Report of NHGRI Community Input: Functional Variant Interpretation February 16-April 12, 2024

Table Of Contents

Overview

As noted in the NHGRI 2020 Strategic Vision, a current barrier to advancing our understanding of human genomics is having the means to determine the functional consequences of genomic variants, both individually and in combination, on human health and disease. Better functional variant interpretation would likely contribute to advances in understanding disease biology, identification of therapeutic targets, understanding gene regulation, and better disease diagnosis. To obtain updated community input four years into our implementation of the 2020 Vision, we engaged members of the research community about future research directions in understanding the effects of variants on function on both short term (2-7 years from now) and long term (7-15 years from now) time frames. Participants with diverse expertise in functional genomics, human genetics, rare disease, gene regulation, technology development, computational genomics, and related areas were invited to provide written input and participate in one of four virtual sessions.

More details on the organization are included in Appendix 1, and the participant list is included in Appendix 2. All participants were asked to share their thoughts on six topic areas: Testing and Assessing Function of Individual Variants; Testing Variants in Context to Assess Function; Technology Development Needs; Predictive Modeling; Resource Building; and Additional Topics (Other). Across all six areas they were asked to discuss gaps, challenges, and opportunities for this field. Specifically, they addressed:

- An overarching question- how can we facilitate short term and long term scientific advances in this area with an eye towards making clinical progress?
- Short term (2-7 years from now) advances needed to obtain better understanding of how human variants affect traits and diseases, especially in areas that today are at the limit of our understanding.
- Long term (7-17 years from now) advances needed to implement these scientific developments to medical practice (e.g. diagnosis, treatment).

This report is a summary of community input received by the National Human Genome Research Institute (NHGRI), organized by themes that emerged in written comments or discussion. We are deeply grateful for participants' investment of time and effort and for the thoughtful and insightful input provided. Please note the views expressed in this document reflect both individual and collective opinions of contributors as captured by NHGRI, and we did not seek consensus; thus, there is no expectation that any individual endorses all the suggestions presented here. As variant interpretation has many applications (understanding gene regulation, understanding disease biology, identifying therapeutic targets, improving disease diagnosis, etc.) and different strategies may be appropriate for these applications, it is perhaps unsurprising that we received some divergent suggestions. Suggestions in each topic area that received the strongest community support are highlighted in Box 1.

Box 1: Highest Priority Topic Suggestions

- Continue testing and assessing function of individual variants at scale
- Increase testing of variants in biologically-relevant and disease-relevant contexts at scale, emphasizing cellular, environmental, and dynamic contexts
- Prioritize diversity in source genetic ancestries when prioritizing or testing variants
- Improve technology development in this space, including developing large-scale synthesis approaches at lower cost for the testing of variant combinations, improving methods to detect and test structural variation, and improving and scaling up of in vivo assays
- Optimize current assays such that they work in more diverse cell types and tissues with high sensitivity and at scale
- Develop more versatile assays and methods for measuring gene and variant effects at higher phenotypic scales such as cellular, tissue, and organismal phenotypes
- Leverage predictive models to improve experimental design
- Improve synergy between modelers and data generators through iterative data collection, AI/ML based modeling, and model testing with targeted experiments
- Develop a resource with a large collection of well-curated and standardized datasets (e.g., epigenomic, transcriptomic and perturbation) that can be used for both computational analyses and experimental design.
- Develop standards for assays, protocols, metadata, analyses, and predictions that would enable data sharing and re-use.

Testing and Assessing Function of Individual Variants

Overall, there was strong support for large scale testing of individual variants. Opinions differed on the balance of testing common/rare variants, non-coding/coding variants, and on the balance of testing at the level of variants, functional elements of the genome (herein referred to as "elements") or genes. There were also differing opinions on what assays might be most informative. Several short-term goals were suggested, including linking variants to measures of genome function, to affected genes and to pathways, in part to identify causal variants for disease. Suggestions for implementing these goals included a map linking genes to regulatory elements, transcription factors, and variants impacting their expression; developing a catalog of gene pathways; testing every Genome-Wide Association Study (GWAS) variant; and testing every clinical Variant of Unknown Significance (VUS).

Prioritizing variants: Many suggestions were made for prioritizing variants for testing. Some advocated for prioritizing large effect, rare or coding variants as they were most likely to be clinically actionable. Some advocated for prioritizing common or non-coding variants as the greatest need for improved understanding. Some advocated that a comprehensive

approach of studying these classes of variants together would be needed to fully understand the role of genetic variation in disease. There was also support for increased testing of structural variants and indels in addition to single nucleotide variants. Some saw potential in studying structural variants using emerging technologies like long-read sequencing, though others questioned the utility or generalizability of such approaches.

For non-coding variants, there were specific suggestions to prioritize variants in promoters and untranslated regions (UTRs), as well as highly conserved variants. There were suggestions to focus on GWAS variants to understand disease, as well as suggestions to focus on balanced variants for training artificial intelligence/machine learning (AI/ML) models, and for deeper exploration of the causality of non-coding variants for unexplained monogenic disease.

For coding variants, specific suggestions were to prioritize variants expected to affect covalent modification (e.g., phosphorylation, glycosylation) or to prioritize variants by predicted effects on protein structure as suggested by deep-learning or AI-based computational tools.

Many sources of information were suggested for variant prioritization. Greater diversity in the source genetic ancestries should be used in selecting variants for testing. There was strong support for using genomic datasets indicative of cell-type-specific genome function (transcriptome, chromatin accessibility, nuclear architecture, epigenomic), estimates of heritability enrichment, statistical or AI/ML fine-mapping approaches, as well as variants in genes with disease relevance (GWAS, ClinVar, druggability, etc.). Testing variants across the spectrum of conservation was suggested, including primate specific variants, human specific variants, and human population specific variants. Health systems' sequencing cohorts could also provide important variants.

Types of assays and systems used to assess function: Input was divided on whether to prioritize breadth/systematic testing or specialized, deeper testing when considering ways of assessing function. Some suggested a catalog of variant effects would be most useful if it systematically focused on a limited number of tractable cellular systems and a limited number of generic, highly scalable readouts because this would support data integration, benchmarking, and predictive models. Others were not sure these generic assays would be fully generalizable and argued for more variant or gene specific assays. Particularly, some thought that clinically actionable results might need disease/phenotype specific assays rather than a more generic framework. There was broad agreement that assays must take biological context into account (as detailed in the next section). It was also suggested that a metric for robustness or variance relative to context might be useful. Calibrating test results against human clinical data, when available, allows an assessment of real-world model performance. Considering the effects of drug perturbations was suggested. Finally, for diseases with unknown/incomplete etiology it may be important to test variants at different scales; cellular, organ, organism, and organism over time.

For non-coding variants, there was strong support for testing using high-throughput screens (e.g. massively parallel reporter assays (MPRA), epigenome editing (CRISPRi/CRISPRa), and base editing) in candidate cell types. Some suggested improving efficiency by systematizing a two-tiered approach using MPRAs for screening variants with "potential effects" followed by CRISPR screens to verify effects in context. Others thought testing variants in situ was more reliable than using reporter assays. There was support for saturation mutagenesis of enhancers harboring GWAS single nucleotide polymorphisms (SNPs). There was also support for a better understanding of regulatory grammar. There was support for using measurements beyond transcription for detection of the effects of non-coding variants, such as splicing, imaging, proteomics, metabolomics, pathway activity reporters, RNA stability, translation, and chromatin accessibility. The possibility was raised that perhaps there are assays that will be informative for classes of genes. Sequence modeling (using AI/ML) is likely to be important in extracting results from experimental tests. Finally, there is growing awareness that some variant effects are only apparent during changes in dynamic systems (changes in cell type, changes in cell state, developmental changes).

For coding variants, there was strong support for deep mutational scanning. The value of generic assays for variant effects, such as protein abundance or cell survival, were recognized. There was also support for more specific assays to detect variant effects, such as the ability to activate/repress transcription or cellular morphology. Some noted that our current screening methods tend to be biased toward detecting variation that disrupts gene function, and the need for approaches to better detect variants with increased or altered function was raised. A variety of assays might be most important for genes linked to more than one disease, to distinguish risk for specific diseases.

Testing Variants in Context to Assess Function

There was very strong support for large scale testing of variants in context, especially in different cell types and cell states. Discordant views were expressed about the strength of evidence that variant-by-variant interaction effects that extend beyond additive effects play an important role in human disease. There were also discordant views on the effects of genetic backgrounds, ranging from views that testing across different genetic variants might not be important to views that some variant effects on function might only be visible in particular genetic backgrounds. Different views were expressed on the relative importance of variant by cell type/cell state, variant by environment, and variant by variant tests. Finally, there was concern that we might mistakenly conclude potential genetic explanations for differences in function or organismal phenotype, when the cause may be the result of environmental effects, social determinants of health, or gene-by-environment interactions. There was support for the idea of direct testing of variants in context with study designs that would allow us to deconvolve these effects.

Goals and approaches: One short-term goal suggested in this area was to link variants to pathways, in part to infer how GWAS variant combinations may be causal for disease.

Another short-term goal suggested was a pilot project to determine whether non-additive/epistatic effects and genetic background effects are significant in human disease. This could be approached by testing a panel of variants in a panel of cell lines (perhaps induced pluripotent stem cell (iPSC) derivatives) with different genomic backgrounds. Routinely incorporating models that explicitly consider multiple functional variants in a locus would lead to reduced bias and perhaps detect a subset of variant combinations for testing. Another suggested approach was to jointly assess the function of coding and non-coding variants at the same locus and leverage information from evolutionary and functional analyses of haplotypic combinations of rare and common variants.

Context considerations: It was suggested that for some variants phenotypic effects will only be apparent when examined in the correct context. Context could be a cell type, a cell state, cell-cell interactions, a developmental state, a dynamic transition, or any combination of these. This is a challenge for experimental design and for interpreting negative results. This may be especially important for non-coding variants because regulatory effects can be cell-type specific.

For diseases with unknown/incomplete etiology it was pointed out that it will be important to test variants effects at different scales; cellular, organ, organism, and organism over time. As we do not generally know how molecular phenotypes are connected to disease, we do not necessarily know which observed phenotypes matter. It was suggested that one way to limit the number of potential tests is to consider when stimuli and other environmental effects are channeled to shared response pathways (e.g., inflammation, growth) that might be prioritized for variant testing.

Important contexts include cellular neighborhood, tissue region, age, developmental stage, sex, exposures, and disease status. While some suggested testing more variants with more read outs in more cellular contexts in one or a few genetic backgrounds would have the highest impact, others thought testing variants across a panel of iPSCs with different genetic backgrounds would have the most impact. Testing in primary cells when possible (instead of cell lines) was proposed to increase impact.

Prioritizing the search space for combinatorial variant testing: It was suggested that focused combinatorial screens should be used when there is strong support of variant-by-variant interactions from predictive models or other evidence of epistasis. Examples include variants frequently seen together in the population, common haplotypes, variants in interacting protein partners, protein residues that appear to co-evolve, and predictions from structural data. Network or cell-type information could help to prioritize combinations for testing. Variant combinations of interest could be identified (perhaps using polygenic scores) in existing iPSC lines to identify sensitized backgrounds. Patient lines might also be sensitized for disease, perhaps after correcting one or more disease causing variants. It was suggested to consider the role of nuclear architecture (Topologically Associating Domains (TADs) and enhancer-promoter loops) to prioritize variant combinations.

There was also suggestion not just to look at main effects of variant impact, but also to look at variability in variant effects (variance of variants). For example, for some variants the genetic background and environment have little or no impact, while for others background and environment have large effects. A catalog of variant-by-genetic background or variant-by-environment variance could aid translation. One could quantify how stable or variable each variant is (stability score) with respect to context, assuming a variant could have stable effects along many axes and highly variable effects along one (e.g., drug exposure).

Addressing diversity and health equity in functional interpretation: Several participants highlighted the importance of not conflating social determinants of health with genetic explanations. Difficulties in the sampling of some populations can exacerbate environmental confounds, yet we must construct studies that are meaningful and equitable. There is an urgent need for data-driven testing of when, where and how much population background (genetic ancestry and environment) matters for variant effects; empirical data are sparse. Overemphasizing the importance of genetic background risks implying that human populations are genetically very different from each other, which is not the case. Addressing these questions is critical for being able to direct future investments into variant characterization, and for nuanced data-driven discussion of politicized topics.

The importance of testing variants derived from GWAS/Whole Genome Sequence (WGS) of different genetic ancestries was emphasized. Assays that test all possible variants, rather than the biased collections that exist today, could promote equity. Focusing on variants or genes based on existing sequencing studies or databases (heavily biased towards European ancestry populations in high-income countries) will exacerbate bias.

Technology Development Needs

New or improved experimental and computational approaches that are still needed: One high level suggestion was research to identify the most useful experiments for mapping and functional genomics; for example, are the most accessible assays such as RNA-seq +ATAC-seq the best assays to use or would histone modifications add information or be better suited to use as a primary assay. Improved throughput for precise variant (prime, base editor) CRISPR screens that can be used at a scale of hundreds to thousands of variants coupled with scRNA-seq output or high-content screening was suggested. Higher throughput pairing of CRISPR and single cell multiomics (e.g., transcriptomics + chromatin accessibility) was raised. Another high-level suggestion was more efficient and more types of RNA-targeting programmable perturbations to understand all RNA species, including noncoding RNAs. More versatile assays and methods for measuring gene and variant effects at higher phenotypic scales such as cellular, tissue, and organismal phenotypes are needed. Several suggested that we need to move past molecular phenotypes and towards these higher phenotypic scales.

Spatial perturbation transcriptomics would permit study of functions such as cell-cell communication and non-cell autonomous effects. Technology development that allowed for high throughput approaches for generation of targeted structural variants, deletions, insertions, and copy number variants in specific cell types or multicellular model organisms would allow characterization of these understudied variant categories. Currently, some assays for structural variants exist, but are low throughput. More assays that are disease phenotype specific, and better ability to grow and manipulate primary cells were also raised, as was the optimization of existing assays such that they work for many different cell types and contexts with high sensitivity and at scale. Several suggested these assays be further optimized instead of continuing to develop new technologies that work in one or only a few types of cells or tissues. Additionally, several raised the importance of high throughput synthesis and assembly of big or long DNA to enable the testing of variant combinations. Current approaches are limited by size, cost, and size restrictions for getting them into cells.

Suggestions were also made for improvements in computational methods that analyze multiple types of data (e.g., chromatin accessibility, splicing, gene program data, single-cell datasets) and understanding how these data types can best be related to causal genes, pathogenic variants in rare diseases, and other disease focused features. Computational approaches to reduce complexity of interaction screens (evaluating different genetic and cellular contexts) could better identify variant by variant or variant by environment effects when they exist. Ultimately, better computational predictors of variant effects of transcription are likely to be needed to interpret the deluge of common variants. Additionally, it was suggested that improving interactions between wet-lab technology developers with computational genomicists would enable experiments and computational approaches to be developed together to improve the type of data produced for analysis.

Experimental and computational approaches that appear to work well enough to apply now for testing and analyzing variant function: There was strong support that for non-coding variants, CRISPRi/a screens followed by scRNA-seq detection, MPRA assays, and to an extent precise variant editing (base editing or prime editing), provide a significant amount of information. MPRA assays were considered an approach that should be further scaled for variant studies. For coding variants, CRISPR-based deep mutational scans/saturation genome editing are useful today. Single cell assays for histone modifications and visualization of enhancer-gene loops using Hi-C are also in practice today. Cell village approaches, allowing testing of cell-autonomous effects in pools of cells with different genetic backgrounds, and variant perturb-seq, allowing testing of variants in non-coding regions in situ, are promising and somewhat in use today. Predictive models of regulatory DNA sequence effects for variant screening, prioritization, experiment design, and extracting results are useful today, as are predictive models of proteins based on structure or evolution to prioritize coding variants for testing. The field would benefit from further cost-reductions, increased high-throughput scaling, and other improvements to these technologies.

Predictive Modeling

There was strong support for including predictive modeling efforts in variant interpretation. It was suggested that such efforts could best analyze the large-scale data by extracting maximal sequence-based information while minimizing noise. It was suggested that developing ways to improve predictions about untested variants, or untested contexts, would be important in the short term, as there are too many variants and contexts to test. It was suggested that incorporation of computational modeling could help attain the optimal experimental design, especially in iterative cycles between data collection and modeling. It was also suggested that data collection should be performed, at least in part, with the intention of creating a data set useful for predictive modeling. Suggestions were also made on how to achieve synergy and engagement between modelers and data collection.

Predictions on untested variants: Suggestions were shared on how to make and utilize predictions on untested variants. There was agreement that accurate, generalized predictions of function would be highly valuable in the context of research applications of variant interpretation. In a clinical setting the viewpoints on modeling were mixed. Some advocated for variant testing calibrated against clinical data when available. One suggestion was to focus initial studies on interpretation of non-coding cancer mutations as there was a higher likelihood of mutation recurrence and a higher likelihood of large effect variants. Large-effect developmental variants could be useful in building models incorporating functional predictions and human constraint maps. Others suggested clinicians need data, not predictions, to guide patients thus it would be better to prioritize data collection for clinical applications.

Improving experimental design: The main suggestion for using modeling to improve experimental design was to obtain insight into what to test. For example, models might predict variants to test because they are likely to have an effect, or variants to test because current models fail to make strong predictions. Models might also inform which variants should be followed up, perhaps because of uncertainty or perhaps because they might be informative in more expensive, lower throughput testing. Models could be used to estimate which read outs downstream of perturbations are most informative (e.g. chromatin, transcription, splicing, polyA). Computational strategies to reduce the complexity of interaction screens would be useful. Finally, it was noted that models that provide effect size as part of the output could be very important.

Modeling approaches: There was strong support for incorporating biophysics, thermodynamics, dynamic differential equations and other mechanistic information into models for variant interpretation. There was also support for modeling dynamics and dynamic contexts (including changes in cell types and cell states) instead of static snapshots. Learning rules and predictive models that generalize across cell types and disease was suggested as a realistic goal, that could be powered by comparative datasets across 5-10 cell types/states. Improvements in benchmarking infrastructure were suggested. Such efforts could focus on curating benchmarking datasets, developing

benchmarking pipelines, and comparing different predictive models. It was suggested a causal inference framework could be developed from functional data, functional models and statistical genetics. Foundation models were suggested as an approach to predict function. Such models could aggregate diverse data sources. However, others were concerned that much more data (function-specific and context-specific) might be required before such models performed reliably.

Synergy between modeling/data collection: Approaches were suggested to improve synergy between people engaged in modeling and data collection. There was agreement that iterative loops between experimental data collection, AI/ML based modeling, then testing model predictions with targeted experimental data generation would be the most synergistic approach. One idea was to interest more modelers in the problem with competitions and targeted funding opportunities. Another idea was to offer curated perturbation datasets in a standardized format that is easy to use. The need for more CRISPRi data for modeling was raised. It was suggested models should be made easily accessible to researchers, and resources should be provided to educate researchers on their implementation and use.

Analyzing data: Several suggestions were shared on how to use modeling to analyze data on variant effects. One overall goal proposed was to learn how sequences encode each layer of molecular activity associated with gene and protein regulation in all major cell types. An immediate need is better models of variants effects on regulatory elements and models of which sequence features determine regulatory element activity. A corollary is better understanding of the syntax of non-coding regulatory regions. One suggestion was that we might learn how regulatory function relates to current proxies such as conservation and epigenetic marks. In support of these needs, it was suggested to build tools to predict the TFs active in most cell types. If successful, such models could be used to connect variants to elements to genes to pathways.

Resource Building

Well-integrated and standardized data resources and knowledgebases: There was strong support for a catalog that integrates genetic variation with key aspects of gene function such as gene expression, protein function, and molecular networks. The foundation for this resource could be large collections of epigenomic, transcriptomic and perturbation datasets (ATAC-seq, RNA-seq, CRISPRi, MPRA, and combinations of these) that can be used for both computational analyses and experimental design. Well-curated and standardized data (i.e., standardized assays, standardized processing, rich and standardized metadata) would be necessary to provide variant effects and predictions. To capitalize on the amount of work being done in this area, we need to learn how to integrate information (potentially superficially discordant) from disparate assays and systems. Data resources should describe the context for variant tests and predictions (cell type, cell state, environmental perturbation, or interactions with other cell types). Several pointed out the need for more harmonized data generation efforts for the building of data

resources. Unified knowledgebases and knowledge graphs that deduce relationships across biological landscape from data, literature, and models would provide important contributions to usability and provide a platform for the storage of hypotheses extracted from models.

Infrastructure needs: Infrastructure requirements were suggested for a resource in this area. There is an urgent need for standards for genome editing, epigenome editing, and reporter assays (CRISPR, MPRA). These should range from protocols to metadata to pre-processing, to analysis, and should include consensus pipelines for analyses and computational predictions. Several suggested that more effective visualization tools for genome function and variant interpretation would improve usability.

Other infrastructure suggestions included sharing tools and other generated resources with the community to facilitate work performed by others. Centralized and easily accessible repositories of several hundred iPSC lines were noted as a potential resource by several participants. Participants suggested that these would be most valuable if they represented different genetic backgrounds, were derived from people with different diseases, or stably expressed CRISPR effectors or other engineered mutations. Ideally, for such a resource each line should have T2T genome sequencing, ATAC-seq and RNA-seq, and information about differentiation potential. Participants also noted the importance of improved access to diverse large-scale biobanks. Sharing tools to interconvert among widely used formats could improve the efficiency of using resources. Community metadata formats and standards would accelerate science; this was seen as a potential role for the National Library of Medicine. Centralized model zoos (collection of pre-trained machine learning models) for genomic predictive models could also assist in work being done by the community. Ultimately, infrastructure is needed that enables the use of genetic catalogs and other genetic resources in a federated fashion.

Coordinated Approaches: Guiding principles for building the resource were also suggested. For consortium generated resources, there should be coordination from the start between consortia and community stakeholders, such as scientists outside of consortia doing similar work as well as those who might be downstream users of the resource. Aggregating and integrating community data could be powerful, perhaps using Global Alliance for Genomics and Health (GA4GH) metadata standards. Consortia reagents, such as perturbation libraries, could be shared across labs and tested in many different contexts. Finally, it was suggested that federal funding going towards variant interpretation that is comparable to the funds spent in the past on large-scale sequencing studies would be appropriate for this area.

Existing resources that could be leveraged or repurposed: Participants suggested several resources that could be leveraged or repurposed for the benefit of functional variant interpretation studies. Reasons for these suggestions varied and ranged from content that could contribute to variant studies to resource structures that could be copied for the design of variant data resources. These include:

- The Genotype-Tissue Expression (GTEx) Project
- The Encyclopedia of DNA Elements (ENCODE) Project (project description)
- PsychENCODE Knowledge Portal
- AD (Alzheimer's Disease) Knowledge Portal
- All of Us Research Program
- International Mouse Phenotyping Consortium (IMPC)
- Open Targets (platform)
- MaveDB
- MPRABase
- NHLBI Next Generation iPSC collection, WiCell
- Dependency Map (DepMap) Project (interactive portal)

Additional Topics

Suggestions about high-level scientific goals for variant interpretation: Participants were asked to discuss ideas that were not covered in the topics above. Items raised included:

- Prioritize genome function studies i) that can be performed comprehensively, as a resource across disease categories; (ii) where comprehensive maps will accelerate process of learning mechanisms of disease; (iii) that require large collaborative, coordinated effort.
- While each type of variation (coding, non-coding, structural variants) is different a basic systematic approach should be adopted for each. Currently, progress is slowed because different types of variation are being studied by different groups without a sense of common purpose or vision. This is the biggest gap/challenge.
- Disease progression and treatment response questions remain very poorly understood genetically and in terms of biomarkers. By contrast, the current focus appears to be lifetime risk of a disease.
- The relative importance of studying variant effects or gene effects is not understood at this time. Perhaps characterization of gene dosage/function relationships is an important axis for study. Loss of function variants like PCSK9 already have clinical utility and many are searching for the next similar story.

Suggestions about approach taken by NIH: Some suggested that NIH should provide incentives to scientists that are good citizens, i.e., sharing data, sharing software, sharing protocols, sharing reagents and participating in multidisciplinary, collaborative research. These incentives could be resources, funding and/or recognition. Participants also raised that NIH should also support infrastructure development in research grants. When asked about implementation of the above, some suggestions included developing specific review criteria and funding opportunities to support sharing and collaborative research. Some suggested that more collaboration is needed across NIH institutes and centers (IC) for study of disease variants. Institutes with domain-specific knowledge should collaborate with each other and with generalist ICs to understand the context of variants that lead to

disease. It was noted that while NIH appears to be moving to the cloud, the scientific community is not. Academia functions on fixed cost, while cloud providers try to maximize profit from user inefficiency; could NIH broker a fixed cost model with cloud providers?

Conclusion

Overall, participants provided valuable community input, addressing current knowledge in this area, as well as key gaps and opportunities. The areas that were most consistently highlighted in individual feedback and group discussions are featured in Box 1. Together these points indicate a continued need for investment in high-throughput variant testing, predictions of variant effects, and a resource for sharing these results. Advancing technology and predictive models would also enable large-scale data generation efforts, as would improve synergy between data generation and modeling efforts. There was also consensus that this area would benefit from the development of standards and highly integrated, well-curated data resources to provide variant effects and predictions as well as enable the development of new research directions.

Appendix 1: Organization

Format:

- Written opinions and virtual panel discussion
- NHGRI received input from 22external scientists.
- NHGRI confirmed they consent to having anonymized comments/discussion shared on public-facing NHGRI website.
- External scientists returned written responses.
- NHGRI shared list of anonymized answers grouped by topic with the group of external scientists.
- NHGRI convened 4 discussion panels, ~2 hours in length, each with about 1/4 of the external scientists. Focused discussion on surprising or divergent viewpoints.
- Feedback collected and summarized by NHGRI.
- Feedback posted on NHGRI web site.
- Expertise: Computational genomics, technology development (HT functional assays, analysis tools), genomic variation, population genetics, rare-disease genetics; across career stages; users, producers and developers

Rationale for Format:

Participants were asked to share written opinions in advance, to ensure individuals consider each topic, and to ensure everyone's response was heard.
Written opinions were shared prior to discussion so participants could consider new information that might change their thinking, and so they could consider opinions that might not arise in their discussion session.
Participants were asked to focus discussion especially on opinions that were most important, surprising, or divergent, as well as the strengths and weaknesses of the ideas to add value to the written opinions.
Four discussion sections were held to attempt to limit "dominant voices" during discussion.
Four discussion sections were held to try to make sure there was time to hear from everyone.

Appendix 2: External Participant List

Christine R. Beck
University of Connecticut Health
The Jackson Laboratory for Genomic Medicine

Michael Beer
Johns Hopkins University

Guillaume Bourque
McGill University

Alessandra Chesi
University of Pennsylvania

Judy H. Cho
Mt Sinai School of Medicine

Nancy J. Cox
Vanderbilt University

Barbara Engelhardt
Gladstone Institutes
Stanford University

Jesse Engreitz
Stanford University

Douglas Fowler
University of Washington

Gamze Gürsoy
Columbia University
New York Genome Center

Danwei Huangfu
Memorial Sloan Kettering Cancer Center

Anshul Kundaje
Stanford University

Tuuli Lappalainen
New York Genome Center

KTH Royal Institute of Technology

Francesca Luca
Wayne State University

Prashant Mali
University of California San Diego

Mark McCarthy
Genentech

Kiran Musunuru
University of Pennsylvania

Steven Reilly
Yale University

Kaitlin E. Samocha
Harvard University
Massachusetts General Hospital
Broad Institute

Neville Sanjana
New York University
New York Genome Center

Gosia Trynka
Wellcome Sanger Institute
Open Targets

Jimmie Chun Ye
Gladstone Institute
University of California San Francisco

Appendix 3: Written Input Guidance

Overview of Ideation phase:

Participants (~25 total with varied expertise) will share their thoughts via email regarding the most critical scientific challenges, opportunities, and technology needs in the area of functional variant interpretation. Please see the attached word file with details on the type of information we are requesting. If you participate, your responses will be anonymized, grouped by topic, shared with other participants, and may be shared in a public-facing report on the NHGRI web site.

Topics (word file attached to email)

NHGRI is seeking community input about future research directions towards understanding the effects of variants on function. Through these sessions, we hope to identify gaps, challenges, and opportunities on both short term (2-7 years from now) and long term (7-15 years from now) timeframes. This input will help inform future NHGRI projects that would address these opportunities.

- Overarching or background question- given limited resources, what are critical activities that would facilitate short term and long term scientific advances in this area with an eye towards making clinical advances?
- Short term (2-7 years from now): Obtain better understanding of how human variants affect traits and diseases, especially in areas that today are at the limit of our understanding.
- Long term (7-15 years from now): Implement these scientific advances in medical practice (e.g. diagnosis, treatment).

To prepare for these sessions, we are requesting your thoughts regarding the most critical scientific challenges, opportunities, and technology needs to achieve these goals. We would especially like to hear your thoughts on the topics listed below. (Note, we are asking your feedback on these broad areas of science, as opposed to your work in particular). As you organize your responses, please consider where there are current gaps and challenges, as well as activities underway that are contributing to progress in this field. Please provide the rationale, whenever possible, for your ideas and suggestions.

Testing and Assessing Function of Individual Variants

- Considerations for the type of variants tested and the types of assays used to assess function.
- How to balance comprehensiveness and complexity (e.g. balance of rare and common variants; coding and non-coding variants; generalizable vs disease/phenotype specific assays).

Testing Variants in Context to Assess Function

- Approaches for considering combinations of variants, and when such approaches would be appropriate.

- What types of contexts to consider (cellular backgrounds, genomic backgrounds, environmental perturbations), and how to balance number of variants tested and the number of contexts explored.
- How to best address diversity and health equity in variant prioritization and functional interpretation.

Technology Development Needs

- New or improved wet lab/experimental or computational approaches that are specifically needed for testing and analyzing variant function.
- Approaches that work well enough to apply at scale now.

Predictive Modeling

- Questions or topics that are best addressed via predictive modeling.
- Stimulating synergy between data needs for improving model development and the use of models to improve experimental design.

Resource Building

- Needed resources to accelerate community progress in this area. This could include data, models, standards, protocols, approaches, or other needed resources.
- Existing resources that could be leveraged or repurposed.

Additional Topics

- Please share any ideas that are not included in the topics above.

Appendix 4: Discussion Session Guidance

Discussion phase: March-May 2024.
We will host three two-hour small group virtual discussion sessions in March-May 2024. We will do our best to schedule sessions based on participant availability, with everyone expected to attend one session. A summary of these discussions will be shared as a public-facing report on the NHGRI web site.

Thank you for your efforts to date. They have been very helpful to our NHGRI planning. Community input is very important to the NIH, and to NHGRI. We are near the end; what remains for you is an opportunity to discuss your views with your peers.

Attached please find the word file (not Included in this report). We have attempted to summarize the written comments we received as anonymized bullets grouped by topics. We have highlighted some ideas; these ideas appeared to be either superficially discordant, suggested by a few participants, or ideas we'd like to hear more about.

Some guidance for the upcoming discussion sessions. We want to know what large-scale science is needed in this area to move the field forward. We want to hear your thoughts on what would help the field as a whole; we are asking you to be a representative for the field. That being said, you don't have to agree or reach consensus; honest disagreement about scientific judgement is expected. Especially where you disagree, hearing your rationale is very important to us.

Please prepare for discussion by reading through the summaries and considering what points you think are most important to discuss. Your "most important points" could be ideas you have already raised, ideas raised by others, or new ideas that are not represented in the summary. Likely we will allot some time for each of the 6 topics and moderate your discussion to try to hear from as many participants about as many ideas as possible. As these discussions are for community input (listening), generally we will refrain from responding to your ideas.

Appendix 5: Acknowledgements

NHGRI
Carolyn Hutter
Afia Asare
Sarah Anstice
Dan Gilchrist
Stephanie Morris
Mike Pazin