Heterozygous variants in CTR9, which encodes a major component of the PAF1 complex, are associated with a neurodevelopmental disorder

**Heterozygous variants in *CTR9*, which encodes a major component of the PAF1 complex, are associated with a neurodevelopmental disorder**

Marije Meuwissen[1,*], Aline Verstraeten[1,*], Emmanuelle Ranza[2,*], Justyna Iwaszkiewicz[3], Maaike Bastiaansen[1], Ligia Mateiu[1], Merlijn Nemegeer[1], Josephina A.N. Meester[1], Alexandra Afenjar[4], Michelle Amaral[5], Diana Ballhausen[6], Sarah Barnett[7], Magalie Barth[8], Bob Asselbergh[9,10], Katrien Spaas[9,10], Bavo Heeman[11,10], Jennifer Bassetti[12], Patrick Blackburn[13], Marie Schaer[14], Xavier Blanc[2], Vincent Zoete[3, 15], Kari Casas[16], Thomas Courtin[17], Diane Doummar[18], Frédéric Guerry[2], Boris Keren[17], John Pappas[19], Rachel Rabin[19], Amber Begtrup[20], Marwan Shinawi[21], Anneke T. Vulto-van Silfhout[22], Tjitske Kleefstra[22], Matias Wagner[23,24], Alban Ziegler[8], Elise Schaefer[25], Benedicte Gerard[26], Charlotte I. De Bie[27], Sjoerd J.B. Holwerda[27], Mary Alice Abbot[28], Stylianos E. Antonarakis[28,&], Bart Loeys[1,22,&]

[1] Center for Medical Genetics, Antwerp University Hospital/ University of Antwerp, Edegem, Belgium

[2] Medigenome, Swiss Institute of Genomic Medicine, Geneva, Switzerland

[3] Molecular Modeling Group, Swiss Institute of Bioinformatics, Quartier Sorge - Bâtiment Amphipôle, CH 1015 Lausanne, Switzerland

[4] Sorbonne Université, Centre de Référence Malformations et maladies congénitales du cervelet et déficiences intellectuelles de causes rares, département de génétique et embryologie médicale, Hôpital Trousseau, AP-HP, Paris, France

[5] Hudson Alpha Institute for Biotechnology, Huntsville, Alabama, USA

[6] Pediatric Metabolic Unit, Pediatrics, Woman-Mother-Child Department, University of Lausanne and University Hospital of Lausanne, Switzerland

[7] Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, USA

[8] Biochemistry and Genetics Department, University Hospital of Angers, Angers, France.

[9] Neuromics Support Facility, VIB Center for Molecular Neurology, VIB, Antwerp, Belgium

[10]Department of Biomedical Sciences, University of Antwerp, Antwerp, Belgium

[11] Applied and Translational Neurogenomics, VIB Center for Molecular Neurology, VIB, Antwerp, Belgium.

[12] Division of Medical Genetics, Department of Pediatrics, Weill Cornell Medicine, New York, N.Y, USA

[13] Department of Pathology, St. Jude Children's Research Hospital, Memphis, TN, 38105, USA.

[14] Autism Brain & Behavior Laboratory, Faculty of Medicine, University of Geneva, Geneva, Switzerland

[15] Department of Fundamental Oncology, Lausanne University, Ludwig Institute for Cancer Research, Route de la Corniche 9A, CH-1066 Epalinges, Switzerland

[16] Medical Genetics, Sanford Broadway Clinic, Fargo, ND, USA

[17] Sorbonne Université, Department of Genetics, AP-HP, La Pitié-Salpêtrière hospital, Paris

[18] Sorbonne Université, Neuropédiatrie, AP-HP, Hôpital d'enfants Armand Trousseau, Paris

[19] NYU Langone Medical Center, New York, NY, USA

[20] GeneDx, Gaithersburg, MD, 20877, USA

[21] Dept. of Pediatrics, Division of Genetics and Genomic Medicine, Washington University School of Medicine, St. Louis, MO, USA

[22] Department of Human Genetics, Radboud University Medical Center, Nijmegen, The Netherlands

[23] Institute of Human Genetics, Technical University München, Munich, Germany.

[24] Institute for Neurogenomics, Helmholtz Zentrum München, Neuherberg, Germany

[25] Service de Génétique médicale, Institut de Génétique médicale d'Alsace, Hopitaux universitaires de Strasbourg, Strasbourg, France

[26] Laboratoires de diagnostic génétique, Institut de Génétique médicale d'Alsace, Hopitaux universitaires de Strasbourg, Strasbourg, France

[27]Department of Clinical Genetics, University Medical Center Utrecht, Utrecht, The Netherlands

[28] Medical Genetics, Department of Pediatrics, University of Massachusetts Medical School – Baystate, Springfield, MA 01199


*Equally contributed to this work

[&] Correspondence to either Stylianos E. Antonarakis (stylianos.antonarakis@medigenome.ch) or Bart Loeys (bart.loeys@uantwerpen.be)

**ABSTRACT**

PURPOSE CTR9 is a subunit of the PAF1 complex (PAF1C), that plays a crucial role in transcription regulation by binding of CTR9 to RNA polymerase II. It is involved in transcription-coupled histone modification by promoting H3K4 and H3K36 methylation. We describe the clinical and molecular studies in 13 probands harboring likely pathogenic *CTR9* missense variants collected through GeneMatcher.

METHODS Whole exome sequencing was performed in all individuals. CTR9 variants were assessed by three-dimensional modelling of the activated human transcription complex Pol II-DSIF-PAF-SPT6 and the PAF1/CTR9 complex. H3K4/H3K36 methylation analysis, mitophagy assessment based on tetramethylrhodamine ethyl ester perchlorate immunofluorescence and RNA-sequencing in skin fibroblasts from 4 patients was performed.

RESULTS. Common clinical findings were variable degrees of intellectual disability, hypotonia, joint hyperlaxity, speech delay, coordination problems, tremor and autism spectrum disorder. Mild dysmorphism and cardiac anomalies were less frequent. For eleven *CTR9* variants, *de novo* occurence was demonstrated. Three-dimensional modelling predicted a likely disruptive effect of the variants on local CTR9 structure and protein interaction. Additional studies in fibroblasts did not unveil the downstream functional consequences of the identified variants.

CONCLUSION We describe a neurodevelopmental disorder caused by (mainly) *de novo* variants in *CTR9*, likely affecting PAF1C function.

**KEYWORDS**

*CTR9;* PAF1C; intellectual disability; autism spectrum disorder; neurodevelopmental disorder

**INTRODUCTION**

The Polymerase-Associated Factor 1 (PAF1) complex (PAF1C) is a highly conserved transcriptional regulator that consists of PAF1, CTR9, LEO1, CDC73, RTF1 and WDR61. CTR9 plays a crucial role in the transcriptional function of PAF1C by binding to the RNA polymerase II. PAF1C is also involved in transcription-coupled histone modification, including H2B mono-ubiquitination as well as H3K4 and H3K36 methylation[1,2], and recruits the ATP-dependent chromatin remodeling enzyme CHD1[3]. In yeast, Paf1C, especially through Paf1 and Ctr9, has been demonstrated to be involved in the induction of mitophagy, i.e. a gatekeeper of cellular metabolism that controls degradation of superfluous or damaged mitochondria[4]. In zebrafish, deletion of *Ctr9* leads to abnormal heart and neural crest development by dysregulation of Notch signaling[5,6]. Studies of mutant *Ctr9* in Drosophila demonstrate a key role in nervous system development[7]. Additionally, in rat brain, *Ctr9* was demonstrated to interact with dopamine transporters in dopaminergic neurons, independent of Paf1C, suggesting a dual nuclear and cytoplasmic function[8].

While PAF1C appears crucial for development, developmental disorders have not been associated with genetic variants in most of its components. Only *LEO1* variants were recently linked to neurodevelopmental disease including autism spectrum disorder (ASD), developmental delay and neurobehavioral problems[9].

As heterozygous *CTR9* variants have been reported in four unrelated Wilms tumor families, *CTR9* was previously classified identified as a Wilms tumor predisposition gene[10]. In three of these families the pathogenic variants affected the consensus splice site, resulting in an in-frame exon 9 deletion while in the fourth family a truncating variant in exon 2 was identified. Further studies in affected tissue derived from these respective patients demonstrated absence of the wild-type *CTR9* allele, indicative of a second hit in the tumor[10].

Three independent *de novo CTR9* missense variants c.1405G>A, p(Glu469Lys)[11], c.2488C>T, p.(Arg830Trp)[12] and c.2515C>T, p(Arg839Trp)[13] were previously identified as singletons in three large

intellectual disability (ID)[11], ASD[12] and developmental disorder[13] cohorts respectively, however, without phenotypic details.

Here, we report 13 patients with (predominantly *de novo*) heterozygous *CTR9* missense variants suffering from a neurodevelopmental disorder characterized by varying degrees of ID, neurodevelopmental delay, hypotonia, fatigability, behavioral abnormalities including ASD, anxiety and aggressive behavior, cardiac anomalies and mild facial dysmorphism.

## MATERIALS AND METHODS

### Patients and samples

The patient collection was established using the web-based platform Genematcher[14]. The study was approved by the institutional ethics review boards and informed consents were obtained from the legal guardians of all participating subjects. A separate informed consent for publication of identifiable photographs was provided.

### Whole exome sequencing

Whole exome sequencing (WES), data annotation and variant filtering were performed in different centers. Details on WES methodology for each case are described in the Supplementary data 1. Variants were classified based on ACMG/AMP guidelines for variant classification[15].

### Structural modelling of CTR9 variants

Available structures of human CTR9 were used; i.e. two electron microscopy structures of the activated human transcription complexes Pol II-DSIF-PAF-SPT6 (PDB: 6TED (residues 3-892) [16] and 6GMH (residues 3-892)[17]), both with 3.1 Å resolution. To evaluate the effect of N-terminal CTR9 variants, the

cryo-EM structure of the human PAF1/CTR9 complex (2.53 Å resolution) was used, containing residues 3-244 of CTR9 and residues 57-116 of PAF1 (PDB: 5ZYQ [18]). Variant impact on protein stability was studied with Foldx 5.0[19]. Side chain replacements (wild type to variant) were performed on the 5ZYQ cryo-EM structure and were repeated five times for each variant. Differences in energy between the folded and unfolded state, referred to as protein stability, were calculated for the wild type and mutant protein and averaged over five mutant structures. Visualization and manipulation of the structures was done with the UCSF Chimera software[20].

**Functional analysis of CTR9 variants**

Methodologies used for fibroblast culture, H3K4 and H3K36 methylation analysis, mitochondrial membrane potential analysis, RNA extraction and RNA-sequencing are described in Supplementary data 2.

**RESULTS**

Identification of *CTR9* variants

Trio-based whole exome sequencing in case 1 and both unaffected parents revealed the presence of the heterozygous *de novo* p.(Glu15Lys) variant in *CTR9* (MIM 609366; RefSeq accession number NM_014633.5; hg19) in the patient. This variant is absent from gnomAD v2.1.1, affects an amino acid residue that is highly conserved up to *D. melanogaster* (Supplementary figure 1) and is predicted damaging by multiple *in silico* prediction programs (Supplementary Table 1). Using Genematcher, we identified 12 additional individuals with *CTR9* missense variants and neurodevelopmental phenotypes. In ten of these individuals it concerned a *de novo* variant. In case 3, the variant was inherited from the affected father and in case 10, only absence in maternal DNA was demonstrated. Except for two variants (p.(Asn455Ser) (1/251322) and p.(Arg878Gln) (1/249646); case 9 and 12 respectively), all

variants are absent from gnomAD v2.1.1. Moreover, all affect highly conserved residues and *in silico* predictions largely suggest deleterious effects (Supplementary Figure 1 and Supplementary Table 1). Interestingly, the p.(Glu376Lys) variant, affecting a CpG dinucleotide, was found in three *de novo* cases, suggesting a mutational hotspot. Variant positions are depicted in Figure 1A and Supplementary Figure 2.

## Clinical description of individuals harboring *CTR9* variants

We describe 13 cases with heterozygous *CTR9* missense variants presenting with a variable neurodevelopmental disorder. Clinical features are summarized in Table 1 and individual cases are described in Supplementary data 3.

### *Cognition*

All cases showed impaired cognitive abilities with variable extent. ID was present in all but three cases (4, 10 and 12), ranging from mild to severe. While case 4 and case 10 don't meet the strict criteria for ID (total IQs above 70), both show significant impairment in other neurodevelopmental domains. The least affected case (case 12) has an IQ of 93, but experiences learning difficulties whereas both parents are highly educated.

### *Neurodevelopment*

Motor development was complicated by hypotonia in nine cases, in six cases joint hyperlaxity was obvious. Other recurrent findings included speech delay (n=9), coordination problems or tremor (n=6) and muscular weakness (n=3).

Brain MRI, performed in six cases, showed non-specific abnormalities in two, comprising delayed myelination, short/ thin corpus callosum and ventricular dilatation.

### *Behaviour*

ASD was present in 4 cases, of which 2 showed developmental regression and severe behavioral problems with aggression during late childhood. Psychotic episodes, aggression, auto-mutilation, ADHD, stereotypies, anxiety and mood disorders were also reported in a subset of cases.

*Visceral findings*

Four patients showed cardiac abnormalities comprising infantile thoracic aortic aneurysm, ventricular septal defect, mild pulmonary valve stenosis and supravalvular aortic stenosis presenting in childhood with normal findings on an early echocardiogram.

*Dysmorphic features*

Variable non-specific, subtle dysmorphic features were noted (Figure 2B-M); some recurrent features include hypertelorism, micro/retrognathia and a broad, flat nasal bridge. Nine cases showed feet abnormalities comprising club feet, varus position of the forefoot, flat feet, sandal gap and clinodactyly, although some of these foot features are also common in the general population.

*Other*

Five cases had failure to thrive and/or feeding problems, including infantile feeding difficulties and selective feeding. Fatigability and unexplained pain were recurrently seen.


*In silico* **structural modelling of** *CTR9* **variants**

The positions of the affected residues on the cryo-EM structure of CTR9 are depicted in Figure 1N. A detailed analysis of the variants is presented in Supplementary Data and Supplementary Figure 3, 4 and 5, but in general the variants described here can be divided into three categories based on their probable consequences for the structure of the PAF1C complex. A first group of variants most probably destabilizes the local CTR9 protein structure. The variants p.(Ile17Thr), p.(Pro25Arg), p.(Glu26Gln), p.(Glu37Gln) and p.(Cys85Tyr) most likely destabilize the N-terminal part of CTR9, possibly also influencing the interactions of CTR9 with PAF1 protein (Supplementary Figure 3, Supplementary Table 2). A second group of variants is predicted to perturb the PAF1 interactions in the tetratricopeptide repeats regions (TPRs). The p.(Glu469Lys), p.(Glu376Lys), p.(Asn455Ser) and p.(Thr766Ala) variants most probably influence the interactions of CTR9 with the PAF1 protein either directly, or indirectly, by induced changes in the local conformation of the TPR domain of CTR9

(Supplementary Figure 4). Finally, a third class of variants influences the interactions with other PAF1C complex subunits. The variant p.(Glu15Lys) does not influence the structural stability of CTR9 (Supplementary table 2). Therefore we assume that its phenotype is likely caused by an effect on the yet unknown interactions on the N-terminal part of CTR9. The substitution of Arg878 to Gln weakens the interactions of CTR9 "trestle" with Pol II subunit E (Supplementary Figure 5).

**Functional analysis of CTR9 variants**

H3K4/H3K36 methylation analysis, mitochondrial quality assessment and RNA-sequencing in fibroblasts does not provide conclusive evidence for downstream functional consequences of the likely the disease-causing variants (Supplementary data 4).

**DISCUSSION**

We report heterozygous missense variants in *CTR9*, a main component of the PAF1 complex, in 13 unrelated cases presenting with a neurodevelopmental disorder. *De novo* occurrence of the variants was documented in 11 cases, in one case the variant was inherited from a mildly affected father and in one case only the mother was available for testing. The missense variants are spread over the *CTR9* gene. A hotspot variant (p.Glu376Lys) was recurrent in three cases (case 7,8 and 13). The only consistent finding between these three cases seems to be the developmental delay and ID. Overall, the neurodevelopmental disorder in our cases includes variable ID (ranging from borderline intelligence to mild/moderate) and delayed motor and speech development. In several cases, motor development is hampered by hypotonia and joint hypermobility. Behavioral problems are prominent in the phenotypic spectrum, with reports of ASD, sometimes associated with regression, aggressive and self-injurious behavior, ADHD, anxiety and mood disorders.

Interestingly, only missense variants were identified, suggesting a dominant-negative effect of the variants in our cohort. The constraint z-score for missense variants in CTR9 in the gnomAD v2.1.1 is

4.3, suggesting intolerance to missense variation. The hypothesized dominant-negative effect is supported by the prior identification of loss-of-function variants in Wilms tumor patients in the absence of neurodevelopmental problems[10]. Homozygous knock-out mice from Welcome Trust Sanger Institute consortium (Ctr9[tm1b(EUCOMM)Wtsi]/Ctr9[tm1b(EUCOMM)Wtsi]) showed embryonic lethality (http://www.informatics.jax.org/marker/phenotypes/MGI:109345).

In yeast, Paf1c is a known regulator of H2B monoubiquitylation, mediated by the Rad6/ Bre1 complex, and of H3K4 and H3K36 di-and trimethylation, mediated by Set1 and Set2 respectively[1]. Pathogenic variants in the human orthologues of these genes/ complexes under PAF1C regulation, including *UBE2A* (*Rad6*), *SETD1A*/ *SETD1B* (*Set1*) and *SETD2* (*Set2*) were recently associated with neurodevelopmental disorders[21-23]. Associated phenotypes show an important overlap with our *CTR9* cases, including ID, developmental delay, ASD and other psychiatric and behavioral manifestations, including aggressive outbursts(Supplementary Table 4). We were unable to demonstrate differential H3K4 or H3K36 methylation in *CTR9* fibroblasts compared to controls. However, because of the phenotypic resemblance, we hypothesize a brain-specific effect of *CTR9*-variants on H3K4 and H3K36 methylation, explaining the absence of differential methylation of H3K4 and H3K36 in fibroblasts. The same may be true for mitophagy dysregulation, as was demonstrated for *UBE2A* pathogenic variants but could not be shown in *CTR9* patient fibroblasts. Future studies are needed to investigate this hypothesis of brain-specific differential H3K4 and H3K36 methylation and/or mitophagy dysregulation in *CTR9*-variants.

In 4 cases (case 1, 3, 6 and 10), variable cardiac abnormalities were identified, comprising of thoracic aortic aneurysm, subaortic valve stenosis, ventricular septal defect (VSD) and pulmonary valve stenosis. PAF1C was previously demonstrated to play a role in cardiogenesis in zebrafish by regulating the developmental potential of the lateral plate mesoderm, cardiomyocyte numbers and heart tube elongation[6]. Zebrafish *Ctr9[LA961]* mutants showed severe defects in morphogenesis of the primitive heart tube[6]. Interestingly, PAF1C was demonstrated to regulate NOTCH signaling in zebrafish[5]. While,

due to its role during cardiogenesis[24], impaired NOTCH signaling might contribute to the cardiac phenotype in our cases, this hypothesis could not be supported by our RNA-seq analysis. Altogether, a cardiac evaluation seems warranted in patients harboring *CTR9* pathogenic variants.

*In vivo* studies, e.g. in zebrafish, could be useful to further study the pathogenicity of the identified *CTR9* variants, however are beyond the scope of the present study.

We conclude that heterozygous, mainly *de novo* missense variants in *CTR9* cause a novel neurodevelopmental disorder characterized by variable ID and neurodevelopmental delay that can be associated with hypotonia, fatigability, behavioral abnormalities including ASD, anxiety, aggressive behavior, cardiac anomalies and mild dysmorphism. Future studies to elucidate the underlying pathogenic mechanism are warranted.

## SUPPLEMENTARY INFORMATION

Eight supplementary figures, three supplementary tables and a detailed description of predicted effect of CTR9 variants on protein function, a case-by-case clinical description and details on the variant detection methodology are available online.

disorders (VASCERN - project ID: 769036 partly co-funded by the European Union Third Health Programme). Dr. Meuwissen, Dr. Kleefstra and Dr. Keren are member of the European Reference Network on Developmental Anomalies and Intellectual Disability (ITHACA).

**ETHICS DECLARATION**

Data were obtained in diagnostic setting in the individual centers after IRB approval. The main IRB in this study is the IRB of the University Hospital Antwerp, Belgium. Informed consent to publish individuals' clinical information and photographs was obtained from the (parents of the) individuals reported in this article.

**COMPETING INTERESTS**

Amber Begtrup is an employee of GeneDx, Inc. The other authors declare no conflict of interest.

**AUTHOR CONTRIBUTIONS**

Conceptualization: M.M., A.V., E.R., S.E.A., B.L.; Supervision: M.M., A.V., E.R., S.E.A., B.L.; Data curation: M.M., A.V., E.R.; Methodology: M.M., A.V., E.R., S.E.A., B.L.; Investigation: M.M., A.V., E.R., J.I., M.B., L.M., M.N., J.A.N.M., A.A., M.A., D.B., S.B., M.B., B.A., K.S, B.H., J.B., P.B., A.B., M.S., X.B., V.Z., K.C., T.C., D.D., F.G., B.K., J.P., R.R., M.S., A.T.V.S., T.K., M.W., A.Z., E.S., B.G., C.I.D.B., S.J.B.H., M.A.A.; Formal analysis: M.M., A.V., E.R., J.I., M.B., L.M., M.N., J.A.N.M., A.A., M.A., D.B., S.B., M.B., B.A., K.S, B.H., J.B., P.B., A.B., M.S., X.B., V.Z., K.C., T.C., D.D., F.G., B.K., J.P., R.R., M.S., A.T.V.S., T.K., M.W., A.Z., E.S., B.G., C.I.D.B., S.J.B.H. and M.A.A.; Writing-original draft: M.M., A.V.; Data curation: M.M., A.V., E.R., J.A.N.M., S.E.A. and B.L.; Visualization: M.M., A.V., E.R., J.I., L.M. and B.A.; Funding acquisition: A.V., B.L.; Writing-review and editing: All authors.

**DATA AVAILABILITY**

The authors declare that the data supporting the findings of this study are available within the article and its supplementary material. Raw sequencing data are available from the corresponding authors on request if in line with the provided consent of the families.

Variants have been submitted to ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/):

- c.43G>A, p.(Glu15Lys)          Accession ID: SCV001738054

- c.109G>C, p.(Glu37Gln)          Accession ID: SCV001738055

- c.50T>C, p.(Ile17Thr)          Accession ID: SCV001738056

- c.1405G>A, p.(Glu469Lys)          Accession ID: SCV001738057

- c.254G>A, p.(Cys85Tyr)          Accession ID: SCV001738058

- c.76G>C, p.(Glu26Gln)          Accession ID: SCV001738059

- c.1126G>A, p.(Glu376Lys)          Accession ID: SCV001738060

- c.1364A>G, p.(Asn455Ser)          Accession ID: SCV001738061

- c.2296A>G, p.(Thr766Ala)          Accession ID: SCV001738062

- c.74C>G p.(Pro25Arg)          Accession ID: SCV001738063

- c.2633G>A p.(Arg878Gln)          Accession ID: SCV001738064

**REFERENCES**

1.    Crisucci EM, Arndt KM. The Roles of the Paf1 Complex and Associated Histone Modifications in Regulating Gene Expression. Genet Res Int. 2011;2011.
2.    Wood A, Schneider J, Dover J, Johnston M, Shilatifard A. The Paf1 complex is essential for histone monoubiquitination by the Rad6-Bre1 complex, which signals for histone methylation by COMPASS and Dot1p. J Biol Chem. 2003;278(37):34739-34742.
3.    Tomson BN, Arndt KM. The many roles of the conserved eukaryotic Paf1 complex in regulating transcription, histone modifications, and disease states. Biochim Biophys Acta. 2013;1829(1):116-126.
4.    Zheng L, Shu WJ, Li YM, et al. The Paf1 complex transcriptionally regulates the mitochondrial-anchored protein Atg32 leading to activation of mitophagy. Autophagy. 2020;16(8):1366-1379.

5.  Akanuma T, Koshida S, Kawamura A, Kishimoto Y, Takada S. Paf1 complex homologues are required for Notch-regulated transcription during somite segmentation. EMBO Rep. 2007;8(9):858-863.
6.  Langenbacher AD, Nguyen CT, Cavanaugh AM, Huang J, Lu F, Chen JN. The PAF1 complex differentially regulates cardiomyocyte specification. Dev Biol. 2011;353(1):19-28.
7.  Bahrampour S, Thor S. Ctr9, a Key Component of the Paf1 Complex, Affects Proliferation and Terminal Differentiation in the Developing Drosophila Nervous System. G3 (Bethesda). 2016;6(10):3229-3239.
8.  De Gois S, Slama P, Pietrancosta N, et al. Ctr9, a Protein in the Transcription Complex Paf1, Regulates Dopamine Transporter Activity at the Plasma Membrane. J Biol Chem. 2015;290(29):17848-17862.
9.  Wang T, Hoekzema K, Vecchio D, et al. Large-scale targeted sequencing identifies risk genes for neurodevelopmental disorders. Nat Commun. 2020;11(1):4932.
10. Martins AG, Pinto AT, Domingues R, Cavaco BM. Identification of a novel CTR9 germline mutation in a family with Wilms tumor. Eur J Med Genet. 2018;61(5):294-299.
11. Lelieveld SH, Reijnders MR, Pfundt R, et al. Meta-analysis of 2,104 trios provides support for 10 new genes for intellectual disability. Nat Neurosci. 2016;19(9):1194-1196.
12. De Rubeis S, He X, Goldberg AP, et al. Synaptic, transcriptional and chromatin genes disrupted in autism. Nature. 2014;515(7526):209-215.
13. Deciphering Developmental Disorders S. Prevalence and architecture of de novo mutations in developmental disorders. Nature. 2017;542(7642):433-438.
14. Sobreira N, Schiettecatte F, Valle D, Hamosh A. GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. Hum Mutat. 2015;36(10):928-930.
15. 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.
16. Vos SM, Farnung L, Linden A, Urlaub H, Cramer P. Structure of complete Pol II-DSIF-PAF-SPT6 transcription complex reveals RTF1 allosteric activation. Nat Struct Mol Biol. 2020;27(7):668-677.
17. Vos SM, Farnung L, Boehning M, et al. Structure of activated transcription complex Pol II-DSIF-PAF-SPT6. Nature. 2018;560(7720):607-612.
18. Xie Y, Zheng M, Chu X, et al. Paf1 and Ctr9 subcomplex formation is essential for Paf1 complex assembly and functional regulation. Nat Commun. 2018;9(1):3795.
19. Delgado J, Radusky LG, Cianferoni D, Serrano L. FoldX 5.0: working with RNA, small molecules and a new graphical interface. Bioinformatics. 2019;35(20):4168-4169.
20. Pettersen EF, Goddard TD, Huang CC, et al. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem. 2004;25(13):1605-1612.
21. Hiraide T, Nakashima M, Yamoto K, et al. De novo variants in SETD1B are associated with intellectual disability, epilepsy and autism. Hum Genet. 2018;137(1):95-104.
22. Luscan A, Laurendeau I, Malan V, et al. Mutations in SETD2 cause a novel overgrowth condition. J Med Genet. 2014;51(8):512-517.
23. Singh T, Kurki MI, Curtis D, et al. Rare loss-of-function variants in SETD1A are associated with schizophrenia and developmental disorders. Nat Neurosci. 2016;19(4):571-577.
24. MacGrogan D, Munch J, de la Pompa JL. Notch and interacting signalling pathways in cardiac development, disease, and regeneration. Nat Rev Cardiol. 2018;15(11):685-704.

**Figure 1. CTR9 variant localisation, clinical pictures and modeling of CTR9 structure and mutated residues**

A. Localization of CTR9 variants at protein level. Representation of the CTR9 protein, including the localization of the Tetratricopeptide (TPR) repeats in orange/ blue, numbered 1-16. Previously reported missense variants are indicated in red (ref 13,14, 15); variants reported in this paper are presented in black (RefSeq accession number NM_014633).

B-M. Mild dysmorphism in individuals harboring *CTR9* variants. (B, C, D) Facial features of case 1 showing a high forehead, arched eyebrows and a left-sided ear pit with slightly low-set ears. (E, F) Facial features of case 10, showing arched eyebrows, a full nasal tip, retrognathia and full lips. (G, H) Facial features of case 4, comprising of a high forehead with broad eyebrows, attached ear lobules, full alae nasi, full lips, a high palate and a prominent chin. (I) Facial features of case 5, showing flared eyebrows and a full lower lip. (J, K, L, M) Hands and feet of cases 4 (J, L) and 10 (K, M). Case 10 shows mild hallux valgus and broad feet with short toes.

N. Structure of CTR9 in RNA PAF II complex and the positions of mutated residues. Cryo-EM structures of the activated RNA Pol-II complex (6TED.pdb, 5ZYQ.pdb) with positions of CTR9 mutated residues analyzed in this study. The crystallographic structure of N-terminal part of CTR9 (5ZYQ.pdb) was overlapped with the cryo-EM structure of the complex. The CTR9 protein is presented in cornflower blue, WDR61 protein in salmon, PAF1 protein in orange, CDC73 in magenta, RNA in red, the rest of the complex in yellow. Wild type residues of CTR9, analyzed in this study, are presented in green color in sphere representation. The heteroatoms of analyzed residues are colored – oxygen in red, nitrogen in blue and sulfur in yellow.

**Table 1. Clinical features of patients harboring *CTR9* variants**

| | Case 1 | Case 2 | Case 3 | Case 4 | Case 5 | Case 6 | Case 7 | Case 8 | Case 9 | Case 10 | Case 11 | Case 12 | Case 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *CTR9* variant (NM_014633.5; hg19) | c.43G>A, p.(Glu15Lys) | c.109G>C, p.(Glu37Gln) | c.50T>C, p.(Ile17Thr) | c.1405G>A, p.(Glu469Lys) | c.254G>A, p.(Cys85Tyr) | c.76G>C, p.(Glu26Gln) | c.1126G>A, p.(Glu376Lys) | c.1126G>A, (p.Glu376Lys) | c.1364A>G, p.(Asn455Ser) | c.2296A>G, p.(Thr766Ala) | c.74C>G p.(Pro25Arg) | c.2633G>A p.(Arg878Gln) | c.1126G>A, p.(Glu376Lys) |
| Inheritance | *De novo* | *De novo* | Paternal inheritance | *De novo* | *De novo* | *De novo* | *De novo* | *De novo* | *De novo* | Unknown (not maternal) | *De novo* | *De novo* | *De novo* |
| Sex | M | F | M | F | M | F | M | F | M | F | M | F | M |
| Growth parameters | | | | | | | | | | | | | |
| Birth weight (g) | 3830 g | 3460 g | 3650 g | 3575 g (+1 SD) | 3850 g (+0.1 SD) | ND | 950 g (28 WG) | ND | 2750 g | 2270 g (32 WG) | 3100 g | Average | 2350 g (34 WG) |
| Age at assessment | 6 y 10 m | 9.5 y | 18 m | 18 y | 12 y | 20 y | 3 y 3 m | 15 y | 31 m | 22 y | 17 m | 7 y | 6 y |
| Height (SD) | 109 cm (-2.1 SD) | 138 cm (+0.49 SD) | 78 cm (-1.18 SD) | 160 cm (-1.7 SD) | 151 cm (+0.2 SD) | 150.3 (-1.99 SD) | 98 cm (+0.51 SD)) | ND | 93.7 cm (+0.39 SD) | 153.8 cm (-1.46 SD) | 77.5 cm (-0.49 SD) | 120.1 cm (-1.6 SD) | 121,7 cm (+1 SD) |
| Weight (SD) | 17,5 kg (-2 SD) | 34.7 kg (+ 0.5 SD) | 9.9 kg (-1.55 SD) | 79,2 kg (+3 SD) | 31 kg (-2 SD) | 68.3 kg (+0.67 SD) | 13.2 kg (-0.75 SD) | ND | 13.7 kg (+0.08 SD) | 80.5 kg (+1.41 SD) | 9.7 kg (-1.04 SD) | 24.9 kg (-0.1 SD) | 29 kg (+3 SD) |
| HC (SD) | ND | 54.5 cm (+1.87 SD) | 50 cm (+1.5 SD) | 56,8 cm (+1 SD) | ND | ND | 52 cm (+2 SD) | ND | 47.1 cm (-1.39 SD) | 53.8 cm (-0.47 SD) | 50 cm (+ 2.83 SD) | 53.8 cm (+1.4 SD) | 52,5cm (+0,5 SD) |
| Hypertelorism | - | + | + | - | + | - | + | ND | ND | - | - | + | - |
| Broad and flat nasal bridge | - | - | + | - | + | ND | - | ND | + | - | - | + | - |
| Joint hyperlaxity | + | +++ | - | + | - | ND | + | ND | ND | + | + | - | - |
| Skeletal findings | Overpronation of the forefoot ; slight pectus excavatum, flat feet | Varus of the forefoot, flat feet | - | Flat feet | Flat feet | ND | Flat feet | ND | ND | Hallux valgus, brachydactyly, flat feet | Pronated feet, 2nd toe overlapping 1st and 3rd toes, clinodactyly of the 5th toes, flat feet | Mild clinodactyly of the 5th finger; sandal gap both feet, flat feet | Rigidity of the lower limbs; tapered fingers; clinodactyly of the 5th digits |
| Other dysmorphic features | Slightly velvety skin | Rounded forehead, flat, long philtrum, wide mouth, low-set ears, retrognathia | ND | - | Velvety skin | ND | ND | ND | Epicanthal folds, bulbous nose with hypoplastic nasal tip, retrognathia | Tapered fingers | Deep set eyes, prominent forehead, upslanted palpebral fissures, broad nasal tip | Epicanthal folds; sparse lateral eyebrows; long and flat philtrum; thin upper lip; retrognathia; | Synophris; small mouth; almond eyes; inverted nipples |
| Delayed early milestones | + | + | + | - | - | - | + | + | + | + | + | - | + |
| Hypotonia | + | + | + | - | - | - | + | + | + | + | + | - | + |
| Motor delay | + | + | + | ± | - | - | + | ND | + | + | + | - | + |
| Speech delay | + | + | - | - | + | + | + | - | + | + | + | - | + |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Regression** | - | - | - | - | + | + | - | - | ± | ND | ND | - | ND |
| **Intellectual disability** | Learning difficulties | +; mild | ND yet | Dysharmonic IQ (VIQ 103, PIQ 75) | + (severe) | + | +, mild to moderate | +, mild | + | +; FSIQ 77 | ND | -; IQ 93, learning difficulties | +; IQ between 58-70 |
| **ASD** | - | - | - | + | + | + | ND | ND | ND | ND | ND | - | + |
| **Behavioral abnormalities** | - | - | Tantrums, agitation, ADHD | Psychotic episodes | Agressive behavior, automutilation, ADHD | | - | ADHD | Stereotypies with hand flapping | Mood disorder, OCD, PTSD, reactive attachment disorder, ADHD | ND | Hand flapping when happy; social-emotional delay | Social-emotional delay, anxiety, ADHD |
| **Other neurological findings** | Balance problems | Balance problems with gait difficulties | ND | Balance/coordination problems; muscle weakness | - | ND | ND | Uncoordination, numbness in feet, muscle weakness | Balance difficulties | History of dystonia, tremors, muscle weakness | ND | - | - |
| **Brain MRI** | ND | Normal | Short CC, delayed myelinitation | ND | Normal | Normal | Delayed myelination, thin CC, ventricular dilatation | ND | No MRI | Normal | No MRI; normal head ultrasound | No MRI | No MRI |
| **Cardiovascular findings** | TAA; BAV; brain vessel tortuosity | - | Subaortic valve stenosis | Normal cardiac evaluation | - | VSD | ND | ND | Normal ultrasound at 1 week | History of mild PVS | Normal fetal ultrasound | Normal cardiac evaluation | - |
| **Ocular findings** | - | Astigmatism, hypermetropia | - | Intermittent esotropia, anisometropia | Abnormal ocular movements | - | ND | ND | Esotropia, bilateral; astigmatism | Glasses | ND | - | - |
| **Unexplained pain** | - | Unexplained pain in legs | ND | Unexplained pain (back and legs) | Unexplained pain episodes | ND | ND | ND | ND | ND | ND | - | Unexplained pain in knee |
| **Easy fatigability** | ND | + | ND | + | ND | ND | + | + | + | ND | ND | + | ND |
| **Additional findings** | FTT; eczema; constipation | | ND | Constipation | FTT since 10 y; phases of anorexia/ hyperphagia | Neonatal hepatomegaly and hemolytic anemia, infantile pyknocytosis | FTT in the first 2 years of life; poor muscle mass | Constipation | Neonatal feeding difficulties | IBS type issues, headaches, dizziness, hot/ cold intolerance | FTT in the first few weeks; constipation | - | Infantile feeding difficulties |

Abbreviations: **F**: female; **M**: male; **y**: years; **m**: months; **WG** : weeks of gestation ; **ND**: not documented; **ASD**: autism spectrum disorder; **TAA**: thoracic aortic aneurysm; **BAV**: bicuspid aortic valve; **VSD**: ventricle septum defect; **PVS**: pulmonary valve stenosis; **CC**: corpus callosum; **FTT**: failure to thrive

**Supplementary data 1: Whole exome sequencing methodology**

**Case 1**

WES of peripheral blood-derived DNA of the proband and his unaffected parents involved sequence capture with the Agilent SureSelect Human All Exon v5 enrichment kit (Agilent Technologies), followed by sequencing on the Illumina HiSeq platform (BGI, Copenhagen, Denmark). After read alignment with BWA (GRCh37 reference build) and GATK-based variant calling, VariantDB was used to annotate, filter and manually curate the identified variants[1].

**Case 2**

Family Trio (patient and both parents) exome sequencing was performed. Library preparation, sequencing was performed according to the manufacturer protocol. Exons were captured using SeqCap EZ MedExome kit (Roche Technologies) and sequence was generated on a NextSeq 500 instrument (Illumina Inc.). For data analysis, raw reads were mapped to the human reference genome GRCh37 using the Burrow-Wheeler aligner (v0.717). The resulting binary alignment/map (BAM) files were further processed by Genome Analysis Tool Kit HaplotypeCaller (GATK HC v3.8). The VCF files were then annotated on Snpeff version 4.3T 29. Only coding non-synonymous and splicing variants were considered. Variant prioritization was conducted in line with the transmission mode (*de novo*, autosomal recessive and X-linked), and the frequency of the variants in the gnomAD database.

**Case 3**

Exome sequencing of Case 3, his symptomatic father and asymptomatic mother were performed using DNA extracted from peripheral blood leukocytes by Integragen. The exonic regions and flanking splice junctions of the genome were captured using the SureSelect Clinical Research Exome (Agilent Technologies, Santa Clara, CA). Massively parallel (NextGen) sequencing was done on an Illumina system with 100bp or greater paired-end reads. Median depth of coverage was superior to 60X for the three exomes with a coverage >25X for 98.5% of the exome. Reads were aligned to human genome build GRCh37/UCSC hg19, and analyzed for sequence variants using a custom-developed analysis tool. Analysis of WES data prioritized variants based on the allele frequency in population database (GnomAD, EVS and 1000 genomes), the presence in databases of medically relevant variants (ClinVar, Human Gene mutation Database), the predicted impact on coding sequence and the mode of inheritance.

**Case 4**

Trio-based exome sequencing on DNA from the index patient and her parents was performed following routine diagnostic procedures as described previously[2]. Essentially, DNA was sequenced on an Illumina HiSeq system after exome enrichment using the Agilent SureSelectXT Human All Exon 50Mb Kit. Reads were aligned to the Hg19 reference genome with BWA and variants were called using GATK and annotated using an in-house developed pipeline. Trio-based analysis of a panel of genes known to cause intellectual disability did not identify a likely cause for her phenotype. Subsequently, analysis of the whole exome was done.

**Case 5**

Whole exome sequence of Individual 5 was performed using the following procedure. DNA was captured and coding regions and splice sites were enriched using the capture kit Twist Core Exome + Spike-In (Twist Bioscience). Sequencing was performed on an Illumina NextSeq 500

instrument following the manufacturer's protocol. Sequences were aligned to the reference human genome build GRCh37/hg19 using Sentieon (version 201911) and GATK. Confirmation of candidate pathogenic/causative variants and segregation analysis were performed by PCR amplification and Sanger sequencing. Bioinformatic analysis was performed with the Saphetor pipeline. The evaluation of variants used the following databases: gnomAD (version 2.1.1), BRAVO, ClinVar (version 04-Apr-2020), LOVD, and the local database of variants. Variants were classified according to the recommendations of the American College of Medical Genetics[3], and the Task Force of Medigenome.

**Case 6-10-11**

Patient 6 had GeneDx's Autism/ID Xpanded Panel (2300+ genes, trio approach): https://www.genedx.com/tests/detail/autism-id-xpanded-panel-849. Patient 10 had duo exome sequencing and Patient 11 had trio exome sequencing. Using genomic DNA from the proband and parent(s), when available, the exonic regions and flanking splice junctions of the genome were captured using the IDT xGen Exome Research Panel v1.0 (Integrated DNA Technologies, Coralville, IA). Massively parallel (NextGen) sequencing was done on an Illumina system with 100bp or greater paired-end reads. Reads were aligned to human genome build GRCh37/UCSC hg19, and analyzed for sequence variants using a custom-developed analysis tool. Additional sequencing technology and variant interpretation protocol has been previously described[4]. The general assertion criteria for variant classification are publicly available on the GeneDx ClinVar submission page (http://www.ncbi.nlm.nih.gov/clinvar/submitters/26957/)

**Case 7**

Trio-exome sequencing of the affected child and the parents was performed using a Sure Select Human All Exon 60Mb V6 Kit (Agilent) for enrichment and the NovaSeq6000 (Illumina) platform for sequencing. An average of 135,888,843 reads were produced per sample and aligned to the UCSC human reference assembly (hg19) with BWA v.0.5.8.1 More than 98% of the exome was covered at least 20× and the average coverage was more than 126× in all four samples. Single-nucleotide variants (SNVs) and small insertions and deletions were detected with SAMtools v.0.1.7. Copy number variations (CNVs) were detected with ExomeDepth and Pindel. Variant prioritization was performed based on an autosomal recessive (MAF <0.1%) and autosomal dominant (*de novo* variants, MAF <0.01%) inheritance model.

**Case 8**

Genomic DNA was isolated from peripheral blood samples collected from the proband and both biological parents. Whole genome sequencing (WGS) was conducted to a mean depth of 35X with >80% of bases covered at 20X. WGS was performed on Illumina HiSeqXs. Reads were aligned and variants called according to standard protocols. Using filters related to call quality, allele frequency, and impact predictions, we searched for rare, damaging *de novo* variation or inherited X-linked, recessive or compound heterozygous variation in affected probands. WGS was carried out under a research protocol and not completed within a CAP/CLIA laboratory. Variants found to be medically relevant and returnable were validated by Sanger sequencing in an independent CLIA laboratory (Emory Genetics Laboratory) before being returned to participants.

**Case 9**

Clinical whole exome Sequencing was performed by the Department of Laboratory Medicine and Pathology at Mayo Clinic in Rochester, Minnesota, USA. Genomic DNA was extracted from blood from the proband, biological mother, and biological father; 97% of the exome was covered at a read depth

of 20X or greater. The exome was captured utilizing a custom reagent developed by Mayo Clinic and Agilent Technologies, targeting 19,456 genes and 187,715 exons using 637,923 probes to capture a 54.1Mbp total region. Sequencing was performed on an Illumina HiSeq 2500 Next Generation sequencing instrument, using HapMap Sample NA12878 as an internal control. Paired-end 101 base-pair reads were aligned to a modified human reference genome (GRCh37/hg19) using Novoalign (Novocraft Technologies, Malaysia). Sequencing quality was evaluated using FastQC ([www.bioinformatics.babraham.ac.uk/projects/fastqc/](www.bioinformatics.babraham.ac.uk/projects/fastqc/)). All germline variants were jointly called through GATK Haplotype Caller and GenotypeGVCF[5]. Each variant was annotated using the BioR Toolkit[6] and subsequently evaluated for clinical relevance. Targeted Sanger sequencing was used to confirm the presence of the variant in the proband and absence in in each parent.

## Case 12

Exomes were enriched using the SureSelectXT Clinical Research Exome V2 (Agilent, elid S30409818, genome build GRCh37) and sequenced on a Illumina Novaseq 6000. The sequencing data was processed with an in-house developed pipeline, DxNextflowWES v1.0.1, based on the Genome Analysis Toolkit (GATK v3.8-1-0-gf15c1c3ef) best practices guidelines (http://doi.org/10.5281/zenodo.4552029)[5]. The read pairs were mapped with BWA-MEM v0.7.17 (*arXiv:1303.3997*), marking duplicates and merging lanes using Sambamba v0.7.0[7] and realigning indels using GATK IndelRealigner. GATK Haplotypecaller was used to call single nucleotide polymorphisms and indels, creating variant call formatted (vcf) files. In a first step an intellectual disability gene panel analysis, comprising 981 genes (genes available on request), was performed. Variants were filtered using a population allele frequency of 0.05% or 0.5% (gnomAD database[8]) for the dominant or recessive inheritance model, respectively. In a second step, often performed when the first step did not result in a diagnosis, the whole exome was analyzed, where variant filtering for a *de novo* or recessive inheritance model was done as previously described[9]. All *de novo* variants were analyzed, but variants in putative recessive genes were only analyzed if both alleles showed a potential deleterious effect on the coding regions. Larger deletions/duplications, missense, synonymous, and intronic variants affecting protein function of other genes cannot be excluded. Variants that matched predefined and validated quality criteria were not validated by Sanger sequencing.

## Case 13

Whole genome sequencing was performed in the DEFIDIAG study supported by y The French Minstry of Health in the framework of the Plan France Médecine initiative 2025. Briefly, trio whole genome sequencing was performed using the optimized CNRGH GS protocol onto a Illumina Novaseq 6000 in order to reach a mean coverage of 30X for each sample. The sequencing data was processed with an in-house developed pipeline: the sequences are aligned to the Human reference genome GRCh37 using the Burrows-Wheeler Aligner BWA and Sambamba software. A local realignment of the sequences around insertion and deletion sites and the base quality recalibration is performed using GATK. After sequence quality control and alignment of the reference genome, the CNRGH performs the variant calling on the entire genome for the Single Nucleotide Variants (SNV), small insertion/deletions (indels) and structural variants (including Copy Number Variant, CNV). SNV and indel calling are performed using the Haplotype Caller from the GATK software in "bp resolution" mode to produce gVCF files. Structural variations are called by combining a reads depth analysis and pairing abnormalities/split reads method using CANVAS and MANTA software. VCF files are then collected by the IMAGINE Polyweb platform: additional combined TRIO gVCF analysis (genotypeGvcf) and CNV Wisecondor analysis was also be performed. The quality of the GS is checked and must be over 25X, sex and trio concordance are checked before sequence analysis. Functional annotations of the called

variants will then be subsequently performed by the IMAGINE Polyweb platform. All detected variations (SNV and structural variations) and their corresponding genes (if applicable) are annotated following the guidelines from the Human Genome Variation Society (HGVS, http://varnomen.hgvs.org/) and the Gene Nomenclature Committee (HGNC) from the HUman Genome Organization (HUGO, (http://www.genenames.org/). Annotations also integrate genetic and predictive information including, amongst others, functional impacts, putative effects in protein coding regions, OMIM information (such as the inheritance mode), phenotypic features as well as reference population frequencies (*i.e.*, gnomAD, DGV) described using the reference coordinates systems (*e.g.,* genome, cDNA and protein), the *'de novo'* or 'inherited' status.

**Supplementary data 2: METHODS Functional studies of *CTR9* variants**

**Fibroblast culture**

Skin fibroblasts of patients 1, 5, 7 and 11 and four unrelated healthy control individuals were cultured in RPMI medium supplemented with 15% fetal bovine serum, 1% L-glutamine, 1% sodium pyruvate, 1% Penicillin/Streptomycin, and 0.1% Primocin.

**H3K4 and H3K36 methylation analysis**

Skin fibroblasts in culture were harvested using TrypLE Express, after which total protein was isolated using RIPA+ buffer (10 mL RIPA, protease and phosphatase inhibitors, 10 uL benzoase). The Pierce BCA Protein Assay kit was used to measure the protein concentrations. The lysates were subsequently incubated with NuPage Reducing agent and NuPage LDS sample buffer for 10 min at 70°C. Equal amounts of protein were loaded onto a Bis-Tris 4-12 % mini gel, together with a PageRuler Plus Prestained Protein Ladder. After gel electrophoresis, proteins were transferred to a polyvinylidene difluoride (PVDF) membrane (Millipore) and blocked in 5% milk (Nestlé, dissolved in TBST buffer) for 1 (anti-H3K36 antibodies) or 2 hours. Primary antibody incubation for the methylated histon markers of interest (anti-H3K4me1 (ab8895, Abcam), 1:500; anti-H3K4me2 (ab32356, Abcam), 1:2000; anti-H3K4me3 (ab8580, Abcam), 1:1000; anti-H3K36me1 (ab9048, Abcam), 1:500; anti-H3K36me2 (ab9049, Abcam), 1:1000; anti-H3K36me3 (ab9050, Abcam), 1:1000) and loading control (anti-β-actin (ab8227, Abcam), 1:1000) was done overnight at 4°C. Next, the membranes were washed with TBST and incubated with a secondary antibody for 2 hours at RT (Goat anti-rabbit IgG (1706515, BioRad), 1:1000-1:10000). Pierce ECL Western Blotting reagent was used as a detection substrate and images were acquired using a LAS 400 mini imager (Cytiva). Analysis and quantification of the protein signals was performed with Image J.

**Mitochondrial membrane potential analysis**

Fibroblasts were seeded in 24-well culture plates at 40.000 cells per well (3 replicate wells per fibroblast line) on the day before the assay. To measure the mitochondrial membrane potential, cells were loaded with 5nM of the cell permeable dye tetramethylrhodamine ethyl ester perchlorate (TMRE; Enzo Life Sciences ENZ-52309). After 30 min of incubation, TMRE fluorescence intensity was measured using the IncuCyte S3 (Essen Bioscience) to perform automated microscopic imaging of 25 positions per well (20x magnification, phase contrast and red fluorescence channel) inside the incubator (37°C and 5% CO2). Images were analyzed in batch using the IncuCyte software. For each image, the area covered with cells was extracted after segmentation based on the phase contrast channel and, independently, the mitochondria were segmented using the red TMRE channel. Outlier image fields with a measured confluency level lower than 10% (too few cells) or higher than 95% (over-confluent) were removed from the analysis. The integrated TMRE intensity of the mitochondrial area divided by the area covered with cells was used as a measure for the mitochondrial membrane potential in each individual image. The experiment was performed 3 times. An initial control experiment with pharmacological depolarization of the mitochondrial membrane (20 μM FCCP, Carbonyl cyanide-4-(trifluoromethoxy)phenylhydrazone) confirmed complete loss of the TMRE signal.

**RNA extraction and RNA-sequencing**

Skin fibroblasts were collected with TrypLE Express, after which RNA was isolated with the RNeasy mini kit. RNA concentrations were determined with the Qubit RNA BR Assay Kit, whereas RNA integrity

numbers (RINs; ≥9.5) were defined with the Fragment Analyzer (Advanced analytics). Subsequent mRNA-sequencing was outsourced to Novogene (Cambridge, UK). Sequence capture involved the NEBNext® Ultra RNA Library Prep Kit and sequencing (9M reads) was done on a NovaSeq 6000 system (Illumina). The paired-end reads were preprocessed with trimmomatic 0.39 and aligned to GRCh38 build 102 ENSEMBL with STAR 2.7.5c. Gene expression quantification was done with featureCounts v2.0.1 and for differential gene expression analysis of the protein coding genes (patients versus controls), DESeq2 1.26 was used. Gene set enrichment analysis (GSEA) was performed with fGSEA 1.12.

**Supplementary data 3: Clinical description of patients harboring CTR9 variants**

**Case 1**

Proband 1 is the second son of unrelated parents, born by caesarean section after an uneventful pregnancy at 41 weeks and 4 days of gestation. The length at birth was 57 cm (+2,2 SD), weight 3830 g (+1,1 SD) and the head circumference 36 cm (+1,1 SD). He was diagnosed with eczema at the age of 7 months. A growth delay was noticed at the age of 12 months and at 21 months. Additional investigations were initiated as his weight remained under 9kg, despite eating well. He received the diagnosis of Non Celiac Gluten Sensitivity. Since the start of a strict gluten elimination diet, he has been growing well, but he remains under the 10th percentile (-2.1 SD) in length and around the 15th percentile (-2 SD) in weight. Growth hormone treatment was initiated at the age of 2, but ceased after 7 weeks due to the detection of a significantly dilated aorta and a bicuspid aortic valve. At the moment this is being treated with Atenolol and Irbesartan and followed up by echocardiogram 4 times a year and a yearly MRI. The Z-scores remain stable under this regime. Between the age of 2 and 3, there were severe constipation problems with multiple ER visits, but this resolved with the exclusion of eggs, cow's milk and soya from his diet. He has a mild global developmental delay. He could sit at the age of 17 months, started walking at 21 months and said his first words at 25 months. He receives occupational and speech therapy. At the age of 6 year and 10 months, he still has an unstable broad based gait and is unable to run fast. He can talk full sentences, but remains less fluid in his speech, compared to other children his age. His mother estimates his language development at the level of a 5 year old. He goes to the first year of primary school, but his reading level is lacking in comparison to his classmates. No neuropsychological or developmental assessments have been performed at the moment, but testing for learning difficulties has been planned in the near future.

**Case 2**

Proband 2 is the 3rd child of unrelated and healthy parents. She has 2 healthy sisters. The pregnancy was marked by hydramnios. She was born at 41 weeks of amenorrhea by cesarean section. She weighed 3.460 kg (+1 SD) for 52 cm (+0.5 SD) and a HC of 38 cm (+2 SD). She had hypotonia with feeding difficulties and mild stridor. She sat up around 1 year old and walked around 24 months old. She started to speak at the age of 2 years but she has a low lexical stock and difficulties in verbal comprehension. She has made progress. Growth was normal and the HC was +2 ds. She had chronic constipation. She had a dysmorphia with frontal bumps, a marked nose root, anteverted nostrils, a flat and long philtrum with a thin upper lip, a large mouth with fairly wide teeth, low-set, posteriorly rotated ears. She had balance problems that improved gradually. MRI, cardiac and abdominal ultrasound, karyotype and CGHarray were normal. Hearing was normal and she has astigmatism with hypermetropia. At the last examination at the age of 9 years, her height was 138 cm (+ 1 SD) for a weight of 34.700 kg (+ 1 SD) and her HC was 54.5 cm (+ 2 SD). There is significant hyperlaxity, especially in the elbows. There is a deformity of the feet with a varus of the forefoot. The clinical and neurological examination was otherwise unremarkable. She attends school normally but with assistance

**Case 3**

Proband 3, a boy, is the second child born at 40 weeks of gestation to unrelated parents of French origin after an uneventful pregnancy. Weight and length were at +1 SD at birth: 3650g and 52cm respectively, while head circumference was 52cm (+2 SD). A moderate hypotonia was noted since birth and he presented with feeding difficulties requiring a hospitalization in neonatal care. He was first seen in our center at age one month as he still had poor sucking and truncal hypotonia (no head control) with a normal peripheral tonus. Visual contact and smiling were of good quality. Extensive metabolic screening, CPK dosage, SNP-array, PTEN sequencing and Prader Willi methylation testing were found

to be normal at the cerebral MRI performed at age 6 months. Independent walking was acquired at age 22 months and the child was able to speak the first words at 12 months and spoke short sentences at age 28 months. A subaortic membrane was surgically resected at age 4 years. Height (101 cm) and weight (17 kgs) were normal (0 SD) while OFC was at +3 SD (54cm). Global motricity was impaired with in particular the impossibility of jumping with a bell foot and great difficulties to remain in a static way on only one foot. He also presented difficulties with pen holding and figure reproduction. He required specialized education owing to behavioral issue (tantrum, opposition behavior and hyperactivity disorder). Trio exome sequencing highlighted a heterozygous *CTR9* missense variant inherited from his 39 years old father who had a +3SD macrocephaly and a personal history of developmental delay (neonatal hypotonia, walking at age 22 months) which latter evolved favorably.

**Case 4**

Proband 4 was born after an uncomplicated pregnancy and delivery by caesarean section at 38 weeks of gestation. She had a good start. Her birth weight was 3575 gram. She was born with a hemangioma in the neck. In the neonatal period snapping and subluxation of the cheek was noted. Physiotherapy was started at the age of 1 month because of back arching and frequent crying. She was diagnosed with eczema, cow's milk allergy and asthmatic bronchitis. Her initial development was normal with sitting around the age of 6 months, walking at the age of 14 months and first words at the age of 15 months. Behavioral problems remained with frequent crying, fears (leading to frequent vomiting), contact difficulties and sensitivity to sounds. She started at normal primary school where she doubled the second year. In the fourth year she went to special primary school because she was overburdened. She was diagnosed with multiple complex developmental disorder (MCDD) and delayed gross and fine motor skills. She has difficulty adapting to new situations and needs a lot of structure. She also received feeding therapy because of eating problems (sensitivity to food structure). She started in normal secondary school, but after developing a psychosis after 6 months she went to special secondary school. Two years later she developed another psychosis after being overburdened for which she was admitted to a psychiatric ward for 3 months. Subsequently, she attended daycare activities and performed volunteer work. Neuropsychological testing at the age of 18 years (WAIS-IV) showed a disharmonic intelligence profile with a verbal comprehension 103, perceptual reasoning 75, working memory 74, and processing speed 86. Since the age of 23 years she lived in sheltered housing. She is treated with fluoxetine, haloperidol and oxazepam. During her life she had various medical problems. She had an adenotonsillectomy because of infections at the age of 2 years. She had a cholecystectomy because of cholecystitis and cholelithiasis at the age of 13 years. She has generalized joint hypermobility with rupture of the ankle ligament and subluxations of the hip, knee and cheek for which she is treated with physiotherapy and braces. She has problems in balance and coordination. She is easily fatigued, she sleeps more than 12 hours per day. Excessive transpiration is noted. She has pain in the back and the legs with transient complaints of muscle weakness. Neurologic investigation including an MRI of the lumbar spine and an EMG showed no neurological substrate. Because of constipation she is treated with laxatives. Ophthalmologic investigations because of vision problems showed intermittent esotropia of the right eye with anisometropia. Cardiologic screening at the age of 24 years showed no abnormalities. Physical examination at the age of 18 years revealed a height of 160 cm (-1.7 SD), a weight of 79.2 kg (+3 SD) and a head circumference of 56.8 cm (+1 SD). She has a high forehead with prominent eyebrows, attached ear lobules, full alae nasi, full lips, a high palate and a prominent chin. She had joint hypermobility with a Beighton score of 6 out of 9 and pes planus. She has fungal infections of 9 toe nails (since the age of 1 year) and a hemangioma in the neck. She is the second child of healthy non-consanguineous parents. She has an older sister who followed normal education. She has a positive family history for various psychiatric and learning problems in both the paternal and the maternal family. Joint hypermobility is also present in the sister and multiple maternal

relatives. She was previously tested for chromosomal aberrations using the Affymetrix CytoScan HD array and metabolic screening was performed which showed normal results.

**Case 5**

Proband 5, a male, was born after an uneventful pregnancy from non-consanguineous parents. There is a history of psychiatric disorders in some maternal relatives. In infancy, chewing difficulties were noted. The patient showed normal initial psychomotor development. He spoke his first words at 12 months. At 2 years of age, language regression and autistic symptoms were noted, rendering him practically non-verbal. Now, at 12 years old, he shows severe developmental delay, and uses less than 10 words to express himself, mostly for requests. He presents major behavioral issues, with frequent severe self-injury and aggression behaviors. Episodes of worsening of the behavioral symptoms are noted, when he does not eat and sleep. The patient required many hospitalizations related with the self-injury behaviors; unexplained signs of pain were noted recurrently, as well as a progressive loss of the use of his hands. Several psychotropic medications were attempted, but none was proven to be effective, with continuous aggravation of his behavior and impairment of his functioning level. The only pharmacological attempt that seemed to lead to temporary improvement was Bumetanide, that was associated with increased eye contact and decreased aggressions episodes and increased use of his hands (being for instance able to use a spoon to eat by himself). He was also able to retrieve a better use of language, with an increased vocabulary, and more frequent combinations of two words. The positive effect unfortunately only lasted for 6 months and the patient started to show a worsening of his functioning level despite continued medication, associated with lack of sleep, reduced communication and increased self-harm behaviors. His behavior, however, seems better controlled as compared to without Bumetanide. At age 12, he shows failure to thrive with a low BMI. The brain MRI was normal. He recently presented with paroxysmal episodes of ocular movements, with slow deviation of the two eyes (1-2 times/week). The EEG does not reveal any abnormality. Genetic investigations included a normal array-CGH.

**Case 6**

Proband 6 is a 20-year-old female seen for evaluation/ testing due to history of Autism Spectrum Disorder and Intellectual Disability. She was born full term following a pregnancy complicated by late onset preeclampsia. Genetic testing during pregnancy included amniocentesis for chromosome analysis, which was normal. She was born via induced vaginal delivery with a birth weight of 3033 gram. She was noted to have hepatosplenomegaly and hemolytic anemia shortly after birth and diagnosed with infantile pyknocytosis at 2 weeks old with resolution at 6 months old. She met early motor and verbal milestones on time (walked at 15 months, words at 12 months). However, she experienced developmental regression at ~18 months old, specifically decreased speech, and was diagnosed with autism spectrum disorder at ~23 months old. Due to this history she was evaluated at the genetics department of an outside institution at 3 years old and again at 6 years old with normal chromosomes (46,XX), negative subtelomere FISH, negative fragile X testing, negative *MECP2* testing and normal metabolic screening. At approximately 16 years old, the patient had further developmental regression with decreased speech (to point of mutism) and decreased physical activity noted over a 2-year period. At that time, she was evaluated by Neurology with normal brain imaging (MRI) and EEG. She was again seen by a geneticist at an outside institution at which time genetic testing included a 30 gene Autism spectrum disorder panel, which was negative, and chromosome microarray, which was abnormal with two copy number variants identified: a 46.13kb duplication of Xp22.33 (not maternally inherited, father not tested) and 4.67Mb deletion of 5q14.3 (inherited from healthy, neurotypical mother). On follow up in Genetics at 20 years old she had additional testing via an expanded Autism Spectrum disorder/ Intellectual disability panel, which identified the *de novo CTR9* variant.

**Case 7**

Proband 7 was born in the context of a quadruple pregnancy following IVF in Vietnam (mother from Vietnam, Swiss father). A homozygous brother died at 28 weeks of gestation in utero which was the motivation for a birth by caesarian section at 29 weeks of gestation. Two heterozygous sisters are in good health. At birth, the proband needed intubation and ventilation for some hours, then CPAP during a month. A first brain MRI at 3 months of age (at term) in Vietnam was described as normal. Psychomotor retardation with predominance of motor development and low weight were noted. He was described as a child that ate a lot, but did not gain weight. Brain MRI at the age of 16 months in Switzerland showed very small corpus callosum, ex vacuo dilatation of ventricles, heterogeneous maturation of white substance and normal spectroscopy. At 20 months, he had a global psychomotor retardation with motor predominance, axial and peripheral hypotonia, hyperlaxity of different joints (upper and lower extremities, hips). Independent walking was possible at 3 years. He spoke the first words at 2 years. At 3 years, he was able to say 10 words, showed a good comprehension in Vietnamese and French, but had expressive problems. The main complaint of the parents was that he had less force than his pairs, was rapidly tired and needed more sleep. Last measures at 3 years and 3 months of age showed a weight of 13.2 kg (-0,75 SD), a height of 98 cm (+0,51), a BMI of 13.8 kg/m2 (-1,5 SD, always very low due to fluctuating appetite, mother had to insist a lot on eating) and head circumference of 52 cm (+1 SD), with clinical appearance of relative macrocephaly). At the age of 3 years and 10 months, he frequented an international school. Educational staff observed intellectual disability, expression problems, unusual fatigability and need of more sleep, small muscle volume and very good social contacts. Besides the *CTR9 de novo* variant, an additional *de novo AFF4* variant was identified.

**Case 8**

Female proband presented at age 13 with seizures, mild intellectual disability, epilepsy and generalized muscle weakness. She was uncoordinated, easily fatigued, and had an abnormal voice. She had experienced gross motor delays since infancy. She started walking at the age of 18 months. Previous genetic testing included mtDNA sequencing and depletion, which was negative; microarray testing revealed a 476 kb duplication at chr6q26, which was classified as a variant of unknown significance (VUS).

**Case 9**

Proband 9 is a now 5-year-old male with normal prenatal and birth history. He was hospitalized at 7 days of life for cyanotic spells and required nasogastric feeds; he was discharged at 1 month of age on thickened feeds. At 24 months of age, he began wearing glasses for bilateral esotropia and astigmatism, optic nerves appears normal by exam. Regarding developmental history, patient rolled at 3m, crawled at 12m, and walked at 16m. His first words were between 18-24m and at 2.5y he was not yet speaking in sentences. More recently, he is reported to have regression of language skills. He was saying a couple of words before age 2, but he is reported to be non-verbal now at age 5. He receives physical, occupational and speech therapies. He is reported to have hand-flapping with excitement but no suspicion for seizures at this time. He has reduced sleep. Upon exam at 31 months of age, patient was at 0 SD for weight, +0.5 SD for height and -1,5 SD for head circumference. Patient was noted to have some facial dysmorphism including thick epicanthal folds, bulbous nose with hypoplastic tip and mild retrognathia. Brain MRI and EEG have not been performed. Audiology evaluation was normal. Chromosomal microarray, Angelman syndrome methylation and Rett/Angelman/developmental disorders gene panel were normal. Family history is non-contributory.

**Case 10**

Proband 10 is a 23-year-old woman with mild cognitive disability, ADHD, mild microcephaly (10%), mood disorder, dystonia, spasticity and tremor. She was born "prematurely" via vaginal delivery, following preeclampsia, and weighed 2270g at birth. There is a possible alcohol prenatal exposure. The biological father has ADHD. The proband started rolling over at 5-6 months, sitting up at 11-12 months, walking at 15-16 months. She said "dada" at 18 months and "mama" at 18 months, was using simple sentences by 24 months. She required early intervention since 9 months, received occupation and physical therapy with noted improvement. She completed high school with supports. She then started community college, and then withdrew. The WISC-IV from 2009 showed verbal comprehension score of 89, perceptual reasoning score of 79, working memory score of 94, processing speed of 65, and Full Scale IQ of 77. She was diagnosed with Cognitive Disorder, Pervasive Developmental Disorder - Not Otherwise Specified, Post-traumatic stress disorder (PTSD), and Reactive Attachment Disorder. The EKG and echocardiogram are normal. She receives foot orthotics. There is a history of ovarian cysts and obstructive sleep apnea. Numerous metabolic tests were within normal limits. Karyotype, array-CGH and *FMR1* analyses were normal. WES identified a variant in the *CTR9* gene. Her biological mother did not have this variant ; paternal sample was not available. Initial WES + mitochondrial DNA results showed a VUS, m.1355G>A, in the MT-RNR1 gene at approximately 26% heteroplasmy. The unaffected mother was found to carry the same variant at a greater level of heteroplasmy (40%).

**Case 11**

Proband 11 is a 2 year 9 month old boy born to non-consanguineous parents of Northern European/ Ashkenazi Jewish descent. Pregnancy was complicated by increased nuchal translucency. Fetal echocardiogram was normal. Patient was born at 39 weeks 1 day gestation to a primigravida mother by caesarian section due to breech position. He weighed 3.1 kg (-0,6 SD) and was 50.8cm (-0.5 SD) long at birth. Head circumference was 36.5cm (+0.3 SD) at birth. There were no complications reported after birth. He had difficulty gaining weight due to poor latch in the first few weeks of life. He also had a history of constipation. Hip ultrasound due to breech position was normal. Patient was noted to have hypotonia, weakness, and joint laxity at 15 months. Head ultrasound was normal. Creatine kinase was normal. At 29 months, cardiology evaluation was normal. Still's murmur was noted. Development became a concern at 12 months of age when he was unable to go from a laying position to a sitting position on his own. He started to roll at 2.5 months, sit when placed at 6 months, cruise at 11 months when placed in a standing position, pulled to stand at 14 months, and started to walk independently at 19 months. The proband started to babble at about 10/11 months and first words came at 26 months. At 29 months, he had about 5-8 words. He was receiving developmental therapies. The proband was first evaluated by clinical genetics at 17 months due to concerns of developmental delay. Clinical evaluation was significant for broad tip of nose, deep set eyes, prominent forehead, upslanted palpebral fissures, pronated feet, 2nd toe overlapping 1st and 3rd toes bilaterally, bilateral clinodactyly of the 5th toe, and normal size calves. Head circumference was 50cm (+1,5 SD). Weight was at -1 SD and height was at the -0,5 SD. Previous genetic testing included normal karyotype from chorionic villus sampling (CVS). Microarray from CVS identified a maternally inherited 734 kb interstitial deletion of 15q24.2q24.3. The deletion included three OMIM genes (*ETFA*, *ISL2*, and *SCAPER*). The patient's father had additional screening for *ETFA*-related disorder, but no additional variant was identified. A Noonan syndrome panel (9 genes) from CVS was negative. Postnatal testing for spinal muscular atrophy was negative.

**Case 12**

Proband 12 is the second child of healthy, non-consanguineous Caucasian parents. She was born at term after an uneventful pregnancy and birth. There were no evident problems in the first years of life. Age at walking was 16 months. Speech development was normal, although slower compared to her sister. Learning problems were first noted in first grade of elementary school. She had difficulties with learning to write and read. Her IQ is 93, which is discordant with parents who both have a master degree. Her social emotional development is delayed. There are no hearing or vision problems. Cardiac examination was normal. At the age of 7 years she was referred to our outpatient clinic because of her learning problems and mild facial dysmorphic features. Her height was 120.1 cm (-1.6 SD), weight 24.9 kg (-0.1 SD) and head circumference 53.8 cm (+1.4 SD). She had mild facial dysmorphic features such as epicanthal folds, hypertelorism, a broad and flat nasal bridge, a long and flat philtrum and thin upper lip. There was a mild clinodactyly of the 5th finger of both hands. She had flat feet and a sandal gap at both feet. Hirsutism on both forearms was noticed. Clinical genetic testing included a SNP-array (normal) followed by whole exome sequencing of the family trio. A *de novo* missense variant in *CTR9* was identified. No other relevant variants were reported.

**Case 13**

Proband 13 is the 3rd child of a consanguineous couple, first cousins of Turkish origin. His 12-year-old brother and 9-year-old sister are both in good health and have normal psychomotor development. The mother had, before the birth of her first child, two spontaneous miscarriages before three months of pregnancy and three fetal deaths in utero between 6 and 8 months of pregnancy, for which no etiology has been found. The pregnancy for the proband was considered as normal; the antenatal ultrasounds didn't show any anomalies. The child was born by Caesarean section at 34 weeks gestation due to the history of fetal deaths in utero of the siblings. The weight was 2350 gram (+1 SD), the height 43.5cm (-0.5 SD) and the head circumference 33.5 cm (+1,2 SD). During the first months of life, the child presented difficulties for breastfeeding. At four months, he was not able to either lift the head and shoulders or lean on the forearms. He benefited from physiotherapy at the age of 9 months. He was able to sit up at the age of 11 months and acquired walking at 22 months. The child also exhibited delayed language development, with few words repeated at the age of 20 months and short sentences of 3-4 words at the age of 2.5 years. Concerning the behavior, the child presented delayed social-emotional development, needs rituals in the current life and is described as inhibited and anxious regarding adults or new situations. The schooling is difficult with few interactions with the other children and difficulties in concentration and attention. Currently, the child benefits from speech therapy and psychomotricity. Two neuropsychological evaluations were realized: one in December 2018 showing an IQ of 62, and another in January 2020 showing an IQ of 55. Medically, the child presents with a physiological hyperopia; the audition was tested and was normal. He had recurrent ENT infections and underwent removal of adenoids and tonsils. There was no brain imaging or electroencephalographic recording. Array CGH found no chromosomal imbalance and analysis of the *FMR1* gene revealed one allele of normal size. On clinical examination at the age of 6 years, the child weighs 29 kg (+3SD), for 121cm (+1SD) and HC at 52.5 cm (+0.5SD). There are some facial peculiarities with almond-shaped eyes, a synophris and a small mouth. We also observe bilateral clinodactyly of the 5th digits and inverted nipples. Examination of the lower limbs showed some stiffness. Parents describe pain in his right knee, appearing every two weeks since he was 18 months old, with normal x-rays.

**Supplementary data 4: RESULTS Functional studies of *CTR9* variants**

**H3K4/H3K36 methylation analysis, mitochondrial quality assessment and RNA-sequencing in fibroblasts does not provide conclusive evidence for downstream functional consequences of the likely  the disease-causing variants**

Skin fibroblasts of four patients (1 (p.(Glu15Lys), 5 p.(Cys85Tyr), 7 p.(Glu376Lys) and 11 (p.(Pro25Arg)) and four healthy controls were available for functional studies. First, the effect of the identified *CTR9* variants on histone methylation was investigated. PAF1C mediates H3K4 and H3K36 methylation through SET1 and SET2, respectively[10-12]. Additionally, deletion of Paf1 in yeast was shown to drastically reduce H3K4 di- and trimethylation, while H3K4 monomethylation remained unaffected[13]. Western blotting of the available patient and control fibroblast samples failed to demonstrate a consistent difference in H3K4 or H3K36 methylation (Supplementary Figure 6). Highly variable methylation levels between individuals, and even between different protein extractions of the same fibroblast line (data not shown), were observed, possibly obscuring subtle differences. However, no significant effects were detected by our analysis.

Second, we examined the impact of the identified *CTR9* variants on mitophagy. It was recently demonstrated that Paf1C also activates mitophagy by regulation of the mitochondrial-anchored ATG32-mediated cargo-recognition system[14]. It binds to the ATG32 promotor in glucose-rich conditions, leading to low levels of mitophagy. Upon glucose starvation, PAF1C dissociates from the promotor, causing ATG32 upregulation and increased mitophagy. In yeast, deletion of *Ctr9* as such was also demonstrated to dysregulate mitophagy[14]. Interestingly, emerging evidence suggests a role of mitophagy in neurodevelopmental disorders in general, including autism spectrum disorders[15]. Of specific interest is the observation that pathogenic variants in *UBE2A*, another gene regulated by PAF1C[12], lead to a neurodevelopmental disorder that is molecularly characterized by impaired mitophagy. Microscopic TMRE analysis of the available fibroblasts, however, did not show evidence for increased accumulation of depolarized mitochondria in patients as compared to controls (Supplementary Figure 7). On the contrary, a trend towards more active mitochondria was observed.

Third, as no convincing pathomechanistic insights were gained from the previous two experiments, we performed mRNA-sequencing in the available skin fibroblasts to find novel leads on the downstream functional consequences of neurodevelopmental disease-causing *CTR9* variants. Differential gene expression analysis revealed 75 up- and 40 down-regulated genes ($p_{adj}<0.05$) in patient versus control fibroblasts (Supplementary Table 3). GSEA did not identify any group of genes to be statistically significant over-represented for a specific biological process after multiple testing correction. However, the normalized enrichment score was slightly enriched for a GO term related to positive regulation of locomotion (p<0.04, NES=-1.55), cellular response to lipid (p<0.15, NES=1.64) and cation transmembrane transport (p<0.02, NES=1.61) (Supplementary figure 8).

| | p.Glu15Lys | p.Ile17Thr | p.Pro25Arg | p.Glu26Gln | p.Glu37Gln | p.Cys85Tyr |
|---|---|---|---|---|---|---|
| *H. sapiens* | DTDEMIELD | | DQLPEGDE | | LKQEHTQ | QMTCLDT |
| *P. troglodytes* | DTDEMIELD | | DQLPEGDE | | LKQEHTQ | QMTCLDT |
| *B. taurus* | DTDEMIELD | | DQLPEGDE | | LKQEHTQ | QMTCLDT |
| *C. familiaris* | DTDEMIELD | | DQLPEGDE | | LKQEHTQ | QMTCLDT |
| *M. musculus* | DTDEMIELD | | DQLPEGDE | | LKQEHTQ | QMTCLDT |
| *X. tropicalis* | DTDEMIELD | | DQLPEGDE | | LKQEHTQ | QMTCLDT |
| *D. rerio* | DTDEMIELD | | DQLPEGDE | | LKQEHTQ | QMTCLDT |
| *D. melanogaster* | DTDEMIEVD | | DQLPDCPE | | LKQERAP | LMRALDM |

| | p.Glu376Lys | p.Asn455Ser | p.Glu469Lys | p.Thr766Ala | p.Arg878Gln |
|---|---|---|---|---|---|
| *H. sapiens* | NNYETMK | EILNNVG | NLGEAKK | RLATSVL | LEQRAQY |
| *P. troglodytes* | NNYETMK | EILNNVG | NLGEAKK | RLATSVL | LEQRAQY |
| *B. taurus* | NNYETMK | EILNNVG | NLGEAKK | RLATSVL | LEQRAQY |
| *C. familiaris* | NNYETMK | EILNNVG | NLGEAKK | RLATSVL | LEQRAQY |
| *M. musculus* | NNYETMK | EILNNVG | NLGEAKK | RLATSVL | LEQRAQY |
| *X. tropicalis* | NNYETMK | EILNNVG | NLGEAKK | RLATLVL | LEQRAQY |
| *D. rerio* | NNYETMK | EILNNLG | NLGEAKK | RLATLVL | LEQRAQY |
| *D. melanogaster* | GNYETMK | EIQNNVA | NLKMAKL | RLAMAIL | LAKRQEY |

**Supplementary Figure 1: Conservation.** Representation of the conservation of the different *CTR9* variants identified in this study (Respective species Uniprot codes Q6PD62, K7D8C0, F1N4V2, E2R0D1, Q62018, Q6DU9, A3KDM3, Q7K0X3).
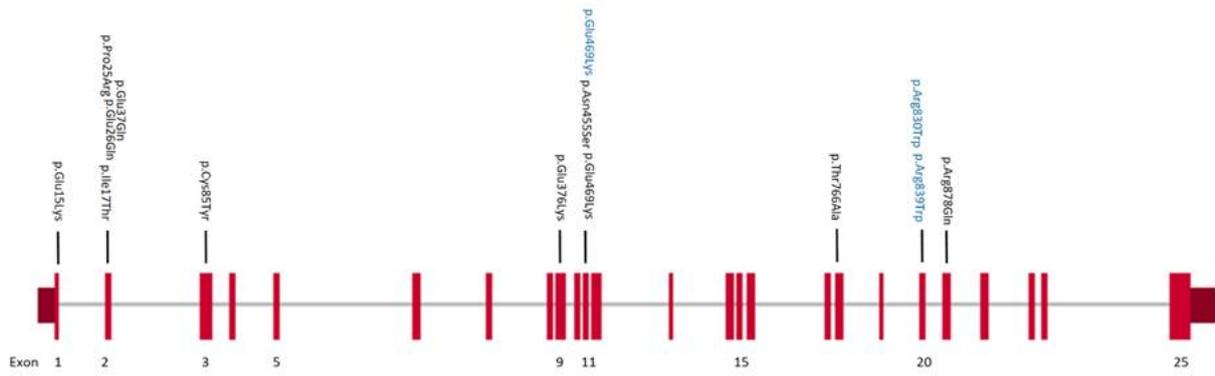
**Supplementary Figure 2. Variants at genomic level.** Representation of the 25 exons of *CTR9* (RefSeq accession number NM_014633). Previously reported missense variants are indicated in blue[16-18]; variants reported in this paper are presented in black.

**Supplementary information and Figure 3, 4 and 5: Modelling of individual *CTR9* variants**

The amino acid residues Glu15, Ile17, Pro25 Glu26, Glu37, Cys85, Glu376 and Asn455 are located in the N-terminal half of CTR9, near or within the 16 tetratricopeptide repeats (TPRs) region that spans positions 41 to 750 (Fig.1N).

The Glu15 is not involved in the formation of intramolecular interactions and the substitution to a lysine is predicted to only slightly increase protein stability (Supplementary table 2). The variant might still cause disease by influencing yet unknown protein interactions occurring at the N-terminus of CTR9 (Supplementary figure 5 C,D).

The residue Ile17 is located in the β-hairpin, with its side chain pointing towards a neighboring α-helix and interacting with the sidechains of Ile6, Ile8, Leu10, Leu19 and Ile33 (Supplementary figure 3C). Substitution to a more polar threonine residue is predicted to perturb these hydrophobic interactions, resulting in destabilization of the protein structure (Supplementary figure 3D).

Residue Pro25 locates to the loop between the N-terminal β-hairpin and the α-helix of CTR9. Its side chain is directed towards inside of the structure and is interacting with Leu24, Leu19 and Val30 residues (Supplementary figure 3A). The p.(Pro25Arg) variant is predicted to have destabilizing effect on the structure (Supplementary table 2), mainly due to the increase of the side chain entropy term (data not shown). Indeed, thanks to its unique structural characteristics the proline residue might rigidify the loop structure, favorably decreasing its entropy. Additionally, the larger Arg side chain might cause the steric hindrance with the neighboring residues (Supplementary figure 3B).

The residue Glu26 is located on the loop preceding the α-helix and its side chain is pointing towards the solvent (Supplementary figure 3G). The substitution of Glu26 to glutamine was predicted to decrease the stability of CT9R N-terminal part (Supplementary table 2, Supplementary figure 3H). Unfavorable electrostatic interactions of Glu26Gln are predicted with the dipole of the nearby helix.

On the contrary to residue 26, the similar change of residue Glu37Gln was predicted to interact favorably with the dipole of the preceding helix, potentially increasing the stability of the protein. Nevertheless, in the wild-type protein the residue Glu37 is forming a hydrogen bond with Arg11 backbone, from the loop of neighboring β-hairpin (Supplementary figure 3E), thus the p.(Glu37Gln) variant might perturb the local structure of CTR9 (Supplementary figure 3F). The local structure changes could potentially perturb the interactions with PAF1 protein with the N-terminus of CTR9.

The side chain of residue Cys85 is involved in the hydrogen bond interaction with the neighboring alpha helix backbone of the first tetratricopeptide repeat (TPR) of the CTR9 (Supplementary figure 3I). The substitution of the Cys side chain to a much larger Tyr will not only disrupt this interaction but also create a steric hindrance with the closest residues (Supplementary figure 3J), weakening the interactions between the helices and destabilizing a local structure of the domain that could influence the interactions with the PAF1 protein, predicted to have a dramatically destabilizing effect (Supplementary table 2).

Glu376 is located after the 7th TPR and in the vicinity to the PAF1 protein interaction interface (Supplementary figure 4A). Although the resolution of the structure does not allow to assess the influence of the p.(Glu376Lys) variant, we might assume that this variant is either directly perturbing the CTR9-PAF1 interaction, or influence it indirectly, by destabilization of the CTR9 α-helix via the replacement of the favorable electrostatic interactions with Lys379 by repulsive interactions (Supplementary figure 4B).

Asn455 is located on 9th TPR of CTR9 and according to 6TED.pdb cryo-EM structure it is in proximity of the PAF1 protein interaction site (Supplementary figure 4C). The p.(Asn455Ser) variant might perturb the CTR9-PAF1 interaction either directly or by destabilization of the local structure (Supplementary figure 4D).

Glu469 is located on 9th TPR of CTR9 protein. The Glu469 role might be to enhance the stability of the α-helix on the TPR9 by interactions with Lys472 and its substitution into Lys might perturb the local structure and influence the TPR interactions with the PAF1 protein (Supplementary figure 4E,F).

The residue Thr766 is part of the α-helix bundle, which precedes the long α-helix structure of CTR9 that is called "trestle", whereas Arg878 is located on that "trestle", which is assumed to play an important role in transcription elongation and histone modification[19]. Thr766 is located in the close vicinity to the Glu25 residue from PAF1 protein (Supplementary figure 4G). This is a first residue of PAF1 represented in the EM complex structure. Presumably, the substitution into smaller and more hydrophobic alanine perturbs the electrostatic/hydrogen bond interactions with PAF1 protein (Supplementary figure 4H). It is not excluded however, that the mutated residue might interact more strongly with the hydrophobic Leu23 and Pro24 of PAF1 that are not present in the current structure.

The Arg878 is making a hydrogen bond with the Glu32 residue of Pol II subunit E protein (Supplementary figure 5A). The substitution of Arg878 into the much shorter Gln might be detrimental for this interaction and in a consequence perturb the position of the CTR9 "trestle" in regard of Pol II complex (Supplementary figure 5B).
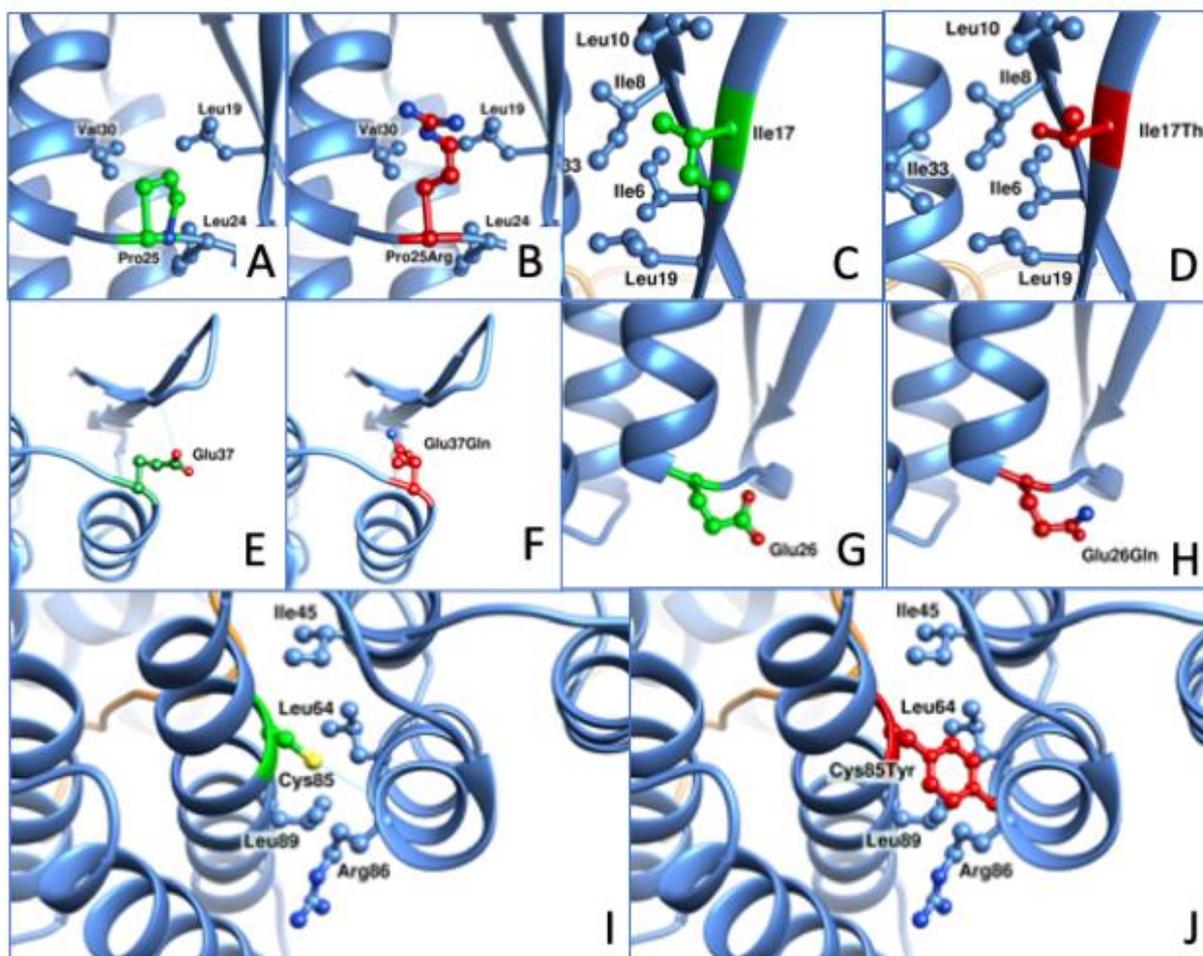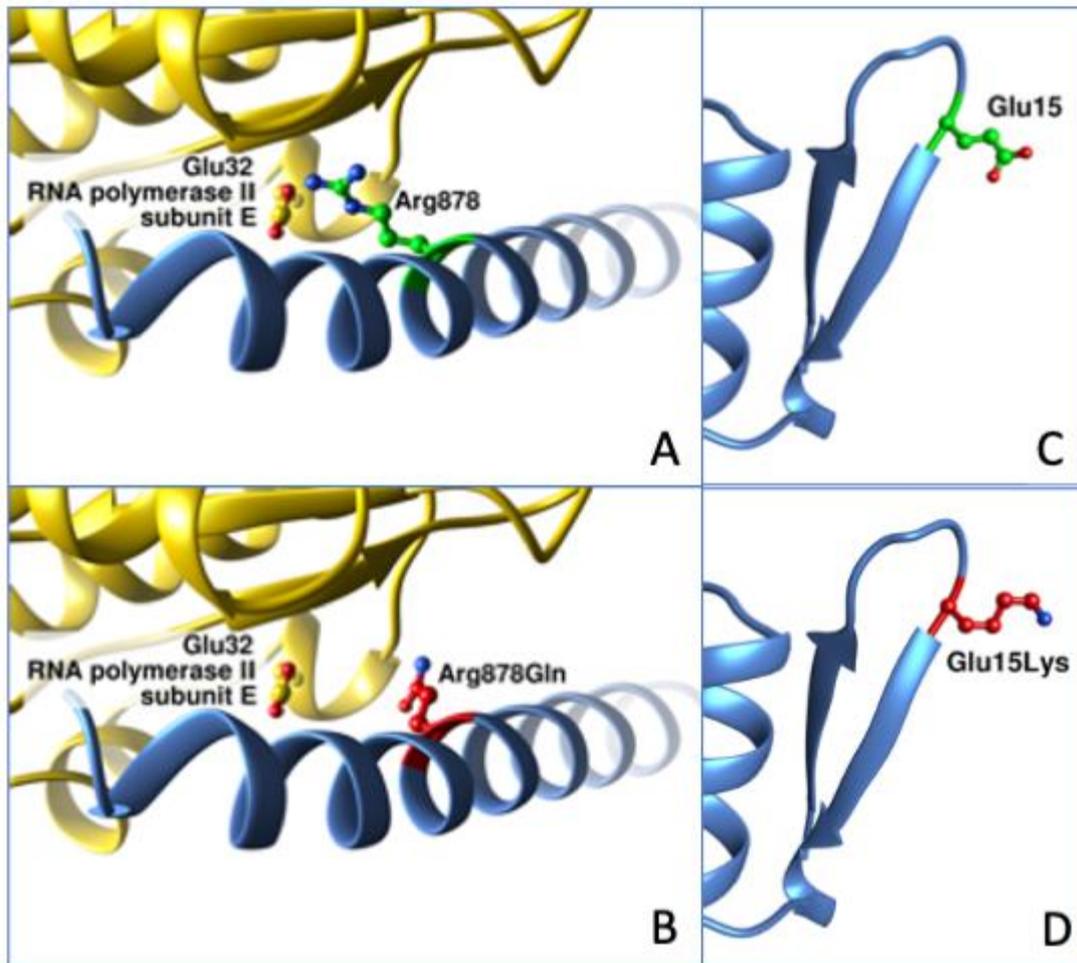
**Supplementary figure 3**. *In silico* **variant modelling, part A. (A, B)** Comparison of interactions of Pro25 and p.Pro25Arg mutant on CTR9 protein structure. The residue is visualized using crystal structure of N-terminus of CTR9 (PDB code 5ZYQ). CTR9 is pictured in blue and PAF1 protein is orange. The Pro25 residues is presented in green and p.Pro25Arg in red and ball and stick representation, the interacting partners in blue in ball and stick representation with the heteroatoms of analyzed residues colored: oxygen in red, nitrogen in blue and sulfur in yellow. **(C, D)** The position and interactions of Ile17 and Ile17Thr mutant on CTR9 protein structure. The residue is visualized using crystal structure of N-terminus of CTR9 (PDB code 5ZYQ). CTR9 is pictured in blue and PAF1 protein is orange. The Ile17 is presented in green and Ile17Thr in red and ball and stick representation, the interacting partners in blue in ball and stick representation with the heteroatoms of analyzed residues colored: oxygen in red, nitrogen in blue and sulfur in yellow. **(E, F)** The position and interaction of Glu37 and p.Glu37Gln side chains. The residues are visualized using crystal structure of the N-terminus of CTR9 (PDB code 5ZYQ). The CTR9 protein is colored in cornflower blue. The Glu37 is presented in green and p.Glu37Gln in red and ball and stick representation, the interacting partners in blue in ball and stick representation with the heteroatoms of analyzed residues colored: oxygen in red, nitrogen in blue and sulfur in yellow. **(G, H)** The position and interaction of Glu26 and p.Glu26Gln side chains. The residues are visualized using crystal structure of the N-terminus of CTR9 (PDB code 5ZYQ). The CTR9 protein is colored in cornflower blue. The Glu26 is presented in green and p.Glu26Gln in red and ball and stick representation, the interacting partners in blue in ball and stick representation with the heteroatoms of analyzed residues colored: oxygen in red, nitrogen in blue and sulfur in yellow. **(I, J)** Comparison between the interactions of the wild type Cys85 and mutated p.Cys85Tyr side chains. The residues are visualized using crystal structure of N-terminus of CTR9 (PDB code 5ZYQ). The CTR9 protein are colored in cornflower blue and

PAF1 protein in orange. The Cys85 is presented in green and p.Cys85Tyr in red and ball and stick representation, the interacting partners in blue in ball and stick representation with the heteroatoms of analyzed residues colored – oxygen in red, nitrogen in blue and sulfur in yellow.
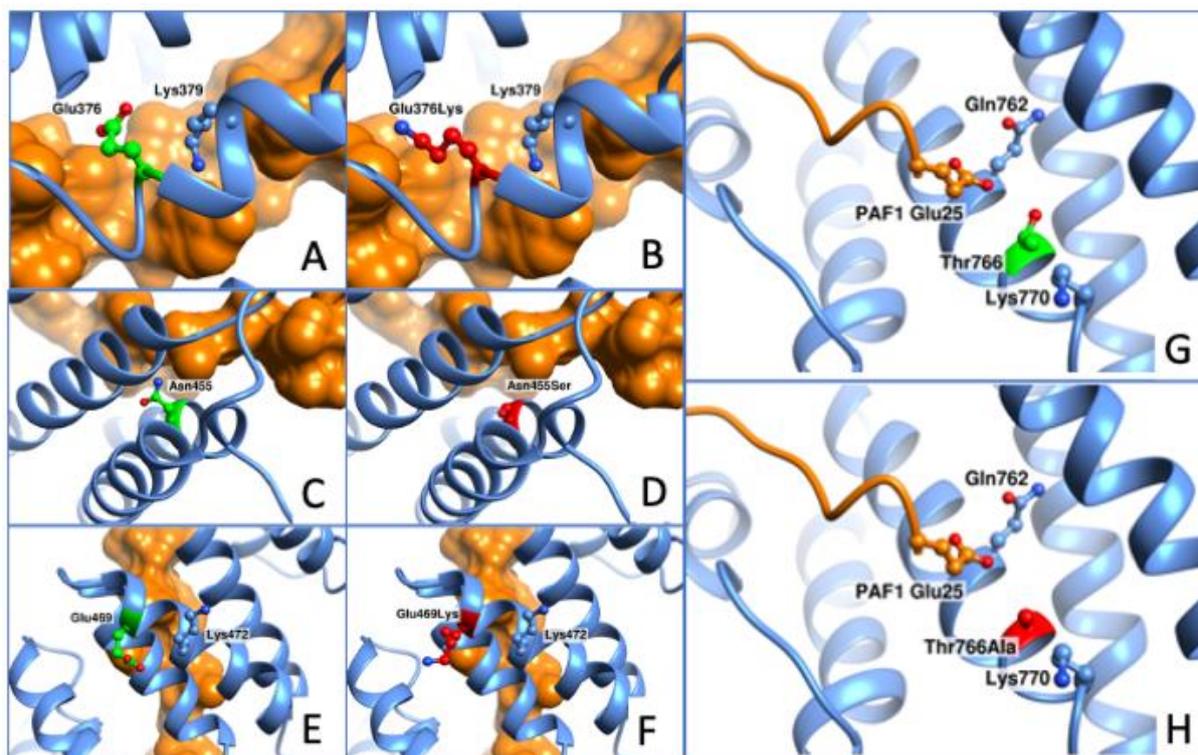
**Supplementary figure 4.** *In silico* **variant modelling, part B.** (**A, B**) Comparison between the interactions of the wild type Glu376 and mutated Glu376Lys side chains. The residues are visualized using EM structure of CTR9 (PDB code 6TED). The CTR9 protein is colored in cornflower blue and PAF1 protein in orange. The analyzed wt residues are presented in green and mutated in red in ball and stick representation, the interacting partners in blue in ball and stick representation with the heteroatoms of analyzed residues colored – oxygen in red, nitrogen in blue and sulfur in yellow. (**C, D**) Comparison between the interactions of the wild type Asn455 and mutated p.Asn455Ser side chains. The residues are visualized using EM structure of CTR9 (PDB code 6TED). The CTR9 protein is colored in cornflower blue and PAF1 protein in orange. The analyzed wt residue is presented in green and mutated in red in ball and stick representation, the interacting partners in blue in ball and stick representation with the heteroatoms of analyzed residues colored – oxygen in red, nitrogen in blue and sulfur in yellow. (**E, F**) Comparison between the interactions of the wild type Glu469 and mutated Glu469Lys side chains. The residues are visualized using EM structure of CTR9 (PDB code 6TED). The CTR9 protein is colored in cornflower blue and PAF1 protein in orange in the surface representation. The analyzed wt residues are presented in green and mutated in red in ball and stick representation, the interacting partners in blue in ball and stick representation with the heteroatoms of analyzed residues colored – oxygen in red, nitrogen in blue and sulfur in yellow. (**G, H**) Position and interactions of Thr766 and Thr766Ala. CTR9 protein is colored blue, protein PAF1 is colored orange. The wild type residue and substitution are visualized using EM structure of CTR9 (PDB code 6TED). The analyzed wt residues are presented in green and mutated in red and ball and stick representation, the interacting partners in ball and stick representation with the heteroatoms of analyzed residues colored – oxygen in red, nitrogen in blue and sulfur in yellow.

**Supplementary figure 5.** *In silico* **variant modelling, part C.** (**A, B**) The position and interactions of Arg878 residue and its p.Arg878Gln mutant. The CTR9 is in cornflower blue, the rest of Pol II complex in yellow. The wild type residue and substitution are visualized using EM structure of CTR9 (PDB code 6MGH). The analyzed wild type residues are presented in green and mutated in red and ball and stick representation, the interacting partners in ball and stick representation with the heteroatoms of analyzed residues colored – oxygen in red, nitrogen in blue and sulfur in yellow. (**C, D**) The positions of Glu15 and p.Glu15Lys on CTR9 protein structure. The residue is presented in green and ball and stick representation using crystal structure of N-terminus of CTR9 (PDB code 5ZYQ). CTR9 is presented in cornflower blue and PAF1 is presented in orange.
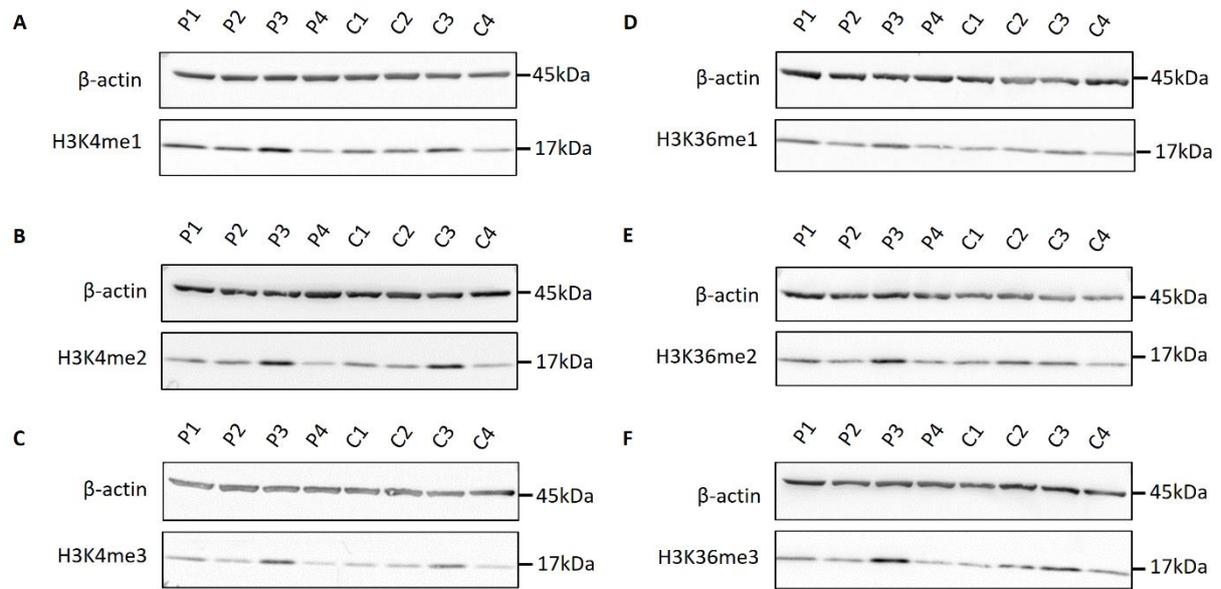
**Supplementary Figure 6**: **H3K4/H3K36 methylation analysis.** Western blotting was done to evaluate the effect of the identified CTR9 missense variants on **A-C**) H3K4 and **D-F**) H3K36 mono-, di- and trimethylation. B-actin was used as a loading control. No consistent differences were observed between patients (P1: case 11; P2: case 7; P3: case 1; P4: case 5) and controls.
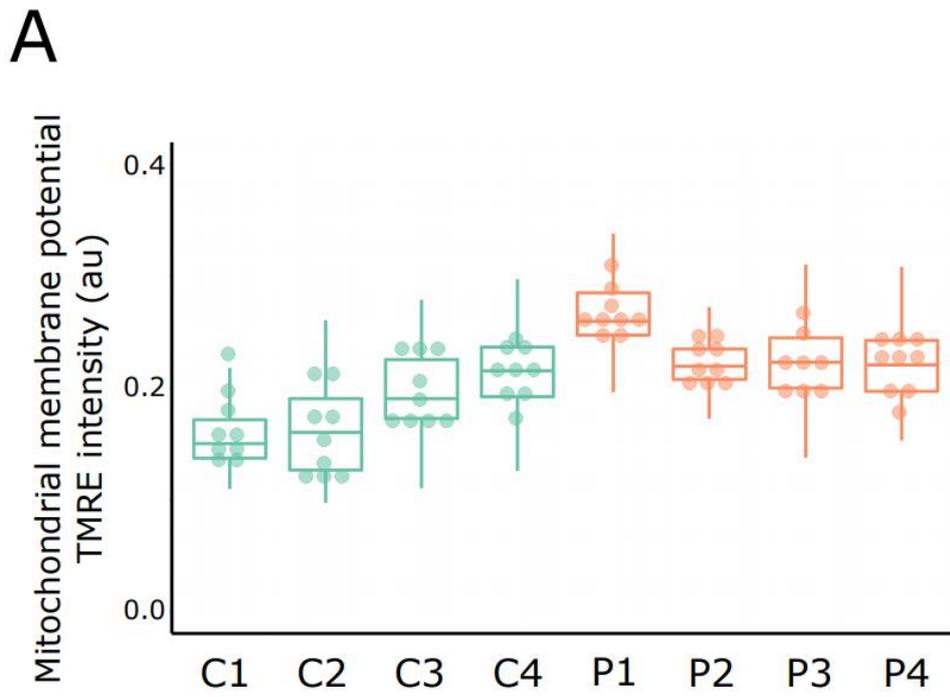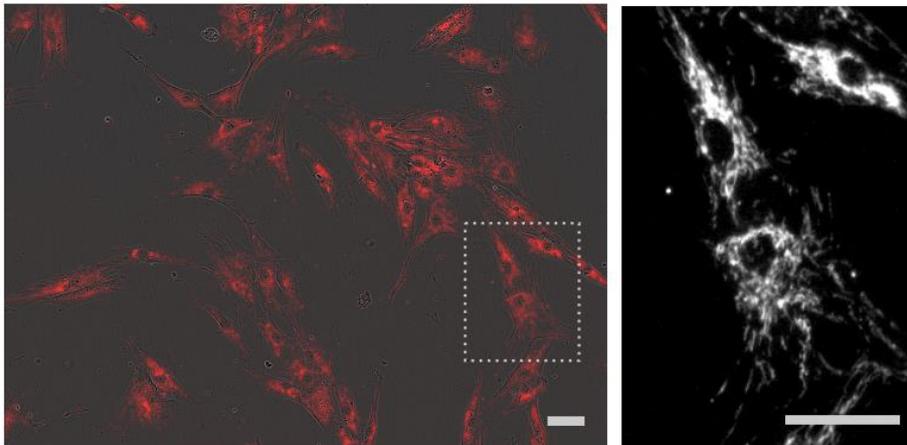
**Supplementary Figure 7. Mitochondrial membrane potential in patient and control fibroblasts.** Fibroblasts labeled with the fluorogenic cell permeable dye TMRE were microscopically imaged to measure the mitochondrial membrane potential. The integrated TMRE intensity corrected for the cell confluency was measured per image. (**A**) Pooled data from 3 independent experiments containing, for each fibroblast line, 3 replicate wells and 25 images per well. The boxplot represents the distribution of individual image measurements (225 images per fibroblast line), and individual plotted data points are the median values per well. (**B**) Example image used for quantification, showing the overlay of the TMRE fluorescence (red) and the phase contrast image (used to correct for confluency differences between image fields). The boxed zoom illustrates the specific mitochondrial localization of the TMRE signal. Scale bar = 50 µm.

**fGSEA BP**
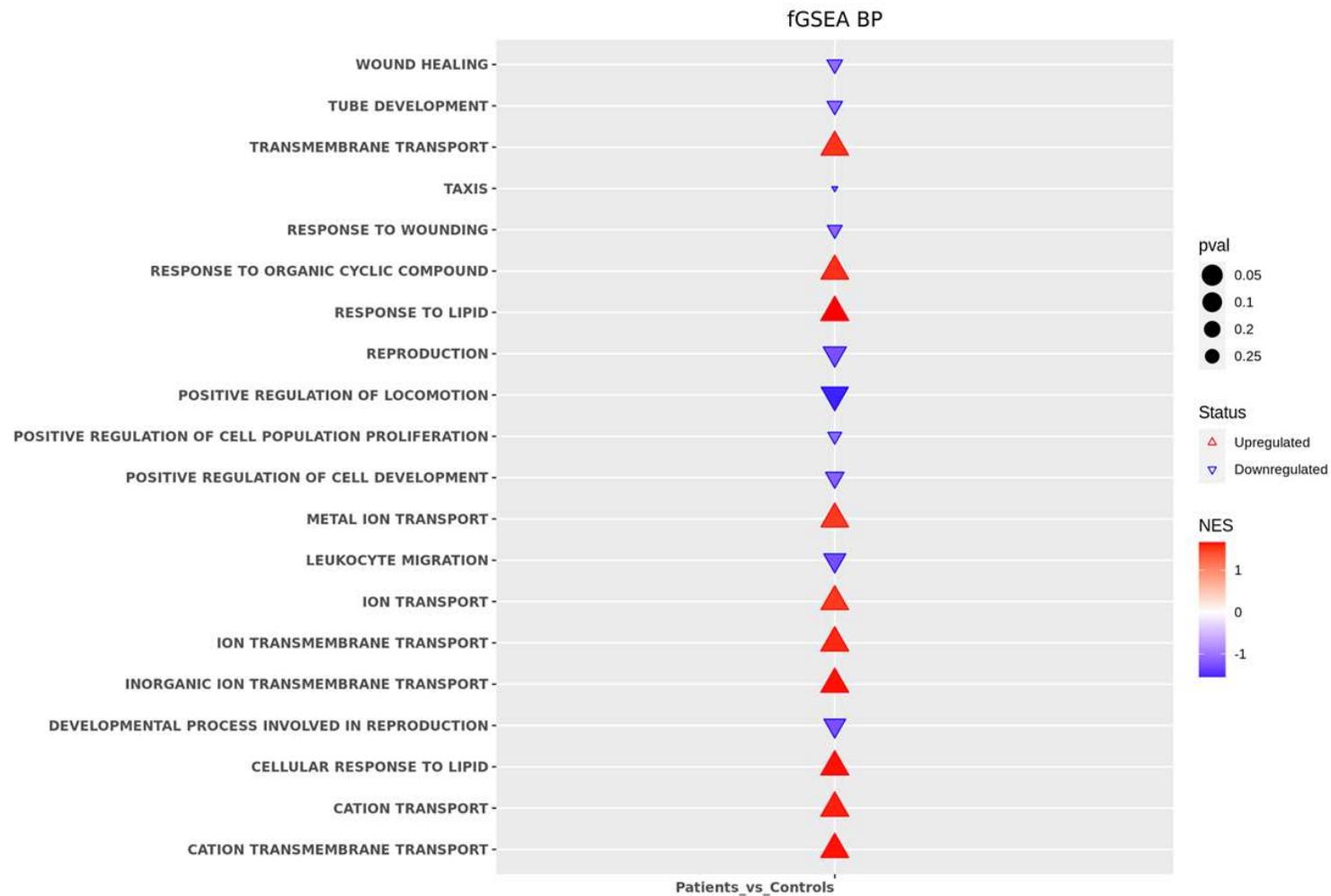
**Supplementary Figure 8: Gene set enrichment.** The genes having a log fold change greater than 0.1 (absolute value) were analyzed with fGSEA (R package). The top 20 enriched biological processes having a normalized enrichment score above 1 are shown together with their corresponding p- values (the p-values did not reach significance after multiple testing correction).

**Supplementary Table 1. Variant details regarding frequency in control populations, prediction scores, ACMG/ AMP classification**

| Cases | Variant (hg19, NM_014633.5) | GnomAd frequency | Bravo frequency | GERP | SIFT | PP2 HDIV | PP2 HVAR | MT | Primate AI | Pathogenic predictions from 21 prediction programs (Varsome) | CADD | ACMG/ AMP |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Case 1 | c.43G>A, p.(Glu15Lys) | Absent | Absent | 5,58 | Damaging 0.002 | Possibly damaging 0.933 | Possibly damaging 0.46 | Disease causing 1 | Damaging 0.9209 | 15/21 | 33 | LP (PS2, PM2, PP3) |
| Case 2 | c.109G>C, p.(Glu37Gln) | Absent | Absent | 5,82 | Damaging 0.002 | Damaging 1 | Damaging 0.992 | Disease causing 1 | Damaging 0.9232 | 18/21 | 27,6 | LP (PS2, PM2, PP3) |
| Case 3 | c.50T>C, p.(Ile17Thr) | Absent | Absent | 5,61 | Damaging 0.001 | Possibly damaging 0.998 | Possibly damaging 0.887 | Disease causing 1 | Damaging 0.9155 | 16/21 | 26,5 | VUS (PM2, PP3) |
| Case 4 | c.1405G>A, p.(Glu469Lys) | Absent | Absent | 5,88 | Tolerated 0.432 | Possibly damaging 0.515 | Benign 0.194 | Disease causing 1 | Damaging 0.9033 | 13/21 | 23,3 | LP (PS2, PM2, PP3) |
| Case 5 | c.254G>A, p.(Cys85Tyr) | Absent | Absent | 5,78 | Damaging 0.026 | Possibly damaging 0.918 | Benign 0.268 | Disease causing 1 | Damaging 0.9059 | 13/21 | 24,6 | LP (PS2, PM2, PP3) |
| Case 6 | c.76G>C, p.(Glu26Gln) | Absent | Absent | 5,82 | Damaging 0.004 | Possibly damaging 0.93 | Benign 0.379 | Disease causing 1 | Damaging 0.8664 | 12/21 | 23,4 | LP (PS2, PM2, PP3) |
| Case 7 | c.1126G>A, p.(Glu376Lys) | Absent | Absent | 5,62 | Damaging 0.004 | Damaging 1 | Damaging 0.943 | Disease causing 1 | Damaging 0.9333 | 16/21 | 28,2 | LP (PS2, PM2, PP3) |
| Case 8 | c.1126G>A, (p.Glu376Lys) | Absent | Absent | 5,62 | Damaging 0.004 | Damaging 1 | Damaging 0.943 | Disease causing 1 | Damaging 0.9333 | 16/21 | 28,2 | LP (PS2, PM2, PP3) |
| Case 9 | c.1364A>G, p.(Asn455Ser) | 1/ 251322 | Absent | 6,07 | Damaging 0.016 | Damaging 0.993 | Possibly damaging 0.863 | Disease causing 1 | Damaging 0.8251 | 18/21 | 25,6 | LP (PS2, PM2, PP3) |
| Case 10 | c.2296A>G, p.(Thr766Ala) | Absent | Absent | 6,02 | Tolerated 0.267 | Benign 0.021 | Benign 0.007 | Disease causing 1 | Tolerated 0.7394 | 5/21 | 22,7 | VUS (PM2) |
| Case 11 | c.74C>G, p.(Pro25Arg) | Absent | Absent | 5,82 | Damaging 0 | Damaging 1 | Damaging 0.99 | Disease causing 1 | Damaging 0.9159 | 16/21 | 31 | LP (PS2, PM2, PP3) |
| Case 12 | c.2633G>A, p.(Arg878Gln) | 1/ 249646 | Absent | 5,93 | Damaging 0.003 | Benign 0.052 | Benign 0.005 | Disease causing 1 | Damaging 0.8657 | 11/21 | 24,4 | LP (PS2, PM2, PP3) |
| Case 13 | c.1126G>A, p.(Glu376Lys) | Absent | Absent | 5,62 | Damaging 0.004 | Damaging 1 | Damaging 0.943 | Disease causing 1 | Damaging 0.9333 | 16/21 | 28,2 | LP (PS2, PM2, PP3) |

**LP**, likely pathogenic; **P**, pathogenic; **VUS**, variant of unknown significance, **PS,** pathogenic strong, **PM,** pathogenic moderate, **PP,** pathogenic supporting,

**Supplementary Table 2. The Foldx 5.0 predictions of the influence of variants on the protein stability of CTR9 N-terminal structure**

| *CTR9* variant | Mean of the difference in stability between the wt and the mutated protein | SD |
|---|---|---|
| p.(Glu15Lys) | -0.85 kcal/mol | 0.002 |
| p.(Ile17Thr) | 1.3 kcal/mol | 0.03 |
| p.(Pro25Arg) | 1.6 kcal/mol | 0.07 |
| p.(Glu26Gln) | 1.9 kcal/mol | 0.18 |
| p.(Glu37Gln) | -1.2 kcal/mol | 0.002 |
| p.(Cys85Tyr) | 17.3 kcal/mol | 2.5 |

The positive values indicate the decrease of stability upon substitution, so the variant is predicted to destabilize the structure, whereas the negative value of the energy indicate predicted stabilization of the structure.

**References**

1.    Vandeweyer G, Van Laer L, Loeys B, Van den Bulcke T, Kooy RF. VariantDB: a flexible annotation and filtering portal for next generation sequencing data. Genome medicine. 2014;6(10):74.
2.    Haer-Wigman L, van Zelst-Stams WA, Pfundt R, et al. Diagnostic exome sequencing in 266 Dutch patients with visual impairment. Eur J Hum Genet. 2017;25(5):591-599.
3.    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.
4.    Retterer K, Juusola J, Cho MT, et al. Clinical application of whole-exome sequencing across clinical indications. Genet Med. 2016;18(7):696-704.
5.    McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010;20(9):1297-1303.
6.    Kocher JP, Quest DJ, Duffy P, et al. The Biological Reference Repository (BioR): a rapid and flexible system for genomics annotation. Bioinformatics. 2014;30(13):1920-1922.
7.    Tarasov A, Vilella AJ, Cuppen E, Nijman IJ, Prins P. Sambamba: fast processing of NGS alignment formats. Bioinformatics. 2015;31(12):2032-2034.
8.    Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature. 2020;581(7809):434-443.
9.    Terhal PA, Vlaar JM, Middelkamp S, et al. Biallelic variants in POLR3GL cause endosteal hyperostosis and oligodontia. Eur J Hum Genet. 2020;28(1):31-39.
10.   Wood A, Schneider J, Dover J, Johnston M, Shilatifard A. The Paf1 complex is essential for histone monoubiquitination by the Rad6-Bre1 complex, which signals for histone methylation by COMPASS and Dot1p. J Biol Chem. 2003;278(37):34739-34742.
11.   Krogan NJ, Dover J, Wood A, et al. The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation. Mol Cell. 2003;11(3):721-729.
12.   Crisucci EM, Arndt KM. The Roles of the Paf1 Complex and Associated Histone Modifications in Regulating Gene Expression. Genet Res Int. 2011;2011.
13.   Xie Y, Zheng M, Chu X, et al. Paf1 and Ctr9 subcomplex formation is essential for Paf1 complex assembly and functional regulation. Nat Commun. 2018;9(1):3795.
14.   Zheng L, Shu WJ, Li YM, et al. The Paf1 complex transcriptionally regulates the mitochondrial-anchored protein Atg32 leading to activation of mitophagy. Autophagy. 2020;16(8):1366-1379.
15.   Wang YM, Qiu MY, Liu Q, Tang H, Gu HF. Critical role of dysfunctional mitochondria and defective mitophagy in autism spectrum disorders. Brain Res Bull. 2021;168:138-145.
16.   De Rubeis S, He X, Goldberg AP, et al. Synaptic, transcriptional and chromatin genes disrupted in autism. Nature. 2014;515(7526):209-215.
17.   Deciphering Developmental Disorders S. Prevalence and architecture of de novo mutations in developmental disorders. Nature. 2017;542(7642):433-438.
18.   Lelieveld SH, Reijnders MR, Pfundt R, et al. Meta-analysis of 2,104 trios provides support for 10 new genes for intellectual disability. Nat Neurosci. 2016;19(9):1194-1196.
19.   Vos SM, Farnung L, Boehning M, et al. Structure of activated transcription complex Pol II-DSIF-PAF-SPT6. Nature. 2018;560(7720):607-612.
20.   Singh T, Kurki MI, Curtis D, et al. Rare loss-of-function variants in SETD1A are associated with schizophrenia and developmental disorders. Nat Neurosci. 2016;19(4):571-577.
21.   Hiraide T, Nakashima M, Yamoto K, et al. De novo variants in SETD1B are associated with intellectual disability, epilepsy and autism. Hum Genet. 2018;137(1):95-104.
22.   Luscan A, Laurendeau I, Malan V, et al. Mutations in SETD2 cause a novel overgrowth condition. J Med Genet. 2014;51(8):512-517.

23.    Stevenson RE, Chudley AE, Srivastava AK, Rodriguez J, Friez MJ, Schwartz CE. UBE2A-related X-linked intellectual disability. Clin Dysmorphol. 2019;28(1):1-6.