UCLA
UCLA Previously Published Works

Title

Prevalence of spinocerebellar ataxia 36 in a US population

Permalink

Journal

ISSN

Authors

Valera, Juliana M
Diaz, Tatyana
Petty, Lauren E
et al.

Publication Date

DOI

Peer reviewed

Prevalence of spinocerebellar ataxia 36 in a US population

OPEN

Juliana M. Valera, MS
Tatyana Diaz
Lauren E. Petty, MS
Beatriz Quintáns, PhD
Zuleima Yáñez, PhD
Eric Boerwinkle, PhD
Donna Muzny, MS
Dmitry Akhmedov, PhD
Rebecca Berdeaux, PhD
Maria J. Sobrido, MD, PhD
Richard Gibbs, PhD
James R. Lupski, MD, PhD, DSc
Daniel H. Geschwind, MD, PhD
Susan Perlman, MD
Jennifer E. Below, PhD
Brent L. Fogel, MD, PhD

Correspondence to Dr. Fogel:
bfogel@ucla.edu

ABSTRACT

Objective: To assess the prevalence and clinical features of individuals affected by spinocerebellar ataxia 36 (SCA36) at a large tertiary referral center in the United States.

Methods: A total of 577 patients with undiagnosed sporadic or familial cerebellar ataxia comprehensively evaluated at a tertiary referral ataxia center were molecularly evaluated for SCA36. Repeat primed PCR and fragment analysis were used to screen for the presence of a repeat expansion in the *NOP56* gene.

Results: Fragment analysis of triplet repeat primed PCR products identified a GGCCTG hexanucleotide repeat expansion in intron 1 of *NOP56* in 4 index cases. These 4 SCA36-positive families comprised 2 distinct ethnic groups: white (European) (2) and Asian (Japanese [1] and Vietnamese [1]). Individuals affected by SCA36 exhibited typical clinical features with gait ataxia and age at onset ranging between 35 and 50 years. Patients also suffered from ataxic or spastic limbs, altered reflexes, abnormal ocular movement, and cognitive impairment.

Conclusions: In a US population, SCA36 was observed to be a rare disorder, accounting for 0.7% (4/577 index cases) of disease in a large undiagnosed ataxia cohort. *Neurol Genet* 2017;3:e174; doi: 10.1212/NXG.0000000000000174

GLOSSARY

ERIS = Error Rate In Sequencing (ERIS); **IBD** = identically by descent; **LOD** = logarithm of the odds; **MAF** = minor allele frequency; **RP-PCR** = repeat primed PCR; **SCA** = spinocerebellar ataxia; **SCA36** = spinocerebellar ataxia 36; **SNP** = single nucleotide polymorphism.
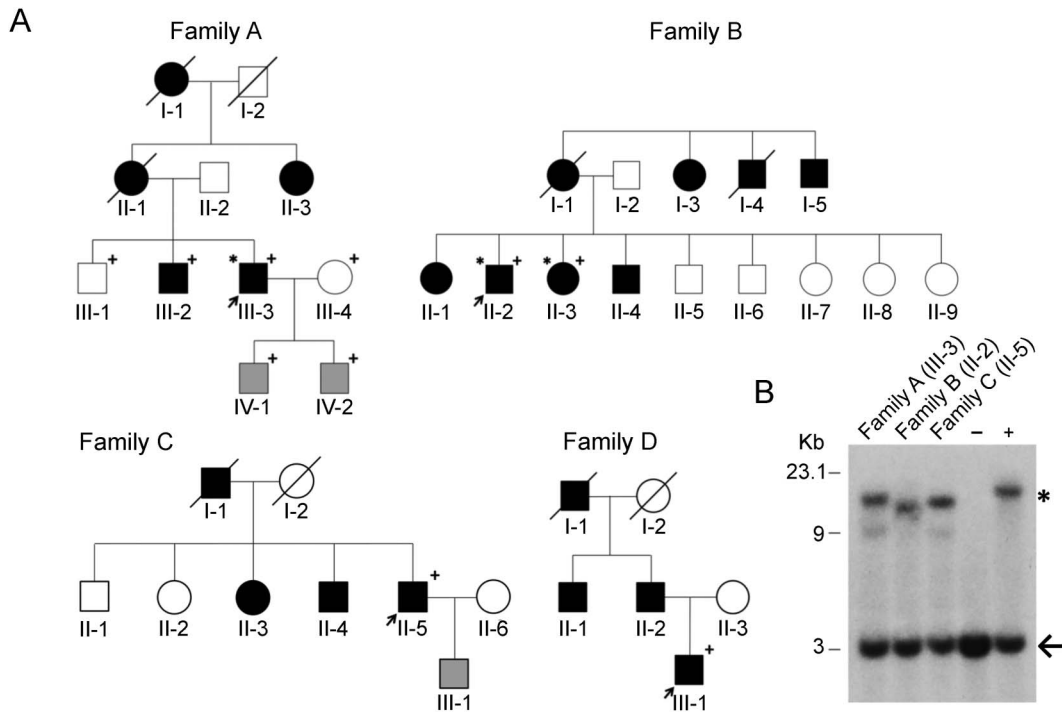
Genetic cerebellar ataxia is a clinically heterogeneous disease that progressively destroys the cerebellum and consequently impairs balance and coordination in the affected individual.[1] The diagnosis of cerebellar ataxia can be challenging because there are more than 500 genes associated with either primary or secondary ataxia, and many affected families remain undiagnosed for decades. The introduction of unbiased genomic diagnostic testing methods, such as clinical exome sequencing, into the diagnostic evaluation has greatly improved time to diagnosis,[2,3] but is limited in its ability to detect certain forms of genetic mutation such as repeat expansion disorders.[4,5] The spinocerebellar ataxias (SCAs) are a diverse group of neurodegenerative disorders characterized by an autosomal dominant pattern of inheritance and cerebellar symptoms, currently consisting of at least 43 distinct clinical entities, of which 11 are repeat expansion disorders.[6] One of these conditions, spinocerebellar ataxia 36 (SCA36), is caused by a GGCCTG hexanucleotide repeat expansion in the first intron of the pre-mRNA processing gene, nucleolar protein 56 (NOP56).[7]

Supplemental data at Neurology.org/ng

(A) Affected individuals (dark fill) and index cases (arrow) are indicated. +Genotyped individuals; *Individuals who had exome sequencing performed; gray fill indicates at-risk individuals, currently asymptomatic. All genotyped individuals in families A and B were included in the linkage analysis. (B) Southern blot analysis of patient DNA from probands of 3 of the identified SCA36 families, a negative control (−), and a positive control (+). The 3 Kb wild-type allele (arrow) and the larger expanded allele (*) are shown. SCA36 = spinocerebellar ataxia 36.

Individuals with this disorder show progressive, late-onset, ataxic symptoms affecting limb, trunk, and/or gait stability.[7] To date, SCA36 prevalence has been reported in regions of Japan,[7,8] Spain,[9] France,[8] Germany,[8] and most recently, in China,[10] but its prevalence in the United States is undetermined. Through combined methods of exome sequencing, linkage analysis, and fluorescent repeat primed PCR (RP-PCR), we molecularly identified SCA36 in 4 index cases, representing less than 1% of the undiagnosed population at a tertiary referral ataxia center.

METHODS **Patient recruitment and diagnostic evaluation.** All patients were initially seen at our tertiary referral ataxia center for disorders of gait and balance. All patients underwent a comprehensive clinical evaluation for acquired causes of ataxia prior to consideration of genetic causes.[1,11] For enrollment in this study, patients were required to have negative testing for the most common genetic ataxias worldwide,[6,12] specifically SCA1, SCA2, SCA3, SCA6, and SCA7, and Friedreich ataxia if clinically appropriate. Age at onset or mode of inheritance was not used as exclusion criteria to prevent bias against phenotypic variability[3] or inaccurate assessment of familial association in relatives who could not be directly assessed clinically.

Standard protocol approvals, registrations, and patient consents. Written informed consent was obtained to collect DNA for genetic analysis. All patients were provided genetic counseling both before and after completion of the study. All study methods were approved by the Institutional Review Board of the University of California, Los Angeles.

Exome sequencing and array genotyping. Whole-exome sequencing was performed for 3 members of 2 families (families A and B, figure 1A) at the Human Genome Sequencing Center (HGSC) at Baylor College of Medicine as part of the Baylor-Hopkins Center for Mendelian Genomics initiative. With 0.5 ng of DNA, an Illumina paired-end pre-capture library was generated as described in the BCM-HGSC protocol (hgsc.bcm.edu/content/protocols-sequencing-library-construction). Four of these libraries were pooled and then hybridized in solution to the HGSC CORE design[13] (52 Mb; NimbleGen, Madison, WI) in accordance with the NimbleGen SeqCap EZ Exome Library SR User's Guide. Each capture pool, containing 4 samples, was sequenced in one lane using the Illumina HiSeq 2000 platform in paired-end mode, with sequencing-by-synthesis reactions extended for 101 cycles from each end and an additional 7 cycles for the index read. On average, each exome yielded 9.6 Gb of data with 93% of the targeted exome bases covered to a depth of 20-fold or greater. Variant calls were generated according to the GATK best practices (Genome Analysis Toolkit 3.4) using the hg19 reference genome and annotated using the SVS Annotation Suite (Golden Helix). Only variants causing non-synonymous protein-coding changes or splice-site variants were included in the analysis. Annotated variants were filtered according to a minor allele frequency (MAF) using the Exome Aggregation Consortium (ExAC) database[14] (r.0.2) (exac.broadinstitute.org) at an MAF of 0.1% for genes containing a single heterozygous variant, and 1% for genes containing multiple heterozygous variants or homozygous variants. Variants remaining after filtering were compared to a list of 2,256 genes identified by the keyword "ataxia" in

the Online Mendelian Inheritance in Man database[15] (omim.org) and assessed for their clinical significance as previously described.[3]

In parallel to the exome workflow, Illumina Infinium Human Exome v1-2 array data were generated for linkage analysis from 6 members of family A and 2 members of family B (figure 1A). Array data quality assessment included orthogonal confirmation of sample identity and purity using the Error Rate In Sequencing (ERIS) pipeline developed at the BCM:HGSC. Using an "e-GenoTyping" approach, ERIS screens all sequence reads for exact matches to probe sequences defined by the variant and position of interest. A successfully sequenced sample must meet quality control metrics of ERIS single nucleotide polymorphism (SNP) array concordance (>90%) and ERIS average contamination rate (<5%). These arrays capture more than 240,000 SNPs, which were carried forward into quality filtering, pedigree validation, and linkage analyses. All insertions/deletions and nonautosomal polymorphisms and all variants without an rs identifier were removed. In addition, SNPs were removed for high missingness (>5%) or mapping to identical locations (duplicates) resulting in 227,103 variants for analysis. Pedigrees were validated using the prePRIMUS QC pipeline to estimate pairwise kinship and were reconstructed in PRIMUS.[16,17]

Linkage analysis. To identify overlapping haplotypes shared identically by descent (IBD) by all genotyped cases and not shared by genotyped controls in families A and B (figure 1A), multifamily parametric linkage analysis was conducted using a fully penetrant dominant model with no phenocopies ($f_{0,1,2}$ = 0,1,1) using ALLEGRO.[18] Logarithm of the odds (LOD) scores generated from linkage in these families are unable to meet genome-wide significance thresholds because of power (max LOD given the information across the pedigrees is 1.2); the purpose of these analyses was to identify candidate regions where IBD sharing was consistent with the model of inheritance for disease in both families for further analysis rather than to identify regions of genome-wide significant linkage. Rare (MAF <0.1% in ExAC), nonsynonymous and splice-site variants from exome sequencing falling within linkage regions (peaks reaching maximum LOD, ~58.6 megabases genome wide) were identified and evaluated as potential disease candidates (figure e-1 at Neurology.org/ng).

SCA36 repeat expansion testing. DNA was isolated from the peripheral blood and purified using the Gentra Puregene Blood Kit (Qiagen, Hilden, Germany). Fluorescent RP-PCR was implemented,[9] and PCR products were separated then analyzed on an ABI Prism 3730 XL Analyzer (Retrogen, San Diego, CA). Peak Scanner 2 software was used to assess for the presence of the SCA36 hexanucleotide repeat expansion in *NOP56*. This method specifically detects expanded alleles and does not assess normal *NOP56* alleles.

Southern blot. For confirmation of repeat expansions, Southern blot analysis was conducted on DNA samples of 3 affected individuals (figure 1B). Genomic DNA (10 μg) was digested overnight with *BsmI* (New England Biolabs, Ipswich, MA). Digested DNA was resolved on a 0.8% agarose gel, depurinated with 0.25 M HCl, denatured with 1.5 M NaCl, 0.5 M NaOH neutralized with 1.5 M NaCl, 0.5 M Tris-HCl (pH 7.0), and transferred with 20× saline-sodium citrate buffer onto a Hybond-N+ nylon membrane (GE Health care, Piscataway, NJ). DNA was immobilized on the membrane by UV cross-linking. A 464-bp probe for hybridization was synthesized from genomic DNA by PCR using primers[10] flanking exons 2 and 3 of *NOP56*. The PCR-amplified probe DNA was used as template to make a 32P-labeled probe using random hexamer primers (#58875; Invitrogen, Carlsbad, CA), Klenow polymerase (#M0210S, NEB) and dNTP's containing [32P]-dCTP.

Hybridization with the labeled probe was performed at 65°C overnight. The membrane was developed by X-ray film exposure.

RESULTS Identification of 2 SCA36 families by exome sequencing and linkage analysis. We independently performed exome sequencing on 3 individuals from 2 undiagnosed families (families A and B, figure 1A) with adult-onset dominant cerebellar ataxia given the high diagnostic yield of this testing in patients previously screened for the most common genetic forms of ataxia.[2,3,19,20] However, we were unable to identify any potentially clinically relevant sequence variation associated with ataxia using standard methods.[3] To address the possibility of a novel gene or a form of genetic variation not detected by exome sequencing (e.g., noncoding, copy number variants, or repeat expansions), we separately performed genotyping using a cytogenomic SNP array and subsequent linkage analysis in both families. Among these data, we identified 2 regions on chromosome 20 that reached maximum theoretical LOD scores across families A and B containing a known clinical ataxia gene, *NOP56*, causing disease through repeat expansion (figure e-1). Exome analysis had indicated that families A and B did not possess any rare protein-coding variants in any genes potentially associated with their phenotype in any of the identified linked regions (figure e-1). The absence of any clinically relevant rare variants in ataxia genes within the sequenced affected family members (figure 1A) prompted us to evaluate these 2 families for the repeat expansion in *NOP56* associated with SCA36.[7]

Identification of an expanded GGCCTG repeat in *NOP56* in families A and B. RP-PCR analysis was used to assess for expansion in the *NOP56* gene. This method specifically detects the presence of an expansion through the generation of multiple products of varying size from different amplification sites, leading to a stuttering effect (figure e-2). RP-PCR identified the pathogenic hexanucleotide repeat expansion in intron 1 of *NOP56* in both families A and B (figure e-2), confirming the diagnosis of SCA36. Because RP-PCR lacks the ability to quantitate expansion size, Southern blot analysis was used as a second confirmatory technique, detecting *NOP56* repeat expansions in the probands from families A and B (figure 1B). Given that SCA36 is considered a rare repeat expansion disorder, but has been observed to have varying prevalence in different ethnic populations,[7–10] the finding of 2 consecutive SCA36 families prompted us to evaluate our entire clinical cohort for the disease to determine the prevalence in a US ataxic population. We analyzed 577 index cases using fluorescent RP-PCR. The demographics of this patient population including age, sex, ethnicity, and phenotype are shown in table 1. Only 2 additional

Table 1   Patient demographics

Phenotype	Percentage of cohort
Spinocerebellar ataxia	36.0
Pure cerebellar ataxia	31.0
Multiple system atrophy	17.0
Spastic ataxia	14.0
Episodic ataxia	4.0
Spastic paraplegia	3.0
Leukodystrophy with ataxia	2.0
Other neurologic phenotype	7.0
Ethnicity	
White, non-Hispanic	76.7
White, Hispanic, or Latino	11.1
Asian	9.1
Chinese	1.2
Indian	1.0
Korean	1.0
Japanese	0.8
Vietnamese	0.6
Taiwanese	0.4
Indonesian	0.2
Unspecified	3.9
Black	3.6
Sex	
Male	48.6
Female	51.4
Average age	54.5 ± 17.1 y

positive cases were identified, thus bringing the total to 4/577 cases (0.7%).

Clinical characteristics of SCA36 in US families. SCA36 was observed in 4 index cases from a cohort of 577 patients (families A–D; figure 1A). An additional 2 positive cases were subsequently identified (1 each from families A and B, respectively). The ancestral

origins of these 4 families were either white, non-Hispanic, European (2 families), or Asian (1 Vietnamese and 1 Japanese family). The clinical features of these 6 individuals (5 men and 1 woman, mean age at onset: 44.5 ± 5.8 years) are summarized in table 2. All affected individuals exhibited an ataxia gait, consistent with cerebellar dysfunction, on examination. Other symptoms included limb ataxia or spasticity and altered reflexes. Saccadic ocular pursuit was present in 5 of 6 individuals, predominantly saccadic pursuit or overshoot, whereas hearing impairment was reported in only 1 of 6 individuals. Two of 6 patients presented with mild cerebellar atrophy detected using neuroimaging.

DISCUSSION We evaluated 577 index cases seen at a tertiary referral ataxia center in the United States and identified the *NOP56* hexanucleotide repeat expansion in 4 families. When compared with other reports, the prevalence of SCA36 in the United States (n = 577, 0.7%) appears to be less frequent than in most other countries evaluated to date, including Spain (n = 160, 6.3%),[9] France (n = 270, 4.4%),[8] Western Japan (n = 251, 3.6%),[7] and Eastern Japan (n = 231, 2.2%)[8] and is more similar to that of China (n = 512, 0.6%)[10] and Germany (n = 175, 0.0%)[8] (table 3). It should be noted that our cohort was of general undiagnosed patients with ataxia, unselected by mode of inheritance or other features, in whom the most frequent genetic etiologies had been previously excluded. As a tertiary referral center, we cannot estimate the number of patients with more common genetic etiologies identified in the community, and therefore not referred, so the overall prevalence in all undiagnosed ataxia cases is likely lower. Consistent with previous reports, SCA36-positive individuals in our cohort descended from European and East Asian backgrounds. Despite the majority of our cohort descending from European or Asian ancestry (table 1), we observe a much lower prevalence than countries such as Spain, France, or Japan.[7–9] This may

Table 2   Clinical findings of SCA36 patients reported in this study

Family	Ethnicity	ID	Sex	Age, y	Age at onset, y	Gait ataxia	Limb ataxia	Saccadic ocular pursuit	Dysarthria	Fasciculation	Reflexes in lower limbs	Hearing impairment	Cerebellar atrophy
A	White, non-Hispanic (English and German)	III-1	M	60	45	+	+	+	−	Face	↑	−	−
		III-4	M	61	50	+	+	+	+	Face, tongue, right bicep	↑	+	−
B	Asian (Vietnamese)	II-2	M	59	46	+	−	+	+	Absent	↑	−	+
		II-3	F	55	41	+	+	−	+	Absent	Normal	−	−
C	Asian (Japanese)	II-5	M	73	50	+	−	+	+	Absent	↑	−	−
D	White, non-Hispanic (European)	III-1	M	42	35	−	−	+	−	Absent	Normal	−	+

+ = present; − = absent; ↑ = increased; SCA36 = spinocerebellar ataxia 36.

Table 3 Comparison of clinical characteristics of US SCA36 patients with other SCA36 populations globally

	Overall US[a]	Han Chinese[10]	Western Japan[7]	Spain[9]	France[8]	Eastern Japan[8]
Sex	5 M/1 F	3 M/2 F	8 M/6 F	20 M/24 F	8 M/12 F	4 M/4 F
Age at onset, y	44.5 ± 5.8	44.8 ± 3.8	53.1 ± 3.5	52.8 ± 7.4	50.0 ± 6.9	52.3 ± 8.6
Age at examination, y	59.3 ± 10.0	55.8 ± 4.5	67.1 ± 8.5	63.8 ± 12.5	61.3 ± 10.0	61.0 ± 8.3
Gait ataxia	100.0%	100.0%	100.0%	98.0%	100.0%	88.0%
Limb ataxia	50.0%	100.0%	93.0%	88.0%	NA	NA
Saccadic ocular pursuit	66.7%	100.0%	93.0%	88.0%	39.0%	63.0%
Dysarthria	66.7%	100.0%	100.0%	64.0%	61.0%	100.0%
Tongue fasciculation	16.7%	20%	71.0%	61.0%	12.0%	63.0%
Muscle fasciculation	33.3%	80.0%	64.0%	NA	NA	NA
Hyperreflexia	66.7%	100.0%	79.0%	41.0%	67.0%	63.0%
Hearing impairment	16.7%	80.0%	NA	74.0%	44.0%	88.0%

Abbreviations: NA = not available; SCA36 = spinocerebellar ataxia 36.
[a] Current study.

stem partially from the inclusion criteria defining our cohort, but likely also reflects regional variation across Europe and Asia, as noted in the studies from China and Germany,[8,10] for example. One of our families (family B) identifies their family origins as emanating from Vietnam, which would be the first reported case of SCA36 from that country.

Clinical examination of those 6 affected SCA36 patients in our cohort confirms many of the previously reported SCA36 symptoms, including gait and limb ataxia, dysarthria, increased lower limb reflexes, polyneuropathy, saccadic pursuit, and cognitive impairment (table 3). Facial or tongue fasciculations are typically seen in more than 50% of reported SCA36-positive cases with the exception of France (12%) and were present in 2 of 6 (33%) SCA36-positive individuals in this study. Hearing impairment is present in more than 40% of previously reported SCA36-positive individuals and was also reported in 2 of 6 individuals from family B (figure 1A and table 2), although detailed audiologic studies were not conducted on all the affected patients in this study. In summary, SCA36 represents a small percentage of the undiagnosed cerebellar ataxia population in the United States. Given that SCA36 is a repeat expansion disorder that would not be detected by the high-yield genomic diagnostic tests performed early in a patient's evaluation,[3] repeat expansion testing should be considered in individuals with appropriate clinical phenotypes and either a sporadic or dominant mode of inheritance.

AUTHOR CONTRIBUTIONS

ACKNOWLEDGMENT

STUDY FUNDING

DISCLOSURE

REFERENCES

1. Shakkottai VG, Fogel BL. Autosomal dominant spinocerebellar ataxia. Neurol Clin 2013;31:987–1007.

2. Pyle A, Smertenko T, Bargiela D, et al. Exome sequencing in undiagnosed inherited and sporadic ataxias. Brain 2015; 138:276–283.

3. Fogel BL, Lee H, Deignan JL, et al. Exome sequencing in the clinical diagnosis of sporadic or familial cerebellar ataxia. JAMA Neurol 2014;71:1237–1246.

4. Fogel BL, Lee H, Strom SP, et al. Clinical exome sequencing in neurogenetic and neuropsychiatric disorders. Ann NY Acad Sci 2016;1366:49–60.

5. Fogel BL, Satya-Murti S, Cohen BH. Clinical exome sequencing in neurologic disease. Neurol Clin Pract 2016;6:164–176.

6. Durr A. Autosomal dominant cerebellar ataxias: polyglutamine expansions and beyond. Lancet Neurol 2010;9: 885–894.

7. Kobayashi H, Abe K, Matsuura T, et al. Expansion of intronic GGCCTG hexanucleotide repeat in NOP56 causes SCA36, a type of spinocerebellar ataxia accompanied by motor neuron involvement. Am J Hum Genet 2011;89:121–130.

8. Obayashi M, Stevanin G, Synofzik M, et al. Spinocerebellar ataxia type 36 exists in diverse populations and can be caused by a short hexanucleotide GGCCTG repeat expansion. J Neurol Neurosurg Psychiatry 2015;86:986–995.

9. García-Murias M, Quintáns B, Arias M, et al. "Costa da Morte" ataxia is spinocerebellar ataxia 36: clinical and genetic characterization. Brain 2012;135:1423–1435.

10. Lee YC, Tsai PC, Guo YC, et al. Spinocerebellar ataxia type 36 in the Han Chinese. Neurol Genet 2016;2:e68. doi: 10.1212/NXG.0000000000000068.

11. Fogel B, Perlman S. Cerebellar disorders: balancing the approach to cerebellar ataxia. In: Tuite PJ, Gálvez-Jiménez N, editors. Uncommon Causes of Movement Disorders. New York: Cambridge University Press; 2011:198–216.

12. Fogel BL, Lee JY, Lane J, et al. Mutations in rare ataxia genes are uncommon causes of sporadic cerebellar ataxia. Mov Disord 2012;27:442–446.

13. Bainbridge MN, Wang M, Yuanqing W, et al. Targeted enrichment beyond the consensus coding DNA sequence exome reveals exons with higher variant densities. Genome Biol 2011;12:R68.

14. Exome Aggregation Consortium (ExAC). ExAC Browser. Cambridge, MA: 2014. Available at: http://biorxiv.org/content/early/2015/10/30/030338. Accessed December 2015.

15. McKusick-nathans Institute of Genetic Medicine; Johns Hopkins Medicine; National Human Genome Research Institute. Online Mendelian Inheritance in Man. Avaiable at: http://omim.org/. Accessed December 2015.

16. Staples J, Qiao D, Cho MH, et al. PRIMUS: rapid reconstruction of pedigrees from genome-wide estimates of identity by descent. Am J Hum Genet 2014;95:553–564.

17. Staples J, Ekunwe L, Lange E, et al. PRIMUS: improving pedigree reconstruction using mitochondrial and Y haplotypes. Bioinformatics 2016;32:596–598.

18. Gudbjartsson DF, Jonasson K, Frigge ML, et al. Allegro, a new computer program for multipoint linkage analysis. Nat Genet 2000;25:12–13.

19. Ohba C, Osaka H, Iai M, et al. Diagnostic utility of whole exome sequencing in patients showing cerebellar and/or vermis atrophy in childhood. Neurogenetics 2013;14:225–232.

20. Sawyer SL, Schwartzentruber J, Beaulieu CL, et al; FORGE Canada Consortium. Exome sequencing as a diagnostic tool for pediatric-onset ataxia. Hum Mutat 2014; 35:45–49.