HGG Advances

ARTICLE

Genome-wide sequencing and the clinical diagnosis of genetic disease: The CAUSES study

Alison M. Elliott,[1,2,3,*] Shelin Adam,[1,2] Christèle du Souich,[1,2] Anna Lehman,[1,2] Tanya N. Nelson,[4] Clara van Karnebeek,[5,6] Emily Alderman,[1] Linlea Armstrong,[1,2] Gudrun Aubertin,[1] Katherine Blood,[1] Cyrus Boelman,[2,7] Cornelius Boerkoel,[1,2] Karla Bretherick,[4] Lindsay Brown,[4] Chieko Chijiwa,[1] Lorne Clarke,[1,2] Madeline Couse,[1] Susan Creighton,[1] Abby Watts-Dickens,[1] William T. Gibson,[1,2] Harinder Gill,[1] Maja Tarailo-Graovac,[2] Sara Hamilton,[1] Harindar Heran,[1] Gabriella Horvath,[2,8] Lijia Huang,[4] Gurdip K. Hulait,[1] David Koehn,[1] Hyun Kyung Lee,[1] Suzanne Lewis,[1,2] Elena Lopez,[1,2] Kristal Louie,[1] Karen Niederhoffer,[1] Allison Matthews,[4] Kirsten Meagher,[1] Junran J. Peng,[2] Millan S. Patel,[1,2] Simone Race,[8] Phillip Richmond,[2] Rosemarie Rupps,[1] Ramona Salvarinova,[2,8] Kimberly Seath,[1] Kathryn Selby,[2,7] Michelle Steinraths,[1] Sylvia Stockler,[2,8] Kaoru Tang,[1] Christine Tyson,[4] Margot van Allen,[1,2] Wyeth Wasserman,[1,2,5] Jill Mwenifumbo,[2] and Jan M. Friedman[1,2]

Summary

Genome-wide sequencing (GWS) is a standard of care for diagnosis of suspected genetic disorders, but the proportion of patients found to have pathogenic or likely pathogenic variants ranges from less than 30% to more than 60% in reported studies. It has been suggested that the diagnostic rate can be improved by interpreting genomic variants in the context of each affected individual's full clinical picture and by regular follow-up and reinterpretation of GWS laboratory results.

Trio exome sequencing was performed in 415 families and trio genome sequencing in 85 families in the CAUSES study. The variants observed were interpreted by a multidisciplinary team including laboratory geneticists, bioinformaticians, clinical geneticists, genetic counselors, pediatric subspecialists, and the referring physician, and independently by a clinical laboratory using standard American College of Medical Genetics and Genomics (ACMG) criteria. Individuals were followed for an average of 5.1 years after testing, with clinical reassessment and reinterpretation of the GWS results as necessary. The multidisciplinary team established a diagnosis of genetic disease in 43.0% of the families at the time of initial GWS interpretation, and longitudinal follow-up and reinterpretation of GWS results produced new diagnoses in 17.2% of families whose initial GWS interpretation was uninformative or uncertain. Reinterpretation also resulted in rescinding a diagnosis in four families (1.9%). Of the families studied, 33.6% had ACMG pathogenic or likely pathogenic variants related to the clinical indication. Close collaboration among clinical geneticists, genetic counselors, laboratory geneticists, bioinformaticians, and individuals' primary physicians, with ongoing follow-up, reanalysis, and reinterpretation over time, can improve the clinical value of GWS.

Introduction

Genome-wide sequencing (GWS; exome sequencing [ES] or genome sequencing [GS]) has transformed the ability to diagnose patients with genetic diseases. Many studies of the diagnostic rate or clinical utility of GWS have used the finding of variants classified as "pathogenic" or "likely pathogenic" according to the American College of Medical Genetics and Genomics and the Association for Molecular Pathology (ACMG/AMP) standards and guidelines[1] to provide a surrogate "molecular diagnosis" of genetic disease in patients,[2–5] although this practice is inconsistent with the ACMG guidelines, which state:

In general, a variant classified as pathogenic using the proposed classification scheme has met criteria informed by empirical data such that a health-care provider can use the molecular testing information in clinical decision making. Efforts should be made to avoid using this as the sole evidence of Mendelian disease; it should be used in conjunction with other clinical information when possible.

A more recent ACMG clinical practice guideline[6] puts this in a clinical, rather than laboratory-focused, context:

Clinical genetic testing by ES/GS can assist clinicians in confirming or establishing a clinical diagnosis that may lead to changes in management, obviate the need for further testing, and/or end the diagnostic odyssey.

[1]Department of Medical Genetics, University of British Columbia, Vancouver, BC, Canada; [2]BC Children's Hospital Research Institute, Vancouver, BC, Canada; [3]Women's Health Research Institute, Vancouver, BC, Canada; [4]Division of Genome Diagnostics, Department of Pathology and Laboratory Medicine, BC Children's and Women's Hospitals, Vancouver, BC, Canada; [5]Department of Pediatrics, Center for Molecular Medicine and Therapeutics, University of British Columbia, Vancouver, BC, Canada; [6]Department of Pediatrics, Emma Children's Hospital, Amsterdam, University Medical Centers, University of Amsterdam, Amsterdam, the Netherlands; [7]Division of Neurology, Department of Pediatrics, University of British Columbia, Vancouver, BC, Canada; [8]Division of Biochemical Diseases, Department of Pediatrics, University of British Columbia, Vancouver, BC, Canada
*Correspondence: aelliott@bcchr.ca
https://doi.org/10.1016/j.xhgg.2022.100108.

Establishing the diagnosis of a genetic disease in an affected individual requires knowledge of that individual's medical history and disease course over time, family history, physical examination, specialist physician evaluations, and imaging studies, as well as the results of GWS and other laboratory tests. GWS is by far the best test for disease-causing genetic variants that has ever been available, but genotype-phenotype correlation is essential for establishing the diagnosis of, rather than just the risk for, a genetic disease in an affected individual. A recent multidisciplinary consensus statement on the use of ES as a first-tier clinical diagnostic test for neurodevelopmental disorders[7] put it this way: "Ultimately, final variant interpretation of the ES test result requires review by expert clinicians who can reassess the patient's phenotype in light of the suggested molecular diagnosis."

To avoid confusion, we shall eschew use of the term "diagnosis" or "molecular diagnosis" to describe a laboratory interpretation based only on standard ACMG variant classification and the limited phenotypic information typically provided on a test requisition. We shall refer to such laboratory-based interpretations as "variant classifications" and reserve use of the term "diagnosis" to mean our multidisciplinary research team's (or, in practice, the affected individual's physician's) interpretation of GWS findings as probably or definitely disease-causing for a particular genetic condition in the context of all of the available clinical information about the affected individual.

Previous studies have found that interpretation of GWS variants by a multidisciplinary team with clinical as well as laboratory expertise and deep knowledge of each affected individual's entire phenotype can produce a higher rate of genetic disease diagnosis than can be obtained through a laboratory report based on variant interpretation with limited clinical information alone.[8–17] Other factors that have been associated with higher rates of genetic disease diagnosis include trio, rather than singleton, GWS;[17,18] the use of GS rather than ES;[11,17] and reanalysis or reinterpretation of GWS data years after the initial analysis. Substantial rates of variant reclassification, both upward to pathogenic/likely pathogenic and downward from pathogenic/likely pathogenic, have been reported with reinterpretation, and these changes may have a substantial impact on clinical diagnosis and management.[17,19,20]

Here we report our experience using GWS in a longitudinal study of 500 families of children with suspected genetic diseases. GWS results were interpreted by a multidisciplinary clinical research team, and individuals were followed for an average of 5.1 years after testing, with clinical reassessment and reinterpretation of the GWS results as necessary.

Material and methods

Ethics and consent

The CAUSES study was approved by the University of British Columbia-BC Children's and Women's Hospital Research Ethics Board (protocol H15-00092). Consent or assent was obtained from probands (when possible) and from their parents.

Recruitment of participants

Individuals were enrolled through a genomic consultation service to identify individuals for whom there was high suspicion of an underlying monogenic disorder that had not been established through conventional genetic testing.[21,22] As the CAUSES GWS Study was trio based, the availability of both parents was required for enrollment. Almost all individuals were <19 years of age at the time of enrollment. An algorithm describing patient and sample flow in the CAUSES Research Study has previously been published.[21] All individuals received pre- and post-test genetic counseling by an experienced genetic counselor, who also obtained informed consent for GWS, opting in or out for incidental findings for parents, genetic counseling research, and health economic evaluation of GWS.[22,23] The approach to incidental findings followed the Canadian College of Medical Geneticists guidelines.[24] Consent for data sharing through DECIPHER was obtained from each family, and parents were offered the option to have whole blood in excess of that needed for DNA extraction banked for future research on their child at the BC Children's Hospital Biobank. For individuals who had not been previously evaluated by a clinical geneticist, a brief clinical examination was performed by a geneticist and phenotypic information was captured. Telemedicine was offered for the pre-test genetic counseling session for families who lived outside of the Vancouver area and had previously been seen by a clinical geneticist.[25]

Individuals who received GWS were classified into one of four major phenotypic categories: isolated intellectual disability (ID), syndromic ID, unexplained disorders of organ dysfunction without ID, and multiple congenital anomalies without ID. Information collected on all individuals included age, sex, self-described ethnicity, relevant family history, consanguinity, phenotypic findings, and previous investigations.

Sequencing and bioinformatic analysis

Genomic DNA was isolated from peripheral blood using standard techniques on each member of the trio/quad at the Division of Genome Diagnostics at BC Children's Hospital. In total, 415 families underwent ES and 85 families underwent GS (81 of whom were selected for ID as part of a collaboration with Genomics England). Hybridization-based capture was used to enrich for exomes on each member of the trio. ES was performed on Illumina platforms at Ambry Genetics, Centogene, or Canada's Michael Smith Genome Sciences Centre. Read alignment and single-nucleotide variant (SNV)/insertion or deletion (indel) calling were performed as previously described.[26,27]

PCR-free genome libraries were sequenced on an Illumina platform at Canada's Michael Smith Genome Sciences Centre or the McGill Genome Centre.[28] Read alignment and SNV/indel calling utilized BWA-0.7.6,[29] Sambamba-0.5.5,[30] GATK HaplotypeCaller 3.5,[31] BCFtools-1.9,[28] and HTSlib-1.9.[28] VarSeq (Golden Helix, Inc., Bozeman, MT, www.goldenhelix.com) was used for SNV/indel annotation and filtering. CNVnator-0.3.3,[32] ERDS-1.1,[33] LUMPY-0.2.13,[34] and Manta-1.1.1[35] were used for SV calling. EverythingSV (https://github.com/FriedmanLab/StructuralVariantAnalysis), Samtools-1.5,[28] SURVIVOR-1.0.2,[36] and ANNOVAR-2018-04-16[37] were used for SV merging, annotation, and filtering. Annotation databases used for SVs included HPO,[38] OMIM,[39] and RefSeq-105.[40]

Interpretation of GWS

GWS FASTQ files were analyzed by a bioinformatician (genomic analyst) who generated a short list of the most promising candidate variants for each family. Criteria used for variant filtration were as follows:

1. Variants were included if their reference allele frequency was consistent with either dominant or recessive disorders.
2. Variants were included if they were predicted to have an impact on the protein's level, structure, or function or were recorded in ClinVar as "pathogenic."
3. Variant prioritization was largely based on the trio family structure. Specifically, those variants identified as *de novo*, compound heterozygous, homozygous, or hemizygous in the proband were the main focus of analysis.
4. Variant prioritization also included a list of all variants with reference frequencies consistent with an autosomal recessive disorder and in genes known to be associated with a disorder. The aim of this list was to identify variants seemingly outside of the trio family structure filter that might still be disease causing as a result of imprinting, incomplete penetrance, variable expressivity, mosaicism in a parent, a parent being affected (but unknown at the time of intake), or a hemizygous variant called as heterozygous.

The complete, fully annotated list of selected variants was provided to the multidisciplinary research team for review; however, the genomic analyst flagged ClinVar "pathogenic" and predicted loss-of-function variants for particular attention. Each of the selected variants was discussed by a multidisciplinary research team that included genome analysts, MD clinical geneticists, genetic counselors, a PhD laboratory geneticist, and the referring physician. The team assigned a diagnostic category ("definitely causal," "probably causal," "uncertain," or "uninformative") for each individual. The categories were assigned by a consensus of clinical judgment based on the full clinical picture, including all of the information available on the variant(s) as well as the complete medical history, disease course over time, family history, physical examination findings, specialist consultations, imaging studies, and other laboratory test results. Individuals who had variants that were judged to be definitely or probably causal of the phenotype were considered to have been diagnosed with a genetic disease in our analysis.

The criteria that our multidisciplinary team used to determine if variants were probably or definitely disease-causing were independent of the ACMG variant classification. In fact, the clinical diagnostic categories were assigned before we obtained Sanger sequencing and standard ACMG classification of the variants through our clinical laboratory. Our bioinformatics analysis included annotation of all of the factors included in the ACMG classification, and this information was discussed by our multidisciplinary research team in the context of each individual's complete clinical picture.

We adopted a convention of $\geq$99% certainty for "definitely causal" and $\geq$90% (but $<$99%) certainty for "probably causal" diagnoses. Individuals who had variants with bioinformatics evidence for pathogenicity but whose clinical features were not clearly compatible with those reported for the genotype in the published literature were categorized as "uncertain." Individuals in whom our bioinformatics analysis found no variant that appeared to be causal for the clinical features were considered to have "uninformative" GWS. A research summary report was sent to the referring physician following the multidisciplinary team discussion.

Orthogonal validation of variants that were interpreted by our multidisciplinary research team as definitely or probably causal or that were uncertain but possibly causal was performed by Sanger sequencing through a clinical laboratory (Division of Genome Diagnostics at BC Children's Hospital). The clinical laboratory independently assigned an ACMG classification to each variant[1] and issued a standard report to the affected individual's chart. Variants classified as "pathogenic" or "likely pathogenic" by the clinical laboratory were uploaded into the ClinVar database.

Definitely or probably causal variants were considered to be responsible for a "partial" diagnosis if the variant(s) appeared to explain only a portion of the individual's phenotype. A dual diagnosis was assigned when probably or definitely causal variants of two different genetic loci were judged to have contributed to the individual's phenotype. Variants that could not be classified as causal but were thought by the multidisciplinary team to be interesting candidates for further research were entered into GeneMatcher and considered for further study.

Post-test genetic counseling and reanalysis

The referring physician and CAUSES genetic counselor met with families of individuals in whom a diagnosis of genetic disease was made (i.e., whose GWS findings were judged to be probably or definitely causal of the clinical phenotype) to discuss these results when the clinical Sanger sequencing report was available. Families who did not receive a diagnosis of genetic disease through CAUSES were contacted by the genetic counselor, so informed, and told that reanalysis would occur periodically until the study end.

Variant call format files for GWS datasets from each family in whom no genetic diagnosis or only a partial diagnosis of genetic disease was made were reanalyzed through a Golden Helix VarSeq annotation and filtered workflow with the most up-to-date ClinVar and OMIM annotations every 1–2 years. Reanalysis was a planned part of the CAUSES project; it was not dependent on physician request. The approach used was similar to the primary analysis, with the focus on variants identified as *de novo*, compound heterozygous, homozygous, or hemizygous in the proband. Other variants were considered in the reanalysis only when important new clinical information had become available on the proband or the phenotype associated with the genetic locus had been expanded or characterized more fully in the published literature. The main focus of the routine reanalysis was to find variants for which the ACMG classification had changed to pathogenic or likely pathogenic in ClinVar or that occurred in genes that had been newly associated with a genetic disorder in OMIM or the literature. If the genomic analyst determined that results of the reanalysis might alter the diagnostic category ("definitely causal," "probably causal," "uncertain," or "uninformative") of an individual, it was reconsidered by the multidisciplinary research team and changed by consensus, if necessary. The reason for changing the diagnostic category was determined and recorded as a change in the bioinformatic pipeline, the emergence of a new disorder, a new publication, or the referring physician's reinterpretation of the phenotype-genotype relationship. The genomic variants associated with such changes were Sanger sequenced in our clinical laboratory, reported through standard clinical protocols (including ACMG classification of variants), and returned to the affected individual's medical record, referring physician, and family, as described above.

Referring physicians, who continued to follow their patients clinically; members of the CAUSES research team; or patients' families (through recontact with CAUSES genetic counselors) could also request reanalysis of genomic variants or reinterpretation of CAUSES results by the multidisciplinary research team.

Statistical analysis

The Mantel-Haenszel test was used to compare the multidisciplinary team's interpretation of the GWS results in each proband with the clinical laboratory's ACMG classification of Sanger-sequenced variants. The cumulative probability of reinterpretation was calculated and plotted in IBM SPSS Statistics v.24 using the Kaplan-Meier module.

Results

CAUSES cohort demographics

We enrolled 500 families, including 531 children (probands and affected sibs). Trios, defined as two parents and one affected child, were usually studied. In 31 families, two affected sibs and both parents were studied; different combinations of affected and unaffected relatives were tested in a few other families. The mean age (±standard deviation) at referral of the children who received GWS was 8.0 (±4.9) years. There were 246 females and 285 males. The individuals were ethnically diverse, with 48.5% of European, 16.2% of South Asian, 15.8% of East Asian, 4.3% of Middle Eastern, and 4.1% of First Nations ancestry. The most frequent indication for GWS was syndromic ID (85.1%), followed by multiple congenital anomalies without ID (5.3%), disorders of organ function (5.0%), and isolated ID (4.6%).

Diagnostic rate

ES was performed in 415 families and GS in 85. The CAUSES multidisciplinary research team diagnosed at least one genetic disorder in 261 (52.2%) of the families studied; 105 families were found to have variants that were probably causal and 156 families had variants that were considered to be definitely causal of a genetic disease in the child (Table S1). Of the 261 families diagnosed with at least one genetic disorder, 36 had variants that could not be classified according to the ACMG classification and 65 had variants the were classified as VUS. The rationale our multidisciplinary team used to diagnose a genetic disease in each of the individuals in whom an ACMG variant of uncertain significance or an unclassified variant was found is shown in Table 1.

Considering all families in which the CAUSES multidisciplinary research team diagnosed at least one genetic disorder, the diagnostic rate was 52.3% with ES and 51.8% with GS. In nine families, the probands had dual diagnoses (Table S2). Of the 261 probands who were diagnosed with a genetic disease, 219 had autosomal dominant (184 *de novo*), 27 autosomal recessive (one with isodisomy), and 13 X-linked recessive or X-linked dominant disorders. One proband had a Y-linked disorder, and one inherited an Xp25 genomic duplication from the mother. Partial diagnoses were established in 19 families (Table S3).

Affected sibs were tested in 31 of the families studied, and a diagnosis of genetic disease was established in 17 (55%) of these families by our multidisciplinary research team (Table S4). There were 11 sib pairs who were concordant for a definitely or probably causal variant associated with the individuals' phenotypes. Six of these families exhibited autosomal recessive inheritance and three autosomal dominant inheritance, with parental mosaicism in two families and maternal inheritance in one. In addition, one sib pair was concordant for orofacial clefts and a paternally inherited *PLEKHA7* variant and discordant for congenital NAD deficiency disorder 2/vertebral, cardiac, renal, and limb defects syndrome 2 (VCRL2; MIM: 617661) caused by compound heterozygosity for *KYNU* variants. In three families, the sibs were discordant for phenotype, and the diagnosis of a different genetic disease was established in each sib. In three other phenotypically discordant sib pairs, definitely or probably causal genetic variants were found in only one sib.

We reported incidental findings in one parent in 21 (4.4%) of the 478 families who opted for return of these results. Eight were pharmacogenomic variants (*DPYD*), and seven were cancer predisposition genes (*BRCA2*, *BRCA1*, *BAP1*, or *CDK4*). Single individuals had incidental findings in *G6PD*, *LDLR*, or *APOB*.

Diagnostic reinterpretation by the multidisciplinary team

In 4 (1.9%) of the 215 families initially diagnosed as having a genetic condition associated with a definitely or probably disease-causing genomic variant, our multidisciplinary research team reinterpreted the GWS findings as uncertain or uninformative as a result of additional information on the individual, gene, or variant that became available during the period of follow-up (Table 2). For individual G483, an ACMG pathogenic *ALPL* variant initially interpreted as definitely causal was reinterpreted as uncertain on the basis of a normal serum alkaline phosphatase level. In G369, a *TBX1* variant initially interpreted as probably causal in a child with neurodevelopmental abnormalities consistent with velocardiofacial syndrome (VCFS; MIM: 192430) was reinterpreted as uncertain after an evaluation by a cardiologist was normal. In G550, a *USP7* variant originally considered to be probably causal was reinterpreted as uncertain on the basis of a detailed physical examination by the referring physician.

An *ATP2A2* variant in a child with global developmental delay, mild dysmorphic features, generalized hypotonia, and intention tremor (G103) was initially interpreted as definitely disease causing but subsequently reinterpreted by our multidisciplinary research team as uncertain. This reinterpretation occurred when a *de novo* missense variant in *DDX23*, a gene that was not known to be disease associated at the time of initial analysis, was reinterpreted as probably disease-causing because of G103's phenotypic

Table 1. Rationale for "causal" or "likely causal" diagnostic interpretation by multidisciplinary research team in individuals with variants that could not be given an ACMG classification or that were classified as ACMG VUS

CAUSES ID no.	Sex	Gene	Variant	Mechanism	Disease	ACMG variant classification	Diagnostic interpretation by multidisciplinary research team	Rationale for diagnostic interpretation
G001-1	F	*ASXL3*	NP_085135.1:p.Asn1224Ter	*de novo* heterozygous	Bainbridge-Ropers syndrome	VUS	definitely causal	phenotype consistent with Bainbridge-Ropers syndrome; bioinformatics predict that variant is damaging
G004-1	F	*TBCK*	NP_001156907.2:p.Arg261Ser NM_001163435.3:c.2060-2A>G	compound heterozygous	infantile hypotonia with psychomotor retardation and characteristic facies 3	likely pathogenic VUS	probably causal	phenotype characteristic of reported cases; VUS predicted as damaging and allelic to likely pathogenic variant
G005-1	F	*CAMK2G*	NP_001354463.1:p.Arg297Trp	*de novo* heterozygous	mental retardation, autosomal dominant 59	unclassified	probably causal	phenotype consistent with reported cases; bioinformatics predict that variant is a damaging, missense variant close to that in previously reported cases
G010-1	M	*SUZ12*	NP_056170.2:p.Arg654Ter	*de novo* heterozygous	Imagawa-Matsumoto syndrome	unclassified	probably causal	phenotype consistent with reported cases; bioinformatics predict that variant is damaging
G022-1	M	*NRXN1*	NP_001317007.1:p.Ile382Met	*de novo* heterozygous	affective psychosis with severe obsessive compulsive disorder, onset at about 9 years of age, refractive to treatment; cognitive deterioration	VUS	probably causal	*de novo* variant predicted to cause haploinsufficiency; very unusual phenotype consistent with that observed in some individuals with *NRXN1* haploinsufficiency reported with deletions
G025-1	M	*CLTC*	NP_004850.1:p.Pro890Leu	*de novo* heterozygous	mental retardation, autosomal dominant 56	unclassified	probably causal	phenotype consistent with reported cases; bioinformatics predict that variant is damaging
G027-1	M	*SMC1A*	NP_006297.2:p.Asp982Val	hemizygous (inherited from mosaic mother)	Cornelia de Lange syndrome 2	VUS	probably causal	phenotype consistent with Cornelia de Lange syndrome; bioinformatics predict that variant is damaging

(*Continued on next page*)

Table 1. ***Continued***

CAUSES ID no.	Sex	Gene	Variant	Mechanism	Disease	ACMG variant classification	Diagnostic interpretation by multidisciplinary research team	Rationale for diagnostic interpretation
G035-1	M	*MED12*	NP_005111.2:p.Arg91Leu	hemizygous (maternally inherited)	Ohdo syndrome	VUS	probably causal	phenotype consistent with Ohdo syndrome; similarly affected maternal uncle also carries variant; bioinformatics predict that variant is damaging
G036-1	M	*SATB2*	NM_001172509.2:c.474-3C>G	*de novo* heterozygous	Glass syndrome	VUS	probably causal	phenotype consistent with reported cases; bioinformatics predict that variant alters splicing
G044-1	F	*SCN8A*	NP_001317189.1:p.Cys324Tyr	*de novo* heterozygous	early infantile epileptic encephalopathy	VUS	probably causal	phenotype consistent with reported cases; bioinformatics predict that variant is damaging
G050-1	M	*EP300*	NP_001420.2:p.Gly2350HisfsTer52	heterozygous (maternally inherited)	Rubinstein-Taybi syndrome 2	VUS	definitely causal	phenotype in both proband and mother consistent with Rubinstein-Taybi syndrome; bioinformatics predict that variant is damaging
G053-1	M	*KCNQ5*	NP_062816.2:p.Val145Gly	*de novo* heterozygous	autosomal dominant mental retardation 46	unclassified	definitely causal	individual included in first published report of "new" genotype-phenotype association
G059-1	M	*SLC16A2*	NP_006508.2:p.Arg371Leu	hemizygous (maternally inherited)	Allan-Herndon-Dudley syndrome	VUS	definitely causal	phenotype characteristic of Allan-Herndon-Dudley syndrome
G066-1	M	*UBE2A*	NP_003327.2:p.Arg95Cys	hemizygous (maternally inherited)	mental retardation, X-linked syndromic, Nascimento type	VUS	probably causal	phenotype in both affected brothers consistent with reported cases; bioinformatics predict that variant is damaging
G066-4	M	*UBE2A*	NP_003327.2:p.Arg95Cys	hemizygous (maternally inherited)	mental retardation, X-linked syndromic, Nascimento type	VUS	probably causal	phenotype in both affected brothers consistent with reported cases; bioinformatics predict that variant is damaging

(*Continued on next page*)

Table 1. ***Continued***

CAUSES ID no.	Sex	Gene	Variant	Mechanism	Disease	ACMG variant classification	Diagnostic interpretation by multidisciplinary research team	Rationale for diagnostic interpretation
G070-1	M	*C2orf69*	NP_710156.3:p.Lys 282GlnfsTer55	homozygous	combined oxidative phosphorylation deficiency 53	VUS	probably causal	phenotype consistent with reported cases; consanguineous family with two early infant deaths in cousins and two additional early childhood deaths in other relatives; bioinformatics predict that variant is damaging
G073-1	M	*SOD1*	NP_000445.1:p. Ala124del	homozygous	progressive spastic tetraplegia and axial hypotonia	unclassified	probably causal	very characteristic phenotype consistent with reported cases; similarly affected sib died in childhood; bioinformatics predict that variant is damaging
G075-1	F	*MAST1*	NP_055790.1:p. Gly98Val	*de novo* heterozygous	mega-corpus-callosum syndrome with cerebellar hypoplasia and cortical malformations	unclassified	probably causal	individual included in first published report of "new" genotype-phenotype association
G077-1	F	*BPTF*	NP_872579.2:p. Arg653Ter	heterozygous (paternally inherited)	neurodevelopmental disorder with dysmorphic facies and distal limb anomalies	unclassified	probably causal	phenotype consistent with reported cases; bioinformatics predict that variant is damaging
G089-1	M	*NF1*	NP_001035957.1:p. Gly1190Val	heterozygous (maternally inherited)	neurofibromatosis 1	VUS	probably causal	phenotype consistent with neurofibromatosis 1; bioinformatics predict that variant is damaging
G091-1	F	*TKT*	NP_001055.1:p. Arg401His NP_001055.1:p. Tyr564del	compound heterozygous	short stature, developmental delay, and congenital heart defects	VUS VUS	definitely causal	phenotype in both sibs consistent with reported cases; bioinformatics predict that both variants are damaging; biochemical assay demonstrates transketolase deficiency

(*Continued on next page*)

Table 1. ***Continued***

CAUSES ID no.	Sex	Gene	Variant	Mechanism	Disease	ACMG variant classification	Diagnostic interpretation by multidisciplinary research team	Rationale for diagnostic interpretation
G091-4	M	*TKT*	NP_001055.1:p.Arg401His NP_001055.1:p.Tyr564del	compound heterozygous	short stature, developmental delay, and congenital heart defects	VUS VUS	definitely causal	phenotype in both sibs consistent with reported cases; bioinformatics predict that both variants are damaging; biochemical assay demonstrates transketolase deficiency
G092-1	M	*PAK3*	NP_002569.1:p.Ser105Cys	hemizygous (maternally inherited)	mental retardation, X-linked 30/47	VUS	probably causal	phenotype consistent with reported cases; bioinformatics predict that variant is damaging
G103-1	F	*DDX23*	NP_004809.2:p.Arg754Cys	*de novo* heterozygous	DDX23-related disorder	unclassified	probably causal	phenotype consistent with reported cases; bioinformatics predict that variant is damaging
G105-1	M	*SMS*	NP_004586.2:p.Met233Ile	hemizygous (maternally inherited)	X-linked mental retardation; Snyder-Robinson type	VUS	definitely causal	phenotype characteristic of Snyder-Robinson syndrome; bioinformatics predict that variant is damaging
G107-1	M	*SMARCA2*	NP_003061.3:p.Asp1571GlufsTer46	*de novo* heterozygous	Nicolaides-Baraister syndrome	VUS	probably causal	phenotype characteristic of Nicolaides-Baraister syndrome; bioinformatics predict that variant is damaging
G114-1	F	*REST*	NP_005603.3:p.Gln827Ter	*de novo* heterozygous	gingival fibromatosis	VUS	probably causal	phenotype characteristic of gingival fibromatosis; variant in last exon, where other variants that cause gingival fibromatosis lie; bioinformatics predict that variant is damaging
G117-1	F	*PIGG*	NP_001120650.1:p.Asn138Ser	homozygous	autosomal recessive mental retardation 53	VUS	probably causal	both sibs included in first published report of "new" genotype-phenotype association
G117-4	M	*PIGG*	NP_001120650.1:p.Asn138Ser	homozygous	autosomal recessive mental retardation 53	VUS	probably causal	both sibs included in first published report of "new" genotype-phenotype association

(*Continued on next page*)

Table 1. ***Continued***

CAUSES ID no.	Sex	Gene	Variant	Mechanism	Disease	ACMG variant classification	Diagnostic interpretation by multidisciplinary research team	Rationale for diagnostic interpretation
G122-1	F	*SRCAP*	NP_006653.2:p.Val1835ProfsTer13	*de novo* heterozygous	non-Floating-Harbor syndrome SRCAP-related neurodevelopmental disability	unclassified	probably causal	phenotype consistent with non-Floating-Harbor syndrome SRCAP-related neurodevelopmental disability; variant lies in exon 25, outside of exons 33–34 involved in Floating-Harbor syndrome; bioinformatics predict that variant is damaging
G125-1	M	*ZFYVE26*	NP_056161.2:p.Lys741ArgfsTer3	compound heterozygous	autosomal recessive spastic paraplegia 15	likely pathogenic	probably causal	phenotype characteristic of reported cases; compound heterozygote with one likely pathogenic variant and allelic very rare VUS with CADD = 26
			NP_056161.2:p.Arg2140Gln			VUS		
G134-1	F	*KCNQ2*	NP_742105.1:p.Arg144Trp	*de novo* heterozygous	early infantile epileptic encephalopathy 7	unclassified	definitely causal	phenotype consistent with reported cases; bioinformatics predict that variant is damaging
G141-1	F	*KMT2C*	NP_733751.2:p.Pro4843AlafsTer12	*de novo* heterozygous	Kleefstra syndrome 2	unclassified	probably causal	phenotype consistent with Kleefstra syndrome; bioinformatics predict that variant is damaging
G160-1	M	*CIC*	NP_001373227.1:p.Ala2056ProfsTer3	*de novo* heterozygous	autosomal dominant mental retardation 45	unclassified	probably causal	phenotype consistent with reported cases; bioinformatics predict that variant is damaging
G169-1	M	*HSD17B3*		compound heterozygous	pseudohermaphroditism, male, with gynecomastia	pathogenic	probably causal	phenotype consistent with reported cases; compound heterozygote with one likely pathogenic variant and allelic rare VUS predicted to be damaging
			NP_000188.1:p.Ala275Val			VUS		
		CTLA4	NP_005205.2:p.Ala54Thr	heterozygous (maternally inherited)	autoimmune lymphoproliferative syndrome, type V	VUS	probably causal	phenotype characteristic of autoimmune lymphoproliferative syndrome; bioinformatics predict that variant is damaging

(*Continued on next page*)

Table 1. ***Continued***

CAUSES ID no.	Sex	Gene	Variant	Mechanism	Disease	ACMG variant classification	Diagnostic interpretation by multidisciplinary research team	Rationale for diagnostic interpretation
G174-1	M	*SET*	NP_003002.2:p.Arg44LeufsTer10	heterozygous (maternally inherited)	autosomal dominant mental retardation 58	unclassified	probably causal	individual included in first published report of "new" genotype-phenotype association
G175-1	M	*ACTB*	NP_001092.1:p.Leu110ArgfsTer10	*de novo* heterozygous	Baraister-Winter syndrome 1	VUS	probably causal	phenotype characteristic of Baraister-Winter syndrome; bioinformatics predict that variant is damaging
G194-1	M	*MN1*	NP_002421.3:p.Trp1248Ter	*de novo* heterozygous	CEBALID syndrome	VUS	definitely causal	phenotype characteristic of CEBALID syndrome; bioinformatics predict that variant is damaging
G198-1	M	*ARID2*	NC_000012.11:g.46298857_46302229del	*de novo* heterozygous	Coffin-Siris syndrome 6	VUS	definitely causal	phenotype consistent with Coffin-Siris syndrome; bioinformatics predict that deletion is damaging
G202-1	M	*DLG4*	NP_001308004.1:p.Asn187ThrfsTer3	*de novo* heterozygous	intellectual developmental disorder 62	unclassified	probably causal	phenotype consistent with reported cases; bioinformatics predict that variant is damaging
G205-1	F	*NBEA*	NP_001371941.1:p.Glu2433ArgfsTer3	*de novo* heterozygous	neurodevelopmental disorder with or without early-onset generalized epilepsy	unclassified	probably causal	individual included in first published report of "new" genotype-phenotype association
G216-1	F	*COL12A1*	NM_004370.6:c.8319+1G>T	heterozygous (maternally inherited)	Bethlem myopathy	VUS	probably causal	phenotype consistent with Bethlem myopathy; bioinformatics predict that deletion is damaging; variant segregates with disease in family
G217-1	M	*KAT6B*	NP_036462.2:p.Arg153Gln	*de novo* heterozygous	SBBYSS syndrome	VUS	probably causal	phenotype consistent with SBBYSS syndrome; bioinformatics predict that variant is damaging
G218-1	M	*ASH1L*	NP_060959.2:p.Arg2691Ter	*de novo* heterozygous	autosomal dominant mental retardation 52	unclassified	probably causal	phenotype consistent with reported cases; bioinformatics predict that variant is damaging

(*Continued on next page*)

Table 1. ***Continued***

CAUSES ID no.	Sex	Gene	Variant	Mechanism	Disease	ACMG variant classification	Diagnostic interpretation by multidisciplinary research team	Rationale for diagnostic interpretation
G231-1	F	*WDR26*	NP_001366332.1:p.Ala150Val	*de novo* heterozygous	Skraban-Deardorff syndrome	VUS	probably causal	phenotype consistent with Skraban-Deardorff syndrome; bioinformatics predict that variant is damaging
G235-1	M	*TRAF7*	NP_115647.2:p.Arg524Trp	*de novo* heterozygous	cardiac, facial, and digital anomalies with developmental delay	unclassified	probably causal	phenotype consistent with reported cases; bioinformatics predict that variant is damaging
G239-1	M	*HNRNPU*	NP_114032.2:p.His451Pro	*de novo* heterozygous	early infantile epileptic encephalopathy 54	unclassified	probably causal	phenotype consistent with reported cases; bioinformatics predict that variant is damaging
G241-1	M	*SZT2*	NP_001352928.1:p.Glu2560SerfsTer92 NP_001352928.1:p.Thr3330Met	compound heterozygous	early infantile epileptic encephalopathy 18	likely pathogenic VUS	probably causal	phenotype consistent with reported cases; likely pathogenic variant allelic to VUS; bioinformatics predict that both variants are damaging
G248-1	F	*CRYBA2*	NP_476434.1:p.Gly65Arg	heterozygous (paternally inherited)	autosomal dominant cataract 42	unclassified	probably causal	phenotype and family history characteristic of autosomal dominant congenital cataracts; variant segregates with cataracts in family; bioinformatics predict that variant is damaging
G256-1	F	*KAT6A*	NP_006757.2:p.Pro933Ser	*de novo* heterozygous	autosomal dominant mental retardation 32	VUS	probably causal	phenotype consistent with reported cases; bioinformatics predict that variant may be damaging
G259-1	F	*WDR45*	NM_001029896.2:c.436+5G>C	*de novo* heterozygous	neurodegeneration with brain iron accumulation 5	VUS	definitely causal	phenotype characteristic of neurodegeneration with brain iron accumulation; bioinformatics predict that variant is damaging
G260-1	M	*MECP2*	NP_001104262.1:p.Gln256Leu	*de novo* hemizygous	X-linked intellectual disability disorder, Lubs type	VUS	probably causal	phenotype consistent with reported cases; bioinformatics predict that variant is damaging

(*Continued on next page*)

Table 1. ***Continued***

CAUSES ID no.	Sex	Gene	Variant	Mechanism	Disease	ACMG variant classification	Diagnostic interpretation by multidisciplinary research team	Rationale for diagnostic interpretation
G272-1	M	*SETD1B*	NM_001353345.2:c.5589+1G>A	*de novo* heterozygous	intellectual developmental disorder with seizures and language delay	unclassified	probably causal	phenotype consistent with reported cases; bioinformatics predict that variant is damaging
G280-1	F	*CWF19L1*	NP_060764.3:p.Glu384Ter	compound heterozygous	autosomal recessive spinocerebellar ataxia 17	likely pathogenic	probably causal	phenotype consistent with autosomal recessive spinocerebellar ataxia; likely pathogenic variant allelic to VUS; bioinformatics predict that both variants are damaging
			NP_060764.3:p.Glu519del			VUS		
G284-1	F	*ABL1*	NP_005148.2:p.Thr117Met	*de novo* heterozygous	congenital heart defects and skeletal malformation syndrome	VUS	probably causal	phenotype consistent with reported cases; bioinformatics predict that variant is damaging; functional studies support effect
G285-1	F	*GLRX5*	NP_057501.2:p.Met128Thr	homozygous	childhood-onset spasticity with hyperglycinemia	VUS	definitely causal	phenotype consistent with reported cases; bioinformatics predict that variant is damaging; biochemical studies consistent
G286-1	M	*IQCE*	NP_689771.3:p.Asp112ValfsTer2	compound heterozygous	post-axial polydactyly type A7	unclassified	probably causal	phenotype characteristic of post-axial polydactyly; bioinformatics predict that both variants are damaging
			NP_689771.3:p.Leu507AlafsTer10			unclassified		
G289-1	F	*CLCN4*	NP_001821.2:p.Gly182Ser	*de novo* heterozygous	Raynaud-Claes syndrome	VUS	probably causal	phenotype characteristic of Raynaud-Claes syndrome; bioinformatics predict that variant is damaging
G291-1	F	*KDM5A*	NC_000012.11:g.460661_470642del	homozygous	autosomal recessive mental retardation 65	VUS	probably causal	phenotype consistent with reported cases; bioinformatics predict that homozygous variant is damaging; homozygous variant also found in similarly affected sib

(*Continued on next page*)

Table 1. ***Continued***

CAUSES ID no.	Sex	Gene	Variant	Mechanism	Disease	ACMG variant classification	Diagnostic interpretation by multidisciplinary research team	Rationale for diagnostic interpretation
G291-4	M	*KDM5A*	NC_000012.11:g.460661_470642del	homozygous	autosomal recessive mental retardation 65	VUS	probably causal	phenotype consistent with reported cases; bioinformatics predict that homozygous variant is damaging; homozygous variant also found in similarly affected sib
G292-1	M	*SYT1*	NP_005630.1:p.Pro180Leu	*de novo* heterozygous	Baker-Gordon syndrome	unclassified	probably causal	phenotype consistent with Baker-Gordon syndrome; bioinformatics predict that variant is damaging
G297-1	M	*JARID2*	NP_004964.2:p.Arg1127Ter	*de novo* hemizygous	*JARID2*-neurodevelopmental syndrome	unclassified	probably causal	individual included in first published report of "new" genotype-phenotype association
G312-1	F	*NAA10*	NP_003482.1:p.Asn101Lys	*de novo* heterozygous	Ogden syndrome	VUS	probably causal	phenotype consistent with Ogden syndrome; bioinformatics predict that variant is damaging
G323-1	M	*GABBR2*	NP_005449.5:p.Pro282Leu	*de novo* heterozygous	neurodevelopmental disorder with poor language and loss of hand skills	VUS	probably causal	phenotype consistent with reported cases; bioinformatics predict that variant is damaging
G336-1	M	*KCNB1*	NP_004966.1:p.Glu71Ter	heterozygous (maternally inherited)	early infantile epileptic encephalopathy 26	VUS	probably causal	phenotype and family history consistent with reported cases; bioinformatics predict that variant is damaging
G338-1	F	*KYNU*	NP_003928.1:p.Lys121del NP_003928.1:p.Ser345Arg	compound heterozygous	vertebral, cardiac, renal, and limb defects syndrome 2	VUS VUS	probably causal	phenotype characteristic of vertebral, cardiac, renal, and limb defects syndrome; variants are allelic and bioinformatics predict that both are damaging; functional studies demonstrated significant reduction in NAD levels
		PLEKHA7	NP_001316559.1:p.Asp191Asn	heterozygous (paternally inherited)	Mendelian non-syndromic cleft lip with or without cleft palate	unclassified	probably causal	phenotype consistent with reported cases; both sibs have orofacial clefting and variant; bioinformatics predict that variant is damaging

(*Continued on next page*)

Table 1. ***Continued***

CAUSES ID no.	Sex	Gene	Variant	Mechanism	Disease	ACMG variant classification	Diagnostic interpretation by multidisciplinary research team	Rationale for diagnostic interpretation
G338-4	F	*PLEKHA7*	NP_001316559.1:p.Asp191Asn	heterozygous (paternally inherited)	Mendelian non-syndromic cleft lip with or without cleft palate	unclassified	probably causal	phenotype consistent with reported cases; both sibs have orofacial clefting and variant; bioinformatics predict that variant is damaging
G345-1	M	*HISG1H1E*	NP_005312.1:p.Gly124ArgfsTer71	*de novo* heterozygous	Rahman syndrome	VUS	definitely causal	phenotype consistent with Rahman syndrome; bioinformatics predict that variant is damaging
G350-1	F	*TSC2*	NP_000539.2:p.His1543Arg	heterozygous (paternally inherited)	tuberous sclerosis 2	VUS	definitely causal	phenotype characteristic of tuberous sclerosis; bioinformatics predict that variant is damaging
G356-1	F	*COL4A3BP*	NP_001365958.1:p.Thr251Ala	*de novo* heterozygous	autosomal dominant mental retardation 34	unclassified	probably causal	phenotype consistent with reported cases; bioinformatics predict that variant is damaging
G363-1	M	*FLNA*	NM_001110556.2:c.4475-1G>T	hemizygous (maternally inherited)	periventricular nodular heterotopia I	VUS	probably causal	phenotype consistent with periventricular nodular heterotopia; bioinformatics predict that variant is damaging
G368-1	M	*KDM5C*	NP_004178.2:p.Ser285Leu	hemizygous (maternally inherited)	X-linked syndromic mental retardation, Claes-Jensen type	VUS	probably causal	phenotype consistent with X-linked syndromic mental retardation, Claes-Jensen type; bioinformatics predict that variant may be damaging
G370-1	M	*RYR1*	NP_000531.2:p.Glu2987Gly NP_000531.2:p.Asp4505His	compound heterozygous	King-Denborough syndrome	VUS VUS	probably causal	phenotype characteristic of King-Denborough syndrome; variants are allelic and bioinformatics predict that both are damaging
G385-1	M	*CAMK2*	NP_057065.2:p.Ser341Thr	*de novo* heterozygous	autosomal dominant mental retardation 53	VUS	probably causal	phenotype consistent with reported cases; bioinformatics predict that variant is damaging

(*Continued on next page*)

Table 1. ***Continued***

CAUSES ID no.	Sex	Gene	Variant	Mechanism	Disease	ACMG variant classification	Diagnostic interpretation by multidisciplinary research team	Rationale for diagnostic interpretation
G392-1	M	*KCNK9*	NP_001269463.1:p.Tyr205Cys	heterozygous (maternally inherited)	Birk-Barel mental retardation dysmorphism syndrome	VUS	probably causal	phenotype and family history consistent with Birk-Barel mental retardation dysmorphism syndrome; bioinformatics predict that variant is damaging
G393-1	M	*ERCC8*	NP_000073.1:p.Val362PhefsTer20	homozygous	Cockayne syndrome, type A	VUS	definitely causal	phenotype characteristic of Cockayne syndrome; parental consanguinity; bioinformatics predict that variant is damaging
G396-1	M	*CHD3*	NP_001005273.1:p.Arg1172Gln	*de novo* heterozygous	Snijders Blok-Campeau syndrome	unclassified	definitely causal	phenotype characteristic of Snijders Blok-Campeau syndrome; bioinformatics predict that variant is damaging
G401-1	F	*CTCF*	NP_006556.1:p.Asp357Asn	*de novo* heterozygous	autosomal dominant mental retardation 21	VUS	probably causal	phenotype consistent with reported cases; bioinformatics predict variant is damaging
G402-4	F	*SMARCC2*	NP_001317217.1:p.Tyr679Ter	*de novo* heterozygous	Coffin-Siris syndrome 8	VUS	probably causal	phenotype consistent with reported cases; bioinformatics predict variant is damaging
G404-1	M	*HNRNPU*	NP_114032.2:p.Pro506Leu	*de novo* heterozygous	early infantile epileptic encephalopathy 54	VUS	probably causal	phenotype consistent with reported cases; bioinformatics predict variant is damaging
G406-1	F	*H3F3A*	NP_002098.1:p.Thr23Ile	*de novo* heterozygous	Bryant-Li-Bhoj neurodevelopmental syndrome 1	VUS	probably causal	individual included in first published report of "new" genotype-phenotype association
G407-1	M	*JARID2*	NP_004964.2:p.Ile557ArgfsTer34	*de novo* heterozygous	JARID2-neurodevelopmental disorder	unclassified	probably causal	individual included in first published report of "new" genotype-phenotype association
G421-1	F	*NEUROD2*	NP_006151.3:p.Glu130Lys	*de novo* heterozygous	early infantile epileptic encephalopathy 72	unclassified	probably causal	phenotype consistent with reported cases; bioinformatics predict variant is damaging

(*Continued on next page*)

Table 1. ***Continued***

CAUSES ID no.	Sex	Gene	Variant	Mechanism	Disease	ACMG variant classification	Diagnostic interpretation by multidisciplinary research team	Rationale for diagnostic interpretation
G422-1	M	*DLL1*	NP_005609.3:p.Arg509Ter	heterozygous (paternally inherited)	neurodevelopmental disorder with non-specific brain abnormalities and with or without seizures	unclassified	definitely causal	individual included in first published report of "new" genotype-phenotype association
G447-1	M	*SLC6A8*	NP_005620.1:p.Phe248del	*de novo* hemizygous (mosaic)	cerebral creatine deficiency syndrome 1	VUS	definitely causal	phenotype consistent with reported cases; bioinformatics predict variant is damaging
G462-1	F	*ASH1L*	NP_060959.2:p.Glu1956Lys	*de novo* heterozygous	autosomal dominant mental retardation 52	VUS	probably causal	phenotype consistent with reported cases; bioinformatics predict variant is damaging
G465-1	F	*ANKRD17*	NP_115593.3:p.Gln1787ArgfsTer5	*de novo* heterozygous	Chopra-Amiel-Gordon syndrome	unclassified	definitely causal	individual included in first published report of "new" genotype-phenotype association
G468-1	F	*GNAO1*	NP_066268.1:p.Asp151Asn	heterozygous (inherited from mosaic mother)	early infantile epileptic encephalopathy 17	VUS	probably causal	phenotype consistent with reported cases; bioinformatics predict variant is damaging
		PTCH1	NM_000264.5:c.654+3A>G	*de novo* heterozygous	basal cell nevus syndrome	VUS	probably causal	phenotype consistent with basal cell nevus syndrome; bioinformatics predict variant is damaging
G472-1	F	*GNB1*	NP_002065.1:p.Gly282Arg	*de novo* heterozygous	autosomal dominant mental retardation 42	VUS	probably causal	phenotype consistent with reported cases; bioinformatics predict variant is damaging
G480-1	M	*MSX1*	NP_002439.2:p.Trp219Arg	heterozygous (paternally inherited)	orofacial cleft 5	VUS	probably causal	phenotype and family history typical of hereditary orofacial clefting; bioinformatics predict variant is damaging; variant segregates with phenotype in family
G482-1	F	*SETD1B*	NP_001340274.1:p.Gln1322Ter	*de novo* heterozygous	intellectual disability, epilepsy, and autism	unclassified	probably causal	phenotype consistent with reported cases; bioinformatics predict variant is damaging

(*Continued on next page*)

Table 1. ***Continued***

CAUSES ID no.	Sex	Gene	Variant	Mechanism	Disease	ACMG variant classification	Diagnostic interpretation by multidisciplinary research team	Rationale for diagnostic interpretation
G487-1	F	*DDX3X*	NP_001347.3:p.Ile190Ser	*de novo* heterozygous	X-linked mental retardation 102	VUS	definitely causal	phenotype consistent with reported cases; bioinformatics predict variant is damaging
G494-1	M	*EDEM3*	NP_079467.3:p.Arg314Ter	homozygous (maternally inherited isodisomy)	EDEM3-related disorder	unclassified	probably causal	individual included in first published report of "new" genotype-phenotype association
G498-1	F	*TANC2*	NP_079461.2:p.Arg1770Gly	*de novo* heterozygous	global developmental delay, cerebellar atrophy, and dysmorphic features (non-clinome)	unclassified	probably causal	phenotype consistent with reported cases; bioinformatics predict variant is damaging
G504-1	F	*TRRAP*	NP_001362453.1:p.Thr10Met	*de novo* heterozygous	developmental delay with or without dysmorphic facies and autism	VUS	probably causal	individual included in first published report of "new" genotype-phenotype association
G508-1	M	*ERCC2*	NP_000391.1:p.Leu581Pro NP_000391.1:p.Arg658Cys	Compound heterozygous	trichothiodystrophy	pathogenic VUS	definitely causal	phenotype consistent with trichothiodystrophy; VUS allelic to pathogenic variant and predicted to be damaging
G536-1	M	*NF1*	NM_001042492.3:c.5609+1G>T	*de novo* heterozygous	neurofibromatosis 1	unclassified	probably causal	phenotype consistent with neurofibromatosis 1; variant predicted as damaging; RNA studies demonstrated disruption of canonical splice site
G553-1	F	*TRIP12*	NP_001335252.1:p.Tyr1744Asp	heterozygous (inherited from mosaic father)	autosomal dominant mental retardation 49	VUS	probably causal	phenotype consistent with reported cases; bioinformatics predict variant is damaging
G553-4	M	*TRIP12*	NP_001335252.1:p.Tyr1744Asp	heterozygous (inherited from mosaic father)	autosomal dominant mental retardation 49	VUS	probably causal	phenotype consistent with reported cases; bioinformatics predict variant is damaging
G558-1	F	*CDH2*	NP_001783.2:p.Asp627Tyr	*de novo* heterozygous	syndromic neurodevelopmental disorder	unclassified	definitely causal	phenotype consistent with newly described disorder, bioinformatics predict variant is damaging

(*Continued on next page*)

Table 1. ***Continued***

CAUSES ID no.	Sex	Gene	Variant	Mechanism	Disease	ACMG variant classification	Diagnostic interpretation by multidisciplinary research team	Rationale for diagnostic interpretation
G559-1	F	*MPZL2*	NP_005788.1:p.Ile24MetfsTer22	compound heterozygous	autosomal recessive deafness 111	pathogenic	probably causal	phenotype consistent with reported cases; VUS allelic to pathogenic variant and predicted to be damaging
			NP_005788.1:p.Asp93Val			VUS		
G561-1	F	*ETV6*	NP_001978.1:p.Lys409Glu	heterozygous (maternally inherited)	thrombocytopenia 5	VUS	probably causal	phenotype and family history consistent with thrombocytopenia 5; bioinformatics predict that variant is damaging; variant segregates with disease in family
G563-1	F	*MTHFS*	NP_006432.1:p.Ala9GlyfsTer42	homozygous	neurodevelopmental disorder with microcephaly, epilepsy, and hypomyelination	unclassified	probably causal	phenotype consistent with neurodevelopmental disorder with microcephaly, epilepsy, and hypomyelination; bioinformatics predict variant is damaging
G575-1	M	*COL5A1*	NP_000084.3:p.Pro1566Leu	heterozygous (maternally inherited)	Ehlers-Danlos syndrome type 1	VUS	probably causal	phenotype consistent with Ehlers-Danlos syndrome; bioinformatics predict variant is damaging

Table 2. Diagnostic reinterpretation[a]

CAUSES ID no.	Sex	Exome or genome sequencing	Gene	Variant (HGVS cDNA nomenclature)	Variant (HGVS protein nomenclature)	Diagnostic interpretation by multidisciplinary research team		ACMG variant classification	Mechanism	Disease (phenotype OMIM no.)	Reason for reinterpretation
						Initial	Final				
G004-1	F	exome	*TBCK*	NM_001163435.3:c.783G>T; NM_001163435.3:c.2060-2A>G	NP_001156907.2:p.Arg261Ser	uninformative	probably causal	likely pathogenic; VUS	compound heterozygous	infantile hypotonia with psychomotor retardation and characteristic facies 3 (616900)	newly described disorder
G007-1	M	exome	*TUBA1A*	NM_006009.4:c.1177C>T	NP_006000.2:p.His393Tyr	uninformative	probably causal	likely pathogenic	*de novo* heterozygous	lissencephaly, AD 3 (611603)	improvement in bioinformatics pipeline
G010-1	M	exome	*SUZ12*	NM_015355.4:c.1960C>T	NP_056170.2:p.Arg654Ter	uncertain	probably causal	unclassified	*de novo* heterozygous	Imagawa-Matsumoto syndrome (618786)	newly described disorder
G025-1	M	exome	*CLTC*	NM_004859.4:c.2669C>T	NP_004850.1:p.Pro890Leu	uncertain	probably causal	unclassified	*de novo* heterozygous	mental retardation, autosomal dominant 56 (617814)	newly described disorder
G036-1	M	exome	*SATB2*	NM_001172509.2:c.474-3C>G	–	uninformative	probably causal	VUS	*de novo* heterozygous	Glass syndrome (612313)	improvement in bioinformatics pipeline
G040-1	M	exome	*ABCB7*	NM_001271696.3:c.1235T>C	NP_001258625.1:p.Met412Thr	uninformative	definitely causal	likely pathogenic	*de novo* hemizygous	sideroblastic anemia and spinocerebellar ataxia (301310)	expansion of phenotype
G063-1	M	exome	*POLR2A*	NM_000937.5:c.3373_3375del	NP_000928.1:p.Lys1125del	uninformative	probably causal	likely pathogenic	*de novo* heterozygous	neurodevelopmental disorder with hypotonia and variable intellectual and behavioral abnormalities (618603)	newly described disorder
G067-1	M	exome	*EBF3*	NM_001375380.1:c.616C>T	NP_001362309.1:p.Arg206Ter	uninformative	definitely causal	likely pathogenic	heterozygous (inherited from mosaic parent)	hypotonia, ataxia, and delayed development syndrome (617330)	new publication
G067-4	F	exome	*EBF3*	NM_001375380.1:c.616C>T	NP_001362309.1:p.Arg206Ter	uninformative	definitely causal	likely pathogenic	heterozygous (inherited from mosaic parent)	hypotonia, ataxia, and delayed development syndrome (617330)	new publication
G070-1	M	exome	*C20rf69*	NM_153689.6:c.843_847del	NP_710156.3:p.Lys282GlnfsTer55	uninformative	probably causal	VUS	homozygous	combined oxidative phosphorylation deficiency 53 (619423)	new publications
G073-1	M	exome	*SOD1*	NM_000454.5:c.371_373del	NP_000445.1:p.Ala124del	uninformative	probably causal	unclassified	homozygous	progressive spastic tetraplegia and axial hypotonia (618598)	new publication
G075-1	F	exome	*MAST1*	NM_014975.3:c.293G>T	NP_055790.1:p.Gly98Val	uninformative	probably causal	unclassified	*de novo* heterozygous	mega-corpus-callosum syndrome with cerebellar hypoplasia and cortical malformations (618273)	new publication (includes this individual)
G077-1	F	exome	*BPTF*	NM_182641.4:c.1957A>T	NP_872579.2:p.Arg653Ter	uninformative	probably causal	unclassified	heterozygous (paternally inherited)	neurodevelopmental disorder with dysmorphic facies and distal limb anomalies (617755)	new publication
G078-1	F	exome	*ASH1L*	NM_018489.3:c.3664_3667del	NP_060959.2:p.Lys1222GlyfsTer10	uninformative	definitely causal	pathogenic	*de novo* heterozygous	autosomal dominant mental retardation 52 (617796)	new publication

(*Continued on next page*)

Table 2. ***Continued***

CAUSES ID no.	Sex	Exome or genome sequencing	Gene	Variant (HGVS cDNA nomenclature)	Variant (HGVS protein nomenclature)	Diagnostic interpretation by multidisciplinary research team		ACMG variant classification	Mechanism	Disease (phenotype OMIM no.)	Reason for reinterpretation
						Initial	Final				
G082-1	M	exome	*ASH1L*	NM_018489.3:c.6803_6804delinsTTCTCA	NP_060959.2:p.Cys2268PhefsTer7	uninformative	definitely causal	pathogenic	*de novo* heterozygous	autosomal dominant mental retardation 52 (617796)	new publication
G089-1	M	exome	*NF1*	NM_001042492.3:c.3569G>T	NP_001035957.1:p.Gly1190Val	uninformative	probably causal	VUS	heterozygous (maternally inherited)	neurofibromatosis 1 (162200)	improvement in bioinformatics pipeline
G092-1	M	exome	*PAK3*	NM_002578.5:c.314C>G	NP_002569.1:p.Ser105Cys	uncertain	probably causal	VUS	hemizygous (maternally inherited)	intellectual developmental disorder, X-linked 30 (300558)	referring physician's interpretation based on patient phenotype
G103-1	F	exome	*DDX23*	NM_004818.3:c.2260C>T	NP_004809.2:p.Arg754Cys	uninformative	probably causal	unclassified	*de novo* heterozygous	DDX23-related disorder	new publication
			ATP2A2	NM_170665.4:c.2684C>T	NP_733765.1:p.Pro895Leu	definitely causal	uncertain	likely pathogenic	*de novo* heterozygous	Darier disease (124200)	*DDX23* variant provides better fit to phenotype
G117-1	F	exome	*PIGG*	NM_001127178.3:c.413A>G	NP_001120650.1:p.Asn138Ser	uncertain	probably causal	VUS	homozygous	autosomal recessive mental retardation 53 (616917)	new publication (includes this individual)
G117-4	M	exome	*PIGG*	NM_001127178.3:c.413A>G	NP_001120650.1:p.Asn138Ser	uncertain	probably causal	VUS	homozygous	autosomal recessive mental retardation 53 (616917)	new publication (includes this individual)
G122-1	F	exome	*SRCAP*	NM_006662.3:c.5503_5515del	NP_006653.2:p.Val1835ProfsTer13	uncertain	probably causal	unclassified	*de novo* heterozygous	developmental delay, hypotonia, musculoskeletal defects, and behavioral abnormalities (619595)	new publication
G164-1	F	exome	*MECP2*	NM_001110792.2:c.1200_1243del	NP_001104262.1:p.Pro401Ter	uninformative	probably causal	pathogenic	*de novo* heterozygous	Rett syndrome (312750)	improvement in bioinformatics pipeline
G169-1	M	exome	*HSD17B3*	NM_000197.2:c.277+4A>T; NM_000197.2:c.824C>T	NP_000188.1:p.Ala275Val	uninformative	probably causal	pathogenic; VUS	compound heterozygous	pseudohermaphroditism, male, with gynecomastia (264300)	improvement in bioinformatics pipeline
			CTLA4	NM_005214.5:c.160G>A	NP_005205.2:p.Ala54Thr	uninformative	probably causal	VUS	heterozygous (maternally inherited)	autoimmune lymphoproliferative syndrome, type V (616100)	referring physician's interpretation based on patient phenotype
G174-1	M	exome	*SET*	NM_003011.4:c.128_131del	NP_003002.2:p.Arg44LeufsTer10	uninformative	probably causal	unclassified	heterozygous (maternally inherited)	autosomal dominant mental retardation 58 (618106)	new publication (includes this individual)
G194-1	M	exome	*MN1*	NM_002430.3:c.3743G>A	NP_002421.3:p.Trp1248Ter	uncertain	definitely causal	VUS	*de novo* heterozygous	CIEBALID syndrome (618774)	new publication
G198-1	M	genome	*ARID2*	NC_000012.11:g.46298857_46302229del	–	uncertain	definitely causal	VUS	*de novo* heterozygous	Coffin-Siris syndrome 6 (617808)	referring physician's interpretation based on patient phenotype

(*Continued on next page*)

Table 2. ***Continued***

CAUSES ID no.	Sex	Exome or genome sequencing	Gene	Variant (HGVS cDNA nomenclature)	Variant (HGVS protein nomenclature)	Diagnostic interpretation by multidisciplinary research team		ACMG variant classification	Mechanism	Disease (phenotype OMIM no.)	Reason for reinterpretation
						Initial	Final				
G205-1	F	exome	*NBEA*	NM_001385012.1:c.7294_7295dup	NP_001371941.1:p.Glu2433ArgfsTer3	uninformative	probably causal	unclassified	*de novo* heterozygous	neurodevelopmental disorder with or without early-onset generalized epilepsy (619157)	new publication (includes this individual)
G231-1	F	exome	*WDR26*	NM_001379403.1:c.449C>T	NP_001366332.1:p.Ala150Val	uninformative	probably causal	VUS	*de novo* heterozygous	Skraban-Deardorff syndrome (617616)	new publication
G235-1	M	exome	*TRAF7*	NM_032271.3:c.1570C>T	NP_115647.2:p.Arg524Trp	uninformative	probably causal	unclassified	*de novo* heterozygous	cardiac, facial, and digital anomalies with developmental delay (618164)	new publication
G292-1	M	genome	*SYT1*	NM_005639.3:c.539C>T	NP_005630.1:p.Pro180Leu	uncertain	probably causal	unclassified	*de novo* heterozygous	Baker-Gordon syndrome (618218)	new publication
G297-1		exome	*JARID2*	NM_004973.4:c.3379C>T	NP_004964.2:p.Arg1127Ter	uncertain	probably causal	unclassified	*de novo* hemizygous	*JARID2*-neurodevelopmental syndrome	new publication (includes this individual)
G328-1	F	exome	*RRAS2*	NM_012250.6:c.68G>A	NP_036382.2:p.Gly23Asp	uninformative	definitely causal	pathogenic	*de novo* heterozygous	Noonan syndrome 12 (618624)	new publication
G338-1	F	exome	*PLEKHA7*	NM_001329630.2:c.571G>A	NP_001316559.1:p.Asp191Asn	uninformative	probably causal	unclassified	heterozygous (paternally inherited)	Mendelian non-syndromic cleft lip with or without cleft palate	referring physician's interpretation based on patient phenotype
G338-4	F	exome	*PLEKHA7*	NM_001329630.2:c.571G>A	NP_001316559.1:p.Asp191Asn	uninformative	probably causal	unclassified	heterozygous (paternally inherited)	Mendelian non-syndromic cleft lip with or without cleft palate	referring physician's interpretation based on patient phenotype
G363-1	M	exome	*FLNA*	NM_001110556.2:c.4475-1G>T	–	uncertain	probably causal	VUS	hemizygous (maternally inherited)	periventricular nodular heterotopia I (300049)	referring physician's interpretation based on patient phenotype
G368-1	M	genome	*KDM5C*	NM_004187.5:c.854C>T	NP_004178.2:p.Ser285Leu	uncertain	probably causal	VUS	hemizygous (maternally inherited)	X-linked syndromic mental retardation, Claes-Jensen type (300534)	referring physician's interpretation based on patient phenotype
G369-1	M	exome	*TBX1*	NM_001379200.1:c.901G>A	NP_001366129.1:p.Ala301Thr	probably causal	uncertain	VUS	*de novo* heterozygous	tetralogy of Fallot (187500)	referring physician's interpretation based on patient phenotype
G402-4	F	genome	*SMARCC2*	NM_001330288.2:c.2037C>A	NP_001317217.1:p.Tyr679Ter	uncertain	probably causal	VUS	*de novo* heterozygous	Coffin-Siris syndrome 8 (618362)	new publication
G406-1	F	exome	*H3F3A*	NM_002107.7:c.68C>T	NP_002098.1:p.Thr23Ile	uninformative	probably causal	VUS	*de novo* heterozygous	H3F3A-related disorder	new publication (includes this individual)
G407-1	M	exome	*JARID2*	NM_004973.4:c.1668_1669dup	NP_004964.2:p.Ile557ArgfsTer34	uncertain	probably causal	unclassified	*de novo* heterozygous	JARID2-neurodevelopmental disorder	new publication
G421-1	F	exome	*NEUROD2*	NM_006160.4:c.388G>A	NP_006151.3:p.Glu130Lys	uninformative	probably causal	unclassified	*de novo* heterozygous	developmental and epileptic encephalopathy 72 (618374)	new publication

(*Continued on next page*)

Table 2. ***Continued***

CAUSES ID no.	Sex	Exome or genome sequencing	Gene	Variant (HGVS cDNA nomenclature)	Variant (HGVS protein nomenclature)	Diagnostic interpretation by multidisciplinary research team		ACMG variant classification	Mechanism	Disease (phenotype OMIM no.)	Reason for reinterpretation
						Initial	Final				
G422-1	M	exome	*DLL1*	NM_005618.4:c.1525C>T	NP_005609.3:p.Arg509Ter	uninformative	definitely causal	unclassified	heterozygous (paternally inherited)	variable neurodevelopmental disorder with multisystem features (618709)	new publication (includes this individual)
G465-1	F	exome	*ANKRD17*	NM_032217.5:c.5360_5363del	NP_115593.3:p.Gln1787ArgfsTer5	uncertain	definitely causal	unclassified	*de novo* heterozygous	Chopra-Amiel-Gordon syndrome (619504)	new publication (includes this individual)
G469-1	M	exome	*BCL11B*	NM_138576.4:c.726_727ins CGCAGCAC	NP_612808.1:p.Thr243ArgfsTer41	uncertain	definitely causal	likely pathogenic	heterozygous (inherited from mosaic father)	intellectual developmental disorder with dysmorphic facies, speech delay, and T cell abnormalities (618092)	new publication
G483-1	F	exome	*ALPL*	NM_000478.5:c.407G>A	NP_000469.3:p.Arg136His	definitely causal	uncertain	pathogenic	heterozygous (paternally inherited)	hypophosphatasia, adult (146300)	referring physician's interpretation based on patient phenotype
G494-1	M	exome	*EDEM3*	NM_025191.4:c.940A>T	NP_079467.3:p.Arg314Ter	uncertain	probably causal	unclassified	homozygous (maternally inherited chromosome 1 uniparental disomy)	congenital disorder of glycosylation, type 2V (619493)	new publication (includes this individual)
G495-1	M	exome	*BRD4*	NM_001379291.1:c.1339C>T	NP_001366220.1:p.Gln447Ter	uncertain	probably causal	pathogenic	*de novo* heterozygous	BRD4-related disorder	newly described disorder
G498-1	F	exome	*TANC2*	NM_025185.3:c.5308A>G	NP_079461.2:p.Arg1770Gly	uncertain	probably causal	unclassified	*de novo* heterozygous	intellectual developmental disorder with autistic features and language delay, with or without seizures (618906)	new publication
G504-1	F	exome	*TRRAP*	NM_001375524.1:c.29C>T	NP_001362453.1:p.Thr10Met	uninformative	probably causal	VUS	*de novo* heterozygous	developmental delay with or without dysmorphic facies and autism (618454)	new publication (includes this individual)
G536-1	M	exome	*NF1*	NM_001042492.3:c.5609+1G>T	–	uncertain	probably causal	unclassified	*de novo* heterozygous	neurofibromatosis 1 (162200)	referring physician's interpretation based on patient phenotype
G550-1	F	exome	*USP7*	NM_003470.2:c.963delC	NP_003461.2:p.Lys322AsnfsTer16	probably causal	uncertain	unclassified	*de novo* heterozygous	Hao-Fountain syndrome (616863)	referring physician's interpretation based on patient phenotype
G558-1	F	exome	*CDH2*	NM_001792.5:c.1879G>T	NP_001783.2:p.Asp627Tyr	uninformative	definitely causal	unclassified	*de novo* heterozygous	agenesis of corpus callosum, cardiac, ocular, and genital syndrome (618929)	new publication

(*Continued on next page*)

Table 2. ***Continued***

CAUSES ID no.	Sex	Exome or genome sequencing	Gene	Variant (HGVS cDNA nomenclature)	Variant (HGVS protein nomenclature)	Diagnostic interpretation by multidisciplinary research team		ACMG variant classification	Mechanism	Disease (phenotype OMIM no.)	Reason for reinterpretation
						Initial	Final				
G559-1	F	exome	*MPZL2*	NM_005797.4:c.72del; NM_005797.4:c.278A>T	NP_005788.1: p.Ile24MetfsTer22; NP_005788.1: p.Asp93Val	uninformative	probably causal	pathogenic; VUS	compound heterozygous	autosomal recessive deafness 111 (618145)	newly described disorder
G563-1	F	exome	*MTHFS*	NM_006441.4:c.10_25dup	NP_006432.1: p.Ala9GlyfsTer42	uninformative	probably causal	unclassified	homozygous	neurodevelopmental disorder with microcephaly, epilepsy, and hypomyelination	newly described disorder
G575-1	M	exome	*COL5A1*	NM_000093.5:c.4697C>T	NP_000084.3: p.Pro1566Leu	uncertain	probably causal	VUS	heterozygous (maternally inherited)	Ehlers-Danlos syndrome, classic type 1 (130000)	referring physician's interpretation based on patient phenotype
G578-1	F	genome	*SPG7*	NM_003119.3:c.1226A>C	NP_003110.1: p.Glu409Ala	uncertain	probably causal	likely pathogenic	heterozygous (inherited from mosaic mother)	"autosomal recessive" spastic paraplegia 7 (607259)	referring physician's interpretation based on patient phenotype

[a]Variants that were interpreted as "uninformative" were not Sanger sequenced. Sanger sequencing and ACMG classification of variants were obtained when the individual was diagnosed by our multidisciplinary research team as having a genetic disease associated with a variant that was probably or definitely causal or of uncertain, but suspected, relationship to the phenotype.

similarity to a recently published series of nine individuals with similar *de novo* missense variants of the highly conserved DEAD box domain of DDX23.[41]

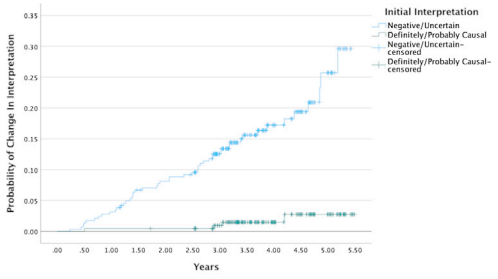
In 49 (17.2%) of the 285 families in whom our multidisciplinary research team initially considered the GWS results to be either uninformative or uncertain, a genetic condition was diagnosed during follow-up when the associated variant was reinterpreted as probably or definitely disease causing (Table 2). One of these families (G103, discussed above) was classified as definitely causal initially, but for a different gene. The probability of obtaining the diagnosis of genetic disease in uninformative or uncertain cases that were subsequently reinterpreted as having a definitely or probably causal variant averaged 4.8% per year and continued at a similar rate throughout the study (Figure 1).

A total of 27 families that were initially interpreted as uninformative or uncertain were subsequently diagnosed with a genetic disease, and the associated genomic variants were reinterpreted as definitely or probably disease-causing on the basis of new publications that described genetic disorders that were unrecognized at the time of initial analysis. Eight of these publications included one or more CAUSES individuals. For nine individuals, clinical reassessment by the referring physician after learning of the GWS result led to the diagnosis of a genetic condition and reinterpretation of the variant as probably or definitely disease-causing. For seven individuals, improvement in the bioinformatics pipeline identified a variant on routine reanalysis that had not been flagged initially but was interpreted as probably or definitely causal for a genetic disease in the individual by our multidisciplinary research team. Five individuals were diagnosed as having a genetic disease that had recently been listed in OMIM when definitely or probably causal variants were identified on routine GWS data reanalysis. For the remaining individual, a genetic disorder was diagnosed after routine reanalysis identified a variant in a locus that had recently been reported to be associated with a broader phenotype than initially recognized.

"False negative"/uninformative diagnoses

In seven families who underwent trio ES, a diagnosis was not established through CAUSES but was identified by means of other testing initiated by the affected individual's physician (Table 3). In three families, additional clinical testing led to the diagnosis: clinical ES with deletion/duplication (del/dup) analysis in two families (G012, G139) and a clinical multigene panel analysis that found a disease-causing variant in exon 1 of *SEPN1* in the third family (G410). Del/dup testing was not performed as part of exome analysis in the CAUSES study, and *SEPN1* exon 1 was poorly covered in the CAUSES exome dataset.

Three individuals who had uninformative ES in CAUSES had disease-causing variants identified by research GS. In one individual, a deep intronic variant was found in the second allele of a recessive locus that appeared to have



Initial Interpretation	Years	0.0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	5.5
Negative/ Uncertain	Number of families	285	281	276	262	257	252	203	128	87	60	29	0
	Number changed to probably/ definitely causal		4	9	19	23	27	35	41	43	45	48	49
	Number censored		0	0	4	5	6	47	116	155	180	208	236
Probably/ Definitely Causal	Number of families	216	216	215	215	215	215	189	131	90	60	36	0
	Number changed to negative/uncertain		0	1	1	1	1	2	3	3	4	4	4
	Number censored		0	0	0	0	0	25	82	123	152	176	212

Figure 1. Follow-up and reclassification of CAUSES participants
Probability of reclassification of 285 cases that were initially uninformative or uncertain (blue line) and of 216 cases that were initially classified as positive (probably or definitely causal variants). For the purposes of this analysis, a "case" is defined as one gene in one subject (proband or sib) that was classified as positive initially or at the end of the study (or both) or one subject in whom no gene was classified as positive at any time during the study.

only one disease-causing variant in the CAUSES exome data (G013). A tandem duplication of two exons of *CASK* was found in another individual (G222), and a 92 kb deletion of the entire *CDNKA* gene and part of the *CDNK2B* gene was identified in a third individual (G199). Short tandem repeat analysis identified a pathogenic expansion in *DMPK*, establishing a diagnosis of myotonic dystrophy type 1 (DM1; MIM: 160900) in the remaining individual (G351), but this did not account for his entire phenotype.

Comparison of ACMG variant classification to multidisciplinary diagnosis of individuals

Over the course of this study, our multidisciplinary research team diagnosed 261 families with one or more genetic diseases caused by variants discovered on CAUSES GWS. If identification of genetic disease in the CAUSES families had been based on the ACMG classification alone (i.e., one pathogenic or likely pathogenic variant allele of an autosomal dominant or X-linked disease locus in a hemizygous male or two pathogenic/likely pathogenic alleles or one pathogenic/likely pathogenic allele and one VUS/unclassified allele for an autosomal recessive or X-linked disease in a female), a genetic disease rate would have been recognized in 33.6% instead of the 52.2% of families diagnosed by our multidisciplinary team (Table S1). This difference in diagnostic rates is highly statistically significant ($p < 0.000001$).

Overall, 89 (74%) of 120 selected variants that were Sanger sequenced and classified as VUS or unclassified according to ACMG criteria were interpreted as probably or definitely disease-causing by our multidisciplinary research team by the end of the study. One variant classified as pathogenic according to the ACMG criteria was interpreted by our research team to be uncertain with respect to disease causation (G483). This individual had a single ACMG pathogenic variant in *ALPL* inherited from an unaffected parent, a normal serum alkaline phosphatase level, and a phenotype (ID, autism, seizure disorder, poor growth, and inflammatory bowel disease) that was not typical of hypophosphatasia (HPPA; MIM: 146300).

Compelling research candidates

Variants in compelling research candidate genes were identified in 113 families (Table S5), with four sets of siblings sharing candidate variants. In 14 cases, a variant in another gene was identified as probably or definitely disease causing by our multidisciplinary research team, raising the possibility of dual diagnoses in these individuals.

Discussion

The heterogeneous ancestry of the pediatric individuals referred to the CAUSES study reflects the marked diversity of the population of British Columbia. Families of European descent were most frequent, but made up a little less than half of the total. South and East Asian families constituted almost 1/3 of the total, with Middle Eastern, First Nations, and other groups accounting for the rest. A Canadian GWS study done in Toronto reported European ancestry in 61% of individuals enrolled.[42] Apart from English, the most frequently spoken languages in British Columbia include Punjabi, Cantonese, and Mandarin (www.statcan.gc.ca). Most of the families seen in the CAUSES study were fluent in English, but translator services were used for non-English-speaking families. The ethnic diversity represented in the CAUSES cohort poses challenges for variant interpretation owing to the lack of adequate representation of non-European ethnic groups, particularly Indigenous populations, in reference databases.[43] The frequency of *de novo* autosomal dominant variants identified in our cohort supports a trio-based approach to GWS, especially for individuals whose ethnicity is poorly represented in reference databases.

Neurology, medical genetics, and biochemical diseases were the clinical services that referred most individuals to the CAUSES study and are the most frequent users of GWS in the province. The largest cohort by indication was individuals with syndromic ID, who may be followed by medical specialists in any (or all) of these three clinical services.

The overall diagnostic rate of 52.2% in the CAUSES study is higher than that reported in many other series,[2,18] but we used the term “diagnosis” to mean clinical identification of a genetic disease in an affected individual rather than the surrogate “molecular diagnosis” used in many published reports. “Molecular diagnosis” is usually based only on standard ACMG variant classification, zygosity, allelism, and the limited phenotypic information provided

Table 3. "False negatives": missed diagnoses

CAUSES ID no.	Sex	Age (years)	Phenotype	CAUSES finding	Follow-up investigations	Gene	Variant (HGVS cDNA nomenclature)	Variant (HGVS protein nomenclature)	Mechanism	Disease (phenotype OMIM no.)	Comment
G012-1	M	12	moderate intellectual disability; mild dysmorphic facial features; mild webbed neck; mild brachydactyly with short distal phalanx of the finger; right lower limb vascular skin abnormality	uninformative exome	clinical exome sequencing with deletion/duplication analysis identified a heterozygous 25 kb likely pathogenic duplication involving exons 4–20 of *CUL4B*	*CUL4B*	NC_000023.10: (c.610+1_611−1)_(2493+1_2494−1)dup	–	hemizygous (inheritance unknown)	X-linked syndromic mental retardation 15 (Cabezas type) (300354)	CAUSES exome sequencing analysis did not include assessment of copy number
G013-1	F	9	short stature (−5 SD); craniofacial dysmorphism; bilateral hip deformities; feeding difficulties; abnormal growth hormone level; congenital adrenal hypoplasia; nephrocalcinosis; osteopenia	uninformative exome	targeted research sequencing found biallelic variants in *POLE*	*POLE*	NM_006,231.4:c.3265G>C; NM_006231.3:c.1682+32C>G	NP_006222.2:p.Val1089Leu; NP_006222.2:p.Asn563Valfs*16	compound heterozygous	IMAGe syndrome (618336)	second variant is deep intronic and was not captured in exome sequencing
G139-1	M	10	developmental delay; microcephaly; cataracts; myopia; hearing loss; renal cysts; cysts of spleen	uninformative exome	clinical exome sequencing with deletion/duplication analysis identified a heterozygous 8,928 bp likely pathogenic deletion involving exons 13–18 of *COL11A1*	*COL11A1*	NC_000001.10:g.103471300_103480228del	–	heterozygous (inheritance unknown)	Stickler syndrome, type II (604841)	CAUSES exome sequencing analysis did not include assessment of copy number
G199-1	F	15	learning disability; astrocytoma; neurofibroma in a muscle	uninformative exome	research genome sequencing identified a heterozygous 92 kb pathogenic deletion involving all of *CDNKA* and part of *CDNK2B*	*CDNK2A/CDNK2B*	NC_000009.11:g.21915312_22006909del	–	*de novo* heterozygous	melanoma and neural system tumor syndrome (155755)	CAUSES exome sequencing analysis did not include assessment of copy number

(*Continued on next page*)

Table 3. ***Continued***

CAUSES ID no.	Sex	Age (years)	Phenotype	CAUSES finding	Follow-up investigations	Gene	Variant (HGVS cDNA nomenclature)	Variant (HGVS protein nomenclature)	Mechanism	Disease (phenotype OMIM no.)	Comment
G222-1	F	4	global developmental delay; cerebellar hypoplasia; microcephaly; cleft lip; cleft palate; Angelman syndrome phenotype with normal *UBE3A* methylation	uninformative exome	research genome sequencing identified a heterozygous 13.9 kb tandem duplication involving exons 11–12 of *CASK*; the deletion was classified as a VUS by ACMG criteria but was considered to be disease causing by the patient's clinicians	*CASK*	NC_000023.10:g.41468838-41482710dup1.8	–	*de novo* heterozygous	mental retardation and microcephaly with pontine and cerebellar hypoplasia (300749)	CAUSES exome sequencing analysis did not include assessment of copy number
G351-1	M	12	myotonic dystrophy; mild intellectual disability; dysmorphic facial features; strabismus; joint hypermobility; inguinal hernias	compound heterozygous variants of *PYCR1* considered to be definitely causal of autosomal recessive cutis laxa type IIB	clinical short tandem repeat analysis of *DMPK* prior to enrollment in CAUSES identified a patho genic 150 repeat expansion in the proband and a 430 repeat expansion in the mother, both of whom were diagnosed with myotonic dystrophy	*DMPK*	NM_004409.5:c.*224_*226CTG150	–	heterozygous (maternally inherited)	myotonic dystrophy type 1 (160900)	CAUSES exome sequencing analysis did not include assessment for expansions of short tandem repeats
G410-1	F	12	facial muscle weakness; velopharyngeal insufficiency; muscular hypotonia of the trunk; proximal muscle weakness; failure to thrive	uninformative exome	clinical testing with a multigene muscle disorder panel; two pathogenic variants of *SEPN1*	*SEPN1*	NM_206926.1:c.1213C>T; NM_020451.2:c.13_22dup	NP_996809.1:p.Arg405Ter; NP_065184.2:p.Gln8Profs*78	compound heterozygous	SEPN1-related myopathy (602771)	second variant is a 10 bp insertion in exon 1; *SEPN1* exon 1 was poorly covered in CAUSES exome dataset

on the test requisition. This siloing of the data available to the laboratory from the rest of the information the clinical team knows about the affected individual makes it difficult or impossible for the laboratory to assess the genotype-phenotype correlation fully, and genotype-phenotype correlation is essential to genetic disease diagnosis. Our use of "diagnosis" to mean interpretation of the GWS laboratory report in the context of all available clinical information about the affected individual and family is consistent with ACMG and UK practice guidelines[1,6,44] and the recommendations of other expert groups.[7,45] High rates of genetic disease diagnosis have also been reported in other GWS studies that have based genetic disease diagnosis on multidisciplinary interpretation and access to patients' complete medical records.[8–10,17,26]

Discordance between ACMG variant classification and clinical diagnosis of genetic disease has been noted in studies comparing ACMG classifications of disease genes to ClinVar[46] or the Human Gene Mutation Database.[47] Some of these differences are probably attributable to incomplete penetrance of many genetic diseases, a problem that is recognized in the ACMG guidelines.[1] Late-onset and variable expressivity of genetic disease also confound "molecular diagnosis" based solely on ACMG variant classification.

We found no difference in the diagnostic rate between 415 families studied with ES (52.3%) and 85 families studied with GS (51.8%) in the CAUSES study. Most previous studies have found similar diagnostic rates in patients who received GS and those who received ES,[2,18] but there is also evidence that GS identifies some disease-causing variants that are not found by ES.[11,17] The high diagnostic rate in our study may also reflect patient selection and our use of trio-based, rather than singleton, GWS, as well as our use of a multidisciplinary research team to make diagnoses in all cases, including those in which variants were unclassified or classified as ACMG VUS.

CAUSES participants were followed clinically after initial interpretation of their GWS results and with periodic bioinformatic reanalysis and reinterpretation of variants as indicated. This follow-up and reclassification increased our diagnostic rate from 43.0% at the time of initial interpretation to 52.2% at the end of the study. Reinterpretation by the CAUSES team resulted in an average increase in diagnosis of 4.8% per year in families initially interpreted as uninformative or uncertain. This rate of additional diagnoses is consistent with other studies that reanalyzed GWS data after a shorter period of time.[48,49] We found that additional genetic diagnoses continued to be made at a similar rate for at least 5 years after the test was done, with no sign of decreasing over this period (Figure 1).

Dual diagnoses

The rate of dual diagnosis in the CAUSES study (eight families; 1.6%) is lower than the ~5%–7% reported in most other studies.[50,51] We did not count variants that had an uncertain relationship to the disease among individuals with a diagnosis, and we did not routinely reanalyze bioinformatic datasets once one diagnosis had been established for an affected individual. However, there are 14 individuals in the CAUSES study who were diagnosed with a genetic disease and also have compelling research variants (Table S5).

"False negatives" and false positives

Seven (2.9%) of the 241 CAUSES families who underwent trio ES and had a result that was interpreted as either uninformative (n = 226) or uncertain (n = 15) were subsequently diagnosed with a specific genetic disorder after a disease-causing genetic variant was found by another clinical or research test (Table 3). Our failure to identify these variants in the CAUSES study reflected either technical limitations of the ES platform or the fact that we did not test for copy number variants in ES data.

It is interesting to note that in six of the seven families in which a "false negative" CAUSES ES result occurred, a disease-causing variant either was found on research GS (G199 and G222) or probably would have been found had GS been done on the family (G012, G013, G139, G410). There were no "false negatives" among individuals who underwent GS, but trio GS was performed in only 85 of the 500 families included in the CAUSES study.

Although rescinding the clinical diagnosis of a genetic disease was frequent prior to the advent of routine genetic testing and often occurs as a result of genetic testing, few studies deal with the occurrence of false positive diagnoses made on the basis of GWS results. We had four such cases (G103, G369, G483, and G550) in our study, all resulting from lack of clinical concordance with the phenotype expected for the observed genotype. Our finding of an occasional false positive clinical diagnosis after GWS is consistent with observations in the DDD,[52] UK 100,000 Genomes Pilot studies,[17] and a recent review of medical records on 130 patients for whom the laboratory and clinical interpretations of sequencing test results were compared.[53]

The difference between variant classification and clinical diagnosis

Genotype-phenotype correlation is the core principle of genetic disease diagnosis. Physicians diagnose genetic disease on the basis of all of the available information about an affected individual, including the medical history and disease course over time, family history, physical examination, specialist consultations, imaging studies, and all of the laboratory results, including reports of ES or GS. The UK practice guideline for variant interpretation[44] advocates use of a genomic multidisciplinary team "to assess the gene variant(s) identified in the context of the patient's phenotype data and ascertain their contribution to the clinical presentation," but genotype-phenotype correlation plays only a minor role (as the phenotypic specificity criterion, PP4) in the ACMG classification. It is, therefore, not surprising that the ACMG classification of an affected

individual's variant(s) as "likely pathogenic" or of "uncertain significance" and the clinician's interpretation of that report in the context of the affected individual's overall clinical picture may differ. The ACMG classification and the clinician are describing two different things: genetic variants on one hand and patients on the other.

The ACMG variant classification alone is not sufficient to diagnose a genetic disease. The information about a genetic variant identified by GWS and categorized by the ACMG guidelines must be interpreted in the context of an affected individual's complete medical history, disease course, family history, physical examination findings, specialist consultations, imaging studies, and other laboratory test results. Having an ACMG pathogenic or likely pathogenic variant does not necessarily mean that the variant is causing an affected individual's genetic disease. For example, heterozygous carriers of an autosomal recessive condition or nonpenetrant carrier of an autosomal dominant disorder have pathogenic variants that do not affect their own health. The ACMG pathogenic classification means that a variant is capable of causing genetic disease in some circumstances. The clinician who ordered the test must decide whether those circumstances exist in a particular affected individual; it is the physician (or multidisciplinary research team in CAUSES) who makes the diagnosis.

Many clinicians diagnose autosomal recessive genetic diseases in patients with characteristic phenotypes and biochemistry results and one ACMG pathogenic or likely pathogenic allele and a VUS of the second allele if the phenotype and other laboratory results are characteristic of the disease.[54–56] Similarly, autosomal dominant genetic disease may be diagnosed in patients with an ACMG VUS of the associated gene, a classical phenotype, and ancillary supporting data.[56–58] Discordance between ACMG variant classification and clinical diagnosis of genetic disease has also been observed in studies of large reference databases.[46,47]

The CAUSES study was a translational research project that provided trio-based ES or GS to 500 families of children with suspected genetic diseases. We diagnosed a specific genetic condition in 52.2% of the individuals enrolled in this study, a high diagnostic rate that we attribute largely to (1) close collaboration between clinical geneticists, genetic counselors, laboratory geneticists, and clinical bioinformaticians on our research team and the affected individual's clinical team in interpreting the variants found and (2) continuing follow-up of GWS results, with reanalysis and reinterpretation pursued over many years to take advantage of technical improvements and new knowledge that have accumulated.

We learned that pre-test genetic counseling involves much more than just "consenting" the family, that "false negative" and false positive results occasionally do occur with clinical GWS, that genetic counseling is valuable in preparing families for possible changes in interpretation that may take place over time, and that follow-up is important for families with uninformative as well as those with positive GWS results. Finally, we were reminded that patient-oriented research is essential to the provision of high-quality genetic health care.

Data and code availability

The scripts used for this analysis are available at https://github.com/FriedmanLab/StructuralVariantAnalysis. Additional clinical data may be available upon request from the corresponding author, subject to privacy or ethical restrictions. The variants reported in Table S1 have been deposited in the ClinVar database.

Supplemental information

Supplemental information can be found online at https://doi.org/10.1016/j.xhgg.2022.100108.

Acknowledgments

We are truly grateful to the families who participated in the CAUSES study and all the referring physicians and genetic counselors. We would also like to thank the secretarial staff of the Provincial Medical Genetics Program (in particular Margaret Sankowski and Ashley Wallace), Sheryl Atkinson, the BCCHR IT Support Team, Rhea Beauchesne, Lynne Beszant, Patricia Birch, Jocelyn Carter-Sim, Rachel Coe, Courtney B. Cook, Alivia Dey, Adrienne Elbert, Jane Gillis, Liza Mak, Anne MacDougall, Patricia Power, Dawn Siciliano, Angela Siemens, Emma Strong, Tracy Tucker, Tasha Wainstein, and the Provincial Medical Genetics Program, Department of Pediatrics, Department of Pathology and Laboratory Medicine, BC Women's Hospital and BC Children's Hospital.

Investigators in the CAUSES study (Clinical Assessment of the Utility of Sequencing as a Service) include Shelin Adam, Christèle du Souich, Alison M. Elliott, A.L., J.M., T.N.N., C.K., and J.M.F. (PI). The bioinformatic pipeline used in part of the CAUSES study was developed in the laboratory of W.W. The CAUSES project was funded by the Mining for Miracles (BCCH Foundation) and Genome British Columbia, with support from the British Columbia Provincial Health Services Authority and British Columbia Women's Hospital.

Author contributions

A.M.E., S.A., C.D.S., A.L., T.N., C.V.K., and J.M.F. contributed to the conception and study design. All authors contributed to data acquisition. A.M.E. and J.M.F. drafted the manuscript and all authors contributed to critical review and editing of the manuscript. A.M.E., S.A., C.D.S., A.L., T.N., and J.M.F. accessed and verified the data.

Declaration of interests

The authors declare no competing interests.

Received: September 16, 2021
Accepted: April 11, 2022

Web resources

ClinVar, https://www.ncbi.nlm.nih.gov/clinvar/
DECIPHER, https://www.deciphergenomics.org/

GeneMatcher, https://genematcher.org/
Github, https://github.com/FriedmanLab/StructuralVariantAnalysis
Golden Helix, www.goldenhelix.com
OMIM, https://www.omim.org/

References

1. Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W.W., Hegde, M., Lyon, E., Spector, E., et al. (2015). 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. *17*, 405–424. https://doi.org/10.1038/gim.2015.30.
2. Ontario Health (Quality) (2020). Genome-Wide Sequencing for Unexplained Developmental Disabilities or Multiple Congenital Anomalies: A Health Technology Assessment. Vol 20. https://www.hqontario.ca/evidence-to-improve-.
3. Wright, C.F., FitzPatrick, D.R., and Firth, H.V. (2018). Paediatric genomics: diagnosing rare disease in children. Nat. Rev. Genet. *19*, 253–268. https://doi.org/10.1038/nrg.2017.116.
4. Shickh, S., Mighton, C., Uleryk, E., Pechlivanoglou, P., and Bombard, Y. (2021). The clinical utility of exome and genome sequencing across clinical indications: a systematic review. Hum. Genet. *140*, 1403–1416. https://doi.org/10.1007/s00439-021-02331-x.
5. Muriello, M. (2022). Exome and whole genome sequencing in the neonatal intensive care unit. Clin. Perinatol *49*, 167–179. https://doi.org/10.1016/j.clp.2021.11.018.
6. Manickam, K., McClain, M.R., Demmer, L.A., Biswas, S., Kearney, H.M., Malinowski, J., Massingham, L.J., Miller, D., Yu, T.W., and Hisama, F.M. (2021). Exome and genome sequencing for pediatric patients with congenital anomalies or intellectual disability: an evidence-based clinical guideline of the American College of Medical Genetics and Genomics (ACMG). Genet. Med. *23*, 2029–2037. https://doi.org/10.1038/s41436-021-01242-6.
7. Srivastava, S., Love-Nichols, J.A., Dies, K.A., Ledbetter, D.H., Martin, C.L., Chung, W.K., Firth, H.V., Frazier, T., Hansen, R.L., Prock, L., et al. (2019). Meta-analysis and multidisciplinary consensus statement: exome sequencing is a first-tier clinical diagnostic test for individuals with neurodevelopmental disorders. Genet. Med. *21*, 2413–2421. https://doi.org/10.1038/s41436-019-0554-6.
8. Fokstuen, S., Makrythanasis, P., Hammar, E., Guipponi, M., Ranza, E., Varvagiannis, K., Santoni, F.A., Albarca-Aguilera, M., Poleggi, M.E., Couchepin, F., et al. (2016). Experience of a multidisciplinary task force with exome sequencing for Mendelian disorders. Hum. Genomics *10*, 24. https://doi.org/10.1186/S40246-016-0080-4.
9. Baldridge, D., Heeley, J., Vineyard, M., Manwaring, L., Toler, T.L., Fassi, E., Fiala, E., Brown, S., Goss, C.W., Willing, M., et al. (2017). The Exome Clinic and the role of medical genetics expertise in the interpretation of exome sequencing results. Genet. Med. *19*, 1040–1048. https://doi.org/10.1038/gim.2016.224.
10. Mak, C.C., Leung, G.K., Mok, G.T., Yeung, K.S., Yang, W., Fung, C.W., Chan, S.H., Lee, S.L., Lee, N.C., Pfundt, R., et al. (2018). Exome sequencing for paediatric-onset diseases: impact of the extensive involvement of medical geneticists in the diagnostic odyssey. NPJ Genomic Med. *3*, 19. https://doi.org/10.1038/S41525-018-0056-5.
11. Splinter, K., Adams, D.R., Bacino, C.A., Bellen, H.J., Bernstein, J.A., Cheatle-Jarvela, A.M., Eng, C.M., Esteves, C., Gahl, W.A., Hamid, R., et al. (2018). Effect of genetic diagnosis on patients with previously undiagnosed disease. N. Engl. J. Med. *379*, 2131–2139. https://doi.org/10.1056/NEJMoa1714458.
12. Basel-Salmon, L., Orenstein, N., Markus-Bustani, K., Ruhrman-Shahar, N., Kilim, Y., Magal, N., Hubshman, M.W., and Bazak, L. (2019). Improved diagnostics by exome sequencing following raw data reevaluation by clinical geneticists involved in the medical care of the individuals tested. Genet. Med. *21*, 1443–1451. https://doi.org/10.1038/s41436-018-0343-7.
13. Schmitz-Abe, K., Li, Q., Rosen, S.M., Nori, N., Madden, J.A., Genetti, C.A., Wojcik, M.H., Ponnaluri, S., Gubbels, C.S., Picker, J.D., et al. (2019). Unique bioinformatic approach and comprehensive reanalysis improve diagnostic yield of clinical exomes. Eur. J. Hum. Genet. *27*, 1398–1405. https://doi.org/10.1038/S41431-019-0401-X.
14. Zastrow, D.B., Kohler, J.N., Bonner, D., Reuter, C.M., Fernandez, L., Grove, M.E., Fisk, D.G., Yang, Y., Eng, C.M., Ward, P.A., et al.; Undiagnosed Diseases Network (2019). A toolkit for genetics providers in follow-up of patients with non-diagnostic exome sequencing. J. Genet. Couns. *28*, 213–228. https://doi.org/10.1002/JGC4.1119.
15. Krenn, M., Tomschik, M., Rath, J., Cetin, H., Grisold, A., Zulehner, G., Milenkovic, I., Stogmann, E., Zimprich, A., Strom, T.M., et al. (2020). Genotype-guided diagnostic reassessment after exome sequencing in neuromuscular disorders: experiences with a two-step approach. Eur. J. Neurol. *27*, 51–61. https://doi.org/10.1111/ENE.14033.
16. Klee, E.W., Cousin, M.A., Pinto e Vairo, F., Morales-Rosado, J.A., Macke, E.L., Jenkinson, W.G., Ferrer, A., Schultz-Rogers, L.E., Olson, R.J., Oliver, G.R., et al. (2020). Impact of integrated translational research on clinical exome sequencing. Genet. Med. *23*, 498–507. https://doi.org/10.1038/s41436-020-01005-9.
17. 100,000 Genomes Project Pilot Investigators, Smedley, D., Smith, K.R., et al. (2021). 100,000 Genomes Pilot on Rare-Disease Diagnosis in Health Care - preliminary Report. N. Engl. J. Med. *385*, 1868–1880. https://doi.org/10.1056/NEJMOA2035790.
18. Clark, M.M., Stark, Z., Farnaes, L., Tan, T.Y., White, S.M., Dimmock, D., and Kingsmore, S.F. (2018). Meta-analysis of the diagnostic and clinical utility of genome and exome sequencing and chromosomal microarray in children with suspected genetic diseases. Npj Genomic Med. *3*, 16. https://doi.org/10.1038/s41525-018-0053-8.
19. Sorelle, J.A., Thodeson, D.M., Arnold, S., Gotway, G., and Park, J.Y. (2019). Clinical utility of reinterpreting previously reported genomic epilepsy test results for pediatric patients. JAMA Pediatr. *173*, e182302. https://doi.org/10.1001/JAMAPEDIATRICS.2018.2302.
20. Deignan, J.L., Chung, W.K., Kearney, H.M., Monaghan, K.G., Rehder, C.W., and Chao, E.C. (2019). Points to consider in the reevaluation and reanalysis of genomic test results: a statement of the American College of Medical Genetics and Genomics (ACMG). Genet. Med. *21*, 1267–1270. https://doi.org/10.1038/S41436-019-0478-1.
21. Elliott, A.M., du Souich, C., Adam, S., Dragojlovic, N., van Karnebeek, C., Nelson, T.N., Lehman, A., Lynd, L.D., Friedman, J.M.; and The CAUSES Study (2018). The Genomic Consultation Service: a clinical service designed to improve patient

selection for genome-wide sequencing in British Columbia. Mol. Genet. Genomic Med. *6*, 592–600. https://doi.org/10.1002/mgg3.410.

22. Dragojlovic, N., Elliott, A.M., Adam, S., van Karnebeek, C., Lehman, A., Mwenifumbo, J.C., Nelson, T.N., du Souich, C., Friedman, J.M., and Lynd, L.D. (2018). The cost and diagnostic yield of exome sequencing for children with suspected genetic disorders: a benchmarking study. Genet. Med. *20*, 1–9. https://doi.org/10.1038/gim.2017.226.
23. Dragojlovic, N., van Karnebeek, C.D.M., Ghani, A., Genereaux, D., Kim, E., Birch, P., Adam, S., du Souich, C., Elliott, A.M., Lehman, A., et al. (2020). The cost trajectory of the diagnostic care pathway for children with suspected genetic disorders. Genet. Med. *22*, 292–300. https://doi.org/10.1038/s41436-019-0635-6.
24. Boycott, K., Hartley, T., Adam, S., Bernier, F., Chong, K., Fernandez, B.A., Friedman, J.M., Geraghty, M.T., Hume, S., Knoppers, B.M., et al. (2015). The clinical application of genome-wide sequencing for monogenic diseases in Canada: position Statement of the Canadian College of Medical Geneticists. J. Med. Genet. *52*, 431–437. https://doi.org/10.1136/jmedgenet-2015-103144.
25. Elliott, A.M., Dragojlovic, N., Campbell, T., Adam, S., Souich, Cd, Fryer, M., Lehman, A., Karnebeek, Cv, Lynd, L.D., and Friedman, J.M. (2021). Utilization of telehealth in paediatric genome-wide sequencing: health services implementation issues in the CAUSES Study. J. Telemed. Telecare. Published online January 20. https://doi.org/10.1177/1357633X20982737.
26. Tarailo-Graovac, M., Shyr, C., Ross, C.J., Horvath, G.A., Salvarinova, R., Ye, X.C., Zhang, L.H., Bhavsar, A.P., Lee, J.J., Drogemoller, B.I., et al. (2016). Exome sequencing and the management of neurometabolic disorders. N. Engl. J. Med. *374*, 2246–2255. https://doi.org/10.1056/NEJMoa1515792.
27. Myers, A., du Souich, C., Yang, C.L., Borovik, L., Mwenifumbo, J., Rupps, R., Study, C., Lehman, A., and Boerkoel, C.F. (2017). FOXP1 haploinsufficiency: phenotypes beyond behavior and intellectual disability? Am. J. Med. Genet. Part. A. *173* (12), 3172–3181. https://doi.org/10.1002/ajmg.a.38462.
28. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R.; and 1000 Genome Project Data Processing Subgroup (2009). The sequence alignment/map format and SAMtools. Bioinformatics *25*, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352.
29. Li, H. (2021). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. Preprint at arXiv. http://arxiv.org/abs/1303.3997.
30. Tarasov, A., Vilella, A.J., Cuppen, E., Nijman, I.J., and Prins, P. (2015). Sambamba: fast processing of NGS alignment formats. Bioinformatics *31*, 2032–2034. https://doi.org/10.1093/bioinformatics/btv098.
31. Poplin, R., Ruano-Rubio, V., DePristo, M.A., Fennell, T.J., Carneiro, M.O., Van der Auwera, G.A., Kling, D.E., Gauthier, L.D., Levy-Moonshine, A., Roazen, D., et al. (2018). Scaling accurate genetic variant discovery to tens of thousands of samples. Preprint at bioRxiv. https://doi.org/10.1101/201178.
32. Abyzov, A., Urban, A.E., Snyder, M., and Gerstein, M. (2011). CNVnator: an approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. Genome Res. *21*, 974–984. https://doi.org/10.1101/gr.114876.110.
33. Zhu, M., Need, A.C., Han, Y., Ge, D., Maia, J., Zhu, Q., Heinzen, E., Cirulli, E., Pelak, K., He, M., et al. (2012). Using ERDS to infer copy-number variants in high-coverage genomes. Am. J. Hum. Genet. *91*, 408–421. https://doi.org/10.1016/j.ajhg.2012.07.004.
34. Layer, R.M., Chiang, C., Quinlan, A.R., and Hall, I.M. (2014). LUMPY: a probabilistic framework for structural variant discovery. Genome Biol. *15*, R84. https://doi.org/10.1186/gb-2014-15-6-r84.
35. Chen, X., Schulz-Trieglaff, O., Shaw, R., Barnes, B., Schlesinger, F., Kallberg, M., Cox, A.J., Kruglyak, S., and Saunders, C.T. (2016). Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. Bioinformatics *32*, 1220–1222. https://doi.org/10.1093/bioinformatics/btv710.
36. Jeffares, D.C., Jolly, C., Hoti, M., Speed, D., Shaw, L., Rallis, C., Balloux, F., Dessimoz, C., Bahler, J., and Sedlazeck, F.J. (2017). Transient structural variations have strong effects on quantitative traits and reproductive isolation in fission yeast. Nat. Commun. *8*, 14061. https://doi.org/10.1038/ncomms14061.
37. Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. *38*, e164. https://doi.org/10.1093/nar/gkq603.
38. Köhler, S., Carmody, L., Vasilevsky, N., Jacobsen, J.O., Danis, D., Gourdine, J.P., Gargano, M., Harris, N.L., Matentzoglu, N., McMurry, J.A., et al. (2019). Expansion of the human phenotype ontology (HPO) knowledge base and resources. Nucleic Acids Res. *47*, D1018–D1027. https://doi.org/10.1093/nar/gky1105.
39. McKusick-Nathans Institute of Genetic Medicine JHU (2021). Omim - online mendelian inheritance in man. https://omim.org/.
40. O'Leary, N.A., Wright, M.W., Brister, J.R., Ciufo, S., Haddad, D., McVeigh, R., Rajput, B., Robbertse, B., Smith-White, B., Ako-Adjei, D., et al. (2016). Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res. *44*, D733–D745. https://doi.org/10.1093/nar/gkv1189.
41. Burns, W., Bird, L.M., Heron, D., Keren, B., Ramachandra, D., Thiffault, I., Del Viso, F., Amudhavalli, S., Engleman, K., Parenti, I., et al. (2021). Syndromic neurodevelopmental disorder associated with de novo variants in DDX23. Am. J. Med. Genet. Part A *185*, 2863–2872. https://doi.org/10.1002/AJMG.A.62359.
42. Lionel, A.C., Costain, G., Monfared, N., Walker, S., Reuter, M.S., Hosseini, S.M., Thiruvahindrapuram, B., Merico, D., Jobling, R., Nalpathamkalam, T., et al. (2018). Improved diagnostic yield compared with targeted gene sequencing panels suggests a role for whole-genome sequencing as a first-tier genetic test. Genet. Med. *20*, 435–443. https://doi.org/10.1038/gim.2017.119.
43. Caron, N.R., Chongo, M., Hudson, M., Arbour, L., Wasserman, W.W., Robertson, S., Correard, S., and Wilcox, P. (2020). Indigenous genomic databases: pragmatic considerations and cultural contexts. Front. Public Heal. *8*, 111. https://doi.org/10.3389/fpubh.2020.00111.
44. Ellard, S., Baple, E.L., Callaway, A., et al. (2020). ACGS best practice guidelines for variant classification in rare disease 2020 recommendations ratified by ACGS quality subcommittee on 4 th. https://doi.org/10.1101/531210.
45. Friedman, J.M., Jones, K.L., and Carey, J.C. (2020). Exome sequencing and clinical diagnosis. JAMA *324*, 627–628. https://doi.org/10.1001/jama.2020.11126.

46. van Rooij, J., Arp, P., Broer, L., Verlouw, J., van Rooij, F., Kraaij, R., Uitterlinden, A., and Verkerk, A.J. (2020). Reduced penetrance of pathogenic ACMG variants in a deeply phenotyped cohort study and evaluation of ClinVar classification over time. Genet. Med. *22*, 1812–1820. https://doi.org/10.1038/s41436-020-0900-8.
47. Park, K.-J., Lee, W., Chun, S., and Min, W.-K. (2021). The frequency of discordant variant classification in the human gene mutation database: a comparison of the American College of Medical Genetics and Genomics guidelines and ClinVar. Lab. Med. *52*, 250–259. https://doi.org/10.1093/labmed/lmaa072.
48. Costain, G., Jobling, R., Walker, S., Reuter, M.S., Snell, M., Bowdin, S., Cohn, R.D., Dupuis, L., Hewson, S., Mercimek-Andrews, S., et al. (2018). Periodic reanalysis of whole-genome sequencing data enhances the diagnostic advantage over standard clinical genetic testing. Eur. J. Hum. Genet. *26*, 740–744. https://doi.org/10.1038/s41431-018-0114-6.
49. Wenger, A.M., Guturu, H., Bernstein, J.A., and Bejerano, G. (2017). Systematic reanalysis of clinical exome data yields additional diagnoses: implications for providers. Genet. Med. *19*, 209–214. https://doi.org/10.1038/gim.2016.88.
50. Posey, J.E., Rosenfeld, J.A., James, R.A., Bainbridge, M., Niu, Z., Wang, X., Dhar, S., Wiszniewski, W., Akdemir, Z.H., Gambin, T., et al. (2016). Molecular diagnostic experience of whole-exome sequencing in adult patients. Genet. Med. *18*, 678–685. https://doi.org/10.1038/gim.2015.142.
51. Posey, J.E., Harel, T., Liu, P., Rosenfeld, J.A., James, R.A., Coban Akdemir, Z.H., Walkiewicz, M., Bi, W., Xiao, R., Ding, Y., et al. (2017). Resolution of disease phenotypes resulting from multilocus genomic variation. N. Engl. J. Med. *376*, 21–31. https://doi.org/10.1056/NEJMoa1516767.
52. Wright, C.F., Eberhardt, R.Y., Constantinou, P., Hurles, M.E., Fitzpatrick, D.R., and Firth, H.V. (2021). Evaluating variants classified as pathogenic in ClinVar in the DDD Study. Genet. Med. *23*, 571–575. https://doi.org/10.1038/S41436-020-01021-9.
53. Berrios, C., Hurley, E.A., Willig, L., Thiffault, I., Saunders, C., Pastinen, T., Goggin, K., and Farrow, E. (2021). Challenges in genetic testing: clinician variant interpretation processes and the impact on clinical care. Genet. Med. *23*, 2289–2299. https://doi.org/10.1038/s41436-021-01267-x.
54. Cecchi, A.C., Vengoechea, E.S., Kaseniit, K.E., Hardy, M.W., Kiger, L.A., Mehta, N., Haque, I.S., Moyer, K., Page, P.Z., Muzzey, D., and Grinzaid, K.A. (2019). Screening for Tay-Sachs disease carriers by full-exon sequencing with novel variant interpretation outperforms enzyme testing in a pan-ethnic cohort. Mol. Genet. Genomic Med. *7*, e836. https://doi.org/10.1002/MGG3.836.
55. Zhang, J., Chen, H., Kornreich, R., and Yu, C. (2019). Prenatal diagnosis of tay-sachs disease. Methods Mol. Biol. *1885*, 233–250. https://doi.org/10.1007/978-1-4939-8889-1_16.
56. Davieson, C.D., Joyce, K.E., Sharma, L., and Shovlin, C.L. (2021). DNA variant classification–reconsidering "allele rarity" and "phenotype" criteria in ACMG/AMP guidelines. Eur. J. Med. Genet. *64*, 104312. https://doi.org/10.1016/J.EJMG.2021.104312.
57. Golubeva, V.A., Nepomuceno, T.C., and Monteiro, A.N.A. (2019). Germline missense variants in BRCA1: new trends and challenges for clinical annotation. Cancers (Basel) *11*, 522. https://doi.org/10.3390/CANCERS11040522.
58. Monteiro, A.N., Bouwman, P., Kousholt, A.N., Eccles, D.M., Millot, G.A., Masson, J.Y., Schmidt, M.K., Sharan, S.K., Scully, R., Wiesmuller, L., et al. (2020). Variants of uncertain clinical significance in hereditary breast and ovarian cancer genes: best practices in functional analysis for clinical annotation. J. Med. Genet. *57*, 509–518. https://doi.org/10.1136/JMEDGENET-2019-106368.