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 Blackburn, et al.

Version of Record: https://www.sciencedirect.com/science/article/pii/S0002929719300102
Manuscript_86c7d7ac48e45aef71b07d84a95dca31

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 Sanders,[10] Kévin Uguen,[11,12] Jacqueline Harris,[13,14] Julie S. Cohen,[13] Moira Blyth,[15] Anna Lehman,[16] Jonathan Berg,[17] Mindy H. Li,[18] Usha Kini,[19] Shelagh Joss,[20] Charlotte von der Lippe,[21] Christopher T. Gordon,[22,23] Jennifer B. Humberson,[24] Laurie Robak,[25] Daryl A. Scott,[25,26,27] Vernon R. Sutton,[25,26,28] Cara M. Skraban,[29,30] Jennifer J. Johnston,[31] Annapurna Poduri,[32] Magnus Nordenskjöld,[33,34] Vandana Shashi,[35] Erica H. Gerkes,[36] Ernie M. Bongers,[37] Christian Gilissen,[37] Yuri A. Zarate,[38] Malin Kvarnung,[33,34] Kevin P. Lally,[39] Peggy A. Kulch,[40] Brina Daniels,[38] Andres Hernandez-Garcia,[25] Nicholas Stong,[41] Julie McGaughran,[42] Kyle Retterer,[7] Kristian Tveten,[43] Jennifer Sullivan,[35] Madeleine R. Geisheker,[44] Asbjorg Stray-Pedersen,[45] Jennifer M. Tarpinian,[46] Eric W. Klee,[8,9,47,48] Julie C. Sapp,[31] Jacob Zyskind,[7] Øystein L. Holla,[43] Emma Bedoukian,[46] Francesca Filippini,[22,23] Anne Guimier,[22,23,49] Arnaud Picard,[23,50] Øyvind L. Busk,[43] Jaya Punetha,[25] Rolph Pfundt,[37] Anna Lindstrand,[33,34] Ann Nordgren,[33,34] Fayth Kalb,[51] Megha Desai,[7] Ashley Harmon Ebanks,[39] Shalini N. Jhangiani,[52] Tammie Dewan,[16] Zeynep H. Coban Akdemir,[25] Aida Telegrafi,[7] Elaine H. Zackai,[29,30] Amber Begtrup,[7] Xiaofei Song,[25] Annick Toutain,[53,54] Ingrid M. Wentzensen,[7] Sylvie Odent,[55,56] Dominique Bonneau,[57,58] Xénia Latypova,[1,2] Wallid Deb,[1,2] CAUSES Study,[16] Sylvia Redon,[11,12] Frédéric Bilan,[59,60] Marine Legendre,[59,60] Caitlin Troyer,[24] Kerri Whitlock,[61] Oana Caluseriu,[61] Marine I. Murphree,[47] Pavel N. Pichurin,[47] Katherine Agre,[47] Ralitza Gavrilova,[47,62] Tuula Rinne,[37] Meredith Park,[63] Catherine Shain,[64] Erin L. Heinzen,[41] Rui Xiao,[25,65] Jeanne Amiel,[22,23,49] Stanislas Lyonnet,[22,23,49] Bertrand Isidor,[1,2] Leslie G. Biesecker,[31] Dan Lowenstein,[66] Jennifer E. Posey,[25] Anne-Sophie Denommé-Pichon,[57,58] Deciphering Developmental Disorder study,[67] Claude Férec,[11,12] Xiang-Jiao Yang,[68,69] Jill A. Rosenfeld,[25] Brigitte Gilbert-Dussardier,[59,60] Séverine Audebert-Bellanger,[70] Richard Redon,[2] Holly A. F. Stessman,[71] Christoffer Nellaker,[72,73,74] Yaping Yang,[25,65] James R. Lupski,[25,26,52,75] David B. Goldstein,[41] Evan E. Eichler,[44,76] Francois Bolduc,[61,77,78] Stéphane Bézieau,[1,2] Sébastien Küry,[1,2,81,*] Philippe M. Campeau[3,79,81,*]

1. CHU Nantes, Service de Génétique Médicale, 9 quai Moncousu, 44093 Nantes, France;
2. l'Institut du Thorax, INSERM, CNRS, UNIV Nantes, 44007 Nantes, France;
3. Centre Hospitalier Universitaire Sainte-Justine Research Centre, University of Montreal, Montreal, QC H3T 1C5, Canada;
4. Department of Medicine, The 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, Dept. of Engineering Science, University of Oxford, Oxford OX1 3PJ, UK;
7. GeneDx, 207 Perry Parkway, Gaithersburg, Maryland 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, UCSF Weill Institute for Neurosciences, University of California, San Francisco, San Francisco, California 94158, USA;
11. UMR1078 "Génétique, Génomique Fonctionnelle et Biotechnologies", Inserm, EFS, Université de Brest, ISBAM, 29200 Brest, France;
12. Laboratoire de Génétique Moléculaire et d'Histocompatibilité, CHRU Brest, 29200 Brest, France;
13. Division of Neurogenetics and Hugo W. Moser Research Institute, Kennedy Krieger Institute, Baltimore, MD 21205, USA;
14. McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA;
15. Yorkshire Regional Genetics Service, Leeds Teaching Hospitals NHS Trust, Department of Clinical Genetics, Chapel Allerton Hospital, Chapeltown Road, Leeds LS7 4SA, UK;
16. Department of Medical Genetics, Children's and Women's Health Centre of British Columbia, Vancouver, BC V6H 3N1, Canada;
17. Molecular and Clinical Medicine, School of Medicine, University of Dundee, Ninewells Hospital & Medical School, Dundee DD1 9SY, UK;
18. Rush University Medical Center, Department of Pediatrics, Division of Genetics, Chicago, IL, 60612, USA;
19. Oxford Centre for Genomic Medicine, Oxford University Hospitals NHS Trust, Oxford OX3 7LE, UK;
20. West of Scotland Regional Genetics Service, Queen Elizabeth University Hospital, Glasgow G51 4TF, UK;
21. Department of Medical Genetics, St. Olav’s Hospital, Trondheim University Hospital, 7006 Trondheim, Norway;
22. 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;
23. Paris Descartes-Sorbonne Paris Cité University, Institut Imagine, 75015 Paris, France;
24. Division of Genetics, Department of Pediatrics, University of Virginia Children's Hospital, Charlottesville, Virginia 22903, USA;
25. Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA;
26. Texas Children’s Hospital, Houston, TX 77030, USA;
27. Department of Molecular Physiology and Biophysics, Baylor College of Medicine, Houston, TX 77030, USA;
28. Baylor Genetics, Houston, TX 77030, USA;
29. Division of Human Genetics, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA;
30. Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA;
31. Medical Genomics and Metabolic Genetics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland 20892-4472, USA;
32. Division of Epilepsy and Clinical Neurophysiology and Epilepsy Genetics Program, Boston Children's Hospital, Harvard Medical School, Boston, MA 02115, USA;

33. Department of Molecular Medicine and Surgery, Center for Molecular Medicine, Karolinska Institutet, 17176 Stockholm, Sweden;
34. Department of Clinical Genetics, Karolinska University Hospital, 17176 Stockholm, Sweden;
35. Division of Medical Genetics, Department of Pediatrics, Duke University Medical Center, Durham, NC 27710, USA;
36. University of Groningen, University Medical Center Groningen, Department of Genetics, Groningen, 9700 RB, The Netherlands;
37. Department of Human Genetics, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen, 6525 GA, the Netherlands;
38. Section of Genetics and Metabolism, University of Arkansas for Medical Sciences, Little Rock, Arkansas 72202, USA;
39. 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;
40. Division of Genetics and Metabolism, Phoenix Children's Hospital, Phoenix, AZ 85006, USA;
41. Institute for Genomic Medicine, Columbia University, New York, NY 10032, USA;
42. Genetic Health Queensland, Royal Brisbane and Women's Hospital, Brisbane, Queensland and School of Medicine, The University of Queensland, Brisbane, Queensland 4029, Australia;
43. Department of Medical Genetics, Telemark Hospital Trust, 3710 Skien, Norway;
44. Department of Genome Sciences, University of Washington School of Medicine, Seattle, Washington 98195, USA;
45. Division of Pediatric and Adolecent Medicine, Oslo University Hospital - Rikshospitalet, Pb 4950 Nydalen, N-0424 Oslo, Norway;
46. Roberts Individualized Medical Genetics Center, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania 19104, USA;
47. Department of Clinical Genomics, Mayo Clinic, Rochester, MN, 55905, USA;
48. Department of Biomedical Informatics, Mayo Clinic, Rochester, MN, 55905, USA;
49. Service de Génétique, Hôpital Necker-Enfants Malades, Assistance Publique - Hôpitaux de Paris (APHP), 75015 Paris, France;
50. Service de Chirurgie Maxillofaciale et Plastique, Centre de référence MAFACE, Hôpital Necker-Enfants Malades, Assistance Publique - Hôpitaux de Paris (APHP), 75015 Paris, France;
51. Division of Genetics, Birth Defects, and Metabolism, Ann & Robert H. Lurie Children's Hospital of Chicago, Chicago, IL 60611, USA;
52. Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX 77030, USA;
53. CHU Tours, Service de Génétique, 2 Boulevard Tonnellé, 37044 Tours, France;
54. INSERM U1253, ibrain, Université de Tours, 37032 Tours, France;
55. Service de Génétique Clinique, Centre Référence “Déficiences Intellectuelles de causes rares” (CRDI), Centre de référence anomalies du développement CLAD-Ouest, CHU Rennes, 35203 Rennes, France;
56. CNRS UMR 6290, Université de Rennes, 2 Avenue du Professeur Léon Bernard, 35043 Rennes, France;
57. CHU Angers, Département de Biochimie et Génétique, 49933 Angers Cedex 9, France;
58. MITOVASC, UMR CNRS 6015-INSERM 1083, Université d'Angers, 49933 Angers, France;
59. CHU Poitiers, Service de Génétique, BP577, 86021 Poitiers, France;
60. EA 3808, Université Poitiers, 86034, France;
61. Department of Medical Genetics, University of Alberta, Edmonton, AB T6G 2H7, Canada;
62. Department of Neurology, Mayo Clinic, Rochester, MN, 55905, USA;

63. Epilepsy Genetics Program, Department of Neurology, Boston Children's Hospital, Boston, Massachusetts 02115, USA;
64. Department of Neurology, Boston Children's Hospital, Harvard Medical School, Boston, Massachusetts 02115, USA;
65. Baylor Genetics, Houston, TX, 77021, USA;
66. Department of Neurology, University of California, San Francisco, San Francisco, California 94143, USA;
67. Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton CB10 1SA, UK;
68. Rosalind & Morris Goodman Cancer Research Center and Department of Medicine, McGill University, Montreal, QC H3A 1A3, Canada;
69. Department of Biochemistry, McGill University and McGill University Health Center, Montreal, QC H3A 1A3, Canada;
70. CHRU Brest, Génétique médicale, 29609 Brest, France;
71. Department of Pharmacology, Creighton University Medical 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, Washington 98195, USA;
77. Division of Pediatric Neurology, University of Alberta, Edmonton, Alberta, Canada;
78. Neuroscience and Mental Health Institute, University of Alberta, Edmonton, Alberta, Canada;
79. Department of Pediatrics, University of Montreal, Montreal, QC H3T1J4, Canada;

80. These authors contributed equally to this work.

81. These authors contributed equally to this work.

*Corresponding authors: p.campeau@umontreal.ca, sebastien.kury@chu-nantes.fr

Abstract

Acetylation of the lysine residues in histones and other DNA-binding proteins play 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 characterized by a wide range of intellectual functioning, including a number of individuals with intellectual disability (ID) and markedly impaired basic life functions and associated with various malformations of the brain, heart, kidneys, and genitourinary system. 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 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 on affected individual-derived skin fibroblasts showed significant changes in 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 suggest a genotype-phenotype correlation useful for clinical evaluation of the pathogenicity of the variants.

Keywords:

TRRAP, histone acetylation, *de novo* variants, intellectual disability, congenital malformations, autism spectrum disorder

Main text

Post-translational modifications of core histones including acetylation, methylation, phosphorylation, and ubiquitination directly alter DNA-histone and histone-histone interactions, thus influencing 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 (HAT) and reversed by histone deacetylases (HDAC). There are three major families of HAT: Gcn5-related N-acetyltransferase (GNAT), MOZ, YBF2/SAS3, SAS2, TIP60 (MYST), and EP300-CREBBP[5]. The activity and localization of most HATs like 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 3859 amino acids that is conserved from yeast to humans. It is an ataxia-telangiectasia mutated (ATM)-related member of the phosphatidylinositol (PI) 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 possess 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 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 resulting in brain atrophy and microcephaly[16]. *TRRAP* has previously been associated with neuropsychiatric disorders such as schizophrenia in 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 variants in *TRRAP* 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. 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 with a variant inherited 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) identified in five individuals and p.(Glu1106Lys) and p.(Gly1883Arg) identified in two individuals each. All the variants were predicted to be deleterious by CADD[23] (scaled C-scores over 20) and variously predicted pathogenic by SIFT[24] and PolyPhen-2 HVAR[25]. As shown in **figure 2A,** the 17 variants seen in our subjects had significantly increased CADD scores compared to 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, assessing 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 may be deleterious but were not clearly pathogenic, since either the inheritance pattern could not be determined, the variant was present in gnomAD or lead to another missense change at the same residue as a variant reported in gnomAD, or were 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=4.2x10-6) based on denovolyzer[26]. Nevertheless, the current number of 22 detected *de novo* variants in *TRRAP* is not genome-wide significant (p=0.08) after correcting for: 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 modelling 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 thirteen 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 not yet characterized. We performed a statistical clustering analysis comparing the mean distance in observed variants to ten million permutations of random variants, as previously described[27]. This analysis revealed a significant clustering of variants along the primary sequence of *TRRAP* (p value = $9x10^{-8}$), suggesting a model where specific domains are affected, haploinsufficiency being 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 visualisation[28,29], which highlights several of these features,

particularly for individuals with variants clustering with the recurrent p.(Ala1043Thr) variant (**Figure 3R**). All the individuals had developmental delay, with a highly variable severity of intellectual disability (ID). 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**) with the highest incidence of malformations being seen in thirteen 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 apparently aggregated in another region, 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).

Thirteen 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 10/13 individuals and showed structural brain anomalies in seven of them (70%), 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 myelination (1/10), and corpus callosum hypoplasia (1/10).

Neurological examination revealed hypotonia in 31% (4/13) of individuals. Only one individual was reported with epilepsy. Seven individuals (54%) were reported to have feeding difficulties and were fed exclusively by gastrostomy tube. Among the 10 individuals who were examine by echocardiography, 70% (7/10) had abnormal results with 50% (5/10) having ventricular septal defect, 30% (3/10) patent ductus arteriosus, 30% (3/10) patent foramen ovale, 20% (2/10) pulmonary hypertension, and 20% (2/10) 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, duplicate kidney and/or small kidney, was described in 60% (6/10) 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 anomalies of external genitalia, including microphallus, hypoplastic scrotum, and cryptorchidism, while two females (2/7; 29%) had a duplicated vagina and/or uterus. Other observed anomalies included dysplastic nails (8/13; 62%), cleft lip/palate (5/13; 38%), clinodactyly of the 5th finger (4/13; 31%), laryngo/tracheomalacia (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 in individuals 1 and 19). Four individuals (4/13; 31%) had visual impairment and three (3/13; 23%) had hearing impairment which has associated with inner ear malformations in 2 cases. Recurrent infections, mainly respiratory and urinary tract infections, affected 3/13 (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 autism spectrum disorder (ASD), and another three individuals (3/11;

27%) had some findings of ASD but no formal diagnosis. Eight of 11 (73%) had developmental delay and mild to severe intellectual disability and three had speech delay but whose IQ was measured above 70, including two that 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 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 embryonic stem cell pluripotency[31,32] which 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 Supplementary notes. 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 differentially expressed genes (DEGs), with a Log_2 Fold Change (Log_2FC) higher than 2 or lower than -2, and a p-value adjusted for 10%FDR lower than 0.01 (padj) (**supplementary notes, 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 using the GOrilla web application indicates an enrichment for the adrenergic receptor signalling pathway, genes important for neurological function, and potassium and ATP-sensitive ion transporters (**Figure S2B, Supplementary notes, Table S5**). The two individuals tested carried variants outside the cluster associated with the more syndromic ID; assuming distinct effects on gene regulation, it will be worth comparing gene expression between the two groups. Finally, we cannot exclude that the transcriptome alteration may be caused by a mechanism other than impaired HAT activity, since it has been shown that TRRAP as direct interactions with different partners not related to HAT complex. Thus, we highlighted candidate pathways that may be useful to uncover the pathomechanism of *TRRAP* variants in future studies.

TRRAP acts as a scaffold in HAT complexes. Although it has not a direct role in acetylation, we hypothesize that pathogenic effect of variants may 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 genes encoding histone acetyltransferase complexes, 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 alterations in several lysine acetyltransferase (KAT) and lysine deacetylase (KDAC) genes including *BRD1*, *HDAC4*, *HDAC6,* and HDAC9 and non-syndromic ASD and/or schizophrenia[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 with 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 intellectual disability. Based on the ExAC dataset without studies

on neuropsychiatric disorders, *TRRAP* is in the top five genes most intolerant to missense variants in the human genome, with a missense z-score of 10.1 (Ref. [22]). While 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 that milder phenotypes may be underrepresented in current exomes 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 expect to identify at least some LOF variants if haploinsufficiency of *TRRAP* was the causal mechanism. In DECIPHER (accessed 2018-05-14), no small or intragenic deletions involving *TRRAP* have been identified. Thus, taking into account the significant clustering, it suggests that missense variants may act either as gain of function or dominant–negative and that haploinsufficiency of *TRRAP* is likely to be prenatally lethal, although we cannot exclude a LOF effect of non-clustering variants associated with a milder phenotype.

TRRAP participates in embryonic development, as demonstrated by its binding with proteins regulating signaling pathways such as Notch in fruit fly[51], Ras in *C. elegans*[52], or Wnt by influencing beta-catenin ubiquitinylation[53]. We suspect therefore that TRRAP variants, more especially those falling within the 1031-1159 region, perturb the interactions with at least one of these developmental signaling pathways, which would explain the multiple malformation 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 less malformations (codons 1859-1932) were viable and may help explain the milder clinical phenotype associated with variants within this cluster. In mice, *Trrap* knock-out leads to early embryonic lethality[14] and a neural cell-specific conditional *Trrap* knockout line[16] revealed premature differentiation of neural progenitors, with 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 that 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 to dissect which gene expression changes are connected to which *TRRAP*-related specific phenotypes.

Supplemental Data

Supplemental Data include two figures and five tables and can be found with this article online.

Acknowledgments

We would like to thank all families for participating in this study. We acknowledge HUGODIMS consortium, which was supported by a grant from the French Ministry of Health and from the Health Regional Agency from Poitou-Charentes (HUGODIMS, 2013, RC14_0107). We are grateful to 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. We would like to thank the members of the Canadian Center for Computational Genomics and the McGill University and Génome Québec Innovation Center for their help in bioinformatics analysis. This work was also supported by funds from 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. and from the National Institutes of Health/Eunice Kennedy Shriver National Institute of Child Health and Human Development grant (HD064667) to D.A.S. This work was supported in part by NINDS R35 NS105078 to JRL and jointly funded National Human Genome Research Institute (NHGRI), and National Heart, Lung, and Blood Institute (NHLBI) grant to the Baylor-Hopkins Center for Mendelian Genomics (UM1 HG006542). JEP was supported by NHGRI K08 HG008986. VS and JS were supported by the Duke Genome Sequencing Clinic grant. We thank the CIHR and FRSQ for clinician-scientist awards to PMC. We would also like to thank the Mayo Clinic Center for Individualized Medicine (CIM) for supporting this research through the CIM Investigative and Functional Genomics Program. E.E.E. is an investigator of the Howard Hughes Medical Institute. The CAUSES Study is funded by Mining for Miracles, British Columbia Children's Hospital Foundation, and Genome British Columbia. CAUSES Study investigators include Shelin Adam, Christele Du Souich, Alison Elliott, Anna Lehman, Jill Mwenifumbo, Tanya Nelson, Clara Van Karnebeek, and Jan Friedman. LGB, JJJ, and JCS were supported by the intramural research program of the National Human Genome Research Institute, grant HG200328 12. We acknowledge also the DECIPHER Consortium which contributed to the exchange of genetic and clinical data between the teams. This work

was supported, in part, by US National Institute of Mental Health grant R01MH101221 to E.E.E.

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

Declaration of interest

E.E.E. is on the scientific advisory board (SAB) of DNAnexus, Inc. The Department of Molecular and Human Genetics at Baylor College of Medicine receives revenue for clinical genetic testing completed at Baylor Genetics laboratory. KMc, KR, JZ, MD, AT, AB, and IMW 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.

Web Resources

ExAC Browser, http://exac.broadinstitute.org/

gnomAD, http://gnomad.broadinstitute.org/

Ensembl VEP, http://grch37.ensembl.org/Homo_sapiens/Tools/VEP

GenBank, http://www.ncbi.nlm.nih.gov/genbank/

OMIM, http://www.omim.org/

Phyre2, http://www.sbg.bio.ic.ac.uk/phyre2/

UniProt, http://www.uniprot.org/uniprot/

DECIPHER, https://decipher.sanger.ac.uk/

Tables

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

Table 1. *De novo* variants in *TRRAP* identified in 24 individuals. RefSeq transcript used for *TRRAP* is 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.

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 (<-2.5SD)	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/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 impairement	3/24 - 12%	3/13 - 23%	0/11 - 0%
Visual impairement	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%

Table 2: Clinical description of individuals with variants inside or outside the 1031-1159 cluster.

Figure titles and legends

Figure 1. Genotype-phenotype correlation associated with *TRRAP* variants. (**A**) Predicted *de novo* and apparently *de novo* variants in affected subjects represented on the TRRAP protein. The variants in red represent individuals with apparent intellectual disability (ID) and malformations, in purple individuals with isolated ID with or without autism spectrum disorder (ASD), and in blue 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 (NP_001231509.1) predicted by PHYRE2 Protein Fold Recognition Server[56] represented by UCSF Chimera[57]. Mutated residues in the cluster 1031-1159 cluster are shown. FAT: FRAP, ATM and TRRAP; PIKK-like: phosphatidylinositol 3-kinase-related protein kinase-like; FATC: FRAP, ATM, TRRAP C-terminal.

Figure 2. *TRRAP* sequence is intolerant to missense variants. (**A**) CADD scores of the 17 variants identified in affected individuals compared to gnomAD singleton missense variants. The individuals' variants were compared to variants seen once in gnomAD, in order to avoid CADD training circularity. (**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 using exome-sequencing standing variation data in the gnomAD database, version 2.0. MTR data downloaded from http://mtr-viewer.mdhs.unimelb.edu.au/. An MTR = 1 (blue dashed line) represents neutrality (i.e., observing the same proportion of missense variants in the window as expected based on the underlying sequence context).

Red segments of MTR plot have achieved exome-wide FDR <0.10 for a significance test of a window's deviation from MTR = 1. Black dashed line (gene-specific median MTR), brown dashed line (gene-specific 25th centile MTR) and orange dashed line (gene-specific 5th centile MTR). The location of our 23 case-ascertained *de novo* variants is 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); 4: p.(Arg1035Gln); 5: p.(Ser1037Arg); 6: p.(Ala1043Thr); 7: p.(Glu1104Gly); 8: p.(Glu1106Lys); 9: p.(Gly1111Trp); 10: p.(Gly1159Arg); 11: p.(Arg1859Cys); 12: p.(Trp1866Arg); 13: p.(Trp1866Cys); 14: p.(Gly1883Arg); 15: p.(Pro1932Leu); 16: p.(Arg3757Gln). We found that case *de novo* variants were significantly enriched in the intolerant 50% of *TRRAP*'s protein-coding sequence with 18 (78%) of the 23 *de novo* events affecting the 50% most intolerant *TRRAP* sequence (Binomial exact test p=0.01). Strikingly, only the most recurring *de novo* missense variant (NM_001244580.1 p.(Ala1043Thr)) resided outside of the intolerant *TRRAP* sequence. (**C**) Localization of the TRRAP residues affected in the individuals in 3D protein models, including 14 out of 17 likely pathogenic variants and 2 out of 6 additional variants of unknown significance. The representation of the structure of human TRRAP (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 and NuA4 coactivator subunit Tra1 present in the protein data bank (PDB ID: 5OJS). Variants in regions non-homologous to Tra1 are not represented. Structure representation was made with UCSF Chimera.

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. (**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 single transverse palmar crease on the fourth picture from the left. (**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 age 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 age 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 eyebrow, 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 visualisation of 9 healthy age and gender matched controls on left, on right 9 individuals with variants in the 1031-1159 cluster. Facial images flipped and aligned to preserve bi-lateral asymmetry.

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. (2018). TRRAP is essential for regulating the accumulation of mutant and wild-type p53 in lymphoma. Blood.

11. Wei, X., Walia, V., Lin, J.C., Teer, J.K., Prickett, T.D., Gartner, J., Davis, S., NISC Comparative Sequencing Program, Stemke-Hale, K., Davies, M.A., et al. (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., et al. (2013). Histone acetyltransferase cofactor Trrap maintains self-renewal and restricts differentiation of embryonic stem cells. Stem Cells Dayt. Ohio *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., 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.

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. (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. Editor. Board Jonathan Haines Al *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., 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. (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.-H., 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. EJHG *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*, D1016-1022.

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.A., Clair, D.S., Blackwood, D.H., International Schizophrenia Consortium, McCaig, C.D., and Shen, S. (2012). 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.

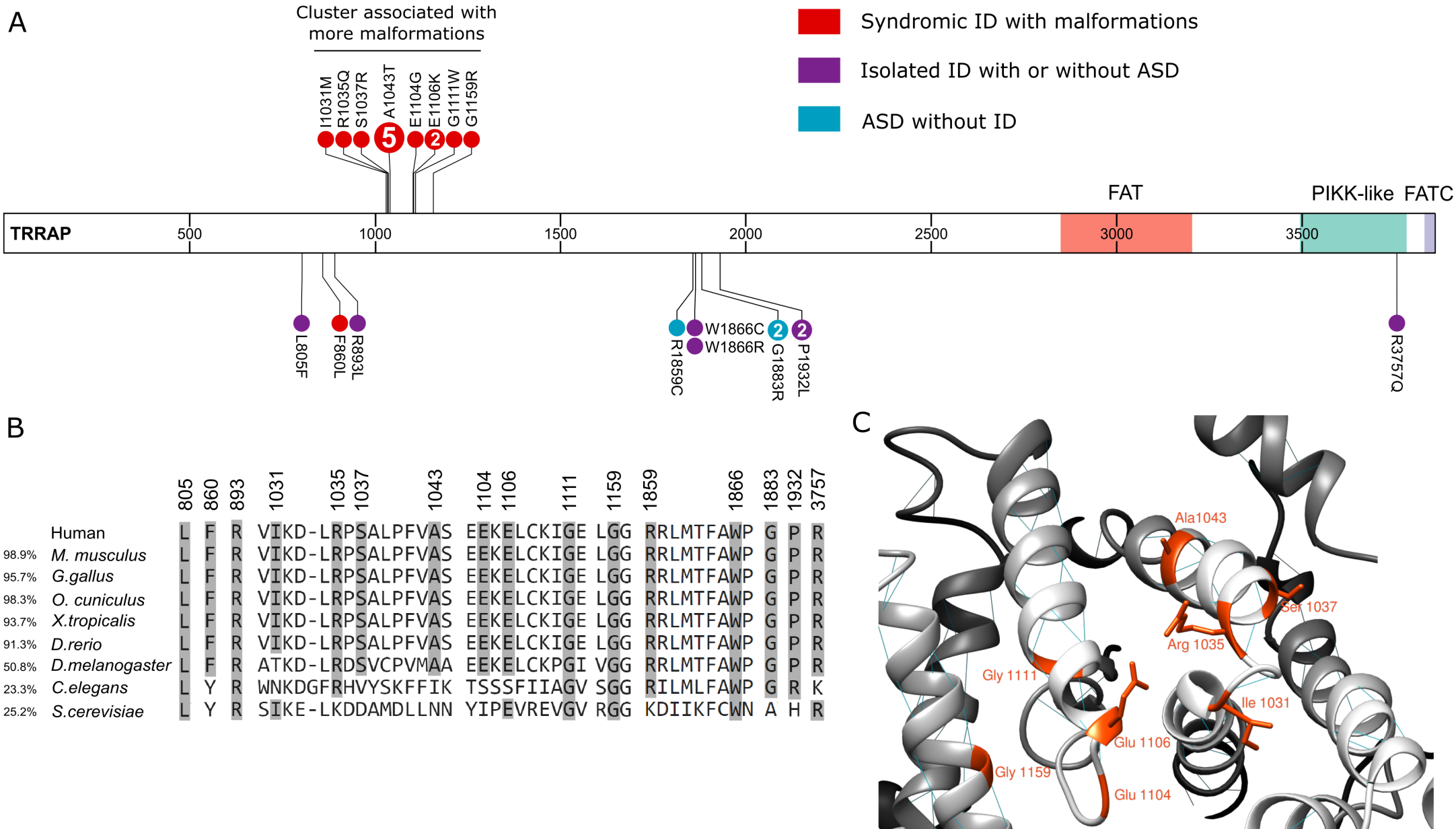
53. Finkbeiner, M.G., Sawan, C., Ouzounova, M., Murr, R., and Herceg, Z. (2008). HAT cofactor TRRAP mediates beta-catenin ubiquitination on the chromatin and the regulation of the canonical Wnt pathway. Cell Cycle Georget. Tex *7*, 3908–3914.

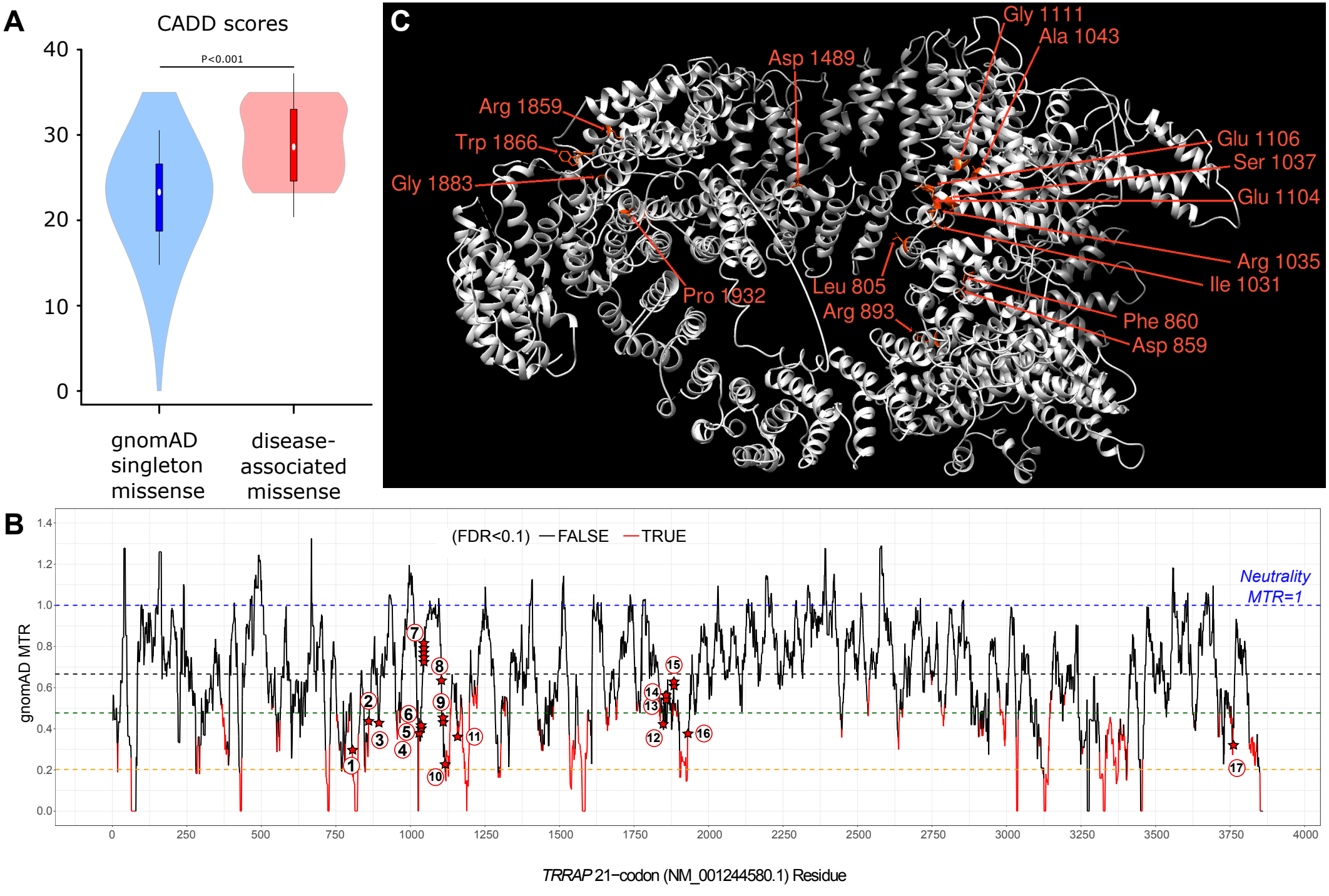
54. Knutson, B.A., and Hahn, S. (2011). Domains of Tra1 important for activator recruitment and transcription coactivator functions of SAGA and NuA4 complexes. Mol. Cell. Biol. *31*, 818–831.

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.

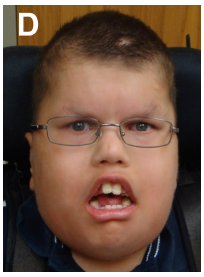

A
Cluster associated with more malformations
I1031M
R1035Q
S1037R
A1043T
E1104G
E1106K
G1111W
G1159R
5
2
Syndromic ID with malformations
Isolated ID with or without ASD
ASD without ID
FAT
PIKK-like
FATC
TRRAP
500
1000
1500
2000
2500
3000
3500
L805F
F860L
R893L
R1859C
W1866C
W1866R
G1883R
P1932L
R3757Q
B
805 860 893 1031 1035 1037 1043 1104 1106 1111 1159 1859 1866 1883 1932 3757
Human L F R VIKD-LRPSALPFVAS EEKELCKIGE LGG RRLMTFAWP G P R
98.9% M. musculus L F R VIKD-LRPSALPFVAS EEKELCKIGE LGG RRLMTFAWP G P R
95.7% G.gallus L F R VIKD-LRPSALPFVAS EEKELCKIGE LGG RRLMTFAWP G P R
98.3% O. cuniculus L F R VIKD-LRPSALPFVAS EEKELCKIGE LGG RRLMTFAWP G P R
93.7% X.tropicalis L F R VIKD-LRPSALPFVAS EEKELCKIGE LGG RRLMTFAWP G P R
91.3% D.rerio L F R VIKD-LRPSALPFVAS EEKELCKIGE LGG RRLMTFAWP G P R
50.8% D.melanogaster L F R ATKD-LRDSVCPVMAA EEKELCKPGI VGG RRLMTFAWP G P R
23.3% C.elegans L Y R WNKDGFRHVYSKFFIK TSSSFIIAGV SGG RILMLFAWP G R K
25.2% S.cerevisiae L Y R SIKE-LKDDAMDLLNN YIPEVREVGV RGG KDIIKFCWN A H R
C
Ala1043
Ser 1037
Arg 1035
Gly 1111
Ile 1031
Glu 1106
Gly 1159
Glu 1104


A
CADD scores
P<0.001
40
30
20
10
0
gnomAD singleton missense
disease-associated missense
C
Gly 1111
Ala 1043
Asp 1489
Arg 1859
Trp 1866
Gly 1883
Glu 1106
Ser 1037
Glu 1104
Arg 1035
Ile 1031
Leu 805
Arg 893
Pro 1932
Phe 860
Asp 859
B
(FDR<0.1) FALSE TRUE
Neutrality
MTR=1
gnomAD MTR
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
250
500
750
1000
1250
1500
1750
2000
2250
2500
2750
3000
3250
3500
3750
4000
TRRAP 21−codon (NM_001244580.1) Residue

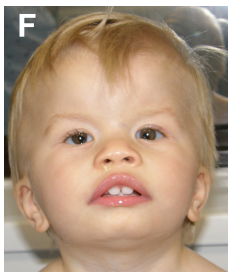

A

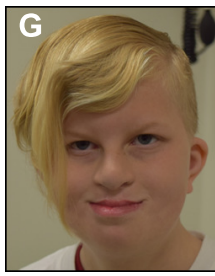

B


C


D


E


F


G


H


I


J


K


L


M


N


O


P


Q


R