Pyrosequencing was shown to be very sensitive to MCC detecting as low as 1%.

In the second study we performed tiered molecular genetic analysis of *DMD* gene in the case of a 14-month-old male patient. The initial screen for deletions/duplications with MLPA method showed deletion of exon 4, which could not be confirmed by another method. A novel point mutation (c.227A>T, p.Asn76Ile) affecting the N-ABD of dystrophin protein was detected by Sanger DNA sequencing. No other potentially pathogenic mutation was detected by sequencing the whole coding region of *DMD* gene. Based on *in silico* analyses the detected missense mutation was predicted to be highly destabilizing on the N-ABD. Based on dystrophin immunohistochemistry studies the detected genetic alteration most likely leads to BMD.

In our third study we summarized molecular genetic, biochemical and clinical data of 13 Hungarian Smith-Lemli-Opitz syndrome patients. 7-DHC concentration was extremely elevated in the patients' sera. The mutation detection rate was 95.8%, in one case we detected only one mutation in heterozygous form in spite of tiered molecular genetic analysis. In

Hungarian SLOS patients 3 variants (p.Trp151*, c.964–1G>C and p.Val326Leu) accounted for 54% of mutations. One patient was homozygous for a null mutation, while all other patients were compound heterozygotes. One novel missense mutation, c.374A>G (p.Tyr125Cys) was also identified, which was classified as „likely pathogenic".

7. Main new scientific findings and their clinical significance

1) We determined significancy level of maternal cell contamination for three molecular genetic methods in a comprehensive way applying mixing experiments simulating different MCC levels (1, 5, 10, 20, 30, and 40%). Sensitivity of Sanger DNA sequencing proved to be variable depending on the mutation type, significant MCC levels were in the range of 5-30%. For MLPA method interference was shown only if MCC was ≥40%. Pyrosequencing was very sensitive detecting as low as 1% MCC. We also compared our results to previously published data about the analytical sensitivity of these methods evaluating them in the context of MCC and prenatal molecular testing.

2) Missense mutations are very rare cause of dystrophinopathies and interestingly, about half of them localize to the N-ABD of dystrophin protein. We detected a novel missense mutation in the *DMD* gene affecting N-ABD, which, based on dystrophin immunohistochemistry analysis, possibly leads to BMD (c.227A>T, p.Asn76Ile, LOVD: https://databases.lovd.nl/shared/variants/0000171133) and is highly destabilizing according to *in silico* analyses.

3) We summarized the mutational spectrum, biochemical and clinical features of 13 Hungarian SLOS patients diagnosed in our laboratory between 2008-2012 broadening the reported European mutational and phenotypic data. The detected mutations and genotypes were submitted to the gene/disease-specific database (https://databases.lovd.nl/shared/genes/DHCR7). In addition a novel, likely pathogenic missense mutation was also identified (c.374A>G (p.Tyr125Cys). In Hungary three *DHCR7* gene mutations (p.Trp151*, c.964–1G>C and p.Val326Leu) comprise about half of the SLOS alleles. Based on our results peripheral blood leukocytes-derived *DHCR7* mRNA is suitable for diagnostic purposes.

8. Funding

This work was supported by the Ministry of National Economy, Hungary, Grant/Award Number: GINOP-2.3.2-15-2016-00039.

9. List of publications

UNIVERSITY of DEBRECEN

UNIVERSITY AND NATIONAL LIBRAY
UNIVERSITY OF DEBRECEN
H-4002 Egyetem tér 1, Debrecen
Phone: +3652/410-443, email: publikaciok@lib.unideb.hu

Registry number: DEENK/11/2019.PL
Subject: PhD Publikációs Lista

Candidate: Katalin Koczok
Neptun ID: GKQM6L
Doctoral School: Doctoral School of Molecular Cellular and Immune Biology
MTMT ID: 10037481

List of publications related to the dissertation

1. **Koczok, K.**, Merő, G., P. Szabó, G., Madar, L., Gombos, É., Ajzner, É., Mótyán, J. A., Hortobágyi, T., Balogh, I.: A novel point mutation affecting Asn76 of dystrophin protein leads to dystrophinopathy.
   Neuromusc. Disord. 28 (2), 129-136, 2018.
   IF: 2.487 (2017)

2. **Koczok, K.**, Gombos, É., Madar, L., Török, O., Balogh, I.: Interfering effect of maternal cell contamination on invasive prenatal molecular genetic testing.
   Prenat. Diagn. 38 (9), 713-719, 2018.
   DOI: http://dx.doi.org/10.1002/pd.5319
   IF: 2.779 (2017)

3. Balogh, I., **Koczok, K.**, P. Szabó, G., Török, O., Hadzsiev, K., Csábi, G., Balogh, L., Dzsudzsák, E., Ajzner, É., Szabó, L., Csákváry, V., Oláh, A.: Mutational spectrum of Smith-Lemli-Opitz syndrome patients in Hungary.
   Mol Syndromol. 3 (5), 215-222, 2012.
   DOI: http://dx.doi.org/10.1159/000343923

DEBRECENI EGYETEM Egyetemi és Nemzeti Könyvtár

UNIVERSITY AND NATIONAL LIBRAY
UNIVERSITY OF DEBRECEN
H-4002 Egyetem tér 1, Debrecen
Phone: +3652/410-443, email: publikaciok@lib.unideb.hu

List of other publications

4. **Koczok, K.**, Gurumurthy, C. B., Balogh, I., Korade, Z., Mirnics, K.: Subcellular localization of sterol biosynthesis enzymes.
J. Mol. Histol. 50 (1), 63-73, 2019.
DOI: http://dx.doi.org/10.1007/s10735-018-9807-y
IF: 2.412 (2017)

5. Ivády, G., Madar, L., Dzsudzsák, E., **Koczok, K.**, Kappelmayer, J., Krulisova, V., Macek, J. M., Horváth, A., Balogh, I.: Analytical parameters and validation of homopolymer detection in a pyrosequencing-based next generation sequencing system.
BMC Genomics. 19, 1-8, 2018.
DOI: http://dx.doi.org/10.1186/s12864-018-4544-x
IF: 3.73 (2017)

6. Korade, Z., Genaro-Mattos, T. C., Tallman, K. A., Liu, W., Garbett, K. A., **Koczok, K.**, Balogh, I., Mirnics, K., Porter, N. A.: Vulnerability of DHCR7+/- mutation carriers to aripiprazole and trazodone exposure.
J. Lipid Res. 58 (11), 2139-2146, 2017.
DOI: http://dx.doi.org/10.1016/j.nmd.2017.12.003
IF: 4.505

7. Korade, Z., Kim, H. Y. H., Tallman, K. A., Liu, W., **Koczok, K.**, Balogh, I., Xu, L., Mirnics, K., Porter, N. A.: The Effect of Small Molecules on Sterol Homeostasis: Measuring 7-Dehydrocholesterol in Dhcr7-Deficient Neuro2a Cells and Human Fibroblasts.
J. Med. Chem. 59 (3), 1102-1115, 2016.
DOI: http://dx.doi.org/10.1021/acs.jmedchem.5b01696
IF: 6.259

8. **Koczok, K.**, Oláh, A., Szabó, G., Oláh, É., Török, O., Balogh, I.: A koleszterin-bioszintézis veleszületett zavara: a Smith-Lemli-Opitz-szindróma.
Orv. hetil. 156 (42), 1695-1702, 2015.
DOI: http://dx.doi.org/10.1556/650.2015.30256
IF: 0.291

9. Oláh, A., **Koczok, K.**, Szabó, G., Varga, J., Balogh, I., Harangi, J.: A veleszületett koleszterinhiány jellemző biomarkerei Smith-Lemli-Opitz-szindrómás betegekben és hordozókban.
Metabolizmus. 13 (3), 196-201, 2015.

10. Ivády, G., **Koczok, K.**, Madar, L., Gombos, É., Tóth, I., Győri, K., Balogh, I.: Molecular Analysis of Cystic Fibrosis Patients in Hungary - an Update to the Mutational Spectrum.
J. Med. Biochem. 34, 1-6, 2015.
IF: 0.742

UNIVERSITY AND NATIONAL LIBRAY
UNIVERSITY OF DEBRECEN
H-4002 Egyetem tér 1, Debrecen
Phone: +3652/410-443, email: publikaciok@lib.unideb.hu

11. Bubán, T., **Koczok, K.**, Földesi, R., Szabó, G., Sümegi, A., Tanyi, M., Szerafin, L., Udvardy, M., Kappelmayer, J., Antal-Szalmás, P.: Detection of internal tandem duplications in the FLT3 gene by different electrophoretic methods.
Clin. Chem. Lab. Med. 50 (2), 301-310, 2011.
DOI: http://dx.doi.org/10.1515/CCLM.2011.762
IF: 2.15

12. Szodoray, P., **Koczok, K.**, Szántó, A., Horváth, I. F., Nakken, B., Molnár, I., Zeher, M.: Autoantibodies to novel membrane and cytosolic antigens of the lachrymal gland in primary Sjögren's syndrome.
Clin. Rheumatol. 27 (2), 195-199, 2008.
DOI: http://dx.doi.org/10.1007/s10067-007-0678-y
IF: 1.559

Total IF of journals (all publications): 26,914
Total IF of journals (publications related to the dissertation): 5,266

The Candidate's publication data submitted to the iDEa Tudóstér have been validated by DEENK on the basis of Web of Science, Scopus and Journal Citation Report (Impact Factor) databases.

17 January, 2019