

Supplement 2: Next-generation sequencing strategies for SYNE1 mutation screening

A) Targeted sequencing by HaloPlex gene panel.

N=179 patients were screened for SYNE1 mutations as part of a high coverage (>94% mean coverage above 20 X / base) HaloPlex gene panel kit (Agilent, Santa Clara, CA, USA) which included >120 known ataxia genes. Library preparation started with fragmentation of DNA by digestion with multiple restriction enzymes. A set of gene target-specific HaloPlex selector probes was hybridized to the fragmented DNA, thereby inducing the target fragments to form circular DNA molecules that also contain method-specific sequence motifs (PBS) as well as a biotin label. Probe/Target DNA molecules were captured and purified using magnetic streptavidin beads. Probe/Target circles were then ligated to ensure that only perfectly hybridized fragments were circularized. A PCR amplification step provided an enriched library ready for sequencing. Molecular barcodes (MID) were used for precise sample tracking. Finally, the enriched library was checked for quality and quantity prior to sequencing on an Illumina MiSeq sequencing device by means of a 2 X 150 bp paired end approach (MiSeq sequencing kit V2). Reads were mapped (BWA-mem) against the hg19 standard reference genome to detect SNPs, SNV, short deletions and insertions (SAMtools, Annovar).

B) Targeted panel sequencing by Agilent Sureselect kit.

N=139 patients were screened for SYNE1 mutations as part of a targeted exon-capture strategy (Agilent Sureselect kit) coupled with multiplexing and high-throughput sequencing of 57 genes causing ataxia when mutated. The Agilent system consists in a capture of the genomic regions harboring the exons, based on the hybridization of complementary designed biotinylated cRNA oligonucleotides. (with a capture size of 285kb). Reads were mapped against the hg19 standard reference genome to detect SNPs, SNV, short deletions and insertions (SAMtools, Annovar).

C) Targeted sequencing by Illumina customized high-coverage panel

N=21 patients were screened as part of an amplicon-based customized panel (Illumina TrueSeq Custom Amplicon, TSCA) (Illumina Inc., San Diego, CA, USA) run on an Illumina MiSeq apparatus covering 76 known ataxia genes (n=21). *SYNE1* coverage was $\geq 20X$ for $93.5\% \pm 2.8$ SD of the coding sequence. Reads were mapped against the hg19 standard reference genome to detect SNPs, SNVs, short deletions and insertions (software used: MiSeq Reporter and VariantStudio [Illumina Inc.], CLC Genomics Workbench [Qiagen, Venlo, Netherlands]).

D) Targeted sequencing by Illumina Nextera Rapid Capture Custom kit

N=54 patients were screened as part of a targeted exon-capture resequencing strategy (Illumina Nextera Rapid Capture Custom kit (Illumina Inc., San Diego, CA, USA) which included 107 known ataxia genes. *SYNE1* coverage was $\geq 20X$ for $98.2\% \pm 0.8$ SD of the coding sequence. Reads were mapped against the hg19 standard reference genome to detect SNPs, SNVs, short deletions and insertions (software used: MiSeq Reporter and VariantStudio [Illumina Inc.], CLC Genomics Workbench [Qiagen, Venlo, Netherlands]).

E) *Whole exome sequencing.*

N= 41 patients were screened for *SYNE1* mutations as part of a whole exome-sequencing study. WES was performed using a SureSelect Human All Exon 50Mb kit (Agilent, Santa Clara, CA, USA) for in-solution enrichment and a HiSeq2000 instrument for exome sequencing (Illumina, San Diego, CA, USA). Paired-end reads of 100 bp length were produced. BWA and GATK software packages (Li and Durbin, 2009, McKenna *et al.*, 2010, DePristo *et al.*, 2011) were used to align sequence reads to the reference and call variant positions, respectively. All data were then annotated and imported into GEnomes Management Application (GEM.app), a web-based tool for next generation sequencing data analysis (Gonzalez *et al.*, 2013) (genomics.med.miami.edu). Using the GEM.app analysis module 'Genes Across Families', we then filtered for non-synonymous homozygous or compound heterozygous truncating variants in

SYNE1 (frame-shift, insertions, deletions, and stop mutations), with low frequency in public databases (minor allele frequency in dbSNP137 and NHLBI ESP6500 < 0.5%), and moderate genotype quality (GATK quality index > 30 and genotype quality GQ > 30).