UCLA UCLA Previously Published Works

Title

Looking beyond the exome: a phenotype-first approach to molecular diagnostic resolution in rare and undiagnosed diseases

Permalink <https://escholarship.org/uc/item/1jn52197>

Journal Genetics in Medicine, 20(4)

ISSN 1098-3600

Authors

Pena, Loren DM Jiang, Yong-Hui Schoch, Kelly [et al.](https://escholarship.org/uc/item/1jn52197#author)

Publication Date 2018-04-01

DOI

10.1038/gim.2017.128

Peer reviewed

HHS Public Access

Author manuscript Genet Med. Author manuscript; available in PMC 2018 April 11.

Published in final edited form as: Genet Med. 2018 April ; 20(4): 464–469. doi:10.1038/gim.2017.128.

Looking beyond the exome: a phenotype-first approach to molecular diagnostic resolution in rare and undiagnosed diseases

Loren DM Pena, MD, PhD1, **Yong-Hui Jiang, MD PhD**1, **Kelly Schoch, MS**1, **Rebecca C. Spillmann, MS**1, **Nicole Walley, MS**1, **Nicholas Stong, PhD**2, **Sarah Rapisardo Horn, PhD**3, **Jennifer A. Sullivan, MS**1, **Allyn McConkie-Rosell, PhD**1, **Sujay Kansagra, MD**4, **Edward C. Smith, MD**4, **Mays El-Dairi, MD**5, **Jane Bellet, MD**6, **Martha Ann Keels, DDS, PhD**7, **Joan Jasien, MD**4, **Peter G. Kranz, MD**8, **Richard Noel, MD, PhD**9, **Shashi K. Nagaraj, MD**10, **Robert K. Lark, MD**11, **Daniel SG Wechsler, MD, PhD**12, **Daniela del Gaudio, PhD**13, **Marco L. Leung, PhD**13, **Laura G. Hendon, MS**14, **Collette C. Parker, MD**15, **Kelly L. Jones, MD**16, **UDN Members**, **David B. Goldstein, PhD**2, and **Vandana Shashi, MD**1,*

¹Division of Medical Genetics, Department of Pediatrics, Duke University Medical Center

2 Institute of Genomic Medicine, Columbia University, New York, N.Y.

³Department of Pathology, Duke University Medical Center

⁴Division of Neurology, Department of Pediatrics, Duke University Medical Center

⁵Department of Ophthalmology, Duke University Medical Center

⁶Departments of Dermatology and Pediatrics, Duke University Medical Center

⁷Departments of Pediatrics and Surgery, Duke University Medical Center

⁸Division of Neuroradiology, Department of Radiology, Duke University Medical Center

⁹Division of Gastroenterology, Department of Pediatrics, Duke University Medical Center

¹⁰Division of Pediatric Nephrology, Department of Pediatrics, Duke University Medical Center

¹¹Department of Orthopedic Surgery, Duke University Medical Center

¹² Division of Pediatric Hematology-Oncology, Department of Pediatrics, Duke University Medical **Center**

13Department of Human Genetics, University of Chicago, Chicago III

¹⁴Division of Maternal Fetal Medicine, University of Mississippi Medical Center, Jackson, Miss

¹⁵Division of Child Neurology, Department of Pediatrics, University of Mississippi Medical Center

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:http://www.nature.com/authors/editorial_policies/license.html#terms

^{*}Corresponding Author: Vandana Shashi, MD, Professor of Pediatrics, Division of Medical Genetics, Duke University Medical Center, Durham NC 27710, vandana.shashi@duke.edu, Fax: 919-668-0414, Tel: 919-681-2616.

Conflicts of Interest

The rest of the authors declare no conflicts of interest related to this manuscript.

Supplementary information is available at the Genetics in Medicine website.

¹⁶Division of Medical Genetics, Department of Pediatrics, University of Mississippi Medical Center

Abstract

Purpose—To describe examples of missed pathogenic variants on whole exome sequencing (WES) and the importance of deep phenotyping for further diagnostic testing.

Methods—Guided by phenotypic information, three children with negative WES underwent targeted single gene testing.

Results—*Individual 1* had a clinical diagnosis consistent with infantile systemic hyalinosis, although WES and an NGS-based ANTXR2 test were negative. Sanger sequencing of ANTXR2 revealed a homozygous single base pair insertion, previously missed by the WES variant caller software. Individual 2 had neurodevelopmental regression and cerebellar atrophy, with no diagnosis on WES. New clinical findings prompted Sanger sequencing and copy number testing of PLA2G6. A novel homozygous deletion of the non-coding exon 1 (not included in the WES capture kit) was detected, with extension into the promoter, confirming the clinical suspicion of infantile neuroaxonal dystrophy. *Individual 3* had progressive ataxia, spasticity and MRI changes of vanishing white matter leukoencephalopathy. An NGS leukodystrophy gene panel and WES showed a heterozygous pathogenic variant in *EIF2B5*; no deletions/duplications were detected. Sanger sequencing of EIF2B5 showed a frameshift indel, likely missed due to failure of alignment.

Conclusions—These cases illustrate potential pitfalls of WES/NGS testing, and the importance of phenotype-guided molecular testing in yielding diagnoses.

Keywords

ANTXR2; PLA2G6; EIF2B5; whole exome sequencing; undiagnosed diseases network; infantile systemic hyalinosis; infantile neuroaxonal dystrophy; leukoencephalopathy with vanishing white matter

Introduction

Whole exome sequencing (WES) has revolutionized clinical genetics by providing a comprehensive and agnostic method for patient evaluation¹. Diagnostic rates vary from 25– 50% and WES has allowed new disease-gene identification and insights into the phenotypic and genetic heterogeneity of Mendelian disorders $2-4$. WES has quickly become part of the standard repertoire of genetic testing, with a prevailing sense that a negative result indicates that disorders in the differential diagnoses have been effectively excluded. We describe three individuals in whom WES and targeted next generation sequencing (NGS)-based testing were non-diagnostic. Phenotype reassessment and use of additional data, such as the single nucleotide polymorphism (SNP) microarray data, helped determine the next steps in the diagnostic process. Targeted single-gene Sanger sequencing and deletion/duplication analyses identified pathogenic variants for the clinically suspected genetic disorder in all three individuals. We provide insights into the reasons for negative WES results and, in an era when genomic technology tends to drive the diagnostic process, we highlight the importance of revisiting clinical information for additional targeted testing.

Patients, Methods, and Results

Individuals 1 and 2 were evaluated at the Duke clinical site of the NIH Undiagnosed Diseases Network (UDN) [\(https://undiagnosed.hms.harvard.edu](https://undiagnosed.hms.harvard.edu)) and individual 3 at the Duke Genome Sequencing Clinic.

Individual 1

An 18-month-old Mexican female with progressive joint contractures and related morbidity was evaluated due to extensive skin plaques, subcutaneous and gingival nodules, a biopsy demonstrated dermal accumulation of amorphous hyaline material. Review of a duodenal biopsy showed dilated lymphatics and mucosal edema, consistent with clinical symptoms of protein-losing enteropathy. She had intact cognitive skills, ruling out alternative diagnoses such as Farber and Winchester syndromes. The parents reported a common ancestor in Mexico.

This extended phenotype was consistent with infantile systemic hyalinosis (ISH, OMIM #228600), an autosomal recessive disorder due to loss of function variants in *ANTXR2*, leading to widespread progressive accumulation of hyaline material and childhood death⁵.

Pertinent Previous Genetic Testing—A SNP microarray identified 64.2 Mb regions of homozygosity (ROH), including the *ANTXR2* locus. An initial NGS-based sequencing of the exons and flanking splice junctions of ANTXR2, followed by a proband-only WES through a commercial laboratory were negative.

Results of post-WES Genetic Testing in UDN—Review of the SNP microarray did not identify deletions in the regions of ROH. Review of the NGS-based ANTXR2 sequencing and WES data confirmed that 98.5% of the coding regions of *ANTXR2* gene were covered at $>10\times$, except exon 1 (85.6%). Sanger sequencing revealed a homozygous pathogenic variant in exon 13 (c.1073dupC). Parental studies confirmed trans configuration. Upon discussion with the commercial laboratory that had performed the WES and NGSbased testing, it appeared that their variant calling software had not reported the variant. Its location adjacent to a homopolymeric repeat region and a common SNP could have contributed to low mapping quality, leading to failure of variant calling (Figure 1). Subsequently, we obtained the BAM files and manual inspection of the data by the UDN bioinformatician confirmed the presence of the variant.

Individual 2

A 3.5-year-old girl of Pakistani origin exhibited developmental regression at 16 months of age, cerebellar atrophy and a negative trio WES (proband and parents). She had lost the ability to cruise, crawl, sit, speak and eat by mouth. The parents were first cousins, and the proband had two first cousins once removed who died at age two years after neurodevelopmental regression.

The patient had optic atrophy, profound generalized hypotonia, minimal spontaneous movements, tongue fasciculations and diminished Achilles reflexes, in contrast to previously observed hypertonia with generalized hyperreflexia at age 2.5 years. Review of brain MRIs

obtained at 2 and 2.5 years of age revealed stable white matter volume loss of the vermis and cerebellar hemispheres, a normal pons and no iron accumulation (Supplementary Materials and Methods, Figure 1S). The new clinical finding of peripheral nerve involvement led us to consider infantile neuroaxonal dystrophy (IND) (OMIM# 256600), a disorder of neurodevelopmental regression in childhood and early death, caused by biallelic variants in PLA2G6^{6,7}.

Pertinent Previous Genetic Testing—Metabolic laboratory tests and an ataxia gene panel (42 genes) were negative. Trio WES through a commercial laboratory detected a homozygous missense variant of unknown significance in RPGRIP1L, but the clinical course and brain MRI findings were not consistent with Joubert syndrome.

Results of Post- WES Genetic Testing in UDN—A review of the SNP microarray identified several ROH in >4.6% of the genome; 12 genes within these regions, including PLA2G6, were associated with cerebellar atrophy and developmental regression. Manual inspection of the WES BAM files found no functionally significant single nucleotide or copy number variants in these 12 genes. Sanger sequencing and multiplex ligation-dependent probe amplification (MLPA) for deletions/duplications was performed for the five genes within the ROH with the greatest phenotypic overlap: PLA2G6, AC02, BCS1L, NDUFA11 and ADSL. Sanger sequencing was normal for all. MLPA showed a novel homozygous deletion of the non-coding exon 1 in PLA2G6 (Figures 2A and 2S). Follow-up MLPA analysis of the parents confirmed that they were carriers of the deletion.

The breakpoint junction of the PLA2G6 deletion was amplified by long-range PCR (Supplementary Methods). The deletion included 2431-bp in the 5'UTR region of PLA2G6 and revealed a 7-bp insertion at the breakpoint junction (c.−545_

−46+1931delinsCGATCTC) (Figure 3S). Fine mapping analysis⁸ demonstrated that the deletion encompassed a portion of the promoter region of the gene. Quantitative RT-PCR showed that mRNA expression of PLA2G6 was significantly lower in the patient's blood compared with unaffected controls $(p<0.01)$ (Figure 2B). Review of the WES data revealed that the capture kit did not include the non-coding exon 1.

Individual 3

An 8-year-old Caucasian female was evaluated for symptoms of ataxia, seizures and white matter disease. At age 3 years, she developed frequent falls and progressive decline in fine motor skills, speech and short-term memory. Generalized seizures started at 7.5 years. Exam revealed dysarthria, lower extremity spasticity, hyperreflexia, clonus and a wide-based gait. Brain MRI showed diffuse symmetric non-enhancing signal abnormalities involving both cerebral hemispheres with volume loss (Figure 4S). The features were consistent with leukoencephalopathy with vanishing white matter (VWM) (OMIM #603896), a disorder that presents with neurological regression, ataxia, spasticity, epilepsy and progressively vanishing white matter in brain MRI⁹. Variants in five genes ($EIF2B1-B5$) can cause the disorder, most frequently biallelic variants in EIF2B5¹⁰.

Pertinent Previous Genetic Testing—SNP microarray analysis detected a paternallyinherited 416 Kb deletion at 10p12.31, interpreted as a benign variant. An NGS panel of 62 genes for VWM leukodystrophy showed a heterozygous c.338G>A, p. Arg113His pathogenic variant in the EIF2B5 gene. Subsequent deletion/duplication testing for the EIF2B5 gene via exon-targeted array-CHG was normal. Trio WES re-identified the heterozygous p. Arg113His variant. Manual inspection of the reads for the *EIF2B5* gene did not reveal any additional variants, with coverage of $>10\times$ for 100% of this gene.

Results of Post-WES Genetic Testing—Due to the continued clinical suspicion of VWM leukodystrophy, and the detection of one pathogenic variant in the gene, Sanger sequencing of *EIF2B5* was pursued. A heterozygous insertion, c.1694delAins45; p.Lys565Ilefs*38 was detected. Although not previously reported in patients with VWM, the insertion had been detected once by the commercial laboratory, in an affected individual. Subsequent parental testing confirmed trans configuration for each variant. Difficulty in alignment of indels larger than 20–50 bp was likely the reason for missing this variant by WES, since retrospective manual inspection of the BAM files failed to detect it.

Discussion

WES is increasingly used as the premier and first-line test for rare and undiagnosed Mendelian disorders^{1,2,4,11–14}. The vast majority (>97%) of variants detected by Sanger sequencing can also be detected by WES and with increased detection of mosaicism, WES is a practical tool for comprehensive molecular evaluation¹⁵. When WES is negative, reanalysis of the data can provide resolution in $10\% - 30\%$ of cases¹⁶ and this is likely to increase with improvements in technology and new disease associations. However, there is little clarity on the diagnostic options if WES and reanalysis remain negative. Although whole genome sequencing (WGS) may be an option, it is not currently widely available clinically¹⁷. Therefore, in instances when WES is negative, clinicians may conclude that no diagnostic options remain for these patients.

WES is a complex high-throughput method, and data loss is possible at each step. Pathogenic variants in known disease causing genes may be missed because of decreased coverage, locus-specific features such as GC-rich regions, homopolymeric repeats, sequencing biases and indels that are $>20-50$ nucleotides¹⁸. The most common reason for variants being missed is a lack of sufficient sequence coverage depth¹⁹. Laboratories performing WES and NGS panels may use alternative methods to capture these²⁰. Clinicians do reconsider the fit of a phenotype when interpreting variants of uncertain significance on WES, but when WES is negative, adequate coverage of selected genes or exons may lead to a belief that the WES was truly comprehensive. However, in all three of our individuals, coverage was adequate ($10\times$ at 98– 100% of the bases), yet the pathogenic variants were missed. Thus, a negative WES result should be interpreted in the clinical context of the individual patient, to determine further testing.

In individual 1, the well-characterized phenotype was consistent with only one diagnosis (ISH), and the ROH on the SNP array included ANTXR2, but molecular confirmation remained elusive despite repeated sequencing and adequate coverage. The single nucleotide

insertion detected by Sanger sequencing is located adjacent to a complex repetitive region and a SNP, both of which could have decreased the mapping quality of the region around the insertion, and resulted in failure of the variant caller program to detect the insertion. Updated variant calling software and/or manual inspection of the reads would have identified this variant. Manual inspection is of particular importance when a limited set of specific genes is under consideration, but it is not standard practice in commercial laboratories.

Individual 2 had clinical features that changed over a year, leading us to strongly consider a diagnosis of infantile neuroaxonal dystrophy. In this instance, the WES did not capture the deletion because the non-coding exon 1 was not included in the capture kit, as is commonly the case with WES capture kits. Furthermore, structural variants of this size (2.3 kb) would not be detectable by WES. The clinical phenotype and the ROH containing the PLA2G6 gene led us to pursue Sanger sequencing and MLPA, which detected the deletion. It is possible that WGS might have detected this structural variant, but its limited availability and lack of validation of WGS structural variant callers make this an impractical option.

Individual 3 had features consistent with VWM leukoencephalopathy. Although indels <50bp are below the resolution of exon-level deletion/duplication analyses, we would expect detection by WES, if alignment works well. WES missed the 46 bp indel in the EIF2B5 gene completely, since the detection rate of indels decreases with sizes >20–50 bp.

Clinical geneticists are aware of WES being unsuitable for trinucleotide repeat disorders, mitochondrial DNA variants, epigenetic disorders and large structural variants. However, for Mendelian disorders in which SNVs and small indels are possible, the prevalent thinking is that if coverage of the genes of interest by WES is adequate, the disorders have been effectively excluded. Indeed, an increased depth of coverage would not have detected the missed variants in all three of our cases. The cases presented here underscore the importance of further testing if the clinical phenotype is strongly indicative of a specific condition when WES is negative.

In conclusion, these three case examples illustrate the importance a multi-pronged approach when WES is negative. These include: 1) Obtaining detailed clinical phenotyping to create an accurate differential diagnosis; 2) Reconsidering the family history and mode of inheritance; 3) Reassessing SNP microarray data to identify potential causal genes; 4) Manual inspection of the WES reads for genes that are of interest and obtaining information on capture kits and coverage; 5) Pursuing alternative sequencing methodologies such as Sanger sequencing and deletion/duplication testing to detect SNVs and indels that might have been missed with WES. Although WES is comprehensive, its limitations must be considered when negative results are obtained, and alternative diagnostic approaches should be pursued if the phenotype is compelling.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Research reported in this manuscript was supported by the NIH Common Fund, through the Office of Strategic Coordination/Office of the NIH Director under Award Number(s) [1U01HG007672-01 to Shashi V and Goldstein DB]. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

David Goldstein is a founder of and holds equity in Pairnomix and Praxis, serves as a consultant to AstraZeneca, and has research supported by Janssen, Gilead, Biogen, AstraZeneca, and UCB.

Undiagnosed Diseases Network Members

Author Manuscript Author Manuscript

References

1. Chong JX, Buckingham KJ, Jhangiani SN, et al. The Genetic Basis of Mendelian Phenotypes: Discoveries, Challenges, and Opportunities. Am J Hum Genet. 2015

- 2. Need AC, Shashi V, Hitomi Y, et al. Clinical application of exome sequencing in undiagnosed genetic conditions. J Med Genet. 2012; 49(6):353–361. [PubMed: 22581936]
- 3. Yang Y, Muzny DM, Xia F, et al. Molecular findings among patients referred for clinical wholeexome sequencing. JAMA. 2014
- 4. Lee H, Deignan JL, Dorrani N, et al. Clinical exome sequencing for genetic identification of rare mendelian disorders. JAMA. 2014
- 5. El-Kamah GY, Fong K, El-Ruby M, et al. Spectrum of mutations in the ANTXR2 (CMG2) gene in infantile systemic hyalinosis and juvenile hyaline fibromatosis. Br J Dermatol. 2010; 163(1):213– 215. [PubMed: 20331448]
- 6. Kurian MA, Morgan NV, MacPherson L, et al. Phenotypic spectrum of neurodegeneration associated with mutations in the PLA2G6 gene (PLAN). Neurology. 2008; 70(18):1623–1629. [PubMed: 18443314]
- 7. Crompton D, Rehal PK, MacPherson L, et al. Multiplex ligation-dependent probe amplification (MLPA) analysis is an effective tool for the detection of novel intragenic PLA2G6 mutations: implications for molecular diagnosis. Mol Genet Metab. 2010; 100(2):207–212. [PubMed: 20226704]
- 8. Larsson Forsell PK, Kennedy BP, Claesson HE. The human calcium-independent phospholipase A2 gene multiple enzymes with distinct properties from a single gene. Eur J Biochem. 1999; 262(2): 575–585. [PubMed: 10336645]
- 9. Bugiani M, Boor I, Powers JM, Scheper GC, van der Knaap MS. Leukoencephalopathy with vanishing white matter: a review. J Neuropathol Exp Neurol. 2010; 69(10):987–996. [PubMed: 20838246]
- 10. van der Knaap MS, Leegwater PA, Konst AA, et al. Mutations in each of the five subunits of translation initiation factor eIF2B can cause leukoencephalopathy with vanishing white matter. Ann Neurol. 2002; 51(2):264–270. [PubMed: 11835386]
- 11. Bowdin S, Gilbert A, Bedoukian E, et al. Recommendations for the integration of genomics into clinical practice. Genet Med. 2016
- 12. Gilissen C, Hoischen A, Brunner HG, Veltman JA. Unlocking Mendelian disease using exome sequencing. Genome Biol. 2011; 12(9):228. [PubMed: 21920049]
- 13. Yang Y, Muzny DM, Reid JG, et al. Clinical whole-exome sequencing for the diagnosis of mendelian disorders. N Engl J Med. 2013; 369(16):1502–1511. [PubMed: 24088041]
- 14. Need AC, Shashi V, Schoch K, Petrovski S, Goldstein DB. The importance of dynamic reanalysis in diagnostic whole exome sequencing. J Med Genet. 2017; 54(3):155–156. [PubMed: 27899421]
- 15. Hamilton A, Tetreault M, Dyment DA, et al. Concordance between whole-exome sequencing and clinical Sanger sequencing: implications for patient care. Molecular genetics & genomic medicine. 2016; 4(5):504–512. [PubMed: 27652278]
- 16. Wenger AM, Guturu H, Bernstein JA, Bejerano G. Systematic reanalysis of clinical exome data yields additional diagnoses: implications for providers. Genet Med. 2017; 19(2):209–214. [PubMed: 27441994]
- 17. Meynert AM, Ansari M, FitzPatrick DR, Taylor MS. Variant detection sensitivity and biases in whole genome and exome sequencing. BMC bioinformatics. 2014; 15:247. [PubMed: 25038816]
- 18. Ross MG, Russ C, Costello M, et al. Characterizing and measuring bias in sequence data. Genome Biol. 2013; 14(5):R51. [PubMed: 23718773]
- 19. 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):815–822. [PubMed: 25973577]
- 20. Tiwari A, Lemke J, Altmueller J, et al. Identification of Novel and Recurrent Disease-Causing Mutations in Retinal Dystrophies Using Whole Exome Sequencing (WES): Benefits and Limitations. PLoS One. 2016; 11(7):e0158692. [PubMed: 27391102]

Figure 1.

Integrated Genome Variant browser showing condensed read alignment for internally realigned data (A), and the BAM file provided by the clinical laboratory (B). Reads are shown stacked together, with colors indicating a mismatch to the reference sequence. Almost all reads show the pathogenic insertion (c.1073dupC, 22:g.80905986dupG) as a purple mark. The insertion call may be considered low quality because of the adjacent homopolymer repeat and nearby variant. The nearby SNP (c.1069G>C, 22:g.80905990C>G) can be seen by the majority of reads containing the orange G nucleotide change. This variant can also contribute to lower mapping quality and can affect variant quality at the insertion.

Pena et al. Page 14

Β.

Figure 2.

A. MLPA analysis of PLA2G6. Normalized MLPA data showing the homozygous deletion of the non-coding exon 1 of the PLA2G6 gene leading to total absence of amplification and hence a ratio of zero for the probe covering this region

B. qRT-PCR analysis of PLA2G6 mRNA expression in blood. Results are expressed as means \pm SD; mRNA levels were quantitated with real-time PCR and normalized to the level of GAPDH. The figure represents real-time PCR quantification of PLA2G6 gene expression in the patient compared to controls. All results were done in triplicates. P-value was calculated by student t-test.