Deleterious Variation in BRSK2 Associates with a Neurodevelopmental Disorder

Susan M. Hiatt,1 Michelle L. Thompson,1 Jeremy W. Prokop,2 James M.J. Lawlor,1 David E. Gray,1 E. Martina Bebin,3 Tuula Rinne,4 Marlies Kempers,5,6 Rolph Pfundt,4 Bregie W. van Bon,4 Cyril Mignot,5,6,7 Caroline Nava,8,9 Christel Depienne,8,9 Louisa Kalsner,10 Anita Rauch,11,12 Pascal Joset,13 Ruxandra Bachmann-Gagescu,11 Ingrid M. Wentzensen,13 Kirsty McWalter,13 and Gregory M. Cooper1,*

Developmental delay and intellectual disability (DD and ID) are heterogeneous phenotypes that arise in many rare monogenic disorders. Because of this rarity, developing cohorts with enough individuals to robustly identify disease-associated genes is challenging. Social-media platforms that facilitate data sharing among sequencing labs can help to address this challenge. Through one such tool, GeneMatcher, we identified nine DD- and/or ID-affected probands with a rare, heterozygous variant in the gene encoding the serine/threonine-protein kinase BRSK2. All probands have a speech delay, and most present with intellectual disability, motor delay, behavioral issues, and autism. Six of the nine variants are predicted to result in loss of function, and computational modeling predicts that the remaining three missense variants are damaging to BRSK2 structure and function. All nine variants are absent from large variant databases, and BRSK2 is, in general, relatively intolerant to protein-altering variation among humans. In all six probands for whom parents were available, the mutations were found to have arisen de novo. Five of these de novo variants were from cohorts with at least 400 sequenced probands; collectively, the cohorts span 3,429 probands, and the observed rate of de novo variation in these cohorts is significantly higher than the estimated background-mutation rate (p = 2.46 × 10⁻⁶). We also find that exome sequencing provides lower coverage and appears less sensitive to rare variation in BRSK2 than does genome sequencing; this fact most likely reduces BRSK2’s visibility in many clinical and research sequencing efforts. Altogether, our results implicate damaging variation in BRSK2 as a source of neurodevelopmental disease.

Developmental delay and intellectual disability (DD and ID), attention-deficient/hyperactivity disorder (ADHD), schizophrenia, language communication disorders, autism spectrum disorders (ASDs), and motor and tic disorders lie under a more general umbrella of neurodevelopmental disorders (NDDs).1,2 Although these are traditionally categorized into discrete disease entities, many symptoms are not unique to a single NDD. Furthermore, many genes have been associated with multiple NDDs,3 and new genetic associations continue to be discovered. This is particularly true given the recent acceleration in large-scale sequencing and cross-site genotype-phenotype “matchmaking” efforts.4,5

Through a Clinical Sequencing Exploratory Research (CSER) project focused on sequence-driven diagnoses for probands with unexplained DD and/or ID,6 we identified variation likely to be deleterious in BRSK2 (MIM: 609236) in four unrelated probands. BRSK2 encodes a serine/threonine-protein kinase, which is involved in axonogenesis and the polarization of cortical neurons.7 BRSK2 is predicted to be relatively intolerant to protein-altering variation in the general population (%ExAC v2 residual variation intolerance score [RVIS] = 4.9462%,8 pLI score (probability that a gene is intolerant to a loss of function mutation) = 0.788). In each proband, the BRSK2 variant was prioritized, after filtering and manual curation, as the most compelling disease-candidate variant of interest (see details by Bowling and colleagues for additional information about the cohort and analytical methods6). Although these observations suggest BRSK2 as a strong candidate NDD-associated gene, we sought additional cases via GeneMatcher to support pathogenicity. GeneMatcher is a database developed as part of the MatchMaker Exchange and has been shown to facilitate rare-disease-gene discovery.10 Information about five additional affected probands who were found by research or diagnostic sequencing (Table 1) and who had variants likely to be deleterious in BRSK2 was independently submitted to GeneMatcher. Informed consent to publish de-identified data was obtained from all affected individuals and/or families (see Supplemental Material and Methods). Altogether, the affected probands ranged in age from 3 years and

1HudsonAlpha Institute for Biotechnology, Huntsville, AL 35806, USA; 2Department of Pediatrics and Human Development, Michigan State University, East Lansing, MI 48824, USA; 3Department of Neurology, University of Alabama Birmingham, Birmingham, AL 35294, USA; 4Department of Human Genetics, Radboud University Medical Center, 6500 HB Nijmegen, the Netherlands; 5Département de Génétique, Groupe Hospitalier Pitié-Salpêtrière, Assistance Publique – Hôpitaux de Paris, Paris 75013, France; 6Centres de Référence Maladies Rares, Défections Intellectuelles de Causes Rares, Paris 75013, France; 7Groupes de Recherche Clinique Paris Sorbonne Déficiences Intellectuelles et Autisme, Paris 75013, France; 8Faculté de Médecine, Institut du Cerveau et de la Moelle épinière, Sorbonne Université, Paris 75013, France; 9Institute of Human Genetics, University Hospital Essen, University of Duisburg-Essen, Essen 45147, Germany; 10Connecticut Children’s Medical Center, Farmington, CT 06012, USA; 11Institute of Medical Genetics, University of Zurich, Schlieren 8952, Switzerland; 12Radiz-Rare Disease Initiative Zurich, Clinical Research Priority Program, University of Zurich, Zurich 8032, Switzerland; 13GeneDx, Gaithersburg, MD 20877, USA
*Correspondence: gcooper@hudsonalpha.org
© 2019 American Society of Human Genetics.
9 months old to 19 years old and had a mean age of just under 8 years (Table 2). All probands described here display ID except for one. All probands present with developmental delays, including speech delay (9/9) and motor delay (7/9). Eight of nine probands, one of whom (proband 9) was considered borderline, present with autism, and seven were reported to have behavioral abnormalities, including stereotypies (4/7), temper tantrums (3/7), and/or ADHD (3/7). Two probands reported sleep issues that were treatable with melatonin. Although most probands were reported to have facial dysmorphism, we did not observe a consistent set of features (see Supplemental Note). Additional details of each case are provided in the supplement (See Table S1 and Supplemental Note).

Parents were available for testing for six of the nine probands, and in all six of these probands, the variants were found to be \textit{de novo} (Table 1 and Figure 1). Six of the nine described variants, including two frameshift variants, one nonsense variant, and three variants affecting canonical splice sites, are predicted to result in loss of function. The remaining three variants were missense. All nine variants are absent from gnomAD\textsuperscript{9} and the Bravo TOPMed database. Although gnomAD does contain variant data aggregated from several disease cohorts, there has been an effort to remove any variants found in individuals with severe pediatric disease.\textsuperscript{9} All variants were computationally predicted to be deleterious and had Combined Annotation Dependent Depletion (CADD) scores\textsuperscript{11} ranging from 24.8 to 38; these scores indicate that they rank among the most highly deleterious variants possible in the human genome reference assembly, similar to most variants previously reported to cause Mendelian diseases.\textsuperscript{11}

The canonical protein encoded by the \textit{BRSK2} locus (GenBank: NP_001243556.1, UniProt: Q8IWQ3, 736 aa) contains several domains, including a protein kinase domain (aa 19–270), a ubiquitin-associated domain (UBA; aa 297–339), a proline-rich domain (aa 424–468), and a kinase-associated domain (KA1; aa 530–653) that contains a KEN box (a degradation signal, aa 603–605) (Figure 1B). An analysis of conservation along the protein identified several regions with elevated measures of conservation (Figure 1C). Two missense variants (c.194G>A [p.Arg65Gln] and c.635G>A [p.Gly212Glu]) are located within the protein kinase domain, and one (c.1861C>T [p.Arg621Cys]) is within the KA1 domain.

We assessed the potential structural effects of the three missense variants by performing computational modeling.\textsuperscript{12} All three missense variants lie within conserved linear motifs (Figure 1C) and affect residues that are conserved across many species (Figures 2 and S1). Arg65 lies within the protein kinase domain and has been found to coordinate intramolecularly with Glu330 to form a salt bridge.\textsuperscript{13} Arg65 also lies within a mitogen-activated protein kinase (MAPK) docking motif\textsuperscript{14} (Figure 2). Thus, p.Arg65Gln is predicted to disrupt both the structure and functional activity of BRSK2. Gly212 lies in the C-lobe of the protein kinase domain,
Table 2. Clinical Characteristics of Individuals with BRSK2 Variation

<table>
<thead>
<tr>
<th>Proband</th>
<th>Gender</th>
<th>Age at last examination</th>
<th>Intellectual disability</th>
<th>Developmental delay(s)</th>
<th>Autism</th>
<th>Behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Male</td>
<td>5 years, 9 months</td>
<td>moderate</td>
<td>speech and fine motor</td>
<td>mild</td>
<td>reported</td>
</tr>
<tr>
<td>2</td>
<td>Male</td>
<td>14 years</td>
<td>severe</td>
<td>speech and social interaction</td>
<td>severe</td>
<td>yes</td>
</tr>
<tr>
<td>3</td>
<td>Male</td>
<td>5 years, 7 months</td>
<td>moderate</td>
<td>speech and fine motor</td>
<td>moderate</td>
<td>yes</td>
</tr>
<tr>
<td>4</td>
<td>Male</td>
<td>6 years, 10 months</td>
<td>severe</td>
<td>speech and social interaction</td>
<td>speech and fine motor</td>
<td>reported</td>
</tr>
<tr>
<td>5</td>
<td>Male</td>
<td>4 years</td>
<td>mild</td>
<td>speech and fine motor</td>
<td>mild</td>
<td>reported</td>
</tr>
<tr>
<td>6</td>
<td>Male</td>
<td>4 years</td>
<td>moderate</td>
<td>speech and fine motor</td>
<td>moderate</td>
<td>yes</td>
</tr>
<tr>
<td>7</td>
<td>Male</td>
<td>7 months</td>
<td>mild</td>
<td>speech and fine motor</td>
<td>severe</td>
<td>yes</td>
</tr>
<tr>
<td>8</td>
<td>Male</td>
<td>10 months</td>
<td>severe</td>
<td>speech and fine motor</td>
<td>mild</td>
<td>reported</td>
</tr>
<tr>
<td>9</td>
<td>Male</td>
<td>11 months</td>
<td>mild</td>
<td>speech and fine motor</td>
<td>mild</td>
<td>reported</td>
</tr>
</tbody>
</table>

Abbreviations as follows: ABD = attention deficit disorder and ADHD = attention-deficit/hyperactivity disorder.

We assessed the degree of enrichment of observed de novo variation in BRSK2 in the sequenced DD- and/or ID-affected cohorts that underlie this study. Two probands (1 and 5) were a part of a cohort of 2,418 DD- and/or ID-affected probands sequenced as trios. An additional proband (2) was sequenced as a trio among a cohort of 550 affected probands, and two others (probands 3 and 4) were among a cohort of 461. In aggregate, these cohorts include five de novo variants in 3,429 affected, sequenced individuals. We compared this observed rate to the expected rate estimated by Samocha et al. (2.97 × 10⁻³ variants per chromosome) of de novo missense, nonsense, splice, and frameshift variation in BRSK2. The observed rate of de novo variation in the DD- and/or ID-affected cohorts considered here is significantly greater than the background mutation rate (five de novo variants observed versus 0.20 expected, p = 2.46 × 10⁻⁶), and this observation remained significant even after a Bonferroni correction for 20,000 genes (p = 0.0492). We note that one proband (7) with a de novo variant was sequenced clinically as a trio, but a cohort size was not available for this proband; furthermore, one or both parents were unavailable for testing for three of the nine observed variants. Thus, although these four additional variants add to the evidence supporting a disease role for BRSK2, they are excluded from the preceding enrichment calculations.

BRSK2 and its homolog BRSK1 (MIM: 609235) encode kinases required for neuronal polarization. These two kinases, along with 11 other kinases, form the AMPK-related family of protein kinases. Although knockouts of either Brsk1 or Brsk2 alone in mice were healthy and fertile, double knockouts of Brsk1 and Brsk2 resulted in pups that exhibited reduced spontaneous movement and little response to tactile stimulation and that died within 2 h of birth. Expression patterns of BRSK2 also support its role in neurodevelopment. BRSK2 is most highly expressed in the brain in human, and Brsk1 and Brsk2 are restricted to the nervous system in mice.
and/or ID. For example, BRSK2 has been shown to phosphorylate TSC2 and suppress mTORC1 activity. The tuberous sclerosis complex (TSC) signaling pathway is one of the pathways associated with autophagy during early axonal growth, and TSC2, specifically, is a regulator of cellular size and growth. BRSK2 has also been shown to interact with PTEN, which has been associated with various developmental disorders (see MIM: 601728,24) including autism. PTEN-deficient mice display malformation of neuronal structures and autistic features that result from aberrant TSC-mTORC1 signaling. WDR45, also known as WIPI4, is a scaffold protein that controls autophagy and has recently been shown to depend on BRSK2 activity. Variation in WDR45 is associated with an X-linked dominant disorder: neurodegeneration with brain iron accumulation (MIM: 300894). The numerous genetic and biochemical interactions between BRSK2 and well-established NDD genes further strengthen the conclusion that damaging variation in BRSK2 underlies an NDD.

Across six recent publications reporting on de novo variation in large cohorts with DD and/or ID or autism, two protein-altering BRSK2 variants were reported: GenBank: NM_001256627.1 (c.992_994del, [p.Lys331del]) was found in a cohort of 2,500 probands with autism, and GenBank: NM_001256627.1 (c.770G>A, [p.Arg257His]) was found in a cohort of 4,293 DD- and/or ID-affected probands. Interestingly, this second variant has been observed as a heterozygote seven times in gnomAD, suggesting it is not a highly penetrant allele contributing to DD and/or ID. These data raise an interesting question, namely as to why the frequency of observed BRSK2 variation in this study is markedly higher than that found in previous studies. This is particularly true for the HudsonAlpha CSER study, in which four variants were found among 581 affected probands (461 of whom were sequenced as trios). Some of the discrepancy is probably due to stochastic variability in observing a small number of rare events. However, one potential systematic explanation is that BRSK2 is less deeply covered in exomes, and the observed enrichment, in part, reflects the effects of the genome sequencing that was used for the HudsonAlpha probands described here. It has been shown previously that genome sequencing provides better coverage, in general, over coding exons than exome sequencing does, and that some exons, including among clinically relevant genes, tend to be more poorly covered by exomes.

We find that BRSK2 is less well covered by exomes than by genomes in gnomAD (Figure 3). For example, when requiring a minimum depth of 20× among exonic bases (plus 10 bp on either side of each exon), we found that 76% of gnomAD exome samples, compared to 93% of genome samples, have half of all BRSK2 bases covered (Figure 3A). Furthermore, we assessed rare-variant detection rates, in particular the rate at which singletons (i.e., variants for which only one alternative allele is observed across the combined set of exomes and genomes) are
observed. There are 46 singletons detected among 15,708 genomes (0.29%) in gnomAD and 189 singletons detected among 125,748 exomes (0.15%); this difference is significant \( p = 1.5 \times 10^{-4} \), Fisher’s exact test) and suggests increased rare-variant sensitivity in genomes relative to exomes. Additionally, considering only exomes, we compared coverage of BRSK2 exons to exons in other disease-associated genes that are annotated in the Development Disorder Genotype-Phenotype Database (DDG2P). Although, again, only 76% of samples have at least half of BRSK2 bases covered at 20X in gnomAD exomes, 99% of samples have half or more of the bases in previously reported DDG2P genes covered in gnomAD exomes (Figure 3B). Thus, we find it likely that the lower rates of BRSK2 variation found in other DD and/or ID studies reflects, at least in part, reduced variant sensitivity of exome sequencing in BRSK2.

We have identified nine individuals harboring rare, heterozygous BRSK2 variants that are likely to be deleterious, and we provide detailed clinical descriptions of the

---

Figure 2. Computational Modeling of BRSK2 Missense Variants
A full model of BRSK2 was created with I-TASSER modeling using PDB: 4YOM, 4YNZ, and 4IW0. This model of BRSK2 was combined via ConSurf mapping with sequences for BRSK2 from 99 species. Amino acid coloring is as follows: gray = not conserved, yellow = conserved hydrophobic, green = conserved hydrophilic, red = conserved polar acidic, blue = conserved polar basic, and magenta = conserved human variants of interest. Zoomed-in views of the three locations are shown, along with codon usage throughout evolution. The conservation score is defined as an additive metric of amino acid conservation and codon selection as previously defined.\(^1\) For example, a conservation score of 2 indicates 100% conservation with >2 standard deviations above the mean for codon selection.\(^1\) s/n indicates synonymous mutations versus non-synonymous mutations observed at the same position in other species; differences are indicated. All three sites are under high selection and have multiple synonymous (s) amino acids in 99 open reading frames (ORFs) of BRSK2 and only a single nonsynonymous (n) change observed at G212. Linear motifs mapped with the Eukaryotic Linear Motif (ELM) tool are shown below each site.
phenotypes observed in these individuals, who all present with varying degrees and manifestations of developmental disorders. We believe these observations strongly support the conclusion that damaging variation in \textit{BRSK2} is causally related to an NDD. The key points of evidence are as follows: (1) we observe a statistically significant enrichment of \textit{de novo} variants in affected individuals relative to the estimated background mutation rate \( p = 2.46 \times 10^{-6} \); (2) although one or both parents were unavailable in three cases, none of the variants described here were found to be inherited, and all observed variants are absent from gnomAD and TopMed; (3) \textit{BRSK2} is relatively intolerant to protein-altering variation in the general population;\(^8,9\) (4) all variants in affected probands are either predicted to result in loss of function or are missense variants at highly conserved residues; (5) all variants are computationally predicted to be evolutionarily deleterious and have, for example, CADD\(^{11}\) scores that are typical for mutations previously reported to underlie Mendelian disease; (6) model organism evidence suggests a role for \textit{BRSK2} in neurodevelopment; and (7) \textit{BRSK2} is known to genetically and/or biochemically interact with several genes that are robustly associated with developmental disease. In summary, these data collectively implicate \textit{BRSK2} as an NDD-related gene.

Accession Numbers

All relevant variant data are supplied within the paper or in supporting files. Complete genome data for probands 3, 4, 8, and 9 are available via dbGAP (accession number phs001089.v3.p1). Complete exome data for other probands is not available for privacy and institutional review board (IRB) reasons.

Supplemental Data

Supplemental Data can be found with this article online at https://doi.org/10.1016/j.ajhg.2019.02.002.