SUPPLEMENTARY MATERIAL

METHODS

NGS panel analysis of known HMN/CMT2 genes

Genetic screening of known disease genes associated with peripheral hereditary neuropathies was performed using a custom NGS probe-based gene panel (Nextera Rapid Capture Custom kit , Illumina) which included 94 disease genes (*AARS, ABHD12, ARHGEF10, ATL1, ATP7A, BSCL2, CCT5, CTDP1, DCTN1, DHTKD1, DNAJB2, DNM2, DNMT1, DST, DYNC1H1, EGR2, FAM134B, FBLN5, FBXO38, FGD4, FIG4, FLVCR1, GALC, GAN, GARS, GDAP1, GJB1, GJB3, GLA, GNB4, HARS, HINT1, HK1, HSPB1, HSPB3, HSPB8, IFRD1, IGHMBP2, IKBKAP, INF2, KARS, KIF1A, KIF1B, KIF5A, KLHL9, LITAF, LMNA, LRSAM1, MARS, MED25, MFN2, MPZ, MTMR2, MYH14, NDRG1, NEFL, NGF, NTRK1, OPA1, PDK3, PEX1, PEX7, PHYH, PLA2G6, PLEKHG5, PLP1, PMM2, PMP22, POLG, PRPS1, PRX, RAB7A, REEP1, RNF170, SBF1, SBF2, SCN11A, SCN9A, SEPT9, SETX, SH3TC2, SLC12A6, SLC5A7, SOX10, SPTLC1, SPTLC2, SURF1, TDP1, TFG, TRIM2, TRPV4, TTR, WNK1, YARS*). Libraries were sequenced with an Illumina MiSeq system (Illumina). Data were analyzed using MiSeq Reporter (Illumina), VariantStudio (Illumina) and CLC Genomics Workbench (CLCbio-Qiagen) softwares. Identified variants were filtered according to quality criteria, functional consequence and frequency in the genetic databases (dbSNP, 1000Genome, ESP, ExAC, HGMD). Candidate variants were confirmed by Sanger sequencing.

Whole exome sequencing (WES)

A total of 50 ng of each gDNA sample, based on Qubit (Thermo Fisher) quantification processed using Agilent SureSelect-QXT (Agilent Technologies), was used for library preparation. Sequencing was performed on High output flow cell, for 2x151 cycles, based on SBS (Sequencing by Synthesis) v2 chemistry. For each run, twelve samples (all

Previtali SC, *et al. J Neurol Neurosurg Psychiatry* 2019; 0:1171–1179. doi: 10.1136/jnnp-2019-320717

members for each family were run together on the same flow cell) were loaded in order to obtain ~30M of clusters, enough to obtain 20X coverage for at least 90% of the target regions.

The raw sequencing reads were mapped to the reference human genome (hg19) using the proprietary NovoAlign software (Novocraft Technologies, Selangor, Malaysia). Duplicate reads were marked and excluded from further analysis using Picard Tools (https://broadinstitute.github.io/picard/). The Genome Analysis Toolkit (GATK) HaplotypeCaller was used to call variant.[1] Low quality variants were removed from further analysis based on thresholds recommended by GATK.

Measures of deleteriousness of each variant: GERP++ calculates the difference between the observed and expected number of substitutions at a genomic location based on comparative sequence analysis with 33 other mammalian species.[2] GERP++ scores range from -12.3 to 6.17, with higher scores indicating higher evolutionary constraint. A score greater than 2 can be considered constrained. CADD uses machine learning to detect variants that occur less frequently than expected by comparing actual and simulated variants at sites where the human genome differs from the predicted human-chimpanzee ancestral genome.[3] Scaled CADD scores range from 1 to 99, where the top 10%, 1%, 0.1%, etc. most deleterious variants have scores of 10, 20, 30, etc. respectively.

The variants were annotated using SnpEff to investigate functional impact and Gemini to annotate each variant for its frequency in the general population and specifically in populations of European ancestry (Genome Aggregation Database).[4 5] A soft threshold of MAF (Minor Allele Frequency) <0.01 was set, with the exception of compound heterozygous inheritance patterns, in which case the threshold was more lenient (MAF<0.10). Gemini was also used to retrieve several measures of the deleteriousness of each variant, including GERP++ (Genome Evolutionary Rate Profiling, version 2), and CADD (Combined Annotation-Dependent Depletion) (Supplementary material).

Previtali SC, *et al. J Neurol Neurosurg Psychiatry* 2019; 0:1171–1179. doi: 10.1136/jnnp-2019-320717

Selected variants were confirmed by Sanger sequencing. Candidate variants were defined on the basis of ACMG guidelines[6] using the following criteria: allele frequencies AF=0 (novel variants) or AF (allele frequency) <0.001 (novel disease associations of very rare variants); CADD scores between 10 and 30 (most deleterious changes); Gerp++ scores >2;, the prediction of the potential effect on gene splicing; the presence of a pathogenetic variant on the other allele for compound heterozygous variants, and functional studies supportive of a pathogenetic effect. In Table 1: class 5 refers to pathogenic variants; class 4 to likely pathogenic, and class 3 refers to variants of uncertain significance.

Human Splicing Finder (http://www.umd.be/HSF/, Marseille, France), Splice Site Prediction by Neural Network Site (http://www.fruitfly.org/seq_tools/splice.html, Berkley, CA, USA), MaxEntScan (http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html) and Gene Splicer (http://www.cbcb.umd.edu/software/GeneSplicer/gene_spl.shtml) were used to evaluate the potential effects on gene splicing.

Custom NGS panel analysis

A targeted Haloplex enrichment system (Agilent Technologies, Santa Clara, USA) has been used for the preparation of the NGS-DNA libraries from each genomic DNA. The following panel of genes has been defined with capture system and designed with the Sure Design tool (Agilent Technologies): *AGRN, ARHGEF28, GNE, KBTBD13, PNKP, SIGMAR1,* and *VRK1.* Sequencing was performed on a MiSeq platform (Illumina Inc) and using the MiSeq V3 Sequencing kit (2x150 cycles). The mean coverage obtained per sample was 508X, with 99% of the target region sequenced ≥20X. Sequence data analysis and calling of variants was performed with the Sure Call software (Agilent Technologies).

Previtali SC, *et al. J Neurol Neurosurg Psychiatry* 2019; 0:1171–1179. doi: 10.1136/jnnp-2019-320717

Constructs, cell culture and transfection

A C-terminal Flag-tagged human *KBTBD13* full length cDNA was cloned into the pCMV6 vector (kindly provided by Dr. Sambuughin, Bethesda, USA). The p.Ala55Gly mutation was inserted into the *KBTBD13* cDNA sequence using the Site directed Mutagenesis kit (Agilent Technologies) following manufacturer's conditions. The presence of the c.164C>G mutation was assessed by Sanger sequencing on selected clones. Either wild type or mutated *KBTBD13* constructs were transfected into 293T cells, which endogenously express Cullin 3, using calcium phosphate. 293T cells were grown in Dulbeccos's modified eagle medium (DMEM) containing 10% fetal bovine serum and supplemented with sodium pyruvate and non-essential amino acids.

Immunoprecipitation and immunoblot analysis

Immunoprecipitation of wild type and mutated KBTBD13 was performed in NP40 1% lysis buffer using a monoclonal anti-Flag antibody (M2 clone, SIGMA) and Dynabeads Protein G (Life Technologies). Cullin 3 was detected in the immunoprecipitated sample using a polyclonal anti-Cullin 3 antibody (Novusbio). Dynabeads Protein G and mouse IgG were used as a negative control for immunoprecipitation, incubated with cell lysates containing either overexpressed wild type KBTBD13 or mutated KBTBD13. Western blot analysis was performed as already reported.[7]

SUPPLEMENTARY REFERENCES

1. Van der Auwera GA, Carneiro MO, Hartl C, et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics* 2013;43:11 10 1-33. doi: 10.1002/0471250953.bi1110s43 [published Online First: 2014/11/29]

2. Davydov EV, Goode DL, Sirota M, et al. Identifying a high fraction of the human genome to be under selective constraint using GERP++. *PLoS Comput Biol* 2010;6(12):e1001025. doi: 10.1371/journal.pcbi.1001025 [published Online First: 2010/12/15]

3. Kircher M, Witten DM, Jain P, et al. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet* 2014;46(3):310-5. doi: 10.1038/ng.2892 [published Online First: 2014/02/04]

4. Cingolani P, Platts A, Wang le L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly (Austin)* 2012;6(2):80-92. doi: 10.4161/fly.19695 [published Online First: 2012/06/26]

5. Paila U, Chapman BA, Kirchner R, et al. GEMINI: integrative exploration of genetic variation and genome annotations. *PLoS Comput Biol* 2013;9(7):e1003153. doi: 10.1371/journal.pcbi.1003153 [published Online First: 2013/07/23]

6. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* 2015;17(5):405-24. doi: 10.1038/gim.2015.30 [published Online First: 2015/03/06]

7. Noseda R, Guerrero-Valero M, Alberizzi V, et al. Kif13b Regulates PNS and CNS Myelination through the Dlg1 Scaffold. *PLoS Biol* 2016;14(4):e1002440. doi: 10.1371/journal.pbio.1002440 [published Online First: 2016/04/14]