Biallelic variants in PCDHGC4 cause a novel neurodevelopmental syndrome with progressive microcephaly, seizures, and joint anomalies

Maria Iqbal1,2,3,39, Reza Maroofian4,39, Büşranur Çavdarlı5,39, Florence Riccardi6,7,39, Michael Field8,39, Siddharth Banka9,10, Dalal K. Bubshait11, Yun Li12, Jozef Hertecant13, Shahid Mahmood Baig3,14,15, David Dyment16, Stephanie Ethymiou4, Uzma Abdullah17, Ehtisham Ul Haq Makhdoom18, Zafar Ali19, Tobias Scherf de Almeida12, Florence Molinari6, Cécile Mignon-Ravix6, Brigitte Chabrol20, Jayne Antony21, Lesley Ades22,23, Alistair T. Pagnamenta24, Adam Jackson9,10, So Shima Imannezhad30, Ehsan Ghayoor Karimiani31,32, Yasra Sarwar3, Sheraz Khan3, Muhammad Jameel3, Angelika A. Noegel2,33, Mashhad University of Medical Sciences, Mashhad, Iran.31Molecular and Clinical Sciences Institute, St. George Ain, United Arab Emirates.14Department of Biological and Biomedical Sciences, Aga Khan University, Karachi, Pakistan.15Pakistan Science Foundation (PSF), Islamabad, Pakistan. Tübingen Hearing Research Centre (THRC), Eberhard Karls University Tübingen, Tübingen, Germany.27Department of Bioinformatics & Biotechnology, Faculty of Basic and Hospital, Ankara, Turkey.6Aix Marseille Univ, INSERM, MMG, Marseille, France.7Assistance Publique-Hôpitaux de Marseille, Hôpital La Timone Enfants, Département de Génétique Médicale, Marseille, France. Genomics England Research Consortium, Christian Beetz25, Vasiliki Karageorgou25, Barbara Vona26, Aboulfazl Rad26, Jamshaid Mahmood Baig27, Tipu Sultan28, Javeria Raza Ali29, Shazia Maqbool30, Fatima Rahman30, Mehran Beiraghi Toosi30, Farah Ashrafzadeh30, Shima Imannezhad30, Ehsan Ghayoor Karimiani31,32, Yasra Sarwar, Sheraz Khan, Muhammad Jameel3, Angelika A. Noegel2,33 Birgit Budde1, Janine Altmüller1, Susanne Motamney1, Wolfgang Höhne1, Henry Houlden4, Peter Nürnberg1,33, Bernd Wollnik12,34, Brigitte Chabrol20, Jayne Antony21, Lesley Ades22,23, Alistair T. Pagnamenta24, Adam Jackson9,10, So Shima Imannezhad30, Ehsan Ghayoor Karimiani31,32, Yasra Sarwar, Sheraz Khan, Muhammad Jameel3, Angelika A. Noegel2,33

PURPOSE: We aimed to define a novel autosomal recessive neurodevelopmental disorder, characterize its clinical features, and identify the underlying genetic cause for this condition.

METHODS: We performed a detailed clinical characterization of 19 individuals from nine unrelated, consanguineous families with a neurodevelopmental disorder. We used genome/exome sequencing approaches, linkage and cosegregation analyses to identify disease-causing variants, and we performed three-dimensional molecular in silico analysis to predict causality of variants where applicable.

RESULTS: In all affected individuals who presented with a neurodevelopmental syndrome with progressive microcephaly, seizures, and intellectual disability we identified biallelic disease-causing variants in Protocadherin-gamma-C4 (PCDHGC4). Five variants were predicted to induce premature protein truncation leading to a loss of PCDHGC4 function. The three detected missense variants were located in extracellular cadherin (EC) domains EC5 and EC6 of PCDHGC4, and in silico analysis of the affected residues showed that two of these substitutions were predicted to influence the Ca2+-binding affinity, which is essential for multimerization of the protein, whereas the third missense variant directly influenced the cis-dimerization interface of PCDHGC4.

CONCLUSION: We show that biallelic variants in PCDHGC4 are causing a novel autosomal recessive neurodevelopmental disorder and link PCDHGC4 as a member of the clustered PCDH family to a Mendelian disorder in humans.

Genetics in Medicine (2021) 23:2138–2149; https://doi.org/10.1038/s41436-021-01260-4

*Author for correspondence.

1Cologne Center for Genomics (CCG), University of Cologne and University Hospital Cologne, Cologne, Germany. 2Institute of Biochemistry I, Medical Faculty, University of Cologne, Cologne, Germany. 3Human Molecular Genetics Laboratory, Health Biotechnology Division, National Institute for Biotechnology and Genetic Engineering (NIBGE) College, PIEAS, Faisalabad, Pakistan. 4Department of Neuromuscular Disorders, UCL Institute of Neurology, London, UK. 5Department of Medical Genetics, Ankara Bilikent City Hospital, Ankara, Turkey. 6Aix Marseille Univ, INSERM, MMG, Marseille, France. 7Assistance Publique–Hôpitaux de Marseille, Hôpital La Timone Enfants, Département de Génétique Médicale, Marseille, France. 8Genetics of Learning Disability Service, Hunter Genetics, Waratah, NSW, Australia. 9Manchester Centre for Genomic Medicine, St Mary’s Hospital, Manchester University NHS Foundation Trust, Health Innovation Manchester, Manchester, UK. 10Division of Evolution and Genomic Sciences, School of Biological Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK. 11Department of Pediatrics, College of Medicine, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia. 12Institute of Human Genetics, University Medical Center Göttingen, Göttingen, Germany. 13Paediatric Genetic and Metabolic Service, Tawam Hospital, Al Ain, United Arab Emirates. 14Department of University Research and Biomedical Sciences, Aga Khan University, Karachi, Pakistan. 15Pakistan Science Foundation (PSF), Islamabad, Pakistan. 16Children’s Hospital of Eastern Ontario Research Institute, University of Ottawa, Ottawa, Canada. 17University Institute of Biochemistry and Biotechnology (UIBB), PAMAS-And Agriculture University, Rawalpindi, Pakistan. 18Neurochemistry and Genetics Laboratory (NGL), Department of Physiology, Faculty of Life Sciences, Government College University, Faisalabad, Pakistan. 19Centre for Biotechnology and Microbiology, University of Swat, Swat, Pakistan. 20Assistance Publique–Hôpitaux de Marseille, APHM, Hôpital Timone Enfants, Service de Neurologie Pédiatrique, Marseille, France. 21T.Y. Nelson Department of Neurology and Neurosurgery, The Children’s Hospital at Westmead, Sydney, Australia. 22Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK. 23CENTOGENE GmbH, Rostock, Germany. 24Department of Otolaryngology, Head and Neck Surgery, Tübingen Hearing Research Centre (THRC), Eberhard Karls University Tübingen, Tübingen, Germany. 25Department of Bioinformatics & Biotechnology, Faculty of Basic and Applied Sciences, International Islamic University, Islamabad, Pakistan. 26Department of Pediatric Neurology, Children’s Hospital and Institute of Child Health, Lahore, Pakistan. 27Development and Behavioural Pediatrics Department, Institute of Child Health and The Childrens Hospital, Lahore, Pakistan. 28Pediatric Neurology Department, Ghaem Hospital, Mashhad University of Medical Sciences, Mashhad, Iran. 29Molecular and Clinical Sciences Institute, St George’s, University of London, Cranmer Terrace, London, UK. 30Innovative Medical Research Center, Mashhad Branch, Islamic Azad University, Mashhad, Iran. 31Center for Molecular Medicine Cologne (CMMIC), University of Cologne, Faculty of Medicine, University Hospital Cologne, Cologne, Germany. 32Center of Excellence “Multiscale Imaging: from Molecular Machines to Networks of Excitable Cells” (MBExC), University of Göttingen, Göttingen, Germany. 33Department of Translational Genomics, Center for Genomic Medicine, King Faisal Specialist Hospital and Research Center, Riyadh, Saudi Arabia. 34Department of Anatomy and Cell Biology, College of Medicine, Alfaaf University, Riyadh, Saudi Arabia. 35Department of Anatomy and Cell Biology, College of Medicine, Alfaaf University, Riyadh, Saudi Arabia. These authors contributed equally: Maria Iqbal, Reza Maroofian, Büşranur Çavdarlı, Florence Riccardi, Michael Field. *A list of authors and their affiliations appears at the end of the paper. **email: mhussain@uni-koeln.de; goekhan.yigit@med.uni-goettingen.de
INTRODUCTION
Protocadherins (PCDHs) comprise a large family of over 80 cell surface receptors that are mainly expressed during the development of the vertebrate nervous system and play a crucial role in the discrimination between self and nonself cell surface identities in the course of establishment and generation of neuronal circuits [1, 2]. Based on their genomic organization, human PCDHs can be divided into two families which are either encoded by genes distributed across the genome (nonclustered PCDHs) or genes clustered in a 1-Mb region on human chromosome 5 [3]. Clustered PCDHs (cPCDH) are encoded by a total of 53 genes arranged in three subclusters (PCDHA, PCDHB, and PCDHG) within this region [4–6]. All cPCDHs have a similar structure. They are type I transmembrane proteins containing six extracellular cadherin (EC) domains, a transmembrane region, and, in case of α- and γ-PCDH, an intracellular domain (ICD) [1]. In the PCDHA and PCDHG subclusters, multiple “variable” exons, that encode for the entire extracellular region, the transmembrane domain and a variable part of the intracellular region, are tandemly arranged upstream of three “constant” exons, which are shared within a subcluster and code for a common C-terminal intracellular domain [4, 7]. cPCDHs are widely expressed in the developing and mature nervous system including the spinal cord, cerebellum, and hippocampus [8–11]. They have been shown to form homophilic cis- and trans-interactions inducing the formation of multimeric protein complexes [12–14]. Neurons have been suggested to create a unique “barcode” by the expression of different combinations of these proteins that results in the generation of neuron-specific sets of cis-dimers and allows self–nonself discrimination based on the formation of trans-homophilic interactions [2, 15]. Recent functional studies have linked numerous cPCDHs to critical neuronal processes such as regulation of neuronal survival, axon outgrowth and targeting, dendrite arbor complexity, self-avoidance of sister axon and dendrite branches, and synaptogenesis [8, 16–18]. Whereas knockout mice of the α-Pcdh cluster are viable and fertile and show only abnormal axonal projections of serotonergic and olfactory sensory neurons [16, 19], disruption of the γ-Pcdh locus leads to neonatal lethality [8, 20, 21]. Recent studies revealed that Pcdhg3, Pcdhg4, and Pcdhg5 are crucial for the observed lethality [22, 23].

Hitherto, rare variants in nonclustered PCDH have been identified in individuals with different neurodevelopmental disorders. Rare biallelic variants in PCDH12 (OMIM 605622) and PCDH15 (OMIM 605514) have been reported in patients with diencephalic–mesencephalic junction dysplasia syndrome 1 (DMJD51; OMIM 251280), Usher syndrome type 1F (USH1F, OMIM 602083), and nonsyndromic hearing loss (DFNB23; OMIM 609533), respectively [24, 25]. Furthermore, more than 100 disease-causing variants have been described in PCDH19 (OMIM 300460) in developmental and epileptic encephalopathy 9 (DEE9, OMIM 300088), making it one of the clinically relevant genes in epilepsy [26]. So far, no disease-causing variant has yet been identified in any of the cPCDHs to be causative for a Mendelian disorder in humans, despite their important role during neurodevelopment and in neural circuit assembly. In this study, we report the identification of biallelic disease-causing variants in Protocadherin-gamma-C4 (PCDHGC4) in 19 individuals from nine unrelated families. Affected individuals presented with progressive microcephaly, global developmental delay, intellectual disability, seizures, joint anomalies, and additional dysmorphic features. These findings establish biallelic PCDHG4 variants as genetic cause for a novel neurodevelopmental disorder in humans, and elucidate the associated phenotype.

MATERIALS AND METHODS
Subjects
Individuals who participated in this study were clinically characterized in several clinics across the world (see Supplemental Information), and we used the GeneMatcher tool [27] to connect centers in which genetic analyses were performed. All individuals reported herein are born to consanguineous families of different geographic origin, and respective families were not related to each other. Subjects or their legal representatives gave written informed consent for the molecular analyses, publication of the results and clinical information, including photographs. All studies were performed in accordance with the Declaration of Helsinki protocols and were reviewed and approved by the local institutional ethics board. DNA from participating family members was extracted from peripheral blood lymphocytes by standard extraction procedures.

Genome/exome sequencing and linkage analysis
Genome and exome sequencing was performed on patient/parent trios (family 8), single (family 9), or multiple affected family members (families 1–7). Details on sequencing and variant screening as well as genome-wide linkage analysis (family 1) are provided as Supplementary Information.

Variant verification and Sanger sequencing
Verification of identified nonsense and missense variants was performed using standard methods for polymerase chain reaction (PCR) amplification and Sanger sequencing. Primer sequences are available on request. The coding sequence of PCDHG4 (NM_018928.2) was analyzed and variants were confirmed by a second PCR on an independent DNA sample and analyzed for cosegregation within the respective families.

Prediction programs
In silico prediction of the mutational effect for all missense variants was performed using Combined Annotation Dependent Depletion (CADD; https://cadd.gs.washington.edu), MutationTaster (www.mutationtaster.org), PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2), and SIFT (https://sift.bia-star.edu.sg). Variants with potential effects on splicing were characterized using ESEfinder and RESCUE ESE (see Supplemental Information).

Structural analysis of mouse Pcdhgb7 and in silico analysis of the mutational effect
Crystal structure of the Ca2+-bound form of mouse Pcdhgb7 was obtained from the Protein Data Bank (www.wwpdb.org; PDB ID 5v5x). Structural analysis, data visualization, and figure preparation were carried out with the program PyMOL 2.3 (www.pymol.org; Schrödinger, LLC) and WebLab viewerPro (Molecular Simulations Inc.).

RESULTS
Clinical presentation of individuals with a novel neurodevelopmental phenotype
In a national and international collaboration, we recruited 19 individuals from nine unrelated families with a clinical diagnosis of a neurodevelopmental disorder. Clinical findings on all affected individuals are summarized in Table 1, with pedigrees and clinical photographs shown in Fig. 1. Comprehensive clinical information on families (1–4, 7, 8) is provided as Supplemental Information. For five individuals (families 5, 6, and 9), no extensive clinical descriptions are available.

Common features in our patient cohort were developmental delay (DD)/intellectual disability (ID) (18/19), microcephaly (12/19), seizures (10/19), hypotonia (10/19), and skeletal/joint anomalies (10/19). Occipital–frontal circumferences (OCFs) at birth ranged from 1.7 SD (individual IV-3, family 6) to −3 SD (individual VI-2, family 5), and we observed microcephaly at birth (OCF ≤ –2 SD) in 2/19 patients. However, at follow up examinations, 12/19 individuals showed progressive mild to severe microcephaly with values from −2 SD to −5.5 SD. Neuroimaging was available for 12 individuals. Brain magnetic resonance image/computed
Table 1. Summary of genetic data and clinical features of affected individuals.

| Pedigree ID | Family 1 | Family 2 | Family 3 | Family 4 |
|-------------|----------|----------|----------|----------|
|             | IV-3     | V-1      | II-1     | II-2     | II-3     | II-2     | II-3     | II-2     | II-3     | II-1     |
| Agea (years) | 10       | 8        | 20       | 11       | 30       | 27       | 24       | 22       | 6/12b    | 14 |
| Gender       | Male     | Female   | Female   | Male     | Female   | Male     | Female   | Male     | Female   | Male |
| Geographic origin | Pakistan | Pakistan | Turkey   | Iraq     | Iraq     | Iraq     | Morocco  | Morocco  | Morocco  | Morocco |
| Parental consanguinity | ++       | ++       | +        | +        | +        | +        | +        | +        | +        | +   |
| PCDHG4C variant | c.1449C>G; p. (Asp483Glu) | c.1410C>T; p. (Gln40*) c.1463C>T; p. (Ala488Val) | c.1449C>G; p. (Asp483Glu) c.1463C>T; p. (Ala488Val) c.1463C>T; p. (Ala488Val) c.1463C>T; p. (Ala488Val) |
| Birth | Gestation (weeks) | 38 | 38 | 40 | 39 | 40 | 40 | 40 | 41 | NR | 41 |
|         | Weight (g) | 3,000 (−0.4 SD) | 2,500 (−1.3 SD) | 3,850 (0.5 SD) | 3,500 (0.5 SD) | Within normal limits | Within normal limits | Within normal limits | 3,020 (−1 SD) | NR | 3,300 (−1 SD) |
|         | Height (cm) | NR | NR | 45 (−2.4 SD) | 50 (0.2 SD) | Within normal limits | Within normal limits | Within normal limits | 45 (−3 SD) | NR | 48 (−2 SD) |
|         | Head circumference (cm) | NR | NR | 34 (−0.5 SD) | 35 (0.6 SD) | NR | NR | NR | 34 (−1 SD) | NR | 34 (−1 SD) |
| Clinical characteristics | | | | | | | | | | |
| Facial dysmorphism | – | – | Sloping forehead, triangular asymmetric face, high nasal bridge | – | – | – | – | – | – | – |
| Acral anomalies | – | – | Clinodactyly, hallux valgus | Hypoplasia of toes | Broad thumbs, swan neck deformity | Broad thumbs, swan neck deformity | Swan neck deformity | Bilateral ulnar clubhand, valgus deformities | Bilateral ulnar clubhand | Bilateral ulnar clubhand |
| Seizures | – | + | +; at 2, 4 and 7 years of age | +; at 1.5 and 6.5 years of age | – | – | +; (singular) febrile seizure | +; status epilepticus starting at 5 years | +; generalized seizures starting at 18 months | +; generalized seizures starting at 18 months |
| Brain MRI anomalies | Thin cerebral cortex | NR | Cerebral atrophy | – | NR | NR | NR | Mild cortical atrophy | NR | Cortical atrophy |
| Hearing impairment | – | – | – | – | – | – | – | – | NR | – |
| Ocular anomalies | Strabismus | – | – | – | – | – | – | – | NR | – |
| Additional findings | Gait abnormalities | Gait abnormalities, recurrent infections | Kyphoscoliosis | Kyphosis | Hypotonic at birth, unstable gait, hyperextensible joints | Hypotonic at birth, unstable gait, hyperextensible joints, white hypopigmented patch | Hypotonic at birth, hyperextensible joints | Hypotonic at birth, lumboscoliosis | Hypotonic at birth | Arthrogryposis |
| Intellectual disability | Moderate Speech impairment, started walking at 4 years of age | Severe Aggressive behavior, poor self-care | Moderate Hyperactivity | Mild Hyperactivity | Severe Global developmental delay, able to speak 2–3 words, able to toilet and dress with assistance; started walking between 7 and 10 years | Severe Global developmental delay, nonverbal, able to toilet and dress with assistance; started walking between 7 and 10 years | Severe Global developmental delay, able to toilet and dress with assistance; started walking between 7 and 10 years | Severe Global developmental delay, no language development at the age of 2 years; able to speak with a lot of words and to copy a text, able to toilet and to dress | Moderate | NR | Severe Global developmental delay, nonverbal, uses signs and gestures to communicate; walked at 10 years of age, able to dress with assistance |
| Pedigree ID | Family 1 | Family 2 | Family 3 | Family 4 |
|------------|---------|---------|---------|---------|
|            | IV-3    | V-1     | II-1    | II-2    | II-3    | II-1    | II-2    | II-3    |
|            |         |         |         |         |         |         |         |         |
| **Current measurements** |         |         |         |         |         |         |         |         |
| Head circumference (cm) | 45 (−4.7 SD) | 45 (−4.5 SD) | 53     | 49     | 51 (−2.5 SD) | 51 (−3.7 SD) | 49 (−3.9 SD) | 52 (−2 SD) | NR | 51 (−2 SD) |
| Height (cm)   | 127 (−2 SD) | 113 (−2.5 SD) | 153   | 149   | 150 (−2.1 SD) | 164 (−1.8 SD) | 150 (−2.1 SD) | 152 (−2 SD) | NR | 138 (−2 SD) |
| Weight (kg)   | NR      | NR      | NR      | NR      | 39      | 38.5     | 32.4     | 48      | NR | 38      |

| Pedigree ID | Family 5 | Family 6 | Family 7 | Family 8 | Family 9 |
|------------|---------|---------|---------|---------|---------|
|            | VI-1    | VI-2    | IV-1    | IV-3    |         |
|            |         |         |         |         |         |
| **Clinical characteristics** |         |         |         |         |         |
| Facial dysmorphism | Mild metopic craniosynostosis, everted inferior lips | Mild metopic craniosynostosis | Epicanthic fold, high nasal bridge, low set ears, overlapping teeth | Epicanthic fold, high nasal bridge, low set ears, overlapping teeth | Dolichocephaly, mild synophrys, upslanting palpebral fissures, long eyelashes, high nasal bridge, hooked nose, overhanging columella, short philtrum, full lips, high narrow palate, mild malar hypoplasia | Mild synophrys, long eyelashes, high nasal bridge, broad nasal bridge and nasal tip, slightly overhanging columella, short philtrum, full lips | Mild synophrys, long eyelashes, high nasal bridge, short philtrum, full lips | Microcephaly, receding forehead | Mildly dysmorphic chin, low nasal bridge, small, upturned nose, long philtrum, metopic craniosynostosis |

| Acral anomalies | Broad thumbs | – | – | Persistent fetal fingertip pads, mild 2nd and 5th clinodactyly | Mild 2nd and 5th finger clinodactyly, small broad feet | – | – | Bilateral thumb-in-palm deformity, rigid equinovarus deformity (right foot), fixed rocker bottom deformity with severe calcaneovalgus (left foot) |

| Seizures | +; tonic seizures starting at 18 months, last seizure at 8 3/12 years | – | – | – | – | – | – | +; febrile seizures |
| Pedigree ID | Family 5 | Family 6 | Family 7 | Family 8 | Family 9 |
|------------|----------|----------|----------|----------|----------|
| VI-1       |          |          |          |          |          |
| Brain MRI anomalies | Mild extra axial widening (at 4/6 years of age) | Left-sided hypodensities within the temporal lobe and left cerebral hemisphere as results of previous microhemorrhages; CSF space enlargement | Acute ischemic insult within the left thalamus and left parietal region | Left temporal fossa aschmoid cyst, mild ventriculomegaly and thinning of body and genu of corpus callosum | Mild ventriculomegaly and thinning of anterior body of corpus callosum |
| Hearing impairment | – | – | – | – | – |
| Ocular anomalies | Microphthalmus, ptosis of the right eye | Esotropia, mild oculomotor apraxia | Esotropia | Oculomotor apraxia | Oculomotor apraxia |
| Additional findings | DDH | DDH | Spasticity and truncal hypotonia, bilateral DDH | Spasticity and truncal hypotonia, hip dislocation (left) | Neonatal hypertonia, 15° contractures at elbows, hypoplastic calves, mild genu valgum, brisk reflexes |
| Intellectual disability | Severe | Moderate | Moderate | Mild | Moderate |
| Development | Global developmental delay, started walking at the age of 4 years, only babbling | Global developmental delay, can stand with support | Global developmental delay, severe receptive and expressive language delay | Global developmental delay | Severe speech delay, global developmental delay |
| Current measurements | | | | | |
| Head circumference (cm) | 48 (−3.6 SD) | 44 (−2.5 SD) | 41 (−2.5 SD) | 56.5 (0.3 SD) | 47 (−1.2 SD) |
| Height (cm) | 127 (−0.5 SD) | 67 (−4.7 SD) | 81 (−2.5 SD) | 155.5 (−2.9 SD) | 95 (−1.3 SD) |
| Weight (kg) | 35 (0.8 SD) | 9 (−1.0 SD) | 28 (−3.4 SD) | 7.3 (−0.6 SD) | 13 |

**Table 1 continued**
Fig. 1 Pedigrees and clinical characteristics of individuals harboring biallelic disease-causing variants in *PCDHGC4*. (a) Pedigrees of nine unrelated families with disease-causing variants in *PCDHGC4*. All affected siblings (solid symbols) in each family carry homozygous disease-causing variants in *PCDHGC4* while unaffected parents are heterozygous for identified *PCDHGC4* variants (white symbols). (b) Upper panel: facial features of subjects IV-3 and V-1 from family 1 (left), clinical characteristics of subjects II-1 and II-2 from family 2 showing kyphoscoliosis, clinodactyly and hallux valgus (subject II-1), and kyphosis and hypoplasia of the toes (subject II-2). Lower panel (from left to right): facial features and hand anomalies observed in subjects II-1 (22 years) and II-3 (14 years) from family 4, clinical characteristics of subjects VI-1 and VI-2 from family 5, and subjects IV-2 and IV-3 from family 6, and facial features and foot anomalies observed subject VI-1 from family 8.
tomography (MRI/CT) abnormalities (11/12 patients) were rather nonspecific (Fig. 2, Table 1). Microcephaly, thin cerebral cortex, mild ventriculomegaly, and cortical atrophy were the commonest features. Seizure types ranged from singular febrile seizures (family 3, subject II-3), recurring events (family 2, subjects II-1 and II-2) to generalized tonic, clonic–focal to multifocal seizures (families 4 and 5; Table 1). Electroencephalogram (EEG) data was available for four subjects (family 2, subjects II-1 and II-2; family 5, subject VI-2; family 7, subjects III-1) and showed no abnormalities. ID and DD were present in all our patients, and we observed motor and speech developmental delay as well as mild to severe cognitive impairment. Three individuals presented with kyphosis and/or scoliosis, hyperextensible joints were observed in three individuals and contractures as well as arthrogryposis were present in four individuals (Fig. 1b, Table 1). Dysmorphic facial features were rather nonspecific and did not reveal a common, recognizable facial presentation within our patient cohort (Fig. 1b, Table 1).

Identification of biallelic truncating and missense PCDHGC4 variants

We performed linkage analysis (family 1) and/or genome/exome sequencing in probands and proband/parent trios. Based on parental consanguinity, autosomal recessive inheritance was considered likely, and we prioritized homozygous, rare exonic, and splice site variants (see Supplemental Information). We identified three different missense variants and five protein truncating variants in the Protocadherin-gamma family member PCDHGC4 (OMIM 606305; NM_018928.2) in all affected individuals (Fig. 3a, b, S1, Table 2). All variants fully cosegregated with the phenotype in the respective families and are absent or very rare in the general human population with minor allele frequencies (MAFs) ranging from 0 to 4 × 10⁻⁶, in line with an autosomal recessive pattern of inheritance (Table 2). We identified four homozygous loss-of-function variants in PCDHGC4, c.118C>T (p.[Gln40*]), c.324del (p.[Phe108Leufs*14]), c.1243C>T (p.[Arg415*]), c.1724dup (p.[Leu575Phefs*63]), that were predicted to lead to an early stop and premature protein truncation, and were absent from the gnomAD database (Fig. 3b, Table 2). In family 9, we found the homozygous variant c.2443-1G>A at the acceptor splice site of intron 1, and by employing an exon-trapping approach we could show that this variant leads to a loss of the acceptor splice-site recognition resulting in severe splicing defects such as whole-exon skipping or usage of a cryptic exonic acceptor splice site, which both are predicted to induce a frameshift and premature protein truncation (Fig. 3b, S3). Within the family of γ-PCDHs, PCDHGC4 is the only member that is not only highly conserved across species, but also under strict mutational constraint [23]. Truncating variants in PCDHGC4 are rarely observed in healthy control individuals. For the canonical transcript of PCDHGC4 (ENST00000306593.1, NM_018928.2) only 12 alleles with non-sense variants, all in heterozygous state, were reported in the gnomAD database in contrast to 29.6 that were expected to be observed in the >240,000 alleles (probability of loss of function intolerance [pLI] = 0.98). Furthermore, biallelic copy-number variants (CNVs) encompassing PCDHGC4 have not been reported so far in the DECIPHER database, the Database of Genomic Variants (DGV), and the structural variant (SV) data set of gnomAD with only two (DGV) and six (gnomAD) heterozygous alterations enlisted in these data sets that affect PCDHGC4. Interestingly, genetic disruption of the entire γ-Pcdh cluster as well as singular knockout of Pcdhg4c in mice also cause a severe neurodevelopmental phenotype, both resulting in neurodegeneration in late embryonic stages and leading to early neonatal lethality [8, 20–23].

Furthermore, we identified three different homoyzgous missense variants, c.1449C>G (p.[Asp483Glu]), c.1463C>T (p.[Ala488Val]), and c.1817T>G (p.[Val606Gly]), in affected individuals of four additional consanguineous families (Fig. 3, Table 2). In silico prediction of the pathogenic effect of these missense variants by different prediction tools leads to the classification as damaging (SIFT), probably damaging (PolyPhen-2), and a Combined Annotation Dependent Depletion (CADD) score of 24.1 to 26.9, indicating deleteriousness of these variants (Table 2). Two missense variants, p.[Asp483Glu] and p.[Ala488Val], were classified as polymorphisms by a single in silico prediction tool, MutationTaster. In two families, families 3 and 6 from Iraq and Saudi Arabia, respectively, we identified the identical missense variant, c.1463C>T (p.[Ala488Val]), in PCDHGC4. In affected individuals of both families, this variant was within a shared haplotype of approximately 309 kb between chr5:140,750,044 and chr5:141,059,868 suggesting a founder.
which enables cis changes in domains, which upon Ca\textsuperscript{2+} conformation and rigidity of these segments is controlled, motifs at the junctions of the EC repeats of PCDHs, the human Genomics in Medicine (2021) 23:2138 – 2149 binding of Ca\textsuperscript{2+} binding is a crucial process for correct PCDH function. Upon PCDHG4 (Fig. 3c). EC domains are extracellular Ca\textsuperscript{2+} substitution of phylogenetically highly conserved amino acids in extracellular cadherin (EC) domain and are predicted to lead to the nature of the variant. On protein level, the three missense variants are located in the extracellular domain of PCDHGC4, within the fifth (p.Asp483Glu) and p.Ala488Val) or sixth (p.Val606Gly) extracellular cadherin (EC) domain and are predicted to lead to the substitution of phylogenetically highly conserved amino acids in PCDHG4 (Fig. 3c). EC domains are extracellular Ca\textsuperscript{2+}-binding domains, which upon Ca\textsuperscript{2+} binding can mediate conformational changes influencing the rigidity of the EC domains of PCDHG4, which enables cis- and trans-homophilic interactions [2]. Ca\textsuperscript{2+} binding is a crucial process for correct PCDH function. Upon binding of Ca\textsuperscript{2+}, which is mediated by several calcium-binding motifs at the junctions of the EC repeats of PCDHs, the conformity and rigidity of these segments is controlled, allowing formation of cis- as well as trans-dimerizations [28, 29]. Whereas EC1 to EC4 contribute to the formation of head-to-tail trans interactions between different cells, EC5 and EC6 are involved in cis-dimerization processes. To gain further insights into the pathogenic effects of the missense variants, we performed an in silico analysis of the mutational effect on the protein structure using the crystal structure of mouse Pcdhgb7, a close homologue of PCDHGC4. All three missense variants were located in or directly adjacent to a Ca\textsuperscript{2+}-binding motif. The p.Asp483Glu variant affects an aspartate that is part of the highly conserved DXD motif in the EC5 repeat of PCDHGC4 directly involved in calcium coordination (Fig. 3d, e). Although this variant does not change the charge of the coordinating residue, it alters
Table 2. In silico prediction and population allele frequencies of PCDHGC4 (NM_018928.2; ENST00000306593.1) variants identified in this study.

| Family | Genomic location (GRCh37/hg19) | HGVS cDNA | HGVS protein | Allele frequency (gnomAD database<sup>a</sup>) | Prediction scores |
|--------|--------------------------------|-----------|--------------|-----------------------------------------------|------------------|
|        |                                 |           |              |                                               | SIFT<sup>b</sup> | PolyPhen-2<sup>c</sup> | MutationTaster<sup>d</sup> | CADD             |
| 1      | chr5:140866189                   | c.1449C>G  | p.(Asp483Glu) | 0                                              | D 0.000 | PD 1.00 | Polymorphism 0.932 | 24.1             |
| 2      | chr5:140864858                   | c.118C>T   | p.(Gln40*)    | 0                                              | NA     | NA     | NA                 | NA               |
| 3;6    | chr5:140866203                   | c.1463C>T  | p.(Ala488Val) | 0.000004                                       | D 0.003 | PD 0.971 | Polymorphism 0.528 | 25.2             |
| 4      | chr5:140866557                   | c.1817T>G  | p.(Val606Gly) | 0                                              | D 0.000 | PD 0.968 | DC 0.999           | 26.9             |
| 5      | chr5:140865064                   | c.324del   | p.(Phe108Leufs*14) | 0                                              | NA     | NA     | DC 1.00           | NA               |
| 7      | chr5:140865983                   | c.1243C>T  | p.(Arg415*)   | 0                                              | NA     | NA     | NA                 | NA               |
| 8      | chr5:140866464                   | c.1724dup  | p.(Leu575Phefs*63) | 0                                              | NA     | NA     | DC 1.00           | NA               |
| 9      | chr5:140874373                   | c.2443-1G>A| NA           | 0                                              | NA     | NA     | DC 1.00           | NA               |

cDNA complementary DNA, D deleterious, DC disease causing. NA not applicable, PD probably damaging.

<sup>a</sup>Accessed in January 2021.

<sup>b</sup>Score 1–0.

<sup>c</sup>HumVar prediction, score 0–1.
progenitor cells which migrate to the developing cortex. Upon arrival, ~40% of these cells are eliminated by endogenously triggered programmed cell death [33–35]. Interestingly, except Pcdhga9, all 21 γ-Pcdhs are expressed in clNs. Expression of four isoforms, Pcdhga1, Pcdhga2, Pcdhgc4, and Pcdhgc5, increases significantly between P8 and P15, corresponding to the period in which programmed cell death of clNs takes place [36]. Recent studies showed that Pcdhgc3, Pcdhgc4, and Pcdhgc5 are crucial components in the regulation of this programmed cell death. Loss of these isoforms enhances the number of clNs undergoing apoptosis, which results in a reduced cortical density of clNs [36]. A similar function of γ-PCDHs in controlling programmed cell death has also been described for neuronal cells of the spinal cord and the retina [8, 9, 22]. Still, further molecular and cellular studies are required to determine whether disease-causing variants in PCDHG4 are alone sufficient to increase programmed cell death in neuronal cells and to give rise to the clinical presentation observed in our patients via this pathway. Interestingly, genetic disruption of the entire γ-Pcdh cluster as well as singular knockout of Pcdhgc4 in mice both result not only in neurodegeneration in late embryonic stages but also lead to early neonatal lethality [8, 20–23]. Currently, it is unclear why disruption of Pcdhgc4 in mice leads to neonatal lethality, whereas biallelic loss-of-function variants in human, as observed in our patients, result in a milder neurodevelopmental disorder comprising progressive microcephaly, seizures, and intellectual disability, especially when considering that both humans and mice share the same set of 22 members within the γ-PCDH cluster. But the difference between the observed phenotypes suggests that the human brain might compensate for the functional failure of PCDHG4 resulting in a higher tolerance of loss-of-function variants in terms of lethality.

So far, to the best of our knowledge, no member of the clustered PCDH family has been shown to be involved in the pathogenesis of a congenital human disorder. In recent years, disease-causing variants in several nonclustered δ-PCDH family members have been described and closely linked to different neurodevelopmental diseases. This includes biallelic loss-of-function variants in PCDH12 and PCDH15, which were identified as cause of diencephalic–mesencephalic junction dysplasia syndrome type 1 (DMUDS1), Usher syndrome type 1F, and nonsyndromic hearing loss, respectively, as well as PCDH19, in which 100 different missense and nonsense variants have been reported to underlie X-linked developmental and epileptic encephalopathies 9 (DEE9) highlighting the importance of cell–cell communication via PCDH19 at the early stages of brain development [24–26, 37]. Although recent studies indicate that complete or partial epigenetic dysregulation of the clustered PCDH occurs in cells of patients with Williams–Beuren syndrome or Down syndrome, and hypermethylation of all three PCDH clusters is detectable in Wilms tumors, a direct link to a monogenic, congenital human disorder has not been established before [38–40]. Currently, it is unclear whether this is due to functional redundancy of the encoded PCDHs. Mice lacking the α- or β-Pcdh cluster are viable and fertile [16], whereas knockout of the whole γ-Pcdh cluster results in neonatal lethality [8, 20, 21]. Similar consequences were observed when only the γC3 to γC5 isoforms within this cluster were disrupted, indicating that one of these three isoforms has a critical function [22]. Very recent results based on the generation of single knockouts of γ-Pcdh members suggest that Pcdhgc4 is the crucial isoform within the gamma cluster, which is required for neuronal survival and responsible for neonatal lethality [23]. The unique role of PCDHG4 is further supported by genetic data indicating that PCDHG4 is the only member within the γ-PCDH cluster that is under strict mutational constraint [23]. We can only speculate about the molecular basis of the distinct role of PCDHG4, especially as its overall structure is similar to other γ-PCDHs.

In conclusion, we show that biallelic truncating and missense variants in PCDHG4 cause a specific human phenotype characterized by neurodevelopmental delay, progressive microcephaly with mild to severe intellectual disability, global developmental delay, joint anomalies, and seizures, providing evidence that disease-causing variants in a single member of the clustered PCDH family are involved in the pathogenesis of a congenital disorder in humans.

DATA AVAILABILITY
The data that support the findings of this study are available on request from the corresponding authors. The genetic data are not publicly available due to privacy or ethical restrictions.

Received: 19 March 2021; Revised: 14 June 2021; Accepted: 15 June 2021;Published online: 9 July 2021

REFERENCES
1. Sano K, et al. Protocadherins: a large family of cadherin-related molecules in central nervous system. EMBO J. 1993;12:2249–2256.
2. Rubinstein R, et al. Molecular logic of neuronal self-recognition through protocadherin domain interactions. Cell. 2015;163:629–642. https://doi.org/10.1016/j.cell.2015.09.026.
3. Hulpiau P, van Roy F. Molecular evolution of the cadherin superfamily. Int J Biochem Cell Biol. 2009;41:349–369. https://doi.org/10.1016/j.biocel.2008.09.027.
4. Wu Q, Maniatis T. A striking organization of a large family of human neural-likelihood cell adhesion genes. Cell. 1999;97:779–790. https://doi.org/10.1016/S0092-8674(00)80789-8.
5. Sugino H, et al. Genomic organization of the family of CNR cadherin genes in mice and humans. Genomics. 2000;63:73–87. https://doi.org/10.1006/geno.1999.0666.
6. Wu Q, et al. Comparative DNA sequence analysis of mouse and human protocadherin gene clusters Genome Res. 2001;11:389–404. https://doi.org/10.1101/gr.167301.
7. Wu Q, Maniatis T. Large exons encoding multiple ectodomains are a characteristic feature of protocadherin genes. Proc Natl Acad Sci USA. 2000;97:3124–3129. https://doi.org/10.1073/pnas.000627397.
8. Wang X, Weiner JA, Levi S, Craig AM, Bradley A, Sanes JR. Gamma protocadherins are required for survival of spinal interneurons. Neuron. 2002;36:843–854. https://doi.org/10.1016/S0896-6273(02)01090-5.
9. Lefebvre JL, Kostadinov D, Chen WW, Maniatis T, Sanes JR. Protocadherins mediate dendritic self-avoidance in the mammalian nervous system. Nature. 2012;488:517–521. https://doi.org/10.1038/nature11305.
10. Chen WW, et al. Pcdhca2 is required for axonal tiling and assembly of serotonergic circuitries in mice. Science. 2017;356:406–411. https://doi.org/10.1126/science.aal3231.
11. Mountoufaris G, et al. Multicluster Pcdh diversity is required for mouse olfactory neural circuit assembly. Science. 2017;356:411–414. https://doi.org/10.1126/science.aal8801.
12. Murata Y, Hamada S, Morishita H, Mutoh T, Yagi T. Interaction with protocadherin-gamma regulates the cell surface expression of protocadherin-alpha. J Biol Chem. 2004;279:49508–49516. https://doi.org/10.1074/jbc.M408771200.
13. Fernández-Monreal M, Kang S, Phillips GR. Gamma-protocadherin homophilic interaction and intracellular trafficking is controlled by the cytoplasmic domain in neurons. Mol Cell Neurosci. 2009;40:344–353. https://doi.org/10.1016/j.mcn.2008.12.002.
14. Schreiner D, Weiner JA. Combinatorial homophilic interaction between gamma-protocadherin multimers greatly expands the molecular diversity of cell adhesion. Proc Natl Acad Sci USA. 2010;107:14893–14898. https://doi.org/10.1073/pnas.1004526107.
15. Thu CA, et al. Single-cell identity generated by combinatorial homophilic interactions between α, β, and γ protocadherins. Cell. 2014;158:1045–1059. https://doi.org/10.1016/j.cell.2014.07.012.
16. Hasegawa S, et al. The protocadherin-alpha family is involved in axonal coalescence of olfactory sensory neurons into glomeruli of the olfactory bulb in mouse. Mol Cell Neurosci. 2008;38:66–79. https://doi.org/10.1016/j.mcn.2008.01.016.
17. Garrett AM, Schreiner D, Lobas MA, Weiner JA. γ-protocadherins control cortical dendrite arborization by regulating the activity of a FAK/PKC/MARCKS signaling pathway. Neuron. 2012;74:269–276. https://doi.org/10.1016/j.neuron.2012.01.028.
18. Molumby MJ, Keefer AB, Weiner JA. Homophilic protocadherin cell-cell interactions promote dendrite complexity. Cell Rep. 2016;15:1037–1050. https://doi.org/10.1016/j.celrep.2016.03.093.
2148

We are grateful to all family members that participated in this study. This work was supported by the Higher Education Commission (HEC) of Pakistan to M.J. and E.U.H. M.; the Center for Molecular Medicine Cologne (CMMC) (Projects 38-RP and C12; 2635/8029/01 and 2635/8326/01) to P.N. and M.S.H.; the Koeln Fortune Program (Faculty of Medicine, University of Cologne; 2011/2020) to M.S.H.; Intramural Funding from the Fortune Program (Faculty of Medicine, University of Tubingen; 2545-1-0) and the Ministry of Science, Research and Art Baden-Württemberg to B.V.; the German Research Foundation (DFG, Deutsche Forschungsgemeinschaft) under Germany’s Excellence Strategy—EXC 2067/1–390729940 to B.W.; the Medical Research Council (MRC) (MR/S01165X/1, MR/S005021/1, G0601943), the National Institute for Health Research University College London Hospitals Biomedical Research Centre, the Rosetree Trust, Ataxia UK, MSA Trust, Brain Research UK, Sparks GOSH Charity, Muscular Dystrophy UK (MDUK), Muscular Dystrophy Association (MDA USA) to H.H.; sequencing of family 3 was performed under the Care4Rare Canada Consortium funded by Genome Canada and the Ontario Genomics Institute (OGI-147), the Canadian Institutes of Health Research, Ontario Research Fund, Genome Alberta, Genome British Columbia, Genome Quebec, and Children’s Hospital of Eastern Ontario Foundation. Three of the authors of this publication are members of the European Reference Network for Intellectual Disability, Telehealth and Congenital Anomalies (ERN-ITHACA; Project ID No 739543) (F.Riccardi; S.Banka; S.Dougou). Family 5 was collected as part of the SYNAPS Study Group collaboration funded by The Wellcome Trust and strategic award (Synaptopathies) (WT93205 MA and WT104033AA) funding and research was conducted as part of the Queen Square Genomics group at University College London, supported by the National Institute for Health Research University College London Hospitals Biomedical Research Centre. Genomic data for family 8 was generated by the 100,000 Genomes Project. The 100,000 Genomes Project is managed by Genomics England Limited (a wholly owned company of the Department of Health and Social Care). The 100,000 Genomes Project is funded by the National Institute for Health Research and NHS England. The Wellcome Trust, Cancer Research UK and the Medical Research Council have also funded research infrastructure. The 100,000 Genomes Project uses data provided by patients and collected by the National Health Service as part of their care and support. A.J. is supported by Solve-RD. The Solve-RD project has received funding from the European Union’s Horizon 2020 research and innovation program under grant agreement number 779257.

FUNDING
Open Access funding enabled and organized by Projekt DEAL.

ETHICS DECLARATION
Subjects or their legal representatives gave written informed consent for the molecular analyses, publication of the results and clinical information, including photographs. All studies were performed in accordance with the Declaration of Helsinki protocols and were reviewed and approved by the local institutional ethics boards (National Institute for Biotechnology and Genetic Engineering [NIBGE], Faisalabad, Pakistan; Department of Medical Genetics, Ankara Bilkent City Hospital, Ankara, Turkey; Children’s Hospital of Eastern Ontario Research Institute, University of Ottawa, Ottawa, Canada; CPP [Comité de Protection des Personnes] I Sud Méditerranée, France; University College London Hospitals, London, UK; HRA Committee East of England [Cambridge South REC; 14/EE/1112], King Faisal Specialist Hospital and Research Center, Riyadh, Saudi Arabia; 100,000 Genomes Project, Genomics England Limited, UK).

COMPETING INTERESTS
V.K. and C.B. are employees of Centogene GmbH (Rostock, Germany). The other authors declare no competing interests.

ADDITIONAL INFORMATION
Supplementary material The online version contains supplementary material available at https://doi.org/10.1038/s41436-021-01260-4.

Correspondence and requests for materials should be addressed to M.S.H. or G.Y.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
