Festivar: An In-House Developed Interpretation Tool for Whole-Exome Sequencing (WES) Data, is Fast, Efficient and Cost-Effective for Diagnosis of Rare Disorders.

Radhakrishna Bettadapura, Ashraf U Mannan, Forona B, Ketaki Karmalkar, Nisha K, Aditi Mhaske, Somanna A N, Laxmi Girigiri, Jitesh Adwani, Rakshita Dave, Swetha N S N, Shanmukh Katragadda, Sameer Phalke, Vamsi Veeramachaneni, Anand Janakiraman, Ramesh Hariharan **Strand Life Sciences, Bangalore, India**

VISIT OUR **WEBSITE**

Abstract

Background: An accurate diagnosis of rare disorders is essential to design appropriate treatment and management strategies. However, establishing a diagnosis for many of these rare disorders is a complex, lengthy and expensive process, which starts with recognition of specific phenotypic features and may involve multiple tests followed by consultation with multiple medical specialists.

Methods: We sequenced an Indian cohort with suspected rare disorders, using whole exome panel. Genetic variations were identified using the Strand NGS software and interpretation was done by using in-house tool, FestiVAR (Fast estimation of variants for automated reporting). FestiVAR prioritizes the variant based on HPO terms assigned to a subjects based on the clinical indication. The genes and variants are ranked according to their HPO matches and the variant label as per ACMG (The American College of Medical Genetics and Genomics) guidelines, which take into account the predicted impact of the variant on the gene/protein function, mode of inheritance/zygosity, presence in the database (ClinVar, gnomAD, dbSNP etc.) and literature. The short-listed are reviewed and relevant variant/s are selected for reporting in the StrandOmics platform.

Results: The diagnostic yield in our cohort was 38% (pathogenic and likely pathogenic variants) and in 22% of subjects, we detected VUS (variant of uncertain significance). We detected all types of variants, such as 21% indel (small deletion, duplication, insertion, or insertion/deletion), 30% missense, 22% nonsense, 12% splice site and 14% copy number variation (CNV).

Conclusion: Our study showed that FestiVAR tool is fast and efficient, which can minimize the time required to perform interpretation and it can be very cost-effective in identifying causative genes/variants in complex rare disorder subjects.

Material and Methods

Subjects and Sample Collection: The study included 7,400 unrelated subjects with suspected rare genetic disorders, primarily neurological. Genomic DNA (gDNA) was extracted from blood, saliva, or tissue using magnetic bead or column-based methods and used for library preparation.

Sequencing & Analysis: WES was performed on the Illumina NovaSeq platform (average coverage >70x) using xGen™ DNA Library Prep EZ Kits (IDT). FASTQ files were generated via bcl2fastq and processed using the DRAGEN Enrichment pipeline (v4.3.16) and Strand NGS. The hg19-hs37d5 genome build was used for analysis, which involved mapping/alignment, sorting, duplicate marking, and variant calling. Variants called by the DRAGEN and Strand NGS (strand-ngs.com) pipelines were combined and uploaded to StrandOmics for variant annotation and prioritization.

Variant Interpretation: Variants were prioritized using our in-house tool FestiVAR (v0.98.9), which ranks variants based on HPO terms, inheritance mode, zygosity, functional impact, population frequency (e.g., gnomAD), and clinical databases (e.g., ClinVar). Variants prioritized in FestiVAR are selected for reporting in StrandOmics (v6.38.0), which is a clinical genomics interpretation and reporting platform.

Copy Number Variant (CNV) Analysis: CNVs were identified through consensus calls from Strand NGS (v3.3.5) and DRAGEN, detecting single exon to large chromosomal events.

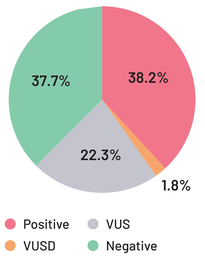
Results

- A total of **7,400 subjects** were analyzed using whole or clinical exome sequencing. The **overall diagnostic yield** (pathogenic and likely pathogenic variants) was **38%**. **Variants of Uncertain Significance (VUS)** were reported in **22%**, while **1.8%** were classified as **VUS with probable damaging effect (VUSD)** using an in-house scoring system.
- **Neurological disorders** made up the majority of subjects **(51.7%),** followed by **immunodeficiency (8%)**, **metabolic (6.9%)**, **nephrological (6.1%)**, **skeletal (4.9%)**, and **cardiac (4.8%)** disorders. **Eye (1.4%)** and **dermatological (1.2%)** disorders were the least represented.
- Despite lower numbers, **eye disorders** had the **highest diagnostic yield (52.21%)**, likely due to well characterized phenotypes and monogenic causes. In contrast, **neurological disorders** had a lower yield **(39.88%)** but the highest **VUS rate (27.8%)**, reflecting their genetic complexity.
- In our cohort, we detected all types of pathogenic/likely pathogenic variants. The mutation spectrum showed 30.5% missense, 22.2% nonsense, 20.8% frameshift, 11.7% splice-site, and 1.2% of various other mutation types such as- inframe insertion/deletion, synonymous, stop loss and start loss.
- The remaining 13.6% were copy number variants (CNV), out of which, 59.5% were multi-genic or band level events and 40.5% were intragenic events.

Table 1: Mutation detection rate (in %) across different disorder groups in a cohort of 7,400 cases analyzed by WES in this study

DisorderGroup	No.of cases	P/LP	VUSD	VUS	Negative
Neurological disorders	3744	39.88	1.68	27.38	31.06
Inborn errors of metabolism	518	39.77	1.93	13.71	44.59
Skeletal disorders	360	44.44	3.89	18.06	33.61
Dermatologic disorders	89	43.82	3.37	19.10	33.71
Nephrological disorders	463	31.10	2.16	16.85	49.89
Cardiac disorders	350	35.43	0.86	20.00	43.71
Eye disorders	113	52.21	4.42	22.12	21.24
Immunodeficiency and related disorders	600	41.17	2.17	17.50	39.17
Others	1163	30.78	1.12	16.42	51.68

Mutation Detection Rate



Mutation Spectrum

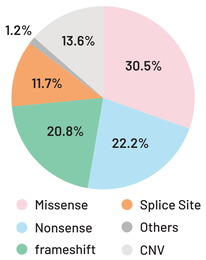
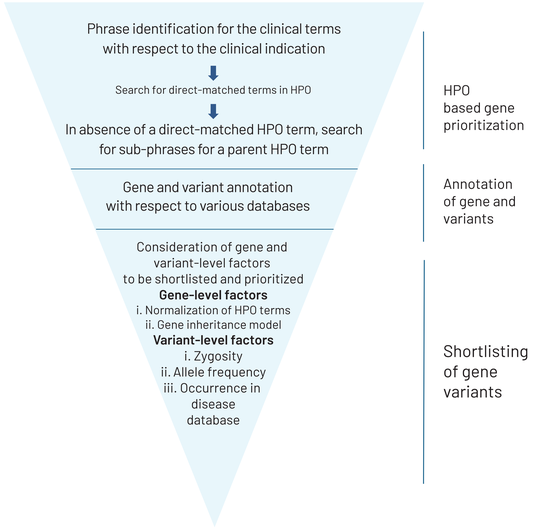


Fig. 1 Mutation detection rate and mutation spectrum in rare genetic disorders cohort. The pie chart shows mutation detection rate of 40% [which includes 'pathogenic' (P), 'likely pathogenic' (LP) variants and 'variants of uncertain significance with probable damaging effect' (VUSD)]. VUS was identified in 22.3% of cases, while no mutation was detected in 37.7% of cases. The mutation types detected in our cohort were missense, frameshift, nonsense, splices site, CNVs and others which include inframe deletions/duplication and synonymous.



Contact

Radhakrishna (RK) Bettadapura
Vice President, Business Development
+1 (415) 917-9605 rk@strandls.com

VISIT **LINKEDIN**

Conclusion

- Among 7,400 rare disease subjects, the overall diagnostic yield was 38%, (which was inclusive of both whole exome and clinical exome subjects).
- Detection rate was highest in eye disorders (>50%), however, lowest efficiency but largest volume were observed in neurological disorders, where uncertain results (VUS) were most common, maybe due to multifactorial causes and detection rate of pathogenic and likely pathogenic variants for this was only 39.88%
- Mutation types were diverse, with missense variants being the most common, followed by frameshift and splice-site changes. A small proportion included in-frame indels, synonymous, stop-loss, and start-loss variants.
- Copy number variants (CNVs) also contributed a significant share, accounting for 13.6% of subjects.
- In conclusion, our tool, FestiVAR is reliable, time saving, reduces the margin of error and helps in accelerated interpretation of complex rare disorder subjects.