Molecular Phenotypes of Null Alleles in Cells (MorPhiC)

Birth of a New Program at the National Human Genome Research Institute (NHGRI)

Adam Felsenfeld, Colin Fletcher, Ajay Pillai, Riley Wilson

Presentation for Research America 11022022

NIH National Human Genome Research Institute

The Forefront of Genomics

“Program Director, Extramural Research Programs, Division of Genome Sciences, National Human Genome Research Institute, National Institutes of Health”

1. Congress gives NHGRI $$$. Most of that pays for grants to independent investigators
2. Usually: Applicant Idea→Grant Proposal→Review→Funding

- PD evaluates award progress, policy compliance, funding level, etc.
- PD’s usually have a “portfolio” of grants within a broadly related topic; maintain knowledge in scientific area

NIH NHGRI

Other Things PD's Do (more fun)

- Strategic planning
- Propose coordinated program concepts to address gaps/opportunities
- Manage programs – goals/ milestones/deliverables, standards, "human factors", policies, facilitate collaborations, ***change***

Molecular Phenotypes of Null Alleles in Cells (MorPhiC)

Birth of a New NHGRI Program

https://www.genome.gov/research-funding/Funded-Programs-Projects/Molecular-Phenotypes-of-Null-Alleles-in-Cells

National Human Genome Research Institute

The Forefront of Genomics

Genome Sequencing is now Cheap! But What does the Sequence Mean?

Wouldn't it be great to:

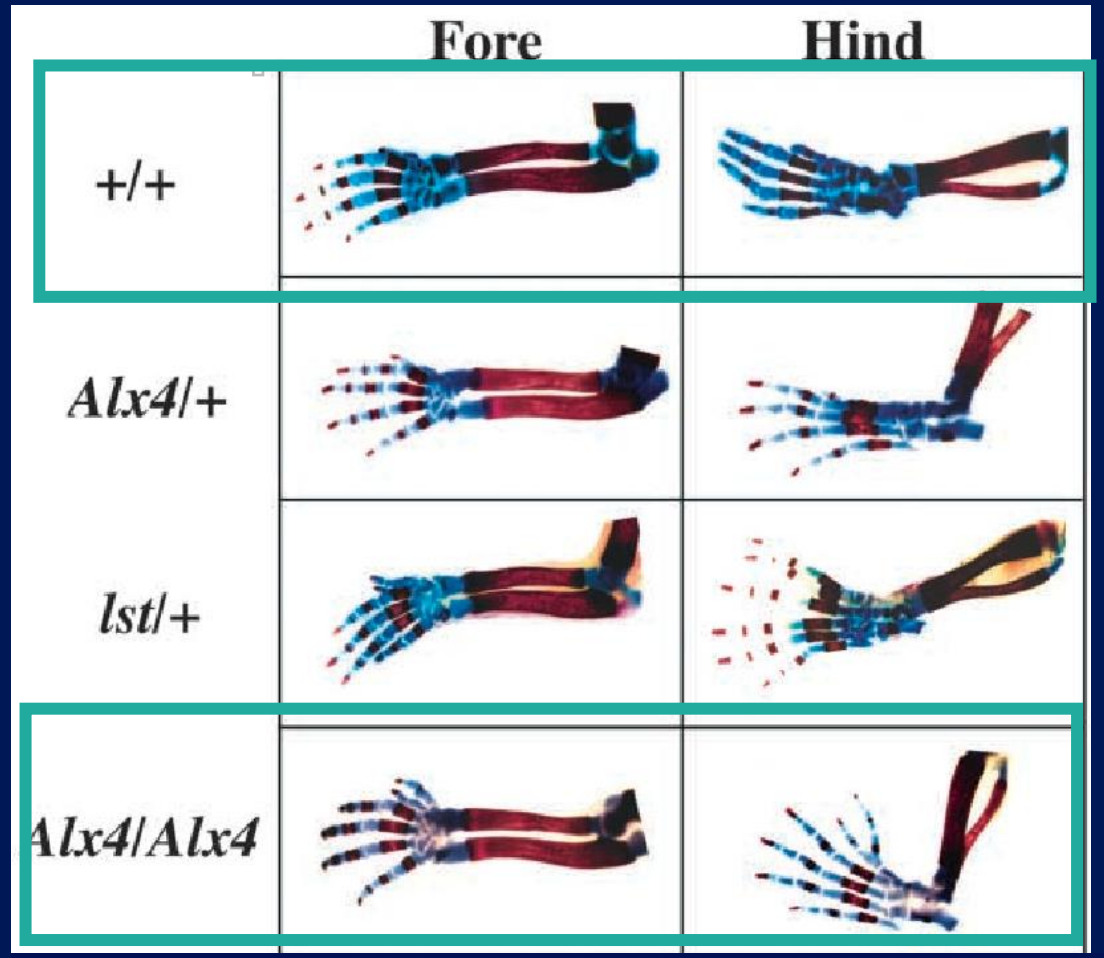
- Know what each/every gene does?
- Learn, or predict what variants or mutations in each gene will do?

- Understand biology: Sequence ("genotype") to form, function ("phenotype")

- Understand disease risk from gene variants

- Clues for how to intervene

Example of “Genotype to Phenotype” in Mouse: Alx4 gene-- extra fingers and toes



Qu et al. doi: 10.1242/dev.125.14.2711

Genome Sequencing is now Cheap! But What does the Sequence Mean?

Wouldn't it be great to:

- Know what each/every gene does?
- Learn, or predict what variants or mutations in each gene will do?

- Understand biology: Sequence ("genotype") to form, function ("phenotype")

- Understand disease risk from gene variants

- Clues for how to intervene

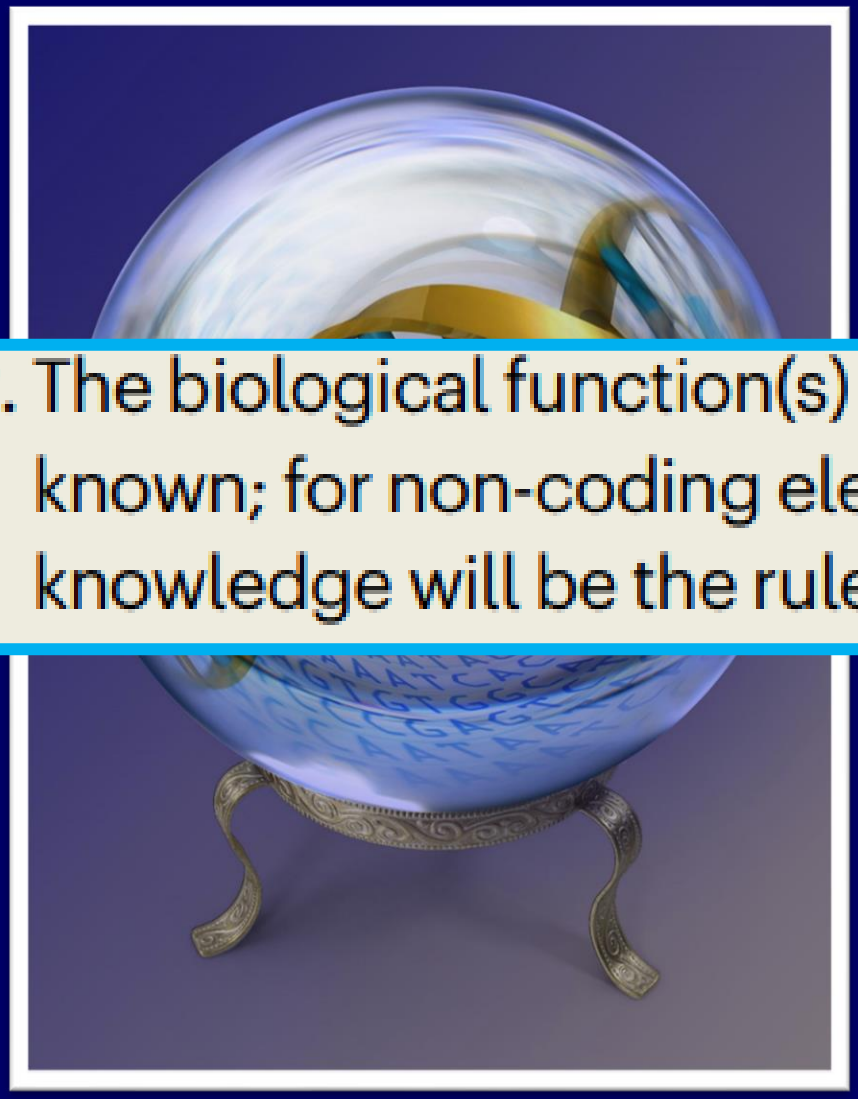
NHGRI Strategic Vision: Compelling Genomics Research Projects in Biomedicine

- **Comprehensive views of genes and regulatory elements**

"…an unprecedented opportunity to decipher the individual and combined roles of each gene and regulatory element. **This must start with establishing the function of each human gene, including the phenotypic effects of human gene knockouts**."

https://doi.org/10.1038/s41586-020-2817-4

NHGRI Bold Predictions



Bold predictions for human genomics by 2030

Some of the most impressive genomics achievements, when viewed in retrospect, could hardly have been imagined ten years earlier. Here are ten bold predictions for human genomics that might come true by 2030. Although most are unlikely to be fully attained, achieving one or more of these would require individuals to strive for something that currently seems out of reach. These predictions were crafted to be both inspirational and aspirational in nature, provoking discussions about what might be possible at The Forefront of Genomics in the coming decade.

1. Generating and analysing a complete human genome sequence will be routine for any research laboratory, becoming

2. The biological function(s) of every human gene will be known; for non-coding elements in the human genome, such knowledge will be the rule rather than the exception.

associated phenotypic information for millions of human participants will be regularly featured at school science fairs.

6. The regular use of genomic information will have transitioned from boutique to mainstream in all clinical settings, making genomic testing as routine as complete blood counts.
7. The clinical relevance of all encountered genomic variants will be readily predictable, rendering the diagnostic designation 'variant of uncertain significance (VUS)' obsolete.
8. An individual's complete genome sequence along with informative annotations will, if desired, be securely and readily accessible on their smartphone.
9. Individuals from ancestrally diverse backgrounds will benefit equitably from advances in human genomics.
10. Breakthrough discoveries will lead to curative therapies involving genomic modifications for dozens of genetic diseases.

Human Genome Sequence Assemblies Are Now Cheap! But what does the sequence mean?

Wouldn't it be great to:

- Know what each/every gene does?
- Learn, or predict what genetic variants in each gene will do?

- Understand biology: genotype to phenotype
- Understand disease risk from gene variants
- Clues for intervention

Concept proposal

MorPhiC Long Term Goal

Create a catalog:

- **Make null alleles (=~ “gene knockouts”)**
- **Of every gene (but start with ~1000 as a pilot)**
- **Of the molecular and cellular effects**
- **In human cells in culture systems that are faithful to biology**

NIH NHGRI

MorPhiC Choices

- **Why null alleles?**
- **Why molecular and cellular effects (phenotypes)?**
- **Why multicellular systems?**

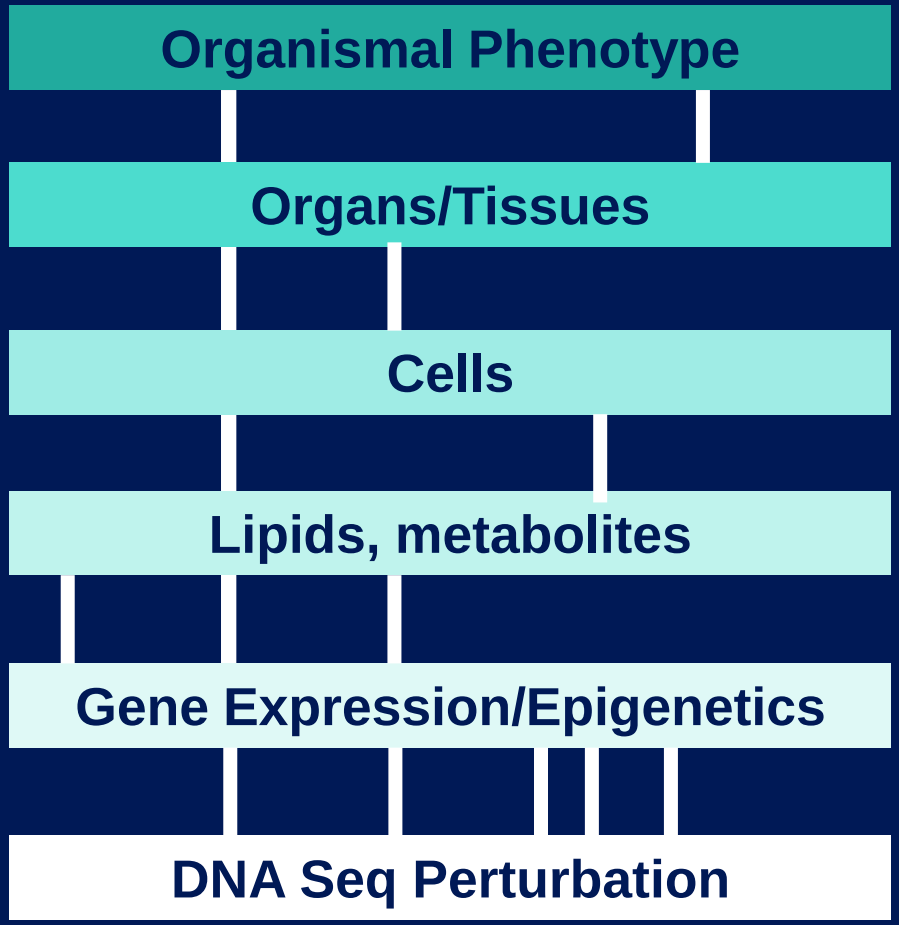
NIH NHGRI

Benefits

- Provide basic, consistent cellular/molecular information for all genes
- Fill gap between sequence/gene expression, and anatomical/physiological phenotypes
- Foothold to mechanism
- Useful data for computational methods
- Cell lines are tools for more detailed study

NIH NHGRI

Perturbation X Function ideal

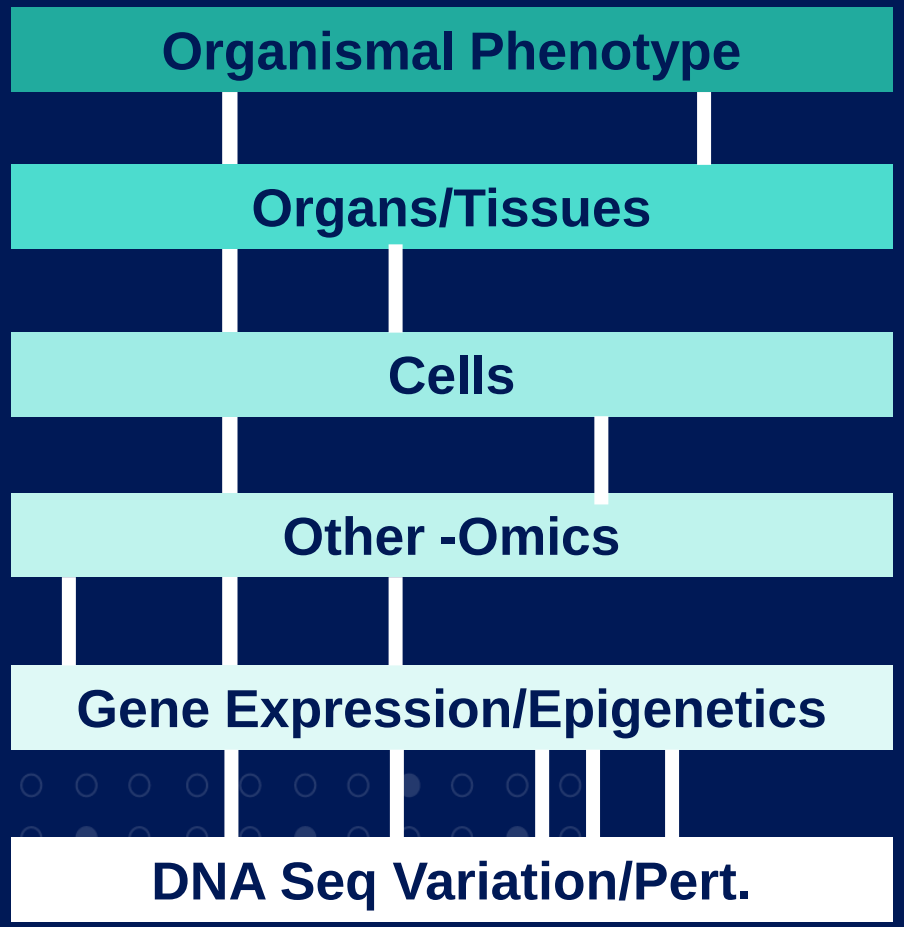


NIH NHGRI

Challenges

- Can we knock out every gene? *Pretty sure*
- Can we do ***enough*** informative tests of cellular and molecular phenotypes? Do we have the right assays?
- Null may kill the cell (hard to interpret). Or may have multiple different effects on different cells (“pleiotropy”).
- How to deal with biological variability

MorPhiC Depends on Other Function Studies



Organismal Phenotype	Assoc. studies; KOMP/IMPC
Organs/Tissues	MorPhiC, Others
Cells	MorPhiC; Others
Other -Omics	MorPhiC; IGVF; Others
Gene Expression/Epigenetics	MorPhiC (trans); IGVF (cis); GTEx
DNA Seq Variation/Pert.	Assoc. studies, KOMP, MorPhiC, Others

NIH NHGRI

Molecular Phenotypes of Null Alleles in Cells (MorPhiC) Phase 1

Data Production Reseach and Development Center Awards

The Jackson Laboratory

Paul Robson

Memorial Sloan Kettering Cancer Center

Danwei Huangfu

Northwestern University

Mazhar Adli

UNIVERSITY OF CALIFORNIA

Luke Gilbert

Data Resource and Administrative Coordinating Center Award

UNIVERSITY OF MIAMI

Stephan Schürer

Still in process

- RFA-HG-22-019: Data Analysis and Validation Centers

Program management

Phase 1 (2022)– Develop pipeline

Can high-quality data be produced at scale?

-Null alleles for 1000 genes (protein coding)
-Test multiple mutagenesis methods; assays
-Data QC, interoperability, infrastructure, comparisons
-Prospects for throughput and cost improvement
-Refine scientific challenges
-Produce initial high-quality data for analysis
-Evaluate value of data/use cases

Phase 2 (2026)

Contingent on lessons of Phase 1, ***to be evaluated prior to a new Concept***

NIH NHGRI

Thanks to colleagues for input on ideas and presentation

Ajay Pillai
Colin Fletcher
Riley Wilson
Carolyn Hutter

END

Phase 1 Systems and Assays

Samples

- iPSC
- hPSC
- Primary (islet cells)
- Background diversity (M/F; populations)

Systems

- Brain organoids
- Embryoids - ectoderm, endoderm, mesoderm
- Other multicellular differentiation models

  (cardiac, neuronal)

Assays

- Transcriptomes (bulk and single cell); Spatial Tx
- Epigenomics (bulk and single cell)
- Metabolomics
- Lipidomics
- Imaging (growth, shape, viability)
- Differentiation/Pluripotency
- Ca++
- Electrophys

MorPhiC Data Flow