SUPPLEMENTARY MATERIAL 1

Clinical exome sequencing and statistical analysis

Clinical exome sequencing

Clinical exome sequencing (CES) was performed as part of evaluation in a commercial genetic lab using a standard pipeline for CES analysis. Customized exome sequencing of all clinically relevant genes was performed to identify relevant variations. A custom capture kit for targeted gene capture was used to assemble DNA library from the patients' genomic DNA. The libraries were sequenced to a mean >80−100× coverage on Illumina sequencing platform. The sequences obtained were aligned to human reference genome (GRCh37/hg19) and analyzed to identify variants. The best practices framework of Genome Analysis Tool Kit (Broad Institute, Cambridge, MA, USA) was followed for identification of variants in the sample. Gene annotation of the variants was performed against Ensembl human gene model (release 9−11) (EMBL-EBI, Cambridge, UK). Clinically relevant mutations were annotated using published variants in literature and a set of publicly available diseases databases (ClinVar, Online Mendelian Inheritance in Man, Genome Wide Association Study, Human Gene Mutation Database, and SwisVar). Common variants were filtered based on allele frequency in publicly available database of normal human variation (1000 Genome, Exome Aggregation Consortium, The Genome Aggregation Database, EVS, dbSNP, 1000 Japanese Genome) and internal Indian population database. Nonsynonymous variants effect was calculated using multiple algorithms such as PolyPhen-2, Scale-Invariant Feature Transform, MutationTaster2, and likelihood ratio test. Synonymous variants found in the CES panel consisting of customized gene panel were not used for clinical interpretation. The clinical effects of identified variants were classified based on the American College of Medical Genetics and Genomics (ACMG) standards and guidelines. A positive genetic diagnosis was considered if the variant was classified as likely pathogenic or pathogenic according to ACMG guidelines.[S1] A variant was classified as novel if they are not published in literature and in publicly available disease database so far.

Statistical analysis

SPSS Statistics (Version 25, IBM Corp., Armonk, NY, USA) was used for statistical analysis. Descriptive statistical analysis (mean ± standard deviation for continuous variables and frequency and percentages for categorical variables) was performed for the demographic and clinical features. Appropriate statistical test (Fisher's exact for categorical variables, Student *t*-test for continuous variables of normal distribution and Mann-Whitney U test for continuous variables of non-normal distribution with Bonferroni correction) was applied for group comparison between individuals with positive genetic diagnosis and those with a negative genetic diagnosis. A threshold of p-values < 0.05 were considered as statistically significant.

REFERENCES

S1. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, 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;17:405-424.