BRAIN
ORIGINAL ARTICLE

In-depth analysis reveals complex molecular aetiology in a cohort of idiopathic cerebral palsy

Na Li,[1,†] Pei Zhou,[1,†] Hongmei Tang,[2,†] Lu He,[2,†] Xiang Fang,[1,†] Jinxiang Zhao,[3] Xin Wang,[3] Yifei Qi,[4] Chuanbo Sun,[1] Yunting Lin,[5] Fengying Qin,[1] Miaomiao Yang,[1] Zhan Zhang,[1] Caihua Liao,[1] Shuxin Zheng,[1] Xiaofang Peng,[1] Ting Xue,[1] Qianying Zhu,[1] Hong Li,[1] Yan Li,[1] Liru Liu,[2] Jingyu Huang,[2] Li Liu,[5] Changgeng Peng,[6] Angela M. Kaindl,[7,8,9] Jozef Gecz,[10] Dingding Han,[1,‡] Dong Liu,[3,‡] Kaishou Xu[2,‡] and Hao Hu[1,11,12,13,‡]

[†,‡]These authors contributed equally to this work.

Cerebral palsy is the most prevalent physical disability in children; however, its inherent molecular mechanisms remain unclear.

In the present study, we performed in-depth clinical and molecular analysis on 120 idiopathic cerebral palsy families, and identified underlying detrimental genetic variants in 45% of these patients. In addition to germline variants, we found disease-related post-zygotic mutations in ∼6.7% of patients with cerebral palsy. We found that patients with more severe motor impairments or a comorbidity of intellectual disability had a significantly higher chance of harbouring disease-related variants. By a compilation of 114 known cerebral palsy-related genes, we identified characteristic features in terms of inheritance and function, from which we proposed a dichotomous classification system according to the expression patterns of these genes and associated cognitive impairments. In two patients with both cerebral palsy and intellectual disability, we revealed that the defective TYW1, a tRNA hypermodification enzyme, caused primary microcephaly and problems in motion and cognition by hindering neuronal proliferation and migration. Furthermore, we developed an algorithm and demonstrated in mouse brains that this malfunctioning hypermodification specifically perturbed the translation of a subset of proteins involved in cell cycling. This finding provided a novel and interesting mechanism for congenital microcephaly. In another cerebral palsy patient with normal intelligence, we identified a mitochondrial enzyme GPAM, the hypomorphic form of which led to hypomyelination of the corticospinal tract in both human and mouse models. In addition, we confirmed that the aberrant *Gpam* in mice perturbed the lipid metabolism in astrocytes, resulting in suppressed astrocytic proliferation and a shortage of lipid contents supplied for oligodendrocytic myelination.

Taken together, our findings elucidate novel aspects of the aetiology of cerebral palsy and provide insights for future therapeutic strategies.

1 Laboratory of Medical Systems Biology, Guangzhou Women and Children's Medical Center, Guangzhou Medical University, 510623, Guangzhou, China
2 Department of Rehabilitation, Guangzhou Women and Children's Medical Center, Guangzhou Medical University, 510120, Guangzhou, China

3   Key Laboratory of Neuroregeneration of Jiangsu and Ministry of Education, Nantong University, 226001, Nantong, China
4   Division of Uterine Vascular Biology, Guangzhou Women and Children's Medical Center, Guangzhou Medical University, 510623, Guangzhou, China
5   Department of Genetics and Endocrinology, Guangzhou Women and Children's Medical Center, Guangzhou Medical University, 510623, Guangzhou, China
6   The First Rehabilitation Hospital of Shanghai, Tongji University School of Medicine, 200029, Shanghai, China
7   Institute of Cell Biology and Neurobiology, Charité-Universitätsmedizin, 13353, Berlin, Germany
8   Department of Pediatric Neurology, Charité-Universitätsmedizin, 13353, Berlin, Germany
9   Center for Chronically Sick Children, Charité-Universitätsmedizin, 13353, Berlin, Germany
10 Adelaide Medical School, University of Adelaide, SA5005, Adelaide, Australia
11 Guangdong Provincial Key Laboratory of Research in Structural Birth Defect Disease, Guangzhou Women and Children's Medical Center, Guangzhou Medical University, 510623, Guangzhou, China
12 Third Affiliated Hospital of Zhengzhou University, 450052, Zhengzhou, China
13 School of Medicine, South China University of Technology, 510641, Guangzhou, China

Correspondence to: Hao Hu, PhD
Laboratory of Medical Systems Biology, Guangzhou Women and Children's
Medical Center, Guangzhou Medical University, Guangzhou, China
E-mail: huh@cougarlab.org

Correspondence may also be addressed to: Kaishou Xu, MD
Department of Rehabilitation, Guangzhou Women and Children's Medical Center
Guangzhou Medical University, Guangzhou, China
E-mail: xksyi@126.com

Dong Liu, PhD
Key Laboratory of Neuroregeneration of Jiangsu and Ministry of Education
Nantong University, Nantong, China
E-mail: liudongtom@gmail.com

Dingding Han, PhD
Laboratory of Medical Systems Biology
Guangzhou Women and Children's Medical Center
Guangzhou Medical University, Guangzhou, China
E-mail: handd@cougarlab.org

Introduction

Cerebral palsy is an umbrella term spanning a group of disorders with compromised ambulant performances, and is attributed to non-progressive disturbances in developing foetal and infant brains.[1] Cerebral palsy can be categorized according to motor signs (spasticity, dyskinesia, ataxia and hypotonia, among which spasticity manifests in ~80% of cerebral palsy cases), involvement of extremities (hemiplegia, monoplegia, diplegia, triplegia and quadriplegia), and anatomical sites of brain lesions (cerebral cortex, pyramidal tract, extrapyramidal system and cerebellum).[2,3] In cerebral palsy children aged 2 years or older, severity is reliably classified via the five-level Gross Motor Function Classification System (GMFCS).[4] The common comorbidities of cerebral palsy include intellectual disability, epilepsy, vision and hearing problems, language deficits, autism spectrum disorders, sleep disorders and secondary musculoskeletal defects such as scoliosis and hip dislocation.[2,3] Cerebral palsy is the most prevalent physical disability of childhood, with a prevalence of

2.0–3.5 cases per 1000 live births in countries with varying degrees of socioeconomic development (>20 million patients around the world), and the incidence of cerebral palsy at term has remained consistent for the past 50 years.[1–3,5–7] Cerebral palsy poses a critical problem for rehabilitation and welfare systems, and the lifetime costs for a child with cerebral palsy in the USA are estimated to be $1 million.[5] The conventional risk factors of cerebral palsy consist of gestational and perinatal insults, including birth asphyxia, neonatal infections and teratogens.[1,5] However, two-thirds of patients with cerebral palsy are born at term, and birth asphyxia happens in <10% of cerebral palsy cases.[2,6,7] Because the estimated upper limit of unconfirmed causality of cerebral palsy comprises 80% of cases, it is reasonable that unknown mechanisms exist accounting for most cerebral palsy events.[3]

A growing body of evidence has indicated prevailing genetic contributors to the aetiology of cerebral palsy, including reports of familial cases and the high concordance of monozygotic twins; furthermore, the positive correlation between cerebral palsy incidence and parental

ages suggests a genetic component, as does the prevalence of congenital malformation observed in patients with cerebral palsy.[8–12] Large cohorts of case-control studies assessing polymorphisms in a variety of candidate genes did not identify any significantly associated variants.[2,13,14] Therefore, the genetic underpinnings of cerebral palsy are probably similar to those of other neurodevelopmental disorders such as intellectual disability and autism spectrum disorder, where causative variants are rarely detected in association studies based on common variants but with large effect sizes.[15–17] In line with this hypothesis, clinically relevant copy number variation (CNV), both *de novo* and inherited, were recently identified in four collections of cerebral palsy patients, resolving 4–31% of the cryptogenic cases.[18–21] Meanwhile, gene-panel sequencing and whole-exome sequencing revealed pathogenic variants, mostly *de novo*, in an aggregate of six cerebral palsy studies, yielding molecular diagnoses for ~15% of cases.[22–27]

So far, the known genetic and molecular information on cerebral palsy aetiology has been insufficient, which impedes improvements in prophylaxis, prognosis and treatment of cerebral palsy patients in the era of precision medicine. Large cohorts need to be recruited to fully decipher the corresponding gene spectrum due to the high genetic heterogeneity of cerebral palsy, and geographic and ethnic biases should be addressed. Whole-genome sequencing (WGS), in comparison with gene-panel sequencing or whole-exome sequencing, can render better coverage of predisposing alleles, and somatic mosaicism or post-zygotic mutations in patients with cerebral palsy must be highlighted as in other neurodevelopmental disorders.[28,29] Hence, targeted investigation into the common characteristics of cerebral palsy-related genes, in addition to detailed mechanistic studies, is warranted before the introduction of a gene-based classification system that may have a profound impact on clinical practice when treating patients with cerebral palsy.

Materials and methods

Cohort recruitment

The cerebral palsy families were recruited by the Department of Rehabilitation of the Guangzhou Women and Children's Medical Center (GWCMC) from August 2014 to June 2017 in a cohort study. This study was approved by the ethics committee of GWCMC (approval no. 2016061601), and written informed consents were signed by all families involved in this study. The family history and clinical characteristics of individuals with cerebral palsy were documented regarding the cerebral palsy phenotypic profiles. The motor signs (spastic, ataxic, dyskinetic etc.) and anatomical distribution (monoplegia, diplegia, triplegia, hemiplegia and quadriplegia) of movement disorders were evaluated and categorized. Motor severity was measured by the GMFCS on five levels.[30] The common comorbidities, i.e. intellectual disability, vision/hearing impairments, epilepsy, dyslexia etc., were recorded for each patient. The fine motor skills were evaluated by the Manual Ability Classification System (MACS).[31] Rehabilitation training outcome was determined by the Activity of Daily Living (ADL),[32] and the Gross Motor Function Measure (GMFM),[33] both of which were measured twice before and after treatments.

MRI and diffusion tensor imaging

All MRIs were performed on a 3.0 T Siemens scanner (Magnetom Systems, Skyra). For diffusion tensor imaging (DTI), an echoplanar sequence with diffusion gradients ($b = 1000$ s/mm^2) applied in 64 non-collinear directions was used. Basic parameters included: repetition time (10 000 ms)/echo time (92 ms), voxel (2.0 × 2.0 × 2.0 mm), field of view (256 × 256) and layer thickness (2.0 mm). DTI measurement was 3D reconstruction of colour-coded fractional anisotropy graphics through DTI Studio (DWI standard deviation).

Analysis pipeline of deep sequencing data

The variant prioritization and evaluation of aetiological involvement were performed using two pipelines, namely MERAP and ANNOVAR. In brief, the MERAP pipeline first filters all identified variants through comparison with the disease-associated variants in the Human Gene Mutation Database (HGMD, 2020.2) and the Online Mendelian Inheritance in Man (OMIM) database to collect known disease-causing variants. To filter out neutral variants, MERAP uses entries from the dbSNP143 (http://www.ncbi.nlm.nih.gov/projects/SNP/), 1000 Genomes (http://www.1000genomes.org/), NHLBI Exome Sequencing Project (ESP, http://evs.gs.washington.edu/EVS/), ExAc (http://exac.broadinstitute.org/) and the Chinese Millionome Database (CMDB; https://db.cngb.org/cmdb/) as screening databases. In principle, candidate variants causing recessive traits should not occur in healthy controls as homozygotes, and the frequency of respective heterozygotes should not exceed 0.1%. MERAP uses the RefSeq genes (http://www.ncbi.nlm.nih.gov/refseq/) as a reference, and non-synonymous changes are described in terms of gene ID, base change, protein change, genomic coordinate, transcript coordinate, protein coordinate, protein length, affiliated with gene description from the Human Gene Nomenclature Committee (http://www.genenames.org/). Changes destroying conventional splice sites or introducing novel splice sites were identified by the MERAP module SSFinder. To assess the pathogenicity of missense mutations, MERAP generates a single score integrating the results of seven different algorithms, including the Grantham score, PhyloP, GERP, SIFT, PolyPhen2, MutationTaster and the Conserved Domains Database (http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml). With empirical false discovery rate cut-offs, this score serves as dichotomized pathogenicity predictions even if any two of the seven algorithms might not coincide, as is often the case. MERAP rules out candidate genes reported to harbour homozygous loss-of-function variants in healthy individuals, which applies to >1% of human genes. Typically, if >3 independent truncating variants are observed in >10 of the exomes listed in the 1000 Genomes, Exome Variant Server (ESP6500), NHLBI GO ESP (http://evs.gs.washington.edu/EVS/) and ExAC (http://exac.broadinstitute.org/) databases, the relevant gene is flagged as loss-of-function tolerant. To facilitate the choice between few remaining candidate genes, MERAP also provides a list of ~4500 known disease genes extracted from OMIM (http://www.ncbi.nlm.nih.gov/omim), ClinVar (http://www.ncbi.nlm.nih.gov/clinvar/) and HGMD (http://www.hgmd.org/), as well as >8000 associated disorders and their symptoms. For novel candidate genes without a previous link to disease, MERAP offers information on their interaction with known disease genes, based on data from Biogrid (http://thebiogrid.org/) and IntAct (http://www.ebi.ac.uk/intact/), assuming that genes implicated in clinically similar disorders tend to cluster in gene or protein interaction networks. Variants were considered *de novo* if neither parent had the variant, and candidate variants were selected by segregation analysis. The pathogenic and probably pathogenic genes/variants were defined according to the standards and guidelines of ACMG.[34]

The ANNOVAR pipeline was also used, with additional considerations from the Residual Variation Intolerance Score (RVIS) and the Combined Annotation-Dependent Depletion (CADD) score.[35] The former ranks genes in terms of intolerance to functional genetic variation and the latter integrates several well-known tools. We empirically set a cut-off RVIS score of 50th percentile for known and novel genes and a cut-off CADD score of 20 for novel

candidate genes. When defining likely causal variants/genes, we followed the guidelines designated by ACMG.

All variants of putative clinical relevance were confirmed by the conventional PCR and Sanger sequencing. Parent-child relationships were confirmed by using the PLINK with single nucleotide polymorphisms (SNPs) drawn from WGS that matched the CytoScan SNP repertoire.[36] The false positive rate of WGS was generated by the comparison between Sanger sequencing and WGS, and we found no false positive variant called by our pipeline. The false negative rate of WGS was evaluated by the SNP genotyping comparison between WGS and CytoScan high-density microarray, which resulted in an average false negative rate of <0.1%.

Zebrafish modelling

All zebrafish experimentation was carried out in accordance with the NIH Guidelines for the care and use of laboratory animals (http://oacu.od.nih.gov/regs/index.htm) and ethically approved by the Administration Committee of Experimental Animals, Jiangsu Province, China [Approval ID: SYXK (SU) 2007–0021].

Zebrafish lines and whole-mount *in situ* hybridization

The zebrafish embryos and adults were maintained at the Zebrafish Center of Nantong University under conditions in accordance with our previous protocols.[37,38] The fish lines *Tg(mnx1:GFP)*[ml2] *Tg(huC:egfp)*, *Tg(huC:mcherry)* and *Tg(gfap:egfp)* have been described in previous work.[37,38] The whole-mount *in situ* hybridization (WISH) was performed as previously described.[39] The templates for generating antisense RNA probes were amplified from the cDNA library, using specific primers targeting *tyw1* and *gpam* as follows: *tyw1*-forward: 5′ AAG GCC AGC TGC AAG AAT AA; *tyw1*-reverse: 5′ GTG TGG TGT CTC CAG CAG AA; *gpam*-forward: 5′ TCC TGT TCA CAA GTC CCA CA; *gpam*-reverse: 5′ AGT TCT TGC GGA GCA TCC TA.

Morpholino and mRNA injections

The *tyw1* splice-blocking morpholino (Gene Tools) sequence was 5′ AAC CTT ATT CCC ACT TAA TGT TAC C. The *gpam* splice-blocking morpholino sequence was 5′ GGT GCT ACT TTT CTC CAA GCT TAC C. The sequence of a standard control morpholino oligo was 5′ CCT CTT ACC TCA GTT ACA ATT TAT A. The *tyw1* translation-blocking morpholino sequence was 5′ CAG CAT CTC ATG TAC TCT CTC CAT C. The *gpam* translation-blocking morpholino sequence was 5′ ACG TCC ATC CCC TCT CTT CAA ACC A. The morpholinos were diluted to 0.3 mM with RNase-free water and injected into the yolk of one to two-cell stage embryos and then raised in E3 medium at 28.5°C. The wild-type and mutated cDNAs [*TYW1* (NM_018264) mut1: p.R389Q; *TYW1* mut2: p.R206C; *GPAM* (NM_001244949) mut1: p.G499R; *GPAM* mut2: p.P669S] containing the open reading frame of the zebrafish and human *tyw1* or *gpam* genes were cloned into pCS2⁺ vector, respectively, and then were transcribed *in vitro* by using the mMESSAGE mMACHINE® Kit (Thermo Fisher Scientific) after the recombinant plasmids linearized with NotI restriction enzyme (NEB), and then the capped mRNAs were purified by RNeasy® Mini Kit (Qiagen). Around 2 nl of mRNA were injected at 50 ng/µl into half-cell stage embryos.

RNA isolation, reverse transcription and PCR

Total RNA was extracted from zebrafish embryos by TRIzol® reagent according to the manufacturer's instructions (Invitrogen, Thermo Fisher Scientific). The reverse transcription was carried out according to the standard protocols from manufacturer (Fermentas, Thermo Fisher Scientific). The PCR was carried out as

previously described to validate the efficiency of splicing-blocking effect of specific morpholino[37] using the following primers: *tyw1*-forward: 5′ TTA TCG GTG TTG TCG GGT TT; *tyw1*-reverse: 5′ CCT CTG CCA ATC TGT CTT CC; *gpam*-forward: 5′ ACA GTG CGC AAA AAG AGG TC; *gpam*-reverse: 5′ TAA GCA CGT TCT CCA CCA CA.

Locomotion analysis in zebrafish larvae, microscopy and statistical analysis

The locomotion analysis of 5 dpf zebrafish larvae was carried out by using the DanioVision system (Noldus Information Technology, Wageningen, Netherlands). The confocal imaging was performed by using a TCS-SP8 LSM confocal imaging system (Leica, Wetzlar, Germany). The zebrafish embryos were embedded as we previously did.[37] Photographs of *in situ* hybridization results were taken by using a DP70 camera on an Olympus stereomicroscope MVX10. Statistical comparisons of the data were carried out by Student's *t*-test, and $P < 0.05$ was considered statistically significant.

Mouse modelling

All mice experiments in this study were approved by the institutional animal care and use committee in the Guangzhou Medical University (registration no. 2019–436, 2019–694). In this study, all mice, either wild-type or mutant, were generated from the C57BL/6J strain and provided by the Cyagen Biosciences.

Generation of knockout mice by CRISPR–Cas9 system

Knockout mouse lines were created by CRISPR–Cas9-mediated genome engineering. To create a *Tyw1* knockout mouse model, exon 2–3 was selected as the target site. Guide RNAs (gRNA) were designed with the following sequences: gRNA1 (matching reverse strand of gene): GCCTAAGGACCACGTTTCGATGG; gRNA2 (matching reverse strand of gene): CATGTAAGACCTCACGATCAAGG.

To create a *Gpam* knockout mouse model, exon 2–4 was selected as the target site. Guide RNA was designed as the following sequences: gRNA1 (matching reverse strand of gene): CTCCCACGGGAAAACCACGCAGG; gRNA2 (matching reverse strand of gene): CCCAAGGGGTCACGCACCACAGG.

Cas9 mRNA and gRNA were co-injected into the cytoplasm of fertilized eggs. The microinjected zygotes were cultured in medium until the two-cell stage, and then were transferred into the oviducts of pseudopregnant ICR females at 0.5 days after mating with vasectomized males. The mutant F0 mice were identified by genomic PCR and the DNA sequence was confirmed by Sanger sequencing. Subsequent breeding of the mutant F0 mice generated offspring with desired genotypes for experiments. Gene knockout efficiency was identified using quantitative PCR (qPCR) (Supplementary Fig. 10).

Behavioural tests

All behavioural tests were carried out on both male and female mice at 8 weeks of age, and the balanced gender was kept for both wild-type and *Tyw1*⁻/⁻ or *Gpam*⁻/⁻ mice. Each test was conducted at fixed time of day (between 08:30 and 18:30) on each training day. Mice were moved to the testing room 1 h before behavioural testing for acclimation, and those participating in multiple tests were allowed to rest for at least 3 days between two tests. All experimental areas were cleaned with 70% ethanol before the tests and between participants. All behavioural tests were carried out with the presence of two researchers blinded to the genotype.

Morris water maze test

The Morris water maze test was performed to investigate the learning and memory ability of mice. Briefly, a platform was

placed at the central zone of one quadrant of the pool below the surface of water. The mice were trained to learn the position of the hidden platform for 4 days. At Day 5, the platform was removed. The mice were released into water and allowed to swim for 60 s to search a virtual quadrant centred on the location of the platform. Duration in zone of platform refers to amount of time mice remained in this virtual zone. Frequency of crossing platform refers to number of times mice crossed the virtual zone, while latency to first arrival refers to the time needed for mice to reach the virtual zone at first time, while distance travelled in target quadrant refers to the distance mice swam in the virtual quadrant. Videos were recorded and analysed by using the SuperMaze V2.0 software (XinRuan, Shanghai, China).

Rotarod test

Rotarod test was performed to investigate motor coordination of mice in each group. First, mice were trained three times at the rate of 30 rpm before test, and each training maintained for 5 min. In the test, the rotation rate increased gradually to reach 40 rpm within 300 s. The test was over when a mouse fell off. Each mouse was tested three times and the mean latency time at rotarod was recorded.

Grip strength test

Grip strength of mouse forelimbs was measured with a grip strength test meter (BIOSEB, EB Instruments). During the test, the mice were placed over the grid allowing only forepaws to attach to the grid. After the paws had grasped the grid, the tails were pulled horizontally until they completely released hold. Each mouse was tested three times. The readings of grip strength and duration of grasping the grid were recorded and analysed by using the SuperGSM software (XinRuan, Shanghai, China).

Tissue preparation

Mice were anaesthetized with 50 mg/kg sodium pentobarbital. After being perfused with 0.9% NaCl and 4% PFA, brains were extracted and postfixed for 24 h, followed by dehydration using sucrose. Brains were cut into two hemispheres, and then serial sagittal sections (30-μm thick) were sliced using a freezing microtome (CM3050S, Leica).

Nissl staining

Sections were immersed into 75% ethanol (30 s), dH$_2$O (30 s) and cresyl violet (2 min). Then, sections were dehydrated with gradient ethanol (75%, 95%, 100%) for 30 s, followed by incubation with xylene and mounted with neutral resins. The images were taken under an inverted microscope.

5-Ethynyl-2′-deoxyuridine labelling

For proliferation assays, pregnant dams on embryonic Day (E)13.5 were intraperitoneally injected with 5-ethynyl-2′-deoxyuridine (EdU, 20 mg/kg body weight) 2 h before euthanasia. Embryonic brains were harvested, fixed in PFA and immersed in 30% sucrose for dehydration. Brain slices were then stained with Click-iT EdU Imaging Kits (C10339, Invitrogen, Thermo Fisher Scientific). Cells incorporated with EdU were determined under a fluorescence microscope. Percentage of EdU$^+$ fluorescence area (indicating the cell number of EdU$^+$) of E13.5 mouse brain cortex was calculated in three random fields per coverslip. For migration assays, pregnant dams at E15.5 were intraperitoneally injected with EdU. Brains were harvested after 72 h and stained as before.

Immunostaining

To detect expression of TYW1 and GPAM in human brain tissue, the normal brain tissue microarrays (BNC17011, Biomax) were prepared. According to the instruction from Biomax official website, BNC17011 arrays were derived from normal brain tissue of male or female at the age of 2 to 50, including brain regions such as frontal lobe, apical lobe, occipital lobe, temporal lobe, midbrain, pons, medulla oblongata, thalamus opticus, cerebellum, hippocampus, callositas, optic nerve and spinal cord. The slides were immersed in xylene and graded ethanol to deparaffinize and rehydrate. After antigen retrieval using microwave with sodium citrate buffer, 0.3% Triton X-100 was added to permeabilize the tissue. The slides were blocked with goat serum and incubated with primary antibodies at 4°C overnight. After washing with PBS, the sections were incubated with secondary antibodies.

For mouse brain tissues, the perfused brain samples were fixed within 4% PFA in 0.01 M PBS at 4°C for 24 h, then washed with PBS twice, cryoprotected with sucrose gradients, snap frozen and sectioned with a cryostat (Leica) to thickness of 30 μm. Then the sections were immuno-stained by using aforementioned methods, except deparaffinization and rehydration.

The number of Cux1$^+$ or Foxp2$^+$ cells in frontal cortex and motor cortex was counted on three sagittal cerebrum sections. Gfap$^+$ and NeuN$^+$ cells in different brain regions were counted in three random fields per section. Data were obtained from at least three independent experiments and analysed with ImageJ software.

To visualize and measure the cholesterol level, filipin staining was performed in tissue sections (10-μm thick) using a cell-based Cholesterol Assay Kit (ab133116, Abcam). Briefly, the tissue sections were washed (3 × 5 min) with Cholesterol Detection wash buffer, and then the Filipin III was added to each section and incubated in the dark for 60 min at room temperature. After washing, the fluorescence images were obtained by a Leica DMi8 (Leica Microsystems) fluorescence microscope using a 20× HC PL FLUOTAR objective.

Western blot analysis

Total protein was extracted from peripheral blood using cell lysate containing RIPA and protease inhibitor cocktail. And the brain tissues were lysed with 2% SDS in PBS with PMSF and proteinase inhibitor cocktail. The BCA assay was used to determine protein concentration. Proteins were resolved on 7.5, 10 or 15% Tris-glycine gels based on different molecular weight and transferred to PVDF membranes. After blocking, the membranes were incubated with primary antibodies, followed by horseradish peroxidase (HRP)-labelled secondary antibodies. Then, blots were visualized by West Pico Plus Chemiluminescent Substrate (Thermo Fisher Scientific) and scanned using ChemiDoc^{TM} MP system (Bio-Rad). Densitometries of individual blot signals were quantified using ImageJ software.

Quantitative PCR with reverse transcription analysis

Total RNA was extracted from the mice brain tissues by using TRIzol® reagent (Invitrogen, Thermo Fisher Scientific), and cDNA was synthesized from 1 μg total RNA by using PrimeScript^{TM} RT Master Mix (RR036Q, Takara). Quantitative PCR was carried out using PowerUp^{TM} SYBR^{TM} Green Master Mix (A25778, Applied Biosystems, Thermo Fisher Scientific) on Biosystems QuantStudio 6 Flex Real-time PCR system (Applied Biosystems, Thermo Fisher Scientific). β-Actin (*Actb*) was used as a reference gene. For comparison, mRNA expression levels of knockout mice were normalized to those of wild-type mice for each participant. The primers

were: *Tyw1*-forward: 5′ GTG GGA CTT GTC GCC TTT G; *Tyw1*-reverse: 5′ GGG AAC GAG TGA CCT GCT T; *Stil*-forward: 5′ GAC ACA ATT CAG GAC TGG TAG AC; *Stil*-reverse: 5′ GGC ATG ATC CAC TTT CTG TTC A; *Sass6*-forward: 5′ ATT CCT TTA CGC GGA CTT AGC; *Sass6*-reverse: 5′ AAG TAG GCT GAA GAC GAG GAG; *Ncapd2*-forward: 5′ AGC CAG ACA AGC CTC ATT GAC; *Ncapd2*-reverse: 5′ TCC ATA GGT GAC GGA TGT CCA; *Cenpe*-forward: 5′ CTT CAG TGG CTG TCT GTG TTC; *Cenpe*-reverse: 5′ CCA TCG CTC TGA TAA ATA GCG TT; β-Actin-forward: 5′ GGC TGT ATT CCC CTC CAT CG; β-Actin-reverse: 5′ CCA GTT GGT AAC AAT GCC ATG T.

Culture of primary neurons

E13.5 mouse embryo brains were taken out from a *Tyw1*$^{+/-}$ pregnant dam. Cortex in each brain was dissected separately and collected in Hibernate-E supplemented with 2% B27™ on ice. Single cells were obtained by using 0.05% trypsin (containing 0.2 mM EDTA) digestion for 10 min at 37°C. After filtration with a 70-μm strainer and centrifugation, cells resuspended in Neurobasal™ medium with 2% B27™ Plus Supplement, 0.25% GlutaMAX™ and 25 μM glutamate were placed in poly-D-lysine coated plates. Adherent neurons were prepared for EdU and terminal deoxynucleotidyl transferase dUTP nick end labelling (TUNEL) staining. To prepare the migration assay, isolated cells were initially cultured in ultra-low attachment plates to form neurospheres, which could be digested using 0.05% trypsin without damage of neurites.

Proliferation and apoptosis assays of primary neurons

At 2 days *in vitro*, cells were labelled with 10 μM EdU solution and incubated for 4 h. Cells were treated by using 4% PFA, followed by 0.5% Triton X-100 permeabilization. After being washed twice with 3% BSA, cells were incubated with Click-iT reaction cocktail according to protocol of Click-iT™ EdU Imaging Kits (C10339, Invitrogen, Thermo Fisher Scientific) and stained with DAPI. Dead cells were labelled using the *In Situ* Cell Death Detection Kit, POD (11684817910, Roche).

Migration test of primary neurons

Neurospheres were digested with 0.05% trypsin to obtain single neurons and resuspended in Neurobasal™ medium with 1% B27™ Plus Supplement. Cells were seeded on the upper layer of a cell culture insert with PET track-etched membrane (8 μm pore size, 353097, Corning) at density of 10^5 cells per well. Neurobasal™ medium with 2% B27™ Plus Supplement was added into the bottom of the lower chamber. After 16 h, the culture insert was taken out and the medium was removed carefully. Cells on the upper surface were wiped, while migrating cells in pore and on the lower surface of membrane were fixed, stained with 0.1% crystal violet and observed.

Flow cytometry

Whole-brain tissues from *Gpam*$^{-/-}$ mice on P1 (postnatal Day 1) were collected, including brain tissues of their wild-type littermate. After digestion with 4 mg/ml papain and 0.1 mg/ml DNase in 37°C shaker for 30 min, suspension was diluted using DMEM/F12 medium containing 5% FBS and filtered with 70-μm strainers. Then, suspension was centrifuged and the cells were fixed in 4% PFA. 0.3% Triton X-100 was used for permeabilization. After being blocked with 5% goat serum, cells were incubated with mouse anti-Gfap antibody and rabbit anti-Ki67 antibody or rabbit anti-Caspase-3 antibody at 4°C for 1 h. Cells were washed with DPBS and incubated with Alexa Fluor 488 goat anti-mouse IgG and Alexa Fluor 647 goat anti-rabbit IgG at

4°C for 30 min. After being washed twice, cells were suspended by using DPBS and analysed on the flow cytometer (BD FACSCanto, BD Biosciences).

Measurement of phosphatidic acid

Phosphatidic acid was measured by using the PicoProbe™ Phosphatidic Acid Assay Kit (BioVision). In brief, medulla of wild-type and *Gpam*$^{-/-}$ mice was homogenized in phosphatidic acid assay buffer. Lipid extraction was obtained according to the protocol and solubilized in 5% Triton X-100 solution. The 'Sample background control' and 'Sample' were prepared in parallel. A standard curve was generated by using phosphatidic acid standard solution. The converter mix was only added in sample and standard wells. All wells were incubated at 45°C for 1 h. Reaction mix was added in each well and incubate at 37°C for 30 min. Fluorescence was recorded at excitation/emission = 535/587 nm and phosphatidic acid concentration was expressed as nanomoles phosphatidic acid per mg tissue weight. The value of phosphatidic acid concentration was normalized.

Measurement of phosphatidylcholine

Phosphatidylcholine was measured by using the phosphatidylcholine assay kit (Abcam). For each individual, ∼10 mg brain tissues were washed with cold PBS, resuspended in the assay buffer and homogenized on ice. After 10 min incubation, samples were centrifuged for 5 min at 4°C at 16 000*g*. The supernatant was incubated with the reaction mix including OxiRed Probe supplemented with hydrolysis enzyme for 30 min. The colorimetric reading was measured at an optical density of 570 nm (OD$_{570}$ nm) on a microplate reader.

Measurement of phosphatidylethanolamine

Phosphatidylethanolamine was measured by using the phosphatidylethanolamine assay kit (Abcam) according to the manufacturer's instructions. Briefly, solubilized lipids were extracted from brain tissues with 5% Triton X-100, incubated with converter mix at 45°C for 1 h, and then with reaction mix at 40°C for 3 h. Fluorescence was recorded at excitation/emission 535/587 nm.

Modelling the effect of TYW1 knockout on ribosomal frameshift

The modelling of TYW1-knockout effect on ribosomal frameshift was based on the fact that TYW1 is a critical enzyme involved in the synthesis of wybutosine, so the hypomorphic and null alleles of *TYW1* could reduce and even remove the production of wybutosine accordingly.[40,41] It is known that wybutosine at the 37 position of tRNA^{Phe} near the 3′ base of an anticodon can stabilize the codon–anticodon interaction thus drastically reduce the ribosomal frameshift on a specific UUU codon.[42] Therefore, we hypothesized that the density and distribution of UUU codons along a specific mRNA should determine the extent of influence exerted by wybutosine, and in parallel, the relative abundance of wybutosine was another major effector during this process. Our modelling procedure was as follows:

(i) Download the UCSC RefSeq (refGene) gene models of human, mouse and zebrafish from https://genome.ucsc.edu/cgi-bin/hgTables. Extract the coding domain sequences (CDS) of each protein-coding gene. If multiple isoforms exist, choose the longest ones.

(ii) Label the location of UUU codons, normalize the location values by the length of CDS. By doing this, we harvested a series of normalized location values, i.e. $L(1)$, $L(2)$, ..., for each gene, where $0 < L(n) < 1$, $n = 1, 2, \ldots$.

(iii) We made the following two assumptions: (a) for a UUU codon at any location, there is a fixed chance of ribosomal frameshift designated by Pf, given a fixed level of wybutosine; and (b) for a given UUU codon located at L(n), after frameshifting, the remaining activity of a protein is proportionate to L(n). Thus, the consequence of translating a mRNA holding multiple UUU codons, i.e. the remaining activity of a protein, designated by R, can be estimated as:

$$R = \sum \left(L(n) \times Pf \times Pr(n) \right) + Pr(n), \text{ where } Pr(1) \tag{1}$$
$$= 1 \text{ and } Pr(n) = Pr(n-1) \times (1 - Pf), \ n = 1, 2, \ldots.$$

(iv) For the wild-type (WT) and TYW1-knockout mice (KO), we defined the R-values for wild-type and KO to be R(WT) and R(KO), respectively. The R-value is determined by the density and distribution of codon UUU in a given mRNA, also by Pf (the chance of ribosomal frameshift at a specific location of mRNA, given a fixed level of wybutosine). We set the values of Pf for R(WT) and R(KO) to be 0.12 and 0.35, respectively.[42] Subsequently, we defined the attenuation coefficient to be R(KO)/R(WT). An attenuation coefficient reflected the extent of protein activity reduction with Tyw1-knockout, and the expected values of attenuation coefficient ranged from 0 to 1. The lower the value of attenuation coefficient, the more severe the ribosomal frameshift and protein degradation when TYW1 is knocked out.

(v) After generating a list of attenuation coefficients for each human/mouse/zebrafish protein, we focused on proteins meeting with these criteria: (a) a protein with ortholog in all three species; (b) a protein with attenuation coefficient <0.34 (i.e. 0.12/0.35); (c) a protein with its corresponding gene matching TYW1 in terms of expression profiles in normal brains; and (d) a protein known to be disease-related, with records in HGMD or OMIM.

Data availability

The authors confirm that the data supporting the findings of this study are available within the article and its Supplementary material. This study's samples and data are available to the scientific community on request.

Results

Cerebral palsy cohort recruitment, medical evaluation and genetic screening

A total of 120 cerebral palsy families, including 118 trios and two quartets, were recruited by the GWCMC from August 2014 to June 2017 in a cohort study. This study was approved by the ethics committee of GWCMC, and written informed consents were signed by the participating families. The following inclusion criteria were adopted: (i) a diagnosis of cerebral palsy made by a board of paediatric neurologists, based on early-onset non-progressive motor disability; (ii) aged 2 years and older at the time of symptom ascertainment; and (iii) uneventful pregnancy and delivery (109 cerebral palsy patients were born at term, six at 36 gestational weeks, two at 35 weeks, three at 34 weeks, one at 33 weeks and one at 30 weeks). The following exclusion criteria were implemented: (i) perinatal insults including potential asphyxia with an Apgar score <7, traumatic brain injury,[43] postencephalitis brain lesions[44] and brain tumours; (ii) late-onset hereditary spastic paraplegia or DOPA-responsive dystonia, due to the distinct progressive trajectory and featured fluctuation of manifestation; (iii) maternal infections with a history of fever during pregnancy; or (iv) hyperbilirubinaemia of the newborn with high bilirubin levels as a risk for kernicterus.[45,46] The classification and summary of the cerebral palsy cohort are shown in Fig. 1A and B. As expected, the majority of cohort patients belonged to spastic cerebral palsy, followed by dyskinetic and ataxic cerebral palsy. There were fewer patients with more

severe motor impairments (GMFCS levels IV and V) than those who were mildly disabled in terms of motion (GMFCS levels I–III). There was a significant bias for males in this cerebral palsy cohort, as observed previously in intellectual disability and autism spectrum disorder cohorts.[47] The prevalent comorbidities included intellectual disability and microcephaly. More details of the cerebral palsy patients can be found in Supplementary Table 1 and Supplementary Fig. 1.

As shown in Fig. 1C and Supplementary Table 2, we initiated a comprehensive genetic assessment via multiple platforms, including targeted PCR for fragile-X, mitochondrial genome sequencing (6000×), a high-density cytogenetic microarray, whole-exome sequencing (600×) and WGS (36×). In principle, we defined the disease-related variants by adopting the ACMG-recommended standards, reinforced by an in-house variation database of 2247 ethnic-matched unaffected individuals, and another database (https://db.cngb.org/cmdb/) of >50 000 individuals of East Asian descent.[34,48]

In sum, we identified cerebral palsy-related candidate genes and variants in 54 families (~45% of the 120 cerebral palsy families), a comparable rate with that of other neurodevelopmental disorders.[16,49–52] Interestingly, the unexpected abundance of recessive variants (~1/4) in the candidate genes, was notably higher than that of previous neurodevelopmental disorder investigations in outbred populations.[50,53,54] Whether this was attributed to the specific genetic structure of the ethnic group in the present study or was due to the characteristic aetiology of cerebral palsy *per se*, remains to be solved in future studies.[55]

Of the 75 cerebral palsy-related variants identified in this cohort, germline point substitutions accounted for ~68% of instances, followed by germline indels (~18.7%) and CNVs (~5.3%) (Fig. 1D). Post-zygotic mutations were, for the first time, described in a cerebral palsy cohort, in which ~6.7% of the cerebral palsy-related variants were classified as post-zygotic mutations. This finding met with the expectation derived from other neurodevelopmental disorder studies, and fills a gap in our understanding of cerebral palsy aetiology.[28]

A closer look at the phenotype-genotype correlation in the patients revealed the following. Patients with more severe physical impairments (GMFCS level IV or V) had a higher chance of harbouring cerebral palsy-related variants (P = 0.00613; Fisher's exact test) than that of patients less disabled in movement (GMFCS levels I–III); patients with comorbidity of intellectual disability were more likely to have deleterious variants (P = 0.01065; Fisher's exact test) compared with those without intellectual disability (Fig. 1E).

In this cohort, we found a total of 129 protein-changing *de novo* mutations in the cerebral palsy patients, the rate of which was significantly higher than that of the in-house 2247 ethnic-matched unaffected individuals (P = 0.0083; bootstrap 1000 times). When counting private mutations of the cerebral palsy families as the data input of a burden analysis, we concluded that the 146 private mutations found in the cerebral palsy patients were significantly more than what we expected by counting the rate of private mutations in the in-house 2247 ethnic-matched controls (P = 0.0142; bootstrap 1000 times). Paternal ages were revealed to correlate positively with *de novo* mutation counts of the patients (r = 0.63, P = 0.025; Pearson correlation test). We also detected 27 post-zygotic mutations in 66 patients with cerebral palsy, but no significant correlation with paternal or maternal ages was established. The aforementioned statistical analyses were based on the protein-coding exons and conventional splice sites with >10× non-redundant coverage (31 856 576 bp) in all the cerebral palsy families and the in-house controls.

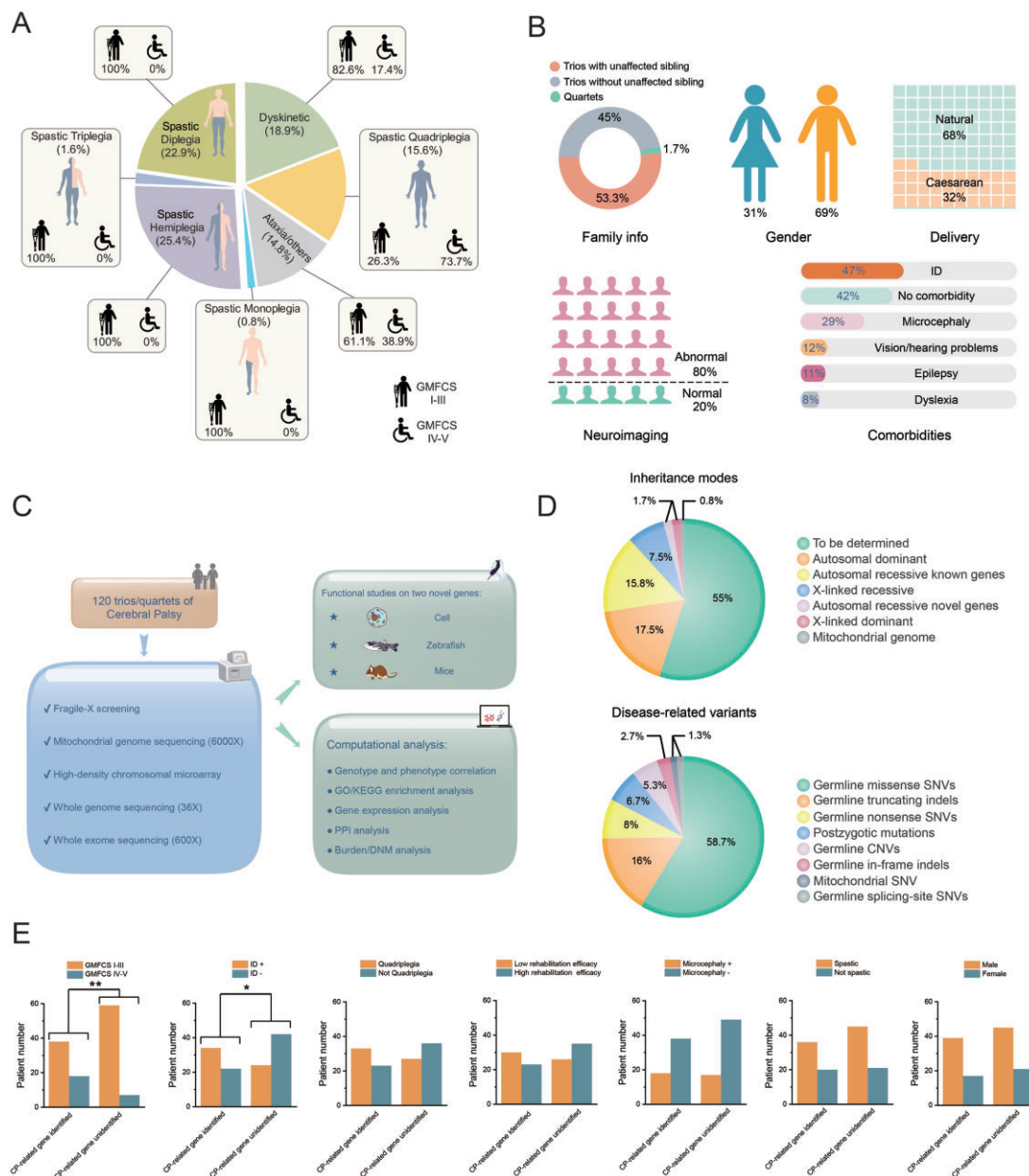The details of disease-related variants identified in this study are shown in Supplementary Tables 3 and 4.

Figure 1 Cerebral palsy cohort study information. (**A**) Classification of the cohort patients based on motor signs, anatomical distributions and motor severity. Spastic quadriplegia (19 patients, GMFCS levels I–III: five patients, GMFCS levels IV–V: 14 patients), dyskinetic (23 patients, GMFCS levels I–III: 19 patients, GMFCS levels IV–V: four patients), spastic diplegia (28 patients, GMFCS levels I–III: 28 patients, GMFCS levels IV–V: zero patients), spastic triplegia (two patients, GMFCS levels I–III: two patients, GMFCS levels IV–V: zero patients), spastic hemiplegia (31 patients, GMFCS levels I–III: 31 patients, GMFCS levels IV–V: zero patients), spastic monoplegia (one patient, GMFCS levels I–III: one patient, GMFCS levels IV–V: zero patients), ataxia and others (18 patients, GMFCS levels I–III: 11 patients, GMFCS levels IV–V: seven patients). (**B**) Summary of the cohort patients in terms of family information (trios with unaffected sibling: 54 patients, trios without unaffected sibling: 64 patients, quartets: two patients), gender (male: 84 patients, female: 38 patients), delivery (natural: 81 patients, caesarean: 39 patients), neuroimaging (abnormal: 98 patients, normal: 24 patients), and comorbidities (no cormorbidity: 51 patients, intellectual disability: 57 patients, microcephaly: 35 patients, vison/hearing problems: 15 patients, epilepsy: 13 patients, dyslexia: 10 patients). (**C**) Project schematics, including cohort recruitment, genetic assessment, gene-oriented bioinformatics and functional analysis. (**D**) Inheritance modes and disease-related variants identified in this study. To be determined: 66, autosomal dominant genes: 21, autosomal recessive known genes: 19, autosomal recessive novel genes: 2, X-linked recessive: 9, X-linked dominant: 2, mitochondrial genome: 1. Germline missense single nucleotide variations (SNVs): 44, germline truncating insertions-deletions (indels): 12, germline nonsense SNVs: 6, postzygotic mutations: 5, germline CNVs: 4, germline in-frame indels: 2, germline splicing-site SNV: 1, mitochondrial SNV: 1. (**E**) Correlations between diagnostic rates and clinical parameters. ID +/− = with/without intellectual disability; Microcephaly +/− = with/without microcephaly.

Assessment of the role of *de novo* mutations

On aggregate, we characterized 26 cerebral palsy-related *de novo* events in ~1/5 of the cohort individuals, which included 18 autosomal dominant mutations (two post-zygotic), two X-linked dominant mutations (one post-zygotic), three *de novo* CNVs spanning multiple disease-related genes and one *de novo* mutation in the mitochondrial genome (Supplementary Figs 2 and 3).

It was noted that there was a post-zygotic mutation observed on a male patient's (Patient CP_102_1) X chromosome while the affected gene, *TAF1* (OMIM: 313650), is known to be related to X-linked recessive intellectual disability. Also, interestingly, one

patient (Patient CP_098_1) had compound heterozygous variants in *ATP8A2* (OMIM: 605870), albeit with one allele of post-zygotic origin.

We found in two spastic diplegia patients (Patients CP_050_1 and CP_095_1) *de novo* mutations of *SPAST* (OMIM: 604277), a gene frequently identified in cerebral palsy patients that is related to early-onset autosomal dominant spastic paraplegia. The hereditary spastic paraplegias comprise a large group of inherited neurological disorders, of which there have been multiple related genes or loci reported.[56,57] When symptoms begin before the age of 2 years, the non-progressive spastic gait (toe walking) of early-onset hereditary spastic paraplegia, such as SPG4 caused by defective *SPAST*, may closely resemble that of spastic diplegic cerebral palsy, thus included in this cohort.

We identified in three cerebral palsy patients (Patients CP_039_1, CP_067_1 and CP_094_1) *de novo* mutations of *KIF1A* (OMIM: 601255), which is known to cause autosomal dominant mental retardation and motor delay. These patients showed variable phenotypes (spastic diplegia and quadriplegia, GMFCS levels I–V). Nevertheless, all of the three patients manifested intellectual disability and severe microcephaly. Another recurrent gene in this cohort was *COL4A1* (OMIM: 120130), in which we identified *de novo* mutations in two cerebral palsy patients (Patients CP_008_1 and CP_033_1). Both patients presented spastic cerebral palsy, epilepsy and severe intellectual disability. In addition, one patient (Patient CP_008_1) showed severe microcephaly and congenital blindness.

Some of the cerebral palsy patients harboured disease-related variants in genes that are functionally relevant. For example, Patient CP_107_1 (dyskinetic cerebral palsy with intellectual disability) had a post-zygotic *de novo* mutation of *BCL11A* (OMIM: 606557), while Patient CP_096_1 [mixed-type cerebral palsy (spasticity and dyskinesia) with intellectual disability] had a *de novo* truncating mutation of *BCL11B* (OMIM: 606558). *BCL11A* and *BCL11B* encode zinc-finger proteins that regulate transcription, highly expressed in brain and bone marrow, and were reported to be related to autosomal dominant developmental disorders. A similar scenario occurred in two cerebral palsy patients (Patients CP_055_1 and CP_061_1), where the former (spastic hemiplegia with severe intellectual disability and microcephaly) had a *de novo* missense mutation of *TUBA1A* (OMIM: 602529), and the latter [spastic hemiplegia with mild motor impairment (GMFCS level I)] had a post-zygotic *de novo* mutation of *TUBB2B* (OMIM: 612850). *TUBA1A* and *TUBB2B* code for brain-enriched tubulin components as subunits of microtubules, and are known to be associated with autosomal dominant brain malformations.

Identification of variants involved in recessive disease

Altogether, we confirmed recessive-disease-related variants in 30 cerebral palsy families (1/4 of the 120 cohort families), including 21 families with autosomal recessive disease, eight families with X-linked recessive disease, and one CNV located at Xq28 of a male patient (Patient CP_119_1) inherited from the unaffected mother, where a variety of CNVs have been reported previously in patients with neuropsychiatric disorders and, in some cases, with gender bias.[58,59]

We noticed that, among the 21 cerebral palsy families with autosomal recessive disease, the zygotic ratio of compound heterozygosity versus homozygosity was 3:1, which was acceptable in a largely outbred country. Interestingly, among the six homozygous variants, only one was found to be autozygous via cytogenetic microarray-based genotyping and family history consulting (Patient CP_036_1, second-cousin consanguinity). This observation may indicate the existence of cerebral palsy-related hotspot

variants in a specific population, although other explanations exist and this hypothesis requires further evidence.

Defects of *RARS2* (OMIM: 611524), a gene coding for mitochondrial arginyl-tRNA synthetase, led to manifestation in two patients (Patients CP_007_1 and CP_110_1). The former patient had severe spastic quadriplegia with intellectual disability, epilepsy, strabismus and amblyopia, probably caused by the highly pathogenic compound heterozygous variants of *RARS2*; whereas the latter patient, with mild dyskinetic cerebral palsy, moderate intellectual disability and hearing problem, was probably associated with a homozygous missense variant that may be less damaging.

Patient CP_105_1 presented ataxic cerebral palsy with mild intellectual disability and microcephaly. A homozygous truncating variant was found in the gene of *SQSTM1* (OMIM: 601530), encoding a ubiquitin-binding protein. According to a recent report, four families with autosomal recessive childhood-onset neurodegenerative disorders were found to have detrimental variants in *SQSTM1*.[60] However, the onset age in our case was much younger (in the first year since birth) compared with that of the previous report (onset between 7 and 15 years of age); therefore, this finding expands the phenotype spectrum related to the gene of *SQSTM1*.

TARS (OMIM: 187790) codes for a threonyl-tRNA synthetase and plays dual roles both as an assembly scaffold of translation initiation components and as a target mRNA selector.[61] A recent report introduced two individuals carrying *TARS* variants with trichothiodystrophy,[62] while we found a spastic hemiplagia patient (Patient CP_014_1) with biallelic TARS variants, who presented an overlapping phenotype including developmental delay, but without a manifestation in the hair or skin (Supplementary Fig. 4).

PTK7 (OMIM: 601890) encodes a tyrosine kinase of the Wnt-signalling pathway, and altered PTK7 activity in mouse models induces perturbation of neural tube development.[63] Recent studies have found an association of PTK7 variants with neural tube defects and foetal anomaly in humans.[64,65] We found that the female Patient CP_052_1 had compound heterozygous deleterious variants of this gene, and the patient showed spastic diplegia with mild motion impairment (GMFCS level I) without other comorbidities reported, although a possibility of spina bifida still remained due to the absence of spinal cord MRI (Supplementary Fig. 5).

Another two novel candidate genes, namely, *TYW1* (OMIM: 611243) and *GPAM* (OMIM: 602395), were identified in two cerebral palsy families with distinct phenotypes (Families CP_012 and CP_063, respectively). The former family had multiple patients, manifesting both cerebral palsy and intellectual disability, while the latter family had a patient with cerebral palsy but without intellectual disability. Because this dichotomous manifestation epitomized the two major subgroups of cerebral palsy (with or without intellectual disability), we carried out detailed functional studies and present the results in the subsequent parts of this paper.

Common features of cerebral palsy-related genes

Combining cerebral palsy-related genes from our present cohort and from previous studies, a total of 114 cerebral palsy-related genes have now been identified so far (Supplementary Table 5). This number is far less than the 700 + genes involved in intellectual disability and related neurodevelopmental disorders.[47] About half of these cerebral palsy-related genes are defined to be intellectual disability-related genes, and this is not unexpected since cerebral palsy is on a diagnostic continuum with intellectual disability, although cerebral palsy and intellectual disability are regarded as distinct disorders (Fig. 2A). The inheritance modes of the 114 cerebral palsy-related genes included autosomal dominant (43.9%), autosomal recessive (41.2%), X-linked recessive (12.3%) and X-linked dominant (2.6%). Compared with those of intellectual
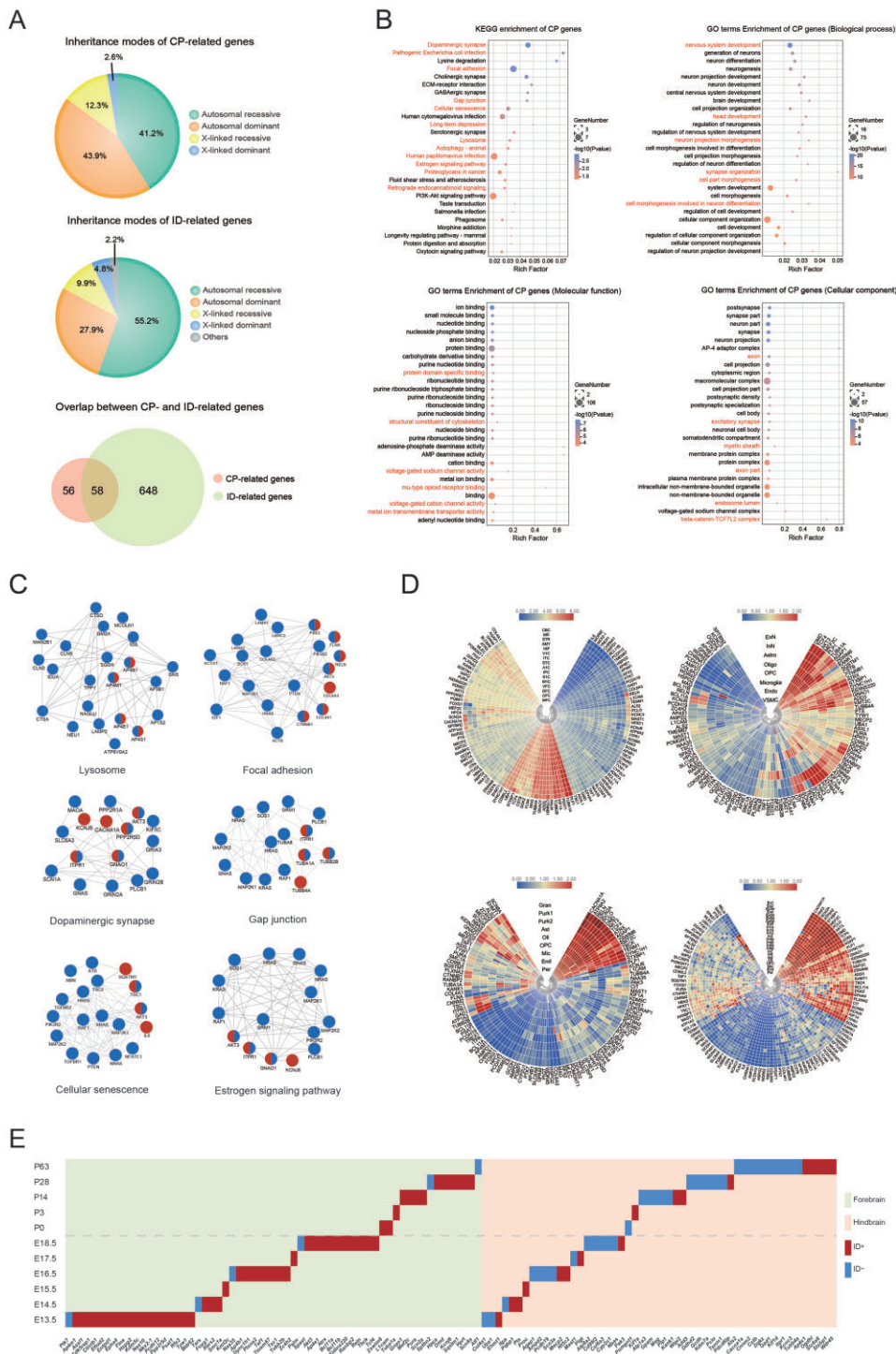
Figure 2 Cerebral palsy-related gene summary. (**A**) Inheritance-mode comparisons between cerebral palsy (CP)-related genes and intellectual disability (ID)-related genes. The data of cerebral palsy-related genes were from this study and a collection of studies from the literature. The data of intellectual disability-related genes were retrieved from the literature.[47] Inheritance modes of 114 cerebral palsy-related genes included 47 autosomal recessive, 50 autosomal dominant, 14 X-linked recessive and three X-linked dominant. Inheritance modes of 706 intellectual disability-related genes included 390 autosomal recessive, 197 autosomal dominant, 70 X-linked recessive, 33 X-linked dominant and 16 others (AD/AR, XLD/XLR). (**B**) Enriched KEGG pathways and GO terms of the cerebral palsy-related genes. *Top left*: Enriched KEGG pathways. *Top right*: Enriched GO terms of Biological Process. *Bottom left*: Enriched GO terms of molecular function. *Bottom right*: Enriched GO terms of Cellular Component. The red colour indicates pathways and GO terms that are also enriched in the intellectual disability-related genes. (**C**) Protein–protein interaction network modules linking cerebral palsy-related genes (red), intellectual disability-related genes (blue) and shared genes (half red, half blue). (**D**) Expression of cerebral palsy-related genes in different human brain regions and cell types. *Top left*: Expression heat map in 16 human brain regions.[127] A1C = primary auditory cortex; AMY = amygdala; CBC = cerebellar cortex; DFC = dorsolateral prefrontal cortex; HIP = hippocampus; IPC = posterior inferior parietal cortex; ITC = inferior temporal cortex; M1C = primary motor cortex; MD = mediodorsal nucleus of thalamus; MFC = medial prefrontal cortex; OFC = orbital prefrontal cortex; S1C = somatosensory cortex; STC = superior temporal cortex; STR = striatum; V1C = primary visual cortex; VFC = ventrolateral prefrontal cortex. *Top right*: Expression heat map in eight cell types of human dorsolateral prefrontal cortex.[127] Astro = astrocyte; Endo = endothelial cell; ExN = excitatory neuron; InN = interneuron; Oligo = oligodendrocyte; OPC = oligodendrocyte progenitor

disability-related genes, cerebral palsy-related genes significantly lack autosomal recessive genes but have more autosomal dominant genes ($P = 0.0006$; Fisher's exact test) (Fig. 2A). We thus speculate that there is a great chance for novel autosomal recessive cerebral palsy-related genes to be identified in future studies.

We found that the 114 cerebral palsy-related genes were enriched in a variety of pathways and protein–protein interaction modules, involving multiple aspects of brain development and functioning (Fig. 2B). We noticed that there were at least six functional modules that reflected the essential and coordinated roles of cerebral palsy-related genes in living organisms (i.e. lysosomes, gap junctions, dopaminergic synapses, focal adhesions, cellular senescence and oestrogen signalling modules) (Fig. 2C). Oestrogen signalling pathways regulate a plethora of physiological processes in mammals, including cellular homeostasis and behaviour. Given the significant gender bias in populations of cerebral palsy, intellectual disability and autism spectrum disorders, it may be worthwhile to explore the involvement of oestrogen signalling pathways in the aetiology of neurodevelopmental disorders. Furthermore, we identified the pathogen-related modules (pathogenic *Escherichia coli* infection, human cytomegalovirus infection, and human papillomavirus infection) enriched in cerebral palsy/intellectual disability/neurodevelopmental disorder genes, comprising ~11% of the 114 cerebral palsy-related genes, which occurred in 4% of cerebral palsy cohort patients, such as *TUBA1A*, *TUBB2B*, *CTNNB1* and *COL4A1*. Although it is well known that cerebral palsy and intellectual disability can be sequelae of intracranial pathogenic infections, whether the variants of these genes increase relevant disease risk remains unclear.

When data-mining in the public databases and literature of gene expression profiles in the brain, we were aware that cerebral palsy-related genes are expressed in a variety of anatomical structures, including the prefrontal cortex, motor cortex, cerebellar cortex, hippocampus, striatum and somatosensory cortex; in addition, these cerebral palsy-related genes are expressed in a variety of brain cells, including excitatory neurons, inhibitory neurons, interneurons, astrocytes, oligodendrocytes, cerebellar granule cells, Purkinje neurons and microglia (Fig. 2D). The top three brain regions for cerebral palsy-related genes were cerebellum, primary motor cortex and striatum; the top three cell types were Purkinje cells, astrocytes and oligodendrocytes. For example, *TYW1* showed enriched expression in neurons within each brain region, but not in pericytes or microglia. *GPAM* showed exclusively high expression in astrocytes, whereas it was untraceable in other brain cells. *PTK7* showed higher expression in pericytes compared with that of other cell types, indicating a role in capillary formation. These specific expression patterns may serve as useful guidelines for further mechanistic studies on these genes.

Cerebral palsy-related genes reached their highest expression levels in different brain regions (forebrain, 57%; hindbrain, 43%) and at sequential developmental stages (embryonic, 58%; postnatal, 42%) (Fig. 2E). Interestingly, we noticed that cerebral palsy-related genes could be roughly divided into two subgroups with distinctive features, namely, forebrain-expressed genes and hindbrain-expressed genes. The former genes, such as *TYW1*, were highly expressed in the forebrain, including the motor cortex and

striatum; furthermore, they had a significantly higher chance to have comorbidity of intellectual disability than that of the latter ones, such as *GPAM*, which were specifically expressed in the hindbrain, including the brainstem and cerebellum ($P = 0.002$; Fisher's exact test). Another observation was that forebrain-expressed cerebral palsy-related genes preferentially had their highest expression levels during embryonic periods, while those of hindbrain-expressed cerebral palsy-related genes occurred throughout embryonic and postnatal periods, but with an inclination for postnatal days ($P = 0.05$; Fisher's exact test). Hence, these results suggest that forebrain-expressed genes may participate in neurogenesis that is generally completed before birth, while hindbrain-expressed genes may take part in myelination that is not complete until 2 years after birth in humans.[66,67] In the next subsection, we illustrate this dichotomous classification system by delving into two typical cerebral palsy-related genes newly identified in this study, namely *TYW1* and *GPAM*.

Null and hypomorphic alleles of *TYW1* cause primary microcephaly and impairment in motion and cognition

The male proband Patient CP_012_1, an ataxic cerebral palsy patient, was the second of two children of a non-consanguineous, healthy Chinese couple with an unremarkable family history. He was born by normal delivery at 40 weeks of gestational age after an uneventful pregnancy with a birth weight of 3.8 kg (+1.17 SD). At conception, the mother was 31 years old and the father was 34 years old. By clinical examination at an age of 6 years, the patient's height, body weight and head circumference were 103 cm (–3.20 SD), 16 kg (–2.06 SD) and 47.5 cm (–3.26 SD), respectively. At an age of 8 years and 7 months, the cranial MRI of the patient showed a slightly widened cerebral subarachnoid space, enlarged ventricle and prominence of cerebellar sulci. Cognitive evaluation (Wechsler Preschool and Primary Scale of Intelligence) was performed at an age of 7 years and 6 months, with a full-scale IQ of 40, indicating a moderate intellectual disability. The level of GMFCS was III, indicating that the patient could walk using a hand-held mobility device. His fine motor skills were poor; the MACS was evaluated at the level of IV (i.e. handles a limited selection of easily managed objects in adapted situations). Patient CP_012_2, the 3-year-older sister of the proband, was diagnosed with mixed-type cerebral palsy (spasticity and dyskinesia). The mother had a normal delivery after 40 weeks of uneventful gestation, and the birth weight of the patient was 3.2 kg (–0.02 SD). The parameters of clinical examination at an age of 14 years of the patient were as follows: height 150 cm (–1.51 SD), weight 41 kg (–1.03 SD) and a head circumference of 49 cm (–3.64 SD). Cognitive evaluation (Estimated Cognitive Level of Children with Cerebral Palsy) was performed at an age of 13 years and 1 month with an IQ of <50, indicating a moderate intellectual disability. The level of the GMFCS was IV, meaning self-mobility with limitations and needing powered mobility. Her fine motor skills were poor; the MACS was at level III (i.e. handles objects with difficulty, needs help to prepare and/or modify activities). The medical details of the two patients are presented in Fig. 3A and Supplementary Tables 1 and 6.

Figure 2 Continued
cell; VSMC = vascular smooth muscle cell. *Bottom left*: Expression heat map in nine cell types of human cerebellum.[128] Ast = astrocyte; End = endothelial cell; Gran = cerebellar granule cell; Mic = microglia; Oli = oligodendrocyte; OPC = oligodendrocyte precursor cell; Per = pericyte; Purk1, 2 = Purkinje neuron, subtype 1, 2. *Bottom right*: Expression heat map in 24 cell types of human frontal cortex.[128] Ast = astrocyte; End = endothelial cell; Ex1, 2, 3e, 4, 5b, 6a, 6b, 8 = excitatory neuron, subtype 1, 2, 3e, 4, 5b, 6a, 6b, 8; In1a, 1b, 1c, 3, 4a, 4b, 6a, 6b, 7, 8 = inhibitory neuron, subtype 1a, 1b, 1c, 3, 4a, 4b, 6a, 6b, 7, 8; Mic = microglia; Oli = oligodendrocyte; OPC = oligodendrocyte precursor cell; Per = pericyte. (**E**) Expression climax of cerebral palsy-related genes in mice, with regard to brain regions (forebrain and hindbrain), developmental stages (11 time points before and after birth), and comorbidity (with or without intellectual disability). *x*-coordinate: cerebral palsy-related genes; *y*-coordinate: embryonic and postnatal days. Gene expression data were retrieved from https://www.ebi.ac.uk/gxa/home.
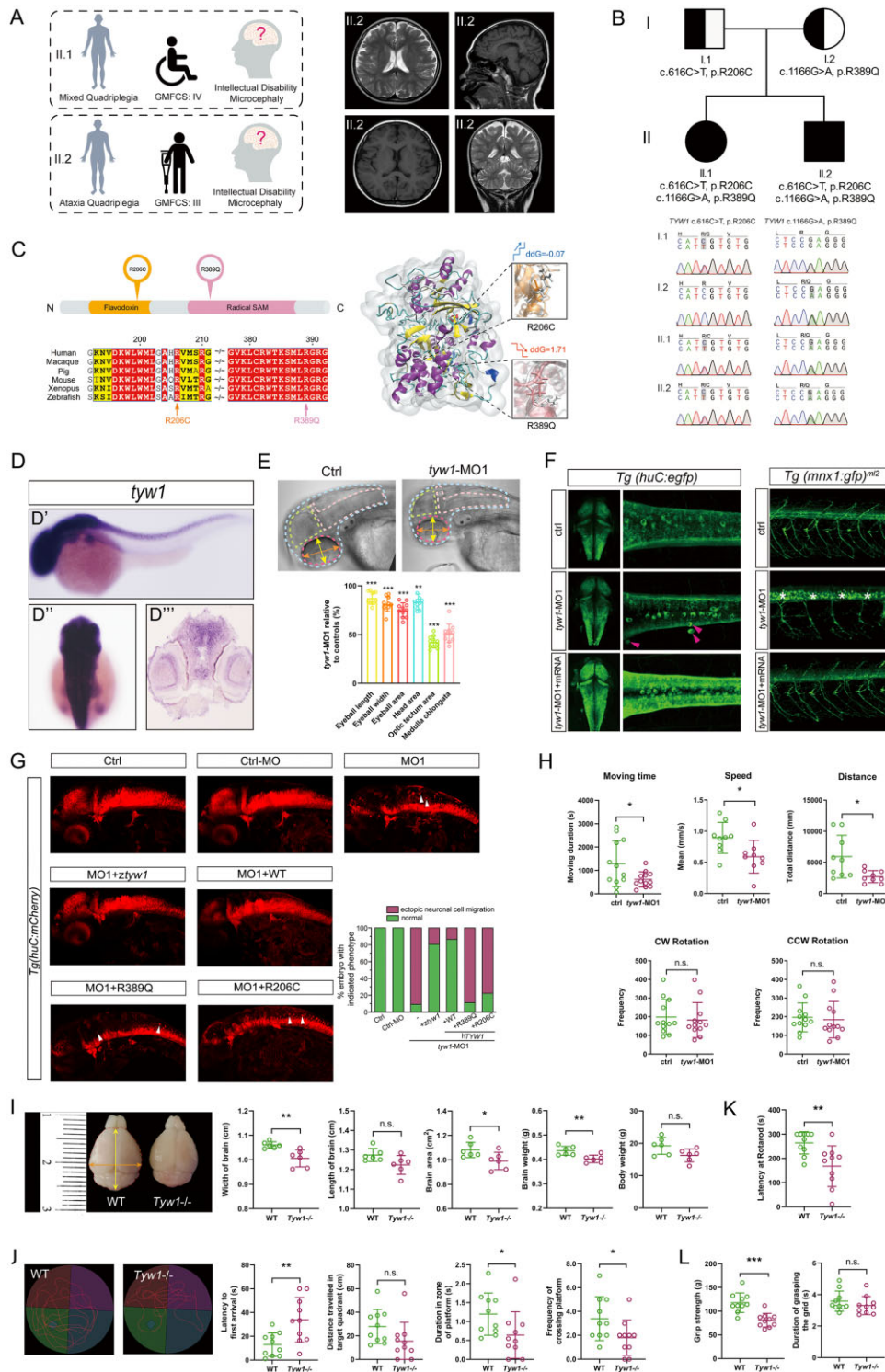
Figure 3 Correlation between the phenotype and genotype of TYW1. (A) Clinical summary of the two index patients harbouring *TYW1* variants, and the cranial MRI of Patient II.2 (CP_012_1). **(B)** Compound heterozygous variants of *TYW1* were identified in the two index patients, which cosegregated with the phenotype in the family members. **(C)** The two identified amino-acid changes (i.e. R206C and R389Q) were located inside the protein domains of flavodoxin and radical SAM, respectively. The two changes took over the highly conserved amnio-acid positions throughout a series of species. A 3D structural model of the TYW1 protein displays the stability alterations (ddG, Gibbs free energy) and local structural changes, before and after the introduction of two amino acid changes. In the two enlarged views, the white bars denote wild-type amino acids and the colour bars denote mutated amino acids. **(D–H)** Morphological and behavioural assays in zebrafish models. **(D)** WISH of zebrafish embryos at 36 hours post-fertilization (hpf) using a zebrafish antisense probe of *tyw1* showing specific expression of *tyw1* in the CNS, including the brain and spinal cord. **D'**: lateral view. **D''**: dorsal view. **D'''**: cross section. **(E)** Head area, optic tectum area, medulla oblongata area, eyeball area, eyeball length and eyeball width of the *tyw1* knockdown (*tyw1*-MO1, a splice-blocking morpholino) zebrafish larvae at 5 days post-fertilization (dpf) were significantly decreased compared with those of the controls. *n* = 10 for each test group. **P < 0.01, ***P < 0.001, unpaired *t*-test. **(F)** Confocal imaging of *egfp*-positive cells in transgenic zebrafish *Tg(huC:egfp)* embryos at 48 hpf showing ectopic neuronal cells (indicated by red arrowheads) in the *tyw1*-MO1 zebrafish group, which could

In each of the two patients, we identified compound heterozygous variants in *TYW1* (OMIM: 611243), which were cosegregating with the phenotype in the family and were confirmed by targeted PCR and Sanger sequencing (Fig. 3B). The pRecessive value of *TYW1* (http://exac.broadinstitute.org/) was 0.9945, which was a strong indicator of a recessive-disease-related gene. The two variants (NM_018264:c.616C>T, p.R206C; c.1166G>A, p.R389Q) were absent from an in-house database with 2247 healthy controls, and were very rare in the public databases (http://exac.broadinstitute.org/; https://db.cngb.org/cmdb/) with allele frequencies below 10^{-5}. Pathogenicity prediction classified the two variants as damaging or disease-causing (Supplementary Table 4). Multiple protein-sequence alignments among different species showed high conservation of both positions, i.e. R206 and R389, which were located in the functional domains (Fig. 3C). Protein structural modelling revealed the functional locations of both amino acids, and showed reduced protein stability (R389Q) and disturbed substrate binding (R206C) (Fig. 3C). Compatible with the prediction, we found a dramatically lower protein level of TYW1 in the patient's peripheral blood sample, as compared with those of the parents and healthy children of the same age and gender (Supplementary Fig. 6).

Using WISH in zebrafish, we found that *tyw1* expression was highly enriched in the developing CNS, including the brain and spinal cord (Fig. 3D). Morpholino-mediated *tyw1* knockdown (*tyw1*-MO) zebrafish models were established by blocking mRNA splicing (MO1) and protein translation (MO2) of *tyw1* (Supplementary Figs 7–9). We observed significant head size reduction in the *tyw1*-MO zebrafish compared with that of wild-type zebrafish (Fig. 3E). It was revealed that *tyw1* deficiency in zebrafish resulted in ectopic neuronal cell migration in the brain and undifferentiated motor neuronal cells in the spinal cord (Fig. 3F and Supplementary Fig. 9). This phenotypic anomaly was effectively rescued by zebrafish *tyw1* mRNA or human *TYW1* mRNA, but not by *TYW1* mRNAs with variants identified in the cerebral palsy patients (Fig. 3G). This result substantiated the pathogenicity of the two variants identified in these patients. By performing swimming behaviour tests, we noticed that the swimming capacity of *tyw1*-MO zebrafish larvae was compromised significantly, specifically in terms of swimming speed, time and distance, but not in rotation (Fig. 3H).

By using CRISPR–Cas9 genome editing technology, we constructed a *Tyw1*-knockout mouse model (Supplementary Fig. 10). In the *Tyw1*$^{-/-}$ mice, the brain size and weight were significantly below the level of those of wild-type littermates, with abnormal intracranial morphology (Fig. 3I and Supplementary Fig. 11). All behavioural tests (i.e. the Morris water maze, rotarod test, and grip strength test) revealed significantly reduced performances of the *Tyw1*$^{-/-}$ group compared with those of the wild-type controls (Fig. 3J–L). The deterioration of both motion and cognition was compatible with the manifestation of the cerebral palsy patients with hypomorphic *TYW1* alleles, and reflected the underlying cerebral regions responsible for both functions.

Defective *Tyw1* hinders neuronal proliferation and migration due to an increased ribosomal frameshift in a subset of cell-cycling-related proteins

The gene of *TYW1* (OMIM: 611243) encodes a tRNA-wybutosine (tRNA-yW) synthesizing protein, localized on the cytosolic surface of the endoplasmic reticulum and which is responsible for producing a hypermodified guanosine (i.e. wybutosine) at position 37 of phenylalanine tRNA (tRNA^{Phe}) adjacent to the 3′ of the anticodon in archaea and eukaryotes.[40,41] This hypermodification at position 37 of tRNA^{Phe} is known to play a critical role in stabilizing the appropriate interactions between codons and anticodons during protein translation, an absence of which promotes a ribosomal frameshift, thus reducing translational accuracy.[68–71] Interestingly, although two codons exist for phenylalanine, i.e. UUU and UUC, it appears that wybutosine exerts an influence only on the UUU codon.[42]

The expression of TYW1 is enriched in neurons, especially during the embryonic development (Figs 2D, E and 4A and B). We observed significantly decreased neurogenesis in the E13.5 brains of *Tyw1*$^{-/-}$ mice compared with those of wild-type mice (Fig. 4C). In *Tyw1*$^{-/-}$ mouse brains, the migration of cortical neuronal precursors between E15.5 and E18.5 was hindered, resulting in an obvious reduction of superficial-layer neurons and a noticeable accumulation of deep-layer neurons (Fig. 4D). Consequently, in both the frontal cortex and motor cortex of adult mouse brains, the thickness of superficial layers (layers II–III) was significantly less in the *Tyw1*$^{-/-}$ group compared with that of wild-type, while there was a non-significant reverse trend in the deep layer (layer VI) (Fig. 4E). The simultaneous abnormalities in both frontal and motor cortices in this mouse model provide mechanistic insights into the cerebral palsy-plus-intellectual disability phenotype of patients with hypomorphic *TYW1* alleles in the present study.

Primary neurons extracted from E13.5 mouse brain cortices showed significantly decreased proliferation in the *Tyw1*$^{-/-}$ group compared with that in the wild-type group, while the apoptotic intensity showed no obvious increase in the *Tyw1*$^{-/-}$ group (Fig. 4F). Migration tests of primary neurons revealed a reduced migratory

Figure 3 Continued

be rescued by the wild-type *tyw1* mRNA. Confocal imaging of GFP-positive cells in the transgenic zebrafish *Tg(mnx1:gfp)*^{ml2} embryos at 72 hpf showed undifferentiated motor neuronal cells (indicated by white stars) in the *tyw1*-MO zebrafish group, which could be rescued by wild-type *tyw1* mRNA. (**G**) The *tyw1* knockdown (*tyw1*-MO1) brought about ectopic neuronal cells in the zebrafish *Tg(huC:mCherry)* embryos at 48 hpf, and there were only 9% of participants that remained normal. This abnormal phenotype could be rescued by zebrafish *tyw1* mRNA (*ztyw1*), human *TYW1* mRNA (wild-type), mutated human *TYW1* leading to R389Q or mutated human *TYW1* leading to R206C, to variable extents (81%, 86%, 11% and 22%, respectively). Ectopic neuronal cells are indicated by white arrowheads. The number of fish used ranged from 63 to 88 in the different test groups. (**H**) Swimming behaviour tests of *tyw1*-MO1 zebrafish larvae at 5 dpf compared with the controls, in terms of moving time, speed, distance, clockwise (CW) rotation and counterclockwise (CCW) rotation. The former three parameters revealed significant reductions in the *tyw1*-MO1 group compared with those in the control group, while the latter two parameters showed no significant change. n = 9 to 12 for each test group. *P < 0.05, n.s. = no significant difference, unpaired *t*-test. (**I–L**) Morphological and behavioural assays on *Tyw1* knockout (*Tyw1*$^{-/-}$) mouse models. (**I**) The *Tyw1*$^{-/-}$ mice at postnatal 8 weeks showed significantly reduced brain size and weight compared with those of wild-type (WT) mice. The orange arrow, yellow arrow and red-dashed lines indicate the width, length and area of brain, respectively. In parallel, measurements of body weight revealed no significant difference between the *Tyw1*$^{-/-}$ mice and the wild-type mice. n = 6 per genotype with equal numbers of male and female mice. *P < 0.05, **P < 0.01, n.s. = no significant difference, unpaired *t*-test. (**J**) Track plots depicting the traces of mice during the probe test in the Morris water maze. The *Tyw1*$^{-/-}$ mice showed worse performance compared with that of wild-type mice in the test. n = 10 per genotype with equal numbers of male and female mice. *P < 0.05, **P < 0.01, n.s. = no significant difference, unpaired *t*-test. (**K**) Rotarod tests showed significantly reduced motor coordination and balance of the *Tyw1*$^{-/-}$ mice compared with those of wild-type mice. n = 10 per genotype with equal numbers of male and female mice. **P < 0.01, unpaired *t*-test. (**L**) Grip strength tests revealed significantly weaker grip strength of the *Tyw1*$^{-/-}$ mice compared with that of the wild-type mice. n = 10 per genotype with equal numbers of male and female mice. ***P < 0.001, n.s. = no significant difference, unpaired *t*-test.
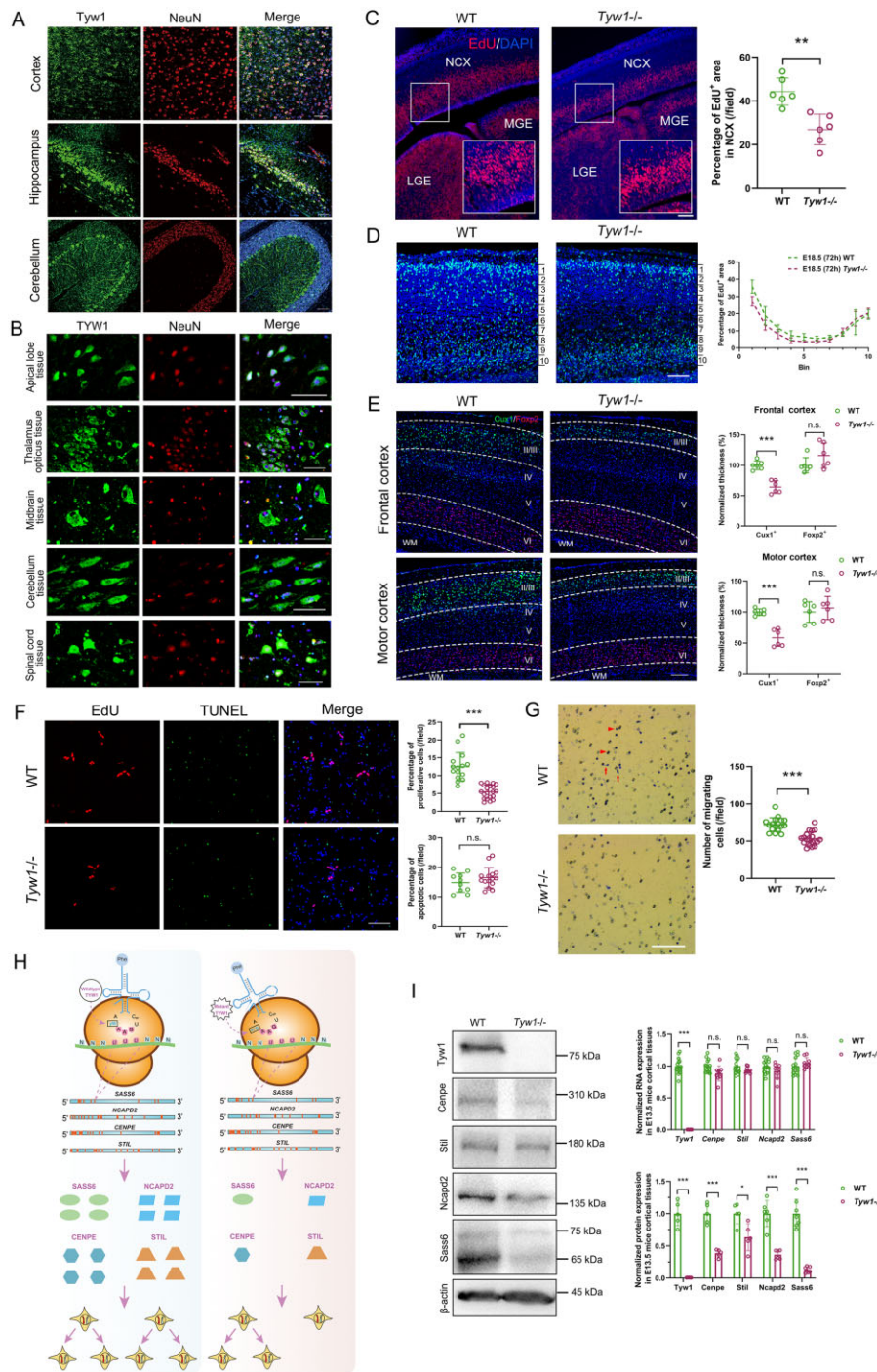
Figure 4 Functional mechanism of defective TYW1. (**A**) Brain section immunostaining of wild-type mice at postnatal 8 weeks, where TYW1 (green) and NeuN (red) are co-labelled, showing TYW1 expression in neurons within the cortex, hippocampus and cerebellum. Blue = DAPI. Scale bar = 75 μm. (**B**) Immunostaining of normal human brain tissue microarrays, where TYW1 (green) and NeuN (red) are co-labelled, revealing TYW1 expression in neurons within a variety of brain regions. Blue = DAPI. Scale bar = 50 μm. (**C**) EdU was intraperitoneally injected at 2 h into pregnant mice before euthanizing them to harvest embryos at E13.5. Cells incorporated with EdU were determined in the embryonic brain sections under fluorescent microscopy and were analysed with ImageJ software. Compared with those of wild-type littermates, significantly reduced percentages of EdU$^+$ areas were observed in the neocortex (NCX) of $Tyw1^{-/-}$ embryonic brains, as well as in the medial ganglionic eminence (MGE) and lateral ganglionic eminence (LGE). Red = EdU; blue = DAPI. $n = 3$ per genotype. For a specific sample, each measurement was performed twice in separate fields. $^{**}P < 0.01$, unpaired t-test. Scale bar = 100 μm. (**D**) EdU was injected into the pregnant mice carrying embryos at E15.5, which were observed after 72 h. The embryonic cortex was divided into 10 equally spaced bins along the vertical axis. The percentage of EdU$^+$ area in each bin was measured. In $Tyw1^{-/-}$ mouse cortices, there were fewer EdU$^+$ cells in superficial layers (layers II–III) than those in wild-type mice, while the number of EdU$^+$ cells in the deep layer (layer VI) was greater than that in wild-type mice. Light blue = EdU; blue = DAPI. $n = 3$ per genotype. Scale bar = 100 μm. (**E**) Brain-section immunostaining of mice at postnatal 8 weeks, where Cux1$^+$ and Foxp2$^+$ cells were labelled for neurons in superficial layers (layers II–III) and the deep layer (layer VI), respectively. Compared with those of wild-type mice, $Tyw1^{-/-}$ mouse brains showed a significantly reduced thickness of superficial layers at both frontal and motor cortices, while in the deep layers, the reverse phenomenon occurred, albeit without statistical significance. WM = white matter. Green = Cux1; red = Foxp2; Blue = DAPI. $n = 3$ per genotype. For a specific sample, each measurement was performed twice

capacity of the $Tyw1^{-/-}$ group compared with that of the wild-type group (Fig. 4G). In parallel, we evaluated the performance of SH-SY5Y cells after introducing a TYW1 knockout (Supplementary Fig. 12). Interestingly, as observed in the primary cortical neurons from mice, TYW1-knockout SH-SY5Y cells presented significantly reduced abilities of proliferation, adhesion and migration, compared with those of wild-type cells. However, there was no observable change in neuron differentiation (Supplementary Fig. 13).

To explore the mechanism of epigenetic regulation by TYW1, we constructed a model to simulate how the levels of wybutosine affected the ribosomal frameshift (see the 'Materials and methods' section). We defined a parameter, namely, the attenuation coefficient, for each protein to depict the vulnerability of a specific protein-coding gene to the reduced level of wybutosine. Typically, if the mRNA sequence of a specific gene does not contain a UUU codon, this gene will not be affected by the ribosomal frameshift related to the wybutosine level. On the contrary, if the mRNA sequence of a gene holds many UUU codons, its protein translation will be sensitive to the wybutosine level, with the extent determined by the density and distribution of UUU codons. This algorithm thus generated a protein list with attenuation coefficients lower than expectation, due to a malfunctioning TYW1. Subsequently, we filtered the list by two criteria (Supplementary Table 7). First, since we observed microcephaly and neurological manifestation in humans, mice and zebrafish, we chose proteins with low attenuation coefficients in all three species. Second, we focused on proteins that are known to be disease-related and that bear resemblance to TYW1 in terms of gene expression patterns, especially in developing brains (Supplementary Fig. 14). In this way, four candidate proteins emerged, namely CENPE (OMIM: 117143), STIL (OMIM: 181590), NCAPD2 (OMIM: 615638) and SASS6 (OMIM: 609321) (Fig. 4H). Measurements in E13.5 mouse brain samples showed significantly reduced levels of these four proteins in the $Tyw1^{-/-}$ group compared with those in the wild-type group, while the corresponding RNA levels did not change significantly (Fig. 4I). Therefore, it is reasonable that the defective $Tyw1$ hindered the production of these four proteins in the brain, with an ensuing reduction of neuronal proliferation and migration. Interestingly, these four proteins are all involved in cell cycling and mitosis, the defects of which have been reported to be related to primary microcephaly. STIL and SASS6 are necessary for centriole duplication during the cell cycle.[72–77] CENPE is required for spindle microtubule capture and attachment at the kinetochore during cell division.[78–80] NCAPD2 is a component of the condensin multiprotein complex that participates in mitotic chromosomal condensation.[81,82]

Null and hypomorphic alleles of GPAM induce drastically reduced motor ability with unaffected cognition

The female proband Patient CP_063_1, a spastic quadriplegic cerebral palsy patient, was the only child of unrelated healthy parents without a relevant family history. She was born by caesarean section at 39 gestational weeks when her mother was 30 years old and father was 32 years old. Her birth weight was 3.05 kg (–0.44 SD). The clinical examination at an age of 2 years and 4 months showed a body height of 83 cm (–2.45 SD), weight of 10 kg (–1.7 SD) and head circumference of 47 cm (–1.31 SD). The cranial neuroimaging taken at 4 years old revealed no apparent anatomical abnormality except for plagiocephaly, which was due to the elongated period of remaining in a supine position after birth. However, there was a reduction of white matter fibre tracts, especially in the corticospinal tract. The level of GMFCS was IV, meaning 'self-mobility with limitations and may use powered mobility'. Her fine motor skills were poor, as the MACS was evaluated at a level of III (i.e. handles objects with difficulty, need help to prepare and/or modify activities). The cognitive evaluation (Wechsler Preschool and Primary Scale of Intelligence) was performed at an age of three years and seven months, revealing a full-scale IQ of 116. This patient epitomized a major subgroup, comprising half of the cerebral palsy population, who were impaired in motion albeit normal in intelligence. The medical information of this patient is presented in Fig. 5A and Supplementary Tables 1 and 6.

Compound heterozygous variants of GPAM (OMIM: 602395) were identified in this patient, the association between which and the cerebral palsy phenotype in this family was confirmed by using targeted PCR and Sanger sequencing (Supplementary Table 3 and Fig. 5B). The pRecessive value 0.9998 of GPAM (http://exac. broadinstitute.org/) rendered a very strong indicator for a recessive-disease-related gene. Allele frequencies of the two variants (NM_001244949:c.1495G>C, p.G499R; c.2005C>T, p.P669S) were both below 10^{-5}, in the large-scale databases (https://db.cngb.org/cmdb/ and http://exac.broadinstitute.org/), and with no match in the in-house database of >2247 healthy individuals. Pathogenicity prediction labelled the two variants damaging or disease-causing (Supplementary Table 4). Multiple protein-sequence alignments in a variety of species showed high conservation of both positions (i.e. G499 and P669), which were located inside the important protein domains. Three-dimensional structural models revealed that G499 and P669 were positioned at the pivotal regions, and both alterations of amino acids led to a dramatic reduction of protein structural stability (Fig. 5C).

Figure 4 Continued

in separate fields. ***$P < 0.001$, n.s. = no significant difference, unpaired t-test. Scale bar = 200 μm. (**F** and **G**) Proliferation, apoptosis and migration analysis of primary neurons from the cortices of E13.5 mice. (**F**) Quantification of the proliferative and apoptotic primary neurons from E13.5 mouse brain cortices. The primary neurons from the $Tyw1^{-/-}$ mice showed significantly decreased proliferation (as indicated by EdU staining), while the intensity of apoptosis did not change significantly (as determined by TUNEL staining). ***$P < 0.001$, n.s. = no significant difference, unpaired t-test. Scale bar = 100 μm. (**G**) Quantification of migrating primary neurons from the E13.5 mouse brain cortex. In the transwell assay, the primary neurons from the $Tyw1^{-/-}$ mice showed significantly decreased migration. Red arrows indicate cells that already migrated through pores, while arrowheads indicate cells that were migrating in pores. ***$P < 0.001$, n.s. = no significant difference, unpaired t-test. Scale bar = 250 μm. (**H**) Schematics of the mechanism of ribosomal frameshift under the conditions of wild-type and mutant alleles of TYW1. GAA is the anticodon of tRNA^{Phe} for phenylalanine, and UUU is the codon of mRNA for phenylalanine. In the four mRNAs, with normalized lengths, of SASS6, NCAPD2, CENPE and STIL, the positions of UUU codons are indicated by red bars. The introduction of mutated TYW1 led to promotion of a ribosomal frameshift at the interaction between tRNA^{Phe} and mRNAs, resulting in the reduced production of a subset of proteins involved in cell cycling. (**I**) Measurements of mRNA and protein levels in the E13.5 mouse brain cortex. In the $Tyw1^{-/-}$ group, the mRNA levels of Cenpe, Stil, Ncapd2 and Sass6 did not change significantly compared with those in the wild-type group. However, the protein levels of CENPE, STIL, NCAPD2 and SASS6 showed significant reductions compared with those in the wild-type group, with remaining levels of 0.38, 0.63, 0.36 and 0.14, respectively. β-Actin was used as the internal control for both RNA and protein measurements. Neither RNA nor protein of TYW1 could be detected in the $Tyw1^{-/-}$ group. Molecular weights: TYW1 = 84 kDa; CENPE = 312 kDa; STIL = 143 kDa; NCAPD2 = 157 kDa; SASS6 = 74 kDa and 65 kDa; β-actin = 43 kDa. Quantitative PCR with reverse transcription and western blotting data were from three independent experiments. *$P < 0.05$, ***$P < 0.001$, unpaired t-test.
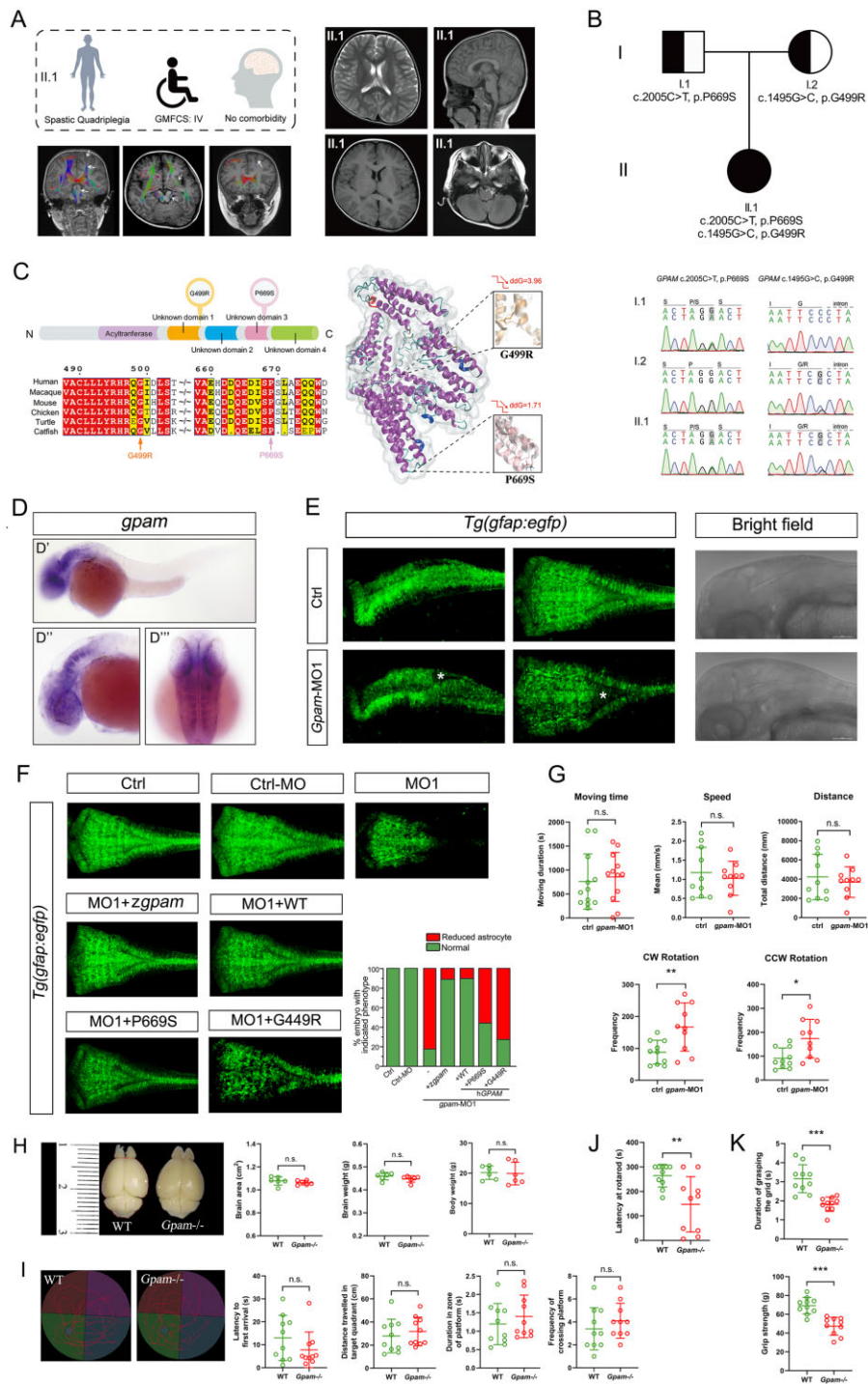
Figure 5 Correlation between the phenotype and genotype of GPAM. (**A**) Brief medical description of the index Patient II.1 (CP_063_1) harbouring *GPAM* variants, and the results of MRI and diffusion tensor imaging (DTI). No gross anatomical abnormality was observed except for plagiocephaly. In comparison with features of the DTI scan of a normal brain, there were relatively fewer white matter fibre bundles observed in the patient brain, especially the left hemisphere. The most obvious change was at the left corticospinal tract, left front limb of the internal capsule and left white-matter junction bundle. Meanwhile, the fractional anisotropy values (including forward/backward, left/right and up/down directions) were lower on the left side than on the right side. These DTI findings were consistent with the clinical manifestations of the patient, such as high muscle tone of the limbs, hyperextension of the knee, scissoring gait and mirror movement of the hands. Furthermore, the right limb function of the patient was worse than that of the left. Red represents fibres from left to right; green represents fibres from front to back; blue represents fibres from top to bottom. The three images represent the distributions of fibres in different layers of the child with cerebral palsy. Deterioration of the corticospinal tract is indicated by white arrows. (**B**) Compound heterozygous variants of *GPAM* were identified in the index patient, which cosegregated with the phenotype in the family members. (**C**) The two identified amino acid changes (i.e. G499R and P669S) were both located inside the conserved protein domains. The two changes took over the amino acid positions that were highly conserved among different species. A 3D structural model of the GPAM protein displayed the reduced protein stability (ddG, Gibbs free energy), as well local structural alterations before and after the introduction of two amino acid changes. In the two enlarged views, white bars denote wild-type amino acids and coloured bars denote mutated amino acids. (**D–G**) Morphological and behavioural assays in zebrafish models. (**D**) WISH of zebrafish embryos at 36 hpf by using a zebrafish antisense *gpam* probe showing specific expression

(continued)

Using WISH analysis in zebrafish, we found that *gpam* was specifically expressed in the developing CNS, mainly in the hindbrain (Fig. 5D). Morpholino-mediated *gpam* knockdown (*gpam*-MO) zebrafish models were established by blocking mRNA splicing (MO1) and protein translation (MO2) of *gpam* (Supplementary Figs 7 and 15). There was no apparent head size difference between *gpam*-MO and wild-type zebrafish, but we observed an obvious reduction of astrocytes in *gpam*-MO zebrafish, especially in the hindbrain (Fig. 5E and Supplementary Fig. 16). This reduction could be effectively rescued by the wild-type human *GPAM* or zebrafish *gpam* mRNA, but not by the *GPAM* mRNAs with variants causing G499R and P669S (Fig. 5F). This result corroborated the pathogenicity of the two variants identified in the patient. Furthermore, swimming behaviour tests showed that the swimming capacity of *gpam*-MO zebrafish was compromised significantly (Fig. 5G), mimicking the motor impairment of the respective cerebral palsy patient. Interestingly, we noticed that the swimming capacity patterns of *tyw1*-MO and *gpam*-MO zebrafish were complementary to each other, such that the swimming distance, speed and moving time were abnormal in *tyw1*-MO zebrafish but normal in *gpam*-MO zebrafish; in contrast, the rotational movement was normal in *tyw1*-MO zebrafish but abnormal in *gpam*-MO zebrafish. This finding may reflect differential underlying mechanisms of these two genes, although more experiments are needed to confirm or refute this hypothesis.

Meanwhile, we constructed *Gpam* knockout mice models via CRISPR–Cas9 genome editing technology (Supplementary Fig. 10). In *Gpam* knockout (*Gpam*$^{-/-}$) mice, there was no significant alteration of brain size or weight compared with that of wild-type littermates (Fig. 5H), compatible with our observations in humans and zebrafish. Similarly, results from the Morris water maze showed no significant differences between *Gpam*$^{-/-}$ mice and wild-type mice (Fig. 5I). However, the rotarod test and the grip strength test revealed significant reductions in motor capacities of *Gpam*$^{-/-}$ mice compared with those of wild-type mice (Fig. 5J and K).

Defective *Gpam* interferes with astrocytic proliferation and oligodendrocytic myelination by perturbing lipid metabolism

The gene of GPAM encodes a mitochondrial enzyme that catalyses the production of lysophosphatidic acid using glycerol-3-phosphate (G3P) as a substrate, which is the rate-limiting step in the synthesis of glycerolphospholipids and triacylglycerol.[83] There are four isoenzymes in this step, among which GPAM occupies most of the activity in the liver, and hepatic knockdown of *Gpam* in mice reduces levels of triacylglycerol, diacylglycerol, fatty acids and cholesterol.[83–85] In the CNS, GPAM was specifically expressed in astrocytes, and reached its highest expression level on P14 in mouse hindbrain (Figs 2D, E and 6A and B). Enzymatic activity in astrocytes is maintained mainly by GPAM, supplemented by GPAT4, an isoenzyme ubiquitously expressed in a variety of brain cells including neurons and oligodendrocytes.[86,87]

In *Gpam*$^{-/-}$ mice, we observed a reduced number of astrocytes, especially in the medulla regions (Fig. 6C). Whole-brain single-cell suspension of P1 mice showed significantly reduced proliferation of astrocytes in *Gpam*$^{-/-}$ mice, but no apparent increase in the intensity of apoptosis was observed (Fig. 6D). Interestingly, brain-section staining revealed an obviously decreased white-matter thickness of the motor cortex in *Gpam*$^{-/-}$ mice (Fig. 6E). In parallel with the decrease of astrocytic density, a myelin marker, namely myelin binding protein (MBP), was revealed to be significantly reduced in the medulla regions of *Gpam*$^{-/-}$ mouse brains (Fig. 6F and G). Given the phenotypic similarities between the respective cerebral palsy patient and the *Gpam*$^{-/-}$ mice, we hypothesize that the defective *Gpam* may lead to hypomyelination of the corticospinal tract that originates from the primary motor cortex, passes through and modulates in the medulla and terminates in the spinal cord.[88] However, this interpretation requires further testing in future studies.

Myelin is a specialized form of oligodendrocytic membrane that tightly wraps and insulates neuronal axons to accelerate conduction of action potentials.[66,89] Compared with the properties of other membranes, myelin has a much higher lipid content that consists of up to 80% of the total dry weight, and the most abundant lipid groups in myelin are cholesterol, phospholipids and glycolipids.[89–91] It has been reported that in mouse models, lipid synthesis blockage in oligodendrocyte causes instant onset of severe demyelination and neurological symptoms; however, this damage can be spontaneously remedied to a great extent after a period of months.[92] Because the blood–brain barrier shields the CNS from circulating lipids, an exogenous lipid supply in the parenchyma of brain may account for this phenomenon; fortuitously, recent studies have discovered that the major suppliers of lipids in the brain are astrocytes.[93–95] Actually, it has been revealed that horizontal lipid flux from astrocytes to oligodendrocytes is a major feature of myelination, and that disruption of lipid synthesis in astrocytes leads to lasting and more severe hypomyelination than inactivation of that in oligodendrocytes.[96–98] In astrocytes, GPAM is committed to a rate-limiting step from G3P to lysophosphatidic acid, the subsequent lipid products of which include phosphatidic acid, diacylglycerol, triacylglycerol, phosphatidylinositol, phosphatidylcholine, phosphatidylethanolamine and phosphatidylserine. Phosphatidylinositol, phosphatidylethanolamine, phosphatidylcholine

Figure 5 Continued
of *gpam* in the brain, especially in the hindbrain. **D'**: lateral view. **D''**: lateral view with magnification of head region. **D'''**: dorsal view. (**E**) Confocal imaging of transgenic zebrafish *Tg(gfap:egfp)* embryos at 48 hpf showing a reduced number of gfap$^+$ cells in the *gpam* knockdown group (*gpam*-MO1, a splicing-blocking morpholino), particularly in the hindbrain, labelled by white stars. Brightfield lateral imaging showing no apparent general change in the brain. *Left*: lateral view; *middle*: dorsal view; *ight*: lateral view. (**F**) The *gpam* knockdown (*gpam*-MO1) caused a reduction of astrocytes in *Tg(gfap:egfp)* embryos at 48 hpf, and there were only 17% of participants that remained normal. This abnormal phenotype could be rescued by zebrafish *gpam* mRNA (z*gpam*), human *GPAM* mRNA (wild-type, WT), mutated human *GPAM* leading to P669S, or mutated human *GPAM* leading to G449R, to variable extents (86%, 87%, 44% and 28%, respectively). The number of fish used ranged from 36 to 72 in the different test groups. (**G**) Swimming behaviour tests of the *gpam*-MO1 zebrafish larvae at 5 dpf compared with the controls, in terms of moving time, speed, distance, clockwise (CW) rotation and counterclockwise (CCW) rotation. The former three parameters showed no significant changes, while the latter two parameters revealed significant changes between the two groups. $n = 10$ to 12 for each test group. $^*P < 0.05$, $^{**}P < 0.01$, n.s. = no significant difference, unpaired *t*-test. (**H–K**) Morphological and behavioural assays of *Gpam* knockout (*Gpam*$^{-/-}$) mouse models. (**H**) The *Gpam*$^{-/-}$ mice at postnatal 8 weeks showed no significant alterations of brain size, brain weight or body weight compared with those of wild-type mice. The red-dashed line indicates the area of brain. $n = 6$ per genotype with equal number of males and females. n.s. = no significant difference, unpaired *t*-test. (**I**) Track plots depicting the traces of mice during the probe test in the Morris water maze. The *Gpam*$^{-/-}$ mice showed no difference in performance compared with that of wild-type mice. $n = 10$ per genotype with equal numbers of male and female mice. n.s. = no significant difference, unpaired *t*-test. (**J**) Rotarod tests showed significantly reduced motor coordination and balance of the *Gpam*$^{-/-}$ mice compared with those of wild-type mice. $n = 10$ per genotype with equal numbers of male and female mice. $^{**}P < 0.01$, unpaired *t*-test. (**K**) Grip strength tests revealed significantly shorter grip duration and weaker grip strength of the *Gpam*$^{-/-}$ mice compared with those of wild-type mice. $n = 10$ per genotype with equal numbers of male and female mice. $^{***}P < 0.001$, unpaired *t*-test.
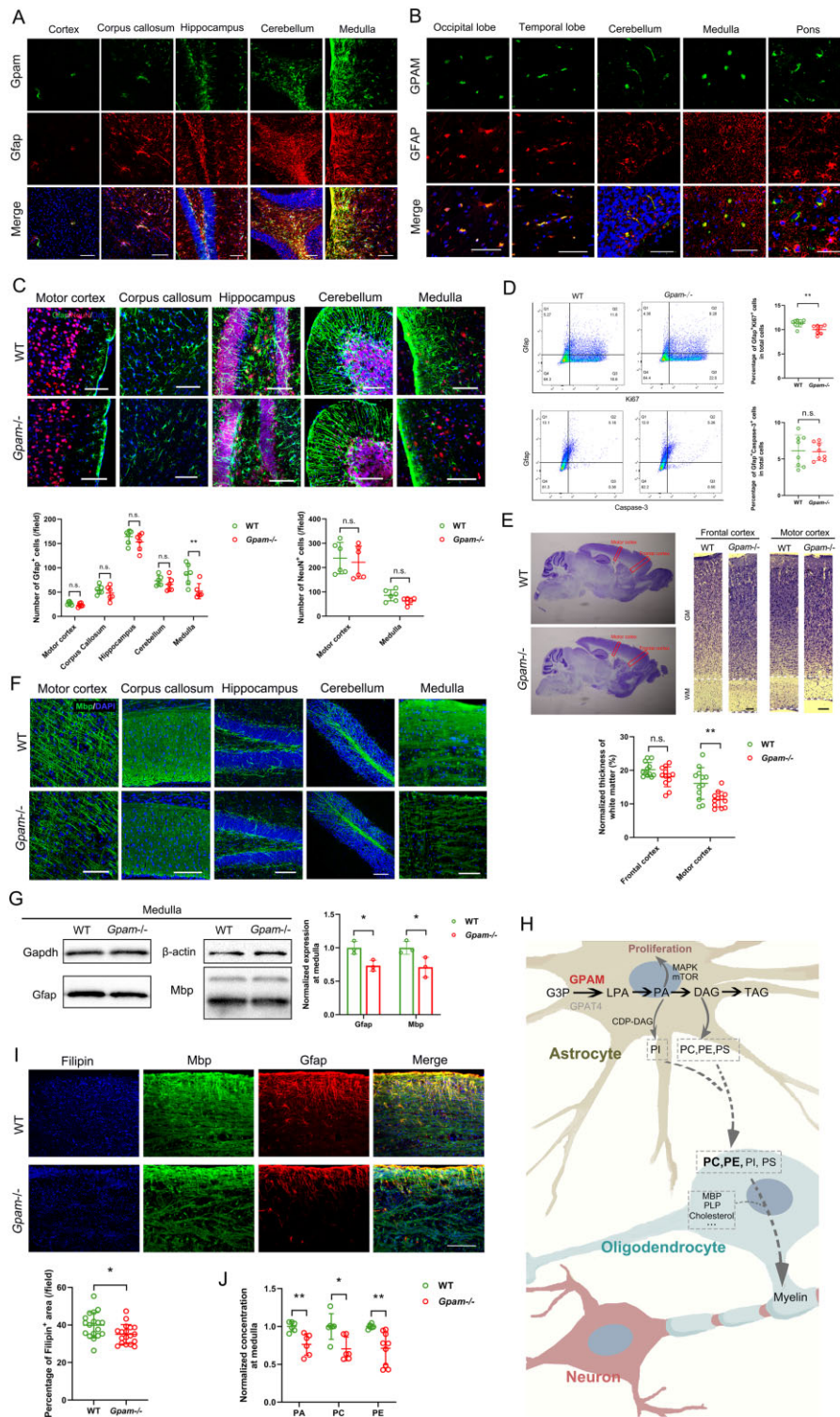
Figure 6 Functional mechanism of defective GPAM. (**A**) Brain-section immunostaining of wild-type mice at postnatal 8 weeks, where GPAM (green) and GFAP (red) are co-labelled, showing GPAM expression in astrocytes within many brain regions. Blue = DAPI. Scale bar = 100 μm. (**B**) Immunostaining of the normal human brain tissue microarrays, where GPAM (green) and GFAP (red) are co-labelled, revealing GPAM expression in astrocytes within a variety of brain regions. Blue = DAPI. Scale bar = 50 μm. (**C**) Brain-section immunostaining of the *Gpam*⁻/⁻ mice at postnatal 8 weeks revealing a significantly reduced number of GFAP⁺ cells in the medulla, as compared with that in wild-type (WT) mice. The same tendency was observed in other brain regions albeit without statistical significance. Green = GFAP; red = NeuN; blue = DAPI. Scale bar = 100 μm. *n* = 3 per genotype. For a specific sample, each measurement was performed twice in separate fields. **P < 0.01, n.s. = no significant difference, unpaired *t*-test. (**D**) Flow cytometry of whole-brain cell suspension from P1 mice stained with GFAP, Ki67 and Caspase-3. Percentages of the GFAP⁺Ki67⁺ cells were significantly lower in the *Gpam*⁻/⁻ mice brain compared with those in wild-type mice, indicating inhibited proliferation of astrocytes. Percentages of GFAP⁺Caspase-3⁺ cells showed no difference between the two groups, indicating no activated apoptosis. *n* = 8 per genotype with equal numbers of male and female mice. **P < 0.01, n.s. = no significant difference, unpaired *t*-test. (**E**) Nissl staining of sagittal brain sections of mice at postnatal 8 weeks showing a significantly thinner white matter in the motor cortex of the *Gpam*⁻/⁻ group, as compared with that of the wild-type group. The

(continued)

and phosphatidylserine are the major components of myelin[99,100] (Fig. 6H). Noticeably, a large body of evidence has implicated an association between *GPAM* polymorphisms and plasma levels of cholesterol.[101–106] In our present study, experiments in brain sections and extracts confirmed this hypothesis, with significantly reduced contents of lipids detected in medullary tissues of *Gpam*[−/−] mice (Fig. 6I and J). Interestingly, we also observed a significant reduction of 1,2-diacyl-G3P (phosphatidic acid), which is a downstream product of GPAM-involved reactions (Fig. 6H and J). Phosphatidic acid is not only a precursor of phosphatidylinositol, a component of myelin, but also promotes cell proliferation in astrocytes by activating the MAPK and mTOR signalling pathways (Fig. 6H).[107–109] In addition, inhibition of phosphatidic acid is known to reduce the number of astrocytes and to cause neurodevelopmental disorders.[110–113]

Discussion

This work is a large-scale, multi-dimensional, clinical and genetic investigation on the aetiology of cerebral palsy. We identified, in 45% of 120 cerebral palsy families, detrimental variants of established and novel genes related to cerebral palsy and other neurodevelopmental disorders. Importantly, future studies should address whether our adjustment of clinical exclusions in the selection of this cohort may have affected the diagnostic yield. The meta-analysis on the 114 cerebral palsy-related genes found in both the present study and previous studies created an updated compendium spanning a variety of attributes of cerebral palsy-related genes, which may be a useful reference for subsequent research. Intriguingly, we proposed a novel classification system for cerebral palsy-related genes according to their dichotomous spatiotemporal expression patterns. This may serve as a useful guideline for exploring the underlying mechanisms of cerebral palsy-related genes and for designing targeted treatments. In future studies, identification of more cerebral palsy-related genes from larger cohorts is needed to substantiate this classification system.

TYW1 is a novel cerebral palsy-related gene, defects of which were confirmed in the present study to result in primary microcephaly and deterioration of both motion and cognition by disturbing proliferation and migration of neurons in the developing brain, including the motor cortex and frontal cortex. Blockage of neuronal proliferation and/or migration during brain development is a common theme in the aetiology of primary microcephaly.[114,115] Interestingly, most of the known MCPH (i.e. microcephaly primary hereditary) genes encode components of machineries of cell division and cell cycling, while our present results of the mechanisms by which TYW1 influences a subset of MCPH genes is in line with previous findings. TYW1 plays a critical role in generating

wybutosine, a hypermodification at position 37 of tRNAPhe responsible for stabilizing codon–anticodon binding, and this epigenetic procedure can be grouped into one of the three aspects of tRNA activation, namely, biogenesis, charging and modification of tRNAs.[70,71] We noticed that a large proportion of genes involved in tRNA charging have been found to be associated with disease, especially neurodevelopmental disorders; however, merely ∼10% of genes participating in tRNA modifications are related to human illness.[68] Because of the enduring existence of most tRNA modification throughout species evolution, we posit that tRNA modifications may be underestimated in their roles in disease onset and progression. Future studies in this field can be modelled after this work on TYW1, and since most tRNA modifications were identified at position 34 and 37 in the anticodon stem loop, the algorithm that we developed for wybutosin at position 37 may be useful for application in a multitude of scenarios.

GPAM encodes a mitochondrial enzyme specifically expressed in astrocytes and plays a crucial role in lipid metabolism in the brain. Aberrant GPAM found in the present study reduced a variety of lipid contents in astrocytes and oligodendrocytes, causing hypomyelination in the corticospinal tract and suppression of proliferation of astrocytes, the latter of which is known to deteriorate the interplay between oligodendrocytes and astrocytes thus aggravating the hypomyelination.[98,116] Dysfunction of lipid metabolism is present in a variety of neurodevelopmental disorders, including intellectual disability, autism spectrum disorders, attention deficit hyperactivity disorder and schizophrenia.[117–121] Additionally, genetic defects in astrocytes, such as *GFAP* and *NPC1*, are known to be related to lesions in brain white matter.[122–125] Our present work helped to elucidate the complex relationship between astrocytes and myelination, which may guide the development of novel therapeutic strategies. Intriguingly, there have been reports in mouse models that a high-fat diet promotes circulating lipid contents and ameliorates hypomyelination in the brain if lipid synthesis in astrocytes is blocked; therefore, it is tantalizing to apply this high-fat diet intervention to *Gpam*[−/−] mice and human patients with similar lipid metabolic problems.[97,126]

Funding

Figure 6 Continued
same tendency was observed in the frontal cortex, albeit without statistical significance. Scale bar = 200 μm. *n* = 3 per genotype. For a specific sample, each measurement was performed four times in separate fields. **$P < 0.01$, n.s. = no significant difference, unpaired *t*-test. (**F**) Brain-section immunostaining of mice at postnatal 8 weeks showing a lower density of myelin binding protein (MBP) in the medulla of the *Gpam*[−/−] mice, compared with that of wild-type mice. Green = MBP; blue = DAPI. Scale bar = 100 μm. (**G**) Western blot of the mouse brain medulla showing significant reductions of GFAP and Mbp in the *Gpam*[−/−] group compared with those in the wild-type group. GAPDH and β-actin were used as controls. Molecular weights: GFAP = 50 kDa; Mbp = 19 and 26 kDa. For a specific sample, each measurement was performed three times independently. *$P < 0.05$, unpaired *t*-test. (**H**) Schematics of the mechanism of GPAM involvement in lipid metabolism and its influence on astrocytic proliferation and oligodendrocytic myelination, by regulating the production of phosphatidic acid (PA), phosphatidylinositol (PI), phosphatidylcholine (PC), phosphatidylethanolamine (PE) and phosphatidylserine (PS). CDP-DAG = cytidine diphosphate diacylglycerol; DAG = diacylglycerol; G3P = glycerol-3-phosphate; LPA = lysophosphatidic acid; PLP = proteolipid protein; TAG = triacylglycerol. (**I**) Brain-medulla section immunostaining showing significant reductions of Filipin, as well as MBP and GFAP, in *Gpam*[−/−] mice compared with those in wild-type mice. The percentage of the Filipin[+] area per imaging field was proportional to the cholesterol abundance. Scale bar = 100 μm. *n* = 3 per genotype. For a specific sample, each measurement was performed four times in separate fields. *$P < 0.05$, unpaired *t*-test. (**J**) Quantification of PA, PC and PE levels in the brain medullas of mice at postnatal 8 weeks revealing a significant reduction in the *Gpam*[−/−] group compared with those in the wild-type group. *n* = 3 per genotype. For a specific sample, each measurement was performed twice or thrice. *$P < 0.05$, **$P < 0.01$, unpaired *t*-test.

Competing interests

The authors report no competing interests.

Supplementary material

Supplementary material is available at *Brain* online.

References

1. Colver A, Fairhurst C, Pharoah PO. Cerebral palsy. *Lancet*. 2014; 383(9924):1240–1249.
2. Moreno-De-Luca A, Ledbetter DH, Martin CL. Genetic [corrected] insights into the causes and classification of [corrected] cerebral palsies. *Lancet Neurol*. 2012;11(3):283–292.
3. Novak I, Morgan C, Adde L, et al. Early, accurate diagnosis and early intervention in cerebral palsy: Advances in diagnosis and treatment. *JAMA Pediatr*. 2017;171(9):897–907.
4. Rosenbaum PL, Palisano RJ, Bartlett DJ, Galuppi BE, Russell DJ. Development of the gross motor function classification system for cerebral palsy. *Dev Med Child Neurol*. 2008;50(4):249–253.
5. Korzeniewski SJ, Slaughter J, Lenski M, Haak P, Paneth N. The complex aetiology of cerebral palsy. *Nat Rev Neurol*. 2018;14(9): 528–543.
6. Graham HK, Rosenbaum P, Paneth N, et al. Cerebral palsy. *Nat Rev Dis Primers*. 2016;2:15082.
7. MacLennan AH, Thompson SC, Gecz J. Cerebral palsy: Causes, pathways, and the role of genetic variants. *Am J Obstet Gynecol*. 2015;213(6):779–788.
8. Petterson B, Stanley F, Henderson D. Cerebral palsy in multiple births in Western Australia: Genetic aspects. *Am J Med Genet*. 1990;37(3):346–351.
9. Fletcher NA, Foley J. Parental age, genetic mutation, and cerebral palsy. *J Med Genet*. 1993;30(1):44–46.
10. Costeff H. Estimated frequency of genetic and nongenetic causes of congenital idiopathic cerebral palsy in west Sweden. *Ann Hum Genet*. 2004;68(Pt 5):515–520.
11. Hemminki K, Li X, Sundquist K, Sundquist J. High familial risks for cerebral palsy implicate partial heritable aetiology. *Paediatr Perinat Epidemiol*. 2007;21(3):235–241.
12. Garne E, Dolk H, Krageloh-Mann I, Holst Ravn S, Cans C, Group SC. Cerebral palsy and congenital malformations. *Eur J Paediatr Neurol*. 2008;12(2):82–88.
13. O'Callaghan ME, MacLennan AH, Haan EA, Dekker G. South Australian Cerebral Palsy Research G. The genomic basis of cerebral palsy: A HuGE systematic literature review. *Hum Genet*. 2009;126(1):149–172.
14. Fahey MC, Maclennan AH, Kretzschmar D, Gecz J, Kruer MC. The genetic basis of cerebral palsy. *Dev Med Child Neurol*. 2017; 59(5):462–469.
15. Turner TN, Coe BP, Dickel DE, et al. Genomic patterns of de novo mutation in simplex autism. *Cell*. 2017;171(3):710–722.e12.
16. Hu H, Kahrizi K, Musante L, et al. Genetics of intellectual disability in consanguineous families. *Mol Psychiatry*. 2018;24(7):1027–1039.
17. Satterstrom FK, Kosmicki JA, Wang J, et al. Large-scale exome sequencing study implicates both developmental and functional changes in the neurobiology of autism. *Cell*. 2020;180(3): 568–584.e23.
18. Segel R, Ben-Pazi H, Zeligson S, et al. Copy number variations in cryptogenic cerebral palsy. *Neurology*. 2015;84(16):1660–1668.
19. Oskoui M, Gazzellone MJ, Thiruvahindrapuram B, et al. Clinically relevant copy number variations detected in cerebral palsy. *Nat Commun*. 2015;6:7949.
20. Zarrei M, Fehlings DL, Mawjee K, et al. De novo and rare inherited copy-number variations in the hemiplegic form of cerebral palsy. *Genet Med*. 2018;20(2):172–180.
21. Corbett MA, van Eyk CL, Webber DL, et al. Pathogenic copy number variants that affect gene expression contribute to genomic burden in cerebral palsy. *NPJ Genom Med*. 2018;3:33.
22. McMichael G, Bainbridge MN, Haan E, et al. Whole-exome sequencing points to considerable genetic heterogeneity of cerebral palsy. *Mol Psychiatry*. 2015;20(2):176–182.
23. Parolin Schnekenberg R, Perkins EM, Miller JW, et al. De novo point mutations in patients diagnosed with ataxic cerebral palsy. *Brain*. 2015;138(Pt 7):1817–1832.
24. Matthews AM, Blydt-Hansen I, Al-Jabri B, et al.; TIDE BC, United for Metabolic Diseases and the CAUSES Study. Atypical cerebral palsy: Genomics analysis enables precision medicine. *Genet Med*. 2018;21(7):1621–1628.
25. Takezawa Y, Kikuchi A, Haginoya K, et al. Genomic analysis identifies masqueraders of full-term cerebral palsy. *Ann Clin Transl Neurol*. 2018;5(5):538–551.
26. van Eyk CL, Corbett MA, Frank MSB, et al. Targeted resequencing identifies genes with recurrent variation in cerebral palsy. *NPJ Genom Med*. 2019;4:27.
27. Jin SC, Lewis SA, Bakhtiari S, et al. Mutations disrupting neuritogenesis genes confer risk for cerebral palsy. *Nat Genet*. 2020; 52(10):1046–1056.
28. D'Gama AM, Walsh CA. Somatic mosaicism and neurodevelopmental disease. *Nat Neurosci*. 2018;21(11):1504–1514.
29. Campbell IM, Shaw CA, Stankiewicz P, Lupski JR. Somatic mosaicism: Implications for disease and transmission genetics. *Trends Genet*. 2015;31(7):382–392.
30. Palisano R, Rosenbaum P, Walter S, Russell D, Wood E, Galuppi B. Development and reliability of a system to classify gross motor function in children with cerebral palsy. *Dev Med Child Neurol*. 1997;39(4):214–223.
31. Eliasson AC, Krumlinde-Sundholm L, Rosblad B, et al. The Manual Ability Classification System (MACS) for children with cerebral palsy: Scale development and evidence of validity and reliability. *Dev Med Child Neurol*. 2006;48(7):549–554.
32. Xu K, Wang L, Mai J, He L. Efficacy of constraint-induced movement therapy and electrical stimulation on hand function of children with hemiplegic cerebral palsy: A controlled clinical trial. *Disabil Rehabil*. 2012;34(4):337–346.
33. Russell DJ, Rosenbaum PL, Cadman DT, Gowland C, Hardy S, Jarvis S. The gross motor function measure: A means to evaluate the effects of physical therapy. *Dev Med Child Neurol*. 1989; 31(3):341–352.
34. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17(5):405–424.
35. Wang K, Li M, Hakonarson H. ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010;38(16):e164.
36. Purcell S, Neale B, Todd-Brown K, et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81(3):559–575.
37. Xu M, Liu D, Dong Z, et al. Kinesin-12 influences axonal growth during zebrafish neural development. *Cytoskeleton (Hoboken)*. 2014;71(10):555–563.
38. Gong J, Wang X, Zhu C, et al. Insm1a regulates motor neuron development in Zebrafish. *Front Mol Neurosci*. 2017;10:274.

39. Gong J, Chai L, Xu G, Ni Y, Liu D. The expression of natriuretic peptide receptors in developing zebrafish embryos. *Gene Expr Patterns*. 2018;29:65–71.

40. Young AP, Bandarian V. TYW1: A radical SAM enzyme involved in the biosynthesis of wybutosine bases. *Methods Enzymol*. 2018;606:119–153.

41. Noma A, Kirino Y, Ikeuchi Y, Suzuki T. Biosynthesis of wybutosine, a hyper-modified nucleoside in eukaryotic phenylalanine tRNA. *EMBO J*. 2006;25(10):2142–2154.

42. Carlson BA, Kwon SY, Chamorro M, Oroszlan S, Hatfield DL, Lee BJ. Transfer RNA modification status influences retroviral ribosomal frameshifting. *Virology*. 1999;255(1):2–8.

43. Gelineau-Morel RN, Zinkus TP, Le Pichon JB. Pediatric head trauma: A review and update. *Pediatr Rev*. 2019;40(9):468–481.

44. Fraley CE, Pettersson DR, Nolt D. Encephalitis in previously healthy children. *Pediatr Rev*. 2021;42(2):68–77.

45. Lauer BJ, Spector ND. Hyperbilirubinemia in the newborn. *Pediatr Rev*. 2011;32(8):341–349.

46. Wu YW, Kuzniewicz MW, Wickremasinghe AC, et al. Risk for cerebral palsy in infants with total serum bilirubin levels at or above the exchange transfusion threshold: A population-based study. *JAMA Pediatr*. 2015;169(3):239–246.

47. Vissers LE, Gilissen C, Veltman JA. Genetic studies in intellectual disability and related disorders. *Nat Rev Genet*. 2016;17(1):9–18.

48. South ST, Lee C, Lamb AN, et al. ACMG Standards and Guidelines for constitutional cytogenomic microarray analysis, including postnatal and prenatal applications: Revision 2013. *Genet Med*. 2013;15(11):901–909.

49. Najmabadi H, Hu H, Garshasbi M, et al. Deep sequencing reveals 50 novel genes for recessive cognitive disorders. *Nature*. 2011;478(7367):57–63.

50. Doan RN, Lim ET, De Rubeis S, et al. Recessive gene disruptions in autism spectrum disorder. *Nat Genet*. 2019;51(7):1092–1098.

51. Ruzzo EK, Perez-Cano L, Jung JY, et al. Inherited and de novo genetic risk for autism impacts shared networks. *Cell*. 2019;178(4):850–866.e26.

52. Gilissen C, Hehir-Kwa JY, Thung DT, et al. Genome sequencing identifies major causes of severe intellectual disability. *Nature*. 2014;511(7509):344–347.

53. Vissers LE, de Ligt J, Gilissen C, et al. A de novo paradigm for mental retardation. *Nat Genet*. 2010;42(12):1109–1112.

54. Martin HC, Jones WD, McIntyre R, et al. Quantifying the contribution of recessive coding variation to developmental disorders. *Science*. 2018;362(6419):1161–1164.

55. Kahrizi K, Hu H, Hosseini M, et al. Effect of inbreeding on intellectual disability revisited by trio sequencing. *Clin Genet*. 2019;95(1):151–159.

56. Kara E, Tucci A, Manzoni C, et al. Genetic and phenotypic characterization of complex hereditary spastic paraplegia. *Brain*. 2016;139(Pt 7):1904–1918.

57. Pearson TS, Pons R, Ghaoui R, Sue CM. Genetic mimics of cerebral palsy. *Mov Disord*. 2019;34(5):625–636.

58. El-Hattab AW, Schaaf CP, Fang P, et al. Clinical characterization of int22h1/int22h2-mediated Xq28 duplication/deletion: New cases and literature review. *BMC Med Genet*. 2015;16:12.

59. Yon DK, Park JE, Kim SJ, Shim SH, Chae KY. A sibship with duplication of Xq28 inherited from the mother; genomic characterization and clinical outcomes. *BMC Med Genet*. 2017;18(1):30.

60. Haack TB, Ignatius E, Calvo-Garrido J, et al. Absence of the autophagy adaptor SQSTM1/p62 causes childhood-onset neurodegeneration with ataxia, dystonia, and gaze palsy. *Am J Hum Genet*. 2016;99(3):735–743.

61. Jeong SJ, Park S, Nguyen LT, et al. A threonyl-tRNA synthetase-mediated translation initiation machinery. *Nat Commun*. 2019;10(1):1357.

62. Theil AF, Botta E, Raams A, et al. Bi-allelic TARS mutations are associated with brittle hair phenotype. *Am J Hum Genet*. Aug 1 2019;105(2):434–440.

63. Williams M, Yen W, Lu X, Sutherland A. Distinct apical and basolateral mechanisms drive planar cell polarity-dependent convergent extension of the mouse neural plate. *Dev Cell*. 2014;29(1):34–46.

64. Lei Y, Kim SE, Chen Z, et al. Variants identified in PTK7 associated with neural tube defects. *Mol Genet Genomic Med*. 2019;7(4):e00584.

65. Meier N, Bruder E, Lapaire O, et al. Exome sequencing of fetal anomaly syndromes: Novel phenotype-genotype discoveries. *Eur J Hum Genet*. 2019;27(5):730–737.

66. Nave KA, Werner HB. Myelination of the nervous system: Mechanisms and functions. *Annu Rev Cell Dev Biol*. 2014;30:503–533.

67. Bercury KK, Macklin WB. Dynamics and mechanisms of CNS myelination. *Dev Cell*. Feb 23 2015;32(4):447–458.

68. El Yacoubi B, Bailly M, de Crecy-Lagard V. Biosynthesis and function of posttranscriptional modifications of transfer RNAs. *Annu Rev Genet*. 2012;46:69–95.

69. Landgraf BJ, McCarthy EL, Booker SJ. Radical S-adenosylmethionine enzymes in human health and disease. *Annu Rev Biochem*. 2016;85:485–514.

70. Rak R, Dahan O, Pilpel Y. Repertoires of tRNAs: The couplers of genomics and proteomics. *Annu Rev Cell Dev Biol*. 2018;34:239–264.

71. Schaffer AE, Pinkard O, Coller JM. tRNA metabolism and neurodevelopmental disorders. *Annu Rev Genomics Hum Genet*. 2019;20:359–387.

72. Izraeli S, Lowe LA, Bertness VL, et al. The SIL gene is required for mouse embryonic axial development and left-right specification. *Nature*. 1999;399(6737):691–694.

73. Kumar A, Girimaji SC, Duvvari MR, Blanton SH. Mutations in STIL, encoding a pericentriolar and centrosomal protein, cause primary microcephaly. *Am J Hum Genet*. 2009;84(2):286–290.

74. Vulprecht J, David A, Tibelius A, et al. STIL is required for centriole duplication in human cells. *J Cell Sci*. 2012;125(Pt 5):1353–1362.

75. Leidel S, Delattre M, Cerutti L, Baumer K, Gonczy P. SAS-6 defines a protein family required for centrosome duplication in *C. elegans* and in human cells. *Nat Cell Biol*. 2005;7(2):115–125.

76. Strnad P, Leidel S, Vinogradova T, Euteneuer U, Khodjakov A, Gonczy P. Regulated HsSAS-6 levels ensure formation of a single procentriole per centriole during the centrosome duplication cycle. *Dev Cell*. 2007;13(2):203–213.

77. van Breugel M, Hirono M, Andreeva A, et al. Structures of SAS-6 suggest its organization in centrioles. *Science*. 2011;331(6021):1196–1199.

78. Putkey FR, Cramer T, Morphew MK, et al. Unstable kinetochore-microtubule capture and chromosomal instability following deletion of CENP-E. *Dev Cell*. 2002;3(3):351–365.

79. Yao X, Abrieu A, Zheng Y, Sullivan KF, Cleveland DW. CENP-E forms a link between attachment of spindle microtubules to kinetochores and the mitotic checkpoint. *Nat Cell Biol*. 2000;2(8):484–491.

80. Yen TJ, Li G, Schaar BT, Szilak I, Cleveland DW. CENP-E is a putative kinetochore motor that accumulates just before mitosis. *Nature*. 1992;359(6395):536–539.

81. Schmiesing JA, Gregson HC, Zhou S, Yokomori K. A human condensin complex containing hCAP-C-hCAP-E and CNAP1, a homolog of Xenopus XCAP-D2, colocalizes with phosphorylated histone H3 during the early stage of mitotic chromosome condensation. *Mol Cell Biol*. 2000;20(18):6996–7006.

82. Martin CA, Murray JE, Carroll P, et al. Mutations in genes encoding condensin complex proteins cause microcephaly through decatenation failure at mitosis. *Genes Dev*. 2016;30(19):2158–2172.

83. Karasawa K, Tanigawa K, Harada A, Yamashita A. Transcriptional regulation of acyl-CoA:glycerol-sn-3-phosphate acyltransferases. *Int J Mol Sci*. 2019;20(4):964.

84. Xu H, Wilcox D, Nguyen P, et al. Hepatic knockdown of mitochondrial GPAT1 in ob/ob mice improves metabolic profile. *Biochem Biophys Res Commun*. 2006;349(1):439–448.

85. Hammond LE, Gallagher PA, Wang S, et al. Mitochondrial glycerol-3-phosphate acyltransferase-deficient mice have reduced weight and liver triacylglycerol content and altered glycerolipid fatty acid composition. *Mol Cell Biol*. 2002;22(23):8204–14.

86. Papatheodorou I, Moreno P, Manning J, et al. Expression Atlas update: From tissues to single cells. *Nucleic Acids Res*. 2020;48(D1):D77–D83.

87. Tabula Muris C, Overall C, Logistical C, et al. Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. *Nature*. 2018;562(7727):367–372.

88. Iwasaki Y, Saito Y, Mori K, et al. An autopsied case of adult-onset bulbospinalform Alexander disease with a novel S393R mutation in the GFAP gene. *Clin Neuropathol*. 2015;34(4):207–214.

89. Aggarwal S, Yurlova L, Simons M. Central nervous system myelin: Structure, synthesis and assembly. *Trends Cell Biol*. 2011;21(10):585–593.

90. O'Brien JS. Stability of the myelin membrane. *Science*. 1965;147(3662):1099–1107.

91. Ozgen H, Baron W, Hoekstra D, Kahya N. Oligodendroglial membrane dynamics in relation to myelin biogenesis. *Cell Mol Life Sci*. 2016;73(17):3291–3310.

92. Saher G, Brugger B, Lappe-Siefke C, et al. High cholesterol level is essential for myelin membrane growth. *Nat Neurosci*. 2005;8(4):468–475.

93. Saher G, Stumpf SK. Cholesterol in myelin biogenesis and hypomyelinating disorders. *Biochim Biophys Acta*. 2015;1851(8):1083–1094.

94. Pfrieger FW, Ungerer N. Cholesterol metabolism in neurons and astrocytes. *Prog Lipid Res*. 2011;50(4):357–371.

95. Wang H, Eckel RH. What are lipoproteins doing in the brain? *Trends Endocrinol Metab*. 2014;25(1):8–14.

96. Molofsky AV, Deneen B. Astrocyte development: A guide for the perplexed. *Glia*. 2015;63(8):1320–1329.

97. Camargo N, Goudriaan A, van Deijk AF, et al. Oligodendroglial myelination requires astrocyte-derived lipids. *PLoS Biol*. 2017;15(5):e1002605.

98. Molina-Gonzalez I, Miron VE. Astrocytes in myelination and remyelination. *Neurosci Lett*. 2019;713:134532.

99. Tracey TJ, Steyn FJ, Wolvetang EJ, Ngo ST. Neuronal lipid metabolism: Multiple pathways driving functional outcomes in health and disease. *Front Mol Neurosci*. 2018;11:10.

100. Chrast R, Saher G, Nave KA, Verheijen MH. Lipid metabolism in myelinating glial cells: Lessons from human inherited disorders and mouse models. *J Lipid Res*. 2011;52(3):419–434.

101. Teslovich TM, Musunuru K, Smith AV, et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature*. 2010;466(7307):707–713.

102. Johansen CT, Wang J, Lanktree MB, et al. An increased burden of common and rare lipid-associated risk alleles contributes to the phenotypic spectrum of hypertriglyceridemia. *Arterioscler Thromb Vasc Biol*. 2011;31(8):1916–1926.

103. Tikkanen E, Tuovinen T, Widen E, et al. Association of known loci with lipid levels among children and prediction of dyslipidemia in adults. *Circ Cardiovasc Genet*. 2011;4(6):673–680.

104. Qi Q, Liang L, Doria A, Hu FB, Qi L. Genetic predisposition to dyslipidemia and type 2 diabetes risk in two prospective cohorts. *Diabetes*. 2012;61(3):745–752.

105. Surakka I, Horikoshi M, Magi R, et al. The impact of low-frequency and rare variants on lipid levels. *Nat Genet*. 2015;47(6):589–597.

106. Yang S, Yin RX, Miao L, Zhang QH, Zhou YG, Wu J. Association between the GPAM rs1129555 SNP and serum lipid profiles in the Maonan and Han populations. *Int J Clin Exp Pathol*. 2018;11(3):1484–1498.

107. Brandenburg LO, Pufe T, Koch T. Role of phospholipase d in g-protein coupled receptor function. *Membranes (Basel)*. 2014;4(3):302–318.

108. Foster DA, Salloum D, Menon D, Frias MA. Phospholipase D and the maintenance of phosphatidic acid levels for regulation of mammalian target of rapamycin (mTOR). *J Biol Chem*. 2014;289(33):22583–22588.

109. Guizzetti M, Costa LG. Effect of ethanol on protein kinase Czeta and p70S6 kinase activation by carbachol: A possible mechanism for ethanol-induced inhibition of glial cell proliferation. *J Neurochem*. 2002;82(1):38–46.

110. Samson HH. Microcephaly and fetal alcohol syndrome: Human and animal studies. *Alcohol Brain Dev*. 1986;167–183.

111. Perez-Torrero E, Duran P, Granados L, Gutierez-Ospina G, Cintra L, Diaz-Cintra S. Effects of acute prenatal ethanol exposure on Bergmann glia cells early postnatal development. *Brain Res*. 1997;746(1-2):305–308.

112. Guizzetti M, Zhang X, Goeke C, Gavin DP. Glia and neurodevelopment: Focus on fetal alcohol spectrum disorders. *Front Pediatr*. 2014;2:123.

113. Burkhardt U, Klein J. Ethanol impairs phospholipase D signaling in astrocytes. In: Patel VB, ed. *Molecular aspects of alcohol and nutrition*. Academic Press; 2016:325–335.

114. Hu WF, Chahrour MH, Walsh CA. The diverse genetic landscape of neurodevelopmental disorders. *Annu Rev Genomics Hum Genet*. 2014;15:195–213.

115. Jayaraman D, Bae BI, Walsh CA. The genetics of primary microcephaly. *Annu Rev Genomics Hum Genet*. 2018;19:177–200.

116. Kiray H, Lindsay SL, Hosseinzadeh S, Barnett SC. The multifaceted role of astrocytes in regulating myelination. *Exp Neurol*. 2016;283(Pt B):541–549.

117. Irons M, Elias ER, Salen G, Tint GS, Batta AK. Defective cholesterol biosynthesis in Smith-Lemli-Opitz syndrome. *Lancet*. 1993;341(8857):1414.

118. Brookes KJ, Chen W, Xu X, Taylor E, Asherson P. Association of fatty acid desaturase genes with attention-deficit/hyperactivity disorder. *Biol Psychiatry*. 2006;60(10):1053–1061.

119. Bell JG, MacKinlay EE, Dick JR, MacDonald DJ, Boyle RM, Glen AC. Essential fatty acids and phospholipase A2 in autistic spectrum disorders. *Prostaglandins Leukot Essent Fatty Acids*. 2004;71(4):201–204.

120. Pinto D, Delaby E, Merico D, et al. Convergence of genes and cellular pathways dysregulated in autism spectrum disorders. *Am J Hum Genet*. 2014;94(5):677–694.

121. Jungerius BJ, Hoogendoorn ML, Bakker SC, et al. An association screen of myelin-related genes implicates the chromosome 22q11 PIK4CA gene in schizophrenia. *Mol Psychiatry*. 2008;13(11):1060–1068.

122. Namekawa M, Takiyama Y, Aoki Y, et al. Identification of GFAP gene mutation in hereditary adult-onset Alexander's disease. *Ann Neurol*. 2002;52(6):779–785.

123. Patel SC, Suresh S, Kumar U, et al. Localization of Niemann-Pick C1 protein in astrocytes: Implications for neuronal degeneration in Niemann- Pick type C disease. *Proc Natl Acad Sci U S A*. 1999;96(4):1657–1662.

124. Schmitt A, Gofferje V, Weber M, Meyer J, Mossner R, Lesch KP. The brain-specific protein MLC1 implicated in megalencephalic leukoencephalopathy with subcortical cysts is expressed in glial cells in the murine brain. *Glia*. 2003;44(3):283–295.

125. Dietrich J, Lacagnina M, Gass D, et al. EIF2B5 mutations compromise GFAP+ astrocyte generation in vanishing white matter leukodystrophy. *Nat Med.* 2005;11(3):277–283.

126. Yazdi M, Ahnmark A, William-Olsson L, et al. The role of mitochondrial glycerol-3-phosphate acyltransferase-1 in regulating lipid and glucose homeostasis in high-fat diet fed mice. *Biochem Biophys Res Commun.* 2008;369(4):1065–1070.

127. Li M, Santpere G, Imamura Kawasawa Y, et al. Integrative functional genomic analysis of human brain development and neuropsychiatric risks. *Science.* 2018;362(6420): eaat7615.

128. Lake BB, Chen S, Sos BC, et al. Integrative single-cell analysis of transcriptional and epigenetic states in the human adult brain. *Nat Biotechnol.* 2018;36(1):70–80.