HHS Public Access

Author manuscript

Am J Med Genet A. Author manuscript; available in PMC 2023 September 01.

Published in final edited form as:

Am J Med Genet A. 2022 February ; 188(2): 556–568. doi:10.1002/ajmg.a.62554.

NMDA receptor genetics: The power of paralog homology and protein dynamics in defining dominant genetic variants.

Jacob G Charron[1,2], **Angel Hernandez**[3], **Stephanie M Bilinovich**[1,4], **Daniel L Vogt**[1,4], **Laura A Bedinger**[5], **Laurie H Seaver**[1,5], **Michael Williams**[1,4], **Seth Devries**[3], **Daniel B Campbell**[1,4], **Caleb P Bupp**[1,5], **Jeremy W Prokop**[1,4,6,*]

[1]Department of Pediatrics & Human Development, Michigan State University, Grand Rapids, MI. 49503

[2]Department of Biology, Calvin University, Grand Rapids MI 49506, USA

[3]Pediatric Neurology, Helen DeVos Children's Hospital, Grand Rapids, MI 49503, USA

[4]Center for Research in Autism, Intellectual, and other Neurodevelopmental Disabilities, Michigan State University, East Lansing, MI, 48824, USA

[5]Division of Medical Genetics, Helen DeVos Children's Hospital, Grand Rapids, MI 49503, USA

[6]Department of Pharmacology and Toxicology, Michigan State University, East Lansing, MI, 48824, USA

Abstract

Predicting genotype-to-phenotype correlations from genomic variants has been challenging, particularly for genes that have a complex balance of dominant and recessive inheritance for phenotypes. Variants in NMDA receptor components *GRIN1*, *GRIN2A*, and *GRIN2B* cause a myriad of dominant disease phenotypes, with the most common being epilepsy and autism spectrum disorder. Starting from the analysis of a Variant of Uncertain Significance (VUS, GRIN2A G760S), we realized the need for tools to map dominant variants for the components of the NMDA receptor. Some variants within *GRIN1*, *GRIN2A*, and *GRIN2B* exert dominant epilepsy and developmental delay, yet other amino acid variants are conserved and predicted to alter protein function but do not have dominant phenotypes. Common variant annotation tools are not powered to determine pathogenic dominant outcomes. To address this gap, we Integrated sequence and structural analyses for *GRIN1*, *GRIN2A*, and *GRIN2B.* Using this approach, we determined that paralog homology mapping and topology can segregate dominant variants, with an elevation of intermolecular contacts between the subunits. Further, demonstrating the general utility of our methodology, we show that 25 VUS within ClinVar also reach a dominant variant annotation, including the GRIN2A G760S variant. Our work suggests paralog homology and

[*]**Corresponding Author**: Jeremy W Prokop (jprokop54@gmail.com) (330)573-3176, 400 Monroe Ave NW, Grand Rapids MI 49503 USA.

Author Contribution

JGC, SMB, DLV, MW, DBC, JWP processed and analyzed data. AH, LAB, LHS, SD, CPB assessed patient records and clinical case review. AH, CPB and JWP oversaw project completion. JGC and JWP wrote the manuscript and compiled figures. All authors approved of the final manuscript version.

Conflicts of Interest

None of the authors have any conflicts of interest to declare.

protein topology as a powerful strategy within the receptor complex to resolve dominant genetic variants relative to variants that would fit a recessive inheritance, requiring two damaging variants. These strategies should be tested in additional dominant genetic disorders to determine the broader utility.

Introduction

The N-methyl-D-aspartate (NMDA) receptors are a class of neuronal ion channels regulated by glutamate [MacDermott et al., 1986]. The main core of the receptor consists of two molecules of the GluN1 subunit (*GRIN1*) and two of the GluN2 subunit (from either *GRIN2A*, *GRIN2B*, *GRIN2C*, or *GRIN2D*). The complex structure has provided remarkable insights into topology [Karakas and Furukawa, 2014], and evolutionary insights [Ryan et al., 2008] have yielded critical regions. Yet, insights into the role of all amino acids within the complex has lacked the high-resolution needed to rapidly segregate functional genetic variants needed for precision medicine.

Individuals with NMDA receptor dysfunction, primarily *GRIN1* (RefSeq NM_007327.3,UniProt Q05586), *GRIN2A* (NM_000833.4, Q12879), and *GRIN2B* (NM_000834.3, Q13224), can experience epilepsy, learning disorders, brain atrophy, and hyperkinetic movement disorders [Lemke et al., 2016] commonly beginning in infancy [Ohba et al., 2015]. Variants of *GRIN1*, *GRIN2A* and *GRIN2B* are significantly associated with Autism Spectrum Disorder (ASD) according to the 2020 SFARI database (Category 1 or 2), with >100 individuals carrying variants within these genes [Abrahams et al., 2013]. In some cases there have been reports that NMDA variants can follow familial inheritance for ASD without signs of epilepsy [Yoo et al., 2012]. The phenotypic spectrum of associated NMDA receptor variants is broad, both between the gene on which variants occur and within different regions of each gene [Yuan et al., 2015]. The ClinGen tools annotate *GRIN1*, *GRIN2A*, *GRIN2B*, and *GRIN2D* (NM_000836.2, O15399) as autosomal dominant genes, yet there is some evidence of recessive disorders within *GRIN1* [Lemke et al., 2016]. A survey of reported variants within the clinical ClinVar [Landrum et al., 2016] and the population gnomAD [Karczewski et al., 2020] tools identifies hundreds of variants within the NMDA receptors, yet many of these genetic variants remain classified as Variants of Uncertain Significance (VUS) or are not associated with clinical phenotypes. While assessing the GRIN2A G760S variant, it became clear that tools are needed to separate dominant variants from other variants that are predicted to be functional yet are not associated with dominant conditions. In this work, we have developed a sequence-to-structure dataset for NMDA receptor proteins that resolves an amino acid matrix for segregating functional outcomes, which can be used to identify dominant disease associated variants.

Results

ClinVar Variant Extraction

Clinical neurogenetic evaluation of a child with epilepsy identified a VUS in GRIN2A, denoted G760S, that required further characterization for potential pathogenicity. This

variant is not reported in any common sequencing database (ClinVar, gnomAD, TOPMed, or Geno2MP). Analysis of ClinVar indicates that a large portion of protein coding clinical variants are classified as Variants of Uncertain Significance (VUS, 55%), with 24% of variants annotated in the pathogenic to likely pathogenic groups (Figure 1A). These values are similar to the rest of the ClinVar database, which annotates ~45% of all variants to be VUS. As of January 2020, there were 1,489 ClinVar deposits for *GRIN1*, *GRIN2A*, *GRIN2B*, *GRIN2C* (NM_000835.4, Q14957), or *GRIN2D* with 611 protein coding changes (Figure 1B). *GRIN2A* ranks the highest in protein coding deposits with 291, while *GRIN2B* (223) and *GRIN1* (81) are well reported. *GRIN2C* (7) and *GRIN2D* (9) have few reported protein coding variants. *GRIN1* has an even annotation of pathogenic relative to VUS (37/34), while *GRIN2A* (53/180) and *GRIN2B* (53/120) are much heavier in VUS annotation, suggesting need for variant ranking and assessment. The majority of annotated ClinVar protein coding variants are those of epilepsy and seizure disorders with additional variants associated with intellectual disability (Figure 1C) matching expected phenotypes. It should be noted that ClinVar is not a powerful tool for phenotype extractions as many of the records lack included information. Additionally, we have extracted terms from ClinVar as they are listed and would encourage the community to change the annotation of disease for *GRIN2A* away from “mental retardation” and towards the more appropriate developmental delay / intellectual disability (DD/ID) terminology. With the large number of variants for the NMDA receptor, we began a systematic assessment of protein amino acids, integrating structural and evolutionary biology, to aid in the classifications of protein variant interpretations, including the *GRIN2A* G760S variant of interest.

Structural Amino Acid Mapping

Homology modeling was utilized to merge four known PDB structures (4PE5, 4TLL, 5UOW, 4TLM) [Karakas and Furukawa, 2014; Lee et al., 2014; Lü et al., 2017] into a single complex containing two subunits of GluN1 and GluN2. To simplify our molecular dynamic simulations (mds) into a single run we made one molecule of GluN2 GRIN2A and one GRIN2B. The complex was then built into a physiological environment containing protonation at pH 7.4, explicit water, a lipid membrane, and NaCl (Figure 2A). A total of 30 nanoseconds of mds was performed on this complex, resolving the average amino acid movement throughout (Figure 2B). The raw simulation data is available at https://doi.org/10.6084/m9.figshare.14991558. The lower the movement of an amino acid the more stabile the amino acids, while high movement corresponds to dynamic loops and regions (Figure 2B). The proteins all show relatively similar average movement of amino acids with GRIN2B having the highest per amino acid root mean squared fluctuation (RMSF) of 2.73Å followed by GRIN1 (2.49Å) and then GRIN2A (2.36Å). Additionally, the trajectory of each amino acid was correlated to all the other amino acids, calculating dynamic cross correlation matrix (DCCM) sites for intra and inter protein dynamics (Figure 2C). All three of the proteins have similar per amino acid correlations with GRIN1 the highest (48.43 amino acids per residue), followed by GRIN2B (44.46) and GRIN2A (38.81) with 99% of amino acids correlating in movement to another amino acid. For protein-protein correlations, a surprisingly elevated GRIN2B intermolecular correlations were observed, averaging 4.66 amino acid correlations per residue (GRIN1 = 1.79, GRIN2A = 1.77) with GRIN2B having 27.28% of amino acids correlating to amino acids within another protein (GRIN1 = 20.83%,

GRIN2A = 17.72%). Overall, we suggest a remarkable overlap in molecular movement between the three proteins of the NMDA receptor complex, bringing insights into structure for all of the amino acids of GRIN1, GRIN2A, and GRIN2B that can be used to screen genomic variants.

Evolutionary Amino Acid Mapping

To map conservation and functionality we employed a high-density map of conservation across open reading frame sequences for *GRIN1* (133 species sequences), *GRIN2A* (199), *GRIN2B* (181), *GRIN2C* (152), and *GRIN2D* (54) for a total of 719 sequences used. All the raw data for fasta alignments can be found at https://doi.org/10.6084/m9.figshare.14991537. The evolution of the GluN2 subunits suggest highest divergence of *GRIN2B* from the three others, with *GRIN2D* likely arising from a split from *GRIN2C* (Figure 3A). Using sequences for each of the genes we calculated codon selection and amino acid conservation at each amino acid on a scale of 0–2, based on our previous metrics [Prokop et al., 2017, 2018] where a value of 2 means the amino acid is 100% conserved with >2 standard deviations of codon selection based on dN-dS statistics. With a map of conservation for each of the genes for every amino acid, we built a 21-codon sliding window of conservation, a sum of ten amino acids before and after each amino acid added together (Figure 3B). This metric allows us to identify highly conserved linear motifs within each gene, bringing conservation motif scores for each amino acid, and thus variant, analyzed. Finally, we built conservation scores for all NMDA receptor proteins and from GluN2 subunits. Aligning all 719 sequences we calculated conservation of GRIN1 amino acids relative to all sequences on a scale of 0 to 2 (Figure 3C). Six total amino acids have a score of 2 (499, 639, 683, 732, 734, 815) that is indicative of 100% conservation and >2 standard deviations of codon selection. An additional 17 and 109 amino acids have a score of between 1.5 and 1.25 suggestive of 100% conservation and codon selection between 0–2 standard deviations. For the GluN2 subunits we calculated the same conservation of each gene relative to all 719 sequences while also calculating amino acids conserved in the 586 GluN2 sequences (removing the 133 *GRIN1* sequences) and those amino acids only conserved within the individual gene (Figure 3D). Mapping these conserved amino acids onto the structural complex reveals a high level of paralog conservation at the contact sites of the subunits (Figure 3E). For every amino acid, we thus have calculated amino acid movement data from the structural simulation that is combined with conservation of each gene, a 21-codon linear motif score for each gene, conservation relative to all NMDA receptor sequences, a 21-codon linear motif score for all NMDA receptor sequences, and for GluN2 subunits a conservation score for these 586 sequences. All these metrics were compiled into the amino acid matrix that can be used to assess genomic variants (https://doi.org/10.6084/m9.figshare.14991582). This supplemental file contains a readme file that describes every tab and details of each column within the document.

Variant Classifications

In addition to the extraction of ClinVar variants (Figure 1) we also extracted GRIN1, GRIN2A, GRIN2B, GRIN2C, and GRIN2D protein coding variants from gnomADv2 [Karczewski et al., 2020], allowing for the addition of population level allele frequencies. All variants were classified based on ClinVar into benign, conflicting-np (not provided),

pathogenic (all pathogenic and likely pathogenic), VUS, and if not found in ClinVar curated as gnomAD. Each variant was assessed using our mds insights (Figure 2), conservation (Figure 3), and any annotations within the UniProt database. Notably, pathogenic variants are enriched within the extracellular (1.23 times expected), the discontinuously helical (4.56 times expected), and helical (3.86) UniProt annotated topology (Figure 4A), while the cytoplasmic region is highly deficient in pathogenic variants (0.19). Variant analysis using PolyPhen2, Provean, and SIFT resulted in classification of variants into categories, with all three tools overwhelmingly classifying the known pathogenic effects to be harmful to protein function and the benign variants as being benign (Figure 4B). These tools however also rank many of the gnomAD variants not reported in ClinVar as damaging to protein function, suggesting that the tools do not have the power to segregate the non-dominant variants that would be within the gnomAD population variants. It is likely that this level of functional annotation is the difference between dominant linked genetics and other forms of inheritance or more subtle phenotypes. Our calculated conservation scores (Figure 3) with values ≥1 are highly predictive of pathogenic variants as well.

The most powerful prediction of pathogenic variants was the use of the conservation within all NMDA receptor genes or within the GluN2 subunits (green box and star, Figure 4B). Particularly important to note is the lack of gnomAD annotated variants conserved in all NMDA receptors, a key separator of these variants from functional annotation and conservation tools. This is indicative of genetics that likely separate dominant functional outcomes from that of recessive or subtle non-epilepsy associated variant impacts. A combined score from Figure 4B was generated, with a scale of 0 (no predicted impact) to 7 (conserved in all sequences with high selection and all functional tools predicted damage). A value ≥3.5 was highly indicative of pathogenic variant scores (p-value 2e-72 relative to all other variants), with no benign ClinVar annotations and very few gnomAD, conflicting-np, and VUS annotations (Figure 4C).

Next, we assessed our mds data for each amino acid (Figure 4D). Benign annotated variants had the highest RMSF of all variant groups, with pathogenic variants slightly lower than the rest (p-value 1e-4). The intramolecular contacts were lowest in the pathogenic group (p-value 6e-8) while the intermolecular contacts are the highest in the pathogenic group but not significant (p-value 0.06). This further defends our conservation data that pathogenic variant annotations are likely conserved sites involved in the intermolecular contacts that would be dominant negative variants, i.e. heterozygous variants that result in complex inactivation due to impairment of multimerization.

UniProt annotations that overlap the variants have details defined in the UniProt Functional Variants tab of the supplemental file (https://doi.org/10.6084/m9.figshare.14991582). A total of 9 variants in our database occur at UniProt annotated sites (Functional determinant, Glycine Binding, or Glutamate Binding) with one in GRIN1 (S688P), four in GRIN2A (R518H, R518C, R518L, N614S), and one in GRIN2B (N615I) identified as ClinVar pathogenic and three within the gnomAD database for GRIN2C (R516C, R516H, D759Y), all with combined impact score of function (Figure 4E). A total of 34 variants occur at amino acids with known posttranslational modification sites (Phospho S/T, Phospho Y, Glycosylation, Disulfide bond, Omega-N-methylarginine) with three having pathogenic

annotations and high functionality (GRIN1 C744Y, GRIN2A C436R, GRIN2B C456Y), six as VUS with highly variable scores, and 25 from gnomAD (Figure 4E). A total of 31 amino acid variants within our analysis have been assessed in a lab through mutagenesis and shown to have functional changes to GRIN1 [Ohba et al., 2015; Lemke et al., 2016; Ogden et al., 2017; Zehavi et al., 2017], GRIN2A [Lemke et al., 2013; Swanger et al., 2016; Lesca et al., 2013; Gao et al., 2017], and GRIN2B [Adams et al., 2014; Swanger et al., 2016; de Ligt et al., 2012; Ogden et al., 2017]. A total of 25 of the pathogenic variants with confirmed wetlab validation of protein function have an average combined score of 4.8±1.4, while five VUS have altered protein function with an averaged combined impact of 3.7±2.1. It should be noted that a variant can be shown to impact protein function and not result in dominant inheritance, such as the VUS GRIN2A V506A, which is the only VUS with a combined score below 3.

Mapping the combined scores for all the variants within GRIN1, GRIN2A, or GRIN2B relative to the amino acid position revealed a clustering of pathogenic variants with depletion of high scoring gnomAD variants (green box, Figure 4F). These sites correspond to extracellular and helical regions. For GRIN1 positions 550–850 showed an increased number of variants with higher pathogenicity scores and increased localization of pathogenic variants. For GRIN2A/B positions 500–900 showed an increased number of variants with higher pathogenicity scores and increased localization of pathogenic variants. This region is the same mapped for conservation across all 719 sequences (Figure 3C–D).

Integrating gnomAD population allele count annotations with variant combined scores for each of the genes shows only GRIN2C to have variants with allele counts >100 and combined scores >2 (Figure 4G). A total of 143 variants had a combined score >3.5 and found conserved throughout all 719 sequences. All the 62 pathogenic variants in this list have no observed gnomAD allele counts, while 6/8 of the conflicting-np had 0 gnomAD occurrences and two of the variants have 1 count, and 28 VUS had 0 gnomAD variants with one having one and two having two gnomAD counts (Figure 4H). A total of 43 variants found in gnomAD but not within ClinVar fall into this category, with the highest allele number of 4 in one gnomAD observation suggesting even these variants to be very rare (Figure 4H). This suggests that the filtering of variants using combined impact score, paralog conservation, topology, and gnomAD allele frequency are of high utility when confidently mapping pathogenic impact of NMDA receptor proteins.

Geno2MP Variants for Dominant Inheritance

To test this model of dominant genetics to phenotype insights, we surveyed the Geno2MP database [Geno2MP] for GRIN1, GRIN2A, and GRIN2B variants. A total of 278 variants were observed with 29 in GRIN1, 157 in GRIN2A, and 92 in GRIN2B. From these, 41 have a combined score >3, ten have the predicted dominant annotation observed (topology, paralog conservation, conserved >3.5, heterozygous in Geno2MP, Table 1), and two have a potential recessive inheritance structure (conserved >3, not conserved in paralog, and homozygous in Geno2MP, Table 1). Both individuals with predicted recessive homozygous variants (GRIN2A D166E, GRIN1 G318R) had abnormality of the nervous system. Of the ten individuals with dominant variants, seven had abnormality of the nervous

system including agenesis of corpus callosum, seizures, intellectual disability, microcephaly, abnormality of movement, and epileptic encephalopathy. There were two individuals with dominant variants with abnormality of the musculature (myopathy, distal arthrogryposis) and one with abnormality of the skeletal system (skeletal dysplasia), where the musculature and skeletal phenotypes were previously associated with epileptic phenotypes. Thus, using the dominant variant mechanisms of topology and paralog conservation has strong accuracy in identifying phenotypes previously connected to the NMDA receptor proteins.

VUS of High Impact

The 28 ClinVar annotated VUS of the NMDA receptor proteins not present within gnomAD and having >3.5 combined impact and conserved in multiple NMDA receptor proteins potentially fit a dominant mechanism. Three of these variants fall within the cytoplasmic topology of the protein and therefore do not fit our dominant mechanism. Of the 25 VUS identified to cluster similar to pathogenic variants, 11 were GRIN2A and 14 were GRIN2B (Table 2). A total of 16/25 of the VUS were conserved in all 719 sequences analyzed while 9/25 of the sites are conserved within the GluN2 586 sequences. Of these variants 11 have no clinical description and 10 are annotated with the term epilepsy somewhere in the description. It should be noted that upon further analysis of ClinVar, several of these VUS have been additionally noted in other individuals and annotated as pathogenic or likely pathogenic [Yang et al., 2014]. Helen DeVos Children's Hospital Pediatric Neurology identified two missense VUS in NMDA receptors (GRIN2B A1315V, GRIN2A G760S) in two separate individuals. GRIN2B A1315V has since initial report been reclassified as likely pathogenic by the clinical sequencing company. However, that variant does not reach our pathogenic classification by our filtering tools. While A1315V has not been observed within gnomAD, it is predicted to have minimal changes to the protein function based on conservation, PolyPhen2, Provean, and SIFT with a combined impact score of 0.25. Contrary, the GRIN2A G760S VUS is within our top 25 identified (Table 2) with no gnomAD occurrence, damaging calls from PolyPhen2, Provean, and SIFT, conserved as a G in all 719 sequences of NMDA receptor proteins, has a low molecular movement, and a sum score of 6.25. The G760S variant was inherited from the father who has pervasive developmental disorder and a noted history of seizures in childhood, and the variant has been reported in a separate study as *de novo* and associated with epilepsy [Yang et al., 2018]. Thus, our combined score has the resolution to distinguish potentially protein-damaging versus pathogenic NMDA receptor variants such as GRIN2A G760S.

Discussion

Accurate variant classification related to pathogenicity is critically important for clinical care. An analysis of ClinVar variants showed that over 55 percent of all variants for the NMDA receptor proteins are VUS. This fact indicates there is much work to be done in categorizing VUS in these proteins. From the ACMG standards, VUS cannot be clinically actionable, requiring further assessments of variants before a genetic diagnosis can be made. While computational tools used for variant interpretation have been hampered by circularity [Grimm et al., 2015], the role of training variants on other known variants rather than on mechanistic studies, they often annotate variants seen in common populations as

damaging without the ability to segregate disorders. Not all functional variants are the same within a protein. For example, some amino acid variants exert dominant negative changes where the variant protein copy also modifies the wild type copy [Herskowitz, 1987], often where one variant results in the loss of multimer contacts in dimers such that ~75% of the complex is altered. These forms of dominant negative variants are the primary interest in mapping dominance for this paper. Other variants in proteins can exert different protein functional outcomes, with the most well studied example in the CFTR protein for cystic fibrosis, where variants can change protein abundance, localization, receptor activation, or ion transport based on the amino acid altered [Sanders et al., 2021]. Gain-of-function changes further complicate interpretation because they can either alter functional motifs that control protein abundance or can add entirely new functional sites that are not predicable based on functional tools. Machine/deep learning tools have been applied to these problems [Quang et al., 2015; Zou et al., 2019], but these tools still struggle to achieve disease association filtering of variants based on the mechanisms of dominant negative changes within a protein complex.

Here we utilized pathogenic variants in combination with a sequence-to-structure data analysis to build a functional matrix of the amino acids involved in forming the NMDA receptor complex. This elucidated that amino acids conserved throughout the entire protein family (paralogs) are often conserved for packing of the receptors through extracellular and helical contacts. With two units of each protein required for forming a functional receptor, a dominant negative variant yields only ~25% of functional complexes to form. Other variants within the proteins that still allow for the complex to form, but decreases the glutamate/glycine binding, ion transport, intracellular activation, or protein interactions not within the core of the complex can result in more subtle changes to the receptor activity, such that 50% or more of the complexes can still function at a critical level. These more subtle changes can be less likely to result in more severe phenotypes or can require both copies of the alleles (recessive) to have functional changes to drive disease. The gnomAD data suggests that many variants within the NMDA receptor have predicted impact but are not currently associated with disease state, likely because these are not dominant negative changes and subtlety impact the receptor complex.

A summary of variant classes for the NMDA receptor can be found in Figure 5. Amino acids conserved throughout the GluN1 and GluN2 subunits are those amino acids critical to the proper packing of the macromolecular complex. Heterozygous variants at these sites result in a dominant negative impact, such that amino acid changes inhibit the binding of proteins to each other. With two of each subunit, a heterozygous variant results in removal of ~75% of the functional complexes, 25% that would have both proteins with the variants, and 50% that would have a single copy with the variants. Thus, only the remaining 25% of the protein complex would have two functional copies of the protein. These variants that alter amino acid contacts at fully conserved amino acids result in dominant negative outcomes and are not present in the gnomAD database. Contrary, we identify multiple gnomAD variants with computational prediction from tools like PolyPhen2, Provean and SIFT for damaging outcomes, but these amino acids are unique to each protein of the NMDA receptor complex. This suggests a potential recessive role of genetic variants, where they can occur within gnomAD and would likely result in decreased receptor function (~25–50%). These variants

would likely only give rise to disease if an individual carried two damaging copies, i.e., recessive genetics. This observation is supported by evidence of recessive disorders within the genes based on sequencing [Lemke et al., 2016].

These strategies should be tested in additional dominant genetic disorders to determine the applicability to larger genomic datasets. For all proteins with available structures, we need cohesive efforts in converting the structure files that are qualitative into quantitative amino acid insights using tools such as molecular dynamic simulations. With the advancements of cryo-EM determining larger and more complex multiunit structures that are commonly associated with dominant negative outcomes, the quantitative approach we take here could have many future applications. The tools of NCBI for ortholog sequence extractions means that paralog mapping and sequence alignments could be readily applied to any gene family. Therefore, these tools could be broadly applied in the future for dominant negative genetic mechanisms.

Methods

Genomic variants from ClinVar [Landrum et al., 2016] and gnomAD (v2.1.1_non-TOPMed) [Karczewski et al., 2020] were extracted on 1/3/2020. The Geno2MP [Geno2MP - Home] variants and all UniProt annotations for each protein were extracted on 7/10/2021. All analysis of genomic variants can be found at https://doi.org/10.6084/m9.figshare.14991582. Homology modeling of the NMDA receptor proteins were done first by modeling GRIN2A (Q12879) using YASARA modeling [Krieger et al., 2009], where four known PDB structures (4PE5, 4TLL, 5UOW, 4TLM) were merged into a single structure of four units. From this base structure homology modeling for GRIN1 (Q05586) and GRIN2B (Q13224) sequences was performed using YASARA (version 17). Structural alignment of two GRIN1 and one GRIN2B molecules was performed onto the GRIN2A model followed by energy minimization using the YASARA2 force field. The complex was modeled into a lipid membrane using the YASARA md_runmembrane macro using the AMBER14 force field [Duan et al., 2003] using a phosphatidyl-ethanolamine (PEA) membrane with 0.997g/mL explicit water, 0.9% NaCl, and protein protonation based on pH 7.4. Following energy minimization with water and the membrane the complex was run for 30 nanoseconds of simulation captured in 300 atomic snapshots, one every 100 picoseconds. The simulation consisted of 52,451 protein atoms, 423 Cl, 433 Na, 149,057 water for a total of 593,366 atoms. Analysis of the trajectory was performed using the md_analyze and md_analyzeres macros. Correlations for DCCM were determined for values >0.8. All structure and dynamics data can be found at https://doi.org/10.6084/m9.figshare.14991558.

Sequences for the open reading frame of each gene was extracted from NCBI paralogs for vertebrate species. Open reading frames were extracted using TransDecoder [Haas et al., 2013]. Sequences were aligned using ClustalW codon [Larkin et al., 2007], removing any sequences with ambiguity or missing exons found in >90% of the other sequences. Following alignment codons were assessed for selection using dN-dS with a Maximum likelihood Muse-Gaut model [Muse and Gaut, 1994] for Tamura-Nei nucleotide substitutions [Tamura and Nei, 1993] using HyPhy [Pond et al., 2005] and MEGA (version 6) [Tamura et al., 2011]. A 21-codon sliding window for conservation was used [Prokop et

al., 2017]. Genomic sequences were aligned using all 719 extracted, placing each human protein in the first slot followed by removal of all gaps within that protein, followed by calculation of conservation as done above. An additional alignment for GRIN2A, GRIN2B, GRIN2C, and GRIN2D was done, placing each human protein sequence at the reference statistical analysis. Sequence alignments and the top tree are available at https://doi.org/10.6084/m9.figshare.14991537.

All extracted variants were assessed with PolyPhen2 (Adzhubei et al. 2010), Provean (Choi and Chan 2015), and SIFT (Ng and Henikoff 2003). These predictions were converted to binary such that 0 was not bad prediction and 1 bad prediction. These scores were added to the conservation score generated above (0–2 scale), and a binary calculation (0 or 1 with 1 being conserved) of conservation in all sequences or GRIN2 sequences, yielding a maximum score of 7. Any category where statistics were calculated they were based on a t-test of the pathogenic annotation relative to all others. Allele counts were extracted from the gnomAD data. Top variants were identified using the cutoff of combined score of >3.5, extracellular/helical topology, and conservation selection ≥1 in all paralog sequences.

Acknowledgements

We would like to thank Ryan M. Bebej (Calvin University) for the assistance he provided in critiquing this manuscript. Funding for this study were from NIH Office of the Director K01ES025435 (to JWP) and Michigan State University.

Data Availability

All protein structure files (https://doi.org/10.6084/m9.figshare.14991558), evolutionary alignments (https://doi.org/10.6084/m9.figshare.14991537), and variant data (https://doi.org/10.6084/m9.figshare.14991582) are available at their respective FigShare deposits.

References

Abrahams BS, Arking DE, Campbell DB, Mefford HC, Morrow EM, Weiss LA, Menashe I, Wadkins T, Banerjee-Basu S, Packer A. 2013. SFARI Gene 2.0: a community-driven knowledgebase for the autism spectrum disorders (ASDs). Mol. Autism 4: 36. [PubMed: 24090431]

Adams DR, Yuan H, Holyoak T, Arajs KH, Hakimi P, Markello TC, Wolfe LA, Vilboux T, Burton BK, Fajardo KF, Grahame G, Holloman C, Sincan M, Smith ACM, Wells GA, Huang Y, Vega H, Snyder JP, Golas GA, Tifft CJ, Boerkoel CF, Hanson RW, Traynelis SF, Kerr DS, Gahl WA. 2014. Three rare diseases in one Sib pair: RAI1, PCK1, GRIN2B mutations associated with Smith-Magenis Syndrome, cytosolic PEPCK deficiency and NMDA receptor glutamate insensitivity. Mol. Genet. Metab 113: 161–170. [PubMed: 24863970]

Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. 2010. A method and server for predicting damaging missense mutations. Nat. Methods 7: 248–249. [PubMed: 20354512]

Choi Y, Chan AP. 2015. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. Bioinforma. Oxf. Engl 31: 2745–2747.

Duan Y, Wu C, Chowdhury S, Lee MC, Xiong G, Zhang W, Yang R, Cieplak P, Luo R, Lee T, Caldwell J, Wang J, Kollman P. 2003. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. J. Comput. Chem 24: 1999–2012. [PubMed: 14531054]

Gao K, Tankovic A, Zhang Y, Kusumoto H, Zhang J, Chen W, XiangWei W, Shaulsky GH, Hu C, Traynelis SF, Yuan H, Jiang Y. 2017. A de novo loss-of-function GRIN2A mutation associated

with childhood focal epilepsy and acquired epileptic aphasia. PloS One 12: e0170818. [PubMed: 28182669]

Geno2MP, NHGRI/NHLBI University of Washington-Center for Mendelian Genomics (UW-CMG), Seattle, WA (URL: http://geno2mp.gs.washington.edu [7/2021].

Grimm DG, Azencott C-A, Aicheler F, Gieraths U, MacArthur DG, Samocha KE, Cooper DN, Stenson PD, Daly MJ, Smoller JW, Duncan LE, Borgwardt KM. 2015. The evaluation of tools used to predict the impact of missense variants is hindered by two types of circularity. Hum. Mutat 36: 513–523. [PubMed: 25684150]

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, MacManes MD, Ott M, Orvis J, Pochet N, Strozzi F, Weeks N, Westerman R, William T, Dewey CN, Henschel R, LeDuc RD, Friedman N, Regev A. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat. Protoc 8: 1494–1512. [PubMed: 23845962]

Herskowitz I 1987. Functional inactivation of genes by dominant negative mutations. Nature 329: 219–222. [PubMed: 2442619]

Karakas E, Furukawa H. 2014. Crystal structure of a heterotetrameric NMDA receptor ion channel. Science 344: 992–997. [PubMed: 24876489]

Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, Collins RL, Laricchia KM, Ganna A, Birnbaum DP, Gauthier LD, Brand H, Solomonson M, Watts NA, Rhodes D, Singer-Berk M, England EM, Seaby EG, Kosmicki JA, Walters RK, Tashman K, Farjoun Y, Banks E, Poterba T, Wang A, Seed C, Whiffin N, Chong JX, Samocha KE, Pierce-Hoffman E, Zappala Z, O'Donnell-Luria AH, Minikel EV, Weisburd B, Lek M, Ware JS, Vittal C, Armean IM, Bergelson L, Cibulskis K, Connolly KM, Covarrubias M, Donnelly S, Ferriera S, Gabriel S, Gentry J, Gupta N, Jeandet T, Kaplan D, Llanwarne C, Munshi R, Novod S, Petrillo N, Roazen D, Ruano-Rubio V, Saltzman A, Schleicher M, Soto J, Tibbetts K, Tolonen C, Wade G, Talkowski ME, Genome Aggregation Database Consortium, Neale BM, Daly MJ, MacArthur DG. 2020. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 581: 434–443. [PubMed: 32461654]

Krieger E, Joo K, Lee J, Lee J, Raman S, Thompson J, Tyka M, Baker D, Karplus K. 2009. Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. Proteins 77 Suppl 9: 114–122. [PubMed: 19768677]

Landrum MJ, Lee JM, Benson M, Brown G, Chao C, Chitipiralla S, Gu B, Hart J, Hoffman D, Hoover J, Jang W, Katz K, Ovetsky M, Riley G, Sethi A, Tully R, Villamarin-Salomon R, Rubinstein W, Maglott DR. 2016. ClinVar: public archive of interpretations of clinically relevant variants. Nucleic Acids Res. 44: D862–868. [PubMed: 26582918]

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23: 2947–2948. [PubMed: 17846036]

Lee C-H, Lü W, Michel JC, Goehring A, Du J, Song X, Gouaux E. 2014. NMDA receptor structures reveal subunit arrangement and pore architecture. Nature 511: 191–197. [PubMed: 25008524]

Lemke JR, Geider K, Helbig KL, Heyne HO, Schütz H, Hentschel J, Courage C, Depienne C, Nava C, Heron D, Møller RS, Hjalgrim H, Lal D, Neubauer BA, Nürnberg P, Thiele H, Kurlemann G, Arnold GL, Bhambhani V, Bartholdi D, Pedurupillay CRJ, Misceo D, Frengen E, Strømme P, Dlugos DJ, Doherty ES, Bijlsma EK, Ruivenkamp CA, Hoffer MJV, Goldstein A, Rajan DS, Narayanan V, Ramsey K, Belnap N, Schrauwen I, Richholt R, Koeleman BPC, Sá J, Mendonça C, de Kovel CGF, Weckhuysen S, Hardies K, De Jonghe P, De Meirleir L, Milh M, Badens C, Lebrun M, Busa T, Francannet C, Piton A, Riesch E, Biskup S, Vogt H, Dorn T, Helbig I, Michaud JL, Laube B, Syrbe S. 2016. Delineating the GRIN1 phenotypic spectrum: A distinct genetic NMDA receptor encephalopathy. Neurology 86: 2171–2178. [PubMed: 27164704]

Lemke JR, Lal D, Reinthaler EM, Steiner I, Nothnagel M, Alber M, Geider K, Laube B, Schwake M, Finsterwalder K, Franke A, Schilhabel M, Jähn JA, Muhle H, Boor R, Van Paesschen W, Caraballo R, Fejerman N, Weckhuysen S, De Jonghe P, Larsen J, Møller RS, Hjalgrim H, Addis L, Tang S, Hughes E, Pal DK, Veri K, Vaher U, Talvik T, Dimova P, Guerrero López R, Serratosa JM, Linnankivi T, Lehesjoki A-E, Ruf S, Wolff M, Buerki S, Wohlrab G, Kroell J, Datta AN, Fiedler B, Kurlemann G, Kluger G, Hahn A, Haberlandt DE, Kutzer C, Sperner J, Becker F, Weber

YG, Feucht M, Steinböck H, Neophythou B, Ronen GM, Gruber-Sedlmayr U, Geldner J, Harvey RJ, Hoffmann P, Herms S, Altmüller J, Toliat MR, Thiele H, Nürnberg P, Wilhelm C, Stephani U, Helbig I, Lerche H, Zimprich F, Neubauer BA, Biskup S, von Spiczak S. 2013. Mutations in GRIN2A cause idiopathic focal epilepsy with rolandic spikes. Nat. Genet 45: 1067–1072. [PubMed: 23933819]

Lesca G, Rudolf G, Bruneau N, Lozovaya N, Labalme A, Boutry-Kryza N, Salmi M, Tsintsadze T, Addis L, Motte J, Wright S, Tsintsadze V, Michel A, Doummar D, Lascelles K, Strug L, Waters P, de Bellescize J, Vrielynck P, de Saint Martin A, Ville D, Ryvlin P, Arzimanoglou A, Hirsch E, Vincent A, Pal D, Burnashev N, Sanlaville D, Szepetowski P. 2013. GRIN2A mutations in acquired epileptic aphasia and related childhood focal epilepsies and encephalopathies with speech and language dysfunction. Nat. Genet 45: 1061–1066. [PubMed: 23933820]

de Ligt J, Willemsen MH, van Bon BWM, Kleefstra T, Yntema HG, Kroes T, Vulto-van Silfhout AT, Koolen DA, de Vries P, Gilissen C, del Rosario M, Hoischen A, Scheffer H, de Vries BBA, Brunner HG, Veltman JA, Vissers LELM. 2012. Diagnostic exome sequencing in persons with severe intellectual disability. N. Engl. J. Med 367: 1921–1929. [PubMed: 23033978]

Lü W, Du J, Goehring A, Gouaux E. 2017. Cryo-EM structures of the triheteromeric NMDA receptor and its allosteric modulation. Science 355.

MacDermott AB, Mayer ML, Westbrook GL, Smith SJ, Barker JL. 1986. NMDA-receptor activation increases cytoplasmic calcium concentration in cultured spinal cord neurones. Nature 321: 519–522. [PubMed: 3012362]

Muse SV, Gaut BS. 1994. A likelihood approach for comparing synonymous and nonsynonymous nucleotide substitution rates, with application to the chloroplast genome. Mol. Biol. Evol 11: 715–724. [PubMed: 7968485]

Ng PC, Henikoff S. 2003. SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res. 31: 3812–3814. [PubMed: 12824425]

Ogden KK, Chen W, Swanger SA, McDaniel MJ, Fan LZ, Hu C, Tankovic A, Kusumoto H, Kosobucki GJ, Schulien AJ, Su Z, Pecha J, Bhattacharya S, Petrovski S, Cohen AE, Aizenman E, Traynelis SF, Yuan H. 2017. Molecular Mechanism of Disease-Associated Mutations in the Pre-M1 Helix of NMDA Receptors and Potential Rescue Pharmacology. PLoS Genet. 13: e1006536. [PubMed: 28095420]

Ohba C, Shiina M, Tohyama J, Haginoya K, Lerman-Sagie T, Okamoto N, Blumkin L, Lev D, Mukaida S, Nozaki F, Uematsu M, Onuma A, Kodera H, Nakashima M, Tsurusaki Y, Miyake N, Tanaka F, Kato M, Ogata K, Saitsu H, Matsumoto N. 2015. GRIN1 mutations cause encephalopathy with infantile-onset epilepsy, and hyperkinetic and stereotyped movement disorders. Epilepsia 56: 841–848. [PubMed: 25864721]

Pond SLK, Frost SDW, Muse SV. 2005. HyPhy: hypothesis testing using phylogenies. Bioinforma. Oxf. Engl 21: 676–679.

Prokop JW, Lazar J, Crapitto G, Smith DC, Worthey EA, Jacob HJ. 2017. Molecular modeling in the age of clinical genomics, the enterprise of the next generation. J. Mol. Model 23: 75. [PubMed: 28204942]

Prokop JW, Yeo NC, Ottmann C, Chhetri SB, Florus KL, Ross EJ, Sosonkina N, Link BA, Freedman BI, Coppola CJ, McDermott-Roe C, Leysen S, Milroy L-G, Meijer FA, Geurts AM, Rauscher FJ, Ramaker R, Flister MJ, Jacob HJ, Mendenhall EM, Lazar J. 2018. Characterization of Coding/Noncoding Variants forSHROOM3in Patients with CKD. J. Am. Soc. Nephrol. JASN

Quang D, Chen Y, Xie X. 2015. DANN: a deep learning approach for annotating the pathogenicity of genetic variants. Bioinforma. Oxf. Engl 31: 761–763.

Ryan TJ, Emes RD, Grant SG, Komiyama NH. 2008. Evolution of NMDA receptor cytoplasmic interaction domains: implications for organisation of synaptic signalling complexes. BMC Neurosci. 9: 6. [PubMed: 18197970]

Sanders M, Lawlor JMJ, Li X, Schuen JN, Millard SL, Zhang X, Buck L, Grysko B, Uhl KL, Hinds D, Stenger CL, Morris M, Lamb N, Levy H, Bupp C, Prokop JW. 2021. Genomic, transcriptomic, and protein landscape profile of CFTR and cystic fibrosis. Hum. Genet 140: 423–439. [PubMed: 32734384]

Swanger SA, Chen W, Wells G, Burger PB, Tankovic A, Bhattacharya S, Strong KL, Hu C, Kusumoto H, Zhang J, Adams DR, Millichap JJ, Petrovski S, Traynelis SF, Yuan H. 2016. Mechanistic

Insight into NMDA Receptor Dysregulation by Rare Variants in the GluN2A and GluN2B Agonist Binding Domains. Am. J. Hum. Genet 99: 1261–1280. [PubMed: 27839871]

Tamura K, Nei M. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol. Biol. Evol 10: 512–526. [PubMed: 8336541]

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol 28: 2731–2739. [PubMed: 21546353]

Yang X, Qian P, Xu X, Liu X, Wu X, Zhang Y, Yang Z. 2018. GRIN2A mutations in epilepsy-aphasia spectrum disorders. Brain Dev 40:205–210. [PubMed: 29056244]

Yang Y, Muzny DM, Xia F, Niu Z, Person R, Ding Y, Ward P, Braxton A, Wang M, Buhay C, Veeraraghavan N, Hawes A, Chiang T, Leduc M, Beuten J, Zhang J, He W, Scull J, Willis A, Landsverk M, Craigen WJ, Bekheirnia MR, Stray-Pedersen A, Liu P, Wen S, Alcaraz W, Cui H, Walkiewicz M, Reid J, Bainbridge M, Patel A, Boerwinkle E, Beaudet AL, Lupski JR, Plon SE, Gibbs RA, Eng CM. 2014. Molecular findings among patients referred for clinical whole-exome sequencing. JAMA 312: 1870–1879. [PubMed: 25326635]

Yoo HJ, Cho IH, Park M, Yang SY, Kim SA. 2012. Family based association of GRIN2A and GRIN2B with Korean autism spectrum disorders. Neurosci. Lett 512: 89–93. [PubMed: 22326929]

Yuan H, Low C-M, Moody OA, Jenkins A, Traynelis SF. 2015. Ionotropic GABA and Glutamate Receptor Mutations and Human Neurologic Diseases. Mol. Pharmacol 88: 203–217. [PubMed: 25904555]

Zehavi Y, Mandel H, Zehavi A, Rashid MA, Straussberg R, Jabur B, Shaag A, Elpeleg O, Spiegel R. 2017. De novo GRIN1 mutations: An emerging cause of severe early infantile encephalopathy. Eur. J. Med. Genet 60: 317–320. [PubMed: 28389307]

Zou J, Huss M, Abid A, Mohammadi P, Torkamani A, Telenti A. 2019. A primer on deep learning in genomics. Nat. Genet 51: 12–18. [PubMed: 30478442]

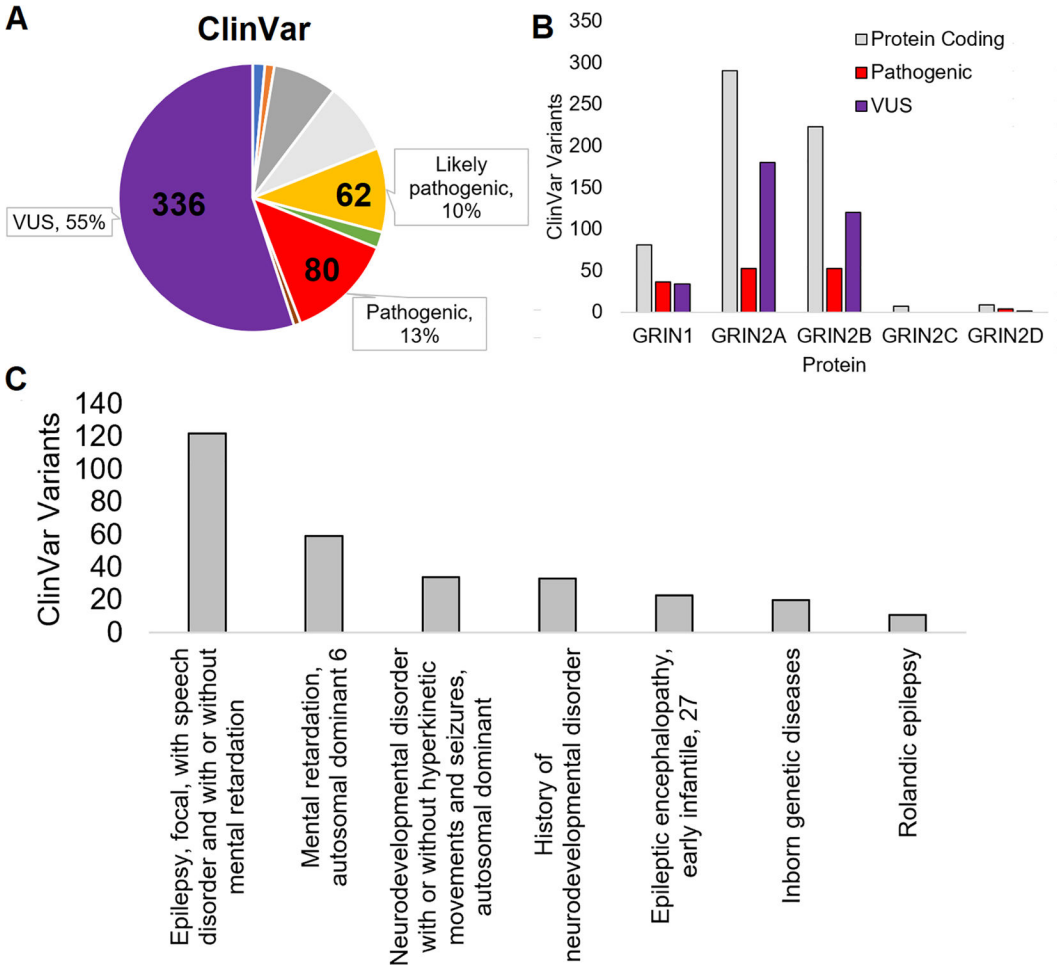

Figure 1. ClinVar analysis of the NMDA receptor units.

A) Percent of variant groups (VUS-magenta, likely pathogenic-yellow, pathogenic-red) from ClinVar annotations for GRIN1, GRIN2A, GRIN2B, GRIN2C, and GRIN2D. **B**) Number of variants in each gene from ClinVar (gray) broken down into the groups of A. **C**) Top phenotypes observed for variants of the receptor units.

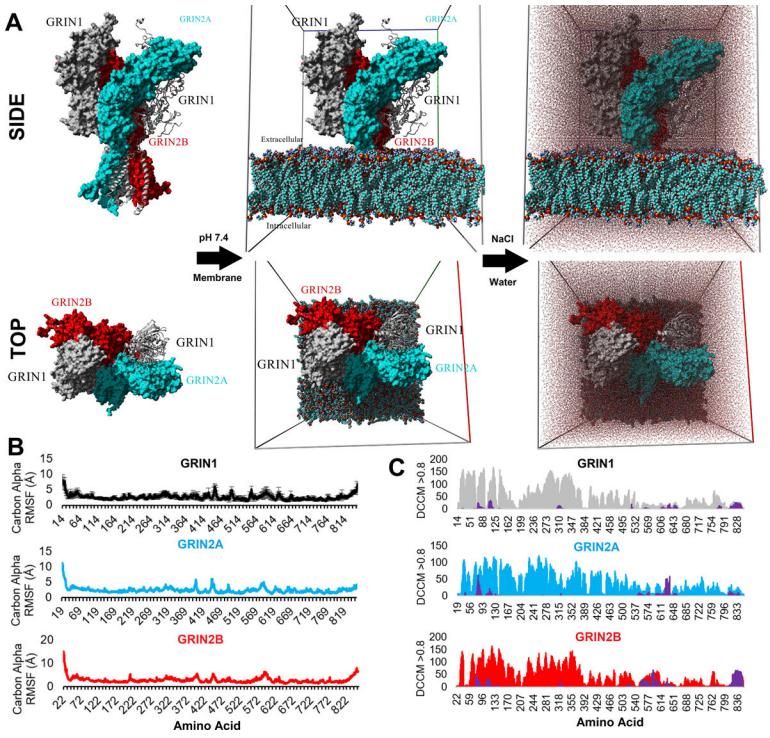

Figure 2. Structure and dynamics of NMDA receptor complex.

A) Model of two units of GRIN1 (gray), one GRIN2A (cyan), and one GRIN2B (red). The model is shown as the side view or the top view with model alone, model embedded into a lipid membrane, and model filled with water. **B**) Movement of amino acids in the three proteins throughout 30 nanoseconds of molecular dynamics simulations. **C**) The number of amino acids that correlate to each amino acid within a protein (GRIN1=gray, GRIN2A=cyan, GRIN2B=red) and to the other proteins (magenta).

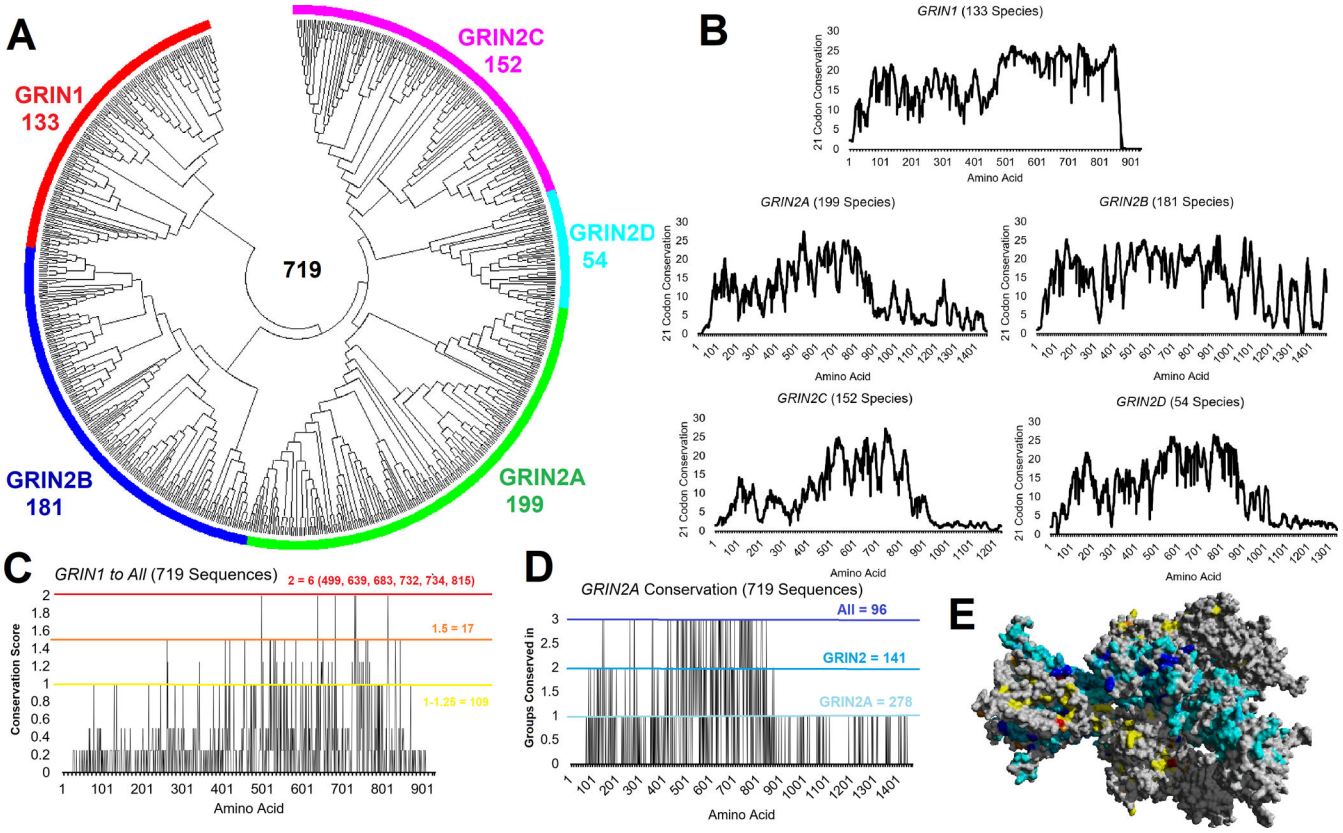

Figure 3. Evolution of NMDA receptor proteins.

A) Phylogenetic tree generated by maximum likelihood for 719 sequences of GRIN1 (red- 133 species), GRIN2B (blue- 181 species), GRIN2A (green- 199 species), GRIN2D (cyan- 54 species), and GRIN2C (magenta- 152 species). **B**) Conservation analysis for each of the genes using a 21-codon sliding window of conservation. **C**) Conservation of GRIN1 sequence using all 719 sequences aligned with number of amino acids labeled for 1–1.25 (yellow), 1.5 (orange), and 2 (red). **D**) Conservation of GRIN2A sequence compared to all 719 sequences. Amino acids with a value of 3 (blue) are conserved in all sequences, those with a value of 2 (light blue) are conserved in all GRIN2 sequences (GRIN2A, GRIN2B, GRIN2C, and GRIN2D), and those with a value of 1 (cyan) are conserved in GRIN2A. **E**) Amino acids colored from panel C-D onto the protein model from figure 2.

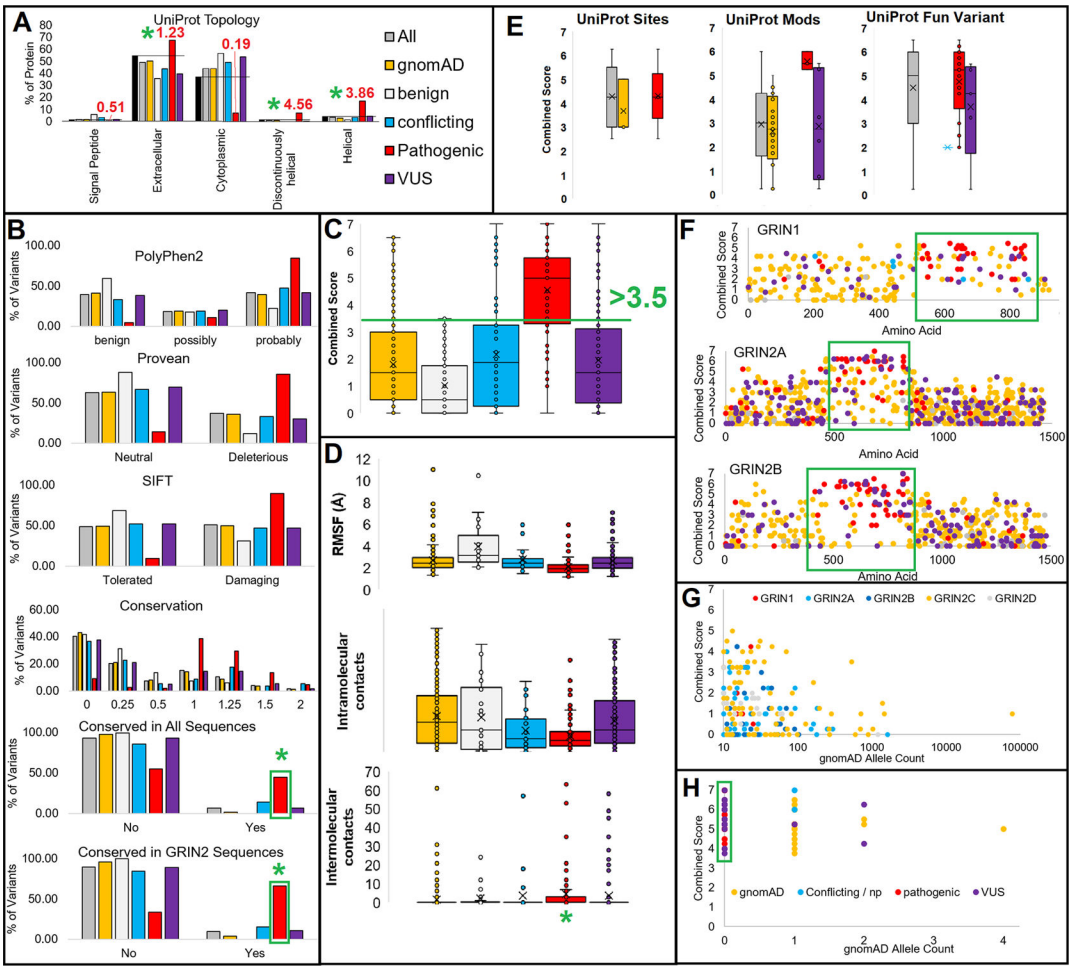

Figure 4. Human genetic variants from NMDA receptor proteins.

A) Following extraction of variants from ClinVar and gnomAD for GRIN1, GRIN2A, GRIN2B, GRIN2C, and GRIN2D, the UniProt topology was annotated for various groups of variants. Throughout all panels of Figure 4, colors are based on sequences for all pooled (gray), gnomAD (yellow), benign (light gray), conflicting interpretation (cyan), pathogenic (red), or VUS (magenta). The black bar represents the percent of amino acids in the proteins with each annotation. In red text is the enrichment of pathogenic variants relative to the expected frequency. **B)** All variants were assessed using PolyPhen2, Provean, SIFT, Conservation, conservation in all species, conservation in GRIN2 sequences. **C**) Scores from panel A were integrated for each of the groups of variants as shown in a box and whisker plot. **D**) Molecular dynamics data for each of the groups shown for RMSF, DCCM Intramolecular contacts, or DCCM Intermolecular contacts. **E)** Functional annotation of variants from UniProt for active sites of proteins (sites, with 9 annotations), sites with posttranslational modifications (Mods, 34 annotations), or functional variants with laboratory insights for altered cellular biology (Fun Variant, 31 annotations). **F**) Variants within GRIN1, GRIN2A, or GRIN2B. The green box identifies regions with low gnomAD variants and high pathogenic variants. **G**) gnomAD allele count (x-axis) relative to combined scores for the five proteins. **H**) gnomAD allele count (x-axis) for variants reaching 3.5

combined impact and found conserved in either all sequences or GRIN2 sequences. Green box identifies the clustering of pathogenic (red) and VUS (magenta) at 0 counts in gnomAD.

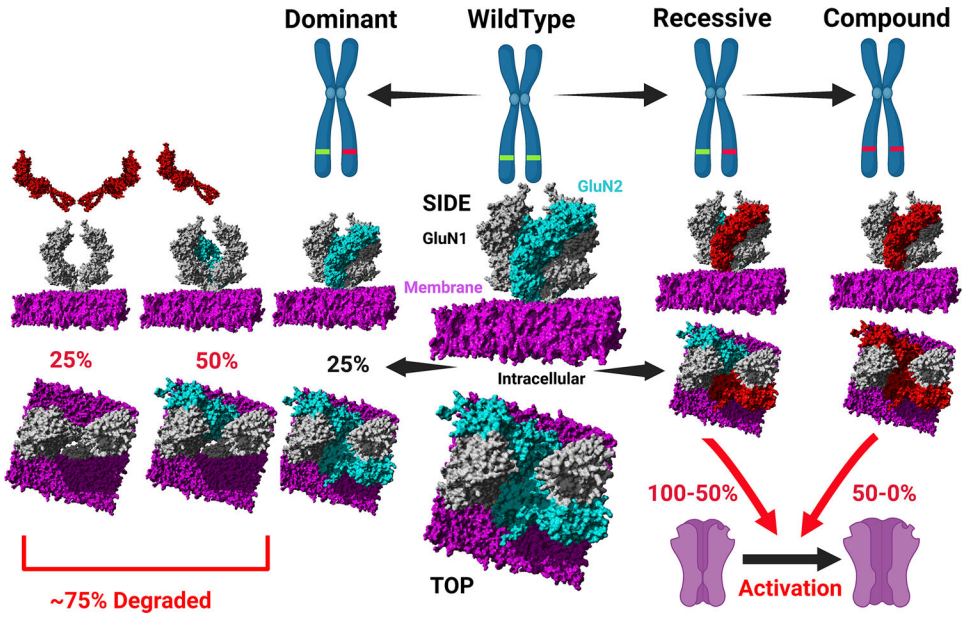

Figure 5. NMDA receptor model of dominant negative vs recessive variants.

On the top of the figure is the chromosome model with wildtype allele in green and mutant allele in red. The wildtype protein complex is shown below with two units of GluN1 (gray) and two units of GluN2 (cyan) in a membrane (magenta). The complex is shown as both the side view and the top view. To the left is the dominant variant model (red proteins) and to the left is the recessive model as either heterozygous or compound heterozygous / homozygous. The figure was generated using Biorender software in combination with YASARA generated images.

Table 1

Phenotypes of the Geno2MP variants fitting models.

All variants are predicted to be damaging in PolyPhen2, Provean, and SIFT.

Protein	Coded	% 719 sequence conservation	Sum	UniProt Topology	Predicted Model	HPO profiles	homozygous profiles	HPO term: Broad	HPO term: narrow/ medium
GRIN1	E522D	100.00	5.25	Extracellular	Dominant	1	0	Abnormality of the nervous system	Agenesis of corpus callosum
GRIN2A	G510C	81.36	5.25	Extracellular	Dominant	1	0	Abnormality of the nervous system	Seizures
GRIN2B	D661A	81.48	5	Extracellular	Dominant	1	0	Abnormality of the nervous system	Intellectual disability
GRIN2B	N616S	81.48	5	Discontinuously helical	Dominant	1	0	Abnormality of the musculature	myopathy
GRIN2A	I654T	81.36	5	Extracellular	Dominant	1	0	Abnormality of the musculature	Distal arthrogryposis
GRIN1	D461G	100.00	4.75	Extracellular	Dominant	1	0	Abnormality of the skeletal system	Skeletal dysplasia
GRIN1	S688Y	100.00	4.5	Extracellular	Dominant	1	0	Abnormality of the nervous system	Microcephaly
GRIN2A	H702Y	81.22	4.25	Extracellular	Dominant	2	0	Abnormality of nervous system	Abnormality of movement
GRIN2A	V417F	99.72	4	Extracellular	Dominant	1	0	Abnormality of the nervous system	Speech apraxia
GRIN2A	G250R	99.44	3.5	Extracellular	Dominant	1	0	Abnormality of the nervous system	Epileptic encephalopathy
GRIN2A	D166E	77.61	4	Extracellular	Recessive	1	1	Abnormality of the nervous system	Abnormality of the cerebellum
GRIN1	G318R	17.79	3.25	Extracellular	Recessive	1	2	Abnormality of the nervous system	Intellectual disability

Table 2

Top VUS that fit the dominant variant predictions based on Pathogenic variant analysis.

Protein	#	AA	Var	Conservation	Conserved in All	GRIN2 Conserved	RMSF	Intra Contacts	Inter Contacts	PolyPhen2	Provean	SIFT	Sum
GRIN2A	498	G	D	1.25	Yes	Yes	1.676	40	0	1	1	1	6
GRIN2A	531	T	M	1.5		Yes	1.474	40	0	1	1	1	6
GRIN2A	638	A	V	1.25	Yes	Yes	1.944	2	0	1	1	1	6
GRIN2A	648	N	S	1.5	Yes	Yes	1.369	19	5	1	1	1	7
GRIN2A	701	M	I	1	Yes	Yes	1.657	20	0	1	1	1	6
GRIN2A	716	A	D	1.25		Yes	1.977	36	0	1	1	1	5
GRIN2A	728	F	L	1	Yes	Yes	1.738	42	0	1	1	1	6
GRIN2A	731	D	H	1.5	Yes	Yes	1.527	54	0	1	1	1	7
*GRIN2A	760	G	S	1.25	Yes	Yes	1.227	20	0	1	1	1	6
GRIN2A	763	I	T	1.5	Yes	Yes	1.436	41	0		1	1	6
GRIN2A	815	D	E	1		Yes	2.961	29	5	1	1		4
GRIN2B	436	C	R	1.25		Yes	1.9	57	0	1	1	1	5
GRIN2B	550	F	S	1.25	Yes	Yes	1.915	2	0	1	1	1	6
GRIN2B	558	V	I	1.25		Yes	2.667	16	28	0.5		1	4
GRIN2B	646	Y	C	1	Yes	Yes	1.562	17	2	1	1	1	6
GRIN2B	649	N	S	1.25	Yes	Yes	1.459	11	2	1	1	1	6
GRIN2B	649	N	T	1.25	Yes	Yes	1.459	11	2	1	1	1	6
GRIN2B	688	N	K	1.5		Yes	2.027	13	0	1	1	1	6
GRIN2B	691	T	A	1.25		Yes	1.57	25	0	1	1	1	5
GRIN2B	726	L	P	1		Yes	1.993	22	0	1	1	1	5
GRIN2B	734	A	T	1.25	Yes	Yes	1.431	9	0	1	1		5
GRIN2B	754	G	R	1		Yes	2.695	1	0		1	1	4
GRIN2B	818	M	R	1	Yes	Yes	2.501	13	18	1	1	1	6
GRIN2B	818	M	T	1	Yes	Yes	2.501	13	18	1	1	1	6
GRIN2B	820	G	E	2	Yes	Yes	2.016	12	18	1	1	1	7

The * next to GRIN2A G760S is the individual from the Helen DeVos Children's Hospital. For each variant is annotated the protein it is found in, the # of the amino acid, the wild type amino acid (AA), the variant (Var) amino acid, the conservation score (0–2), whether the variant is conserved in all sequence alignments or in the GRIN2 sequences (GRIN2A, GRIN2B, GRIN2C, GRIN2D), the root mean squared fluctuation (RMSF) of the molecular dynamics for the amino acid, the intra and inter molecular contacts based on DCCM, the predictions scores for damaging variants based on PolyPhen2/Provean/SIFT (score of 1 means its bad), and the summed variant score (0–7, with 7 high functional outcome prediction).