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; Stanley F. Nelson, MD

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

**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% (2 of 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.
Clinical Exome Sequencing for Genetic Identification of Rare Mendelian Disorders

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 Sequencing Test**

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 exonic 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 highly 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.
Table 1. Overall Molecular Diagnosis Rate

<table>
<thead>
<tr>
<th></th>
<th>Total (N = 814)</th>
<th>Proband (n = 338)</th>
<th>Trio (n = 410)</th>
<th>Other (n = 66)*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No. of Patients</strong></td>
<td>% (95% CI)</td>
<td>% (95% CI)</td>
<td>% (95% CI)</td>
<td>% (95% CI)</td>
</tr>
<tr>
<td>Diagnosis</td>
<td>213</td>
<td>26 (23-29)</td>
<td>74</td>
<td>22 (18-27)</td>
</tr>
<tr>
<td>Potential diagnosis</td>
<td>228</td>
<td>28 (25-31)</td>
<td>121</td>
<td>36 (31-41)</td>
</tr>
<tr>
<td>No significant variant</td>
<td>342</td>
<td>42 (39-46)</td>
<td>139</td>
<td>41 (36-46)</td>
</tr>
<tr>
<td>Other†</td>
<td>31</td>
<td>4 (3-5)</td>
<td>4</td>
<td>1 (0-3)</td>
</tr>
</tbody>
</table>

Abbreviation: CES, clinical exome sequencing.

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

† 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.

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 of 294 patients, 51% [95% CI, 46%-57%]).

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.

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.

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. All other comparisons were done by a 2-tailed Fisher exact test. P value of .05 was used as a significance threshold. The 95% CI for proportion was calculated using an online calculator. The odds ratio and 95% CI for the significance of difference in diagnostic rate were also calculated using an online calculator.

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 (eTable 4 in the Supplement). In the childhood 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 (eTable 4).
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 of 127 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 (P = .03). 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 = .02). 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

### Table 3. Overall Molecular Diagnosis Rate of Phenotypic Subgroups by Clinical Exome Sequencing Test Type

<table>
<thead>
<tr>
<th>Phenotypic Subgroup</th>
<th>All No. of Cases</th>
<th>% (95% CI)</th>
<th>Proband No. of Cases</th>
<th>% (95% CI)</th>
<th>Trio No. of Cases</th>
<th>% (95% CI)</th>
<th>Other No. of Cases</th>
<th>% (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DD</td>
<td>83/298</td>
<td>28 (23-33)</td>
<td>12/61</td>
<td>19 (11-31)</td>
<td>68/211</td>
<td>32 (26-39)</td>
<td>3/24</td>
<td>13 (4-32)</td>
</tr>
<tr>
<td>DD + hypotonia</td>
<td>28/111</td>
<td>25 (18-34)</td>
<td>2/27</td>
<td>7 (1-24)</td>
<td>26/80</td>
<td>33 (23-43)</td>
<td>0/4</td>
<td>0 (0-55)</td>
</tr>
<tr>
<td>DD + epilepsy or seizures</td>
<td>32/108</td>
<td>30 (22-39)</td>
<td>4/26</td>
<td>15 (6-34)</td>
<td>27/73</td>
<td>37 (27-48)</td>
<td>1/9</td>
<td>11 (4-60)</td>
</tr>
<tr>
<td>DD + autism</td>
<td>11/69</td>
<td>16 (9-27)</td>
<td>0/10</td>
<td>0 (0-32)</td>
<td>10/47</td>
<td>21 (12-35)</td>
<td>1/12</td>
<td>8 (0-38)</td>
</tr>
<tr>
<td>DD + heart disorder</td>
<td>11/36</td>
<td>31 (18-47)</td>
<td>1/4</td>
<td>25 (3-71)</td>
<td>9/30</td>
<td>30 (17-48)</td>
<td>1/2</td>
<td>50 (9-91)</td>
</tr>
<tr>
<td>Ataxia and related neurological disorders</td>
<td>11/86</td>
<td>13 (7-22)</td>
<td>10/77</td>
<td>13 (7-22)</td>
<td>1/6</td>
<td>17 (1-58)</td>
<td>0/3</td>
<td>0 (0-62)</td>
</tr>
<tr>
<td>Muscular dystrophy and related disorders</td>
<td>22/74</td>
<td>30 (20-41)</td>
<td>14/57</td>
<td>25 (15-37)</td>
<td>8/15</td>
<td>53 (30-75)</td>
<td>0/2</td>
<td>0 (0-71)</td>
</tr>
<tr>
<td>Cardiomyopathy and arrhythmia</td>
<td>10/39</td>
<td>26 (14-41)</td>
<td>7/23</td>
<td>30 (15-51)</td>
<td>3/14</td>
<td>21 (7-48)</td>
<td>0/2</td>
<td>0 (0-71)</td>
</tr>
<tr>
<td>Cancer predisposition</td>
<td>7/36</td>
<td>19 (9-35)</td>
<td>5/15</td>
<td>33 (15-59)</td>
<td>1/16</td>
<td>6 (0-30)</td>
<td>1/5</td>
<td>20 (2-64)</td>
</tr>
<tr>
<td>Disorder of sexual development</td>
<td>6/35</td>
<td>17 (8-33)</td>
<td>5/18</td>
<td>28 (12-51)</td>
<td>0/14</td>
<td>0 (0-25)</td>
<td>1/3</td>
<td>33 (6-80)</td>
</tr>
<tr>
<td>Retinal disorders</td>
<td>15/31</td>
<td>48 (32-65)</td>
<td>5/12</td>
<td>42 (19-68)</td>
<td>7/11</td>
<td>64 (35-85)</td>
<td>3/8</td>
<td>38 (13-70)</td>
</tr>
</tbody>
</table>

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

* The trio-CES group includes 12 quartets (trio + sibling) and 1 quintet (trio + 2 siblings).
Table 4. Molecular Diagnosis Rate of Phenotypic Subgroups by Age Group

<table>
<thead>
<tr>
<th>Phenotypic Subgroup</th>
<th>Proband-CES</th>
<th>Trios-CES</th>
<th>Mean</th>
<th>Rate</th>
<th>95% CI</th>
<th>Rate</th>
<th>95% CI</th>
<th>Rate</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>&lt;5 y</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DD</td>
<td>2/23</td>
<td>9 (95%)</td>
<td>45/109</td>
<td>41   (32-51)</td>
<td>2/7</td>
<td>29 (8-65)</td>
<td>6/34</td>
<td>18 (8-34)</td>
<td>2/10</td>
</tr>
<tr>
<td>DD + hypotonia</td>
<td>8/17</td>
<td>47 (95%)</td>
<td>45/109</td>
<td>45   (32-60)</td>
<td>8/14</td>
<td>57 (7-49)</td>
<td>3/11</td>
<td>27 (9-57)</td>
<td>1/3</td>
</tr>
<tr>
<td>DD + epilepsy or seizures</td>
<td>6/13</td>
<td>46 (95%)</td>
<td>25/57</td>
<td>44 (32-57)</td>
<td>6/14</td>
<td>7 (0-34)</td>
<td>11/38</td>
<td>29 (17-45)</td>
<td>1/3</td>
</tr>
<tr>
<td>DD + dysmorphic features</td>
<td>2/13</td>
<td>15 (93%)</td>
<td>25/57</td>
<td>25 (11-47)</td>
<td>0/6</td>
<td>0 (0-44)</td>
<td>4/23</td>
<td>17 (6-38)</td>
<td>0/1</td>
</tr>
<tr>
<td>DD + heart disorder</td>
<td>6/34</td>
<td>18 (95%)</td>
<td>7/18</td>
<td>39 (20-61)</td>
<td>0/1</td>
<td>0 (0-83)</td>
<td>2/11</td>
<td>18 (4-49)</td>
<td>1/2</td>
</tr>
<tr>
<td>Ataxia and related neurological disorders</td>
<td>0/3</td>
<td>0 (62)</td>
<td>0/4</td>
<td>0 (55)</td>
<td>0/104</td>
<td>14 (7-23)</td>
<td>1/2</td>
<td>50 (9-91)</td>
<td>0/1</td>
</tr>
<tr>
<td>Muscular dystrophy and related disorders</td>
<td>0/1</td>
<td>0 (83)</td>
<td>0/5</td>
<td>0 (36)</td>
<td>0/11</td>
<td>27 (9-57)</td>
<td>1/4</td>
<td>25 (3-71)</td>
<td>11/45</td>
</tr>
<tr>
<td>Cardiomyopathy and arrhythmia</td>
<td>1/6</td>
<td>16 (95)</td>
<td>0/4</td>
<td>0 (55)</td>
<td>0/2</td>
<td>0 (7-11)</td>
<td>0/6</td>
<td>0 (4-44)</td>
<td>5/13</td>
</tr>
<tr>
<td>Cancer predisposition</td>
<td>1/3</td>
<td>33 (95)</td>
<td>0/8</td>
<td>0 (37)</td>
<td>0/5</td>
<td>0 (40-40)</td>
<td>0/5</td>
<td>0 (4-49)</td>
<td>2/3</td>
</tr>
<tr>
<td>Disorder of sexual development</td>
<td>3/8</td>
<td>38 (95)</td>
<td>0/8</td>
<td>0 (37)</td>
<td>0/7</td>
<td>0 (40-40)</td>
<td>0/5</td>
<td>0 (4-49)</td>
<td>2/3</td>
</tr>
<tr>
<td>Retinal disorders</td>
<td>3/5</td>
<td>60 (23-88)</td>
<td>5/12</td>
<td>42 (13-68)</td>
<td>4/6</td>
<td>67 (30-91)</td>
<td>4/6</td>
<td>67 (30-91)</td>
<td>4/6</td>
</tr>
</tbody>
</table>

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

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 premature adult teeth in whom trio-CES identified a novel 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 missense de 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 homozgyous 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 pater-
 modifications, we allowed for the patient to opt out of making decisions by the genomic data board. In recent guidelines, the final decision to report was based on the American College of Medical Genetics sequence coverage information for each gene, which does not sample all protein-coding bases: the average of these findings and to explore the effect of this approach on clinical practices. Thus, we do not have a full accounting of diagnostic yield for 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 spino-cerebellar 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. The relative coverage of any given considered gene list by referring physicians is an important consideration for practice. 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 PGD 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.
Clinical Exome Sequencing for Genetic Identification of Rare Mendelian Disorders

Robert A. Kujirakwe, MD, PhD
Sharon F. Lench, MD
David A. Brown, MD, PhD

ARTICLE INFORMATION
Published Online: October 18, 2014, doi:10.1001/jama.2014.14604.

Author Affiliations: 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, Strom, Kantarci, Quintero-Rivera, Das, Toy, Grody, Vilain); Department of Pediatrics, David Geffen School of Medicine, University of California, Los Angeles (Dorrani, Youshaw, 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).

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, Youshaw, 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.

Funding/Sponsor: Part of the study was supported by the National Institutes of Health’s National Center for Advancing Translational Science (UCLA CT5 ULTRRO0024) and National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIH-P30-SP30AR057230), the California Institute of Regenerative Medicine (RT2-01985), the K12 Child Health Research Career Development Award (K21DH034610-16), 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