AJHG The American Journal of Human Genetics

# A recurrent missense variant in AP2M1 impairs clathrin-mediated endocytosis and causes developmental and epileptic encephalopathy

--Manuscript Draft--

| | |
|---|---|
| **Manuscript Number:** | AJHG-D-19-00045R3 |
| **Full Title:** | A recurrent missense variant in AP2M1 impairs clathrin-mediated endocytosis and causes developmental and epileptic encephalopathy |
| **Article Type:** | Article |
| **Keywords:** | clathrin-mediated endocytosis; developmental and epileptic encephalopathy; synaptic transmission; computational phenotypes; Human Phenotype Ontology; neurodevelopmental disorders |
| **Corresponding Author:** | Ingo Helbig<br>Children's Hospital of Philadelphia<br>Philadelphia, UNITED STATES |
| **First Author:** | Ingo Helbig, MD |
| **Order of Authors:** | Ingo Helbig, MD |
| | Tania Lopez-Hernandez |
| | Oded Shor |
| | Peter Galer |
| | Shiva Ganesan |
| | Manuela Pendziwiat |
| | Annika Rademacher |
| | Colin A Ellis |
| | Nadja Hümpfer |
| | Niklas Schwarz |
| | Simone Seiffert |
| | Joseph Peeden |
| | Joseph Shen |
| | Katalin Štěrbová |
| | Trine Bjørg Hammer |
| | Rikke S Møller |
| | Deepali N Shinde |
| | Sha Tang |
| | Lacey Smith |
| | Annapurna Poduri |
| | Roland Krause |
| | Felix Benninger |
| | Katherine L Helbig, MS |
| | Volker Haucke |
| | Yvonne G Weber |
| **Abstract:** | The developmental and epileptic encephalopathies (DEE) are heterogeneous disorders with a strong genetic contribution, but the underlying genetic etiology |

| | |
|---|---|
| | remains unknown in a significant proportion of individuals. To explore whether statistical support for genetic etiologies can be generated based on phenotypic features, we analyzed whole exome sequencing data and phenotypic similarities using Human Phenotype Ontology (HPO) in 314 individuals with DEE. We identified a de novo c.508C>T; p.(Arg170Trp) variant in AP2M1 in two individuals with a phenotypic similarity that was higher than expected by chance (p=0.003) and a phenotype related to Epilepsy with Myoclonic-Atonic Seizures. We subsequently found the same de novo variant in two individuals with neurodevelopmental disorders and generalized epilepsy in a cohort of 2,310 individuals who underwent diagnostic whole-exome sequencing. AP2M1 encodes the μ-subunit of the adaptor protein complex 2 (AP-2), which is involved in clathrin-mediated endocytosis (CME) and synaptic vesicle recycling. Modeling of protein dynamics indicated that the p.Arg170Trp variant impairs the conformational activation and thermodynamic entropy of the AP-2 complex. Functional complementation of the μ-subunit carrying the p.Arg170Trp variant in human cells and astrocytes derived from AP-2μ conditional knockout mice revealed a significant impairment of CME of transferrin. In contrast, stability, expression levels, membrane recruitment, and localization were not impaired, suggesting a functional alteration of the AP-2 complex as the underlying disease mechanism. We establish a recurrent pathogenic variant in AP2M1 as a cause of DEE with distinct phenotypic features and implicate dysfunction of the early steps of endocytosis as a disease mechanism in epilepsy. |

# A recurrent missense variant in *AP2M1* impairs clathrin-mediated endocytosis and causes developmental and epileptic encephalopathy

Ingo Helbig,[1,2,3,4,20#] Tania Lopez-Hernandez,[5,20] Oded Shor,[6,7,20] Peter Galer,[1,2] Shiva Ganesan,[1,2] Manuela Pendziwiat,[3] Annika Rademacher,[3] Colin A. Ellis,[4] Nadja Hümpfer,[5,8] Niklas Schwarz,[9] Simone Seiffert,[9] Joseph Peeden,[10] Joseph Shen,[11] Katalin Štěrbová,[12] Trine Bjørg Hammer,[13] Rikke S. Møller,[13,14] Deepali N. Shinde,[15] Sha Tang,[15] Lacey Smith,[16] Annapurna Poduri,[16,17] Roland Krause,[18] Felix Benninger,[6,7,21] Katherine L. Helbig,[1,2,21] Volker Haucke,[5,8,21] Yvonne G. Weber,[9,19,21] the EuroEPINOMICS-RES Consortium, and the GRIN Consortium

[1] Division of Neurology, Children's Hospital of Philadelphia, Philadelphia, PA, 19104 USA

[2] Department of Biomedical and Health Informatics (DBHi), Children's Hospital of Philadelphia, Philadelphia, PA, 19104 USA

[3] Department of Neuropediatrics, Christian-Albrechts-University of Kiel, 24105 Kiel, Germany

[4] Department of Neurology, University of Pennsylvania, Perelman School of Medicine, Philadelphia, PA, 19104 USA

[5] Leibniz-Forschungsinstitut für Molekulare Pharmakologie (FMP), 13125 Berlin, Germany

[6] Department of Neurology, Rabin Medical Center, Petach Tikva, 4941492 Israel

[7] Sackler School of Medicine, Tel Aviv University, Tel Aviv, 6997801 Israel

[8] Freie Universität Berlin, Faculty of Biology, Chemistry, Pharmacy, 14195 Berlin, Germany

[9] Department of Neurology and Epileptology, Hertie Institute for Clinical Brain Research, University of Tübingen, 72076 Tübingen, Germany

[10] East Tennessee Children's Hospital, University of Tennessee Department of Medicine, Knoxville, TN, 37916 USA

[11] Division of Genetics, Department of Pediatrics, University of California San Francisco, Fresno, CA, 93701 USA

[12] Department of Child Neurology, Charles University 2nd Faculty of Medicine and University Hospital Motol, 150 06 Prague, Czech Republic

[13] Danish Epilepsy Centre Filadelfia, 4293 Dianalund, Denmark

[14] Institute for Regional Health Services, University of Southern Denmark, 5230 Odense, Denmark

[15] Division of Clinical Genomics, Ambry Genetics, Aliso Viejo, CA, 92656 USA

[16] Epilepsy Genetics Program, Department of Neurology, Boston Children's Hospital, Boston, MA, 02115 USA

[17] Department of Neurology, Harvard Medical School, Boston, MA, 02115 USA

[18] Luxembourg Centre for Systems Biomedicine, University of Luxembourg, 4365 Esch-sur-Alzette, Luxembourg

[19] Department of Neurosurgery, University of Tübingen, 72076 Tübingen, Germany

[20]These authors contributed equally to this work.

[21]Senior authors contributed equally to this work.

# corresponding author

| | |
|---|---|
| Corresponding author: | **Ingo Helbig, MD** |
| Corresponding author's address: | Division of Neurology<br>Children's Hospital of Philadelphia<br>Perelman School of Medicine<br>University of Pennsylvania<br>Philadelphia, PA 19104 |
| Corresponding author's phone/fax: | phone: +1 215-590-1719; fax: +1 215-590-1771 |
| Corresponding author's email: | helbigi@email.chop.edu |

| | |
|---|---|
| Number of words in abstract: | 250 |
| Number of words in main text: | 5861 |
| Number of references: | 47 |
| Number of figures: | 4 |
| Number of tables: | 1 |
| Number of supplemental figures: | 7 |
| Number of supplemental tables: | 2 |

**ABSTRACT**

The developmental and epileptic encephalopathies (DEE) are heterogeneous disorders with a strong genetic contribution, but the underlying genetic etiology remains unknown in a significant proportion of individuals. To explore whether statistical support for genetic etiologies can be generated based on phenotypic features, we analyzed whole exome sequencing data and phenotypic similarities using Human Phenotype Ontology (HPO) in 314 individuals with DEE. We identified a *de novo* c.508C>T; p.(Arg170Trp) variant in *AP2M1* in two individuals with a phenotypic similarity that was higher than expected by chance (p=0.003) and a phenotype related to Epilepsy with Myoclonic-Atonic Seizures. We subsequently found the same *de novo* variant in two individuals with neurodevelopmental disorders and generalized epilepsy in a cohort of 2,310 individuals who underwent diagnostic whole-exome sequencing. *AP2M1* encodes the μ-subunit of the adaptor protein complex 2 (AP-2), which is involved in clathrin-mediated endocytosis (CME) and synaptic vesicle recycling. Modeling of protein dynamics indicated that the p.Arg170Trp variant impairs the conformational activation and thermodynamic entropy of the AP-2 complex. Functional complementation of the μ-subunit carrying the p.Arg170Trp variant in human cells and astrocytes derived from AP-2μ conditional knockout mice revealed a significant impairment of CME of transferrin. In contrast, stability, expression levels, membrane recruitment, and localization were not impaired, suggesting a functional alteration of the AP-2 complex as the underlying disease mechanism. We establish a recurrent pathogenic variant in *AP2M1* as a cause of DEE with distinct phenotypic features and implicate dysfunction of the early steps of endocytosis as a disease mechanism in epilepsy.

**Keywords:** clathrin-mediated endocytosis; developmental and epileptic encephalopathy; synaptic transmission; computational phenotypes; Human Phenotype Ontology; neurodevelopmental disorders

## INTRODUCTION

A substantial proportion of childhood epilepsies present as developmental and epileptic encephalopathies (DEE), characterized by intractable epilepsy with associated cognitive comorbidities.[1] DEE often occur in the absence of explanatory imaging or metabolic findings, and genetic causes are increasingly implicated.[2-4] Recent progress through massively parallel sequencing technologies has enabled the discovery of an increasing number of associated genes in the last decade,[5-7] most commonly due to pathogenic *de novo* variants.[5] The genetic landscape of DEE is heterogeneous, with pathogenic variants in single genes often explaining fewer than one percent of all individuals.[7] While a clear gene-phenotype association is seen for some genetic etiologies such as Dravet Syndrome and pathogenic variants in *SCN1A* (MIM: 607208),[8; 9] many genetic epilepsies demonstrate significant phenotypic heterogeneity, with overlapping clinical presentations associated with a wide spectrum of genetic etiologies.[10; 11]

Discovery of underlying genetic causes in the epilepsies and neurodevelopmental disorders has primarily advanced through the ability to process and analyze large genomic datasets.[12; 13] In contrast, phenotypic data are frequently collected in non-standard formats and therefore cannot be used for a systematic analysis across larger cohorts of affected individuals.[14] The Human Phenotype Ontology (HPO) has been developed as a standardized format to provide both terminology and semantics to a broad range of phenotypic features, including neurological features.[15; 16] This standardized vocabulary has already been used to identify individuals with rare monogenic diseases in large cohorts[17; 18] and is frequently applied in a diagnostic setting to define phenotypic overlap for variant interpretation. In addition, methods to determine phenotypic similarity have been developed that incorporate the hierarchical structure of the ontology.[19; 20] Given their phenotypic complexity, the childhood epilepsies lend themselves to novel analysis methods that capitalize on available phenotypic information in addition to genomic data.

Here, we analyzed trio whole exome sequencing data in 314 individuals phenotyped with HPO terms. We assessed exome sequencing data for potential *de novo* variants and determined observed and predicted phenotypic similarity in individuals with *de novo* variants in the same gene. Two individuals with a *de novo* c.508C>T; p.(Arg170Trp) *AP2M1* (OMIM: 601024; NM_004068.3) variant had a higher phenotypic similarity than expected by chance, and more detailed phenotyping identified a clinical phenotype consistent with Epilepsy with Myoclonic-Atonic Seizures, also known as Doose Syndrome. We subsequently identified two additional individuals with the identical *de novo* c.508C>T; p.(Arg170Trp) *AP2M1* variant and comparable phenotypes in a large diagnostic cohort. Functional analysis revealed that the p.Arg170Trp variant in *AP2M1* encoding the μ-subunit of the clathrin adaptor complex AP2 impairs the early stages of clathrin-mediated endocytosis (CME), thereby identifying defective CME as a disease mechanism for neurodevelopmental disorders.

## MATERIAL AND METHODS

### Participant recruitment

Informed consent for participation in this study was obtained from parents of all probands in agreement with the Declaration of Helsinki. All studies were completed per protocol with local approval by institutional review boards (IRB). For study inclusion, all probands underwent clinical data review of medical history information including developmental and seizure history, neurological findings, and morphological details. Available EEG and brain imaging data were reviewed for all individuals. Epilepsy syndromes and seizure types were classified according to the International League Against Epilepsy (ILAE) classification criteria.[1; 21] The initial discovery cohort includes individuals from four major cohorts: the EuroEPINOMICS-RES cohort (RES, n=135), a cohort of individuals with epileptic encephalopathies recruited through a study by the German research foundation (DFG, n=109), a cohort of individuals recruited through the Genomics Research and Innovation Network (GRIN, n=48), and a cohort of

individuals recruited through the Epilepsy Genetics Research Project at the Children's Hospital of Philadelphia (EGRP, n=8). The confirmation cohort includes 2310 individuals with epilepsy who underwent diagnostic whole exome sequencing at Ambry Genetics (Aliso Viejo, CA, USA).

**Genetic analysis**

Trio-based whole exome sequencing (WES) on all probands and parents was performed in a research or diagnostic context with various platforms and enrichment kits. All variants of interest were confirmed by Sanger sequencing. For the RES cohort, WES was performed at the Wellcome Trust Sanger Institute (Hinxton, UK) within the EuroEPINOMICS-RES project using the Illumina TruSeq DNA Sample Preparation Kit, the Agilent Technologies SureSelect Human All Exon 50Mb Kit, and the Illumina HiSeq2000 per manufacturer's protocols and as previously described.[12; 22; 23] For the DFG cohort, WES was performed at the Institute of Clinical Molecular Biology at the University of Kiel and the Cologne Center for Genomics, using NimbleGen SeqCap EZ Human Exome Library v2.0, Nextera Rapid Capture Exome, Nextera Rapid Capture Expanded Exome, Agilent SureSelect Human All Exon V5, Agilent SureSelect Human All Exon 50Mb. For the GRIN cohort, WES was performed at the Broad Institute using Nextera Rapid Capture Exome kit. For the EGRP cohort, exome sequencing was performed in diagnostic setting at GeneDx (n=6) using SureSelect Human All Exon V4 (50Mb) kit, the Division of Genomic Diagnostics at the Children's Hospital of Philadelphia (n= 2) using SureSelect Clinical Research Exome kits.

For the confirmation cohort, including Individual 3 and Individual 4, trio-based diagnostic whole exome sequencing using IDT xGen Exome Research Panel v1.0 was performed on the proband and unaffected parents at Ambry Genetics (Aliso Viejo, CA, USA). Genomic DNA extraction, exome library preparation, sequencing, bioinformatics, and data analyses were performed as previously described.[24; 25] Identified candidate variants were confirmed using Sanger sequencing.

All genetic data on individuals included in the initial discovery cohort were re-analyzed using a standardized pipeline. Burrows Wheeler Alignment (v 0.7.12) MEM algorithm was used to align the raw data to the HS37d5 human reference genome. After alignment for each sample, Samblaster (v 0.1.20) was used to add mate tags (MC and MQ) to the paired-end lines. GATK tools (v4.0.0.0) was used to perform Base Quality Score Recalibration (BQSR) before SNP and indel calling using HaplotypeCaller with interval lists specific to exome enrichment kit used for each sample. PICARD tools (v2.0.1) was used to combine gVCF files for each trio, followed by genotyping with the genotypeGVCF tool implemented in GATK. Variant selection and filtration were done using GATK tools before generating a merged variant file (VCF) using the MergeVcfs functionality of PICARD tool. Annotation of the VCF file was performed using a customized version of ANNOVAR. *De novo*, homozygous, heterozygous, and rare variants were extracted from the annotated file if passing the following quality criteria: (1) read depth in proband and parents ≥ 10x; (2) genotype quality in proband and parents should be ≥ 20, (3) allele frequencies < 1% in all population databases, (4) RVIS percentile < 70, (5) read ratio ≥ 25% of the alternate allele in the proband of the trio.

**Phenotypic similarity analysis**

For the 314 individuals included in the discovery cohort, 3529 HPO terms were assigned with a median of 9 terms per individual. 2579 terms had been assigned manually by the clinicians of the EuroEPINOMICS-RES project while 950 terms were translated from the EuroEPINOMICS phenotype database where categorical and free-text terms were added. For the final analysis, all terms per individual were merged, duplicates removed, and obsolete terms replaced by synonyms in HPO version used for this study (HPO version 1.2; release format-version: 1.2; data-version: releases/2017-12-12; downloaded on 3/10/18). This resulted in a combined set of 11146 HPO terms, including all initially assigned HPO terms and all

unique ancestral terms per individual, which was used to assess the frequency *p* of each assigned or ancestral HPO term in the cohort of 314 individuals (**Table S1**). This allowed for computation of the Information Content (IC) for each term, defined as $-\log_2(p)$.[19]

**Phenotypic similarity between two individuals**

Phenotypic similarity was assessed by summing over the most informative common ancestor for all pairs of HPO terms between two individuals according to Resnik.[26] For two individuals ($P_1$ and $P_2$) a matrix *s* holds all HPO terms in individual $P_1$ (n terms as rows) and all HPO terms in individual $P_2$ (m terms as columns). Each position $s_{ij}$ represents a pair of HPO terms; the common parent term with the highest Information Content (IC) is chosen as the most informative common ancestor (MICA) as a score for $s_{ij.}$ (**Figure 1**).

Two different methods, referred to as $sim_{max}$ and $sim_{av}$, generate a symmetric score for the similarity for s by summing over either the maximum or the average of all rows and columns with appropriate normalization respectively.[27] For our primary analysis, we used $sim_{max}$ (Equation 1).

$$sim_{max}(P_1, P_2) = \frac{1}{2}\left(\sum_{i=1}^{m} \max_{1 \le i \le n} s_{ij} + \sum_{j=1}^{n} \max_{1 \le i \le m} s_{ji}\right) \quad (1)$$

The $sim_{av}$ has initially been suggested for semantic similarities in the gene ontology[27] and has been applied to Human Phenotype Ontology research by other authors.[28]

$$sim_{av}(P_1, P_2) = \frac{1}{2}\left(\frac{1}{m}\sum_{j=1}^{m} \max_{1 \le i \le n} s_{ij} + \frac{1}{n}\sum_{j=1}^{n} \max_{1 \le i \le m} s_{ji}\right) \quad (2)$$

The $sim_{max}$ method is the more conservative method to assess semantic similarity in our hands.

**Observed versus expected phenotypic similarity score per gene**

The expected similarity scores for genes with *n de novo* variants were assessed by determining the distribution of median similarities of *n* individuals randomly selected from the overall cohort with 100,000 permutations. Exact p-values for the observed similarities for all genes with *n de novo* variants were determined based on the distribution of similarity scores in 100,000 permutations. Custom computer code used to generate results for the phenotypic similarity analysis used in this manuscript is publicly available.

**Structural modeling and normal mode analysis**

The structures of the AP-2 complex molecule were taken from The Protein Data Bank (PDB-101; accession number 2XA7 and 4UQI for the active and inactive state, respectively).[29] Each *in silico* missense variant was created by mutagenesis plugin in PyMol Molecular Graphics System Version 1.8 (Schrödinger, LLC., Cambridge, MA). Wild-type (WT) and missense structures were analyzed by an ENCoM coarse grained normal mode analysis method to evaluate the effect of variants on the stability of the protein. This method is based on an entropic considerations C package of ENCoM[30] available at ENCoM development website, compiled and used on a Ubuntu platform (Canonical Group, UK). The calculation of the entropic difference (ΔG) between the WT AP-2 complex and missense variants was done using Matlab software (Mathworks, Natick, MA). For each variant the entropic change (ΔG) was calculated by subtracting WT from variant entropy. ΔG was normalized to the maximum absolute values and cluster analysis was performed using Matlab software.

**Clustering of entropy changes for *APM21* and *AP2S1* variants**

To assess whether the AP2M1 p.Arg170Trp variant results in entropy changes in the AP-2 complex comparable with other variants known to cause disease in AP-2 subunits, we compared the effect of pathogenic variants in *AP2S1* (MIM: 602242) and rare population variants in *AP2M1* (**Table S2**). Pathogenic variants in *AP2S1* cause autosomal dominant hypocalciuric hypercalcemia (MIM: 600740), a rare genetic kidney disease. No other human diseases have been associated with variants in genes encoding AP-2 subunit so far. Entropic difference (ΔG) between the wildtype (WT) AP-2 complex and missense variants were assessed using Matlab software (Mathworks, Natick, MA). Each variant was normalized to WT values by calculating a delta between them (ΔG).

**Cell lines and primary astrocytes**

HeLa and HEK293T cells were obtained from ATCC. Cells were cultured in DMEM with 4.5g/L glucose (Lonza) containing 10% heat-inactivated fetal bovine serum (FBS) (Gibco) and 1% penicillin/streptomycin (Gibco) during experimental procedures. Cells were routinely tested for mycoplasma contamination. Conditional AP-2µ knockout (KO) mouse (AP-2lox/lox × inducible CAG-Cre was described previously.[31] Primary mouse astrocytes were prepared from 1- to 3-day old pups. Cerebral cortices were dissected and the meninges were carefully removed in cold sterile HBSS. The tissue was trypsinized for 10 min at 37°C and mechanically dissociated in complete DMEM medium with 10% FBS, 1% penicillin/streptomycin plus 40 U/ml DNase I (Sigma) 10 times through a small pore fire-polished Pasteur pipette. The cell suspension was pelleted and resuspended in fresh complete DMEM, filtered through a 100 µm nylon membrane (BD Falcon) and plated into 10 $cm^2$ cell culture dishes. When cells reached confluence, astrocytes were trypsinized, plated in poly-L-lysine-coated (Sigma) glass coverslips in 12-well plates at about 100 000 cells/$cm^2$ and cultured for another 7-10 days. To deplete AP-2µ, culture astrocytes from floxed animals expressing a tamoxifen-inducible Cre recombinase were treated with 0.1 µM (Z)-4-hydroxytamoxifen

(Sigma) the day after plating. Astrocytes derived from littermate floxed littermates that were Cre negative were used as controls and treated with equal amounts of (Z)-4-hydroxytamoxifen. All animal experiments and procedures have been approved by the Landesamt für Gesundheit und Soziales (LaGeSo) Berlin according to §8.1 German Animal Welfare Act.

**Molecular biology**

AP-2µ p.Arg170Trp variant was first generated by overlap/extension PCR using the plasmid AP2u2-mCherry (Addgene plasmid # 27672) and the plasmid AP-2µ IRES mRFP in an AAV-HBA-EWB vector. siRNA-resistant AP-2µ WT and p.Arg170Trp variants were generated using the plasmid AP-2µ p.Arg170Trp-mCherry by overlap/extension PCR (primers available upon request). The integrity of all cloned constructs was confirmed by DNA sequencing.

**siRNA and plasmid transfections**

HeLa cells were transfected with siRNA using jetPRIME® (Polyplus) according to the manufacturer's protocol. To achieve optimal knockdown efficiency, two rounds of silencing were performed. Thus, cells were consecutively transfected on day 1 and day 3 and the experiment was performed on day 5. For transient overexpression of proteins in knockdown cells, plasmids were transfected on day 3 together with the second round of siRNA using also jetPRIME®. For AP-2µ silencing, the siRNA used was µ2-adaptin 5'-GUGGAUGCCUUUCGGGUCA. Transfection of MISSION® siRNA universal negative control #1 (Sigma) served as control siRNA. Primary astrocytes were transfected with lipofectamin 2000 (Invitrogen) at 1:2 ratio (DNA:lipofectamin) in Opti-MEM medium (as described in manufacturer's instructions) and after 4 h, medium was replaced with complete DMEM and incubated 48 h at 37°C.

**Antibodies**

Immunoblotting was performed using alpha-adaptin (AP-2α) (BD transduction 610502, 1:500, mouse), mu-adaptin (AP-2µ) (BD transduction 611351, 1:500, mouse), clathrin heavy chain (clone TD.1, IgG from tissue culture supernatant, 1:500, mouse), Rab5 (BD transduction 610724, 1:250, mouse), GAPDH (Sigma G8795, 1:5000, mouse), SNAP25 (Synaptic Systems 111011, 1:1000, mouse) as primary antibodies, and LI-COR 800CW and 680RD infrared as secondary antibodies. Western blot development was done using a LI-COR Odyssey Fc imager, and western blot bands were quantified using Image Studio Lite Version 4.0 software (LI-COR). For immunostaining, AP-2α (Homebrew, 1:100, mouse), clathrin heavy chain (clone X22, IgG from tissue culture supernatant, 1:250, mouse), RFP (MBL PM005, 1:500, rabbit) were used as primary antibodies and Alexa-568 goat anti-rabbit and Alexa-488 goat anti-mouse (1:500, Invitrogen) as secondary antibodies. RFP-Trap®_M (Chromotek) beads were used for immunoprecipitation the Cherry-tagged AP-2µ WT and p.Arg170Trp protein variants.

**Immunocytochemistry and confocal imaging**

HeLa cells seeded on coverslips were fixed for 13 min with 4% paraformaldehyde (w/v, PFA) in phosphate-buffered saline solution (PBS) on ice and washed three times with PBS. Cells were permeabilized and blocked in blocking solution (PBS, 10% goat serum and 0.3% Triton X-100) for 30 min, and incubated with primary antibodies (diluted in blocking solution) for 1 h. After three washes with PBS secondary antibodies diluted in blocking solution were incubated for 1 h, followed by three washes in PBS. Coverslips were mounted in Immu-Mount (Thermo Fisher) with 1.5 mg/ml DAPI (Sigma) and visualized using a Zeiss laser scanning confocal microscope LSM780. Co-localization experiments were analyzed using ImageJ software. Data are presented as mean values ± standard error of the mean (SEM) from 5 independent experiments (*N*). Statistical testing was performed using paired *t*-test.

**Cell lysates, co-immunoprecipitation and membrane fractionation experiments**

HEK293T cells were transfected with calcium phosphate to express mCherry or the mCherry-tagged versions of AP-2µ WT or p.Arg170Trp variant. Cells were washed 3 times in ice-cold PBS and harvested in lysis buffer (20 mM HEPES pH 7.4, 100 mM KCl, 2 mM $MgCl_2$, 1 mM PMSF, 0.1% protease inhibitor cocktail, 1% Triton X-100). Lysates were incubated on a rotating wheel at 4°C for 30 min, followed by centrifugation at 17,000 *g* for 10 min at 4 °C. Protein levels were quantified using the BCA Kit (Pierce, Thermo Scientific) and equally concentrated lysates were boiled for 5 min in Laemmli sample buffer. Between 15 and 40 µg of protein was loaded onto a 10% acrylamide gel for SDS–polyacrylamide gel electrophoresis (SDS–PAGE) and analyzed via immunoblot using LI-COR 800CW and 680RD infrared secondary antibodies. For co-immunoprecipitation experiments, cells were harvested 48 h post-transfection in lysis buffer (20 mM HEPES pH 7.4, 100 mM KCl, 2 mM $MgCl_2$, 1 mM PMSF, 0.1% protease inhibitor cocktail, 1% Triton X-100) and incubated on ice for 30 min, followed by centrifugation at 17,000 *g* for 10 min at 4 °C. Proteins in supernatants were quantified using the BCA Kit (Pierce, Thermo Scientific) and 1 mg of protein was mixed with RFP-Trap beads for 1h at 4 °C on a rotating wheel. Beads were pelleted, washed four times in lysis buffer, and bound protein was eluted in 40 µl of Laemmli sample buffer. Eluates were loaded onto a 10% acrylamide gel for SDS–PAGE followed by immunoblotting. In the case of membrane fractionation studies, 48 h post-transfection cells were harvested in homogenization buffer (20 mM HEPES pH 7.4, 10 mM KCl, 2 mM $MgCl_2$, 1 mM PMSF, 0.1% protease inhibitor cocktail) and homogenized using a 1 ml syringe through a 25 G needle 10-20 times. Total cell lysate was collected after centrifugation at 720 *g* for 5 min at 4 °C. To obtain the cytosolic fraction, total cell lysate was centrifuged at 100,000 *g* for 1 hour at 4 °C, the supernatant was collected, and protein concentration and volume determined. The membrane pellet was washed twice in homogenization buffer and resuspended in lysis buffer (20 mM HEPES pH 7.4, 100 mM KCl, 2 mM $MgCl_2$, 1 mM PMSF, 0.1% protease inhibitor cocktail, 1% Triton X-100) by pipetting and pass through a 25 G needle in half of the volume corresponding to the cytosolic fraction with lysis buffer (20 mM HEPES pH 7.4, 100 mM KCl, 2 mM $MgCl_2$, 1 mM PMSF, 0.1% protease inhibitor cocktail, 1% Triton X-

100). Protein detection in both cytosolic and membrane fractions was determined by using the BCA Kit (Pierce, Thermo Scientific). Equal amounts of membrane and cytosolic fractions were loaded onto a 10% acrylamide gel for SDS–PAGE followed by immunoblotting. Protein levels of AP-2α, AP-2μ and CHC in the total cell lysates were normalized to Rab5, protein levels in the membrane fraction were normalized to SNAP25. Data are presented as mean values ± SEM from 3-8 independent experiments (*N*). Statistical testing was performed using a one-sample *t*-test.

**Transferrin uptake**

For HeLa cells, AP-2μ depleted cells were rescued by re-introducing siRNA resistant AP-2μ WT and p.Arg170Trp variants both fused to mCherry. For AP-2μ KO astrocytes, re-expression of AP-2μ WT or p.Arg170Trp variants containing RFP after an internal ribosomal entry site (IRES) were used. Transferrin endocytosis assays were essentially done as described earlier.[32; 33] Briefly, HeLa cells seeded on coverslips were serum-starved for 1 h and treated with 25 μg ml−1 Tf-Alexa647 (Life Technologies) for 10 min at 37 °C. Primary astrocytes were seeded on coverslips, serum-starved overnight and treated with 50 μg ml−1 Tf-Alexa647 (Life Technologies) for 5 min at 37 °C. Cells were washed twice with ice-cold PBS and acid washed at pH 5.3 (0.1 M Na-acetate, 0.2 M NaCl) for 1 min on ice. The coverslips were washed twice with ice-cold PBS and fixed with 4% PFA for 30 min at room temperature. Cells were washed three times with PBS, followed by immunocytochemistry staining as described in Immunocytochemistry. Transferrin uptake was analyzed using a Zeiss laser scanning confocal microscope LSM780 and quantified using ImageJ software. At least ten images were taken per condition. Cells positive for mCherry (HeLa cells) or RFP (primary astrocytes) were selected manually and regions of interested (ROIs) were drawn around them. After background subtraction (by employing the rolling circle method with a radius of 50 pixels), the integrated density for AF647 was measured. Data are presented as mean values ± SEM. *N* = number of independent experiments. *n* = total number of analyzed cells.

## RESULTS

### HPO-based similarity analysis demonstrates significant phenotypic overlap in individuals with the recurrent *AP2M1* c.508C>T; p.(Arg170Trp) *de novo* variant

We identified 11 genes with *de novo* variants in two or more individuals in our cohort of 314 individuals with DEE (**Figure 1**). To assess whether *de novo* variants in the identified 11 genes are associated with a gene-phenotype relationship, we performed a semantic similarity analysis of phenotypic features across all 314 individuals, using quantitative phenotypic similarity based on the Information Content of HPO terms (**Figure 1**). We calculated the median phenotypic similarity among all individuals with *de novo* variants in the same gene and compared the observed value per gene to the expected phenotypic similarity scores for groups of the same size, computed by permutation analysis (**Figure S1**). Based on this distribution, we determined whether individuals with *de novo* variants in the same gene had a higher phenotypic similarity than expected by chance (**Figure 1**). *AP2M1* was the only gene in this group not previously associated with DEE and was identified in two individuals with an identical c.508C>T; p.(Arg170Trp) *de novo* variant (GRCh37 chr3: g.183898715C>T). Both individuals with the identical *AP2M1* c.508C>T; p.(Arg170Trp) *de novo* variant had a significant phenotypic similarity ($p = 0.003$), suggesting a significant gene-phenotype relationship and a narrow phenotypic spectrum. As a next step, we queried a diagnostic cohort of 2310 individuals with epilepsy who underwent diagnostic whole exome sequencing. We identified two additional individuals with the same *AP2M1 de novo* variant, resulting in a total of four individuals with the recurrent *AP2M1* NM_004068.3 c.508C>T; p.(Arg170Trp) *de novo* variant.

### Detailed phenotypic review of individuals with the recurrent *AP2M1* c.508C>T; p.(Arg170Trp) *de novo* variant suggests a recognizable phenotype

We manually reviewed the phenotypes of the four individuals with the recurrent *AP2M1* NM_004068.3 c.508C>T; p.(Arg170Trp) *de novo* variant in detail. All affected individuals presented with global developmental delay apparent in the first six months of life with seizure onset between 21 months and 4 years. Two out of four individuals had received a diagnosis of autism spectrum disorder and all affected individuals were female. Three out of four individuals had atonic seizures and generalized epileptiform discharges on EEG (**Table 1, see Supplemental Note**). Accordingly, *AP2M1*-related epilepsies share a recognizable phenotype reminiscent of Epilepsy with Myoclonic-Atonic Seizures, demonstrating that the phenotypic similarity assessed by HPO for the first two individuals (Individuals 1, 2) was recapitulated in the second two individuals (Individuals 3, 4), and together the key clinical features conform to a reasonably recognizable electroclinical syndrome. Within our first cohort, 64/314 individuals had a diagnosis of epilepsy with myoclonic-atonic seizures, also referred to as Myoclonic-Astatic Epilepsy (MAE) or Doose Syndrome, suggesting that up to 3% of individuals with MAE may have *de novo* variants in *AP2M1* (point estimate 3.125%; 95% CI 2.3-14.0). Of note, both individuals with *AP2M1 de novo* variants in the first cohort had been independently phenotyped at two different centers prior to sequencing, excluding that assigned HPO terms were influenced by knowledge of the underlying genetic etiology or phenotyping bias within a single contributing center.

**The *AP2M1* c.508C>T; p.(Arg170Trp) variant affects thermodynamic entropy of the AP-2 complex**

Given the important roles of AP-2 in the brain, we hypothesized that impaired AP-2 function in the presence of the p.Arg170Trp variant may underlie DEE in humans. Arg170 is part of a basic phospholipid-binding patch within AP-2μ that, based on structural data, is postulated to stabilize the active open conformation of AP-2 and its association with cargo membrane proteins.[34] To test whether the p.Arg170Trp variant may affect AP-2 activation, we assessed protein thermodynamics in the inactive (closed) and active (open) conformation of AP-2 wildtype (WT) or the AP-2 p.Arg170Trp variant by

molecular modeling (**Figure 2**). We found that the p.Arg170Trp variant showed a significant increase of entropy compared with WT. In the inactive (closed) state of the protein, the p.Arg170Trp variant caused an increased entropy mainly located in the α, β, and μ subunits. In the active (open) state of the protein, the p.Arg170Trp variant caused a significant increase in entropy compared to WT in the μ and σ subunits as well as at the cargo protein.

In order to estimate the functional effect of the entropy changes described above, we compared the entropy changes of *AP2M1* p.Arg170Trp variant with rare population variants in *AP2M1* and pathogenic variants in *AP2S1*, the causative genetic etiology for autosomal dominant hypocalciuric hypercalcemia. Regarding the inactive (closed) state of the protein, cluster analysis of entropic changes indicated a clear separation between all tested rare population variants and the pathogenic variants in the *AP2S1.* The cluster analysis showed that entropy change regarding the p.Arg170Trp variant was separated significantly from both (**Figure S2**). In the active (open) state of the protein the entropy change due to the p.Arg170Trp variant clustered with the pathogenic variants in the *AP2S1* and was separated from the rare population variants (**Figure S2**). Given that only active AP-2 is capable of associating with cargo membrane proteins, these results suggest that the *AP2M1* p.Arg170Trp variant may functionally impair AP-2-mediated cargo recognition and, thereby, CME.

**Clathrin-mediated endocytosis is reduced in human HeLa cells and AP-2μ KO mice astrocytes expressing the pathogenic *AP2M1* c.508C>T; p.(Arg170Trp) variant**

Based on the molecular modeling data, we hypothesized that the recurrent *AP2M1* c.508C>T; p.(Arg170Trp) variant may affect CME. We probed this prediction experimentally by quantitatively determining the efficacy of CME in AP-2μ-depleted human HeLa cells rescued by plasmid-based re-expression of siRNA-resistant mCherry-labelled wild-type (WT) or p.Arg170Trp variant AP-2μ. HeLa cells

expressing the pathogenic p.Arg170Trp AP-2µ variant showed reduced levels of internalized AlexaFluor647-labeled transferrin after 10 minutes compared with cells expressing the WT AP-2 complex (**Figure 3, Figure S4**). A similar endocytic defect was observed in primary astrocytes derived from conditional AP-2µ knockout (KO) mice upon re-expressing the p.Arg170Trp variant of AP-2µ (**Figure 3, Figure S5**). These results suggest that the pathogenic *AP2M1* c.508C>T; p.(Arg170Trp) variant is affecting AP-2-mediated cargo recognition and, thereby, CME.

**The *AP2M1* c.508C>T; p.(Arg170Trp) variant does not affect AP-2 complex stability, expression, membrane recruitment, or localization**

Given the functional alterations in CME mediated by the p.Arg170Trp variant, we next assessed whether the impaired CME may be due to stability, expression levels, membrane recruitment, or localization of the p.Arg170Trp variant-containing AP-2 complex (**Figure 4, Figure S3**). However, we found that AP-2µ wildtype (WT) and p.Arg170Trp variants colocalize with clathrin equally well in HeLa cells suggesting that localization of the pathogenic AP-2µ variant in clathrin-coated pits is not altered. Likewise, the p.Arg170Trp variant of AP-2µ does not affect the expression levels, the stability, or the membrane recruitment of AP-2 complexes (**Figure S3**). Taken together, these results suggest that the p.Arg170Trp variant impairs AP-2 function in CME in human cells and primary brain astrocytes at an early step, likely by affecting the recognition of cargo membrane proteins (**Figure S6)**.

**DISCUSSION**

In our study we identified four individuals with a homogeneous phenotype of a developmental and epileptic encephalopathy due to a recurrent *AP2M1 de novo* c.508C>T; p.(Arg170Trp) variant. *AP2M1* is highly expressed in the central nervous system and has been previously studied extensively in a functional context.[31; 35] *AP2M1* codes for the essential µ-subunit of the endocytic clathrin adaptor complex AP-2

involved in clathrin-mediated endocytosis (CME) at the plasma membrane in neurons and non-neuronal cells. CME is a major mechanism for the recycling of synaptic vesicle (SV) components at mammalian central synapses.[36; 37] AP-2 plays a dual role in this process by integrating the sorting of SV protein cargo with the reformation of release-ready SVs.[31] AP-2 also regulates the neuronal surface levels of GABA and glutamate receptors, and thereby contributes to long-term plastic changes in neurotransmission and to excitatory/inhibitory balance.[38; 39] Heterozygous mutant mice with a targeted disruption of the *AP2M1* gene do not have an apparent phenotype. Homozygous mutant mice, however, die before day 3.5 postcoitus, indicating that the μ2 subunit of the AP-2 complex is critical for early embryonic development.[40]

*AP2M1* is highly intolerant to variation in the general population with a pLI of 0.99 and missense z-score of 5.82 in the ExAC database[41] and an RVIS ExAC score of -0.31 (30$^{th}$ percentile).[42] Only a single loss-of-function variant in *AP2M1* is observed in the ExAC database, compared to more than 19 expected loss-of-function variants. The *AP2M1* c.508C>T; p.(Arg170Trp) variant is absent from all population databases, including ExAC and gnomAD.[41] Pathogenic variants in several other genes involved in neurotransmission are known to cause epilepsy and neurodevelopmental disorders, including *STXBP1* (MIM: 602926), *SNAP25* (MIM: 600322), *STX1B* (MIM: 601485), *CLTC* (MIM: 118955), *DNM1* (MIM: 602377), and *PPP3CA* (MIM: 114105).[5; 10; 43-46] Within this group of genetic etiologies, *AP2M1* stands out, as the AP-2 complex mediates endocytic sorting of both presynaptic vesicle proteins and postsynaptic ion channels. *AP2M1*-related dysfunction may therefore provide a link between disorders of synaptic function and ion channelopathies, the two major groups of genetic etiologies identified in human epilepsy.

We present several lines of experimental evidence that support the view that the encoded p.Arg170Trp variant in *AP2M1* is causal for the observed epilepsy phenotype in the individuals in our study due to

impaired AP-2 complex function. First, using molecular modeling *in silico* we found a significant impact of the p.Arg170Trp *AP2M1* variant not only on the μ-subunit but also on other subunits of the AP-2 complex, suggesting that this variant affects overall AP-2 complex function both in the inactive (closed) and active (open) state. Pathogenic variants in *AP2S1* cause autosomal dominant hypocalciuric hypercalcemia, a rare genetic kidney disease. No other human disease has been associated with disease-causing variants in genes encoding AP-2 subunits so far. In the active (open) state, the type of thermodynamic changes predicted from the p.Arg170Trp *AP2M1* variant are more closely related to known pathogenic variants in *AP2S1* than rare population variants in *AP2M1*. While the modeling data can only provide indirect predictions about the overall functional alterations of the AP-2 complex, our normal mode analysis indicates a pronounced instability of the bound cargo protein. This instability likely relates to the fact that Arg170 is part of the N-terminal phosphatidylinositol (4,5)-bisphosphate ($PIP_2$)-binding site of the μ2 subunit formed by residues Lys167, Arg169, Arg170 and Lys421. It was previously reported that substituting three positively charged residues (Lys167, Arg169, Arg170) to Glu (KRR>E) had little effect on AP-2 binding to $PIP_2$-containing membranes *in vitro*, while additional mutations in the second $PIP_2$-binding site of the μ2 subunit resulted in a 4-fold reduction in binding to $PIP_2$.[34] The substitution of the charged Arg170 by a large hydrophobic side chain (Trp; of the p.Arg170Trp) may therefore conceivably lead to more unstable association of AP-2 with cargo membrane proteins at $PIP_2$-containing membranes.[34; 47] In summary, our *in silico* modeling using normal mode analysis suggests that the p.Arg170Trp *AP2M1* variant not only affects the function of the μ-subunit, but also globally interferes with the thermodynamic stability of the AP-2 complex and, thereby with cargo membrane protein recognition.

Second, we explored the impact of the p.Arg170Trp variant both in AP-2μ-depleted human HeLa cells and primary astrocytes derived from conditional AP-2μ KO mice. Analysis in both model systems suggested that the AP-2 μ-subunit with the p.Arg170Trp variant is expressed at levels comparable to wildtype and is

integrated into functional AP-2 complexes. In contrast, the uptake of the CME cargo transferrin was significantly reduced in p.Arg170Trp *AP2M1*-expressing HeLa cells and astrocytes. These results indicate that the effect of the p.Arg170Trp is due to an alteration of AP-2 complex function rather than haploinsufficiency. Based on these data, we therefore postulate that defective endocytic sorting of either one or several AP-2-dependent cargo membrane proteins in neuronal cells underlies developmental and epileptic encephalopathies, possibly due to imbalance between excitatory and inhibitory neurotransmission. It is well established that AP-2 critically regulates both excitatory and inhibitory neurotransmission by controlling the surface number of postsynaptic glutamate and $GABA_A$ receptors as well as the endocytic sorting of the presynaptic vesicular glutamate (vGLUT) and GABA transporters (vGAT) to reform synaptic vesicles during activity-dependent neurotransmission.[38] Given the various known links between the GABA-ergic system and hyperexcitability, small changes in the efficacy of endocytic sorting of presynaptic vGAT may result in reduced GABA content of inhibitory synaptic vesicles and thereby cause excitatory/inhibitory imbalance and epilepsy. Alternatively, it is possible that excitatory/inhibitory imbalance and epilepsy result from impaired clathrin/AP-2-mediated endocytosis of postsynaptic glutamate receptors or associated factors at excitatory synapses, and thereby, elevated excitatory transmission, in addition to other possibilities. Future studies, for example in p.Arg170Trp AP2M1 knock-in mice, will be needed to address these hypothetical scenarios in detail.

Within our study, statistical evidence for an involvement of *AP2M1* in genetic epilepsies was generated by a phenotypic similarity analysis based on Human Phenotype Ontology. As moderately sized cohorts are underpowered to provide statistically significant evidence for uncharacterized gene-disease relationships on a genomic level, we leveraged the existing rich phenotypic information in this cohort to evaluate the role of *de novo* variants in human epilepsy. We compared predicted versus observed phenotypic similarity in groups of individuals with *de novo* variants in shared genes. Four out of the 11 identified genes had an

uncorrected p-value of <0.05, including genes known to be associated with neurodevelopmental disorders with a clinically distinct phenotype such as *DNM1* (**Figure 1**). In contrast, genes known to be associated with a broader neurodevelopmental phenotype such as *SCN2A* (MIM: 182390) were not significant in this analysis. Results for *SCN1A*, the gene linked to Dravet Syndrome, were borderline significant. Within the list of 11 genes with *de novo* variants in two or more trios, *AP2M1* was the only previously undescribed etiology for genetic epilepsies, and the phenotypic similarity between both individuals was significant (**Table 1**). Given that HPO terms were assigned manually in our study, this method is prone to bias, possibly resulting both in false positive and false negative findings. For example, both individuals with pathogenic *de novo* variants in *DNM1* were phenotyped by clinicians from the same center, which makes it impossible to distinguish a phenotyping bias from a true disease association. While the *DNM1* phenotype is known to be relatively homogeneous, we cannot exclude that at least some of the observed similarity in both individuals was due to a phenotyping bias within a single contributing group. However, both individuals with *AP2M1 de novo* variants were independently phenotyped by clinicians at two different centers prior to sequencing. This makes it unlikely that assigned HPO terms were influenced by knowledge of the underlying genetic etiology or phenotyping bias within a single contributing enter. When comparing semantic phenotypic similarity to the probability of *de novo* variants in each of the genes in 314 individuals, we find that the statistical evidence based on phenotypes does not correlate with the statistical evidence from genotypes (**Figure S7**). This suggests that phenotypic similarity may provide statistical support for the involvement of a gene in disease independent of the probability of observed *de novo* variants. Taken together, we demonstrate that semantic phenotypic similarity can be used to provide statistical support for genetic etiologies in human epilepsy and identify recognizable disease entities. Prior studies have applied related strategies in large datasets of syndromic diseases or hospital-wide cohorts.[17; 18] In addition to the success of studies in heterogeneous phenotypes, we show that HPO-based approaches can also be used in cohorts with relatively homogeneous phenotypes such as the non-

lesional pediatric epilepsies, provided that phenotypic information is sufficiently deep for underlying gene-disease relationships to be identified.

In summary, we identify a recurrent pathogenic variant in *AP2M1* as a cause of genetic epilepsies resulting in a recognizable electroclinical phenotype with features of Epilepsy with Myoclonic-Atonic Seizures or Doose Syndrome. By demonstrating impairment of CME despite intact integration into functional AP-2 complexes, we establish dysfunction of the early step of CME as a disease mechanism in neurodevelopmental disorders and epilepsy.

## DESCRIPTION OF SUPPLEMENTAL DATA

Supplemental data includes detailed phenotypic description of individuals, seven figures, and two tables.

## DECLARATION OF INTERESTS

D.N.S. and S.T. are full time employees of Ambry Genetics. Exome sequencing is one of Ambry's commercially available diagnostic tests.

## CONSORTIA

### EuroEPINOMICS-RES Consortium Members

Rudi Balling, Nina Barisic, Stéphanie Baulac, Hande Caglayan, Dana Craiu, Peter De Jonghe, Christel Depienne, Renzo Guerrini, Helle Hjalgrim, Dorota Hoffman-Zacharska, Johanna Jähn, Karl Martin Klein, Bobby P.C. Koeleman, Vladimir Komarek, Eric Leguern, Anna-Elina Lehesjoki, Johannes R. Lemke, Holger Lerche, Tarja Linnankivi, Carla Marini, Patrick May, Hiltrud Muhle, Deb K. Pal, Aarno Palotie, Felix Rosenow, Susanne Schubert-Bast, Kaja Selmer, Jose M. Serratosa, Sanjay Sisodiya, Ulrich Stephani, Pasquale Striano, Arvid Suls, Tiina Talvik, Sarah von Spiczak, Sarah Weckhuysen, Federico Zara

### GRIN Consortium Members

Paul Avillach, Anna Bartels, Sawona Biswas, Florence Bourgeois, Batsal Devkota, Tracy Glauser, Barbara Hallinan, Allison Heath, Joel Hirschhorn, Judson Kilbourn, Sek Won Kong, Ian Krantz, In-Hee Lee, Kenneth Mandl, Eric Marsh, Kristen Sund, Deanne Taylor, Peter White

## ACKNOWLEDGMENTS

We thank the participants and their family members for taking part in the study. We would like to thank Mahgenn Cosico and Margaret O'Brien for their support in enrolling research participants and administrative assistance. The current study has been supported by the EuroEPINOMICS-Rare Epilepsy Syndrome (RES) Consortium, which provided capacity for exome sequencing. This work was supported by intramural funds of the Children's Hospital of Philadelphia, the German Research Foundation (HE5415/3-1 to I.H.) within the EuroEPINOMICS framework of the European Science Foundation, the German Research Foundation (DFG, HE5415/5-1, HE5415/6-1 to I.H., HA2686/13-1 (Koselleck Program) to V.H., WE4896/3-1 to Y.G.W.), by a DFG Research Unit FOR2715 (Le1030/16-1, We4896/4-1 and He5415/7-1 to I.H. and Y.G.W.), the Luxembourg National Research Fund (INTER/DFG/17/11583046 to R.K.), by the German Society for Epileptology (to Y.G.W.), and the Genomics Research and Innovation Network (GRIN, grinnetwork.org), a collaboration amongst the Children's Hospital of Philadelphia, Boston Children's Hospital, Cincinnati Children's Hospital Medical Center. I.H. also received support through the International League Against Epilepsy. This work was also partially supported by NIH grant NS108874. O.S. and F.B. were supported by the Israel Science Foundation (grant No. 1010/16). C.A.E. was supported in part by a Ruth L. Kirschstein National Research Service Award (NRSA) Institutional Research Training Grant, T32 NS091008-01. N.S. was in part supported from the intramural research funding program of the Faculty of Medicine Tübingen (Fortüne 2381-0-0). K.S. was supported by the Ministry of Health of Czech Republic (AZV15–33041 and DRO 00064203). A.P. is supported by the Boston Children's Hospital Translational Research Program.

**WEB RESOURCES**

Human Phenotype Ontology (https://hpo.jax.org/app/)

ExAC Database (http://exac.broadinstitute.org)

gnomAD Browser (http://gnomad.broadinstitute.org/)

Residual Variation Intolerance Score (RVIS) (http://genic-intolerance.org)

ENCoM development (https://github.com/NRGlab/ENCoM)

Online Mendelian Inheritance in Man (http://www.omim.org)

Computer code used for phenotypic similarity analysis (https://github.com/galerp/helbig_lab)

**ACCESSION NUMBERS**

The EuroEPINOMICS-RES exome-sequencing data are deposited in the European Genome-Phenome Archive, accession numbers EGAS00001000190, EGAS00001000386, and EGAS00001000048.

**REFERENCES**


1. Scheffer, I.E., Berkovic, S., Capovilla, G., Connolly, M.B., French, J., Guilhoto, L., Hirsch, E., Jain, S., Mathern, G.W., Moshe, S.L., et al. (2017). ILAE classification of the epilepsies: Position paper of the ILAE Commission for Classification and Terminology. Epilepsia *58*, 512-521.

2. McTague, A., Howell, K.B., Cross, J.H., Kurian, M.A., and Scheffer, I.E. (2016). The genetic landscape of the epileptic encephalopathies of infancy and childhood. Lancet Neurol *15*, 304-316.

3. Helbig, I., and Tayoun, A.A. (2016). Understanding Genotypes and Phenotypes in Epileptic Encephalopathies. Mol Syndromol *7*, 172-181.

4. Helbig, I., Heinzen, E.L., Mefford, H.C., and International League Against Epilepsy Genetics Commission. (2018). Genetic literacy series: Primer part 2-Paradigm shifts in epilepsy genetics. Epilepsia *59*, 1138-1147.

5. Heyne, H.O., Singh, T., Stamberger, H., Abou Jamra, R., Caglayan, H., Craiu, D., De Jonghe, P., Guerrini, R., Helbig, K.L., Koeleman, B.P.C., et al. (2018). De novo variants in neurodevelopmental disorders with epilepsy. Nat Genet *50*, 1048-1053.

6. Helbig, K.L., Farwell Hagman, K.D., Shinde, D.N., Mroske, C., Powis, Z., Li, S., Tang, S., and Helbig, I. (2016). Diagnostic exome sequencing provides a molecular diagnosis for a significant proportion of patients with epilepsy. Genet Med *18*, 898-905.

7. Lindy, A.S., Stosser, M.B., Butler, E., Downtain-Pickersgill, C., Shanmugham, A., Retterer, K., Brandt, T., Richard, G., and McKnight, D.A. (2018). Diagnostic outcomes for genetic testing of 70 genes in 8565 patients with epilepsy and neurodevelopmental disorders. Epilepsia *59*, 1062-1071.

8. Djemie, T., Weckhuysen, S., von Spiczak, S., Carvill, G.L., Jaehn, J., Anttonen, A.K., Brilstra, E., Caglayan, H.S., de Kovel, C.G., Depienne, C., et al. (2016). Pitfalls in genetic testing: the story of missed SCN1A mutations. Mol Genet Genomic Med *4*, 457-464.

9. Claes, L., Del-Favero, J., Ceulemans, B., Lagae, L., Van Broeckhoven, C., and De Jonghe, P. (2001). De novo mutations in the sodium-channel gene SCN1A cause severe myoclonic epilepsy of infancy. Am J Hum Genet *68*, 1327-1332.

10. Stamberger, H., Nikanorova, M., Willemsen, M.H., Accorsi, P., Angriman, M., Baier, H., Benkel-Herrenbrueck, I., Benoit, V., Budetta, M., Caliebe, A., et al. (2016). STXBP1 encephalopathy: A neurodevelopmental disorder including epilepsy. Neurology *86*, 954-962.

11. Wolff, M., Johannesen, K.M., Hedrich, U.B.S., Masnada, S., Rubboli, G., Gardella, E., Lesca, G., Ville, D., Milh, M., Villard, L., et al. (2017). Genetic and phenotypic heterogeneity suggest therapeutic implications in SCN2A-related disorders. Brain *140*, 1316-1336.

12. EuroEPINOMICS-RES Consortium, Epilepsy Phenome/Genome Project, and Epi4K Consortium. (2017). De Novo Mutations in Synaptic Transmission Genes Including DNM1 Cause Epileptic Encephalopathies. Am J Hum Genet *100*, 179.

13. Allen, A.S., Berkovic, S.F., Cossette, P., Delanty, N., Dlugos, D., Eichler, E.E., Epstein, M.P., Glauser, T., Goldstein, D.B., Han, Y., et al. (2013). De novo mutations in epileptic encephalopathies. Nature *501*, 217-221.

14. Helbig, I., and Lindhout, D. (2017). Advancing the phenome alongside the genome in epilepsy studies. Neurology *89*, 14-15.

15. Robinson, P.N., Kohler, S., Bauer, S., Seelow, D., Horn, D., and Mundlos, S. (2008). The Human Phenotype Ontology: a tool for annotating and analyzing human hereditary disease. Am J Hum Genet *83*, 610-615.

16. Kohler, S., Vasilevsky, N.A., Engelstad, M., Foster, E., McMurry, J., Ayme, S., Baynam, G., Bello, S.M., Boerkoel, C.F., Boycott, K.M., et al. (2017). The Human Phenotype Ontology in 2017. Nucleic Acids Res *45*, D865-D876.

17. Bastarache, L., Hughey, J.J., Hebbring, S., Marlo, J., Zhao, W., Ho, W.T., Van Driest, S.L., McGregor, T.L., Mosley, J.D., Wells, Q.S., et al. (2018). Phenotype risk scores identify patients with unrecognized Mendelian disease patterns. Science *359*, 1233-1239.

18. Akawi, N., McRae, J., Ansari, M., Balasubramanian, M., Blyth, M., Brady, A.F., Clayton, S., Cole, T., Deshpande, C., Fitzgerald, T.W., et al. (2015). Discovery of four recessive developmental disorders using probabilistic genotype and phenotype matching among 4,125 families. Nat Genet *47*, 1363-1369.

19. Kohler, S., Schulz, M.H., Krawitz, P., Bauer, S., Dolken, S., Ott, C.E., Mundlos, C., Horn, D., Mundlos, S., and Robinson, P.N. (2009). Clinical diagnostics in human genetics with semantic similarity searches in ontologies. Am J Hum Genet *85*, 457-464.

20. Pesquita, C., Faria, D., Falcao, A.O., Lord, P., and Couto, F.M. (2009). Semantic similarity in biomedical ontologies. PLoS Comput Biol *5*, e1000443.

21. Fisher, R.S., Cross, J.H., French, J.A., Higurashi, N., Hirsch, E., Jansen, F.E., Lagae, L., Moshe, S.L., Peltola, J., Roulet Perez, E., et al. (2017). Operational classification of seizure types by the International League Against Epilepsy: Position Paper of the ILAE Commission for Classification and Terminology. Epilepsia *58*, 522-530.

22. Suls, A., Jaehn, J.A., Kecskes, A., Weber, Y., Weckhuysen, S., Craiu, D.C., Siekierska, A., Djemie, T., Afrikanova, T., Gormley, P., et al. (2013). De novo loss-of-function mutations in CHD2 cause a fever-sensitive myoclonic epileptic encephalopathy sharing features with Dravet syndrome. Am J Hum Genet *93*, 967-975.

23. Vogtle, F.N., Brandl, B., Larson, A., Pendziwiat, M., Friederich, M.W., White, S.M., Basinger, A., Kucukkose, C., Muhle, H., Jahn, J.A., et al. (2018). Mutations in PMPCB Encoding the Catalytic Subunit of the Mitochondrial Presequence Protease Cause Neurodegeneration in Early Childhood. Am J Hum Genet.

24. Farwell Hagman, K.D., Shinde, D.N., Mroske, C., Smith, E., Radtke, K., Shahmirzadi, L., El-Khechen, D., Powis, Z., Chao, E.C., Alcaraz, W.A., et al. (2017). Candidate-gene criteria for clinical reporting: diagnostic

exome sequencing identifies altered candidate genes among 8% of patients with undiagnosed diseases. Genet Med *19*, 224-235.

25. Farwell, K.D., Shahmirzadi, L., El-Khechen, D., Powis, Z., Chao, E.C., Tippin Davis, B., Baxter, R.M., Zeng, W., Mroske, C., Parra, M.C., et al. (2015). Enhanced utility of family-centered diagnostic exome sequencing with inheritance model-based analysis: results from 500 unselected families with undiagnosed genetic conditions. Genet Med *17*, 578-586.
26. Resnik, P. (1995). Using information content to evaluate semantic similarity in a taxonomy. In 14th International Joint Conference on Artificial Intelligence. (Montreal, Morgan Kaufmann Publishers Inc. San Francisco, CA, USA), pp 448-453
27. Schlicker, A., Domingues, F.S., Rahnenfuhrer, J., and Lengauer, T. (2006). A new measure for functional similarity of gene products based on Gene Ontology. BMC Bioinformatics *7*, 302.
28. Deng, Y., Gao, L., Wang, B., and Guo, X. (2015). HPOSim: an R package for phenotypic similarity measure and enrichment analysis based on the human phenotype ontology. PLoS One *10*, e0115692.
29. Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. (2000). The Protein Data Bank. Nucleic Acids Res *28*, 235-242.
30. Frappier, V., Chartier, M., and Najmanovich, R.J. (2015). ENCoM server: exploring protein conformational space and the effect of mutations on protein function and stability. Nucleic Acids Res *43*, W395-400.
31. Kononenko, N.L., Puchkov, D., Classen, G.A., Walter, A.M., Pechstein, A., Sawade, L., Kaempf, N., Trimbuch, T., Lorenz, D., Rosenmund, C., et al. (2014). Clathrin/AP-2 mediate synaptic vesicle reformation from endosome-like vacuoles but are not essential for membrane retrieval at central synapses. Neuron *82*, 981-988.
32. Posor, Y., Eichhorn-Gruenig, M., Puchkov, D., Schoneberg, J., Ullrich, A., Lampe, A., Muller, R., Zarbakhsh, S., Gulluni, F., Hirsch, E., et al. (2013). Spatiotemporal control of endocytosis by phosphatidylinositol-3,4-bisphosphate. Nature *499*, 233-237.

33. Schoneberg, J., Lehmann, M., Ullrich, A., Posor, Y., Lo, W.T., Lichtner, G., Schmoranzer, J., Haucke, V., and Noe, F. (2017). Lipid-mediated PX-BAR domain recruitment couples local membrane constriction to endocytic vesicle fission. Nat Commun *8*, 15873.

34. Jackson, L.P., Kelly, B.T., McCoy, A.J., Gaffry, T., James, L.C., Collins, B.M., Honing, S., Evans, P.R., and Owen, D.J. (2010). A large-scale conformational change couples membrane recruitment to cargo binding in the AP2 clathrin adaptor complex. Cell *141*, 1220-1229.

35. Kononenko, N.L., Classen, G.A., Kuijpers, M., Puchkov, D., Maritzen, T., Tempes, A., Malik, A.R., Skalecka, A., Bera, S., Jaworski, J., et al. (2017). Retrograde transport of TrkB-containing autophagosomes via the adaptor AP-2 mediates neuronal complexity and prevents neurodegeneration. Nat Commun *8*, 14819.

36. Dittman, J., and Ryan, T.A. (2009). Molecular circuitry of endocytosis at nerve terminals. Annu Rev Cell Dev Biol *25*, 133-160.

37. Saheki, Y., and De Camilli, P. (2012). Synaptic vesicle endocytosis. Cold Spring Harb Perspect Biol *4*, a005645.

38. Kittler, J.T., Chen, G., Honing, S., Bogdanov, Y., McAinsh, K., Arancibia-Carcamo, I.L., Jovanovic, J.N., Pangalos, M.N., Haucke, V., Yan, Z., et al. (2005). Phospho-dependent binding of the clathrin AP2 adaptor complex to GABAA receptors regulates the efficacy of inhibitory synaptic transmission. Proc Natl Acad Sci U S A *102*, 14871-14876.

39. Froemke, R.C. (2015). Plasticity of cortical excitatory-inhibitory balance. Annu Rev Neurosci *38*, 195-219.

40. Mitsunari, T., Nakatsu, F., Shioda, N., Love, P.E., Grinberg, A., Bonifacino, J.S., and Ohno, H. (2005). Clathrin adaptor AP-2 is essential for early embryonal development. Mol Cell Biol *25*, 9318-9323.

41. Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., et al. (2016). Analysis of protein-coding genetic variation in 60,706 humans. Nature *536*, 285-291.

42. Petrovski, S., Wang, Q., Heinzen, E.L., Allen, A.S., and Goldstein, D.B. (2013). Genic intolerance to functional variation and the interpretation of personal genomes. PLoS Genet *9*, e1003709.

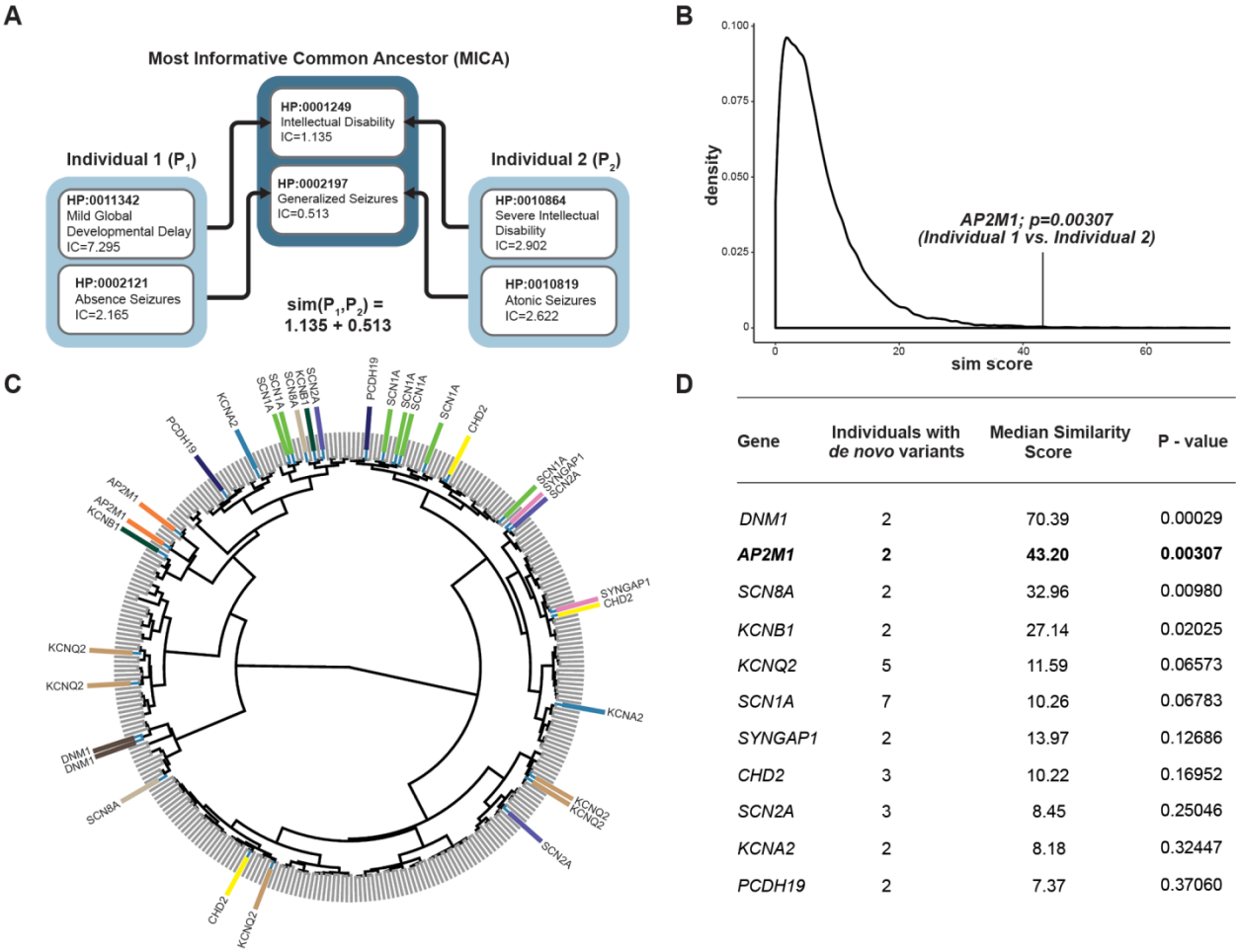
43. Schubert, J., Siekierska, A., Langlois, M., May, P., Huneau, C., Becker, F., Muhle, H., Suls, A., Lemke, J.R., de Kovel, C.G., et al. (2014). Mutations in STX1B, encoding a presynaptic protein, cause fever-associated epilepsy syndromes. Nat Genet *46*, 1327-1332.

44. Hamdan, F.F., Myers, C.T., Cossette, P., Lemay, P., Spiegelman, D., Laporte, A.D., Nassif, C., Diallo, O., Monlong, J., Cadieux-Dion, M., et al. (2017). High Rate of Recurrent De Novo Mutations in Developmental and Epileptic Encephalopathies. Am J Hum Genet *101*, 664-685.

45. von Spiczak, S., Helbig, K.L., Shinde, D.N., Huether, R., Pendziwiat, M., Lourenco, C., Nunes, M.E., Sarco, D.P., Kaplan, R.A., Dlugos, D.J., et al. (2017). DNM1 encephalopathy: A new disease of vesicle fission. Neurology *89*, 385-394.

46. Myers, C.T., Stong, N., Mountier, E.I., Helbig, K.L., Freytag, S., Sullivan, J.E., Ben Zeev, B., Nissenkorn, A., Tzadok, M., Heimer, G., et al. (2017). De Novo Mutations in PPP3CA Cause Severe Neurodevelopmental Disease with Seizures. Am J Hum Genet *101*, 516-524.

47. Kadlecova, Z., Spielman, S.J., Loerke, D., Mohanakrishnan, A., Reed, D.K., and Schmid, S.L. (2017). Regulation of clathrin-mediated endocytosis by hierarchical allosteric activation of AP2. J Cell Biol *216*, 167-179.




| Gene | Individuals with *de novo* variants | Median Similarity Score | P - value |
|---|---|---|---|
| *DNM1* | 2 | 70.39 | 0.00029 |
| ***AP2M1*** | **2** | **43.20** | **0.00307** |
| *SCN8A* | 2 | 32.96 | 0.00980 |
| *KCNB1* | 2 | 27.14 | 0.02025 |
| *KCNQ2* | 5 | 11.59 | 0.06573 |
| *SCN1A* | 7 | 10.26 | 0.06783 |
| *SYNGAP1* | 2 | 13.97 | 0.12686 |
| *CHD2* | 3 | 10.22 | 0.16952 |
| *SCN2A* | 3 | 8.45 | 0.25046 |
| *KCNA2* | 2 | 8.18 | 0.32447 |
| *PCDH19* | 2 | 7.37 | 0.37060 |

**Figure 1. HPO-based analysis demonstrating phenotypic similarity in individuals with *de novo* variants**

(A) Illustration of similarity score calculation using the Most Informative Common Ancestor (MICA) approach. (B) Distribution of phenotypic similarities between n=2 individuals in the cohort of 314 individuals, using 100,000 permutations. The vertical line indicates the observed value for two individuals with *de novo* variants in *AP2M1*. The observed value of 43.20 is in the top 0.3 percentile of the distribution, translating into an exact p-value of 0.00307. (C) Dendrogram of 314 individuals clustered by phenotypic similarity, using a ward.D2 algorithm for clustering of the similarity matrix. Gene labels refer to individuals with *de novo* variants in genes with two or more *de novo* variants in the entire cohort. (D) Exact p-values for observed versus predicted phenotypic similarity for all 11 genes with two or more individuals with *de novo* variants in the cohort of 314 individuals. P-values are uncorrected referring to the distribution of expected similarities for each number of individuals.

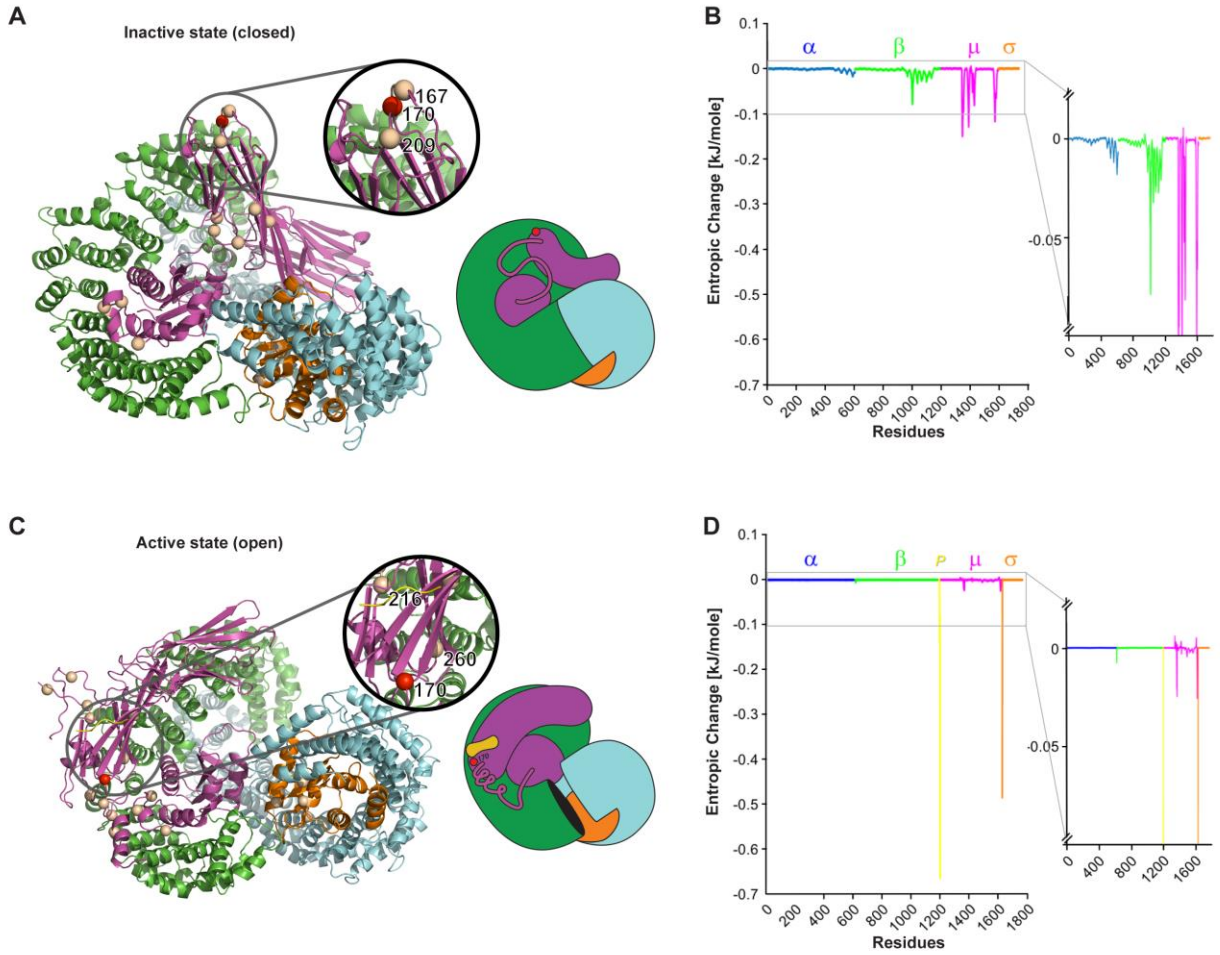



**Figure 2. Effect of the *AP2M1* p.Arg170Trp variant on the thermodynamic entropy of the AP-2 complex**

(A) The entire structure and a simplified cartoon model of the AP-2 complex is shown in the inactive closed state. Color-coding: AP-2α (blue), AP-2β (green), AP-2μ (magenta), AP-2 σ (orange), AP-2-bound cargo peptide P (yellow). The p.Arg170Trp variant is depicted as a red sphere. Further pathogenic variants in *AP2S1* (AP-2σ) and rare population variants in *AP2M1* (AP-2μ) variants are shown as golden spheres. (B) Difference in entropy (entropic change; ΔG) between the wildtype AP-2 complex and the AP-2 complex containing the *AP2M1* p.Arg170Trp variant are graphically depicted as ΔG for each residue across the entire AP-2 complex for the inactive closed state. The inlay shows an enlarged y-axis to emphasize the

differences in ΔG for each subunit. (C) The structure of the AP-2 complex and a simplified cartoon model including the bound cargo peptide P (yellow) is shown in the active state with labelling of *AP2S1* and *AP2M1* variants. (D) Difference in entropy between the wildtype AP2 complex and the AP-2 complex containing the *AP2M1* p.Arg170Trp variant for the active open state.

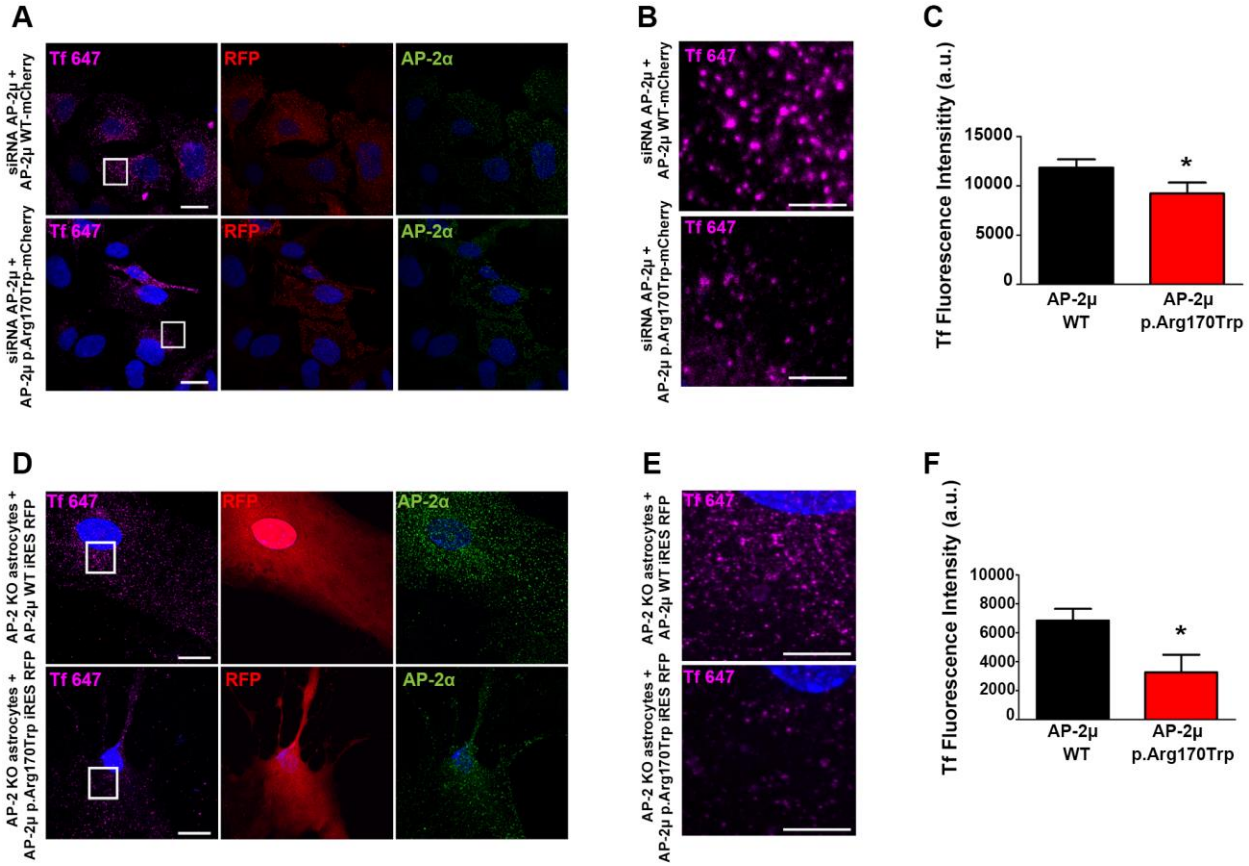



**Figure 3. Defective clathrin-mediated endocytosis (CME) in cells expressing the AP2M1 p.Arg170Trp variant** (A) Representative images of HeLa cells depleted of endogenous AP-2µ rescued by re-expression of siRNA-resistant mCherry-AP-2µ wildtype (WT) or p.Arg170Trp mutant and allowed to internalize AlexaFluor$^{647}$-labeled transferrin (Tf) for 10 min at 37ºC. Cells were fixed and immunostained for endogenous AP-2$\alpha$ and RFP. RFP was labeled to amplify the signal for mCherry and identify transfected cells. Scale bars: 20µm. (B) Zoom of the marked area in (A) illustrates reduced Tf endocytosis in cells expressing the p.Arg170Trp variant. Scale bars: 5µm. Note the punctate distribution of WT or Arg170Trp mutant mCherry-AP-2µ consistent with its proper targeting to endocytic pits (see also Figure S4). (C) Quantification of data shown in (A). Data represent mean ± SEM, $N$ = 3 independent experiments (with n= 198 for AP-2µ WT, n= 172 for AP-2µ p.Arg170Trp total cells analyzed). *$P < 0.05$, paired two-tailed $t$-test. (D) Representative images of primary astrocytes from WT or AP-2µ knockout (KO) mice rescued by re-expression of untagged AP-2µ WT or p.Arg170Trp together with soluble RFP and allowed to internalize

AlexaFluor$^{647}$-labeled transferrin (Tf) for 5 min at 37ºC. Cells were fixed and immunostained for endogenous AP-2$\alpha$ and cytoplasmic RFP. RFP expressed from the same construct after an internal ribosomal entry site (IRES) was labeled to identify transfected cells. Scale bars: 20µm. (E) Zoom of the marked area in (D) shows less Tf in astrocytes expressing the mutant variant of AP-2µ. Scale bars: 10µm. (F) Quantification of data shown in (D). Data represent mean ± SEM, N = 5 independent experiments (with n= 74 for AP-2µ WT, n= 72 for AP-2µ p.Arg170Trp total cells analyzed). *$P < 0.05$, Unpaired *t*-test.

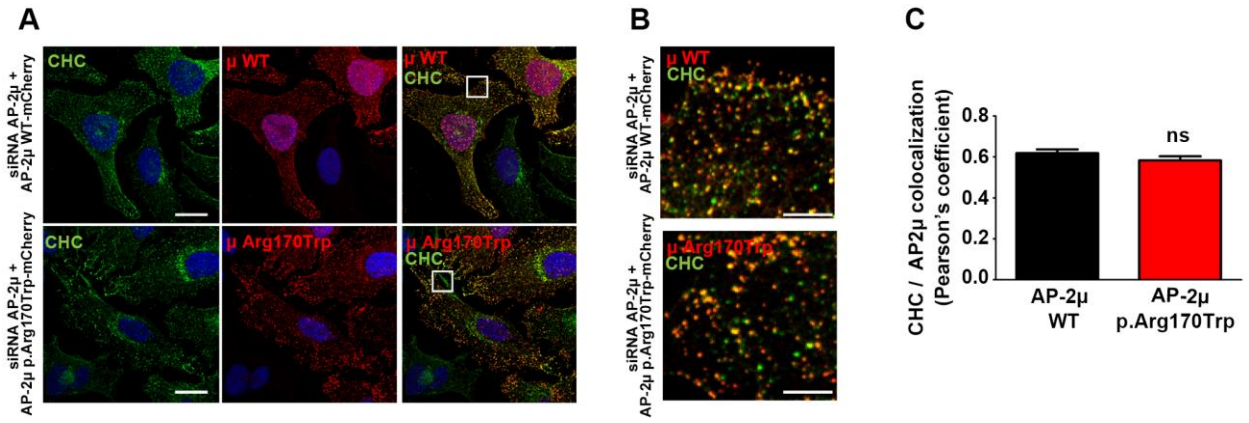


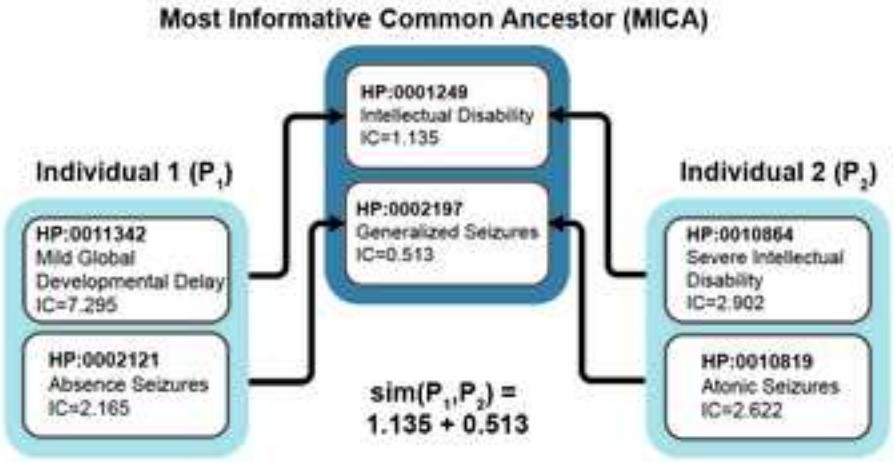

**Figure 4. Intact localization of AP-2 complex carrying the p.Arg170Trp variant** (A) Representative confocal images (maximum intensity projections) of HeLa cells depleted of endogenous AP-2µ rescued by re-expression of siRNA-resistant mCherry-AP-2µ WT or p.Arg170Trp mutant and immunostained with clathrin heavy chain (CHC) and RFP antibodies. RFP was labeled to amplify the signal for mCherry-tagged variants and identify transfected cells. Scale bars: 20µm. (B) Merged magnified views of boxed area in (A). Scale bars: 5µm. (C) Pearson's correlation coefficient for the co-localization of AP-2µ WT or p.Arg170Trp with clathrin heavy chain (CHC). Data represent mean ± SEM, N = 5 independent experiments. Paired *t*-test.

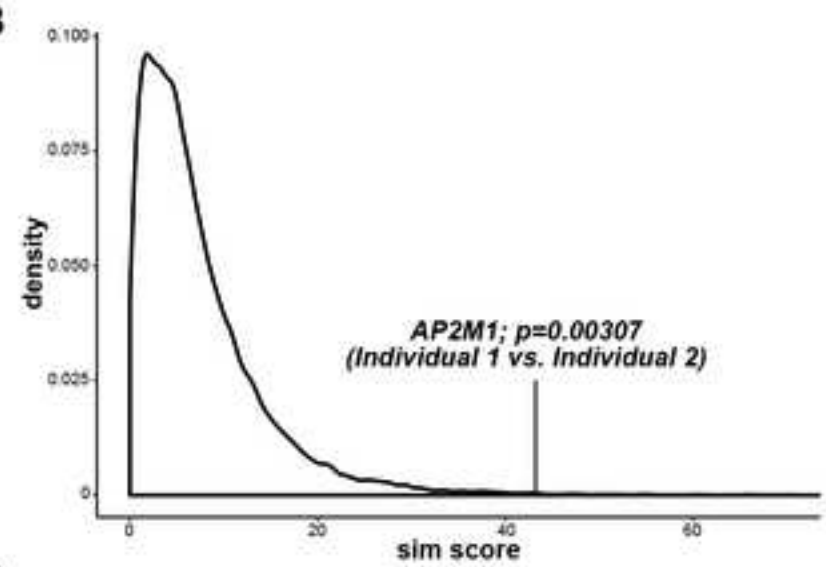
**Table 1. Phenotypic details of individuals with the recurrent *AP2M1* c.508C>T; p.(Arg170Trp) variant**

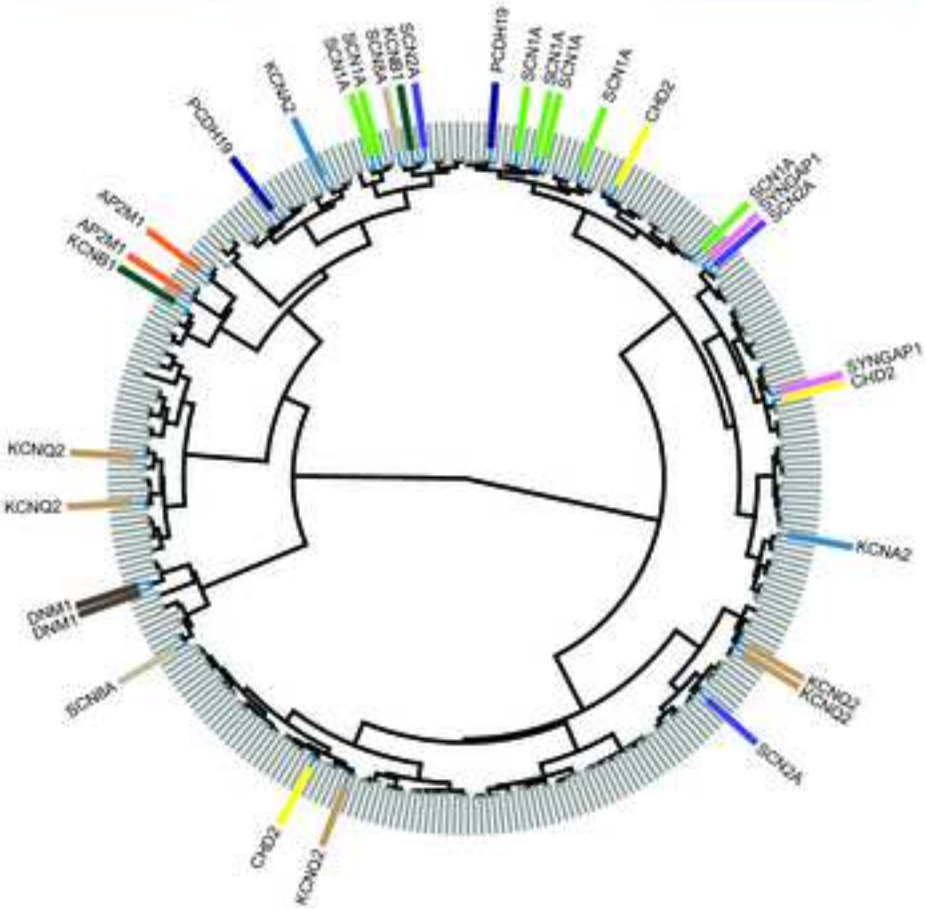
| | PROBAND ID | | | |
|---|---|---|---|---|
| | **Individual 1** | **Individual 2** | **Individual 3** | **Individual 4** |
| **Age/Sex** | 7y/F | 15y/F | 4y/F | 8y/F |
| **Development** | globally delayed | globally delayed | globally delayed | globally delayed |
| **Age at seizure onset** | 1y9m | 1y3m | 3y | 4y |
| **Seizure types** | atypical absence; myoclonic atonic; absence with eyelid myoclonia | atonic; atypical absence; absence with eyelid myoclonia | atonic; bilateral tonic-clonic | focal impaired-awareness |
| **Seizure outcome** | drug-responsive | drug-resistant | drug-resistant | partially responsive |
| **Intellectual disability (severity)** | moderate | moderate | severe | severe |
| **Autism Spectrum Disorder** | absent | absent (aggressive and self-harming behaviors) | present | present |
| **Ataxia** | truncal and gait ataxia | absent | gait ataxia only | truncal and gait ataxia |
| **Other exam findings** | hypotonia | hypotonia | hypotonia; chorea and myoclonus; prominent maxilla and thin upper lip | hypotonia; tremor; long thin hands and feet |
| **MRI findings** | parieto-occipital white matter abnormalities | normal | normal | normal |
| **EEG findings** | generalized polyspike-wave discharges | 3-4 Hz generalized spike-wave discharges | generalized spike-wave discharges | multifocal epileptiform activity |

Figure 1

Click here to access/download;Figure;Figure 1.tif




D

| Gene | Individuals with *de novo* variants | Median Similarity Score | P - value |
|---|---|---|---|
| *DNM1* | 2 | 70.39 | 0.00029 |
| ***AP2M1*** | **2** | **43.20** | **0.00307** |
| *SCN8A* | 2 | 32.96 | 0.00980 |
| *KCNB1* | 2 | 27.14 | 0.02025 |
| *KCNQ2* | 5 | 11.59 | 0.06573 |
| *SCN1A* | 7 | 10.26 | 0.06783 |
| *SYNGAP1* | 2 | 13.97 | 0.12686 |
| *CHD2* | 3 | 10.22 | 0.16952 |
| *SCN2A* | 3 | 8.45 | 0.25046 |
| *KCNA2* | 2 | 8.18 | 0.32447 |
| *PCDH19* | 2 | 7.37 | 0.37060 |

Figure 2 revised

Click here to access/download;Figure;Figure 2_rev3_v2.tif

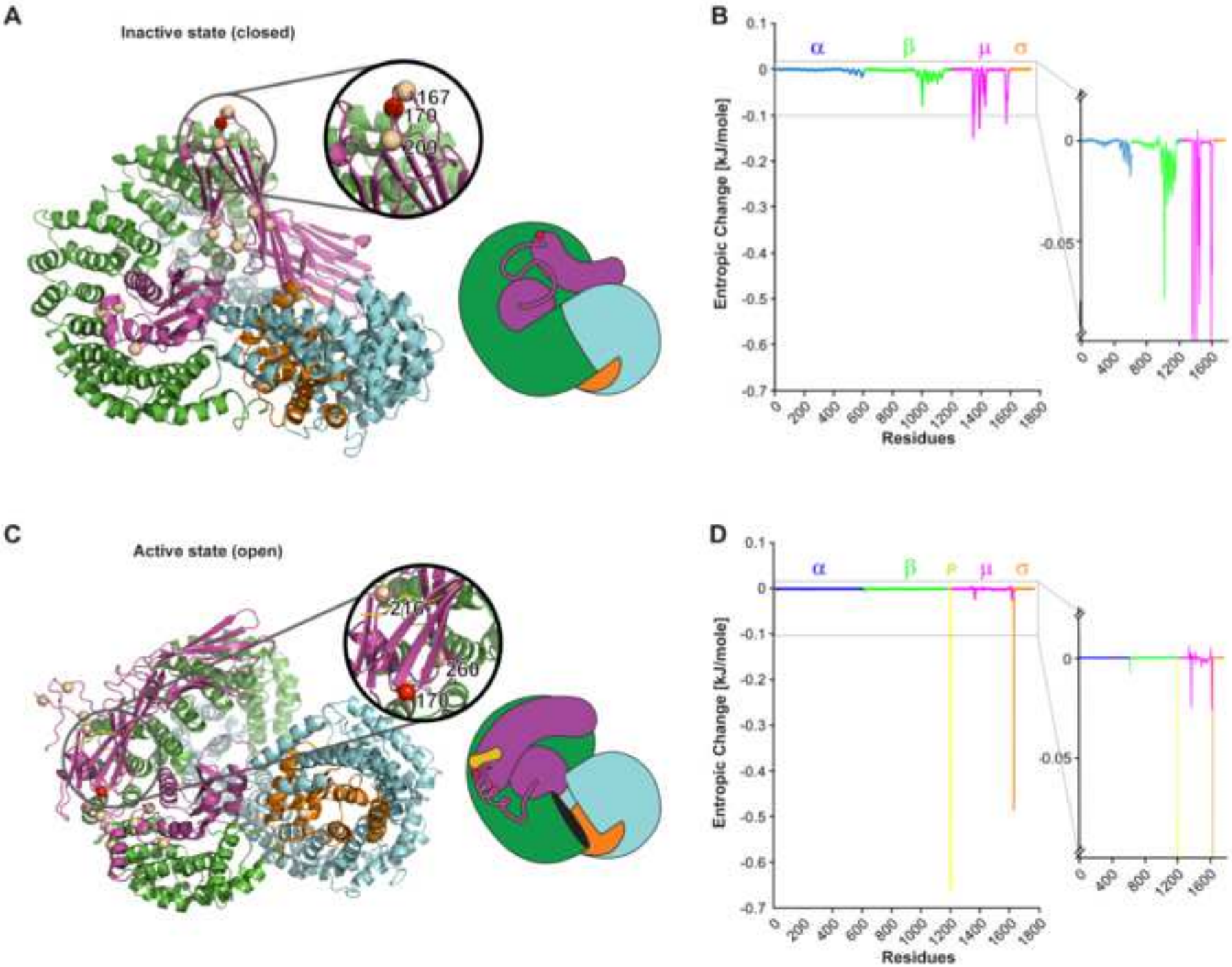


Figure 3



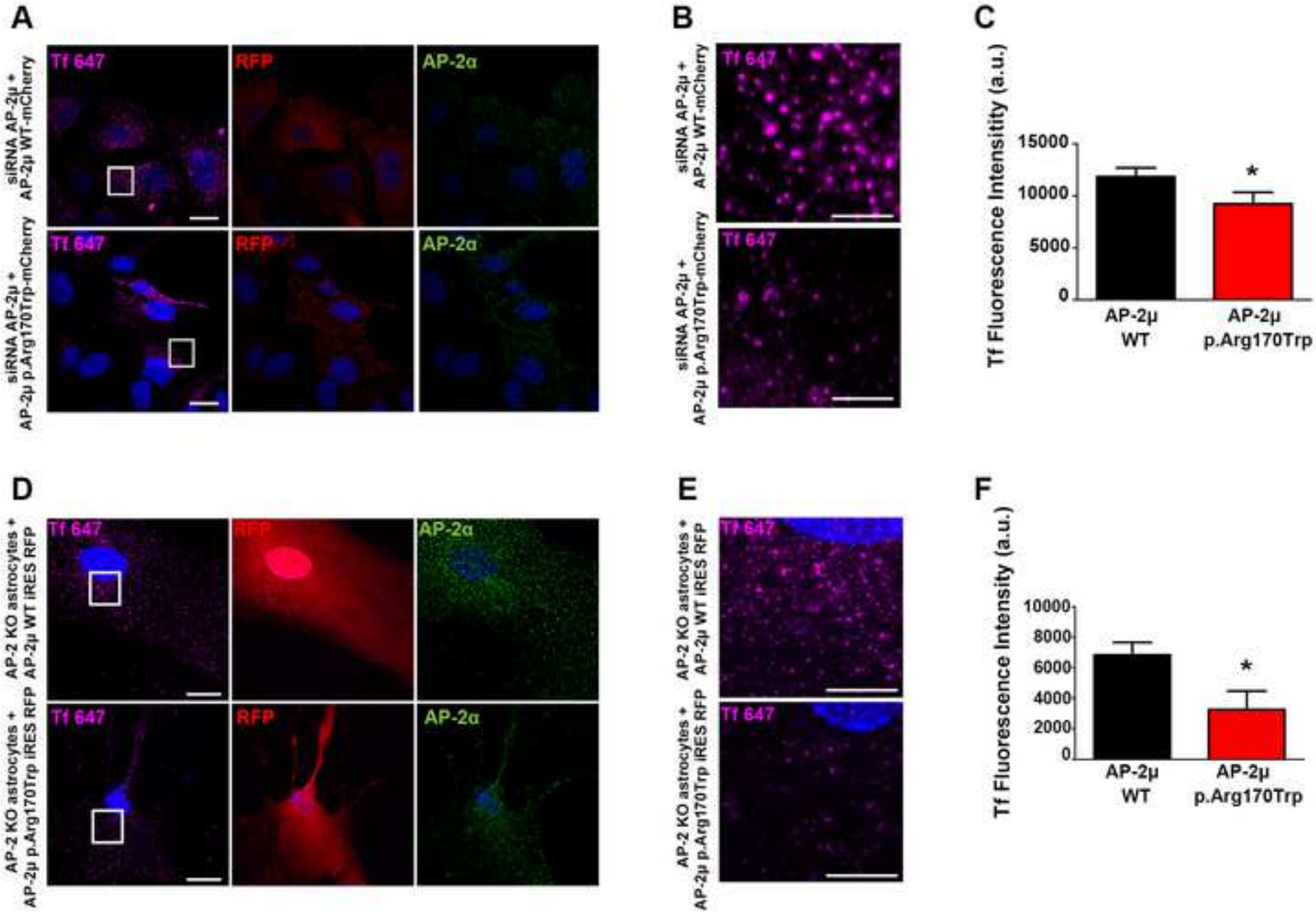
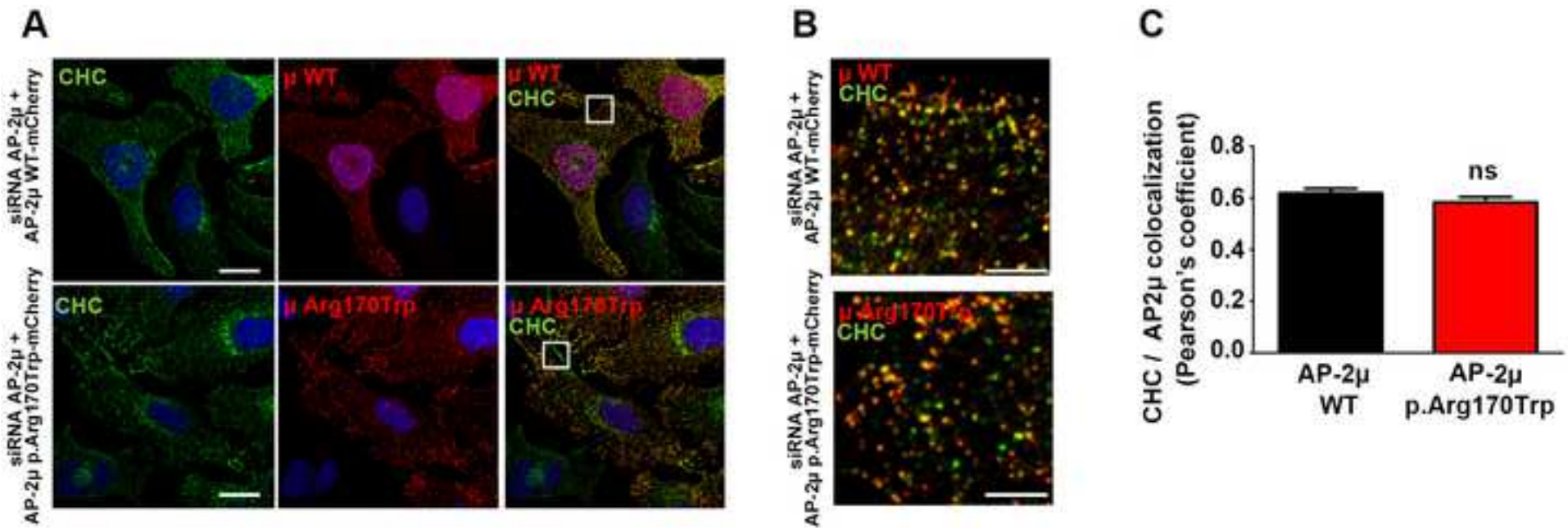


Figure 4





# AJHG DECLARATION OF INTERESTS POLICY

Transparency is essential for a reader's trust in the scientific process and for the credibility of published articles. At AJHG, we feel that disclosure of competing interests is a critical aspect of transparency. Therefore, we ask that all authors disclose any financial or other interests related to the submitted work that (1) could affect or have the perception of affecting the author's objectivity, or (2) could influence or have the perception of influencing the content of the article, in a "Declaration of Interests" section.

***What types of articles does this apply to?***

We ask that you disclose competing interests for all submitted content, including research articles as well as front matter (e.g., Reviews, Previews, etc.) by completing and submitting the "Declaration of Interests" form below. We also ask that you include a "Declaration of Interests" section in the text of all research articles even if there are no interests declared. For front matter, we ask you to include a "Declaration of Interests" section only when you have information to declare.

***What should I disclose?***

We ask that you and all authors disclose any personal financial interests (examples include stocks or shares in companies with interests related to the submitted work or consulting fees from companies that could have interests related to the work), professional affiliations, advisory positions, board memberships, or patent holdings that are related to the subject matter of the contribution. As a guideline, you need to declare an interest for (1) any affiliation associated with a payment or financial benefit exceeding $10,000 p.a. or 5% ownership of a company or (2) research funding by a company with related interests. You do not need to disclose diversified mutual funds, 401ks, or investment trusts.

***Where do I declare competing interests?***

Competing interests should be disclosed on the "Declaration of Interests" form as well as in the last section of the manuscript before the "References" section, under the heading "Declaration of Interests". This section should include financial or other competing interests as well as affiliations that are not included in the author list. Examples of "Declaration of Interests" language include:

> "AUTHOR is an employee and shareholder of COMPANY."
> "AUTHOR is a founder of COMPANY and a member of its scientific advisory board."

*NOTE:* Primary affiliations should be included on the title page of the manuscript with the author list and do not need to be included in the "Declaration of Interests" section. Funding sources should be included in the "Acknowledgments" section and also do not need to be included in the "Declaration of Interests" section. (A small number of front-matter article types do not include an "Acknowledgments" section. For these articles, reporting of funding sources is not required.)

***What if there are no competing interests to declare?***

For *research* articles, if you have no competing interests to declare, please note that in a "Declaration of Interests" section with the following wording:

> "The authors declare no competing interests."

*Front-matter* articles do not need to include this section when there are no competing interests to declare.

# AJHG DECLARATION OF INTERESTS FORM

If submitting materials via Editorial Manager, please complete this form and upload with your final submission. Otherwise, please e-mail as an attachment to the editor handling your manuscript.

***Please complete each section of the form and insert any necessary "Declaration of Interest" statement in the text box at the end of the form. A matching statement should be included in a "Declaration of Interest" section in the manuscript.***

***<u>Institutional Affiliations</u>***

We ask that you list the current institutional affiliations of all authors, including academic, corporate, and industrial, on the title page of the manuscript. ***Please select one of the following:***

- (●) All affiliations are listed on the title page of the manuscript.
- ( ) I or other authors have additional affiliations that we have noted in the "Declaration of Interests" section of the manuscript and on this form below.

***<u>Funding Sources</u>***

We ask that you disclose all funding sources for the research described in this work. ***Please confirm the following:***

- (●) All funding sources for this study are listed in the "Acknowledgments" section of the manuscript.*

*A small number of front-matter article types do not include an "Acknowledgments" section. For these, reporting funding sources is not required.

***<u>Competing Financial Interests</u>***

We ask that authors disclose any financial interests, including financial holdings, professional affiliations, advisory positions, board memberships, receipt of consulting fees etc., that:

(1) could affect or have the perception of affecting the author's objectivity, *or*
(2) could influence or have the perception of influencing the content of the article.

***Please select one of the following:***

- ( ) The authors have no financial interests to declare.
- (●) I or other authors have noted any financial interests in the "Declaration of Interests" section of the manuscript and on this form below.

***Advisory/Management and Consulting Positions***

We ask that authors disclose any position, be it a member of a Board or Advisory Committee or a paid consultant, that they have been involved with that is related to this study. ***Please select one of the following:***

- (●) The authors have no positions to declare.
- ( ) I or other authors have management/advisory or consulting relationships noted in the "Declaration of Interests" section of the manuscript and on this form below.

***Patents***

We ask that you disclose any patents related to this work by any of the authors or their institutions. ***Please select one of the following:***

- (●) The authors have no related patents to declare.
- ( ) I or one of my authors have a patent related to this work, which is noted in the "Declaration of Interests" section of the manuscript and on this form below.

***Please insert any "Declaration of Interests" statement in this space.*** This exact text should also be included in the "Declaration of Interests" section of the manuscript. If no authors have a competing interest, please insert the text, "The authors declare no competing interests."

D.N.S. and S.T. are full time employees of Ambry Genetics. Exome sequencing is one of Amb ry's commercially available diagnostic tests.

**On behalf of all authors, I declare that I have disclosed all competing interests related to this work. If any exist, they have been included in the "Declaration of Interests" section of the manuscript.**

**Name:** Ingo Helbig

**Manuscript Number (if available):** AJHG-D-19-00045