Supplementary Data

Supplemental File 1

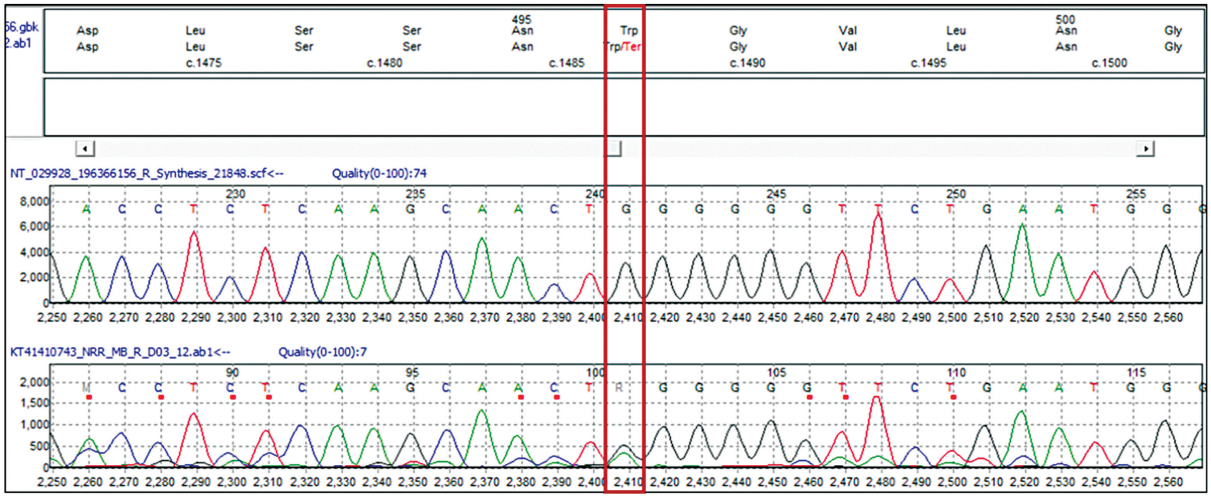
Describes the methodology for genetic testing and shows the electropherograms of the Sanger sequencing of the parents showing the variation in *NRROS* gene.

Methodology

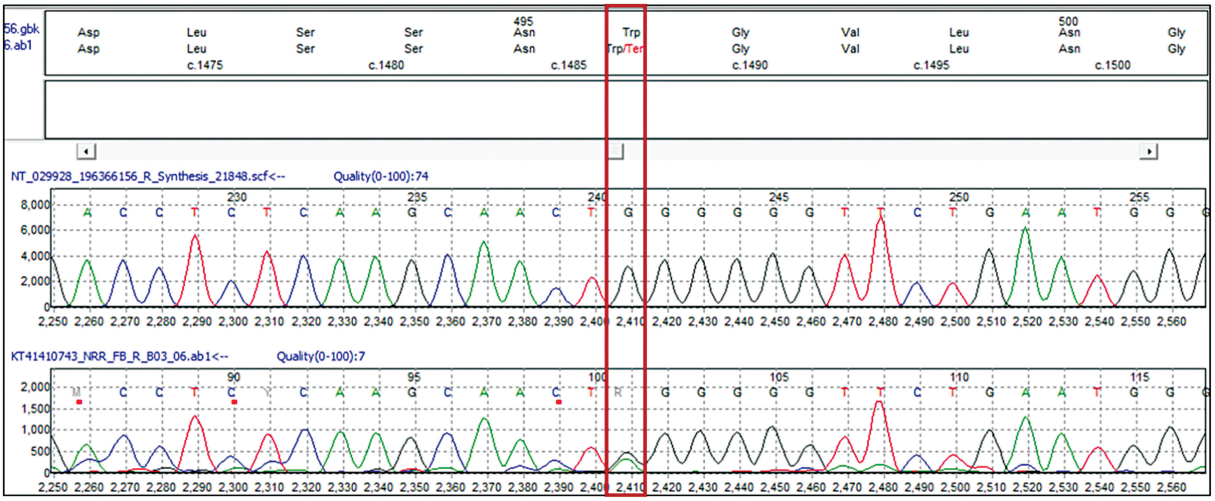
Sequencing of the protein coding regions of approximately 30 Mb of the human exome (targeting approximately 99% of regions in CCDS and RefSeq) was performed using Illumina next-generation sequencing systems at a mean depth of 80–100× and % of bases covered at 20× depth >90% in the target region. In some cases, due to the complexity of the sequence, not all variants in the flanking regions were able to be analyzed. A base was considered to have sufficient coverage at 20× and an exon was considered fully covered if all coding bases plus three nucleotides of flanking sequence on either side are covered at 20× or more. GATK best practice framework was followed for variant identification. BWA-mem aligner was used to align the obtain sequences to human reference genome (GRCh37/hg19). Duplicate reads identification and removal. Base quality recalibration and re-alignment of reads based on indels were done using inbuilt Sentieon modules.[1] Sention's Haplotypecaller module had been used to identify the variants that were relevant to the clinical indications. Along with this, Deep variant analysis pipeline on Google cloud platform was used as a secondary pipeline to call genetic variants.[2] Quality checks (QC) were performed on all VCF files to exclude variants where sequencing is of poor quality. Additional QC metrics includes total homozygous and heterozygous calls (SNVs and indels), proportion of variant calls that were common, number of variants falling into different annotated consequence categories, and number of extreme heterozygotes (alternate allele proportion 0.8). Variant annotations were done using published databases such as OMIM, GWAS, GNOMAD, 1000Genome, etc.[3–6] Nonsynonymous and splice site variants were used for clinical interpretation. Silent variations that do not result in any change in amino acid in the coding region were not reported. The data quality statistics are as follows: total data generated (Gb), 11; panel coverage %, 99; Q30 bases, >85%. For Sanger sequencing, targeted sequencing and mutation analysis was performed by polymerase chain reaction (PCR) followed by automated DNA sequencing of the amplicon using BigDye ABI Genetic Analyzer 3500DX platform. The raw data obtained are subsequently analyzed for the nucleotide variants.

References

1 Freed D, Aldana R, Weber JA, Edwards JS. The Sentieon Genomics Tools-A fast and accurate solution to variant calling from next-generation sequence data. BioRxiv 2017; . Doi: 10.1101/115717
2 https://cloud.google.com/life-sciences/docs/tutorials/deepvariant
3 Landrum MJ, Lee JM, Benson M, et al. ClinVar: public archive of interpretations of clinically relevant variants. Nucleic Acids Res 2016;44(D1):D862–D868
4 Welter D, MacArthur J, Morales J, et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. Nucleic Acids Res 2014;42(Database issue):D1001–D1006
5 Auton A, Brooks LD, Durbin RM, et al; 1000 Genomes Project Consortium. A global reference for human genetic variation. Nature 2015;526(7571):68–74
6 https://gnomad.broadinstitute.org/

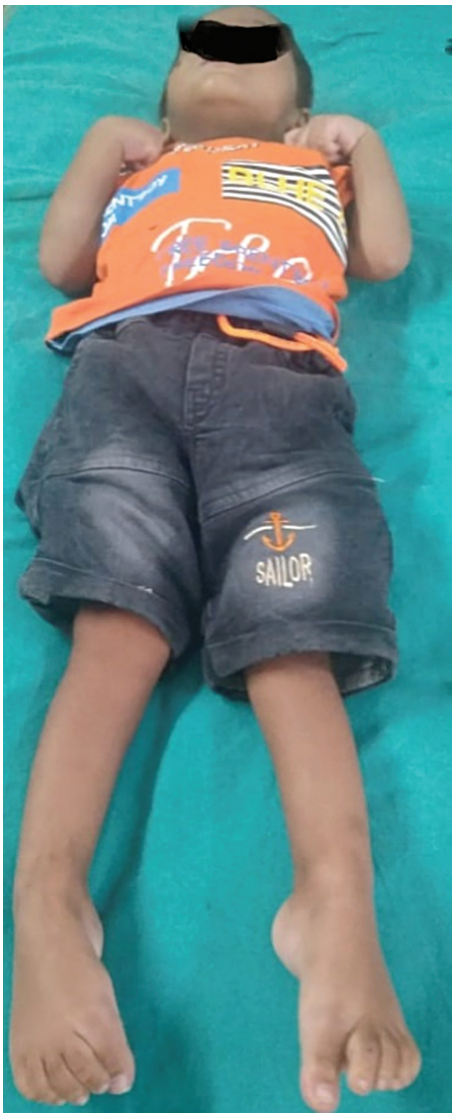


1) Sanger sequencing data (electropherogram) for the mother's sample showing nucleotide change at chr3: c.1487G > A, (p.Trp496Ter) in *NRROS* gene.



2) Sanger sequencing data (electropherogram) for the father's sample showing nucleotide change at chr3: c.1487G > A, (p.Trp496Ter) in *NRROS* gene.

Clinical picture 1 showing the proximal hypotonia in the bilateral hip girdle muscles with frog-legged posturing.



Clinical picture 2 showing the decorticate posturing of the child.

