Missense Variants in the Histone Acetyltransferase Complex Component Gene *TRRAP* Cause Autism and Syndromic Intellectual Disability

Benjamin Cogné,[1,2,80] Sophie Ehresmann,[3,80] Eliane Beauregard-Lacroix,[3,80] Justine Rousseau,[3] Thomas Besnard,[1,2] Thomas Garcia,[3] Slavé Petrovski,[4,5] Shiri Avni,[6] Kirsty McWalter,[7] Patrick R. Blackburn,[8,9] Stephan J. Sanders,[10] Kévin Uguen,[11,12] Jacqueline Harris,[13,14,15] Julie S. Cohen,[13,14] Moira Blyth,[16] Anna Lehman,[17] Jonathan Berg,[18] Mindy H. Li,[19] Usha Kini,[20] Shelagh Joss,[21] Charlotte von der Lippe,[22] Christopher T. Gordon,[23,24] Jennifer B. Humberson,[25] Laurie Robak,[26,27] Daryl A. Scott,[26,27,28] Vernon R. Sutton,[26,27,29] Cara M. Skraban,[30,31] Jennifer J. Johnston,[32] Annapurna Poduri,[33] Magnus Nordenskjöld,[34,35] Vandana Shashi,[36] Erica H. Gerkes,[37] Ernie M.H.F. Bongers,[38] Christian Gilissen,[38] Yuri A. Zarate,[39] Malin Kvarnung,[34,35] Kevin P. Lally,[40] Peggy A. Kulch,[41] Brina Daniels,[39] Andres Hernandez-Garcia,[26]

(Author list continued on next page)

Acetylation of the lysine residues in histones and other DNA-binding proteins plays a major role in regulation of eukaryotic gene expression. This process is controlled by histone acetyltransferases (HATs/KATs) found in multiprotein complexes that are recruited to chromatin by the scaffolding subunit transformation/transcription domain-associated protein (TRRAP). *TRRAP* is evolutionarily conserved and is among the top five genes intolerant to missense variation. Through an international collaboration, 17 distinct *de novo* or apparently *de novo* variants were identified in *TRRAP* in 24 individuals. A strong genotype-phenotype correlation was observed with two distinct clinical spectra. The first is a complex, multi-systemic syndrome associated with various malformations of the brain, heart, kidneys, and genitourinary system and characterized by a wide range of intellectual functioning; a number of affected individuals have intellectual disability (ID) and markedly impaired basic life functions. Individuals with this phenotype had missense variants clustering around the c.3127G>A p.(Ala1043Thr) variant identified in five individuals. The second spectrum manifested with autism spectrum disorder (ASD) and/or ID and epilepsy. Facial dysmorphism was seen in both groups and included upslanted palpebral fissures, epicanthus, telecanthus, a wide nasal bridge and ridge, a broad and smooth philtrum, and a thin upper lip. RNA sequencing analysis of skin fibroblasts derived from affected individuals skin fibroblasts showed significant changes in the expression of several genes implicated in neuronal function and ion transport. Thus, we describe here the clinical spectrum associated with *TRRAP* pathogenic missense variants, and we suggest a genotype-phenotype correlation useful for clinical evaluation of the pathogenicity of the variants.

Post-translational modifications including acetylation, methylation, phosphorylation, and ubiquitination, of core histones directly alter DNA-histone and histone-histone interactions and thus influence nucleosome dynamics.[1] Tight regulation of these marks is required by cells for proper gene transcription,[2] DNA repair,[3] and DNA replication. One major activator of transcription is the acetylation of histone tails, which act by neutralizing the positive charges of lysine residues or by recruiting chromatin remodelers and transcription factors.[4] This tightly regulated process is performed by histone acetyltransferases (HATs) and reversed by histone deacetylases

[1]Centre Hospitalier Universitaire de Nantes, Service de Génétique Médicale, 9 quai Moncousu, 44093 Nantes, France; [2]INSERM, CNRS, UNIV Nantes, l'institut du thorax, 44007 Nantes, France; [3]Centre Hospitalier Universitaire Sainte-Justine Research Centre, University of Montreal, Montreal, QC H3T 1C5, Canada; [4]Department of Medicine, University of Melbourne, Austin Health and Royal Melbourne Hospital, Melbourne, VIC 3010, Australia; [5]AstraZeneca Centre for Genomics Research, Precision Medicine and Genomics, IMED Biotech Unit, AstraZeneca, Cambridge CB2 0AA, UK; [6]Visual Geometry Group, Department of Engineering Science, University of Oxford, Oxford OX1 3PJ, UK; [7]GeneDx, 207 Perry Parkway, Gaithersburg, MD 20877, USA; [8]Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN 55905, USA; [9]Center for Individualized Medicine, Health Sciences Research, Mayo Clinic, Rochester, MN 55905, USA; [10]Department of Psychiatry, Weill Institute for Neurosciences, University of California, San Francisco, San Francisco, CA 94158, USA; [11]UMR 1078, Génétique, Génomique Fonctionnelle et Biotechnologies, Inserm, L'Etablissement Français du Sang, Institut Brestois Santé Agro Matière, Université de Brest Occidentale, 29200 Brest, France; [12]Service de Génétique médicale et de biologie de la reproduction, Centre Hospitalier Regional Universitaire Brest, 29200 Brest, France; [13]Division of Neurogenetics, Kennedy Krieger Institute, Baltimore, MD 21205, USA; [14]Hugo W. Moser Research Institute, Kennedy Krieger Institute, Baltimore, MD 21205, USA; [15]McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA; [16]Department of Clinical Genetics, Chapel Allerton Hospital, Yorkshire Regional Genetics Service, Leeds Teaching Hospitals National Health Service Trust, Chapeltown Road, Leeds LS7 4SA, UK; [17]Department of Pediatrics, University of British Columbia, Vancouver, BC V6H 3N1, Canada; [18]Molecular and Clinical Medicine, School of Medicine, University of Dundee, Ninewells Hospital and Medical School, Dundee DD1 9SY, UK; [19]Rush University Medical Center, Department of Pediatrics, Division of Genetics, Chicago, IL 60612, USA; [20]Oxford Centre for Genomic Medicine, Oxford University Hospitals National Health Service Trust, Oxford OX3 7LE, UK; [21]West of Scotland Regional Genetics Service, Queen Elizabeth University Hospital, Glasgow G51 4TF, UK; [22]Department of Medical Genetics, St. Olav's Hospital, Trondheim University Hospital, 7006 Trondheim, Norway; [23]Laboratory of Embryology and Genetics of Human Malformations, Institut National de la Santé et de la Recherche Médicale (Inserm), UMR 1163, Institut Imagine, 75015 Paris, France; [24]Paris Descartes-Sorbonne Paris Cité University, Institut Imagine, 75015 Paris, France; [25]Division of Genetics,

(Affiliations continued on next page)

Nicholas Stong,[42] Julie McGaughran,[43,44] Kyle Retterer,[7] Kristian Tveten,[45] Jennifer Sullivan,[36] Madeleine R. Geisheker,[46] Asbjorg Stray-Pedersen,[47] Jennifer M. Tarpinian,[48] Eric W. Klee,[8,9,49,50] Julie C. Sapp,[32] Jacob Zyskind,[7] Øystein L. Holla,[45] Emma Bedoukian,[48] Francesca Filippini,[23,24] Anne Guimier,[23,24,51] Arnaud Picard,[24,52] Øyvind L. Busk,[45] Jaya Punetha,[26] Rolph Pfundt,[38] Anna Lindstrand,[34,35] Ann Nordgren,[34,35] Fayth Kalb,[53] Megha Desai,[7] Ashley Harmon Ebanks,[40] Shalini N. Jhangiani,[54] Tammie Dewan,[17] Zeynep H. Coban Akdemir,[26] Aida Telegrafi,[7] Elaine H. Zackai,[30,31] Amber Begtrup,[7] Xiaofei Song,[26] Annick Toutain,[55,56] Ingrid M. Wentzensen,[7] Sylvie Odent,[57,58] Dominique Bonneau,[59,60] Xénia Latypova,[1,2] Wallid Deb,[1,2] CAUSES Study,[17] Sylvia Redon,[11,12] Frédéric Bilan,[61,62] Marine Legendre,[61,62] Caitlin Troyer,[25] Kerri Whitlock,[63] Oana Caluseriu,[63] Marine I. Murphree,[49] Pavel N. Pichurin,[49] Katherine Agre,[49] Ralitza Gavrilova,[49,64] Tuula Rinne,[38] Meredith Park,[65] Catherine Shain,[66] Erin L. Heinzen,[42] Rui Xiao,[26,29] Jeanne Amiel,[23,24,51] Stanislas Lyonnet,[23,24,51] Bertrand Isidor,[1,2] Leslie G. Biesecker,[32] Dan Lowenstein,[67] Jennifer E. Posey,[26]

(Author list continued on next page)

(HDACs). There are three major families of HATs: Gcn5-related N-acetyltrasnferase (GNAT), MYST (MOZ, SAS2, SAS3—also known as YBF2—and TIP60), and p300 (EP300-CREBBP).[5] The activity and localization of most HATs, such as TIP60 or GCNL5, depend on a multiprotein assembly that contains the scaffolding protein transformation/transcription domain-associated protein (TRRAP).

TRRAP is a large protein of 3,859 amino acids and is conserved from yeast to humans. It is an ataxia-telangiectasia mutated (ATM) related member of the phosphatidylinositol 3-kinase-related kinase (PIKK) family.[6] Like other ATM-related members, it contains FAT (FRAP, ATM, and TRRAP) and FATC (FRAP, ATM, and TRRAP, C terminus) domains flanking a PI3/PI4-kinase domain. The kinase domain of TRRAP does not engage in catalytic activity[7] but is required for the proper recruitment of HAT complexes.[8] TRRAP has been shown to be involved in P53-, E2F-, and c-MYC-dependent gene transcription and oncogenic transformation.[6,9,10] As stressed in cancer studies, TRRAP plays an important role in cell-cycle regulation. A recurrent somatic *TRRAP* variant, c.2165C>T p.(Ser722Phe),[11] has been identified in melanoma, and the oncogenic potential of TRRAP has been identified in glioblastoma multiforme,[12] pancreatic adenocarcinoma,[13] and lymphoma.[10] Furthermore, *Trrap* knockout

Department of Pediatrics, University of Virginia Children's Hospital, Charlottesville, VA 22903, USA; [26]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA; [27]Texas Children's Hospital, Houston, TX 77030, USA; [28]Department of Molecular Physiology and Biophysics, Baylor College of Medicine, Houston, TX 77030, USA; [29]Baylor Genetics, Houston, TX 77021, USA; [30]Division of Human Genetics, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA; [31]Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA; [32]Medical Genomics and Metabolic Genetics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD 20892-4472, USA; [33]Division of Epilepsy and Clinical Neurophysiology and Epilepsy Genetics Program, Boston Children's Hospital, Harvard Medical School, Boston, MA 02115, USA; [34]Department of Molecular Medicine and Surgery, Center for Molecular Medicine, Karolinska Institutet, 17176 Stockholm, Sweden; [35]Department of Clinical Genetics, Karolinska University Hospital, 17176 Stockholm, Sweden; [36]Division of Medical Genetics, Department of Pediatrics, Duke University Medical Center, Durham, NC 27710, USA; [37]Department of Genetics, University Medical Center Groningen, University of Groningen, Groningen, 9700 RB, the Netherlands; [38]Department of Human Genetics, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen, 6525 GA, the Netherlands; [39]Section of Genetics and Metabolism, University of Arkansas for Medical Sciences, Little Rock, AR 72202, USA; [40]Department of Pediatric Surgery, The McGovern Medical School at the University of Texas Health Science Center and Children's Memorial Hermann Hospital, Houston, TX 77030, USA; [41]Division of Genetics and Metabolism, Phoenix Children's Hospital, Phoenix, AZ 85016, USA; [42]Institute for Genomic Medicine, Columbia University, New York, NY 10032, USA; [43]Genetic Health Queensland, Royal Brisbane and Women's Hospital, Brisbane, Queensland 4029, Australia; [44]School of Medicine, The University of Queensland, Brisbane, Queensland 4029, Australia; [45]Department of Medical Genetics, Telemark Hospital Trust, 3710 Skien, Norway; [46]Department of Genome Sciences, University of Washington School of Medicine, Seattle, WA 98195, USA; [47]Norwegian National Unit for Newborn Screening, Division of Pediatric and Adolecent Medicine, Oslo University Hospital, Rikshospitalet, Pb 4950 Nydalen, N-0424 Oslo, Norway; [48]Roberts Individualized Medical Genetics Center, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA; [49]Department of Clinical Genomics, Mayo Clinic, Rochester, MN 55905, USA; [50]Department of Health Sciences Research, Mayo Clinic, Rochester, MN 55905, USA; [51]Service de Génétique, Hôpital Necker—Enfants Malades, Assistance Publique—Hôpitaux de Paris (APHP), 75015 Paris, France; [52]Service de Chirurgie Maxillofaciale et Plastique, Centre de référence des Malformations de la Face et de la Cavité Buccale, Hôpital Necker-Enfants Malades, Assistance Publique—Hôpitaux de Paris (APHP), 75015 Paris, France; [53]Division of Genetics, Birth Defects, and Metabolism, Ann and Robert H. Lurie Children's Hospital of Chicago, Chicago, IL 60611, USA; [54]Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX 77030, USA; [55]Service de Génétique, Centre Hospitalier Universitaire de Tours, 2 Boulevard Tonnellé, 37044 Tours, France; [56]Inserm U1253, Ibrain, Université de Tours, 37032 Tours, France; [57]Service de Génétique Clinique, Centre Référence Déficiences Intellectuelles de Causes Rares, Centre de Référence Anomalies du Développement, Centre Labellisé pour les Anomalies du Développement (CLAD) Ouest, Centre Hospitalier Universitaire de Rennes, 35203 Rennes, France; [58]Institut de Génétique et Développement de Rennes, UMR 6290, Université de Rennes, 2 Avenue du Professeur Léon Bernard, 35043 Rennes, France; [59]Centre Hospitalier Universitaire de Angers, Département de Biochimie et Génétique, 49933 Angers Cedex 9, France; [60]Mitochondrial and Cardiovascular Pathophysiology (MITOVASC), Unité mixte de Recherche, Centre National de la Recherche Scientifique 6015, Inserm 1083, Université d'Angers, 49933 Angers, France; [61]Centre Hospitalier Universitaire de Poitiers, Service de Génétique, BP577, 86021 Poitiers, France; [62]Equipe d'accueil 3808, Université Poitiers, Poitiers 86034, France; [63]Department of Medical Genetics, University of Alberta, Edmonton, AB T6G 2H7, Canada; [64]Department of Neurology, Mayo Clinic, Rochester, MN 55905, USA; [65]Epilepsy Genetics Program, Department of Neurology, Boston Children's Hospital, Boston, MA 02115, USA; [66]Department of Neurology, Boston Children's Hospital, Harvard Medical School, Boston, MA 02115, USA; [67]Department of Neurology, University of California, San Francisco, San Francisco, CA 94143, USA; [68]Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton CB10 1SA, UK; [69]Rosalind & Morris Goodman Cancer Research Center and Department of Medicine, McGill University, Montreal, QC H3A 1A3, Canada; [70]Department of Biochemistry, McGill University and McGill University Health Center, Montreal, QC H3A 1A3, Canada; [71]Department of Pharmacology, Creighton University Medical

(Affiliations continued on next page)

Anne-Sophie Denommé-Pichon,[59,60] Deciphering Developmental Disorders study,[68] Claude Férec,[11,12] Xiang-Jiao Yang,[69,70] Jill A. Rosenfeld,[26] Brigitte Gilbert-Dussardier,[61,62] Séverine Audebert-Bellanger,[12] Richard Redon,[2] Holly A.F. Stessman,[71] Christoffer Nellaker,[72,73,74] Yaping Yang,[26,29] James R. Lupski,[26,27,54,75] David B. Goldstein,[42] Evan E. Eichler,[46,76] Francois Bolduc,[63,77,78] Stéphane Bézieau,[1,2] and Sébastien Küry,[1,2,81,*] and Philippe M. Campeau[3,79,81,*]

leads to early embryonic lethality in mice through errors in the cell cycle and a failure to arrest at the mitotic checkpoint.[14] In mouse embryonic stem cells (ESCs), *Trrap* is indispensable for self-renewal as well as correct differentiation,[15] suggesting an essential role in embryonic development and morphogenesis. Moreover, brain-specific *Trrap* knockout in mice leads to premature differentiation of neural progenitors and abnormal brain development through a decrease in the expression of cell-cycle regulators. This decreased expression results in brain atrophy and microcephaly.[16] *TRRAP* has previously been associated with neuropsychiatric disorders such as schizophrenia in a few patients.[17–20] We herein provide data showing that *TRRAP* pathogenic variants are associated with a variable neurodevelopmental disorder.

Through an international collaboration and aided by the web-based tool GeneMatcher,[21] we identified 17 distinct missense *TRRAP* variants with strong clinical and/or molecular evidence for pathogenicity in 24 individuals with neurodevelopmental disorders (Table 1, Figure 1A). These variants were identified either by trio or solo exome sequencing (ES) from research and clinical cohorts. All affected individuals or their guardians gave appropriate consent for research procedures. This study was approved by the CHU de Nantes ethics committee (comité consultatif sur le traitement de l'information en matière de recherche no. 14.556). Methods are described in Table S1.

These 17 variants were absent from ExAC and gnomAD[22] and were found *de novo* or apparently *de novo* (maternity and paternity not checked) in all individuals, except for two sisters who had inherited a variant from a mother with low-level mosaicism (Figure S1) and an individual whose father was unavailable but whose paternal grandparents did not carry the variant. Three variants were recurrently observed: p.Ala1043Thr was identified in five individuals, and p.Glu1106Lys and p.Gly1883Arg were each identified in two individuals. All the variants were predicted to be deleterious by CADD[23] (scaled C scores were over 20), and they were variously predicted to be pathogenic by SIFT[24] and PolyPhen-2 HVAR.[25] As shown in Figure 2A, the 17 variants seen in the individuals we studied had significantly increased CADD scores compared to the scores for singleton missense variants reported in gnomAD.

The 17 variants all occurred at residues conserved among vertebrates (Figure 1B) and in regions depleted in missense variants in gnomAD. Indeed, when we assessed missense tolerance ratios for *TRRAP*, we observed that most of the 17 variants were in regions intolerant to missense variants (Figure 2B). Nine out of the 17 variants occurred at highly mutable CpG sites, including one within the codon that leads to the recurrent p.Ala1043Thr variant observed in five individuals. Six missense variants with lesser evidence for pathogenicity were found in another six unrelated individuals (individuals 25 to 30 in Table S1). These variants might be deleterious but were not clearly pathogenic: perhaps the inheritance pattern could not be determined; the variant was present in gnomAD or led to another missense change at the same residue as a variant reported in gnomAD; or the variant was located in a less conserved region of *TRRAP* (Table S2).

Given the number of *de novo* variants identified, the enrichment for *TRRAP de novo* variants in our study was calculated as ($p = 4.2 \times 10^{-6}$) on the basis of denovolyzer.[26] Nevertheless, the current number of 22 detected *de novo* variants in *TRRAP* is not of genome-wide significance ($p = 0.08$) after correction for the following: (a) ~19,000 protein-coding genes, (b) 22,898 trios studied, and (c) the underlying mutability of the full-length protein-coding *TRRAP* transcript. However, this statistical calculation does not take into account the spatial distribution of the variants. Indeed, three-dimensional modeling of human TRRAP structure inferred from the orthologous *Saccharomyces cerevisiae* protein Tra1 (Figure 2C) suggested a clustering of the variants in different regions of TRRAP. The most important clustering was observed for 13 variants between codons 1031 and 1159. Interestingly, when visualized in 3D, these variants localized near one another (Figure 1C), revealing a domain of TRRAP with a potentially novel specific function, although this domain has not yet been characterized. We performed a statistical clustering analysis comparing the mean distance between observed variants to ten million permutations of random variants, as previously described.[27] This analysis revealed

School, Omaha, NE 68178, USA; [72]Nuffield Department of Women's and Reproductive Health, University of Oxford, Oxford, UK; [73]Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, Oxford, UK; [74]Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, University of Oxford, Oxford OX3 7FZ, UK; [75]Department of Pediatrics, Baylor College of Medicine, Houston, TX 77030, USA; [76]Howard Hughes Medical Institute, Seattle, WA 98195, USA; [77]Division of Pediatric Neurology, University of Alberta, Edmonton, AB, Canada; [78]Neuroscience and Mental Health Institute, University of Alberta, Edmonton, AB, Canada; [79]Department of Pediatrics, University of Montreal, Montreal, QC H3T1J4, Canada
[80]These authors contributed equally
[81]These authors contributed equally
*Correspondence: sebastien.kury@chu-nantes.fr (S.K.), p.campeau@umontreal.ca (P.M.C.)
https://doi.org/10.1016/j.ajhg.2019.01.010.

Table 1. *De Novo TRRAP* Variants Identified in 24 Individuals

cDNA	Protein	Inheritance	CpG	gnomAD	CADD Phred Score (v1.3)	SIFT	PolyPhen2 HVAR	Number of Individuals
c.2413C>T	p.Leu805Phe	*de novo*	no	absent	28.2	deleterious (0)	probably damaging (0.998)	1
c.2580C>G	p.Phe860Leu	*de novo*	no	absent	27.6	deleterious (0.03)	possibly damaging (0.867)	1
c.2678G>T	p.Arg893Leu	apparently *de novo*	yes	absent	34	deleterious (0)	probably damaging (0.986)	1
c.3093T>G	p.Ile1031Met	*de novo*	no	absent	23.4	deleterious (0.02)	benign (0.308)	1
c.3104G>A	p.Arg1035Gln	*de novo*	yes	absent	23.9	tolerated (0.09)	benign (0.404)	1
c.3111C>A	p.Ser1037Arg	*de novo*	yes	absent	23.7	tolerated (0.14)	possibly damaging (0.656)	1
c.3127G>A	p.Ala1043Thr	*de novo*	yes	absent	23.2	tolerated (0.27)	benign (0.066)	5
c.3311A>G	p.Glu1104Gly	*de novo*	no	absent	24.6	deleterious (0.04)	probably damaging (0.91)	1
c.3316G>A	p.Glu1106Lys	*de novo*[a]	no	absent	27.7	deleterious (0)	possibly damaging (0.816)	2
c.3331G>T	p.Gly1111Trp	apparently *de novo*	yes	absent	34	deleterious (0)	probably damaging (0.999)	1
c.3475G>A	p.Gly1159Arg	*de novo*	no	absent	33	deleterious (0)	probably damaging (0.999)	1
c.5575C>T	p.Arg1859Cys	*de novo*	yes	absent	34	deleterious (0)	probably damaging (0.997)	1
c.5596T>A	p.Trp1866Arg	*de novo*	no	absent	28.7	deleterious (0)	probably damaging (0.999)	1
c.5598G>T	p.Trp1866Cys	*de novo*	no	absent	33	deleterious (0)	probably damaging (0.999)	1
c.5647G>A	p.Gly1883Arg	*de novo*	yes	absent	33	deleterious (0)	probably damaging (1)	2
c.5795C>T	p.Pro1932Leu	germline mosaicism	yes	absent	35	deleterious (0)	probably damaging (0.997)	2
c.11270G>A	p.Arg3757Gln	*de novo*	yes	absent	28.6	deleterious (0.01)	benign (0.269)	1

The RefSeq transcript used for *TRRAP* is RefSeq: NM_001244580.1. Apparently *de novo* was mentioned when paternity and maternity were not checked. a. For one individual with p.(Glu1106Lys), father was unavailable, paternal grandparents were tested and did not carry the variant.

a significant clustering of variants along the primary sequence of *TRRAP* (p value = 9×10^{-8}), suggesting a model in which specific domains are affected and haploinsufficiency is unlikely, at least for clustering variants.

Among the 24 individuals who carried pathogenic variants, 19 presented with facial dysmorphisms. Recurrent features that were noted among these individuals included upslanted palpebral fissures, epicanthus, telecanthus, a wide nasal bridge and ridge, a broad and smooth philtrum, and a thin upper lip (Figure 3). We performed a computer-assisted facial gestalt visualization,[28,29] which highlighted several of these features, particularly for individuals with variants clustering with the recurrent p.Ala1043Thr variant (Figure 3R). All the individuals had developmental delay, although the severity of intellectual disability (ID) was highly variable. Whereas most individuals had apparent ID with markedly impaired basic life functions, some of them presented with mild ID or even no cognitive deficits (Table 2 and Table S3). Peripheral neuropathy was also noted; it was severe in one individual and consisted of lower-limb hyperreflexia in five other individuals.

In addition to alteration in cerebral function, some individuals showed brain, cerebellum, heart, kidney, or urogenital malformations. We observed a strong genotype-phenotype correlation (Figure 1A, Table 2); the highest incidence of malformations was seen in 13 individuals whose variants cluster in the region of the predicted protein from codons 1031 to 1159: c.3093T>G (p.Ile1031Met), c.3104G>A (p.Arg1035Gln), c.3111C>A (p.Ser1037Arg), c.3127G>A (p.Ala1043Thr), c.3311A>G (p.Glu1104Gly), c.3316G>A (p.Glu1106Lys), c.3331G>T (p.Gly1111Trp), and c.3475G>A (p.Gly1159Arg). In contrast, individuals with variants residing outside of this region had less malformation and presented mainly with autism spectrum disorder (ASD) and/or ID, sometimes associated with epilepsy. Variants in these individuals were more dispersed along the protein, although some, including c.5575C>T (p.Arg1859Cys), c.5596T>A (p.Trp1866Arg), c.5598G>T (p.Trp1866Cys), c.5647G>A (p.Gly1883Arg), and c.5795C>T (p.Pro1932Leu), apparently aggregated in another region.

13 individuals with variants in the codon 1031–1159 region had global developmental delay and apparent ID, ranging from speech delay and learning difficulties to markedly impaired basic life functions (Table 2 and Table S3). The last available occipitofrontal-circumference measurements revealed microcephaly (ranging from −2.8 to −5 standard deviations [SDs]) in 46% (6/13) of individuals. Cerebral magnetic resonance imaging (MRI) had been performed in 10 out of 13 individuals, and seven of those 10 (70%) showed structural brain anomalies, including cerebellar vermis hypoplasia (6/10), ventricular enlargement (3/10), cortical atrophy (2/10), brainstem atrophy (2/10), polymicrogyria (1/10), focal gliosis (1/10), delayed

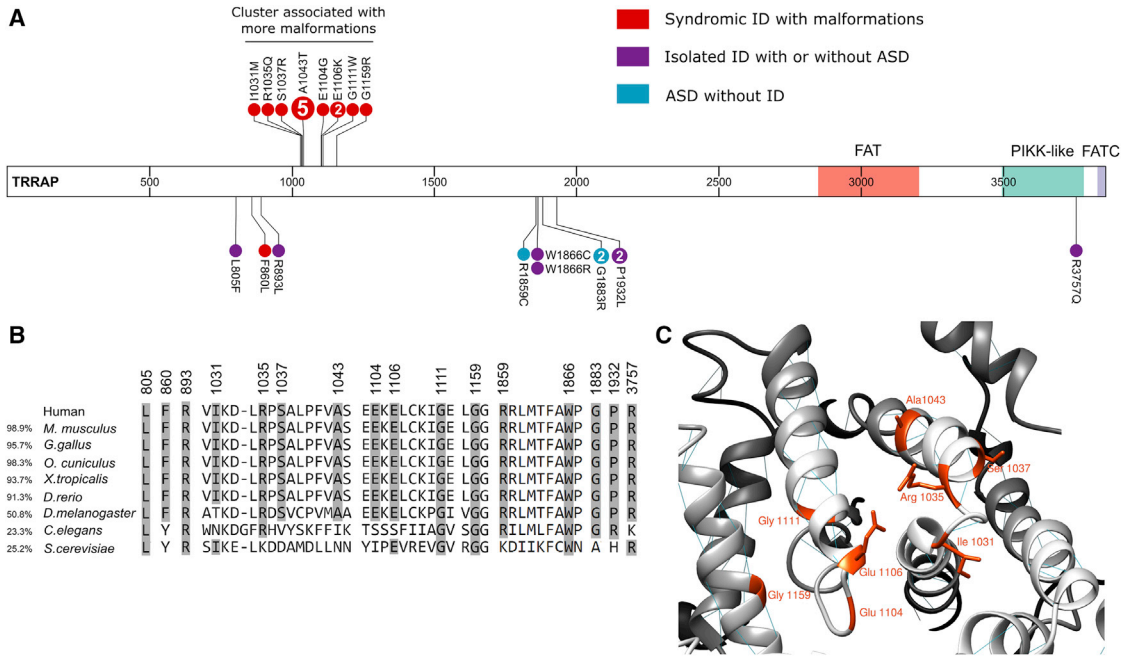

Figure 1. Genotype-Phenotype Correlation Associated with *TRRAP* Variants
(A) Predicted *de novo* and apparently *de novo* variants in affected individuals are represented on the TRRAP protein. The variants in red represent individuals with apparent ID and malformations, the variants in purple represent individuals with isolated ID with or without ASD, and the variants in blue represent individuals with only ASD and an IQ above 70. If more than one individual was heterozygous for the variant, the number of affected individuals is indicated in the circle. Adapted from ProteinPaint.[55]
(B) Amino acid conservation of each mutated residue. The overall amino acid similarity with the human sequence is shown on the left.
(C) Homology model of human TRRAP (GenBank: NP_001231509.1) predicted by PHYRE2 Protein Fold Recognition Server[56] represented by UCSF Chimera.[57] Mutated residues in the 1031–1159 cluster are shown. Abbreviations are as follows: FAT—FRAP, ATM, and TRRAP; PIKK-like—phosphatidylinositol 3-kinase-related protein kinase-like; and FATC—FRAP, ATM, and TRRAP C-terminal.

myelination (1/10), and corpus callosum hypoplasia (1/10). Neurological examination revealed hypotonia in 31% (4/13) of individuals. Only one individual was reported to have epilepsy. Seven individuals (54%) were reported to require feeding exclusively by gastrostomy tube. Among the 10 individuals who were examined by echocardiography, 70% (7/10) had abnormal results, 50% (5/10) had ventricular septal defects, 30% (3/10) had patent ductus arteriosus, 30% (3/10) had patent foramen ovale, 20% (2/10) had pulmonary hypertension, and 20% (2/10) had aortic coarctation. Abdominal ultrasound revealed anomalies in 70% (7/10) of individuals in which it was performed. Abnormal renal morphology, namely multicystic dysplastic kidney, hydronephrosis, a duplicate kidney, and/or a small kidney, was described in 60% (6/10) of individuals, and vesicoureteral reflux was also observed in 30% (3/10) of these individuals. Individual 15 presented with a large left-sided posterolateral congenital diaphragmatic hernia (Table S3). Hernias of the abdominal wall were also found in 23% (3/13) of individuals and included an umbilical hernia, an omphalocele, and an inguinal hernia. Three males (3/6; 50%) had external-genitalia anomalies, including microphallus, hypoplastic scrotum, and cryptorchidism, and two females (2/7; 29%) had a duplicated vagina and/or uterus. Other observed anomalies included dysplastic nails (8/13; 62%), cleft lip and palate (5/13; 38%), clinodactyly of the fifth finger (4/13; 31%), laryngotracheomalacia (3/13), accessory nipple (3/13; 23%), bilateral cutaneous syndactyly of the second and third toe (2/13; 15%), and anomalies of the lacrimal glands (1/13; 8%; see also below with regard to individuals 1 and 19). Four individuals (4/13; 31%) had visual impairment, and three (3/13; 23%) had hearing impairment. Hearing impairment was associated with inner-ear malformations in two cases. Recurrent infections, mainly respiratory and urinary-tract infections, affected three out of 13 (23%) individuals. Individual 9 died at 12 years of age in the context of multiple co-morbidities, including renal failure with acute fluid fluctuations, tracheostomy for severely obstructive laryngotracheomalacia, intermittent supraventricular tachycardia, arterial insufficiency, and polyendocrinopathy (insulin-dependent diabetes, adrenal insufficiency, and hypothyroidism).

Among individuals with variants falling outside of the 1031–1159 region, 5/11 (45%) were diagnosed with ASD, and another three individuals (3/11; 27%) had some findings of ASD but no formal diagnosis. 8/11 (73%) had developmental delay and mild-to-severe ID, and three had speech delay, but their IQs were measured above 70, and two of these IQs were in the normal range. Four individuals (4/11; 36%) had various types of epilepsy, namely absence and tonic-clonic seizures, or Lennox-Gastaut syndrome. The age of seizure onset ranged from 2 to 10 years old. Malformations were infrequent in this group overall, although

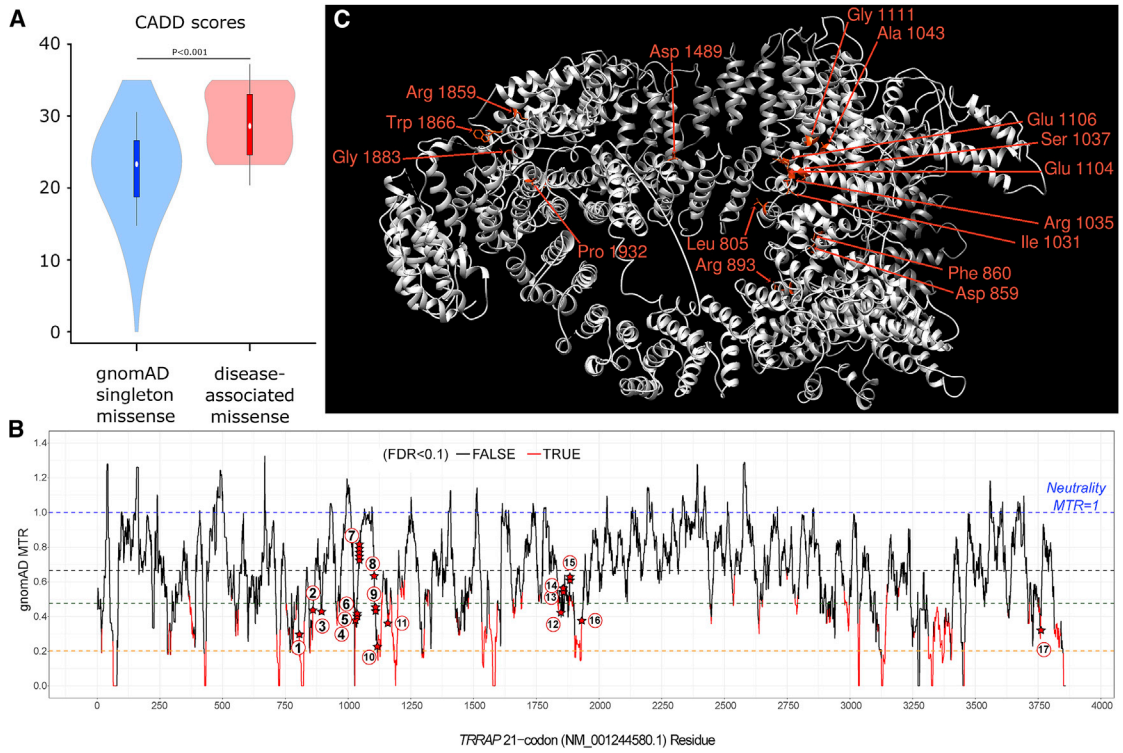

Figure 2. *TRRAP* Sequence Is Intolerant to Missense Variants

(A) CADD scores of the 17 variants identified in affected individuals are compared to scores for gnomAD singleton missense variants. In order to avoid CADD training circularity, we compared the individuals' variants to variants seen once in gnomAD.

(B) *TRRAP* missense tolerance ratio (MTR) plot. The MTR is a statistic that quantifies the extent of purifying selection that has been acting specifically against missense variants in the human population. For *TRRAP*, we adopted the 21-codon sliding window and used exome-sequencing standing-variation data in the gnomAD database, version 2.0. MTR data were downloaded from Missense Tolerance Ratio (MTR) Gene Viewer (see Web Resources). An MTR = 1 (blue dashed line) represents neutrality (i.e., observing the same proportion of missense variants in the window as expected on the basis of the underlying sequence context). Red segments of the MTR plot have achieved exome-wide FDR<0.10 for a significance test of a window's deviation from MTR = 1. The black dashed line signifies gene-specific median MTR, the brown dashed line signifies gene-specific 25th centile MTR, and the orange dashed line signifies gene-specific fifth centile MTR. The locations of our 23 case-ascertained *de novo* variants are denoted by red stars along *TRRAP*'s MTR plot. The 17 different variants are numbered within circles as follows: (1) p.Leu805Phe; (2) p.Phe860Leu; (3) p.Arg893Leu; (4) p.Ile1031Met; (5) p.Arg1035Gln; (6) p.Ser1037Arg; (7) p.Ala1043Thr; (8) p.Glu1104Gly; (9) p.Glu1106Lys; (10) p.Gly1111Trp; (11) p.Gly1159Arg; (12) p.Arg1859Cys; (13) p.Trp1866Arg; (14) p.Trp1866Cys; (15) p.Gly1883Arg; (16) p.Pro1932Leu; and (17) p.Arg3757Gln. We found that *de novo* variants were significantly enriched in the intolerant 50% of *TRRAP*'s protein-coding sequence; 18 (78%) of the 23 *de novo* events affected the most intolerant 50% of the *TRRAP* sequence (binomial exact test p = 0.01). Strikingly, only the most recurring *de novo* missense variant (GenBank: NM_001244580.1 p. Ala1043Thr) resided outside of the intolerant *TRRAP* sequence.

(C) Localization of the mutated TRRAP residues on 3D protein models including 14 out of 17 likely pathogenic variants and two out of six additional variants of unknown significance are shown. The representation of the structure of human TRRAP (GenBank: NP_001231509.1) was predicted by PHYRE2 Protein Fold Recognition Server by comparison to its *Saccharomyces cerevisiae* ortholog, according to the cryo-EM structure of the SAGA (Spt-Ada-Gcn5-acetyltransferase) and NuA4 coactivator subunit Tra1 present in the protein data bank (PDB: 5OJS). Variants in regions non-homologous to Tra1 are not represented. Structure representation was made with UCSF Chimera.

individual 2 had microcephaly and heart malformations, individual 1 had lacrimal duct aplasia, individual 19 had lacrimal duct aplasia and optic disc colobomas, and individual 21 had a postaxial polydactyly of one hand.

TRRAP-associated chromatin remodeling complexes are generally associated with gene activation,[30] which is consistent with their HAT activity. Nevertheless, the NuA4 complex has been shown to have a gene-repression activity necessary for ESC pluripotency.[31,32] This gene-repression activity seems to be independent from its lysine acetyltransferase activity.[33] To test the hypothesis that *TRRAP* variants alter gene expression, we obtained skin fibroblasts from two individuals, individual 1, with p.Leu805Phe, and individual 19, with p.Trp1866Cys and performed next-generation sequencing with technical replicates of RNA (i.e., separately prepared libraries from the same samples). The RNA library preparation and sequencing as well as bioinformatics analysis methods can be found in the Supplemental Data. We found that, in comparison to two typically developing individuals (controls), both individuals with *TRRAP* variants had remarkably different gene expression patterns (Figure S2A). Interestingly, most differentially expressed genes (DEGs) analyzed with DESeq2 were upregulated in affected individuals compared to controls (Figure S2B). Moreover, the individual with p.Leu805Phe had 619

Figure 3. Photographs of Individuals with *TRRAP* Variants
(A) Individual 1 at the age of 8 years. Note the telecanthus, broad nasal bridge, widely spaced eyes, relatively thin upper vermilion, flared eyebrows, and ectropion.

(legend continued on next page)

DEGs; the Log_2 fold change (Log_2FC) was higher than 2 or lower than −2, and the p value was adjusted for 10% false discovery rate lower than 0.01 (padj) (Supplemental Data, Table S5).

To identify genes with significant expression differences, we performed differential gene expression analysis between the two individuals with *TRRAP* variants (combined as biological replicates) and two unaffected controls. Gene ontology (GO) enrichment analysis of these genes with the GOrilla web application indicates an enrichment for the adrenergic receptor signaling pathway, genes important for neurological function, and potassium and ATP-sensitive ion transporters (Figure S2B, Supplemental Data, Table S5). The two individuals who were tested carried variants outside the cluster associated with the more syndromic ID; if there are distinct effects on gene regulation, it will be worth comparing gene expression between the two groups. Finally, because it has been shown that TRRAP has direct interactions with different partners not related to the HAT complex, we cannot exclude the possibility that the transcriptome alteration might be caused by a mechanism other than impaired HAT activity. Thus, we highlighted candidate pathways that might be useful for uncovering the pathomechanism of *TRRAP* variants in future studies.

TRRAP acts as a scaffold in HAT complexes. Although it does not have a direct role in acetylation, we hypothesize that pathogenic effects of variants might be due to dysregulation of acetylation, a major process that has been associated with several neurodevelopmental disorders.[34] Pathogenic variants of *KAT6B* (MIM: 605880) cause both Say-Barber-Biesecker-Young-Simpson syndrome (SBBYSS [MIM: 603736])[35–37] and genitopatellar syndrome (GPS [MIM: 606170]),[38,39] and pathogenic variants in *KAT6A* and *BRPF1* mutations have also been associated with a neurodevelopmental disorder.[40–42] Rubinstein-Taybi syndrome (MIM: 180849 and 613684) is associated with variants in HAT-complex-encoding genes, namely *CREBBP* and *EP300*.[43–46] In addition to cognitive impairment, abnormal histone acetylation can also result in behavioral disorders, as evidenced by the associations found between non-syndromic ASD and/or schizophrenia and alterations in several lysine acetyltransferase and lysine deacetylase genes, including *BRD1*, *HDAC4*, *HDAC6,* and *HDAC9*.[34,47–50]

Variants in *TRRAP* were associated with neuropsychiatric disorders, including childhood disintegrative disorder,[17] schizophrenia,[18,19] and ASD.[20] The ASD report included individuals 18 and 19, who had p.Trp1866Arg and p.Trp1866Cys, respectively. We thus confirmed the association with ASD and provide evidence that it can be found either isolated or associated with ID. On the basis of the ExAC dataset alone without studies on neuropsychiatric disorders, *TRRAP* is in the top five human genes that are most intolerant of missense variants: it has a missense z-score of 10.1.[22] Although this study includes only the first 24 identified individuals, a strength of the study is that it was primarily ascertained by sequencing, reducing phenotypic ascertainment bias. Given the highly constrained region of the observed variants coupled with the population constraint and evolutionary conservation, we hypothesize that variants outside of these regions are likely to be associated with prenatal lethality, although we cannot exclude the possibility that milder phenotypes might be underrepresented in current exome datasets. It is worth noting that we exclusively identified missense variants in the affected individuals. Given the loss-of-function (LoF) intolerance of *TRRAP* in ExAC (pLI = 1.00), we would

(B) Individual 5 at the age of 8.5 years. Note the wide mouth, thin upper lip, and widely spaced eyes with a wide and depressed nasal bridge.
(C) Individual 6 at the age of 29 years. Note the sparse eyebrows, upslanting palpebral fissures, smooth philtrum, thin upper lip, and low columella.
(D) Individual 9 at the age of 11 years. Note the deeply set eyes, sparse eyebrows, and wide nasal bridge.
(E) Individual 8. Note the telecanthus, low-set ears with upturned earlobes, and, on the fourth picture from the left, the single transverse palmar crease.
(F) Individual 12 at the age of 5 years. Note the prominent forehead, arched eyebrows, short palpebral fissures, epicanthal folds, depressed nasal bridge, and thick upper vermilion.
(G) Individual 13 at the age of 14 years. Note the upslanted palpebral fissures and prominent forehead.
(H) Individual 10 at the ages of 1 month, 16 months, and 3 years. Note the cleft lip and palate, wide mouth, epicanthic fold, prognathism, and supernumerary nipples.
(I) Individual 15 at the age of 12 years. Note the wide nasal bridge and upslanting palpebral fissures.
(J) Individual 19 at the ages of 2.5 years and 8 years. Note folded-down upper eyelid and sparse medial eyebrows.
(K) Individual 16 at the age of 2 years. Note the prominent forehead, epicanthic fold, telecanthus, flat nasal bridge, and low-set ears.
(L) Individual 20 at the age of 10 years. Note the widely spaced eyes, telecanthus, wide nasal bridge and ridge, and thin upper vermilion.
(M) Individual 18. Note the narrow nose, flared eyebrows, almond-shaped eyes with hypoplastic infraorbital ridges, telecanthus, smooth philtrum, and small, low-set, and posteriorly rotated ears.
(N) Individual 21. Note the short palpebral fissures, epicanthal folds, and thin upper vermilion.
(O) Individual 22 at the age of 24 years. Note the broad nasal bridge, deeply set eyes, upslanted palpebral fissures, widely spaced eyes, and posteriorly rotated ears.
(P) Individual 23 at the age of 19 years. Note the deeply set eyes, upslanted palpebral fissures, widely spaced eyes, epicanthal folds, and posteriorly rotated ears.
(Q) Individual 24. Note the smooth philtrum and wide nasal ridge.
(R) Average facial gestalt visualization of nine healthy age- and gender-matched controls on the left; on the right, nine individuals with variants in the 1031–1159 cluster. Facial images are flipped and aligned to preserve bilateral asymmetry.

Table 2. Clinical Description of Individuals with Variants Inside or Outside the 1031–1159 Cluster

Symptoms	All Individuals	Cluster 1031–1159	Variants Outside the Cluster
Global developmental delay	24/24 (100%)	13/13 (100%)	11/11 (100%)
Intellectual disability	17/20 (85%)	11/11 (100%)	6/9 (67%)
Facial dysmorphisms	19/24 (79%)	11/13 (85%)	8/11 (73%)
Autism spectrum disorder	5/24 (21%)	0/13 (0%)	5/11 (45%)
Microcephaly (<-2.5 SD)	7/24 (29%)	6/13 (46%)	1/11 (9%)
Short stature	7/23 (30%)	4/12 (33%)	3/11 (27%)
Hypotonia	8/24 (33%)	4/13 (31%)	4/11 (36%)
Feeding difficulties	8/24 (33%)	7/13 (54%)	1/11 (9%)
Seizures	5/24 (21%)	1/13 (8%)	4/11 (36%)
Cleft lip and palate	5/24 (21%)	5/13 (38%)	0/11 (0%)
Cerebellar hypoplasia	6/18 (33%)	6/11 (55%)	0/7 (0%)
Cerebral abnormalities	6/18 (33%)	6/11 (55%)	0/7 (0%)
Cardiac malformations	10/15 (66%)	9/12 (75%)	1/3 (33%)
Renal malformations	5/17 (29%)	5/13 (38%)	0/4 (0%)
Genital malformations	5/24 (21%)	5/13 (38%)	0/11 (0%)
Hearing impairment	3/24 (12%)	3/13 (23%)	0/11 (0%)
Visual impairment	4/24 (17%)	3/13 (23%)	1/11 (9%)
Scoliosis	3/24 (12%)	3/13 (23%)	0/11 (0%)
Dysplastic nails	8/24 (33%)	8/13 (62%)	0/11 (0%)
Lower-limb hyperreflexia	5/24 (21%)	1/13 (8%)	4/11 (36%)
Lacrimal-duct aplasia	3/24 (12%)	1/13 (8%)	2/11 (18%)
Accessory nipple	4/24 (17%)	3/13 (23%)	1/11 (9%)

expect to identify at least some LoF variants if haploinsufficiency of *TRRAP* was the causal mechanism. In DECIPHER (accessed May 14, 2018), no small or intragenic deletions involving *TRRAP* have been identified. Thus, when the significant clustering is taken into account, our results suggest that missense variants might act either as gain-of-function or dominant-negative variants and that haploinsufficiency of *TRRAP* is likely to be prenatally lethal, although we cannot exclude the possibility that an LoF effect of non-clustering variants is associated with a milder phenotype.

TRRAP participates in embryonic development, as demonstrated by its binding with proteins regulating the Notch signaling pathway in fruit fly[51], the Ras signaling pathway in *C. elegans*[52], or the Wnt[53] signaling pathway in 293T cells.[53] Therefore we suspect that TRRAP variants, more especially those falling within the 1031–1159 region, perturb the interactions with at least one of these developmental signaling pathways; such a perturbance would explain the multiple malformations observed in about half of the affected individuals.

In yeast, a series of ~100 codon deletion mutants in the ortholog tra1 showed reduced or complete loss of viability.[54] Most deletions impaired coactivator complex assembly, notably the ones encompassing the homologous 1031–1159 cluster (mutants Δ13–Δ14), as well as the regions homologous to those containing variants p.Leu805Phe, p.Phe860Leu, and p.Arg893Leu (mutants Δ11–Δ12) and the p.Arg3757Gln variant (mutant Δ39). In contrast, mutants Δ21–Δ22 encompassing the region homologous to the cluster associated with fewer malformations (codons 1859–1932) were viable, which might help explain the milder clinical phenotype associated with variants within this cluster. In mice, *Trrap* knockout leads to early embryonic lethality,[14] and a neural-cell-specific conditional *Trrap* knockout line[16] revealed premature differentiation of neural progenitors, depletion of progenitor pools, and a significant reduction in cortical thickness. These mice exhibited striking microcephaly, in agreement with what we observed in half of the individuals in our study cohort, primarily those with variants in the 1031–1159 cluster.

In summary, we report evidence that variants in *TRRAP* are associated with a pleiotropic neurodevelopmental syndrome with a potential genotype-phenotype correlation. Our functional data highlight an enrichment of genes related to neuronal function and ion transport. This enrichment could underline the pathophysiology of the

disease. Future *in vitro* and *in vivo* studies on variants inside and outside the main cluster will be required if we are to determine which gene expression changes are connected to which *TRRAP*-related specific phenotypes.

Supplemental Data

Supplemental Data can be found with this article online at https://doi.org/10.1016/j.ajhg.2019.01.010.

Acknowledgments

We would like to thank all families for participating in this study. This work was supported in part by grants from: the French Ministry of Health and the Health Regional Agency from Poitou-Charentes (HUGODIMS, 2013, RC14_0107) to S.B.; the National Institute of Neurological Disorders and Stroke (The Epilepsy Phenome/Genome Project NS053998; Epi4K NS077364, NS077274, NS077303, and NS077276) to D.L. and D.B.G.; the National Institutes of Health/Eunice Kennedy Shriver National Institute of Child Health and Human Development grant (HD064667) to D.A.S.; NINDS R35 NS105078 to J.R.L.; the National Human Genome Research Institute (NHGRI) and National Heart, Lung, and Blood Institute (NHLBI) to the Baylor-Hopkins Center for Mendelian Genomics (UM1 HG006542); NHGRI K08 HG008986 to J.E.P.; the Duke Genome Sequencing Clinic to V.S. and J.S.; the intramural research program of the NHGRI (grant HG200328 12) to L.G.B., J.J.J., and J.C.S.; the US National Institute of Mental Health grant R01MH101221 to E.E.E.; the Kids Brain Health Network and Dart NeuroScience to F.B.; and Mining for Miracles, British Columbia Children's Hospital Foundation, and Genome British Columbia to the CAUSES Study. We thank the Canadian Institutes of Health Research (CIHR) and Fonds de la recherche en santé du Québec (FRSQ) for clinician-scientist awards to P.M.C.; and the Mayo Clinic Center for Individualized Medicine (CIM) for supporting this research through the CIM Investigative and Functional Genomics Program. We are grateful to the members of the Canadian Center for Computational Genomics and the McGill University and Génome Québec Innovation Center for their help in bioinformatics analysis.

Declaration of Interests

E.E.E. is on the scientific advisory board of DNAnexus. The Department of Molecular and Human Genetics at Baylor College of Medicine receives revenue for clinical genetic testing completed at Baylor Genetics Laboratory. K.M., K.R., J.Z., M.D., A.T., A.B., and I.M.W. are employees of GeneDx, Inc. Dr. Goldstein is Founder and holds equity in Pairnomix and Praxis Therapeutics. Dr. Goldstein is not aware of any overlap with Pairnomix or Praxis Therapeutics.

Received: July 16, 2018
Accepted: January 18, 2019
Published: February 28, 2019

Web Resources

DECIPHER, https://decipher.sanger.ac.uk/
Ensembl VEP, http://grch37.ensembl.org/Homo_sapiens/Tools/VEP
ExAC Browser, http://exac.broadinstitute.org/
GenBank, http://www.ncbi.nlm.nih.gov/genbank/
gnomAD, http://gnomad.broadinstitute.org/
Missense Tolerance Ratio (MTR) Gene Viewer, http://mtr-viewer.mdhs.unimelb.edu.au/
OMIM, http://www.omim.org/
Phyre2, http://www.sbg.bio.ic.ac.uk/phyre2/
UniProt, http://www.uniprot.org/uniprot/

References

1. Bowman, G.D., and Poirier, M.G. (2015). Post-translational modifications of histones that influence nucleosome dynamics. Chem. Rev. *115*, 2274–2295.
2. Venkatesh, S., and Workman, J.L. (2015). Histone exchange, chromatin structure and the regulation of transcription. Nat. Rev. Mol. Cell Biol. *16*, 178–189.
3. Hunt, C.R., Ramnarain, D., Horikoshi, N., Iyengar, P., Pandita, R.K., Shay, J.W., and Pandita, T.K. (2013). Histone modifications and DNA double-strand break repair after exposure to ionizing radiations. Radiat. Res. *179*, 383–392.
4. Legube, G., and Trouche, D. (2003). Regulating histone acetyltransferases and deacetylases. EMBO Rep. *4*, 944–947.
5. Berndsen, C.E., and Denu, J.M. (2008). Catalysis and substrate selection by histone/protein lysine acetyltransferases. Curr. Opin. Struct. Biol. *18*, 682–689.
6. McMahon, S.B., Van Buskirk, H.A., Dugan, K.A., Copeland, T.D., and Cole, M.D. (1998). The novel ATM-related protein TRRAP is an essential cofactor for the c-Myc and E2F oncoproteins. Cell *94*, 363–374.
7. Vassilev, A., Yamauchi, J., Kotani, T., Prives, C., Avantaggiati, M.L., Qin, J., and Nakatani, Y. (1998). The 400 kDa subunit of the PCAF histone acetylase complex belongs to the ATM superfamily. Mol. Cell *2*, 869–875.
8. Park, J., Kunjibettu, S., McMahon, S.B., and Cole, M.D. (2001). The ATM-related domain of TRRAP is required for histone acetyltransferase recruitment and Myc-dependent oncogenesis. Genes Dev. *15*, 1619–1624.
9. Zhao, L.-J., Loewenstein, P.M., and Green, M. (2017). Enhanced MYC association with the NuA4 histone acetyltransferase complex mediated by the adenovirus E1A N-terminal domain activates a subset of MYC target genes highly expressed in cancer cells. Genes Cancer *8*, 752–761.
10. Jethwa, A., Słabicki, M., Hüllein, J., Jentzsch, M., Dalal, V., Rabe, S., Wagner, L., Walther, T., Klapper, W., Bohnenberger, H., et al.; MMML Network Project (2018). TRRAP is essential for regulating the accumulation of mutant and wild-type p53 in lymphoma. Blood *131*, 2789–2802.
11. Wei, X., Walia, V., Lin, J.C., Teer, J.K., Prickett, T.D., Gartner, J., Davis, S., Stemke-Hale, K., Davies, M.A., Gershenwald, J.E., et al.; NISC Comparative Sequencing Program (2011). Exome sequencing identifies GRIN2A as frequently mutated in melanoma. Nat. Genet. *43*, 442–446.
12. Wurdak, H., Zhu, S., Romero, A., Lorger, M., Watson, J., Chiang, C.-Y., Zhang, J., Natu, V.S., Lairson, L.L., Walker, J.R., et al. (2010). An RNAi screen identifies TRRAP as a regulator of brain tumor-initiating cell differentiation. Cell Stem Cell *6*, 37–47.
13. Loukopoulos, P., Shibata, T., Katoh, H., Kokubu, A., Sakamoto, M., Yamazaki, K., Kosuge, T., Kanai, Y., Hosoda, F., Imoto, I., et al. (2007). Genome-wide array-based comparative genomic

hybridization analysis of pancreatic adenocarcinoma: identification of genetic indicators that predict patient outcome. Cancer Sci. *98*, 392–400.
14. Herceg, Z., Hulla, W., Gell, D., Cuenin, C., Lleonart, M., Jackson, S., and Wang, Z.Q. (2001). Disruption of Trrap causes early embryonic lethality and defects in cell cycle progression. Nat. Genet. *29*, 206–211.
15. Sawan, C., Hernandez-Vargas, H., Murr, R., Lopez, F., Vaissière, T., Ghantous, A.Y., Cuenin, C., Imbert, J., Wang, Z.-Q., Ren, B., and Herceg, Z. (2013). Histone acetyltransferase cofactor Trrap maintains self-renewal and restricts differentiation of embryonic stem cells. Stem Cells *31*, 979–991.
16. Tapias, A., Zhou, Z.-W., Shi, Y., Chong, Z., Wang, P., Groth, M., Platzer, M., Huttner, W., Herceg, Z., Yang, Y.-G., and Wang, Z.Q. (2014). Trrap-dependent histone acetylation specifically regulates cell-cycle gene transcription to control neural progenitor fate decisions. Cell Stem Cell *14*, 632–643.
17. Gupta, A.R., Westphal, A., Yang, D.Y.J., Sullivan, C.A.W., Eilbott, J., Zaidi, S., Voos, A., Vander Wyk, B.C., Ventola, P., Waqar, Z., et al. (2017). Neurogenetic analysis of childhood disintegrative disorder. Mol. Autism *8*, 19.
18. Xu, B., Ionita-Laza, I., Roos, J.L., Boone, B., Woodrick, S., Sun, Y., Levy, S., Gogos, J.A., and Karayiorgou, M. (2012). De novo gene mutations highlight patterns of genetic and neural complexity in schizophrenia. Nat. Genet. *44*, 1365–1369.
19. Takata, A., Xu, B., Ionita-Laza, I., Roos, J.L., Gogos, J.A., and Karayiorgou, M. (2014). Loss-of-function variants in schizophrenia risk and SETD1A as a candidate susceptibility gene. Neuron *82*, 773–780.
20. Geisheker, M.R., Heymann, G., Wang, T., Coe, B.P., Turner, T.N., Stessman, H.A.F., Hoekzema, K., Kvarnung, M., Shaw, M., Friend, K., et al. (2017). Hotspots of missense mutation identify neurodevelopmental disorder genes and functional domains. Nat. Neurosci. *20*, 1043–1051.
21. Sobreira, N., Schiettecatte, F., Valle, D., and Hamosh, A. (2015). GeneMatcher: A matching tool for connecting investigators with an interest in the same gene. Hum. Mutat. *36*, 928–930.
22. Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., et al.; Exome Aggregation Consortium (2016). Analysis of protein-coding genetic variation in 60,706 humans. Nature *536*, 285–291.
23. Kircher, M., Witten, D.M., Jain, P., O'Roak, B.J., Cooper, G.M., and Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. Nat. Genet. *46*, 310–315.
24. Kumar, P., Henikoff, S., and Ng, P.C. (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat. Protoc. *4*, 1073–1081.
25. Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S., and Sunyaev, S.R. (2010). A method and server for predicting damaging missense mutations. Nat. Methods *7*, 248–249.
26. Ware, J.S., Samocha, K.E., Homsy, J., and Daly, M.J. (2015). Interpreting de novo variation in human disease using denovolyzeR. Curr. Protoc. Hum. Genet. *87*, 7.25.1–7.25.15.
27. Lelieveld, S.H., Wiel, L., Venselaar, H., Pfundt, R., Vriend, G., Veltman, J.A., Brunner, H.G., Vissers, L.E.L.M., and Gilissen, C. (2017). Spatial clustering of de novo missense mutations identifies candidate neurodevelopmental disorder-associated genes. Am. J. Hum. Genet. *101*, 478–484.
28. Ferry, Q., Steinberg, J., Webber, C., FitzPatrick, D.R., Ponting, C.P., Zisserman, A., and Nellåker, C. (2014). Diagnostically relevant facial gestalt information from ordinary photos. eLife *3*, e02020.
29. Reijnders, M.R.F., Janowski, R., Alvi, M., Self, J.E., van Essen, T.J., Vreeburg, M., Rouhl, R.P.W., Stevens, S.J.C., Stegmann, A.P.A., Schieving, J., et al. (2018). PURA syndrome: clinical delineation and genotype-phenotype study in 32 individuals with review of published literature. J. Med. Genet. *55*, 104–113.
30. Murr, R., Vaissière, T., Sawan, C., Shukla, V., and Herceg, Z. (2007). Orchestration of chromatin-based processes: Mind the TRRAP. Oncogene *26*, 5358–5372.
31. Chen, P.B., Hung, J.-H., Hickman, T.L., Coles, A.H., Carey, J.F., Weng, Z., Chu, F., and Fazzio, T.G. (2013). Hdac6 regulates Tip60-p400 function in stem cells. eLife *2*, e01557.
32. Fazzio, T.G., Huff, J.T., and Panning, B. (2008). An RNAi screen of chromatin proteins identifies Tip60-p400 as a regulator of embryonic stem cell identity. Cell *134*, 162–174.
33. Acharya, D., Hainer, S.J., Yoon, Y., Wang, F., Bach, I., Rivera-Pérez, J.A., and Fazzio, T.G. (2017). KAT-independent gene regulation by Tip60 promotes ESC self-renewal but not pluripotency. Cell Rep. *19*, 671–679.
34. Tapias, A., and Wang, Z.-Q. (2017). Lysine acetylation and deacetylation in brain development and neuropathies. Genomics Proteomics Bioinformatics *15*, 19–36.
35. Clayton-Smith, J., O'Sullivan, J., Daly, S., Bhaskar, S., Day, R., Anderson, B., Voss, A.K., Thomas, T., Biesecker, L.G., Smith, P., et al. (2011). Whole-exome-sequencing identifies mutations in histone acetyltransferase gene KAT6B in individuals with the Say-Barber-Biesecker variant of Ohdo syndrome. Am. J. Hum. Genet. *89*, 675–681.
36. Szakszon, K., Salpietro, C., Kakar, N., Knegt, A.C., Oláh, É., Dallapiccola, B., and Borck, G. (2013). De novo mutations of the gene encoding the histone acetyltransferase KAT6B in two patients with Say-Barber/Biesecker/Young-Simpson syndrome. Am. J. Med. Genet. A. *161A*, 884–888.
37. Yilmaz, R., Beleza-Meireles, A., Price, S., Oliveira, R., Kubisch, C., Clayton-Smith, J., Szakszon, K., and Borck, G. (2015). A recurrent synonymous KAT6B mutation causes Say-Barber-Biesecker/Young-Simpson syndrome by inducing aberrant splicing. Am. J. Med. Genet. A. *167A*, 3006–3010.
38. Campeau, P.M., Kim, J.C., Lu, J.T., Schwartzentruber, J.A., Abdul-Rahman, O.A., Schlaubitz, S., Murdock, D.M., Jiang, M.-M., Lammer, E.J., Enns, G.M., et al. (2012). Mutations in KAT6B, encoding a histone acetyltransferase, cause Genitopatellar syndrome. Am. J. Hum. Genet. *90*, 282–289.
39. Simpson, M.A., Deshpande, C., Dafou, D., Vissers, L.E.L.M., Woollard, W.J., Holder, S.E., Gillessen-Kaesbach, G., Derks, R., White, S.M., Cohen-Snuijf, R., et al. (2012). De novo mutations of the gene encoding the histone acetyltransferase KAT6B cause Genitopatellar syndrome. Am. J. Hum. Genet. *90*, 290–294.
40. Millan, F., Cho, M.T., Retterer, K., Monaghan, K.G., Bai, R., Vitazka, P., Everman, D.B., Smith, B., Angle, B., Roberts, V., et al. (2016). Whole exome sequencing reveals de novo pathogenic variants in KAT6A as a cause of a neurodevelopmental disorder. Am. J. Med. Genet. A. *170*, 1791–1798.
41. Murray, C.R., Abel, S.N., McClure, M.B., Foster, J., 2nd, Walke, M.I., Jayakar, P., Bademci, G., and Tekin, M. (2017). Novel causative variants in *DYRK1A, KARS*, and *KAT6A* associated with intellectual disability and additional phenotypic features. J. Pediatr. Genet. *6*, 77–83.

42. Yan, K., Rousseau, J., Littlejohn, R.O., Kiss, C., Lehman, A., Rosenfeld, J.A., Stumpel, C.T.R., Stegmann, A.P.A., Robak, L., Scaglia, F., et al.; DDD Study; and CAUSES Study (2017). Mutations in the chromatin regulator gene BRPF1 cause syndromic intellectual disability and deficient histone acetylation. Am. J. Hum. Genet. *100*, 91–104.
43. Roelfsema, J.H., White, S.J., Ariyürek, Y., Bartholdi, D., Niedrist, D., Papadia, F., Bacino, C.A., den Dunnen, J.T., van Ommen, G.-J.B., Breuning, M.H., et al. (2005). Genetic heterogeneity in Rubinstein-Taybi syndrome: Mutations in both the CBP and EP300 genes cause disease. Am. J. Hum. Genet. *76*, 572–580.
44. Bartholdi, D., Roelfsema, J.H., Papadia, F., Breuning, M.H., Niedrist, D., Hennekam, R.C., Schinzel, A., and Peters, D.J.M. (2007). Genetic heterogeneity in Rubinstein-Taybi syndrome: Delineation of the phenotype of the first patients carrying mutations in EP300. J. Med. Genet. *44*, 327–333.
45. Tsai, A.C., Dossett, C.J., Walton, C.S., Cramer, A.E., Eng, P.A., Nowakowska, B.A., Pursley, A.N., Stankiewicz, P., Wiszniewska, J., and Cheung, S.W. (2011). Exon deletions of the EP300 and CREBBP genes in two children with Rubinstein-Taybi syndrome detected by aCGH. Eur. J. Hum. Genet. *19*, 43–49.
46. Negri, G., Milani, D., Colapietro, P., Forzano, F., Della Monica, M., Rusconi, D., Consonni, L., Caffi, L.G., Finelli, P., Scarano, G., et al. (2015). Clinical and molecular characterization of Rubinstein-Taybi syndrome patients carrying distinct novel mutations of the EP300 gene. Clin. Genet. *87*, 148–154.
47. Xu, L.-M., Li, J.-R., Huang, Y., Zhao, M., Tang, X., and Wei, L. (2012). AutismKB: An evidence-based knowledgebase of autism genetics. Nucleic Acids Res. *40*, D1016–D1022.
48. Piton, A., Jouan, L., Rochefort, D., Dobrzeniecka, S., Lachapelle, K., Dion, P.A., Gauthier, J., and Rouleau, G.A. (2013). Analysis of the effects of rare variants on splicing identifies alterations in GABAA receptor genes in autism spectrum disorder individuals. Eur. J. Hum. Genet. *21*, 749–756.
49. Lang, B., Alrahbeni, T.M., Clair, D.S., Blackwood, D.H., International Schizophrenia Consortium, McCaig, C.D., and Shen, S. (2011). HDAC9 is implicated in schizophrenia and expressed specifically in post-mitotic neurons but not in adult neural stem cells. Am. J. Stem Cells *1*, 31–41.
50. Severinsen, J.E., Bjarkam, C.R., Kiaer-Larsen, S., Olsen, I.M., Nielsen, M.M., Blechingberg, J., Nielsen, A.L., Holm, I.E., Foldager, L., Young, B.D., et al. (2006). Evidence implicating BRD1 with brain development and susceptibility to both schizophrenia and bipolar affective disorder. Mol. Psychiatry *11*, 1126–1138.
51. Gause, M., Eissenberg, J.C., Macrae, A.F., Dorsett, M., Misulovin, Z., and Dorsett, D. (2006). Nipped-A, the Tra1/TRRAP subunit of the Drosophila SAGA and Tip60 complexes, has multiple roles in Notch signaling during wing development. Mol. Cell. Biol. *26*, 2347–2359.
52. Ceol, C.J., and Horvitz, H.R. (2004). A new class of C. elegans synMuv genes implicates a Tip60/NuA4-like HAT complex as a negative regulator of Ras signaling. Dev. Cell *6*, 563–576.
53. Finkbeiner, M.G., Sawan, C., Ouzounova, M., Murr, R., and Herceg, Z. (2008). HAT cofactor TRRAP mediates beta-catenin ubiquitination on the chromatin and the regulation of the canonical Wnt pathway. Cell Cycle *7*, 3908–3914.
54. Knutson, B.A., and Hahn, S. (2011). Domains of Tra1 important for activator recruitment and transcription coactivator functions of SAGA and NuA4 complexes. Mol. Cell. Biol. *31*, 818–831.
55. Zhou, X., Edmonson, M.N., Wilkinson, M.R., Patel, A., Wu, G., Liu, Y., Li, Y., Zhang, Z., Rusch, M.C., Parker, M., et al. (2016). Exploring genomic alteration in pediatric cancer using ProteinPaint. Nat. Genet. *48*, 4–6.
56. Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J.E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. Nat. Protoc. *10*, 845–858.
57. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera—A visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605–1612.

The American Journal of Human Genetics, Volume *104*

Supplemental Data

Missense Variants in the Histone Acetyltransferase Complex Component Gene *TRRAP* Cause Autism and Syndromic Intellectual Disability

Benjamin Cogné, Sophie Ehresmann, Eliane Beauregard-Lacroix, Justine Rousseau, Thomas Besnard, Thomas Garcia, Slavé Petrovski, Shiri Avni, Kirsty McWalter, Patrick R. Blackburn, Stephan J. Sanders, Kévin Uguen, Jacqueline Harris, Julie S. Cohen, Moira Blyth, Anna Lehman, Jonathan Berg, Mindy H. Li, Usha Kini, Shelagh Joss, Charlotte von der Lippe, Christopher T. Gordon, Jennifer B. Humberson, Laurie Robak, Daryl A. Scott, Vernon R. Sutton, Cara M. Skraban, Jennifer J. Johnston, Annapurna Poduri, Magnus Nordenskjöld, Vandana Shashi, Erica H. Gerkes, Ernie M.H.F. Bongers, Christian Gilissen, Yuri A. Zarate, Malin Kvarnung, Kevin P. Lally, Peggy A. Kulch, Brina Daniels, Andres Hernandez-Garcia, Nicholas Stong, Julie McGaughran, Kyle Retterer, Kristian Tveten, Jennifer Sullivan, Madeleine R. Geisheker, Asbjorg Stray-Pedersen, Jennifer M. Tarpinian, Eric W. Klee, Julie C. Sapp, Jacob Zyskind, Øystein L. Holla, Emma Bedoukian, Francesca Filippini, Anne Guimier, Arnaud Picard, Øyvind L. Busk, Jaya Punetha, Rolph Pfundt, Anna Lindstrand, Ann Nordgren, Fayth Kalb, Megha Desai, Ashley Harmon Ebanks, Shalini N. Jhangiani, Tammie Dewan, Zeynep H. Coban Akdemir, Aida Telegrafi, Elaine H. Zackai, Amber Begtrup, Xiaofei Song, Annick Toutain, Ingrid M. Wentzensen, Sylvie Odent, Dominique Bonneau, Xénia Latypova, Wallid Deb, CAUSES Study, Sylvia Redon, Frédéric Bilan, Marine Legendre, Caitlin Troyer, Kerri Whitlock, Oana Caluseriu, Marine I. Murphree, Pavel N. Pichurin, Katherine Agre, Ralitza Gavrilova, Tuula Rinne, Meredith Park, Catherine Shain, Erin L. Heinzen, Rui Xiao, Jeanne Amiel, Stanislas Lyonnet, Bertrand Isidor, Leslie G. Biesecker, Dan Lowenstein, Jennifer E. Posey, Anne-Sophie Denommé-Pichon, Deciphering Developmental Disorders study, Claude Férec, Xiang-Jiao Yang, Jill A. Rosenfeld, Brigitte Gilbert-Dussardier, Séverine Audebert-Bellanger, Richard Redon, Holly A.F. Stessman, Christoffer Nellaker, Yaping Yang, James R. Lupski, David B. Goldstein, Evan E. Eichler, Francois Bolduc, Stéphane Bézieau, Sébastien Küry, and Philippe M. Campeau

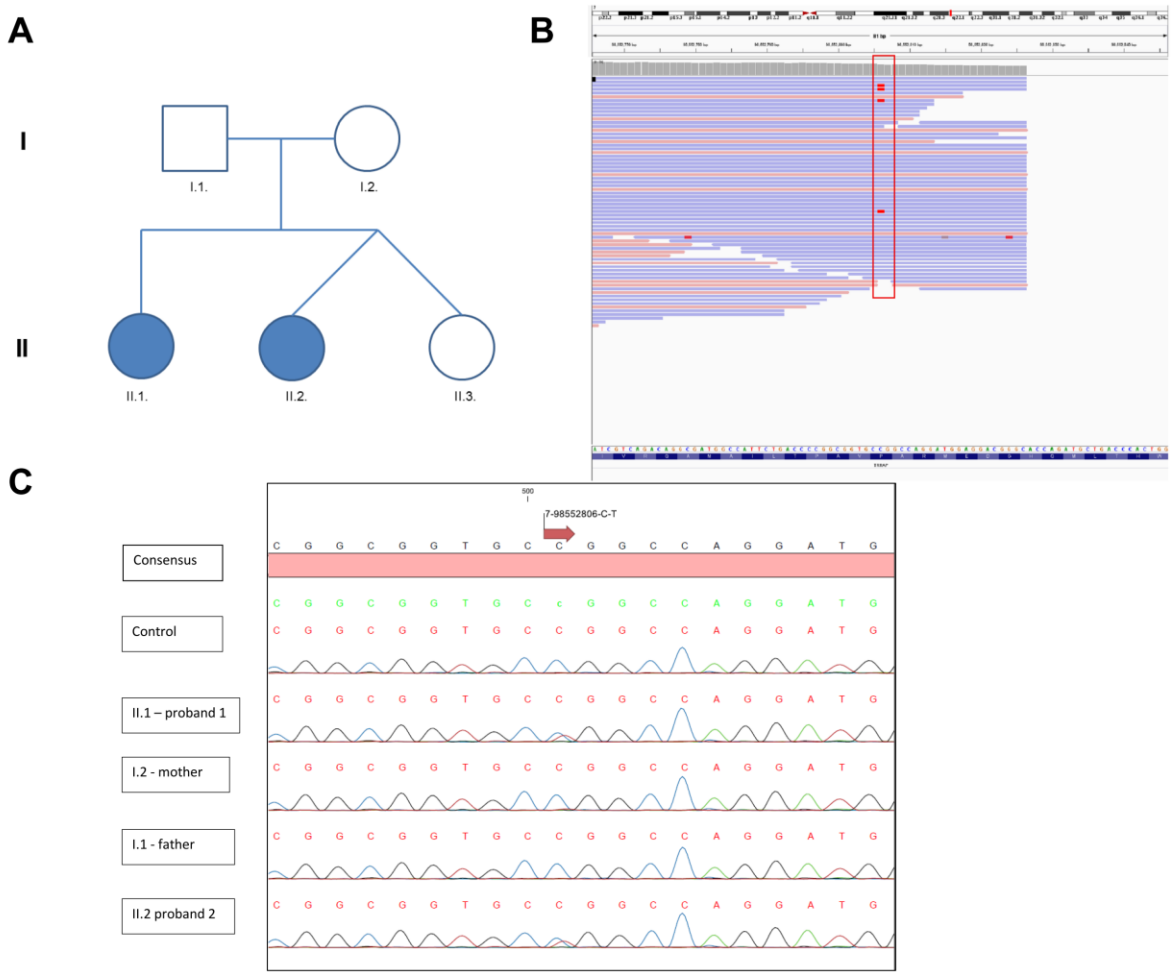

Figure S1. Maternal mosaicism of variant c.5795C>T (p.Pro1932Leu) identified in individuals 22 and 23. (A) Pedigree of the family. (B) Integrative Genome Viewer snapshot of BAM emitted by GATK haplotype caller in the mother (individual I.2.). Reads are shown aligned to the genome in blue for forward strand and red for reverse strand. The C to T variant is shown in red on 4/55 reads. This low 7% representation is significantly deviated from the expected 50% and is indicative of mosaicism. (C) Segregation analysis of the variant by Sanger sequencing confirms the heterozygous variant in the siblings and the mosaicism in the mother.

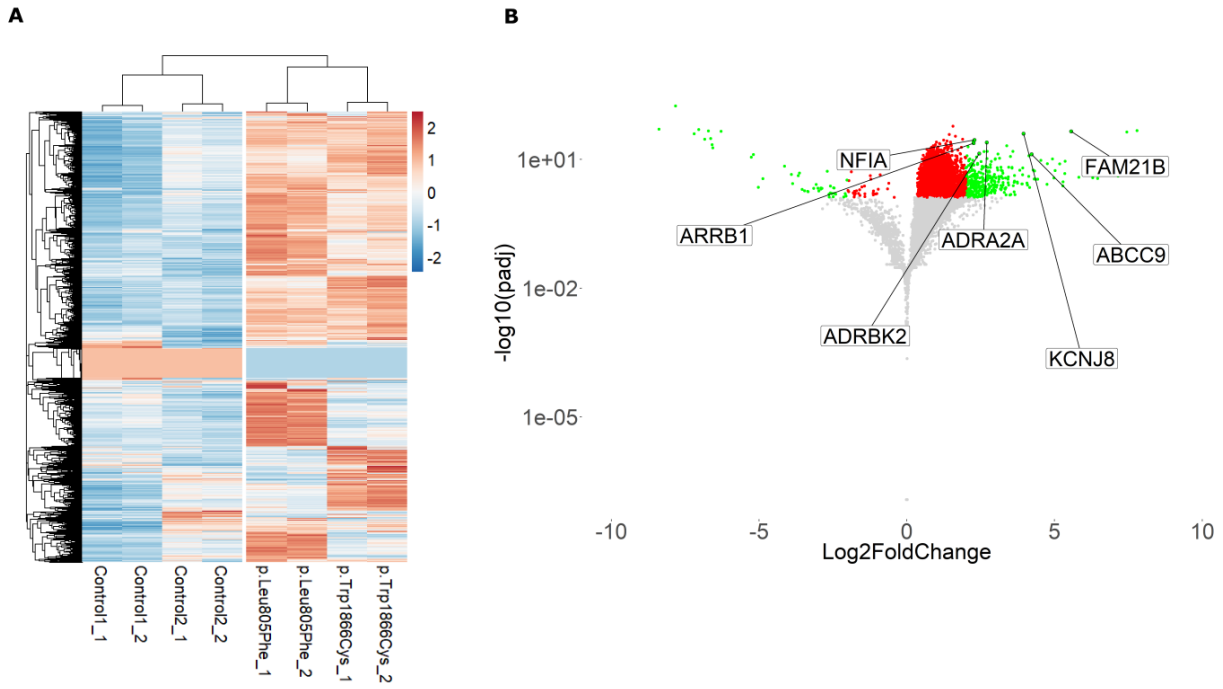

Figure S2. Fibroblasts harboring the p.(Leu805Phe) or p.(Trp1866Cys) variants have differential gene expression patterns.

RNAseq was performed on two healthy controls and two individual fibroblasts to assay gene expression with technical duplicates. (A) Heatmap of genes expressed with at least 10 counts in at least one condition. Normalized count values from "DESeq2" were plotted and scaled row-wise using the "pheatmap" R package. (B) Log_2 Fold Change (Log_2FC) and p-values adjusted for 10% False Discovery Rate (padj) were calculated for the two affected individuals and two controls respectively pooled together using the "DESeq2" R package then plotted using the "ggplot2" R package as a volcano plot. Green dots represent genes with a padj lower than 0.01, and an absolute Log_2FC higher than 2 (padj < 0.01, abs(Log_2FC) > 2) and are referred to as Differentially Expressed Genes (DEGs). Red dots represent genes with a padj lower than 0.01 but an absolute Log_2FC lower than 2 (padj < 0.01, abs(Log_2FC) < 2).

Supplementary methods

RNAseq

Human primary fibroblasts for patient were cultured in DMEM (ThermoFisher cat# 11995-065), 10% FBS, 1mM GlutaMax (ThermoFisher cat# 35050-061) and antibiotics-antimycotics (ThermoFisher cat# 15240-062). Fibroblasts were plated at 1 million cells per 150 mm dish and allowed to grow until they reached 80% confluency. Cells were washed twice with D-PBS, resuspended in QIAzol (Qiagen cat# 79306) and stored at -80°C until all samples were ready for RNA extraction. RNA isolation was performed using the RNeasy mini kit (Qiagen cat# 74104), according to the manufacturer's protocol. Samples were treated with the Turbo DNA free kit (ThermoFisher cat# AM1907) and quality was assessed using the Agilent 2100 Bioanalyzer. Sequencing was performed at the CHU Sainte-Justine and Génome Québec Integrated Clinical Genomic Centre in Pediatrics (CIGCP). mRNA Libraries were prepared using the TruSeq Stranded mRNA kit for 48 samples (Illumina), according to the manufacturer's instructions. Samples were run on the Illumina HiSeq 4000 PE100 with 7 samples per lane. Output files were analyzed using the MUGQIC RNAseq pipeline (MUGQIC) steps 1 through 14 on the Guillimin Génome Québec HPC. In summary, BAM files were converted to FASTQ using Picard (BROAD Institute), sequences were trimmed using Trimmomatic (Bolger, 2014) then aligned to the GRCh37 genome using STAR (Dobin, 2013); duplicate and misaligned reads were discarded using Picard, and raw counts were called using HTseq (Anders S, 2015). Differential expression analysis was performed using the DESeq2 R package (Love, 2014), with default parameters. GO annotation analysis was performed using the GOrilla web application (Eden, 2009). Control and patient cell lines of the same type were respectively pooled together into two groups for differential gene expression analysis, so as to find genes which were significantly differentially expressed in all patients compared to all controls. Significantly expressed genes were selected with an adjusted p value (10% False Discovery Rate, padj)

lower than 0.01, and a log2 Fold Change (log2FC) higher than 2 or lower than -2; corresponding to an overall fold change of at least 4 or -4. Genes of interest were all significantly differentially expressed in individual patient analyses with DESeq2 compared to controls.

qPCR

Total RNA from human cell lines was extracted as described above. Affected individuals were compared to 5 controls for ABCC9 and 6 controls for KCNJ8. Among the controls used for ABCC9and KCNJ8, the following cell lines were obtained from the NIGMS Human Genetic Cell Repository at the Coriell Institute for Medical Research: AGO8498A; GM00041 and GM07532. Reverse transcription was performed with 1 µg- of total RNA in a 20 µl reaction volume using qScript cDNA SuperMix (QuantaBio; CA101414-102) according to the manufacturer's specifications. For Real-Time PCR, cDNA samples were diluted 2-fold (*HMBS* and *KCNJ8*) or 35-fold (*GAPDH*, *ABCC9*), respectively. Gene-specific transcription level was measured in triplicates using the PowerUp SYBR Green Master Mix (Thermo Fisher Scientific ; A25741) and the LightCycler® 96 System from Roche Life Science). Primers (amplified sequences 85 to 200 bp) were designed using PrimerBank (https://pga.mgh.harvard.edu/primerbank/) and Primer-Blast (http://www.ncbi.nlm.nih.gov/tools/primer-blast/). *KCNJ8* (FW : GCTCTTCGCTATCATGTGGT; REV : GAAGAGAAAAGCAGAAGTGAAAGAC E=2.07), *ABCC9* (FW: TCGCTTCCTTTTGAGTCCTG; REV: ATGTCCTCTGTTTCTGCGG E= 2.15), and controls: *HMBS* (FW : GGCAATGCGGCTGCAA; REV : GGGTACCCACGCGAATCAC E= 1.93), *GABDH* (FW : AGCCACATCGCTCAGACA; GCCCAATACGACCAAATCC E=1.92). Melting curves for each primer sets showed a single symmetrical amplicon, and no primer-dimer peaks were observed in the no-reverse transcription-control (NoRT) reactions. PCR efficiency for each set of primers was calculated based on the slope of the standard curve from a 5-fold serial dilution and taken into account for the calculation of the relative quantification ratio (RQ) as described by (Pfaffl, 2001).

Statistics

Outliers were identify and deleted with the Grubbs' test, $\alpha = 0.05$ (Prism 6 – GraphPad). Results are expressed as the means ± standard errors of the means (SEMs). Unpaired parametric t-test (Prism 6 - GraphPad) was performed unless otherwise indicated.

Bibliography for methods section

1. Anders S, P. P. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics, 31(2), 166-9.
2. Bolger, A. M. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics, 30(15), 2114–2120.
3. BROAD Institute. (n.d.). Picard Tools. Retrieved 2017, from https://github.com/broadinstitute/picard
4. Dobin, A. D. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics, 29(1), 15–21.
5. Eden, E. N. (2009). GOrilla: A Tool For Discovery And Visualization of Enriched GO Terms in Ranked Gene Lists. BMC Bioinformatics, 10(48).
6. Love, M. I. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biolog, 15(12), 550.
7. MUGQIC. (n.d.). MUGQIC pipelines RNAseq, 2.2.1-beta. Retrieved 2017, from https://bitbucket.org/mugqic/mugqic_pipelines/src/master/pipelines/rnaseq/
8. Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Research, 29(9), 45e–45. https://doi.org/10.1093/nar/29.9.e45

Supplementary discussion on RNAseq analysis

Apart from common DEGs, each patient had a majority of DEGs which weren't shared. Of those, the p.Leu805Phe individual had upregulation of more than 80 cell cycle regulation genes (supplementary). GO analysis of DEGs in this patient showed a significant enrichment in chromosome segregation, chromosome organization, cell division, mitotic cell cycle process, and other related GOs. *TRRAP* has been shown to be necessary for cell-cycle regulation, and *Trrap* conditional knock-out in mouse brain leads to premature neural progenitor differentiation, microcephaly and disorganized neuronal layers[1]. Moreover, *Trrap* knockout mouse ESCs accumulate chromosomal aberrations because of chromosome fragmentation and lagging, and fail to arrest at the mitotic checkpoint[2], consistent with the genes upregulated in this individual (*NDE1, MIS18A, NEK2, CENPE, CENPF,* and others – Supplementary). Thus, it is possible that overexpressing *TRRAP*-related genes can lead to cell-cycle deregulation and abnormal neural progenitor differentiation, leading to the neural phenotype observed in *TRRAP* patients. The p.Trp1866Cys individual did not have the same enrichment in cell cycle regulation genes. Nevertheless, the DEGs with the lowest adjusted P values were associated mostly with neuronal function (*EMB*, *PDE11A*, *STMN2*, *DYSF*, *SGIP1*, *PLEKHG5,* and others - Supplementary). In addition, members of the *HOXC* cluster were significantly upregulated (*HOXC9*, *HOXC10*, *HOXC11*, *HOXC-AS1*, *HOXC-AS2*, *HOXC-AS3, HOTAIR*), as well as *HOXB9* and *HOXD13*, suggesting a role for TRRAP in regulating 5' Hox gene signaling. Homeobox containing transcription factors (*Hox* genes) have been implicated in embryonic morphogenesis and a multitude of disorders[3], and the spatio-temporal regulation of their expression has been thoroughly studied in a model of "temporal collinearity" [4; 5]. Each cluster of *Hox* genes is globally repressed, then progressively activated from the 3' to the 5' end during embryogenesis through the addition of active histone marks (H3K27Ac, H3K4m3) and the removal of repressive histone marks (H3K27m3) at the same time in all the clusters (*Hox* A, B,

C, and D). An irregular early expression of these *Hox* genes can cause abnormal embryonic development in mouse and drosophila[6; 7].

1. Tapias, A., Zhou, Z.W., Shi, Y., Chong, Z., Wang, P., Groth, M., Platzer, M., Huttner, W., Herceg, Z., Yang, Y.G., et al. (2014). Trrap-dependent histone acetylation specifically regulates cell-cycle gene transcription to control neural progenitor fate decisions. Cell stem cell 14, 632-643.
2. Herceg, Z., Hulla, W., Gell, D., Cuenin, C., Lleonart, M., Jackson, S., and Wang, Z.Q. (2001). Disruption of Trrap causes early embryonic lethality and defects in cell cycle progression. Nature genetics 29, 206-211.
3. Quinonez, S.C., and Innis, J.W. (2014). Human HOX gene disorders. Molecular genetics and metabolism 111, 4-15.
4. Mallo, M., and Alonso, C.R. (2013). The regulation of Hox gene expression during animal development. Development 140, 3951-3963.
5. Montavon, T., and Duboule, D. (2013). Chromatin organization and global regulation of Hox gene clusters. Philosophical transactions of the Royal Society of London Series B, Biological sciences 368, 20120367.
6. Coiffier, D., Charroux, B., and Kerridge, S. (2008). Common functions of central and posterior Hox genes for the repression of head in the trunk of Drosophila. Development 135, 291-300.
7. Kmita, M., van Der Hoeven, F., Zakany, J., Krumlauf, R., and Duboule, D. (2000). Mechanisms of Hox gene colinearity: transposition of the anterior Hoxb1 gene into the posterior HoxD complex. Genes & development 14, 198-211.

Supplemental Acknowledgments

We would like to thank Frédérique Allaire from the Health Regional Agency of Poitou-Charentes for supporting this project. We acknowledge Léa Ferrand and Emilie Le Blanc's assistance for grant and data management. E.E.E. is an investigator of the Howard Hughes Medical Institute. CAUSES Study investigators include Shelin Adam, Christele Du Souich, Alison Elliott, Anna Lehman, Jill Mwenifumbo, Tanya Nelson, Clara Van Karnebeek, and Jan Friedman. We acknowledge also the DECIPHER Consortium which contributed to the exchange of genetic and clinical data between the teams.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.