MaveMD: A functional data resource for genomic medicine

Abbye E. McEwen[1,2,3], Jeremy Stone[2], Malvika Tejura[3], Pankhuri Gupta[3,4], Benjamin J. Capodanno[2], Estelle Y. Da[5], Sally B. Grindstaff[2], Nick Moore[5], David Reinhart[2], Ashley E. Snyder[2], Andrew B. Stergachis[2,3,4], Lea M. Starita[2,3], Douglas M. Fowler[2,3,6*], Alan F. Rubin[5,7,8*]

Affiliations:
[1]Department of Laboratory Medicine and Pathology, University of Washington, Seattle, WA, USA
[2]Brotman Baty Institute for Precision Medicine, University of Washington, Seattle, WA, USA
[3]Department of Genome Sciences, University of Washington, Seattle, WA, USA
[4]Division of Medical Genetics, Department of Internal Medicine, University of Washington, Seattle, WA, USA
[5]Bioinformatics Division, Walter and Eliza Hall Institute of Medical Research, Parkville, VIC, Australia
[6]Department of Bioengineering, University of Washington, Seattle, WA, USA
[7]Department of Medical Biology, University of Melbourne, Melbourne, VIC, Australia
[8]Collaborative Centre for Genomic Cancer Medicine, University of Melbourne, Melbourne, VIC, Australia

*Correspondence to: alan.rubin@unimelb.edu.au and dfowler@uw.edu

Abstract

Variants of uncertain significance (VUS) undermine genetic medicine implementation because they have an unknown relationship to disease and cannot be used for clinical decision-making. While evidence from multiplexed assays of variant effect (MAVEs) can help resolve VUS, major barriers prevent routine clinical use, including data fragmentation and assay calibration. To address these challenges, we present MaveMD (MAVEs for MeDicine), a new interface for the MaveDB database that displays clinical evidence calibrations, provides intuitive visualizations, integrates with ClinVar and ClinGen, and exports clinical evidence compatible with ACMG/AMP guidelines. MaveMD currently contains 476,076 variant effect measurements curated from 82 MAVE datasets spanning 39 disease-associated genes, enabling classification of 75% of ClinVar VUS and 62% of future variants in these genes. MaveMD is designed to support and facilitate future data generation efforts and the use of MAVE evidence in clinical practice, thereby reducing the VUS burden and improving genetic medicine outcomes.

Introduction

Clinical interpretation of genetic variants is challenging, often because of variants of uncertain significance (VUS) that cannot be used for diagnosis or to guide clinical decision-making. The widely-used American College of Medical Genetics and Genomics (ACMG) and Association for

Molecular Pathology (AMP) variant classification framework integrates evidence including population frequency, familial segregation, case-control studies, computational predictions, and functional assays to classify variants as pathogenic, likely pathogenic, VUS, likely benign, or benign[1]. Each piece of evidence can support pathogenicity or benignity and is weighted as very strong, strong, moderate, supporting, or stand-alone. The final variant classification is determined by combining evidence using a point-based system[1,2]. If the point total does not meet the threshold for likely pathogenic or likely benign, the variant is classified as VUS. Up to 40% of individuals undergoing genetic testing receive VUS results[3–5], disproportionately affecting individuals of non-European ancestry[6,7].

Data from multiplexed assays of variant effect (MAVEs) are a powerful source of evidence for classifying genetic variants and resolving VUS[8]. These high-throughput experiments simultaneously characterize thousands of variants in a single assay. Each variant receives a functional score from the assay, generating a comprehensive variant effect map for the target gene[9–12]. Incorporating MAVE data into variant classification workflows allows reclassification of ~55% of VUS across well-studied genes[8,13], and can play a major role in reducing classification disparities for individuals from non-European ancestries[14]. Although MAVE-derived evidence is valuable for clinical variant classification, its use has been constrained because of barriers to discovering and accessing MAVE data, as well as the lack of necessary expertise amongst clinicians to evaluate data quality, assay utility, and evidence strength[15–17].

MAVE data has been difficult to access because it is fragmented across supplemental tables of publications, websites maintained by individual laboratories, gene-specific databases, general repositories like GitHub and Zenodo, and the MaveDB community database. This dispersion creates an untenable burden for clinical laboratories compounded by data loss as supplemental tables, websites and databases become unavailable over time[18–22]. Even if clinical users can locate MAVE data, they need key metadata to evaluate its clinical utility and inform its use, including information about the model system and experiment design, the molecular phenotype assessed, and the types of variants the assay detects (e.g., splicing, dominant negative, loss of function, or gain of function). Additionally, clinicians need to calibrate the data, a process that transforms MAVE-derived functional scores into evidence for clinical variant classification. Calibration relies on clinical control variants with previously-established pathogenic or benign classifications from sources like ClinVar[23] to serve as benchmarks for determining evidence strength[24]. Many MAVE datasets have been calibrated, but these

evidence assignments are often also buried in supplementary tables and therefore are not easily discoverable or usable. Moreover, multiple calibration methods are available, including methods that assign variant-specific evidence strength[25,26] and the ClinGen-specified OddsPath method[24], a useful quantitative metric that reflects an assay's ability to distinguish between benign and pathogenic variants.

To promote data discoverability, MaveDB has been established as the community database for MAVE data with 2,752 total datasets, including variant effect measurements in 713 human genes, 487 of which have a known relationship to human disease. In addition to variant functional scores and basic information about the assay target, MaveDB stores text metadata designed to meet the needs of data scientists and researchers, including a short description of the assay, an abstract, and abbreviated methods[27,28]. However, MaveDB lacked features needed by clinical users, including individual variant search, standardized human genomic coordinate mapping, clinically-oriented metadata, and clinical evidence strength for individual variants. Thus, even if clinical users downloaded a dataset from MaveDB, they would then need to merge information from diverse resources before using the MAVE data (**Figure 1A**).

The clinical utility of MAVE data depends on understanding the relationship between the specific assay molecular phenotype (*i.e.*, what the assay is measuring) and the gene-disease relationship being evaluated. An assay must measure a molecular function directly relevant to the disease pathophysiology to provide meaningful evidence. Therefore, clinicians must determine the molecular phenotype measured, the variant types the assay reliably detects (*e.g.*, loss-of-function, gain-of-function, and dominant negative), and if an assay can detect splicing variants or variants predicted to undergo nonsense mediated decay (NMD). Without this key information, even well-calibrated MAVE data may lead to incorrect variant classifications if applied to inappropriate gene-disease contexts, underscoring the importance of displaying easily understandable assay metadata. Furthermore, MAVE datasets have grown in scale, with some assessing more than 10,000 variant effects[29], and newer calibration methods are sufficiently complex[26] that a centralized repository is needed to provide calibrated data to clinical laboratories.

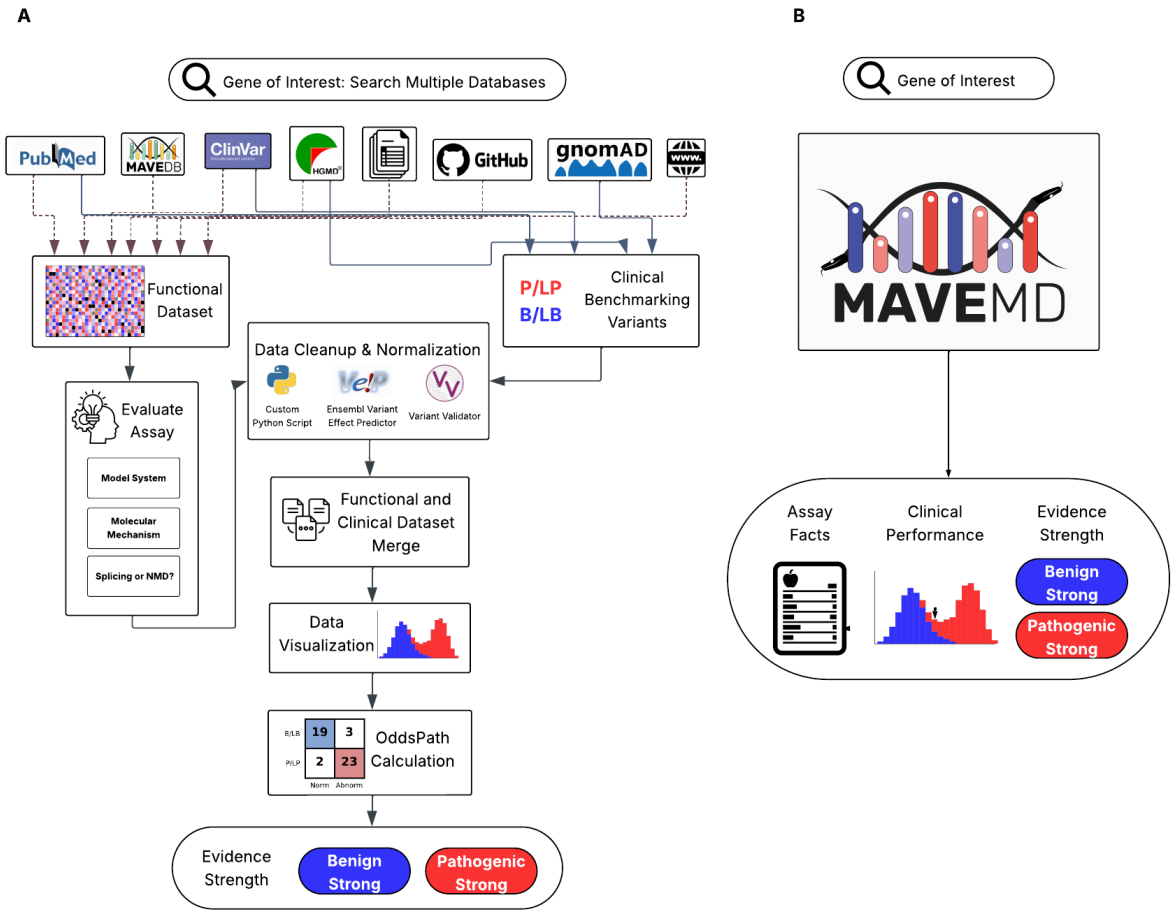


Figure 1: Comparison of workflows for using MAVE data in a clinical context. A) Currently clinical users must retrieve data from multiple databases, perform data cleanup and normalization, merge MAVE datasets with clinical control variants, assess clinical performance, and calibrate the functional/MAVE data to determine evidence strength. (B) The streamlined MaveMD interface displays the results of those steps in a single platform, providing direct access to key assay metadata, visualization of clinical performance, and evidence strength classifications in an accessible, clinician-friendly format.

To bridge this gap and bring functional assay data into medicine at scale, we present MaveMD (MAVEs for MeDicine), a resource built on MaveDB and informed by surveys of clinical users to help clinicians discover, evaluate and use MAVE data (**Figure 1B**). MaveMD includes a new variant search function powered by ClinGen Allele IDs to ensure accuracy and portability between information systems[30]. We also developed and implemented a standardized, clinically-relevant metadata model based on existing minimum information standards[31], and adopted modern Global Alliance for Genomics and Health (GA4GH) standards to power sophisticated API-based data integration[32,33]. MaveMD links MAVE functional scores to known pathogenic and benign clinical control variants from ClinVar, and exposes evidence strength assignments from two calibration methods[24,26]. The clinical interface is accessible via REST API for programmatic access or the web front-end with standardized and intuitive visualizations

allowing users to evaluate assay results, explore key metadata, and examine calibrations. We populated MaveMD with an initial set of 82 curated and calibrated MAVE datasets with 476,076 distinct variant effect measurements for 39 clinically-relevant genes, including 12 actionable genes on the ACMG secondary findings list[34]. This includes 251,573 (53%) variant effect measurements with evidence strength assignments from the ClinGen-recommended OddsPath method[24] and 452,304 (95%) from the research-use-only ExCALIBR method[26], and 475,199 (99.8%) having evidence strength from either or both methods. The curations and calibrations currently in MaveMD enable reclassification of 75% of ClinVar VUS and 62% of future variants in these genes, demonstrating a scalable approach to overcoming the VUS challenge[35]. Lastly, MaveMD is integrated with the Impact of Genetic Variation on Function (IGVF) Consortium's Catalog, providing access to MAVE data as it is generated[36]. Thus, MaveMD builds upon MaveDB by adding extensive metadata curation, evidence calibration, and visualizations designed to support routine integration of functional evidence into variant classification workflows.

Methods

Data curation and selection criteria

To establish a comprehensive collection of clinically relevant MAVE datasets, we performed systematic curation of published functional studies. We identified candidate datasets by searching MaveDB, querying ClinVar for publications citing foundational MAVE methodology or clinical assay calibration papers[13,24,37], and soliciting expert recommendations from the genetics community. To ensure focus on genes with sufficient evidence for clinical interpretation, we only included genes with moderate or greater disease associations as determined by either the ClinGen Gene-Disease Validity Working Group classifications or the GenCC database[38].

A multidisciplinary curation team led by a molecular genetic pathologist (A.E.M.) and including a senior graduate student (M.T.) and genetic counselor (P.G.) met weekly to review curations, resolve ambiguities, and ensure consistency. Systematic metadata was extracted across six key domains encompassing over 180 individual fields. Dataset identification and availability metadata included gene identifiers (HGNC symbols and IDs), publication details (PMID, year, first author), data locations (MaveDB URNs, supplemental tables, repositories), and accessibility status. Assay design parameters captured experimental methodology (arrayed vs. pooled, saturation vs. targeted variants), model system specifications (organism, cell line or

yeast details), library strategies (variant types, mutagenesis approach, delivery method), phenotypes measured, and molecular processes investigated. Technical performance metadata included replicate structure, variance analysis between replicates, and comparisons with other functional assays. Functional scoring included methods used for calculating variant effect scores and associated statistics. Score classification described methods for assigning functional classifications and assay thresholds for each functional class. Clinical performance metrics captured the use and source of pathogenic and benign clinical control variants, evidence strength assignments per ACMG/AMP guidelines (OddsPath), and performance statistics (sensitivity, specificity, PPV, ROC-AUC). This comprehensive metadata collection ensured standardized data capture across all curated datasets and directly informed the design requirements for MaveMD.

Data wrangling and standardization

The heterogeneous nature of published MAVE data necessitated extensive standardization before datasets could be submitted to MaveDB, which underlies MaveMD. Source data, including variant functional scores, were collected in diverse supplemental table formats (Excel, CSV, TSV, PDF) with inconsistent column naming conventions, variant nomenclature, and score representations. We developed a systematic pipeline to transform this information into MaveDB-compatible formats.

For each curated dataset, we performed manual inspection to identify the data's structure, including the variant representation format (amino acid substitutions, nucleotide changes, mixed nomenclature, custom notation), score columns and their meanings (raw scores, normalized scores, confidence intervals, standard errors), replicate structure when present, and any additional metadata columns requiring preservation. Converting variant descriptions to the HGVS-based format[39] used by MaveDB was the most complex part of the process. Source datasets employed various nomenclature systems: single-letter amino acid codes, three-letter amino acid codes, nucleotide positions without reference sequences, genomic coordinates in multiple genome builds, and laboratory-specific custom notations. Following standardization and deposition into MaveDB, all datasets were mapped to the human reference genome using sequence alignment as previously described[40].

Data counting and summarization

Datasets in MaveDB and elsewhere are counted as unique combinations of a functional assay and assay target. In MaveDB, this is tabulated at the level of the experiment record. For cases

where the same raw sequencing data was analyzed in multiple ways to produce different sets of variant effect measurements, these were not considered unique datasets.

After data wrangling and standardization, the 82 curated datasets encompassed 476,076 variant effect measurements deposited in MaveDB. This count reflects all assay-level measurements, where unique genetic variants tested in multiple assays contribute multiple counts to the total. After collapsing variants measured in multiple assays, the dataset contained 268,862 unique variants (**Supplemental Table 1**). In this manuscript, counts reported at the unique variant level are based on this collapsed dataset, while counts at the variant effect measurement level are taken from the complete set of assay-level measurements.

Interface prototype development and assessment

To inform the development of the MaveMD interface, we utilized findings from a comprehensive needs assessment survey[16]. The survey captured responses from 190 genetics professionals between February and June 2024, who evaluated a prototype MAVE data dashboard featuring ~4,000 *BRCA1* variants. Participants evaluated eight specific interface features using Likert-scale questions, assessing each component's potential utility for clinical variant interpretation. The dashboard included a histogram displaying MAVE functional score distributions colored by ClinVar significance with functional thresholds, OddsPath calculations, ACMG/AMP evidence codes, variant identifiers, functional consequences, scores, and associated errors.

Software implementation

MaveMD is built on the MaveDB open source codebase and shares many key components[28]. Briefly, the software is implemented in Python and JavaScript using FastAPI and Vue.js, respectively, and a Postgres database. MaveDB and MaveMD source code is available under the AGPL-3.0 license (see **Data and code availability**). They are deployed on Amazon Web Services as a series of Docker containers and other services (**Supplemental Figure S1**).

Results

Understanding MaveMD requirements

Since MaveDB's launch, clinical users have consistently requested features such as variant search that would help them leverage MAVE data in their practice. We surveyed 190 respondents, predominantly laboratory medical geneticists (23%) and variant review scientists

(23%), using a mockup MaveMD interface with different proposed features: metadata about how the assay was performed, functional classification thresholds (e.g., loss-of-function or functionally normal), standardized variant identifiers, a visualization comparing pathogenic and benign control variants to assay results, and calibrated evidence strength using ACMG/AMP v3 evidence codes[1,16].

Response to the mockup was overwhelmingly positive, with strong support for all the proposed features[16]. However, while 94% of respondents preferred having access to standardized ACMG evidence codes (PS3/BS3), only 68% stated they would directly use these codes if displayed. This gap revealed that clinical users prefer transparency and want to review both the evidence codes and the underlying supporting data before making decisions. These findings translated into clear requirements for MaveMD: robust mapping of MAVE variants to the reference genome, variant search functionality, inclusion of metadata essential for clinical interpretation, and presentation of calibrations with underlying clinical control variants.

Contextualizing MAVE data for clinical application

Inspired by the results of our survey, we systematically curated 82 clinically relevant MAVE datasets across 39 disease-associated genes[13,25,29,35,41–81]. Our curation data model included all metadata outlined in the MAVE minimum information standards[31] augmented with additional fields critical for clinical use, such as the molecular mechanism measured by the assay, variant consequences detected by the assay, and whether the assay could detect splicing variants or variants that would be subject to NMD in a physiologic context. Ultimately, each dataset underwent comprehensive metadata curation across six domains: dataset identification, assay design parameters, technical performance metrics, functional classifications, score calculations, and clinical performance statistics (**Supplemental Table 2**).

To understand the coverage of our MAVE data, we analyzed the overlap between variants characterized in our curated datasets and existing clinical resources. Of the 268,862 unique total variants across all curated MAVEs, 11.55% (31,065) were present in ClinVar, with 4.25% (11,417) having confident clinical classifications (pathogenic/likely pathogenic or benign/likely benign; **Supplemental Table 1**).

Curated datasets reflected diverse experimental approaches, contributing to the difficulties clinicians face when trying to evaluate assays. Arguably the most essential challenge is understanding the type of assay being performed, and this diversity emphasizes the importance of organizing and displaying interpretable metadata (**Figure 2A**). Of the 82

datasets we curated, six combined multiple dissimilar functional assays and were excluded. We found that assays were evenly split between those that measured overall protein function and those that measured a specific function. Assay types included cell fitness (54%), reporters (41%), and direct protein function (5%), with assay types further split into subtypes to help guide application of each dataset.

The assay systems also differed substantially across our curated datasets, with most experiments performed in various mammalian cell lines (68%) and the remainder in yeast cells. Variants were mostly generated using cDNA-based approaches that required introduction of a synthetic target sequence (82%), with the remaining minority using endogenous genome editing techniques such as saturation genome editing. The overwhelming majority (93%) of datasets used saturation mutagenesis, measuring the effects of all possible variants across a region or gene. Of the assays we curated, most were pooled assays, as expected for deep mutational scanning and other MAVE methods (82%), but a few used arrayed formats for technical or historic reasons.

As part of curation, we ensured that each dataset was uploaded to MaveDB. Of the 42 publications analyzed, 17 (41%) already had data deposited, 17 (41%) only had data in supplemental tables on journal websites, and one (2.4%) could only be found on a project-specific academic website, highlighting the discoverability challenges facing clinical users (**Figure 2B**). One curated dataset was associated with links that were no longer functional but available through other sources, and one dataset could not be curated because the only data source was a non-functional website. Likewise, two datasets were excluded because they were in PDF files that could not be easily converted.

For MAVE data to be translated into functional evidence, clinical control variants must be compared to functional classifications derived from variant effect measurements. While most datasets provided a functional classification for each variant or included information such as clearly defined score intervals (55%), the remaining studies did not. Additionally, only a minority of datasets included clinical evidence strength calculations (26%) with 25 using the OddsPath method and four using a variant-specific log-likelihood ratio method.

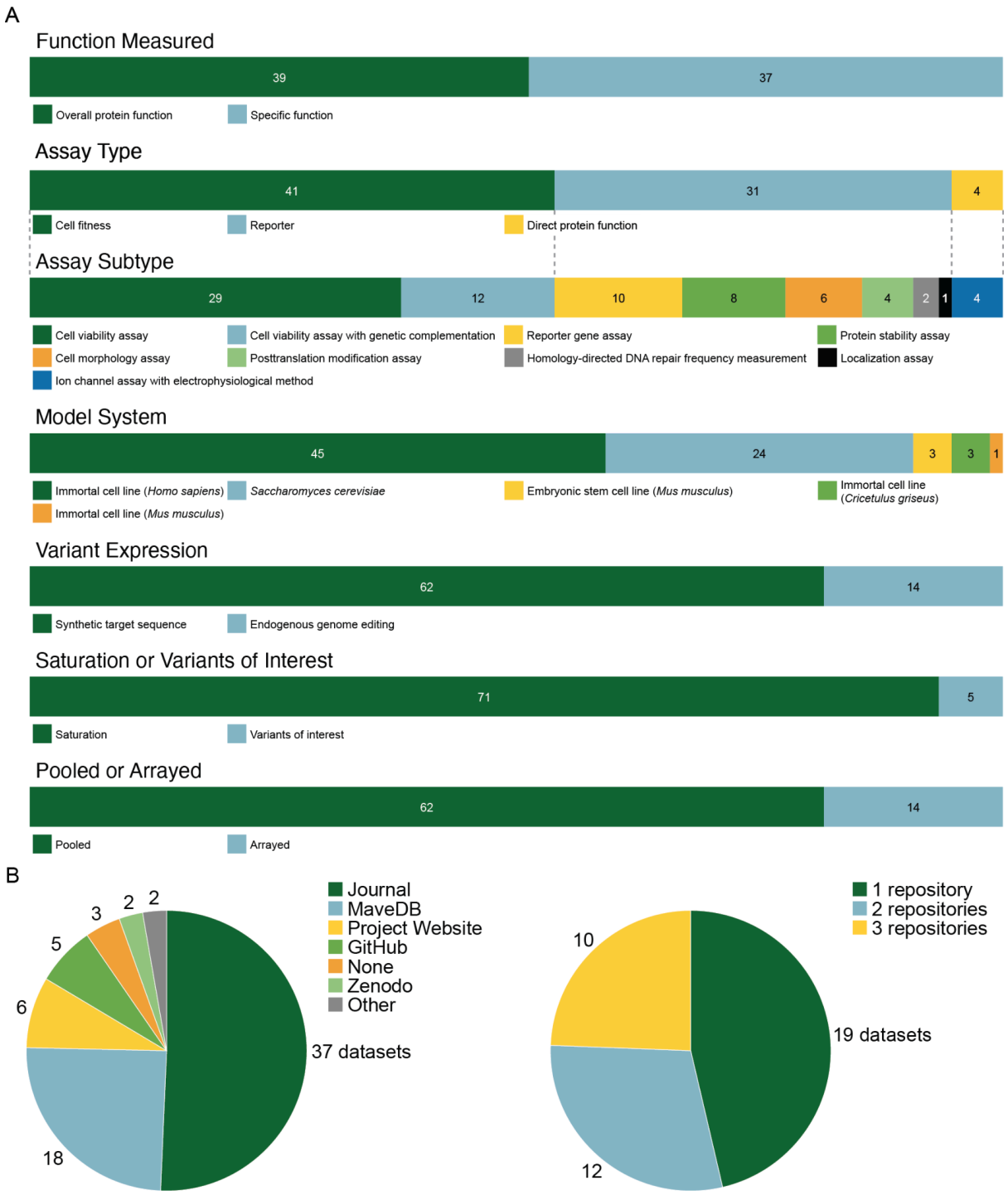


Figure 2: Summary of curated clinically-relevant MAVE datasets. A) The stacked bar plots show the proportion of datasets split across several metadata categories. The legend for each bar is shown below. B) The pie chart demonstrates the location of MAVE dataset deposition at the time of publication.

MaveMD clinical interface implementation

In developing the MaveMD interface, we were guided by three key priorities: variant search functionality, representation of metadata essential for clinical interpretation, and presentation

of calibrations with underlying clinical control variants[16]. Variant search presented significant challenges due to the diversity of transcripts, variant identifiers, variant formats, and coordinate systems. To address this complexity, we utilized the ClinGen Allele Registry, which provides and maintains universal identifiers for genetic variants[30]. We first implemented an automatic mapping process that determines the corresponding human genomic variant for every variant in MaveDB that comes from a human sequence, regardless of the assay system [40], allowing us to create or retrieve associated ClinGen Allele IDs from the Registry, as well as display annotations from other resources such as gnomAD (**Figure 3**).

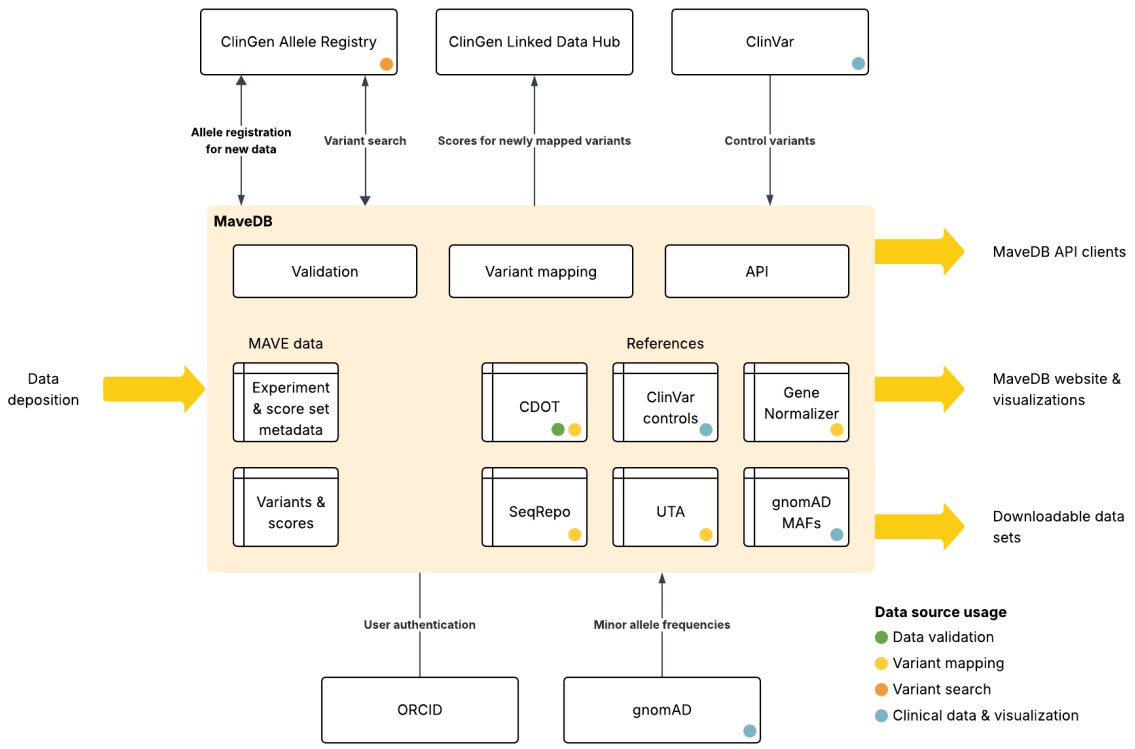


Figure 3: Data flow for clinically-relevant datasets in MaveDB. This schematic shows the various data sources and services used by MaveDB and MaveMD as part of the mapping and annotation process. Sources of reference information include the cdot transcript resolver (https://github.com/SACGF/cdot), classified variants from ClinVar, Gene Normalizer for resolving ambiguous sequences (https://github.com/cancervariants/gene-normalization), SeqRepo for storing a local collection of sequences[82], the Universal Transcript Archive (UTA) for storing aligned transcripts (https://github.com/biocommons/uta/), and population minor allele frequencies from gnomAD[83].

Having mapped variants and obtained ClinGen Allele IDs, we were able to implement two types of search designed to support commonly-used variant nomenclatures from clinical reporting systems. Users can search for variants with standard HGVS strings or through a structured "fuzzy search" interface that requires a gene symbol, variant type (protein or

cDNA), position, and reference and alternate alleles (**Supplemental Figure S2**). Search results display available MAVE datasets and provide direct access to a purpose-built clinical dashboard, including the option to choose between different functional datasets when multiple measurements are available (**Supplemental Figure S3**). The top panel displays variant identifiers and external links, including ClinGen Allele ID and ClinVar Variant ID, enabling integration and exploration in other clinical resources. Users are also able to link directly to a MaveMD variant from the corresponding variant page in the ClinGen Allele Registry via the ClinGen Linked Data Hub. The variant's functional consequence is also prominently displayed with a color-coded functional classification.

Upon searching for a specific variant, MaveMD provides a clinical view designed to help users access and interpret MAVE data and clinical metadata for variant classification. We developed a standardized presentation format that transforms MAVE data into actionable information, reducing the workload for clinical users while maintaining scientific rigor and transparency about assay capabilities and limitations. The interface contains three major features: assay fact labels, interactive visualizations displaying functional scores, and evidence calibration information.

The first major feature, assay fact labels, addresses the need for rapid assessment of a functional assay's relevance to a given clinical application (**Figure 4A**). They display key characteristics including assay type (cell fitness, reporter, direct protein function) and molecular phenotype assessed, model system (*e.g.*, human cells or yeast), variant functional consequences detected (loss of function, gain of function, dominant negative), the ability to detect splicing or NMD variants, and the total number of variants tested. A clinical performance section in each assay fact label provides at-a-glance evidence strength assessments using color-coded ACMG/AMP-style evidence codes and OddsPath values for both normal and abnormal function or a note if OddsPath values are not available. Thus, our assay fact labels highlight critical information for clinicians, allowing rapid comparison between different MAVE datasets using a consistent format, and reveal assays meeting specific clinical needs.

Figure 4A presents two examples: the left panel illustrates *BRCA1* saturation genome editing results that measures overall BRCA1 function related to cell survival and detects both splicing and NMD variants[46]. The right panel shows a cell survival based assay of DNA mismatch repair for MSH2, which, due to its cDNA-based approach, cannot detect variants affecting splicing or

causing NMD[56]. This distinction is clinically relevant since a laboratory investigating a putative splice variant needs to quickly identify that cDNA-based experimental methods might yield falsely normal results.

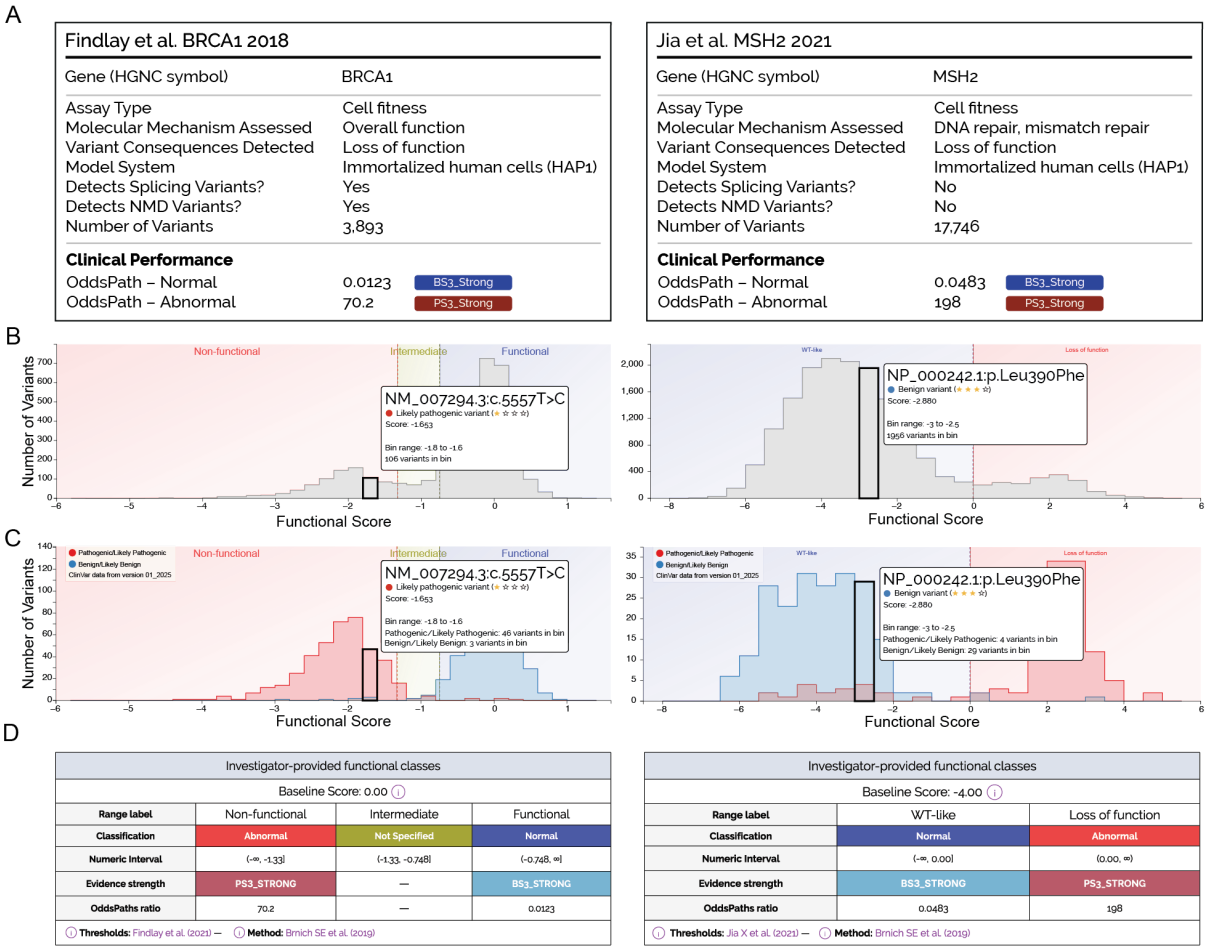


Figure 4: MaveMD data display elements. The MaveMD interface features multiple different visualizations and visual summaries. A) Each dataset has its own assay facts box that summarizes the key details that will help a user evaluate its clinical utility. Where needed, additional references are denoted with a superscript number that links to the full reference on the associated MaveDB or MaveMD page. B) Interactive histograms show the full distribution of variant scores in an assay and the relative position of a selected variant. Vertical bars also denote any evidence- or function-based cutoffs. C) Another interactive histogram displays the distribution of variant scores for variants from ClinVar. This view is intended to help users evaluate the clinical relevance of an assay based on the separation between independently-classified pathogenic and benign variants. Users can choose different snapshots of the ClinVar database. D) OddsPath and the associated evidence codes and score ranges are shown in a table below the histogram.

The second major feature facilitates the assessment of assay performance via an interactive histogram displaying the distribution of functional scores across all tested variants. This visualization highlights user-selected variants within the broader context of all variant effect measurements in the assay. In the "Overall Distribution" view, the histogram features clear demarcation of author-specified functional classes with colored backgrounds (for example, blue for "wild type-like" or "normal activity", red for "loss of function" or "abnormal activity", yellow for "intermediate") (**Figure 4B**). This color coding is especially helpful for a dataset like MSH2, as in contrast to most MAVE readouts, lower functional scores actually indicate normal protein activity.

The third major feature provides transparent access to evidence calibrations. By toggling to a "Clinical Controls View", the interactive histogram displays only the functional scores for classified ClinVar variants and shows OddsPath calculations and evidence strength assignments (**Figure 4C and D**). Variants are color-coded by their ClinVar status, with red for pathogenic/likely pathogenic and blue for benign/likely benign, alongside statistics showing the number of clinically classified variants in each functional score bin. This visualization immediately reveals whether functional scores are concordant with previously-classified clinical variants and highlights potential discordances requiring further investigation. Recognizing that multiple calibration algorithms already exist and continue to be developed, MaveMD allows users to toggle between the results of different methods. For example, we currently provide calibrations from the OddsPath and ExCALIBR methods (**Supplemental Figure S4**)[24,26].

Users can also download the MAVE datasets, calibrations, and metadata they need to make their own assessments. To preserve the rich contextual metadata required for interpretation, we use the GA4GH Variant Representation Specification (VRS) for describing genomic variants[33] and the GA4GH Variant Annotation Specification (VA-Spec) for describing assay scores, functional annotations, and clinical evidence strength (https://github.com/ga4gh/va-spec).

Discussion

Our systematic curation and centralization of MAVE datasets and the new MaveMD interface directly confronts the data dispersion crisis plaguing clinical genetics. The fragmentation we documented, with datasets scattered across supplemental tables, institutional websites, and various repositories, creates a substantial burden for clinical laboratories already managing

heavy caseloads. Prior to this work, clinical laboratories faced the daunting task of independently evaluating each functional assay, requiring both deep domain expertise and computational skills often unavailable in clinical settings. By establishing MaveMD as a central hub with persistent identifiers and standardized formats, we eliminate the need for exhaustive multi-platform searches while enabling long-term data discovery and preservation.

Fundamental challenges remain that both limit the current implementation and point toward future solutions. The inconsistent representation of measurement uncertainty across datasets complicates the incorporation of variant-level confidence into clinical interpretation. Much of this heterogeneity is driven by the diverse landscape of scoring methods available to researchers[84], as well as the distinct sources of error present in different experimental designs. Assembling a harmonized dataset as we have done here is the first step in tackling this systematically by further improving metadata annotations and identifying candidates for large-scale reanalysis.

The most complex unresolved issue is dealing with conflicting MAVE data. When multiple assays for the same variant yield discordant results, clinical users lack a clear framework for reconciliation. Because assays may measure different molecular properties, a system like ClinVar's transparent display of conflicting interpretations is not suitable. A variant could show normal protein abundance in one MAVE and demonstrate impaired enzymatic activity in another; both results may be correct, but their integration requires sophisticated understanding of disease mechanisms and the relative importance of different molecular functions.

Our current approach of displaying all available MAVE data with associated metadata promotes transparency, but leaves the task of reconciliation to clinical users, and the required expertise may exceed what is reasonable to expect from clinical laboratories. Datasets assembled in MaveMD can be used to train machine learning models generating composite functional scores that leverage the complementary strengths of different experimental approaches, as has been demonstrated for TP53[13,85]. MaveDB already supports multiple score calculations for a single assay, as well as model-based combinations of multiple assays in its data model. While large-scale reanalysis and model training is out of scope for MaveDB's current role as a data sharing platform, disseminating such results through MaveMD would be straightforward.

While MaveMD significantly improves access to MAVEs for clinical use, important limitations remain in connecting assays to specific gene-disease relationships. Currently, we provide comprehensive assay metadata including molecular phenotypes measured and variant types detected, but clinical users must determine whether a given assay is appropriate. Furthermore, our current implementation displays all pathogenic and benign clinical controls from ClinVar regardless of disease association, as disease annotations are often incomplete, inconsistent, or unreliable. This approach may include variants related to diseases with different molecular mechanisms than the condition of interest, potentially affecting calibration accuracy. Future development should prioritize integration with a centralized, well-curated resource that systematically maps gene-disease relationships to their underlying molecular mechanisms. Such a resource would enable MaveMD to automatically flag which assays are most relevant for specific clinical indications. Until this information is standardized and computationally accessible, clinical users must exercise careful judgment in selecting and applying MAVE evidence, considering both the molecular basis of the disease in question and the specific functional readout of each assay. Our curation experience illustrates the intensive effort required to extract, standardize, and calibrate each dataset, and scaling this manual process to accommodate the accelerating pace of MAVE publications is a fundamental challenge. By reporting crucial insights into what information is essential for clinical interpretation, our expanded data model will help shift this work from post-publication curation to data deposition, while also streamlining the process for data submitters by moving away from free text towards more intuitive structured metadata.

MaveMD and our associated expert curations demonstrate that the primary barriers to implementing functional evidence in the clinic are logistical rather than scientific. By addressing data centralization, metadata and format standardization, and calibration complexity, we have created a foundation for routine use of MAVE and other functional data that we hope will satisfy clinical needs now and into the future. However, challenges around measurement uncertainty, conflicting results, and sustainable maintenance require continued community effort. The long-term success of functional evidence in clinical practice depends on establishing sustainable infrastructure for data generation, data sharing, and platform maintenance. As functional assays become increasingly sophisticated and widespread, parallel advances in computational infrastructure, experimental and analytical methods, and community standards will be essential to realize their full potential for resolving variants of uncertain significance and reducing disparities in genetic medicine.

Declaration of interests

DMF is a scientific advisory board member of Alloz Bio. The remaining authors declare no competing interests.

Acknowledgments

This work was supported by an NIH NHGRI Advancing Medical Genomics Research award (R01HG013025), and L.M.S. and D.M.F. were supported by NHGRI IGVF grant (UM1HG011969). A.B.S. holds a Career Award for Medical Scientists from the Burroughs Wellcome Fund and is a Pew Biomedical Scholar. A.E.M. was supported by an Early Career Award from the Alex's Lemonade Stand for Childhood Cancer and RUNX1 Foundation (21-25037) and a Brotman Baty Institute Catalytic Collaborations Grant (CC28). P.G. was supported by the Career Ladder Education Program for Genetic Counselors grant from the Warren Alpert Foundation (WAF-CLEP-PD 10089501-01). This project received grant funding from the Australian Government.

Data and code availability

MaveDB, MaveMD, and associated online documentation are available at https://mavedb.org. Source code for MaveDB and MaveMD is available at https://github.com/VariantEffect/mavedb-api (for the back-end application) and https://github.com/VariantEffect/mavedb-ui (for the website). The present paper describes version 2026.1.0.

References

1. Richards, 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* **17**, 405–424 (2015).

2. Tavtigian, S. V. *et al.* Modeling the ACMG/AMP variant classification guidelines as a Bayesian classification framework. *Genet Med* **20**, 1054–1060 (2018).

3. Chen, E. *et al.* Rates and Classification of Variants of Uncertain Significance in Hereditary Disease Genetic Testing. *JAMA Netw Open* **6**, e2339571 (2023).

4. Chrysafi, P. *et al.* Prevalence of Variants of Uncertain Significance in Patients Undergoing Genetic Testing for Hereditary Breast and Ovarian Cancer and Lynch Syndrome. *Cancers (Basel)* **15**, 5762 (2023).

5. Rehm, H. L. *et al.* The landscape of reported VUS in multi-gene panel and genomic testing: Time for a change. *Genet Med* **25**, 100947 (2023).

6. Martin, B. E. *et al.* Comparing the frequency of variants of uncertain significance (VUS) between ancestry groups in a paediatric epilepsy cohort. *J Med Genet* **61**, 645–651 (2024).

7. Ndugga-Kabuye, M. K. & Issaka, R. B. Inequities in multi-gene hereditary cancer testing: lower diagnostic yield and higher VUS rate in individuals who identify as Hispanic, African or Asian and Pacific Islander as compared to European. *Fam Cancer* **18**, 465–469 (2019).

8. McEwen, A. E., Tejura, M., Fayer, S., Starita, L. M. & Fowler, D. M. Multiplexed assays of variant effect for clinical variant interpretation. *Nat Rev Genet* https://doi.org/10.1038/s41576-025-00870-x (2025) doi:10.1038/s41576-025-00870-x.

9. Findlay, G. M. Linking genome variants to disease: scalable approaches to test the functional impact of human mutations. *Hum Mol Genet* **30**, R187–R197 (2021).

10. Starita, L. M. *et al.* Variant Interpretation: Functional Assays to the Rescue. *The American Journal of Human Genetics* **101**, 315–325 (2017).

11. Tabet, D., Parikh, V., Mali, P., Roth, F. P. & Claussnitzer, M. Scalable Functional Assays for the Interpretation of Human Genetic Variation. *Annu Rev Genet* **56**, 441–465 (2022).

12. Weile, J. & Roth, F. P. Multiplexed assays of variant effects contribute to a growing genotype-phenotype atlas. *Hum Genet* **137**, 665–678 (2018).

13. Fayer, S. *et al.* Closing the gap: Systematic integration of multiplexed functional data resolves variants of uncertain significance in BRCA1, TP53, and PTEN. *Am J Hum Genet* **108**, 2248–2258 (2021).

14. Dawood, M. *et al.* Using multiplexed functional data to reduce variant classification inequities in underrepresented populations. *Genome Med* **16**, 143 (2024).

15. Allen, S. *et al.* Validating data from multiplex assays of variant effect: A CanVIG-UK national survey of NHS clinical scientists. *Am J Hum Genet* **112**, 1479–1488 (2025).

16. Park, M. S. *et al.* Insights on improving accessibility and usability of functional data to unlock their potential for variant interpretation. *Am J Hum Genet* **112**, 1468–1478 (2025).

17. Villani, R. M. *et al.* Consultation informs strategies for improving the use of functional evidence in variant classification. *Am J Hum Genet* **112**, 1489–1495 (2025).

18. Briney, K. A. Measuring data rot: An analysis of the continued availability of shared data from a Single University. *PLoS One* **19**, e0304781 (2024).

19. Costa, M., García S, A. & Pastor, O. The consequences of data dispersion in genomics: a comparative analysis of data sources for precision medicine. *BMC Med Inform Decis Mak* **23**, 256 (2023).

20. Federer, L. M. Long-term availability of data associated with articles in PLOS ONE. *PLoS One* **17**, e0272845 (2022).

21. Lemaire, B. *et al.* Web references are not eternal: time-trend and qualitative impact of the loss of access to online resources cited in peer-reviewed medical journals. *Curr Med Res Opin* **41**, 543–548 (2025).

22. Mighton, C. *et al.* Data sharing to improve concordance in variant interpretation across laboratories: results from the Canadian Open Genetics Repository. *J Med Genet* **59**, 571–578 (2022).

23. Landrum, M. J. *et al.* ClinVar: updates to support classifications of both germline and somatic variants. *Nucleic Acids Res* **53**, D1313–D1321 (2025).

24. Brnich, S. E. *et al.* Recommendations for application of the functional evidence PS3/BS3 criterion using the ACMG/AMP sequence variant interpretation framework. *Genome Medicine* **12**, 3 (2019).

25. van Loggerenberg, W. *et al.* Systematically testing human HMBS missense variants to reveal mechanism and pathogenic variation. *Am J Hum Genet* **110**, 1769–1786 (2023).

26. Zeiberg, D. *et al.* Gene-based calibration of high-throughput functional assays for clinical variant classification. 2025.04.29.651326 Preprint at https://doi.org/10.1101/2025.04.29.651326 (2026).

27. Esposito, D. *et al.* MaveDB: an open-source platform to distribute and interpret data from multiplexed assays of variant effect. *Genome Biology* **20**, 223 (2019).

28. Rubin, A. F. *et al.* MaveDB 2024: a curated community database with over seven million variant effects from multiplexed functional assays. *Genome Biol* **26**, 13 (2025).

29. Radford, E. J. *et al.* Saturation genome editing of DDX3X clarifies pathogenicity of germline and somatic variation. *Nat Commun* **14**, 7702 (2023).

30. Pawliczek, P. *et al.* ClinGen Allele Registry links information about genetic variants. *Hum Mutat* **39**, 1690–1701 (2018).

31. Claussnitzer, M. *et al.* Minimum information and guidelines for reporting a multiplexed assay of variant effect. *Genome Biol* **25**, 100 (2024).

32. Rehm, H. L. *et al.* GA4GH: International policies and standards for data sharing across genomic research and healthcare. *Cell Genom* **1**, 100029 (2021).

33. Wagner, A. H. *et al.* The GA4GH Variation Representation Specification: A computational framework for variation representation and federated identification. *Cell Genom* **1**, 100027 (2021).

34. Lee, K. *et al.* ACMG SF v3.3 list for reporting of secondary findings in clinical exome and genome sequencing: A policy statement of the American College of Medical Genetics and Genomics (ACMG). *Genet Med* **27**, 101454 (2025).

35. Tejura, M. *et al.* A scalable approach to resolving variants of uncertain significance. 2026.02.14.705848 Preprint at https://doi.org/10.64898/2026.02.14.705848 (2026).

36. IGVF Consortium. Deciphering the impact of genomic variation on function. *Nature* **633**, 47–57 (2024).

37. Fowler, D. M. & Fields, S. Deep mutational scanning: a new style of protein science. *Nat Meth* **11**, 801–807 (2014).

38. DiStefano, M. T. *et al.* The Gene Curation Coalition: A global effort to harmonize gene–disease evidence resources. *Genetics in Medicine* **24**, 1732–1742 (2022).

39. den Dunnen, J. T. *et al.* HGVS Recommendations for the Description of Sequence Variants: 2016 Update. *Hum Mutat* **37**, 564–569 (2016).

40. Arbesfeld, J. A. *et al.* Mapping MAVE data for use in human genomics applications. *Genome Biol* **26**, 179 (2025).

41. Adamovich, A. I. *et al.* The functional impact of BRCA1 BRCT domain variants using multiplexed DNA double-strand break repair assays. *Am J Hum Genet* **109**, 618–630 (2022).

42. Biar, C. G. *et al.* An integrated, scaled approach to resolve TSC2 variants of uncertain significance. 2026.01.16.699909 Preprint at https://doi.org/10.64898/2026.01.16.699909 (2026).

43. Boettcher, S. *et al.* A dominant-negative effect drives selection of TP53 missense mutations in myeloid malignancies. *Science* **365**, 599–604 (2019).

44. Bolognesi, B. *et al.* The mutational landscape of a prion-like domain. *Nat Commun* **10**, 4162 (2019).

45. Buckley, M. *et al.* Saturation genome editing maps the functional spectrum of pathogenic VHL alleles. *Nat Genet* **56**, 1446–1455 (2024).

46. Findlay, G. M. *et al.* Accurate classification of BRCA1 variants with saturation genome editing. *Nature* **562**, 217–222 (2018).

47. Fortuno, C. *et al.* An updated quantitative model to classify missense variants in the TP53 gene: A novel multifactorial strategy. *Hum Mutat* **42**, 1351–1361 (2021).

48. Gebbia, M. *et al.* A missense variant effect map for the human tumor-suppressor protein CHK2. *Am J Hum Genet* **111**, 2675–2692 (2024).

49. Gersing, S. *et al.* Characterizing glucokinase variant mechanisms using a multiplexed abundance assay. *Genome Biol* **25**, 98 (2024).

50. Gersing, S. *et al.* A comprehensive map of human glucokinase variant activity. *Genome Biol* **24**, 97 (2023).

51. Giacomelli, A. O. *et al.* Mutational processes shape the landscape of TP53 mutations in human cancer. *Nature Genetics* 1 (2018) doi:10.1038/s41588-018-0204-y.

52. Gilbert, M. A. *et al.* Functional characterization of 2,832 JAG1 variants supports reclassification for Alagille syndrome and improves guidance for clinical variant interpretation. *Am J Hum Genet* **111**, 1656–1672 (2024).

53. Glazer, A. M. *et al.* Deep Mutational Scan of an SCN5A Voltage Sensor. *Circ Genom Precis Med* **13**, e002786 (2020).

54. Grønbæk-Thygesen, M. *et al.* Deep mutational scanning reveals a correlation between degradation and toxicity of thousands of aspartoacylase variants. *Nat Commun* **15**, 4026 (2024).

55. Hu, C. *et al.* Functional analysis and clinical classification of 462 germline BRCA2 missense variants affecting the DNA binding domain. *Am J Hum Genet* **111**, 584–593 (2024).

56. Jia, X. *et al.* Massively parallel functional testing of MSH2 missense variants conferring Lynch syndrome risk. *Am J Hum Genet* **108**, 163–175 (2021).

57. Jiang, C. *et al.* A calibrated functional patch-clamp assay to enhance clinical variant interpretation in KCNH2-related long QT syndrome. *Am J Hum Genet* **109**, 1199–1207 (2022).

58. Kato, S. *et al.* Understanding the function–structure and function–mutation relationships of p53 tumor suppressor protein by high-resolution missense mutation analysis. *PNAS* **100**, 8424–8429 (2003).

59. Kozek, K. A. *et al.* High-throughput discovery of trafficking-deficient variants in the cardiac potassium channel KV11.1. *Heart Rhythm* **17**, 2180–2189 (2020).

60. Li, C. *et al.* Comprehensive functional characterization of SGCB coding variants predicts pathogenicity in limb-girdle muscular dystrophy type R4/2E. *J Clin Invest* **133**, e168156 (2023).

61. Lo, R. S. *et al.* The functional impact of 1,570 individual amino acid substitutions in human OTC. *Am J Hum Genet* **110**, 863–879 (2023).

62. Ma, J. G. *et al.* Multisite Validation of a Functional Assay to Adjudicate SCN5A Brugada Syndrome-Associated Variants. *Circ Genom Precis Med* **17**, e004569 (2024).

63. Ma, K. *et al.* Saturation mutagenesis-reinforced functional assays for disease-related genes. *Cell* **187**, 6707-6724.e22 (2024).

64. Matreyek, K. A. *et al.* Multiplex assessment of protein variant abundance by massively parallel sequencing. *Nature Genetics* **50**, 874–882 (2018).

65. McDonnell, A. F. *et al.* Deep mutational scanning quantifies DNA binding and predicts clinical outcomes of PAX6 variants. *Mol Syst Biol* **20**, 825–844 (2024).

66. Meitlis, I. *et al.* Multiplexed Functional Assessment of Genetic Variants in CARD11. *Am J Hum Genet* **107**, 1029–1043 (2020).

67. Mighell, T. L., Evans-Dutson, S. & O'Roak, B. J. A Saturation Mutagenesis Approach to Understanding PTEN Lipid Phosphatase Activity and Genotype-Phenotype Relationships. *The American Journal of Human Genetics* **102**, 943–955 (2018).

68. Muhammad, A. *et al.* High-throughput functional mapping of variants in an arrhythmia gene, KCNE1, reveals novel biology. *Genome Med* **16**, 73 (2024).

69. Olvera-León, R. *et al.* High-resolution functional mapping of RAD51C by saturation genome editing. *Cell* **187**, 5719-5734.e19 (2024).

70. O'Neill, M. J. *et al.* Multiplexed Assays of Variant Effect and Automated Patch Clamping Improve KCNH2-LQTS Variant Classification and Cardiac Event Risk Stratification. *Circulation* **150**, 1869–1881 (2024).

71. Popp, N. A. *et al.* Multiplex and multimodal mapping of variant effects in secreted proteins via MultiSTEP. *Nat Struct Mol Biol* https://doi.org/10.1038/s41594-025-01582-w (2025) doi:10.1038/s41594-025-01582-w.

72. Sahu, S. *et al.* Saturation genome editing-based clinical classification of BRCA2 variants. *Nature* **638**, 538–545 (2025).

73. Sahu, S. *et al.* Saturation genome editing of 11 codons and exon 13 of BRCA2 coupled with chemotherapeutic drug response accurately determines pathogenicity of variants. *PLoS Genet* **19**, e1010940 (2023).

74. Shepherdson, J. L. *et al.* Mutational scanning of CRX classifies clinical variants and reveals biochemical properties of the transcriptional effector domain. *Genome Res* **34**, 1540–1552 (2024).

75. Sun, S. *et al.* A proactive genotype-to-patient-phenotype map for cystathionine beta-synthase. *Genome Medicine* **12**, 13 (2020).

76. Sung, A. Y. *et al.* Systematic analysis of NDUFAF6 in complex I assembly and mitochondrial disease. *Nat Metab* **6**, 1128–1142 (2024).

77. Wan, A., Place, E., Pierce, E. A. & Comander, J. Characterizing variants of unknown significance in rhodopsin: A functional genomics approach. *Hum Mutat* **40**, 1127–1144 (2019).

78. Waters, A. J. *et al.* Saturation genome editing of BAP1 functionally classifies somatic and germline variants. *Nat Genet* **56**, 1434–1445 (2024).

79. Weile, J. *et al.* A framework for exhaustively mapping functional missense variants. *Molecular Systems Biology* **13**, 957 (2017).

80. Woo, I. *et al.* Saturation genome editing of BARD1 resolves VUS and provides insight into BRCA1-BARD1 tumor suppression. 2025.11.03.25339440 Preprint at https://doi.org/10.1101/2025.11.03.25339440 (2025).

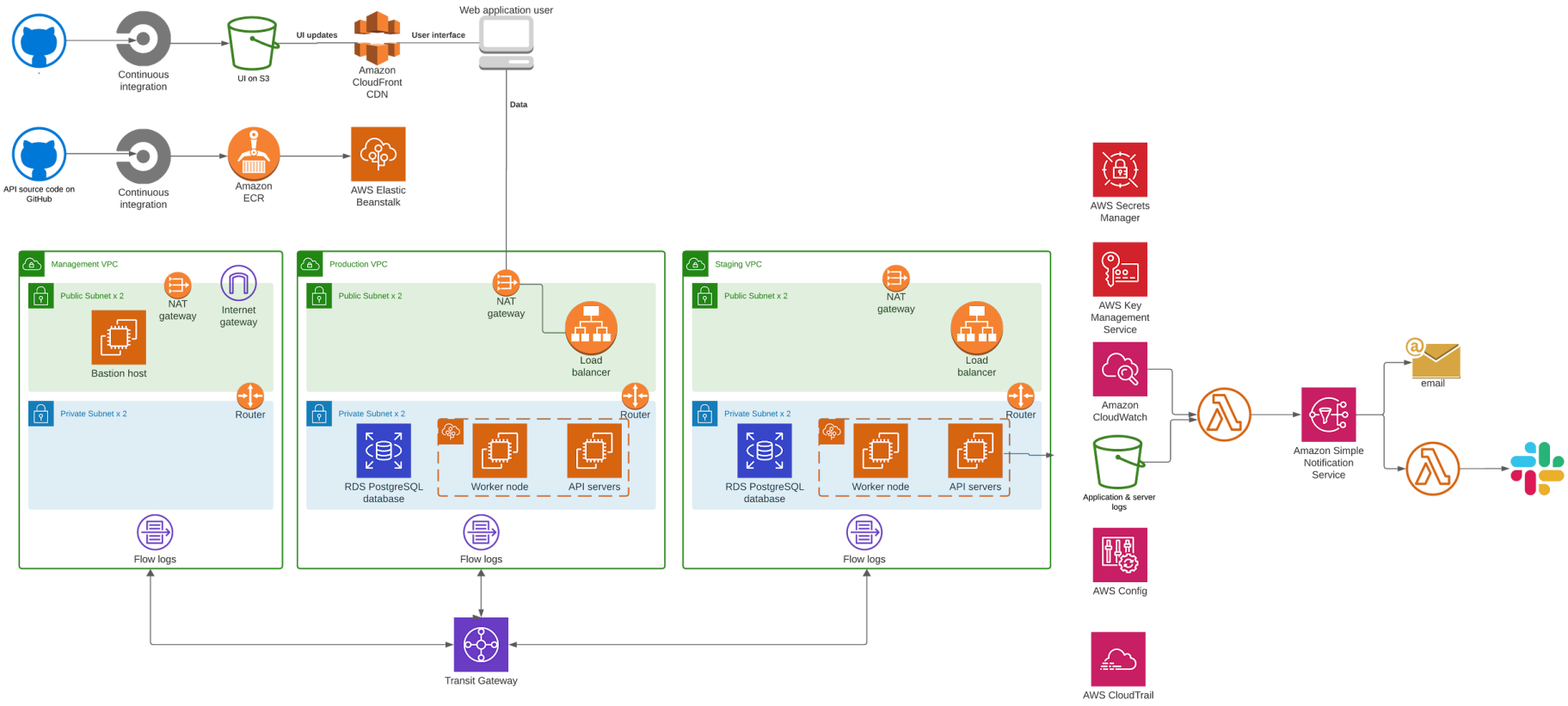
81. Zheng, H. *et al.* Proactive functional classification of all possible missense single-nucleotide variants in KCNQ4. *Genome Res* **32**, 1573–1584 (2022).

82. Hart, R. K. & Prlić, A. SeqRepo: A system for managing local collections of biological sequences. *PLoS One* **15**, e0239883 (2020).

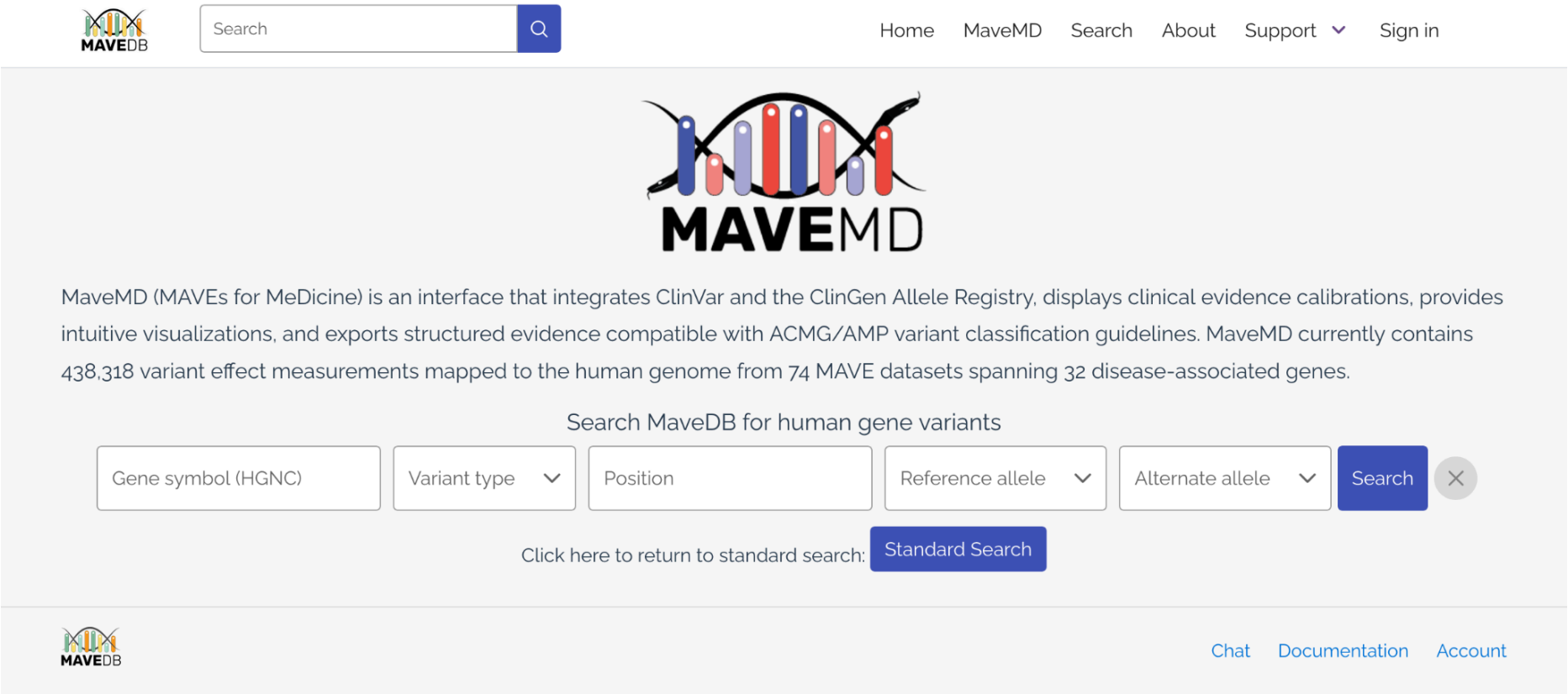
83. Chen, S. *et al.* A genomic mutational constraint map using variation in 76,156 human genomes. *Nature* **625**, 92–100 (2024).

84. Çubuk, H., Jin, X., Phipson, B., Marsh, J. A. & Rubin, A. F. Variant scoring tools for deep mutational scanning. *Mol Syst Biol* https://doi.org/10.1038/s44320-025-00137-x (2025) doi:10.1038/s44320-025-00137-x.

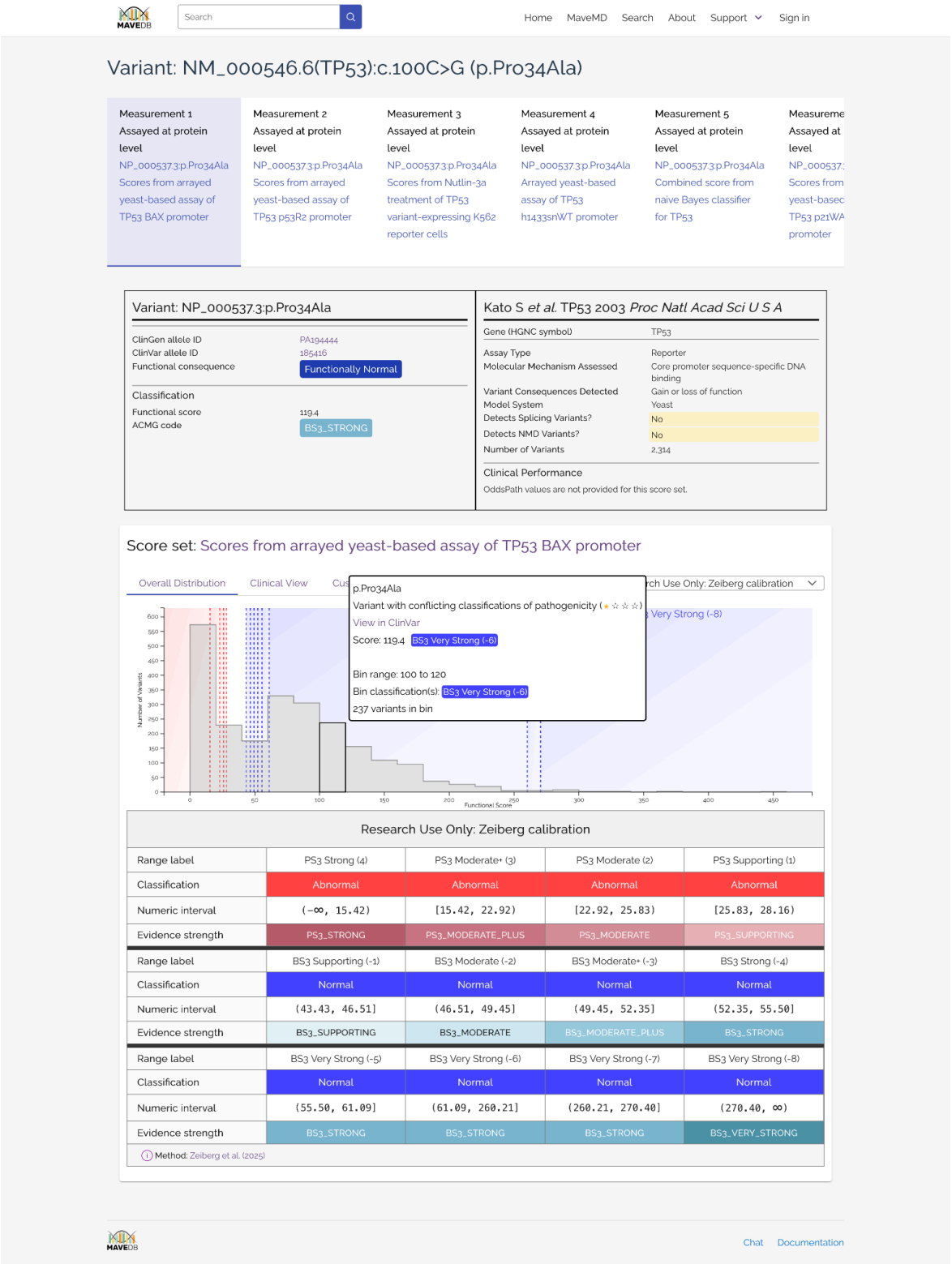
85. Calhoun, J. D. *et al.* Combining multiplexed functional data to improve variant classification. *ArXiv* arXiv:2503.18810v1 (2025).



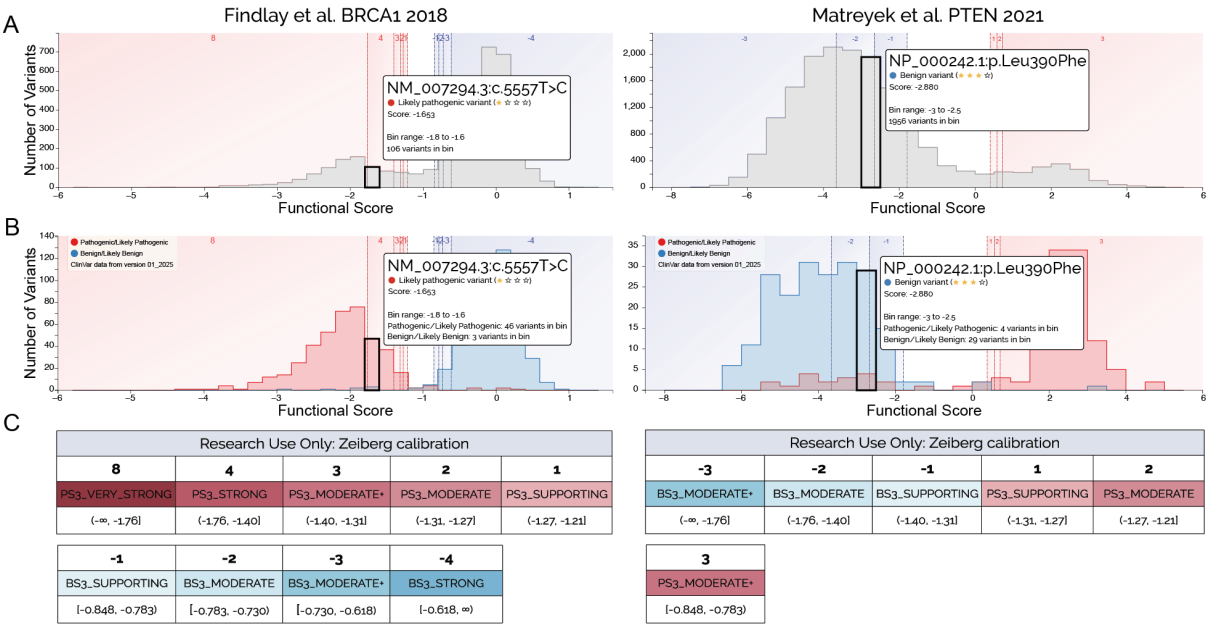
Supplemental Figure S1: Amazon Web Services (AWS) architecture diagram. MaveDB and MaveMD are hosted in a cloud computing environment provided by AWS. This schematic shows the various processes and containers involved in running, updating, and monitoring MaveDB and MaveMD.



Supplemental Figure S2: Variant fuzzy search form. This screenshot shows the search interface for the variant fuzzy search implemented in MaveMD. Variant type, reference allele, and alternate allele are dropdown selections, and gene symbol and position are free text. Input to this form is validated and passed to the ClinGen Allele Registry via API to resolve variant search queries.



Supplemental Figure S3: Variant search results page. This screenshot shows the results page for the variant search implemented in MaveMD. Available measurements (assay results) are available at the top of the page and the user can select which one should be used for visualization. The variant is shown in HGVS format, along with alternate identifiers. The variant's position in the assay score histogram is also highlighted.



Supplemental Figure S4: MaveMD support for alternative calibration methods. A) Interactive histograms show the full distribution of variant scores in an assay and the relative position of a selected variant. Vertical bars also denote evidence thresholds. B) Another interactive histogram displays the distribution of variant scores for variants from ClinVar. C) Clinical calibration details and the associated evidence codes and score ranges are shown in a table below the histogram. The histogram shows the calibrations based on the ExCALIBR method[26].