J Biochem Technol (2026) 17(4): 84-91
https://doi.org/10.51847/rzu9JnoYIJ
ISSN: 0974-2328

From Genetic Variants to Cellular Function: Integrating Functional Genomics and Quantitative Proteomics to Resolve Mutation Effects

Sarah Thompson*, Mariana Silva, Rafael Duarte, Emily Carter

Received: 08 September 2026 / Received in revised form: 12 October 2026, Accepted: 15 October 2026, Published online: 05 December 2026

Abstract

A sequence variant can alter RNA processing, the amount of protein produced, its molecular properties, or a downstream cellular phenotype. These effects may coexist, yet measurements at different levels need not agree. Functional-genomic screens measure such endpoints as transcript inclusion, growth, or reporter activity, whereas quantitative proteomics measures sampled peptides and proteins under particular analytical conditions. This integrative methodological review examines how complementary assays can be interpreted together without equating a change in signal with a single disease mechanism. Targeted literature selection prioritizes experimental variant-effect maps, RNA analyses, quantitative protein measurements, and methodological guidance on assay calibration and missing observations. The proposed interpretation approach distinguishes directly observed molecular layers from inferred intermediates, accounts for endogenous versus engineered expression, and uses discordance to select discriminating experiments. Reduced protein abundance may explain impaired activity in some variants; other variants preserve abundance while changing catalysis, interactions, localization, or pathway output. Missing peptides, tissue-restricted transcripts, saturation of cellular readouts, and background-dependent effects further limit apparent agreement. Integration is most defensible when samples, isoforms, controls, normalization, biological context, and uncertainty are explicitly aligned. The resulting evidence can prioritize mechanistic hypotheses and validation experiments, but does not independently establish clinical pathogenicity.

Sarah Thompson*
Department of Molecular Genetics, University of Toronto, Toronto, Canada.

Mariana Silva
Department of Genetics, University of Campinas, Campinas, Brazil.

Rafael Duarte
Department of Biochemistry, University of São Paulo, São Paulo, Brazil.

Emily Carter
Department of Computational Biology, University of Toronto, Toronto, Canada.

***E-mail:** sarah.thompson@utoronto.ca

Keywords: Variant effect mapping, Transcript splicing, Protein abundance, Multiplexed functional assays, Quantitative proteomics, Assay discordance

Introduction

The functional effect of a genetic change is often summarized as "loss of function," although this label can conceal markedly different biological events. A variant may lower transcript output, promote protein degradation, alter catalytic activity, disrupt an interaction, misdirect intracellular localization, or modify a phenotype through more than one route. Conversely, a preserved signal in one experiment cannot establish normality at another layer. As Starita *et al*. (2017) argued in surveying functional assays for variant interpretation, the usefulness of an experimental score depends on whether the assay interrogates a biologically relevant function and is interpretable against appropriate controls.

This distinction has become urgent as multiplexed assays of variant effect (MAVEs) scale beyond a few hand-selected alleles. A single experimental design can now assess large variant libraries, but assay scalability does not make measured endpoints interchangeable. Protein abundance is a plausible mediator of reduced activity, yet steady-state abundance is influenced by synthesis, degradation, cell state, and expression design. A growth phenotype may reflect the protein of interest while also integrating compensatory pathways and assay-specific selection. Even a strong molecular association is therefore not, without additional tests, proof of the intervening mechanism.

The question addressed here is how functional genomics and quantitative proteomics can jointly distinguish transcript, abundance, molecular activity, and cellular effects while preserving context and uncertainty. The central strategy is to specify the quantity directly measured by each method, identify which mechanistic link is only inferred, and choose further experiments according to the disagreements that remain. The focus is variant-specific inference rather than a comprehensive survey of omics platforms or a clinical classification algorithm.

Review Approach and Integration Strategy

This work is a targeted critical and integrative methodological review, not a systematic review or meta-analysis. A question-driven literature search undertaken on 1 November 2026 used publicly accessible scholarly indices, PubMed records, and

© 2026 The Author(s). This is an **Open Access** article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0).
https://creativecommons.org/licenses/by/4.0/deed.en

publisher pages, with searches combining genetic variant effects, deep mutational scanning, RNA splicing, protein abundance, quantitative proteomics, assay calibration, missingness, and variant interpretation. Relevant experimental papers were identified alongside methodological articles and interpretation recommendations; follow-up searches examined competing molecular explanations. The final bibliography is a purposively assembled set of journal publications, not the total number of eligible studies. No exhaustive database export, formal dual screening, pooled quantitative estimate, or risk-of-bias score was produced.

The unit of synthesis is an evidence claim tied to a defined variant, assay, model, and measurement scale—not a paper's general conclusion. For each candidate source, a source-use ledger records article identity, DOI, available content, system and perturbation, observed endpoint, claim limitation, and intended section. The comparison deliberately separates (i) sequence or transcript predictions, (ii) observed RNA processing, (iii) protein/peptide measurements, (iv) biochemical or interaction assays, and (v) cellular phenotypes. Database harmonization is equally important: Esposito *et al*. (2019) developed MaveDB to make multiplexed variant-effect data more accessible with associated assay information, but standardized storage cannot itself establish comparability across distinct biological endpoints.

Evidence was assessed for directness, appropriate calibration, assay coverage, expression dosage, and reproducibility, rather than placed into a single universal ranking. A reported experimental result is distinguished from its authors' mechanistic interpretation and from the conditional cross-assay inferences developed here. Where the accessible source establishes only a broad result, this review does not reconstruct unverified effect sizes or thresholds. The integrative logic is consequently explanatory and hypothesis-generating; **Tables 1-3** translate it into measurement, assay, and discordance questions without implying that all source studies used the same protocols.

From Variant to Molecular Consequence

A variant-to-phenotype explanation contains several potential links: the sequence alters a transcript or encoded protein; the altered molecular species changes its concentration or biochemical properties; those changes influence cellular networks; and a phenotype emerges in a particular tissue and time window. Assays typically observe selected links rather than the whole chain. A transcript readout and a protein abundance readout can be mutually informative without measuring the same biological quantity. Similarly, a variant score from a survival selection is an empirical cellular endpoint, not a direct enzymatic rate constant.

The practical consequence is that evidence should be organized by mechanistic layer before it is combined. Parallel findings can strengthen a mechanistic account when they connect through calibrated intermediate observations; inconsistent findings can also be informative when they expose a missing intermediate or an inappropriate context. The next sections first distinguish transcript-to-protein quantity from protein-to-cellular function.

Transcript Processing, Dosage, and Protein Abundance

Variant annotation as missense does not exclude a splice effect. Soemedi *et al*. (2017) demonstrated that disease-associated coding substitutions can disrupt splicing, making an amino-acid-based explanation incomplete when transcript consequences have not been assessed. RNA sequencing in affected or informative cells can expose abnormal junctions, allelic imbalance, and expression changes that are invisible in DNA annotation alone. Studies applying transcriptomics to unresolved Mendelian disease, including Kremer *et al*. (2017) and Cummings *et al*. (2017), illustrate the diagnostic value of measured RNA phenotypes while also emphasizing their dependence on the tissue and transcript actually sampled.

Prediction offers a way to direct those experiments, rather than replace them. Jaganathan *et al*. (2019) developed sequence-based modeling of splice outcomes, and Wagner *et al*. (2023) demonstrated the importance of tissue context in predicting aberrant splicing. An apparent null RNA result in an easily obtained cell type may therefore reflect tissue-restricted expression, isoform competition, or inadequate coverage instead of absence of a splice consequence. RNA abundance itself is not a translation-rate measurement: two transcripts with similar counts may differ in translational efficiency, while decay and transcript surveillance may alter the apparent isoform distribution.

Protein measurements add a necessary but nonidentical layer. Matreyek *et al*. (2018) used variant abundance by massively parallel sequencing to map steady-state abundance effects in cells; these reporter-linked values reveal abundance phenotypes under their assay configuration, not direct measurements of enzymatic activity. Deep mass-spectrometry studies show both the reach and the limitation of native peptide evidence. Sinitcyn *et al*. (2023) used extensive peptide coverage to improve detection of human sequence variants and isoforms, whereas ordinary protein-group summaries often cannot distinguish all isoforms or alleles. Targeted pathway proteomics in genetically characterized mouse material, as illustrated by Q. Yu *et al*. (2023), adds pathway-level abundance information but inherits the species, tissue, and sample-specific limits of the model.

These layers should be joined by matched identifiers, isoform definitions, allele representation, sample provenance, and compatible time points. For example, a lower peptide intensity may reflect lower protein quantity, loss of a detectable peptide after a sequence change, or modified digestion and ionization. Interpreting the change requires peptide-level corroboration and a reference measurement. **Table 1** specifies what each layer can support, and which inference remains unobserved.

Table 1. Observable molecular consequences and the inference gap between RNA and protein layers

Molecular layer	Primary measurement and unit	Permitted interpretation	Key limitation / discriminating test	Source
Splice usage	Junction or isoform fraction; percent or proportion	Altered transcript processing in sampled tissue	Compare matched tissue, alleles and isoforms	Soemedi *et al*. (2017)

RNA expression	Normalized counts or allelic ratio	Transcript quantity or imbalance	Test decay, cell composition and relevant tissue	Kremer *et al*. (2017)
Variant-linked abundance	Relative cellular reporter score	Steady-state abundance phenotype in assay system	Measure endogenous protein and activity separately	Matreyek *et al*. (2018)
Native proteomic evidence	Peptide intensity, sequence coverage	Detected peptides / possible proteoform support	Verify unique peptides, variant-dependent detectability	Sinitcyn *et al*. (2023)
Targeted pathway protein levels	Relative or normalized peptide/protein intensity	Protein differences in measured pathway and material	Account for tissue, genotype, depth and peptide mapping	Q. Yu *et al*. (2023)

Note. The categories and inference boundaries are the authors' cross-assay synthesis, not a common measurement protocol. Units describe possible reported formats, not pooled outcomes.

Protein Activity, Interactions, and Cellular Readouts

A reduction in protein concentration is one route to impaired function, but it does not establish that the molecules remaining are defective. Cagiada *et al*. (2021) compared variant effects on abundance and activity across large-scale data to distinguish mechanisms that cannot be recovered from either score alone. Chiasson *et al*. (2020) likewise examined vitamin K epoxide reductase variants with separate abundance and activity measurements. A variant with near-reference abundance but strongly altered activity deserves investigation of catalytic, binding, topological, or regulatory effects rather than automatic assignment to degradation.

Other assays observe additional properties. The LABEL-seq approach of Simon *et al*. (2024) profiled intracellular abundance together with functions such as interactions and drug-relevant properties, demonstrating the analytical value of multiple readouts for the same variant system. Protein localization can matter even where total quantity remains measurable; Lacoste *et al*. (2024) systematically studied mislocalization across disease-associated coding variants. Such phenotypes require imaging or compartment-sensitive assays and should not be silently inferred from bulk mass-spectrometric abundance.

The genetic perturbation also shapes what the endpoint means. Endogenous saturation genome editing of BRCA1 variants provides an allele-sensitive cellular readout within the edited setting (Findlay *et al*., 2018). Protein domain libraries, by contrast, can enable extraordinarily broad assessment of abundance effects: Beltran *et al*. (2025) profiled more than 500 human domains. These designs are complementary, not substitutes; domain behavior, edited endogenous alleles, and ectopically expressed constructs place different constraints on dosage, splicing, compartmentalization, and cellular adaptation. An apparently normal engineered-protein score can coexist with an abnormal endogenous transcript, and a growth defect need not isolate the immediate protein mechanism.

Table 2 separates these assay questions. It intentionally avoids a single "functional effect" column that would conceal differences between a biochemical property, a molecular localization event, and a selected cellular phenotype.

Table 2. Functional assays distinguished by the phenotype that can legitimately be inferred

Assay context	What is observed	What it cannot alone establish	Orthogonal check	Source
Paired abundance–activity screen	Relative abundance and defined activity	That lower abundance caused an activity change	Match active protein amount or test rescue	Chiasson *et al*. (2020)
Biochemical/interactome multiplexing	Defined binding, activity or interaction signals	Organism-level disease effect	Endogenous model and interaction specificity	Simon *et al*. (2024)
Protein localization imaging	Intracellular distribution of variant construct	Net stability or catalytic competence	Subcellular fractionation and functional assay	Lacoste *et al*. (2024)
Saturation genome editing	Fitness / selected cellular endpoint	Universal molecular mechanism	RNA, abundance and biochemical follow-up	Findlay *et al*. (2018)
Domain-scale variant abundance	Domain-associated abundance or stability proxy	Full-length protein phenotype in all tissues	Full-length endogenous validation	Beltran *et al*. (2025)

Note. Orthogonal checks are proposed discriminating tests, not experiments necessarily conducted in each cited study. A screen's score is meaningful only in its reported expression and cellular context.

Cross-Platform Integration and Discordant Findings

Cross-assay agreement is meaningful only after establishing what should agree. A low RNA level and low total protein level may suggest a dosage-linked mechanism, but cannot distinguish reduced transcription, abnormal RNA decay, diminished translation, or faster degradation without additional evidence. In contrast, normal abundance with poor pathway output directs attention to catalytic activity, interactions, localization, or pathway compensation. Grønbæk-Thygesen *et al*. (2024) reported relationships between degradation and toxicity across aspartoacylase variants; the association is mechanistically informative but cannot be generalized as a uniform causal relationship for other proteins and cellular settings.

Simple averaging of standardized assay scores is rarely justified. Scores can have opposite directions, different reference ranges, nonlinear saturation, distinct technical variance, or different

functional meanings. A more defensible procedure first names each endpoint and its uncertainty, then asks which causal explanations are compatible with the complete pattern. If an assay is missing, its score should not be imputed as "normal." Agreement can be stronger evidence of a specific intermediate mechanism when an intervention perturbs that intermediate and modifies the downstream phenotype in the predicted direction.

Assay Calibration, Controls, and Context Dependence

A variant score is interpretable only relative to the controls and dynamic range of its own experiment. Weile *et al*. (2017) described approaches for constructing large maps of missense function and calibrating their readouts against experimentally informative reference classes. Minimum reporting guidance from Claussnitzer *et al*. (2024) further emphasizes assay description, data and metadata conventions, and reproducible interpretation. Across platforms, a reference-like score should therefore be treated as an assay-specific result, not a universal certificate of unaffected function.

Calibration must address reference alleles, presumed damaging controls, replicate variation, batch effects, assay direction, and saturation. If deleterious variants cluster at a floor, differing molecular severities may become indistinguishable; if highly active variants exceed a ceiling, apparent agreement with wild type may be instrumental rather than biological. Smith and Kitzman (2023) benchmarked computational splice predictions against massively parallel splicing assays, illustrating why algorithmic scores require experimental context and why accuracy varies by variant class and benchmark. The comparison is pertinent to integrated interpretation: agreement between two correlated predictors is weaker corroboration than agreement between differently constructed direct measurements.

Model choice deserves the same scrutiny as assay choice. Allelic dosage, cell lineage, genetic background, cofactor availability, stress conditions, and sampling time can change the consequence of a variant. An overexpressed fusion protein may bypass endogenous transcriptional regulation or alter complex stoichiometry, while a genome-edited cell may compensate over passages. Consequently, the unit of replication should include biological preparations and relevant experimental batches rather than merely sequencing reads or technical injections. Calibration information should accompany every transferred variant score so that context mismatch remains visible.

Missingness, Identifiability, and Uncertainty Propagation

In quantitative proteomics, an unobserved peptide is not equivalent to a peptide concentration of zero. Kong *et al*. (2022) reviewed missing-data mechanisms and the dependence of imputation performance on assumptions. Variant-sensitive cleavage, abundance below detection, chromatography, ion suppression, and incomplete sequence databases are distinct causes with different implications. Suhre *et al*. (2024) showed how protein-altering variants matter for mass-spectrometry-based protein quantitative trait locus analysis; an apparent protein-level association may require scrutiny of peptide sequence and enrichment behavior before it is interpreted as changed total abundance.

An inference may also be underdetermined even when every platform produces a number. A reduced activity score alongside reduced abundance could represent a purely dosage effect, a specific activity defect, or both. Modeling stability–fitness correspondence can prioritize explanations, but Gerasimavicius *et al*. (2023) examined why deep-mutational-scanning functional scores and predicted stability effects are not universally equivalent. Similarly, imputation methods such as the FUSE approach described by T. Yu *et al*. (2024) estimate effects for incomplete functional screens; predicted scores should retain prediction intervals or uncertainty labels rather than be merged invisibly with observed measurements. ProteoCast, reported by Abakarova *et al*. (2026), provides another computational prioritization approach whose predictions remain distinct from directly measured protein activity.

The proposed response to discordance is a conditional experimental decision rather than a forced consensus score. **Table 3** identifies competing explanations and the smallest informative cross-check. A concordant pattern deserves similar scrutiny: shared technical bias, matched misannotation, or a common expression artifact can produce agreement across nominally different assays. Only evidence linked at the level of variant, isoform, perturbation, and biological system can narrow these alternatives.

Table 3. Interpreting discordant variant signals without assuming a single mechanism

Observation pattern	Competing explanations	Discriminating measurement	Status of inference
Altered RNA splicing; near-reference bulk protein	Isoform compensation; stable preexisting protein; peptide aggregation across isoforms	Allele-aware isoform protein evidence at a matched time	Conditional synthesis from Wagner *et al*. (2023)
Low abundance; low activity	Dosage-mediated loss; intrinsic defect; both	Normalize activity to active protein; controlled rescue	Conditional synthesis from Cagiada *et al*. (2021)
Normal abundance; changed pathway output	Altered catalytic specificity; partner interaction; mislocalization	Direct activity, interaction and localization measures	Conditional synthesis from Simon *et al*. (2024)
Variant peptide not detected	Low quantity; sequence-aware peptide loss; limited coverage	Alternative peptide/protease and targeted assay	Conditional synthesis from Sinitcyn *et al*. (2023)
Different experimental and predicted scores	Context-specific biology; model extrapolation; calibration bias	Replication in an appropriate biological model	Conditional synthesis from Gerasimavicius *et al*. (2023)

Note. Patterns and proposed cross-checks are original qualitative synthesis, not frequency estimates, pooled outcomes, or universal rules. "Near-reference" means within the source assay's relevant comparator range, not demonstrated normality across all functions.

Evidence Requirements for Variant Interpretation

An integrated variant assessment becomes defensible only when it distinguishes a measured abnormality from the explanation assigned to it. Three questions organize the evidence: Is the abnormal result technically reliable? Does the assay measure a function relevant to the gene's established disease mechanism? Does an orthogonal result narrow the route from molecular alteration to cellular consequence? The ACMG/AMP framework of Richards *et al*. (2015) treats functional evidence as one component of variant classification rather than a stand-alone diagnosis. Brnich *et al*. (2020) further specified the importance of assay validation, controls representing known benign and pathogenic variants, reproducibility, and calibrated interpretation of functional assays. Those requirements remain necessary when a persuasive multi-omics narrative is available.

The proposed first step is to record a variant's exact genome and transcript identifiers, zygosity, relevant isoform, and disease mechanism before attaching any assay score. For splice-related findings, one should measure abnormal junction usage or allelic expression in a relevant context and distinguish transcript degradation from altered transcript production. Smirnov *et al*. (2022) proposed guidance on the evidentiary strength of RNA phenotypes; importantly, a negative RNA result in an unsuitable tissue is not functional benign evidence. For protein effects, assay controls should test whether low signal reflects reduced abundance, poor peptide coverage, altered epitope recognition, or localization away from the sampled compartment.

The second step is to separate protein quantity from protein quality. A paired abundance–activity experiment with controlled dosage can test whether a loss of function is mostly explained by less protein, while partner-interaction, localization, and rescue studies can probe alternative mechanisms. Comparative measurements should use matched cell backgrounds, expression levels, isoform definitions, and time windows. An apparent effect observed only during engineered overexpression requires confirmation at a physiologically relevant dosage. Likewise, concordant growth and proteomics scores may share a confounder rather than demonstrate a causal chain. Replicate uncertainty and the calibration range must be preserved throughout integration.

Table 4 operationalizes these rules as conditional research judgments, not automatic clinical labels. There is evidence that well-calibrated multiplexed assays can contribute to resolving uncertain variants: Fayer *et al*. (2021) integrated functional results for BRCA1, TP53, and PTEN variants, within gene-specific clinical interpretation processes. Its applicability must not be generalized to unvalidated assays or unrelated disease mechanisms. Evidence of altered molecular function merits clinical interpretation only alongside gene–disease validity, inheritance, phenotype, segregation, and independent relevant information; separate readouts that measure the same process must not be double-counted.

Table 4. Conditional evidence combinations and validation priorities for variant interpretation

Current evidence	Research-level inference	Most informative validation	Interpretation boundary
Abnormal allele-aware RNA processing, with matched protein change	Transcript-mediated consequence is plausible	Confirm junctions, isoform, and effect in relevant cells	Tissue- and transcript-dependent; Smirnov *et al*. (2022)
Decreased abundance and decreased activity	Reduced dosage may contribute; intrinsic dysfunction remains possible	Assay activity per measured protein; controlled expression and rescue	Requires independently calibrated controls; Cagiada *et al*. (2021)
Near-reference abundance with impaired activity or interaction	A protein-quality mechanism is plausible	Direct biochemical, partner-binding, or localization test	A normal abundance value is endpoint-specific; Simon *et al*. (2024)
Cellular phenotype with no validated molecular intermediate	Functional effect is observed, mechanism unresolved	Orthogonal molecular readouts, rescue, and genetic-background replication	Clinical classification needs validated gene-specific evidence; Brnich *et al*. (2020)
Proteomic non-detection or assay disagreement	Neither molecular absence nor null effect is established	Targeted alternate peptides and independent method, including missingness model	Avoid treating missing signal as zero; Kong *et al*. (2022)

Note. Rows are this review's proposed interpretation choices grounded in the cited assay and guidance literature, not extracted clinical performance estimates. "Near-reference" describes one assay comparison; it is not a universal benign threshold.

Figure 1 summarizes why integrated measurements may converge, diverge, or remain uninformative. The central scientific distinction is between direct observations at particular layers and conditional explanatory connections between those layers. The return path from orthogonal validation represents a test of the proposed mechanism, not automatic accumulation of clinical evidence.

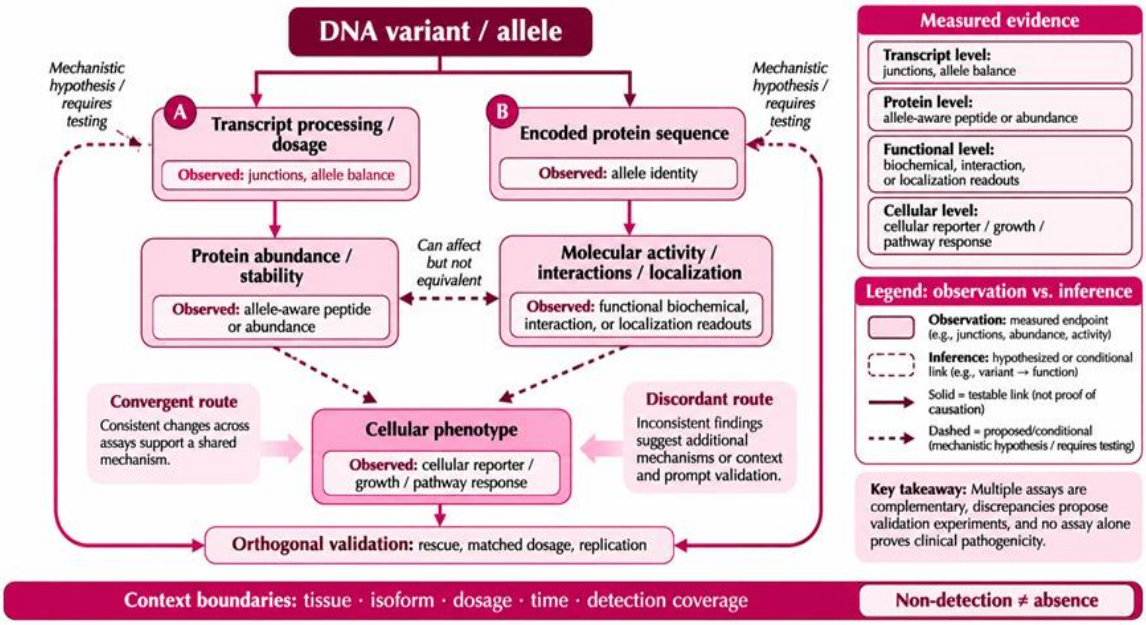


Figure 1. Variant-to-phenotype evidence routes and the limits of cross-assay inference

Note. A multilayer qualitative evidence map separates sequence identity, transcript processing, protein abundance, molecular activity or localization, and cellular phenotype. Solid connectors indicate experimentally testable relationships, not demonstrated causal effects for every variant. Dashed connectors indicate conditional mechanistic hypotheses requiring additional validation. A return arrow represents orthogonal validation and context reassessment. Measurement windows, cell type, isoform, dosage, and missing data constrain each comparison. This is an original synthesis, not a quantitative model or clinical classification algorithm.

Results and Discussion

The most useful outcome of cross-assay integration is often not agreement but a better-defined unresolved question. Reduced abundance and reduced function may support a dosage-mediated explanation, provided activity is evaluated relative to the protein actually present. Preserved abundance with reduced activity instead shifts attention toward catalytic sites, macromolecular partners, localization, and cell-specific compensation. Discordant RNA and protein results justify examining isoform representation, protein lifetime, sampling time, or measurement coverage before invoking a novel biological mechanism. This approach converts assay discrepancies into experiments that discriminate among rival explanations.

There is also a translational asymmetry. Experimental maps can rank variants efficiently, whereas disease attribution depends on whether the mapped function is causally relevant to the phenotype in a particular genetic and clinical context. Deep mutational scanning and proteome-scale prediction provide valuable prioritization, but computational prediction, reporter abundance, and clinical evidence should retain distinct status. The proposed evidence map is especially helpful for designing validation because it exposes unsupported links; it should not be collapsed into a single numerical pathogenicity score without independent calibration. Its usefulness will ultimately depend on prospective tests across variants, laboratories, cellular backgrounds, and disease mechanisms.

Limitations

This targeted review was purposively curated and does not estimate a pooled effect, systematic study coverage, or the frequency of discordance across all assays. Source accessibility differed, and not every report underwent comprehensive full-text re-extraction. Experimental platforms vary in expression dosage, genetic background, isoform coverage, reporter architecture, peptide detectability, and statistical treatment of missingness, constraining direct cross-study comparisons. Strong effects in a tractable cell line may not extrapolate to a disease-relevant tissue. The conditional validation scheme and figure are conceptual synthesis; they have not been prospectively evaluated as a classification system.

Conclusion

Functional genomics and quantitative proteomics can clarify variant mechanisms when each assay is tied to the molecular quantity it actually measures. RNA changes, protein abundance, biochemical performance, and cellular phenotypes provide related but non-equivalent evidence. Matched context, calibrated controls, explicit treatment of missingness, and orthogonal testing make convergence informative and discordance experimentally useful. Integrated evidence can refine mechanistic hypotheses and guide next experiments; clinical pathogenicity still requires separately validated, disease-relevant interpretation.

Acknowledgments: None

Conflict of interest: None

Financial support: None

Ethics statement: None

References

Abakarova, M., Freiberger, M. I., Liehrmann, A., Rera, M., & Laine, E. (2026). Proteome-wide prediction of the functional impact of missense variants with ProteoCast. *Nature Communications*, *17*(1), 3813. doi:10.1038/s41467-026-72140-1

Beltran, A., Jiang, X., Shen, Y., & Lehner, B. (2025). Site-saturation mutagenesis of 500 human protein domains. *Nature*, *637*(8047), 885–894. doi:10.1038/s41586-024-08370-4

Brnich, S. E., Abou Tayoun, A. N., Couch, F. J., Cutting, G. R., Greenblatt, M. S., Heinen, C. D., Kanavy, D. M., Luo, X., McNulty, S. M., Starita, L. M., et al. (2020). Recommendations for application of the functional evidence PS3/BS3 criterion using the ACMG/AMP sequence variant interpretation framework. *Genome Medicine*, *12*(1), 3. doi:10.1186/s13073-019-0690-2

Cagiada, M., Johansson, K. E., Valanciute, A., Nielsen, S. V., Hartmann-Petersen, R., Yang, J. J., Fowler, D. M., Stein, A., & Lindorff-Larsen, K. (2021). Understanding the origins of loss of protein function by analyzing the effects of thousands of variants on activity and abundance. *Molecular Biology and Evolution*, *38*(8), 3235–3246. doi:10.1093/molbev/msab095

Chiasson, M. A., Rollins, N. J., Stephany, J. J., Sitko, K. A., Matreyek, K. A., Verby, M., Sun, S., Roth, F. P., DeSloover, D., Marks, D. S., et al. (2020). Multiplexed measurement of variant abundance and activity reveals VKOR topology, active site and human variant impact. *eLife*, *9*, e58026. doi:10.7554/eLife.58026

Claussnitzer, M., Parikh, V. N., Wagner, A. H., Arbesfeld, J. A., Bult, C. J., Firth, H. V., Muffley, L. A., Nguyen Ba, A. N., Riehle, K., Roth, F. P., et al. (2024). Minimum information and guidelines for reporting a multiplexed assay of variant effect. *Genome Biology*, *25*(1), 100. doi:10.1186/s13059-024-03223-9

Cummings, B. B., Marshall, J. L., Tukiainen, T., Lek, M., Donkervoort, S., Foley, A. R., Bolduc, V., Waddell, L. B., Sandaradura, S. A., O'Grady, G. L., et al. (2017). Improving genetic diagnosis in Mendelian disease with transcriptome sequencing. *Science Translational Medicine*, *9*(386), eaal5209. doi:10.1126/scitranslmed.aal5209

Esposito, D., Weile, J., Shendure, J., Starita, L. M., Papenfuss, A. T., Roth, F. P., Fowler, D. M., & Rubin, A. F. (2019). MaveDB: An open-source platform to distribute and interpret data from multiplexed assays of variant effect. *Genome Biology*, *20*(1), 223. doi:10.1186/s13059-019-1845-6

Fayer, S., Horton, C., Dines, J. N., Rubin, A. F., Richardson, M. E., McGoldrick, K., Hernandez, F., Pesaran, T., Karam, R., Shirts, B. H., et al. (2021). Closing the gap: Systematic integration of multiplexed functional data resolves variants of uncertain significance in BRCA1, TP53, and PTEN. *The American Journal of Human Genetics*, *108*(12), 2248–2258. doi:10.1016/j.ajhg.2021.11.001

Findlay, G. M., Daza, R. M., Martin, B., Zhang, M. D., Leith, A. P., Gasperini, M., Janizek, J. D., Huang, X., Starita, L. M., & Shendure, J. (2018). Accurate classification of BRCA1 variants with saturation genome editing. *Nature*, *562*(7726), 217–222. doi:10.1038/s41586-018-0461-z

Gerasimavicius, L., Livesey, B. J., & Marsh, J. A. (2023). Correspondence between functional scores from deep mutational scans and predicted effects on protein stability. *Protein Science*, *32*(7), e4688. doi:10.1002/pro.4688

Grønbæk-Thygesen, M., Voutsinos, V., Johansson, K. E., Schulze, T. K., Cagiada, M., Pedersen, L., Clausen, L., Nariya, S., Powell, R. L., Stein, A., et al. (2024). Deep mutational scanning reveals a correlation between degradation and toxicity of thousands of aspartoacylase variants. *Nature Communications*, *15*(1), 4026. doi:10.1038/s41467-024-48481-0

Jaganathan, K., Kyriazopoulou Panagiotopoulou, S., McRae, J. F., Darbandi, S. F., Knowles, D., Li, Y. I., Kosmicki, J. A., Arbelaez, J., Cui, W., Schwartz, G. B., et al. (2019). Predicting splicing from primary sequence with deep learning. *Cell*, *176*(3), 535–548.e24. doi:10.1016/j.cell.2018.12.015

Kong, W., Hui, H. W. H., Peng, H., & Goh, W. W. B. (2022). Dealing with missing values in proteomics data. *Proteomics*, *22*(23–24), e2200092. doi:10.1002/pmic.202200092

Kremer, L. S., Bader, D. M., Mertes, C., Kopajtich, R., Pichler, G., Iuso, A., Haack, T. B., Graf, E., Schwarzmayr, T., Terrile, C., et al. (2017). Genetic diagnosis of Mendelian disorders via RNA sequencing. *Nature Communications*, *8*, 15824. doi:10.1038/ncomms15824

Lacoste, J., Haghighi, M., Haider, S., Reno, C., Lin, Z.-Y., Segal, D., Qian, W. W., Xiong, X., Teelucksingh, T., Miglietta, E., et al. (2024). Pervasive mislocalization of pathogenic coding variants underlying human disorders. *Cell*, *187*(23), 6725–6741.e13. doi:10.1016/j.cell.2024.09.003

Matreyek, K. A., Starita, L. M., Stephany, J. J., Martin, B., Chiasson, M. A., Gray, V. E., Kircher, M., Khechaduri, A., Dines, J. N., Hause, R. J., et al. (2018). Multiplex assessment of protein variant abundance by massively parallel sequencing. *Nature Genetics*, *50*(6), 874–882. doi:10.1038/s41588-018-0122-z

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W. W., Hegde, M., Lyon, E., Spector, E., et al. (2015). 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. *Genetics in Medicine*, *17*(5), 405–424. doi:10.1038/gim.2015.30

Simon, J. J., Fowler, D. M., & Maly, D. J. (2024). Multiplexed profiling of intracellular protein abundance, activity, interactions and druggability with LABEL-seq. *Nature Methods*, *21*(11), 2094–2106. doi:10.1038/s41592-024-02456-7

Sinitcyn, P., Richards, A. L., Weatheritt, R. J., Brademan, D. R., Marx, H., Shishkova, E., Meyer, J. G., Hebert, A. S., Westphall, M. S., Blencowe, B. J., et al. (2023). Global detection of human variants and isoforms by deep proteome

sequencing. *Nature Biotechnology*, *41*(12), 1776–1786. doi:10.1038/s41587-023-01714-x

Smirnov, D., Schlieben, L. D., Peymani, F., Berutti, R., & Prokisch, H. (2022). Guidelines for clinical interpretation of variant pathogenicity using RNA phenotypes. *Human Mutation*, *43*(8), 1056–1070. doi:10.1002/humu.24416

Smith, C., & Kitzman, J. O. (2023). Benchmarking splice variant prediction algorithms using massively parallel splicing assays. *Genome Biology*, *24*(1), 294. doi:10.1186/s13059-023-03144-z

Soemedi, R., Cygan, K. J., Rhine, C. L., Wang, J., Bulacan, C., Yang, J., Bayrak-Toydemir, P., McDonald, J., & Fairbrother, W. G. (2017). Pathogenic variants that alter protein code often disrupt splicing. *Nature Genetics*, *49*(6), 848–855. doi:10.1038/ng.3837

Starita, L. M., Ahituv, N., Dunham, M. J., Kitzman, J. O., Roth, F. P., Seelig, G., Shendure, J., & Fowler, D. M. (2017). Variant interpretation: Functional assays to the rescue. *The American Journal of Human Genetics*, *101*(3), 315–325. doi:10.1016/j.ajhg.2017.07.014

Suhre, K., Venkataraman, G. R., Guturu, H., Halama, A., Stephan, N., Thareja, G., Sarwath, H., Motamedchaboki, K., Donovan, M. K. R., Siddiqui, A., et al. (2024). Nanoparticle enrichment mass-spectrometry proteomics identifies protein-altering variants for precise pQTL mapping. *Nature Communications*, *15*(1), 989. doi:10.1038/s41467-024-45233-y

Wagner, N., Çelik, M. H., Hölzlwimmer, F. R., Mertes, C., Prokisch, H., Yépez, V. A., & Gagneur, J. (2023). Aberrant splicing prediction across human tissues. *Nature Genetics*, *55*(5), 861–870. doi:10.1038/s41588-023-01373-3

Weile, J., Sun, S., Cote, A. G., Knapp, J., Verby, M., Mellor, J. C., Wu, Y., Pons, C., Wong, C., van Lieshout, N., et al. (2017). A framework for exhaustively mapping functional missense variants. *Molecular Systems Biology*, *13*(12), 957. doi:10.15252/msb.20177908

Yu, Q., Liu, X., Keller, M. P., Navarrete-Perea, J., Zhang, T., Fu, S., Vaites, L. P., Shuken, S. R., Schmid, E., Keele, G. R., et al. (2023). Sample multiplexing-based targeted pathway proteomics with real-time analytics reveals the impact of genetic variation on protein expression. *Nature Communications*, *14*(1), 555. doi:10.1038/s41467-023-36269-7

Yu, T., Fife, J. D., Bhat, V., Adzhubey, I., Sherwood, R., & Cassa, C. A. (2024). FUSE: Improving the estimation and imputation of variant impacts in functional screening. *Cell Genomics*, *4*(10), 100667. doi:10.1016/j.xgen.2024.100667