

The role of next-generation sequencing in the diagnosis of lysosomal storage disorders

Katalin Komlosi^{1,2}, Alexander Sólyom³, Michael Beck²

¹Institute of Medical Genetics, University of Pécs, Pécs, Hungary

²Institute of Human Genetics, University Medical Center, Johannes Gutenberg University
Mainz, Mainz, Germany

³Clinical Research, Plexcera Therapeutics LLC, New York, NY, USA

Corresponding author:

Katalin Komlosi, MD, PhD

Institute of Medical Genetics

University of Pécs

Szigeti street 12.

H-7624 Pécs

Hungary

Phone: +36 (72) 535-978

Fax: +36 (72) 535-978

e-mail: komlosi.katalin@pte.hu

Key words: NGS, LSD, target enrichment, coverage, clinical utility

Abbreviations:

LSD-lysosomal storage disorders

NGS-next-generation sequencing

WES-whole exome sequencing

WGS-whole genome sequencing

Abstract

Next-generation sequencing (NGS) panels are used widely in clinical diagnostics to identify genetic causes of various monogenic disease groups including neurometabolic disorders and, more recently, lysosomal storage disorders. Many new challenges have been introduced through these new technologies, both at the laboratory level and at the bioinformatics level, with consequences including new requirements for interpretation of results, and for genetic counseling. We review some recent examples of the application of NGS technologies, with purely diagnostic and with both diagnostic and research aims, for establishing a rapid genetic diagnosis in lysosomal storage disorders. Given that NGS can be applied in a way that takes into account the many issues raised by international consensus guidelines, it can have a significant role even early in the course of the diagnostic process, in combination with biochemical and clinical data. Besides decreasing the delay in diagnosis for many patients, a precise molecular diagnosis is extremely important as new therapies are becoming available within the LSD disease spectrum for patients who share specific types of mutations. A genetic diagnosis is also the prerequisite for genetic counseling, family planning and the individual choice of reproductive options in affected families.

Introduction

Next-generation sequencing, also referred to as massively parallel sequencing, is a high-throughput DNA sequencing technology that enables the fast generation of data on thousands to millions of base pairs of DNA from an individual patient by sequencing large numbers of genes in a single reaction.¹ The rapid emergence and huge success of these technologies in the research field paved the way for a new era in genetic diagnostics. Next-generation sequencing (NGS) panels are used widely in clinical diagnostics to identify genetic causes of various monogenic disease groups, such as epilepsy,² intellectual disability,^{3,4} neurodevelopmental disorders,⁵ and neurometabolic disorders,⁶ among others. However, many new challenges have been introduced through these new methods, both at the technical level and at the bioinformatic level, with consequences including new requirements for interpretation of results, and for genetic counseling. In order to define the precise role of NGS in diagnostics, consensus guidelines have been proposed which mostly deal with NGS testing for rare and monogenic diseases.⁷

Lysosomal storage diseases (LSDs) are a group of rare diseases in which a genetic mutation leads to the accumulation of specific macromolecules which impair the function of the

lysosome often causing secondary disturbances in other cellular and metabolic processes. With a combined incidence of ~1 in 1500 to 7000 live births, they encompass >50 different entities.⁸ Unfortunately, there is considerable clinical variability within each disease phenotype, and also overlapping symptomatology among LSDs which hampers the precise diagnosis of these disorders.⁹ This makes a diagnosis solely based on the clinical presentation difficult in many cases.

In the past, many methods have been used to approach a diagnosis in the LSDs.¹⁰ Laboratory assays based on identification or quantification of the storage products have been one of the most favored methods. Tests for elevated secreted substrate material are routinely used to screen for glycosaminoglycans and oligosaccharides in patients suspected of having mucopolysaccharidoses and disorders that present with oligosacchariduria. The urine screens are usually very sensitive, however, affected individuals with normal urine screens and healthy individuals with elevated excretions have been reported.¹¹ The second main pillar of establishing a diagnosis in LSDs are enzyme activity measurements. Enzyme activity detected in dried blood spots is very useful in the diagnosis of a small number of LSDs, however, this method usually requires verification with a second type of assay. Measurement of enzyme activity in leukocytes, plasma or fibroblasts is often a logistic and technical problem, with particular storage and transport requirements as well as laboratory expertise. Enzyme activities are usually not reliable in detecting heterozygous carriers of a disease. While biochemical methods are still the gold-standard of diagnosis, they are often very laborious, time-consuming and require a pre-selection by clinical phenotype to reduce the number of biochemical tests used for each patient suspected of LSDs.¹² Diagnosis of LSDs is still a clinical challenge and can often take several years.

Even having achieved a diagnosis with traditional assays, establishing a molecular genetic diagnosis is still important for several reasons. The precise molecular diagnosis is of great importance in the LSDs as new therapies are becoming available for specific types of LSDs, and for patients who share specific types of mutations.^{9,13} The complexity underlying the pathogenesis of LSDs and the small population of patients, who are many times children, make the development of therapies for these diseases challenging. Current treatments are only available for a small subset of LSDs and have, for the most part, not been effective at treating or preventing neurological symptoms. Clinical trials have demonstrated some clinical benefit of enzyme replacement therapy in Gaucher disease, Fabry disease, mucopolysaccharidoses type I, II, IV, VI, VII and in Pompe disease. However, the usefulness of enzyme replacement therapy is limited due to the fact that a given enzyme preparation does not have beneficial

effects on all aspects of a disorder to the same degree. The effectiveness of each therapy is determined by whether the symptoms it affects are those which are most important to the patients' overall health and quality of life. Additionally, clinical studies have shown that the pathologic changes underlying many symptoms of a lysosomal storage disorder are not reversible, even with long-term treatment.⁹ Thus, a precise molecular diagnosis early in the course of the disease can be of great significance. A definitive diagnosis is also the prerequisite for patients to have access to studies of newly developed therapeutic strategies, such as small molecules for substrate reduction or for enzyme enhancing therapy.^{9,13}

The application of NGS in the diagnosis of LSDs: overview of the literature

Studies of purely diagnostic targeted NGS panels

As outlined above, there are many factors hampering the diagnosis of LSDs, among others the phenotypic and penetrance variability, common signs and symptoms between certain disease groups, the genetic heterogeneity in some forms and the difficulties of biochemical diagnostics. The application of a broad diagnostic tool could enhance the sometimes very long diagnostic process for the affected families and thereby grant quicker access to current and investigated therapeutic options while providing the basis for more appropriate genetic counseling. There is only scarce published data on the application of next-generation sequencing in the diagnostics of lysosomal storage disorders.

The recent study by Fernandez-Marmiesse et al¹² reports for the first time the results of a pilot project to evaluate the application of NGS to mutation screening in the diagnosis of lysosomal storage disorders. With the help of a next-generation sequencing-based approach for the simultaneous analysis of 57 lysosomal genes they were able to correctly diagnose 18 positive blinded controls and provide genetic diagnosis to 25 patients with a high to moderate clinical suspicion index for LSDs from Spain and Portugal. In their study they analyzed 84 probands, including 18 positive controls in whom biochemical testing and Sanger sequencing had already ascertained a diagnosis of LSD and 66 patients with a suspected LSD. They used a custom Sure Select oligonucleotide probe library to capture all coding exons and exon-intron boundaries of 57 genes known to be associated with LSDs according to GeneReviews.¹² Sequencing was performed on SOLiD4 and HiSeq2000 Illumina platforms, with the latter yielding better coverage (99,97% of bases covered by at least 20 reads and only 0,03% of target bases with a coverage of less than 20 reads). As to the technical difficulties encountered, the authors emphasize that two factors limit enrichment. The local sequence

architecture has a strong effect on the efficiency of DNA enrichment for individual exons,¹⁴ so that exons close to repetitive regions are not fully covered. For exons located in CpG islands (high frequency of CpG sites: regions of DNA where a cytosine nucleotide is followed by a guanine nucleotide in the linear sequence of bases along its 5'→3' direction), coverage also decreased dramatically, producing gaps in coverage of certain exons of LSD genes, mostly in *IDUA*, *GBA* and *GAA* genes. The problem was overcome, however, with a greater overall coverage achieved on the HiSeq platform. To establish an NGS-LSD assay as a diagnostic tool for routine use it is essential to eliminate sequence gaps in order not to miss mutations in certain exons.¹²

The authors also point out the difficulties encountered with data filtering and interpretation in the NGS assay. One of the main challenges of using NGS for diagnostic applications is the interpretation of a massive number of genomic variants detected by sequencing platforms. Coverage irregularities contributed to false-positive variants. Thus, by improving enrichment efficacy or increased coverage, false-positive variants can be largely avoided. Efficient and reliable identification of causative variants is crucial to the implementation of this technology in routine diagnostics. The authors conclude that they would not propose their assay as the sole diagnostic tool, but as a useful adjunct to diagnosis for specialists in the clinical management of LSD patients especially given its advantage to provide accurate information in a short time.¹²

Another example for the clinical utility of gene panels in the diagnosis of lysosomal storage disorders is the diagnosis of late-onset Pompe disease by a next-generation sequencing approach published by Lévesque et al. in 2016.¹⁵ The authors developed a gene panel to analyze the coding sequences and splice site junctions of the *GAA* gene along with 77 other genes causing muscle disorders with overlapping phenotypes. They achieved a median coverage of 200X with all *GAA* exons successfully covered with >20X and only 0,3% of exons across all genes in the panel were <20X. Using known variants in Pompe patients and controls the authors could demonstrate an excellent sensitivity (100%) and specificity (98%) across all selected genes. Analyzing 34 patients with suspected muscle disorders of unknown etiology the study could show a detection rate of 32%. Especially in patients with an atypical presentation, including one late-onset Pompe disease case, the gene panel was instrumental in reaching a diagnosis. In all diagnosed patients biochemical testing (acid alpha-glucosidase activity) confirmed the molecular results. The NGS panel was shown to facilitate the diagnosis in patients showing non-specific muscle weakness or atypical phenotypes.¹⁵

Diagnostic and research NGS panels to detect LSDs

According to the „Guidelines for diagnostic next-generation sequencing” on behalf of the EuroGentest and the European Society of Human Genetics⁷: „for diagnostic purpose, only genes with a known (ie, published and confirmed) relationship between the aberrant genotype and the pathology should be included in the analysis” of NGS assays. However, several applications with a research focus have ventured to apply NGS technologies to identify sequence variations in known, but also in potentially new, candidate genes for lysosomal storage disorders. Among the first publications of an extensive research panel for the identification of LSDs was the study by Di Fruscio et al. in 2015 aiming to detect DNA sequence variation in the autophagy-lysosomal pathway.¹⁶ The autophagy-lysosomal pathway regulates cell homeostasis and plays an important role in lysosomal storage disorders and common neurodegenerative diseases.¹⁷ The authors developed Lysoplex, a targeted NGS approach enabling the parallel analysis of 891 genes involved in lysosomal (194 genes), endocytic (627 genes) and autophagic (106 genes) pathways. Lysoplex was successfully validated on 14 different types of LSDs and used to analyze 48 patients with unknown mutations and a clinical phenotype of neuronal ceroid lipofuscinosis. The average coverage depths for 95% of the target sequences was 40X and for 80% of the target sequences 100X was achieved. With a training set of 16 genomic DNAs from patients affected by 14 different forms of LSDs (including MPSI, II, IIB, VI, Fabry disease, NPC1, MLD, Pompe disease among others), whose molecular diagnosis was already known except for 2 cases, Di Fruscio and coworkers were able to detect all of the previously known mutations and to identify disease-causing mutations in the two previously undiagnosed LSD patients.¹⁶

The authors also demonstrate the efficacy of Lysoplex in the analysis of patients with unknown molecular defect with a clinical diagnosis of neuronal ceroid lipofuscinosis (NCL). The NCLs are a family of autosomal recessive neurodegenerative disorders that annually affect 1:100,000 live births worldwide. This family of diseases results from mutations in one of 14 different genes that share common clinical and pathological etiologies. Though the disease phenotypes may vary in their age and order of presentation, all typically include progressive visual deterioration and blindness, cognitive impairment, motor deficits and seizures. Advances have been made in genetic diagnosis and counseling for families. However, comprehensive treatment programs that delay or halt disease progression have been elusive. Current disease management is primarily targeted at controlling the symptoms rather than "curing" the disease.¹⁸ A definite molecular diagnosis in such patients is the prerequisite for the enrollment in ongoing and future therapeutic trials. Di Fruscio and coworkers analyzed

48 individuals with the clinical diagnosis of NCL and detected in 29/48 patients homozygous or compound heterozygous variants in known NCL genes. In 3 additional cases only a single heterozygous variant in known NCL genes was found, postulating that the second variant was not detectable because of a deep intronic or a large copy number variation (overall detection rate of 67% for causative mutations in NCL genes). In addition, the research-based NGS approach allowed the identification of recessive mutations in three potential novel NCL candidate genes. However, further functional analyses and in vivo studies in animal models are needed to confirm the causative nature of those variants.¹⁶

Clinical exome sequencing to detect LSDs

In the last few years whole exome sequencing (WES) or whole genome sequencing (WGS) approaches have gained unlimited consideration as universal tests for the identification of most Mendelian disorders^{19,20,21} with the exceptions of those caused by complex structural variations. As a more focused alternative to WGS, WES has been demonstrated to be a cost-effective solution with a higher throughput on coding sequences that favors the identification of novel disease genes.

A recent publication demonstrates the complexity and clinical variability of LSDs, as well as the potential role of WES in exposing gaps in clinical knowledge due to the rarity of many of these diseases. Three siblings, between 40 and 60 years of age, developed severe osteolysis of bones in the hands and feet. Known skeletal dysplasia genes were sequenced, but no mutation found. WES eventually revealed deleterious mutations in the *ASAHI* gene, causing acid ceramidase deficiency or Farber disease, which was confirmed by enzymatic testing.²² The patients did have „typical” symptoms of the disease in childhood, but they were so mild that an LSD was not suspected. These became the oldest Farber disease patients ever diagnosed (by 40 years!). In addition, the *ASAHI* gene can now be added to those tested in patients with osteolysis.

Discussion

Recent reports of the application of NGS technologies in the diagnosis of LSDs highlight the high clinical utility of gene panels in patients with clinical suspicion of lysosomal storage disorders and the potential to reduce diagnostic delay in a group of patients where a definitive diagnosis can be delayed for years. With decreasing sequencing costs and increasing availability of tests fulfilling guidelines for diagnostic NGS testing, there are several advantages of using gene panels early in the investigation of patients with LSDs.

However, there are several issues raised by the application of high-throughput sequencing that needs to be considered and harmonized across different laboratories.

Since the technology and applications change constantly and rapidly, the available NGS platforms are not yet stable enough to ensure that all results obtained are equivalent. However, the technical variability cannot hamper the implementation of NGS assays, only the implementation of insufficiently validated and poor quality tests should be prevented.⁷ Before implementing a diagnostic test, the aim of the assay should be clearly stated. It should be clear whether the test may be used to exclude a diagnosis, or to confirm a diagnosis. The 'diagnostic yield' of a test has to be certain before introducing an assay in the routine clinical diagnostic setting. The 'diagnostic yield' refers to the chance that a disease-causing variant is identified and a definitive molecular diagnosis can be made.⁷ The value should always refer to a patient cohort. This defines the performance of NGS from clinical point of view, and indicates the efficiency of the test. Currently, most laboratories offer their own selection of genes in the frame of a diagnostic gene panel. From the standpoint of the patients and medical practitioners, it would be important to harmonize genetic testing. For the diagnosis of LSDs it would make sense to have 'core disease gene lists' established by clinical and laboratory experts of the field. Adding additional genes to standard core lists will increase the diagnostic yield, but it should not compromise the detection rate and the identification of disease-causing mutations. A simple rating system on the basis of coverage and diagnostic yield has already been proposed on behalf of EuroGentest and the ESHG,⁷ with a type A test offering >99% variant calls of the coding region and flanking intron sequences and fills all the gaps with Sanger-sequencing or another complementary analysis. Type B tests would clearly state which regions are sequenced at >99% reliable reference and fills some of the gaps with resequencing. Type C NGS tests only rely on the quality of NGS sequencing without offering additional resequencing. The ordering physician has to be fully informed about the limitations of the NGS assay, but also about the potential unanticipated effects of a particular type of genetic testing. Thus, the laboratory should specify the diseases not relevant to the clinical phenotype of the patient that could be caused by mutations in the tested genes. Especially in a broad panel of LSDs only a part of the analyzed genes will be relevant to the disease phenotype of the patient. However, heterozygous mutations in recessive conditions might be detected, leading to the unsought detection of disease carriers.⁷ This, in turn, can have consequences for counselling and reproductive choices. Laboratories, therefore, should provide information of the chance of unsolicited findings.

A pitfall of using gene panels as first-tier analysis in the place of biochemical assays for identification of LSDs include the identification of variants of unknown significance (VUS) or Unclassified Variants (UVs or class 3 pathogenicity variants).^{7,15} The VUS rates will greatly vary from gene to gene, but large collaborative efforts to share variant data, such as ClinVar, disease-specific databases and other public databases are attempting to minimize this limitation.²³ In patients harbouring a VUS, biochemical testing is essential to attempt to define the functional consequence of the sequence variant, however, this is not possible for pathologies that have no biochemical marker and in which confirmation will need to be obtained by other methods.¹² Another important policy for laboratories is to set up a local variant database to manage disease variants and to collect data on VUS with the aim to eventually classify these variants definitively.⁷

The detection of a single heterozygous mutation in a gene for an autosomal recessive disorder (as in most lysosomal storage disorders) and the limited performance of bioinformatic tools to detect deletions and duplications from targeted NGS data^{15,24} is another limitation encountered in the application of NGS assays in the diagnosis of LSDs. Currently, bioinformatic identification of deletions and duplications in many targeted NGS-based approaches is of limited sensitivity, especially for fewer than three exons.²⁴

Due to the characteristics of the technology, NGS also creates uncertainties and limitations of an order never encountered before in genetic testing. Uncertainties in genetic testing have always been existed, and this prompts us to constantly develop new tools to deal with those issues. Targeted NGS assays can have a role as support genetic tools in the diagnosis of lysosomal storage disorders always in combination with biochemical and clinical data. NGS assays have proven to be very powerful for making diagnoses that are particularly challenging, e.g. due to an atypical or late-onset presentation. Another strength of the method is to bring diagnostic odysseys to a more rapid conclusion than with approaches used previously.

The pilot study of Fernandez-Marmiesse et al¹² ended with 40 patients out of the 84 analyzed probands still undiagnosed, although it must be stated that 64% of these patients had a low or moderate index of suspicion of LSD. The considerable phenotypic overlap between certain LSDs and non-LSD neurometabolic or neurodevelopmental conditions pose a great limitation to the application of targeted NGS panels with genes specific only for LSDs. To address this

problem several approaches are being used. The application of broad-range genetic panels that encompass most known neurometabolic disorders¹² or the application of the Mendeliome, an NGS panel encompassing all known genes for Mendelian disorders²⁵ might be a solution to enhance the detection rate in those patient groups. On the other hand, many reports already underline the diagnostic use of whole-exome sequencing as an ultimate test and the subsequent application of virtual panels in the stepwise diagnostic process of monogenic disorders with phenotypic and genetic heterogeneity.^{19,20,21}

Conclusions

Genetic analysis has not been the primary screening tool used in the diagnosis of LSDs due to the cost and time requirements of the sequential genetic tests necessary to diagnose most of the disorders.^{10,12} However, with the wide availability of NGS technologies, a genetically based diagnosis can be established in 4-6 weeks, while providing cost reductions compared to the sequential Sanger sequencing of several single genes. Although the advantages of applying NGS in the diagnosis of LSDs are numerous, validation of a genetic diagnosis often still needs biochemical testing to confirm the functional consequences of a sequence variant. Given that NGS can be applied in a way that takes into account the many issues raised by international consensus guidelines, it can have an important role in the diagnostic process of lysosomal storage disorders. We emphasize the importance of a precise molecular genetic diagnosis early in the disease course of patients affected by LSDs, particularly as new therapies are becoming available for patients who share specific types of mutations. Further significance of NGS technologies in establishing a genetic diagnosis in patients, is due to the benefit derived by the opportunity for genetic counseling, family planning and the individual choice of reproductive options (including preimplantation genetic diagnosis) for affected families. It will be imperative to continually optimize this powerful technology with the potential benefits of patients and their families in mind.

References

1. Chambers C, Jansen LA, Dhamija R. Review of commercially available epilepsy genetic panels. *J Genet Counsel.* 2016;25:213-217.
2. de Koning TJ., Jongbloed JD, Sikkema-Raddatz B, Sinke RJ. Targeted next-generation sequencing panels for monogenetic disorders in clinical diagnostics: the opportunities and challenges. *Expert Rev Mol Diag.* 2015;15(1):61-70.

3. Najmabadi H, Hu H, Garshasbi M et al. Deep sequencing reveals 50 novel genes for recessive cognitive disorders. *Nature*. 2011;21;478(7367):57-63.
4. Hu H, Wienker TF, Musante L et al. Integrated sequence analysis pipeline provides one-stop solution for identifying disease-causing mutations. *Hum Mutat*. 2014;35(12):1427-35.
5. Okamoto N, Miya F, Tsunoda T et al. Targeted next-generation sequencing in the diagnosis of neurodevelopmental disorders. *Clin Genet*. 2015;88(3):288-92.
6. Ciara E, Rokicki D, Halat P. Difficulties in recognition of pyruvate dehydrogenase complex deficiency on the basis of clinical and biochemical features. The role of next-generation sequencing. *Mol Genet Metab Rep*. 2016;18(7):70-76.
7. Matthijs G, Souche E, Alders M et al. Guidelines for diagnostic next-generation sequencing. *EJHG* 2016;24:2-5.
8. Xu M, Motabar O, Ferrer M, et al. Disease models for the development of therapies for lysosomal storage diseases. *Ann N Y Acad Sci*. 2016 May 4. doi:10.1111/nyas.13052.
9. Beck M. New therapeutic options for lysosomal storage disorders: enzyme replacement, small molecules and gene therapy. *Hum Genet*. 2007;121(1):1-22.
10. Winchester B. Lysosomal diseases: diagnostic update. *J Inherit Metab Dis*. 2014;37(4):599-608.
11. Peelen GO, de Jong JG, Wevers RA. HPLC analysis of oligosaccharides in urine from oligosaccharidosis patients. *Clin Chem*. 1994;40(6):914-21.
12. Fernández-Marmiesse A, Morey M, Pineda M et al. Assessment of a targeted resequencing assay as a support tool in the diagnosis of lysosomal storage disorders. *Orphanet J Rare Dis*. 2014;25:9:59.
13. Geraets RD, Koh Sy, Hastings ML et al. Moving towards effective therapeutic strategies for Neuronal Ceroid Lipofuscinosis. *Orphanet J Rare Dis*. 2016;16:11:40.
14. Hoischen A, van Bon BW, Gilissen C et al. De novo mutations of SETBP1 cause Schinzel-Giedion syndrome. *Nat Genet*. 2010;42(6):483-5.
15. Lévesque S, Auray-Blais C, Gravel E et al. Diagnosis of late-onset Pompe disease and other muscle disorders by next-generation sequencing. *Orphanet J Rare Dis*. 2016;11:8.
16. Di Fruscio G, Schulz A, De Cegli R et al. Lysoplex: An efficient toolkit to detect DNA sequence variations in the autophagy-lysosomal pathway. *Autophagy*. 2015;11(6):928-938.

17. Settembre C, Fraldi A, Medina DL et al. Signals from the lysosome: a control centre for cellular clearance and energy metabolism. *Nat Rev Mol Cell Biol.* 2013;14(5):283-96.
18. Schulz A, Kohlschütter A, Mink J et al. NCL diseases - clinical perspectives. *Biochim Biophys Acta.* 2013;1832(11):1801-1806.
19. Ng SB, Nickerson DA, Bamshad MJ, Shendure J. Massively parallel sequencing and rare disease. *Hum Mol Genet.* 2010;19(R2):R119-124.
20. Lelieveld SH, Spielmann M, Mundlos S, Veltman JA, Gilissen C. Comparison of Exome and Genome Sequencing Technologies for the Complete Capture of Protein-Coding Regions. *Hum Mutat.* 2015;36(8):815-22.
21. Meienberg J, Bruggmann R, Oexle K, Matyas G. Clinical sequencing: is WGS the better WES? *Hum Genet.* 2016;135(3):359-362.
22. Bonafé L, Kariminejad A, Li J et al. Peripheral osteolysis in adults linked to *ASAHI* (acid ceramidase) mutations: A new presentation of Farber disease. *Arthritis Rheumatol.* 2016 Mar 4. doi: 10.1002/art.39659.
23. Landrum MJ, Lee, JM, Riley GR et al. ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acid Res.* 2014;42: D980-985.
24. Tan R, Wang Y, Kleinstein SE, et al. An evaluation of copy number variation detection tools from whole-exome sequencing data. *Hum Mutat.* 2014;15:661.
25. Moosa S, Obregon MG, Altmüller J et al. Novel IFT122 mutations in three Argentinian patients with cranioectodermal dysplasia: Expanding the mutational spectrum. *Am J Med Genet A.* 2016;170(5):1295-301.

Acknowledgements

This work was supported by the Janos Bolyai Research Scholarship of the Hungarian Academy of Sciences to KK.