LETTER TO THE EDITOR

Reply: Evaluation of exome sequencing variation in undiagnosed ataxias

Angela Pyle, Helen Griffin, Michael J. Keogh, Rita Horvath and Patrick F. Chinnery

Institute of Genetic Medicine, Newcastle University, NE1 3BZ, UK
Correspondence to: Patrick F. Chinnery,
E-mail: p.f.chinnery@ncl.ac.uk

Sir,

Erin Sandford and colleagues raise three main concerns that they believe undermine the likely causal nature of some of the genetic variants we detected by exome sequencing undiagnosed patients with ataxia (Pyle et al., 2015):

(i) Some of our proposed pathogenic variants have a minor allele frequency (MAF) of >0.01 in the population. Although useful as an initial guide, the 1% MAF is an arbitrary cut-off designed to rapidly filter-out common polymorphic genetic variants from a large exome data set. However, this must be done with care for several reasons. First, allele frequencies vary from database to database, and throughout the world. To choose their example, c.709C>T p.Pro237Ser in Patients P15 and P16 has a MAF = 0.0153 in the US-based ESP6500 (their Table 1), which contrasts with our local population (MAF = 0.0052), and the UK-based 1000 Genomes project (MAF = 0.0064). To our mind it would be ridiculous to reject this variant as non-pathogenic because one database documents a MAF > 0.01. Surely the important point here is simply that the variant is rare? Many well-known pathogenic alleles have a much greater carrier frequency than MAF 0.01, such as the DeltaF508 variant in CFTR, which is the most common cause of cystic fibrosis, or pathogenic expansions in FXN, which cause Friedreich’s ataxia. A MAF threshold of <0.01 is only a guide, and if rigidly enforced, will ‘throw the baby out with the bathwater’, causing further delays in diagnosis at additional cost.

(ii) Sandford et al. (2015) state that ‘multiple independent predictive and experimental approaches should be utilized to corroborate predicted pathogenicity’. We wholeheartedly agree with this point, we did so in the less certain cases, and we continue to do so for some of the patients described in our paper. However, we would argue that it is unwise to doggedly follow a stepwise algorithm of functional validation ‘in series’, before searching for other patients with the same disorder and a similar genetic variant. Given the many novel genetic variants being identified in each individual, it is a major challenge to decide which variants to evaluate further with functional tests, and more importantly, which can be safely ignored. Given the myriad of different mechanisms implicated in inherited ataxia (Anheim et al., 2012), it would be a colossal undertaking to carry out functional analyses for every possible pathogenic variant in each patient. In many instances, the clinico-pathological relevance of a specific functional test is not well understood, and the strongest evidence for pathogenicity comes from observing the same genetic variant in patients with a similar clinical phenotype on a different genetic background (MacArthur et al., 2014). It is therefore sensible to carry out functional analyses and further genetic studies ‘in parallel’, carefully interpreting emerging data to accelerate the pace towards a confident diagnosis. This can only be done by sharing findings in a form that others can interpret, which boils down to an accurate description of the genetic and clinical data. This was the explicit aim of our study (see our Discussion, p. 282, first paragraph). We precisely defined our classification of variants (p. 277), which was in keeping with the American College of Medical Genetics Guidelines (Richards et al., 2008). In this way, readers can make up their own mind about the likely pathogenic role of a particular variant in a particular disease. Surely by publishing our findings, we have reduced the chance that this will lead to ‘false clinical tests in the future’, and not the contrary, which is the tenet of the letter by Sandford et al.? (iii) The final point made by Sandford et al. is that: ‘multiple unrelated cases with similar presentation should be reported before assigning a gene as causative for a particular phenotype’, citing this as a reason for not assigning pathogenicity in Patients P5/P6 (SACS), P15/P16 (NPC1), and P17/P18 (SLC1A3). However, similar phenotypes have been previously described for all three of these genes (Imrie et al., 2007, Sevin et al., 2007, de Vries et al., 2009; Guernsey et al., 2010). Had these authors not published their findings,
we would not be in a position to provide confirmatory observations in our study, adding further weight to the importance of sharing clinical and genetic data for variant phenotypes. We absolutely agree that reporting multiple cases of particular genotypes is paramount to understand phenotypic diversity, but for rare diseases, the sensible approach must be to anonymously report clinical descriptions in association with specific genetic variants, and build the publically available data. There are several ways of doing this, one of which is publication in specialist journals—which is exactly what we have done.

As stated at the beginning of their letter, the issues raised by Sandford *et al.* stem from their concern regarding the higher rate of ‘confirmed pathogenic’ and ‘possible pathogenic’ variants in our study (9/22) when compared to two other studies [9/50 (Nemeth *et al.*, 2013); and 16/76 (Fogel *et al.*, 2014)]. No doubt they realize that none of these frequencies are significantly different when compared statistically—so perhaps our findings are actually endorsed by the observations of others.

**References**


