medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Distances from ligands as main predictive features for pathogenicity and functional effect of variants in NMDA receptors.

Ludovica Montanucci[1§], Tobias Brünger[2§], Nisha Bhattarai[1], Christian M Boßelmann[1], Sukhan Kim[3,4], James P. Allen[3], Jing Zhang[3], Chiara Klöckner[5], Piero Fariselli[6], Patrick May[7], Johannes R. Lemke[5], Scott J. Myers[3,4], Hongjie Yuan[3,4], Stephen F. Traynelis[3,4*], Dennis Lal[1,8-10*]

[1] Department of Neurology, McGovern Medical School, The University of Texas Health Science Center at Houston, Houston, TX, USA

[2] Cologne Center for Genomics, University of Cologne, University Hospital Cologne, Cologne, Germany.

[3] Department of Pharmacology and Chemical Biology, Emory University School of Medicine, Atlanta, GA 30322, USA.

[4] Center for Functional Evaluation of Rare Variants (CFERV), Emory University School of Medicine, Atlanta, GA 30322, USA.

[5] Institute of Human Genetics, University of Leipzig Hospitals and Clinics, Leipzig, Germany

[6] Department of Medical Sciences, University of Torino, Torino (Italy).

[7] Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Esch-sur-Alzette, Luxembourg.

[8] Program in Medical and Population Genetics, Broad Institute of Massachusetts Institute of Technology (M.I.T.) and Harvard, USA.

[9] Stanley Center for Psychiatric Research, Broad Institute of Harvard and M.I.T, Cambridge, MA, USA

[10] Epilepsy Center, Neurological Institute, Cleveland Clinic, Cleveland, OH, US.

[§] Joint first Authors

[*]Corresponding Author

Corresponding authors:

Dennis Lal, PhD
Department of Neurology, McGovern Medical School UTHealth Houston, Houston, TX, USA.
Center for Neurogenetics, The University of Texas Health Science Center at Houston, Houston, TX, USA
Email: Dennis.Lal@uth.tmc.edu
Tel: 713-500-8916
1133, John Freeman Blvd, Houston, TX, 77030

Stephen F. Traynelis
Department of Pharmacology and Chemical Biology, Emory University School of Medicine, Atlanta, GA 30322, USA.
Email: strayne@emory.edu
Tel: 404.727.0357 (Office); 404.727.1375 (Lab); 404.727.0365 (Fax)
Emory University School of Medicine, 1510 Clifton Rd. NE, Atlanta, GA 30322, USA.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Abstract

Genetic variants in genes *GRIN1*, *GRIN2A*, *GRIN2B*, and *GRIN2D*, which encode subunits of the N-methyl-D-aspartate receptor (NMDAR), have been associated with severe and heterogeneous neurologic diseases. Missense variants in these genes can result in gain or loss of the NMDAR function, requiring opposite therapeutic treatments. Computational methods that predict pathogenicity and molecular functional effects are therefore crucial for accurate diagnosis and therapeutic applications. We assembled missense variants: 201 from patients, 631 from general population, and 159 characterized by electrophysiological readouts showing whether they can enhance or reduce the receptor function. This includes new functional data from 47 variants reported here, for the first time. We found that pathogenic/benign variants and variants that increase/decrease the channel function were distributed unevenly on the protein structure, with spatial proximity to ligands bound to the agonist and antagonist binding sites being key predictive features. Leveraging distances from ligands, we developed two independent machine learning-based predictors for NMDAR missense variants: a pathogenicity predictor which outperforms currently available predictors (AUC=0.945, MCC=0.726), and the first binary predictor of molecular function (increase or decrease) (AUC=0.809, MCC=0.523). Using these, we reclassified variants of uncertain significance in the ClinVar database and refined a previous genome-informed epidemiological model to estimate the birth incidence of molecular mechanism-defined GRIN disorders. Our findings demonstrate that distance from ligands is an important feature in NMDARs that can enhance variant pathogenicity prediction and enable functional prediction. Further studies with larger numbers of phenotypically and functionally characterized variants will enhance the potential clinical utility of this method.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Introduction

Pathogenic variants in the GRIN family of genes encoding the N-methyl-D-aspartate receptor (NMDAR) subunits have been found in patients with various neuropsychiatric disorders, including autism spectrum disorders, epilepsy, intellectual disability, attention-deficit/hyperactivity disorder, and schizophrenia [1–8]. NMDARs are tetrameric ligand-gated ion channels permeable to Na^+, K^+, and Ca^{2+}, composed of two glycine-binding GluN1 subunits and two glutamate-binding GluN2 subunits, which can be a combination of any two of GluN2A, GluN2B, GluN2C, or GluN2D[9,10]. GluN1 subunits are encoded by the gene *GRIN1* and GluN2 subunits are encoded by the genes *GRIN2A, GRIN2B, GRIN2C*, and *GRIN2D*. Among the pathogenic variants identified in the GRIN gene family, those in *GRIN2A* (46%) and *GRIN2B* (38%) account for the vast majority, followed by *GRIN1* variants (14%)[11]. Variants in these genes have been associated with a spectrum of neurodevelopmental disorders[12]. For example, *GRIN1* and *GRIN2B* patients can present with mild or severe intellectual disabilities[1,3,13]. While some *GRIN2A* patients have severe intellectual disabilities, roughly half have no intellectual impairment[8,14]. Most patients with variants in *GRIN2A* have seizures whereas the majority of patients with variants in *GRIN2B* do not have seizures[4]. In addition, low muscle tone is rare among *GRIN2A* patients while common among patients involving other GRIN genes[15,16]. All GRIN patients present with speech impairment, even those without intellectual disabilities [17]. *GRIN2D* patients appear to have the most severe phenotype, although there is not yet enough data to understand the full clinical spectrum[15,16]. More recently, patients with protein-truncating variants in *GRIN2A* have been associated with schizophrenia[18] and are susceptible to seizures and delayed maturation of parvalbumin interneurons, both of which resolve after adolescence[19].

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

NMDARs serve many cellular functions, including both pre-synaptically to influence neurotransmitter release and long-term plasticity[20–22] and post-synaptically to mediate the slow component of postsynaptic currents and synaptic plasticity[23,24]. NMDARs function as signal coincidence detectors, as their activation requires changes in membrane potential to relieve the pore blocking by Mg^{2+} ions as well as the synaptic release of glutamate [10,25,26]. Thus, NMDARs are precisely regulated by the binding of glutamate and its co-agonist glycine, of extracellular Mg^{2+} that blocks the channel, and of other endogenous extracellular modulators such as zinc ions (Zn^{2+})[27]. Genetic variants in GRIN genes can cause a heterogeneous spectrum of alterations of the NMDAR function, which can be grouped into two main types: gain of the NMDAR function (or gain-of-function effect or GoF) and partial or complete loss of the NMDAR function (or loss-of-function effect or LoF) [28]. Because of the many functions and modulators of NMDARs, a large number of NMDAR-based molecules have been developed as therapeutic options designed to mitigate dysfunction of the glutamatergic system [29,30].

Variant interpretation in the GRIN genes, both for pathogenicity and molecular functions, is still challenging. Currently, more than 65% of missense variants in GRIN genes are classified as variant of unknown significance according to the ClinVar database (accessed July 2022)[31]. To improve variant interpretation, several exome-wide bioinformatic approaches have been developed that can identify clusters of patient variants or population variant depletion across genomic, protein sequences or 3D structures[32–37]. These methods and targeted clinical-genetic studies for *GRIN2A* and *GRIN2B* showed enrichment of pathogenic over population variants in several UniProt defined domains[13,37,38]. Although these approaches can identify important regions, they typically don't explore the underlying structure-to-function relationship. However, in the absence of functional characterization of every possible genetic variant in NMDAR

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

encoding genes, prediction models are needed to enable precision care since different disease mechanisms have contraindicated treatment needs[39–42]. For example, among patients with developmental and epileptic encephalopathy, those with a variants that cause gain of NMDAR function represent candidates for potential treatment with NMDAR blockers, such as memantine[43–45] or GluN2B-selective inhibitors[42,46], while those with complete or partial loss of the NMDAR function may potentially respond to positive allosteric modulators of the NMDAR [47–49].

However, the functional consequences are not known for most of the variants. Electrophysiological studies that experimentally determine the molecular functional consequences introduced by missense variants are expensive and time-consuming, and it is difficult to envision how all the possible GRIN missense variants can be functionally assessed. Machine learning (ML)-based methods may be able to take advantage of the limited available experimental data to predict the molecular functional consequences of the NMDAR variants that have not been experimentally tested, as has been proven successful for example in voltage-gated potassium channels[50] and sodium channels[51]. These methods are based on sequence and structural protein features, as discriminative gene-level and protein-level features have been found to be associated with the GoF/LoF effects of variants [52–54]. To date, few predictors are available for variant functional effects (VPatho[55], LoGoFunc https://www.biorxiv.org/content/10.1101/2022.06.08.495288v1.full.pdf) and none are specifically designed for GRIN variants. Here we seek to identify structural features of missense variants that are predictive of variant pathogenicity and of increased or decreased functional effects that are specific for the GRIN genes to develop a ML-based method to predict pathogenicity and Increased/Decreased consequences in GRIN genes. Therefore, we aggregated

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

a unique set of 201 expert-curated patient variants and 631 population variants from the gnomAD database together with 159 functionally characterized missense variants from electrophysiological readouts across *GRIN1*, *GRIN2A* and *GRIN2B*. Some of these variants were functionally characterized in this study for the first time. Since previous work[12,56–60] described how individual missense variants in NMDARs alter their ligand-induced regulation, we first sought whether spatial distance from agonists that are bound to their binding sites in the NMDAR protein structure correlates with pathogenicity and increased/decreased functional effects. With the identified distances from ligands and additional biophysical and evolutionary scores, we have built a ML-based predictor specifically trained to classify variants in GRIN genes as pathogenic or benign, and a ML-based predictor specifically trained to classify the molecular consequences of missense variants into increased or decreased effect, thus providing a valuable resource for clinical genetics.

Material and Methods

Clinical data set of GRIN missense variants

Clinical dataset. The variant data set comprises manually curated patient missense variants collected through patient registries. Clinical cases were collected using a REDCap survey with 391 fields on genetic and clinical data (REDCap version 10.9.3 https://www.project-redcap.org/) and offer the integration of retrospective and longitudinal data. All clinical cases were manually reviewed and classified in accordance with American College of Medical Genetics (ACMG) guidelines. These variants were found in the genes *GRIN1, GRIN2A*, and *GRIN2B*. Missense variants from the general population in the same genes were derived from the gnomAD database[61]. In order to map the variants on the NMDAR structure, we restricted our analysis to

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

missense variants that were localized in the domains of the NMDAR structure that are atomically resolved in the PDB: the amino-terminal (ATD), agonist-binding (ABD), and transmembrane (TMD) domains. Missense variants that are localized in C-terminal domain (CTD) could not be considered as the CTD is not present in any resolved NMDAR structure. This comprises a set of 832 missense variants, of which 201 from patients and 631 from controls. The full list of the considered variants along with their clinical annotation is available and interactively accessible at https://GRIN-portal.broadinstitute.org and and also reported in Supplemental Table S1.

Functional data set of GRIN missense variants

We identified 127 missense variants in *GRIN1*, *GRIN2A* and *GRIN2B* for which functional testing through electrophysiological and biochemical assays were published and completed to an extent to allow GoF and LoF determination by the criteria of Myers et al. 2023[28] (Supplemental Table S2). We also identified additional variants in the literature for which some functional data exists, but which lack completion of all assays needed to classify by criteria described in Myers et al. (2023; Supplemental Table S3). Among these two sets of variants, we present new data allowing completion of functional and biochemical assessment of 34 known missense variants according to the criteria of Myers et al. (2023; Supplemental Table S4, S5). From these determinations, we categorized a total of 159 variants as either having an increased or decreased ion channel function of the response for our implementation here by evaluating the criteria described in Myers et al. (2023)[28], omitting consideration of surface expression, since our algorithms explore receptor function as it relates to protein structure, not factors that influence trafficking. Thus, functional data from 159 variants is the starting point for our functional analyses.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Functional analysis of missense variants. cDNAs encoding human NMDAR subunits GluN1-1a (hereafter GluN1; GenBank accession codes: NP_015566), GluN2A (GenBank accession codes: NP_000824), and GluN2B (GenBank accession codes: NM_000825) were used and site-directed mutagenesis used to introduce human variants (QuikChange; Stratagene, La Jolla, CA, USA); all mutant cDNAs were verified by dideoxy DNA sequencing (Eurofins MWG Operon, Huntsville, AL, USA). The cDNA for WT and mutant NMDAR subunits was linearized using the appropriate restriction enzyme and cRNA was synthesized *in vitro* using the mMessage mMachine T7 kit (Ambion, Austin, TX, USA).

Xenopus laevis ovaries with unfertilized oocytes (Stage V-VI) were obtained from Xenopus One Inc (Dexter, MI, USA) and digested with Collagenase Type 4 (Worthington-Biochem, Lakewood, NJ, USA; 850 µg/ml, 15 ml for a half ovary) in Ca^{2+}-free Barth's solution that contained (in mM) 88 NaCl, 2.4 $NaHCO_3$, 1 KCl, 0.82 . maintained at 16°C and injected with cRNA encoding either WT or variant NMDAR subunits (GluN1:GluN2A or GluN2B ratio 1:2, 5-10 ng total in 50 nl of RNAase-free water per oocyte). Injected oocytes were maintained in normal Barth's solution at 16-19°C.

Two-electrode voltage clamp (TEVC) current recordings from Xenopus oocytes expressing NMDARs were performed as previously described [62,63]. Oocytes were transferred to a recording chamber and were perfused with extracellular recording solution composed of (in mM) 90 NaCl, 1 KCl, 0.5 $BaCl_2$, 10 HEPES, and 0.01 EDTA (23°C, pH 7.4 with NaOH, EDTA omitted in experiment measuring $Mg^{2+}IC_{50}$). Current responses to glutamate and glycine were recorded under voltage clamp at a holding potential of -40 mV; current and voltage electrodes were filled with 3 and 0.3 M KCl, respectively. Maximally effective concentrations of agonists (100 µM

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

glutamate and 100 μM glycine) were used unless stated otherwise. The reagent 2-aminoethyl methanethiol sulfonate hydrobromide (MTSEA; Toronto Research Chemicals, Ontario, Canada) was made fresh and used within 30 min.

HEK293 cells (ATCC CRL-1573) were plated on glass coverslips coated with 0.1 mg/ml poly-D-lysine and maintained in Dulbecco's modified Eagle medium (DMEM) with 10% fetal bovine serum and 10 U/ml streptomycin at 37°C (5% CO_2). The cells were transfected with cDNA encoding GluN1, GluN2A, and eGFP at a ratio of 1:1:5, or GluN1, GluN2B, and eGFP at a ratio of 1:1:3 using the calcium phosphate method[62]. 12-24 hrs post transfection the cells were transferred to the recording chamber and perfused with recording solution composed of (in mM) 150 NaCl, 3 KCl, 1.0 $CaCl_2$, 10 HEPES, 0.01 EDTA, and 2.10 D-mannitol (the pH was adjusted to 7.4 with NaOH). 3-5 MΩ fire-polished patch electrodes were made from thin-walled micropipettes (TW150F-4, World Precision Instruments, Sarasota, FL, USA) and filled with internal solution composed of (in mM) 110 D-gluconate, 110 CsOH, 30 CsCl, 5 HEPES, 4 NaCl, 0.5 $CaCl_2$, 2 $MgCl_2$, 5 BAPTA, 2 NaATP and 0.3 NaGTP (the pH was adjusted to 7.4 with CsOH; osmolality was 300-305 mOsmol/kg). Whole cell currents in response to application of 1.0 mM glutamate and 100 μM glycine (V_{HOLD} -60 mV, 23°C) were recorded by an Axopatch 200B patch-clamp amplifier (Molecular Devices, Union City, CA, USA). The current responses were filtered at 8 kHz (-3 dB) with an 8-pole Bessel filter (Frequency Devices) and digitized at 20 kHz on a Digidata 1440A system controlled by Clampex 10.3 (Molecular Devices). The position of double-barreled theta-glass tubing was controlled by a piezoelectric translator to obtain rapid solution exchange (Burleigh Instruments, Newton, NJ, USA). Large current responses were corrected off-line for series resistance errors[64].

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

To measure receptor surface expression, HEK 293 cells grown in 96-well plates (50,000 cells/well) were transfected with cDNA encoding beta-lactamase (β-lac) fused in frame to the N-terminus of GluN1 with WT or mutant GluN2, or similarly constructed β-lac-GluN2 with WT or mutant GluN1 using Fugene6 (Promega, Madison, WI)[56]. Wells treated with Fugene6 alone without cDNA were used to determine background absorbance. NMDAR antagonists (200 μM DL-APV and 200 μM 7-CKA) were added to cultures when transfected. Six wells were transfected for each variant to determine surface and total protein levels (3 wells each). After 24 hours, cells were rinsed with Hank's Balanced Salt Solution (HBSS) that was composed of (in mM) 140 NaCl, 5 KCl, 0.3 Na_2HPO_4, 0.4 KH_2PO_4, 6 glucose, 4 $NaHCO_3$ with 10 mM HEPES added. Subsequently, 100 μl of a 100 μM nitrocefin (Millipore, Burlington, MA, USA) solution in HBSS plus HEPES was added to each of the wells and extracellular enzymatic activity was determined. The cells in the three wells were lysed by 50 μl H_2O (30 min) prior to the addition of 50 μl of 200 μM nitrocefin to determine total enzymatic activity. The absorbance at 468 nm was read every min for 30 min at 30°C, and the rate of increase in absorbance was determined from the slope of a linear fit to the data.

Structural localization of variants

In order to investigate correlations between functional effects (pathogenicity and molecular function) and structural features of NMDAR missense variants, we localized each variant onto the 3-dimensional (3D) protein structure using SIFTS tools[65] to cross reference amino acid positions between protein sequences and protein structures. We used two crystal structures of the human NMDAR available in the PDB: 7EU7[66] (3.50 Å of resolution) which comprises two *GRIN1*-encoded and two *GRIN2A*-encoded chains, and 7EU8[66] (4.07 Å of resolution) which

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

comprises two *GRIN1*-encoded chains and two *GRIN2B*-encoded chains. We used chain A and B of 7EU7[66] to localize *GRIN1* and *GRIN2A* variants respectively, and chain B of 7EU8 to localize *GRIN2B* variants. Of the variants in our collected clinical dataset, 832 could be mapped onto the heterotetrametric protein structure. Of these, 201 are from patients and 631 from population.

Computation of distance features for missense variants

For each variant that could be mapped onto the 3D structure, we computed features that synthesize information about their position in the 3D structure with respect to the functionally important regions of the protein complex. In particular, we computed their distances from: the membrane center, the pore axis, and the four ligands known to primarily regulate NMDAR function (glutamate, glycine, magnesium ion Mg^{2+}, and zinc ion Zn^{2+}). We used the protein complex 7EU7[66] to compute distances for *GRIN1* and *GRIN2* variants and 7EU8[66] for *GRIN2B* variants. We annotated the membrane through the PPM server (https://opm.phar.umich.edu/ppm_server, version PPM 2.0)[67] and calculated the minimum distance of the wild-type residue of each variant from the membrane center as defined by the PPM server. We annotated the pore using the Mole2.5 webserver[68] (https://mole.upol.cz/, access 24.03.2022) and calculated the minimum distance of the wildtype residue of each variant from the pore axis using R-script and the "bio3D" package[69]. To calculate the minimum distance of the wildtype residue of each variant from the ligands (glutamate, glycine, Mg^{2+}, Zn^{2+}) we first mapped these ligands to the corresponding reference protein structures 7EU7 and 7EU8. Two ligands, glutamate and glycine, were already crystallized in the protein complex 7EU7. To map these two ligands to the 7EU8 protein complex, we performed a structural alignment of the two protein complexes (7EU7 and 7EU8) using mTM-align[70] (root mean square error (RMSD):

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

3.22Å). To place Zn^{2+}, we carried out two structural alignments of the amino terminal domain of the NMDAR crystallized with two Zn^{2+} ions bound to GluN2A (PDB-ID: 5TQ2, resolution of 3.29 Å) with 7EU7[66] and 7EU8[66], respectively. The structural alignments were performed with the mTM-align[70] program ($RMSD_{7EU7}$ = 3.01Å, $RMSD_{7EU8}$ = 2.22Å). Since no NMDAR structure has been crystallized with Mg^{2+} inside yet, we calculated the estimated the location of the Mg^{2+} ions by calculating the center of the described magnesium binding residues (N615 *GRIN2A*, N616 *GRIN2B* and N616 *GRIN1)*[2,71]. For each missense variant, we then calculated the minimum distance of its wild-type residue from each ligand. The distance calculation was restricted to residues located in the same domain. Domain annotations are detailed in Supplementary Table S1 and are: amino terminal domain (ATD), agonist binding domain (ABD), ABD-TMB linkers (S1 and S2), and transmembrane domain (TMD) comprising M1, M2, M3, and M4 helices. All distances were computed considering all atoms. The considered distance features are tabulated in Table 1.

Computation of biophysical and evolutionary features for missense variants

In addition to distance features, we also computed biophysical and evolutionary scores. For each variant, we computed three biophysical features: the relative solvent accessibility of the wild-type residue computed through the DSSP program, the difference in hydrophobicity between the wild-type and the substituted residues according to the Kyte-Dolittle scale and the difference in the interaction energy (computed as the Bastolla-Vendruscolo statistical potential[72]) between the wild-type and mutate residue with their structural environment (5 Å sphere centered in the mutated residue). These two latter scores have been weighted through the sequence profile as in Montanucci *et al.*, 2022[69]. The biophysical features are listed in Table 1. In addition to

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

biophysical scores, we computed, for each variant in our clinical and functional dataset the BLOSUM62[75] and EVE[76] evolutionary conservation scores. These scores for each variant are tabulated in Supplementary Table 1.

Table 1 List of the considered distance, biophysical, and evolutionary features for missense variants.

Feature	Short name	Description
		Distance features
Distance from glutamate	dist_glu	Distance (Å) of the wild-type residue from glutamate ligand using all-atoms.
Distance from glycine	dist_gly	Distance (Å) of the wild-type residue from glycine ligand using all-atoms.
Distance from Mg^{2+}	dist_mg	Distance (Å) of the wild-type residue from magnesium ligand using all-atoms.
Distance from Zn^{2+}	dist_zn	Distance (Å) of the wild-type residue from zinc ligand using all-atoms.
Distance from pore	dist_pore	Distance (Å) of the wild-type residue from the pore axis using all-atoms.
Distance from membrane	dist_mem	Distance (Å) of the wild-type residue from membrane using all-atoms.
		Biophysical features
Relative solvent accessibility	RSA	Relative solvent accessibility of the wilt-type residue computed through the DSSP program (Kabsch and Sander 1983) [77].
Hydrophobicity difference	KD	Hydrophobicity score accounting for the difference in hydrophobicity between wildtype and mutate residue as in Montanucci *et al.*, 2019[73].
Energy difference	E3D	Difference in the interaction energy between the wild-type and mutate residue with their structural environment (5 Å sphere) as in Montanucci *et al.*, 2019[73]
		Evolutionary features
Conservation difference	S_BLOSUM62	Difference between the wild-type and variant residue in the BLOSUM62 substitution matrix as in Montanucci *et al.*, 2019[73].
EVE evolutionary model	EVE	Score from the deep learning based and unsupervised EVE model[78].

Enrichment analysis

We performed an enrichment analysis to reveal the distribution of pathogenic and population variants and of variants with increased and decreased functional consequence on the NMDAR

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

structure. To perform the enrichment analysis, we applied the Wilcoxon rank sum test to investigate differences in patient and population variants and their 3D distances from functionally relevant protein region sites (pore axis and membrane) and the four selected ligands. Similarly, we tested differences for variants associated with an increased or decreased functional effect. We applied Bonferroni correction to account for multiple testing.

Developing of ML-based binary classifiers for pathogenicity and functional prediction

Pathogenicity predictor In order to develop a binary ML-based predictor which classifies missense variants in the GRIN genes as benign or pathogenic, we trained each of four binary predictors, each one on a different set of features, listed in Table 2. PP-dist, PP-evo, PP-biophys trained on only distances from ligands and pore axis, evolutionary scores, biophysical features, respectively. Finally, PP-dist&evo is trained on the combination of features of the two best performing predictors, which are distances and evolutionary features. All the predictors have been trained on the subset of the clinical dataset comprising variants that could be mapped on the 3D structure. A 5-fold cross-validation procedure was applied.

Functional predictor In order to develop a binary ML-based predictor which classifies missense variants in the GRIN genes as increasing or decreasing functional effect, we trained four binary predictors, each one trained on a different feature set, listed in Table 2. FP-dist is trained on distances from the four ligands: glutamate, glycine, Mg^{2+}, and Zn^{2+}; FP-evo is trained on evolutionary scores; FP-biophys is trained on the three biophysical features of the variant. Finally, FP-dist&evo is trained on distance and evolutionary features. The predictors have been trained on the functional dataset comprised of 159 variants, of which 89 with increased and 71 with decreased functional effect (see Supplementary Table 2). Details about machine learning

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

implementation, cross-validation and indexes to evaluate performances are described in the Supplementary Notes. In order to compare the performance with existing variant functional effects predictors, we retrieved the prediction of our 159 functionally characterized variants of LoGoFunc. At the moment of submission VPatho was not available.

Table 2 Description of the input features for each pathogenicity and functional effect predictor.

Predictor	Input Features*
	Pathogenicity Predictor (PP)
PP-dist	dist_glu, dist_gly, dist_mg, dist_zn, dist_pore
PP-evo	S_BLOSUM62, EVE
PP-biophys	RSA, KD, E3D
PP-dist&evo	dist_glu, dist_gly, dist_mg, dist_zn, dist_pore, S_BLOSUM62
	Functional (Increase/Decrease) Predictor (FP)
FP-dist	dist_glu, dist_gly, dist_mg, dist_zn
FP-evo	S_BLOSUM62, EVE
FP-biophys	RSA, KD, E3D
FP-dist&evo	dist_glu, dist_gly, dist_mg, dist_zn, S_BLOSUM62, EVE

*Short names of the biophysical features are derived from Table 1.

Results

Spatial proximity to ligands and pore is different in pathogenic versus benign variants

Here we asked whether the spatial distance from important ligands, such as glutamate, glycine, magnesium and Zn^{2+} and functional important sites-such as the pore-axis correlate with variant pathogenicity in the *GRIN1*, *GRIN2A* and *GRIN2B* determined using ACMG criteria. We calculated the 3D distances of the 832 variants that could be mapped on the deposited NMDAR structures from the pore, the membrane and the four ligands that regulate the NMDAR activity and we compared the residue distances from these sites between the two groups of variants, an expert-curated set of patient variants (see Supplemental Tables S2-6).

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

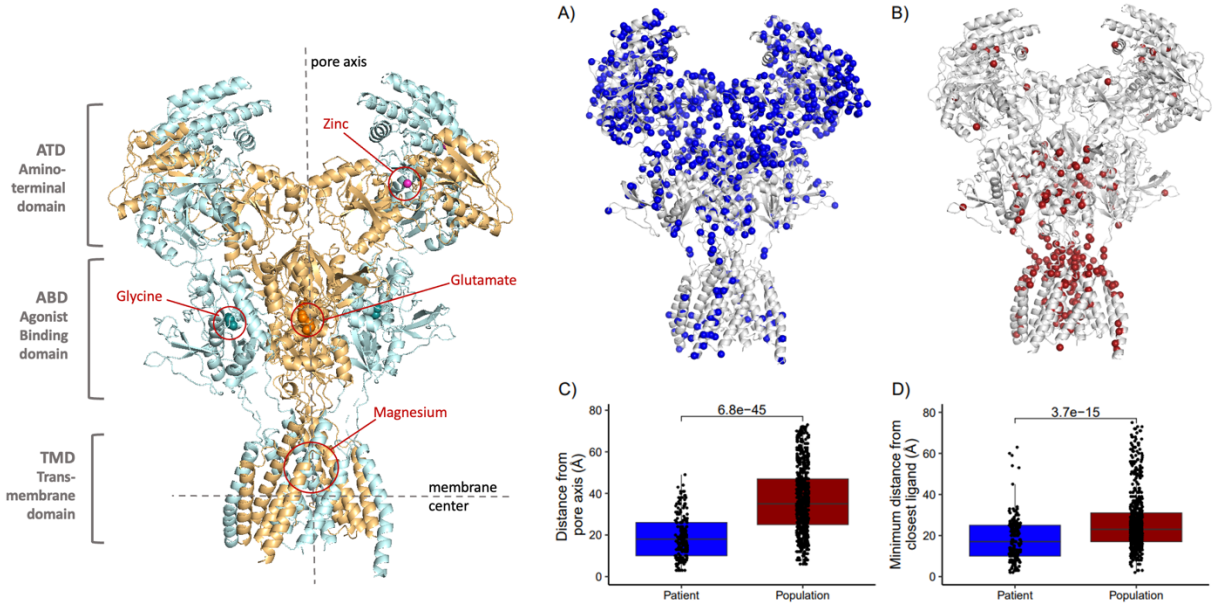

Figure 1 Variant distance from the pore axis and from ligands significantly differs between patient and population variants in the GRIN-genes. Left image: GluN1-GluN2A heterotetramer protein complex (PDB-ID: 7eu7). GluN1 subunits are colored in cyan, GluN2A subunits are colored in orange. The four ligand binding sites are highlighted. **A)** NMDAR structure with the 631 population variants of this study highlighted in blue. **B)** NMDAR structure with the 201 patient variants of this study highlighted in red. **C)** Boxplot of the distance from the pore axis for 201 patients and 631 population missense variants in *GRIN1*, *GRIN2A*, and *GRIN2B*. **D)** Boxplot of the minimum distance from the closest ligand (glutamate, glycine, $Zn^{2+}$and Mg^{2+}) for 203 patients and 631 population missense variants in *GRIN1*, *GRIN2A*, and *GRIN2B*. To quantify the differences in the distances to the ligands and the pore axis we performed the Wilcoxon-rank sum test and corrected for eight tests using Bonferroni correction.

Compared to the spatial distribution of population variants, patient variants where closer to the pore (Population variants $_{Median\ distance}$ = 35 Å, Patient variants $_{Median\ distance}$ = 18 Å, P =6.8e-45, Wilcoxon rank sum test, Figure 1B) and to the closest ligand-(Population variants $_{Median\ distance}$ = 23 Å, Patient variants $_{Median\ distance}$ = 17 Å, P = 3.7e-15, Wilcoxon rank sum test, Figure 1C). When performing the same enrichment analysis for each domain separately, we found that in the agonist binding domain, *GRIN2A* and *GRIN2B* patient variants in the agonist binding domain are located closer to glutamate, which is bound in the cleft of the bilobed agonist binding domain,

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

compared to population variants (P_{GRIN2A}= 4.3e-14, and P_{GRIN2B}= 4.1e-03, Figure 2A), while no significant difference in proximity to glycine was found for variants in *GRIN1*. In the amino terminal domain, no significant difference is found in proximity to ligands between patient variants. However, this result could be due to the small number of population variants in this domain. In the transmembrane binding site, we observed that patient variants are closer to the Mg^{2+} binding site compared to population variants in *GRIN2A* (P = 2.6e-02, Figure2D).

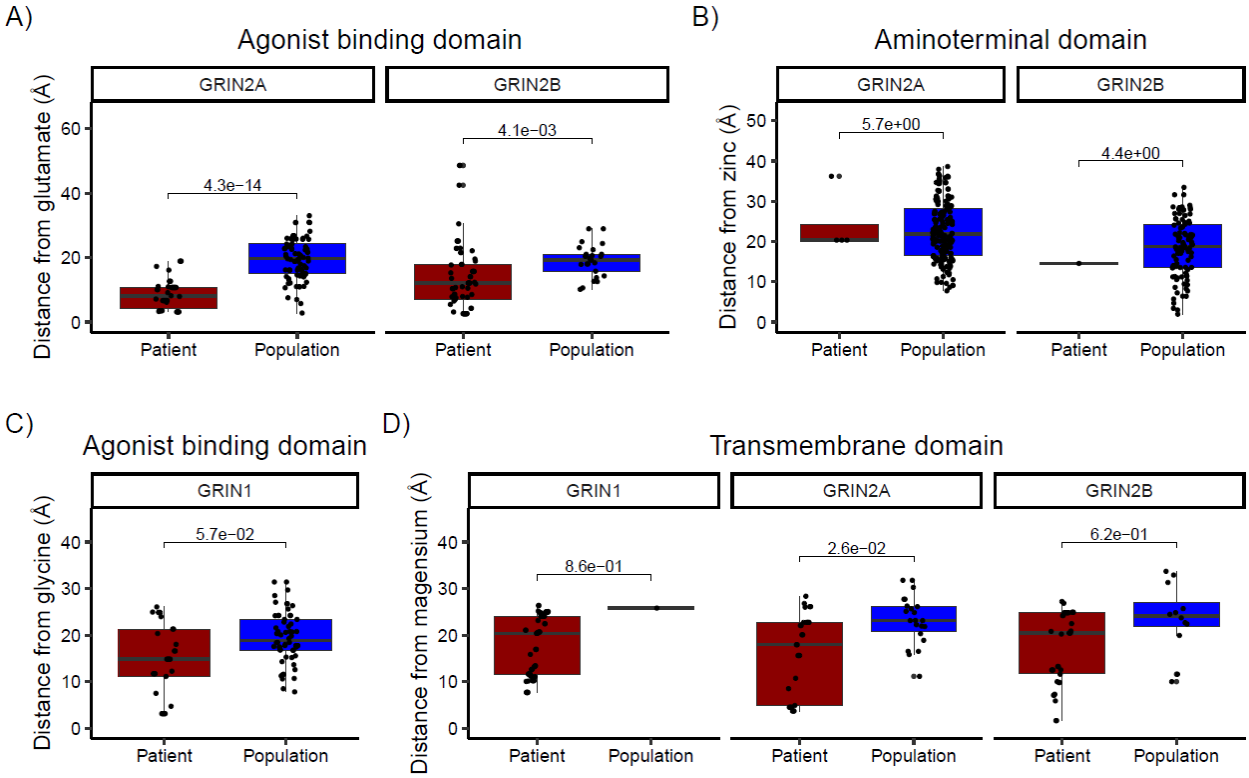

Figure 2 Variant distance from NMDAR ligands significantly differs between patient and population variants in the GRIN-genes. Boxplot of the distances of patient and population variants from four different NMDAR ligands: **A)** glutamate, **B)** Zn^{2+}, **C)** glycine, and **D)** Mg^{2+}. Only variants that are located in the same domain where the ligand is bound are shown and the distance is only computed for the protein subunits that are involved into the ligand binding. **A)** Distances from glutamate of patient and population variants that are located in the agonist binding domain (ABD). **B)** Distances from the Zn^{2+} of patient and population variants that are located in the amino terminal domain. **C)** Distances from glycine of patient and population variants that are located in the agonist binding domain (ABD). **D)** Distances from Mg^{2+} of patient and population variants that are located in the transmembrane

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

domain (TMD). To quantify the differences in the distances to the ligands, we performed the Wilcoxon-rank sum test and corrected for eight test using Bonferroni correction.

Machine learning method based on ligand and pore proximity to predict variant pathogenicity

Given the strong variant position to pathogenicity associations that we observed, we explored whether our generated distance features could be used to develop a method for the prediction of variant pathogenicity in the GRIN genes using our collected clinical dataset of 201 patient and 631 population variants. We developed a ML-based method, PP-dist, based on only distance features (distance from pore axis, glutamate, glycine, Mg^{2+}ion and Zn^{2+}) to classify a GRIN variant as pathogenic or benign. As a comparison, we trained two additional pathogenicity predictors on the same clinical dataset, PP-evo trained with only evolutionary features, and PP-biophys trained on biophysical properties of the amino-acid substitution (see Table 2 for input features of each predictor). The binary classifier based on only distances from ligands, PP-dist, reaches high prediction performances, with an overall accuracy of 0.892, an area under the ROC curve (AUC) of 0.9237 and a Matthews correlation coefficient (MCC) of 0.698 (see Table 3 and Figure 3). The predictive power of these distances is therefore very high, considering that these input features are based on only the position of the substituted residue and do not contain any information on the properties of the alternative residue. This indicates that the distance from pore and ligands, are major features in determining pathogenicity of variants in the GRIN genes. While PP-biophys shows a low MCC of 0.156 indicative of a poor predictor, PP-evo shows a high MCC (0.534) and an overall accuracy of 0.832 (Table 3 and Figure 3). This suggests that biophysical features provide much less information about variant pathogenicity than evolutionary scores. When a last predictor PP-dist&evo was trained with features from the two classes of distances from pore and ligands and evolutionary, the prediction performances improved by 1%

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

in respect to the individual predictors PP-dist and PP-evo, reaching an overall accuracy of 0.903, a MCC of 0.726 and an AUC of 0.945.

To compare our developed predictors with existing methods, we also report the performances of other three among the main pathogenicity scores and predictors on the clinical dataset of this study (see Table 3 and Figure 3). This comparison shows that the simple distance measurements from ligands and pore allow a prediction accuracy that is ~10% higher than the best non-gen-specific pathogenicity predictor.

Table 3 Performances of pathogenicity predictors on 838 variants in the NMDAR.

Predictor*	**Q_2**	**MCC**	**AUC**	**TNR**	**NPV**	**TPR**	**PPV**	**$N_{Patient}$**	**$N_{Population}$**
Predictors trained on NMDAR-specific dataset									
PP-dist	0.892	0.698	0.937	0.942	0.918	0.736	0.801	159	498
PP-evo	0.832	0.534	0.873	0.925	0.895	0.565	0.724	200	572
PP-biophys	0.748	0.156	0.658	0.929	0.780	0.179	0.444	201	631
PP-dist&evo	**0.903**	**0.726**	**0.945**	**0.954**	**0.920**	**0.741**	**0.837**	**201**	**631**
Protein-unspecific Scores and Predictors									
AlphaMissense[79]	0.685	0.482	0.910	0.596	0.982	0.965	0.432	201	631
EVE[78]	0.825	0.550	0.884	0.876	0.887	0.680	0.657	200	572
REVEL[80]	0.686	0.287	0.760	0.709	0.850	0.613	0.405	199	614

* The performances of all the predictors are computed with a threshold of 0.5. The Q_2 is the overall accuracy of the predictor; MCC is the Matthews correlation coefficient; AUC is the area under the ROC curve; TNR is the true negative rate; NPV is the negative predicted value; TPR is the true positive rate; PPV is the positive predicted value (See Supplementary material for the corresponding equations). $N_{Patient}$ and $N_{Population}$ are the number of variants found in patients and in the general population in the training dataset. The best predictor is highlighted in bold.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

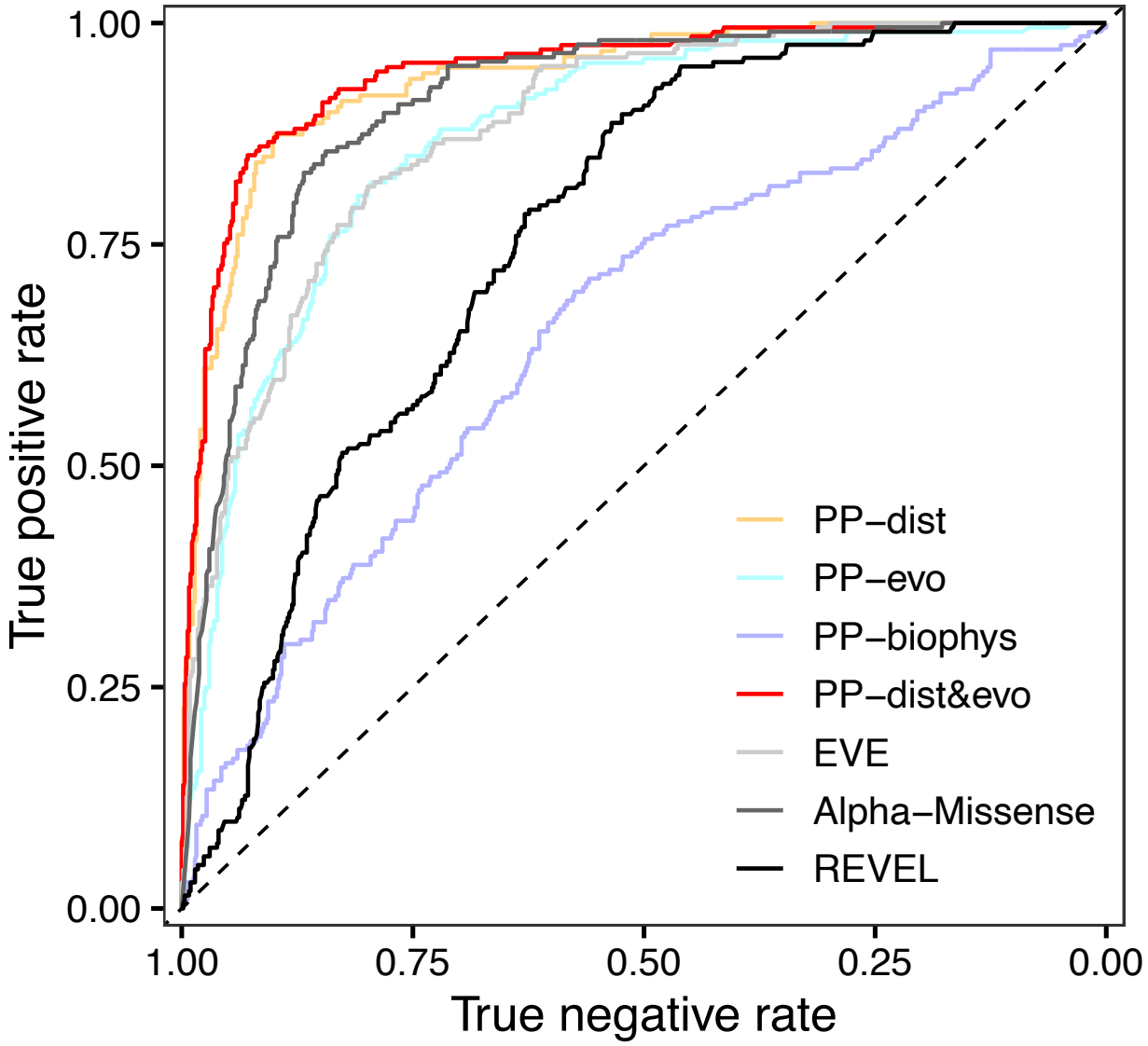

Figure 3 ROC curve for the ML-based pathogenicity predictors (PP-dist, PP-evo, PP-biophys and PP-dist&evo) and the three additional pathogenicity predictors and scores (AlphaMissense, EVE, and REVEL) on the dataset composed of the 832 GRIN variants of the clinical dataset.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Validation of the pathogenicity predictor and pathogenicity prediction for variants of uncertain significance

So far, we only used expert curated patient variants. As a further validation, we applied our best pathogenicity predictor, PP-dist&evo, to classify variants that were in ClinVar (July 2022) and, at the same time, were not part of our expert curated dataset (for prediction scores see Supplementary Table 1). The total number of ClinVar variants that were not in our clinical dataset and were classified as (likely-) benign and (likely-) pathogenic and located in our 75% most confident class assignments is 44. Of these variants, 39 were correctly classified by our method, reaching a prediction accuracy of 0.89 (Figure 4A). We then used our best model, PP-dist&evo, to classify 100 ClinVar variants that were of uncertain or conflicting significance (VUS). Predictions for VUS were distributed across the whole spectrum of the pathogenicity score (Figure 4B). We reclassified 95 VUS that were assigned with a prediction score within the 75% most confident class assignments. Out of these 95 VUS, we predicted 19% (n = 18 VUS) as pathogenic and 81% (n = 77 VUS) as benign.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

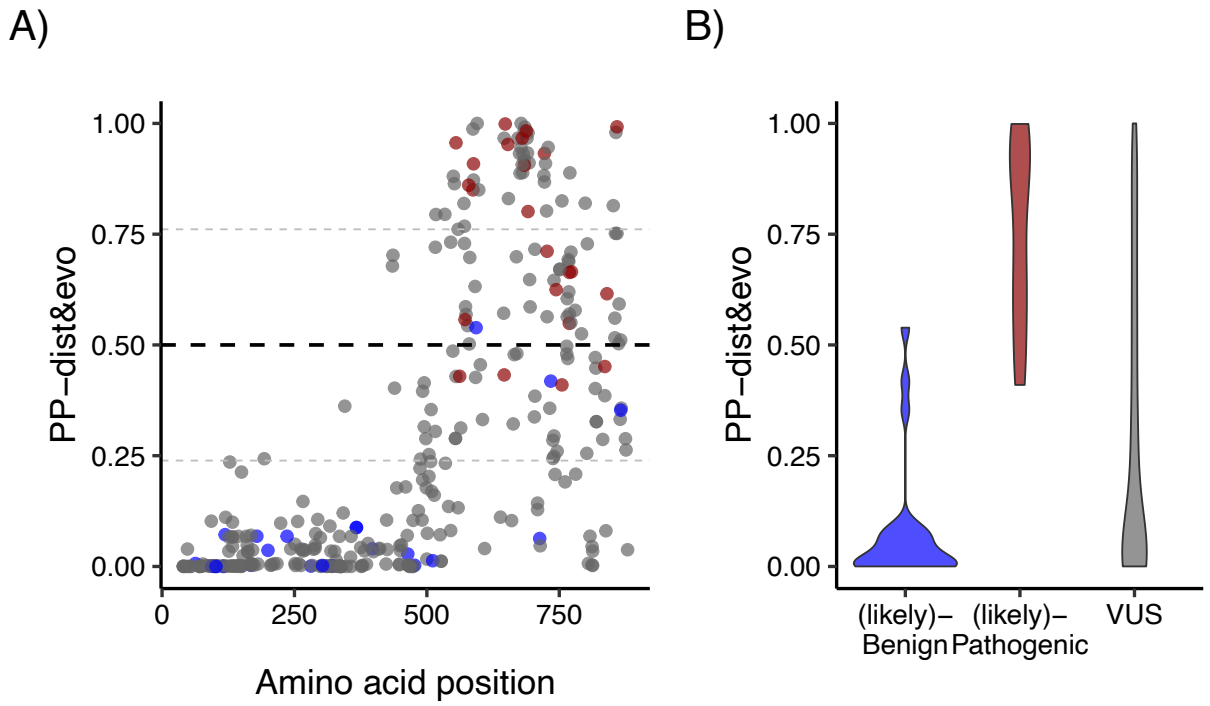

Figure 4 Pathogenicity prediction for ClinVar variants that were not part of our clinical dataset. A) Pathogenicity prediction for ClinVar variants that were not comprised in our clinical dataset. Prediction scores >0.5 indicate pathogenic variants, while predictor scores <0.5 indicate benign variants. The small grey line above and below 0.5 corresponds to the 75% most confident class assignments (over all possible amino acid variants). The prediction has been done with the best pathogenicity prediction method, PP-dist&evo. Prediction scores are displayed along the multiple sequence alignment of *GRIN1, GRIN2A & GRIN2B* (see alignment in the Supplementary Table 4). Blue, red, and grey dots correspond to ClinVar benign, pathogenic and VUS variants, respectively. **B)** Distribution of our prediction scores for ClinVar variants classified as benign, pathogenic and variants of unknown significance (VUS).

Spatial proximity to ligands and pore is different in variants which cause an increased versus decreased functional consequence.

After exploring the relationship between residue localization on the 3D-structure and pathogenicity of variants in the GRIN genes, we explored the relationship between the variant localization on the 3D-structure and the differential functional outcomes of the GRIN variants. We used data from 159 variants combining data from the peer reviewed literature and 47 variants

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

generated in this study (Supplementary Table 2). In Figure 5 the boxplot of the distance of variants with different molecular effect is shown. Variants in *GRIN2A* and *GRIN2B* with decreasing effect are significantly closest (*P*=7.2e-3) to the glutamate in respect to variants with an increasing effect. Conversely, variants with increasing effect are significantly closest (*P*=4.3e-2) to Mg^{2+} in respect to variants with a decreasing effect. We evaluated function based on potency measurements for glutamate, glycine, Mg^{2+}, and Zn^{2+}. Because there is no reason a priori to assume that distance to the pore or agonist binding site will necessarily correlate with trafficking, which involves other parts of the receptor, we did not include an assessment of surface expression. This means that our categorization as Increasing and Decreasing function will not necessarily predict patient Gain- or Loss-of-Function as defined by Myers et al. (2023)[28]. Rather, this predictor will indicate whether a missense variant is likely to alter function of a receptor once it reaches the cell surface. In addition, for a subset of 69 variants, we did not have a measure of variant actions on deactivation time course, however this correlates with agonist EC_{50} and thus this effect is in part captured by our measure of potency in these variants (Xu et al., 2024) [81,82]. We aggregated electrophysiological readouts from published and newly recorded data (Supplemental Table S1-3) in this study for 159 variants and classified all variants according to their molecular effect as described in the methods. We found the number of variants could be categorized as follows: $N_{Increase} = 89$, $N_{Decrease} = 71$. The results for this set of variants are summarized in Supplementary Tables S2-5. We observed that proximity to the pore, and in particular to the Mg^{2+} binding sites are associated with variants that “increased” function effects and depleted for variants that “decreased" function in all the GRIN genes (Figure 5D). In *GRIN2A* (P-value = 0.01), variants that “decreased” function are associated with close proximity to glutamate binding sites, compared to variants that “increased” function (Figure 5A).

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

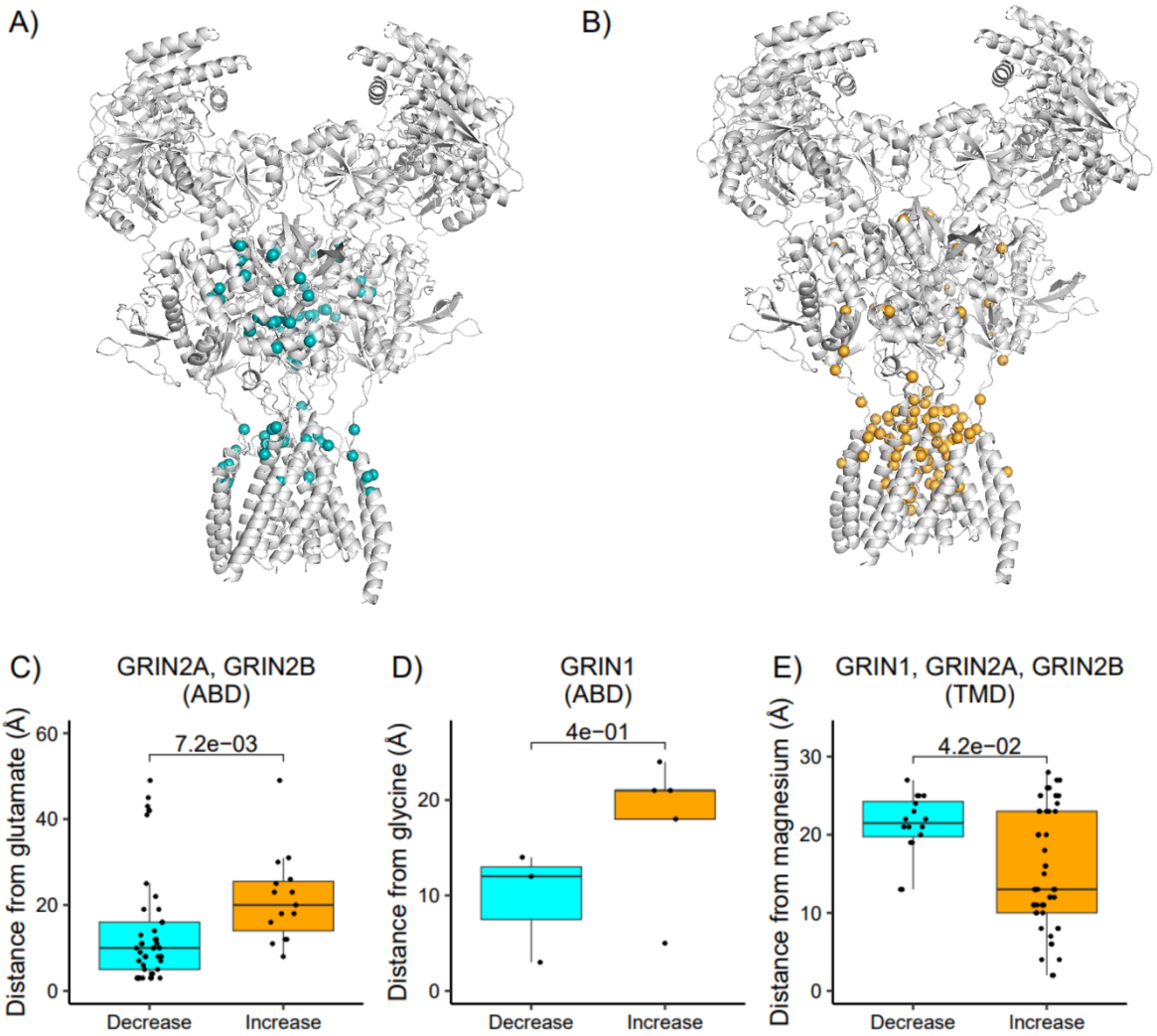

Figure 5: Variant distances from NMDAR ligands significantly differ between variants with different electrophysiological readouts in the GRIN-genes. In plot **A)** and **B)** variants whose molecular effect was classified into either decreasing **(A)** or increasing **(B)** effect are shown on the NMDA structure. In plots **C)**, **D)** and **E)** the boxplot of the distances between the variant and the ligands for the variants with different molecular effect are shown. Only variants that are located in the domain where the target ligand binds are considered. **C)** distance from glutamate, **D)** distance from glycine, and E) distance from Mg^{2+}.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

ML-based method using ligand and pore proximity to predict functional effect of variants

With a similar procedure as that for pathogenicity predictor, we independently developed four different binary classifiers to predict the functional consequences of variants in *GRIN1*, *GRIN2A* and *GRIN2B* described in Table 2. All performances are shown in Table 4 and Figure 6. While the predictors based on only evolutionary or biophysical features (FP-evo and FP-biophys, respectively) are poor predictors (with MCC of 0.269 and 0.404, respectively), the FP-dist predictor, based only on the 3D distances of variants from each of the four considered ligands, reaches an overall accuracy of 0.740 and a MCC of 0.482. This shows here, for the first time, that distances of the variant from ligands contain information about the direction ("increasing" or "decreasing" activity) of the functional effect of the variant and that can be used for functional effect prediction.

If we combine distance and evolutionary scores, our final best functional predictor for GRIN variants, FP-dist&evo reaches an overall accuracy of 0.765 and a MCC of 0.523 (for predictions see Supplementary Table 1). As a comparison, are also reported in Table 4 the performances of the pathogenicity predictor Alphamissense and of the only other available non-protein specific functional (LoF/GoF) predictor LoGoFunc (https://www.biorxiv.org/content/10.1101/2022.06.08.495288v1.full.pdf). These performances show that pathogenicity prediction scores like Alphamissense have little to no predictive power for the functional effects of variants and FP-dist&evo clearly outperforms, on the GRIN variants, the only other available functional predictor.

Table 4 Performances of the binary predictors for the functional effect of variants (either increased or decreased function) on the data set of 159 experimentally characterized NMDAR variants.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Predictor	Q_2	MCC	AUC	TNR	NPV	TPR	PPV	$N_{increase}$	$N_{decrease}$
Functional (Increased/Decreased function) predictors specifically trained on GRIN variants									
FP-dist	0.740	0.482	0.800	0.644	0.776	0.828	0.716	64	59
FP-evo	0. 6841	0.269	0.635	0.551	0.613	0.714	0.659	84	69
FP-biophys	0.703	0.404	0.740	0.500	0.761	0.871	0.679	85	70
FP-dist&evo	**0.765**	**0.523**	**0.809**	**0.667**	**0.780**	**0.845**	**0.755**	**84**	**69**
LoGoFunc predictor on the data set of 159 functionally tested GRIN variants									
LoGoFunc	0.522	0.217	0.510	0.971	0.482	0.161	0.875	70	87
AlphaMissense evolutionary score on the data set of 159 functionally tested GRIN variants									
AlphaMissense	0.491	0.193	0.651	1.000	0.467	0.080	1.000	71	88

The Q_2 is the overall accuracy of the predictor; MCC is the Matthews correlation coefficient; AUC is the area under the ROC curve; TNR is the true negative rate; NPV is the negative predicted value; TPR is the true positive rate; PPV is the positive predicted value (see Supplementary Material for details). $N_{increase}$ and $N_{decrease}$ are the number of variants experimentally determined as increased/decreased function in the training dataset for which all the features for each predictor were available. The best predictor is highlighted in bold.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

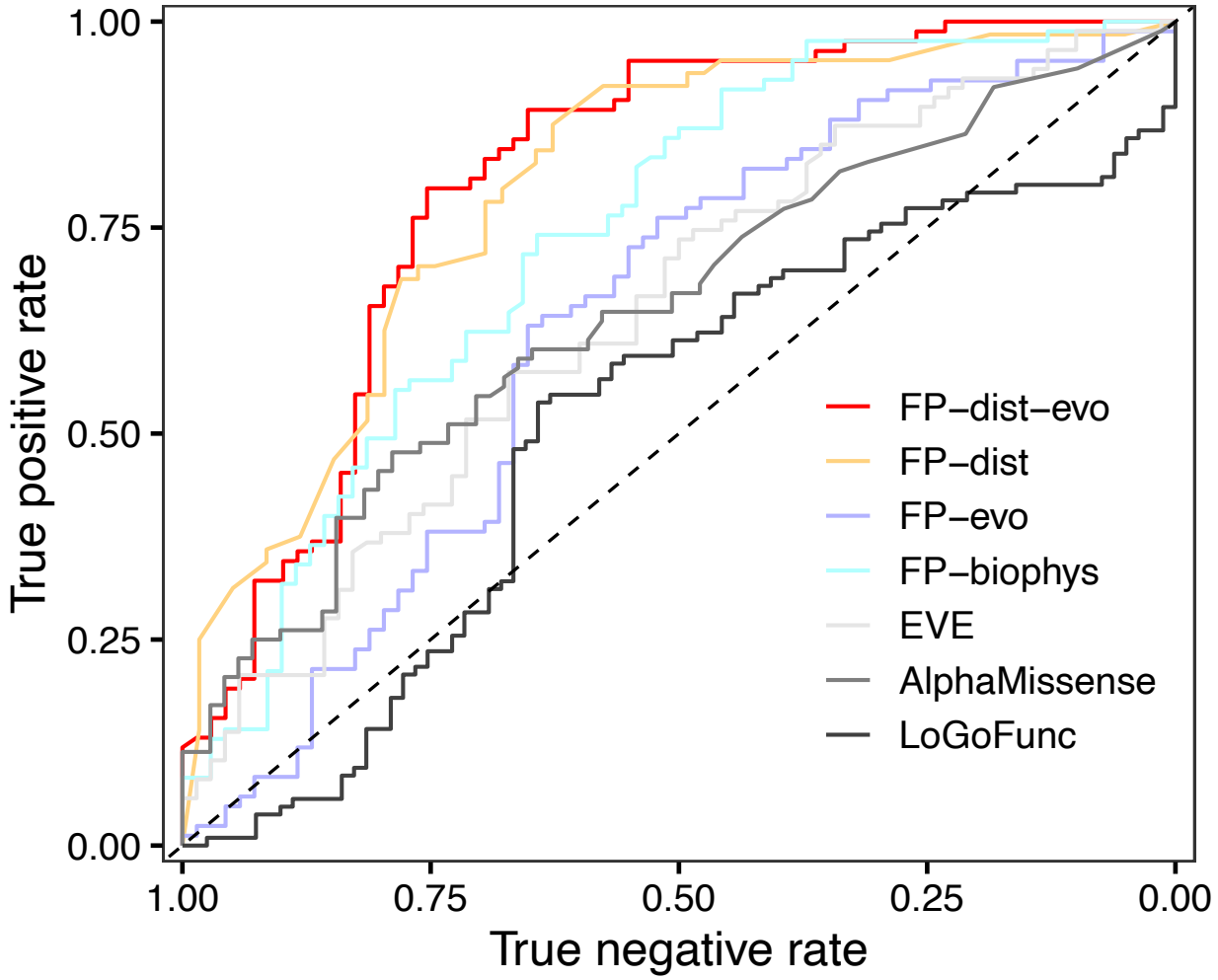

Figure 6 ROC curve for the four ML-based functional predictors (PP-dist, PP-evo, PP-biophys, PP-dist&evo), the EVE score and AlphaMissense pathogenicity predictor, and the LoGoFunc LoF/GoF predictor on the dataset composed of 159 variants in the GRIN genes.

Global incidence estimates of increase/decrease subdisorders

We previously published a mutation-rate-informed prediction model to predict the birth incidence of *de novo* variant-associated disorders such as GRIN disorders. However, we were not able to separate GoF vs LoF disorder subtypes. In the final set of analyses, we estimated the fraction of molecular mechanism defined disorder subtypes using two approaches. In the first,

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

we predicted for all observed variants in our expert curated patient cohort the "increasing" and "decreasing" function (Supplementary Table 1) status (75% most confident class assignments). The fraction for either classification was used as factor for the previous gene level birth incidences. Among our patient cohort, using our increased/decreased functional predictor, we predicted an increased functional effect in 52% of all cases in *GRIN1*. In contrast, in patient variants observed in *GRIN2A* and *GRIN2B* the more recurrent predicted functional effect was "decreasing" function (Decreasing $_{GRIN2A}$= 61%, Decreasing $_{GRIN2B}$= 56%, Figure 7B). Since the first approach might be confounded by a not yet recognized ascertainment bias of our cohort, we used a second approach to estimate the ratio of variants that "increase"/"decrease" function in the patient population. We applied the functional prediction algorithm to all possible variants along *GRIN1*, *GRIN2A* and *GRIN2B* which we classified as pathogenic using our pathogenicity prediction algorithm (N = 4,630). "Decreasing" effects are suggested as the dominant functional effect in each of the three GRIN genes (Figure 7C). Notably, this model doesn't consider that some variants might occur de novo more frequent that others due to CpG rich codons. Therefore, our two approaches, that each have advantages and disadvantages, we estimate the incidence for "increasing" function variant associated disorders for *GRIN1*as 2.15-2.81, for *GRIN2A* as 0.32-1.08 and for *GRIN2B* 0.50-2.31 in 100,000 births. Similarly, for "Decreasing" variant associated disorders we obtained an incidence estimate for *GRIN1* of 2.63-3.33, for *GRIN2A* as 1.7-2.46 and for *GRIN2B* 3.03-4.85 in 100,000 births. Inclusion of measures of surface localization together with parameters used to define "Increasing" and "Decreasing" function are needed before LoF and GoF can formally be used.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

A)

	Predicted incidence in the US* (per 100,000 births)			
Reference	Patient cohort		All variants predicted as pathogenic considering the 75% most confident class assignments (N = 4,630)	
Genes	Newborns with a missense variant causing a **increased** function	Newborns with a missense variant causing a **decreased** function	Newborns with pathogenic missense variant** predicted to cause a **increased** function	Newborns with a pathogenic missense variant** predicted to cause a **decreased** function
GRIN1	148 (4.03)	52 (142)	185 (5.07)	14 (0.39)
GRIN2A	44 (1.21)	58 (1.57)	52 (1.46)	49 (1.35)
GRIN2B	84 (2.28)	112 (3.07)	92 (2.51)	104 (2.84)

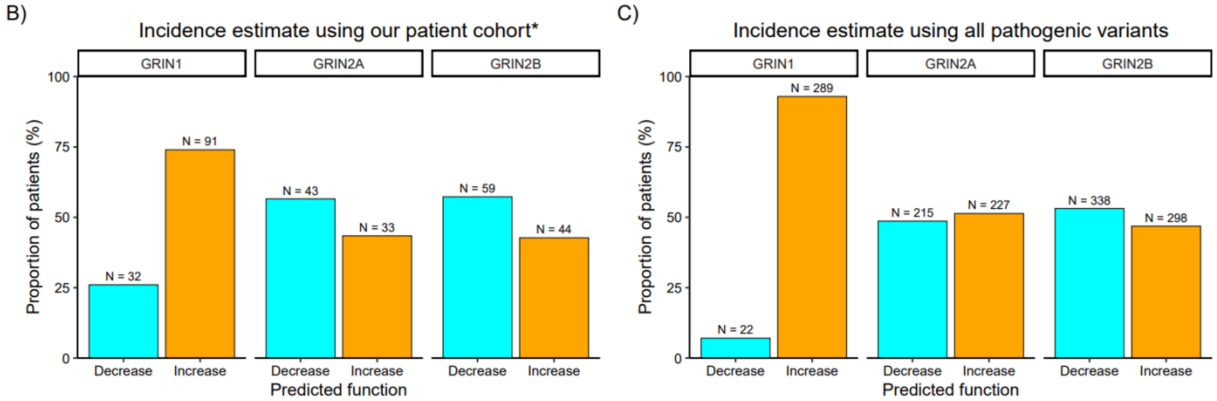

Figure 7 Prediction of the missense variant effects across our cohort of patient variants. **A)** Incidence estimates for the functional effect of NMDAR encoding genes. Overall incidences were obtained from (López-Rivera *et al., 2020*)[83]. Function specific incidences were calculated based on the gene-wise observed ratio of variants with an Increasing/Decreasing effect. *according to 3,659,289 births in USA. ** predicted with PP-dist&evo pathogenicity model. **B)** Gene wise fraction of patient variants from our patient cohort (see methods for details) predicted to cause an increasing or decreasing function considering the 75% most confident class assignments. **C)** As in **B** but predicting the function of all variants in *GRIN1*, *GRIN2A* and *GRIN2B* that were predicted to be pathogenic using our best pathogenicity prediction method, PP-dist&evo (considering the 75% most confident class assignments).

Discussion

Tailoring treatment to individual patients' genetic variants has made significant progress in many fields of medicine in recent years [84]. For disorders caused by genetic variants in GRIN genes, the possibility of successful drug treatment critically depends on the knowledge of the change in function caused by the pathogenic variant, that is gain or loss (partial or complete) of the NMDAR function. Indeed, the knowledge of the molecular mechanism affected by a variant can

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

guide safe and effective targeted treatment[45]. To date, only 20% of all the single amino acid exchanges in GRIN genes have been characterized by electrophysiological and biochemical readouts[85]. In this work we collected the most comprehensive dataset of GRIN variants located in the structurally resolved regions of the NMDAR, with 201 patient variants curated by clinical and genetics experts together and 631 population variants and 159 variants whose molecular consequences have been characterized by in vitro electrophysiology. Taking advantage of this comprehensive clinical and functional dataset of GRIN variants and using machine learning, we identify distance from ligands as a main predictive feature for both pathogenicity and functional prediction and we provide the most accurate pathogenicity prediction specifically developed for GRIN variants and the first computational method for the prediction of variant functional consequences in the GRIN genes (for prediction scores see Supplementary Table 1).

On the one hand, protein-unspecific pathogenicity predictors take the advantage of large training datasets and of capturing general principles of protein functioning though the use of massive evolutionary information[78,86–88]. On the other hand, protein-specific pathogenicity predictors that can incorporate knowledge on protein-specific structure and function, have been shown to enhance the accuracy of pathogenicity prediction[50,51,89,90]. When allowed by data availability, protein-specific pathogenicity predictors can improve the accuracy of pathogenicity prediction, by capturing protein-specific structural patterns and constraints. Here we increase the accuracy of pathogenicity prediction for missense variants in the GRIN genes by integrating structural information on the distance of variant residues to functionally important sites, ligands and the pore.

We also showed that distance from ligands and the pore has predictive power also for functional prediction. A limitation of this work, however, is that we explicitly omit effects of variants on

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

trafficking, which almost certainly have structural determinants beyond the ion channel pore and agonist binding pocket. Surface expression is clinically important, and a topic we will explore separately. So, our functional predictor is predicting increased and decreased function and not GoF/LoF. In line with previously reported observations in which variants with a LoF effect are predominantly located in the ABD domain, due to a disturbance of the agonist binding sites of GluN1 and GluN2A/B[4,14] we also find an enrichment of variants with increased functional effect. For a couple of variants close to the ligand binding site it has been proposed that the amino acid substitutions lead to a reduced agonist binding[59,60]. In contrast most variants in the TMD have been shown to have a GoF (or complex) effect[91–93]. Here we quantify for the first time the correlation of the functional effects of variants in NMDAR proteins with their spatial distance from the ligands. Hence these 3D distances can be used as a proxy to estimate the functional consequence of yet untested variants. Although this is the first step towards an accurate model specifically designed to predict the functional effect of NMDAR variants, we show that protein-unspecific models trained for pathogenicity were not sufficient to develop a strong prediction model for variant effects in NMDAR genes, and only including our newly generated distance features sufficiently boosts the performances to allow functional prediction.

This work is affected by the limitations that classification of the functional consequence of a variant in the two classes of LoF and GoF is an over-simplification of the real biophysical modifications which take place on a molecular level. Still, we could separately show that the distance of variant residues to ligands that regulate the NMDAR in particular correlate with the ligand-specific fold change potency, demonstrating its validity also on the level of individual electrophysiological readouts. Consequently, once more specific electrophysiological readouts

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

will become available our spatial distance to functional site annotation will become even more helpful to train models that predict the molecular consequence with higher granularity.

In summary, we introduced a potentially powerful approach to predict the directionality of the functional effects of likely pathogenic missense variants in GRIN genes. In a clinical setting like, treatment decisions must often be made before functional studies of disease-causing variants can be done. In the future, our prediction method could be adapted and benchmarked for use in conjunction with best current clinical practices, for example, to predict which individuals with pathogenic variants may be likely to benefit from a particular treatment based on their variants' LoF or GoF effects. Our method could potentially be refined with large-scale experimental data, for example, by introducing more specific types of predictions than the mere binary LoF and GoF classification, such as directly predicting a specific change of potency (e.g., glutamate). Because most GRIN genes are depleted for functional variants in the general population, it is likely that more GRIN genes could contribute to disease for which disease associations or mechanisms have not yet been elucidated and to which our method could potentially be applied – such as LoF variants in *GRIN2D*. In future iterations, also clinical and phenotypic data might be incorporated to enhance predictions for the underlying molecular defect.

Acknowledgements

We thank Rui Song, Wenshu XiangWei, Yuchen Xu, Daniel Teuscher, Lingling Xie, Ruth K. Mizu, and Wenjuan Chen for valuable assistance and sharing unpublished data.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Disclosures

H.Y. and S.F.T. are co-inventors of Emory-owned intellectual property. S.F.T. is a member of the SAB for Sage Therapeutics, Eumentis Therapeutics, Neurocrine, the GRIN2B Foundation, the CureGRIN Foundation, and CombinedBrain. S.F.T. is a consultant for GRIN Therapeutics. H.Y. is the PI on a research grant from Sage Therapeutics and GRIN Therapeutics to Emory. S.F.T. is PI on a research grant from GRIN Therapeutics to Emory. S.F.T. is cofounder of NeurOp, Inc. and Agrithera. DL recieved funds from the Simons Foundation and GRIN Therapeutics.

Grant Support

This work was supported by NS111619 (ST), HD082373 (HY), and AG072142 (SJM).

References

1. Lemke, J.R., Geider, K., Helbig, K.L., Heyne, H.O., Schütz, H., Hentschel, J., Courage, C., Depienne, C., Nava, C., Heron, D., et al. (2016). Delineating the GRIN1 phenotypic spectrum: A distinct genetic NMDA receptor encephalopathy. Neurology *86*, 2171–2178. 10.1212/WNL.0000000000002740.

2. Endele, S., Rosenberger, G., Geider, K., Popp, B., Tamer, C., Stefanova, I., Milh, M., Kortüm, F., Fritsch, A., Pientka, F.K., et al. (2010). Mutations in GRIN2A and GRIN2B encoding regulatory subunits of NMDA receptors cause variable neurodevelopmental phenotypes. Nat Genet *42*, 1021–1026. 10.1038/ng.677.

3. Hamdan, F.F., Gauthier, J., Araki, Y., Lin, D.-T., Yoshizawa, Y., Higashi, K., Park, A.-R., Spiegelman, D., Dobrzeniecka, S., Piton, A., et al. (2011). Excess of de novo deleterious mutations in genes associated with glutamatergic systems in nonsyndromic intellectual disability. Am J Hum Genet *88*, 306–316. 10.1016/j.ajhg.2011.02.001.

4. Myers, S.J., Yuan, H., Kang, J.-Q., Tan, F.C.K., Traynelis, S.F., and Low, C.-M. (2019). Distinct roles of GRIN2A and GRIN2B variants in neurological conditions. F1000Res *8*, F1000 Faculty Rev-1940. 10.12688/f1000research.18949.1.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

5. Carvill, G.L., Regan, B.M., Yendle, S.C., O'Roak, B.J., Lozovaya, N., Bruneau, N., Burnashev, N., Khan, A., Cook, J., Geraghty, E., et al. (2013). GRIN2A mutations cause epilepsy-aphasia spectrum disorders. Nat Genet *45*, 1073–1076. 10.1038/ng.2727.

6. Lemke, J.R., Lal, D., Reinthaler, E.M., Steiner, I., Nothnagel, M., Alber, M., Geider, K., Laube, B., Schwake, M., Finsterwalder, K., et al. (2013). Mutations in GRIN2A cause idiopathic focal epilepsy with rolandic spikes. Nat. Genet. *45*, 1067–1072. 10.1038/ng.2728.

7. Lemke, J.R., Hendrickx, R., Geider, K., Laube, B., Schwake, M., Harvey, R.J., James, V.M., Pepler, A., Steiner, I., Hörtnagel, K., et al. (2014). GRIN2B mutations in West syndrome and intellectual disability with focal epilepsy. Ann Neurol *75*, 147–154. 10.1002/ana.24073.

8. Reutlinger, C., Helbig, I., Gawelczyk, B., Subero, J.I.M., Tönnies, H., Muhle, H., Finsterwalder, K., Vermeer, S., Pfundt, R., Sperner, J., et al. (2010). Deletions in 16p13 including GRIN2A in patients with intellectual disability, various dysmorphic features, and seizure disorders of the rolandic region. Epilepsia *51*, 1870–1873. 10.1111/j.1528-1167.2010.02555.x.

9. Laube, B., Kuhse, J., and Betz, H. (1998). Evidence for a tetrameric structure of recombinant NMDA receptors. J Neurosci *18*, 2954–2961.

10. Paoletti, P., Bellone, C., and Zhou, Q. (2013). NMDA receptor subunit diversity: impact on receptor properties, synaptic plasticity and disease. Nat Rev Neurosci *14*, 383–400. 10.1038/nrn3504.

11. XiangWei, W., Jiang, Y., and Yuan, H. (2018). De Novo Mutations and Rare Variants Occurring in NMDA Receptors. Curr Opin Physiol *2*, 27–35. 10.1016/j.cophys.2017.12.013.

12. Benke, T.A., Park, K., Krey, I., Camp, C.R., Song, R., Ramsey, A.J., Yuan, H., Traynelis, S.F., and Lemke, J. (2021). Clinical and therapeutic significance of genetic variation in the GRIN gene family encoding NMDARs. Neuropharmacology *199*, 108805. 10.1016/j.neuropharm.2021.108805.

13. Platzer, K., Yuan, H., Schütz, H., Winschel, A., Chen, W., Hu, C., Kusumoto, H., Heyne, H.O., Helbig, K.L., Tang, S., et al. (2017). GRIN2B encephalopathy: novel findings on phenotype, variant clustering, functional consequences and treatment aspects. J Med Genet *54*, 460–470. 10.1136/jmedgenet-2016-104509.

14. Strehlow, V., Heyne, H.O., Vlaskamp, D.R.M., Marwick, K.F.M., Rudolf, G., de Bellescize, J., Biskup, S., Brilstra, E.H., Brouwer, O.F., Callenbach, P.M.C., et al. (2019). GRIN2A-related disorders: genotype and functional consequence predict phenotype. Brain *142*, 80–92. 10.1093/brain/awy304.

15. Li, D., Yuan, H., Ortiz-Gonzalez, X.R., Marsh, E.D., Tian, L., McCormick, E.M., Kosobucki, G.J., Chen, W., Schulien, A.J., Chiavacci, R., et al. (2016). GRIN2D Recurrent De Novo Dominant Mutation Causes a Severe Epileptic Encephalopathy Treatable with NMDA Receptor Channel Blockers. Am J Hum Genet *99*, 802–816. 10.1016/j.ajhg.2016.07.013.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

16. Platzer, K., Krey, I., and Lemke, J.R. (2022). GRIN2D-Related Developmental and Epileptic Encephalopathy. In GeneReviews®, M. P. Adam, D. B. Everman, G. M. Mirzaa, R. A. Pagon, S. E. Wallace, L. J. Bean, K. W. Gripp, and A. Amemiya, eds. (University of Washington, Seattle).

17. Myers, K.A., and Scheffer, I.E. (1993). GRIN2A-Related Speech Disorders and Epilepsy. In GeneReviews®, M. P. Adam, D. B. Everman, G. M. Mirzaa, R. A. Pagon, S. E. Wallace, L. J. Bean, K. W. Gripp, and A. Amemiya, eds. (University of Washington, Seattle).

18. Liu, R., Dang, W., Du, Y., Zhou, Q., Liu, Z., and Jiao, K. (2015). Correlation of functional GRIN2A gene promoter polymorphisms with schizophrenia and serum D-serine levels. Gene *568*, 25–30. 10.1016/j.gene.2015.05.011.

19. Camp, C.R., Vlachos, A., Klöckner, C., Krey, I., Banke, T.G., Shariatzadeh, N., Ruggiero, S.M., Galer, P., Park, K.L., Caccavano, A., et al. (2023). Loss of Grin2a causes a transient delay in the electrophysiological maturation of hippocampal parvalbumin interneurons. Commun Biol *6*, 1–16. 10.1038/s42003-023-05298-9.

20. Wang, H., Lv, S., Stroebel, D., Zhang, J., Pan, Y., Huang, X., Zhang, X., Paoletti, P., and Zhu, S. (2021). Gating mechanism and a modulatory niche of human GluN1-GluN2A NMDA receptors. Neuron *109*, 2443-2456.e5. 10.1016/j.neuron.2021.05.031.

21. Bouvier, G., Larsen, R.S., Rodríguez-Moreno, A., Paulsen, O., and Sjöström, P.J. (2018). Towards resolving the presynaptic NMDA receptor debate. Curr Opin Neurobiol *51*, 1–7. 10.1016/j.conb.2017.12.020.

22. Dore, K., Stein, I.S., Brock, J.A., Castillo, P.E., Zito, K., and Sjöström, P.J. (2017). Unconventional NMDA Receptor Signaling. J Neurosci *37*, 10800–10807. 10.1523/JNEUROSCI.1825-17.2017.

23. Iacobucci, G.J., and Popescu, G.K. (2017). NMDA receptors: linking physiological output to biophysical operation. Nat Rev Neurosci *18*, 236–249. 10.1038/nrn.2017.24.

24. Hunt, D.L., and Castillo, P.E. (2012). Synaptic plasticity of NMDA receptors: mechanisms and functional implications. Curr Opin Neurobiol *22*, 496–508. 10.1016/j.conb.2012.01.007.

25. Attwell, D., and Gibb, A. (2005). Neuroenergetics and the kinetic design of excitatory synapses. Nat Rev Neurosci *6*, 841–849. 10.1038/nrn1784.

26. Traynelis, S.F., Wollmuth, L.P., McBain, C.J., Menniti, F.S., Vance, K.M., Ogden, K.K., Hansen, K.B., Yuan, H., Myers, S.J., and Dingledine, R. (2010). Glutamate receptor ion channels: structure, regulation, and function. Pharmacol Rev *62*, 405–496. 10.1124/pr.109.002451.

27. Hansen, K.B., Wollmuth, L.P., Bowie, D., Furukawa, H., Menniti, F.S., Sobolevsky, A.I., Swanson, G.T., Swanger, S.A., Greger, I.H., Nakagawa, T., et al. (2021). Structure, Function,

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

and Pharmacology of Glutamate Receptor Ion Channels. Pharmacol Rev *73*, 298–487. 10.1124/pharmrev.120.000131.

28. Myers, S.J., Yuan, H., Perszyk, R.E., Zhang, J., Kim, S., Nocilla, K.A., Allen, J.P., Bain, J.M., Lemke, J.R., Lal, D., et al. (2023). Classification of missense variants in the N-methyl-d-aspartate receptor GRIN gene family as gain- or loss-of-function. Hum Mol Genet *32*, 2857–2871. 10.1093/hmg/ddad104.

29. Hanson, J.E., Ma, K., Elstrott, J., Weber, M., Saillet, S., Khan, A.S., Simms, J., Liu, B., Kim, T.A., Yu, G.-Q., et al. (2020). GluN2A NMDA Receptor Enhancement Improves Brain Oscillations, Synchrony, and Cognitive Functions in Dravet Syndrome and Alzheimer's Disease Models. Cell Rep *30*, 381-396.e4. 10.1016/j.celrep.2019.12.030.

30. Kemp, J.A., and McKernan, R.M. (2002). NMDA receptor pathways as drug targets. Nat Neurosci *5 Suppl*, 1039–1042. 10.1038/nn936.

31. Landrum, M.J., and Kattman, B.L. (2018). ClinVar at five years: Delivering on the promise. Human Mutation *39*, 1623–1630. 10.1002/humu.23641.

32. Silk, M., Petrovski, S., and Ascher, D.B. (2019). MTR-Viewer: identifying regions within genes under purifying selection. Nucleic Acids Res *47*, W121–W126. 10.1093/nar/gkz457.

33. Silk, M., Pires, D.E.V., Rodrigues, C.H.M., D'Souza, E.N., Olshansky, M., Thorne, N., and Ascher, D.B. (2021). MTR3D: identifying regions within protein tertiary structures under purifying selection. Nucleic Acids Res *49*, W438–W445. 10.1093/nar/gkab428.

34. Quinodoz, M., Peter, V.G., Cisarova, K., Royer-Bertrand, B., Stenson, P.D., Cooper, D.N., Unger, S., Superti-Furga, A., and Rivolta, C. (2022). Analysis of missense variants in the human genome reveals widespread gene-specific clustering and improves prediction of pathogenicity. Am J Hum Genet *109*, 457–470. 10.1016/j.ajhg.2022.01.006.

35. Sivley, R.M., Dou, X., Meiler, J., Bush, W.S., and Capra, J.A. (2018). Comprehensive Analysis of Constraint on the Spatial Distribution of Missense Variants in Human Protein Structures. Am J Hum Genet *102*, 415–426. 10.1016/j.ajhg.2018.01.017.

36. Pérez-Palma, E., May, P., Iqbal, S., Niestroj, L.-M., Du, J., Heyne, H.O., Castrillon, J.A., O'Donnell-Luria, A., Nürnberg, P., Palotie, A., et al. (2020). Identification of pathogenic variant enriched regions across genes and gene families. Genome Res *30*, 62–71. 10.1101/gr.252601.119.

37. Perszyk, R.E., Kristensen, A.S., Lyuboslavsky, P., and Traynelis, S.F. (2021). Three-dimensional missense tolerance ratio analysis. Genome Res *31*, 1447–1461. 10.1101/gr.275528.121.

38. Strehlow, V., Heyne, H.O., Vlaskamp, D.R.M., Marwick, K.F.M., Rudolf, G., de Bellescize, J., Biskup, S., Brilstra, E.H., Brouwer, O.F., Callenbach, P.M.C., et al. (2019).

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

GRIN2A-related disorders: genotype and functional consequence predict phenotype. Brain *142*, 80–92. 10.1093/brain/awy304.

39. Chiron, C., Marchand, M.C., Tran, A., Rey, E., d'Athis, P., Vincent, J., Dulac, O., and Pons, G. (2000). Stiripentol in severe myoclonic epilepsy in infancy: a randomised placebo-controlled syndrome-dedicated trial. STICLO study group. Lancet *356*, 1638–1642. 10.1016/s0140-6736(00)03157-3.

40. Jen, J.C., Graves, T.D., Hess, E.J., Hanna, M.G., Griggs, R.C., Baloh, R.W., and CINCH investigators (2007). Primary episodic ataxias: diagnosis, pathogenesis and treatment. Brain *130*, 2484–2493. 10.1093/brain/awm126.

41. Schoonjans, A.-S., Lagae, L., and Ceulemans, B. (2015). Low-dose fenfluramine in the treatment of neurologic disorders: experience in Dravet syndrome. Ther Adv Neurol Disord *8*, 328–338. 10.1177/1756285615607726.

42. Mullier, B., Wolff, C., Sands, Z.A., Ghisdal, P., Muglia, P., Kaminski, R.M., and André, V.M. (2017). GRIN2B gain of function mutations are sensitive to radiprodil, a negative allosteric modulator of GluN2B-containing NMDA receptors. Neuropharmacology *123*, 322–331. 10.1016/j.neuropharm.2017.05.017.

43. Pierson, T.M., Yuan, H., Marsh, E.D., Fuentes-Fajardo, K., Adams, D.R., Markello, T., Golas, G., Simeonov, D.R., Holloman, C., Tankovic, A., et al. (2014). GRIN2A mutation and early-onset epileptic encephalopathy: personalized therapy with memantine. Ann Clin Transl Neurol *1*, 190–198. 10.1002/acn3.39.

44. Chidambaram, S., and Manokaran, R.K. (2022). Favorable Response to "Memantine" in a Child with GRIN2B Epileptic Encephalopathy. Neuropediatrics *53*, 287–290. 10.1055/s-0041-1739130.

45. Xu, Y., Song, R., Chen, W., Strong, K., Shrey, D., Gedela, S., Traynelis, S.F., Zhang, G., and Yuan, H. (2021). Recurrent seizure-related GRIN1 variant: Molecular mechanism and targeted therapy. Ann Clin Transl Neurol *8*, 1480–1494. 10.1002/acn3.51406.

46. Han, W., Yuan, H., Allen, J.P., Kim, S., Shaulsky, G.H., Perszyk, R.E., Traynelis, S.F., and Myers, S.J. (2022). Opportunities for Precision Treatment of GRIN2A and GRIN2B Gain-of-Function Variants in Triheteromeric N-Methyl-D-Aspartate Receptors. J Pharmacol Exp Ther *381*, 54–66. 10.1124/jpet.121.001000.

47. Tang, W., Liu, D., Traynelis, S.F., and Yuan, H. (2020). Positive allosteric modulators that target NMDA receptors rectify loss-of-function GRIN variants associated with neurological and neuropsychiatric disorders. Neuropharmacology *177*, 108247. 10.1016/j.neuropharm.2020.108247.

48. Zhu, S., and Paoletti, P. (2015). Allosteric modulators of NMDA receptors: multiple sites and mechanisms. Curr Opin Pharmacol *20*, 14–23. 10.1016/j.coph.2014.10.009.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

49. Addis, L., Virdee, J.K., Vidler, L.R., Collier, D.A., Pal, D.K., and Ursu, D. (2017). Epilepsy-associated GRIN2A mutations reduce NMDA receptor trafficking and agonist potency - molecular profiling and functional rescue. Sci Rep *7*, 66. 10.1038/s41598-017-00115-w.

50. Boßelmann, C.M., Hedrich, U.B.S., Müller, P., Sonnenberg, L., Parthasarathy, S., Helbig, I., Lerche, H., and Pfeifer, N. (2022). Predicting the functional effects of voltage-gated potassium channel missense variants with multi-task learning. EBioMedicine *81*, 104115. 10.1016/j.ebiom.2022.104115.

51. Heyne, H.O., Baez-Nieto, D., Iqbal, S., Palmer, D.S., Brunklaus, A., May, P., Epi25 Collaborative, Johannesen, K.M., Lauxmann, S., Lemke, J.R., et al. (2020). Predicting functional effects of missense variants in voltage-gated sodium and calcium channels. Sci Transl Med *12*, eaay6848. 10.1126/scitranslmed.aay6848.

52. Sevim Bayrak, C., Stein, D., Jain, A., Chaudhary, K., Nadkarni, G.N., Van Vleck, T.T., Puel, A., Boisson-Dupuis, S., Okada, S., Stenson, P.D., et al. (2021). Identification of discriminative gene-level and protein-level features associated with pathogenic gain-of-function and loss-of-function variants. Am J Hum Genet *108*, 2301–2318. 10.1016/j.ajhg.2021.10.007.

53. Jung, S., Lee, S., Kim, S., and Nam, H. (2015). Identification of genomic features in the classification of loss- and gain-of-function mutation. BMC Med Inform Decis Mak *15 Suppl 1*, S6. 10.1186/1472-6947-15-S1-S6.

54. Liu, M., Watson, L.T., and Zhang, L. (2015). HMMvar-func: a new method for predicting the functional outcome of genetic variants. BMC Bioinformatics *16*, 351. 10.1186/s12859-015-0781-z.

55. Ge, F., Li, C., Iqbal, S., Muhammad, A., Li, F., Thafar, M.A., Yan, Z., Worachartcheewan, A., Xu, X., Song, J., et al. (2023). VPatho: a deep learning-based two-stage approach for accurate prediction of gain-of-function and loss-of-function variants. Brief Bioinform *24*, bbac535. 10.1093/bib/bbac535.

56. Swanger, S.A., Chen, W., Wells, G., Burger, P.B., Tankovic, A., Bhattacharya, S., Strong, K.L., Hu, C., Kusumoto, H., Zhang, J., et al. (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. 10.1016/j.ajhg.2016.10.002.

57. Hansen, K.B., Clausen, R.P., Bjerrum, E.J., Bechmann, C., Greenwood, J.R., Christensen, C., Kristensen, J.L., Egebjerg, J., and Bräuner-Osborne, H. (2005). Tweaking agonist efficacy at N-methyl-D-aspartate receptors by site-directed mutagenesis. Mol Pharmacol *68*, 1510–1523. 10.1124/mol.105.014795.

58. Chen, P.E., Geballe, M.T., Stansfeld, P.J., Johnston, A.R., Yuan, H., Jacob, A.L., Snyder, J.P., Traynelis, S.F., and Wyllie, D.J.A. (2005). Structural features of the glutamate binding site in recombinant NR1/NR2A N-methyl-D-aspartate receptors determined by site-directed mutagenesis and molecular modeling. Mol Pharmacol *67*, 1470–1484. 10.1124/mol.104.008185.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

59. Maier, W., Schemm, R., Grewer, C., and Laube, B. (2007). Disruption of interdomain interactions in the glutamate binding pocket affects differentially agonist affinity and efficacy of N-methyl-D-aspartate receptor activation. J Biol Chem *282*, 1863–1872. 10.1074/jbc.M608156200.

60. Laube, B., Hirai, H., Sturgess, M., Betz, H., and Kuhse, J. (1997). Molecular determinants of agonist discrimination by NMDA receptor subunits: analysis of the glutamate binding site on the NR2B subunit. Neuron *18*, 493–503. 10.1016/s0896-6273(00)81249-0.

61. Karczewski, K.J., Francioli, L.C., Tiao, G., Cummings, B.B., Alföldi, J., Wang, Q., Collins, R.L., Laricchia, K.M., Ganna, A., Birnbaum, D.P., et al. (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. Nature *581*, 434–443. 10.1038/s41586-020-2308-7.

62. Chen, W., Tankovic, A., Burger, P.B., Kusumoto, H., Traynelis, S.F., and Yuan, H. (2017). Functional Evaluation of a De Novo GRIN2A Mutation Identified in a Patient with Profound Global Developmental Delay and Refractory Epilepsy. Mol Pharmacol *91*, 317–330. 10.1124/mol.116.106781.

63. XiangWei, W., Kannan, V., Xu, Y., Kosobucki, G.J., Schulien, A.J., Kusumoto, H., Moufawad El Achkar, C., Bhattacharya, S., Lesca, G., Nguyen, S., et al. (2019). Heterogeneous clinical and functional features of GRIN2D-related developmental and epileptic encephalopathy. Brain *142*, 3009–3027. 10.1093/brain/awz232.

64. Traynelis, S.F. (1998). Software-based correction of single compartment series resistance errors. J Neurosci Methods *86*, 25–34. 10.1016/s0165-0270(98)00140-x.

65. Dana, J.M., Gutmanas, A., Tyagi, N., Qi, G., O'Donovan, C., Martin, M., and Velankar, S. (2019). SIFTS: updated Structure Integration with Function, Taxonomy and Sequences resource allows 40-fold increase in coverage of structure-based annotations for proteins. Nucleic Acids Research *47*, D482–D489. 10.1093/nar/gky1114.

66. Zhang, Y., Ye, F., Zhang, T., Lv, S., Zhou, L., Du, D., Lin, H., Guo, F., Luo, C., and Zhu, S. (2021). Structural basis of ketamine action on human NMDA receptors. Nature *596*, 301–305. 10.1038/s41586-021-03769-9.

67. Lomize, M.A., Pogozheva, I.D., Joo, H., Mosberg, H.I., and Lomize, A.L. (2012). OPM database and PPM web server: resources for positioning of proteins in membranes. Nucleic Acids Res *40*, D370-376. 10.1093/nar/gkr703.

68. Pravda, L., Sehnal, D., Toušek, D., Navrátilová, V., Bazgier, V., Berka, K., Svobodová Vařeková, R., Koča, J., and Otyepka, M. (2018). MOLEonline: a web-based tool for analyzing channels, tunnels and pores (2018 update). Nucleic Acids Research *46*, W368–W373. 10.1093/nar/gky309.

69. Grant, B.J., Rodrigues, A.P.C., ElSawy, K.M., McCammon, J.A., and Caves, L.S.D. (2006). Bio3d: an R package for the comparative analysis of protein structures. Bioinformatics *22*, 2695–2696. 10.1093/bioinformatics/btl461.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

70. Dong, R., Pan, S., Peng, Z., Zhang, Y., and Yang, J. (2018). mTM-align: a server for fast protein structure database search and multiple protein structure alignment. Nucleic Acids Res *46*, W380–W386. 10.1093/nar/gky430.

71. Fedele, L., Newcombe, J., Topf, M., Gibb, A., Harvey, R.J., and Smart, T.G. (2018). Disease-associated missense mutations in GluN2B subunit alter NMDA receptor ligand binding and ion channel properties. Nat Commun *9*, 957. 10.1038/s41467-018-02927-4.

72. Bastolla, U., Farwer, J., Knapp, E.W., and Vendruscolo, M. (2001). How to guarantee optimal stability for most representative structures in the Protein Data Bank. Proteins *44*, 79–96. 10.1002/prot.1075.

73. Montanucci, L., Capriotti, E., Frank, Y., Ben-Tal, N., and Fariselli, P. (2019). DDGun: an untrained method for the prediction of protein stability changes upon single and multiple point variations. BMC Bioinformatics *20*, 335. 10.1186/s12859-019-2923-1.

74. Montanucci, L., Capriotti, E., Birolo, G., Benevenuta, S., Pancotti, C., Lal, D., and Fariselli, P. (2022). DDGun: an untrained predictor of protein stability changes upon amino acid variants. Nucleic Acids Res *50*, W222–W227. 10.1093/nar/gkac325.

75. Henikoff, S., and Henikoff, J.G. (1992). Amino acid substitution matrices from protein blocks. Proc Natl Acad Sci U S A *89*, 10915–10919. 10.1073/pnas.89.22.10915.

76. Hopf, T.A., Ingraham, J.B., Poelwijk, F.J., Schärfe, C.P.I., Springer, M., Sander, C., and Marks, D.S. (2017). Mutation effects predicted from sequence co-variation. Nat Biotechnol *35*, 128–135. 10.1038/nbt.3769.

77. Kabsch, W., and Sander, C. (1983). How good are predictions of protein secondary structure? FEBS Lett *155*, 179–182. 10.1016/0014-5793(82)80597-8.

78. Frazer, J., Notin, P., Dias, M., Gomez, A., Min, J.K., Brock, K., Gal, Y., and Marks, D.S. (2021). Disease variant prediction with deep generative models of evolutionary data. Nature *599*, 91–95. 10.1038/s41586-021-04043-8.

79. Cheng, J., Novati, G., Pan, J., Bycroft, C., Žemgulytė, A., Applebaum, T., Pritzel, A., Wong, L.H., Zielinski, M., Sargeant, T., et al. (2023). Accurate proteome-wide missense variant effect prediction with AlphaMissense. Science *381*, eadg7492. 10.1126/science.adg7492.

80. Ioannidis, N.M., Rothstein, J.H., Pejaver, V., Middha, S., McDonnell, S.K., Baheti, S., Musolf, A., Li, Q., Holzinger, E., Karyadi, D., et al. (2016). REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. Am J Hum Genet *99*, 877–885. 10.1016/j.ajhg.2016.08.016.

81. Lester, R.A., and Jahr, C.E. (1992). NMDA channel behavior depends on agonist affinity. J Neurosci *12*, 635–643. 10.1523/JNEUROSCI.12-02-00635.1992.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

82. Vance, K.M., Simorowski, N., Traynelis, S.F., and Furukawa, H. (2011). Ligand-specific deactivation time course of GluN1/GluN2D NMDA receptors. Nat Commun *2*, 294. 10.1038/ncomms1295.

83. López-Rivera, J.A., Pérez-Palma, E., Symonds, J., Lindy, A.S., McKnight, D.A., Leu, C., Zuberi, S., Brunklaus, A., Møller, R.S., and Lal, D. (2020). A catalogue of new incidence estimates of monogenic neurodevelopmental disorders caused by de novo variants. Brain *143*, 1099–1105. 10.1093/brain/awaa051.

84. Ashley, E.A. (2016). Towards precision medicine. Nat Rev Genet *17*, 507–522. 10.1038/nrg.2016.86.

85. Hansen, K.B., Wollmuth, L.P., Bowie, D., Furukawa, H., Menniti, F.S., Sobolevsky, A.I., Swanson, G.T., Swanger, S.A., Greger, I.H., Nakagawa, T., et al. (2021). Structure, Function, and Pharmacology of Glutamate Receptor Ion Channels. Pharmacol Rev *73*, 298–487. 10.1124/pharmrev.120.000131.

86. Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S., and Sunyaev, S.R. (2010). A method and server for predicting damaging missense mutations. Nat Methods *7*, 248–249. 10.1038/nmeth0410-248.

87. Rentzsch, P., Witten, D., Cooper, G.M., Shendure, J., and Kircher, M. (2019). CADD: predicting the deleteriousness of variants throughout the human genome. Nucleic Acids Research *47*, D886–D894. 10.1093/nar/gky1016.

88. Sim, N.-L., Kumar, P., Hu, J., Henikoff, S., Schneider, G., and Ng, P.C. (2012). SIFT web server: predicting effects of amino acid substitutions on proteins. Nucleic Acids Res *40*, W452–W457. 10.1093/nar/gks539.

89. Stead, L.F., Wood, I.C., and Westhead, D.R. (2011). KvSNP: accurately predicting the effect of genetic variants in voltage-gated potassium channels. Bioinformatics *27*, 2181–2186. 10.1093/bioinformatics/btr365.

90. Li, B., Mendenhall, J.L., Kroncke, B.M., Taylor, K.C., Huang, H., Smith, D.K., Vanoye, C.G., Blume, J.D., George, A.L., Sanders, C.R., et al. (2017). Predicting the Functional Impact of KCNQ1 Variants of Unknown Significance. Circ Cardiovasc Genet *10*, e001754. 10.1161/CIRCGENETICS.117.001754.

91. Amador, A., Bostick, C.D., Olson, H., Peters, J., Camp, C.R., Krizay, D., Chen, W., Han, W., Tang, W., Kanber, A., et al. (2020). Modelling and treating GRIN2A developmental and epileptic encephalopathy in mice. Brain *143*, 2039–2057. 10.1093/brain/awaa147.

92. Li, J., Zhang, J., Tang, W., Mizu, R.K., Kusumoto, H., XiangWei, W., Xu, Y., Chen, W., Amin, J.B., Hu, C., et al. (2019). De novo GRIN variants in NMDA receptor M2 channel pore-forming loop are associated with neurological diseases. Hum Mutat *40*, 2393–2413. 10.1002/humu.23895.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.06.24306939; this version posted May 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

93. Xu, Y., Song, R., Perszyk, R.E., Chen, W., Kim, S., Park, K.L., Allen, J.P., Nocilla, K.A., Zhang, J., XiangWei, W., et al. (2024). De novo GRIN variants in M3 helix associated with neurological disorders control channel gating of NMDA receptor. Cell Mol Life Sci *81*, 153. 10.1007/s00018-023-05069-z.