Identification of Novel ADGRV1 and KCNC2 Variants Using Whole-Exome Sequencing in Two Colombian Patients with Usher and Encephalopathy Syndromes

Estephania Candelo (ecandelo@icesi.edu.co)
Fundación Valle del Lili
Lorena Diaz-Ordoñez
Universidad Icesi
Rafael Pacheco
Universidad Icesi
Emelina Ruiz
Fundación Valle del Lili
Harry Pachajoa
Fundación Valle del Lili

Research Article

Keywords: mutation, whole-exome sequencing, usher syndrome, hearing loss, epilepsy.

Posted Date: October 27th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-923411/v1

Abstract

Introduction: Usher syndrome has a broad phenotypic and genotypic spectrum. Developmental and epileptic encephalopathy-52 (DEE52) is a sever autosomal recessive seizure disorder that is characterized by infantile onset of refractory seizures, consequently resulting in delayed global development. This study aimed to describe the clinical features and to investigate the four variants identified in a Colombian family with Usher syndrome and *KCNC2* encephalopathy syndrome.

Methods and Results: We present a case of a family with two clinically relevant phenotypes: a mother with a compound heterozygous mutation causing Usher Syndrome, type IIC (*USH2C*) and her 15-year-old son who carried one heterozygous variant in the *KCNC2* gene (p.P470S) and two cis mutations (p.V2927I and p.Q4955EfsTer10) in the *ADGRV1* gene segregated from his mother, and a second non-disrupted allele. Owing to this, the boy did not present with *USH2C* but presented a developmental epilepsy syndrome. His younger sibling was unaffected, although he did inherit the trans mutation in a single pathogenic allele from his mother.

Discussion and Conclusion: Whole-exome sequencing helps detect genes related to known and novel hearing loss and seizure syndrome. However, familiar segregation studies are an excellent method to clarify genotype-phenotype correlation in families, where multiple genes of clinically relevant have been identified. This method helps determine the genotype-phenotype relationship of a disease, which is associated with the clinical presentation and determines the pathogenicity of variants that are classified as variants of uncertain clinical significance.

Introduction

ADGRV1 encodes the adhesion G-protein coupled receptor (GPCR) V1, which comprises a signal peptide located at the N-terminal domain and two subunits: alpha and beta. The alpha subunit contains 35 Calx-β motifs, which are related with calcium-mediated protein–protein interaction [McMillan et al., 2002], and six ethylene-responsive element binding factor-associated amphiphilic repression or EAR repeats, a conserved repeated region of about 44 amino acid residues, which possibly play a role in ligand recognition by the receptor [Scheel et al., 2002]. One GPCR proteolytic site (GPS) motif is located at the end of the alpha subunit. This GPS motif is a conserved sequence of ~40 amino acids containing canonical cysteine and tryptophan residues and represents an autoproteolytic site that is likely relevant for GPCR signaling [Sugita et al., 1998]. Moreover, the beta subunit is located at the C-terminal and contains seven transmembrane domains, such as all GPCRs, which participate in the transduction of physiological signals from the outside of the cell to the inside via G proteins (1,4,5) (shown in Fig. 1).

KCNC2 encodes to a potassium voltage-gated potassium channel subunit (Kv3.2), which is essential to kinetically optimize a high-frequency firing of fast spiking GABAergic interneurons and retinal ganglion cells. The voltage-gated potassium channel mediates transmembrane potassium transport in excitable membranes, primarily in the brain. It also contributes to firing sustained trains of short action potential at

high-frequency in thalamocortical and suprachiasmatic nucleus neurons as well as in hippocampal and neocortical interneurons. Thus, this channel plays a role in maintaining the fidelity of synaptic transmission in neocortical GABAergic interneurons by generating action potential repolarization at nerve terminals and thus reduces calcium influx and GABA neurotransmitter release [Rudy and McBain, 2001; Vetri et al., 2020].

Materials And Methods

Clinical report - mother

A 36-year-old woman who visited our clinic was the first child of parents who had a non-consanguineous marriage. Her family history was positive for epilepsy, and she had deafness since childhood. Her hearing loss started when she was 5 years old with occasional dizziness, frequent falls, loss of balance, otalgia, and frequent episodes of acute otitis media, which more frequently affected the right ear. After a year, at the age of 6 years, she presented with profound hearing loss and underwent a hearing device implantation. However, she showed no improvement and experienced loss in sound discrimination until the age of 25 years. Thus, she received supportive therapy with bilabial pronunciation during 6–7 years of age to preserve spoken language. She used to have frequent upper tract respiratory infections and asthma. She had never presented with nyctalopia or visual acuity decline (shown in Fig. 2). Ophthalmological studies showed that all anterior segment findings were within normal limits, and she had normal pupils with no afferent pupillary defect. Fundus examination revealed normal findings. Audiometry findings revealed a profound hearing loss in both the ears with responses at 67.5 dB and 63.75 dB for right and left ears, respectively. Visual field examination showed pattern II with bilateral inferotemporal loss, more markedly in the right eye. The patient has two sons, aged 15 and 11 years, respectively. The oldest son presented multiple seizures episodes, whose description is provided in the subsequent section. After obtaining consent from the patient and her family members to participate in the study, peripheral blood sample was collected for whole-exome sequencing (WES) and Sanger sequencing.

Clinical report - child

The 15-year-old boy was the first child of non-consanguineous parents. His family history was positive for epilepsy. In addition, his mother had neurosensorial progressive hearing loss since 6 years of age with audiometry findings revealing bilateral hearing loss with responses at 67 and 85 dB in right and left ear, respectively, with 100% discrimination and refractive disorder. The patient’s younger brother aged 12 years is undergoing schooling and shows no cognitive impairment. The patient was born at term by vaginal delivery after an uneventful pregnancy with a birth weight of 2,700 g (body length and head circumference at birth were not reported). At 30 months of life, the patient started to have clusters of tonic spasms at night, on average 10 episodes each night, with gaze deviation on some occasions and without sphincter relaxation, which were more severe at the beginning of the sleep cycle and persisted in following days. At the age of 3 years, the patient developed cardiorespiratory arrest, which required an

extended stay in the intensive care unit (ICU) which was followed by global developmental. The developmental milestones and the ages at which the patient achieved them were as follows: crawling at 11 months, walking at 18 months, and poor speech at 2 years. During the stay in the ICU, the patient was diagnosed to have clusters of tonic spasms and seizures. Subsequently, he was treated with valproic acid and clonazepam until he reached the age of 5 years. However, epilepsy remained substantially drug-resistant. During this time, the patient recovered standing and walking and started saying a few words. At the age of 7 years, he developed epileptic encephalopathy requiring ICU and adjustment of the epileptic drugs with levetiracetam, clobazam, and topiramate, which resulted in partial control of the epileptogenic activity. In the following years, seizures occurred several times a week on some occasions associated with fever requiring the adjustment of anticonvulsant therapy. Additionally, he required ICU hospitalization owing to infectious disease.

The patient was additionally assessed by the genetics department for infantile epileptic encephalopathy, swallowing disorder, cognitive deficits, and psychomotor developmental delay with speech delay. Physical examination revealed the following: at 14 years, anthropometric parameters such as head circumference was 49 cm ($p < 1$; z-score: −4.71), weight was 35.1 kg ($p = 1$; z-score: −3.26), and length was 143.5 cm ($p < 1$; z-score: −3.63). Dysmorphisms included microcephaly, long face, bilateral epicanthal folds, overfolded helix, high set nasal root implantation and prominent nasal bridge, retrognathia, several hyperpigmented lesions in the chest, and four café-au-lait macules (shown in Fig. 2). The neurological examination revealed that the patient had a childish attitude, poor speech with prominent restrictive vocabulary (12 to 15 words), mainly guttural sounds, and echolalic syllables. He could nominate some words after his parents assigned them. He communicated with his mother by gestures and followed simple commands. On few occasions, he had an aggressive conduct. At the last observation, the patient was 15 years old and occasionally presents generalized tonic-clonic seizures. The child was treated with levetiracetam every 8 hours. The last severe seizure episode was in January 2021, and he has not presented any new sleep spasm episodes in the previous years. The patient presented an auditive evoke potential score of 30 dB with v wave, representing a normal register.

He had previously undergone normal brain magnetic resonance imaging. In addition, previous assessments for plasma amino acids and acyl carnitines, array Comparative Genomic Hybridization (CGH), and electroencephalography were all normal.

WES

WES was conducted as trio-WES, wherein both parents and their affected child underwent sequencing simultaneously. The exome was captured using a SureSelectXT Human All Exon V5 capture kit (Agilent) of 51 Mb, and sequencing was performed using an Illumina Hiseq 2000 sequencing system (Illumina, San Diego, CA). In addition, the mother underwent WES to identify any germline mutations that could explain her Usher syndrome. During the sequencing, paired readings of 101 nucleotides in length were obtained. Subsequently, the different variants were analyzed with a focus on genes related to the patient's phenotype. Only variants in the coding region and those within ± 20 bp of flanking intron regions with

minor allele frequency <1% were evaluated and compared using datasets obtained from the 1000 Genomes Project Consortium, dbSNP, Exome Variant Server, and Exome Aggregation Consortium databases. Parental carrier status of the clinically relevant variants was confirmed by Sanger sequencing.

Sanger sequencing

Sanger sequencing was performed to verify and validate variants within the family. Genomic DNA was extracted and amplified by polymerase chain reaction (PCR) using oligonucleotide primers designed to amplify exons 19 and 73 of *ADGRV1* (GeneBank NG_007083) and exon 3 of *KCNC2* (GeneBank NC_000012.12). Primer sequences and PCR conditions will be provided upon request. PCR products were sequenced using the BigDye Terminator Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) and were analyzed using the ABI Prism 3500 Genetic Analyzer (Applied Biosystems).

Protein structure modeling and stability analysis

To identify the potentially pathogenic genetic variants in mother and her child, the amino acid sequence of human *ADGRV1* (residues 1108–1208 and 2814–3048) and *KCNC2* (residues 453–502) were obtained from the Uniprot entries Q8WXG9 and Q96PR1, respectively. Three models were generated *ab initio* using I-Tasser [Roy et al., 2010; Yang et al., 2014], Robetta [Kim et al., 2004], and SwissModel [Bienert et al., 2017; Waterhouse et al., 2018]. These models were validated using the structure validation algorithms Verify3D [Eisenberg et al., 1997], ProSA-web [Wiederstein and Sippl, 2007], and ERRAT [Colovos and Yeates, 1993]. Then, the best *in silico* models were chosen to predict the change in protein stability following single point mutation using MAESTRO web [Laimer et al., 2015], Rhapsody [Ponzoni et al., 2020], I-Mutant [Capriotti et al., 2006], DynaMut2 [Rodrigues et al., 2021], PremPS [Chen et al., 2020], CUPSAT [Parthiban et al., 2006], and SDM and mCSM-membrane (only for KCNC2 model) [Pires et al., 2020].

Mutational analysis

The pathogenic effects of three missense mutations were analyzed using 10 *in silico* prediction tools: Polyphen-2 [Ramensky et al., 2002], DANN [Quang et al., 2015], EIGEN [Dong et al., 2015; Ionita-Laza et al., 2016], FATHMM-MLK [Shihab et al., 2013], MutationTaster [Schwarz et al., 2014], BayesDel noAF [Feng, 2017], LRT [Chun and Fay, 2009], MetaLR [Dong et al., 2015], SIFT [Kumar et al., 2009], and SIFT4G [Sim et al., 2012].

Results

Genetic variants

Genetic variants identified using trio-WES in the mother and her affected child were prioritized based on the characteristics of the variant, *in silico* predictions of pathogenicity, and the phenotype of the patient. Filtering and Sanger sequencing validation revealed genetic variants in *ADGRV1* in both the mother and child, and one *de novo* missense variant in *KCNC2* in the child only (shown in Table 1 and Fig. 3).

Protein modeling and stability analysis

The structure of the domains Calx-beta 9 and Calx-beta 20-21 of *ADGRV1* and a region of 273 amino acids of *KCNC2* were modeled by SwissModel, I-TASSER, and Robetta. All models were validated using Verify3D [Eisenberg et al., 1997], ProSA-web [Wiederstein and Sippl, 2007], and ERRAT [Colovos and Yeates, 1993] (shown in Table 2). Subsequently, VERIFY3D was used to determine the compatibility of an atomic model (three-dimensional) with its own amino acid sequence (one-dimensional). A higher score indicated high-quality of the structure [Eisenberg et al., 1997]. In the ProSA-web tool, the z-score of a protein was defined as the energy separation between the native fold and the average of an ensemble of the misfolds in standard deviation units of the database. A z-score outside a range characteristic for native proteins of similar sizes indicated an erroneous structure [Al-Khayyat and Al-Dabbagh, 2016]. Finally, ERRAT was used to calculate the error function based on the statistics of nonbonded atom-atom interactions in the reported structure [Colovos and Yeates, 1993].

The ERRAT quality factor for all models was high (>80%). As the generally accepted range for a high-quality model is >50, this analysis revealed that the backbone conformation and nonbonded interactions of all models were within the scope of a high-quality model. The Z-scores obtained using ProSA-web for all Robetta models were within the range of scores typically observed for proteins of similar size, indicating highly reliable structures. The evaluation data from Verify3D showed that models built by Rosetta software were the best structures for *ADGRV1* Calx-beta 9 and *ADGRV1* Calx-beta 20-21, with 95.54% and 80.00% residues having an average 3D-1D score ≥0.2 in the 3D/1D profile, respectively. Nevertheless, all models of *KCNC2* residues 346–368 had <80% of the amino acids with score ≥0.2 in the 3D/1D profile. The Robetta model had a value near to 80% (79,12), and thus, it was selected. The protein stability change was predicted using the Robetta models (shown in Table 3).

Mutational analysis

The *in silico* tools for pathogenicity prediction were used to further validate the mutations identified, and the analysis has been described in Table 4. Of the predictor bioinformatic tools that identified the P470S variant in *KCNC2* as pathogenic, eight considered it as damaging or disease causing, whereas only two considered it as tolerable. Regarding the variants located in *ADGRV1*, all ten bioinformatic tools classified the variant G1148 as damaging. In contrast, V2729I mutation was predicted as a tolerated or neutral by five of the ten bioinformatic tools (shown in Table 4).

Discussion/conclusion

The massive parallel sequencing of the entire genome and exome has facilitated the direct assessment of causative genetic variation, which allows the co-occurrence of more than one rare disease in a single family. Here we describe the illustrative case of a single family where this strategy allowed the identification of two distinct etiologies. We used segregation studies and phenotypic manifestation to diagnose the co-occurrence of hereditary encephalopathy and USH2 in the same family. The elucidation of genetic causes involved in rare diseases can also be hampered by factors such as variability in the

phenotypic manifestation with regard to the age of onset, clinical features, and severity of symptoms and from difficulties in distinguishing two etiologically distinct disorders with partially overlapping symptoms [Hoefsloot et al., 2014; Lal et al., 2016]. The successful identification of disease-associated genetic loci critically depends on precise phenotype information as much as on accurate inheritance models. Indeed, problems in delineating the exact phenotype are generally correlated with the failure in discriminating the affected family members from the unaffected ones and the segregation model [Lal et al., 2016].

We report the case of a 36-year-old mother of two sons. The older son had a history of severe developmental delay and epilepsy. In addition, the mother had a family history of epilepsy and a personal account of deafness since childhood and retinopathy from adulthood. Thus, a trio-WES was performed, which revealed three variants in *ADGRV1* and one variant in *KCNC2* in the boy with epilepsy, without more genetic variants of clinical interest. *ADGRV1* encodes the adhesion GPCR V1, a large calcium-binding protein that is highly expressed in the embryonic central nervous system (CNS). During the mid-gestation period (18–28 weeks), its expression is prominent in the ventricular zone. Still, its expression declines as neurogenesis completes, suggesting that *ADGRV1* could play a role in the development of CNS [McMillan et al., 2002].

GPCR V1 is required for hair bundle ankle link formation, which connects growing stereocilia in the developing cochlear hair cell in the inner ear. In response to extracellular calcium, it activates kinases such as protein kinase A and C to regulate myelination by inhibiting the ubiquitination of the myelin-associated glycoprotein (MAG), which enhances the stability of this protein in myelin-forming cells of the auditory pathway. In retinal photoreceptors, it is required for the maintenance of the periciliary membrane complex that possibly plays a role in regulating intracellular protein transport [Shin et al., 2013]. Reportedly, homozygous or compound heterozygous mutation in *ADGRV1* cause USH2C (OMIM: 605472), a heterogeneous autosomal recessive disorder characterized by deafness and blindness in adults [Weston et al., 2004; Besnard et al., 2012]. In the present case report, the mother was found to have three genetic variants. She had a compound heterozygous mutation, where variants c.8779G>A and c.14862_14866delACAAA were located on the same allele (in *cis*) (shown in Fig. 2 and Table 1). Although the *cis* configuration of these genetic variants would contribute to a single pathogenic allele, the third variant (c.3443G>A) affected the other allele (in *trans)*, resulting in two pathogenic alleles in the affected patient. In itself, Sanger sequencing cannot differentiate whether two genetic variants are cis or trans. However, in this case, familiar segregation studies of the variants confirmed this finding (shown in Fig. 2 and Table 1).

The c.3443G>A (p.G1148D) mutation is located at the Calx-beta 9 repeat of the GPCR V1. The Calx-beta 9 repeat domain is related to the calcium-binding and specific protein–protein interactions, perhaps calcium-mediated, which allow cell adhesion [McMillan et al., 2002]. However, in ClinVar, the p.G1148D mutation has been reported as a variant of uncertain clinical significance (VUS). Furthermore, most *in silico* predictions performed in this study suggest that this variant reduces protein stability. These findings propose that p.G1148D might be a likely pathogenic variant and a cause of USH2C in our patient (shown in Fig. 2, Table 2, and Table 3). The c.8779G>A (p.Val2927Ile) variant is located close to the Calx-

beta 20 domain of the GPCR V1 protein. Although this variant is reported as a VUS in clinical databases, three of 10 *in silico* predictions showed that this variant might not affect the protein stability (shown in Table 4). However, p.Val2927Ile might not be considered a likely benign variant owing to it being in cis with frameshift variant c.14862_14866delACAAA. Additional functional studies are required to test the pathogenicity of this variant. The novel frameshift mutation c.14862_14866delACAAA (p.Gln4955GlufsTer10) creates a premature stop codon before the Calx-beta 33 domain leading to the loss of GPS motif and all C-terminal transmembrane domains. It results in a shortened and likely nonfunctional GPCR V1 protein owing to the loss of cell membrane anchoring domains.

In this study, we present a family case with two clinically relevant phenotypes: (1) in a mother with a compound heterozygous mutation that resulted in USH2C and (2) in her 15-year-old son that carried one *de novo* heterozygous variant in *KCNC2* and two cis mutations (p.Val2927Ile and p.Gln4955GlufsTer10) in *ADGRV1* inherited from his mother as well as a second non-disrupted allele. Owing to this, the son was not affected with USH2C. The mother's younger child was not affected as he only inherited a single mutated allele (shown in Fig. 2B). In addition, WES identified a heterozygous missense mutation in *KCNC2* (p.Pro470Ser) in the older son with developmental epilepsy, which results from a single amino acid substitution of proline with serine at position 470 of the protein. *KCNC2* encodes a voltage-gated potassium channel that mediates transmembrane potassium transport in excitable membranes, primarily in the brain, and contributes to the regulation of the fast action potential repolarization and in sustained high-frequency firing in neurons of the CNS.

Here, we report a novel compound heterozygous mutation, c.3443G>A and c. 14862_14866delACAAA, in *ADGRV1* and a heterozygous mutation c.1408C>T in *KCNC2*, which was identified by WES. The two missense mutations in *ADGRV1* and *KCNC2* result, according to *in silico* predictions, in a decreased protein stability, while the frameshift mutation in *ADGRV1* results in a shorter aberrant protein. We broadened the mutational spectrum of these two diseases and have segregated variants in two genes in one family, generating a broad genotypic and phenotypic range of the diseases and providing new loci for both disorders. The combination of molecular diagnostic techniques such as WES with bioinformatic tools as well as clinical information can help patients with USH2 and *KCNC2* epileptic syndrome obtain a more accurate diagnosis and understand the complexity of disease segregation in affected families. WES helps identify the copy number and structural variations as well as detect genes related to novel hearing loss and seizure syndromes. However, familiar segregation studies are an excellent method to clarify genotype–phenotype correlation in families where multiple clinically relevant gene variants have been identified. They also help in VUS reclassification, as in the present case. Post-lingual hearing loss is a heterogenous disorder caused by environmental and genetic factors. It also has a wide inheritance pattern owing to its broad heterogeneity and the indistinguishable phenotypes for many of the genes. Thus, an appropriate approach is the massive parallel sequencing technique which target a large set of genes such as in exome-wide analysis.

Declarations

Acknowledgement

We thank the patient and his parents for agreeing to the publication of this report. We also thank the people who have contributed to this study's development and execution and Tobias Yates for his language editing.

Statement of Ethics

This research was conducted in accordance with the Declaration of Helsinki. The subjects and his/her parents provided written informed consent to the reporting of these cases (including the publication of images). Information revealing the subject's identity was not included in the manuscript. The patients were identified by number and not by his/her real name. *This study protocol was reviewed and approved by Comite de Etica e Investigaciones Medicas Fundacion Valle del Lili, approval number [1504].*

Consent to participate statement: written informed consent was obtained from participants and his parent to participate in the study.

Conflict of Interest Statement

The authors have no conflicts of interest to declare

Funding Sources

All forms of support and funding for study design, data collection, data analysis and manuscript writing were provided by authors' employers (Universidad Icesi and Fundación Valle del Lili).

Author Contributions

All authors made substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data; took part in drafting the article or revising it critically for important intellectual content; agreed to submit to the current journal; gave final approval of the version to be published; and agree to be accountable for all aspects of the work.

Data Availability Statement

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

References

1. Al-Khayyat MZS, Al-Dabbagh AGA. In silico Prediction and Docking of tertiary Structure of LuxI, an Inducer synthase of Vibrio fischeri. Rep Biochem Mol Biol. 2016;4(2):66-75.
2. Besnard T, Vaché C, Baux D, Larrieu L, Abadie C, Blanchet C, et al. Non-USH2A mutations in USH2 patients. Hum Mutat. 2012;33(3):504-10.

3. Bienert S, Waterhouse A, De Beer TAP, Tauriello G, Studer G, Bordoli L, et al. The SwissModel Repository-new features and functionality. Nucleic Acids Res. 2017;45(D1):D313-9.
4. Capriotti E, Calabrese R, Casadio R. Predicting the insurgence of human genetic diseases associated to single point protein mutations with support vector machines and evolutionary information. Bioinformatics. 2006;22(22):2729-34.
5. Chen Y, Lu H, Zhang N, Zhu Z, Wang S, Li M. PremPS: predicting the impact of missense mutations on protein stability. PLOS Comp Biol. 2020;16(12):e1008543.
6. Chun S, Fay JC. Identification of deleterious mutations within three human genomes. Genome Res. 2009;19(9):1553-61.
7. Colovos C, Yeates TO. Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci. 1993;2(9):1511-9.
8. Dong C, Wei P, Jian X, Gibbs R, Boerwinkle E, Wang K, et al. Comparison and integration of deleteriousness prediction methods for non-synonymous SNVs in whole-exome sequencing studies. Hum Mol Genet. 2015;24(8):2125-37.
9. Eisenberg D, Lüthy R, Bowie JU. VERIFY3D: assessment of protein models with three-dimensional profiles. Methods Enzymol. 1997;277:396-404.
10. Feng BJ. PERCH: A unified framework for disease gene prioritization. Hum Mutat. 2017;38(3):243-51.
11. Hoefsloot LH, Feenstra I, Kunst HPM, Kremer H. Genotype-phenotype correlations for hearing impairment: approaches to management. Clin Genet. 2014;85(6):514-23.
12. Ionita-Laza I, Mccallum K, Xu B, Buxbaum JD. A spectral approach integrating functional genomic annotations for coding and noncoding variants. Nat Genet. 2016;48(2):214-20.
13. Kim DE, Chivian D, Baker D. Protein structure prediction and analysis using the Robetta server. Nucleic Acids Res. 2004;32(Web Server issue):W526-31.
14. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc. 2009;4(7):1073-81.
15. Laimer J, Hofer H, Fritz M, Wegenkittl S, Lackner P. Maestro - multi agent stability prediction upon point mutations. BMC Bioinformatics. 2015;16:116.
16. Lal D, Neubauer BA, Toliat MR, Altmüller J, Thiele H, Nürnberg P, et al. Increased probability of co-occurrence of two rare diseases in consanguineous families and resolution of a complex phenotype by next generation sequencing. PLOS ONE. 2016;11(1):e0146040.
17. Parthiban V, Gromiha MM, Schomburg D. CUPSAT: prediction of protein stability upon point mutations. Nucleic Acids Res. 2006;34(Web Server issue):W239-42.
18. Pires DEV, Rodrigues CHM, Ascher DB. mCSM-membrane: predicting the effects of mutations on transmembrane proteins. Nucleic Acids Res. 2020;48(W1):W147-53.
19. Ponzoni L, Peñaherrera DA, Oltvai ZN, Bahar I. Rhapsody: predicting the pathogenicity of human missense variants. Bioinformatics. 2020;36(10):3084-92.

20. Quang D, Chen Y, Xie X. DANN: A deep learning approach for annotating the pathogenicity of genetic variants. Bioinformatics. 2015;31(5):761-3.
21. Ramensky V, Bork P, Sunyaev S. Human non-synonymous SNPs: server and survey. Nucleic Acids Res. 2002;30(17):3894-900.
22. McMillan DR, Kayes-Wandover KM, Richardson JA, White PC. Very large G-protein-coupled receptor-1, the largest known cell surface protein, is highly expressed in the developing central nervous system. J Biol Chem. 2002;277(1):785-92.
23. Rodrigues CHM, Pires DEV, Ascher DB. DynaMut2: assessing changes in stability and flexibility upon single and multiple point missense mutations. Protein Sci. 2021;30(1):60-9.
24. Roy A, Kucukural A, Zhang Y. I-TASSER: A unified platform for automated protein structure and function prediction. Nat Protoc. 2010;5(4):725-38.
25. Rudy B, McBain CJ. Kv3 channels: voltage-gated K+ channels designed for high-frequency repetitive firing. Trends Neurosci. 2001;24(9):517-26.
26. Scheel H, Tomiuk S, Hofmann K. A common protein interaction domain links two recently identified epilepsy genes. Hum Mol Genet. 2002;11(15):1757-62.
27. Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the deep-sequencing age. Nat Methods (Vol. 11, Issue 4, pp. 361–362). 2014:361-2.
28. Shihab HA, Gough J, Cooper DN, Stenson PD, Barker GLA, Edwards KJ, et al. Predicting the functional, molecular, and phenotypic consequences of amino acid substitutions using hidden markov models. Hum Mutat. 2013;34(1):57-65.
29. Shin D, Lin ST, Fu YH, Ptácek LJ. Very large G-protein-coupled receptor 1 regulates myelin-associated glycoprotein via Gαs/Gαq-mediated protein kinases A/C. Proc Natl Acad Sci U S A. 2013;110(47):19101-6.
30. Sim NL, Kumar P, Hu J, Henikoff S, Schneider G, Ng PC. SIFT web server: predicting effects of amino acid substitutions on proteins. Nucleic Acids Res. 2012;40:W452-7.
31. Sugita S, Ichtchenko K, Khvotchev M, Südhof TC. α-Latrotoxin Receptor CIRL/latrophilin 1 (CL 1) Defines an Unusual Family of Ubiquitous G-protein-linked Receptors: G-protein COUPLING NOT REQUIRED FOR TRIGGERING EXOCYTOSIS. J Biol Chem. 1998;273(49):32715-24.
32. Vetri L, Cali F, Vinci M, Amato C, Roccella M, Granata T, et al. A de novo heterozygous mutation in KCNC2 gene implicated in severe developmental and epileptic encephalopathy. Eur J Med Genet. 2020;63(4):103848.
33. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. Swissmodel: homology modeling of protein structures and complexes. Nucleic Acids Res. 2018;46(W1):W296-303.
34. Weston MD, Luijendijk MWJ, Humphrey KD, Möller C, Kimberling WJ. Mutations in the VLGR1 gene implicate G-protein signaling in the pathogenesis of Usher syndrome Type II. Am J Hum Genet. 2004;74(2):357-66.

35. Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res. 2007;35(Web Server issue):W407-10.
36. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER Suite: protein structure and function prediction. Nat Methods. 2015;12(1):7-8.

Tables

Table 1. Whole Exome Sequencing findings

	Gene			
	ADGRV1	*ADGRV1*	*ADGRV1*	*KCNC2*
Variant	c.3443G>A	c.8779G>A	c.14862_14866delACAAA	c.1408C>T
Protein	p.Gly1148Asp	p.Val2927Ile	p.Gln4955GlufsTer10	p.Pro470Ser
Type	Missense	Missense	Frameshift	Missense
Mother	X	X	X	
Child		X	X	X
ACMG Criteria	n/n	n/n	n/n	n/n
Reported Pathogenicity*	VUS	VUS	PP	VUS

Table 2. Model validation of predicted 3D protein structures

Region	Model	Verify3D	ProSA-web	ERRAT
ADGRV1				
Calx-beta 9 (amino acid 1108 to 1208)	SwissModel	47.52	-4.19	81.57
	I-Tasser	45.54	-4.88	93.54
	Robetta*	90.10	-5.54	88.64
Calx-beta 20-21 (amino acid 2814 to 3048)	SwissModel	69.33	-3.86	82.25
	I-Tasser	72.34	-4.34	85.13
	Robetta*	80.00	-4.83	85.71
KCNC2				
Amino acid 346 to 618	SwissModel	34.06	-1.81	91.53
	I-Tasser	45.79	-2.5	86.41
	Robetta *	79.12	-5.31	82.12

*Model selected for predict protein stability

Table 3. Protein stability prediction of models build by Robetta

Predictor \ Protein variant	ADGRV1		KCNC2
	p.Gly1148Asp	p.Val2927Ile	p.Pro470Ser
Maestro Web	ΔΔGpred 1.8104 cpred 0.799	ΔΔGpred 0.261 cpred 0.896	ΔΔGpred 1.533 cpred 0.859
mCSM-membrane (stability)	(ΔΔG): -0.219 Kcal/mol (Destabilizing)	(ΔΔG): -0.29 Kcal/mol (Destabilizing)	(ΔΔG): -0.259 Kcal/mol (Destabilizing)
mCSM-membrane (pathogenicity)	Pathogenic	Pathogenic	Pathogenic
CUPSAT (Stability thermal)	Destabilizing Unfavorable -4.75	Stabilizing Favorable 0.94	Stabilizing Favorable 0.47
CUPSAT (Stability denaturants)	Destabilizing Unfavorable -0.06	Stabilizing Favorable 2.2	Stabilizing Favorable 0.98
Site Directed Mutator	-2.46 (Reduced stability)	-0,57 (Reduced stability)	0.49 (Increased stability)
PremPS	ΔΔG 1,37	ΔΔG 0,18	ΔΔG -0,25
Rhapsody	0.83 (deleterious)	0.42 (neutral)	0.60 (deleterious)
I-Mutant	Decrease DDG -1.22 Kcal/mol	Decrease DDG -0.72 Kcal/mol	Decrease DDG -1.23 Kcal/mol
DynaMut2	(ΔΔG Stability) -1.58 kcal/mol (Destabilizing)	(ΔΔG Stability) -0,46 kcal/mol (Destabilizing)	(ΔΔG Stability) -0.43 kcal/mol (Destabilizing)

Table 4. *In silico* analysis of *ADGRV1* G1148D and V2729I, and *KCNC2* P470S

Predictor	ADGRV1		KCNC2
	G1148D prediction (score)	V2729I prediction (score)	P470S prediction (score)
Polyphen-2	Probably damaging (1.000)	Probably damaging (0.976)	Probably damaging (1.000)
DANN	(0.9984)	(0.9972)	(0.9986)
EIGEN	Pathogenic (0.6389)	Pathogenic (0.5386)	Pathogenic (0.9336)
FATHMM-MKL	Damaging (0.9944)	Damaging (0.9881)	Damaging (0.9911)
MutationTaster	Disease causing (1.0)	Disease causing (0.9752)	Disease causing (1.0)
BayesDel noAF	Damaging (0.1649)	Tolerated (-0.4346)	Damaging: (0.2718)
LRT	Deleterious (0)	Neutral (0.00004)	Deleterious (0.000001999)
MetaLR	Damaging (0.5051)	Tolerated (0.1898)	Damaging (0.9573)
SIFT	Damaging (0)	Tolerated (0.098)	Tolerated (0.204)
SIFT4G	Damaging (0)	Tolerated (0.053)	Tolerated (0.186)

Figures

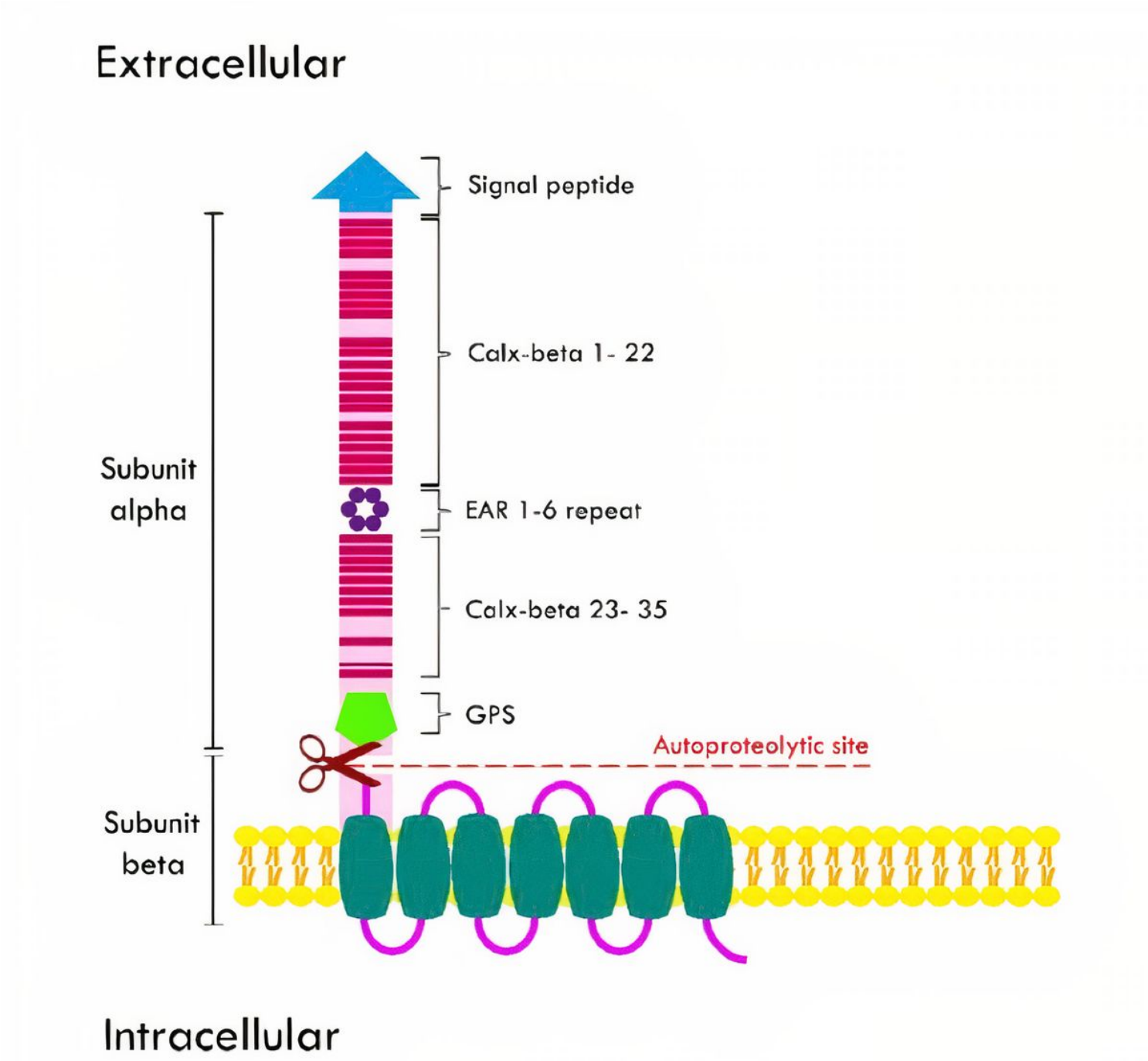

Figure 1

ADGRV1 protein structure.

Figure 2

Facial features of patients. The 36-year-old mother of two (a) presenting with high anterior hair implantation, upslanting palpebral fissures, and posteriorly rotated ears. Her 15-year-old son presenting (b) with microcephaly, long face, bilateral epicanthal folds, prominent maxillae, high set nasal root implantation, prominent nasal bridge, and retrognathia.

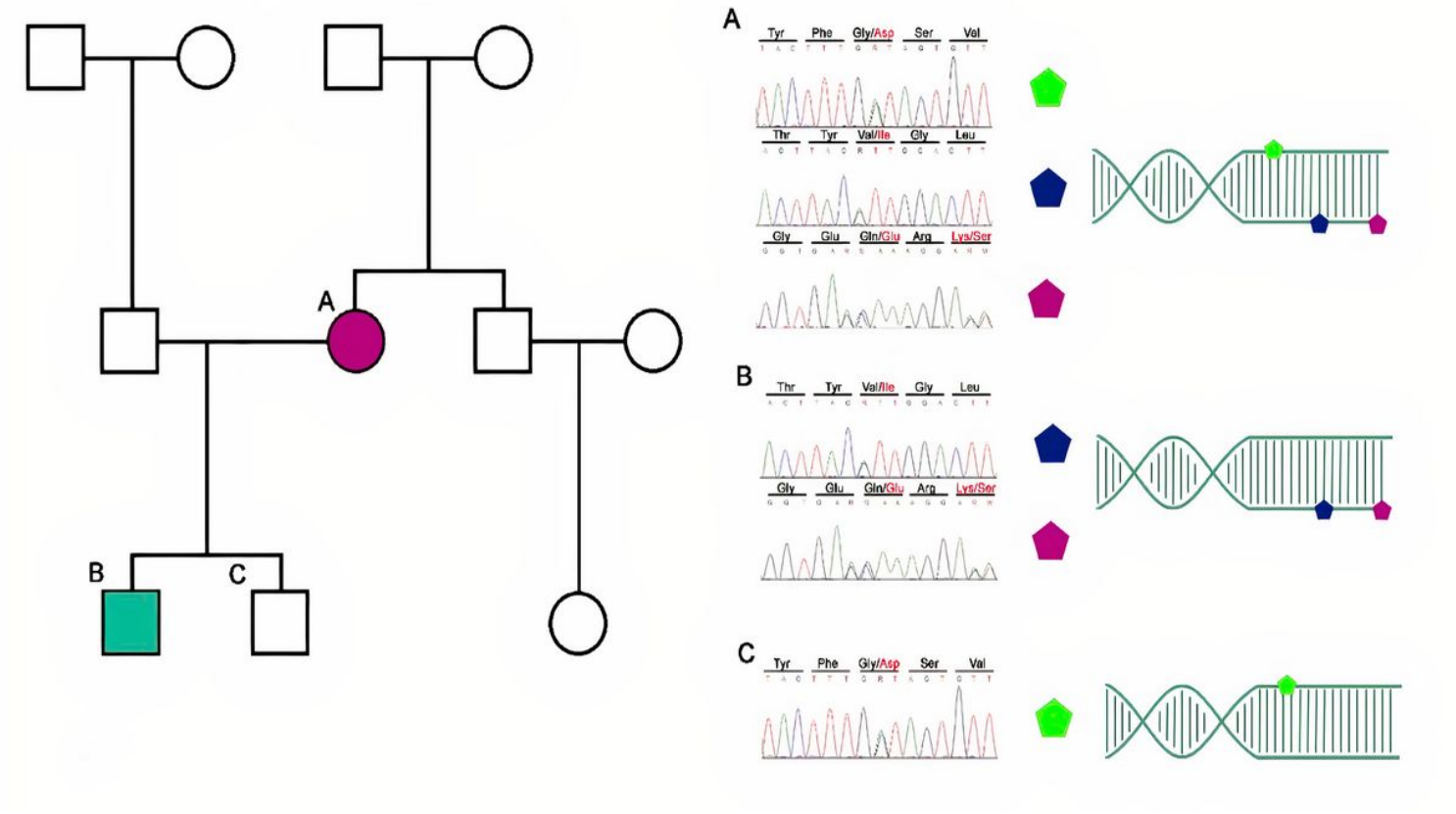

Figure 3

Familiar pedigree and Sanger sequencing electropherogram. A. Segregation of the identified ADGRV1 variants in the family. B. Sequence electropherograms of the mother (A), the affected child (B), and the healthy child (C). C. Representation of ADGRV1 variants; green pentagon indicates c.3443G>A variant, blue pentagon indicates c.8779G>A variant, and purple pentagon indicates c.14862_14866delACAAA variant.