Rapporteurs: Elian Silverman
and Anne Niehaus
Program Analysts, NHGRI

NHGRI Genomic Medicine IX: NHGRI's Genomic Medicine Portfolio – Bedside to Bench
April 19-20, Silver Spring, MD

Session 1: Introduction and Background

Welcome, Introductions, and Goals of the Meeting

NHGRI's genomic medicine meetings bring together key individuals and institutions involved in genomic medicine. These meetings began in 2011 and have led to the development of many of NHGRI's Genomic Medicine programs and efforts, including ClinGen, eMERGE-PGx, and IGNITE, as well as enhanced partnerships with international groups and additional stakeholders such as the Centers for Medicare and Medicaid Services, the Food and Drug Administration, and the National Academy of Medicine.

NHGRI's research portfolio touches on many aspects of genomic medicine, from the basic sciences and technology development that support clinically relevant discoveries and applications to patient-focused research and improving clinical care. However, difficulty remains in translating basic science research to clinical research, and in bringing clinical discoveries back to the bench for investigation. This meeting focused on one particular challenge within the translational gap—characterizing and interpreting the clinical significance of variants of uncertain significance (VUS).

Convincing Clinicians to Use Functionalized Genomic Information (Howard Jacob)

A case study of a patient with chronic kidney disease and a VUS in *SHROOM3* highlighted the challenge of demonstrating causality. Eleven GWAS have reported *SHROOM3* variants as being associated with markers of chronic kidney disease. QTL mapping with rat models showed that multiple variants within *SHROOM3* were predicted to have functional effects on the protein. Zebrafish assays aided in the discovery of variants with a causal link to renal failure. While these data seemed to implicate *SHROOM3* as causing this patient's disease, few clinicians were ready to cite it as such in a patient's medical record. Yet it was not clear what additional evidence would make this link more convincing to clinicians.

This case study was used to foster discussions of the level and types of evidence needed to prompt a clinician to implicate a gene as causal in the medical record of a particular patient. Randomized trials, the gold standard of evidence for many clinicians, cannot be conducted on all potentially pathogenic variants. A new focus is needed on establishing criteria *upfront* to create thresholds for proving causality with functional and experimental data. There are many additional types of data, such as allele frequency and segregation data, that are needed to interpret these functional data. Without these additional evidence types, clinicians may not feel that the functional data are convincing enough to classify a variant as causal. These challenges are heightened for multifactorial common diseases, such as chronic kidney disease, because genetics is only one of the many potentially casual factors.

Magnitude of the Problem – Basic Science Perspective on Need for Integration (Monte Westerfield)

Two case studies were presented to illustrate successful collaborations between basic scientists and clinicians. The first demonstrated how positive results can help validate candidate genes and enhance gene discovery for Usher syndrome. Usher syndrome is the leading cause of deaf-blindness, and 11 genes have been shown to be involved through a classic Mendelian recessive inheritance pattern. In a large patient cohort of Usher syndrome patients, exome sequencing identified mutations in *PDZD7*, a gene of unknown function. The ortholog was cloned in zebrafish to demonstrate causality, demonstrating that *PDZD7* co-localized in the retina and inner ear, similar to other known Usher syndrome proteins, and thus could lead to retinal degeneration. Segregation data supported the role of *PDZD7* as a genetic modifier, leading to *PDZD7* being added to the Usher syndrome testing panel.

The second case study demonstrated how negative results can reveal incorrect diagnoses. Mapping homozygosity by descent in a consanguineous family with deafness failed to identify gene candidates. Subsequent whole exome sequencing (WES) for homozygous SNPs identified a mutation in *AHI1*, a well-known gene responsible for Joubert syndrome, but although these patients were deaf they had no other Joubert syndrome abnormalities. Zebrafish models demonstrated that the mutation blocks expression and produces a Joubert-like syndrome. Subsequent segregation analysis in the family revealed a sibling who was homozygous for this allele but not deaf. Thus, this variant was ruled not associated with Joubert syndrome—illustrating the importance of properly interpreting exome data and pairing sequence analysis with clinical evaluation and follow-up studies.

While both these case studies demonstrate successful collaborations between the 'bench and bedside,' they also reveal gaps. NIH's Undiagnosed Disease Network (UDN) was highlighted as a model for moving forward. The UDN includes seven clinical sites, a coordinating center, two DNA sequencing cores, a metabolomics core, a model organism screening center, and a central biorepository. All of these groups work together to diagnose patients and validate VUSs and all have access to clinical and functional data, but there are only two basic labs with full IRB-approved access to UDN clinical records. Making these data available to larger numbers of basic labs would speed discovery and understanding. A broader network of geneticists could be developed based on the UDN model, otherwise interactions tend to be haphazard and one-off. There are significant barriers to investigators' accessing clinical data and not all are regulatory or HIPAA-related; some appear to be cultural or related to pressures to publish and lead. The field should think about innovative ways to gather rich phenotypic information from patients and to include these descriptions within the medical record.

Magnitude of the Problem – Clinical Perspective on Need for Integration (Gail Herman)

Much of the clinical utility achieved through genomics thus far has been in rare, Mendelian disorders and cancer diagnosis. In addition to having a high diagnostic yield, WES provides a unique opportunity to identify actionable secondary findings. In 2013, ACMG published a "minimum list" of 56 actionable genes and specific mutations that should be returned to patients regardless of their indication for clinical sequencing. Although clinicians should aim to order the most specific genetic test possible, there are some conditions in which new potentially causal genes are routinely discovered and WES is more appropriate and cost-effective. Clinical exome sequencing has revolutionized clinical genetics as it has a higher diagnostic yield than anything previously used. At Nationwide Children's, all WES must be must be approved by a clinical geneticist, which promotes communication with medical specialists. They have had a 44% diagnostic yield, and nearly half the diagnoses resulted in a change in management beyond reproductive risk; three led to identification of novel genes. Mutations were *de novo* in nearly half the diagnosed cases, which is critical when counseling families as the recurrence risk of these is under 1%.

The 2013 ACMG/AMP Interpreting Sequence Variants paper (Richards *et al*.) helped standardize the process for classifying variants and includes a detailed scheme for how laboratories should grade evidence. A variation of the ClinGen Pathogenicity Calculator can help grade each of these evidence categories and come to a final pathogenicity call. It is the responsibility of the *laboratory director* to integrate this functional information into the report. Then, the *clinician* with his/her clinical expertise can come to a conclusion for the patient.

Discussion

Clinical genetics should begin to adopt more of a Bayesian-inference approach. While randomized trials are ideal, they are not possible for all genes. Thus, we need to compile the evidence, consider prior and

conditional probabilities, and come to a clinical decision. There is a distinction between labs weighing evidence to assert pathogenicity and clinicians weighing evidence to diagnose a patient.

The level of evidence for a variant depends on the case and clinical context. Even if there are no actions taken as a result the variant classification, evidence should be captured in the medical record. The community will have to accept that once information is put in the medical record, there is a chance it will be misinterpreted. One can ask oneself whether a clinician is more likely to make a wrong decision or a right one with the additional information, and decide accordingly on placing it in the record. While clinicians are familiar with ambiguous findings and the need to place one piece of data in an entire clinical context, genetic data are frequently over-interpreted as being definitive. The field needs to do a better job of conveying the uncertainty. Infrastructure is needed to integrate genetic and clinical information. A genetics community is needed in every medical specialty so a nephrologist, for example, can turn to a colleague and request an interpretation. It would also be helpful to encourage a centralized way to collect outcomes of genetic diagnostic searches. The ClinGen frameworks incorporate the functional data for a specific gene-disease relationship and pair this information with case-control data and other clinical data. It would also be helpful for journals to require authors to submit their functional data into a database.

Specific guidelines to dictate the frequency or depth of variant re-evaluation do not exist. ClinGen is developing a Maintenance of Certification (MOC) module for clinicians to go back and re-evaluate five cases in which a VUS was reported on a patient. However, most clinicians do not have time to complete this re-assessment in a routine or systematic manner.

Session 2: Vexing Clinical Problems Needing Basic Input

Speeding Functional Assessment to Benefit Patients (Stephen Kingsmore)

Implementing genomics in the 47 level III NICUs and PICUs across California is estimated to be favorable economically and would address the state's entire level III NICU/PICU population of roughly 30,000 children per year. Implementing genome and exome sequencing on this scale would be a comparatively small cost in the care of critically ill children. Challenges to be faced in large scale implementation include that one may have cost-effectiveness or speed in sequencing, but not both. Gaps to be bridged include early identifications of patients in need of sequencing through predictive modeling, similar to existing algorithms for assessing patients for sepsis. The shortage of genetic counselors will continue to impede genomic medicine implementation. There is also a need for improvement of EHR features and data-driven models of genetic disease topologies. Lastly, there remains a variant interpretation bottleneck. To better implement genomic medicine in the NICU/PICU populations and others, large data sets and modeling will be needed to build predictive tools.

***De Novo* Variants that Inform Clinical Phenotypes (Kricket Seidman)**

Congenital heart disease (CHD) is common among children and 9% of the children born with CHD have a profoundly severe congenital heart defect. Prognosis is compromised by associated lifelong issues such as neurodegenerative disorder (NDD). The severity of the CHD correlates with the NDD, suggesting they are linked. NHLBI's Pediatric Cardiac Genomics Consortium (PCGC) of 10,000 children includes about 1200 patients with severe critical CHD and their parents, which allows searching for damaging *de novo* mutations through WES of the trio.

Damaging *de novo* mutations were enriched in children with CHD compared to a control population, and were more enriched when limited to known CHD genes. Recurrent damaging mutations were found in

21 of the genes identified. Some of the identified variants were associated with known CHD genes and others with syndromic genes, but variants were also identified in genes not previously implicated in CHD. Significant enrichment in pathways related to NDD, CHD, and chromatin modification was found. To determine whether a *de novo* mutation is causative in a single patient one needs to know when in development the gene is expressed, not just that it's expressed in the heart. This can be done by studying human iPSC-derived cardiomyocytes and defining cardiac developmental transcription in single cells. Measuring force of contraction with various mutations edited in makes a valuable functional assay.

Caution is needed in diagnosing single patients using *de novo* mutations because there are clearly other factors that affect phenotype. To address these issues, there should be focus on better models to functionally annotate CHD mutations. We could start to think about modeling human mutations in isogenic iPSC-derived cells similar to this work on cardiomyocytes.

Discussion

The ClinGen Dosage-Sensitivity Map is a useful tool because certain genes are more intolerant of missense variation. In many genes there is no significant difference between synonymous and missense mutations but in others, there is an extreme intolerance of missense mutations. There are 2-3,000 genes with almost no loss of function (LoF) mutations in large population databases; many are enriched for CHD, and for 80% their function is unknown. Specific protein domains are important for functional characterization as well; LoF variants in some parts of the titin gene have no effect at all.

De novo mutations are a driver of disease irrespective of phenotype in a NICU environment. Having a phenotype or diagnosis to pursue makes it easier to be more confident in causal inferences, while function is harder to infer in healthy individuals. Rapid, cost-effective results would make exome sequencing more useful clinically. If the causal variant is discovered, there has to be a nuanced interpretation for what it means to that particular child.

Concern was expressed about rising uptake of direct-to-consumer offerings, particularly in pregnancy. The importance of keeping genetic testing from escaping the context of medical care is a key reason for the scientific community to accelerate efforts in large-scale implementation of genomic medicine. The field might learn from the success of private companies at communicating results to patients. To create a standard assessment, all of the variants in ClinVar could be run through functional assays and data could be aggregated. Defining a minimal set of assays would also be helpful. It could also be useful to have a "functionizer," analogous to a "Phenomizer," through Matchmaker Exchange.

Large-scale implementation is not possible in all settings because genomic medicine is not yet applicable everywhere, so at present large-scale implementation needs to be done in research settings. While the field has been successful in defining disease genes, there is a huge gap between this knowledge and broad implementation into hospitals. Community physicians are not yet ready to interpret exomes and need to be able to hand them off to those who can interpret them. Eventually the process needs almost to be "shrink-wrapped" to something deployable to any hospital.

Session 3: From Variant to Disease Mechanisms - Specific Examples of How Model Systems have used Genetic / Genomic Approaches to Lend Insight into Human Disease that had Clinical Relevance

Integrating Model Organism Data around Clinical Genomics (Calum MacRae)

Case studies have shown that clinical evaluation and genetic testing have had limited predictive value, in large part due to phenotypic heterogeneity and other limitations of phenotypes such as their focus on

morphology, semi-subjective nature, late- or even end-stage appearance, and need for (and frequent lack of) provocative stimuli. In addition to these challenges, clinical genetics has focused on extreme phenotypes, lacked prospective cohorts, and failed to adequately incorporate selection pressures and environmental contributions. A major limitation of genetic studies is not knowing who is unaffected because the resolution of phenotyping is so low, particularly for phenotypes presenting later in life. Limited dimensionality of phenotyping with a deliberate reduction in complexity is another problem, contrasting the 10 to the 9^{th} (theoretically) genomic variants, 10 to the 16^{th} transcripts, etc., with only 10 to the 4^{th} recognized and measured phenotypes (=total billable tests and their clinical components).

Model organism studies have the potential to identify many gaps in our current genetic or phenotypic architecture. For example, saturation screens can be used to identify all or most of the genes contributing to a given trait and reverse genetics can be used to manipulate each gene and explore the resulting phenotypic effects. In addition to harnessing the model organism studies for scalable parallel phenotyping, clinical genetics needs to move towards 'next generation phenotyping' involving patient-entered data, symptom ontologies, structured information, therapeutic responses, rigorous probability estimates on a population scale, and comprehensive multi-scale dynamic phenotyping, all integrated into clinical care. A shared lexicon for translation and a minimal clinical dataset to maximize information content would be crucial in complementing clinical care and uniting it with genomic discovery.

Leveraging Congenital Heart Disease Mouse Model Findings to Improve Clinical Outcomes (Cecilia Lo)

Intrinsic factors play a significant role in determining the long-term outcomes of patients with CHD, but genetic factors lead to differential outcomes. A large-scale, forward genetic screen provided a non-biased way to identify genes that contribute to CHD. Noninvasive phenotyping of 100,000 murine fetuses produced 300 mutant mouse lines with a wide range of CHD phenotypes. Detailed curation of the phenotypes and WES sequence analysis revealed 147 pathogenic mutations in 98 genes, leading to the discovery of 47 novel CHD genes.

This work demonstrates the potential of systems genetics with mutagenesis to investigate the complex genetics of human disease. Selecting an animal model with (a) similar anatomy/physiology of the human disease of study, (b) available inbred strains to adequately assess genetic heterogeneity, and (c) a phenotype ontology that parallels the human phenotype ontology is critical. Public databases are needed to disseminate these rich phenotype and genotype data.

Discussion

Ontologies should be improved by standardizing traits that are routinely measured, gathering rich patient-entered phenotypic information, and analyzing existing genotypic and phenotypic datasets with machine learning algorithms. In addition to standardizing the elements gathered in EHRs, we need to gather more longitudinal information. With enough EHR data, unsupervised machine learning algorithms could be used to mine the data and look for new phenotypes. Current data mining efforts within EHRs are not in parallel with genomic efforts or large enough in scale to 'curate the Phenome.'

Model organism studies allow the exploration of biomarkers that could replace or supplement an original human biomarker. To facilitate the translation of basic science research to clinical improvements, scalable ontologies that allow for systematic cross-referencing are necessary. EHRs should have platforms that allow busy clinicians to enter phenotypic information easily, as well as terminologies that don't rely on billing codes. Additionally, we should develop ways for patients to review and add information into their EHR, as well as structured ways to capture family history, drug responses, socioeconomic and cultural factors, and environmental/exposure data.

The current phenotypic-based taxonomy of disease and phenotypic-ascertainment approach should be supplemented with genotypic-ascertainment studies. This approach, paired with family history information, would help us learn what phenotypes are associated with variants in an unbiased manner. In addition to moving towards 'genome-first' research and longitudinal phenotyping, the field should encourage sequencing and phenotyping healthy individuals. Studies with large control populations are often underfunded and are crucial for understanding the genetic drivers of disease. GM10 could focus on how to facilitate this research and define a 'healthy genome.' Given the substantial interest from patients and the private sector, it would be prudent to involve these individuals at GM10.

Session 4: Computational Approaches to Variant Function Prediction Methods for Predicting Functional Consequences of Variants

Leveraging Massive-Scale Databases of Human Genetic Variation (Daniel MacArthur)

The Exome Aggregation Consortium (ExAC) was initiated in 2013 to aggregate exome data to benefit the broader rare disease community. The first release started with raw data from 92,000 patient germline exomes aggregated and called together, producing a final call set of over 60,000 high quality and consented samples. ExAC has high ancestral diversity and is the largest catalog of protein coding genetic information of over 10 million variants with most being novel. Frequency data are publicly available on all discovered variants. Caveats include that ExAC is essentially a convenience sample, has limited phenotype data, and has few samples that have been re-consented for re-contact.

ExAC improves VUS analysis by filtering variants that are too common to be causal, assessing penetrance by comparing to large case series, and identifying genes that are depleted for specific classes of variation (particularly LoF variants). Over 2600 genes severely depleted in LoF are currently of unknown function and would be excellent targets for functional studies. Hundreds of seemingly healthy individuals in ExAC carry dominant, severe disease-causing variants. This could be due to false positive assertions of pathogenicity, undiagnosed disease, somatic mosaicism, and incomplete penetrance.

Comparing ExAC frequencies to over 10K prion disease cases provided a model to assess incomplete penetrance. Prion disease variants are 30 times more common than expected based on the prevalence of prion disease phenotypes. Penetrance can be quantitatively assessed by comparing frequencies in cases and controls. ExAC can go beyond the gene level to identify variation constrained within specific gene regions. Next steps for ExAC include increasing the number of samples (120K projected for next release), moving to genomes (20K projected this year), and performing genotype-based recall of consented participants. Needed resources include: larger, harmonized, centralized repositories of variants linked ethically to phenotypes; regulatory support for data aggregation and reuse; increased focus on sequencing samples that are consented for recontact, deeper phenotyping and data sharing; and large, uniformly ascertained rare disease case series to assess penetrance.

Empowering Variant Effect Prediction with Large Scale Mutagenesis Data (Douglas Fowler)

While it can be effective to use a model system to interrogate a gene variant of particular interest, the sequence space is vast with a typical protein having over 7,000 possible amino acid exchanges. An alternative is to test large numbers of single mutations in a gene simultaneously through deep mutational scanning and assessment of protein function. Starting with a coding sequence of interest, a library of mutations is produced and inserted into the model organism, and a selection function (or series of selections) for the protein of interest is imposed. Variants with the selected properties will be enriched and variants without them depleted. Deep sequencing can then be applied to capture the

frequency of each variant after selection and enrichment based on function, or a functional score, can be calculated, producing a sequence-function map of all (or most) possible mutations. Models for predicting disease based on function can be trained on variants with known disease-causing effects. Unfortunately these methods are not currently scalable to hundreds or thousands of genes.

Functional consequences of mutations generated by deep mutational scanning are assessed by a variety of tools to predict variant effect but these have known limitations, especially when used to predict disease. Predicting the effect of a variant on a protein or protein domain is a more constrained and proximate problem and so may be more accurate, and may also allow assessment of activity-enhancing variants. Combining multiple tools into a single regression model ("Envision") improved the variance explained in training sets of six proteins and was modestly generalizable to a seventh test-set; this should improve as more proteins are added. To predict the categorical effects of mutations, one can discretize functional scores and classify damaging mutations in test-set proteins with reasonable accuracy that is higher the more similar the test protein is to the training set.

The "Envision" predictor compared favorably with other predictors in biochemical effects though it wasn't as good as ClinVar for pathogenicity predictions, but combining predictions of biochemical effect and pathogenicity can improve accuracy. Additionally, Envision can be trained to predict function-enhancing mutations, probably by predicting stabilizing mutations. The overall goal is to build a predictor model that accurately predicts protein variant effects for any variant in any protein, using generalized assays for molecular and cellular phenotyping.

Discussion

Identifying constrained non-coding regions in the whole genome version of ExAC will be very difficult, requiring both the constrained non-coding region and the population size to be very large. With the current 5.5K whole genomes, ExAC can identify non-coding variants above 0.1% frequency which is useful but not optimal. Joint calling of all these genomes is needed, and current exome-based calling algorithms may need to be optimized for genomes. NHGRI-funded programs could contribute more value at present by obtaining much larger number of exomes rather than small numbers of whole genomes. Regulatory guidance is also needed on what can be done with European and other international samples. Better approaches for linking variants to phenotypes are needed, particularly a way to push out phenotype data for a given set of variants given the right cohort and consent.

In returning information to a patient with a VUS, three assessments are needed: 1) Has the variant ever been seen before in a collection like ExAC and how common is it? 2) Has it been seen before in disease patients? 3) If the variant is novel, what functional impact has been identified through tabulated high throughput functional assays such as Envision? Capturing feedback from the end-user is essential for improving the tools and better aligning basic efforts to functionalize variants with high priority clinical variants. A more systematic effort to identify clinical need across all genes is needed; modeling impact of VUS is less of a problem once one knows what the gene does, but variants in genes of unknown function are difficult to model. Prioritizing such genes could be based on medical actionability of variants, which ClinGen is addressing. ClinGen should map the function of a gene and the authoritative evidence that links it to a disease. ClinGen is evaluating evidence in the literature but more input from basic scientists would be helpful in generating the evidence to be incorporated. Reference populations are also challenging; Middle Eastern populations are almost totally missing from reference populations but are vastly over-represented in Mendelian disease studies due to consanguinity.

Panel 1

Many cohorts planned for sequencing are phenotyped in limited axes— the PCGC, for example, has focused on cardiac phenotyping but includes clinical centers with smaller collections of patients more deeply phenotyped for non-cardiac traits. The Common Disease Centers may also be emphasizing patients with deep phenotyping in multiple domains.

A significant challenge is the need for participants to be re-phenotyped often and deeply. IPS cells provide the potential for generating patient-specific organoids, which would allow for the evolving of rich cellular and tissue phenotypes. Liver and adipose tissues removed during bariatric surgery could be made available with the correct consenting. The PMI consent hasn't been written yet but it would be good for investigators such as those gathered in GM9 to weigh in on its design.

Examples of quantifying evidence include the "inferential distance" of a phenotype from clinical disease, and estimates of how impenetrant a variant would have to be to be causal in datasets such as ExAC. Empirical null models of many variants and models across genes will give an empirical sense of the probability of a given observation, such as comparing it to tests across several millions or even billions of genomic variants. High-throughput functional assays that are meaningfully scalable, such as whether transcription is up- or down-regulated or protein structure is affected, would be very useful; a generic quantifiable pre-generated resource such as the Envision tool and resource will enable estimation not only of a specific variant's effect but also discovery of variant effects not previously hypothesized. Even currently imperfect estimates of how well a protein is working with or without a certain variant can be used to modify estimates of variant pathogenicity. A model in which basic science labs determine the basic principles and are then connected to relevant phenotypes would enable design of high-throughput assays to understand the function of as many variants as possible, similar to high-throughput drug screens. This would facilitate assessing pathogenicity in days, as is critical for clinical care, rather than the years currently needed to design the models and grow the model organisms, etc.

Genomics often falls prey to the "Nirvana Fallacy," or comparing what is actually happening to what could be happening in the field, but clinically one should focus on comparing the decision that would be made in the absence of variant information with the decision made with it. In Bayesian terms, added information improves the accuracy of decision-making, especially in clinical settings where there may be nothing else to go on. Currently, knowledge often doesn't go beyond an individual institution or even lab, but efforts to include basic scientists in clinical forums such as tumor boards can be very helpful in assessing what a particular VUS may be doing in a specific patient. Scaling of clinical decision-making is needed with the increasing availability of ever-richer characterizations of a patient's genome. Genomics as a field also fixates on using variant data to make a definitive diagnosis, in contrast to considering it a biomarker which is recognized to improve decision-making but not necessarily to be deterministic *per se*. This is another legacy of exceptionalism, expecting genetic predictions to be perfect, which is not expected of other clinical fields. As a result, papers published in genomics describing uncertainty are used by insurance groups to demonstrate why they do not want to pay for genomics.

Clinicians and basic scientists need to come together to determine the most useful large-scale datasets for interpreting a VUS. ENCODE data are useful for understanding gene function but cannot be used to interpret a VUS directly; they may however be informative in combination with other databases. While there are limitations to testing variants in iPS cells, there will be organoids that mimic tissue and can be tested in microfluidic systems under different physiological conditions. The new round of ENCODE is not only amplifying a catalog of potential regulatory signals in particular tissues but working to do that in

patients with diseases, as well as developing methods to determine the promoters and enhancers in a particular tissue. While basic scientists are moving to bridge the gap with clinicians, there is still limited contact and we need to do better to meet in the middle. Basic scientists also need better understanding of what would be credible to clinicians if no number of fly and mouse studies will convince them.

Relatedly, functional data are needed on the clinical side; clinical labs sometimes interpret the variants with no functional data at all. Dialogue is needed between the clinical lab and the clinician but currently little infrastructure for these bidirectional conversations exists. The UDN is a good model for the way interactions between the sequencing lab (as well as basic labs) and clinicians should be taking place. For patient reporting, there should be some consideration given to crowd-source data for patient-entered information in disease-based groups. Patients often become experts on their condition and can provide details doctors often cannot.

Session 5: Functionalizing VUS's

Massively Parallel Functional Analysis of Missense Mutations in *BRCA1* for Interpreting Variants of Uncertain Significance (Lea Starita)

Traditionally, the clinical field has interpreted the impact of genetic variation through classical genetic experiments, like retrospective or familial studies that cannot be scaled. Computational predictions of variant pathogenicity are scalable, and rapidly improving, but they still don't produce completely reliable results. To address this, a toolkit of methods is being developed for functionally assessing hundreds to thousands of genetic variants in each experiment and applied to genes such as *BRCA1*. Deep mutational scans of the *BRCA1* RING domain and saturation genome editing of exon 18 were used to understand the effect of missense variants on splicing, a particularly challenging problem.

Challenges for scaling up this work include the amount of manual work to construct a library and deliver variants, developing parallelizable assays for protein function, sequencing the variants, developing computational variant scoring pipelines, and calculating likelihood estimates for pathogenicity. While libraries can be constructed and sequencing performed in a few months, building the protein function assays and analyzing the data are rate limiting steps. As a first step forward, we could create a gene list like the ACMG-56 with certain criteria and prioritize resolving VUSs in genes returned to patients.

CRISPR-Cas9 Mediated Mouse Model Creation and Transcription Regulation (Haoyi Wang)

In comparison to conventional gene targeting and gene editing technologies (such as Zinc-finger Nucleases, ZFN, and transcription Activator–Like Effector Nucleases, TALEN), CRISPR-based systems take significantly less time, cost less, and have more flexibility. One significant bottleneck in CRISPR-based procedures is the microinjection process, which has been improved with the development of Zygote Electroporation of Nuclease (ZEN) technology. The recently developed Casilio platform, a combination of Cas9 and Pumilio, has many applications for modeling epigenetic abnormalities and gene regulation networks, as well as a unique olimerization feature for labeling centromeres, telomeres, and unique sequences with a small number of guide RNAs. A list of actionable genes would also be useful to gene-editing research. Gene-editing systems such as CRISPR-Cas9 should be explored to develop cost-effective models for the development of rare disease therapeutics.

Discussion

The community should develop a list of criteria to guide the development of functional assays for various conditions. For instance, we could focus on 'common, treatable, doable' first (i.e., genes for

common disorders like intellectual disability; genes that have an actionable treatment; and genes for which assays could be developed in a reasonable and cost-effective manner).

Lack of standardization is a significant challenge in promoting translation of basic to clinical research. A standard nomenclature for functional variation and prediction would greatly improve systematic mining of the literature. Resources such as the Variation Ontology (VariO) of functional terminology and NCBI's Allele Registry should help with literature mining and labeling of unique variants. Vanderbilt is creating a 'gene by medical phenome' catalog using genetically predicted expression of genes (from GTEx data) to collate the up and down regulation of every gene in the genome.

The community would benefit from a large-scale repository to store functional data. This likely extends beyod the scope of ClinGen's curation interface, which accommodates and promotes a systematic assessment of functional data, and focuses on clinical validity and actionability. KOMP, RegulomeDB, HaploReg, and the Sanger Institute's zebrafish database are existing resources that could facilitate the interaction between the clinic and basic research. As a product of GM9, it would be helpful to aggregate a list of resources (as well as a brief explanation) to facilitate work between the bench and bedside. A long-term goal would be for these individual resources to connect to each other with a standard API, enabling users to query and cross reference all of the databases. While many of the model organism databases already have human disease orthology pages, there is a need for continued investment, greater linkage, and more complete aggregation.

We should also consider utilizing the sequence and clinical data from the upcoming PMI cohort. These data could be used to prioritize the development of different functional assays. While difficult to ascertain, it will also be important to gather data from people who are homozygous knock-outs or control populations, such as people who should have developed diabetes but did not.

Session 6: Biomedical phenotype ontologies and data integration

Translating Human to Models and Back Again: Phenotype Ontologies for Data Integration and Discovery (Melissa Haendel)

The Monarch Initiative is a fairly new consortium which seeks to have clinical, computational, and basic science leading large-scale integration efforts for genotype to phenotype. In OMIM, there are three thousand Mendelian diseases without a known genetic basis. If the orthologs of every coding gene in the human genome are compared to the top five most common model organisms, there is 80% coverage, so there is a lot of information to be extracted from model organisms, but vocabulary descriptions of phenotypes differ across species. To make phenotypes computable across organisms, we need a "universal converter box" or suite of ontologies. Logical axioms could then be used to decompose a term and make it computable with standardized terms that are equivalent across species.

The Human Phenotype Ontology (HPO) is a graph structure representing a suite of clinical phenotypes. Unsurprisingly, clinical vocabularies do not represent a patient's phenotype as though he/she were a biological specimen. Additionally, anatomical systems are represented in terms of other ontologies with a great amount of data. Large-scale data integration would bring all this together. While ontologies do well with genes, they are much poorer with environment and phenotypes. A newly-developed phenotype exchange format called "PhenoPacket" is useful in different kinds of contexts. Deep phenotyping within and across species can aid diagnosis, discovery, and translational matchmaking. An exchange standard is needed to facilitate distributed phenotype data sharing for patients and across species. Lastly, a computable genotype to phenotype evidence model could aid variant interpretation.

Data Integration: Genome X Transcriptome X EMR (Nancy Cox)

The Genotype-Tissue Expression project (GTEx) is a way of using information from the transcriptome to understand the relationship between gene and phenotype. Using GTEx as a reference panel, SNP-based predictors can be created for each gene and each GTEx tissue. The gene-based test entails imputing transcript levels for each gene and tissue and examining the association of that endophenotype with disease. More than 18,000 genes have been assayed and show a correlation between the genetically predicted and directly measured transcript level of at least 0.2, in at least one tissue.

The biobank associated with the Vanderbilt's EHR, "BioVU," permits Phenome-wide Association Studies (PheWAS) of disease and hospitalization codes at the gene level in >200K patients with DNA samples. PheWAS was used to identify the medical phenome associated with genetic variants predicting the knockdown of each gene in each tissue, as well as variants associated with up-regulation. Large numbers of phenotypes are associated with expression variation in Mendelian genes such as the glutamate transporter *GRIK5*, discovered to be related to a multitude of eye phenotypes in BioVu which were then rapidly validated in a zebrafish model. Genetically predicted gene expression is a bell-shaped curve due largely to natural variation, with a "healthy genome" less likely to be in the tails of the curve. The number of genes a person has in the tails of the expression distribution is associated with the number of PheWAS codes accumulated in a lifetime. Across BioVU, variants predicting reduced expression of Mendelian (LoF) disease genes were associated with the phenotypes that make up that Mendelian disease and seemed to contribute disproportionately to disease risk. There will be more people with increased risk of disease due to reduced expression of these genes than there are people with a Mendelian disease due to true LoF variants, and many of the former may respond to reasonably innocuous therapies for Mendelian diseases such as dietary modification. BioVU data are currently being used to develop a database of Mendelian disease genes and associated phenotypes, and to identify "Mendelian genes in waiting" associated with multiple congenital anomalies and intellectual disability.

Mendelian genes tend to collect on certain axes of disease risk, such as wound healing, *TGF-beta* signaling, and apoptosis, just as common disease genes do. By identifying phenotypes associated with reduced expression in progressively larger numbers of people, functional assays will be useful in looking at one or more ends of the axis. The big picture is that there is a continuum of variation in "Mendelian" genes spanning from mouse embryonic lethal to loss-of-function tolerance.

Discussion

If translatable phenotypes were available, a massively parallel mutagenesis scan could be done through massively parallel clinical phenotyping with the same endophenotypes. To promote broader adoption by the community, enriching the ability for clinicians to see other phenotypes that might associate with a known Mendelian disease is key. Crossing the translational divide between basic scientists and clinicians will require ontology tools that account for differences in the way clinicians and basic scientists record information. Additionally, bioinformaticians need to share the process of making these pipelines instead of individuals making similar and separate pipelines. For those who don't have access to BioVU, expression can be imputed in many tissues that have been surveyed by GTEx. This enables learning about potentially causal mechanisms for which transcriptome data are not available.

The medical genetics community is interested and willing to provide phenotypic information in a way that is useful to these databases but they need to know how—this could be facilitated by working with ACMG to make links to physicians. A challenge to clinical engagement is avoiding additional work in the context of ordering genetic tests, though that may be the best time to capture standardized phenotypic

information. Integration of test ordering with clinical databases that simultaneously pull relevant phenotype information would facilitate bidirectional communication. The data could then be deposited in Monarch or ClinVar as patient level data, with appropriate consent. If all that could be done by filling out one form, it would be efficient for the clinician and highly valuable for research.

Making Sausage: Who Decides if Genetic Testing is Reimbursed (Robert Nussbaum)

The four main players in reimbursement decisions are the patients, providers (employers), purveyors (labs), and payers ("4 Ps"). Payers are heterogeneous in their knowledge of genetic testing and often feel that they have been burned in the past by new technology and the costly misuse of testing which may trigger unnecessary and expensive downstream testing.

Payers respond to professional guidelines, as well as to a large number of unpaid claims. Employers wish to keep insurance premiums down and have market power with the payers, who want to keep these employers in their system and therefore, have to be responsive to them. Payers look to partner with purveyors or labs and others to provide utilization management. Labs feel uncomfortable in this role and want to stay closer to the providers and the patients.

Labs survive by billing for tests but have limited resources to generate the evidence supporting reimbursement. While there are attempts to provide evidence for clinical utility of various genes and tests, few studies demonstrate economic value to the payer community. It often falls on the purveyor to demonstrate economic viability of testing, which is not always in the purveyor's scope. Among labs, competition fuels the tendency to claim low VUS rates and over-call, which feeds payer paranoia. Academic researchers play an important role in evidence development and assessment.

Providers are crucial partners for implementing clinical utility studies, which labs are not in a position to implement themselves. Providers also play key roles in updating professional guidelines and promoting consensus. Involving patients in negotiations for new insurance products may help to reduce adversarial interactions with the reimbursement system when claims or services are denied.

Panel 2

Clinicians and the broader healthcare system should be mobilized to provide useful phenotypic information that is both clinically valuable and accessible in EHR systems. Given that parsing and curating data from publications is incredibly time consuming, journals could mandate that functional data be deposited into a standard database prior to publication. The field also needs to gather more quantitative data sets so that probabilistic inferences can be made.

A criterion for prioritizing genes for functional studies should include discrepancies in variant pathogenicity calls. Promoting communication between the clinician and the laboratory, as well as providing the laboratory with more phenotypic information, would help reduce variant discrepancies. One challenge is that most terms clinicians currently use are based on therapy and reimbursement, rather than on fundamental biology. The Undiagnosed Disease Program (UDP) found it helpful to give feedback to clinicians on their clinical evaluation—based on a 5-star system in PhenoTips, clinicians were shown how 'meaningful' and 'detailed' their phenotyping was. This simple assessment substantially improved the quality of their phenotyping. Hiring additional genetic counselors and having them act as intermediaries and partners with specialty clinicians (e.g., neurologists) would be incredibly helpful.

Patients could be given the opportunity to add phenotypic data into their medical record, but such data would need to be collected in a standard format. The Monarch Initiative is moving to translate HPO terms to patient-friendly language, and the Global Alliance for Genomics and Health (GA4GH) has translated HPO into several major world languages. Patients could provide phenotypic information before their appointments so the physician could review and 'co-sign' their report.

While there are large, existing datasets with valuable genomic data (e.g., ExAC), it is challenging to re-contact these individuals and re-phenotype them. Ideally, there would be standardized phenotyping and follow-up questionnaires for the individuals in these datasets. PhenX would be a good resource for this, with over 450 standard data collection protocols across 20 different disease domains, all mapped to LOINC terms and formatted for RedCap and other resources.

To facilitate interaction between clinicians and basic scientists, a platform could be built to allow individuals to deposit either a 'clinical problem' or a 'research platform'—sort of a "matchmaker exchange" for clinicians and researchers to exchange priorities. Holding joint case-conferences would also promote communication between clinicians and basic scientists.

Summary Discussion: Promoting Bedside-Back-to-Bench Research

Overarching themes from the two-day meeting can be framed in terms of variants, phenotypes, and people. The first identified theme was prioritizing functionalization to determine on which genes to focus. ClinVar would be a good place to start, by identifying variants that have conflicts across pathogenicity classifications. It might be useful to have a dedicated small group of both clinical and basic scientists to prioritize genes, as well as other groups to work together on data integration. CSER and UDN may be good opportunities to collect priority genes and use cases from clinicians. The Inter-Society Coordinating Committee (ISCC), a group of professional societies focused on education of physicians, could also advise on prioritization. The ClinGen Gene Curation Working Group effort is identifying and curating genes that have varying levels of clinical validity and additional functional assessments. For gene prioritization, the ACMG 56 is a good place to start now.

It may not be the role of NHGRI to develop and implement criteria for classifying VUSs; this seems more the domain of professional groups such as ACMG. Energy is better spent on phenotype data being captured systematically and facilitating data exchange for patients who receive a VUS. Engaging with ACMG will speed up the process; basic scientists and computer scientists should join ACMG and attend their working groups and meetings.

Phenotyping as a second theme repeatedly underscored the need for deep phenotyping in those with unusual genotypes. Interested journals might be approached to publish papers on the minimum phenotype dataset for a gene or common condition. Since the degree of deep phenotyping needed may change with clinical context, the field should consider more iterative phenotyping because it allows questions to be answered as they arise for the clinician.

EHR phenotypes could be designed for more for data-driven models of clinical features The Monarch Initiative's effort to translate Monarch to patient language in a pilot phase by the end of this year will provide an important resource for patient-driven phenotyping. There is a clear need for common vocabularies and mapping across vocabularies, as well as common databases and a quantitative standard. If a model organism and variant associated with the phenotype is reported, then basic, standardized, minable information should accompany it into the database. Genomic Medicine Meeting X could invite editors to help determine data mining standards and use of standardized terminology.

For the third theme of bridging people in basic and clinical sciences, many resources already exist and the challenge may be in raising mutual awareness. Lack of user-friendly interfaces is a common problem. Physicians are accustomed to working with ambiguity and we need to be clear that many genetic test reports contain ambiguous information that requires clinical correlation. If GM9 were to develop a paper highlighting this ambiguity it should highlight how ambiguity pervades all aspects of medicine, not just genomics. Large clinical laboratories could convene at a future GM meeting to discuss standards for lab reports, though ACMG and the Association of Molecular Pathologists (AMP) are actively tackling this.

Creating a 'matchmaker exchange' for clinicians and basic scientists would be helpful for fostering interactions and collaboration. Workshops at specialized medical conferences would help engage the medical community. Similarly, clinicians should be invited to basic science conferences such as ENCODE and model organism meetings. Non-traditional venues and industrial conferences could be worthwhile to target. The National Science Foundation is interested in developing phenotype exchange standards, and NIH could leverage these connections. Webinars and case conferences, particularly those offering continuing medical education credits, are also helpful for facilitating communication.

Since there are many different players involved (basic scientists, clinicians, industry), NHGRI could convene groups to establish shared infrastructure, standards, and goals. These groups could create a list of 100 'clinically-impactful' genes and diseases, and then support basic scientists and encourage industry to develop high-throughput assays and animal models for these genes. Resulting data should be made freely available, with special emphasis on sequencing underrepresented populations, studying complex conditions, and evaluating noncoding regions. NHGRI could also incentivize interactions between clinicians and basic scientists by recommending partnerships in funding opportunity announcements or developing dedicated partnership programs similar to the sequencing-ELSI partnerships in CSER.

Next Steps and Collaborations:

Genes prioritized for functionalization should include genes implicated as causal in CSER, UDN, and similar studies; LoF intolerant genes implicated in disease but of unknown function; and genes with practical functional assays. Mapping electronic phenotypes to the Monarch Initiative will be another important collaboration. Basic scientists and computer scientists should be encouraged to join more clinically related groups, such as ACMG, and vice versa. To further bridge the gap, a 'matchmaker exchange' for clinicians and basic scientists would be helpful for fostering interactions and collaboration.

Ideas for future Genomic Medicine meetings include having large clinical laboratories discuss and compare standards for lab reports. Involving clinicians as actual consumers of those reports would be valuable. Other ideas include an invitation to editors to determine data mining standards or to patients to promote patient-derived phenotyping and data exchange.

Next steps from this meeting will involve drafting a meeting summary and recommendations, and a white paper for publication co-authored by all presenters and moderators who comply with International Committee of Medical Journal Editors authorship guidelines.