ANNALS of Clinical and Translational Neurology
Open Access
AMERICAN NEUROLOGICAL ASSOCIATION

BRIEF COMMUNICATION

Genome sequencing in persistently unsolved white matter disorders

Guy Helman[1,2], Bryan R. Lajoie[3], Joanna Crawford[2], Asako Takanohashi[4], Marzena Walkiewicz[1], Egor Dolzhenko[3], Andrew M. Gross[3], Vladimir G. Gainullin[3], Stephen J. Bent[5], Emma M. Jenkinson[6], Sacha Ferdinandusse[7], Hans R. Waterham[7], Imen Dorboz[8], Enrico Bertini[9,10], Noriko Miyake[11], Nicole I. Wolf[12], Truus E. M. Abbink[12], Susan M. Kirwin[13], Christina M. Tan[13], Grace M. Hobson[13], Long Guo[14], Shiro Ikegawa[14], Amy Pizzino[4], Johanna L. Schmidt[4], Genevieve Bernard[15,16,17], Raphael Schiffmann[18], Marjo S. van der Knaap[12,19], Cas Simons[1,2], Ryan J. Taft[3] & Adeline Vanderver[4,20]

[1]Murdoch Children's Research Institute, The Royal Children's Hospital Melbourne, Parkville, Melbourne, Australia
[2]Institute for Molecular Bioscience, The University of Queensland, Brisbane, Australia
[3]Illumina Inc., San Diego, California
[4]Division of Neurology, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania
[5]Data61, Commonwealth Scientific and Industrial Research Organisation, Brisbane, Australia
[6]Faculty of Biology, Medicine and Health, School of Biological Sciences, Division of Evolution and Genomic Sciences, University of Manchester, Manchester, United Kingdom
[7]Laboratory Genetic Metabolic Diseases, Department of Clinical Chemistry, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, The Netherlands
[8]INSERM UMR 1141, DHU PROTECT, Université Paris Diderot- Sorbonne, Paris Cité, France
[9]Unit of Neuromuscular and Neurodegenerative Disorders, Laboratory of Molecular Medicine, Bambino Gesu' Children's Hospital, Rome, Italy
[10]Genetics and Rare Diseases Research Division, Ospedale Pediatrico Bambino Gesù, 00146, Rome, Italy
[11]Department of Human Genetics, Yokohama City University Graduate School of Medicine, Fukuura, Kanazawa-ku, Yokohama, 236-0004, Japan
[12]Department of Child Neurology, Emma Children's Hospital, Amsterdam University Medical Centers, Vrije Universiteit Amsterdam and Amsterdam Neuroscience, Amsterdam, The Netherlands
[13]Molecular Diagnostics Laboratory, Nemours Biomedical Research, Nemours/Alfred I. duPont Hospital for Children, Wilmington, Delaware
[14]Laboratory of Bone and Joint Diseases, RIKEN Center for Integrative Medical Sciences, Tokyo, Japan
[15]Departments of Neurology and Neurosurgery, Pediatrics, and Human Genetics, McGill University, Montreal, Canada
[16]Division of Medical Genetics, Montreal Children's Hospital, McGill University Health Center, Montreal, Canada
[17]Child Health and Human Development Program, Research Institute of the McGill University Health Center, Montreal, Canada
[18]Institute of Metabolic Disease, Baylor Scott & White Research Institute, Dallas, Texas
[19]Department of Functional Genomics, Amsterdam Neuroscience, VU University, Amsterdam, the Netherlands
[20]Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania

Correspondence
Adeline Vanderver, Division of Neurology, Children's Hospital of Philadelphia, 34th and Civic Center Blvd, Philadelphia, PA 19010. Tel: +1 215 590 3068; Fax: +1 215 590 1234; E-mail: vandervera@email.chop.edu
Cas Simons, Murdoch Children's Research Institute, Royal Children's Hospital, Parkville, VIC 3052, Australia. Tel: +61 7 9936 6488; Fax: +61 3 8341 6212; E-mail: cas.simons@mcri.edu.au
Ryan J. Taft, VP Scientific Research, 5200 Illumina Way, San Diego, CA 92122. Tel: +1 (858) 943 5775; Fax: +1 858 202 4766; E-mail: rtaft@illumina.com

Funding Information
AP, JLS, AT, and AV are supported by the Myelin Disorders Bioregistry Project. GB has received the New Investigator Salary Award

Abstract

Genetic white matter disorders have heterogeneous etiologies and overlapping clinical presentations. We performed a study of the diagnostic efficacy of genome sequencing in 41 unsolved cases with prior exome sequencing, resolving an additional 14 from an historical cohort (n = 191). Reanalysis in the context of novel disease-associated genes and improved variant curation and annotation resolved 64% of cases. The remaining diagnoses were directly attributable to genome sequencing, including cases with small and large copy number variants (CNVs) and variants in deep intronic and technically difficult regions. Genome sequencing, in combination with other methodologies, achieved a diagnostic yield of 85% in this retrospective cohort.

from the Canadian Institutes of Health Research (2017-2022). Sequencing was provided by Illumina, Inc. The Myelin Disorders Bioregistry Project was supported by the Children's Research Institute at the Children's Hospital of Philadelphia. This study was in part financed by the Australian National Health and Medical Research Council (NHMRC 1068278). The research conducted at the Murdoch Children's Research Institute was supported by the Victorian Government's Operational Infrastructure Support Program.

Received: 1 October 2019; Revised: 5 November 2019; Accepted: 5 November 2019

Annals of Clinical and Translational Neurology **2020; 7(1): 144–152**

doi: 10.1002/acn3.50957

Guy Helman and Bryan R. Lajoie contributed equally to this work

Cas Simons, Ryan J. Taft and Adeline Vanderver contributed equally to this work

Introduction

Next-generation sequencing (NGS) including both targeted gene panels and exome sequencing has become a central component of the diagnostic evaluation of individuals with a neurologic disorder of unknown origin.[1,2] Exome sequencing has shown success in phenotypically diverse cohorts, including the leukodystrophies, a broad class of genetic disorders that affect the white matter of the central nervous system (CNS).[2,3]

Leukodystrophies are diverse in origin and highly heterogeneous in presentation and disease course, making diagnosis challenging. Until recently, the probability of obtaining a definitive molecular diagnosis was less than 50% over 5 years[2] and despite increases of up to threefold in diagnostic efficacy, many cases remain unsolved. Previously, we reported on a cohort of 191 families with a suspected leukodystrophy, of which 71 were persistently unsolved despite targeted molecular and enzymatic testing and 19 were lost to follow-up.[1] Trio-or-greater exome sequencing analysis resolved 42% (30/71) of unsolved cases, which, in combination with standard of care approaches, yielded a 77% overall diagnostic rate from 172 families available for testing. This was a substantial increase over the historical norms,[2] but left at least 23% persistently unsolved.[1]

Genome sequencing has the potential to detect a wide variety of variant types, including single-nucleotide variants and indels (out-performing exome sequencing in protein-coding regions), copy number variants (CNV),[4] repeat expansions,[5] and detection of pathogenic nonprotein-coding variation that is missed by other NGS approaches. We pursued genome sequencing on the remaining 41 persistently unsolved families to assess the potential value of genome sequencing diagnostics in a pediatric neurological disease cohort.

Methods

Recruitment

Affected individuals were referred to the Myelin Disorders Bioregistry Project (MDBP) for unsolved leukoencephalopathy of presumed genetic etiology between 1 August 2009 to 31 July 2013, as previously described.[1] The study had approval from the institutional review boards at collaborating institutions.

Clinical and neuroimaging descriptions are available on request. All individuals had abnormal white matter identified by neuroimaging, suggestive of a leukodystrophy. Symptom onset ranged from birth to 19 years. Ethnicities varied and included individuals of mixed and northern

European descent, as well as African American, Arabian, African, Asian, and Latin American origin.[1]

Genome sequencing and analysis

Genome sequencing was performed at Illumina Inc., San Diego on an Illumina 2000 using 2 x 125 nucleotide paired-end reads. Dual analyses were run in parallel for each sample. In one, reads were aligned to the reference human genome (GRCh37) using the Burrows-Wheeler Aligner (BWA) software package and pedigree informed variant calling was performed using the GATK HaplotypeCaller v3.7.[6] Variant annotation was performed using SnpEff v4.3m.[7] This analysis used a custom variant annotation and interpretation interface to identify possible causal variants. In the other analysis, samples were processed using the Illumina Secondary Analysis Software v5.11.0 (Northstar v5 release), aligned to the reference human genome (GRCh37). Candidate variants were identified using a custom-built variant interpretation engine. In both analyses, variants were triaged and prioritized by minor allele frequency, conservation, genotype, inheritance, disease-association, consequence, and predicted pathogenicity. Candidate splice-altering variants were validated using a minigene splicing assay (Appendix S1).[8]

Assessment of Variants

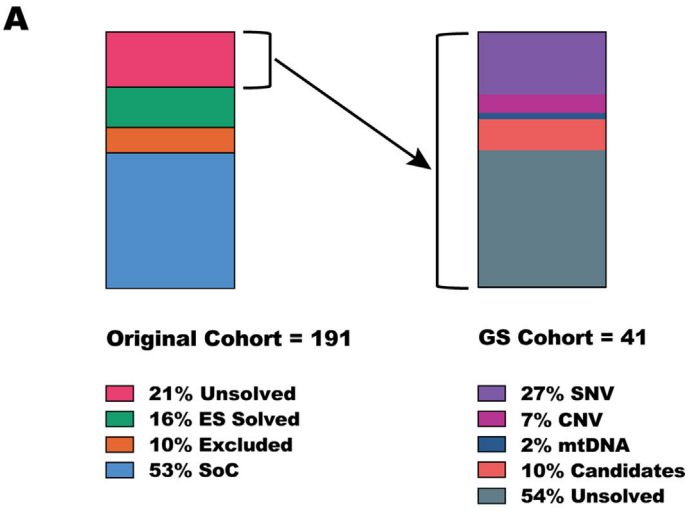
Putatively causal variants were assessed as per the American College of Medical Genetics (AGMC) guidelines.[9] Cases with variants in known disease genes meeting the ACMG criteria for pathogenic or likely pathogenic, and whose clinical features were concordant with the established gene–disease relationship (including magnetic resonance imaging (MRI) patterns) were classified as resolved. Candidate variants in potentially novel disease genes were submitted to GeneMatcher.

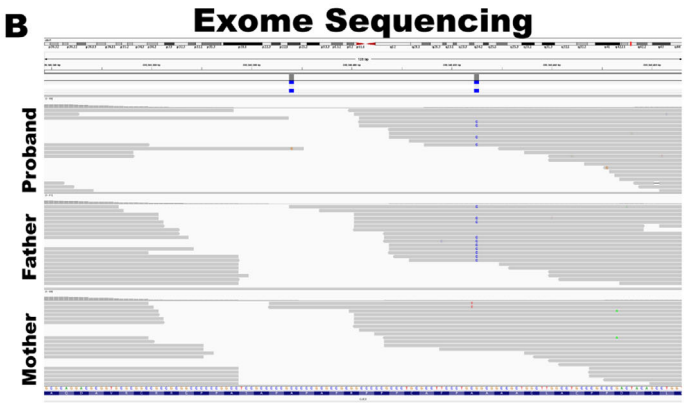
Results

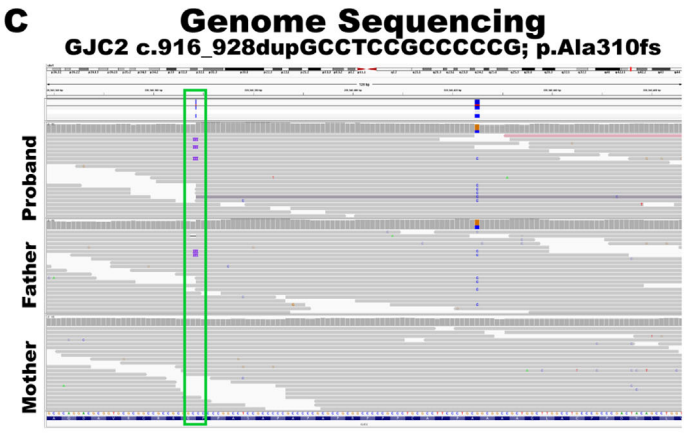
Forty-one families that remained without a molecular etiology after exome sequencing received genome sequencing (Fig. 1A). The mean read depth in probands was 34X and on average, 91% of the genome had coverage depth greater than 20X (Appendix S1). Genome sequencing resulted in a molecular diagnosis for 14 families (34%) (Table 1). Nine diagnoses were achieved through improvements in variant curation or novel disease-associated genes described since the exome sequencing analysis. Five diagnoses were achieved through identification of CNVs (three cases), deep intronic variants (one case), and variants in technically difficult regions (one case). One case was considered clinically resolved following multidisciplinary review.

Since our initial study, more than 1,200 new gene–disease relationships have been described, and studies have shown increases of up to 10% in diagnostic yield with reanalysis within 24–36 months.[10] Reanalysis in the context of recently published literature (and associated annotation pipelines) allowed resolution of an additional eight cases. These included a de novo missense variant in *H3F3B* (OMIM:601058) (LD_0246),[11] a hemizygous, synonymous variant in *AIFM1* (OMIM:300169) (LD_0500),[12] and biallelic variants in *HIKESHI* (OMIM:614908) (LD_0162),[13] *NKX6-2* (OMIM:605955) (LD_0527),[14] and *SPATA5* (OMIM:613940) (LD_0808),[15] all previously associated with prominent white matter disease in multiple affected individuals. Affected individuals in families LD_0579 and LD_0587 had variants in genes previously implicated in neurologic syndromes where improvements to our analysis pipeline or improved phenotypic understand permitted a diagnosis or prioritized a high confidence candidate, in *L1CAM* (OMIM:308840)[16] and *KDM5C* (OMIM:314690),[17] respectively. White matter abnormalities are a rare association with *L1CAM*-related disorder.[16] LD_0587 has a clinical presentation with intellectual disability, epilepsy, aggressive behavior, and macrocephaly, all consistent features of *KDM5C*-related disorder and is classified as clinically resolved but the identified variant lacks definitive proof to be classified as likely pathogenic or pathogenic.[17] This case was not included in overall numbers of definitively resolved cases as this variant remains a variant of uncertain significance per the ACMG criteria. Finally, causal nonprotein-coding variants were identified in *SNORD118* (OMIM:616663), confirming a clinical diagnosis of leukoencephalopathy with calcifications and cysts (OMIM:614561) in LD_0807.[18] *SNORD118* variants were found as part of a targeted cohort study and concomitantly found in our genome sequencing cohort.[18]

In two individuals, genome sequencing revealed variants not identified on exome sequencing analysis due to lack of variant annotation or stringent filtering. For LD_0725, genome sequencing revealed a mitochondrial DNA variant, m.3243A>G in *MT-TL1* (OMIM:590050), carried on 42% of reads in the affected individual and 15% of the maternal reads. This variant has been previously associated with mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes (MELAS [OMIM:540000]). LD_0315 had biallelic variants in *DARS2* (OMIM:610956), associated with leukoencephalopathy with brainstem and spinal cord involvement and lactate elevation (LBSL [OMIM:611105]). An intronic variant, c.228-21delTinsCC, p.(Arg76Serfs*5) that has been shown to affect mRNA splicing of *DARS2* with skipping of exon 3[19] was found in *trans* with a previously unreported missense variant, c.294G>T, p.(Glu98Asp).







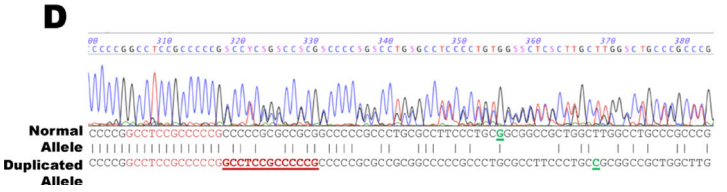


Figure 1. Myelin Disorders Bioregistry Project Unsolved Cohort and Challenging Variants in Known Leukodystrophy Genes. (A) The initial "unsolved" cohort consisted of 191 families. 101 were diagnosed by standard of care (SoC) methodologies, 71 of the remaining families underwent exome sequencing, while 19 were excluded due to lack of DNA or inability to access the family trio for NGS testing. The genome sequencing cohort consisted of 41 families, 14 of which were solved. (B) Variants in *GJC2* (OMIM:608803) were identified in LD_0617 and are causative of Pelizaeus-Merzbacher-like disease (OMIM:608804). The 13bp paternally inherited duplication, c.916_928dupGCCTCCGCCCCCG, p.Ala310fs in *GJC2*, was previously undetected by exome sequencing due to a lack of coverage in this region.[1] (C) The duplication lies in the middle of a 110 nucleotide block of a ~ 95% GC-rich region and was identified in the affected individual and father on genome sequencing, where it is indicated by a green rectangle. (D) Sanger chromatogram confirming the presence of the c.916_928dupGCCTCCGCCCCCG variant in the LD_0617 proband. The duplicated sequence is indicated in red with the second copy in LD_0617 bold and underlined.

Stringent variant filtering performed during the prior exome sequencing analysis excluded this intronic variant and precluded a diagnosis.[1] To confirm the predicted splicing change of intronic and exonic variants based on SpliceAI annotations,[20] we generated a minigene splicing reporter assay (Appendix S1), demonstrating each variant results in the skipping of exon 3 (Appendix S1).

The remaining five cases were resolved due to the identification of variants not typically identifiable by exome sequencing. The first is related to a region that is technically difficult to assess. Biallelic variants were identified in LD_0617 in *GJC2* (OMIM:608803), causative of Pelizaeus-Merzbacher-like disease (PMLD [OMIM:608804]), a maternally inherited missense variant c.203A>G, p.(Tyr68Cys), and a 13 base paternally inherited duplication c.916_928dupGCCTCCGCCCCCG, p.(Ala310fs). The duplication was previously undetected by exome sequencing due to a lack of coverage as it lies in the middle of a 110 nucleotide block of a ~ 95% GC-rich region (Fig. 1B).[1] The c.916_928dupGCCTCCGCCCCCG *GJC2* variant called by genome sequencing was supported by six of 19 genome sequencing reads covering the position (Fig. 1C), and was confirmed by Sanger sequencing (Fig. 1D).

In family LD_0393, genome sequencing revealed a homozygous intronic variant, c.1969+115_1969+116delAG in *CSF1R* (OMIM:164770). Heterozygous *CSF1R* variants are associated with another leukodystrophy, hereditary diffuse leukoencephalopathy with spheroids (OMIM:221820). This biallelic variant affects a cryptic splice site and results in the inclusion of a novel 99nt pseudo-exon, with an in-frame stop codon p.(Pro658-Serfs*24) resulting in nonsense-mediated decay of the truncated protein. These individuals were found to have a novel syndrome due to biallelic variants in *CSF1R* characterized by brain malformation and skeletal dysplasia, reported in Guo et al. (2019).[21]

Genome sequencing CNV analysis provided a diagnosis in three affected individuals (Table 1). LD_0498 was found to have impaired peroxisomal β-oxidation of C26:0 and genome sequencing revealed a homozygous, 7.5kb deletion in *ACBD5* (OMIM:616618) causing loss of exons 7 and 8, confirmed by Sanger sequencing.[22] LD_0594 was found to have an X-linked 396kb duplication covering *ARX* (OMIM: 300382), previously associated with a broad spectrum of neuroimaging abnormalities and early infantile epileptic encephalopathy (OMIM:308350). LD_0671 was found to harbor a 21kb de novo deletion encompassing a region in chromosome 2 covering *SATB2* (OMIM:608148), previously associated with Glass syndrome (OMIM:612313).[23]

Discussion

Of the original 191 families examined in this cohort, 101 received a definitive diagnosis using enzymatic, biochemical, or single gene molecular approaches (53%). Nineteen families were subsequently lost to follow-up, leaving 71 families able to be tested by NGS. Thirty cases were solved by exome sequencing,[1] and in this study, genome sequencing established a diagnosis in an additional 14 families yielding a 20% improvement, and an overall diagnostic efficacy of NGS of genome or exome sequencing increased to 44/71 (62%) and in total resolution of 145/172 (84%) of the original testable cohort.

Reanalysis by genome sequencing also led to the identification of variants in recently described disease genes *AIFM1*,[12] *HIKESHI*,[13] *H3F3B*,[11] *NKX6-2*,[14] and *SPATA5*,[15] and genes missed in the previous analysis (*L1CAM*). This is consistent with multiple recent studies that have shown increases of up to 10% in diagnostic yield with reanalysis within 24–36 months.[10] Improvements in variant annotation and filtering detected variants in mitochondrial DNA and intronic regions proximal to captured exons. It is expected that these would have been identified on exome sequencing if performed with current testing modalities.

Genome sequencing findings improved the overall diagnostic yield in this cohort by ~11% (5/44 cases solved by NGS), through the detection of variants in GC-rich regions, deep intronic variants, and CNVs. As sequencing assays and associated informatic pipelines improve, some of these variant types may be detected by exome sequencing. In our exome analysis, however, these variants were not found due to limitations of either sequencing or analysis, some of which will be difficult to overcome due to

Table 1. Clinically significant single-nucleotide, indel, and copy number variants.

Family ID	Gender	Current Age	Age at Initial MRI	Gene	Zygosity	Variant (GRCh37) & Transcript ID	cDNA	Protein	rsID	gnomAD AF[1]	Max AF[2]	SIFT/ Mutation Taster[3]	Variant Classification	ACMG Criteria[4]
Novel Gene (*n* = 4)														
LD_0246	M	8 Y	11 M	*H3F3B*	Het., *de novo*	17-73774722-G-C ENST00000254810	c.365C> G	Pro122Arg	N/A	0.0000	0.0000	N/A/ DC	Pathogenic	PS2, PM2, PP3
LD_0500	M	6 Y	23 M	*AIFM1*	Hemi., *de novo*	X-129274569-G-A ENST00000287295	c.720C> T	Asp240Asp	N/A	0.0000	0.0000	N/A/ N/A	Pathogenic	PS2, PS3, PM2
LD_0527	M	8 Y	UNK	*NKX6-2*	Hom.	10-134598648-C-TA ENST00000368592	c.606delGinsTA	Lys202fs	N/A	0.0000	0.0000	N/A/ N/A	Likely pathogenic	PVS1, PM2
LD_0807	F	17 Y	12 Y	*SNORD118*	Het. (m)	17-8076848-A-C ENST00000363593	n.58 T> G	N/A	N/A	0.0000	0.0000	N/A/ N/A	VUS	PM2, PM3, PP4, BP6
					Het. (p)	17-8076762-G-A ENST00000363593	n. *9C> T	N/A	N/A	0.001907	0.001907	N/A/ N/A	VUS	PM2, PM3, PP4, BP6
Intronic (*n* = 1)														
LD_0393	F (3) & M (1)	N/A	N/A	*CSF1R*	Hom.	5-149440309-CT-ENST00000286301	c.1969 + 115_1969+ 116delAG[5]	Pro658Serfs*24	N/A	0.0000	0.0000	N/A/ N/A	Likely pathogenic	PS3, PM2, PP1
MT-DNA (*n* = 1)														
LD_0725	F	21 Y	1 Y	*MT-TL1*	Het. (m)	MT-3243-A-G ENST00000386347	n.14A> G	N/A	rs199474657	N/A	N/A	N/A/ N/A	Pathogenic	PS1, PS3, PP3, PP4
Challenging Variants (*n* = 2)														
LD_0315	M	7 Y	19 M	*DARS2*	Het. (p)	1-173797450-T-CC ENST00000361951	c.228-21delTinsCC	N/A	rs528772984	0.000021	0.002	N/A/ N/A	Likely pathogenic	PS3, PM1
					Het. (m)	1-173797537-G-T ENST00000361951	c.294G>T[6]	Glu98Asp	N/A	0.0000	0.0000	T/ DC	Likely pathogenic	PS3, PM2, PP2, PP3
LD_0617	F	7 Y	2 Y	*GJC2*	Het. (p)	1-228346388-C-G CCTCCGCCCCCG ENST00000366714	c.916_ 928dupGCCT CCGCCCCCG	Ala310fs	N/A	0.0000	0.0000	N/A/ N/A	Pathogenic	PVS1, PM2, PP4
					Het. (m)	1- 228345662-A-G ENST00000366714	c.203A> G	Tyr68Cys	rs1031720654	0.0000	0.0000	D/ DC	Likely pathogenic	PM2, PM3, PP3, PP4
Improved Pipeline/ Separate study novel gene (*n* = 4)														
LD_0162	M & M	Deceased	3 Y	*HIKESHI*	Hom.	11-86017416-G-C ENST0000027843	c.160G> C	Val54Leu	rs202003795	0.0001299	0.0002	D/ DC	Likely pathogenic	PS1, PM1, PP1, PP4, PP5
LD_0579	M	5 Y	5 M	*L1CAM*	Hemi. (m)	X-153135278-A-T ENST00000264161	c.1103T> A	Ile368Asn	N/A	0.0000	0.0000	D/ DC	Likely pathogenic	PM1, PM2, PP2, PP3
LD_0808	M	6 Y	18 M	*SPATA5*	Het. (p)	4-123855300-G-A ENST00000274008	c.554G> A	Gly185Glu	rs753587518	0.0000253	0.0002	D/ DC	Likely pathogenic	PM2,PM3, PP2, PP3, PP4
					Het. (m)	4-123949348-G-C ENST00000274008	c.1877G> C	Trp626Ser	rs375343753	0.0000122	0.0000	D/ DC	Likely pathogenic	PM2, PM5, PP3, PP4, PP5
Clinically Resolved														
LD_0587	M	9 Y	9 M	*KDM5C*	Hemi. (m)	X-53226186-C-A ENST00000375401	c.2663G> T	Arg888Leu	rs375850872	0.0000144	0.0001	D/ P	VUS	PM2, PP3

(*Continued*)

Table 1. (Continued)

Copy Number Variants									
Family ID	Gender	Current Age	Age at initial MRI	CNV Coordinates	Size (in nucleotides)	CNV Type	Zygosity	Genes Covered	Known Syndrome (OMIM)
LD_0498	F	9 Y	4 Y	Chr10:27500219-27507820 Precise breakpoints as determined by Sanger: c.626-689_937-235delins936 + 1075_936+1230inv	7,601	Deletion	Homozygous	*ACBD5*	N/A
LD_0594	M	11 Y	2 M	ChrX:24642417-25038408	395,991	Duplication	Het., maternally inherited	*ARX, POLA1, PCYT1B*	N/A
LD_0671	M	5 Y	18 M	Chr2:200222207-200243660	21,453	Deletion	De novo	*SAT82*	Glass Syndrome (OMIM:612313)

Key: M – Month; Y – Years; UNK – Unknown; MT– mitochondrial; Het. – heterozygous; Hemi. – Hemizygous; Hom. - Homozygous; N/A – Not available; AF – allele frequency.
[1]Total allele frequency in Genome Aggregation Database (gnomAD) dataset.
[2]Maximum reported allele frequency for any single subpopulation represented in ExAC, 1000 Genomes, or Exome Variant Server datasets.
[3]DC – Disease-Causing; DA – Disease-Causing Automatic; D – Damaging; T – Tolerated
[4]ACMG criteria upon which classification is based
[5]Variant shown in Guo et al. 2019 to result in inclusion of a novel 99nt pseudo-exon, with an in-frame stop codon p.(P658Sfs*24) resulting in nonsense-mediated decay.
[6]Variant shown to result in *DARS2* Exon 3 skipping. See Appendix S1.

limitations of the assay (e.g. GC-rich regions) or the ability of exome to accurately detect large genomic alterations in single cases.

Our study lends support to a previous study that suggested that currently up to 80% of white matter disorders may be able to be solved.[2] Notably, no recurring diagnoses were made in the current cohort of $n = 41$, and only a minority across all tiers of NGS testing (20%) were associated with canonical leukodystrophy genes. These data indicate that only a fraction of affected individuals would achieve a diagnosis using targeted testing approaches, including panels of classically defined leukodystrophy genes now in widespread use. An established diagnostic workflow for laboratory and genetic testing in combination with MRI pattern recognition[24] suggests that if initial biochemical and enzymatic testing is not confirmatory and if the MRI does not fit an established leukodystrophy, broad-based NGS testing should be implemented. Our data support this approach, favoring genome sequencing over exome sequencing.

Acknowledgments

We thank the patients and their families. AP, JLS, AT, and AV are supported by the Myelin Disorders Bioregistry Project. GB has received the New Investigator Salary Award from the Canadian Institutes of Health Research (2017-2022). Sequencing was provided by Illumina, Inc. The Myelin Disorders Bioregistry Project was supported by the Children's Research Institute at the Children's Hospital of Philadelphia. This study was in part financed by the Australian National Health and Medical Research Council (NHMRC 1068278). The research conducted at the Murdoch Children's Research Institute was supported by the Victorian Government's Operational Infrastructure Support Program. We thank Elizabeth Bhoj and Yanick J. Crow for their consultation and work with regard to cases undergoing genome sequencing. We also thank Brian P. Brooks, Sara Zondag, Lisa Green, Sudeshna Mitra, Lucy Civitello, Natasha Shur, Valeria Zincke, Silvia Delgado, Janice E. Brunstrom-Hernandez, Celia Chang, Robert Keating, Jessica Carpenter, Jayne Antony, Shekeeb Mohammad, Marc C. Patterson, Tarannum Lateef, Taeun Chang, James Reese, Shaaron Towns, Diego Preciado, Dewi Depositario-Cabacar, Meganne Leach, Catherine Zorc, Jenny Wilson, Eileen Walters, Steven Leber, Srikanth Muppidi, Kimberly Chapman, Amy Waldman, Lindsey Scussel, Andrea Gropman, Tena Rosser, Phillip Pearl, Eva Fung, Sumit Parikh, Bruce Cohen, James Reggin, Cengiz Yalcinkaya, Yuval Shafrir, Mark DiFazio, Emily Freilich, Charles Lourenco, Carolina Tesi-Rocha, Brent Fogel, Jay Desai, Hernan Amartino, Nicole Weaver, Valynne Long, Michael Gambello, Melissa Cirillo, Ilana Kahn, Deepak Gill, Maria Gieron, Emily de

Los Reyes, Bennet Lavenstein, Brendan Lanpher, Gerhard Kurleman, Dan Miller, Amirah Khouzam, Vani Rajan, Erin Ramos, Shimul Chowdhury, Tina Hambuch, Kelin Ru, Greg Baillie, Sean Grimmond, Ljuba Caldovic, Joe Devaney Miriam Bloom, Sarah Evans, Jennifer Murphy, and Nathan McNeill for providing care to the affected individuals and families, referral to the Myelin Disorders Bioregistry Project, or previous contributions to this longstanding project.

Conflict of Interest

BRL, ED, AG, VG, and RJT are/were employees of Illumina, Inc. Nemours receives revenue from diagnostic testing performed in the Nemours Molecular Diagnostics Laboratory. GB has received compensation for traveling to meetings and advisory boards from Ionis, Shire/Takeda, Actelion Pharmaceuticals and Children's Hospital of Philadelphia. She served on the scientific advisory board for Ionis (2019) and has received research grants from Shire/Takeda and Bluebird Bio. AV receives support from Shire, Gilead, Eli Lilly and Illumina for research activities. Otherwise the authors report no conflict of interest.

Author Contributions

GH, BRL, CS, RJT, and AV designed and managed the project and wrote the manuscript. JC, AT, MW, EMJ, YJC, SF, HRW, ID, EB, NM, NIW, TEMA, SMK, CT, GMH, LG, and SI provided functional analysis of variants and confirmation of variants and contributed to figures. BRL, SJB, CS, and RJT contributed bioinformatics expertise. GH, BRL, AP, JLS, GB, RS, MSvdK, GB, CS, RJT, and AV performed analysis of or provided consultation for cases from the manuscript. All authors reviewed the manuscript.

References

1. Vanderver A, Simons C, Helman G, et al. Whole exome sequencing in patients with white matter abnormalities. Ann Neurol 2016;79:1031–1037.
2. Kevelam SH, Steenweg ME, Srivastava S, et al. Update on leukodystrophies: a historical perspective and adapted definition. Neuropediatrics 2016;47:349–354.
3. van der Knaap MS, Bugiani M. Leukodystrophies: a proposed classification system based on pathological changes and pathogenetic mechanisms. Acta Neuropathol 2017;134:351–382.
4. Roller E, Ivakhno S, Lee S, et al. Canvas: versatile and scalable detection of copy number variants. Bioinformatics 2016;32:2375–2377.
5. Dolzhenko E, van Vugt J, Shaw RJ, et al. Detection of long repeat expansions from PCR-free whole-genome sequence data. Genome Res 2017;27:1895–1903.
6. Van der Auwera GA, Carneiro MO, Hartl C, et al. FastQ data to high confidence variant calls: the genome analysis toolkit best practices pipeline. Curr Protoc Bioinformatics 2013;43:11.0.1-33.
7. Cingolani P, Platts A, le Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin) 2012;6:80–92.
8. Gaildrat P, Killian A, Martins A, et al. Use of splicing reporter minigene assay to evaluate the effect on splicing of unclassified genetic variants. Methods Mol Biol 2010;653:249–257.
9. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 2015;17:405–424.
10. Costain G, Jobling R, Walker S, et al. Periodic reanalysis of whole-genome sequencing data enhances the diagnostic advantage over standard clinical genetic testing. Eur J Hum Genet 2018;26:740–744.
11. Farwell Hagman KD, Shinde DN, Mroske C, et al. Candidate-gene criteria for clinical reporting: diagnostic exome sequencing identifies altered candidate genes among 8% of patients with undiagnosed diseases. Genet Med 2017;19:224–235.
12. Miyake N, Wolf NI, Cayami FK, et al. X-linked hypomyelination with spondylometaphyseal dysplasia (H-SMD) associated with mutations in AIFM1. Neurogenetics 2017;18:185–194.
13. Edvardson S, Kose S, Jalas C, et al. Leukoencephalopathy and early death associated with an Ashkenazi-Jewish founder mutation in the Hikeshi gene. J Med Genet 2016;53:132–137.
14. Dorboz I, Aiello C, Simons C, et al. Biallelic mutations in the homeodomain of NKX6-2 underlie a severe hypomyelinating leukodystrophy. Brain 2017;140:2550–2556.
15. Tanaka AJ, Cho MT, Millan F, et al. Mutations in SPATA5 Are Associated with microcephaly, intellectual disability, seizures, and hearing loss. Am J Hum Genet 2015;97:457–464.
16. Tsurusaki Y, Okamoto N, Suzuki Y, et al. Exome sequencing of two patients in a family with atypical X-linked leukodystrophy. Clin Genet 2011;80:161–166.
17. Vallianatos CN, Farrehi C, Friez MJ, et al. Altered gene-regulatory function of KDM5C by a novel mutation associated with autism and intellectual disability. Front Mol Neurosci 2018;11:104.
18. Jenkinson EM, Rodero MP, Kasher PR, et al. Mutations in SNORD118 cause the cerebral microangiopathy leukoencephalopathy with calcifications and cysts. Nat Genet 2016;48:1185–1192.

19. van Berge L, Hamilton EM, Linnankivi T, et al. Leukoencephalopathy with brainstem and spinal cord involvement and lactate elevation: clinical and genetic characterization and target for therapy. Brain 2014;137(Pt 4):1019–1029.
20. Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, et al. Predicting splicing from primary sequence with deep learning. Cell 2019;176:535–548.e24.
21. Guo L, Bertola DR, Takanohashi A, et al. Bi-allelic CSF1R mutations cause skeletal dysplasia of dysosteosclerosis-pyle disease spectrum and degenerative encephalopathy with brain malformation. Am J Hum Genet 2019;104:925–935.
22. Ferdinandusse S, Falkenberg KD, Koster J, et al. ACBD5 deficiency causes a defect in peroxisomal very long-chain fatty acid metabolism. J Med Genet 2016;54:330–337.
23. Zarate YA, Smith-Hicks CL, Greene C, et al. Natural history and genotype-phenotype correlations in 72 individuals with SATB2-associated syndrome. Am J Med Genet A 2018;176:925–935.
24. Parikh S, Bernard G, Leventer RJ, et al. A clinical approach to the diagnosis of patients with leukodystrophies and genetic leukoencephelopathies. Mol Genet Metab 2015;114:501–515.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Appendix S1. A supplemental text PDF file is provided which includes methods and results for the minigene splicing reporter assay used in the case of LD_0315 (Section 1) and sequencing coverage achieved for proband samples (Table S2). Case reports are for each of the 41 families with diagnoses associated with single-nucleotide variants and small insertion-deletion events (Section 2), copy number variants (Section 3), or remaining unsolved after genome sequencing and analysis (Section 3). The supplemental text of the original study can be found at: https://imb.uq.edu.au/download/Vanderver_AON_2016.case_reports.pdf

