GeNeXpreSS

High Throughput Functional Assay Platforms to Screen Multiple Variants

Ch Neeraja, Ashwin Dalal

Diagnostics Division, Centre for DNA Fingerprinting ad Diagnostics, Hyderabad, India

Correspondence to: Dr Ch Neeraja *Email:* cneeraja22@gmail.com

Technological advances have led to the development of high-throughput sequencing platforms enabling human genome sequencing to be used in clinical practice. Several genomic variants are being identified across individuals of diverse populations, using high-throughput sequencing studies. Inability to ascertain clinical relevance to the identified pool of genetic variants continues to be a critical roadblock towards the development of precision medicine.

Variant Interpretation: Functional Assays to the Rescue (Starita et al., 2017)

Next generation sequencing has revolutionized the field of genetic diagnostics. However, ability to sequence large parts of the human genome has thrown up new challenges too. Notably, many of the variants are categorized as Variants of Unknown Significance (VUS) based on the ACMG/AMP criteria. Functional analysis of variants is quite challenging owing to the limitations of screening multiple variants using biochemical methods or computational predictions. Multiplexed assays for variant effects (MAVEs) is a powerful method to screen thousands of variants in a single experiment. The result of MAVE is a variant effect map which reveals the functional relevance of single variants in the genetic element. MAVEs are a family of methods that includes Deep Mutational Scanning (DMS) experiments to study protein sequence-function relationships and Massively Parallel Reporter Assays (MPRA) on gene regulatory sequences. The authors have reviewed various high throughput functional assays and have highlighted their utility in reclassification of variants of uncertain significance.

Massively parallel reporter assays for characterization of de novo promoter variants (Koesterich J et al., 2023)

Autism spectrum disorder (ASD) is a heritable and complex neurodevelopmental disorder that includes both common and rare/de novo variants (DNVs) in the coding and non-coding genome. Koesterich and team studied the DNVs of non-coding regions, in view of the limited insight regarding variations in the promoter and enhancer regulatory regions. Defects in neural progenitor cells (NPCs) have been implicated in ASD and similar neurodevelopmental disorders. In the present study, the transcriptional impacts of DNVs were examined in NPCs. NPCs are also tractable for the lentivirus-based massively parallel reporter assay (lentiMPRA), which can simultaneously test thousands of sequences in a single experiment. These sequences included both wildtype (reference) and mutated (alternate) forms upstream of a minimal promoter and reporter gene, so that changes in the expression of the reporter gene can be detected for each sequence variant. Using this approach, 3600 promoter DNVs were characterized in ASD cases and sibling controls. A subset of 165 high confidence DNVs (HcDNVs) was identified. These HcDNVs were enriched with transcriptionally related genomic annotations including transcription factor binding and epigenetic markers of active transcription, suggesting their role in gene regulation.

Homology-directed repair (HDR) reporter assay to evaluate *BRCA1* variants (Nagy G et al., 2023)

Identification of variants which are functionally abnormal in tumor suppressor proteins is critical

for cancer surveillance, prognosis, and treatment options. *BRCA1* is an essential gene owing to its tumor suppression activity by regulating the repair of DNA double strand breaks *via* the homology directed repair (HDR) mechanism. Knowledge on the impact of pathogenic variants in ''actionable'' genes (e.g., *BRCA1* and breast cancer) provides evidence for medical management. Multiplexed functional assay includes testing of hundreds of protein variants simultaneously and determines their functional impact. MAVE studies reveal functional importance of residues in the *BRCA1* coiled-coil and serine cluster domains. Libraries of *BRCA1* mutated at single amino acid residues from 1280–1576 were generated and function of these variants was analysed in the homology-directed repair (HDR) reporter assay. Nagy and team have employed a HDR Reporter assay to evaluate over 300 missense and nonsense *BRCA1* variants between amino acid residues 1280 and 1576, which encompasses the coiled-coil and serine cluster domains. It was inferred that the functionally abnormal variants tended to cluster in residues known to interact with *PALB2*, which is critical for homology-directed repair. Multiplexed results were confirmed by singleton assay and by ClinVar database variant interpretations.

Comprehensive functional characterization of *SGCB* coding variants predicts pathogenicity in limb–girdle muscular dystrophy type R4/2E (Li C et al., 2023)

Limb-girdle muscular dystrophy (LGMD) type R4/2E is caused by mutations in β-sarcoglycan (*SGCB*), which is a key component of the dystrophin-associated protein complex. In muscle cells, the dystrophin-associated protein complex localizes to the membrane and connects the intracellular cytoskeleton to the extracellular matrix, allowing for coordinated force production in muscle. The sarcoglycan subcomplex (SGC) is composed of 4 single-pass transmembrane proteins: α-sarcoglycan, β-sarcoglycan, γ-sarcoglycan, and δ-sarcoglycan. Biallelic loss-of-function mutations in any subunit can lead to LGMD. More than 50% of patients clinically diagnosed with a myopathy carry a variant of unknown significance in a myopathy gene, often leaving them without a genetic diagnosis. To provide functional evidence for the pathogenicity of missense variants, Li and team performed deep mutational scanning of *SGCB* and assessed SGC cell surface localization for all 6,340 possible amino acid changes. Lentiviral expression of YFP-SGCB-HA-WT plasmid construct in ADG-HEK cells displayed strong cell surface expression. ADG-HEK cells transduced with presumptive pathogenic variants had a significant decrease in cell surface expression of SGCB. Single amino acid saturation mutagenesis was employed to generate libraries comprising every possible missense, synonymous, and nonsense variant. Variant functional scores were bimodally distributed and perfectly predicted pathogenicity of known variants. Variants with less severe functional scores more often appeared in patients with slower disease progression, implying a relationship between variant function and disease severity.

References

1. Koesterich J, et al. Characterization of De Novo Promoter Variants in Autism Spectrum Disorder with Massively Parallel Reporter Assays. Int J Mol Sci. 2023; 24(4): 3509.
2. Li C, et al. Comprehensive functional characterization of SGCB coding variants predicts pathogenicity in limb-girdle muscular dystrophy type R4/2E. J Clin Inves. 2023;133(12): e168156.
3. Nagy G, et al. Multiplexed assay of variant effect reveals residues of functional importance in the BRCA1 coiled-coil and serine cluster domains. PLoS One. 2023; 18(11): e0293422.
4. Starita LM, et al. Variant Interpretation: Functional Assays to the Rescue. Am J Hum Genet. 2017;101: 315–325.