

Systematic reanalysis of clinical exome data yields additional diagnoses: implications for providers

Aaron M. Wenger, PhD¹, Harendra Guturu, PhD¹, Jonathan A. Bernstein, MD, PhD¹ and Gill Bejerano, PhD^{1,2,3}

Purpose: Clinical exome sequencing is nondiagnostic for about 75% of patients evaluated for a possible Mendelian disorder. We examined the ability of systematic reevaluation of exome data to establish additional diagnoses.

Methods: The exome and phenotypic data of 40 individuals with previously nondiagnostic clinical exomes were reanalyzed with current software and literature.

Results: A definitive diagnosis was identified for 4 of 40 participants (10%). In these cases the causative variant is de novo and in a relevant autosomal-dominant disease gene. The literature to tie the causative genes to the participants' phenotypes was weak, nonexistent, or not readily located at the time of the initial clinical exome reports. At

the time of diagnosis by reanalysis, the supporting literature was 1 to 3 years old.

Conclusion: Approximately 250 gene–disease and 9,200 variant–disease associations are reported annually. This increase in information necessitates regular reevaluation of nondiagnostic exomes. To be practical, systematic reanalysis requires further automation and more up-to-date variant databases. To maximize the diagnostic yield of exome sequencing, providers should periodically request reanalysis of nondiagnostic exomes. Accordingly, policies regarding reanalysis should be weighed in combination with factors such as cost and turnaround time when selecting a clinical exome laboratory.

Genet Med advance online publication 21 July 2016

Key Words: diagnostic yield; exome; reanalysis

INTRODUCTION

The reported diagnostic yield for clinical exome sequencing of patients with presumed Mendelian disorders is around 25%.^{1–5} It remains unclear what proportion of undiagnosed patients harbor a causative mutation in the already sequenced exome that is not prioritized in the initial analysis.

There are several reasons that causative mutations in the exome might be unrecognized. First, the processes both to call variants from short sequence reads and to annotate the impact of variants on genes are imperfect.^{6,7} Factors at each step can result in a causative variant being unidentified. Second, exome interpretation is informed by the reported patient phenotype. If key elements of the phenotype are not available to the clinical laboratory, because they either were not recorded or have not yet clinically emerged, this decreases the likelihood that a causative variant would be prioritized. Third, the representation of known gene–disease and variant–disease associations in readily searchable structured databases is incomplete, and the gap between the primary literature and structured databases is growing.⁸ Identifying relevant information in the free-text primary literature relies on imperfect ad hoc heuristic searches such as with search engines, rather than the preferable systematic queries of databases. Fourth, knowledge of variant and gene function is incomplete. Over 1,600 characterized Mendelian phenotypes are not yet tied to a causative gene, some

Mendelian phenotypes exhibit genetic heterogeneity and have multiple causative genes, and many Mendelian conditions are yet to be recognized.^{9,10}

Furthermore, a significant amount of expert time—from 20 to 40 h—is often required to evaluate a clinical exome.¹¹ This introduces potential variability and bias, and makes it difficult to precisely replicate exome analysis. The required manual time also may decrease the frequency of reanalysis. Some clinical laboratories reanalyze exomes upon a provider's request, limit reanalysis requests to one per year, or require a fee to reanalyze.¹²

To evaluate how improvements in genome analysis methods and expanded knowledge of variant and gene function might increase the diagnostic yield of exome sequencing, we recruited 40 individuals with previously nondiagnostic clinical exome sequencing. We reanalyzed exome data with the benefit of up-to-date analysis software and the current literature, and we applied standard American College of Medical Genetics and Genomics guidelines for variant classification.¹³

MATERIALS AND METHODS

Participant recruitment

Beginning in the fall of 2014, patients of the Medical Genetics Service at Stanford Children's Health at Stanford who had nondiagnostic clinical exome reports were offered the opportunity to enroll in an exome data reanalysis study. For participants,

The first two authors contributed equally to this work.

¹Department of Pediatrics, Stanford University, Stanford, California, USA; ²Department of Computer Science, Stanford University, Stanford, California, USA; ³Department of Developmental Biology, Stanford University, Stanford, California, USA. Correspondence: Gill Bejerano (bejerano@stanford.edu)

Submitted 29 February 2016; accepted 16 May 2016; advance online publication 21 July 2016. doi:10.1038/gim.2016.88

the original raw sequencing reads were obtained from the clinical laboratories after consent. The first 40 cases for which data were obtained were included. For all cases, the laboratories had sequenced the exome of the proband and evaluated candidate variants in parents as deemed necessary.

Participant age is reported as of the date on which a sample was collected for the exome sequencing test. On average, the initial nondiagnostic clinical exome reports for the participants had been issued 20 months before reanalysis.

Research was conducted under a protocol approved by the Stanford Institutional Review Board for human subjects research. Written informed consent was obtained for all participants.

Read mapping and variant calling

Sequencing reads were mapped to the GRCh37/hg19 assembly of the human genome using the MEM algorithm of the Burrows-Wheeler Aligner, version 0.7.10-r789, with default parameters.¹⁴ Duplicate reads were marked with Picard Tools, version 1.105.¹⁵ Variants were called using the Genome Analysis Toolkit, version 3.4-46-gbc02625, following the HaplotypeCaller workflow in the Genome Analysis Toolkit Best Practices, including insertion/deletion realignment and base quality score recalibration.¹⁶

Structured encoding of participant phenotypes

Phenotips was used to encode each participant’s medical history in Human Phenotype Ontology nomenclature.^{17,18} The clinical data originally provided to the reference laboratory were used for the encoding. Encoded phenotypes were matched against the gene/phenotype annotations in the Human Phenotype Ontology to evaluate the relevance of each gene to the participant’s phenotype.

Identifying causative variants

ANNOVAR, version 527, was used to annotate variants with a predicted effect on protein-coding genes from the ENSEMBL gene set, version 75, and with an allele frequency in the Exome Aggregation Consortium and 1000 Genomes Project control human populations.^{19–22} Variants were filtered to retain only rare variants (maximum allele frequency <0.1% in any 1000 Genomes Project or Exome Aggregation Consortium population) that are predicted to be missense, truncating, in-frame insertion or deletion, stop codon loss, or splice site disrupting.

A literature review was performed for each candidate causative variant. Variants were classified in accordance with the guidelines of the American College of Medical Genetics and Genomics.¹³ A variant was considered causative if, in the clinical judgment of the evaluating medical geneticist, it is classified as pathogenic or likely pathogenic for a disease that matches the participant’s phenotype.

Once a candidate variant was identified, the clinical laboratory was contacted to perform parental testing for the variant. Following parental testing, updated clinical reports were issued and participants received updated genetic counseling.

Quantifying growth in gene–disease and variant–disease annotations

Data from the OMIM database and the Human Genome Mutation Database (HGMD) Pro 2015.2 were used to quantify the growth in gene–disease and variant–disease associations.^{23,24}

The OMIM gene–disease association count is the “Phenotype description, molecular basis known” statistic reported at <http://omim.org/statistics/entry>. Historical values were obtained from https://web.archive.org/web/*/http://www.ncbi.nlm.nih.gov/Omim/mimstats.html.

The number of HGMD variant–disease associations was calculated using the publication year tied to each variant. Only DM variants (disease-causing; demonstrated to be pathogenic) are included. The HGMD gene–disease annotation counts use the first publication year of any variant tied to a gene. Unique gene symbols are considered distinct genes.

RESULTS

Attributes of the study population

Of the 40 participants, 15 are male and 25 are female. Most are pediatric (age <18 years); at the time of exome sequencing, 31 participants were younger than 10 years of age. The primary indication for testing varied, but the majority of participants (*n* = 28) have a neurologic or neurodevelopmental condition. Participants underwent clinical exome sequencing at the Baylor Miraca Genetics Laboratory, the UCLA Clinical Genomics Center, or Ambry Genetics (Table 1).

Discovery of causative mutations in four participants

Reanalysis of exome data from 40 participants revealed a causative mutation in 4 (Table 2).

Case 1. Case 1 is an 18-year-old female with absent speech, intellectual disability, short stature, hypertrichosis, and dysmorphic facial features. Reanalysis identified a heterozygous rare missense variant in *KMT2A* (NM_005933) c.3464G>A, p.C1155Y. Analysis of the parents confirmed the variant to be *de novo* in the proband. Mutations in *KMT2A* cause Wiedemann-Steiner Syndrome (WDSTS; OMIM 605130), an autosomal-

Table 1 Attributes of the study population

Sex		Primary indication	
Female	25	Neurologic	28
Male	15	Congenital anomalies	4
		Metabolic	2
Age (years)		Musculoskeletal	2
<2	9	Cancer	1
2–5	10	Gastrointestinal	1
5–10	12	Hearing loss	1
>10	9	Vascular	1

Forty study participants were recruited from among patients of the Medical Genetics Service at Stanford Children’s Health who received nondiagnostic clinical exome reports.

Table 2 Diagnosis of 4 of 40 participants (10%) by exome data reanalysis

Case	Causative mutation	Diagnosis from exome data reanalysis	Date of clinical exome report	Publication date of article that ties causative gene to disease
1	<i>KMT2A</i> c.3464G>A p.C1155Y	Wiedemann-Steiner Syndrome (OMIM 605130)	July 2012	August 2012
2	<i>DEAF1</i> c.737G>C p.R246T	Autosomal-dominant mental retardation 24 (OMIM 615828)	November 2013	May 2014; case reports in December 2010 and November 2012
3	<i>IFIH1</i> c.2159G>A p.R720Q	Aicardi-Goutieres syndrome 7 (OMIM 615846)	September 2014	May 2014
4	<i>PIK3R1</i> c.1135C>T p.R379W	SHORT syndrome (OMIM 269880)	October 2012	July 2013

Four participants with nondiagnostic clinical exome sequencing received a diagnosis from exome data reanalysis. All four participants were diagnosed with autosomal-dominant disorders due to missense variants. Analysis of parents confirmed that all mutations are de novo in the probands.

dominant disorder characterized by hypertrichosis cubiti, short stature, and developmental delay.²⁵

Initial nondiagnostic exome sequencing and analysis results for case 1 were reported in July 2012. The first published report of *KMT2A* mutations as the cause of WDSTS is from August 2012 (ref. 25). A June 2015 report identifies the *KMT2A* c.3464G>A variant as the presumed cause of WDSTS in an unrelated male.²⁶ The variant is not listed in HGMD Pro 2015.2.

Case 2. Case 2 is a 10-year-old female with autism, speech delay, intellectual disability, bouts of aggression, self-injurious behavior, and sleep difficulties. Reanalysis identified a heterozygous rare missense variant in *DEAF1* (NM_021008) c.737G>C, p.R246T. Analysis of the parents confirmed the variant to be de novo in the proband. Mutations in *DEAF1* cause autosomal-dominant mental retardation 24 (OMIM 615828), which is characterized by autistic features, developmental delay, poor expressive speech, mood swings, high pain threshold, and sleep difficulties.²⁷

Initial nondiagnostic exome sequencing and analysis results for case 2 were reported in November 2013. Case reports that connect mutations in *DEAF1* to intellectual disability date to December 2010 and November 2012. The definitive report of mutations in *DEAF1* as the cause of autosomal-dominant mental retardation 24 is from May 2014 (refs. 27–29). To our knowledge, the *DEAF1* c.737G>C variant has not been reported in a second case and is not listed in HGMD Pro 2015.2.

Case 3. Case 3 is a 6-year-old female with developmental delay, failure to thrive, spasticity, and cerebral atrophy. Reanalysis identified a heterozygous rare missense variant in *IFIH1* (NM_022168) c.2159G>A, p.R720Q. Analysis of the parents confirmed the variant to be de novo in the proband. Mutations in *IFIH1* cause Aicardi-Goutieres syndrome 7 (OMIM 615846), an autosomal dominant inflammatory disease characterized by severe neurologic problems, growth retardation, axial hypotonia, spasticity, and brain imaging changes.

Initial non-diagnostic exome sequencing and analysis results for case 3 were reported in September 2014. The first published report of *IFIH1* mutation as the cause of Aicardi-Goutieres syndrome 7 is from May 2014 and includes the *IFIH1* c.2159G>A variant.³⁰ The variant was added to the HGMD database in May 2014, soon after publication.

Case 4. Case 4 is a 4-year-old male with short stature, failure to thrive, microcephaly, hearing impairment, recurrent pneumonia, hypertrichosis, carious teeth, and dysmorphic facial features. Reanalysis identified a heterozygous rare missense variant in *PIK3R1* (NM_181504) c.1135C>T, p.R379W. Analysis of the parents confirmed the variant to be de novo in the proband. Mutations in *PIK3R1* cause SHORT syndrome (OMIM 269880), which is characterized by short stature, hyperextensible joints, ocular depression, Rieger anomaly, teething delay, insulin resistance, and hearing deficits.³¹

Initial nondiagnostic exome sequencing and analysis results for case 4 were reported in October 2012. The first published report of *PIK3R1* mutations as the cause of SHORT syndrome is from July 2013 and includes the c.1135C>T variant.³¹ The variant was added to the HGMD database in July 2013, soon after publication.

Gene–disease and variant–disease associations are continually growing

We used the OMIM and HGMD databases to quantify the rate of growth in the number of gene–disease and variant–disease associations in the literature.

As of 21 October 2015, OMIM lists 6,212 Mendelian disorders. The molecular basis (i.e., gene–disease association) is documented for 4,570 of these. Eleven years earlier, on 20 October 2004, OMIM listed the molecular basis of only 1,636 disorders. Since 2004, the number of OMIM disorders with a noted molecular basis (i.e., gene–disease associations) has increased steadily, at an average rate of 266 entries per year (Figure 1a).

Similarly, the number of gene–disease and variant–disease associations in HGMD has increased steadily. HGMD lists 51,790 pathogenic variants in 1,601 genes from 2004 and earlier. Today, HGMD lists 147,728 pathogenic variants in 4,110 genes, an average growth rate of 9,210 variants and 241 genes per year (Figure 1b,c).

DISCUSSION

Exome sequencing is a valuable diagnostic tool for patients with Mendelian disorders. The yield of proband-only exome sequencing is estimated at 25%. Conceptually, cases for which exome sequencing is nondiagnostic fall into two distinct categories: those where the diagnosis lies outside the data produced

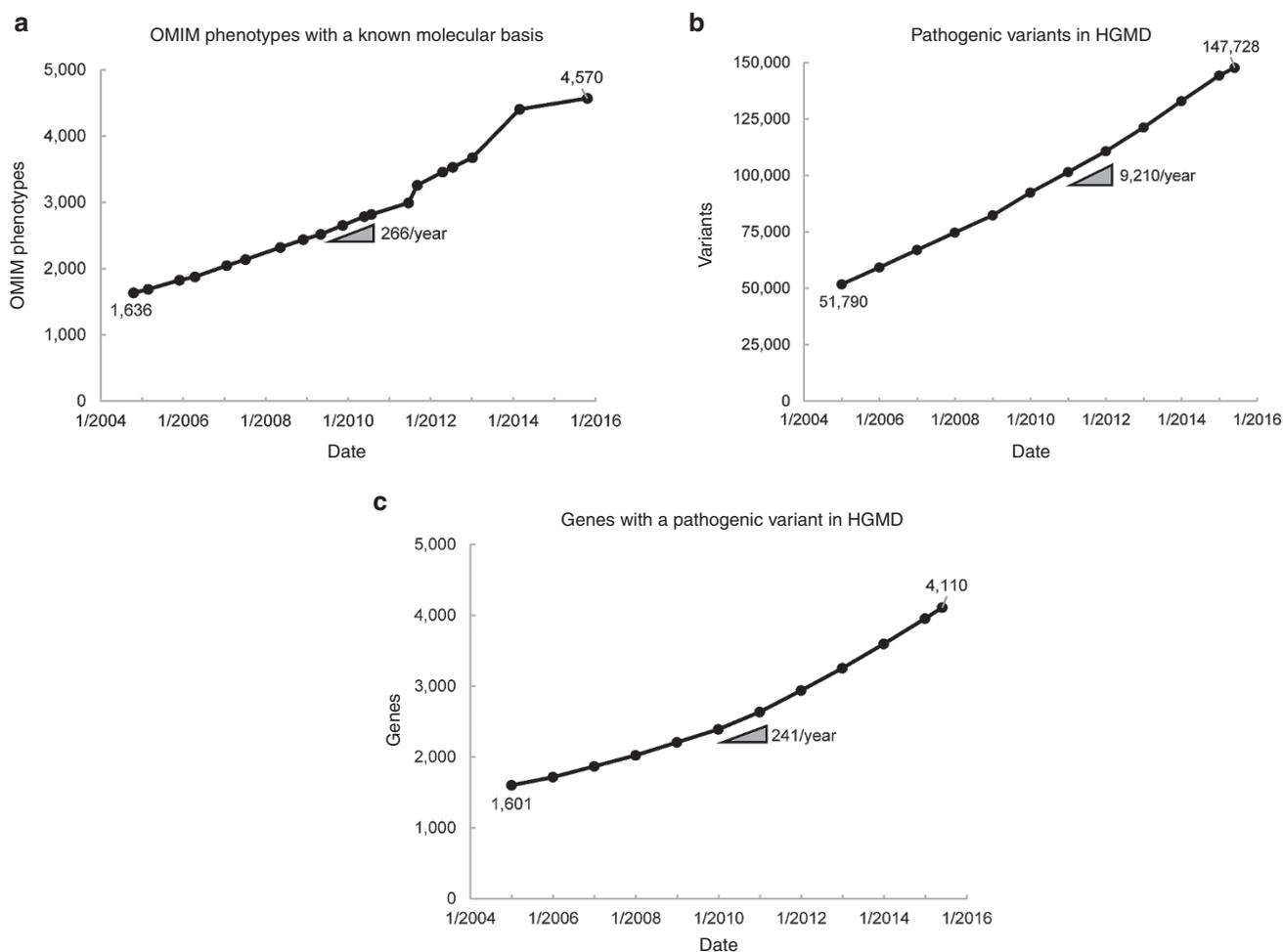


Figure 1 Growth in gene–disease and variant–disease associations. (a) The number of phenotypes with a known molecular basis—that is, a gene–disease association—in OMIM. (b) The number of pathogenic variants—that is, a variant–disease association—in the Human Gene Mutation Database (HGMD). (c) The number of genes with a pathogenic variant in HGMD. The growth rate is an average over the full time span.

(e.g., the causative mutation is in a coding region not covered by the exome, or it is noncoding, or the cause is not genetic); and those where the diagnosis lies in the available genetic data, but incomplete recognition of the phenotype or limitations of current tools or knowledge hinder discovery.

As tools and knowledge related to gene–disease associations improve, we will be able to diagnose more cases in the latter category. We comprehensively reanalyzed 40 unsolved exome cases for which a nondiagnostic exome report was issued, on average, 20 months before reanalysis. We identified a definitive diagnosis in 10% of cases (4/40).

The primary reason that we find new diagnoses is the growing knowledge of gene–disease and variant–disease associations in the literature. A nondiagnostic exome could be solved by any one new publication. We show that the pace of publication is quite rapid, with around 250 new gene–disease associations and 9,200 variant–disease associations curated each year.

The effect of continual growth in the literature is well exemplified by case 1, who was diagnosed with WDSTS caused by a de novo mutation in *KMT2A*. The clinical laboratory issued

a nondiagnostic exome report for case 1 on 25 July 2012. The first paper to link *KMT2A* to WDSTS was published in the *American Journal of Human Genetics* 2 weeks later on 10 August 2012. The information was added to HGMD that same day, and it was logged in OMIM shortly thereafter. Had the clinical exome been ordered a month later, it is probable that the test would have identified the proper diagnosis. Yet, nearly 3 years later, the patient remained undiagnosed. This illustrates the great need to regularly reevaluate nondiagnostic exomes in light of updated knowledge to maximize diagnostic yield.

Frequent reevaluation of exomes is challenging in practice because of the required labor. Clinical laboratories can devote 20 to 40 h of expert labor to issue an initial clinical exome report. Effort may accordingly be prioritized to processing new exomes rather than reanalyzing old ones, as the benefit and likelihood of identifying a new diagnosis must be balanced against the cost of reanalysis. Many clinical laboratories do reanalyze exomes upon a provider's request, but there is often a limit to the number of requests or an associated fee.¹² Increased automation

could lessen the expense and shift the cost-benefit calculation toward more frequent reanalysis.

Efforts should focus on automating as much of the analysis process as is feasible. In particular, standard approaches are needed to encode patient phenotypes and to measure objectively the relevance of a gene or disease to a patient's phenotype.³² Such systems will depend on databases of well-substantiated gene–disease and variant–disease associations represented in structured ontologies. Thus it is also important to develop systems to keep databases up to date with well-curated information from the current literature.

The 40 cases evaluated in this report had proband-only exome sequencing with follow-up assessment of candidate variants in the parents. It is recognized that the yield of trio exome sequencing may be higher.^{3,5} In the four cases diagnosed after reanalysis in this report, the causative variant was *de novo* in the proband and was not listed on the clinical exome report. This is because many laboratories report variants only in genes that are relevant to the patient's phenotype, and the disease associations of the four genes of interest were to varying degrees not well characterized or documented at the time that the clinical reports were issued. Providers should consider ordering trio exomes to facilitate the identification of *de novo* coding variants, of which a typical exome has one or two.³³ The use of trio exome studies should also simplify the task of data reanalysis since segregation can be used to prioritize variants.

Both the initial analysis of exome data and its subsequent reanalysis may benefit from collaboration between the clinical laboratory and the ordering provider. Laboratories rely on providers for phenotypic data and for updates to this information to be considered at reanalysis. Providers rely on laboratories to report relevant variants and to update variant interpretation based on new information. Practical considerations suggest that benefits may follow from ordering providers initiating reanalysis. In requesting reanalysis, the provider may send an update on the patient phenotype to the laboratory. The request may also serve to confirm that the provider (i) is expecting the results of reanalysis and (ii) is in contact with the patient. Nonetheless, there are expected to be occasions when a diagnosis is found on reanalysis by a clinical laboratory without a request from an ordering provider. It may therefore be helpful for undiagnosed patients who have undergone exome sequencing to ensure providers have up-to-date contact information.

Our experience illustrates that a “negative” nondiagnostic result from exome sequencing does not mean that the disease etiology lies outside of the data already produced. For patients with a high suspicion of a Mendelian disorder, providers should periodically request a reevaluation of exome data by clinical laboratories, including in the absence of new phenotypic findings. The findings of this study suggest that reanalysis at a 2- to 3-year interval could result in a 10% diagnostic yield. Larger studies may be helpful to define a standard practice for the timing of reanalysis, taking into account the cost of reanalysis and the evolving rate of discovery of gene–disease relationships. Furthermore, providers should weigh policies regarding

reanalysis, along with cost of testing and turnaround time, in selecting a laboratory for exome sequencing.

ACKNOWLEDGMENTS

The authors thank the participant families for consenting to the study; Hugh O’Brodivich, David K. Stevenson, Gary Shaw, Louanne Hudgins, and Will Talbot for supporting interdepartmental collaboration; David Cooper and Peter Stenson for providing the HGMD data; Kirstie MacMillen and Julia Buckingham for recruiting participants; Allie Newhall for helping with data management; the Stanford Medical Genetics community for referring cases and providing feedback; and the staff at Baylor Miraca Genetics Laboratory, the University of California, Los Angeles Clinical Genomics Center, and Ambry Genetics for facilitating data access. This work was supported by the Stanford Department of Pediatrics, Stanford Discovery Fund, DARPA, and a grant from the National Institutes of Health (U01 MH105949 to G.B.).

DISCLOSURE

The authors declare no conflict of interest.

REFERENCES

1. Yang Y, Muzny DM, Xia F, et al. Molecular findings among patients referred for clinical whole-exome sequencing. *JAMA* 2014;312:1870–1879.
2. Yang Y, Muzny DM, Reid JG, et al. Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *N Engl J Med* 2013;369:1502–1511.
3. Lee H, Deignan JL, Dorrani N, et al. Clinical exome sequencing for genetic identification of rare Mendelian disorders. *JAMA* 2014;312:1880–1887.
4. Iglesias A, Anyane-Yeboah K, Wynn J, et al. The usefulness of whole-exome sequencing in routine clinical practice. *Genet Med* 2014;16:922–931.
5. Retterer K, Juusola J, Cho MT, et al. Clinical application of whole-exome sequencing across clinical indications. *Genet Med* 2014;18:696–704.
6. Cornish A, Guda C. A Comparison of Variant Calling Pipelines Using Genome in a Bottle as a Reference. *Biomed Res Int* 2015;2015:456479.
7. McCarthy DJ, Humburg P, Kanapin A, et al. Choice of transcripts and software has a large effect on variant annotation. *Genome Med* 2014;6:26.
8. Baumgartner WA Jr, Cohen KB, Fox LM, Acquah-Mensah G, Hunter L. Manual curation is not sufficient for annotation of genomic databases. *Bioinformatics* 2007;23:i41–i48.
9. Chong JX, Buckingham KJ, Jhangiani SN, et al.; Centers for Mendelian Genomics. The Genetic Basis of Mendelian Phenotypes: Discoveries, Challenges, and Opportunities. *Am J Hum Genet* 2015;97:199–215.
10. De Rubeis S, He X, Goldberg AP, et al.; DDD Study; Homozygosity Mapping Collaborative for Autism; UK10K Consortium. Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature* 2014;515:209–215.
11. Dewey FE, Grove ME, Pan C, et al. Clinical interpretation and implications of whole-genome sequencing. *JAMA* 2014;311:1035–1045.
12. Jamal SM, Yu JH, Chong JX, et al. Practices and policies of clinical exome sequencing providers: analysis and implications. *Am J Med Genet A* 2013;161A:935–950.
13. Richards S, Aziz N, Bale S, et al.; ACMG Laboratory Quality Assurance Committee. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* 2015;17:405–424.
14. Li H. *Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM*. ArXiv13033997 Q-Bio March 2013. <http://arxiv.org/abs/1303.3997>. Accessed 14 January 2016.
15. Broad Institute. *Picard Tools*. <http://broadinstitute.github.io/picard/>. Accessed 14 January 2016.
16. DePristo MA, Banks E, Poplin R, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 2011;43:491–498.
17. Girdea M, Dumitriu S, Fiume M, et al. PhenoTips: patient phenotyping software for clinical and research use. *Hum Mutat* 2013;34:1057–1065.

18. Köhler S, Doelken SC, Mungall CJ, et al. The Human Phenotype Ontology project: linking molecular biology and disease through phenotype data. *Nucleic Acids Res* 2014;42(Database issue):D966–D974.
19. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 2010;38:e164.
20. Yates A, Akanni W, Amode MR, et al. Ensembl 2016. *Nucleic Acids Res* 2016;44(D1):D710–D716.
21. Exome Aggregation Consortium, Lek M, Karczewski K, et al. Analysis of protein-coding genetic variation in 60,706 humans. *bioRxiv* 2015:030338.
22. 1000 Genomes Project Consortium, Auton A, Brooks LD, et al. A global reference for human genetic variation. *Nature* 2015;526:68–74.
23. Amberger JS, Bocchini CA, Schiettecatte F, Scott AF, Hamosh A. OMIM.org: Online Mendelian Inheritance in Man (OMIM®), an online catalog of human genes and genetic disorders. *Nucleic Acids Res* 2015;43(Database issue): D789–D798.
24. Stenson PD, Ball EV, Mort M, Phillips AD, Shaw K, Cooper DN. The Human Gene Mutation Database (HGMD) and its exploitation in the fields of personalized genomics and molecular evolution. *Curr Protoc Bioinforma Ed Board Andreas Baxevanis AI* 2012;Chapter 1:Unit1.13.
25. Jones WD, Dafou D, McEntagart M, et al. De novo mutations in MLL cause Wiedemann-Steiner syndrome. *Am J Hum Genet* 2012;91:358–364.
26. Bramswig NC, Lüdecke HJ, Alanay Y, et al. Exome sequencing unravels unexpected differential diagnoses in individuals with the tentative diagnosis of Coffin-Siris and Nicolaides-Baraitser syndromes. *Hum Genet* 2015;134:553–568.
27. Vulto-van Silfhout AT, Rajamanickam S, Jensik PJ, et al. Mutations affecting the SAND domain of DEAF1 cause intellectual disability with severe speech impairment and behavioral problems. *Am J Hum Genet* 2014;94:649–661.
28. Vissers LE, de Ligt J, Gilissen C, et al. A de novo paradigm for mental retardation. *Nat Genet* 2010;42:1109–1112.
29. Rauch A, Wieczorek D, Graf E, et al. Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. *Lancet* 2012;380:1674–1682.
30. Rice GI, del Toro Duany Y, Jenkinson EM, et al. Gain-of-function mutations in IFIH1 cause a spectrum of human disease phenotypes associated with upregulated type I interferon signaling. *Nat Genet* 2014;46:503–509.
31. Thauvin-Robinet C, Auclair M, Duplomb L, et al. PIK3R1 mutations cause syndromic insulin resistance with lipoatrophy. *Am J Hum Genet* 2013;93: 141–149.
32. Akawi N, McRae J, Ansari M, et al.; DDD study. Discovery of four recessive developmental disorders using probabilistic genotype and phenotype matching among 4,125 families. *Nat Genet* 2015;47:1363–1369.
33. Iossifov I, O’Roak BJ, Sanders SJ, et al. The contribution of de novo coding mutations to autism spectrum disorder. *Nature* 2014;515:216–221.