IJHS

Original Article

Applying filtration steps to interpret the results of whole-exome sequencing in a consanguineous population to achieve a high detection rate

Ahmed A. Alfares

Department of Pediatrics, College of Medicine, Qassim University, Qassim, Saudi Arabia

Address for correspondence:
Ahmed A. Alfares, Department of Pediatrics, College of Medicine, Qassim University, Qassim, Saudi Arabia. Tel.: 00966163800050. E-mail: fars@qu.edu.sa

WEBSITE: ijhs.org.sa
ISSN: 1658-3639
PUBLISHER: Qassim University

ABSTRACT

Objective: Interpreting whole-exome sequencing (WES) data are challenging, requiring extensive time, and effort to review all the variants in the variant call format. Here, we examined the application of custom filters to narrow the number of candidate variants in a consanguineous population that requires further analysis.

Methods: In 100 cases undergoing WES, we applied a custom filtration process to look primarily for homozygous variants in autosomal recessive (AR) disorders, and second for variants in either autosomal dominant or x-linked disorders.

Results: Most identified disease-causing variants were homozygous in AR disorders. By applying our custom filtration process, we narrowed the number of candidate variants requiring further analysis to 5–15 per case while maintaining a high detection rate and completing analysis in around 45 min.

Conclusion: A custom filtration process and strategy targeting a specific population provide excellent detection rates in less time and should be considered as a first-tier laboratory workflow for analysis.

Keywords: Consanguinity, filtration steps, variant interpretation, whole-exome sequencing

Introduction

Interpreting whole-exome sequencing (WES) data are challenging, with an estimated six hundred thousand rare or novel variants per person spread across the genome, and many variants being nonsynonymous at conversed loci. Furthermore, a high number of these variants are predicted to be deleterious,[1] with high inaccuracy in genome variation annotation databases,[2] altogether creates significant challenges for testing laboratories when implementing and interpreting WES data.

WES is a method to sequences and analyzes the whole coding region of the genome. The reported diagnostic yield of WES ranges from 15% to 20%.[3,4] However, in consanguineous populations, this yield could reach up to 49%,[5,6] a higher detection rate that could simply be related to the simpler interpretation of variants at the genotype level.[5]

Here, we review the implementation of custom filtration steps during analysis of raw data from WES samples of a consanguineous population. Using this method, we narrowed down the number of candidate variants requiring classification and thus the time required to complete the analysis of a single case.

Methods

Molecular sequencing was done at a commercial CAP-accredited laboratory. Raw data, including fastq, BAM, and variant call format (vcf) files, were provided for the analysis. We applied custom-designed filters to assess and analyze vcf files. The design comprises two parts. The first, automated using a bioinformatics pipeline, uses three steps to create a processed vcf file as follows: (1) Alignment, (2) variant calling, and (3) variant classification, these steps could vary based on the sequencing systems and the type of the capturing kits [Figure 1]. The initial filters looked for sequencing quality controls and population allele frequency either from open databases or a local database. The second step involves manual filtration process using commercial VarSeq software from GoldenHelix (http://www.goldenhelix.com/). Using this software, we built custom steps and filter chains [Figure 1], and annotating all identified variants across ClinVar, any previously reported variant in ClinVar was filtered out for further evaluation. However, for the purpose of testing our filtration steps, in these work variants in ClinVar were considered at later stage and after completing the filtration process. Subsequent step involves filtration based on the mode of inheritance as follows:

Autosomal recessive (AR), autosomal dominant (AD), or x-linked (XL). Next, we considered the allele state either homozygous, heterozygous, or hemizygous in cases of XL inheritance. We did not consider other modes of inheritance in this study such as mitochondrial or digenic. Furthermore, as a limitation of WES, we could not assess variants in noncoding regions of the genome.

Finally, we looked for the impact of the variant at the transcript level. During this assessment, we evaluated the most severe impacts (loss-of-function [LOF], missense, silent, and intronic). We classified variants into pathogenic/likely pathogenic based on the American College of Medical Genetics and Genomics (ACMGG) guidelines.[7] Variant assessments included clinical information with physical examination,

Table 1: Average number of identified variants in processed vcf files with hit rates for the whole cohort and separated by mode of inheritance and allele state

Inheritance patterns and zygosity	Number of identified variants per case	Diagnostic rate among positive cases only	Diagnostic rate among 100 solo WES cases
All cohort (100 cases)	65,000–75,000	Not applicable	45%
AR and homozygous	12–25	82%	37%
AD and homozygous/heterozygous	130–150	11%	5%
XL or compound heterozygous	50–70	7%	3%

vcf: Variant call format, AR: Autosomal recessive, AD: Autosomal dominant, XL: X-linked

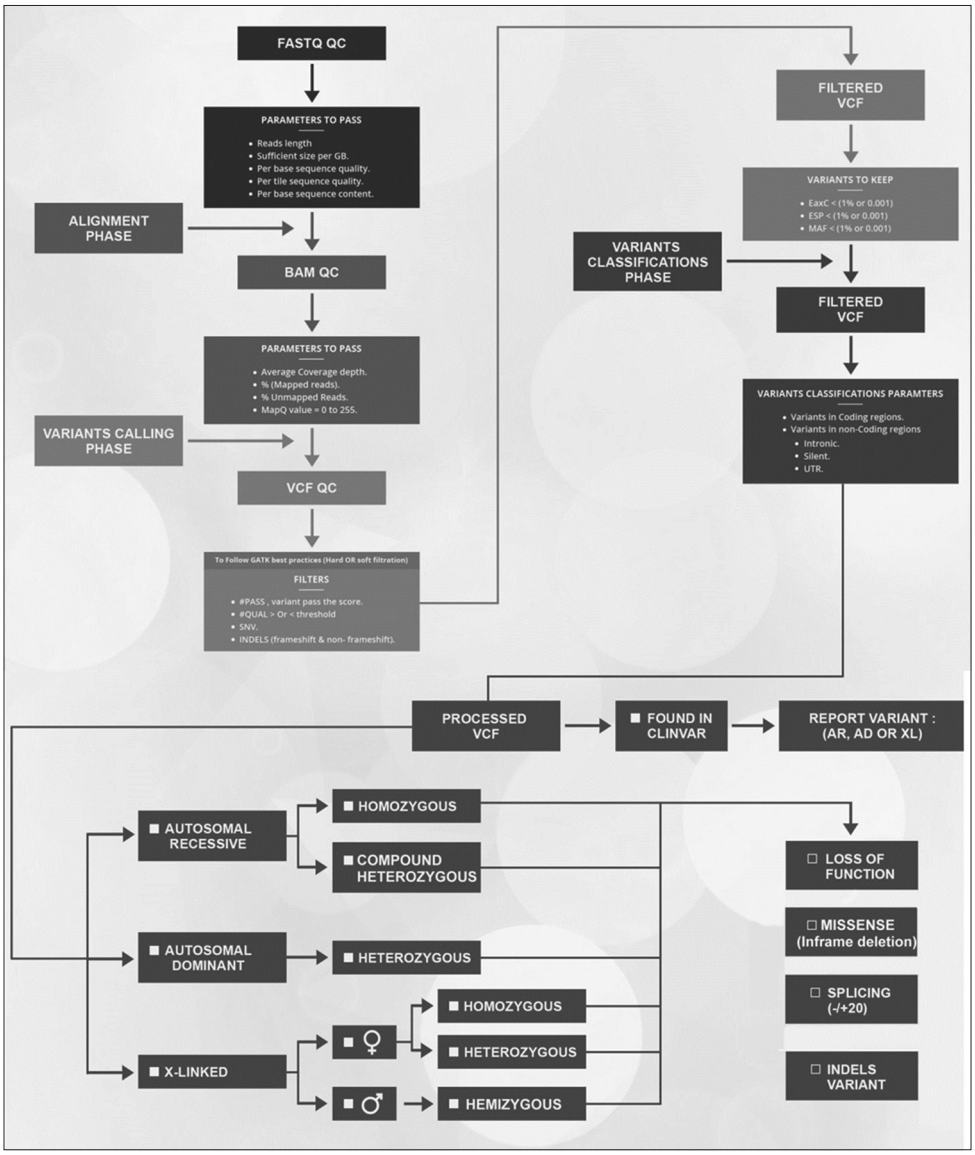

Figure 1: The automated bioinformatics pipeline workflow to generate processed variant call format files followed by manual filtration

Table 2: List of all identified diseases causing variants in processed vcf files from 100 consanguineous samples underwent whole exome sequencing

Case	Gender	Clinical Indication	Diagnosis	OMIM	g.DNA (GRCh37)	Gene, Transcript ID, c.DNA, and amino acid	Frequency in internal database 1500 samples	Frequency in public Databases	ClinVar (RCV)	Predilection tools	References PMID
Homozygous variants in AR disorders											
1	F	Spastic diplegia, developmental delay, subglottic stenosis and laryngomalacia, and hyperpigmentations on extremities	Krabbe disease	245200	Chr14:g.88452941T>C	*GALC,* NM_000153.3, c.334A>G, p.Thr112Ala	Absent	ExAC: ALL:G=0.25%	NA	AG: C0 SIFT D M: D	8687180
2	M	Failure to thrive, developmental and speech delay, and celiac disease	Mental Retardation Autosomal recessive36	602014	Chr19:g.1912476G>A	*ADAT3,* NM_138422.3, c.430G>A, p.Val144Met	0.0002	NA	000162122.2 000254727.1	AG: C0 SIFT: D.	23620220
3	M	Muscular hypotonia, strabismus, brachycephalic triangle long face, pointed chin, long philtrum, thin upper lips, and epicanthal folds with squint and rocker bottom feet	Mental Retardation Autosomal recessive36	615286	Chr19:g.1912476G>A	*ADAT3,* NM_138422.3, c.430G>A, p.Val144Met	0.0002	NA	000162122.2 000254727.1	AG:C0 SIFT: D.	23620220
4	F	Developmental delay, abnormal behavior, shy, elongated face, skin texture and folds, joint hyperlaxity, and webbing neck	Keutel syndrome	245150	Chr12:g.15037146C>T	*MGP,* NM_001190839.2, c.169+1G>A	Absent	NA	00015418.22	NA	15810001
5	F	Encephalocele, ventricular septal defect, and enlarged kidneys.	CC2D2A-related phenotype	612285	Chr4:g.15565047del	*CC2D2A,* NM_001080522.2, c.3084del, p.lys1029Argfs*3	Absent	ExAc:0.00084%	000294687.1	NA	19466712
6	M	Mental retardation	Warburg microsyndrome1	600118	Chr2:g.135887600C>T	*RAB3GAP1,* NM_001172435.1, c.1009C>T, p.Arg337*	Absent	ExAc:T=0.0012%	000171306.1	NA	26421802
7	M	Spastic paraplegia	Spastic paraplegia type 56	615030	Chr4:g.108866582A>T	*CYP2U1,* NM_183075.2, c.947A>T, p.Asp316Val	Absent	NA	000162185.1 000162185.1 000162142.1	P: D AG :C65 SIFT: D	23176821

(*Contd...*)

Table 2: (*Continued*)

Case	Gender	Clinical Indication	Diagnosis	OMIM	g.DNA (GRCh37)	Gene, Transcript ID, c.DNA, and amino acid	Frequency in internal database 1500 samples	Frequency in public Databases	ClinVar (RCV)	Predilection tools	References PMID
Homozygous variants in AR disorders											
8	M	Multiple joint flexion deformity	CACP Syndrome	208250	Chr1:g.186277990 _186277991del	*PRG4,* NM_005807.4, c.3139_3140del, p.Lys1047Aspfs*33	Absent	NA	NA	NA	16429407
9	M	Dilated cardiomyopathy and renal dysfunction	Nephrotic Syndrome type3	610725	Chr10:g.96006340C>T	*PLCE1,* NM_016341.3, c.3058C>T, p.Gln1020*	0.007	NA	NA	NA	23595123
10	F	High-arched palate and inguinal hernia	Homocystinuria	236200	Chr21:g.44482421C>T	*CBS,* NM_001178008.1, c.1039G>A, p.Gly347Ser	Absent	NA	000197584.2 000169132.1	AG: C0 SIFT: D	12124992
11	M	Developmental delay, Speech delay, and hepatomegaly	Polycystic kidney disease	263200	Chr6:g.51889738G>A	*PKHD1,* NM_138694.3, c.4870C>T, p.Arg1624Trp	Absent	1-ExAc:T=0.016%	000479548.1 000168407.4	AG:C0 SIFT: D	11898128
12	M	Cholestatic jaundice and metabolic acidosis	Fructose intolerance	229600	Chr9:g.104190767 _104190770del	*ALDOB,* NM_000035.3, c.360_363del, p.Asn120Lysfs*32	Absent	ExAc: 0.0049%	000169213.4	NA	2339710
13	F	Developmental delay, delayed myelination of the white matter, and poor vision	Galloway–Mowat syndrome	251300	Chr15:g.85191768C>T	*WDR73,* NM_032856.3, c.287G>A, p.Arg96Lys	Absent	NA	000190493.2	AG: C55 SIFT: D M: D	26123727 27001912
14	M	Seizures, choking episodes	Multiple mitochondrial dysfuntion syndrome type 4	616370	Chr14:g.74961032G>A	*ISCA2,* NM_194279.3, c.229G>A, p.Gly77Ser	Absent	NA	000255374.1 000310400.1 000170534.3 000162184.1	AG: C55 SIFT: D	25558065
15	M	Hypotonia and developmental delay	Biotin-responsive basal ganglia disease	607483	Chr2:g.228552932T>C	*SLC19A3,* NM_025243.3, c.1264A>G, p.Thr422Ala	0.005	NA	000004825.2 000489300.1	AG: C0 M: D	607483
16	F	Developmental delay, hypotonia, and seizure disorder	Early infantile epileptic encephalopathy	615833	Chr12:g.8242578C>T	*NECAP1,* NM_015509.3, c.142C>T, p.R48*	Absent	NA	000119841.3	NA	615833
17	M	Developmental delay Cerebellar atrophy	Spinocerebellar ataxia type 17	616127	Chr10:g.102013196dup	*CWF19L1,* NM_018294.5, c.605dup, p.Tyr202*	Absent	NA	NA	NA	25361784

(*Contd...*)

Table 2: (*Continued*)

Case	Gender	Clinical Indication	Diagnosis	OMIM	g.DNA (GRCh37)	Gene, Transcript ID, c.DNA, and amino acid	Frequency in internal database 1500 samples	Frequency in public Databases	ClinVar (RCV)	Predilection tools	References PMID
Homozygous variants in AR disorders											
18	F	Global Developmental delay, hypotonia, and dysmorphic child	Confirmed to have Joubert syndrome type 1	213300	Chr9:g.139327014C>T	*INPP5E,* NM_019892.4, c.1304G>A, p.Arg435Gln	Absent	NA	000022404.3 000201569.1	AG: C0 SIFT: D M: D	19668216
19	M	Neurodegenerative disease, seizure disorder	Sandhoff disease	268800	Chr5:g.74012501 _74012508del	*HEXB,* NM_000521.3, c.1169+3_1169+10del	Absent	NA	000079055.4	NA	21567908
20	M	Hepatomegaly, high liver enzymes, high lactic acid, high ammonia, and direct hyperbilirubinemia	Glycogen storage disease type 1a	232200	Chr17:g.41055964C>T	*G6PC,* NM_000151.3, c.247C>T, p.Arg83Cys	Absent	ExAc:T=0.058%, ESP:: T=0.08% -	000424594.1 000012778.5 000360229.1	AG: C0 SIFT: D M: D	8211187
21	M	Hypotonia and developmental delay	Congenital Muscular Dystrophy- type A1	236670	Chr9:g.134398428 _134398429del	*POMT1,* NM_007171.3, c.2179_2180del, p.Ser727Alafs*3	0.001	NA	000150016.2 000003408.4	NA	17878207
22	M	Poor vision and dysmorphic features	Leber congential amaurosis-6	613826	Chr14:g.21780621del	*RPGRIP1,* NM_020366.3, c.1107del, p.Glu370Asnfs*5	Absent	NA	000171128.2	NA	11283794
23	F	Developmental delay and deafness	Usher syndrome type 1B	276900	Chr11:g.76867138G>A	*MYO7A,* NM_000260.3, c.470+1G>A	Absent	NA	000154316.1	NA	9382091
24	F	Failure to thrive, generalized lipodystrophy, and hypertriglyceridemia	Congenital Generalized lipodystrophy type1	608594	Chr9:g.139571570del	*AGPAT2,* NM_006412.3, c.335del, p.Pro112Argfs*39	Absent	NA	NA	NA	11967537
25	F	Developmental delay, dysmorphic facial features, pulmonary hypertension, and right solitary kidney	Mosaic Variegated aneuploidy Syndrome 2	614114	Chr11:g.95561146C>A	*CEP57,* NM_014679.4, c.1082C>A, p.Ser361*	Absent	NA	NA	NA	NA
26	F	Jaundice and conjugated hyperbilirubinemia	Dubin–Johnson Syndrome	237500	Chr10:g.101578967 _101578968del	*ABCC2,* NM_000392.3, c.2561_2562del, p.Glu854Valfs*3	Absent	NA	NA	NA	NA

(*Contd...*)

Table 2: (*Continued*)

Case	Gender	Clinical Indication	Diagnosis	OMIM	g.DNA (GRCh37)	Gene, Transcript ID, c.DNA, and amino acid	Frequency in internal database 1500 samples	Frequency in public Databases	ClinVar (RCV)	Predilection tools	References PMID
Homozygous variants in AR disorders											
27	M	Developmental regression, myoclonic epilepsy, and basal ganglia calcification in CT scan of the brain	Aicardi–Goutieres Syndrome type2	610181	Chr13:g.51509055A>G	*RNASEH2B,* NM_024570.3, c.356A>G, p.Asp119Gly	Absent	NA	000492016.1 000171218.1	AG: C15 SIFT: D M: D	27435318
28	F	Hepatosplenomegaly, developmental delay, and cherry red spot	Niemann–Pick disease	257200	Chr11:g.6415215C>T	*SMPD1,* NM_000543.4, c.1430C>T, p.Pro477Leu	Absent	ExAc: T=0.0024%	000169478.1	P: D SIFT: D M: D	12369017
29	M	Cholestatic jaundice and elevated liver enzymes	PFIC type 4	615878	Chr9:g.71840267C>T	*TJP2,* NM_001170416.1, c.1093C>T, p.Arg365*	Absent	NA	NA	NA	NA
30	M	Painless injuries, eczema, and generalized abnormal ulcers	Hereditary sensory neuropathy type IID	243000	Chr2:g.167099103C>T	*SCN9A,* NM_002977.3, c.3503G>A, p.Trp1168*	Absent	NA	NA	NA	NA
31	F	Cholestatic jaundice	Bile Acid Synthesis defect, Congenital 1	607765	Chr16:g.30998325del	*HSD3B7,* NM_025193.3, c.694+2del, NA	Absent	NA	000171481.1	NA	NA
32	F	Intellectual disability, spasticity, and neurogenic bladder	Spastic Paraplegia and Psychomotor retardation	616756	Chr6:g.105219824G>A	*HACE1,* NM_020771.3, c.1990C>T, p.Arg664*	Absent	ExAc: T=0.0012%	NA	NA	NA
33	M	Encephalopathy, hypotonia, dysmorphic features, and basal ganglia lesions on MRI	Thiamine Metabolism Dysfunction Syndrome Type 2	607483	Chr2:g.228564240dup	*SLC19A3,* NM_025243.3, c.191dupT, p,Val65Glyfs*160	Absent	NA	NA	NA	NA
34	M	Developmental delay, speech delay, decreased vision, and light sensitivity	Achromatopsia	616517	Chr1:g.161761940del	*ATF6,* NM_007348.3, c.511del, p.Ile171Phefs*3	Absent	NA	NA	NA	NA
35	M	Developmental delay, hypotonia, and recurrent metabolic acidosis	French Canadian Type of Leigh Syndrome	220111	Chr2:g.44201018A>C	*LRPPRC,* NM_133259.3, c.1177T>G, p.Tyr393Asp	Absent	NA	000200464.2	AG: C0 SIFT: D M: D	NA

(*Contd...*)

Table 2: (*Continued*)

Case	Gender	Clinical Indication	Diagnosis	OMIM	g.DNA (GRCh37)	Gene, Transcript ID, c.DNA, and amino acid	Frequency in internal database 1500 samples	Frequency in public Databases	ClinVar (RCV)	Predilection tools	References PMID
Homozygous variants in AR disorders											
36	F	Failure to thrive and pancreatic insufficiency	Diarrhea 5 with Tufting Enteropathy Congenital	613217	Chr2:g.47604160dup	*EPCAM,* NM_002354.2, c.499dup, p.Gln167Profs*21	0.016	NA	00013612.25	NA	19820410
37	F	Sideroblastic anemia and splenomegaly	Pyridoxine-refractory sideroblastic anemia 2	205950	Chr3:g.39431966 _39431967del	*SLC25A38,* NM_017875.2, c.244-245del, p.Leu83Phefs*69	Absent	NA	NA	NA	NA
Heterozygous variants in AD disorders											
1	M	Developmental delay, short stature, macrodontia, triangular face, and large and prominent ears	KBG Syndrome	148050	Chr16:g.89351578G>A	*ANKRD11,* NM_013275, c.1372C>T, p.Arg458*	Absent	NA	NA	Na	21782149
2	F	Large forehead, depressed nasal bridge, bulbous nose, left pulmonary artery stenosis, and speech delay	Noonan Syndrome	163950	Chr12:g.112926887G>A	*PTPN11,* NM_001330437.1, c.1519G>A, p.Gly507Arg	Absent	NA	NA	P: D. AG: C65 SIFT: D	16358218
3	F	Developmental delay and dysmorphic features	Sotos syndrome	117550	Chr5:g.176636892C>T	*NSD1,* NM_022455.4, c.1492C>T, p.Arg498*	Absent	NA	NA	NA	12464997
4	M	Seizures	Benign familial neonatal seizures	121200	Chr20:g.62046439G>A	*KCNQ2,* NM_172107.2, c.1342C>T, p.Arg448*	Absent	NA	000463586.1 000187902.1 000020970.3		11690625
5	M	Short stature, coxa vara, lower tibial bowing, and cupping metaphyseal at proximal phalanges	Metaphyseal Chondrodysplasia, Schmid Type	156500	Chr6:g.116441508A>C	*COL10A1,* NM_000493.3, c.1771T>G, p.Cys591Gly	Absent	NA	NA	AG: C0 SIFT: D M: D	NA

(*Contd...*)

Table 2: (*Continued*)

Case	Gender	Clinical Indication	Diagnosis	OMIM	g.DNA (GRCh37)	Gene, Transcript ID, c.DNA, and amino acid	Frequency in internal database 1500 samples	Frequency in public Databases	ClinVar (RCV)	Predilection tools	References PMID
Hemizygous / heterozygous variants in XL disorders											
1	M	Developmental delay, cardiomyopathy, seizure disorder, elevated liver enzymes, and CPK	Danon disease	300257	ChrX:g.119581768A>C	*LAMP2*, NM_001122606.1, c.669T>G, p.Tyr223*	Absent	NA	NA	NA	NA
2	F	Developmental delay, hypotonia, clinodactyly, downturned corners of the mouth, triangular face, and long philtrum	Cornelia de Lange Syndrome 5	300882	ChrX:g.71684460A>G	*HDAC8*, NM_018468.2, c.859T>C, p.Cys287Arg	Absent	NA	NA	AG: C0 SIFT: D M: D	NA
3	F	Developmental delay, hypotonia, and seizures	MECP2-related disorders	300005	ChrX:g.153297719G>A	*MECP2*, NM_001110792.1, c.352C>T, p.Arg118Trp	Absent	NA	000255874.1 000125852.24	P: D, SIFT: D. M: D	10508514

M: Male, F: Female, AG: Align GVGD (http://AG.iarc.fr/), M: mutation taster: (http://www.mutationtaster.org/), P: PolyPhen2 (http://genetics.bwh.harvard.edu/pph2/), D: Diseases causing, deleterious or damaging. vcf: Variant call format, AR: Autosomal recessive, AD: Autosomal dominant, XL: X-linked

further laboratory testing or imaging, segregation analysis, genotype and phenotype correlation, previous publication, or *de novo* assessment of the variant. Results were considered "positive" (meaning a likely disease-causing variant has been identified based on the ACMGG guidelines). As a proof of concept, we used 100 vcf files from consanguineous cases that were subjected to solo WES study.

Results

After applying the filtration steps [Figure 1], we could narrow the candidate variants to around only 5–10 LOF variants and 10–15 missense variants found in the homozygous state and genes with the AR mode of inheritance.

Many of these variants can be eliminated easily just by looking at the phenotype; in most cases, the gene's clinical phenotype reported in the literature and the patient phenotype have no relation. Further variants can be eliminated as having high allele frequencies in our population or as our laboratory has observed previously, as being homozygous in nonaffected individuals. For each case, the time required to assess the remaining 5–10 variants (LOF and missense) averages 45 min, with 82% probability of identifying disease-causing variants using this filter chain. For cases with unidentified variants after the first filter, we proceeded to look for the next possible filter with the second-highest hit rate. For AD disorders and variants in the heterozygous or homozygous state, the number of identified variants were higher (130–150 variants), both missense and LOF. On average, 50–70 variants were identified in the other possible modes of inheritance in different scenarios such as compound heterozygous or XL. However, the diagnostic yield is low for these variants, accounting for around 7% of all positive cases [Table 1]. Our overall hit rate for the full sample of 100 cases with WES is 45% [Table 2] similar to the existing literature.[5]

Discussion

WES has become a valuable tool in clinical settings for obtaining molecular diagnoses. Designing methods and tools that can facilitate the diagnostic accuracy of WES will certainly facilitate better and improved healthcare by identifying the molecular defects underlying rare disorders. Consanguinity impacts disorder incidence since deleterious and disease allele variations are known to occur as a result of long runs of homozygosity[8] or missense substitutions in a homozygous state.[9] In general, consanguineous marriages are expected to result in a high incidence of AR genetic disorders. The high rate of consanguinity in Saudi Arabia leads to possible founder effects for many genetic disorders, and population-specific AR genetic disorders.[10,11] It is critical to design a custom workflow focusing on the target population, starting from the bioinformatics pipeline, and proceeding to variant analysis and classification. For example, in our population, extensive effort

during pipeline and workflow design focused on homozygous variants, which present higher chances of identifying disease-causing variants due to the high rates of consanguinity. Similar approaches have been applied before looking at autozygosity regions in the genomes.[12] However, with advances in technology, we can achieve better resolution and examining the variant level. Furthermore, a consanguineous population has fewer AD variants requiring less attention.

Furthermore, while specific populations already have custom databases, custom bioinformatics, and filtration steps for populations may enable better and faster interpretation of the results. By applying our custom filters to identify only homozygous variants in AR disorders, we could substantially narrow the number of candidate variants while still achieving a high hit rate toward 82% with an identifiable, disease-causing variant (positive cases) and around 36% of the whole cohort. Given the manageable number of variants requiring additional analysis, we achieved this in around 45 min, compared to 5 h without the filtration, and our hits account for the large percentage of positive cases.

In cases with different modes of inheritance (AD, XL), the number of identified variants is still high and would still require additional time to complete the analysis. This is unsurprising since consanguinity has little-to-no impact on these disorders' underlying genotype. However, given the high rate of consanguinity, analysis of these variants should follow the full and complete analysis of AR disorder variants.

In conclusion, WES is a very useful tool to identify disease-causing variants, particularly in a consanguineous population, where higher detection rates are achieved. In this report, we verified that custom filtration steps and analysis to look primarily for homozygous variants in AR disorders will achieve the higher possible detection rates in less time, and testing laboratories are encouraged to consider this process for the first-tier analysis of WES raw data.

References

1. 1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, *et al.* A global reference for human genetic variation. Nature 2015;526:68-74.
2. Yen JL, Garcia S, Montana A, Harris J, Chervitz S, Morra M, *et al.* A variant by any name: Quantifying annotation discordance across tools and clinical databases. Genome Med 2017;9:7.
3. Lee H, Deignan JL, Dorrani N, Strom SP, Kantarci S, Quintero-Rivera F, *et al.* Clinical exome sequencing for genetic identification of rare mendelian disorders. JAMA 2014;312:1880-7.
4. Wortmann SB, Koolen DA, Smeitink JA, van den Heuvel L, Rodenburg RJ. Whole exome sequencing of suspected mitochondrial patients in clinical practice. J Inherit Metab Dis 2015;38:437-43.
5. Alfares A, Alfadhel M, Wani T, Alsahli S, Alluhaydan I, Al Mutairi F, *et al*. A multicenter clinical exome study in unselected cohorts from a consanguineous population of Saudi Arabia demonstrated a high diagnostic yield. Mol Genet Metab 2017;121:91-5.
6. Monies D, Abouelhoda M, AlSayed M, Alhassnan Z, Alotaibi M, Kayyali H, *et al.* The landscape of genetic diseases in Saudi Arabia based on the first 1000 diagnostic panels and exomes. Hum Genet 2017;136:921-39.
7. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, *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;17:405-24.
8. Szpiech ZA, Xu J, Pemberton TJ, Peng W, Zöllner S, Rosenberg NA, *et al.* Long runs of homozygosity are enriched for deleterious variation. Am J Hum Genet 2013;93:90-102.
9. Xue Y, Chen Y, Ayub Q, Huang N, Ball EV, Mort M, *et al.* Deleterious- and disease-allele prevalence in healthy individuals: Insights from current predictions, mutation databases, and population-scale resequencing. Am J Hum Genet 2012;91:1022-32.
10. al-Abdulkareem AA, Ballal SG. Consanguineous marriage in an urban area of Saudi Arabia: Rates and adverse health effects on the offspring. J Community Health 1998;23:75-83.
11. Al-Owain M, Al-Zaidan H, Al-Hassnan Z. Map of autosomal recessive genetic disorders in Saudi Arabia: Concepts and future directions. Am J Med Genet A 2012;158A:2629-40.
12. Alkuraya FS. Impact of new genomic tools on the practice of clinical genetics in consanguineous populations: The Saudi experience. Clin Genet 2013;84:203-8.