Whole-exome sequencing of a Saudi epilepsy cohort reveals association signals in known and potentially novel loci

Abdulrahman H. Al Anazi1, Ahmed S. Ammar1, Mahmoud Al-Hajj2, Cyril Cyrus3, Danah Aljaafari4, Iname Khoda4, Ahmed K. Abdelfatah1, Abdulllah A. Alsuilaman4, Firas Alanazi1, Rawan Alanazi1, Divya Gandla5, Hetal Lad5, Samar Barayan1, Brendan J. Keating5 and Amein K. Al-Ali3*

Abstract

Background: Epilepsy, a serious chronic neurological condition effecting up to 100 million people globally, has clear genetic underpinnings including common and rare variants. In Saudi Arabia, the prevalence of epilepsy is high and caused mainly by perinatal and genetic factors. No whole-exome sequencing (WES) studies have been performed to date in Saudi Arabian epilepsy cohorts. This offers a unique opportunity for the discovery of rare genetic variants impacting this disease as there is a high rate of consanguinity among large tribal pedigrees.

Results: We performed WES on 144 individuals diagnosed with epilepsy, to interrogate known epilepsy-related genes for known and functional novel variants. We also used an American College of Medical Genetics (ACMG) guideline-based variant prioritization approach in an attempt to discover putative causative variants. We identified 32 potentially causative pathogenic variants across 30 different genes in 44/144 (30%) of these Saudi epilepsy individuals. We also identified 232 variants of unknown significance (VUS) across 101 different genes in 133/144 (92%) subjects. Strong enrichment of variants of likely pathogenicity was observed in previously described epilepsy-associated loci, and a number of putative pathogenic variants in novel loci are also observed.

Conclusion: Several putative pathogenic variants in known epilepsy-related loci were identified for the first time in our population, in addition to several potential new loci which may be prioritized for further investigation.

Keywords: Epilepsy, Neurological conditions, WES, Saudi Arabia, Variants

Background

Epilepsy is a group of chronic neurological diseases characterized by unprovoked seizures due to abnormal neuronal firing in the brain. It is one of the most frequently encountered medical problems in neurology clinics. Epilepsy affects up to 1 in 26 individuals in the USA and approximately 5–7 per 1000 individuals worldwide with reports indicating that at least 50 to 100 million people in the world suffer from this disease, with approximately 85% of these cases living in developing countries [1–3]. However, 20–30% of all epilepsy cases are due to acquired factors, such as infection, stroke, trauma, neoplasms, and autoimmunity, while the remaining cases are thought to be due to genetic factors [4]. Epidemiological studies have observed an increased risk of epilepsy development in the relatives of epileptic individuals [5].

Genome-wide association studies (GWAS) of epilepsy phenotypes using genome-wide genotyping (GWG) arrays have discovered many single-nucleotide polymorphisms (SNPs) associated with the disease. A
meta-analysis conducted by the International League Against Epilepsy in 2014 identified SNPs associated with the development of epilepsy in the sodium voltage-gated channel alpha subunit 1 (SCN1A), Protocadherin 7 (PCDH7), FA Complementation Group L (FANCL) and Vaccinia Related Kinase 2 (VRK2) genes [6]. A subsequent large GWAS encompassing over 15,200 epilepsy subjects and 29,600 controls revealed 16 genome-wide significant signals reaching statistical significance. Genes in these loci have diverse roles in histone modification, ion-channels, pyridoxine metabolism, synaptic transmission as well as transcription factors. Interestingly, functional annotation of almost 500 SNPs, that were observed to be significant at a genome-wide statistical threshold, found that most are in intronic (46%) or intergenic (29%) regions, with only 4 non-synonymous SNP associations observed, of which 2 were missense variants [7]. Furthermore, approximately half of the SNPs were implicated to impact gene transcription from gene regulation database analyses [7].

Second-generation sequencing technologies including whole-exome sequencing (WES) have been used in the diagnostic or research settings to identify genetic mutation(s) which may cause highly variable phenotypic expression in epileptic subjects [8–10]. The very nature of the WES approach favors detection of rarer highly penetrant gene-coding variants which are often hard to find in GWG-based GWAS approaches.

Performing WES in Saudi Arabian epilepsy populations offer a unique opportunity for the discovery of rare genetic variants impacting this disease as there is a high rate of consanguinity among large tribal pedigrees. A study in the Eastern province of Saudi Arabia has reported the prevalence of epilepsy to be around 6–7 per 1000 individuals [11]. In the present study, WES was performed on 144 individuals diagnosed with epilepsy with varying age-of-onset. An American College of Medical Genetics (ACMG) guideline-based variant prioritization approach followed WES which allowed for the discovery of potentially causative variants in our cohort of epilepsy subjects.

**Materials and methods**

**Patient sampling and ethical approval**

Over a period spanning 2018–2020, samples and data from consecutive subjects with epilepsy attending the Neurosurgery Clinics, King Fahd Hospital of the University, Al-Khobar, and King Fahd Hospital, Alhafof, Saudi Arabia, were collected for inclusion in this study. Participants ranged in age from 13–51 and were clinically diagnosed with epilepsy at the point of recruitment. The phenotype data of all subjects were reviewed by a consultant committee to verify uniformity among sites and eligibility consistent with International League Against Epilepsy [12]. In this study, the diagnosis of epilepsy was made by a consultant specializing in epilepsy based on the patients’ clinical history. Moreover, cases with moderate-to-severe intellectual disability, or cancer, were excluded from the study. If there was doubt regarding a patient’s phenotypic eligibility, the individual cases were reviewed by the consultant committee and if needed, additional data were requested prior to a decision being made by the committee regarding inclusion of the patient. However, on completion of the project the medical records of all patients were reviewed again. Among all our epileptic patients included in the study, eight patients also had type 2 diabetes, five patients had hypertension, one patient had cardiovascular disease and stroke, one patient had polycystic ovarian syndrome, and two patients had muscular dystrophy. Table 1 outlines the subjects’ demographic and clinical characteristics.

Ethical approval for the study was obtained from the local Institutional Review Board (IRB) committees (IRB-2015-01-063), and the study was conducted according to the ethical principles of the Declaration of Helsinki and Good Clinical Practice guidelines. All patients included in the study signed a written informed consent.

**DNA sequencing, read alignment, variant calling and quality control**

Blood samples were collected from subjects in EDTA vacutainers and after collection were immediately stored at –80 °C. Standard DNA preparation was performed using DNeasy Blood kits (Qiagen, MD, USA). Whole-exome sequencing libraries were generated using the Agilent SureSelect Human All Exon Kit V5 (Agilent, CA, USA) and sequenced on a HiSeq 2500 instrument (Illumina, CA, USA) using standard paired-end sequencing protocol. Raw sequencing reads were stored as FASTQ files and then aligned to the human reference genome (GRCh37) using Illumina’s Dynamic Read Analysis for

| Parameter                        | Value     |
|----------------------------------|-----------|
| Age (mean ± SD) (years)          | 28.30 ± 11.15 |
| Gender (male %)                  | 50.5      |
| Age of initial diagnosis (years) | 13–51     |
| Duration of the disease (years)  | 13.46 ± 8.70 |
| Percentage of subjects with affected family members (%) | 24.5   |
| Management of disease            |           |
| Monotherapy (%)                  | 57.5%     |
| Poly-therapy (%)                 | 42.5%     |
GENomics (DRAGEN) Pipeline. Resultant BAM files were position-sorted and duplicate reads marked. Single-sample gVCF files were generated by the DRAGEN Germline Pipeline, and joint calling of all samples in the study cohort were performed by DRAGEN Joint Genotyping.

Principal components analysis (PCA) and Kinship
KING was used for relatedness inference based on the genotype of exome SNPs (MAF > 0.01) [13]. Estimated kinship coefficient and number of SNPs with zero shared alleles (IBS0) between a pair of individuals were plotted. Parent–offspring, sibling pairs, and unrelated pairs can be distinguished as separate clusters on the scatterplot. Ancestry and kinship toolkit (AKT) was used to calculate PCAs and plot the results [14].

Variant annotation, filtering and prioritization
Variants were annotated with SnpEff to predict the effects of variants [15]. Rare variants were defined as minor allele frequency (MAF) < 1% in the Genome Aggregation Database (GnomAD) [16]. Intronic, synonymous, 3’ and 5’ UTR, up- and downstream variants were identified and excluded from the analysis. The remaining rare variants were considered to be potentially deleterious variants. Genetic variants classified in ClinVar as “Likely pathogenic” or “Pathogenic,” and in Human Gene Mutation Database (HGMD) as disease-causing mutations (DM) for epilepsy or seizures were collected and curated together with research literatures to serve as the knowledgebase for variant prioritization and classification [17, 18].

Results
Principal component analysis
The common genetic variants of these Saudi epilepsy individuals show a unique cluster when compared to the world’s major populations based on principal component analysis (Fig. 1). The samples demonstrated a genetically matched background which avoided false attribution of associations due to population stratification.

Potentially pathogenic rare variants identified in 44 epilepsy subjects
Based on a combination of variant filtering, and integration of variant databases in ClinVar, and HGMD, we identified a total of 32 potentially causative pathogenic variants across 30 different genes in 44/144 (30%) epilepsy subjects as shown in Table 2. Additional file 1: Supplementary Table S1 outlines additional minor allele frequencies in additional databases and further annotation of these putative pathogenic variants.

The 30 genes harboring the likely pathogenic variants observed in these 44 Saudi epilepsy subjects were then assessed for overlap with 102 previously collated monogenic epilepsy genes [7]. Likely pathogenic variants in 12 of these 102 monogenic epilepsy genes were observed in 44 epileptic subjects: CHRNA4, CLN3, CLN8, DEPDC5, KCNJ10, KCNMA1, POLG, PRICKLE1, SCN1A, SCN2A, SCN8A and SCN9A. Of the 18 additional genes from Table 2 with likely pathogenic variants, a number including SZT2, SCN10A, UBA5 have been reported in whole-exome sequencing in epilepsy subjects [19, 20]. Only one homozygous mutation, a stop-gain, was observed in single individual [21], in Potassium Calcium-Activated Channel Subfamily M Alpha 1 (KCNMA), a calcium-sensitive potassium channel gene which has been shown to have a role in general and early-onset epilepsy-related phenotypes [22–26]. This individual, a female who was diagnosed with childhood epilepsy at 12 years old, has one family member with a diagnosis of epilepsy and was assessed to have first degree of consanguinity.

Most commonly observed likely pathogenic variants
An in-frame insertion mutation in non-imprinted in Prader–Willi/Angelman syndrome region protein 2 (NIPA2) (chr15, GRCh37 position: 23006299) was observed in seven of the study subjects (shown in Table 3 along with age of onset). A missense variant in SH2B Adaptor Protein 3 (SH2B3) (chr12, position:
Table 2 Potentially causative pathogenic variants derived from screening of 144 Saudi epilepsy subjects. Chromosomal position is outlined for the 32 putative pathogenic variants in 30 gene regions along with annotation of the putative pathogenic variant mapped to human reference genome build 37 (GRCh37). Minor allele frequencies (MAF) are shown for the: Saudi epilepsy cohort; Genome Aggregation Database (GnomAD); and Human Gene Mutation Database (HGMD). ClinVar annotation for likely clinical significance is also listed.

| Chromosome position | Patient(s) with Heterozygous or Homozygous** Variant | Variant Type | Gene (NCBI Gene ID) | Cohort Variant MAF* | ExAC MAF | GnomAD MAF | HGMD Mutation Phenotype | Clinical significance in CLINVAR database |
|---------------------|--------------------------------------------------|--------------|---------------------|---------------------|---------|-----------|------------------------|------------------------------------------|
| chr1: 43907014      | 065 Missense                                      | Szt2         | 0.003               | 0.00034             | 0.002581|           | Epileptic encephalopathy early-onset | Conflicting interpretations of pathogenicity |
| chr1: 47746675      | 094, 128, 134 Missense                            | Stil         | 0.01                | 0.00213             | 0.003538|           | Intellectual disability seizures microcephaly | Conflicting interpretations of pathogenicity |
| chr1: 97915614      | 101, 141 Splicing                                 | Dpyd         | 0.007               | 0.007067            |          |           | Dihydro-pyrimidine dehydrogenase deficiency | Conflicting interpretations of pathogenicity drug response |
| chr1: 119683231     | 070, 129, 135 Missense                            | Wars2        | 0.01                |                     | 0.006159|           | Intellectual disability | Conflicting interpretations of pathogenicity |
| chr1: 160011671     | 021, 045 Missense                                 | Kcnj10       | 0.007               |                     | 0.001179|           | Developmental delay failure to thrive ataxia hypotonia and tonic–clonic seizures | Uncertain significance |
| chr2: 166243269     | 019 Missense                                      | Scn2a        | 0.003               | 0.00073             | 0.005002|           | Epileptic encephalopathy | Conflicting interpretations of pathogenicity |
| chr2: 166848930     | 069 Missense                                      | Scn1a        | 0.003               | 0.00056             | 0.004116|           | Intractable epilepsy | Conflicting interpretations of pathogenicity |
| chr2: 167138296     | 085 Missense                                      | Scn9a        | 0.003               | 0.00189             | 0.008274|           | Febrile seizures | Conflicting interpretations of pathogenicity |
| chr3: 38739016      | 007, 039 Frame shift                              | Scn10a       | 0.007               | 0.0001              | 0.00130 |           | Refractory epilepsy & autism spectrum disorder | Uncertain significance |
| chr3: 132378559     | 009, 101 Stop gained                              | Ubas         | 0.007               |                     | 0.009118|           | Pathogenic | |
| chr4: 119736287     | 032, 141 Missense                                 | Sec24d       | 0.007               |                     | Not seen |           | Intellectual disability and epilepsy | Likely pathogenic |
| chr8: 1719594       | 071 Missense                                      | Cln8         | 0.003               | 0.00082             | 0.002221|           | Neuronal ceroid lipofuscinosis late infantile | Conflicting interpretations of pathogenicity |
| chr10: 79396648     | 017** Stop gain                                   | Kcnma1       | 0.008               |                     | Not seen |           | Epilepsy | |
| chr11: 9225637      | 047 Frame shift                                   | Dend5a       | 0.003               |                     | Not seen |           | Epileptic encephalopathy | |
| chr11: 93521218     | 087 Missense                                      | Med17        | 0.003               | 0.000115            |           |           | Seizures and hypoplasia of the corpus callosum | Uncertain significance |
### Table 2 (continued)

| Chromosome Position | Patient(s) with Heterozygous or Homozygous** Variant | Variant Type | Gene (NCBI Gene ID) | Cohort Variant MAF* | ExAC MAF | GnomAD MAF | HGMD Mutation Phenotype | Clinical significance in CLINVAR database |
|---------------------|-----------------------------------------------------|--------------|---------------------|---------------------|----------|-----------|------------------------|-----------------------------------------|
| chr12: 42863325     | 083 Missense PRIKLE1                                 | 0.003        | 1.00E−05            | 0.000641            |          |           | Likely pathogenic       |                                         |
| chr12: 52164462     | 096 Missense SCN8A                                    | 0.003        | 1.00E−05            | 0.000109            |          |           | Paroxysmal kinesigenic dyskinesia | Conflicting interpretations of pathogenicity |
| chr12: 111856571    | 021, 037, 043, 138 Missense SH2B3                     | 0.014        | 0.00099            | 0.004796            |          |           | Erythrocytosis idiopathic     | Uncertain significance                     |
| chr13: 100925464    | 048 Missense PCCA                                     | 0.003        |                      | 0.000976            |          |           | Epilepsy & neurodevelopmental delay | Conflicting interpretations of pathogenicity |
| chr15: 23006299     | 002, 009, 040, 044, 050, 057, 089 In-frame insertion NIPA2 | 0.024        |                      | 0.008426            |          |           | Childhood absence epilepsy    |                                         |
| chr15: 52632432     | 121 Missense MYOSA                                    | 0.003        | Not seen            |                      |          |           | Developmental delay seizures & dystonia | Pathogenic                               |
| chr15: 73617728     | 127 Missense HCN4                                    | 0.003        |                      | 0.000018            |          |           | Myoclonic epilepsy infantile  |                                         |
| chr15: 89866693     | 049 Missense POLG                                     | 0.003        | 0.00036            | 0.001169            |          |           | Depression ataxia and cardiomyopathy | Conflicting interpretations of pathogenicity |
| chr15: 89870429     | 030 Missense POLG                                     | 0.003        | 0.00044            | 0.002358            |          |           | Progressive external ophthalmoplegia | Conflicting interpretations of pathogenicity |
| chr15: 89876827     | 032 In-frame insertion POLG                          | 0.003        |                      | 0.004516            |          |           | Hepatic encephalopathy        | Benign/Likely benign                      |
| chr16: 150392       | 112 Missense NPRL3                                    | 0.003        | 0.000933            |                      |          |           | Focal epilepsy                | Uncertain significance                    |
| chr16: 28500627     | 104 Missense CLN3                                     | 0.003        | 4.00E−05            | 0.000203            |          |           | Retinal degeneration          | Uncertain significance                    |
| chr17: 61791402     | 084 Missense STRADA                                   | 0.003        | Not seen            |                      |          |           |  | Pathogenic                   |
| chr19: 14024452     | 071 Splicing CC2D1A                                   | 0.003        | Not seen            |                      |          |           | Autism spectrum disorder intellectual disability and seizures | Pathogenic                               |
| chr20: 61981924     | 055, 098, 106, 121 Missense CHRNA4                   | 0.014        | Not seen            |                      |          |           | Epilepsy nocturnal frontal lobe | Pathogenic                               |
| chr21: 34003387     | 141 Missense SYNU1                                    | 0.003        | 0.00082            | 0.008603            |          |           | Parkinson disease             | Likely benign                             |
| chr22: 32188751     | 019 Stop gained DEPDCS                                | 0.003        | Not seen            |                      |          |           | Epilepsy familial focal with variable foci | Pathogenic                               |

Chromosome position is based on build 37 of the human genome. ExAC = Exome Aggregation Consortium; MAF = Minor allele frequency; GnomAD = Genome Aggregation Database; HGMD = Human Gene Mutation Database

** homozygous variants
111856571) and in Cholinergic Receptor Nicotinic Alpha 4 Subunit (CHRNA4) (chr20 position: 61981924) were both observed in four individuals. Missense variants in STIL Centriolar Assembly Protein (STIL) (chr1, position: 47746675), and in Tryptophanyl-tRNA Synthetase 1 (WARS2) were both observed in three individuals.

## Variants of unknown significance identified in epilepsy 133 subjects

In highly curated genomic disease databases such as Clin-Var, there are a large number of variants of unknown significance (VUS), where there is unknown or conflicting clinical significance to date for association of such variants with epilepsy-related phenotypes. We identified 232 variants of unknown significance across 101 different genes in 133/144 (92%) subjects as shown in Additional file 2: Supplementary Table S2. Interestingly when the genes harboring these variants of unknown significance were intersected with the 102 previously collated monogenic epilepsy genes it was observed that 43 of these monogenic gene variants were enriched in the 101 loci with VUS from these 133 Saudi epilepsy subjects (Additional file 2: Table S2). We note that two individuals were observed to have homozygous potential pathogenic variants. One individual showed homozygosity for a missense variant in Spermatogenesis Associated 5 (SPATA5), with 3 individuals showing heterozygosity for this mutation (Additional file 2: Supplementary Table S2). SPATA5 has shown clear association with severe childhood epilepsy [27–32]. This female individual was diagnosed with childhood epilepsy at the age of 10 years old and two family members having a diagnosis of epilepsy, and both parents are listed as cousins.

Another homozygous mutation was observed in one individual for a missense mutation in Calcium Voltage-Gated Channel Subunit Alpha1 H (CACNA1H) with three additional individuals also carrying one copy of this mutation. Mutations in this gene have been associated with generalized and severe epilepsies [33, 34], although it has been debated whether this is a bona fide monogenic epilepsy gene [35]. This female individual was diagnosed with epilepsy at 15 years of age and does not have any other family members with a diagnosis of epilepsy, or any consanguinity noted.

## Discussion

We performed the first whole-exome sequencing study in Saudi Arabia epilepsy subjects. Using 144 individuals, we compared putative pathogenic variants as well as variants of unknown significance with population-based

### Table 3

| Mutation type, gene and position | Study subject | Age of onset | Gender | Other family members affected | Management | Response to treatment |
|---------------------------------|---------------|--------------|--------|-------------------------------|------------|-----------------------|
| In-frame insertion in NIPA2 Pathogenic p.N334_E335insD (chr15: 23006299) | 002 | 7 | M | No | Mono | Yes |
| | 009 | 16 | M | No | Mono | Yes |
| | 040 | 2 | F | No | Poly | Partial |
| | 044 | 24 | M | Yes | Mono | Yes |
| | 050 | 15 | F | No | Mono | Yes |
| | 057 | 1 | M | No | Poly | Partial |
| | 089 | 15 | M | Yes | n/a | n/a |
| Missense variant in SH2B3 Pathogenic c.622G > C p.Glu208Gln (chr12: 111856571) | 021 | n/a | M | No | Mono | Yes |
| | 037 | 8 | F | No | Poly | Yes |
| | 043 | 13 | F | No | Mono | Yes |
| | 138 | 26 | F | No | Poly | Yes |
| Missense variant in CHRNA4 Pathogenic c.839C > T p.Ser280Phe (chr20: 61981924) | 055 | 31 | F | Yes | n/a | n/a |
| | 098 | 13 | M | n/a | n/a | n/a |
| | 106 | n/a | F | n/a | n/a | n/a |
| | 121 | 5 | F | No | Mono | Yes |
| Missense variant in STIL Pathogenic c.1455G > C p.Leu485Phe (chr1: 47746675) | 094 | 23 | M | No | n/a | n/a |
| | 128 | 19 | M | No | Poly | Yes |
| | 134 | 15 | F | Yes | Mono | Yes |
| Missense variant in WARS2 Pathogenic c.37 T > G p.Trp123Gly (chr1: 119683231) | 070 | 22 | M | n/a | n/a | n/a |
| | 129 | 16 | M | No | Poly | Yes |
| | 135 | 9 | M | Yes | Poly | Yes |
whole-exome sequencing and whole genome sequencing databases.

The highest number of observed mutations across the 144 subjects were observed in NIPA2, a highly selective magnesium transporter. This in-frame insertion variant (NP_011718.1:p.(Asn334Glu335insAsp)) was observed in 7 subjects from 144 overall (5%). This variant has been previously reported within a population of subjects with childhood absence epilepsy (CAE) [36–38].

A missense variant in CHRNA4 was observed in 4 subjects. CHRNA4 is a nicotinic acetylcholine receptor, belonging to a superfamily of ligand-gated ion channels which play an established role in signal transmission at synapses. Mutations in CHRNA4 have been reported with nocturnal frontal lobe epilepsy type 1. A missense variant in SH2B3 was observed in 4 subjects. This gene is involved in a range of signaling activities by growth factor and cytokine receptors as part of the SH2B adaptor family of proteins. Mutations in this gene have been associated with susceptibility to celiac disease type 13 and susceptibility to insulin-dependent diabetes mellitus. It has low expression in the brain however as evident in the Genotype-Tissue Expression (GTEx) database. Missense mutations in STIL were observed in 3 subjects. STIL is a cytoplasmic protein which plays a role in the regulation of the mitotic checkpoint machinery. It too has low expression levels in all GTEx brain tissues.

There are a number of prioritized signals observed in this Saudi epilepsy study that may be novel or have very limited reports of association. Deficiency of WARS2 was observed in a patient with severe infantile-onset leukencephalopathy, profound intellectual disability, spastic quadriplegia, epilepsy and microcephaly [39]. Rare mutations in DPYD have been implicated in children with unspcific neurological symptoms [40]. Epileptic encephalopathy caused by recessive loss-of-function (LoF) mutations have been reported in DENND5A [41]. A previous report of infantile cerebral and cerebellar atrophy showed association with a mutation in MED17 [42]. A LoF mutation in HCN4 has been reported to be associated with Familial benign myoclonic epilepsy in infancy [43]. Mutations in STRADA, SYNJ1, CACNA1A and NPRL3 have also been reported with severe epilepsy-related disease, but the association has not been reported for heterozygous variants in isolated forms of epilepsy [44–51].

A number of prioritized signals of putative pathogenicity were observed that have no reports of association with epilepsy in the literature including SEC24D, PCCA, MYOSA which may be strong candidates for further functional studies. SEC24D was reported to play a role in vesicle trafficking and mutations in this gene are associated with Cole-Carpenter syndrome, a disorder affecting bone formation [52]. PCCA codes for the alpha subunit of the mitochondrial enzyme Propionyl-CoA carboxylase, and mutations in this gene leads to enzyme deficiency and are associated with propionic acidemia [53]. MYOSA encodes myosin 5A, and mutations in this gene are associated with Griscelli syndrome, which is characterized by hypopigmentation and a primary neurological abnormality [54]. These aforementioned genes may be good candidates for further functional studies.

This study is limited in that incomplete pedigrees and depth of sub-phenotyping are available, although putative pathogenic variants many known epilepsy-related loci are evident, and a number of potential new loci may be prioritized for further investigation. The Saudi population offers a lot of promise for elucidation of the genetic etiology of common diseases such as epilepsy due to consanguinity, with extended homozygosity stretches often observed over several megabases, affording the opportunity to enrich for recessive forms of epilepsy. While this study looked at 144 subjects, we are aiming to expand the cohort to encompass genetic analyses of extended pedigrees with additional phenotyping.

Conclusions
We identified for the first time 32 potentially causative pathogenic variants in Saudi individuals with epilepsy. In addition, several potential new loci have been identified that have no reports of association with epilepsy in other populations. These potentially causative pathogenic variants in these new loci may be prioritized for further functional studies.

Supplementary Information
The online version contains supplementary material available at https://doi.org/10.1186/s40246-022-00444-6.

Additional file 1. Supplementary Table S1: outlines additional minor allele frequencies in additional databases and further annotation of these putative pathogenic variants.

Additional file 2. Supplementary Table S2: identified 232 variants of unknown significance across 101 different genes in 133/144 (92%) subjects.

Acknowledgements
The authors would like to acknowledge the financial supported extended by King Abdulaziz City for Science and Technology, Riyadh, Saudi Arabia (Grant Number 14-MED81-46). We are also grateful to the nurses and technical staff for their work and dedication.

Author contributions
AHA, ASA, MA, DA, IK, AKA, AAA, FA, RA and SB were involved in the design of the work, critically revising the protocol, patient recruitment, analysis, interpretation of data and drafting of the manuscript. CC, DG, HL, BJK and AKA were involved in the design of the work, laboratory work, analysis, interpretation of data, critically revising the protocol and drafting the manuscript. All authors read and approved the final manuscript.
Funding
This work is supported by King Abdulaziz City for Science and Technology, Riyadh, Saudi Arabia, Grant Number 14-MED81-46. This is a government agency. They have no role in the design, collection, analysis, and interpretation of the data and writing of the manuscript.

Availability of data and materials
The datasets generated during the current study are available in the European Nucleotide Archive (ENA) repository, https://eurof2softenkins.protection.outlook.com/#browse?fid=3%2Fwww.ebi.ac.uk%2Fena%2Fbrowser%3Fview=2 2FPRIEB57370&data=0%7C20%7C%7Caalal%40a.ru.edu%7C7f185a%40e6e%40 4cfe9d880582dcb30%00%7C7C26ebbb68d80441f8a3a942b75e060%7C0%7C 7C6380456872872225%7C7CUnknown%7C?TFwpbGb2z3dBeysWIjoMhMC 4wLjAwMDA1LJQiloviLm2zLCJBTi6ik1haWVlLCJXCl6mN0%3D%7C2000% 7C%7C%7C#data=497%7C6OmQ6HTM69EHDODReNEemdK7Eeb2Qr5% 2FXW%2B4v%3Dserialized=0, under the title “Whole-Exome Sequencing of a Saudi Epilepsy Cohort Reveals Association Signals in Known and Potentially Novel Loci” with accession number PRJEB57558. All requests for data can be sent to the corresponding author (AKA), and verified academic investigators will be granted full access.

Declarations

Ethics approval and consent to participate
The study received ethical approval from the Imam Abdulrahman bin Faisal University Ethical Committee (Reference: IRB-2015-01-065). The study was conducted according to the ethical principles of the Declaration of Helsinki and Good Clinical Practice guidelines. A signed informed consent form was obtained from each participating patient and control by the treating physicians prior to inclusion in the study.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interest.

Author details
1. Department of Neurosurgery, King Fahd Hospital of the University, Alkhobar, College of Medicine, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia.
2. Department of Neurosurgery, King Fahd Hospital, Alhafaf, Saudi Arabia.
3. Department of Clinical Biochemistry, College of Medicine, Imam Abdulrahman Bin Faisal University, P.O. Box 1982, 31441 Dammam, Saudi Arabia.
4. Department of Neurology, King Fahd Hospital Bin Faisal University, Dammam, Saudi Arabia.
5. Department of Surgery, University of Pennsylvania School of Medicine, Philadelphia, PA, USA.

Received: 8 September 2022 Accepted: 15 December 2022 Published online: 20 December 2022

References
1. Englund MJ, Liverman CT, Schultz AM, Strawbridge LM. Epilepsy across the spectrum: promoting health and understanding. A summary of the Institute of Medicine report. Epilepsy Behav. 2012;23(2):266–76.
2. Shorvon SD, Farmer PJ. Epilepsy in developing countries: a review of epidemiological, sociocultural, and treatment aspects. Epilepsia. 2005;21(6):577–81.
3. Brundtland GH. Welcoming: The WHO view and launch of the second phase of the Global Campaign Against Epilepsy. Epilepsia. 2002;43(Suppl 6):5–6.
4. Hildebrandt MS, Dahl HH, Damiano JA, Smith RJ, Scheffer IE, Berkovic SF. Recent advances in the molecular genetics of epilepsy. J Med Genet. 1988;25(Suppl):I-36-54.
5. Annergren JF, Hauser WA, Anderson VE, Kurland LT. The risks of seizure disorders among relatives of patients with childhood onset epilepsy. Neurology. 1982;32(2):174–9.
6. International League Against Epilepsy Consortium on Complex Epile¬pies. Electronic address e-auea. Genetic determinants of common epilepsies: a meta-analysis of genome-wide association studies. Lancet Neurol 2014;13:893–905.
7. International League Against Epilepsy Consortium on Complex E. Genome-wide mega-analysis identifies 16 loci and highlights diverse biological mechanisms in the common epilepsies. Nat Commun 2018;9:5269.
8. Muthaffar OY. The utility of whole exome sequencing in diagnosing pediatric neurological disorders. Balkan J Med Genet. 2020;23(2):7–24.
9. Sun H, Shen XR, Fang ZB, Jiang ZZ, Wei XJ, Wang ZY, et al. Next-genera¬tion sequencing technologies and neurogenetic diseases. Life (Basel). 2021;11(4):361.
10. Consortium EK, Project EP. Ultra-rare genetic variation in common epilepsy: a case-control sequencing study. Lancet Neurol. 2017;16(2):340–3.
11. Pottow HH, Alshayban DM, Joseph R, Al-Musaf F, Al-Jabran O, Aljaafar O. Impact of adherence to antiepileptic medications on quality of life of epileptic patients in the Eastern Province of Saudi Arabia: a cross-sectional study. J Med Assoc Saudi Arabia. 2020;20(5):1.
12. Reynolds EH. The ILAE/IBE/WHO epilepsy global campaign history. International League Against Epilepsy. International Bureau for Epilepsy. Epilepsia. 2002;43(Suppl 6):9–11.
13. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. Robust relationship inference in genome-wide association studies. Bioinformatics. 2010;26(22):2867–73.
14. Arthur R, Schulz-Trieglaff O, Cox AJ, O’Connell J. AKT: ancestry and kinship toolkit. Bioinformatics. 2017;33:142–4.
15. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnPEff. SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly. 2012;6(2):80–92.
16. Karczewski KJ, Franchi LC, Tiao G, Cummings BB, Affolot J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature. 2020;581(7809):434–43.
17. Landrum MJ, Lee JM, Benson M, Brown G, Chao C, Chitipiralla S, et al. ClinVar: public archive of interpretations of clinically relevant variants. Nucleic Acids Res. 2016;44(D1):D862–8.
18. Stenson PD, Ball EV, Mort M, Phillips AD, Shiel JA, Thomas NS, et al. Human Gene Mutation Database (HGMD) 2003 update. Hum Mutat. 2003;21(6):577–81.
19. Kambouris M, Thevenon J, Soldatos A, Cox A, Stephen J, Ben-Omran T, et al. Biallelic. Ann Clin Transl Neurol. 2017;4(1):26–35.
20. Hong SY, Yang JJ, Li SY, Lee IC. A wide spectrum of genetic disorders causing severe childhood epilepsy in Taiwan: a case series of ultra‑rare genetic cause and novel mutation analysis in a pilot study. J Pers Med. 2020;10(4):281.
21. Al-Saady ML, Kaiser CS, Waksasquf F, Korenke GC, Waisfisz Q, Polstra D. Impact of adherence to antiepileptic medications on quality of life of epileptic patients in the Eastern Province of Saudi Arabia: a cross-sectional study. J Med Assoc Saudi Arabia. 2020;20(5):1.
28. Zanus C, Costa P, Faletta F, Musante L, Russo A, Grazian L, et al. Description of a peculiar alternating ictal electroclinical pattern in a young boy with a novel SPATS2 mutation. Epileptic Disord. 2020;22(5):659–63.

29. Braun F, Hentschel A, Sickmann A, Marteau T, Hertel S, Förster F, et al. Muscular and molecular pathology associated with SPATS2 deficiency in a child with EHLMRs. Int J Mol Sci. 2021;22(15):7835.

30. Papuc SM, Abela L, Steindl K, Begemann A, Simmons TL, Schmitt B, et al. The role of recessive inheritance in early-onset epileptic encephalopathies: a combined whole-exome sequencing and copy number study. Eur J Hum Genet. 2019;27(3):408–21.

31. Puusseau P, Kovacs-Nagy R, Alhaddad B, Braunisch M, Hoffmann GF, Kotzaeridou U, et al. Compound heterozygous SPATS2 variants in four families and functional studies of SPATS2 deficiency. Eur J Hum Genet. 2018;26(3):407–19.

32. Kurata H, Terashima H, Nakashima M, Okazaki T, Matsumura W, Ohno K, et al. Characterization of SPATS2-related encephalopathy in early childhood. Clin Genet. 2016;90(3):437–44.

33. Stringer RN, Jurkovicova-Tarabova B, Souza IA, Ibrahim J, Vacic T, Fathalla WM, et al. De novo SCN8A and inherited rare CACNA1H variants associated with severe developmental and epileptic encephalopathy. Mol Brain. 2021;14(1):126.34.

34. Wei Z, Liu C, Wu Z, Cao M, Qiao X, Han T, et al. The prognosis of epilepsy patients with CACNA1H missense variants: a longitudinal cohort study. Seizure. 2021;91:52–9.

35. Calhoun JD, Huffman AM, Bellinski I, Kinsley L, Bachman E, Gerard E, et al. CACNA1H variants are not a cause of monogenic epilepsy. Hum Mutat. 2020;41(6):1138–44.

36. Jiang Y, Zhang Y, Zhang P, Sang T, Zhang F, Ji T, et al. NIPA2 located in 15q11.2 is mutated in patients with childhood absence epilepsy. Hum Genet. 2012;131(7):1217–24.

37. Xie H, Zhang Y, Zhang P, Wang J, Wu Y, Wu X, et al. Functional study of NIPA2 mutations identified from the patients with childhood absence epilepsy. PLoS ONE. 2013;9(10): e109749.

38. Jiang Y, Zhang Y, Zhang P, Xie H, Chan P, et al. NIPA2 mutations are corelate with childhood absence epilepsy in the Han Chinese population. Hum Genet. 2014;133(5):675–6.

39. Thesen BE, Rumpantsetva A, Cohen JS, Alcaraz WA, Shinde DN, Tang S, et al. Deficiency of WARS2, encoding mitochondrial tryptophanyl tRNA synthetase, causes severe infantile onset leuкоencephalopathy. Am J Med Genet A. 2017;173(9):2505–10.

40. Schmidt C, Hoffmann U, Kohlmüller D, Mürder T, Zanger UM, Schwab M, et al. Comprehensive analysis of pyrimidine metabolism in 450 children with unspsective neurological symptoms using high-pressure liquid chromatography-electrospray ionization tandem mass spectrometry. J Inherit Metab Dis. 2005;28(6):1109–22.

41. Han C, Alkhater R, Foukh T, Minassian AG, Galati M, Liu RH, et al. Epileptic encephalopathy caused by mutations in the guanine nucleotide exchange factor DENND5A. Am J Hum Genet. 2016;99(6):1359–67.

42. Kaufmann R, Straussberg R, Mandel H, Fattal-Valevski A, Ben-Zeev B, Naamati A, et al. Infantile cerebral and cerebellar atrophy is associated with a mutation in the MED17 subunit of the transcription preinitiation mediator complex. Am J Hum Genet. 2010;87(5):667–70.

43. Campostirini G, DiFrancesco JC, Castellotti B, Milanesi R, Gneccci-Ruscone T, Bonzanni M, et al. A loss-of-function. HCNC4 mutation associated with familial benign myoclonic epilepsy in infancy causes increased neuronal excitability. Front Mol Neurosci. 2018;11:269.

44. Aerden M, Vallaeys L, Holvoet M, De Waeele L, Van Den Bogaert K, Devriendt K. Homozygous missense STRADA mutation in a patient with polyhydramnios, megalencephaly and symptomatic epilepsy syndrome. Clin Dysmorphol. 2021;30(3):121–4.

45. Alsaif HS, Khashab HYEL, Alkuraya FS. Two further cases of polyhydramnios, megalencephaly, and symptomatic epilepsy syndrome, caused by a truncating variant in STRADA. Am J Med Genet A. 2021;185(2):604–7.

46. Ben Romdhan S, Sakka S, Farhat N, Triki S, Dammak M, Mhiri C. A Novel SYNJ1 mutation in a Tunisian family with Juvenile Parkinson’s disease associated with epilepsy. J Mol Neurosci. 2018;66(2):273–8.

47. Pastural E, Barrat FJ, Dufourcq-Lagelouse R, Certain S, Sanal O, Jabado N, et al. Griscelli disease maps to chromosome 15q21.1 and is associated with mutations in the myosin-Va gene. Nat Genet. 1997;16(3):289–92.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.