Title
Clinical Exome Sequencing for Genetic Identification of Rare Mendelian Disorders

Journal
JAMA, 312(18)

ISSN
0098-7484

Authors
Lee, Hane
Deignan, Joshua L
Dorrani, Naghmeh
et al.

Publication Date
2014-11-12

DOI
10.1001/jama.2014.14604

Peer reviewed

Published in final edited form as:
JAMA. 2014 November 12; 312(18): 1880–1887. doi:10.1001/jama.2014.14604.

Clinical Exome Sequencing for Genetic Identification of Rare Mendelian Disorders

Hane Lee, PhD, **Joshua L. Deignan, PhD**, **Naghmeh Dorrani, MS, CGC**, **Samuel P. Strom, PhD**, **Sibel Kantarci, PhD**, **Fabiola Quintero-Rivera, MD**, **Kingshuk Das, MD**, **Traci Toy, BS**, **Bret Harry, BS**, **Michael Yourshaw, PhD**, **Michelle Fox, MS, CGC**, **Brent L. Fogel, MD, PhD**, **Julian A. Martinez-Agosto, MD, PhD**, **Derek A. Wong, MD**, **Vivian Y. Chang, MD, MS**, **Perry B. Shieh, MD, PhD**, **Christina G. S. Palmer, PhD, CGC**, **Katrina M. Dipple, MD, PhD**, **Wayne W. Grody, MD, PhD**, **Eric Vilain, MD, PhD**, and **Stanley F. Nelson, MD**

Department of Pathology and Laboratory Medicine, David Geffen School of Medicine, University of California, Los Angeles (Lee, Deignan, Strom, Kantarci, Quintero-Rivera, Das, Toy, Grody, Nelson); Clinical Genomics Center, David Geffen School of Medicine, University of California, Los Angeles (Lee, Deignan, Dorrani, Strom, Kantarci, Quintero-Rivera, Das, Toy, Grody, Vilain, Nelson); Department of Pediatrics, David Geffen School of Medicine, University of California, Los Angeles (Dorrani, Yourshaw, Fox, Martinez-Agosto, Wong, Chang, Dipple, Grody, Vilain); Institute for Digital Research and Education, University of California, Los Angeles (Harry); Department of Neurology, David Geffen School of Medicine, University of California, Los Angeles (Fogel, Shieh); Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles (Martinez-Agosto, Palmer, Dipple, Grody, Vilain, Nelson); Department of Psychiatry and Biobehavioral Sciences, David Geffen School of Medicine, University of California, Los Angeles (Palmer, Nelson)

Abstract

Importance—Clinical exome sequencing (CES) is rapidly becoming a common molecular diagnostic test for individuals with rare genetic disorders.

Corresponding Author: Stanley F. Nelson, MD, Clinical Genomics Center, David Geffen School of Medicine, University of California, Los Angeles, 695 Charles E. Young Dr S, GONDA 5506A, Los Angeles, CA 90095 (snelson@ucla.edu).

Author Contributions: Drs Lee and Nelson had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.
Study concept and design: Lee, Kantarci, Dipple, Grody, Vilain, Nelson.
Acquisition, analysis, or interpretation of data: All authors.
Drafting of the manuscript: Lee, Quintero-Rivera, Das, Martinez-Agosto, Dipple, Nelson.
Critical revision of the manuscript for important intellectual content: All authors.
Statistical analysis: Lee, Yourshaw, Nelson.
Obtained funding: Chang, Nelson.
Administrative, technical, or material support: Deignan, Quintero-Rivera, Das, Harry, Fox, Fogel, Shieh, Palmer, Dipple, Grody, Vilain, Nelson.
Study supervision: Das, Martinez-Agosto, Grody, Vilain, Nelson.

Conflict of Interest Disclosures: All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Drs Lee, Deignan, Dorrani, Strom, Kantarci, Quintero-Rivera, Das, Toy, Grody, and Nelson are part of the Department of Pathology and Laboratory Medicine, which derives revenue from the clinical use of the clinical exome sequencing test. Dr Fogel reports receiving grant funding from the National Institute of Mental Health, the National Institute for Neurological Disorders and Stroke, and the National Ataxia Foundation. Dr Shieh reports that he is a member of the speakers bureau for Grifols. Dr Vilain reports receiving grant funding from the University of California, Los Angeles.

Objective—To report on initial clinical indications for CES referrals and molecular diagnostic rates for different indications and for different test types.

Design, Setting, and Participants—Clinical exome sequencing was performed on 814 consecutive patients with undiagnosed, suspected genetic conditions at the University of California, Los Angeles, Clinical Genomics Center between January 2012 and August 2014. Clinical exome sequencing was conducted as trio-CES (both parents and their affected child sequenced simultaneously) to effectively detect de novo and compound heterozygous variants or as proband-CES (only the affected individual sequenced) when parental samples were not available.

Main outcomes and Measures—Clinical indications for CES requests, molecular diagnostic rates of CES overall and for phenotypic subgroups, and differences in molecular diagnostic rates between trio-CES and proband-CES.

Results—Of the 814 cases, the overall molecular diagnosis rate was 26% (213 of 814; 95% CI, 23%-29%). The molecular diagnosis rate for trio-CES was 31% (127 of 410 cases; 95% CI, 27%-36%) and 22% (74 of 338 cases; 95% CI, 18%-27%) for proband-CES. In cases of developmental delay in children (<5 years, n = 138), the molecular diagnosis rate was 41% (45 of 109; 95% CI, 32%-51%) for trio-CES cases and 9% (2of 23, 95% CI, 1%-28%) for proband-CES cases. The significantly higher diagnostic yield (*P* value = .002; odds ratio, 7.4 [95% CI, 1.6-33.1]) of trio-CES was due to the identification of de novo and compound heterozygous variants.

Conclusions and Relevance—In this sample of patients with undiagnosed, suspected genetic conditions, trio-CES was associated with higher molecular diagnostic yield than proband-CES or traditional molecular diagnostic methods. Additional studies designed to validate these findings and to explore the effect of this approach on clinical and economic outcomes are warranted.

Over the last few years, advances in next-generation sequencing technologies have decreased the cost of sequencing per base pair about 10-fold, improved accuracy, and greatly increased the speed of generating sequence data. Exome sequencing, which sequences the protein-coding region of the genome, has been rapidly applied to variant discovery in research settings and recent increases in accuracy have enabled development of clinical exome sequencing (CES) for mutation identification in patients with suspected genetic diseases.

Early in 2012, our center launched a CES program with the goal of delivering a more comprehensive method for determining a molecular diagnosis for patients with presumed rare Mendelian disorders that have remained undiagnosed despite exhaustive genetic, biochemical, and radiological testing. We introduced a new test, called trio-CES, in which the whole exome of the affected proband and both parents are sequenced. The trio-CES test has the potential benefit of permitting more sensitive identification of de novo variants and compound heterozygotes and removing from consideration the many heterozygous rare variants observed in each exome from being considered causal in the affected individual because transmission is observed from an unaffected parent. This has not been routinely implemented by other centers due to costs and potential concerns for incidental findings in the unaffected parents. This study reports the first sequential 814 cases tested by our

laboratory and investigates diagnostic yields from different implementations of exome sequencing.

Methods

Our CES test was validated according to the Clinical Laboratory Improvement Amendments regulations and College of American Pathologists guidelines as a single test from DNA extraction to result reporting. The cases were ascertained between January 17, 2012, and August 31, 2014. All work was performed within the University of California, Los Angeles (UCLA), Clinical Genomics Center. This study was approved by the UCLA institutional review boards.

Clinical Exome SequencingTest

The cases are from a consecutive set of clinical cases referred for exome sequencing from clinics at a single university health system as well as from outside referring physicians. All samples required completion of a requisition form, pretest genetic counseling with clinical consent, and provision of a recent clinical note related to the reason that the clinician ordered exome sequencing. Patients were not required to undergo standardized clinical examinations or diagnostic testing prior to referral. Clinical characteristics and prior laboratory investigations of patients reported by referring clinicians were not systematically confirmed by study investigators. Most cases were submitted from geneticists for which substantial prior genetic investigation had been performed and no clear resolution was determined. Patients determined to be appropriate for exome sequencing often presented with clinical symptoms that either involved more than 1 body system or were deemed to be highly genetically heterogeneous. For instance, mutations in more than 1000 different disease genes can manifest as developmental delay. Thus, no specific genetic test is clearly defined for this group of patients. In addition, some cases were referred because other molecular diagnostic testing (such as microarray analysis) or specific gene sequencing (either isolated or panels) were either inconclusive or not available clinically.

Patients were tested after pretest counseling to describe the goals of the test, turnaround time, cost, limitations, the current rate of diagnostic success, the potential to identify variants of uncertain clinical significance and their meaning, and the possibility to opt in or out of receiving the results of incidental findings. This discussion was documented by signature of the referring physician and the patient or patient representative.

On the CES test requisition form, physicians were asked to choose primary clinical indications, report ethnicity, provide any family history, and offer any differential diagnoses, suspected causative genes, or both. Ethnicity information, provided by the physician, was used for quality control and validating genetic findings. For instance, more rare variants will be observed in African American populations. Two test options were offered: a trio-CES and a proband-CES. Proband-CES is exome sequencing of the patient only and trio-CES is exome sequencing of up to 3 family members, including the patient. Typically trio-CES tests include both unaffected parents and the patient, but other combinations of family members were accepted if appropriate, depending on the family history and individual availability.

Sequencing, Data Analysis, and Quality Control

Whole blood collected in tubes with EDTA or purified genomic DNA from tissue sources was accepted for testing. Genomic DNA extraction, library preparation, sequencing, and data analysis were performed using validated protocols (eMethods 1-3 and eFigure 1 in the Supplement). For the patients receiving trio-CES, library preparation and sequencing were only performed after all 3 blood samples had arrived to minimize potential differences due to batch effects. Exome capture was performed using SureSelect Human All Exon V2 Kit (Agilent Technologies) and sequencing was performed using the HiSeq 2000 for a 50-bp paired-end run or HiSeq 2500 for a 100-bp paired-end run (both from Illumina). An average of 60 million independent paired reads or 9.7 Gb of sequence data were generated per sample to provide a mean 100-fold coverage across the RefSeq protein-coding exons and flanking intronic sequence (±2 bp) with more than 93% of these bases and 94% of all reported Human Gene Mutation Database (HGMD) variant positions with a depth of coverage 10 × or more. Thus, we estimated that CES has a more than 93% chance of observing clinically relevant single nucleotide or small indel (insertions and deletions) variant(s). The mitochondrial genome is not specifically captured, but as a byproduct of being present at a high copy number, 99% of the base positions of the protein-coding sequences in the mitochondrial genome are sequenced simultaneously to detect homoplasmic and highly heteroplasmic variants[1] (although we have not reported any clinically significant mitochondrial genome variant, we observed mitochondrial polymorphisms in almost all cases).

Various quality metrics were recorded for each sample for quality control and these metrics were highly consistent from sample to sample (eTable 1 in the Supplement) providing a sensitive indicator of possible sample contamination, errors in library preparation, or inadequate sequencing. Of 1734 samples sequenced, no samples failed these initial quality metrics. Normalized coverage was used to search for evidence of deletions or duplications of an entire chromosomal arm and to confirm the sex of the individuals from the X and the Y chromosome coverage. Because the relative depth of coverage of each exon was generally consistent, variance from the normal coverage was a sensitive means to assess for poor capture quality. Of 1734 samples, 3 samples were rejected due to atypical exon coverage and repeated successfully with a new blood specimen. Consanguinity was estimated and possible uniparental disomy was identified on the basis of long (>5 Mb) stretches of homozygosity.[2,3] Finally, when 1 or more additional family members were sequenced at the same time as the proband, the inheritance pattern and variant sharing across the exome was queried to confirm genetic relationships.

For the first approximately 300 cases, all reported variants were confirmed by Sanger sequencing and more than 99% confirmed. After empirically determining that variants with a QUAL score (a scaled probability of a variant existing at a given site based on the sequencing data calculated by the Genome Analysis Toolkit [GATK] variant caller) 500 or higher to be highly accurate,[4] only clinically significant variants with a QUAL score lower than 500 and all small indels were confirmed by Sanger sequencing before reporting. In current practice, approximately 20% of all reported variants, which are deemed to be of

lower certainty, undergo Sanger sequencing confirmation and of these more than 99% are confirmed.

Variant Analysis and Interpretation

Variants were annotated using Variant Annotator eXtras (VAX, custom annotator)[5] to provide information regarding their effect on protein function, allele frequency in the general population, and prior evidence of disease causality and filtered to select likely pathogenic DNA variants from an average of 21 259 DNA variants per exome (eMethods 4 in the Supplement). Variants with minor allele frequency greater than 1% were removed and deemed to be likely benign (eFigure 2 in the Supplement).[6-8] Of the remaining variants, variants that resulted in amino acid substitutions, microdeletions, microduplications, splice-site changes, or premature protein terminations in the canonical transcript or the most commonly referenced transcript in the literature were selected. For trio-CES cases, variants were further filtered into 4 categories: de novo (new variants not observed in either parent, usually heterozygous in patient, and potentially causing an autosomal dominant condition), homozygous (both parents are heterozygous for the same variant and the child inherited a rare allele from both parents, potentially causing a recessive condition), compound heterozygous (the affected individual has 1 rare variant from the mother and a different rare variant inherited from the father, potentially causing a recessive condition), and inherited variants (this is the largest group of variants and are inherited from a parent and are usually not disease causing; eFigure 2 in the Supplement). Variants were called high quality if each individual in the trio had minimum of 10 × depth of coverage, the parents and unrelated samples in the batch lacked evidence of the alternate allele being present, and the QUAL score was 500 or higher.

Given the heterogeneity of genetic conditions (more than 4000 disease-gene relationships have been described in the Online Mendelian Inheritance in Man [OMIM]), we developed a permissive approach to use the clinician notes to generate a list of phenotypic key words (examples shown in eTables 2-3 in the Supplement). Each phenotypic key word was then searched in the HGMD (professional version) or OMIM database to generate a list of potential disease genes for each patient. From this process we generated a primary gene list (PGL), and the filtered variants were then further annotated as to whether or not they occurred in the PGL. For proband-CES cases, all homozygous and potential compound heterozygous variants with a minor allele frequency less than 1% and all heterozygous variants with a minor allele frequency less than 0.1% within the PGL were examined in depth and prioritized by how well the patient's phenotypes matched to the gene description (eMethods 4-5 and eFigures 2-3 in the Supplement). In addition, heterozygous variants with a minor allele frequency less than 1% that were identified in genes known to cause recessive disorders similar to the patient's phenotype were searched for and highlighted. For trio-CES cases, all de novo, homozygous, and compound heterozygous variants were examined. Inherited heterozygous variants in genes in the PGL that were known to be imprinted or pathogenic (ie, reported in HGMD) were inspected as well. All variants in HGMD are not necessarily disease causing.[9] For this reason, even if a variant was annotated to be present in HGMD, the literature supporting the evidence of each variant was manually curated by our bioinformaticians.

A variant list of all CES cases was presented by the bioinformatician and reviewed at the weekly genomic data board meeting consisting of pathologists, geneticists, genetic counselors, bioinformaticians, and the ordering physician, whenever possible. Non-UCLA physicians had the option of joining the genomic data board meeting in person or through a prescheduled encrypted online conference. Our interpretation process consisted of a brief synopsis of the patient course, followed by a presentation of the annotated filtered variant list; this permitted an informed discussion among the members of the board to evaluate each variant and decide which, if any, to report as pathogenic, likely pathogenic, or variant of uncertain clinical significance (eMethods 5 and eFigure 3 in the Supplement) and determine a molecular diagnosis. Typically, 12 to 16 cases were reviewed within each 2-hour meeting.

Statistical Analysis

To test the significance of a higher diagnostic rate of trio-CES than proband-CES, *P* values were calculated by 1-tailed Fisher exact test.[10] All other comparisons were done by a 2-tailed Fisher exact test. A *P* value of .05 was used as a significance threshold. The 95% CI for proportion was calculated using an online calculator.[11] The odds ratio and 95% CI for the significance of difference in diagnostic rate were also calculated using an online calculator.[12]

Because there were more proband-CES cases in the adult group (due to the lack of availability of older parents), and there were differences in clinical indications based on age, we also assessed the relative diagnostic yield only in the highly genetically heterogeneous group of developmental delay cases and only when CES was performed in children. Age groups of younger than 5 years, 5 to 18 years, and older than 18 years at testing were used for a more direct comparison with published results.[13]

Results

Study Population Characteristics

Patients were referred from a total of 143 different physicians from 42 different institutions. A UCLA clinic ordered 459 of the total 814 cases. Of the 814 patients, 520 patients (64%) were children (18 years or younger) and 254 (49%) of those children were younger than 5 years at testing (eTable 4 in the Supplement). More cases were ordered for trio-CES (353 of 520 patients, 68% [95% CI, 64%-72%]) in the childhood group than in the adult group (57 of 294 patients, 19% [95% CI, 15%-24%], $P < .001$) reflecting parental availability. In total, 453 were males (56%) and 361 were females(44%). The CES was requested for males more than females (*P* value = .003) in the childhood group (male: 310 of 520 patients, 60% [95% CI, 55%-64%]; female: 210 of 520 patients, 40% [95% CI, 36%-45%]) than in the adult group (men: 143 of 294 patients, 49% [95%CI,43%-54%];women: 151 of294 patients, 51% [95% CI, 46%-57%]). The most common clinical indication was developmental delay overall (298 of 814 patients, 37% [95% CI, 33%-40%]) and in the childhood group (274 of 520 patients, 53% [95% CI, 48%-57%]), whereas ataxia was the most common clinical indication in the adult group (77 of 294 patients, 26% [95% CI, 21%-32%]).

Overall Diagnostic Rate

Each individual case from CES generated a PGL and required substantial filtering to determine causality. Of 814 patients, approximately 6% had genetic evidence of consanguinity. For nonconsanguineous families, there were typically fewer than 10 genes with homozygous variants in the proband that were not also homozygous in the parents and fewer than 10 genes with compound heterozygous variants. For proband-CES samples, it was not possible to reliably infer which genes contained de novo or compound heterozygous variants; in a typical case approximately 10 genes with homozygous variants, 40 to 50 genes with potential compound heterozygous variants (in which 2 heterozygous variants in the same gene were observed), and 250 to 300 genes with a single heterozygous variant with minor allele frequency less than 0.1% were considered for initial interpretation (eFigure 2 in the Supplement). In contrast, there were an average of 1.1 amino acid– altering, novel (never observed in the general population), high-quality de novo variants in each trio, which was consistent with prior observations.[14]

Table 1 summarizes the diagnostic rate of our sequential series of cases. Overall, a molecular diagnosis (with the causative variant(s) identified in a well-established clinical gene) was provided for 213 of the 814 total cases (26% [95% CI, 23%-29%]). The breadth of molecular diagnoses was large, and the list of the 213 cases is provided in eTable 2 with detailed variant information and its pathogenicity. Of 264 reported variants in 213 cases, 188 were reported as likely pathogenic and 73 were reported as pathogenic variants (eTable 2 in the Supplement). In addition, 228 of the 814 total cases (28% [95% CI, 25%-31%]) had potential molecular diagnoses based on the results from the CES (Table 1). These are variants identified in clinically relevant genes that were classified as variants of uncertain significance because (1) additional phenotyping (∼ 25%) or segregation analysis (∼ 50%) was needed, or (2) only 1 heterozygous protein-damaging variant was identified in a gene known to cause a recessive disorder consistent with the patient's phenotype but not covered 100% by CES (∼ 25%). Thus, other methods to detect additional variant alleles (exonic deletion or duplication or variants in noncoding regions) were recommended. Some of the cases described in this report have been previously reported.[15]

Diagnostic Rate for Trio-CES vs Proband-CES and in Phenotypic Subgroups

Overall, 410 of the 814 patients (50%) were trio-CES, with both parents sequenced, and 338 patients (42%) were proband-CES. Sixty-six patients (8%) were submitted with only 1 parent or other family members without the parents. There was a significantly higher molecular diagnostic yield from cases performed as trio-CES (127 of 410 cases; 31% [95% CI, 27%-36%]) relative to proband-CES (74 of 338 cases; 22% [95 CI, 18%-27%], $P = .003$) in the overall cohort of cases. Among the 127 trio-CES tests with a conclusive molecular diagnosis, 50% (63 of 127 cases; 95% CI, 41%-58%) had a de novo variant, 16% (20 of127 cases; 95%CI, 10%-23%) had a homozygous variant, 20% (26 of 127 cases; 95% CI, 14%-28%) had compound heterozygous variants, and 8% (10 of 127; 95% CI, 4%-14%) had an X-linked hemizygous variant (Table 2).

The most common phenotype of patients described by the referring physician was developmental delay (298 of 814 cases, 37%). The 5 most common comorbid phenotypes

with developmental delay included hypotonia (111 of 298 cases, 37%), epilepsy or seizures (108 of 298 cases, 36%), dysmorphic features (144 of 298 cases, 48%), autism (69 of 298 cases, 23%), or congenital heart disorder or defect (36 of 298 cases, 12%). The molecular diagnosis rate for each of these phenotypes is shown in Table 3. Although overall 28% (83 of 298 cases; 95% CI, 23%-33%) of developmental delay resulted in a molecular diagnosis, only 16% (11 of 69 cases; 95% CI, 9%-27%) of developmental delay with autism resulted in a molecular diagnosis. The diagnostic rate of developmental delay with autism relative to developmental delay and dysmorphic features of 31% (44 of 144; 95% CI, 24%-39%) indicates differences in diagnostic success with these different presentations ($P = .03$). In order to explore the diagnostic success rate with trio-CES relative to proband-CES accounting for age, we analyzed diagnostic yield in children younger than 5 years at testing (n = 138). For this category, trio-CES was requested for 79% (109 of 138 cases) of all developmental delay cases and 17% (23 of 138 cases) were requested as proband-CES. There was an improvement in the diagnostic yield of trio-CES, with 45 of 109 cases (41% [95% CI: 32%-51%]) receiving a molecular diagnosis relative to the proband-CES cases, which had a molecular diagnostic yield of 9% (2 of 23 cases; 95% CI, 1%-28%, *P* value = .002; odds ratio, 7.4 [95% CI, 1.6-33.1]; Table 4).

Of all diseases for which CES was applied, the group of patients with retinal disorders had the highest molecular diagnostic rate (15 of 31 cases, 48% [95% CI, 32%-65%]), suggesting that a larger fraction of all possible Mendelian disease genes for this family of disorders have been described at this point in the literature. In contrast, ataxia (11 of 86 cases, 13% [95% CI, 7%-22%], *P* value < .001) and disorder of sexual development (6 of 35 cases, 17% [95% CI, 8%-33%], *P* value = .009) cases had much lower diagnostic rates, suggesting a larger fraction of unknown genes or nongenetic underlying mechanisms (Table 3).

Two of the 410 trio-CES cases (0.5%) had mislabeled blood tubes (the blood sample labeled as 1 of the parent's was from the proband from the submitting blood draw facility), which were detected by assessing relatedness of the samples, and both instances were remedied by obtaining new specimens.

Illustrative Cases

A benefit of performing trio-CES is also underscored for identification of de novo variants in genes that are not yet associated with any human disorder, and the significance of identifying these de novo variants has already proven to be critical in a few cases. For instance, an infant (case 75 in eTable 2 in the Supplement) with multifocal complex partial epilepsy and regression of developmental milestones had a novel de novo missense variantin *KCNT1* detected, which was not known at the time to be associated with any human disease. However, soon after completion of CES for this infant, *KCNT1* de novo variants were reported in the literature as a cause of infantile epileptic encephalopathy (OMIM 614959), a condition that matches the patient phenotype well permitting update of the PGL and provision of a conclusive molecular diagnosis.[17] Another infant (case 113 in eTable 2 in the Supplement) with developmental delay, seizures, perisylvian polymicrogyria, and microcephaly had a novel de novo missense variant in *TUBB2A*, recently identified as a causal gene.[18] Furthermore, we reviewed case 107 (eTable 2 in the Supplement), which was

reported 3 months prior with a de novo variant in *TUBB2A* as a variant of uncertain significance, and generated an addendum classifying the variant as likely pathogenic.

Trio-CES is also broadening the clinical phenotype of various rare Mendelian genetic diseases. An example is a 9-year-old girl with developmental delay, mild intellectual disability, hypotonia, dysmorphic features, early tooth eruption, and premature adult teeth in whom trio-CES identified a novel missensede novo variant in the *KMT2A (MLL)* gene that is known to cause autosomal dominant Wiedemann-Steiner syndrome (OMIM 605130; case 130 in eTable 2 in the Supplement). Wiedemann-Steiner syndrome had not been considered by the referring clinicians because the patient had not manifested the hallmark phenotype of the syndrome, namely hairy elbows. However, upon discovery of the de novo variant and another review of the clinical presentation, the ordering physician was able to note the manifestation of the excess growth of terminal hair analogous to hairy elbows.[19]

Although CES is not intended to identify copy number variant, uniparental disomy, somatic mosaic heterozygous variants or variants in the mitochondrial genome, in selected cases we observed evidence of large homozygous or hemizygous exonic deletions and duplications (del and dup). A total of 7 such observations were made, and 5 of 7 were confirmed by an outside laboratory (eTable 3 in the Supplement). Examples include the identification of paternal uniparental disomy in a trio-CES case due to 2 homozygous regions of 5Mb and 19Mb on chromosome 6 with no decrease in coverage and no maternally inherited variants on the entire chromosome 6 (data not shown). We have also observed potential somatic mosaic heterozygous variants with significantly more reads from the reference allele than the alternate allele in 2 cases (case 1, 176 reference reads and 40 alternate reads; case 2, 193 reference reads and 35 alternate reads; *P* value <.001 for both cases). Both were confirmed by Sanger sequencing.

Incidental Findings

Return of “incidental” or “off-target” findings remains an area of debate in medical genetics.[20-24] Although we used published guidelines[20] to determine which genes and diseases may be considered for reporting, we did not restrict ourselves to this gene list nor did we actively search for incidental variants for every case. Instead, we have chosen to create a set of criteria to define an incidental variant in any gene. These criteria include considering only variants that are (1) present in a gene that is unrelated to the primary clinical concern(s) of the patient (typically not in PGL) and are also (2) predicted to be pathogenic or likely pathogenic according to current American College of Medical Genetics sequence interpretation guidelines. The final decision to report was made by the genomic data board. In concert with recent modifications,[25] we allowed for the patient to opt out of receiving such incidental findings via our consent form, though 97% of patients or parents (252 of 260) have chosen to receive them. Incidental variants have only been reported in 5% of cases and include likely pathogenic variants found in *BRCA1* and *BRCA2*, Lynch syndrome genes, and cardiomyopathies or hereditary arrhythmias.

Discussion

Clinical exome sequencing has rapidly become a component of the clinical approach to individuals with rare diseases and is being applied to a wide range of clinical presentations that require a broad search for causal variants across the spectrum of genetically heterogeneous Mendelian disorders. Similarly to the initial description of CES performed at Baylor College of Medicine,[13,26] the current study describes a molecular diagnostic rate of about 26%. Referrals to our center are for a much broader range of potential Mendelian genetic diseases than in the prior study with more than half of the sequencing requested for nonneurological diseases. However, developmental delay remains the most common reason for testing. In the instance of a new presentation of an affected child with multiorgan syndromic features and without any prior family history of a possible genetic syndrome, our data support that trio-CES is more sensitive than proband-CES, especially for genetically heterogeneous conditions such as developmental delay. This is primarily because of increased sensitivity to observe de novo variants and compound heterozygous variants. Because on average only a single de novo variant is observed per trio-CES, de novo variants have the potential to highlight novel disease-causing genes. The de novo mutations in *KCNT1* and *TUBB2A*, prior to their description as clinical genes, serve as examples of this potential. However, trio-CES did not increase the diagnostic rate for cardiomyopathy, cancer predisposition, and disorder of sex development patient groups. This may be due to ordering physicians preferentially selecting proband-CES if they have a suspected gene in mind for these nonsyndromic disorders and reserving trio-CES for more complicated cases, in which there is greater uncertainty.

This study has a number of important limitations. For example, physicians may not order trio-CES due to concerns about increased cost of trio-CES or concerns about incidental findings in unaffected parents. Because the trio-CES and proband-CES are not randomized, other unobserved confounding factors may also be affecting the diagnostic yield. Additionally, diagnostic usage of CES is not able to detect all causal mutation types, and thus specific mutations will be unobserved by this test. For example, CES does not detect pathogenic repeat expansions (for disorders such as spinocerebellar ataxia) or most copy number variants. Also, CES does not sample all protein-coding bases: the average sequence coverage information for each gene is available online.[27] The relative coverage of any given considered gene list by referring physicians is an important consideration for the appropriateness of CES. Our study is reporting a consecutive case series referred for CES from a wide variety of clinical practices. Thus, we do not have a full accounting of prior genetic or other phenotypic testing to allow assessment of cost-effectiveness.

Another challenging part of the CES test is the interpretation of the variants in the context of the phenotypic data provided. Although every effort is made to collect as much clinical information as possible from the referring clinicians, there are instances in which the clinician's input at the genomic data board discussion is useful to assess the plausibility of a given variant in the context of a more complete clinical description. Reanalysis of a negative exome data with updated PGL is only performed upon patient's or physician's request. There was only 1 such request and the additional information did not result in identification of a significant variant. Challenges remain in the interpretation of many cases. Improved

knowledge of rare allele frequencies of healthy individuals, improved coverage of the genome by sequencing, enhanced methods for detecting all types of genetic variation, and more routine use of trio-CES will improve molecular diagnostic success rates as this field matures.

Conclusions

In this sample of patients with undiagnosed, suspected genetic conditions, trio-CES was associated with higher molecular diagnostic yield than proband-CES or traditional molecular diagnostic methods. Additional studies designed to validate these findings and to explore the effect of this approach on clinical and economic outcomes are warranted.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding/Support: Part of the study was supported by the National Institutes of Health's National Center for Advancing Translational Science (UCLA CTSI UL1TR000124) and National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIH-P30-5P30AR057230), the California Institute of Regenerative Medicine (RT2-01985), the K12 Child Health Research Career Development Award (2K12HD034610-16), and the Hyundai Hope on Wheels Scholar Award.

Role of the Funder/Sponsor: The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and decision to submit the manuscript for publication.

Additional Contributions: We thank the staffs at the University of California, Los Angeles, Clinical Genomics Center for their compensated assistance in performing the clinical exome sequencing test, routing samples, and administrative tasks.

References

1. Griffin HR, Pyle A, Blakely EL, et al. Accurate mitochondrial DNA sequencing using off-target reads provides a single test to identify pathogenic point mutations. Genet Med. published online June 5, 2014. 10.1038/gim.2014.66
2. Rehder CW, David KL, Hirsch B, Toriello HV, Wilson CM, Kearney HM. American College of Medical Genetics and Genomics: standards and guidelines for documenting suspected consanguinity as an incidental finding of genomic testing. Genet Med. 2013; 15(2):150–152. [PubMed: 23328890]
3. Papenhausen P, Schwartz S, Risheg H, et al. UPD detection using homozygosity profiling with a SNP genotyping microarray. Am J Med Genet A. 2011; 155A(4):757–768. [PubMed: 21594998]
4. Strom SP, Lee H, Das K, et al. Assessing the necessity of confirmatory testing for exome-sequencing results in a clinical molecular diagnostic laboratory. Genet Med. 2014; 16(7):510–515. [PubMed: 24406459]
5. Yourshaw M, Taylor SP, Rao AR, Martín MG, Nelson SF. Rich annotation of DNA sequencing variants by leveraging the Ensembl Variant Effect Predictor with plugins. Brief Bioinform. published online March 12, 2014. 10.1093/bib/bbu008
6. MacArthur DG, Manolio TA, Dimmock DP, et al. Guidelines for investigating causality of sequence variants in human disease. Nature. 2014; 508(7497):469–476. [PubMed: 24759409]
7. Duzkale H, Shen J, McLaughlin H, et al. A systematic approach to assessing the clinical significance of genetic variants. Clin Genet. 2013; 84(5):453–463. [PubMed: 24033266]
8. Richards CS, Bale S, Bellissimo DB, et al. ACMG recommendations for standards for interpretation and reporting of sequence variations. Genet Med. 2008; 10(4):294–300. [PubMed: 18414213]

9. Stenson PD, Mort M, Ball EV, Shaw K, Phillips A, Cooper DN. The Human Gene Mutation Database. Hum Genet. 2014; 133(1):1–9. [PubMed: 24077912]
10. GraphPad Software. [Accessed October 8, 2014] Data Analysis Resource Center. http://graphpad.com/quickcalcs/contingency1.cfm
11. GraphPad Software. [Accessed October 8, 2014] QuickCalcs. http://graphpad.com/quickcalcs/ConfInterval1.cfm
12. Hutchon, DJR. [Accessed October 8, 2014] Calculator for confidence intervals of odds ratio in an unmatched case control study. http://www.hutchon.net/ConfidOR.htm
13. Yang Y, Muzny DM, Reid JG, et al. Clinical whole-exome sequencing for the diagnosis of mendelian disorders. N Engl J Med. 2013; 369(16):1502–1511. [PubMed: 24088041]
14. O'Roak BJ, Vives L, Girirajan S, et al. Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. Nature. 2012; 485(7397):246–250. [PubMed: 22495309]
15. Fogel BL, Lee H, Deignan JL, et al. Exome sequencing in the clinical diagnosis of sporadic or familial cerebellar ataxia. JAMA Neurol. published online August 18, 2014. 10.1001/jamaneurol.2014.1944
16. DSS Research. [Accessed October 8, 2014] Calculators. https://www.dssresearch.com/KnowledgeCenter/toolkitcalculators/samplesizecalculators.aspx
17. Heron SE, Smith KR, Bahlo M, et al. Missense mutations in the sodium-gated potassium channel gene *KCNT1* cause severe autosomal dominant nocturnal frontal lobe epilepsy. Nat Genet. 2012; 44(11):1188–1190. [PubMed: 23086396]
18. Cushion TD, Paciorkowski AR, Pilz DT, et al. De novo mutations in the β-tubulin gene *TUBB2A* cause simplified gyral patterning and infantile-onset epilepsy. Am J Hum Genet. 2014; 94(4):634–641. [PubMed: 24702957]
19. Strom SP, Lozano R, Lee H, et al. De Novo variants in the *KMT2A* (*MLL*) gene causing atypical Wiedemann-Steiner syndrome in 2 unrelated individuals identified by clinical exome sequencing. BMC Med Genet. 2014; 15(1):49. [PubMed: 24886118]
20. Green RC, Berg JS, Grody WW, et al. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. Genet Med. 2013; 15(7):565–574. [PubMed: 23788249]
21. Holtzman NA. ACMG recommendations on incidental findings are flawed scientifically and ethically. Genet Med. 2013; 15(9):750–751. [PubMed: 24008255]
22. Allyse M, Michie M. Not-so-incidental findings. Trends Biotechnol. 2013; 31(8):439–441. [PubMed: 23664778]
23. Yu JH, Harrell TM, Jamal SM, Tabor HK, Bamshad MJ. Attitudes of genetics professionals toward the return of incidental results from exome and whole-genome sequencing. Am J Hum Genet. 2014; 95(1):77–84. [PubMed: 24975944]
24. Tabor HK, Auer PL, Jamal SM, et al. NHLBI Exome Sequencing Project. Pathogenic variants for Mendelian and complex traits in exomes of 6517 European and African Americans. Am J Hum Genet. 2014; 95(2):183–193. [PubMed: 25087612]
25. American College of Medical Genetics and Genomics. [Accessed May 30, 2014] ACMG updates recommendation on “opt out” for genome sequencing return of results. https://www.acmg.net/docs/Release_ACMGUpdatesRecommendations_final.pdf
26. Jacob HJ. Next-generation sequencing for clinical diagnostics. N Engl J Med. 2013; 369(16):1557–1558. [PubMed: 24088040]
27. UCLA Health. [Accessed October 8, 2014] UCLA Clinical Genomics Center. http://pathology.ucla.edu/genomics

Table 1

Overall Molecular Diagnosis Rate

			CES Test					
	Total (N = 814)		Proband (n = 338)		Trio (n = 410)		Other (n = 66)[a]	
	No. of Patients	% (95% CI)	No. of Patients	% (95% CI)	No. of Patients	% (95% CI)	No. of Patients	% (95% CI)
Diagnosis	213	26 (23-29)	74	22 (18-27)	127	31 (27-36)	12	18 (11-29)
Potential diagnosis	228	28 (25-31)	121	36 (31-41)	84	20 (17-25)	23	35 (24-47)
No significant variant	342	42 (39-46)	139	41 (36-46)	173	42 (38-47)	30	45 (34-57)
Other[b]	31	4 (3-5)	4	1 (0-3)	26	6 (4-9)	1	2 (0-9)

Abbreviation: CES, clinical exome sequencing.

[a] The other CES group includes cases in which only 1 or no parent was sequenced but other family members were sequenced.

[b] Other includes cases in which we confirmed previously reported microarray finding that the genomic data board determined to be likely pathogenic, cases in which we reported variants of uncertain significance in novel genes, and cases in which CES was ordered to rule out a clinical diagnosis.

JAMA. Author manuscript; available in PMC 2015 May 12.

Table 2

Distribution of Mutation Types for Trio–Clinical Exome Sequencing Cases With Conclusive Molecular Diagnosis

Mutation Type	No. of Cases (n = 127)	% (95% CI)
De novo	63	50 (41-58)
Homozygous	20	16 (10-23)
Compound heterozygous	26	20 (14-28)
Inherited heterozygous[a]	5	4 (1-9)
Copy number variant or uniparental disomy	3	2 (1-7)
X-linked hemizygous	10	8 (4-14)

[a] Inherited from 1 of the affected or unaffected parents in an autosomal dominant disorder gene with unknown or lower penetrance.

Table 3

Overall Molecular Diagnosis Rate of Phenotypic Subgroups by Clinical Exome Sequencing Test Type

			CES Test					
	All		Proband		Trio[a]		Other[b]	
Phenotypic Subgroup	**Rate**	**% (95% CI)**	**Rate**	**% (95% CI)**	**Rate**	**% (95% CI)**	**Rate**	**% (95% CI)**
DD	83/298	28 (23-33)	12/63	19 (11-31)	68/211	32 (26-39)	3/24	13 (4-32)
DD + hypotonia	28/111	25 (18-34)	2/27	7 (1-24)	26/80	33 (23-43)	0/4	0 (0-55)
DD + epilepsy or seizures	32/108	30 (22-39)	4/26	15 (6-34)	27/73	37 (27-48)	1/9	11 (0-46)
DD + dysmorphic features	44/144	31 (24-39)	4/30	13 (5-30)	37/99	37 (28-47)	3/15	20 (6-46)
DD + autism	11/69	16 (9-27)	0/10	0 (0-32)	10/47	21 (12-35)	1/12	8 (0-38)
DD + heart disorder	11/36	31 (18-47)	1/4	25 (3-71)	9/30	30 (17-48)	1/2	50 (9-91)
Ataxia and related neurological disorders	11/86	13 (7-22)	10/77	13 (7-22)	1/6	17 (1-58)	0/3	0 (0-62)
Muscular dystrophy and related disorders	22/74	30 (20-41)	14/57	25 (15-37)	8/15	53 (30-75)	0/2	0 (0-71)
Cardiomyopathy and arrhythmia	10/39	26 (14-41)	7/23	30 (15-51)	3/14	21 (7-48)	0/2	0 (0-71)
Cancer predisposition	7/36	19 (9-35)	5/15	33 (15-59)	1/16	6 (0-30)	1/5	20 (2-64)
Disorder of sexual development	6/35	17 (8-33)	5/18	28 (12-51)	0/14	0 (0-25)	1/3	33 (6-80)
Retinal disorders	15/31	48 (32-65)	5/12	42 (19-68)	7/11	64 (35-85)	3/8	38 (13-70)

Abbreviations: CES, clinical exome sequencing; DD, developmental delay.

[a]The trio-CES group includes 12 quartets (trio + sibling) and 1 quintet (trio + 2 siblings).

[b]The other CES group includes cases where only 1 or no parent was sequenced but other family members were sequenced.

JAMA. Author manuscript; available in PMC 2015 May 12.

Table 4

Molecular Diagnosis Rate of Phenotypic Subgroups by Age Group

	Age Groups											
	<5 y				5-18 y				>18 y			
	Proband-CES		Trio-CES		Proband-CES		Trio-CES		Proband-CES		Trio-CES	
Phenotypic Subgroup	**Rate**	**% (95% CI)**	**Rate**	**% (95% CI)**	**Rate**	**% (95% CI)**	**Rate**	**% (95% CI)**	**Rate**	**% (95% CI)**	**Rate**	**% (95% CI)**
DD	2/23	9 (1-28)	45/109	41 (32-51)[a]	8/30	27 (14-45)	22/93	24 (16-33)	2/10	20 (5-52)	1/9	11 (0-46)
DD + hypotonia	0/17	0 (0-22)	20/44	45 (32-60)[a]	2/7	29 (8-65)	6/34	18 (8-34)	0/3	0 (0-62)	0/2	0 (0-71)
DD + epilepsy or seizures	0/8	0 (0-37)	15/29	52 (34-69)[a]	3/14	21 (7-48)	11/39	28 (16-44)	1/4	25 (3-71)	1/5	20 (2-64)
DD + dysmorphic features	2/13	15 (3-43)	25/57	44 (32-57)	1/14	7 (0-34)	11/38	29 (17-45)	1/3	33 (6-80)	1/4	25 (3-71)
DD + autism	0/3	0 (0-62)	5/20	25 (11-47)	0/6	0 (0-44)	4/23	17 (6-38)	0/1	0 (0-83)	1/4	25 (3-71)
DD + heart disorder	0/1	0 (0-83)	7/18	39 (20-61)	0/1	0 (0-83)	2/11	18 (4-49)	1/2	50 (9-91)	0/1	0 (0-83)
Ataxia and related neurological disorders					0/3	0 (0-62)	0/4	0 (0-55)	10/74	14 (7-23)	1/2	50 (9-91)
Muscular dystrophy and related disorders	0/1	0 (0-83)	4/5	80 (36-98)	3/11	27 (9-57)	1/4	25 (3-71)	11/45	24 (14-39)	3/6	50 (19-81)
Cardiomyopathy and arrhythmia	1/4	25 (3-71)	0/4	0 (0-55)	2/4	50 (15-85)	2/8	25 (6-60)	4/15	27 (10-52)	1/2	50 (9-91)
Cancer predisposition			0/5	0 (0-49)	0/2	0 (0-71)	0/6	0 (0-44)	5/13	38 (18-65)	1/5	20 (2-64)
Disorder of sexual development	3/8	38 (13-70)	0/8	0 (0-37)	0/7	0 (0-40)	0/5	0 (0-49)	2/3	67 (20-94)	0/1	0 (0-83)
Retinal disorders							3/5	60 (23-88)	5/12	42 (19-68)	4/6	67 (30-91)

Abbreviations: CES, clinical exome sequencing; DD, developmental delay.

[a] *P* values for testing if the diagnostic rate of trio-CES was higher than proband-CES were .01 or less. The other CES group was not included in the calculation. *P* value was only calculated for these 3 groups that had significant (80%) power based on the observed proportions.[16]