

## Mutations in *PNKP* Cause Recessive Ataxia with Oculomotor Apraxia Type 4

Jose Bras,<sup>1,6</sup> Isabel Alonso,<sup>2,3,4,6</sup> Clara Barbot,<sup>2,3,4</sup> Maria Manuela Costa,<sup>5</sup> Lee Darwent,<sup>1</sup> Tatiana Orme,<sup>1</sup> Jorge Sequeiros,<sup>2,3,4</sup> John Hardy,<sup>1</sup> Paula Coutinho,<sup>2,3,4</sup> and Rita Guerreiro<sup>1,\*</sup>

Hereditary autosomal-recessive cerebellar ataxias are a genetically and clinically heterogeneous group of disorders. We used homozygosity mapping and exome sequencing to study a cohort of nine Portuguese families who were identified during a nationwide, population-based, systematic survey as displaying a consistent phenotype of recessive ataxia with oculomotor apraxia (AOA). The integration of data from these analyses led to the identification of the same homozygous *PNKP* (polynucleotide kinase 3'-phosphatase) mutation, c.1123G>T (p.Gly375Trp), in three of the studied families. When analyzing this particular gene in the exome sequencing data from the remaining cohort, we identified homozygous or compound-heterozygous mutations in five other families. *PNKP* is a dual-function enzyme with a key role in different pathways of DNA-damage repair. Mutations in this gene have previously been associated with an autosomal-recessive syndrome characterized by microcephaly; early-onset, intractable seizures; and developmental delay (MCSZ). The finding of *PNKP* mutations associated with recessive AOA extends the phenotype associated with this gene and identifies a fourth locus that causes AOA. These data confirm that MCSZ and some forms of ataxia share etiological features, most likely reflecting the role of *PNKP* in DNA-repair mechanisms.

Hereditary autosomal-recessive cerebellar ataxias (ARCAs) are rare neurodegenerative disorders that are clinically and genetically very heterogeneous and are characterized by cerebellar ataxia that is frequently associated with peripheral sensorimotor neuropathy. Ataxia with oculomotor apraxia (AOA) is a subgroup involving cerebellar ataxia, sensorimotor axonal neuropathy, oculomotor apraxia, and extrapyramidal features. Chromosomal instability, immunodeficiency, and sensitivity to ionizing radiations, all usually observed in persons with ataxia telangiectasia (AT) and AT-like disorders, are absent in individuals with AOA.<sup>1</sup> Early-onset recessive ataxia (AOA1 or EAOH [MIM 208920]) is a progressive syndrome associated with hypoalbuminemia and elevated levels of cholesterol and is caused by mutations in *APTX* (aprataxin).<sup>1–3</sup> Autosomal-recessive spinocerebellar ataxia 1 (AOA2 or SCAR1 [MIM 606002]), a progressive ataxia,<sup>4</sup> occurring later than AOA1, is characterized by increased alpha-fetoprotein levels and is caused by mutations in *SETX* (senataxin).<sup>5</sup> A third gene, *PIK3R5* (phosphoinositide-3-kinase, regulatory subunit 5; AOA3 [MIM 615217]), has been identified in a consanguineous Saudi Arabian family whose affected members have clinical features similar to those of individuals with AOA2.<sup>6</sup>

Next-generation sequencing technologies have been remarkably useful for the identification of novel genes causing Mendelian diseases. At the same time, these recently developed technologies have allowed the expansion of phenotypes, particularly in various neurological

diseases. The association of the same molecular event with a large-spectrum phenotype or very different phenotypes can be explained by several factors, one being the pleiotropism associated with many genes.<sup>7</sup>

We used exome sequencing and homozygosity mapping to study a series of nine Portuguese families affected by early-onset recessive AOA. These families were identified during a nationwide, population-based, systematic survey of hereditary ataxias and spastic paraplegias. Performed in Portugal from 1994 to 2004, this survey used multiple sources of information to identify affected persons in the community and health-care settings. Detailed methods of the survey are described elsewhere.<sup>8</sup> All individuals studied here presented established clinical criteria for AOA,<sup>1</sup> and samples were collected after receipt of written informed consent from participants. This study used only de-identified, previously collected DNA samples that were stored at the authorized Center for Predictive and Preventive Genetics, Institute for Molecular and Cell Biology biobank and database.

After previous exclusion of mutations in the genes known to be directly associated with this phenotype (*APTX*, *SETX*, and *PIK3R5*), we performed whole-genome genotyping in ten individuals from seven families (nine affected and one unaffected). We used Illumina OmniExpress Beadchips, which assay over 700,000 markers, as per the manufacturer's instructions. We used Illumina GenomeStudio for initial data analysis and quality

<sup>1</sup>Department of Molecular Neuroscience, Institute of Neurology, University College London, Queen Square, London WC1N 3BG, UK; <sup>2</sup>Center for Predictive and Preventive Genetics, Institute for Molecular and Cell Biology, University of Porto, Porto 4150-180, Portugal; <sup>3</sup>Unidade de Investigação Genética e Epidemiológica em Doenças Neurológicas Research Group, Institute for Molecular and Cell Biology, University of Porto, Porto 4150-180, Portugal; <sup>4</sup>Instituto de Investigação e Inovação em Saúde, University of Porto, Porto 4150-180, Portugal; <sup>5</sup>Hospital Pedro Hispano, Unidade Local de Saúde de Matosinhos, Matosinhos 4454-509, Portugal

<sup>6</sup>These authors contributed equally to this work

\*Correspondence: [r.guerreiro@ucl.ac.uk](mailto:r.guerreiro@ucl.ac.uk)

<http://dx.doi.org/10.1016/j.ajhg.2015.01.005>. ©2015 The Authors

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

control, which led to the exclusion of three samples (two from family 2 and one from family 6) from further analysis involving whole-genome genotyping data. Runs of homozygosity (ROH) were identified with PLINK v.1.07<sup>9</sup> on the basis of having a size greater than 1.5 Mb and a minimum of 50 SNPs per region. We scanned the genome for regions of homozygosity by using a sliding window of 50 SNPs and allowing at most two missing genotypes and one heterozygote call per ROH. We focused the initial homozygosity mapping on pedigrees with reported consanguinity (families 5 and 7). The only homozygous region, a region >1.5 Mb and shared by all affected individuals, was on chromosome 19 (49,506,390–51,400,356) (Figure S1). Upon analysis of the other families, we found that the proband of family 8 also shared a large homozygous tract overlapping this region (chr19: 49,425,838–53,225,722).

We performed exome sequencing in 16 individuals from eight families (12 affected and four unaffected). For this analysis, we prepared genomic DNA according to Illumina's TruSeq Sample Preparation v.3 and performed the exome capture with Illumina's TruSeq Exome Enrichment according to the manufacturer's instructions. Sequencing was performed on an Illumina HiSeq2500 with 100-bp paired-end reads. Following quality control procedures, samples yielded between 10.1 and 12.7 Gb of high-quality, aligned data. This amount of data was the result of mean target coverage between 66.39× and 88.5×, 93.5%–94.6% of targets' being covered at greater than or equal to 10×, and less than 0.2% of targets' not being covered even once. We performed sequence alignment and variant calling against the reference human genome (UCSC Human Genome Browser hg19) by using the Burrows-Wheeler Aligner<sup>10</sup> and the Genome Analysis Toolkit.<sup>11,12</sup> Prior to variant calling, PCR duplicates were removed with the Picard software. On the basis of the hypothesis that the mutation was rare, we excluded all common SNPs (MAF > 5%) identified in dbSNP v.137 and in our in-house database of sequencing data for other diseases (n > 2,000). Given the apparent autosomal-recessive mode of inheritance in the three families (families 5, 7, and 8) and the fact that two parental pairs were consanguineous, we focused on homozygous variants located in the shared ROH. One *PNKP* mutation, c.1123G>T (p.Gly375Trp; RefSeq accession number NM\_007254.3), was present in homozygosis in all affected individuals and absent in all unaffected ones. Expanding this analysis to the other studied families, we found the same mutation in homozygosis in family 2; probands in families 1, 3, 4, and 6 were compound heterozygotes for different mutations in the same gene. The proband in family 6 was tested by Sanger sequencing, and no causative mutations were found in this gene in family 9. We also used Sanger sequencing to confirm all mutations identified by exome sequencing, establish compound heterozygosity, and verify intrafamilial segregation. We performed PCR amplifications and purified the resulting products with ExoSAP-IT (USB), then performed direct Sanger sequencing of both

strands with BigDye Terminator v.3.1 chemistry v.3.1 (Applied Biosystems) and an ABI 3730XL Genetic Analyzer (Applied Biosystems). Sequencing traces were analyzed with Sequencher software v.4.2 (Gene Codes).

The main clinical features of the 11 AOA-affected persons (four males and seven females) who were found to have *PNKP* mutations are presented in Table 1. All of them were either homozygotes or compound heterozygotes for *PNKP* mutations (Table 2 and Figure 1). Age at onset ranged from 1 to 9 years; the mean was  $4.3 \pm 2.3$  years. Most individuals exhibited dystonia as their first symptom; this was so prominent that many of these individuals underwent diagnostic procedures for extrapyramidal disorders. In all cases, dystonia spontaneously attenuated during the course of the disease. The second most common symptom was ataxia, followed by oculomotor apraxia. All individuals had signs of polyneuropathy with early, generalized areflexia. Distal-muscle wasting and weakness led to tetraplegia and short atrophic hands and feet. Loss of the ability to walk occurred 7–21 years after disease onset, and most individuals were wheelchair bound by adolescence. Cognitive impairment was present in seven persons, two of whom were severely demented. Alpha-fetoprotein, albumin, and cholesterol levels were highly variable in this cohort; alpha-fetoprotein levels were increased in five individuals, albumin was decreased in six, and cholesterol was elevated in five cases, all of which was determined at later stages of disease progression. In all 11 individuals, brain MRIs revealed cerebellar atrophy.

Mutations in *PNKP* were originally found to cause early infantile epileptic encephalopathy-10 (EIEE10 [MIM 613402]), characterized by microcephaly, seizures, and developmental delay.<sup>13</sup> More recently, this gene was associated with cerebellar atrophy in two Dutch siblings who had progressive debilitating polyneuropathy, microcephaly, severe intellectual disability, and mild epilepsy.<sup>14</sup> No ocular signs were described in those Dutch cases. None of the individuals we studied had microcephaly or epilepsy; cognitive impairment was observed in most but not all individuals (although it progressed to severe dementia in two cases).

These families were identified through a population-based nationwide survey for hereditary ataxias, to the best of our knowledge the largest ever performed. We can thus conclude that AOA4 is the most frequent form of AOA in the Portuguese population (it affects nearly 40% of the families identified) and the most frequent recessive ataxia in this population, after Freidreich ataxia.<sup>8</sup> Mean age at onset (4.3 years) for individuals with AOA4 is closer to that of individuals with AOA1 ( $6.9 \pm 4.7$  years) than to that of those with AOA2 ( $14.6 \pm 3.4$ ), although it is lower than both.<sup>16,17</sup> Clinical presentation, with marked extrapyramidal manifestations and its rapid progression, also more closely resembles that of AOA1 than that of AOA2. Additionally, in individuals with AOA4, albumin may be low or normal and cholesterol levels are normal or high, a similar situation to what is seen in individuals with AOA1 after some years

**Table 1. Clinical, Biochemical, and Imaging Features of the Individuals Carrying *PNKP* Mutations**

	Family 1		Family 2	Family 3	Family 4	Family 5	Family 6	Family 7		Family 8	
	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11
Consanguinity	–		yes	–	–	yes	yes	yes		–	
Gender	male	female	female	female	male	female	female	female	male	female	male
Age at onset (years)	5	9	6	3	1	3	4	4	7	2	3
First sign	dystonia	dystonia	dystonia	ataxia	dystonia	ataxia	ataxia	OMA	OMA	dystonia	dystonia
More prominent sign	neurop.	neurop.	neurop.	neurop.							
OMA	+++	+++	+	++	+	+	++	+++	+++	+++	+++
Age in wheelchair (years)	18	22	18	20	NA	15	25	15	14	20	15
Dystonia	+	+	+	+	++	+	–	–	–	++	++
Cognitive impairment	+	+	++	+	NA	–	–	–	+	++	+
Motor deficit	+++	++	+++	+++	+++	+++	+++	+++	+++	+++	+++
Decreased vibration sense	++	+	NA	++	NA	++	+++	++	++	NA	++
Pyramidal signs	–	–	+	–	–	–	–	–	–	+	–
Obesity	–	–	+	+	–	–	+	–	–	++	–
MRI findings	CA	CA	CA	CA							
Other features	BSA	BSA	dem.	NA	NA	NA	NA	NA	NA	dem.	BSA
$\alpha$ -fetoprotein levels	1.5 N	1.5 N	1.5 N	N	N	1.5 N	4 N	N	N	N	N
Albumin levels	N	N	↓	↓	↓	↓	N	NA	NA	↓	↓
Cholesterol levels	N	N	N	↑	↑	N	↑	NA	NA	↑	↑

Abbreviations and symbols are as follows: P1–P11, individuals 1–11, respectively; +, present; ++, moderate; +++, severe; –, absent; ↑, increased level; ↓, decreased level; N, normal level; NA, not available; CA, cerebellar atrophy; BSA, brainstem atrophy; OMA, oculomotor apraxia; neurop., neuropathy; dem., dementia.

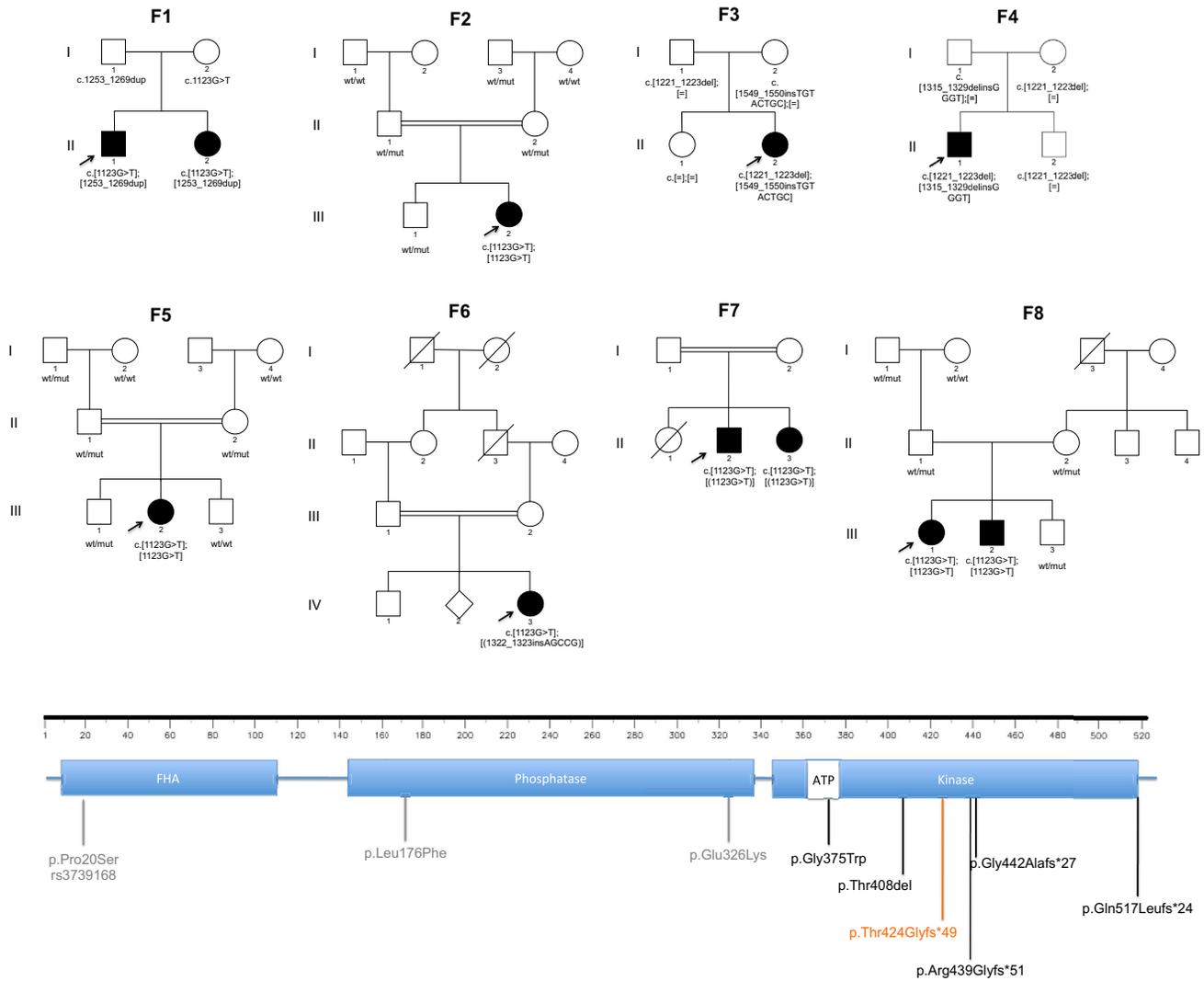
of disorder progression. On the other hand,  $\alpha$ -fetoprotein was elevated in some individuals with AOA4, which seems to always be the case in those with AOA2.

**Table 2. *PNKP* Mutations that Caused AOA in Affected Members of the Eight Portuguese Families in Whom Mutations Were Found**

Family	cDNA (RefSeq NM_007254.3)	Protein (RefSeq NP_009185.2)
1	c.[1123G>T];[1253_1269dupGGGTCGCCATCGACAAC]	p.[(Gly375Trp)];[(Thr424Glyfs*49)]
2	c.[1123G>T];[1123G>T]	p.[(Gly375Trp)];[(Gly375Trp)]
3	c.[1221_1223del];[1549_1550insTGACTGC]	p.[(Thr408del)];[(Gln517Leufs*24)]
4	c.[1221_1223del];[1315_1329delinsGGGT]	p.[(Thr408del)];[(Arg439Glyfs*51)] <sup>a</sup>
5	c.[1123G>T];[1123G>T]	p.[(Gly375Trp)];[(Gly375Trp)]
6	c.[1123G>T(;)1322_1323insAGCCG]	p.[(Gly375Trp(;)Gly442Alafs*27)]
7	c.[1123G>T];[(1123G>T)]	p.[(Gly375Trp)];[(Gly375Trp)]
8	c.[1123G>T];[1123G>T]	p.[(Gly375Trp)];[(Gly375Trp)]

<sup>a</sup>One of the allelic mutations in F4 is formed by a complex allele that was initially not properly identified. Figure S3 shows the detailed view of the alignment at that position, clearly displaying the complex allele.

The different phenotypes so far associated with mutations in *PNKP* do not seem to relate to either the type or the location of the mutation (Figure 1). *PNKP* is a modular protein with three domains: an amino-terminal fork-head-associated (FHA) domain, which mediates interactions with the scaffold proteins XRCC1 and XRCC4; a DNA phosphatase domain; and a DNA kinase domain.<sup>18</sup> The same homozygous variant, p.Thr424Glyfs\*49 (c.1253\_1269dup), located in the kinase domain of the protein, has been found in individuals with progressive cerebellar atrophy and in individuals with MCSZ (an autosomal-recessive syndrome characterized by microcephaly, early onset, intractable seizures, and developmental delay).<sup>13,14</sup> We have now identified this same variant in compound heterozygosity with p.Gly375Trp in family 1. Missense substitutions in the FHA and phosphatase domains of the protein have only been associated with developmental syndromes. In most affected individuals, MCSZ was found to be caused by the p.Glu326Lys (c.1123G>T) variant in homozygosity.<sup>13</sup> In the present study, variants found to be associated with AOA were all located in, or adjacent to, the kinase region of the protein. The position of the most frequent variant, p.Gly375Trp, in this cohort is highly conserved across species (Figure S2) and is located in a potential ATP nucleotide-binding domain (amino acids 372–379) within the kinase region. Additionally, several in-silico-prediction software programs rate this variant as



**Figure 1. Pedigrees of the Portuguese Families In which AOA-Affected Members Had *PNKP* Mutations and Protein Representation with the Location of the Variants Found to Be Associated with Different Phenotypes**  
 Black symbols represent individuals affected by AOA. An arrow indicates the index individual in each family. Mutation segregation with disease is indicated underneath each symbol; all unaffected family members carry either a single variant (wt/mut) or no variants (wt/wt), and all affected individuals harbor mutations in homozygosity or compound heterozygosity. In the lower panel, *PNKP* (RefSeq NP\_009185.2) is represented with the depicted regions as predicted by UniProtKB (Q96T60). Variants found in this study are represented in black and are all located in or adjacent to the kinase region of the protein. Variants previously identified in individuals with MCSZ are represented in gray. An additional 17-bp intron 15 deletion was reported in one family, which is not represented here.<sup>13</sup> The variant p.Thr424Glyfs\*49 (in orange) has been previously associated with MCSZ<sup>13</sup> and with progressive cerebellar atrophy and polyneuropathy.<sup>14</sup> The same change was identified in compound heterozygosity with p.Gly375Trp in family 1. Targeted resequencing has also identified the p.Pro20Ser (c.58G>A) variant in an individual with epileptic encephalopathy;<sup>15</sup> this variant was considered to be pathogenic in the homozygous state but had previously been identified in the Exome Variant Server (overall MAF of 0.9%) and dbSNP databases and classified as benign in ClinVar.

deleterious (SIFT score = 0; PolyPhen score = 0.999; CADD\_phred = 18.51).<sup>19</sup>

A functional assessment of the impact of MCSZ-associated *PNKP* mutations showed that the studied variants reduced the cellular stability and the level of *PNKP*; some ablated cellular DNA 5' kinase activity, and all studied variants reduced cellular 3' phosphatase activity.<sup>20</sup> The structural unit of the wild-type protein is stabilized by the anchoring of the kinase domain to the phosphatase domain via interactions between the five C-terminal resi-

dues, Q517–G521, and the interdomain linker and phosphatase domain.<sup>18,20</sup> The deletion found in family 3, although adjacent to the kinase region of the protein, probably affects these interactions and thereby has consequences for protein conformation and function. It is possible that the different phenotypes associated with the different mutations result from defects in the phosphatase or kinase activities of the protein or from conformational protein alterations that lead to deficient protein-protein interactions or substrate recognition. It

is also plausible that the different phenotypes are the result of an additional genetic-variation load.<sup>21</sup> PNKP has important roles in multiple pathways involved in DNA-damage repair, including single-strand breaks (SSBs) and double-strand breaks (DSB).<sup>14,21,22</sup> The cellular response to these breaks involves a complex signaling network and the key activator of this network is the serine/threonine protein kinase ATM (AT mutated), which has been shown to target PNKP.<sup>23,24</sup> Mutations in *ATM* are known to cause AT, another autosomal-recessive ataxia. Mutations in other ataxia-related genes, such as *TDPI* (tyrosyl-DNA phosphodiesterase-1, associated with spinocerebellar ataxia with axonal neuropathy-1 (SCAN1) [MIM 607250]) and *APTX*, also result in defects in SSB repair. In this way, the molecular involvement of PNKP in these different pathways, and the possible compensation by related proteins, might also contribute to the phenotypic differences associated with mutations in this gene.<sup>22</sup> A STRING analysis supported this possibility, showing that PNKP interacts with proteins involved in neurodegenerative diseases and microcephaly syndromes.<sup>14</sup> The phenotypes associated with PNKP mutations could also be influenced by gene-environment interactions. It has been recently demonstrated that physiologically and environmentally relevant doses of cadmium and copper can inhibit both the phosphatase and kinase activities of PNKP.<sup>25</sup>

In this study, we identified homozygous or compound-heterozygous *PNKP* mutations in eight of the nine Portuguese families we studied (with a total of 11 affected individuals), suggesting that, in Portugal, mutations in *PNKP* are the most frequent cause of AOA. However, additional studies are needed to determine whether this is also the case in other populations. Our results also extend the phenotype associated with mutations in this gene. Future studies clarifying PNKP regulation and interaction with other repair enzymes might help explain this phenotypic variability.

### Supplemental Data

Supplemental Data include three figures and can be found with this article online at <http://dx.doi.org/10.1016/j.ajhg.2015.01.005>.

### Acknowledgments

This work was supported in part by an anonymous donor, the Wellcome Trust/Medical Research Council Joint Call in Neurodegeneration award (WT089698) to the UK Parkinson's Disease Consortium (members are from the University College London Institute of Neurology, the University of Sheffield, and the Medical Research Council Protein Phosphorylation Unit at the University of Dundee), a fellowship from the Alzheimer's Society to J.B., and a fellowship from Alzheimer's Research UK to R.G.

Received: November 5, 2014

Accepted: January 9, 2015

Published: February 26, 2015

### Web Resources

The URLs for data presented herein are as follows:

dbSNP, <http://www.ncbi.nlm.nih.gov/projects/SNP/>

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

Picard, <http://picard.sourceforge.net/>

UCSC Human Genome Browser, <http://genome.ucsc.edu/cgi-bin/hgGateway>

### References

- Barbot, C., Coutinho, P., Chorão, R., Ferreira, C., Barros, J., Fineza, I., Dias, K., Monteiro, J., Guimarães, A., Mendonça, P., et al. (2001). Recessive ataxia with ocular apraxia: review of 22 Portuguese patients. *Arch. Neurol.* 58, 201–205.
- Moreira, M.C., Barbot, C., Tachi, N., Kozuka, N., Mendonça, P., Barros, J., Coutinho, P., Sequeiros, J., and Koenig, M. (2001). Homozygosity mapping of Portuguese and Japanese forms of ataxia-oculomotor apraxia to 9p13, and evidence for genetic heterogeneity. *Am. J. Hum. Genet.* 68, 501–508.
- Moreira, M.C., Barbot, C., Tachi, N., Kozuka, N., Uchida, E., Gibson, T., Mendonça, P., Costa, M., Barros, J., Yanagisawa, T., et al. (2001). The gene mutated in ataxia-ocular apraxia 1 encodes the new HIT/Zn-finger protein aprataxin. *Nat. Genet.* 29, 189–193.
- Le Ber, I., Bouslam, N., Rivaud-Péchéux, S., Guimarães, J., Benomar, A., Chamayou, C., Goizet, C., Moreira, M.C., Klur, S., Yahyaoui, M., et al. (2004). Frequency and phenotypic spectrum of ataxia with oculomotor apraxia 2: a clinical and genetic study in 18 patients. *Brain* 127, 759–767.
- Moreira, M.C., Klur, S., Watanabe, M., Németh, A.H., Le Ber, I., Moniz, J.C., Tranchant, C., Aubourg, P., Tazir, M., Schöls, L., et al. (2004). Senataxin, the ortholog of a yeast RNA helicase, is mutant in ataxia-ocular apraxia 2. *Nat. Genet.* 36, 225–227.
- Al Tassan, N., Khalil, D., Shinwari, J., Al Sharif, L., Bavi, P., Abduljaleel, Z., Abu Dhaim, N., Magrashi, A., Bobis, S., Ahmed, H., et al. (2012). A missense mutation in *PIK3R5* gene in a family with ataxia and oculomotor apraxia. *Hum. Mutat.* 33, 351–354.
- Guerreiro, R., Brás, J., Hardy, J., and Singleton, A. (2014). Next generation sequencing techniques in neurological diseases: redefining clinical and molecular associations. *Hum. Mol. Genet.* 23 (R1), R47–R53.
- Coutinho, P., Ruano, L., Loureiro, J.L., Cruz, V.T., Barros, J., Tuna, A., Barbot, C., Guimarães, J., Alonso, I., Silveira, I., et al. (2013). Hereditary ataxia and spastic paraplegia in Portugal: a population-based prevalence study. *JAMA Neurol.* 70, 746–755.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J., and Sham, P.C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575.
- Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760.
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernysky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M.A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20, 1297–1303.
- DePristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., del Angel, G., Rivas, M.A.,

- Hanna, M., et al. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* *43*, 491–498.
13. Shen, J., Gilmore, E.C., Marshall, C.A., Haddadin, M., Reynolds, J.J., Eyaid, W., Bodell, A., Barry, B., Gleason, D., Allen, K., et al. (2010). Mutations in PNKP cause microcephaly, seizures and defects in DNA repair. *Nat. Genet.* *42*, 245–249.
14. Poulton, C., Oegema, R., Heijnsman, D., Hoozeboom, J., Schot, R., Stroink, H., Willemsen, M.A., Verheijen, F.W., van de Spek, P., Kremer, A., and Mancini, G.M. (2013). Progressive cerebellar atrophy and polyneuropathy: expanding the spectrum of PNKP mutations. *Neurogenetics* *14*, 43–51.
15. Carvill, G.L., Heavin, S.B., Yendle, S.C., McMahon, J.M., O’Roak, B.J., Cook, J., Khan, A., Dorschner, M.O., Weaver, M., Calvert, S., et al. (2013). Targeted resequencing in epileptic encephalopathies identifies de novo mutations in CHD2 and SYNGAP1. *Nat. Genet.* *45*, 825–830.
16. Le Ber, I., Moreira, M.C., Rivaud-Péchoix, S., Chamayou, C., Ochsner, F., Kuntzer, T., Tardieu, M., Saïd, G., Habert, M.O., Demarquay, G., et al. (2003). Cerebellar ataxia with oculomotor apraxia type 1: clinical and genetic studies. *Brain* *126*, 2761–2772.
17. Anheim, M., Monga, B., Fleury, M., Charles, P., Barbot, C., Salih, M., Delaunoy, J.P., Fritsch, M., Arning, L., Synofzik, M., et al. (2009). Ataxia with oculomotor apraxia type 2: clinical, biological and genotype/phenotype correlation study of a cohort of 90 patients. *Brain* *132*, 2688–2698.
18. Bernstein, N.K., Williams, R.S., Rakovszky, M.L., Cui, D., Green, R., Karimi-Busheri, F., Mani, R.S., Galicia, S., Koch, C.A., Cass, C.E., et al. (2005). The molecular architecture of the mammalian DNA repair enzyme, polynucleotide kinase. *Mol. Cell* *17*, 657–670.
19. Liu, X., Jian, X., and Boerwinkle, E. (2013). dbNSFP v2.0: a database of human non-synonymous SNVs and their functional predictions and annotations. *Hum. Mutat.* *34*, E2393–E2402.
20. Reynolds, J.J., Walker, A.K., Gilmore, E.C., Walsh, C.A., and Caldecott, K.W. (2012). Impact of PNKP mutations associated with microcephaly, seizures and developmental delay on enzyme activity and DNA strand break repair. *Nucleic Acids Res.* *40*, 6608–6619.
21. Weinfeld, M., Mani, R.S., Abdou, I., Aceytuno, R.D., and Glover, J.N. (2011). Tidying up loose ends: the role of polynucleotide kinase/phosphatase in DNA strand break repair. *Trends Biochem. Sci.* *36*, 262–271.
22. Caldecott, K.W. (2008). Single-strand break repair and genetic disease. *Nat. Rev. Genet.* *9*, 619–631.
23. Segal-Raz, H., Mass, G., Baranes-Bachar, K., Lerenthal, Y., Wang, S.Y., Chung, Y.M., Ziv-Lehrman, S., Ström, C.E., Helleday, T., Hu, M.C., et al. (2011). ATM-mediated phosphorylation of polynucleotide kinase/phosphatase is required for effective DNA double-strand break repair. *EMBO Rep.* *12*, 713–719.
24. Derheimer, F.A., and Kastan, M.B. (2010). Multiple roles of ATM in monitoring and maintaining DNA integrity. *FEBS Lett.* *584*, 3675–3681.
25. Whiteside, J.R., Box, C.L., McMillan, T.J., and Allinson, S.L. (2010). Cadmium and copper inhibit both DNA repair activities of polynucleotide kinase. *DNA Repair (Amst.)* *9*, 83–89.