Genetics in Medicine (2024) 26, 101076

ELSEVIER

Genetics in Medicine
An Official Journal of the ACMG

www.journals.elsevier.com/genetics-in-medicine

ARTICLE

Narrowing the diagnostic gap: Genomes, episignatures, long-read sequencing, and health economic analyses in an exome-negative intellectual disability cohort

Kerith-Rae Dias[1,2], Rupendra Shrestha[3], Deborah Schofield[3], Carey-Anne Evans[1,4], Emily O'Heir[5], Ying Zhu[1,4,6], Futao Zhang[1,4], Krystle Standen[4], Ben Weisburd[5], Sarah L. Stenton[5], Alba Sanchis-Juan[5], Harrison Brand[5], Michael E. Talkowski[5], Alan Ma[7,8], Sondy Ghedia[9,10], Meredith Wilson[7], Sarah A. Sandaradura[7,11], Janine Smith[7,8], Benjamin Kamien[12,13], Anne Turner[14], Madhura Bakshi[15], Lesley C. Adès[7,11], David Mowat[14,16], Matthew Regan[17], George McGillivray[18], Ravi Savarirayan[18,19,20], Susan M. White[18,20], Tiong Yang Tan[18,19,20], Zornitza Stark[18,20,21], Natasha J. Brown[18,19,20], Luis A. Pérez-Jurado[22,23], Emma Krzesinski[17,24], Matthew F. Hunter[17,24], Lauren Akesson[25,26,27], Andrew Paul Fennell[17,24], Alison Yeung[18,19,20], Tiffany Boughtwood[19,21], Lisa J. Ewans[14,16,28], Jennifer Kerkhof[29], Christopher Lucas[4], Louise Carey[4], Hugh French[30], Melissa Rapadas[28,31], Igor Stevanovski[28,31], Ira W. Deveson[28,31,32], Corrina Cliffe[4], George Elakis[4], Edwin P. Kirk[4,14,16], Tracy Dudding-Byth[6], Janice Fletcher[4], Rebecca Walsh[4], Mark A. Corbett[33], Thessa Kroes[33], Jozef Gecz[33,34], Cliff Meldrum[35], Simon Cliffe[35], Meg Wall[18], Sebastian Lunke[18], Kathryn North[19,20,21,36], David J. Amor[19,20], Michael Field[6], Bekim Sadikovic[29,37], Michael F. Buckley[4], Anne O'Donnell-Luria[5,38], Tony Roscioli[1,2,4,*]

ARTICLE INFO

Article history:
Received 30 June 2023
Received in revised form
12 January 2024
Accepted 16 January 2024
Available online 19 January 2024

ABSTRACT

Purpose: Genome sequencing (GS)–specific diagnostic rates in prospective tightly ascertained exome sequencing (ES)–negative intellectual disability (ID) cohorts have not been reported extensively.

Methods: ES, GS, epigenetic signatures, and long-read sequencing diagnoses were assessed in 74 trios with at least moderate ID.

Results: The ES diagnostic yield was 42 of 74 (57%). GS diagnoses were made in 9 of 32 (28%) ES-unresolved families. Repeated ES with a contemporary pipeline on the GS-diagnosed families identified 8 of 9 single-nucleotide variations/copy-number variations undetected in

*Correspondence and requests for materials should be addressed to Tony Roscioli, New South Wales Health Pathology Randwick Genomics, 4th Floor Campus Centre, Barker Street, Randwick, Sydney, NSW 2031, Australia. *Email address:* tony.roscioli@health.nsw.gov.au

Affiliations are at the end of the document.

doi: https://doi.org/10.1016/j.gim.2024.101076

Keywords:
Episignature
Exome negative
Genome sequencing
Health economics
Intellectual disability

older ES, confirming a GS-unique diagnostic rate of 1 in 32 (3%). Episignatures contributed diagnostic information in 9% with GS corroboration in 1 of 32 (3%) and diagnostic clues in 2 of 32 (6%). A genetic etiology for ID was detected in 51 of 74 (69%) families. Twelve candidate disease genes were identified. Contemporary ES followed by GS cost US$4976 (95% CI: $3704; $6969) per diagnosis and first-line GS at a cost of $7062 (95% CI: $6210; $8475) per diagnosis.

Conclusion: Performing GS only in ID trios would be cost equivalent to ES if GS were available at $2435, about a 60% reduction from current prices. This study demonstrates that first-line GS achieves higher diagnostic rate than contemporary ES but at a higher cost.

Introduction

Intellectual disability (ID) affects 3% of the population, of which 15% are moderate to severe and enriched for Mendelian etiologies.[1] Despite significant molecular diagnostic improvements, approximately half of the individuals with ID remain undiagnosed after exome sequencing (ES).[2] Genome sequencing (GS) increases coding region coverage consistency, has improved detection sensitivity for structural variants (SVs), including copy-number variants (CNVs) and tandem repeat expansions (TREs), and facilitates the detection of noncoding and mitochondrial variation.[3] Short-read sequencing (SRS) trio GS as a first-tier stand-alone test has been proposed to maximize diagnoses in Mendelian disorders when cost is not a consideration.[4]

Diagnostic testing costs from published and commercial sources range from US$3825 to $9304 (trio ES)[5] and from US$4963 to $15,000 (trio GS).[6] Despite the large resources required for the care of people with ID, there is currently limited health economic evidence to support ES versus GS. Studies that evaluate diagnostic rates, costs, sensitivity, and cost-effectiveness of these tests are integral to their implementation into clinical practice.[5] These were therefore assessed in this prospective, tightly ascertained, moderate to severe ID cohort using ES, GS, episignatures, and long-read sequencing (LRS). Contemporary ES was repeated on families diagnosed by GS where the original ES was reported as uninformative to investigate the current diagnostic potential for ID unique to GS. A cost for GS trios as a first-line test was also estimated based on the incremental diagnostic rate of GS over contemporary ES in ID from this study.

Materials and Methods

Cohort ascertainment

Seventy-four trios (proband and unaffected parents) were recruited to this study (Figure 1) funded through Australian Genomics, the Center for Mendelian Genomics (Broad CMG), and New South Wales Statewide Genomic Service. Families were referred from Australian hospitals to the New South Wales Health Pathology Randwick Genomics (RG) and Victorian Clinical Genetics Services (VCGS) laboratories. Probands with moderate or severe ID, noncontributory microarrays, and *FMR1* molecular testing were included. Individuals with autism spectrum disorder or prior ES were excluded. All individuals had genetic counseling, and their parents/guardians consented for diagnostic ES and research GS. The study was approved by HREC/16/MH/251.

ES analysis

For original ES, DNA analyzed at RG was extracted from peripheral blood samples and libraries using the Ion AmpliSeq Exome RDY kit (Thermo Fisher Scientific), sequence on the Ion PI chip kit (v3) and analyzed on a Life Technologies Proton instrument. Reads were aligned to human genome build hg19/GRCh37 and SNVs and indels were identified using TorrentSuite v5.10.1 software and analyzed as reported previously.[7] The Agilent SureSelect QXT CREv2 kit (Agilent Technologies) was at VCGS for library preparation, sequenced on an Illumina NextSeq500 (Illumina), and variants were analyzed using the Illumina DRAGEN Bio-IT Platform after alignment to hg19/GRCh37. The original ES analyses did not include CNV pipelines. Contemporary ES analysis was performed on selected families at RG to determine GS-unique diagnoses, using library building Exome 2.0 kits (Twist Bioscience) with sequencing on an Illumina NovaSeq 6000 (Illumina), mapped to hg38/GRCh38 and analyzed using the Illumina DRAGEN Bio-IT Platform. SNVs and indels were classified as previously reported,[8] utilizing the American College of Medical Genetics and Genomics and the Association for Molecular Pathology (ACMG-AMP) guidelines with modifications and incorporating aspects of the scoring system reported by Karbassi et al.[9] Genomic variants were filtered in the Genomic Annotation and Interpretation Application pipeline[8] at RG. A commercial genomic analysis pipeline, Alissa Clinical Informatics Platform (Agilent Technologies), was used for genomic analysis at VCGS. Potential CNVs were identified using DECoN,[10] XHMM,[11] the Illumina

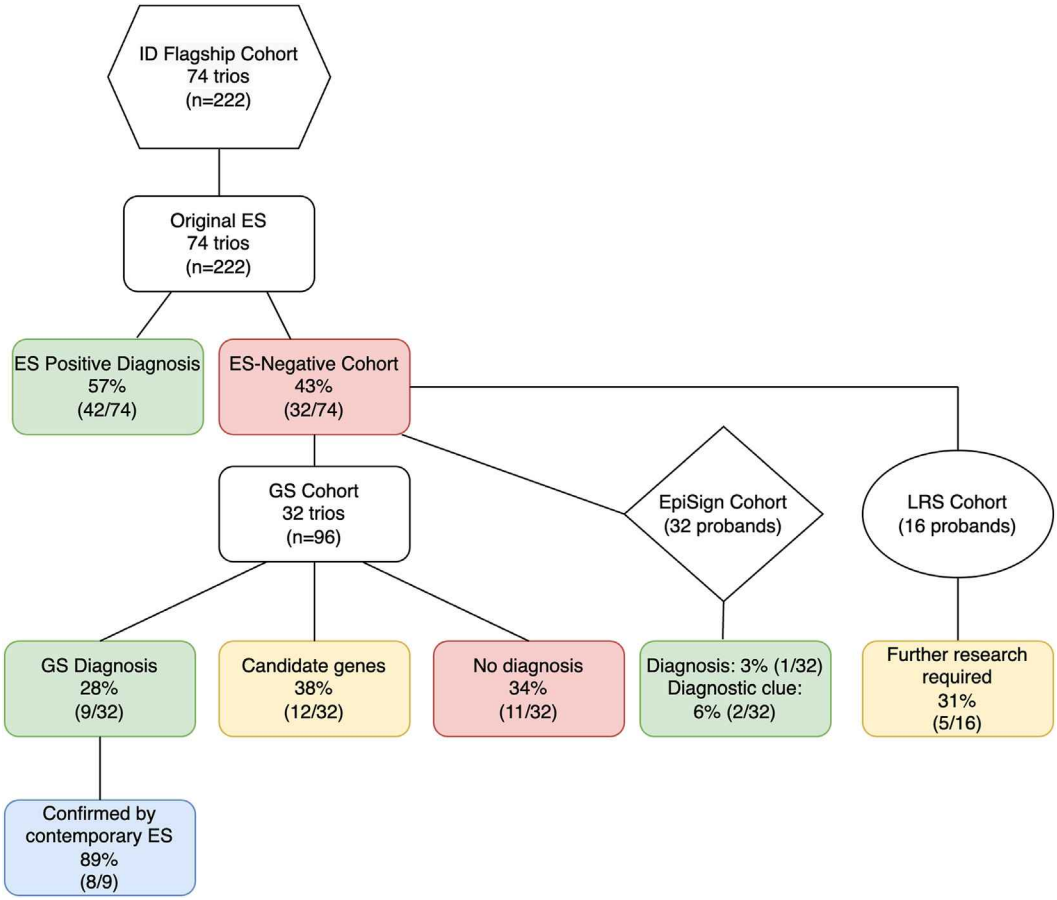

Figure 1 ID study. Seventy-four trios with molecularly undiagnosed ID investigated using ES (74 trios), GS (32 trios), EpiSign (32 probands), and LRS (16 probands). The overall ES/GS diagnostic yield is 69% (51/74) with 31% (23/74) unresolved. ID, intellectual disability; ES, exome sequencing, GS, genome sequencing; LRS, long read sequencing.

DRAGEN Bio-IT Platform at RG, and an internal CNV detection tool CxGo[12] at VCGS.

GS analysis

GS was performed on ES-negative families at the Broad CMG. DNA libraries were prepared using the PCR-free KAPA Hyper Prep kit without amplification module (KK8505; KAPA Biosystems) with palindromic forked adaptors (Roche) and run on the NovaSeq 6000 (Illumina). Reads were mapped to hg38/GRCh38 using BWA aligner and analyzed using the Genome Analysis Toolkit (GATK) HaplotypeCaller package (v4.0). GS data were analyzed at RG using Genomic Annotation and Interpretation Application, aligned to GRCh38. SV, CNV and the mitochondrial genome were analyzed by a number of different platforms in the study laboratories (see Supplemental Methods). GS data were analyzed by the Broad CMG using the open-source genomic analysis platform seqr.[13] Clinical phenotypes correlated with gene names were used in PubMed searches (https://pubmed.ncbi.nlm.nih.gov/). Diagnostic outcomes were assessed based on ACMG-AMP criteria. Candidate research genes were deposited into Matchmaker Exchange through seqr[13] or GeneMatcher.[14]

Episignature analysis

Episignature analysis was conducted on ES-negative families with available DNA using EpiSign, a genome-wide DNA methylation test. DNA was extracted from peripheral blood and processed using the Illumina Infinium EPIC bead chip array (Illumina) as described previously.[15] Data were analyzed using EpiSign (V3) at the Molecular Diagnostics Laboratory, London Health Sciences Centre, Western University, London, Canada.[16]

Long-read sequencing analysis

DNA was subject to LRS using Oxford Nanopore Technologies (ONT) at the Garvan Institute Sequencing Platform. Libraries were prepared from high molecular weight DNA using native ligation-based library kit (SQK-LSK110) and sequenced on PromethION R9.4.1 flowcells at 1 sample

per flowcell. ONT sequencing data were converted from FAST5 to BLOW5 format[17] using slow5tools (v0.3.0).[18] Data were base called using *Guppy*. Targeted genotyping was performed on 48 Short Tandem Repeat (STR) sites associated with neurological disease[19] using a local haplotype aware ONT read assembly spanning each STR site, annotating STR size, motif and other summary statistics using Tandem Repeats Finder (4.09) followed by manual inspection and motif counting.[19]

Health economic analysis

Health economic analysis was undertaken from the health care funder perspective at the time of diagnosis to evaluate the cost-effectiveness of performing GS over ES under 3 different models: (1) performing GS on ES-negative families, (2) GS-unique diagnoses after contemporary ES, and (3) performing GS on all families as a first-line test. Laboratory costs for ES and GS were sourced from RG and VCGS. These were fixed costs at publicly reimbursement rates in Australia and thus have no uncertainty in the unit costs. All costs are Australian prices unless otherwise stated and are converted to US dollars, based on the exchange rate of AU$ 1 = US$ 0.6577 of June 1, 2023 (https://www.xe.com/currencytables/). Unit costs are listed in Supplemental Table 1. The cost-effectiveness of GS relative to ES or contemporary ES was analyzed for the incremental cost per additional diagnosis under 3 diagnostic pathways with uncertainty assessed using a bootstrapping method. This generated 1000 replicated data sets, randomly drawn with replacement. Incremental costs per additional diagnosis were estimated for each data set, and a 95% confidence interval was estimated as an uncertainty related to incremental cost from the distribution of incremental cost based on the percentile method. The price reduction of GS was estimated such that it produced the same cost per diagnosis as contemporary ES in ID. Because this study was performed to cost the diagnostic rate of different technologies, the base results included the costs that were within the scope of laboratory testing, including sample collection, sequencing, and data analysis but not including the costs of clinical review. A sensitivity analysis was conducted assuming 2 clinical geneticist/genetic counselor visits, 1 at the beginning, and 1 at follow-up.

Results

Clinical characteristics of the cohort

This cohort included 74 probands (41 male, 33 female) with a median age of 15 years (6–43 years). ID severity was determined by developmental assessments as moderate (35), severe (37) and profound (2). Nonfamilial facial features were present in 19 of 74 (26%), 12 of 74 (16%) had microcephaly, and 10 of 74 (13%) had seizures (Supplemental Table 2).

ES diagnostic yield

Trio ES analysis was performed in 74 families. Pathogenic or likely pathogenic variants were identified in 42 of 74 families (57%). These diagnoses included 32 SNVs and 10 indels (Supplemental Table 3). No diagnosis was obtained in 43% (32/74).

GS diagnostic yield compared with original ES

The ES-negative cases (n = 96; 32 families) had trio GS at the Broad CMG to investigate the GS-specific diagnostic rate. This analysis resulted in an additional diagnosis for 9 families (9/74; 12%), which included 3 SNVs, 1 indel, and 5 CNVs (Supplemental Figure 1, Table 1[20-22], Supplemental Table 4). Mitochondrial variant, promoter, and TRE analyses did not detect any additional diagnoses. Twelve candidate genes were implicated in ID (Supplemental Table 5). Additional evidence of pathogenicity through collaborative studies facilitated through Matchmaker Exchange via seqr[13] or GeneMatcher[14] resulted in 2 families (3%) with possible diagnoses. Twenty-three families (31%) remained undiagnosed after GS testing.

GS-unique diagnostic yield compared with repeated contemporary ES

Contemporary ES reanalysis including CNV detection was performed at RG to determine the GS-unique diagnostic rate for ID. Three cases with SNVs and 5 cases with 1 indel and 4 CNVs detected by GS were also detected by contemporary ES analysis (Supplemental Figure 2, Supplemental Table 6). Compound heterozygous *YARS1* variants detected in proband 47 in GS were present in prior ES data but were discarded as the only phenotype in OMIM at the time was Charcot-Marie-Tooth syndrome without correlation to neurocognitive disease. A de novo heterozygous SNV in *NR2F1* in proband 43 and a de novo heterozygous indel in *POGZ* in proband 55 were both present in prior ES data, but they were assessed previously as false positives. A de novo heterozygous variant in *CELF2* predicted to result in a missense change was identified by GS and contemporary ES in proband 56 but was not a known Mendelian gene at the time of the original ES. All variants were absent from gnomAD and had in silico scores supportive of pathogenicity. Likewise, 4 of the 5 CNVs detected by GS and not detected in the original ES were also detected by contemporary ES using CNV callers. These included a de novo heterozygous partial *MED13L* 26.6 kb out-of-frame tandem duplication (NM_015335.5, exons 5 - 16 of 31) in proband 44, a paternal mosaic heterozygous partial deletion of *SCN2A* in proband 54, a maternally inherited hemizygous *UPF3B* deletion in proband 58, and a de novo heterozygous *EHMT1* partial duplication in proband 62.

Table 1 Genome sequencing, episignature and long read sequencing findings in ID Flagship cohort

Proband	GS Gene (HGNC ID)	HGVS Nomenclature		GS ACMG Criteria	GS ACMG Interpretation	Episignature	LRS
43	*NR2F1*(7975)	NC_000005.10:g.93585289G>A	NC_000005.10(NM_005654.6):c.266G>A NP_005645.1:p.(Cys89Tyr)	PS2, PM1, PM2, PP3	LP	ND	NT
44	*MED13L* (22962)	NC_000012.12:g.115991285_116027907dup	NC_000012.12(NM_015335.5): c.480-5305_3670dup NP_056150.1:p.(Ser1224Asnfs*8)	PVS1, PS2	P	BWS	ND
45	*XRN1* (30654) (research)	NC_000003.12:g.142425225G>T	NC_000003.12(NM_001282857.2):c.624C>A XP_016862130.1:p.(Asp208Glu)	NA	NA	ND	ND
46	*SF3B1* (10768) (research)	NC_000002.12:g.197402101T>G	NC_000002.12(NM_012433.4):c.2107A>C NP_036565.2:p.(Thr703Pro)	NA	NA	ND	ND
47	*YARS1* (12840)	NC_000001.11:g.32779421_32779423del	NC_000001.11(NM_003680.4):c.1438_1440del NP_003671.1:p.(Glu480del)	PS2, PM4	LP	ND	ND
		NC_000001.11:g.32787071G>A	NC_000001.11(NM_003680.4):c.689C>T NP_003671.1:p.(Ser230Phe)	PS2, PM2, PP3			
48						ND	ND
49	*UBR5* (16806) (research)	NC_000008.11:g.102288189C>A	NC_000008.11(NM_015902.6):c.4791G>T NP_056986.2:p.(Glu1597Asp)	NA	NA	ND	ND
50	*DYRK1A* (3091)	NC_000021.9:g.37512175A>G	NC_000021.9(NM_001347721.2):c.1909A>G NP_001334650.1:p.(Met637Val)	PS2, PP2, BS2	VUS	ND	ND
	POLR3B (30348)	NC_000012.12:g.106433856G>A	NC_000012.12(NM_018082.6):c.1765G>A NP_060552.4:p.(Gly589Arg)	PS2, PM1, PM2, PP2, PP3	LP (no second allele)		
51						ND	ND
52						ND	ND
53						ND	ND
54	*SCN2A* (10588)	NC_000002.12:g.165312617_165332618del	NC_000002.12(NM_001371246.1): c.1034+529_2388+1050del NP_001358175.1:p.(Gln346Serfs*14)	PVS1, PS2, PM2	P	ND	ND
55	*POGZ* (18801)	NC_000001.11:g.151405203del	NC_000001.11(NM_015100.4):c.3837del NP_055915.2:p.(Lys1279Asnfs*31)	PVS1, PS2, PM2	P	ND	Not analyzed because of low coverage
56	*CELF2* (2550) (research)	NC_000010.11:g.11165632A>G	NC_000010.11(NM_001326342.2): c.221A>G NP_001313271.1:p.(Tyr74Cys)	PS2, PM2	Research originally, now published[a]	ND	ND
57						ND	ND
58	*UPF3B* (20439)	NC_000023.11:g.119808858_119816662del	NC_000023.11(NM_080632.3): c.*18216_*26020del NP_542199.1:p.(?)	PVS1, PM4	LP	ND	Not analyzed because of low coverage
59						ND	NT
60						ND	NT

(continued)

Table 1 Continued

Proband	GS Gene (HGNC ID)	HGVS Nomenclature		GS ACMG Criteria	GS ACMG Interpretation	Episignature	LRS
61	*MCM6* (6949) (research)	NC_000002.12:g.135868621T>C	NC_000002.12(NM_005915.6):c.605A>G NP_005906.2:p.(Asp202Gly)	NA	Research originally, now published[b]	NT	NT
62	*EHMT1* (24650)	NC_000009.12:g.137699442_137754877dup	NC_000009.12(NM_024757.5): c.22-11525_1369+586dup NP_079033.4:p.(?)	PVS1, PS2	P	Kleefstra	NT
63	*KANSL3* (25473) (research)	NC_000002.12:g.96609039C>A	NC_000002.12(NM_001115016.3): c.1409G>T NP_001108488.1:p.(Gly470Val)	NA	NA	ND	ND
	KANSL3 (25473) (research)	NC_000002.12:g.96613488T>C	NC_000002.12(NM_001115016.3):c.795A>G NP_001108488.1:p.(Pro265=)	NA	NA		
	FBP2 (3607) (research)	NC_000009.12:g.94593659C>T	NC_000009.12(NM_003837.4):c.68G>A NP_003828.2:p.(Arg23His)	NA	Research originally, now published[c]		
64						ND	NT
65						ND	NT
66						ND	NT
67	*PPP2R5C* (9311) (research)	NC_000014.9:g.101882257T>C	NC_000014.9(NM_001352913.2):c.556T>C NP_001339842.1:p.(Trp186Arg)	NA	NA	ND	NT
68	*ZBTB34* (31446) (research)	NC_000009.12:g.126879889_126879890del	NC_000009.12(NM_001099270.4):c.490_491del NP_001092740.2:p.(Pro164Serfs*19)	NA	NA	ND	NT
	NFE2L3 (7783) (research)	NC_000007.14:g.26152577del	NC_000007.14(NM_004289.7):c.79del NP_004280.5:p.(Arg27Alafs*2)	NA	NA		
	NFE2L3 (7783) (research)	NC_000007.14:g.26152931G>T	NC_000007.14(NM_004289.7):c.433G>T NP_004280.5:p.(Gly145Cys)	NA	NA		
69						UPD7pat /*MRXSCJ*	NT
70	*SHANK2* (14295) / near *CREB3L1* (18856)	NC_000011.10:g.46301244_70530165inv	NC_000011.10(NM_012309.5): c.2062-27234_*24171625inv NP_036441.2:p.(?)	PVS1, PS2	P	ND	NT
71						ND	NT
72	*ARFGEF3* (21213) (research)	NC_000006.12:g.138298628_138298629del	NC_000006.12(NM_020340.5):c.3671_3672del NP_065073.3:p.(His1224Argfs*11)	NA	NA	ND	NT
73						ND	NT
74	*MAP2K4* (6844) (research)	NC_000017.11:g.12129194A>T	NC_000017.11(NM_003010.4):c.947A>T NP_003001.1:p.(Gln316Leu)	NA	NA	ND	NT

ACMG, American College of Medical Genetics; *del*, deletion; *GS*, genome sequencing; *dup*, duplication; *ID*, intellectual disability; *inv*, inversion; *LP*, likely pathogenic; *LRS*, long-read sequencing; *NA*, not applicable; *ND*, not detected; *NT*, not tested; *P*, pathogenic.

[a]Itai et al.[20]

[b]Smits et al.[21]

[c]Gizak et al.[22]

GS-unique diagnoses not detected with contemporary ES analysis

One of the 5 GS-detected CNVs or SVs were not detected by contemporary ES. A large 24 Mb de novo heterozygous copy-neutral chromosome 11 pericentric inversion, with break points in intronic regions of *SHANK2* and *CREB3L1* was detected in proband 70. A subsequent karyotype confirmed the inversion. With known de novo pathogenic loss-of-function variants, the *SHANK2* breakpoint is likely to invert exons, resulting in truncation with nonsense-mediated decay and loss of protein function. A deletion partially spanning exon 8 and a partial intronic region of *ANK3* in an isoform expressed in the brain (NM_020987.5) was detected in proband 73 with the RG GS pipeline, but not with the Broad CMG GS pipeline, contemporary ES pipeline (RG), or ES CxGo (VCGS). This variant was subsequently assessed as a false positive. This brings the total GS-unique diagnoses down from 9 (12%) to 1 (3%) after contemporary ES analysis.

Genome-wide episignature results

Episignatures were investigated in 32 ES-negative probands where DNA was available. The outcomes were a diagnosis in 1 family and unresolved findings in 2 families (Supplemental Table 7). This resulted in a minimum diagnostic yield of 3% and a maximum yield considering all abnormal episignatures of 9%. An *EHMT1*/Kleefstra syndrome (MIM 610253) episignature was detected in proband 62 and correlated with a GS CNV finding of a partial *EHMT1* duplication (Figure 2A). This CNV was also detected through a contemporary ES pipeline with CNV analysis using CxGo. A high-confidence episignature interpreted as mosaic Beckwith-Wiedemann syndrome (BWS) (MIM 130650) IC2 on chromosome 11 was detected in proband 44, where a de novo *MED13L* duplication on chromosome 12 was also identified on GS CNV analysis (Figure 2B). This CNV was assessed as the diagnosis. There is currently no episignature for *MED13L*, and it is unknown whether the mosaic IC2 episignature may be related to the *MED13L* event. There were no clinical features consistent with BWS present. An inconclusive episignature for *KDM5C* was detected in proband 69 (Figure S3A), as well as a likely mosaic uniparental disomy of the paternal chromosome 7 for *MEST*, *SVOPL*, and *HTR5A* (hemi-methylation observed at *GRB10* and *PEG10* regions on chromosome 7 [Supplemental Figure 3B]). A prior single-nucleotide variation (formerly single-nucleotide polymorphism) array did not detect a mosaic isodisomy in this patient, and no rare variants in *KDM5C* were identified.

LRS results

STR analysis of 48 genes on LRS data from 16 probands, where DNA was available, identified no pathogenic TREs, but 5 probands with variants of potential interest were identified (Supplemental Table 7). An approximate 275-repeat intronic *STARD7* nonreference motif expansion, TTCAG/AACTG, also present in the general population (gnomAD [v3.1.2]), was identified in proband 49. This expansion was confirmed to be paternally inherited and is likely benign for an ID phenotype. A different intronic *STARD7* ATTTC pentamer repeat expansion has been reported in Familial Adult Myoclonic Epilepsy.[23]

Health economic analyses results

Cost-effectiveness analysis provided information on the health and cost impacts of trio GS compared with the current standard of care test, trio ES. The average cost per family under the original ES pathway was estimated at $1851. Performing GS on 32 original ES-negative families increased the average cost per family to $3956, with a further 9 families being diagnosed (a total of 51 families diagnosed). Had contemporary ES been performed at the beginning of the study, 50 of 74 families would have been diagnosed by ES. Performing GS on the contemporary ES-negative families, therefore, yielded 1 additional GS-unique diagnosis. Using GS as a first-line test resulted in an average cost of $4867 per family, with 51 of 74 families being diagnosed.

The incremental cost per additional genetic diagnosis using GS on the original ES-negative families was estimated at $17,305 (95% CI: $11,125; $40,153). The incremental cost per additional genetic diagnosis using GS on contemporary ES-negative families was estimated at $116,808 (95% CI: $37,314; Dominated [ie, higher cost but no additional diagnosis]). GS as a first-line test had an incremental cost of $223,184 (95% CI: $74,395; Dominated) per additional genetic diagnosis compared with contemporary ES as a first-line test. Performing GS on contemporary ES-negative families was less costly compared with performing GS as a first-line test, with the contemporary ES followed by GS pathway costing $4976 (95% CI: $3704; $6969) per diagnosis and first-line GS costing $7062 (95% CI: $6210; $8475) per diagnosis (Table 2). Cost-effectiveness acceptability curves were calculated for a range of willingness-to-pay threshold cost for each additional genetic diagnosis using GS (Supplemental Figure 4). Because there was only 1 additional diagnosis for GS in this cohort, the cost-effectiveness acceptability curves show that for a willingness-to-pay threshold of additional $150,000 for each additional diagnosis, there was a 62% probability that GS would be cost-effective if it was performed on contemporary ES-negative families and about 26% if it was performed as a first-line test.

Based on Australian costs ($1851 for ES and $4867 for GS), the cost of GS trios would need to be reduced by 61.2% (95% uncertainty range: 59.4%-61.9%) to be cost equivalent to contemporary ES trios on a cost per diagnosis basis if all families have GS as a first-line test. If GS is to be

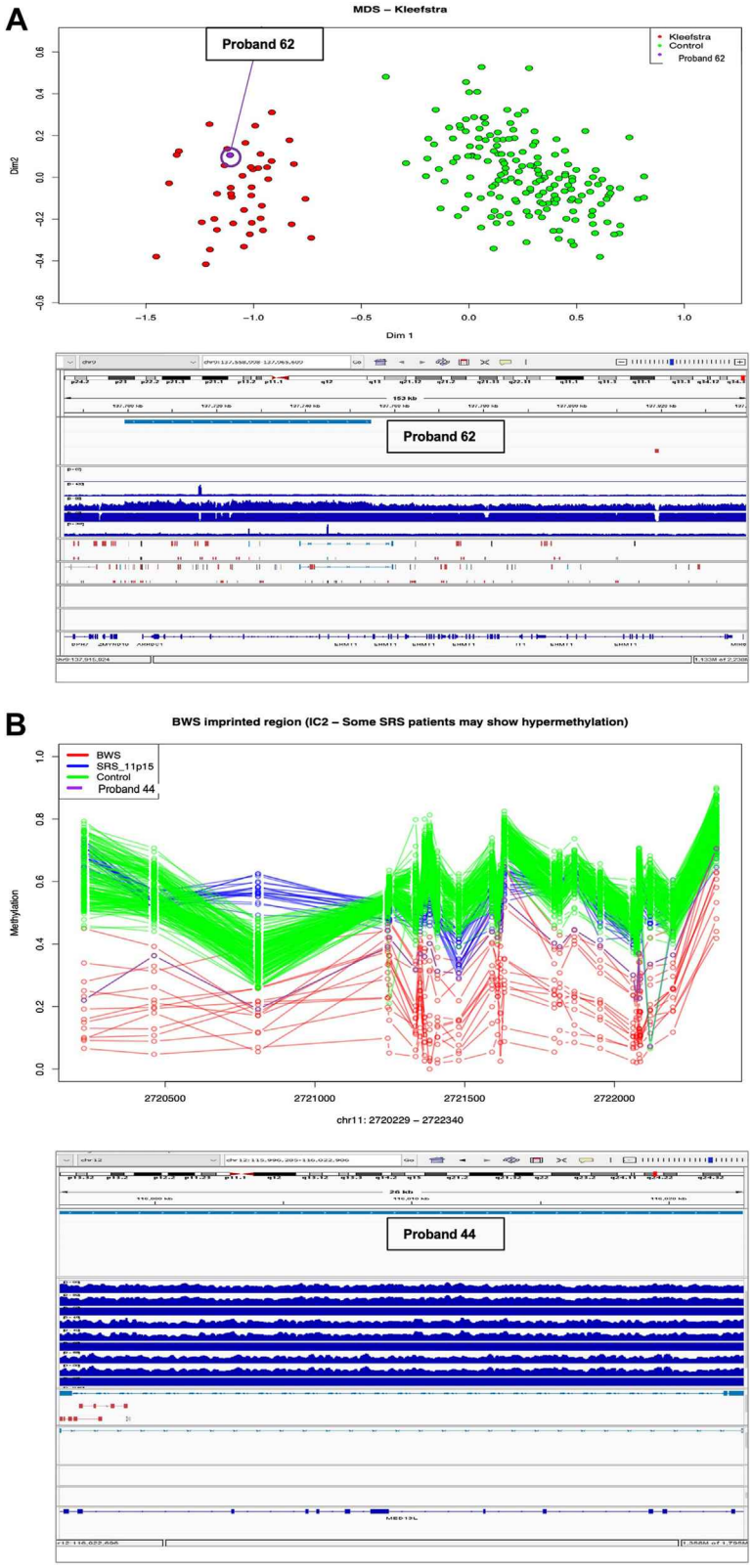

Figure 2 **Concordance between GS and episignature findings.** A. A Kleefstra syndrome episignature in proband 62 was concordant to the finding of a duplication in *EHMT1* using GS. B. A Beckwith-Weidemann syndrome with IC2 episignature in proband 44 and detection of a de novo heterozygous partial out of frame tandem duplication (NM_015335.5, exons 5 - 16 of 31) of *MED13L* using GS might be related by a yet unknown mechanism.

Table 2 Comparisons of average cost, average cost per diagnosis, and incremental cost per additional diagnosis under different diagnostic pathways

	Compared with Original ES			Compared with Contemporary ES		
	Original ES	Original ES followed by GS	GS (as a first line test)	Contemporary ES	Contemporary ES followed by GS	GS (as a first line test)
Total number of diagnosis	42	51	51	50	51	51
Average cost per family	1851	3956	4867	1851	3429	4867
Average cost per diagnosis (95% CI)	3261 (2796; 4029)	5740 (4459; 7680)	7062 (6210; 8475)	2739 (2362; 3301)	4976 (3704; 6969)	7062 (6210; 8475)
Incremental cost per additional diagnosis (95% CI)	Reference diagnostic pathway	17,305 (11,125; 40,153)	24,798 (14,879; 55,796)	Reference diagnostic pathway	116,808 (37,314; Dominated)	223,184 (74,395; Dominated)

CI, confidence interval; *Dominated*, higher cost but no additional diagnosis; *ES*, exome sequencing; *GS*, genome sequencing.

performed only on contemporary ES-negative families (the contemporary ES followed by GS pathway), the cost of GS trios would need to be reduced by 43.7% (95% uncertainty range: 32.2%-51.4%) for it to be cost equivalent. Based on costs more reflective of those in the United States ($2393 for ES and $6598 for GS), the approximate cost of GS of $2435 would be cost equivalent to contemporary ES on a cost per diagnosis basis. This would represent a reduction in the cost of GS of 63.1% (95% uncertainty range: 61.3%-63.7%), to become cost equivalent to contemporary ES.

Sensitivity analysis

Analysis including Australian-specific costs of clinical review and genetic counseling, in addition to laboratory costs, shows that the estimated average cost per family would be $2050 for the contemporary ES pathway, $3634 for GS following contemporary ES, and $5066 for GS as a first-line test. Sensitivity analysis exploring a range of GS diagnostic yields at 5%, 10%, 15%, and 20% more than the contemporary ES diagnostic yield showed that the estimated incremental cost per additional genetic diagnosis using GS on contemporary ES-negative families was $38,936 (95% CI: $17,521; Dominated) if the GS diagnostic yield was 5% more than the contemporary ES diagnostic yield and $11,681 (95% CI: $7964; $23,848) if it was 20% more than the contemporary ES diagnostic yield (Supplemental Figure 5).

Discussion

This study was assembled to obtain health economic evidence for the most rational genomic testing pathways for people with Mendelian forms of ID. Trio ES was performed on 74 prospective, tightly ascertained families with at least moderate ID, followed by trio GS in 32 ES-negative families, episignature analysis in 32 ES-negative individuals, and LRS in 16 ES-negative individuals. A genetic etiology for ID was detected in 51 of 74 families (69%) using a combination of older Illumina and Life Technologies ES and current GS, with the original ES contributing 42 of 74 (57%) and GS-unique diagnoses contributing an additional 1 diagnosis (1%). Although genomic testing has become available and bioinformatic pipelines more standardized, the time required to complete cohort studies is still significant and results in diagnostic outcome biases because of technological advancement. The GS-unique diagnostic yield was therefore refined by repeated contemporary ES sequencing and analysis. Although GS identified 9 events not reported in the original ES, contemporary ES was able to detect 8 of these. The diagnoses were missed because of older sequencing chemistries and unavailability of ES CNV analyses. The 1 GS-unique diagnosis included a large 24 Mb de novo copy-neutral inversion on chromosome 11 with breakpoints within *SHANK2* and near *CREB3L1*. The presence of a de novo etiology for ID enabled accurate low-recurrence risk genetic counseling for this family. Episignature analysis detected an episignature concordant for the *EHMT1* duplication event detected on GS/contemporary ES. Episignature testing also provided further diagnostic information in 2 individuals, consistent with a partial BWS episignature in 1 and a possible paternal UPD7 and/or a partial *MRXSCJ* episignature in the other. The effects of controlling for interim novel gene identifications, refinements to bioinformatic pipelines, and improved sequencing quality have increased the relative utility of ES compared with GS. This study has shown that combining contemporary ES with episignature analyses has narrowed the diagnostic gap between ES and GS to maximize ID diagnoses.

LRS more accurately detects more SVs and STR expansions/contractions across the genome than short-read GS.[24] Much of this variation is highly polymorphic and previously unannotated. Large-scale LRS studies will be required to generate appropriate reference data to assist in interpreting this variation, and functional studies will be required to validate possible pathogenic alleles. Because it is still unknown how much of the additional variation revealed by LRS is clinically relevant, it remains unclear how much of a diagnostic advantage these technologies offer over current first-tier genetic testing.

The GS-specific diagnostic rate for ID appears similar between contemporary ES and GS.[25] A recent Dutch study that performed ES and GS on 150 trios with neurodevelopmental disorders reported a similar diagnostic yield for GS (30.0%) and ES (28.7%).[26] All GS-unique diagnoses reported in that study were CNV events; therefore, it may be concluded that the diagnostic benefit of performing GS in ID is to detect CNVs. The higher diagnostic rate (69%) in this study compared with the diagnostic rate of 30% in the Dutch study may be explained by the selection criteria focusing on families with moderate to severe ID in the study reported here. Most GS findings were identifiable on contemporary ES because of better sequencing quality or higher coverage (22%, 2/9), introduction of ES CNV analysis (44%, 4/9), and interim publication of novel Mendelian gene disease relationships (22%, 2/9). Ongoing improvements in bioinformatic pipelines, coding region coverage, SV/CNV detection, consanguinity assessment, and integration of Human Phenotype Ontology-coded phenotypic information can facilitate the identification of missed diagnoses through data reanalysis.[4,27] A significant impact on diagnostic rates for both ES and GS is novel Mendelian gene disease relationship ascertainment over time. This has been estimated to be approximately 244 novel genes per year at the end of 2018 and includes many conditions with ID.[28] It would be expected that as the rate of novel gene disease relationship discovery plateaus and eventually falls, the relative utility of reanalysis will be reduced.

ES sensitivity has increased through better sequencing and the inclusion of intronic and noncoding regions where definitive pathogenic variants are described. GS has a simpler library building process, provides more uniform genome coverage, and reduces the number of tests for a diagnosis. GS has some drawbacks compared with ES, including decreased sensitivity for detecting mosaic variants because of reduced depth, increased costs for data generation, analysis and storage, and a limited understanding of the functional effects of noncoding variation.[29] Costs for confirmatory tests, including RNA sequencing (RNA-seq) for potential splicing variants, should also be considered. RNA-seq from relevant tissues can increase diagnostic yields by 7% to 36%.[30] RNA-seq costs are relatively low but need to be incorporated into a post-GS cost estimates. Although upfront ES costs can be seen as most cost-effective, the additional costs of GS even with small increases in diagnostic rates and only a proportion of eligible families seeking reproductive modulation, the costs are small compared with life-long care for those families with ID.[31,32] Thus, even a single additional diagnosis may have substantial health and cost implications.

The lifetime costs to families living with ID and society have been estimated to be at least US$172,000 per family/year.[31] It is reasonable to support the larger upfront GS costs because these are recovered over time, even with a small incremental diagnostic rate over contemporary ES. GS funds and those for patient care are, however, often derived from different sources. Consequently, although the general adoption of GS can be argued, the upfront costs need to be reduced for GS to be adopted as a routine first-tier diagnostic test. The potential for lower-cost GS data production is emerging, with costs estimated to be reduced to US$1/Gb (Almogy G, Pratt M, Oberstrass F, et al. Cost-efficient genome sequencing using novel mostly natural sequencing-by-synthesis chemistry and open fluidics platform. bioRxiv. 2022:1-8. https://doi.org/10.1101/2022.05.29.493900) or US$200 per genome[33] for high sample volumes exclusive of data analysis and tertiary reporting. This is well below the estimate for cost-equivalence of GS to ES for achieving a genomic report in ID trios of US$2435 calculated in this study.

The health economic assessments in this study related to ID trios. Other disease types, different sample sizes, gene panels or singleton sequencing will not have the same health economic outcomes and would require independent analysis studies. Despite this, a similar sized cohort ($n = 87$) was reported to have comparable diagnostic yields to this study for ES, which was between 20% and 75% depending on cohort structure, followed by GS performed on the ES-negative cases ($n = 12$) yielded 1 diagnosis (8%) of a missense and an intragenic deletion of 90 Kb.[34] The similarity of ES and GS diagnostic rates is becoming apparent and should be taken into consideration when allocating health budget resources for genomic sequencing versus investment in lower-cost sequencing technologies.

Useful adjunct methodologies with lower costs than GS, such as episignatures, have become available[35] resulting in alternatives to binary decisions between ES and GS. The proportion of genes with episignatures (EpiSign v4) were estimated in this study, the Dutch study,[26] diagnosed ID cases in RG internal databases, and the top 200 DECIPHER ID genes ranked by frequency of diagnoses (Supplemental Figure 6). Of the 51 diagnoses in this study, 9 were genes with episignatures, consistent with 12% (9/74) of cases being diagnosable using a comprehensive ID gene panel at lower cost than ES or GS. Similarly, had episignature testing been performed in the Dutch cohort,[26] it could have provided information about the causative gene in 13 of 150 (7%) cases. The utility of episignature testing is potentially even higher when applied before GS as 3 of 9 (33%) of the GS-identified genes in this study and 13 of 45 (29%) in the Dutch study[26] may have been diagnosable by a panel containing ID genes with known episignatures. Forty-nine of the top 200 (24%) ID genes in DECIPHER database have an episignature, whereas 58 genes (12%) have been assessed as having an episignature out of 930 cases with diagnoses in 469 ID genes in the RG database. These findings demonstrate that the concurrent use of ES and episignatures could assist in prioritizing the choice of the most relevant post-ES methodology to maximize diagnostic rates and minimize costs. Episignatures may have a higher clinical utility with potential cost reductions if performed with ES to prioritize GS. Episignatures have also been shown to have utility in providing a genome-wide functional test to identify likely

genetic etiologies but also in the re-classification of variant of uncertain significance.[36] No variants of uncertain significance were identified in this study that had episignature pathogenicity assessment alterations, but this represents additional potential utility for episignatures.[37]

After extensive testing, 23 of 74 (31%) families in this cohort remained undiagnosed. RNA-seq has been reported to detect an additional 7% to 35% of diagnoses and therefore can be considered to detect abnormally spliced transcripts because of deep intronic or synonymous/missense variants with splicing effects.[30] A splice variant was detected in a similar cohort[26] (0.7%); therefore, it is likely that at least 1 case in this cohort might be diagnosed through RNA-seq. Although the most comprehensive combination of diagnostic techniques could be GS, episignatures, and RNA-seq, these are not yet available routinely in most diagnostic settings. Variation relevant to ID and commonly missed by SRS can also be assessed using LRS.[24] Its utility lies in confirmatory testing and detecting variants inaccessible to SRS, such as large CNVs/SVs and repeat expansions, particularly variation in repeat-rich and high GC-content gene regions, but analysis and interpretation of these data types are still under development.[38]

The narrowing of the diagnostic gap using contemporary ES analysis and episignature informs the most cost-effective combination of diagnostic techniques for ID. Most variants are predicted to be identified using ES trios analyzed with a contemporary ES pipeline inclusive of CNV analysis. Using complementary genetic technologies after contemporary ES and ES reanalysis such as episignatures and then GS will maximize the diagnostic rates in ID. The GS trio price setting cost equivalent to ES for ID trio diagnostic rates was assessed as US$2435, approximately a 60% reduction from current prices. Ongoing technological developments resulting in lower-cost GS will increase its relative value compared with ES, to the point where it is likely to become the preferred first-line genomic test.

URLs

Alamut Visual Plus: https://www.sophiagenetics.com/platform/alamut-visual-plus/

CADD Score: http://cadd.gs.washington.edu/

ClinVar: https://www.ncbi.nlm.nih.gov/clinvar/

Eukaryotic Promoter Database: https://epd.epfl.ch/EPDnew_database.php

GeneMatcher https://genematcher.org

gnomAD: https://gnomad.broadinstitute.org/

OMIM: https://www.omim.org/

PubMed: https://pubmed.ncbi.nlm.nih.gov/

RefSeq: https://www.ncbi.nlm.nih.gov/refseq/

seqr: https://seqr.broadinstitute.org/

UCSC Genome Browser: http://genome.ucsc.edu/

VarCards: http://varcards.biols.ac.cn/

Data Availability

This study generated GS data that are deposited in the NHGRI GREGoR Consortium: Genomics Research to Elucidate the Genetics of Rare Disease - dbGaP Study Accession: phs003047. Proband numbers are mapped to accession numbers (Supplemental Table 8).

Acknowledgments

This study is dedicated to the families who took part and to Dr Anne Turner, who died on August 29, 2022. The authors thank the families that participated in this research and the clinical teams involved in their care. The authors thank the laboratory teams at New South Wales Statewide Genomic Service, Victorian Clinical Genetics Services, and the Broad Institute Center for Mendelian Genomics for their technical expertise.

Funding

ES was funded through the Australian Genomics grant (NHMRC GNT1113531) and the New South Wales Statewide Genomic Service. K.-R.D. and C.-A.E. were supported by the Australian NHMRC Centre for Research Excellence in Neurocognitive Disorders (NHMRC-RG172296). Genome sequencing and analysis were provided by the Broad Institute of Massachusetts Institute of Technology and Harvard Center for Mendelian Genomics as part of the GREGoR (Genomics Research to Elucidate the Genetics of Rare diseases) consortium and was funded by the National Human Genome Research Institute Grants U01HG011755 and R01HG009141.

Author Information

Conceptualization: T.R., D.J.A., A.O.-L.; Data Curation: K.-R.D., K.S., A.M., S.G., M.W., S.A.S., J.S., B.K., A.T., M.B., L.C.A., D.M., M.R., G.M., R.S., S.M.W., T.Y.T., Z.S., N.J.B., L.A.P.-J., E.K., M.F.H., L.A., A.P.F., A.Y., L.J.E., C.L., L.C., H.F., M.R., C.C., G.E., E.K., T.D.-B., J.G., S.C., M.F., T.R.; Formal Analysis: K.R.-D., R.S., C.-A.E., E.O., Y.Z., F.Z., B.W., S.L.S., A.S.-J., H.B., M.E.T., J.K., I.S., J.F., R.W., M.A.C., T.K., M.W., S.L., B.S., M.F.B., A.O.-L., T.R.; Funding Acquisition: T.R., C.M., K.N., B.S., A.O.-L.; Writing-original draft: K.-R.D., R.S., I.W.D.; Writing-review and editing: T.R., A.O.-L., D.S.

Ethics Declaration

The study was approved by the Melbourne Health Human Research Ethics Committee (HREC/16/MH/251) and by the

Mass General Brigham Institutional Review Board (2013P001477).

Conflict of Interest

The authors declare no conflicts of interest.

Additional Information

The online version of this article (https://doi.org/10.1016/j.gim.2024.101076) contains supplementary material, which is available to authorized users.

Affiliations

[1]Neuroscience Research Australia, Sydney, NSW, Australia; [2]Prince of Wales Clinical School, Faculty of Medicine and Health, University of New South Wales, Sydney, NSW, Australia; [3]Centre for Economic Impacts of Genomic Medicine, Macquarie Business School, Macquarie University, Sydney, NSW, Australia; [4]New South Wales Health Pathology Randwick Genomics, Prince of Wales Hospital, Sydney, NSW, Australia; [5]Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, Massachusetts; [6]The Genetics of Learning Disability Service, Waratah, NSW, Australia; [7]Department of Clinical Genetics, Children's Hospital at Westmead, Sydney Children's Hospital Network, Sydney, NSW, Australia; [8]Specialty of Genomic Medicine, Sydney Medical School, University of Sydney, Sydney, NSW, Australia; [9]Department of Clinical Genetics, Royal North Shore Hospital, Sydney, NSW, Australia; [10]Northern Clinical School, Royal North Shore Hospital, Sydney, NSW, Australia; [11]Disciplines of Child and Adolescent Health and Genetic Medicine, University of Sydney, Sydney, NSW 2050, Australia; [12]Genetic Services of Western Australia, Perth, WA, Australia; [13]School of Paediatrics and Child Health, University of Western Australia, Perth, WA, Australia; [14]Centre for Clinical Genetics, Sydney Children's Hospital, Sydney, NSW, Australia; [15]Department of Clinical Genetics, Liverpool Hospital, Sydney, NSW, Australia; [16]Discipline of Paediatrics & Child Health, Faculty of Medicine and Health, School of Clinical Medicine, University of New South Wales, Sydney, NSW, Australia; [17]Monash Genetics, Monash Health, Melbourne, VIC, Australia; [18]Victorian Clinical Genetics Services, Murdoch Children's Research Institute, Melbourne, VIC 3052, Australia; [19]Murdoch Children's Research Institute, Melbourne, VIC, Australia; [20]Department of Paediatrics, University of Melbourne, Melbourne, VIC, Australia; [21]Australian Genomics, Melbourne, VIC, Australia; [22]Genetics Unit, Universitat Pompeu Fabra, Institut Hospital del Mar d'Investigacions Mediques (IMIM), Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Barcelona, Spain; [23]Women's and Children's Hospital, South Australian Health and Medical Research Institute & University of Adelaide, Adelaide, SA, Australia; [24]Department of Paediatrics, Monash University, Melbourne, VIC, Australia; [25]Melbourne Pathology, Melbourne, VIC, Australia; [26]Department of Pathology, The Royal Melbourne Hospital, Melbourne, VIC, Australia; [27]Melbourne Medical School, University of Melbourne, Melbourne, VIC, Australia; [28]Genomics and Inherited Disease Program, Garvan Institute of Medical Research, University of New South Wales Sydney, Sydney, NSW, Australia; [29]Verspeeten Clinical Genome Centre London Health Sciences Centre, London, ON, Canada; [30]Department of Medical Genomics, Royal Prince Alfred Hospital, Sydney, NSW, Australia; [31]Centre for Population Genomics, Garvan Institute of Medical Research and Murdoch Children's Research Institute, Sydney, NSW, Australia; [32]St Vincent's Clinical School, Faculty of Medicine, University of New South Wales, Sydney, NSW, Australia; [33]Adelaide Medical School and Robinson Research Institute, University of Adelaide, Adelaide, SA, Australia; [34]South Australian Health and Medical Research Institute, Adelaide, SA, Australia; [35]State Wide Service, New South Wales Health Pathology, Sydney, NSW, Australia; [36]Global Alliance for Genomics and Health, Toronto, ON, Canada; [37]Department of Pathology and Laboratory Medicine, Western University, London, ON, Canada; [38]Division of Genetics and Genomics, Boston Children's Hospital, Boston, MA

References

1. Maia N, Nabais Sá MJ, Melo-Pires M, de Brouwer APM, Jorge P. Intellectual disability genomics: current state, pitfalls and future challenges. *BMC Genomics*. 2021;22(1):909. http://doi.org/10.1186/s12864-021-08227-4
2. Srivastava S, Love-Nichols JA, Dies KA, et al. Meta-analysis and multidisciplinary consensus statement: exome sequencing is a first-tier clinical diagnostic test for individuals with neurodevelopmental disorders. *Genet Med*. 2019;21(11):2413-2421. http://doi.org/10.1038/s41436-019-0554-6
3. Marshall CR, Chowdhury S, Taft RJ, et al. Best practices for the analytical validation of clinical whole-genome sequencing intended for the diagnosis of germline disease. *NPJ Genom Med*. 2020;5:47. http://doi.org/10.1038/s41525-020-00154-9
4. Ewans LJ, Schofield D, Shrestha R, et al. Whole-exome sequencing reanalysis at 12 months boosts diagnosis and is cost-effective when applied early in Mendelian disorders. *Genet Med*. 2018;20(12):1564-1574. http://doi.org/10.1038/gim.2018.39
5. Schwarze K, Buchanan J, Taylor JC, Wordsworth S. Are whole-exome and whole-genome sequencing approaches cost-effective? A systematic review of the literature. *Genet Med*. 2018;20(10):1122-1130. http://doi.org/10.1038/gim.2017.247
6. Jegathisawaran J, Tsiplova K, Hayeems RZ, et al. Trio genome sequencing for developmental delay and pediatric heart conditions: a comparative microcost analysis. *Genet Med*. 2022;24(5):1027-1036. http://doi.org/10.1016/J.GIM.2022.01.020
7. Evans CA, Pinner J, Chan CY, et al. Fetal diagnosis of Mowat-Wilson syndrome by whole exome sequencing. *Am J Med Genet A*. 2019;179(10):2152-2157. http://doi.org/10.1002/ajmg.a.61295

8. Sundercombe SL, Berbic M, Evans CA, et al. Clinically responsive genomic analysis pipelines: elements to improve detection rate and efficiency. *J Mol Diagn*. 2021;23(7):894-905. http://doi.org/10.1016/j.jmoldx.2021.04.007
9. Karbassi I, Maston GA, Love A, et al. A standardized DNA variant scoring system for pathogenicity assessments in Mendelian disorders. *Hum Mutat*. 2016;37(1):127-134. http://doi.org/10.1002/humu.22918
10. Fowler A, Mahamdallie S, Ruark E, et al. Accurate clinical detection of exon copy number variants in a targeted NGS panel using DECoN. *Wellcome Open Res*. 2016;1:20. http://doi.org/10.12688/WELLCOMEOPENRES.10069.1
11. Fromer M, Moran JL, Chambert K, et al. Discovery and statistical genotyping of copy-number variation from whole-exome sequencing depth. *Am J Hum Genet*. 2012;91(4):597-607. http://doi.org/10.1016/J.AJHG.2012.08.005
12. Sadedin SP, Ellis JA, Masters SL, Oshlack A. Ximmer: a system for improving accuracy and consistency of CNV calling from exome data. *GigaScience*. 2018;7(10):giy112. http://doi.org/10.1093/gigascience/giy112
13. Pais LS, Snow H, Weisburd B, et al. seqr: a web-based analysis and collaboration tool for rare disease genomics. *Hum Mutat*. 2022;43(6):698-707. http://doi.org/10.1002/humu.24366
14. Sobreira N, Schiettecatte F, Valle D, Hamosh A. GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. *Hum Mutat*. 2015;36(10):928-930. http://doi.org/10.1002/humu.22844
15. Aref-Eshghi E, Kerkhof J, Pedro VP, et al. Evaluation of DNA methylation episignatures for diagnosis and phenotype correlations in 42 Mendelian neurodevelopmental disorders. *Am J Hum Genet*. 2020;106(3):356-370. http://doi.org/10.1016/J.AJHG.2020.01.019
16. Sadikovic B, Levy MA, Kerkhof J, et al. Clinical epigenomics: genome-wide DNA methylation analysis for the diagnosis of Mendelian disorders. *Genet Med*. 2021;23(6):1065-1074. http://doi.org/10.1038/s41436-020-01096-4
17. Gamaarachchi H, Samarakoon H, Jenner SP, et al. Fast nanopore sequencing data analysis with SLOW5. *Nat Biotechnol*. 2022;40(7):1026-1029. http://doi.org/10.1038/s41587-021-01147-4
18. Samarakoon H, Ferguson JM, Jenner SP, et al. Flexible and efficient handling of nanopore sequencing signal data with slow5tools. *Genome Biol*. 2023;24(1):69. http://doi.org/10.1186/s13059-023-02910-3
19. Stevanovski I, Chintalaphani SR, Gamaarachchi H, et al. Comprehensive genetic diagnosis of tandem repeat expansion disorders with programmable targeted nanopore sequencing. *Sci Adv*. 2022;8(9):eabm5386. http://doi.org/10.1126/sciadv.abm5386
20. Itai T, Hamanaka K, Sasaki K, et al. De novo variants in CELF2 that disrupt the nuclear localization signal cause developmental and epileptic encephalopathy. *Hum Mutat*. 2021;42(1):66-76. http://doi.org/10.1002/humu.24130
21. Smits DJ, Schot R, Popescu CA, et al. De novo MCM6 variants in neurodevelopmental disorders: a recognizable phenotype related to zinc binding residues. *Hum Genet*. 2023;142(7):949-964. http://doi.org/10.1007/s00439-023-02569-7
22. Gizak A, Diegmann S, Dreha-Kulaczewski S, et al. A novel remitting leukodystrophy associated with a variant in FBP2. *Brain Commun*. 2021;3(2):fcab036. http://doi.org/10.1093/braincomms/fcab036
23. Corbett MA, Kroes T, Veneziano L, et al. Intronic ATTTC repeat expansions in STARD7 in familial adult myoclonic epilepsy linked to chromosome 2. *Nat Commun*. 2019;10(1):4920. http://doi.org/10.1038/S41467-019-12671-Y
24. Miller DE, Sulovari A, Wang T, et al. Targeted long-read sequencing identifies missing disease-causing variation. *Am J Hum Genet*. 2021;108(8):1436-1449. http://doi.org/10.1016/j.ajhg.2021.06.006
25. Kingsmore SF, Cakici JA, Clark MM, et al. A randomized, controlled trial of the analytic and diagnostic performance of singleton and trio, rapid genome and exome sequencing in ill infants. *Am J Hum Genet*. 2019;105(4):719-733. http://doi.org/10.1016/J.AJHG.2019.08.009
26. van der Sanden BPGH, Schobers G, Corominas Galbany J, et al. The performance of genome sequencing as a first-tier test for neurodevelopmental disorders. *Eur J Hum Genet*. 2023;31(1):81-88. http://doi.org/10.1038/s41431-022-01185-9
27. Bullich G, Matalonga L, Pujadas M, et al. Systematic collaborative reanalysis of genomic data improves diagnostic yield in neurologic rare diseases. *J Mol Diagn*. 2022;24(5):529-542. http://doi.org/10.1016/j.jmoldx.2022.02.003
28. Bamshad MJ, Nickerson DA, Chong JX. Mendelian gene discovery: fast and furious with no end in sight. *Am J Hum Genet*. 2019;105(3):448-455. http://doi.org/10.1016/J.AJHG.2019.07.011
29. Schwarze K, Buchanan J, Fermont JM, et al. The complete costs of genome sequencing: a microcosting study in cancer and rare diseases from a single center in the United Kingdom. *Genet Med*. 2020;22(1):85-94. http://doi.org/10.1038/s41436-019-0618-7
30. Montgomery SB, Bernstein JA, Wheeler MT. Toward transcriptomics as a primary tool for rare disease investigation. *Cold Spring Harb Mol Case Stud*. 2022;8(2):a006198. http://doi.org/10.1101/MCS.A006198
31. Doble B, Schofield D, Evans CA, et al. Impacts of genomics on the health and social costs of intellectual disability. *J Med Genet*. 2020;57(7):479-486. http://doi.org/10.1136/jmedgenet-2019-106445
32. Ewans LJ, Minoche AE, Schofield D, et al. Whole exome and genome sequencing in Mendelian disorders: a diagnostic and health economic analysis. *Eur J Hum Genet*. 2022;30(10):1121-1131. http://doi.org/10.1038/s41431-022-01162-2
33. Hale C. *Illumina Pitches $200 Genomes with New Line of DNA Sequencers*. Fierce Biotech; 2022.
34. Álvarez-Mora MI, Sánchez A, Rodríguez-Revenga L, et al. Diagnostic yield of next-generation sequencing in 87 families with neurodevelopmental disorders. *Orphanet J Rare Dis*. 2022;17(1):60. http://doi.org/10.1186/S13023-022-02213-Z
35. Frésard L, Montgomery SB. Diagnosing rare diseases after the exome. *Cold Spring Harb Mol Case Stud*. 2018;4(6):a003392. http://doi.org/10.1101/mcs.a003392
36. Ciolfi A, Foroutan A, Capuano A, et al. Childhood-onset dystonia-causing KMT2B variants result in a distinctive genomic hypermethylation profile. *Clin Epigenetics*. 2021;13(1):157. http://doi.org/10.1186/s13148-021-01145-y
37. Levy MA, McConkey H, Kerkhof J, et al. Novel diagnostic DNA methylation episignatures expand and refine the epigenetic landscapes of Mendelian disorders. *HGG Adv*. 2022;3(1):100075. http://doi.org/10.1016/J.XHGG.2021.100075
38. Pauper M, Kucuk E, Wenger AM, et al. Long-read trio sequencing of individuals with unsolved intellectual disability. *Eur J Hum Genet*. 2021;29(4):637-648. http://doi.org/10.1038/S41431-020-00770-0