

**REPORT****Exome sequencing in undiagnosed inherited and sporadic ataxias****Angela Pyle,<sup>1</sup> Tania Smertenko,<sup>1</sup> David Bargiela,<sup>1</sup> Helen Griffin,<sup>1</sup> Jennifer Duff,<sup>1</sup> Marie Appleton,<sup>2</sup> Konstantinos Douroudis,<sup>1</sup> Gerald Pfeffer,<sup>1</sup> Mauro Santibanez-Koref,<sup>1</sup> Gail Eglon,<sup>1</sup> Patrick Yu-Wai-Man,<sup>1</sup> Venkateswaran Ramesh,<sup>3</sup> Rita Horvath<sup>1,\*</sup> and Patrick F. Chinnery<sup>1,\*</sup>**

\*These authors contributed equally to this work.

Inherited ataxias are clinically and genetically heterogeneous, and a molecular diagnosis is not possible in most patients. Having excluded common sporadic, inherited and metabolic causes, we used an unbiased whole exome sequencing approach in 35 affected individuals, from 22 randomly selected families of white European descent. We defined the likely molecular diagnosis in 14 of 22 families (64%). This revealed *de novo* dominant mutations, validated disease genes previously described in isolated families, and broadened the clinical phenotype of known disease genes. The diagnostic yield was the same in both young and older-onset patients, including sporadic cases. We have demonstrated the impact of exome sequencing in a group of patients notoriously difficult to diagnose genetically. This has important implications for genetic counselling and diagnostic service provision.

- 1 Wellcome Trust Centre for Mitochondrial Research, Institute of Genetic Medicine, Newcastle University, Central Parkway, Newcastle upon Tyne, NE1 3BZ, UK
- 2 Department of Clinical Biochemistry, Royal Victoria Infirmary, Newcastle upon Tyne Foundation Hospitals NHS Trust, Newcastle upon Tyne, NE1 4LP, UK
- 3 Department of Paediatric Neurology, Royal Victoria Infirmary, Newcastle upon Tyne Foundation Hospitals NHS Trust, Newcastle upon Tyne, NE1 4LP, UK

Correspondence to: Patrick F. Chinnery,  
Institute of Genetic Medicine,  
Newcastle University, Central, Parkway,  
Newcastle upon Tyne,  
NE1 3BZ, UK  
E-mail: patrick.chinnery@ncl.ac.uk

**Keywords:** ataxia; whole exome sequencing

**Introduction**

The inherited ataxias are clinically and genetically heterogeneous, presenting at any age, and usually without features pointing to a specific molecular diagnosis (Anheim *et al.*, 2010; Jayadev and Bird, 2013). Defining the genetic aetiology is important because clinically similar disorders can have different recurrence risks, and in some

instances there are treatment implications (Jayadev and Bird, 2013).

The first step in investigating patients is the exclusion of structural, toxic and inflammatory disorders. Following this, testing for the more common trinucleotide repeat neurometabolic disorders provides a diagnosis in up to ~60% of familial cases. After excluding the more common genetic causes, over 40 different loci are implicated. Subsequent

investigation usually proceeds on a gene-by-gene basis, which takes considerable time, and is expensive. Many patients do not receive a genetic diagnosis for many years, if at all, limiting genetic counselling and prenatal diagnosis.

Targeted next-generation sequencing panels have been shown to increase the diagnostic yield in suspected inherited ataxia, particularly in those with adolescent onset and a family history (Hoischen *et al.*, 2010). However, after excluding common forms of spinocerebellar ataxia (SCA) the overall detection rate remains <20% in routine clinical practice (Nemeth *et al.*, 2013). This may be due to the limited portfolio of genes included on custom-designed platforms (Nemeth *et al.*, 2013), but the clinical overlap of different neurogenetic syndromes presents a further challenge. In some patients, ataxia may develop in conjunction with spastic paraplegia or a neuropathy, and the initial clinical presentation may mislead the clinician to select the wrong multi-gene panel. Not based on any prior assumption about the underlying gene defect, whole exome sequencing provides a potential solution to these problems. Here we set out to determine the impact of whole exome sequencing in patients with suspected inherited ataxia who did not have one of the more common forms diagnosed on routine genetic testing.

## Materials and methods

### Inclusion criteria

We studied 35 individuals with unexplained ataxia identified at random through routine referrals to the regional neurogenetics service at Newcastle upon Tyne, England. The clinic serves a region of ~3 million in the north of England. To ensure that the cohort reflected standard clinical practice, we did not select cases based on age, gender or the presence of a family history. Prior to inclusion, all had routine clinical investigations to exclude treatable causes of acquired ataxia, including brain MRI and CSF examination with oligoclonal band analysis. All had negative genetic testing for SCA1, 2, 3, 6, 7, 17 (now known as *ATXN1*, *ATXN2*, *CACNA1A*, *ATXN7*, and *TBP*, respectively), *DRPLA* (now known as *ATN1*), *FA* (now known as *FXN*) and *FMR1* in adult males, which are positive in 17% of routine referrals to our clinic.

### Demographic and clinical characteristics

The 35 affected individuals were from 22 families of white European descent. The mean age was 25 years [standard deviation (SD) 14, range 3–57 years] at the time of the study, and 14 were male. Detailed clinical features and the results of clinical investigations are shown in Table 1 and the Supplementary material. Twenty-five of 35 individuals had a family history, suggestive of dominant inheritance in 11/25, and autosomal recessive in 14/25. There was no known consanguinity.

## Molecular genetics and bioinformatics

Blood genomic DNA was fragmented, exome enriched and sequenced (Illumina TruSeq™ 62 Mb and HiSeq 2000, 100 bp paired-end reads). Coverage data are summarized in Supplementary Tables 1 and 2. In-house bioinformatic analysis included alignment to UCSC hg19, using Burrows-Wheeler Aligner; duplicate removal (Picard v1.85) and variant detection (Varscan v2.2) (Koboldt *et al.*, 2009), Dindel v1.01 (Albers *et al.*, 2011). Further analysis was performed on variants with a minor allele frequency <0.01 in several databases: dbSNP137, 1000 Genomes (April 2012 data release), the National Heart, Lung and Blood Institute (NIH) Exome Sequencing Project (ESP) 6500 exomes, and 286 unrelated in-house controls. Rare homozygous and compound heterozygous variants were defined, and protein altering and/or putative ‘disease causing’ mutations, along with their functional annotation, were identified using ANNOVAR (Wang *et al.*, 2010). Putative pathogenic variants were confirmed by Sanger sequencing using custom-designed primers (<http://frodo.wi.mit.edu>) (ABI BigDye® v3.1 3130xl Genetic Analyzer, Life Technologies). Comparative genomic hybridization was performed in presumed recessive cases where a single likely pathogenic allele was found in a strong candidate gene. Quantitative pyrosequencing was used to determine the proportion of mutated alleles in given tissues (Pyromark v2.0, Qiagen; Supplementary Table 3).

### Variant classification

Variants were defined using *a priori* criteria: (i) confirmed pathogenic—a variant previously shown to be pathogenic, or in a known ataxia disease gene where the variant was predicted to affect protein structure or function, and segregated with at least one additional affected family member; (ii) possible pathogenic variants—variant in a known ataxia gene and predicted to affect protein function but not fulfilling all of the above criteria; and (iii) variants of uncertain significance or no candidate variants found.

## Results

Confirmed pathogenic variants were found in 9/22 probands (41%) (Table 1 and Supplementary Table 4). Three families had novel compound SACS mutations, each found in two affected siblings, including a full deletion of SACS detected from exome coverage (Pyle *et al.*, 2012, 2013). Known compound heterozygous *SPG7* mutations were found in three affected individuals from one family with no spasticity (Casari *et al.*, 1998). Two siblings presenting with adult-onset ataxia had compound heterozygous mutations in *NPC1*, confirmed by subsequent oxysterol analysis (Supplementary Table 5) (Carstea *et al.*, 1997). Likely *de novo* dominant *TUBB4A* mutations were found in two families (Simons *et al.*, 2013). One family showed varying degrees of mosaicism in the mildly affected mother and heterozygosity in the severely affected offspring (Fig. 1 and Supplementary Table 6). A *de novo* dominant

**Table 1 | Clinical presentation, laboratory investigations and exome sequencing results for 35 patients with unexplained ataxia**

ID, Sex	Age of onset (y)	Family history	Symptom at onset	Clinical signs			Investigations				Gene	Variants			
				Cerebellar ataxia	Eyes	Cerebellar dysarthria	Electrophysiology	MRI	Generalized atrophy	Normal			Denyelinating sensorimotor neuropathy	Cerebellar atrophy	Generalized atrophy
<b>Confirmed pathogenic</b>															
P1, M	Teens	Aff sib (P2)	Learning difficulties	+	-	-	-	-	+	-	-	+	-	SACS	c.2076delGp.Thr692Thr fs*713; c.3965_3966delACp.Gly1322Valfs*1343 <sup>s</sup>
P2, F	26	Aff sib (P1)	Upper limb clumsiness	+	+	-	-	-	+	-	-	+	+	-	-
P3, M	Teens	Aff sib (P4)	Upper limb clumsiness	+	+	+	+	+	-	-	+	-	-	SACS	hemizygous c.13048G > T; p.Glu4350 <sup>fs</sup> ; 0.7Mb deletion on Chr13q12.12 <sup>s</sup>
P4, M	Childhood	Aff sib (P3)	Walking delay	+	+	+	+	+	-	-	+	+	+	-	-
P5, M	40	Aff sib (P6)	Gait disturbance	+	-	-	-	-	-	-	-	-	-	SACS	c.1580C > G; p.Ser527 <sup>fs</sup> ; c.6781C > A; p.Leu226 Ile <sup>s</sup>
P6, F	40s	Aff sib (P5)	Gait disturbance	+	+	-	-	-	-	-	-	-	-	-	-
P7, F	23	Aff sib (P8)	Speech, balance	+	+	-	-	-	-	-	-	+	-	KCNK3	het. c.1259G > A; p.Arg420His <sup>s</sup>
P8, F	57	Aff sib (P7)	Upper limb clumsiness	+	+	-	-	-	-	-	-	+	+	-	-
P9, M	30	Aff sib (P10)	Upper limb clumsiness	+	+	+	+	+	-	-	+	-	+	SPG7	c.1529C > T; p.Ala510Val; c.1715C > T; p.Ala572Val
P10, F	29	Aff sib (P9)	Upper limb clumsiness	+	-	+	-	-	-	-	-	+	+	-	-
P11, F	Infancy	Aff sib (P12), mother (P13)	Delayed motor and speech development	+	-	+	-	-	-	-	-	+	-	TUBB4A	het c.900C > A; p.Met300Ile <sup>s</sup>
P12, F	Infancy	Aff sib (P11), mother (P13)	Delayed motor and speech development	+	-	+	-	-	-	-	-	+	-	-	-
P13, F	30	Aff children (P11, P12)	Gait disturbance, slurred speech	+	-	+	-	-	-	-	+	+	-	-	-
P14, F	5	None	Learning difficulties, gait disturbance	+	-	-	-	-	-	-	+	-	-	TUBB4A	het c.1091C > A; p.Ala364Asp <sup>s</sup>
P15, F	21	2 Aff sibs (P17)	Gait disturbance	+	-	+	-	-	-	-	+	-	-	-	-
P16, F	21	2 Aff sibs (P16)	Upper limb clumsiness	+	+	+	+	+	-	-	-	-	-	NPCI	c.467T > C; p.Met156Thr c.709C > T; p.Pro237Ser <sup>s</sup>
P17, M	30	Aff mother, aff cousin (P19)	Speech disturbance	+	+	+	+	+	-	-	-	-	-	SLC1A3	het c.1361G > A; p.Arg454Gln <sup>s</sup>
P18, F	39	Aff cousin (P18), aff aunt	Upper limb clumsiness Speech disturbance	+	+	+	+	+	-	-	-	-	-	-	-
<b>Possibly pathogenic</b>															
P19, M	32	None	Gait disturbance	+	-	-	-	-	+	-	-	-	+	ZFYVE26	c.2338C > T; p.Arg780 <sup>fs</sup> ; c.2450delT; p.Leu817Cysfs*12
P20, F	40	None	Gait disturbance	+	-	+	-	-	-	-	-	-	-	WFS1	c.577A > C; p.Lys193Gln; c.1367G > A; p.Arg456His
P21, F	Infancy	None	Walking delay	+	-	+	+	+	-	-	-	-	-	FASTKD2	c.-66A > G; c.149A > G; p.Lys50Arg <sup>s</sup>
P22, M	37	Aff son (P23)	Upper limb clumsiness	+	-	+	-	-	-	-	-	-	+	ZFYVE27	c.805-2A > G <sup>s</sup>
P23, M	35	Aff father (P22)	Upper limb clumsiness	+	-	+	-	-	-	-	-	-	+	-	-
P24, M	25	Aff father (P22) Aff half-brother (P23)	Upper limb clumsiness	+	+	+	+	+	-	-	-	-	+	-	-
P25, F	11	None	Upper limb clumsiness	+	-	+	-	-	-	-	+	-	-	WNK1	c.1994C > T; p.Thr65Ile <sup>s</sup> ; c.3272C > T; p.Thr109Ile <sup>s</sup>

(continued)

Table 1 Continued

ID, Sex	Age of onset (y)	Family history	Symptom at onset	Clinical signs				Investigations				Gene	Variants	
				Cerebellar ataxia	Eyes		Cerebellar dysarthria	Electrophysiology		MRI				
					Jerky ocular pursuit	Gaze evoked nystagmus		Hypometric saccades	Optic atrophy		Axonal sensorimotor neuropathy			Demyelinating sensorimotor neuropathy
Uncertain significance or no candidate variants found														
P26, F	13	None	Upper limb clumsiness	+	+	+	–	–	–	–	–	–	KCNB2	c.1589C>T: p.Ser530Phe; c.2351T>C: p.Leu784Pro
P27, M	20	2. Aff children	Gait disturbance	+	+	+	–	–	–	+	–	–	ABCB7 POLG WFS1	hom (hemi) c.818G>A: p.Arg273Gln <sup>s</sup> c.2243G>C: p.Trp748Ser; c.3428A>G: p.Glu1143Gly het c.1294C>G: p.Leu432Val
P28, F	Childhood	Aff cousin	Gait disturbance	+	+	–	–	–	–	–	–	–	No candidate variants	No candidate variants
P29, F	40	Aff child (P30)	Gait disturbance	+	+	–	–	–	–	–	–	–	No candidate variants	No candidate variants
P30, M	21	Aff mother (P29)	Gait disturbance	+	–	–	–	–	–	+	–	–	No candidate variants	No candidate variants
P31, F	46	None	Gait disturbance	+	+	–	–	–	–	–	–	–	No candidate variants	No candidate variants
P32, M	31	None	Upper limb clumsiness	+	+	–	–	–	–	–	–	–	No candidate variants	No candidate variants
P33, M	Teens	Aff sib (P34)	Upper limb clumsiness	+	–	–	–	–	–	–	–	–	No candidate variants	No candidate variants
P34, M	11	Aff sib (P33)	Upper limb clumsiness	+	+	+	–	–	–	–	–	–	No candidate variants	No candidate variants
P35, F	20	Mother has hypotonia	Gait disturbance	+	+	–	–	–	–	–	–	–	No candidate variants	No candidate variants

M = male; F = female; Hom = homozygous; Het = heterozygous; + = present; – = absent, not applicable (test not carried out); sib = sibling; aff = affected; N = normal (all results normal/negative).

<sup>s</sup>Segregation analysis performed in the family.

Two have been reported previously (Pyle *et al.*, 2012, 2013).

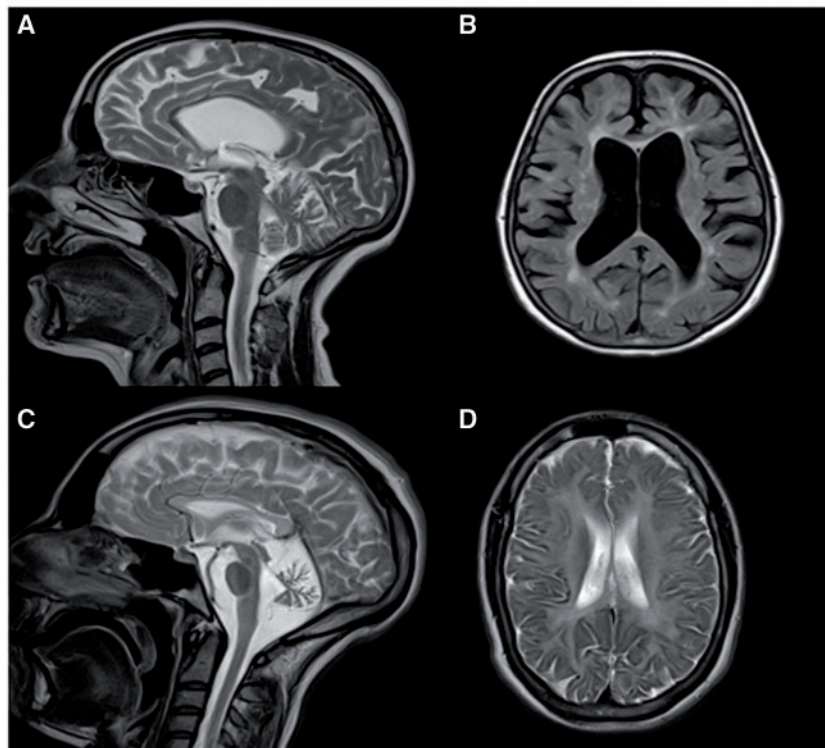
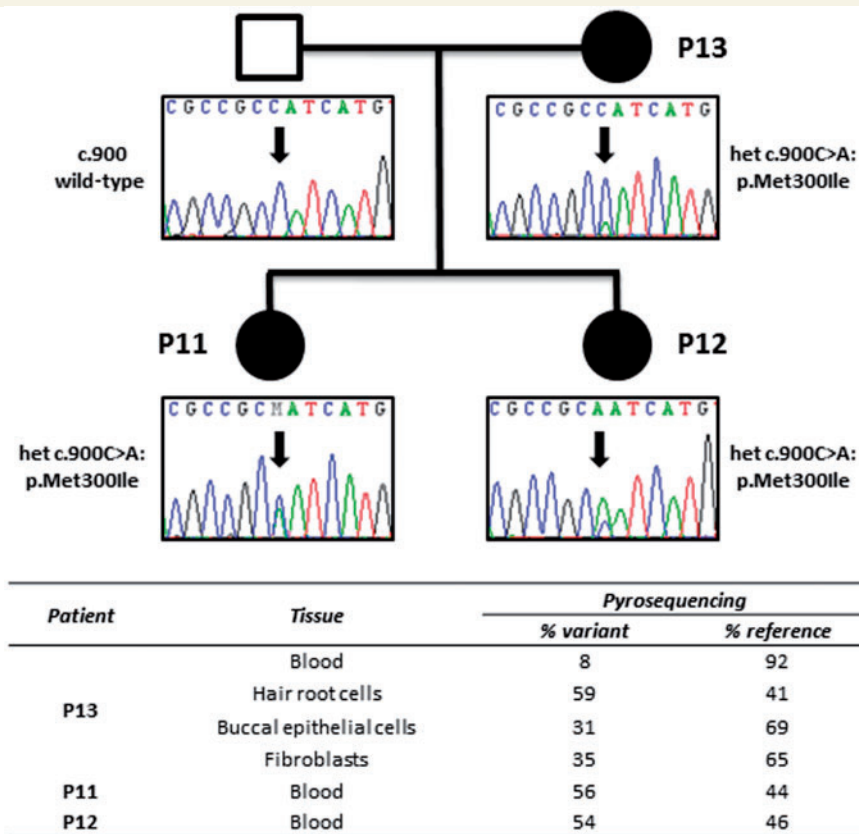
*SLC1A3* mutation segregated with ataxia in three members of a family (de Vries *et al.*, 2009). Finally, a dominant *KCNC3* mutation previously described in a large Filipino kindred and three European index cases (Waters *et al.*, 2006; Subramony *et al.*, 2013) segregated with ataxia in four members of a three-generation autosomal dominant pedigree.

Possible pathogenic variants were identified in 5/22 probands (23%), including compound heterozygous *FASTKD2* variants (c.149A>G: p.Lys50Arg, which is highly conserved; and c.-66A>G predicted to affect exon 1 splicing) (Ghezzi *et al.*, 2008). One proband had *de novo* compound heterozygous mutations in *ZFYVE26* (SPG15) (Fig. 2) (Hanein *et al.*, 2008). A predicted splice site mutation, c.805-2A>G (Supplementary Fig. 1), was detected in three members of an autosomal dominant pedigree in the previously described gene *ZFYVE27* (Mannan *et al.*, 2006). Previously described compound heterozygous *WFS1* mutations (p.Lys193Gln and p.Arg456His) were identified in one family (Cryns *et al.*, 2003). Detailed clinical data and the results of segregation analyses are included in the Supplementary material.

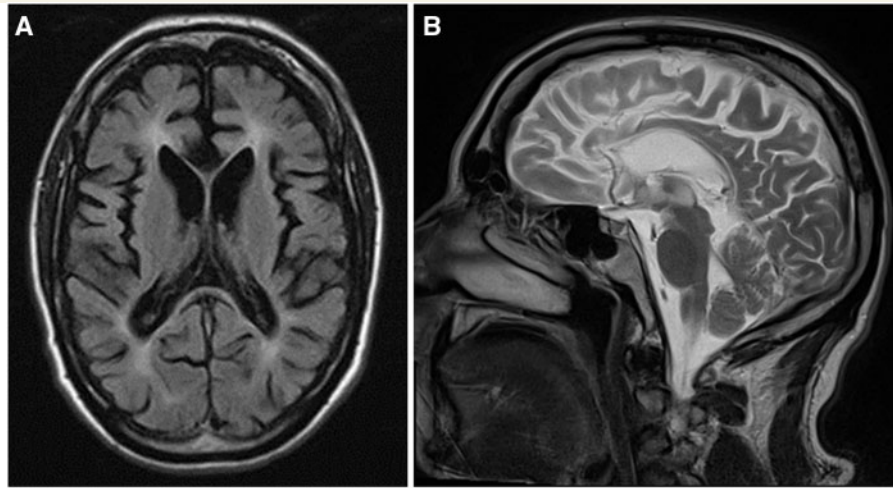
We were unable to confidently identify likely candidates in 10/22 probands but present potential candidate genes that may be recognized by others.

## Discussion

In this study whole exome sequencing identified confirmed pathogenic variants in 9/22 families with unexplained ataxia, with a similar diagnostic yield in both young (<20 years of age) 4/9 (44%) and older-onset (>20 years of age) 5/9 (56%) patients. Likely pathogenic variants were found in 5/22 probands, again with a similar diagnostic yield in both young (<20 years of age) 2/5 (40%) and older-onset (>20 years of age) 3/5 (60%) cases. The likely genetic cause was identified in 2/4 (50%) of adult-onset sporadic cases, and one gene was implicated in 3/6 (50%) of early-onset presumed recessive cases. Taken together these account for the likely molecular diagnosis in 64% of our families. These findings contrast with those of a targeted next generation sequencing approach, which identified the cause in 18% of patients in a similar study cohort (Nemeth *et al.*, 2013). Greater genome coverage provides the likely explanation for the higher diagnostic yield reported here, underscored by our observation that some of the definite mutations were found in genes known to cause ataxia, but not always considered 'ataxia genes', and thus not included on some multi-gene panels (Nemeth *et al.*, 2013). The extra coverage had the greatest impact on the diagnosis of adult-onset patients with recessive ataxia—a disorder generally considered to be of childhood onset. Our findings are endorsed by a recent study in children (Sawyer *et al.*, 2014), but unexpectedly



**Figure 1** *De novo* *TUBB4A* mutation in Patients P11, 12 and 13. *Top*: Segregation of *TUBB4A* c.900C>A (p.Met300Ile) mutation in Patient P13 (mother), and her two daughters. Confirmatory Sanger sequencing and pyrosequencing in different tissues from the mother showing tissue mosaicism in the mother for the presumed *de novo* dominant allele. *Middle*: Pyrosequencing results of this mutation in different tissues. *Bottom*: Brain MRI from Patient P13, the mother (**A** and **B**) and daughter Patient P12 (**C** and **D**). (**A** and **B**) T<sub>2</sub> and T<sub>1</sub> images showing generalized atrophy and periventricular high signal. (**C** and **D**) T<sub>2</sub> images showing marked cerebellar atrophy and diffuse hypomyelination.



**Figure 2** T<sub>1</sub> axial (A) and T<sub>2</sub> sagittal (B) MRI in Patient P19. The MRI shows global cerebral atrophy with relative sparing of the occipital lobes and cerebellum, and marked atrophy of the corpus callosum previously described in patients with mutations in *ZFYVE26* (SPG15) (Goizet *et al.*, 2009).

show the potential impact of exome sequencing in patients with suspected inherited ataxia presenting at any age.

With 11 genes implicated in 14 families, and no recurrent mutations in the same gene, our observations reaffirm the known genetic heterogeneity of inherited ataxia in an outbred European population. Similar genetic heterogeneity was also noted in 46 patients with sporadic and familial cerebellar ataxia studied with exome sequencing (Fogel *et al.*, 2014). Of the 16 patients with a confirmed diagnosis, only one of the disease genes identified by Fogel *et al.* (2014) was also identified in our patients (*SPG7*), and only two of the genes implicated in the remainder (*WFS1*, *ZFYVE26*) were also implicated in the patients we describe here (Table 1). In both our study and the work of Fogel *et al.*, (2014), the non-targeted exome sequencing approach increased the likelihood of detecting causal variants, particularly those in newly described disease genes, or exceptionally rare disease genes not found on targeted capture arrays. Although we did not directly compare exome sequencing to an ‘ataxia multi-gene panel’, our results show that exome sequencing is highly likely to have a greater diagnostic yield. In one of the largest panels produced to date, Nemeth *et al.* (2013) studied 117 known and putative ataxia genes. Eighty-nine per cent of the coding sequence of these genes was covered >20-fold using our exome sequencing approach, so it is highly likely that we would have detected the mutations found in their cohort. On the other hand, only 29% of the disease genes that we identified were included in the multi-gene panel of Nemeth *et al.* (2013), including genes not previously considered to be ‘ataxia genes’, such as *SPG7*. Moreover, there may be limited overlap between different multi-gene panels, which are defined by specific laboratories reflecting ataxia in a given population, and based on ataxia genes known at that time. Sequencing the

entire exome is likely to address these concerns. Given that exome sequencing costs are approximately the same as a single candidate gene test in most diagnostic laboratories, exome sequencing provides a rapid and cost effective means of reaching a diagnosis in a group of patients that have been notoriously difficult to diagnose at the molecular level.

A further advantage of the exome-based approach is the identification of known disease genes in a different clinical context, broadening the phenotype. With the potential for false positives, we also sought corroborative phenotypic or biochemical data when we detected a variant in an unexpected gene. This approach confirmed the diagnosis of Niemann Pick type C in two siblings with adult-onset ataxia, but lacking the characteristic eye movement disorder seen in childhood (Patients P15 and P16), and also prompted careful review of neuroimaging, which confirmed the thin corpus callosum characteristic of *ZFYVE26*/SPG15 (Patient P19, Fig. 2) (Goizet *et al.*, 2009). In these patients, exome sequencing highlighted likely causal variants, which were substantiated by subsequent clinical and biochemical studies. On the other hand, we observed an adult-onset for ataxic disorders generally considered to present in childhood (SACS), and atypical clinical presentations, such as saccinopathy presenting with a Charcot–Marie–Tooth phenotype. Likewise, although ataxia has been described as a presenting feature of *SPG7*, the absence of spasticity was thought to reduce the likelihood of a positive *SPG7* result, moving this specific gene test down the priority list. These observations demonstrate the importance of iterating between the clinical and genetic data to maximize the potential of large-scale sequencing, not only providing a diagnosis for a specific family, but also advancing our understanding of the phenotypic spectrum of specific disease genes.

We did not observe many of the rare autosomal recessive forms of ataxia observed in a similar sized UK cohort (Nemeth *et al.*, 2013), reflecting the genetic heterogeneity of inherited ataxia. We suspect that the extreme genetic heterogeneity has prevented us from identifying any recurrent novel ataxia genes with any degree of confidence. However, publishing our findings will hopefully allow others to refer to the variant data in our undiagnosed cases, and thus reach a final conclusion.

The use of exome sequencing in neurogenetic disorders is endorsed by recent findings in 24 cases (15 families) with undiagnosed suspected inherited polyneuropathies (Charcot–Marie–Tooth disease), where pathogenic mutations were found in five probands and possible causal mutations in three (Klein *et al.*, 2014), giving a total yield of 53%. An even higher yield may be possible for presumed autosomal recessive disorders when mother–father–affected child trios are sequenced, given the recent findings in 55 families with hereditary spastic paraplegia where the likely diagnosis was defined in 75% (Novarino *et al.*, 2014). However, identifying trios is challenging for adult-onset cases where parental samples are not readily available, and the advantage of trios is not so apparent for autosomal dominant disorders, which account for 44% of ataxia patients we describe here.

Although the whole exome approach provides a major advance in likely diagnostic yield in suspected inherited ataxia, the approach may uncover unexpected findings, such as mutations in genes known to cause familial cancer syndromes, or a different neurodegenerative disease such as familial Alzheimer’s disease. These incidental findings are rare, and were not discovered in this study, but must be considered when taking informed consent for exome studies (Bamshad *et al.*, 2011). It should also be borne in mind that our study was based on only 22 probands from one geographic region. These patients were from a largely outbred stable population of white European extraction, and the spectrum of different disease-causing alleles and genes may differ in other parts of the world, particularly if there is evidence of a founder effect. In these contexts a more focused sequencing approach may be more appropriate. Our study also had wide entry criteria, reflecting routine clinical practice. Although it is reassuring that exome sequencing appeared to be effective across this clinical spectrum, larger studies may show that it is more or less effective in specific clinical subgroups. Finally, we were unable to identify the likely genetic basis of the ataxia in 36% families. Exome sequencing the parents may reveal the discordant heterozygous variant responsible for the disorder that is otherwise challenging to identify. This is particularly relevant for congenital or childhood-onset ataxias (Ohba *et al.*, 2013), but many *de novo* dominant mutations can also cause adult-onset ataxia, as shown here (Fig. 1). It remains possible that some of the later-onset cases are not primarily genetically determined; however, given the relevant family history in most, there are also likely to be technical

explanations. Exome capture does not provide complete coverage of all coding regions of the genome, particularly those with GC-rich regions. In addition, large genomic rearrangements and trinucleotide repeat sequences are not reliably detected from exome-capture data. It is therefore possible that some of the undiagnosed cases have pathogenic trinucleotide expansions not routinely tested in our region, including SCA10, 12 and 31, although these are very rare in our experience. Finally, it is likely that some causal variants will reside within non-coding regulatory regions. Some of these issues will be resolved by whole genome sequencing, although not without substantial additional cost and bioinformatics complexity.

## Acknowledgements

We are grateful to the physicians who referred patients for further investigation.

## Funding

P.F.C. is an Honorary Consultant Neurologist at Newcastle upon Tyne Foundation Hospitals NHS Trust, is a Wellcome Trust Senior Fellow in Clinical Science (101876/Z/13/Z), and a UK NIHR Senior Investigator. P.F.C. receives additional support from the Wellcome Trust Centre for Mitochondrial Research (096919Z/11/Z), the Medical Research Council (UK) Centre for Translational Muscle Disease research (G0601943), and EU FP7 TIRCON, and the National Institute for Health Research (NIHR) Newcastle Biomedical Research Centre based at Newcastle upon Tyne Hospitals NHS Foundation Trust and Newcastle University. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. R.H. was supported by the Medical Research Council (UK) (G1000848) and the European Research Council (309548). G.P. is the recipient of a Bisby Fellowship from the Canadian Institutes of Health Research. P.Y.W.M. is the recipient of an MRC Clinician Scientist Fellowship.

Wellcome Trust, Medical Research Council (UK), National Institute for Health Research (NIHR), Newcastle NIHR Biomedical Research Centre, European Research Council. The funding sources were not involved in the in-study design, in the collection, analysis, and interpretation of data, in the writing of the report, or in the decision to submit the paper for publication.

## Supplementary material

Supplementary material is available at *Brain* online.

## References

- Albers CA, Lunter G, MacArthur DG, McVean G, Ouwehand WH, Durbin R. Dindel: accurate indel calls from short-read data. *Genome Res* 2011; 21: 961–73.
- Anheim M, Fleury M, Monga B, Laugel V, Chaigne D, Rodier G, et al. Epidemiological, clinical, paraclinical and molecular study of a cohort of 102 patients affected with autosomal recessive progressive cerebellar ataxia from Alsace, Eastern France: implications for clinical management. *Neurogenetics* 2010; 11: 1–12.
- Bamshad MJ, Ng SB, Bigham AW, Tabor HK, Emond MJ, Nickerson DA, et al. Exome sequencing as a tool for Mendelian disease gene discovery. *Nat Rev Genet* 2011; 12: 745–55.
- Carstea ED, Morris JA, Coleman KG, Loftus SK, Zhang D, Cummings C, et al. Niemann-Pick C1 disease gene: homology to mediators of cholesterol homeostasis. *Science* 1997; 277: 228–31.
- Casari G, De Fusco M, Ciarmatori S, Zeviani M, Mora M, Fernandez P, et al. Spastic paraplegia and OXPHOS impairment caused by mutations in paraplegin, a nuclear-encoded mitochondrial metalloprotease. *Cell* 1998; 93: 973–83.
- Cryns K, Sivakumaran TA, Van den Ouweland JM, Pennings RJ, Cremers CW, Flothmann K, et al. Mutational spectrum of the WFS1 gene in Wolfram syndrome, nonsyndromic hearing impairment, diabetes mellitus, and psychiatric disease. *Hum Mutat* 2003; 22: 275–87.
- de Vries B, Mamsa H, Stam AH, Wan J, Bakker SL, Vanmolkot KR, et al. Episodic ataxia associated with EAAT1 mutation C186S affecting glutamate reuptake. *Arch Neurol* 2009; 66: 97–101.
- Fogel BL, Lee H, Deignan JL, Strom SP, Kantarci S, Wang X, et al. Exome sequencing in the clinical diagnosis of sporadic or familial cerebellar ataxia. *JAMA Neurol* 2014; 71: 1237–46.
- Ghezzi D, Saada A, D'Adamo P, Fernandez-Vizorra E, Gasparini P, Tiranti V, et al. FASTKD2 nonsense mutation in an infantile mitochondrial encephalomyopathy associated with cytochrome c oxidase deficiency. *Am J Hum Genet* 2008; 83: 415–23.
- Goizet C, Boukhris A, Maltete D, Guyant-Marechal L, Truchetto J, Mundwiller E, et al. SPG15 is the second most common cause of hereditary spastic paraplegia with thin corpus callosum. *Neurology* 2009; 73: 1111–9.
- Hanein S, Martin E, Boukhris A, Byrne P, Goizet C, Hamri A, et al. Identification of the SPG15 gene, encoding spastizin, as a frequent cause of complicated autosomal-recessive spastic paraplegia, including Kjellin syndrome. *Am J Hum Genet* 2008; 82: 992–1002.
- Hoischen A, Gilissen C, Arts P, Wieskamp N, van der Vliet W, Vermeer S, et al. Massively parallel sequencing of ataxia genes after array-based enrichment. *Hum Mutat* 2010; 31: 494–9.
- Jayadev S, Bird TD. Hereditary ataxias: overview. *Genet Med* 2013; 15: 673–83.
- Klein CJ, Middha S, Duan X, Wu Y, Litchy WJ, Gu W, et al. Application of whole exome sequencing in undiagnosed inherited polyneuropathies. *J Neurol Neurosurg Psychiatry* 2014; 85: 1265–72.
- Koboldt DC, Chen K, Wylie T, Larson DE, McLellan MD, Mardis ER, et al. VarScan: variant detection in massively parallel sequencing of individual and pooled samples. *Bioinformatics* 2009; 25: 2283–5.
- Mannan AU, Krawen P, Sauter SM, Boehm J, Chronowska A, Paulus W, et al. ZFYVE27 (SPG33), a novel spastin-binding protein, is mutated in hereditary spastic paraplegia. *Am J Hum Genet* 2006; 79: 351–7.
- Nemeth AH, Kwasniewska AC, Lise S, Parolin Schneckenberg R, Becker EB, Bera KD, et al. Next generation sequencing for molecular diagnosis of neurological disorders using ataxias as a model. *Brain* 2013; 136 (Pt 10): 3106–18.
- Novarino G, Fenstermaker AG, Zaki MS, Hofree M, Silhavy JL, Heiberg AD, et al. Exome sequencing links corticospinal motor neuron disease to common neurodegenerative disorders. *Science* 2014; 343: 506–11.
- Ohba C, Osaka H, Iai M, Yamashita S, Suzuki Y, Aida N, et al. Diagnostic utility of whole exome sequencing in patients showing cerebellar and/or vermis atrophy in childhood. *Neurogenetics* 2013; 14: 225–32.
- Pyle A, Griffin H, Duff J, Bennett S, Zwolinski S, Smertenko T, et al. Late-onset saccinopathy diagnosed by exome sequencing and comparative genomic hybridization. *J Neurogenet* 2013; 27: 176–82.
- Pyle A, Griffin H, Yu-Wai-Man P, Duff J, Eglon G, Pickering-Brown S, et al. Prominent sensorimotor neuropathy due to SACS mutations revealed by whole-exome sequencing. *Arch Neurol* 2012; 69: 1351–4.
- Sawyer SL, Schwartzentruber J, Beaulieu CL, Dymont D, Smith A, Warman Chardon J, et al. Exome sequencing as a diagnostic tool for pediatric-onset ataxia. *Hum Mutat* 2014; 35: 45–9.
- Simons C, Wolf NI, McNeil N, Caldovic L, Devaney JM, Takanohashi A, et al. A de novo mutation in the beta-tubulin gene TUBB4A results in the leukoencephalopathy hypomyelination with atrophy of the basal ganglia and cerebellum. *Am J Hum Genet* 2013; 92: 767–73.
- Subramony SH, Advincula J, Perlman S, Rosales RL, Lee LV, Ashizawa T, et al. Comprehensive phenotype of the p.Arg420his allelic form of spinocerebellar ataxia type 13. *Cerebellum* 2013; 12: 932–6.
- Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 2010; 38: e164.
- Waters MF, Minassian NA, Stevanin G, Figueroa KP, Bannister JP, Nolte D, et al. Mutations in voltage-gated potassium channel KCNC3 cause degenerative and developmental central nervous system phenotypes. *Nat Genet* 2006; 38: 447–51.