Cellular evidence and modeling platforms

Invitae's functional modeling platform controls the quality of data from both published and Invitae-generated high-throughput cellular assays to improve variant interpretation.

Background

When available, data from experimental studies that characterize the impact of genetic variants on protein function are considered strong evidence in support of benign or pathogenic classifications by the American College of Medical Genetics and Genomics/Association for Molecular Pathology (ACMG/AMP) guidelines[1] and Sherloc.[1,2]

For example, the ACMG/AMP guidelines for sequence variant interpretation include two functional evidence criteria: well-established in vitro or in vivo functional studies supportive of a damaging effect on the gene or gene product (i.e., criterion PS3) and well-established in vitro or in vivo functional studies show no damaging effect on protein function or splicing (i.e., criterion BS3). At Invitae, we have expanded this class of evidence in the Sherloc variant interpretation system to include more than 25 refined criteria to help improve consistency and reduce subjectivity. However, this type of evidence has been relatively scarce in the scientific literature, as effects on protein function have been experimentally characterized for only a small fraction of genetic variants. Over the past decade, high-throughput sequencing-based cellular assays, collectively termed multiplex assays of variant effect (MAVEs), have been developed to systematically characterize a wide array of molecular functions, including protein-protein interactions,[3] enzymatic activity,[4] regulatory potential,[5] and protein stability.[6] Unlike previous approaches, MAVEs enable the characterization of many DNA variants within a single, pooled experiment (see online appendix, available at invit.ae/cep-appendix).

MAVEs present a useful opportunity to incorporate new, highly informative data into variant interpretation during clinical germline genetic testing. However, many MAVE experiments have been performed for basic research purposes rather than for clinical utility. As such, among the hundreds of MAVE experiments performed over the past decade, only a few dozen focused on genes associated with hereditary monogenic diseases and even fewer characterized the well-established molecular functions associated with those diseases (e.g., signaling, protein stability, cell death). Furthermore, because the technologies behind MAVEs were only emerging when professional practice guidelines for variant interpretation were last published by ACMG/AMP,[1] there are no standardized quality control practices for assessing the clinical value of MAVE data. As a result, the application of MAVE data in variant interpretation by genetic testing laboratories has been inconsistent or has required the generation of gene-specific or publication-specific expert recommendations that are both complex and rigid.[7]

Framework for evaluating high-throughput experimental data and integrating the data into clinical interpretation

We previously described Invitae's functional modeling platform (FMP),[8] which employs a machine learning approach to control the quality of different types of data that are then weighted and used to predict the impact of genetic variants on molecular function. As such, the FMP provides a framework for uniformly analyzing MAVE data and integrating pathogenicity scores from the FMP into the Sherloc variant interpretation system **(Figure 1)**.[2]

Figure 1. Functional modeling platform.

High-throughput assay data and known pathogenic and benign variant labels are used to train a model. The performance of the model is evaluated to determine whether model predictions correlate with known pathogenic and benign variants. This evidence is then incorporated into the larger variant interpretation system within Sherloc, which also includes evidence such as patient phenotype and family segregation.

Using Invitae's FMP, we modeled 49 MAVE datasets from 22 publications (see online appendix). The performance of each model was assessed by calculating the area under the receiver operating characteristic (AUROC) curve using a minimum of five known pathogenic and five known benign variants per gene. The range of possible results, based on three sample datasets, is shown in **Figure 2**.

Of the 49 datasets, 42 showed relatively poor performance in their ability to discriminate between benign and pathogenic variants (AUROC < 0.8). Many factors likely contributed to this observed utility rate. For example, many MAVEs only characterized one aspect of a gene's function, whereas pathogenicity is often associated with multiple molecular functions. Additionally, many of these datasets included low numbers of clinically understood variants with which to gauge performance. Whatever the source, these results caution against naively using this evidence in clinical settings and highlight the importance of rigorous quality control.

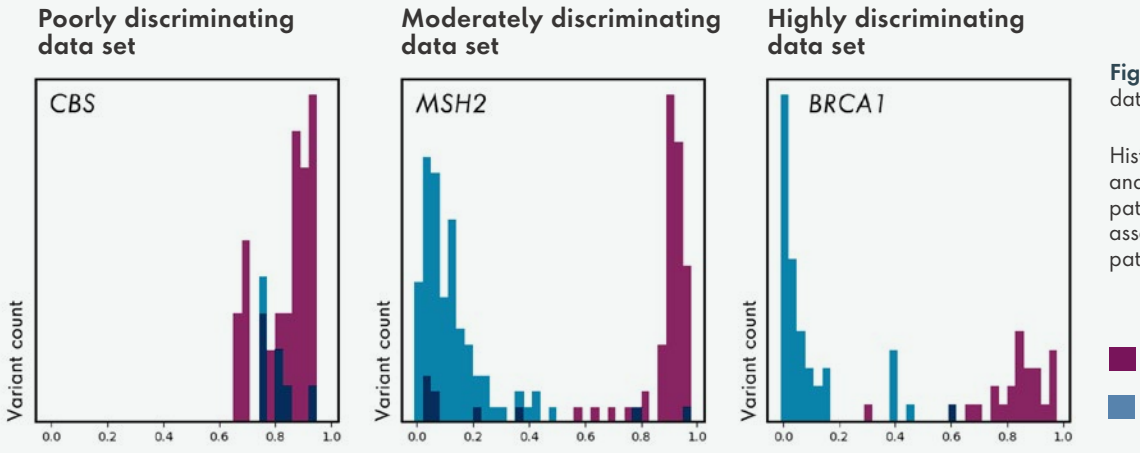

Figure 2. Evaluation of MAVE datasets using Invitae's FMP.

Histograms for *CBS*,[9] *MSH2*,[10] and *BRCA1*[11] show predicted pathogenicity based on the assay data (x-axis) for known pathogenic and benign variants.

The remaining seven datasets, focused on five genes *(BRCA1,*[11] *BRCA2,*[12] *MSH2,*[10] *SCN5A,*[13] *and TP53*[14-16]*)*, met the performance threshold that we set for clinical variant interpretation (AUROC ≥ 0.8) and have been incorporated into Sherloc. In addition to providing a quality control process for evaluating each MAVE dataset, the FMP calculates the degree of confidence for benign predictions (negative predictive value, NPV) and pathogenic predictions (positive predictive value, PPV) on a variant-by-variant basis. This allows a different weight to be assigned to each variant's score within Sherloc (see online appendix). Among the seven datasets that passed quality control, more than 3,000 unique variants achieved sufficiently confident predictions (>80% NPV or >80% PPV) to impact Sherloc variant interpretation scoring **(Figure 3)**. Importantly, this quality control framework can now be used to more rapidly evaluate novel datasets or to re-evaluate previous datasets as additional data become available.

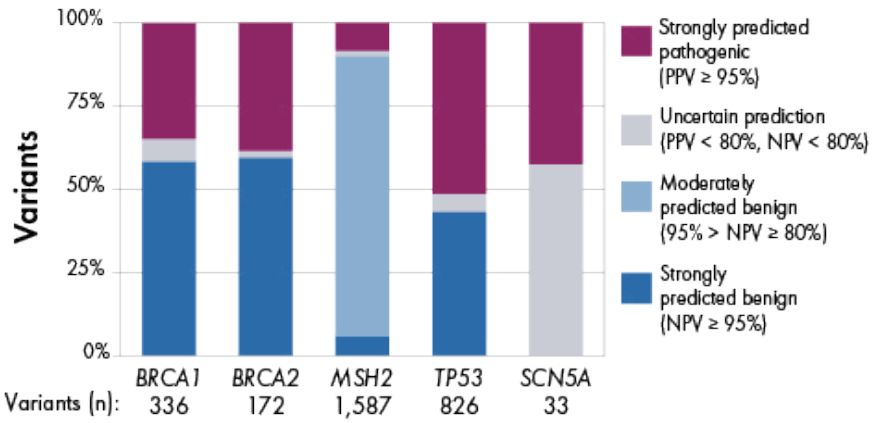

Figure 3. Predictive output of seven performant MAVE datasets from the published literature.

Functional evidence from seven datasets[10-16] for five genes that have been incorporated into clinical variant interpretation at Invitae. Bars depict the fractions of variants with pathogenic and benign predictions at various performance thresholds.

Cellular evidence platform: developing internal cellular data optimized for clinical impact

Invitae's expansive variant interpretation database and clinical testing experience present a unique opportunity to design MAVE studies that prioritize clinical impact for patients. By leveraging our clinical expertise and data, we can target previously overlooked disease genes, prioritize variant(s) of uncertain significance (VUS) that occur most frequently in patients, and refine the clinically understood variant sets used to train the machine-learning models. Moreover, given that functional evidence alone is not sufficient to produce definitive classifications, we can focus on variants for which orthogonal evidence (e.g., clinical observations) already exists in our database, as additional functional evidence may tip the interpretation scales in these cases.

To this end, Invitae has developed an experimental approach, called the Cellular Evidence Platform (CEP), that can generate functional data for genes and variants of clinical interest (**Figure 4**; see online appendix for detailed methodology). The CEP is a scalable system that enables functional characterization of variants in genes associated with either gain-of-function or loss-of-function mechanisms of disease. Importantly, evidence derived from these internal assays are integrated into variant interpretation using the FMP to ensure rigor in quality and scoring that is equivalent to that of externally published MAVE datasets.

Figure 4. Overview of Invitae's Cellular Evidence Platform.

First, a pool of variants are synthesized for a target gene. Next, the pool of variants is introduced into a cell population such that each cell has only one variant copy of the gene and all variants are present in the cell population. Single-cell RNA expression profiles are generated for the cells, such that an average of 100 independent cell profiles are generated per variant in the variant pool. Cells are mapped to variants using variant barcodes, and expression data for high-quality cells are normalized, filtered, and prepared for modeling. Gene expression-based features are selected and data are divided into training and validation sets. Supervised learning algorithms are trained using expression features of known pathogenic and benign variants. Finally, the most performant model for each gene is selected, controlled for quality, and incorporated into Sherloc.

As of February 2023, 44 Invitae-generated MAVE datasets have been evaluated. For these experiments, we targeted single-nucleotide variants that Invitae has observed rather than generating all possible single-nucleotide variants (a common practice in published MAVEs). Because functional evidence must be combined with clinical evidence to reach definitive classifications within Sherloc, this strategy allows us to test more genes without reducing the clinical impact of the functional evidence. Of these 44 datasets, 19 met the aforementioned predictive performance threshold (AUROC > 0.8; **Figure 5**), resulting in a success rate of 43% (19/44). These successful datasets span multiple biologic pathways and include examples of both loss-of-function and gain-of-function disease mechanisms. For the 19 genes with sufficiently predictive CEP-generated MAVE datasets, we observed an average reclassification rate of 4.6% among targeted VUS. For each reclassified VUS, an additional 18 VUS received Sherloc points, moving them one step closer to a future reclassification.

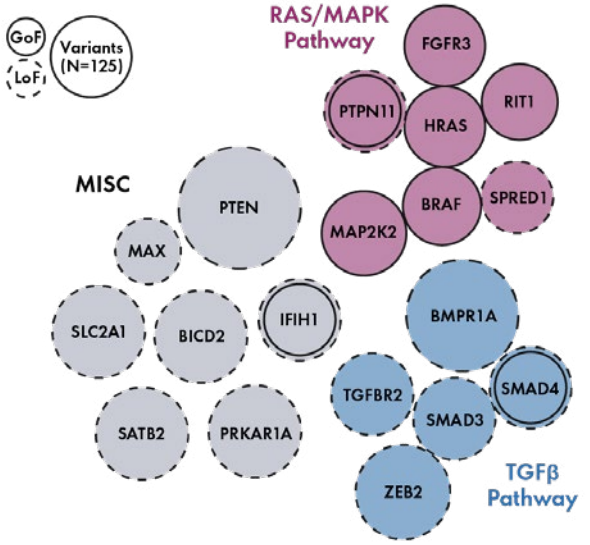

Figure 5. Genes with high-quality CEP-generated MAVE data.

As of February 2023, evidence based on the CEP is available for 19 genes from several large pathways. Circles that are touching represent genes that interact with each other. The size of the circle correlates with the size of the cellular experiment (not necessarily the number of integrated variants). Of these, *MAX* was the smallest experiment with 74 variants, and *PTEN* was the largest with 275 variants.

Summary

As more healthcare providers turn to genetic testing for diagnostic confirmation and treatment decisions, the continued development of scalable and innovative approaches for resolving VUS is critical. Although MAVEs represent such an approach and have been commonly used for academic research purposes, they have rarely been implemented in the diagnostic testing setting. This missed opportunity is largely a consequence of limited guidance from professional societies and expert user groups either on which genes, assays, and datasets qualify as well-established functional evidence, or on how best to validate these studies for clinical variant interpretation.

To address these challenges, Invitae utilizes our FMP to control the quality of MAVE datasets and incorporate them into clinical variant interpretation only if they pass a stringent threshold. This process effectively balances scalability and performance, allowing our interpretation system access to more high-quality evidence. As of February 2023, we have incorporated functional evidence for thousands of variants from 24 unique genes, with both externally (n = 5 genes) and internally (n = 19 genes) generated MAVE datasets. Together, this work highlights the significant potential of MAVE data for reducing VUS rates over time and demonstrates the benefit of generating MAVE datasets at Invitae to accelerate the clinical utility of high-throughput functional assays for variant interpretation.

Looking forward, we believe our CEP will become a key differentiator for Invitae. The advantages extend beyond improved interpretation of germline sequence variants. The rich data generated with the cellular models allow for a deeper understanding of disease mechanisms, including whether genes and variants have a gain-of-function or

loss-of-function mechanism of disease. In addition, functional analysis of multiple genes in the same biochemical pathway may help uncover candidate genes for existing diseases. Finally, we anticipate that the utility of these cellular models will extend to somatic variant interpretation, patient stratification of treatment options, and drug development.

View online appendix at invit.ae/cep-appendix.

References:

1. 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):405-424.
2. Nykamp K, Anderson M, Powers M, *et al.* Sherloc: a comprehensive refinement of the ACMG-AMP variant classification criteria. *Genet Med.* 2017;19(10):1105-1117.
3. Araya CL, Fowler DM, Chen W, Muniez I, Kelly JW, Fields S. A fundamental protein property, thermodynamic stability, revealed solely from large-scale measurements of protein function. *Proc Natl Acad Sci U S A.* 2012;109(42):16858-16863.
4. Romero PA, Tran TM, Abate AR. Dissecting enzyme function with microfluidic-based deep mutational scanning. *Proc Natl Acad Sci U S A.* 2015;112(23):7159-7164.
5. Kwasnieski JC, Mogno I, Myers CA, Corbo JC, Cohen BA. Complex effects of nucleotide variants in a mammalian cis-regulatory element. *Proc Natl Acad Sci U S A.* 2012;109(47):19498-19503.
6. Hasle N, Matreyek KA, Fowler DM. The Impact of Genetic Variants on PTEN Molecular Functions and Cellular Phenotypes. *Cold Spring Harb Perspect Med.* 2019;9(11).
7. Fortuno C, Lee K, Olivier M, *et al.* Specifications of the ACMG/AMP variant interpretation guidelines for germline TP53 variants. *Hum Mutat.* 2021;42(3):223-236.
8. Invitae. Integrating functional modeling to enhance clinical variant interpretation in Mendelian disease. Invitae.
9. Sun S, Weile J, Verby M, et al. A proactive genotype-to-patient-phenotype map for cystathionine beta-synthase. *Genome Med.* 2020;12(1):13.
10. Jia X, Burugula BB, Chen V, *et al.* Massively parallel functional testing of MSH2 missense variants conferring Lynch syndrome risk. *Am J Hum Genet.* 2021;108(1):163-175.
11. Findlay GM, Daza RM, Martin B, *et al.* Accurate classification of BRCA1 variants with saturation genome editing. *Nature.* 2018;562(7726):217-222.
12. Richardson ME, Hu C, Lee KY, *et al.* Strong functional data for pathogenicity or neutrality classify BRCA2 DNA-binding-domain variants of uncertain significance. *Am J Hum Genet.* 2021;108(3):458-468.
13. Glazer AM, Wada Y, Li B, *et al.* High-Throughput Reclassification of SCN5A Variants. *Am J Hum Genet.* 2020;107(1):111-123.
14. Kato S, Han SY, Liu W, et al. Understanding the function-structure and function-mutation relationships of p53 tumor suppressor protein by high-resolution missense mutation analysis. *Proc Natl Acad Sci U S A.* 2003;100(14):8424-8429.
15. Giacomelli AO, Yang X, Lintner RE, *et al.* Mutational processes shape the landscape of TP53 mutations in human cancer. *Nat Genet.* 2018;50(10):1381-1387.
16. Kotler E, Shani O, Goldfeld G, *et al.* A Systematic p53 Mutation Library Links Differential Functional Impact to Cancer Mutation Pattern and Evolutionary Conservation. *Mol Cell.* 2018;71(1):178-190.e8.