


Article

Exploring the Genotype–Phenotype Correlations in a Child with Inherited Seizure and Thrombocytopenia by Digenic Network Analysis

Shuanglong Lu, Zhixiao Niu  and Xiaohong Qiao * 

Department of Pediatrics, Tongji Hospital, Tongji University School of Medicine, 389 Xincun Road, Shanghai 200065, China; 1400456@tongji.edu.cn (S.L.); 2131143@tongji.edu.cn (Z.N.)

* Correspondence: 1911596@tongji.edu.cn; Tel.: +86-21-66111045; Fax: +86-21-56050502

Abstract: Understanding the correlation between genotype and phenotype remains challenging for modern genetics. Digenic network analysis may provide useful models for understanding complex phenotypes that traditional Mendelian monogenic models cannot explain. Clinical data, whole exome sequencing data, in silico, and machine learning analysis were combined to construct a digenic network that may help unveil the complex genotype–phenotype correlations in a child presenting with inherited seizures and thrombocytopenia. The proband inherited a maternal heterozygous missense variant in *SCN1A* (NM_001165963.4:c.2722G>A) and a paternal heterozygous missense variant in *MYH9* (NM_002473.6:c.3323A>C). In silico analysis showed that these two variants may be pathogenic for inherited seizures and thrombocytopenia in the proband. Moreover, focusing on 230 epilepsy-associated genes and 35 thrombopoiesis genes, variant call format data of the proband were analyzed using machine learning tools (VarCoPP 2.0) and Digenic Effect predictor. A digenic network was constructed, and *SCN1A* and *MYH9* were found to be core genes in the network. Further analysis showed that *MYH9* might be a modifier of *SCN1A*, and the variant in *MYH9* might not only influence the severity of *SCN1A*-related seizure but also lead to thrombocytopenia in the bone marrow. In addition, another eight variants might also be co-factors that account for the proband’s complex phenotypes. Our data show that as a supplement to the traditional Mendelian monogenic model, digenic network analysis may provide reasonable models for the explanation of complex genotype–phenotype correlations.

Keywords: digenic network analysis; seizure; thrombocytopenia

Citation: Lu, S.; Niu, Z.; Qiao, X. Exploring the Genotype–Phenotype Correlations in a Child with Inherited Seizure and Thrombocytopenia by Digenic Network Analysis. *Genes* **2024**, *15*, 1004. <https://doi.org/10.3390/genes15081004>

Academic Editors: Albert Jeltsch, Yu An, Junyu Zhang, Yiran Guo and Hong Guo

Received: 23 June 2024

Revised: 24 July 2024

Accepted: 27 July 2024

Published: 31 July 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

A traditional Mendelian monogenic model often correlates the genotype with phenotype at a “single-disorder” paradigm, which means a single genetic condition accounts for all the clinical features of patients [1]. However, with the expanding use of next-generation sequencing (NGS), researchers realized that the real genetic mechanism of some diseases may be more complex than we thought. When phenotypes in a family pedigree do not fit a monogenic segregation pattern, exhibiting non-Mendelian inheritance, incomplete penetrance, or phenotypic variability, it is difficult for us to identify genes implicated in diseases. More genetic models are needed to clarify the complex genotype–phenotype correlations in such cases.

When two variants are implicated to describe the phenotypes of patients, there might be three different types of variant combinations. The first one is true digenic: The simultaneous presence of two pathogenic variants is required to produce the phenotypes of the disease; a variant at one of these loci alone does not result in the phenotype. The second one is monogenic plus modifier: a variant on the major gene induces a disease phenotype while another variant in the modifier gene may relieve or aggravate the phenotype. The third one is dual molecular: two variants at different loci are responsible for either distinct

or overlapping phenotypes for two different diseases [2–7]. Digenic network analysis can predict these candidate disease-causing variant combinations and distinguish digenic inheritance (variants at two different loci result in a simple or syndromic genetic disease) from dual molecular diagnosis (variants at two different loci result in two monogenic disorders), hence revealing us how different genetic conditions and molecular mechanisms could be combined in a single patient. Furthermore, it may also help us find more genetic targets for the treatment of some genetic diseases [8–10].

SCN1A, located at 2q24.3, encodes the α 1 subunit of the sodium channel and is one of the most important epilepsy genes in humans. Different pathogenic variants of *SCN1A* can have different effects on sodium channel function, which may lead to a wide range of phenotypes with different disease severities, including Dravet syndrome (DS, OMIM #607208), developmental and epileptic encephalopathy 6B (DEE, OMIM #619317), generalized epilepsy with febrile seizures plus type 2 (GEFS+, OMIM #604403), and familial hemiplegic migraine (FHM, OMIM #609634) [11]. However, the phenotypic variability in families with the same variant remains unexplained by a traditional Mendelian monogenic model [12,13].

MYH9, located at 22q12.3, encodes the heavy chain of non-muscle myosin IIA, an actin-binding molecular motor that is essential for development. Pathogenic variants of *MYH9* lead to an array of autosomal dominant disorders, including macrothrombocytopenia and granulocyte inclusions with or without nephritis or sensorineural hearing loss (OMIM #155100) and deafness, autosomal dominant 17 (OMIM #603622) [14]. The phenotypes of *MYH9*-related disease (*MYH9*-RD) are also variable among patients in the same family, with different severity and different organ involvement [15,16].

Here, we report a child with pathogenic missense variants in *SCN1A* and *MYH9* simultaneously, presenting with phenotypes of recurrent seizures and thrombocytopenia. By combining clinical data, whole exome sequencing data, in silico, and machine learning analysis, a digenic network map was constructed, which may provide a reasonable explanation for the complex genotype–phenotype correlations in this patient and his family.

2. Materials and Methods

2.1. Patient

Blood samples and clinical data were collected from the proband, his parents, his maternal aunt, and his cousin. Written informed consent was obtained from the proband's parents.

2.2. Whole Exome Sequencing (WES) and Sanger Validation

DNA was extracted from the blood samples using a Genomic DNA Extraction Kit (MyGenostics Kangwei century, Beijing, China). Extracted DNA was fragmented by sonication to an average size of 180 bp using a Bioruptor[®] (Diagenode, Seraing, Belgium) and then used as a template for library preparation. A P039-Exome probe was hybridized with a DNA library to capture the target gene by MyGenostics Technology Co., Ltd., Beijing, China. The enrichment libraries were sequenced on an Illumina HiSeq X ten sequencer for paired read 150 bp. The resulting reads were quality-controlled, trimmed, and mapped against the human reference genome hg19. Variants were further annotated by ANNOVAR, and some variants identified by WES were amplified from genomic DNA by polymerase chain reaction (PCR) and then sequenced by automatic genetic analyzer ABI 3730xl (Applied Biosystems, Foster City, CA, USA).

2.3. Variant Analysis and Interpretation

Variants were annotated through multiple databases, such as the 1000 Genomes, ESP6500, dbSNP, EXAC, in-house (MyGenostics), and HGMD, predicted by MutationTaster 2021 (<https://www.genecascade.org/MutationTaster2021> (accessed on 1th June 2024)), and classified according to the Association for Molecular Pathology/American College of Medical Genetics and Genomics (ACMG) guidelines [17]. Protein mutation analysis,

including the protein residue conservation of mutated positions across different species and the variant's possible impact on the protein's function, was carried out by the VarSite web server (<https://www.ebi.ac.uk/thornton-srv/databases/VarSite> (accessed on 1 June 2024)) [18]. Two core variants were submitted to the Leiden Open Variation Database (LOVD) with individual ID 00,448,671 (<https://databases.lovd.nl/shared/individuals/00448671> (accessed on 1 June 2024)).

2.4. Digenic Network Analysis

The variant call format (VCF) data annotated by ANNOVAR were submitted to the bioinformatics platform ORVAL (Oligogenic Resource for Variant AnaLysis, <https://orval.ibsquare.be> (accessed on 1 June 2024)) [19] for the exploration of disease-causing variant combinations. The genome version was set as GRCh37/hg19, and the patient information was set as male. Variant filtering was set at a Minor Allele Frequency (MAF) lower than or equal to 3.5% in protein-coding genes to remove intergenic and intronic/synonymous variants. Gene sets containing 230 epilepsy-associated genes [20] and 35 thrombopoiesis genes [21] reported so far were used as gene-filtering panels to ensure that the analysis was consistent with the patient's phenotype.

The machine learning tools Variant Combination Pathogenicity Predictor (VarCoPP 2.0) [22] and the Digenic Effect predictor [23] based on the Digenic Diseases Database (DIDA) [24] and Oligogenic Diseases Database (OLIDA) [25] were used to predict bi-locus variant combinations pathogenicity and digenic combination effects. Gene pair pathogenicity score greater than 0.891 was set as the network filter, and the final digenic network was constructed. VarCoPP 2.0 is a balanced Random Forest (RF) predictor that consists of 400 decision trees. It has been trained on the pathogenic variant combinations present in the OLIDA database against a large subset of variant data derived from control individuals of the 1000 Genomes Project (1KGP). Based on the variant pairs' biological features, including Combined Annotation Dependent Depletion (CADD) raw score, gene haploinsufficiency prediction, inheritance-specific pathogenicity prediction, selective pressure, biological distance, biological process similarity, and distance in an in-house developed Knowledge Graph, VarCoPP 2.0 can predict the pathogenicity score (value between 0 and 1) of any bi-locus variant combination. If the gene pair pathogenicity score was greater than 0.891 (genome version hg19), it had a 99.9% probability of being a true positive disease-causing variant combination. The Digenic Effect predictor is a classification Random Forest (RF) predictor trained on 240 pathogenic variant combinations. It uses the variant pairs' biological features, including the CADD raw score, gene recessiveness probability, essential in mice, and the same pathway to predict digenic combination effects, including true digenic combinations, monogenic plus modifier combinations, and dual-molecular diagnosis combinations.

2.5. Protein–Protein Interactions and Pathways Analysis

Those 10 genes passing the >0.891 gene pair pathogenicity score filter were employed for the construction of the digenic network and subsequently uploaded to the STRING database (version 12.0; <https://string-db.org> (accessed on 1 June 2024)) [26] to explore Protein–Protein Interactions, and the interaction type was set as the full STRING network, with the minimum required interaction score set as medium confidence (0.40). These 10 genes were also uploaded to the Reactome Pathway Knowledgebase 2024 (<https://reactome.org> (accessed on 1 June 2024)) [27] to explore the biological pathways of these proteins, and the main and common biological pathways of key genes were mapped.

2.6. Gene Expression and Co-Expression Analysis

To illustrate the expression and co-expression profiles of 10 genes passing the >0.891 gene pair pathogenicity score filter in our digenic network, the transcript expression levels of them in different tissues, based on transcriptomic data from the Human Protein Atlas (HPA) RNA-seq data and Genotype-Tissue Expression (GTEx) project data [28] were

downloaded from the HPA database (<https://www.proteinatlas.org> (accessed on 1 June 2024)). The consensus-normalized expression (nTPM) values of the genes in the cerebral cortex and bone marrow were mapped.

GEPIA2 [29], a web server for large-scale expression profiling and interactive analysis, was used for the co-expression analysis. Pearson correlations of these 10 gene expressions in the human brain of the GTEx database were calculated, and a co-expression matrix was constructed.

3. Results

3.1. Clinical Observations

A 4-year-old boy was referred to our department with a sudden, unexpected onset of fever-associated repeated seizures, without any previous history of seizures, together with a blood finding of thrombocytopenia. Recurrent seizures with fever were observed and manifested as generalized tonic-clonic seizures that lasted up to 10 min. Head computed tomography (CT) showed only a slightly widened supratentorial brain ventricle and extracerebral space (Figure 1D). Electroencephalography (EEG) and cerebrospinal fluid examination results were normal, and the antiepileptic drug levetiracetam was used to control the seizures.

The initial leukocyte count was low at $2.9 \times 10^9/L$, including 60.9% lymphocytes and 31.1% neutrophils. Hemoglobin (Hb) was low at 8.5 gm/dL, and his platelet count was very low at $36 \times 10^9/L$, indicating thrombocytopenia. Bone marrow smear showed hypocellular marrow with normal erythroid and granulocytic precursors and decreased megakaryocytes. Antiplatelet antibodies and platelet-associated immunoglobulins were not detected. HB and leukocyte count recovered to normal after the fever had settled with antibiotics; however, the platelet count remained low, ranging from 30 to $50 \times 10^9/L$. His peripheral blood smear showed Döhle body-like cytoplasmic inclusion bodies in neutrophils (Figure 1D) and some giant platelets, which might be signs of *MYH9*-related disorder.

No special symptoms or signs, including skeletal abnormalities, microcephaly, mental retardation, or visual or hearing difficulties, were observed. Other laboratory data, including liver and kidney function, electrolytes, coagulation, myocardial enzymes, and echocardiography, were normal.

Family history (Figure 1A) revealed that the boy's mother had a single seizure lasting for about 15 s at the age of 28 years old and did not have fever-associated convulsions as a child. The boy's father had no history of seizures. The boy's maternal aunt had a single fever-associated convulsion lasting for about 10 s at the age of 5 years old and had no subsequent recurrence. The boy's maternal cousin had a single fever-associated convulsion lasting for about 10 s at the age of 1 year old and had no subsequent recurrence. All the seizures of the boy, his mother, his maternal aunt, and his cousin were manifested as generalized tonic-clonic, but the boy's seizures were more frequent and lasted much longer than his mother, maternal aunt, and cousin (Table 1).

The boy's mother had no family history of hematological disease. The boy's father once had mild thrombocytopenia ($82 \times 10^9/L$) when he had a respiratory infection, and no other special symptoms were found in him. The relatives of the boy's father had no symptoms of thrombocytopenia.

The boy had been followed up in our department for about 4 years. After the treatment of levetiracetam, the boy's seizure had been well controlled, only reoccurred once for 5 min in the first year of disease onset, and had no subsequent recurrence. Eltrombopag was given to treat his thrombocytopenia, but his platelet counts were constantly low (ranging from 30 to $80 \times 10^9/L$).

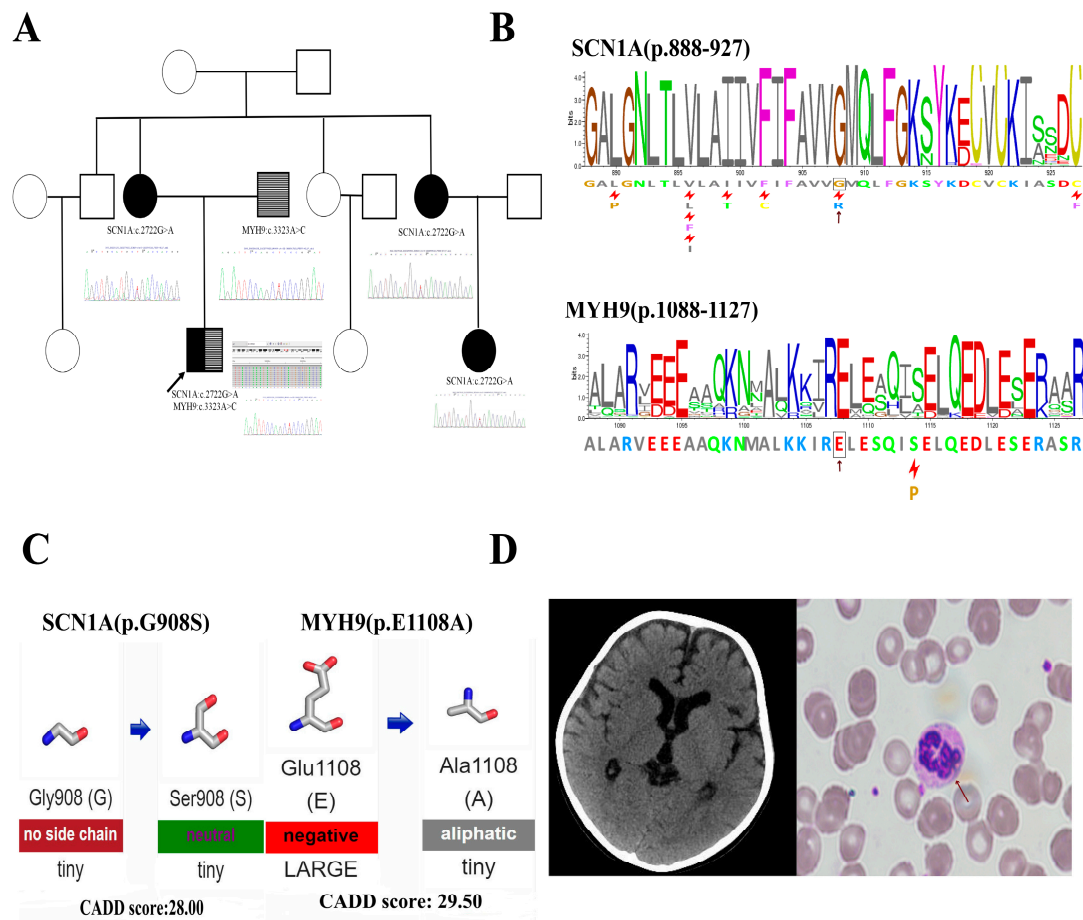


Figure 1. Genetical analysis and clinical data of the patient. (A) Pedigree diagram of patients. The black arrow indicates the proband, the genotypes, and the Sanger sequencing results were placed below or beside each family member; the black filling represents seizure phenotype, and the horizontal stripe filling represents thrombocytopenia phenotype. (B) Forty residue sequence logos centered on Gly908 of SCN1A and Glu1108 of MYH9. Residue conservations at each position from multiple alignments of homologous proteins were depicted; the taller the letter, the more commonly it is found at that position, and residues marked by the red lightning bolts are known pathogenic variants at that position. (C) The property changes of the sidechain for SCN1A (p.G908S) AND MYH9 (p.E1108A). The hydrophobicity, the size of the sidechain, and the Combined Annotation Dependent Depletion (CADD) scores of variants were depicted. (D) Head computed tomography (CT) scan image of the patient and the Romanowsky–Giemsa staining of the patient’s peripheral blood smear. The red arrow indicates Döhle body-like cytoplasmic inclusion bodies in the neutrophils.

Table 1. Clinical features of the proband and his affected family members.

| Person | Gender | Onset Age (Years) | Type of Seizure | Age Range of Seizures (Years) | Fever/Temp (°C) | Duration of Longest Seizure | Duration of Episode/How Many Seizures Per Episode | Frequency of Episodes | On Treatment at the Time |
|---------|--------|-------------------|--------------------------|-------------------------------|-----------------|-----------------------------|---|-----------------------|--------------------------|
| Patient | M | 4 | Generalized tonic-clonic | 4~5 | Yes/39.8 | 10 min | 5 days/20+ times | Twice in one year | Levetiracetam |
| Mother | F | 28 | Generalized tonic-clonic | 28 | Yes/39.5 | 15 s | Once | – | Supportive treatment |
| Aunt | F | 5 | Generalized tonic-clonic | 5 | Yes/– | 10 s | Once | – | Supportive treatment |
| Cousin | F | 1 | Generalized tonic-clonic | 1 | Yes/– | 10 s | Once | – | Supportive treatment |

3.2. WES and Variants Interpretation

WES and Sanger sequencing were performed, and two possible pathogenic variants were identified. The proband had inherited a heterozygous missense variant in *SCN1A*

(NM_001165963.4:c.2722G>A) from his mother (Figure 1A). His maternal aunt and cousin also had this variant (Figure 1A). Evidence analysis showed that the variant located in the hotspot region of the *SCN1A* gene (Pathogenic Moderate 1, PM1) was absent from the normal control database (Pathogenic Moderate 2, PM2), changed an amino acid residue where different pathogenic changes had been reported previously (Pathogenic Moderate 5, PM5), and multiple computational evidence supports it as a deleterious variant (Pathogenic supporting 3, PP3). Protein mutation analysis showed that this variant can cause a Glycine-to-Serine residue change at position 908 of the *SCN1A* protein. Position 908 is highly conserved (Figure 1B, conservation = 1.0, from 200 aligned protein sequences) and was predicted to be very important for the function of the protein. In addition, Glycine is an amino acid with no side chain, and Serine has a neutral side chain; changes from Glycine to Serine may change the flexibility and hydrophobicity of the protein in this position, and the CADD score for this variant is 28, indicating that it might be a deleterious variant (Figure 1C). Finally, the missense variant in *SCN1A* (NM_001165963.4:c.2722G>A) in the proband was classified to be Likely Pathogenic according to the ACMG guideline (PM1 + PM2 + PM5 + PP3).

WES and the Sanger sequencing revealed that the proband inherited another heterozygous missense variant in *MYH9* (NM_002473.6:c.3323A>C) from his father (Figure 1A). Evidence analysis showed this variant had a damaging effect on the gene and protein in vivo (Pathogenic Strong 3, PS3), which was absent from the normal control database (Pathogenic Moderate 2, PM2). Missense variants in *MYH9* have a low rate of benign missense variation. Missense variants are a common mechanism of disease for *MYH9* (Pathogenic supporting 2, PP2), and multiple computational evidence supports it being a deleterious variant (Pathogenic supporting 3, PP3). Protein mutation analysis showed that this variant can cause a Glutamic acid-to-Alanine residue change at position 1108 of the *MYH9* protein; position 1108 is also highly conserved (Figure 1B; conservation = 0.8 from 200 aligned protein sequences). In addition, Glutamic acid is a negatively charged hydrophilic amino acid, and Alanine has an aliphatic hydrophobic side chain. Changes from Glutamic to Alanine may change the hydrophobicity of protein in this position. The CADD score for this variant is 29.5, which means it might also be a deleterious variant (Figure 1C). Finally, the missense variant in *MYH9* (NM_002473.6:c.3323A>C) in the proband was classified to be Likely Pathogenic according to the ACMG guideline (PS3 + PM2 + PP2 + PP3).

3.3. Digenic Network and Combinations Annotation

Focused on 230 epilepsy-associated genes [20] and 35 thrombopoiesis genes [21] (Supplementary Table S1), a digenic network for the patient was constructed by machine learning tool VarCoPP 2.0 (Supplementary Figure S1). A total of 3203 variant combinations in 47 genes were found in the network, of which 2903 combinations were predicted to be neutral; a total of 233 combinations were predicted to be candidate disease-causing with low confidence, 45 combinations were predicted to be disease-causing with 99% confidence, and 22 combinations were predicted to be disease-causing with 99.9% confidence. *SCN1A* and *MYH9* were core genes in the network, with the highest network centrality of 29 (Supplementary Table S2).

To achieve the highest true positive probability results, the gene pair pathogenicity score filter was set at >0.891, and a final digenic network containing 10 genes and disease-causing variant combinations with 99.9% confidence was constructed, including 9 epilepsy-associated genes and 2 thrombopoiesis genes (*DIAPH1* was shared by both phenotypes).

Meanwhile, the Digenic Effect predictor revealed that the phenotype of *SCN1A* might be modified by *MYH9* and *SCN3A* and that the phenotype of *MYH9* might be modified by *DIAPH1* in the network (Figure 2A).

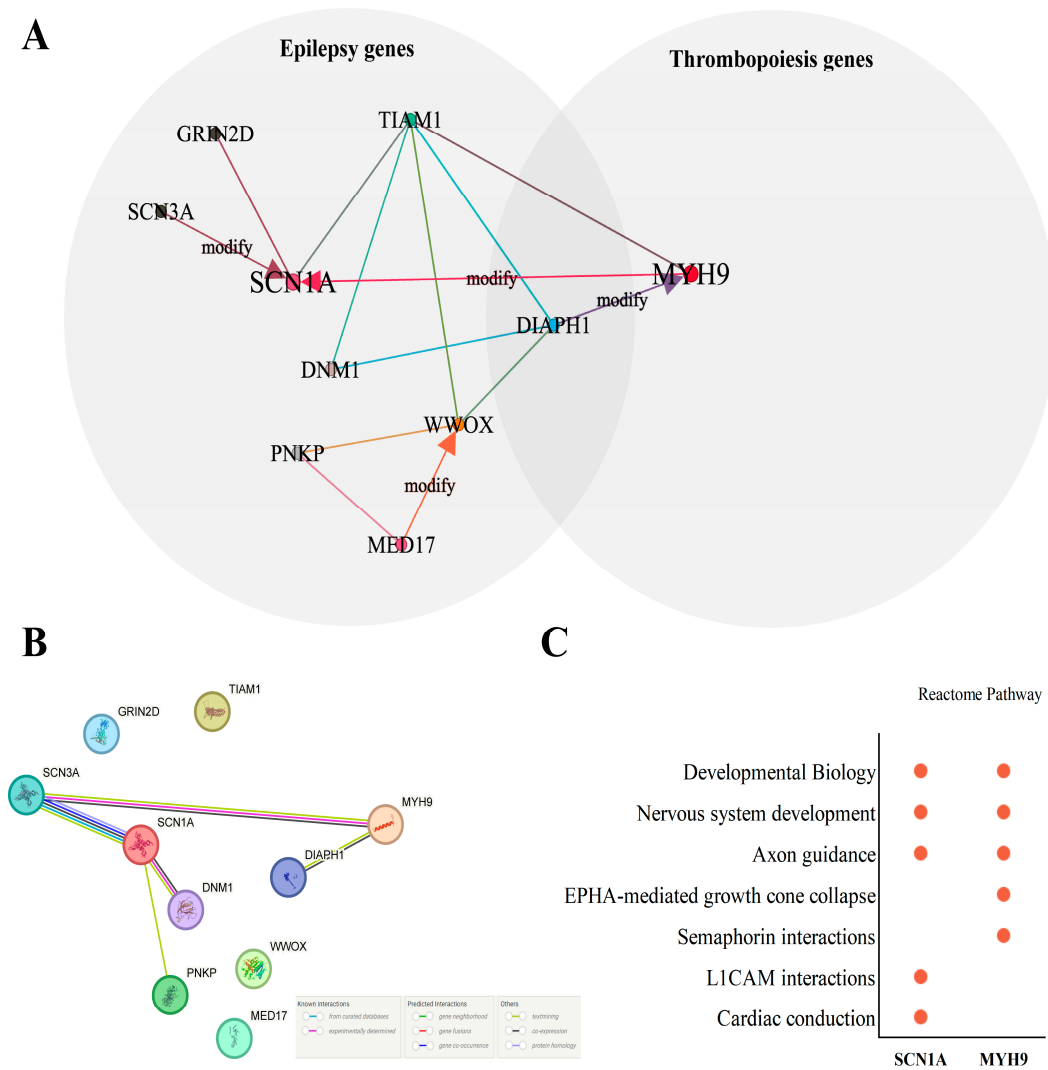


Figure 2. Digenic network of patient and core genes’ relationship analysis. **(A)** Digenic network and Venn map of related genes. Two shadow circles represent types of related genes, lines represent two genes that had digenic combinations with 99.9% true positive probability confidence. The arrow of the line indicates one gene might be a phenotype modifier of another gene. **(B)** Protein–Protein Interactions (PPIs) network of 10 genes in the digenic network. Lines represent the interaction between proteins. **(C)** Reactome pathways of SCN1A and MYH9. The red dot represents the pathways the gene is involved.

3.4. Protein–Protein Interactions and Pathways Analysis of Core Genes

Because there were different relationship types between pairs of genes carrying variants causing digenic disease, including direct interaction, indirect interaction, pathway membership, co-expression, and similar functions [24], to better understand the relationship of gene pairs in the patient’s digenic network, Protein–Protein Interactions (PPIs) network of 10 genes was constructed, which revealed that MYH9 might have indirect interaction with SCN1A through SCN3A. Meanwhile, DIAPH1 was found to have a direct interaction with MYH9, which may help its phenotype modifier function in MYH9 (Figure 2B).

In addition, Reactome Pathway analysis of the 10 genes showed that SCN1A and MYH9 had common pathways: developmental biology–nervous system development–axon guidance (Supplementary Table S3 and Figure 2C).

3.5. Expression, Co-Expression Profiles of Genes, and Digenic Model for Patient

Because most protein-encoding genes were affected by their expression level, to better understand the relationship of gene pairs, the expression profiles of 10 core genes based on HPA RNA-seq data and the GTEx project data were depicted. Nearly all 10 genes have different degrees of expression in the human cerebral cortex; however, only *MYH9*, *DIAPH1*, and *PNKP* have relatively high expression levels in the human bone marrow (Figure 3A). The co-expression matrix in the human brain shows that the expression of all the genes is positively correlated (Figure 3B), including the gene pair *SCN1A* and *MYH9*, which has a Pearson correlation coefficient of 0.17 ($p = 6.5 \times 10^{-9}$) (Figure 3C).

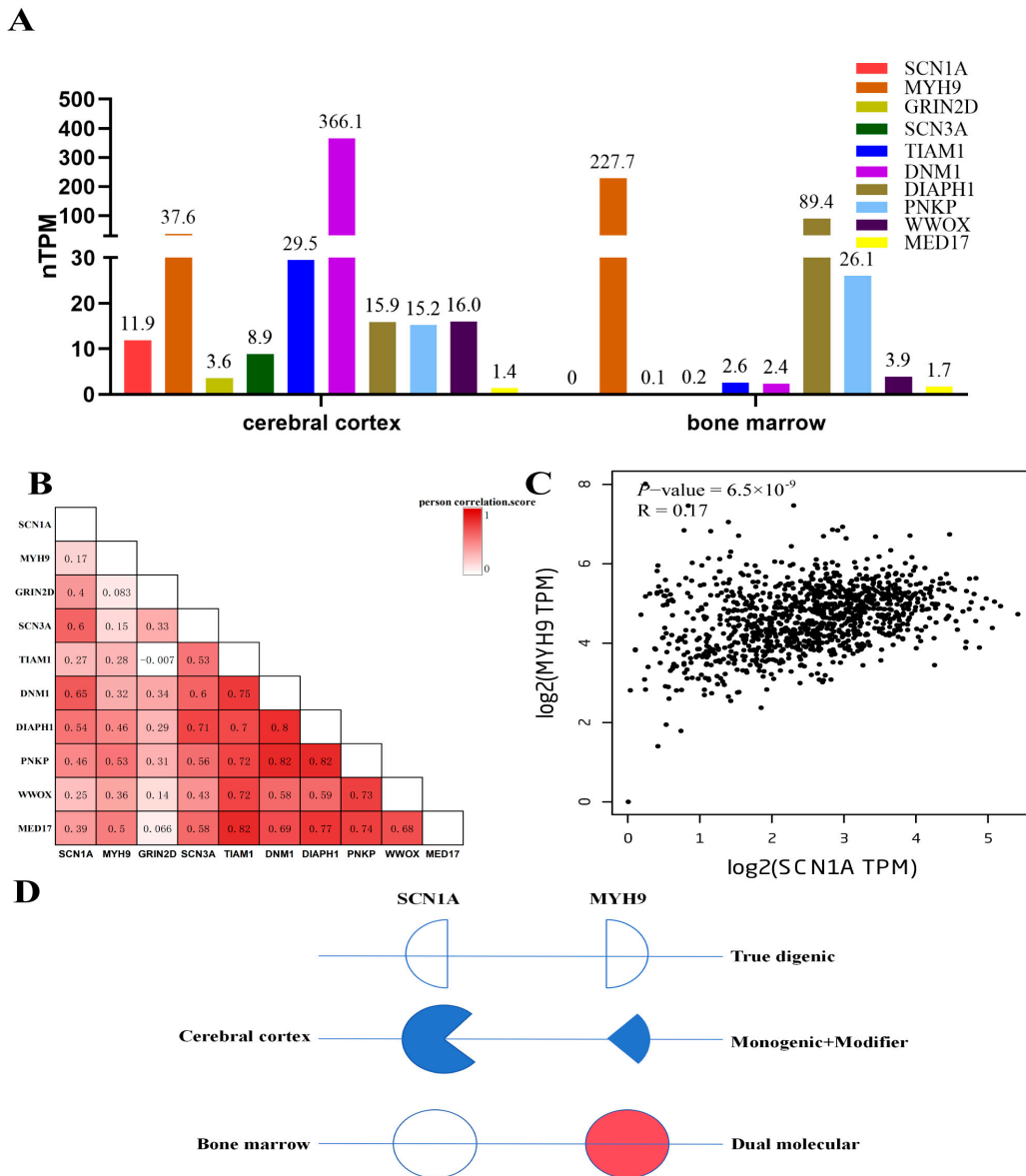


Figure 3. Expression, co-expression profiles of genes, and digenic model for patient. (A) Expression profiles of 10 genes in human cerebral cortex and bone marrow. (B) Co-expression matrix of 10 genes in the human brain. The numbers inside the box were the value of the Pearson correlation coefficient, and the color depth represents the correlation degree. (C) Co-expression map of *SCN1A* and *MYH9* in the human brain. (D) Digenic model for explanation of recurrent seizure and thrombocytopenia in the patient. Three types of variant combinations were depicted. Color fillings represent genes that have pathogenic effects, blue represents effects on seizure, and red represents effects on thrombocytopenia.

Based on the indirect interaction between *MYH9* and *SCN1A*, common pathways, and co-expression, a digenic model was built to explain recurrent seizures and thrombocytopenia in the patient. In the cerebral cortex, the seizure phenotype of *SCN1A* may be modified by *MYH9*, whereas in the bone marrow, *MYH9* may have an independent dual molecular effect, leading to thrombocytopenia (Figure 3D).

4. Discussion

The ultimate goal of medical genetics is to reveal the complex genotype–phenotype relationships of genetic disorders. Based on the philosophic tenet of Occam’s razor (the simplest diagnosis is the most likely to be correct) [30], medical geneticists usually use a single genetic diagnosis to explain all the phenotypes of the patient; however, in this “single-disorder” paradigm, the phenotypes do not always fit into the genotype perfectly. With the wide application of WES, the principle of Occam’s razor has been challenged [31,32], and different variant combination models, including true digenic, monogenic plus modifier, and dual molecular models (Figure 3D) have been used to explain complex phenotypes [33–36]. Here, we reported a patient with complex phenotypes that cannot be diagnosed by the “single-disorder” paradigm. Digenic network analysis was applied, and a reasonable digenic model was established to explain the patient’s seizures and thrombocytopenia.

For the phenotype of recurrent seizure, a maternal heterozygous missense variant in *SCN1A* (NM_001165963.4:c.2722G>A) was found in the patient and was classified to be likely pathogenic according to ACMG guideline, combined with clinical manifestations, the patient and his mother, aunt, and cousin can all be diagnosed as *SCN1A*-related GEFS+ (OMIM #604403). However, the frequency and duration of seizures in the patient were more frequent and much longer than in his mother, aunt, and cousin. Such intra-familial variability in phenotype severity has been observed in some other studies [12,13,37,38] and was explained by mosaicism, variants in regulatory regions, clinical management, or modifier genes. Some genes have been identified as modifiers of *SCN1A*-related seizure before [39–41], but all of these were epilepsy-related genes; no *MYH9*-modified seizure has been reported before.

Our study identified a paternal heterozygous missense variant in *MYH9* (NM_002473.6:c.3323A>C), which was previously considered as a reason for the phenotype of macrothrombocytopenia and granulocyte inclusions, with or without nephritis or sensorineural hearing loss (OMIM #155100). After expression, a PPI network, and pathway analysis, we found that in the cerebral cortex of humans, *MYH9* interacted with *SCN1A* indirectly [42], was co-expressed with *SCN1A*, and had common biological pathways with *SCN1A* for nervous system development [43], especially for modeling synaptic plasticity [44,45]. Hence, we propose that in the cerebral cortex of the patient, the paternal variant in *MYH9* might be a modifier of *SCN1A* and might be an explanation for the seizure phenotype variability observed between the patient and his mother, aunt, and cousin. The *SCN1A*-*MYH9* variant combination in the cerebral cortex fits in the second type of variant combination (monogenic plus modifier).

Meanwhile, for the phenotype of thrombocytopenia, we found that in the bone marrow, the expression of *SCN1A* is very low, indicating that *SCN1A* might have little effect on the patient’s hematological system phenotypes. The thrombocytopenia of the patient can only be explained by the variant of *MYH9*, which means the variant in *SCN1A* accounts for the seizure phenotype, and the variant in *MYH9* leads to thrombocytopenia in the bone marrow of the patient independently. The *SCN1A*-*MYH9* variant combination in bone marrow may also fit in the third type of variant combination (dual molecular).

Additionally, a variant in the epilepsy- and thrombopoiesis-associated gene *DIAPH1* was also identified in our digenic network, but it was not a core gene in our digenic network (gene centrality score was far less than *SCN1A* and *MYH9*, details in Supplementary Table S2) and the patient’s variant in *DIAPH1* was classified as uncertain according to the ACMG guideline. The patient’s symptoms cannot be explained only by the phenotype terms of *DIAPH1* (deafness, autosomal dominant 1, with or without thrombocytopenia,

OMIM #124900, seizures, cortical blindness, microcephaly syndrome, OMIM #616632). Hence, we propose it might be only a modifier gene for the patient's phenotypes in our digenic network, and it may affect *SCN1A*- and *MYH9*-related phenotypes directly or indirectly (Figure 2A). Similarly, another seven genes were identified in our digenic network, which were all classified as uncertain according to the ACMG guidelines; however, as members of the network, they may also be co-factors that account for the proband's complex phenotypes.

Our study had some limitations. First, we only focused on point variants in the patient. Other genetic conditions, such as copy number variation (CNV) and epigenetic changes, may also be involved in the patient's complex phenotypes. This will hinder the digenic model's further application. More database and machine learning tools about CNV and epigenetic changes are needed to improve and validate the digenic model. Second, due to limited conditions, the full pedigree chart of the proband cannot be obtained. This may hinder our interpretation of the digenic network. Third, although the machine learning tool VarCoPP 2.0 predicted the variant combinations in our network and had a 99.9% probability of being a true positive disease-causing variant combination, there were still possibilities of coincidence. The coexistence of two different monogenic disorders cannot be excluded. More experimental studies are needed in the future to validate the disease-causing model of our variant combinations. Similarly, although the Digenic Effect predictor predicted our digenic combination effects into a monogenic plus modifier model, we can only speculate that the modifier effect of *MYH9* may be the reason for the intra-familial variability of phenotype; however, other genetic and environmental factors may also attribute to the severity variability of epilepsy within the family, so more experimental studies are also needed to analyze the functional interaction of *MYH9* and *SCN1A* to support the modifier effect of *MYH9* in our network.

In conclusion, from our digenic network, we proposed that in the cerebral cortex, the variant of *MYH9* might work as a modifier of *SCN1A* and, in this family, led to intra-family seizure phenotype variability, whereas in the bone marrow, the variant of *MYH9* has probably led to thrombocytopenia. Also, another eight genes might be co-factors for our patient's complex phenotypes. Although our digenic network provided a reasonable model for the explanation of genotype–phenotype correlations in our patient, further evidence is needed to validate it in the future.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/genes15081004/s1>, Figure S1: The total digenic network of patient; Table S1: Gene sets for variants filtering; Table S2: Digenic network genes' centrality; Table S3: Reactome pathway of 10 genes in the digenic network.

Author Contributions: X.Q. conceived and designed the study, Z.N. collected the data, and S.L. analyzed the data and wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the ethics committee of Tongji Hospital (K-W-2024-002, 2024-03).

Informed Consent Statement: Written informed consent was obtained from the proband's parents.

Data Availability Statement: Data for this study are available upon request from the corresponding author.

Acknowledgments: We thank all staff at the Department of Pediatrics, Tongji Hospital, Tongji University School of Medicine. The graphical abstract was created with Figdraw.

Conflicts of Interest: The authors have no conflicts of interest to disclose.

References

1. Chong, J.X.; Buckingham, K.J.; Jhangiani, S.N.; Boehm, C.; Sobreira, N.; Smith, J.D.; Harrell, T.M.; McMillin, M.J.; Wiszniewski, W.; Gambin, T.; et al. The Genetic Basis of Mendelian Phenotypes: Discoveries, Challenges, and Opportunities. *Am. J. Hum. Genet.* **2015**, *97*, 199–215. [[CrossRef](#)]
2. Posey, J.E.; Harel, T.; Liu, P.; Rosenfeld, J.A.; James, R.A.; Akdemir, Z.H.C.; Walkiewicz, M.; Bi, W.; Xiao, R.; Ding, Y.; et al. Resolution of Disease Phenotypes Resulting from Multilocus Genomic Variation. *N. Engl. J. Med.* **2017**, *376*, 21–31. [[CrossRef](#)]
3. Gazzo, A.; Raimondi, D.; Daneels, D.; Moreau, Y.; Smits, G.; Van Dooren, S.; Lenaerts, T. Understanding mutational effects in digenic diseases. *Nucleic Acids Res.* **2017**, *45*, e140. [[CrossRef](#)]
4. Papadimitriou, S.; Gazzo, A.; Versbraegen, N.; Nachtegael, C.; Aerts, J.; Moreau, Y.; Van Dooren, S.; Nowé, A.; Smits, G.; Lenaerts, T. Predicting disease-causing variant combinations. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 11878–11887. [[CrossRef](#)]
5. Deltas, C. Digenic inheritance and genetic modifiers. *Clin. Genet.* **2018**, *93*, 429–438. [[CrossRef](#)]
6. Schäffer, A.A. Digenic inheritance in medical genetics. *J. Med. Genet.* **2013**, *50*, 641–652. [[CrossRef](#)]
7. Rahit, K.M.T.H.; Tarailo-Graovac, M. Genetic Modifiers and Rare Mendelian Disease. *Genes* **2020**, *11*, 239. [[CrossRef](#)]
8. van Heyningen, V.; Yeyati, P.L. Mechanisms of non-Mendelian inheritance in genetic disease. *Hum. Mol. Genet.* **2004**, *13*, R225–R233. [[CrossRef](#)]
9. Badano, J.L.; Katsanis, N. Beyond Mendel: An evolving view of human genetic disease transmission. *Nat. Rev. Genet.* **2002**, *3*, 779–789. [[CrossRef](#)]
10. Riordan, J.D.; Nadeau, J.H. From Peas to Disease: Modifier Genes, Network Resilience, and the Genetics of Health. *Am. J. Hum. Genet.* **2017**, *101*, 177–191. [[CrossRef](#)]
11. Scheffer, I.E.; Nabbout, R. SCN1A-related phenotypes: Epilepsy and beyond. *Epilepsia* **2019**, *60* (Suppl. S3), S17–S24. [[CrossRef](#)] [[PubMed](#)]
12. Passamonti, C.; Petrelli, C.; Mei, D.; Foschi, N.; Guerrini, R.; Provinciali, L.; Zamponi, N. A novel inherited SCN1A mutation associated with different neuropsychological phenotypes: Is there a common core deficit? *Epilepsy Behav.* **2015**, *43*, 89–92. [[CrossRef](#)] [[PubMed](#)]
13. Guerrini, R.; Cellini, E.; Mei, D.; Metitieri, T.; Petrelli, C.; Pucatti, D.; Marini, C.; Zamponi, N. Variable epilepsy phenotypes associated with a familial intragenic deletion of the SCN1A gene. *Epilepsia* **2010**, *51*, 2474–2477. [[CrossRef](#)] [[PubMed](#)]
14. Pecci, A.; Ma, X.; Savoia, A.; Adelstein, R.S. MYH9: Structure, functions and role of non-muscle myosin IIA in human disease. *Gene* **2018**, *664*, 152–167. [[CrossRef](#)] [[PubMed](#)]
15. Asensio-Juárez, G.; Llorente-González, C.; Vicente-Manzanares, M. Linking the Landscape of MYH9-Related Diseases to the Molecular Mechanisms that Control Non-Muscle Myosin II-A Function in Cells. *Cells* **2020**, *9*, 1458. [[CrossRef](#)] [[PubMed](#)]
16. Pecci, A.; Klersy, C.; Gresele, P.; Lee, K.J.; De Rocco, D.; Bozzi, V.; Russo, G.; Heller, P.G.; Loffredo, G.; Ballmaier, M.; et al. MYH9-related disease: A novel prognostic model to predict the clinical evolution of the disease based on genotype-phenotype correlations. *Hum. Mutat.* **2014**, *35*, 236–247. [[CrossRef](#)]
17. Richards, S.; Aziz, N.; Bale, S.; Bick, D.; Das, S.; Gastier-Foster, J.; Grody, W.W.; Hegde, M.; Lyon, E.; Spector, E.; et al. Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* **2015**, *17*, 405–424. [[CrossRef](#)]
18. Laskowski, R.A.; Stephenson, J.D.; Sillitoe, I.; Orengo, C.A.; Thornton, J.M. VarSite: Disease variants and protein structure. *Protein Sci.* **2020**, *29*, 111–119. [[CrossRef](#)] [[PubMed](#)]
19. Renaux, A.; Papadimitriou, S.; Versbraegen, N.; Nachtegael, C.; Boutry, S.; Nowé, A.; Smits, G.; Lenaerts, T. ORVAL: A novel platform for the prediction and exploration of disease-causing oligogenic variant combinations. *Nucleic Acids Res.* **2019**, *47*, W93–W98. [[CrossRef](#)]
20. Chi, W.; Kiskinis, E. Integrative analysis of epilepsy-associated genes reveals expression-phenotype correlations. *Sci. Rep.* **2024**, *14*, 3587. [[CrossRef](#)]
21. Kim, B. Diagnostic workup of inherited platelet disorders. *Blood Res.* **2022**, *57*, S11–S19. [[CrossRef](#)]
22. Versbraegen, N.; Gravel, B.; Nachtegael, C.; Renaux, A.; Verkinderen, E.; Nowé, A.; Lenaerts, T.; Papadimitriou, S. Faster and more accurate pathogenic combination predictions with VarCoPP2.0. *BMC Bioinform.* **2023**, *24*, 179. [[CrossRef](#)]
23. Versbraegen, N.; Fouché, A.; Nachtegael, C.; Papadimitriou, S.; Gazzo, A.; Smits, G.; Lenaerts, T. Using game theory and decision decomposition to effectively discern and characterise bi-locus diseases. *Artif. Intell. Med.* **2019**, *99*, 101690. [[CrossRef](#)] [[PubMed](#)]
24. Gazzo, A.M.; Daneels, D.; Cilia, E.; Bonduelle, M.; Abramowicz, M.; Van Dooren, S.; Smits, G.; Lenaerts, T. DIDA: A curated and annotated digenic diseases database. *Nucleic Acids Res.* **2016**, *44*, D900–D907. [[CrossRef](#)]
25. Nachtegael, C.; Gravel, B.; Dillen, A.; Smits, G.; Nowé, A.; Papadimitriou, S.; Lenaerts, T. Scaling up oligogenic diseases research with OLIDA: The Oligogenic Diseases Database. *Database* **2022**, *2022*, baac023. [[CrossRef](#)] [[PubMed](#)]
26. Szklarczyk, D.; Kirsch, R.; Koutrouli, M.; Nastou, K.; Mehryary, F.; Hachilif, R.; Gable, A.L.; Fang, T.; Doncheva, N.T.; Pyysalo, S.; et al. The STRING database in 2023: Protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res.* **2023**, *51*, D638–D646. [[CrossRef](#)] [[PubMed](#)]
27. Milacic, M.; Beavers, D.; Conley, P.; Gong, C.; Gillespie, M.; Griss, J.; Haw, R.; Jassal, B.; Matthews, L.; May, B.; et al. The Reactome Pathway Knowledgebase 2024. *Nucleic Acids Res.* **2024**, *52*, D672–D678. [[CrossRef](#)]
28. GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* **2013**, *45*, 580–585. [[CrossRef](#)]

29. Tang, Z.; Kang, B.; Li, C.; Chen, T.; Zhang, Z. GEPIA2: An enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acids Res.* **2019**, *47*, W556–W560. [[CrossRef](#)]
30. Blaser, S.; Schaye, V.; Hwang, J.; Cocks, P.; Kudlowitz, D. Hickam’s dictum, Occam’s razor, and Crabtree’s bludgeon: A case of renal failure and a clavicular mass. *Diagnosis* **2022**, *9*, 133–139. [[CrossRef](#)]
31. Linscott, K.B.; Cassady, J.A.; Robin, N.H. Occam’s razor dulled: The occurrence of multiple genetic diagnoses. *Curr. Opin. Pediatr.* **2021**, *33*, 545–548. [[CrossRef](#)] [[PubMed](#)]
32. Spedicati, B.; Morgan, A.; Pianigiani, G.; Musante, L.; Rubinato, E.; Santin, A.; Nardone, G.G.; Faletra, F.; Girotto, G. Challenging Occam’s Razor: Dual Molecular Diagnoses Explain Entangled Clinical Pictures. *Genes* **2022**, *13*, 2023. [[CrossRef](#)] [[PubMed](#)]
33. Katsanis, N. The continuum of causality in human genetic disorders. *Genome Biol.* **2016**, *17*, 233. [[CrossRef](#)] [[PubMed](#)]
34. Rosina, E.; Pezzani, L.; Pezzoli, L.; Marchetti, D.; Bellini, M.; Pilotta, A.; Calabrese, O.; Nicastrò, E.; Cirillo, F.; Cereda, A.; et al. Atypical, Composite, or Blended Phenotypes: How Different Molecular Mechanisms Could Associate in Double-Diagnosed Patients. *Genes* **2022**, *13*, 1275. [[CrossRef](#)] [[PubMed](#)]
35. Papadimitriou, S.; Gravel, B.; Nachttegael, C.; De Baere, E.; Loeyls, B.; Vikkula, M.; Smits, G.; Lenaerts, T. Toward reporting standards for the pathogenicity of variant combinations involved in multilocus/oligogenic diseases. *Hum. Genet. Genom. Adv.* **2023**, *4*, 100165. [[CrossRef](#)] [[PubMed](#)]
36. Ameratunga, R.; Woon, S.-T.; Bryant, V.L.; Steele, R.; Slade, C.; Leung, E.Y.; Lehnert, K. Clinical Implications of Digenic Inheritance and Epistasis in Primary Immunodeficiency Disorders. *Front. Immunol.* **2017**, *8*, 1965. [[CrossRef](#)]
37. Suls, A.; Velizarova, R.; Yordanova, I.; Deprez, L.; Van Dyck, T.; Wauters, J.; Guerguelcheva, V.; Claes, L.; Kremensky, I.; Jordanova, A.; et al. Four generations of epilepsy caused by an inherited microdeletion of the *SCN1A* gene. *Neurology* **2010**, *75*, 72–76. [[CrossRef](#)]
38. Depienne, C.; Trouillard, O.; Gourfinkel-An, I.; Saint-Martin, C.; Bouteiller, D.; Graber, D.; Barthez-Carpentier, M.-A.; Gautier, A.; Villeneuve, N.; Dravet, C.; et al. Mechanisms for variable expressivity of inherited *SCN1A* mutations causing Dravet syndrome. *J. Med. Genet.* **2010**, *47*, 404–410. [[CrossRef](#)]
39. de Lange, I.M.; Mulder, F.; van’t Slot, R.; Sonsma, A.C.M.; van Kempen, M.J.A.; Nijman, I.J.; Ernst, R.F.; Knoers, N.V.A.M.; Brilstra, E.H.; Koeleman, B.P.C. Modifier genes in *SCN1A*-related epilepsy syndromes. *Mol. Genet. Genom. Med.* **2020**, *8*, e1103. [[CrossRef](#)]
40. Calhoun, J.D.; Hawkins, N.A.; Zachwieja, N.J.; Kearney, J.A. *Cacna1g* is a genetic modifier of epilepsy in a mouse model of Dravet syndrome. *Epilepsia* **2017**, *58*, e111–e115. [[CrossRef](#)]
41. Hawkins, N.A.; Kearney, J.A. *Hlf* is a genetic modifier of epilepsy caused by voltage-gated sodium channel mutations. *Epilepsy Res.* **2016**, *119*, 20–23. [[CrossRef](#)] [[PubMed](#)]
42. Dash, B.; Han, C.; Waxman, S.G.; Dib-Hajj, S.D. Nonmuscle myosin II isoforms interact with sodium channel α subunits. *Mol. Pain* **2018**, *14*, 1744806918788638. [[CrossRef](#)] [[PubMed](#)]
43. Ma, X.; Adelstein, R.S. The role of vertebrate nonmuscle Myosin II in development and human disease. *BioArchitecture* **2014**, *4*, 88–102. [[CrossRef](#)] [[PubMed](#)]
44. Newell-Litwa, K.A.; Horwitz, R.; Lamers, M.L. Non-muscle myosin II in disease: Mechanisms and therapeutic opportunities. *Dis. Model. Mech.* **2015**, *8*, 1495–1515. [[CrossRef](#)]
45. Javier-Torrent, M.; Saura, C.A. Conventional and Non-Conventional Roles of Non-Muscle Myosin II-Actin in Neuronal Development and Degeneration. *Cells* **2020**, *9*, 1926. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.