Biallelic mutations in valyl-tRNA synthetase gene VARS are associated with a progressive neurodevelopmental epileptic encephalopathy

Jennifer Friedman1,2,3, Desiree E. Smith5,6, Mahmoud Y. Issa7, Valentina Stanley1,8, Rengang Wang1,8, Marisa I. Mendes5,6, Meredith S. Wright4, Kristen Wigby3,4, Amber Hildreth3,4, John R. Crawford1,2,3, Alanna E. Koehler8, Shimul Chowdhury4, Shareef Nahas4, Liting Zhai9, Zhiwen Xu9,10, Wing-Sze Lo9,10, Kiely N. James3,8, Damir Musaev1,8, Andrea Accogli11,12, Kether Guerrero11,13,14, Luan T. Tran11,13,14, Tarek E.I. Omar15, Tawfeg Ben-Omran16, David Dimmock4, Stephen F. Kingsmore4, Gajja S. Salomons5,6, Maha S. Zaki7, Geneviève Bernard11,13,14 & Joseph G. Gleeson1,2,3,4,8

Aminoacyl-tRNA synthetases (ARSs) function to transfer amino acids to cognate tRNA molecules, which are required for protein translation. To date, biallelic mutations in 31 ARS genes are known to cause recessive, early-onset severe multi-organ diseases. VARS encodes the only known valine cytoplasmic-localized aminoacyl-tRNA synthetase. Here, we report seven patients from five unrelated families with five different biallelic missense variants in VARS. Subjects present with a range of global developmental delay, epileptic encephalopathy and primary or progressive microcephaly. Longitudinal assessment demonstrates progressive cortical atrophy and white matter volume loss. Variants map to the VARS tRNA binding domain and adjacent to the anticodon domain, and disrupt highly conserved residues. Patient primary cells show intact VARS protein but reduced enzymatic activity, suggesting partial loss of function. The implication of VARS in pediatric neurodegeneration broadens the spectrum of human diseases due to mutations in tRNA synthetase genes.
The spectrum of genes responsible for pediatric neurodevelopmental and neurodegenerative conditions (PNDDC) is extraordinarily broad, often requiring a whole-exome or whole-genome sequencing approach rather than a targeted gene or panel approach in order to achieve a molecular diagnosis and to identify new causes of disease. This is in part due to the vast number of genetic forms of PNDDC, in addition to wide-ranging phenotypic and genetic heterogeneity among these conditions. Previous publications have addressed this issue by recruitment of large cohorts of patients with PNDDC, or through matchmaker exchanges to identify recurrently mutated genes.

Charged tRNAs are required by the ribosome to meet the cellular requirement for protein synthesis, and their availability regulates many aspects of cellular function. Charging of tRNAs requires the function of aminoacyl-tRNA synthetases (ARSs), which are ubiquitously expressed and highly evolutionarily conserved. Each amino acid has one or more designated ARS enzymes to catalyze a bond with a cognate tRNA. With 37 ARS genes for the 20 typical amino acids, 17 encode a cytoplasmic-localized enzyme, 17 encode a mitochondrial-localized enzyme, and 3 charge tRNA in both cellular locations. Individual ARS proteins that function predominantly in the mitochondria are designated with a ‘m’ after the gene name, for instance the VARS paralogue functions in the cytoplasm and VARS2 paralogue functions in the mitochondria to charge valine tRNAs.

Mutations in 31 different ARS genes are currently implicated in recessive diseases, with mutations in most mitochondrion-localized ARS genes associated with mitochondrial encephalopathies, and cytoplasmic-localized ARS genes associated with a wider range of diseases but affecting primarily the nervous system. Analysis of mutant ARS enzyme activity typically show moderate to severe reduction, but with some activity remaining, consistent with the idea that these are non-redundant and essential enzymes.

Valine is an essential branched chain amino acid, utilized for protein synthesis both in the cytoplasm and mitochondria. VARS2 is nuclear-encoded and is mutated in individuals with recessive mitochondrial encephalomyopathy. Missense mutations in VARS2 found in patients can lead to protein destabilization, which can reduce steady-state protein levels. Karaca et al. identified VARS as a candidate gene among 40 other genes profiled in 128 families with a range of neurologic presentations. They reported two pathogenic variants NM_006295.2:c.2653C>T p.(Leu885Phe) and NM_006295.2:c.3173G>A p. (Arg1058Gln) in two consanguineous families with epileptic encephalopathy; phenotypic spectrum and function were not studied.

Here we expand the findings of Karaca et al. by using exome or genome sequencing to identify and characterize five distinct biallelic VARS variants in seven additional epileptic encephalopathy patients. The VARS phenotype is characterized by a spectrum of global developmental delay, epileptic encephalopathy and primary or progressive microcephaly. Pathogenic variants disrupt highly conserved residues, and localize to the VARS tRNA binding domain and adjacent to the anticodon domain. In patients for whom tissue was available, fibroblasts show intact VARS protein but reduced enzymatic activity, suggesting partial loss of function.

Results
Identification of VARS biallelic mutations in five families.
Patients were referred for neurological or genetic assessment for PNDDC at the National Research Center in Cairo, Rady Children’s Hospital in San Diego, and McGill University Health Center in Montreal. Seven patients from five families were included in this study. Whole-exome (WES) or whole-genome (WGS) sequencing in each family led to the identification of a single genetic variant that met criteria for causality using established protocols. The researchers enrolling subjects from these three locations were matched through collaborative networks and the Matchmaker Exchange and recognized a common core phenotype of neurodevelopmental disorder with microcephaly, seizures, and cortical atrophy (NDMCSA), which contributed to the OMIM entry #617802 linked to VARS. Families 2937, 3308, and 3439, all from Egypt, had documented first-cousin parental consanguinity. Family 3007, originating from Syria, had a history of remote parental consanguinity. Family GB31, from Montreal, Canada, was Caucasian and had no history of consanguinity.

WES was used for analysis of families 2937, 3308, 3439, and GB31, and focused on identification of homozygous or compound heterozygous, rare (low minor allele frequency in the gnomAD variome) and not identified in the GME variome and potentially damaging variants prioritized using the American College of Medical Genetics organization, OMIM identity, gene function, protein expression, phenotypic assessment, and silico prediction tools. WGS was used for analysis of family 3007 with analysis filtered to retain variants with allele frequencies of <0.5% in the Exome Variant Server, 1000 Genomes, and gnomAD databases and were prioritized using the phenotypes extracted from the medical record using Human Phenotype Ontology codes. VAAST and Phevor prioritization algorithms were used to rank variants (Supplementary Methods 1).

For all families we used standard clinical variant filtering (Supplementary Methods 1), identifying a group of potentially pathogenic variants (Supplementary Data 1). For families 2937, 3308, and 3439, with documented consanguinity, we additionally filtered for homozygous variants, and further filtered for segregation within the family, predicted damaging by ‘Mutation Taster’, and Polyphen2 score >0.85; only the variant in VARS remained for each family after this approach. For family 3007, five variants passed filtering for further consideration, and all except homozygous variants in VARS were discarded due to poor phenotypic overlap, failure to find a second variant or deletion for an autosomal recessive condition, or variant inherited from an unaffected parent in an autosomal dominant condition. For this family we were unable to fully exclude impact of heterozygous SCO2 stop-gain on clinical phenotype. For family GB1, four variants met filtering criteria and all except VARS were excluded based upon failure to find a second variant or deletion for an autosomal recessive condition or occurrence of variant in a healthy population. For this family, VARS remained as the final candidate, and bi-allelic inheritance was confirmed following co-segregation analysis.

Families 2937, 3308, and 3439, derived from a cohort of 5000 families that underwent WES, and two families recruited from Cairo, shared the same NM_006295.2:c.3355C>A p.(Arg1119Cys) variant (Families 3308 and 3439), suggesting a founder mutation (Table 1, Fig. 1a). One additional family from Cairo (Family 2937) with a similar phenotype had a NM_006295.2:c.2840G>A p.(Arg947His) variant. Family 3007, identified in San Diego and of remotely consanguineous Syrian descent, was one of 200 families that were studied with WGS, and had a NM_006295.2: c.1981C>A p.(Pro661Thr) variant. Family GB31 from Montreal, Canada had compound heterozygous NM_006295.2:c. (2074G>C;1324C>T) p.(Ala692Pro;Arg442*) variants. Candidate variants were confirmed by Sanger sequencing, with segregation confirmed amongst genetically informative family members according to a strictly recessive inheritance mode.
VARS mutations associate with a severe neurodevelopmental disorder. Family 2937 had a single affected child from an uncomplicated delivery that developed generalized myoclonic and focal seizures at 3 months of age. Examination at 1 yr 7 months showed globally delayed milestones with head circumference of 40 cm (−4.6 SD), axial hypotonia and limb hypertonia, brisk reflexes and extrapyramidal movements. MRI performed at 1 yr showed cerebral cortical atrophy and mild central atrophy, with hypoplastic corpus callosum (Fig. 1b). Feeding difficulties persisted, but seizures had been well controlled on Valproate. Electroencephalogram showed right greater than left epileptiform discharges.

Fig. 1 VARS mutations lead to a neurodevelopmental disorder with microcephaly, seizures, and cortical atrophy. a Pedigrees showing consanguinity in three families (double bars). Family 3007 showed remote consanguinity. Seven affected individuals were identified (black), and one individual was likely affected but DNA was not available for genotyping (3439-IV-1). Multiple miscarriages in family GB31 suggested fetal demise. b Available axial or midline sagittal brain MRIs demonstrated diffuse cortical atrophy (yellow star) and thinned corpus callosum (yellow arrow) compared to control. The brain MRI in 3007-II-1 was acquired at 6 months of age, and thus the patient was probably too young to observe these features. Two individuals from family 3439 had follow-up MRIs that demonstrated progressive cortical atrophy. D—days; M—months; Y—years. c Major domains of the 1264 amino acid protein, showing the GST, aminoacylation (including HIGH, Editing and KMSKS domains) and anticodon-binding domains. Disease-associated variants described in this study (red) and those from Karaca et al.3 (black) are noted. Variants were in the aminoacylation or in or near the anticodon-binding domains, suggesting altered function. d Sequence conservation of mutated amino acids. The variants were either precisely conserved or maintain charge/polarity, with the exception of NM_006295.2:c.2653C>T p.(Leu885Phe), which is absent in T. thermophiles.
Family 3308 presented with a single affected child with seizures beginning at 3 weeks, characterized by generalized myoclonic and focal seizures on a weekly basis, refractory to treatment with Valproate and Levetiracetam. Electroencephalogram showed frequent generalized discharges. Assessment at 6 months showed severe global developmental delay, head circumference of 36 cm (−5.2 SD), axial hypotonia and limb hypertonia with brisk reflexes. There was a ventriculoseptal defect, which was clinically insignificant. The child died by 3 yr from chronic respiratory insufficiency resulting from neuromotor disability.

Family 3439 presented with four affected children, all showing global developmental delay, axial hypotonia, brisk reflexes, and extrapyramidal movements. The oldest (IV-1) died at 8 months followed by seizures and respiratory compromise, but otherwise normal percentiles, and thus molecular diagnosis could not be confirmed in these individuals. The electroencephalogram in 3439-IV-3 showed hypsarrhythmia, and seizures could not be controlled with Valproate, Levetiracetam, and Tiagabine. All other affected individuals have now died.

Family 3007 presented with a single affected child at 4 months with global developmental delay and infantile spasms. Electroencephalogram met criteria for hypsarrhythmia. At 10 months the seizures were controlled on Vigabatrin monotherapy. MRI at 6 months was essentially unremarkable, and head circumference continued along normal percentiles, but global developmental delay was evident. This child displayed feeding difficulties and had persistent mild elevations of liver transaminases (aspartate aminotransferase and alanine aminotransferase, ALT/AST) and elevated liver ultrasound, abn abnormal, echo echogenicity, VSD ventriculoseptal defect, PFO patent foramen ovale.

Family GB31 presented with a severely affected child soon after birth (36 weeks gestation). The mother had a history of recurrent pregnancy loss, including an elective termination at 18 weeks gestation.
We noted that the head circumference of most affected children at 18 months, demonstrating loss of subcortical white matter on both axial and midline sagittal images. Similarly, individual volume, with collapse of subcortical white matter volume, evident in cerebellum were histologically unremarkable. Disorder. Hippocampus, basal ganglia, thalamus, brainstem, and gliotic with minimal myelination but preserved axonal density, corpus callosum. The hemispheric white matter was intensely MRI at 3 months and then repeated at 4 yr (Fig. 1b). Side-by-side fixation, we repeated brain MRI in older children to assess for a loss in cerebral growth and/or atrophy. To confirm this suspicion, we repeated brain MRI in older children to assess for acquired cortical atrophy. Individual 3439-IV-4 had initial brain MRI at 3 months and then repeated at 4 yr (Fig. 1b). Side-by-side comparison demonstrated progressive loss of cerebral cortical volume, with collapse of subcortical white matter volume, evident on both axial and midline sagittal images. Similarly, individual 3439-IV-5 had an initial brain MRI at 4 months and then repeated at 18 months, demonstrating loss of subcortical white matter volume and increased lateral ventricle size. The MRI findings together with brain neuropathology from GB31-II-6 demonstrate that pathogenic variants in VARS also show particular predilection for the cerebral white matter.

**VARS deficiency associates with progressive cortical atrophy.**

We noted that the head circumference of most affected children tended to slow in growth by 1 yr of age (Table 1), which predicted a loss in cerebral growth and/or atrophy. To confirm this suspicion, we repeated brain MRI in older children to assess for acquired cortical atrophy. Individual 3439-IV-4 had initial brain MRI at 3 months and then repeated at 4 yr (Fig. 1b). Side-by-side comparison demonstrated progressive loss of cerebral cortical volume, with collapse of subcortical white matter volume, evident on both axial and midline sagittal images. Similarly, individual 3439-IV-5 had an initial brain MRI at 4 months and then repeated at 18 months, demonstrating loss of subcortical white matter volume and increased lateral ventricle size. The MRI findings together with brain neuropathology from GB31-II-6 demonstrate that pathogenic variants in VARS also show particular predilection for the cerebral white matter.

**Mutations in VARS alter evolutionarily conserved amino acids.**

The VARS mutations occurred throughout the open reading frame, predicting alteration in evolutionarily conserved amino acids. VARS encodes a 1264 amino acid protein consisting of an N-terminal glutathione-S-transferase and aminoacylation enzymatic domains and a C-terminal anticodon-binding domain, used to recognize of valine tRNAs (Fig. 1c). The amino acids that we identified as mutated in our cohort, as well as one of the pathogenic variants identified by Karaca et al. were conserved in all mammals and were further conserved to within the same amino acid type in all species profiled. Furthermore, all of our patients had variants that altered amino acid charge or group (Fig. 1d). The NM_006295.2:c.2653C>T p.(Leu885Phe) Karaca et al. mutation, however, had no counterpart in *T. thermophiles*.

**Variants predicted to affect protein structure or function.**

The class 1a ARS enzymes, which include VARS, utilize an alpha-helix bundle domain to recognize the key residues in the anticodon domain of the cognate tRNA, essential for correct charging [17]. The structure of the VARS orthologue has been determined in the tRNA-bound state for *T. thermophiles* [18], where tRNA residues A35-C36 (the major identity elements), are base-stacked upon each other, and fit into a pocket on the alpha-helix bundle domain of VARS. Hydrogen bonds are formed between VARS and A35-C36 of tRNA [2] in a base-specific manner. The discrimination of valine from the structurally similar isoleucine and threonine in the amino acid binding domain occurs via ‘double sieve’ mechanisms [19], whereby VARS excludes these incorrect bases using distinct mechanisms.

To predict the impact of the mutations on protein structure, we utilized a crystal structure of the *T. thermophiles* VARS-tRNAVal complex because human and *T. thermophiles* protein sequences share high sequence similarity (~90% conservation) [18,19], and it is the only species where structure is available. We constructed a model of human VARS-tRNAVal complex by homology modeling and examined the location relative the alpha-helical and beta-sheets of the protein (Supplementary Fig. 1), as well as how each mutation is predicted to impact local structure and tRNA binding (Fig. 2a). Residues Pro661, Arg947, and Arg1119 are in proximity to the tRNA molecule (≤6 Å). The p.(Pro661Thr) mutant exhibits no obvious change in distance from tRNA or in contacts with adjacent amino acids, but occurs within a conserved beta-sheet. Both Arg947 and Arg1119 are in alpha helices and in close proximity to the C-37 and C-39 of the tRNAVal, respectively. These may create hydrogen-bonding interactions that would be abolished by the mutations, possibly leading to reduced tRNA binding (Figs 2b, d, f). Arg947 in particular has been previously hypothesized to be important in tRNA binding [18].

The other two residues, Ala692 and Arg1058, have a distance of >15 Å from the tRNA, but again occur in alpha helices. The NM_006295.2:c.2074G>C p.(Ala692Pro) introduces the helix breaker proline and is predicted to result in a loss of a hydrogen bond with Ala689 (Fig. 2c), which may cause a conformational change, while the NM_006295.2:c.3173G>A p.(Arg1058Gln) mutation effect is less certain (Fig. 2e). Thus, all patient mutations that we identified occur within highly ordered regions of VARS, and some were additionally predicted to have an impact on tRNA binding or protein structure, possibly affecting the VARS function.

**Patient mutations in VARS reduce aminoacylation activity.**

To evaluate the impact of variants on protein stability and function, we collected skin punches then generated primary fibroblast cultures from an affected member of four of the families, as well as parental healthy controls (obligate carriers) in three families, and compared with results from cultures from unrelated healthy controls. We observed that cells from affected individuals displayed typical growth characteristics under basal conditions. For some of the recessive ARS genes mutated in human disease, there are reduced levels of protein at steady state, including for the mitochondrial-localized VARS213,20,23. Western blot analysis demonstrated VARS protein levels in patient cells that were indistinguishable from controls (Fig. 3a), suggesting that these mutations do not destabilize VARS or alter steady-state levels in detectable ways.

VARS transferase activity was assessed by incubating cytoplasmic extracts with yeast total tRNA, ATP and [15N]-valine. Valine bound by tRNA was subsequently quantified by LC-MS/MS. LARS (leucine) and RARS (arginine) activities were assessed simultaneously as controls. Extracts from affected individuals (2937-III-1, 3308-III-1, 3439-IV-4, GB31-II-6) showed that VARS transferase activity was reduced by about fivefold (i.e., to 20% of levels observed in controls, Fig. 3b) which proved to be statistically significant (*P* < 0.0001 for both comparisons, one tailed Student’s *t*-test). This reduced but not abolished activity is
consistent with reduced activity seen in other recessive ARS diseases. Next we determined if healthy parents showed reduced activity, given that they are carriers (i.e., heterozygous) for the mutation. We found that parents had activities that ranged in the 60–80% of unrelated control levels (Fig. 3c), which is consistent with their heterozygous state. Controls all showed activities within 10% of each other. Activities of LARS and RARS in patients showed some variability in activity but were all within 75–110% of control activities. These experiments point to a specific reduction in VARS activity in patient cells, which likely accounts for their clinical presentations.

Discussion

The process of charging tRNAs with amino acids is an evolutionarily conserved process in all metazoa, critical for protein translation. In mammals there are 37 ARS proteins linking the 20 primary amino acids with the 497 human tRNAs. In this study, we identified seven patients from five families with biallelic mutations in VARS. There were five pathogenic variants in total, with three of these present in a homozygous state, while two were seen in a compound heterozygous state, which associated with the most clinically severe case. All missense variants were absent or reported only rarely in public references databases (EXAC, gnomAD), predicted pathogenic by in silico analysis, disrupted highly conserved amino acid residues in critical regions of the protein, and led to severely reduced transferase activity in the tested patient cells.

The five distinct VARS pathogenic variants reported here and two previously identified were distributed throughout the VARS protein. All variants were encompassed in the full-length transcript subserving tRNA synthetase activity. Five were located in the aminoacylation domain, and two were in or adjacent to the anticodon-binding domain. Interestingly, irrespective of where the mutation occurred, we found reduced aminoacylation activity, suggesting that transferase activity reflects both the ability to bind tRNA as well as transfer the charged valine to the tRNA.

Recent attention has focused on non-translational function of ARS proteins, evidenced by the addition of newly evolved domains that are neither necessary nor sufficient for ARS activity. For the most part, the function of these additional protein domains, some of which are mammalian-specific, remain unknown. The mammalian-specific VARS N-terminal 218 residue GST domain is not essential for ARS activity, but rather binds the elongation factor EF-1a to facilitate association and attachment of EF-1a to the ribosome.
efficient transfer of charged tRNAs. In addition, there are at least six alternatively spliced VARS transcripts, which are catalytic nulls, but display specific biological activities across a spectrum of cell-based assays, particularly 'neutrophil oxidative burst'. Whether any of the mutations we identified additionally interfere with non-catalytic functions of VARS will require further study.

Mutations in both mitochondrial and cytoplasmic-localized ARSs cause a spectrum of recessive disorders, many of which include neurological phenotypes. Pathogenic variants in cytoplasmic ARSs have been associated with microcephaly, seizures, developmental delay, intellectual disability, and sensorineural hearing loss. RARS, DARS, and EPRS bi-allelic mutations cause

---

**Fig. 3** VARS patient cells show normal VARS protein level, but reduced enzymatic ARS activity. a Western blot of patient or control cells showing retained VARS protein level in affected individuals (2937-III-1A and 3308-III-1A), carrier parents (2937-III-1F and 3308-II-2M), and controls (3311-III-2C, 3838-III-1C, and 4936-III-1C). b VARS, LARS (leucine ARS) and RARS (arginine ARS) aminoacylation activity. VARS activity was specifically reduced in patient cells, normalized to control cells at 100% (from \( n = 3 \) replicates, error bars = S.E.M.). c VARS activity in patients compared with carrier parents and controls. Black: patients, light gray: parents of affecteds (obligate carriers of the VARS mutation), dark gray: healthy controls (from \( n = 3 \) replicates, error bars = S.E. M). Differences between patients and controls or carriers were calculated based upon the average of the group (\( P < 0.0001 \) for both comparisons, One tailed Student’s t-test). Full gel in Supplementary Fig. 2.
hypothesized that SARS mutations are implicated in intellectual disability, and HARS has been associated with Usher syndrome. Dominant (mono-allelic) mutations in multiple ARS genes (GARS, VARS, AARS, and HARS) have been associated with the progressive axonal neuropathy Charcot-Marie Tooth type 2 Disease, presumably acting in a dominant-negative capacity. In addition to neurologic phenotypes, other systemic manifestations are associated with recessive cytoplasmic ARS mutations (LARS, MARS), including liver dysfunction and interstitial lung disease. The mechanism driving tissue-specific phenotypes associated with certain ARS genes remains largely unexplored.

The non-syndromic clinical presentation of these affected individuals prevented us from identifying them as the same genetic condition prior to WES or WGS, but in retrospect several clinical aspects are shared. The clinical presentation in the first weeks to months of life with severe epilepsy despite aggressive anti-convulsant treatment, and association with hypsarrhythmia on electroencephalogram can be a clue. Another clue is acquired microcephaly, in the range of minus 5 SDs, observed in all but one subject after 6 months of age. The brain MRI appearance of cortical atrophy with reduced white matter volume without edema, and without involvement of deep brain structures or brainstem volume loss can differentiate VARS from leukodystrophies, mitochondrial, and pontocerebellar disorders. Unfortunately, none of our subjects have yet survived past childhood, and most succumb from medical complications in the first few years of life, suggesting a lethal condition.

The clinical features of neurodevelopmental delay with microcephaly, seizures, and cortical atrophy (MIM#617802) linked with VARS mutations can be compared with other defined conditions, as well as the function of the genes implicated in those disorders.Recessive microcephaly, seizures, and developmental delay (MIM#613402) due to mutations in the DNA-repair enzyme encoded by PNPK has onset in infancy. Recessive microcephaly (postnatal progressive) with seizures and brain atrophy (MIM#613668) is due to mutations in MED17, encoding a subunit of the transcription pre-initiation mediator complex. The VARS phenotype is also similar to AIMP1-related disorders (MIM#603605) which present with microcephaly, cortical cerebral atrophy and early-onset seizures but can be differentiated from VARS based upon presence of hypomyelination leukodystrophy, similar to that of X-linked Pelizaeus–Merzbacher disease. While there are over 55 unique genetic causes for early infantile epileptic encephalopathy (EIEE), many involving neuronal ion channels, we think that the VARS-related syndrome can be distinguished clinically from EIEE.

The etiology of varied features and severity in our cohort are uncertain. These may include environmental factors, although the contribution of other rare variants identified in these patients cannot be excluded at this point. Multiple miscarriages occurring in family GB31 that included two different neural tube defects suggests that the phenotypic spectrum may be even broader, and, at the extreme end, incompatible with life. Unfortunately, the deceased fetuses were not available for genotyping, so it remains unresolved if they shared the same genetic mutations. Notably this family displayed the only protein-termination variant in our cohort, which not surprisingly may have a more severe impact on protein function.

There is emerging evidence that genes encoding factors involved in RNA metabolism including tRNA synthetase, tRNA splicing complex, and RNA helicases, play essential roles in human brain development. Further study including in vitro and in vivo modeling systems will be necessary to determine the mechanism by which pathologic VARS variants disrupt tRNA synthetase activity, whether alternatively splice transcripts have a role in neural development or function, how these abnormalities disrupt normal brain development and how these VARS functions may be related to the function of other genes implicated in neural development.

Methods
Patient ascertainment. All experiments involving human participants or data were conducted in compliance with all relevant ethical regulations. Informed consent was obtained from all participants and their families. Approval for human subjects research was obtained from National Research Center in Cairo and University of California, San Diego Institutional Review Boards, the McGill University Health Center Research Institute Research Ethics Board, and Rady Children’s Hospital Research Compliance. All affected individuals were clinically evaluated by a pediatric neurologist and geneticist. Subjects underwent complete dysmorphic, neurological, and general examinations. Brain MRI was obtained in all living subjects, and longitudinal assessment was performed up until the present time to determine disease course. Venous blood or saliva for DNA isolation was collected from all genetically informative living members of each family at the time of ascertainment. In four families, dural skin punch was obtained after consent from the index individual and genetically informative relatives, and cultured in 10% fetal bovine serum and DMEM to obtain low-passage primary fibroblast cultures.

Whole exome or genome sequencing and variant identification. Details regarding sequencing, filtering and prioritization protocols for each family are outlined in Supplementary Note 1 and Supplementary Data 1. Exome or Genome sequencing was performed on affected member(s) from each family using established methods. All variants were prioritized by allele frequency, conservation, and predicted effect on protein function, and were tested by Sanger sequencing for segregation with disease within the entire family.

Construction of human VARS-tRNA structural model. The human VARS structure model was generated by the protein structure homology-modeling server Phyre V2.0 based on the T. thermopiles VARS structure (PDB: 1GAX). The structure of the T. thermopiles tRNA structural model (PDB: 1GAX) was docked into the human VARS model using the PathDock server. All the structural figures were prepared with Pymol (http://www.pymol.org).

Western blot. Western blots were performed with rabbit polyclonal anti-VARS antibody (Abclonal A14182: 1:1000) and mouse monoclonal anti-a-tubulin (Sigma T6074; 0.2 µg/ml) in 5% non-fat milk in PBST. Full blot in Supplementary Fig. 2.

Combined VARS, LARS, and VARS activity assay. Skin punches were collected at a single timepoint from each patient studied here, then expanded and used for enzymatic analysis. Steady-state aminoacylation assays were performed in technical triplicate in fibroblast lysates [cytosolic fraction, no mitochondrial proteins detected using western blot (Supplementary Fig. 3)] at 37 °C for 10 min in reaction buffer (50 mmol/L Tris buffer pH 7.5, 12 mmol/L MgCl2, 25 mmol/L KCl, 1 mg/ml bovine serum albumin, 0.5 mmol/L spermine, 1 mmol/L ATP, 0.2 mmol/L yeast total tRNA, 1 mmol/L dithiothreitol, 0.3 mmol/L [15N]-valine, [15C]-leucine and [15N]-arginine). Reaction was terminated using trichloroacetic acid. After sample washing with trichloroacetic acid, ammonia was added to release the labeled amino acids from the tRNAs. [15C]-valine, [15C]-leucine, and [15C]-arginine were added as internal standards. Labeled amino acids were quantified by LC-MS/MS. Intra-assay variation was <15%. Student t-test with unequal variances were performed using SPSS version 22 to test significant differences between patients and controls/carriers.

Data availability
VARS variant data that support the findings of this study have been deposited in ClinVar with accession codes [https://www.ncbi.nlm.nih.gov/clinvar/variation/522588/#clinal-assertions] SCV000715349.1 (family 3007), SCV000734850 (c.2074G>C), and SCV000734851 (c.1324C>T) (family GB31). The whole-exome sequencing data from individuals from families 2937, 3308, and 3439 in this study have been deposited in the database of Genotypes and Phenotypes (dbGaP) under accession phs000288.v1.p1. Additional data that support the findings of this study are available from the corresponding author upon reasonable request.

Received: 5 January 2018 Accepted: 3 October 2018
Published online: 12 February 2019

References
1. Sobreira, N. L. M. et al. Matchmaker exchange. Curr. Protoc. Human. Genet. 95, 9.31.1–9.31.15 (2017).
2. Boycott, K. M. et al. International cooperation to enable the diagnosis of all rare genetic diseases. Am. J. Hum. Genet. 100, 695–705 (2017).

NATURE COMMUNICATIONS | DOI: 10.1038/s41467-018-07067-3 | www.nature.com/naturecommunications
3. Karaca, E. et al. Genes that affect brain structure and function identified by rare variant analyses of Mendelian neurologic disease. *Neuron* **88**, 499–513 (2015).

4. Alazami, A. M. et al. Accelerating novel candidate gene discovery in neurogenetic disorders via whole-exome sequencing of prescreened multiplex consanguineous families. *Cell Rep.* **10**, 148–161 (2015).

5. Najmabadi, H. et al. Deep sequencing reveals 50 novel genes for recessive childhood ataxia. *Hum. Mutat.* **33**, 244–253 (2012).

6. Schimmel, P. The emerging complexity of the tRNA world: mammalian tRNAs beyond protein synthesis. *Nat. Rev. Mol. Cell Biol.* **19**, 45–58 (2018).

7. Antonellis, A. & Green, E. D. The role of aminoacyl-tRNA synthetases in mitochondrial disease. *Nat. Rev. Genet.* **14**, 774–788 (2013).

8. Santos-Cortez, R. L. et al. Mutations in KARS, encoding lysyl-tRNA synthetase, cause autosomal-recessive nonsyndromic hearing impairment DFN889. *Am. J. Hum. Genet.* **93**, 132–140 (2013).

9. Pufferberg, E. G. et al. Genetic mapping and exome sequencing identify variants associated with five novel diseases. *PLoS. One* **7**, e82936 (2012).

10. Pierce, S. B. et al. Mutations in mitochondrial histidyl-tRNA synthetase HARS2 cause ovarian dysgenesis and sensorineural hearing loss of Perrault syndrome. *Proc. Natl. Acad. Sci. USA* **108**, 6543–6548 (2011).

11. Baertling, F. et al. Neonatal encephalocardiomyopathy caused by mutations in the aminoacyl-tRNA synthetase TARS. *J. Clin. Invest.* **132**, 777–788 (2012).

12. Diodato, D. et al. VARS2 and TARS2 mutations in patients with Charcot-Marie-Tooth disease type 2D and distal spinal muscular atrophy type V. *Am. J. Hum. Genet.* **72**, 1293–1299 (2003).

13. van Meel, E. et al. Rare recessive loss-of-function methionyl-tRNA synthetase mutations presenting as a multi-organ phenotype. *Bmc. Med. Genet.* **14**, 106 (2013).

14. Feinstein, M. et al. Pelizaeus-Merzbacher-like disease caused by AIMP1/p43 homologous mutation. *Am. J. Hum. Genet.* **87**, 820–828 (2010).

15. MacArthur, D. G. et al. Guidelines for investigating causality of sequence variants in human disease. *Nature* **577**, 651–663 (2014).

16. Feigin, I. et al. Human CLP1 mutations alter tRNA biogenesis, affecting both cerebral and central nervous system function. *Cell* **157**, 636–650 (2014).

17. Alazami, A. M. et al. Accelerating novel candidate gene discovery in neurogenetic disorders via whole-exome sequencing of prescreened multiplex consanguineous families. *Cell Rep.* **10**, 148–161 (2015).

18. Najmabadi, H. et al. Deep sequencing reveals 50 novel genes for recessive childhood ataxia. *Hum. Mutat.* **33**, 244–253 (2012).

19. Schimmel, P. The emerging complexity of the tRNA world: mammalian tRNAs beyond protein synthesis. *Nat. Rev. Mol. Cell Biol.* **19**, 45–58 (2018).

20. Taylor, R. W. et al. Use of whole-exome sequencing to determine the genetic basis of Charcot-Marie-Tooth disease. *Hum. Mutat.* **38**, 197–202 (2016).

21. Schneidman-Duhovny, D., Inbar, Y., Nussinov, R. & Wolfson, H. J. Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858 (2015).

22. Schindelin, D. H., Buhovaty, D., Inbar, Y., Nussinov, R. & Wolfson, H. J. PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Res.* **33**, W363–W367 (2005).
and J.G.G. wrote the manuscript. T.B.-O., G.S.S., G.B., J.F., and J.G.G. supervised the project. All authors critically reviewed the paper.

**Additional information**

Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-018-07067-3.

**Competing interests:** D.D. is a consultant for Biomarin, Ichorian, and Complete Genomics. J.F. holds shares in Illumina and Spouse is Founder and Principal of Friedman Bioventure, which holds a variety of publicly traded and private biotechnology interests. D.D. is on the Scientific Advisory Board for Audentes Therapeutics. The remaining authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

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