

# Increasing the sensitivity of clinical exome sequencing through improved filtration strategy

Hanan E. Shamseldin<sup>1</sup>, Sateesh Maddirevula<sup>1</sup>, Eissa Faqeih<sup>2</sup>, Niema Ibrahim<sup>1</sup>, Mais Hashem<sup>1</sup>, Ranad Shaheen<sup>1</sup> and Fowzan S. Alkuraya<sup>1,3</sup>

**Background:** Clinical exome sequencing (CES) has greatly improved the diagnostic process for individuals with suspected genetic disorders. However, the majority remains undiagnosed after CES. Although understanding potential reasons for this limited sensitivity is critical for improving the delivery of clinical genomics, research in this area has been limited.

**Materials and Methods:** We first calculated the theoretical maximum sensitivity of CES by analyzing >100 families in whom a Mendelian phenotype is mapped to a single locus. We then tested the hypothesis that positional mapping can limit the search space and thereby facilitate variant interpretation by reanalyzing 33 families with “negative” CES and applying positional mapping.

**Results:** We found that >95% of families who map to a single locus harbored genic (as opposed to intergenic) variants that

are potentially identifiable by CES. Our reanalysis of “negative” CES revealed likely causal variants in the majority (88%). Several of these solved cases have undergone negative whole-genome sequencing.

**Conclusion:** The discrepancy between the theoretical maximum and the actual clinical sensitivity of CES is primarily in the variant filtration rather than the variant capture and sequencing phase. The solution to negative CES is not necessarily in expanding the coverage but rather in devising approaches that improve variant filtration. We suggest that positional mapping is one such approach.

*Genet Med* advance online publication 6 October 2016

**Key Words:** autozygome; clinical genomics; clinical exome; clinical genome; variant interpretation

## INTRODUCTION

Whole-exome sequencing (WES) is a powerful application of massively parallel sequencing technology that allows the sequencing of the entire coding genome and its flanking sequence.<sup>1</sup> Since its first application to identify the cause of a Mendelian disorder, it has become a standard tool in the study of these diseases, which are caused by variants in single genes.<sup>2,3</sup> This success has quickly penetrated the clinical arena, where clinical exome sequencing (CES) is increasingly integrated in the diagnostic process of individuals with suspected genetic disorders. The power of CES to identify the causal variants without the need for expensive ancillary tests or subjective clinical evaluation is especially appealing for such heterogeneous disorders as developmental delay and intellectual disability. Recent studies show that CES is economically competitive with the traditional testing strategy, arguing for its potential use as a first-tier standard diagnostic test.<sup>4–6</sup>

Clinical sensitivity of CES, i.e., the percentage of cases for which CES identifies a likely causal variant for the phenotype undergoing investigation, is surprisingly comparable among large and heterogeneous case series. For example, Retterer et al. identified a likely causal variant in 28.8% of 3,040 samples.<sup>7</sup>

Similarly, Farwell and colleagues reported a yield of 30% among their first 500 tested cases.<sup>8</sup> These figures are very similar to those generated by previous large studies such as those by Yang et al. and Lee et al. that reported yields of 25.2% for 2,000 cases and 26% for 814 cases, respectively.<sup>9,10</sup> A lower yield was reported among 486 adults with a suspected genetic diagnosis for various indications (17.5%).<sup>11</sup>

Although very high compared to other tests, this clinical sensitivity leaves >70% of patients with suspected genetic disorders undiagnosed molecularly. Nongenetic or non-Mendelian etiology may explain some of the “negative” CES; however, CES in conditions very likely to be Mendelian, e.g., retinal dystrophy and intellectual disability with positive family history, also suffered from a limited sensitivity.<sup>9,10,12</sup> Another possible explanation is that CES was rarely ordered as a first-tier test in the published cohorts, so it is possible that tested cases were enriched for unlikely mutational mechanisms and novel genes that may be more difficult to uncover by CES. Recent studies of CES as a first-tier test did suggest a higher sensitivity (50–60%); this may explain the limited clinical sensitivity but only partially.<sup>6,13</sup> Finally, there is a growing concern that cases with “negative” CES may harbor variants that are not captured or

The first two authors contributed equally to this work.

<sup>1</sup>Department of Genetics, King Faisal Specialist Hospital and Research Center, Riyadh, Saudi Arabia; <sup>2</sup>Department of Pediatric Subspecialties, Children's Hospital, King Fahad Medical City, Riyadh, Saudi Arabia; <sup>3</sup>Department of Anatomy and Cell Biology, College of Medicine, Alfaisal University, Riyadh, Saudi Arabia. Correspondence: Fowzan S Alkuraya (falkuraya@kfsshr.edu.sa)

Submitted 30 March 2016; accepted 24 August 2016; advance online publication 6 October 2016. doi:10.1038/gim.2016.155

sequenced by the assay, although the magnitude of this is currently unknown.

One major challenge in addressing the issue of clinical sensitivity of CES is lack of data regarding the theoretical maximum potential of CES. In other words, the contribution of the various classes of mutations to the etiology of Mendelian diseases is unknown. For example, although the limited available data regarding WGS suggest that the contribution of nongenic variants to Mendelian diseases is very limited,<sup>14,15</sup> this may have been an underestimate due to bias in interpreting this difficult class of mutations. One approach that has the potential to address this challenge is positional mapping. A Mendelian phenotype can be mapped to a single positional locus with sufficient meiotic events irrespective of the nature of the mutation. Therefore, interrogating a large number of phenotypes that are mapped to a single locus each can provide the much needed benchmark to calculate the maximum potential of CES as a diagnostic test, as we show in this study. After showing that the theoretical maximum yield of CES is much higher than what is experienced in practice, we demonstrate that the causal variants in the majority of “negative” CES can indeed be identified by improved variant filtration rather than increased coverage.

## MATERIALS AND METHODS

### Human subjects

Families who map to single loci have been recruited as part of several positional mapping projects with appropriate consenting and IRB approval. They are all consanguineous and have several members with the same autosomal-recessive phenotype. Patients enrolled in this specific study have all been evaluated clinically for suspected genetic disorders and received a negative CES report, i.e., a report in which no likely causal variant for the indication of the test was identified. CES was performed in major diagnostic laboratories (names are not revealed for confidentiality).

### Positional mapping

All families with single loci that we present in this paper were mapped using autozygosity analysis essentially as described before.<sup>16</sup> Briefly, we first performed genome-wide genotyping using the Axiom SNP Chip platform following the manufacturer's instructions (Affymetrix). We then scanned the generated files for regions of homozygosity  $\geq 2$  Mb as surrogates of autozygosity given the parental consanguinity using AutoSNPa.<sup>17</sup> The entire set of autozygous blocks per genome (autozygome) was compared among family members with the same phenotype. Only families with a single autozygous interval exclusively shared among the affected members are included in **Supplementary Table S1** online. For our reanalysis of “negative” CES, we used the autozygome of the tested individual and, when applicable, other affected relatives to filter the exomic variants, essentially as described before.<sup>3,18</sup>

### Exome sequencing and variant interpretation

Obtaining the raw CES data was not possible, so we had to repeat WES to perform the reanalysis of “negative” CES cases. Exome

capture was performed using the TruSeq Exome Enrichment kit (Illumina) following the manufacturer's protocol. Samples were prepared as an Illumina sequencing library; in the second step, the sequencing libraries were enriched for the desired target using the Illumina Exome Enrichment protocol. The captured libraries were sequenced using Illumina HiSeq 2000 Sequencer. The reads were mapped against UCSC hg19 by BWA. The SNPs and indels were detected by SAMTOOLS. Variants from WES were filtered such that only novel (or very low-frequency 0.1%), coding/splicing, homozygous variants that are within the autozygome of the affected individual (or shared autozygome of the affected individuals when applicable) were considered likely causal variants. Frequency of variants was determined using publically available variant databases (1000 Genomes, Exome Variant Server, and ExAC Browser) as well as a database of 817 in-house ethnically matched exomes as described previously.<sup>19</sup> The ACMG guidelines on variant interpretation and classification were followed.<sup>20</sup> Because the ACMG guidelines apply only to established disease genes, we labeled variants in novel genes as variants of unknown significance.

## RESULTS

### Very high theoretical maximum sensitivity of CES

Our group has been working on mapping recessive Mendelian phenotypes since 2007 by exploiting the special structure of the local population with consanguinity, large family structure, and free access to health care.<sup>21</sup> By surveying the thousands of families recruited in the process, we were able to identify 104, mostly published, families in whom a phenotype that follows the autosomal-recessive mode of inheritance could be mapped to a single autozygous interval (**Supplementary Figure S1** online shows the pedigree of all unpublished families). Using this as a denominator, we found that the underlying variant was identified within the critical interval in all but three (97%). **Supplementary Table S1** online summarizes the classes of variants identified in these families. Although all identified variants were genic, intronic variants deserve a special mention. In aggregate, splicing variants accounted for 17% of the total variants and 2% were  $>50$  nucleotides from the nearest exon. Because the latter group is not typically captured by the design of CES, we conclude that the hypothetical maximum yield of CES in the setting of autosomal recessive phenotypes is  $>95\%$ .

Our own experience with CES performed by the same major reference laboratories from which the literature on clinical utility of CES is derived revealed a much lower sensitivity ( $\sim 30\%$ ) compared to the hypothetical maximum of 95%. This provided an opportunity to investigate the reasons for this discrepancy and whether positional mapping can unlock the full potential of CES. Therefore, we approached families with suspected autosomal recessive phenotypes in whom CES (some additionally had clinical WGS) did not reveal a likely causal variant, i.e., “negative,” and 33 agreed to recruitment under an IRB-approved research protocol (KFSHRC RAC2121053). We then proceeded with our previously published pipeline of autozygosity analysis and WES.<sup>18</sup>

### Reanalysis of “negative” CES using positional mapping markedly improves sensitivity

**Table 1** summarizes the results of our analysis of “negative” CES. Overall, a likely causal variant was identified in 88% of the families. Even if we limited our analysis to established disease genes, a likely causal variant was identified in 48%, although we note that CES laboratories typically do report variants in novel genes; however, none of these was reported. Although the numbers are small, cases in which autosomal-recessive inheritance was clearly implied from the pedigree structure had a higher success rate compared to simplex cases presumed to be recessive on the basis of parental consanguinity only. This cohort of cases that had been “negative” on CES, despite the small numbers, gave us an opportunity to address why the causal variants were originally missed.

One prominent class of mutations was splicing. In none of these cases was the splicing variant deep enough to be missed by the capture design of CES, although they were not in the canonical splice site. These include the +8 variant in *RTTN*,<sup>22</sup> +4 in *SMG9*, and +5 in *ECHS1*. Of note, the +5 variant in *ECHS1* was not called, even by clinical whole-genome sequencing, which was requested by the referring clinician after receiving a negative CES result. Even the deep -24 variant in *COG6* was, in fact, included in the VCF by the reference laboratory but was not included in the final report.<sup>12</sup> These results suggest that splicing variants beyond the canonical 1/2 position are very difficult to call with confidence as disease-causing by CES without supporting evidence from positional mapping. The *CTU2* variant that abolishes the splice-donor site by replacing the last bp of the exon may have been missed because it masquerades as a synonymous change, although CES laboratories are usually aware of this well-established phenomenon.<sup>23</sup> Surprisingly, none of the other “missed” variants could be presumed challenging a priori. For example, the missense variants we identified in *ISCA2*, *ASNS*, *C3orf17*, and *SLC1A4* were all missed by CES, and *VRK1* was also missed on WGS. Similarly, the nonsense variant in *UNC80* is expected to be easily called by all next-generation sequencing platforms. Even the indels we identified in this resequencing study were not particularly challenging and were identified by the same sequencing platform we used in this study, as typically used by the CES laboratories, i.e., Illumina HiSeq2000 (e.g., the four-nucleotide deletion in *GOLGA24*). Again, it is worth highlighting that the latter variant was missed not only on CES but also on clinical WGS.

## DISCUSSION

CES is increasingly becoming the test of choice in clinical genomics, particularly in patients with challenging phenotypes. An important question that has arisen as a result of this widespread use is: how should a “negative” CES report be interpreted? It is tempting to speculate that coverage is the main culprit in the limited sensitivity of CES, i.e., the missed variants in “negative” CES are present outside the ~2% coding/flanking sequences that are targeted by CES. The limited available data for whole-genome sequencing (WGS), however, do not seem

to support this. For example, the application of WGS for 152 cases with various clinical phenotypes that are suspected to be genetic in etiology only revealed one nongenic variant that would have been missed by WES: a small deletion 1.5 kb downstream of *SOX3*.<sup>14</sup> The overall clinical sensitivity of WGS in that study was 21% and was not superior to that of CES.

In this report, we attempted to address this question by first calculating the theoretical maximum sensitivity of CES using an approach that is neutral to the class of variants. In fact, we suggest that the data we obtained by dissecting the various classes of variants identified in the largest published cohort of phenotypes that map to a single locus each can inform the wider question about the contribution of individual mutation classes to the etiology of Mendelian diseases. Unfortunately, our data are limited to autosomal-recessive “disease” phenotypes, so we caution against their generalization to other modes of inheritance but suspect similar data can be collected from laboratories that primarily deal with autosomal-dominant and X-linked phenotypes. Nonetheless, the finding that the overwhelming majority of autosomal-recessive disease variants are within genes is reassuring in that the extreme rarity of reports of nongenic autosomal-recessive disease variants likely reflect their genuine rarity rather than a bias against their detection.<sup>25,26</sup> Even when deep splicing variants that are not covered by the current design of CES are excluded, the sensitivity of CES is much less than the theoretical maximum of 95%. This suggests that the solution to “negative” CES is not necessarily WGS because it is likely that the causal variants have in fact been sequenced but failed to be called by the laboratory director. This is highly consistent with the data we present regarding the resequencing of individuals with “negative” CES, in which we show that, in many instances, the causal variant was “missed” because it was filtered out at the filtration stage rather than because of inherent technical limitation. This highlights the need for improved filtering algorithms used in the analysis of CES. Indeed, we posit that investing in these improved algorithms should be a priority over investing in clinical WGS because the latter will probably miss the same variants when the same algorithms are used, as demonstrated by at least three cases in this cohort. Although positional mapping is only applicable in a subset of cases, our data strongly suggest that it should be incorporated in the filtering algorithms whenever applicable and there are methods that allow this directly on VCF files.<sup>27</sup> Assuming similar quality of WES performed on a research basis by us and CES performed previously (the latter very likely has a much higher quality), the only obvious variable to which we can attribute our higher success rate is the implementation of positional mapping.

In conclusion, we present an innovative method to calculate the theoretical maximum yield of CES in the setting of autosomal-recessive diseases, which we show to be much higher than currently experienced. We demonstrate that positional mapping can minimize this discrepancy by extracting causal variants from most “negative” CES. Our data strongly support the incorporation of positional mapping in the analysis of CES whenever applicable.

**Table 1** Summary of cases with “negative” CES and the results of reanalysis using research-grade WES with autozygome filtering

ID	Phenotype	Status	Variant	Class	FHX	ACMG score	ACMG class	Number of affected	Reference (case)
13DG2280	Primary microcephaly	Potential causal variant identified	ASNS: NM_001178076.1: c.962G>A: p.(Arg321His)	M	Familial	PM1, PM2, PP1, PP3	Likely pathogenic	3	Unpublished
13DG1857	Primary microcephaly	Unsolved	?	?	Familial	NA	NA	3	Unpublished
14DG0714	Meckel-Gruber syndrome	Potential causal variant identified	<sup>a</sup>	FS (in del)	Familial	NA	NA	4	Unpublished
12DG1204	Microcephaly, postnatal progressive, with seizures and brain atrophy	Potential causal variant identified	MED17:NM_004268.4:c.1217G>C: p.(Agr406Thr)	M	Familial	PM1, PM2, PP3, PP1	Likely Pathogenic	3	Unpublished
15DG2303	Bardet-Biedl syndrome	Potential causal variant identified	BB52:NM_031885.3:c.805_1910del: p.(Val269Glufs*12)	FS (LD)	Familial	PVS1, PM2,	Pathogenic	2	Unpublished
14DG1870	Joubert syndrome	Unsolved	?	?	Simplex	NA	NA	2	Unpublished
09DG00744	Microcephaly and short stature	Unsolved	?	?	Familial	NA	NA	2	Unpublished
HMC14	Shaheen syndrome	Potential causal variant identified	COG6:NM_020751.2:c.1167-24A>G: r.1166_1167ins-1167-37_1167-2	S	Familial	PVS1, PM2, PP1	Pathogenic	7	PMID: 26077850
14DG1883	Dysmorphic facies, renal agenesis, ambiguous genitalia, polydactyly	Potential causal variant identified	CTU2:NM_001012762.1:c.873G>A: p.(=):r.738_944del	S	Familial	NA	NA	3	PMID: 26633546
10DG1904	Primary microcephaly	Potential causal variant identified	VRK1:NM_003384.2:c.236C>T: p.(Pro79Leu)	M	Familial	PM2, PP1, PP2, PP3, PP4	Likely Pathogenic	3	PMID: 27431290
15DG2238	Ehlers-Danlos syndrome (EDS) with skeletal dysplasia	Potential causal variant identified	C3ORF17:NM_015412.3:c.280C>T: p.(Arg94Cys)	M	Familial	NA	NA	2	Unpublished
15DG1457	Crisponi-Sohar syndrome	Potential causal variant identified	CRLF1:NM_004750.4:c.983dupG: p.(Ser328Argfs*2)	FS (in del)	Familial	PVS1, PM2, PP1, PP4	Pathogenic	4	PMID: 26804344
15DG2649	Dysmorphism and epilepsy	Potential causal variant identified	<sup>a</sup>	M	Familial	NA	NA	3	Unpublished
15DG2666	Epileptic encephalopathy	Potential causal variant identified	<sup>a</sup>	FS (in del)	Familial	NA	NA	2	Unpublished
14DG0248	Encephalopathy	Potential causal variant identified	SLC13A5:NM_177550.4:c.1227dup: p.(Ile410Hisfs*13)	FS (in del)	Familial	PVS1, PM2	Pathogenic	3	Unpublished
15DG0764	Skeletal dysplasia	Potential causal variant identified	C3ORF17:NM_015412.3:c.280C>T: p.(Arg94Cys)	M	Familial	NA	NA	2	PMID: 26633546
15DG0859	Microcephaly, hypotonia, and failure to thrive	Potential causal variant identified	GOLGA2:NM_004486.4: c.1263_1266del:p.(Glu423Argfs*6)	FS (in del)	Simplex	NA	NA	1	PMID: 26742501
14DG1291	Congenital microcephaly	Potential causal variant identified	R7TM:NM_17363:c.2885+8A>G:r.2885_2886ins2885+1_2885+7	S	Familial	PVS1, PM2, PP1	Pathogenic	3	PMID: 26608784
15DG2249	Mitochondrial disease	Potential causal variant identified	ECH51:NM_004092.3:C.88+5G>A:r.89_90ins89+1_89+11	S	Familial	PVS1, PM2, PP1, PP4	Pathogenic	2	Unpublished
15DG0032	Recurrent metabolic crises and encephalopathy	Potential causal variant identified	TANGO2:NM_152906.5:c.94C>T: p.(Arg32*)	NS	Simplex	PVS1, PM2	Likely pathogenic	1	Unpublished

ACMG, American College of Genetics and Genomics; FS, frameshift; LD, large deletion; M, missense; NA, NA: not applicable (novel gene where ACMG guidelines do not apply); NFS, nonframeshift; NS, nonsense; S, splice site. <sup>a</sup>Novel gene to be submitted as a separate publication.

Table 1 Continued on next page

**Table 1 Continued**

ID	Phenotype	Status	Variant	Class	FHx	ACMG score	ACMG class	Number of affected	Reference (case)
15DG1393	Recurrent acidosis	Potential causal variant identified	<i>FBP1</i> :NM_000507.3:c.114_115insCTGCA:p.(Cys39Leufs*30)	FS (indel)	Familial	PVS1, PM2, PP1	Pathogenic	2	unpublished
15DG2105	Global developmental delay with brain atrophy	Potential causal variant identified	<sup>a</sup>	NFS (indel)	Familial	NA	NA	2	Unpublished
14DG0152	Neurodegenerative disease	Potential causal variant identified	<i>ISCA2</i> :NM_194279.3:c.229G>A:p.(Gly77Ser)	M	Familial	NA	NA	3	PMID: 25558065
15DG0513	Secondary microcephaly, developmental delay, and myoclonic seizure	Potential causal variant identified	<i>SLC14A4</i> :NM_001193493.1:c.463T>C:p.(Trp155Arg)	M	Familial	PM2, PP1, PP2, PP3, PP4	Likely pathogenic	3	Unpublished
15DG0137	Skeletal dysplasia	Potential causal variant identified	<sup>a</sup>	NS	Familial	NA	NA	3	Unpublished
15DG0246	Developmental delay, dysmorphism, and hearing loss	Unsolved	?	?	Familial	NA	NA	2	Unpublished
14DG1661	Global developmental delay, dysmorphic features, and congenital heart disease	Potential causal variant identified	<i>SMG9</i> :NM_019108.2:c.701+4A>G:r.702_813del	S	Familial	PVS1, PS3, PM2	Pathogenic	2	PMID: 27018474
12DG2078	Klippel-Feil syndrome	Potential causal variant identified	<i>MYO18B</i> :NM_032608.5:c.6905C>A:p.(Ser2302*)	NS	Familial	PVS1, PM2, PP1	Pathogenic	2	PMID: 25748484
13DG2107	Psychomotor retardation and seizures	Potential causal variant identified	<i>KCTD3</i> :NM_016121.3:c.1036_1073del p.(P346Tfs*4)	FS (indel)	Familial	NA	NA	4	PMID: 25558065
15DG0797	Secondary microcephaly	Potential causal variant identified	<i>UNC80</i> :NM_182587.3:c.3778C>T:p.(Arg1260*)	NS	Familial	PVS1, PM2, PM3	Pathogenic	3	PMID: 26708753
14DG1221	Stickler-like syndrome	Potential causal variant identified	<i>HYAL2</i> :NM_033158.4:c.749C>T:p.(Pro250Leu)	M	Familial	NA	NA	3	PMID: 26633546
15DG1901	Nonimmune hydrops fetalis	Potential causal variant identified	<i>THSD1</i> :NM_018676.3:c.617G>A:p.(Cys206Tyr)	M	Simplex	PM2, PP1, PP2, PP3, PP4	Likely pathogenic	1	Unpublished
16DG0646	Cerebellar atrophy	Potential causal variant identified	<i>GRID2</i> :NM_001510.3:c.910C>T:p.(Arg304*)	NS	Familial	PVS1, PM2, PP4	Pathogenic	2	Unpublished

ACMG, American College of Genetics and Genomics; FS, frameshift; LD, large deletion; M, missense; NA, NA: not applicable (novel gene where ACMG guidelines do not apply); NFS, nonframeshift; NS, nonsense; S, splice site. <sup>a</sup>Novel gene to be submitted as a separate publication.

## SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper at <http://www.nature.com/gim>

## ACKNOWLEDGMENTS

We thank the families for their enthusiastic participation. We also thank the Sequencing and Genotyping Facilities at KFSHRC for its technical help. This work was supported by KACST grants 13-BIO1113-20 (F.S.A.) and KSCDR (F.S.A.).

## DISCLOSURE

The authors declare no conflict of interest.

## REFERENCES

1. Ng SB, Turner EH, Robertson PD, et al. Targeted capture and massively parallel sequencing of 12 human exomes. *Nature* 2009;461:272–276.
2. Shendure J. Next-generation human genetics. *Genome Biol* 2011;12:408.
3. Alkuraya FS. Discovery of mutations for Mendelian disorders. *Human Genetics*. 2016:1–9.
4. Monroe GR, Frederix GW, Savelberg SMC, et al. Effectiveness of whole-exome sequencing and costs of the traditional diagnostic trajectory in children with intellectual disability. *Genet Med* 2016; 18:949–956.
5. Gomez CM, Das S. Clinical exome sequencing: the new standard in genetic diagnosis. *JAMA Neurol* 2014;71:1215–1216.
6. Anazi S, Maddirevula S, Faqeh E, et al. Clinical genomics expands the morbid genome of intellectual disability and offers a high diagnostic yield. *Mol Psychiatry* 2016. doi: 10.1038/mp.2016.113 (e-pub ahead of print).
7. Retterer K, Juusola J, Cho MT, et al. Clinical application of whole-exome sequencing across clinical indications. *Genet Med* 2016;18:696–704.
8. Farwell KD, Shahmirzadi L, El-Khechen D, et al. Enhanced utility of family-centered diagnostic exome sequencing with inheritance model-based analysis: results from 500 unselected families with undiagnosed genetic conditions. *Genet Med* 2015;17:578–586.
9. Yang Y, Muzny DM, Xia F, et al. Molecular findings among patients referred for clinical whole-exome sequencing. *JAMA* 2014;312:1870–1879.
10. Lee H, Deignan JL, Dorrani N, et al. Clinical exome sequencing for genetic identification of rare Mendelian disorders. *JAMA* 2014;312:1880–1887.
11. Posey JE, Rosenfeld JA, James RA, et al. Molecular diagnostic experience of whole-exome sequencing in adult patients. *Genet Med* 2016;18:678–685.
12. Yavarna T, Al-Dewik N, Al-Mureikhi M, et al. High diagnostic yield of clinical exome sequencing in Middle Eastern patients with Mendelian disorders. *Hum Genet* 2015;134:967–980.
13. Stark Z, Tan TY, Chong B, et al. A prospective evaluation of whole-exome sequencing as a first-tier molecular test in infants with suspected monogenic disorders. *Genet Med*; e-pub ahead of print 3 March 2016.
14. Taylor JC, Martin HC, Lise S, et al. Factors influencing success of clinical genome sequencing across a broad spectrum of disorders. *Nat Genet* 2015;47:717–726.
15. Gilissen C, Hehir-Kwa JY, Thung DT, et al. Genome sequencing identifies major causes of severe intellectual disability. *Nature* 2014;511:344–347.
16. Alkuraya FS. Autozygome decoded. *Genet Med* 2010;12:765–771.
17. Carr IM, Flintoff KJ, Taylor GR, Markham AF, Bonthron DT. Interactive visual analysis of SNP data for rapid autozygosity mapping in consanguineous families. *Hum Mutat* 2006;27:1041–1046.
18. Alkuraya FS. The application of next-generation sequencing in the autozygosity mapping of human recessive diseases. *Hum Genet* 2013;132:1197–1211.
19. Saudi Mendeliome Group. Comprehensive gene panels provide advantages over clinical exome sequencing for Mendelian diseases. *Genome Biol* 2015;16:134.
20. Richards S, Aziz N, Bale S, et al. 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;17:405–423.
21. Alkuraya FS. Genetics and genomic medicine in Saudi Arabia. *Mol Genet Genomic Med* 2014;2:369–378.
22. Shamseldin H, Alazami AM, Manning M, et al.; Care4Rare Canada Consortium. RTTN mutations cause primary microcephaly and primordial dwarfism in humans. *Am J Hum Genet* 2015;97:862–868.
23. Shaheen R, Patel N, Shamseldin H, et al. Accelerating matchmaking of novel dysmorphology syndromes through clinical and genomic characterization of a large cohort. *Genet Med* 2016;18:686–695.
24. Shamseldin HE, Bennett AH, Alfadhel M, Gupta V, Alkuraya FS. GOLGA2, encoding a master regulator of golgi apparatus, is mutated in a patient with a neuromuscular disorder. *Hum Genet* 2016;135:245–251.
25. Weedon MN, Cebola I, Patch AM, et al.; International Pancreatic Agnesis Consortium. Recessive mutations in a distal PTF1A enhancer cause isolated pancreatic agnesis. *Nat Genet* 2014;46:61–64.
26. Bae BI, Tietjen I, Atabay KD, et al. Evolutionarily dynamic alternative splicing of GPR56 regulates regional cerebral cortical patterning. *Science* 2014;343:764–768.
27. Carr IM, Bhaskar S, O’Sullivan J, et al. Autozygosity mapping with exome sequence data. *Hum Mutat* 2013;34:50–56.