Germline De Novo Mutations in *GNB1* Cause Severe Neurodevelopmental Disability, Hypotonia, and Seizures

Slavé Petrovski,[1,2,*] Sébastien Küry,[3] Candace T. Myers,[4] Kwame Anyane-Yeboa,[5] Benjamin Cogné,[3] Martin Bialer,[6] Fan Xia,[7] Parisa Hemati,[1] James Riviello,[1] Michele Mehaffey,[4] Thomas Besnard,[3] Emily Becraft,[8] Alexandrea Wadley,[9] Anya Revah Politi,[1] Sophie Colombo,[1] Xiaolin Zhu,[1] Zhong Ren,[1] Ian Andrews,[10] Tracy Dudding-Byth,[11,12] Amy L. Schneider,[13] Geoffrey Wallace,[14] University of Washington Center for Mendelian Genomics, Aaron B.I. Rosen,[4] Susan Schelley,[8] Gregory M. Enns,[8] Pierre Corre,[15] Joline Dalton,[16] Sandra Mercier,[3] Xénia Latypova,[3] Sébastien Schmitt,[3] Edwin Guzman,[1] Christine Moore,[6] Louise Bier,[1] Erin L. Heinzen,[1] Peter Karachunski,[16] Natasha Shur,[17] Theresa Grebe,[18] Alice Basinger,[19] Joanne M. Nguyen,[20] Stéphane Bézieau,[3] Klaas Wierenga,[9] Jonathan A. Bernstein,[8] Ingrid E. Scheffer,[13,21,22] Jill A. Rosenfeld,[7] Heather C. Mefford,[4] Bertrand Isidor,[3] and David B. Goldstein[1,*]

Whole-exome sequencing of 13 individuals with developmental delay commonly accompanied by abnormal muscle tone and seizures identified de novo missense mutations enriched within a sub-region of *GNB1*, a gene encoding the guanine nucleotide-binding protein subunit beta-1, Gβ. These 13 individuals were identified among a base of 5,855 individuals recruited for various undiagnosed genetic disorders. The probability of observing 13 or more de novo mutations by chance among 5,855 individuals is very low ($p = 7.1 \times 10^{-21}$), implicating *GNB1* as a genome-wide-significant disease-associated gene. The majority of these 13 mutations affect known Gβ binding sites, which suggests that a likely disease mechanism is through the disruption of the protein interface required for Gα-Gβγ interaction (resulting in a constitutively active Gβγ) or through the disruption of residues relevant for interaction between Gβγ and certain downstream effectors (resulting in reduced interaction with the effectors). Strikingly, 8 of the 13 individuals recruited here for a neurodevelopmental disorder have a germline de novo *GNB1* mutation that overlaps a set of five recurrent somatic tumor mutations for which recent functional studies demonstrated a gain-of-function effect due to constitutive activation of G protein downstream signaling cascades for some of the affected residues.

Neurodevelopmental disability represents a collection of clinically and biologically heterogeneous disorders. However, when a genetic disorder lacks distinguishing clinical features, it has proven challenging to stratify populations of affected individuals for phenotypically driven gene discovery. A recent strategy that has proven powerful for discovering disease-associated genes among more clinically heterogeneous populations has been large whole-exome sequencing studies on individuals presenting with relatively non-specific clinical features and identifying genes with excess de novo mutations in the affected population.[1–3] Clinical sequencing labs are finding that a third of individuals recruited for undiagnosed genetic disorders are later diagnosed on the basis of the results of trio whole-exome sequencing and that the majority are explained by pathogenic de novo mutations.[3–6]

Here, we describe 13 individuals (Table 1) with a previously unrecognized genetic condition characterized by global developmental delay (13/13), hypotonia (11/13), seizures (10/13), ophthalmological manifestations (8/13), and growth delay (6/13) accompanied by additional variable symptoms (Table 2). Clinical summaries for all 13 affected individuals are available in the Supplemental Note. These 13 individuals were found to have a de novo missense mutation in *GNB1* (MIM: 139380), the gene encoding guanine nucleotide-binding protein (G protein) subunit beta-1 (Gβ).

GNB1 encodes a ubiquitously[7] present β subunit of heterotrimeric G proteins, formed by its association with

[1]Institute for Genomic Medicine, Columbia University, New York, NY 10032, USA; [2]Department of Medicine, Austin Health and Royal Melbourne Hospital, University of Melbourne, Melbourne, VIC 3050, Australia; [3]Service de Génétique Médicale, Centre Hospitalier Universitaire Nantes, Nantes 44093, France; [4]Division of Genetic Medicine, Department of Pediatrics, University of Washington, Seattle WA, 98195, USA; [5]Division of Clinical Genetics, Department of Pediatrics, Columbia University Medical Center, New York, NY 10032, USA; [6]Division of Medical Genetics, Northwell Health, Manhasset, NY 11030, USA; [7]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA; [8]Division of Medical Genetics, Department of Pediatrics, Stanford University School of Medicine, Stanford, CA 94305, USA; [9]Section of Genetics, Department of Pediatrics, University of Oklahoma, Oklahoma City, OK 73019, USA; [10]School of Women's and Children's Health, University of New South Wales, Kensington, NSW 2052, Australia; [11]Genetics of Learning Disability Service, Hunter Genetics, Waratah, NSW 2298, Australia; [12]Priority Research Centre GrowUpWell, University of Newcastle, Callaghan, NSW 2308, Australia; [13]Department of Medicine, Austin Health, University of Melbourne, Heidelberg, VIC 3081, Australia; [14]Department of Neurosciences, Royal Children's Hospital, Herston School of Medicine, University of Queensland, Brisbane, QLD 4072, Australia; [15]Service de Stomatologie, Centre Hospitalier Universitaire Nantes, Nantes 44093, France; [16]Department of Neurology, University of Minnesota, Minneapolis, MN 55454, USA; [17]Division of Genetics, Department of Pediatrics, Albany Medical Center, Albany, NY 12208, USA; [18]Phoenix Children's Hospital and Department of Child Health, University of Arizona College of Medicine, Phoenix, AZ 85724, USA; [19]Cook Children's Physician Network, Fort Worth, TX 76102, USA; [20]Division of Medical Genetics, Department of Pediatrics, University of Texas Health Science Center at Houston, Houston, TX 77030, USA; [21]Florey Institute for Neuroscience and Mental Health, University of Melbourne, Parkville, VIC 3050, Australia; [22]Department of Paediatrics, Royal Children's Hospital, University of Melbourne, Melbourne, VIC 3050, Australia

*Correspondence: slavep@unimelb.edu.au (S.P.), dg2875@cumc.columbia.edu (D.B.G.)

http://dx.doi.org/10.1016/j.ajhg.2016.03.011.

Table 1. Comparison of Symptoms among 13 Individuals

Symptoms	Individual 1	Individual 2	Individual 3	Individual 4	Individual 5	Individual 6
Mutation (GRCh37; GenBank: NM_002074.4)	chr1: g.1737954T>C	chr1: g.1737948T>C	chr1: g.1737942A>T	chr1: g.1737942A>G	chr1: g.1735987T>C	chr1: g.1718817C>T
	c.227A>G	c.233A>G	c.239T>A	c.239T>C	c.301A>G	c.976G>A
	p.Asp76Gly	p.Lys78Arg	p.Ile80Asn	p.Ile80Thr	p.Met101Val	p.Ala326Thr
Ethnicity	Ashkenazi	mixed Ashkenazi and Sephardic Jewish	mixed Ashkenazi and Sephardic Jewish	North African	European	European
Gender, Age	M, 8.5 years	M, 13 months	M, 5 years	F, 4 years	M, 19 years	F, 17 years
Developmental delay	yes	yes	yes	yes	yes	yes
Abnormal muscle tone	yes (persistent hypotonia)	yes (initially presented with hypotonia, followed by improved tone in extremities)	yes (profound hypotonic quadriparesis, central hypotonia)	severe hypotonia of axis and lower limbs, dystonic hypertonia of upper limbs	no	mild hyper-reflexia, mild hypotonia
Epilepsy	no	infantile spasms	9 months: generalized epilepsy onset	no	infantile spasms, absence seizures, nonconvulsive status epilepticus, focal seizures	1 year: febrile status epilepticus followed by afebrile focal convulsive seizures and absence and atonic seizures; 14 years: focal hyperkinetic sleep-related seizures
EEG	8 years: normal	6 months: hypsarrhythmia, VEEG captured infantile spasms	15 months: generalized subcortical alterations; 3 years: generalized epileptiform process	unknown	7, 10, and 11 months: modified hypsarrhythmia; 2 years: normal; 3 years: multifocal and 2 Hz GSW	14 months: normal; 5 years: L and R posterior temporal epileptiform activity
MRI	3 years: normal	6 months: normal	14 months: bilateral polymicrogyria, delayed myelination	11 months: normal	scattered, non-specific T2 white-matter abnormalities	11 months: normal
Ophthalmological Disorder	no	strabismus	cortical vision impairment	slow ocular pursuit response	no	no
Growth delays: height and weight	height at 10^{th}–25^{th} percentile, weight at 10^{th} percentile	height < 3^{rd} percentile, weight at 12^{th} percentile	failure to thrive: weight and height below 5^{th} percentile	height at 2^{nd} percentile, weight at 1^{st} percentile	height just above 75^{th} percentile, weight just below 50^{th} percentile	height at 10^{th}–25^{th} percentile, weight just above 3^{rd} percentile
Microcephaly	no	no	no	no	no	no
Macrocephaly	OFC at 75^{th}–90^{th} percentile	no	no	no	no	no
Dysmorphic features	no	no	mild hypotonic face, outward flexed and deviated great toes	cleft palate	no	prominent central incisors, short philtrum, long tapering fingers, narrow hands, joint hypermobility
Additional features	vocal tics and ADHD, paternally inherited 259 kb interstitial deletion at 22q11.1	no	delayed hearing potentials	peripheral hypothyroidism, asthma, adducted thumbs	autism spectrum disorder, behavioral problems	maternally inherited 7q11.21 0.3 Mb duplication, aggressive behavior, drooling

Abbreviations are as follows: M, male; F, female; VEEG, video electroencephalogram; GSW, generalized spike-and-wave; L, left; R, right; OFC, occipitofrontal circumference; and ADHD, Attention-deficit hyperactivity disorder.

subunits Gα and Gγ, which are essential to the signaling function of G-protein-coupled receptors (GPCRs). The association between Gα and Gβγ covers the interaction sites on both the α subunit and the Gβγ dimer, thus preventing effector interactions and rendering the G protein inactive. A ligand binding to the GPCR promotes activation where Gα disassociates from the Gβγ dimer and the GPCR. The disassociation makes Gβγ free to regulate various effector proteins and signaling cascades, including interactions with a variety of enzymes and ion channels. Some effectors directly regulated by Gβγ include activation of adenylyl cyclase 2, interactions with β-adrenergic receptor kinase 1, inhibition of calcium channels, activation of potassium channels, activation of phospholipase C-β2, and activation of class IB phosphoinositide 3-kinases.[8] Yoda and colleagues recently used proteomic analysis to show that

Individual 7	Individual 8	Individual 9	Individual 10	Individual 11	Individual 12	Individual 13
chr1: g.1737953A>C	chr1: g.1735987T>C	chr1: g.1737952C>T	chr1: g.1737942A>G	chr1: g.1737942A>G	chr1: g.1737942A>T	chr1: g.1736004A>G
c.228T>G	c.301A>G	c.229G>A	c.239T>C	c.239T>C	c.239T>A	c.284T>C
p.Asp76Glu	p.Met101Val	p.Gly77Ser	p.Ile80Thr	p.Ile80Thr	p.Ile80Asn	p.Leu95Pro
American (European)	mixed European and fractional Filipino	mixed Mexican, German, and Spanish	American (European)	mixed European	American (European)	Hispanic
M, 8 years	F, 20 years	M, 11 years	M, 4 years	M, 10 years	M, 6 years	F, 4 years
yes	yes	yes	yes	yes	yes	yes
hypertonia (hemiplegia)	hypotonia at 2 years, gross motor milestone delay in toddlerhood, mildly low muscle tone in childhood	presented with hypotonia, diagnosed with hypotonic cerebral palsy (slowly improving)	6 months: poor head control, severe hypotonia, and non-ambulatory, severe oropharyngeal dysphagia	diffuse severe hypotonia, remains non-ambulatory	initially presented with hypotonia, has developed hypertonia in the extremities	hypotonia of axis, progression to hypertonia of extremities
4 years: hemiclonic seizures followed by complex partial status	11 years: onset of tonic-clonic seizures	no	2 years: onset of head drops and myoclonic seizures	9 years: onset of focal seizures	4 years: onset of focal seizures	4 months: tonic posturing of the arms and occasional generalized upper-body myoclonus
burst suppression	4 years: normal; 11 years: diffuse slowing and rare right temporal spikes	mild to moderate generalized slowing without focal, lateralizing, or epileptiform features	2 years: generalized paroxysmal epileptiform activity, myoclonic seizure with generalized paroxysmal polyspike slow-wave activity	9 years: multifocal spikes and generalized background slowing	4 years: high-amplitude background with occipital slowing, multifocal epileptiform discharges	4 months: multifocal epileptiform discharges from both posterior quadrants, right temporoparietal sharp waves
abnormal (see Supplemental Note)	3 and 11 years: normal	questionable cortical thickening on left side along sylvian fissures	7 months: normal; 2.5 years: abnormal (see Supplemental Note)	normal	1 month: normal	4 months: normal
strabismus, no fixing or following	strabismus (intermittent left esotropia)	no	strabismus (intermittent left esotropia), multivectorial nystagmus (unspecified)	nystagmus	cortical vision impairment	horizontal and vertical nystagmus, cortical vision loss
failure to thrive	episode of unexplained weight loss at 15 years	height at 50th–75th percentile, weight at 10th–25th percentile	height and weight at 50th percentile at 3 years	height and weight at 1st percentile	height and weight < 1st percentile	height and weight < 3rd percentile
no	no	no	no	no	no	<3rd percentile
no	birth: OFC at 96th percentile	no	2 years: OFC = 54 cm; 3 years: OFC = 55 cm	no	no	no
pectus excavatum, high-arched palate	no	no	no	no	ears with overfolded helices, limited elbow extension, clenched hand	cleft palate
hydronephrosis due to bilateral ureteropelvic junction obstruction	2 years: autism, developmental regression of language; now: non-verbal with significant receptive language	pes planus, attention deficit hyperactivity disorder	severe receptive and expressive language disorder	2 years: mild bilateral hydronephrosis, regression of language; now: non-verbal	bilateral moderate to severe sensorineural hearing loss	maternally inherited 1.5 Mb loss of 4q34–3q35.1

mutant GNB1-expressing cells had increased activation of the AKT, mTOR, and ERK pathways.[9]

GNB1 has additional literature support for its candidacy as a gene associated with neurodevelopmental disease. It is a member of the N-methyl-D-aspartic acid (NMDAR) synaptic transmission gene set[10] and the set of 842 fragile X mental retardation protein (FMRP) target genes.[11] Both the NMDAR- and FMRP-associated gene sets have been repeatedly implicated in neurodevelopmental and neuropsychiatric disorders, such as epileptic encephalopathies,[1,2] autism,[12,13] and schizophrenia.[14] Moreover, *GNB1* directly interacts with multiple genes encoding disease-associated Gα subunits, including *GNAO1* (MIM: 139311; associated with epileptic encephalopathy [MIM: 615473]), *GNAI3* (MIM: 139370; associated with auriculocondylar syndrome 1 [MIM: 602483]), *GNAT2*

Table 2. Clinical Features Shared among Three or More Individuals with a *GNB1* Germline De Novo Mutation

Description of HPO Term	HPO Term	No. of Individuals (of 13)	c.227A>G (p.Asp76Gly)	c.228T>G (p.Asp76Glu)	c.229G>A (p.Gly77Ser)	c.233A>G (p.Lys78Arg)	c.239T>A (p.Ile80Asn) (of two)	c.239T>C (p.Ile80Thr) (of three)	c.284T>C (p.Leu95Pro)	c.301A>G (p.Met101Val) (of two)	c.976G>A (p.Ala326Thr)
Global developmental delay	HP:0001263	13	+	+	+	+	++	+++	+	++	+
Hypotonia	HP:0001252	11	+	–	+	+	++	+++	+	+	+
Seizures	HP:0001250	10	–	+	–	+	++	++	+	++	+
Growth delay	HP:0001510	6	–	–	–	+	++	++	+	–	–
Multifocal epileptiform discharges	HP:0010841	5	–	–	–	–	+	++	+	+	–
Expressive language delay	HP:0002474	4	–	–	–	–	–	++	+	+	–
Failure to thrive	HP:0001508	4	–	+	–	–	+	+	+	–	–
Feeding difficulties	HP:0011968	4	–	–	–	+	++	–	+	–	–
Intellectual disability	HP:0001249	4	+	–	–	–	–	–	+	+	+
Limb hypertonia	HP:0002509	4	–	+	–	–	+	+	+	–	–
Strabismus	HP:0000486	4	–	+	–	+	–	+	+	–	–
Cortical visual impairment	HP:0100704	3	–	–	–	–	++	–	+	–	–
Developmental regression	HP:0002376	3	–	–	–	–	–	+	–	++	–
Early-onset hypotonia	HP:0008947	3	+	–	–	+	+	–	–	–	–
EEG with generalized epileptiform discharges	HP:0011198	3	–	–	–	+	–	+	–	+	–
Focal seizures with impairment of awareness	HP:0002384	3	–	+	–	–	+	–	–	+	–
Generalized tonic-clonic seizures	HP:0002069	3	–	–	–	–	–	–	–	++	+
Inability to walk	HP:0002540	3	–	–	–	–	–	++	+	–	–
Nystagmus	HP:0000639	3	–	–	–	–	–	++	+	–	–

Abbreviations are as follows: HPO, Human Phenotype Ontology; and EEG, electroencephalogram.

(MIM: 139340; associated with achromatopsia 4 [MIM: 613856]), *GNAS* (MIM: 139320; associated with pseudohypoparathyroidism Ia [MIM: 103580), Ib [MIM: 603233], and Ic [MIM: 612462], pseudopseudohypoparathyroidism [MIM: 612463], and somatic-mosaic McCune-Albright syndrome [MIM: 174800]), *GNAL* (MIM: 139312; associated with dystonia 25 [MIM: 615073]), *GNAQ* (MIM: 600998; associated with somatic-mosaic Sturge-Weber syndrome [MIM: 185300]), and *GNA11* (MIM: 139313; associated with hypocalcemia [MIM: 615361] and type II hypocalciuric hypercalcemia [MIM: 145981]).

Consistent with an important developmental role, mice homozygous for a disrupted *Gnb1* Gt(prvSStrap)4B8Yiw (MGI: 4438390) experience partial perinatal lethality and lethality throughout fetal growth and development.[15,16] Moreover, embryos homozygous for mutant *Gnb1* show an excess of neurological phenotypes (including abnormal brain morphology, decreased cell proliferation, exencephaly, decreased brain size, abnormal neural plate morphology, abnormal embryonic neuroepithelium morphology, abnormal neuronal precursor proliferation, abnormal cortical ventricular zone morphology, and abnormal neuron differentiation), whereas pups without neural-tube defects surviving up to postnatal day 2 also display thin cerebral cortex and abnormal suckling behavior.[15]

GNB1 is found deleted in some individuals with chromosome 1p36 deletion syndrome (MIM: 607872), which includes intellectual disability, developmental delay, seizures (including infantile spasms), and muscular hypotonia. Five individuals with 200–823 kb overlapping 1p36.33 deletions were reported to have classic features of the 1p36.33 syndrome.[17] The smallest region of overlap found across the five individuals was a 174 kb stretch including *GNB1*, *CALML6*, *TMEM52*, *C1ORF222*, and *KIAA1751*.[17] The clinical features most commonly shared (by more than two of the five individuals) included straight eyebrows (n = 3), a broad nasal root (n = 3), a pointed chin (n = 3), abnormally low-set ears (n = 4), developmental delay (n = 5), expressive-language problems (n = 4), behavioral problems (n = 4), and neonatal hypotonia (n = 3). Our 13 individuals do not share the dysmorphologies reported among these individuals with 1p36.33 syndrome, but they do share neurodevelopmental and seizure features. Because the disease-associated gene is not yet known, this work further suggests that *GNB1* is a good candidate for playing a causal role in 1p36.33 syndrome. However, unlike the existing mouse knockout models and the human 1p36.33 deletion syndrome, which represent possible *GNB1* haploinsufficient phenotypes, the existing functional studies performed by Yoda and colleagues showed that some of our 13 missense de novo mutations within the Gβγ interaction region act as activating mutations that enhance or continuously activate downstream signaling pathways.[9]

Our 13 unrelated individuals with a de novo germline mutation in *GNB1* range from 13 months to 20 years of age (Table 1 and Supplemental Note). Six of the individuals had trio whole-exome sequencing (WES) performed after written informed consent was obtained through an institutional-review-board-approved research study at the Institute for Genomic Medicine at Columbia University (protocol AAAO8410, individuals 1–3), Centre Hospitalier Universitaire Nantes (individual 4), the Department of Pediatrics at the University of Washington (protocol 48919, individual 5), and the University of Melbourne (H2007/02961, individual 6).

For all six trios, DNA was extracted from maternal, paternal, and proband samples, exome sequenced on a HiSeq 2500 with the Kapa Biosystem's Library Preparation Kit, and whole-exome captured with Nimblegen SeqCap EZ v.3.0 (individuals 1–3), SeqCap EZ v.2.0 (individuals 5 and 6), or Agilent SureSelect v.5 (individual 4). Two paired-end 100 bp reads were used for the exome-capture sequencing. Raw reads were processed with pipelines based on the Genome Analysis Toolkit (GATK) best-practices protocol. For individuals 1–3, trio sequence data were analyzed with an updated version of our established trio sequencing framework,[4] which identifies "qualifying" genotypes not observed in the parents, 4,326 control individuals from the Institute for Genomic Medicine, or two external databases of 6,503 and 60,706 control individuals provided by the National Heart, Lung, and Blood Institute (NHLBI) Exome Sequencing Project (ESP6500SI [March 2013 release]) and the Exome Aggregation Consortium (ExAC Browser v.0.3 [January 2015 release]), respectively (Figure S1A). Individual 4's genotypes were screened against ESP6500SI, ExAC Browser v.0.3, the Genome of the Netherlands (GoNL) database, and the Japanese Human Genetic Variation Database, and a screening similar to that in Figure S1A was used. Genotypes of individuals 5 and 6 were screened against ESP6500SI, ExAC Browser v.0.3, dbSNP, and 1000 Genomes (Figure S1B). For individuals 1–3, qualifying variants were identified on the basis of four main modes of inheritance: compound heterozygous, newly homozygous, X-linked hemizygous, and putatively de novo (Figure S1).[4] We found that individual 1 had three, individual 2 had four, and individual 3 had two qualifying genotypes. Individuals 4 and 5 had two qualifying genotypes, whereas individual 6 had only the single qualifying genotype (Table S2).

For individuals 7–13, only the probands were sequenced with the proband-only exome sequencing approach from Baylor Miraca Genetics Laboratories (BMGL).[5,18] Upon identification of the *GNB1* de novo mutations in individuals 1–6, the probands and parents of trios 7–13 were Sanger sequenced for the *GNB1* missense variants, and this showed that the variants had arisen de novo in these seven additional individuals.

All 13 individuals were found to have a Sanger-validated *GNB1* de novo missense mutation (GenBank: NM_002074.4): c.227A>G (p.Asp76Gly) in individual 1, c.228T>G (p.Asp76Glu) in individual 7, c. 229G>A (p. Gly77Ser) in individual 9, c.233A>G (p.Lys78Arg) in individual 2, c.239T>A (p.Ile80Asn) in individuals 3 and

12, c.239T>C (p.Ile80Thr) in individuals 4, 10, and 11, c.284T>C (p.Leu95Pro) in individual 13, c.301A>G (p.Met101Val) in individuals 5 and 8, and c.976G>A (p.Ala326Thr) in individual 6. We found that the Combined Annotation Dependent Depletion (CADD) score predicted all *GNB1* de novo missense mutations affecting our 13 individuals to be among the 1% most deleterious substitutions in the human genome (Table S3).[19]

A list of Human Phenotype Ontology (HPO)[20] terms were defined for each of the 13 individuals (Table S4). For three GNB1 residues, we were able to use the HPO terms to ask whether individuals sharing a GNB1 substitution also share certain features more often with each other than with individuals affected at other GNB1 residues. We found that four of the five individuals with a p.Ile80 substitution reported growth delay (HP: 0001510), whereas six of the remaining eight individuals did not. Both dystonia and dysphagia were shared among two of five individuals with a p.Ile80 substitution but were not reported among the other eight individuals. The two individuals sharing a p.Met101 substitution were both diagnosed with autism spectrum disorder, which was not assigned among the remaining 11 individuals. These two individuals with a p.Met101 substitution were phenotyped at different recruitment sites. A single person (individual 3) also presented with polymicrogyria, a rare brain malformation frequently resulting from mutations affecting the PI3K-AKT-mTOR pathway.[21] Although these genotype-phenotype correlations are suggestive, reliably determining the phenotypic consequences of the different mutations will require a larger number of individuals with *GNB1*-related neurodevelopmental disease.

The identification of multiple missense de novo mutations in *GNB1* is remarkable given that *GNB1* is significantly depleted of functional variation in the general population—suggesting that *GNB1* might be under strong purifying selections as a result of critical homeostatic roles. Using the ExAC reference cohort, we tallied the unique *GNB1* consensus coding sequence (CCDS) protein-coding variants at 105 and the subset considered non-synonymous at 29 (27.6%). Given the sequence context of *GNB1*, we expect that 77.6% of all simulated (possible) *GNB1* variants would result in a non-synonymous effect. Thus, it is clear from the standing variation in the human population that *GNB1* is under strong purifying selection ($p = 2.1 \times 10^{-27}$, two-tailed binomial exact test). Multiple genic-intolerance metrics consistently support the premise that *GNB1* is intolerant to functional variation.[22–24] However, to formally assess whether we have securely implicated *GNB1* as a gene associated with disease, we adopted the framework established in the large trio sequencing efforts that have focused on estimating the mutation rate in genes and then asking whether the observed total burden of de novo mutations is beyond what could be explained by chance.[1,2,25,26] The total subject base from which our 13 individuals were identified is 5,855 individuals recruited from multiple sequencing clinics that have ascertained individuals with a presumed genetic disorder of unknown cause. Using the underlying mutation rate of *GNB1*,[27] we calculated the probability of observing 13 or more de novo mutations by chance among 5,855 individuals to be $p = 7.1 \times 10^{-21}$, which survives correction for the 18,669 protein-coding genes routinely tested in exome sequencing by approximately 15 orders of magnitude. Importantly, and independently of whether *GNB1* is considered an intolerant gene, this means that *GNB1* is now securely implicated as a genome-wide-significant disease-associated gene after we account for all considered individuals; however, the fact that *GNB1* is intolerant makes the mutational excess more remarkable, as does the clear localization of the mutations that cause disease.

Nine of the 13 *GNB1* mutations affect a residue coded for in exon 6 (chr1: g.1737912–1737979; GRCh37). Exon 6 represents only 6.4% of the total *GNB1* protein-coding sequence, and therefore finding 9 of the 13 mutations in this exon by chance is highly improbable ($p = 1.0 \times 10^{-8}$, two-tailed binomial exact test). The nine de novo *GNB1* (GenBank: NM_002074.4) mutations in exon 6 affect closely spaced residues: p.Asp76Gly, p.Asp76Glu, p.Gly77Ser, p.Lys78Arg, p.Ile80Asn, and p.Ile80Thr (Figure 1). The stretch of sequence that overlaps exon 6 is known to be the sequence relevant for the interaction between Gα and Gβγ in the G protein complex.[8]

Recently, *GNB1* somatic mutations were found to associate with hematological transformation and also with therapeutic resistance to different kinase inhibitors in the presence of additional driver oncogenic mutations.[9] Yoda and colleagues identified a catalog of 113 somatic mutations from tumor sequencing collections (83 unique variants), available in Table S1 from Yoda et al.[9] In particular, among their identified catalog of mutations, they identified five *GNB1* and *GNB2* (MIM: 139390) loci that they highlighted as sites of recurrent (more than two) somatic missense mutations. The five residues affected by somatic mutations were Lys57 (11 mutations), Lys78 (three mutations), Ile80 (eight mutations), Lys89 (four mutations), and Met101 (three mutations). All five of these residues are located along the Gβ protein surface that interacts with Gα subunits and downstream effectors.[8] Remarkably, 8 of our 13 individuals have a germline de novo *GNB1* missense mutation that overlaps one of these five recurrent oncogenic amino acid residues (Figure 1).[9] Through multiple lines of functional evidence, Yoda and colleagues showed that these recurrent somatic missense mutations confer cytokine-independent growth and activate canonical G protein signaling downstream of G proteins through the disruption of the Gα-Gβγ interaction interface and downstream effector interaction interfaces.[9] One of the mutations evaluated by Yoda and colleagues causes a p.Ile80Thr substitution—observed in 3 of our 13 individuals; through tandem affinity-purification and mass-spectrometry analyses using tagged wild-type and p.Ile80Thr

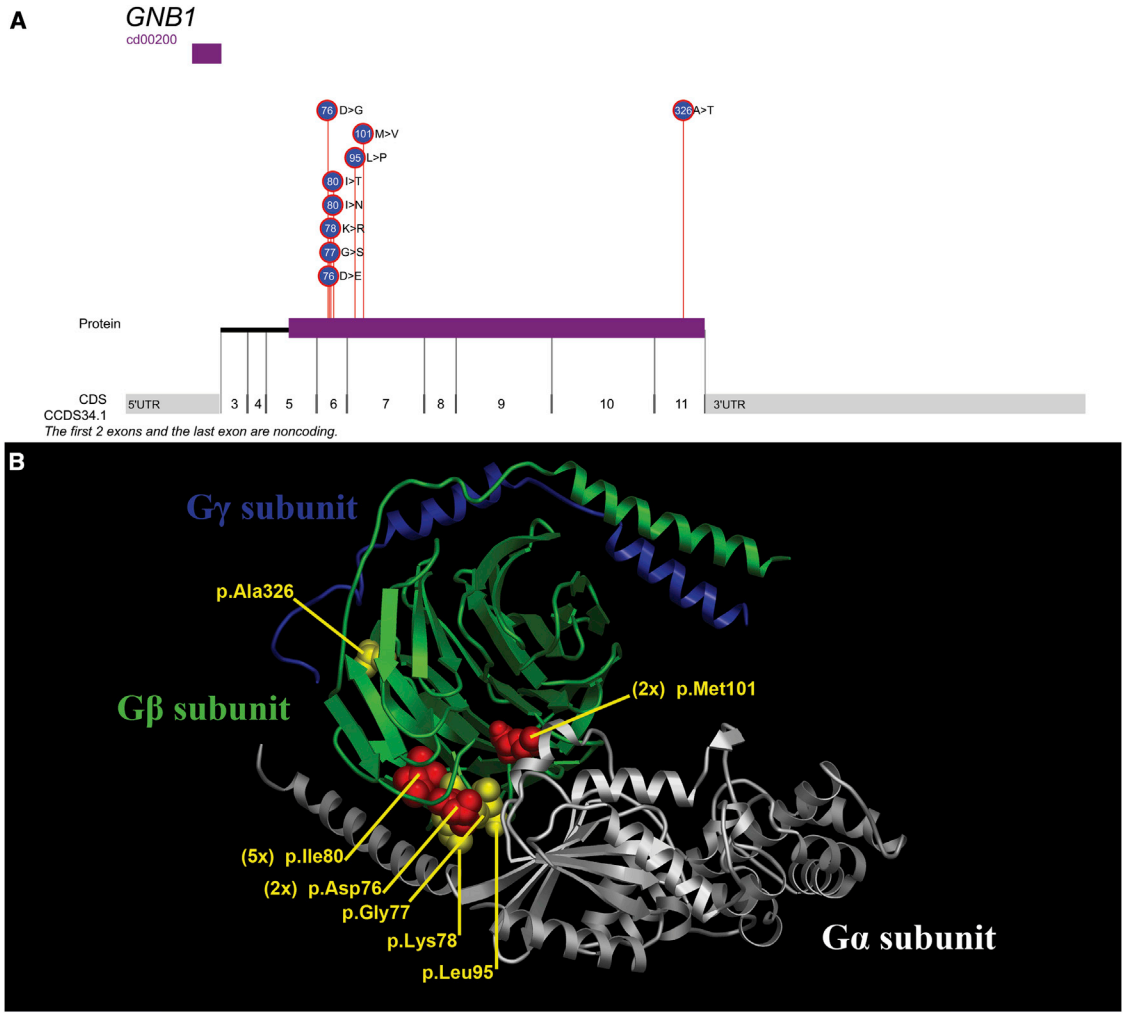

Figure 1. Localization of *GNB1* Mutations
(A) The distribution of the 13 mutations across *GNB1* (GenBank: NM_002074.4) shows a preferential enrichment of de novo mutations affecting exon 6. cd0020 represents the WD40 domain of GNB1.
(B) Molecular representation of a heterotrimeric G protein (Gα, white; Gβ, green; Gγ, blue) is based on a crystal structure (PDB: ID 1GP2; DOI: http://dx.doi.org/10.2210/pdb1gp2/pdb).[28] The Gβ side chains of the four residues affecting a single subject are indicated in yellow, whereas the three recurrently affected residues are indicated in red.

proteins, they demonstrated that the mutant p.Ile80Thr protein had reduced binding to almost all Gα subunits. Interestingly, among their data, they were also able to show that different *GNB1* alleles affected different protein interactors, suggesting variable downstream consequences of different *GNB1* mutations.[9] The secure implication of *GNB1* as a neurodevelopmental-disease-associated gene will enable future research to determine the biology of how mutations in *GNB1* contribute to the pathogenesis of neurodevelopmental disease.

Although none of our 13 individuals have a reported malignancy, risk of malignancy is something that warrants observation as the individuals are prospectively followed. Also, this wouldn't be the first time that an oncogene has been implicated in a developmental disorder. For example, mutations in genes that encode key components or regulators of pathways that can be found constitutively activated in many types of cancer (such as genes associated with RASopathies, the SWI-SNF complex, and the PI3K-Akt-mTOR pathway) are also known to be responsible for multiple groups of developmental disorders. *BRAF* (MIM: 164757) is one example in which somatic mutations are known to cause various cancer types, whereas germline mutations are a known cause of developmental disorders such as cardiofaciocutaneous syndrome (MIM:115150), Noonan syndrome (MIM: 613706), and LEOPARD syndrome (MIM: 613707). Moreover, there is evidence that a subset of *BRAF* mutations can cause both developmental disorders and cancer (c.1789C>G [p.Leu597Val] [ClinVar: 13969; GenBank: NM_004333.4]). To explore this further, we took the list of 53 *BRAF* missense mutations that were found to be recurrent (i.e., more than two observations reported among the COSMIC catalog of somatic mutation in cancer)[29] and found that 11 (~21%) of the somatic mutations found as recurrent mutations across tumor samples have also been linked to one or more of

cardiofaciocutaneous syndrome, Noonan syndrome, and undefined RASopathy within ClinVar or the Human Gene Mutation Database (HGMD) when the search is restricted to pathogenic curated variants.

The seminal work published on the structure of GNB1 provided experimental evidence on 15 GNB1 residues—including the precise residues affected in 8 of our 13 individuals—that mediate interactions with both Gα subunits and effector proteins. According to the crystal structures of heterotrimeric Gαβγ, two Gβγ regions of interaction with Gα are localized: the switch interface (Gβ residues 57, 59, 98, 99, 101, 117, 119, 143, 186, 228, and 332) and the amino-terminal interface (Gβ residues 55, 78, 80, and 89).[8,30] The *GNB1* mutation found in individual 2 (c.233A>G [p.Lys78Arg]) affects a Gβ residue reported to be important for regulating (1) activation of adenylyl cyclase 2, (2) inhibition of calcium channels, and (3) activation of potassium channels. The *GNB1* mutations found in individuals 3, 4, and 10–12 (c.239T>C [p.Ile80Thr] and c.239T>A [p.Ile80Asn]) affect a Gβ residue found important for (1) inhibition of calcium channels, (2) activation of potassium channels, (3) activation of phospholipase C-β2, and (4) Gα binding. The *GNB1* recurrent mutation found in individuals 5 and 8 (c.301A>G [p.Met101Val]) affects a Gβ residue found important for (1) activation of adenylyl cyclase 2, (2) inhibition of calcium channels, (3) activation of phospholipase C-β2, and (4) G-protein-coupled receptors.[8] This existing literature highlights multiple downstream signaling cascades that are potentially affected by the disruption of these specific residues. Although the relevance remains unclear, eight of our identified mutations affect the GNB1 residues found to be important for inhibition of calcium channels: Leu55, Lys78, Ile80, Met101, Asn119, Thr143, Asp186, and Trp332.[8] By plotting the distribution of disease-associated variants on a model of the structure of GNB1, we observed the tertiary clustering of these substitutions in the Gα-Gβγ binding region (Figure 1B).

Interestingly, the recurring *GNB1* germline de novo mutation (c.239T>C [p.Ile80Thr]) found among our 13 individuals has been reported in the ExAC Browser as a low-confidence singleton variant (ExAC variant 1-1737942-A-G). As highlighted by Yoda and colleagues,[9] the mutation causing the p.Ile80Thr substitution is among the most frequent somatic mutations reported in *GNB1* tumor genetics.[9] Indeed, as expected, visually assessing the variant calls of the single low-confidence ExAC observation makes it clear that the variant in the ExAC Browser, if not a technical artifact, is a post-zygotic mutation whereby only 10 variant reads out of 76 overlapping reads support the variant (i.e., ~13% of reads support the p.Ile80Thr variant in contrast to the expected 50%; $p = 3 \times 10^{-11}$, two-tailed binomial exact test).

Although oncogenic mutations that result in strong downstream signaling activation might not be compatible with survival as germline mutations, there is already evidence in Mendelian disease literature that some recurrent oncogenic mutations are viable but result in various developmental disorders.[31,32] It has been suggested that postnatal intervention to reduce Ras-MAPK activity could alleviate symptoms associated with RASopathies.[31] We cautiously speculate that this paradigm might also eventually be applicable to GNB1-related disorders. *GNB1* activating mutations studied by Yoda and colleagues were found to be responsive to the introduction of a small-molecule inhibitor of both PI3K and mTOR pathways (BEZ235), which they found suppressed PI3K-mTOR signaling and improved survival of mice with mutant *GNB1*-induced leukemia.[9]

This study highlights how sequencing large collections of undiagnosed individuals allows us to define the variable phenotypes associated with genes such as *GNB1*. Moreover, many of the *GNB1* mutations identified as germline de novo mutations in neurodevelopmental disease have been described as important sites in tumor genetics,[9] suggesting that these mutations occur at mitotically mutable GNB1 residues enriched among known Gβ interaction residues. Yet, despite the observed mitotic mutability, this *GNB1* sequence remains extremely intolerant to germline functional mutations in the human population.

Accession Numbers

The accession numbers for the *GNB1* variants reported in this paper are ClinVar: SCV000266332, SCV000266333, SCV000266334, SCV000266335, SCV000266336, SCV000266337, SCV000266338, SCV000266339, and SCV000266340.

Supplemental Data

Supplemental Data include a Supplemental Note, Supplemental Acknowledgments, one figure, and four tables and can be found with this article online at http://dx.doi.org/10.1016/j.ajhg.2016.03.011.

Received: December 12, 2015
Accepted: March 15, 2016
Published: April 21, 2016

Web Resources

Analysis Tool for Annotated Variants (ATAV), https://redmine.igm.cumc.columbia.edu/projects/atav/wiki
Burrows-Wheeler Aligner, http://bio-bwa.sourceforge.net/
Cassandra, https://www.hgsc.bcm.edu/cassandra-0/
ClinVar, http://www.ncbi.nlm.nih.gov/clinvar/
Consensus Coding Sequence (CCDS), https://www.ncbi.nlm.nih.gov/CCDS/
COSMIC, http://cancer.sanger.ac.uk/cosmic
Custom BMGL sequencing protocol, https://www.hgsc.bcm.edu/content/protocols-sequencing-library-construction/
Custom HUGODIMS pipeline, https://github.com/Oodnadatta/HUGODIMS-pipeline/
Ensembl genome assembly GRCh37, http://grch37.ensembl.org/Homo_sapiens/Info/Index
Ensembl Variant Effect Predictor (VEP), http://grch37.ensembl.org/Homo_sapiens/Tools/VEP

Exome Aggregate Consortium (ExAC) Browser, http://exac.broadinstitute.org/
GATK, https://www.broadinstitute.org/gatk/
Genic Intolerance, http://genic-intolerance.org/
Genome of the Netherlands (GoNL), http://www.nlgenome.nl/
Human Gene Mutation Database (HGMD), http://www.hgmd.cf.ac.uk/ac/index.php
Human Phenotype Ontology (HPO), http://compbio.charite.de/hpoweb/showterm?id=HP:0000118
Integrated Genomics Viewer (IGV), https://www.broadinstitute.org/igv/
Java-Based Utilities for Bioinformatics (JVarkit), http://dx.doi.org/10.6084/m9.figshare.1425030
Mouse Genome Informatics, http://www.informatics.jax.org/
NHLBI Exome Sequencing Project (ESP) Exome Variant Server, http://evs.gs.washington.edu/EVS/
OMIM, http://omim.org/
OpenAstexViewer, http://openastexviewer.net/web/
Picard, http://broadinstitute.github.io/picard
PolyPhen-2, http://genetics.bwh.harvard.edu/pph2/
Primer3Plus, http://www.primer3plus.com/cgi-bin/dev/primer3plus.cgi/
R: The R Project for Statistical Computing, https://www.r-project.org/
RCSB Protein Data Bank, http://www.rcsb.org/pdb/home/home.do
SAMtools, http://samtools.sourceforge.net/
SnpEff, http://snpeff.sourceforge.net/
STRING, http://string-db.org/
The Human Protein Atlas: Tissue Atlas, http://www.proteinatlas.org/tissue/

References

1. EuroEPINOMICS-RES Consortium; Epilepsy Phenome/Genome Project; Epi4K Consortium (2014). De novo mutations in synaptic transmission genes including DNM1 cause epileptic encephalopathies. Am. J. Hum. Genet. *95*, 360–370.
2. Allen, A.S., Berkovic, S.F., Cossette, P., Delanty, N., Dlugos, D., Eichler, E.E., Epstein, M.P., Glauser, T., Goldstein, D.B., Han, Y., et al.; Epi4K Consortium; Epilepsy Phenome/Genome Project (2013). De novo mutations in epileptic encephalopathies. Nature *501*, 217–221.
3. Deciphering Developmental Disorders Study (2015). Large-scale discovery of novel genetic causes of developmental disorders. Nature *519*, 223–228.
4. Zhu, X., Petrovski, S., Xie, P., Ruzzo, E.K., Lu, Y.F., McSweeney, K.M., Ben-Zeev, B., Nissenkorn, A., Anikster, Y., Oz-Levi, D., et al. (2015). Whole-exome sequencing in undiagnosed genetic diseases: interpreting 119 trios. Genet. Med. *17*, 774–781.
5. Yang, Y., Muzny, D.M., Reid, J.G., Bainbridge, M.N., Willis, A., Ward, P.A., Braxton, A., Beuten, J., Xia, F., Niu, Z., et al. (2013). Clinical whole-exome sequencing for the diagnosis of mendelian disorders. N. Engl. J. Med. *369*, 1502–1511.
6. Need, A.C., Shashi, V., Hitomi, Y., Schoch, K., Shianna, K.V., McDonald, M.T., Meisler, M.H., and Goldstein, D.B. (2012). Clinical application of exome sequencing in undiagnosed genetic conditions. J. Med. Genet. *49*, 353–361.
7. Fong, H.K., Hurley, J.B., Hopkins, R.S., Miake-Lye, R., Johnson, M.S., Doolittle, R.F., and Simon, M.I. (1986). Repetitive segmental structure of the transducin beta subunit: homology with the CDC4 gene and identification of related mRNAs. Proc. Natl. Acad. Sci. USA *83*, 2162–2166.
8. Ford, C.E., Skiba, N.P., Bae, H., Daaka, Y., Reuveny, E., Shekter, L.R., Rosal, R., Weng, G., Yang, C.S., Iyengar, R., et al. (1998). Molecular basis for interactions of G protein betagamma subunits with effectors. Science *280*, 1271–1274.
9. Yoda, A., Adelmant, G., Tamburini, J., Chapuy, B., Shindoh, N., Yoda, Y., Weigert, O., Kopp, N., Wu, S.C., Kim, S.S., et al. (2015). Mutations in G protein β subunits promote transformation and kinase inhibitor resistance. Nat. Med. *21*, 71–75.
10. Kirov, G., Pocklington, A.J., Holmans, P., Ivanov, D., Ikeda, M., Ruderfer, D., Moran, J., Chambert, K., Toncheva, D., Georgieva, L., et al. (2012). De novo CNV analysis implicates specific abnormalities of postsynaptic signalling complexes in the pathogenesis of schizophrenia. Mol. Psychiatry *17*, 142–153.
11. Darnell, J.C., Van Driesche, S.J., Zhang, C., Hung, K.Y., Mele, A., Fraser, C.E., Stone, E.F., Chen, C., Fak, J.J., Chi, S.W., et al. (2011). FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. Cell *146*, 247–261.
12. Iossifov, I., O'Roak, B.J., Sanders, S.J., Ronemus, M., Krumm, N., Levy, D., Stessman, H.A., Witherspoon, K.T., Vives, L., Patterson, K.E., et al. (2014). The contribution of de novo coding mutations to autism spectrum disorder. Nature *515*, 216–221.
13. Iossifov, I., Ronemus, M., Levy, D., Wang, Z., Hakker, I., Rosenbaum, J., Yamrom, B., Lee, Y.H., Narzisi, G., Leotta, A., et al. (2012). De novo gene disruptions in children on the autistic spectrum. Neuron *74*, 285–299.
14. Fromer, M., Pocklington, A.J., Kavanagh, D.H., Williams, H.J., Dwyer, S., Gormley, P., Georgieva, L., Rees, E., Palta, P., Ruderfer, D.M., et al. (2014). De novo mutations in schizophrenia implicate synaptic networks. Nature *506*, 179–184.
15. Eppig, J.T., Blake, J.A., Bult, C.J., Kadin, J.A., and Richardson, J.E.; Mouse Genome Database Group (2015). The Mouse Genome Database (MGD): facilitating mouse as a model for human biology and disease. Nucleic Acids Res. *43*, D726–D736.
16. Georgi, B., Voight, B.F., and Bućan, M. (2013). From mouse to human: evolutionary genomics analysis of human orthologs of essential genes. PLoS Genet. *9*, e1003484.
17. Rosenfeld, J.A., Crolla, J.A., Tomkins, S., Bader, P., Morrow, B., Gorski, J., Troxell, R., Forster-Gibson, C., Cilliers, D., Hislop, R.G., et al. (2010). Refinement of causative genes in monosomy 1p36 through clinical and molecular cytogenetic characterization of small interstitial deletions. Am. J. Med. Genet. A. *152A*, 1951–1959.
18. Bainbridge, M.N., Wang, M., Wu, Y., Newsham, I., Muzny, D.M., Jefferies, J.L., Albert, T.J., Burgess, D.L., and Gibbs, R.A. (2011). Targeted enrichment beyond the consensus coding DNA sequence exome reveals exons with higher variant densities. Genome Biol. *12*, R68.
19. 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.
20. Köhler, S., Doelken, S.C., Mungall, C.J., Bauer, S., Firth, H.V., Bailleul-Forestier, I., Black, G.C., Brown, D.L., Brudno, M., Campbell, J., et al. (2014). The Human Phenotype Ontology project: linking molecular biology and disease through phenotype data. Nucleic Acids Res. *42*, D966–D974.
21. Rivière, J.B., Mirzaa, G.M., O'Roak, B.J., Beddaoui, M., Alcantara, D., Conway, R.L., St-Onge, J., Schwartzentruber, J.A.,

Gripp, K.W., Nikkel, S.M., et al.; Finding of Rare Disease Genes (FORGE) Canada Consortium (2012). De novo germline and postzygotic mutations in AKT3, PIK3R2 and PIK3CA cause a spectrum of related megalencephaly syndromes. Nat. Genet. *44*, 934–940.

22. Petrovski, S., Wang, Q., Heinzen, E.L., Allen, A.S., and Goldstein, D.B. (2013). Genic intolerance to functional variation and the interpretation of personal genomes. PLoS Genet. *9*, e1003709.
23. Samocha, K.E., Robinson, E.B., Sanders, S.J., Stevens, C., Sabo, A., McGrath, L.M., Kosmicki, J.A., Rehnström, K., Mallick, S., Kirby, A., et al. (2014). A framework for the interpretation of de novo mutation in human disease. Nat. Genet. *46*, 944–950.
24. Itan, Y., Shang, L., Boisson, B., Patin, E., Bolze, A., Moncada-Vélez, M., Scott, E., Ciancanelli, M.J., Lafaille, F.G., Markle, J.G., et al. (2015). The human gene damage index as a gene-level approach to prioritizing exome variants. Proc. Natl. Acad. Sci. USA *112*, 13615–13620.
25. Neale, B.M., Kou, Y., Liu, L., Ma'ayan, A., Samocha, K.E., Sabo, A., Lin, C.F., Stevens, C., Wang, L.S., Makarov, V., et al. (2012). Patterns and rates of exonic de novo mutations in autism spectrum disorders. Nature *485*, 242–245.
26. Homsy, J., Zaidi, S., Shen, Y., Ware, J.S., Samocha, K.E., Karczewski, K.J., DePalma, S.R., McKean, D., Wakimoto, H., Gorham, J., et al. (2015). De novo mutations in congenital heart disease with neurodevelopmental and other congenital anomalies. Science *350*, 1262–1266.
27. 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*, 1–15, 15.
28. Wall, M.A., Coleman, D.E., Lee, E., Iñiguez-Lluhi, J.A., Posner, B.A., Gilman, A.G., and Sprang, S.R. (1995). The structure of the G protein heterotrimer Gi alpha 1 beta 1 gamma 2. Cell *83*, 1047–1058.
29. Forbes, S.A., Beare, D., Gunasekaran, P., Leung, K., Bindal, N., Boutselakis, H., Ding, M., Bamford, S., Cole, C., Ward, S., et al. (2015). COSMIC: exploring the world's knowledge of somatic mutations in human cancer. Nucleic Acids Res. *43*, D805–D811.
30. Lambright, D.G., Sondek, J., Bohm, A., Skiba, N.P., Hamm, H.E., and Sigler, P.B. (1996). The 2.0 A crystal structure of a heterotrimeric G protein. Nature *379*, 311–319.
31. Rauen, K.A. (2013). The RASopathies. Annu. Rev. Genomics Hum. Genet. *14*, 355–369.
32. Champion, K.J., Bunag, C., Estep, A.L., Jones, J.R., Bolt, C.H., Rogers, R.C., Rauen, K.A., and Everman, D.B. (2011). Germline mutation in BRAF codon 600 is compatible with human development: de novo p.V600G mutation identified in a patient with CFC syndrome. Clin. Genet. *79*, 468–474.