Published in final edited form as:
Continuum (Minneap Minn). 2011 April ; 17(2 NEUROGENETICS): 347–352. doi:10.1212/01.CON.0000396975.87637.86.

INTERPRETATION OF GENETIC TESTING: VARIANTS OF UNKNOWN SIGNIFICANCE

Brent L. Fogel, MD, PhD

Abstract

As the number of genes available for commercial sequencing increases and the promise of clinical whole-genome sequencing becomes a reality, the interpretation of the results of these tests becomes more challenging for the practicing neurologist as these studies have the potential to detect novel genetic variants. Such reports are becoming more frequent in general practice, and neurologists are often left to puzzle over the relevance of these "variants of unknown significance," as such genetic changes are often described, and how to communicate this information to the patients and their families. This article will briefly illustrate how clinicians can use such results in the care of their patients. Only genetic variants involving coding sequence will be considered, although similar methods may also be applied to changes such as noncoding alterations or copy number variations. It is also important to note that in some cases, particularly those involving tests that only sequence select exons, negative test results may also require special interpretation.

Case

Note: This case has been previously published.[1]

A 40-year-old woman presented with progressive ataxia since the age of 23. She had saccadic visual pursuit, gait and appendicular ataxia, a distal loss of vibration sense, and areflexia. MRI of the brain showed cerebellar atrophy primarily affecting the midline vermis. She had no family history of neurologic disease. A detailed diagnostic evaluation[2, 3] did not reveal an acquired cause for her ataxia. On the basis of her history and clinical examination, an autosomal recessive cerebellar ataxia was suspected. Following diagnostic algorithms,[4] the senataxin gene, *SETX*, was sequenced for suspicion of ataxia with oculomotor apraxia, type 2 (AOA2; Online Mendelian Inheritance in Man [OMIM] # 606002). The patient was found to be a heterozygote for three separate missense sequence variants: c.1304A>G (p.H435R), c.2975A>G (p.K992R), and c.6590A>G (p.H2197R). At the time of original testing, all variants were novel and reported as being of "unknown significance."

Reprinted with permission from Fogel BL, Perlman S. Novel mutations in the senataxin DNA/RNA helicase domain in ataxia with oculomotor apraxia 2. Neurology 2006;67:2083–2084.

Address correspondence to Dr Brent L. Fogel, University of California, Los Angeles, Department of Neurology, 710 Westwood Plaza, Los Angles, CA 90095, bfogel@ucla.edu.

Relationship Disclosure: Dr Fogel receives research support from the National Institutes of Health.
Unlabeled Use of Products/Investigational Use Disclosure: Dr Fogel reports no disclosure.

DISCUSSION

This case highlights the dilemma of evaluating novel genetic testing results in a clinical setting. Three potential interpretations are presented to emphasize the major issues involved in the clinical utilization of these results.

Interpretation 1

The patient has a suspected hereditary ataxia phenotype consistent with AOA2, as well as coding variants in the *SETX* gene, which result in changes to the protein; therefore, the variants are likely causative.

Altough this interpretation may be correct, it does not account for the fact that sequence variants in genes are commonplace and often benign (e.g., single nucleotide polymorphisms), even if they lead to an amino acid change. The relative occurrence of polymorphisms can also vary between genes; *SETX*, for example, has a particularly high frequency of such variants.[5] A closer evaluation of the specific variants may establish their likelihood to be deleterious, but such a determination cannot be established solely by the sequence information provided thus far. It is particularly important to note that AOA2 is autosomal recessive and, by definition, requires mutations in both *SETX* alleles. The sequencing information as reported does not reveal whether the variants are all present on a single allele or how they may be distributed between the two. Further analysis will be required to determine whether these variants are pathogenic.

Interpretation 2

Although the patient has a suspected hereditary ataxia consistent with AOA2 and coding variants in *SETX*, the variants previously have not been described as mutations. Reliably determining whether the variants are causative is therefore not possible.

In contrast to the previous scenario, this interpretation is overly conservative because it does not account for the wealth of publicly available bioinformatics data to aid in assessing these variants. In general, a relatively simple analysis could be quite informative and useful in determining the next steps in the evaluation and management of this patient. In some cases, which may vary depending on the clinician and the practice, this may also be an appropriate stage for a subspecialty consultation.

Interpretation 3

The patient has a suspected hereditary ataxia consistent with AOA2 and coding variants in *SETX*, creating a high level of suspicion for this diagnosis. Although the variants in question have not been described previously, a more detailed analysis is warranted as they may be novel disease-causing mutations.

This interpretation is the most appropriate given this scenario but in practice often prompts a subspecialty consultation because many clinicians are unsure about how to approach the required analysis. Fortunately, publicly available resources greatly simplify the steps required to maximize an estimation of the probability that a mutation is deleterious. Although many strategies exist, a basic method is outlined here and can be modified depending on the clinical situation and the gene of interest. Several useful online resources can be found at the National Center for Biotechnology Information (NCBI) website (www.ncbi.nlm.nih.gov) and in Practice Table 1.

Step 1: Bioinformatics

Has the variant previously been reported in normal or diseased individuals? Commercial genetic testing services generally maintain databases of normal and disease-associated variations in the genes they sequence, but the completeness of such databases may vary. Therefore, a variant reported as having "unknown significance" may still have been identified previously. Depending on the gene, association with disease can be determined by examination of gene-specific mutation databases or, if the gene of interest is not categorized in such a database, by review of the literature. Determining whether the reported change is a benign single nucleotide polymorphism can be further evaluated by querying a database such as the NCBI Single Nucleotide Polymorphism Database (dbSNP) (www.ncbi.nlm.nih.gov/snp).

No published data existed at the time of the initial identification of the variants in this example, nor was a comprehensive gene-specific database for *SETX* variants available. They could now be identified through a literature review[1, 6, 7] or through the online *SETX* Variant Database (*149.142.212.78/LOVD/home.php?select_db=SETX*).[8] A search of the dbSNP for human *SETX* variants identifies only one of the three, p.K992R (rs61742937), as a rare variant found in 1.4% of chromosomes from a sample population of 36 healthy individuals from Utah. No homozygous individuals for the variant were reported, so it cannot be assumed to be benign.

Step 2: Protein Function

Is the variant likely to have biological significance? Because of the difficulty in reliably assessing the potential impact of synonymous mutations on gene expression and regulation aside from those that directly disrupt splice sites, only nonsynonymous mutations will be considered here. The simplest method is to utilize an amino acid scoring matrix designed to compare the evolutionary conservation of amino acid substitutions between related proteins. Although several such matrices exist, the two most utilized are the Point Accepted Mutation (PAM) matrix and the Blocks Substitution Matrix (BLOSUM). A detailed description of these matrices and their use[9] is beyond the scope of this discussion. Briefly, PAM matrices are based on percent divergence between two related protein sequences whereas BLOSUM matrices are based on percent similarity. Therefore, for this type of mutational analysis one should choose a low PAM matrix (eg, PAM1) or a high BLOSUM matrix (eg, BLOSUM90). Such an analysis does not account for the relationship of the amino acid residue in question to the entire protein, so a more detailed analysis can be performed using online resources that make predictions based upon the entire protein structure. Such analysis can be strengthened by examining whether the variant lies within a functional domain or whether the original amino acid is conserved phylogenetically across diverse species.

For this example, the p.H435R and p.H2197R changes are much less conservative than the p.K992R variant using both PAM1 and BLOSUM90 matrices. A conserved domain search identifies residue 2197 as part of a putative DNA helicase, but the other variants are not within known conserved domains.[1] Detailed analysis using PolyPhen-2[10] and SNPs3D[11] suggests p.K992R is benign whereas both p.H435R and p.H2197R are likely harmful. Phylogenic analysis, again using PolyPhen-2 and SNPs3D, identified lysine (K) as the most common residue at position 992; however, arginine (R) is the naturally observed equivalent residue in the horse (*Equus caballus*). Histidine (H) was conserved at positions 435 and 2197 in all mammalian species evaluated (including the rhesus macaque, white-tufted-ear marmoset, horse, giant panda, cow, dog, rat, mouse, rabbit, short-tailed gray opossum, and duck-billed platypus) suggesting both residues are likely functionally essential.

Step 3: Heritability

Is the variant inherited or did it arise *de novo*? For an autosomal recessive disorder, are multiple heterozygous variants on separate alleles? For an autosomal dominant disorder, do other affected family members share the same mutation? To answer these questions, additional family members must be available for genetic analysis. It is preferable that the parents of the proband be tested initially because this will most easily address the first two questions. For a suspected dominant mutation, detection in an affected parent or sibling is consistent with causality but is not conclusive. Conversely, detection in an unaffected sibling or parent may suggest the variant is benign, but could also reflect incomplete penetrance. Absence in either parent suggests the variant arose *de novo* (or may reflect alternate paternity). For compound heterozygous variants in suspected autosomal recessive diseases the variants should segregate between the parents. Failure to do so would not support this mode of inheritance. In some cases it may be cost-effective to contact the company that performed the original testing, as some may offer free testing to parents of the proband to facilitate this analysis.

In this example, genetic testing of the patient's unaffected parents and brother revealed that they were heterozygous for the p.H435R variant (father and brother) and the p.K992R and p.H2197R variants (mother), consistent with the pathogenicity predictions for p.H435R and p.H2197R. Taken together with the available clinical information and the above analysis, the p.H435R and p.H2197R were designated as disease-causing whereas the p.K992R mutation was suspected to be a benign polymorphism, and the patient was diagnosed with AOA2.[1] These predictions have subsequently been further supported for the p.K992R[6] and p.H2197R[7] variants.

In summary, these steps can be applied to novel genetic variants found in the causative genes for many mendelian disorders to help direct further evaluation and management of patients with suspected neurogenetic disease. It is important to note that predictive analysis should not be utilized as a substitute for expert clinical opinion.

This case illustrates some of the complexities involved in pursuing these types of evaluations, which must then be clearly communicated to the patient and family. This is often beyond the training of most practicing neurologists, and referral to a trained medical geneticist or genetic counselor should be considered, even before initiating the evaluation. For example, referral for genetic counseling is recommended for patients with rare genetic causes of stroke.[12] Subspecialty referral may also be appropriate for management of the neurogenetic disease in question.

Acknowledgments

The author wishes to thank Giovanni Coppola for helpful comments regarding this manuscript.

REFERENCES

1. Fogel BL, Perlman S. Novel mutations in the senataxin DNA/RNA helicase domain in ataxia with oculomotor apraxia 2. Neurology. 2006; 67(11):2083–2084. [PubMed: 17159128]
2. Fogel BL, Perlman S. An approach to the patient with late-onset cerebellar ataxia. Nat Clin Pract Neurol. 2006; 2(11):629–635. [PubMed: 17057750]
3. Klockgether T. Sporadic ataxia with adult onset: classification and diagnostic criteria. Lancet Neurol. 2010; 9(1):94–104. [PubMed: 20083040]
4. Fogel BL, Perlman S. Clinical features and molecular genetics of autosomal recessive cerebellar ataxias. Lancet Neurol. 2007; 6(3):245–257. [PubMed: 17303531]

5. Lynch DR, Braastad CD, Nagan N. Ovarian failure in ataxia with oculomotor apraxia type 2. Am J Med Genet A. 2007; 143A(15):1775–1777. [PubMed: 17593543]

6. Bernard V, Stricker S, Kreuz F, et al. Ataxia with oculomotor apraxia type 2: novel mutations in six patients with juvenile age of onset and elevated serum alpha-fetoprotein. Neuropediatrics. 2008; 39(6):347–350. [PubMed: 19569000]

7. Anheim M, Monga B, Fleury M, et al. Ataxia with oculomotor apraxia type 2: clinical, biological and genotype/phenotype correlation study of a cohort of 90 patients. Brain. 2009; 132(pt 10):2688–2698. [PubMed: 19696032]

8. Fokkema IF, den Dunnen JT, Taschner PE. LOVD: easy creation of a locus-specific sequence variation database using an "LSDB-in-a-box" approach. Hum Mutat. 2005; 26(2):63–68. [PubMed: 15977173]

9. Koonin, EV.; Galperin, MY. Sequence - evolution - function: computational approaches in comparative genomics. 1st ed.. Norwell, MA: Kluwer Academic; 2003.

10. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. Nat Methods. 2010; 7(4):248–249. [PubMed: 20354512]

11. Yue P, Melamud E, Moult J. SNPs3D: candidate gene and SNP selection for association studies. BMC Bioinformatics. 2006; 7:166. [PubMed: 16551372]

12. Goldstein LB, Bushnell CD, Adams RJ, et al. American Heart Association Stroke Council, Council on Cardiovascular Nursing, Council on Epidemiology and Prevention, Council for High Blood Pressure Research; Council on Peripheral Vascular Disease; Interdisciplinary Council on Quality of Care and Outcomes Research. Guidelines for the primary prevention of stroke. A guideline for healthcare professionals from the American Heart Association/American Stroke Association. Stroke. [published online ahead of print December 2, 2010].

Practice Table 1

Suggested Internet Resources for General Evaluation of Nonsynonymous Genetic Variants of Unknown Significance[a]

- **Literature Review**
  - PubMed (www.ncbi.nlm.nih.gov/pubmed)
- **Gene-Specific Variant Reporting**
  - GeneTests (www.ncbi.nlm.nih.gov/sites/GeneTests)
  - Human Genome Variation Society, Locus Specific Mutation Databases (www.hgvs.org/dblist/glsdb.html)
  - Leiden Open Variation Database, List of Public Installations (www.lovd.nl)
  - Online Mendelian Inheritance in Man (www.ncbi.nlm.nih.gov/omim)
- **Genome Browser**
  - UCSC Genome Bioinformatics (www.genome.ucsc.edu)
- **Single Nucleotide Polymorphism Database**
  - The Single Nucleotide Polymorphism Database (dbSNP) (www.ncbi.nlm.nih.gov/snp)
- **Protein Domain Analysis**
  - NCBI Conserved Domain Search (www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi)
- **Prediction of Variant Functional Effect/Phylogenetic Analysis**
  - PolyPhen-2 (genetics.bwh.harvard.edu/pph2)
  - SNPs3D (www.snps3d.org)

[a]This list is not comprehensive and is meant as a starting point for development of an individual analysis procedure.