

# Future Opportunities for Genome Sequencing and Beyond

Summary of the NHGRI Workshop of  
July 28-29, 2014  
(Condensed Version)

National Advisory Council for Human Genome  
Research

September 8, 2014



National Human Genome  
Research Institute

# What?

- What questions and opportunities in genomics can be addressed “at scale”\*, starting with sequencing but not limited to it
- Consider options for future NHGRI programs to address

\* “at scale” implies:

- Answers to the scientific questions that require scale to address
- High-quality comprehensive data resources; developed technologies, approaches, project designs, analysis methods, policies, file formats, etc.
- Highly managed efforts

# How?

- **Acknowledge our history, but not be bound by it**
- **Agenda included topics that are proximate to sequencing and could raise new scientific opportunities for exploring genomic problems at scale**

# Why now?

## Genomic Science is Changing

- **Technology**
- **Availability of “infrastructure” including data and analysis tools**
- **Community**
- **Clinical applications**

## Different era

- **Federal budgets**
- **“Central planning” vs “Investigator initiated”**

# Workshop Scope

Structure of  
Genomes

Biology of  
Genomes

Biology of  
Disease

Science of  
Medicine

Effectiveness of  
Healthcare

**Genome  
Function**

**(Related  
Tech Dev)**

**Disease Gene  
and Variant  
Discovery**

**Clinical  
Applications  
of Sequencing**

**(Related  
Informatics)**

# Workshop Topics

## Major topics:

- I. Genetic architecture of health and disease at scale:
  - Discovering variants conferring risk for common disease;
  - Discovering the genomic basis of Mendelian disease
- II. Integrating genomic variant discovery with function
- III. Clinical genome sequencing at scale
- IV. Comparative and Evolutionary Genomics

Possible features/implementation of future NHGRI initiatives

# Watch the Meeting and Read the Tweets

The screenshot shows the homepage of the National Human Genome Research Institute (NHGRI) website. At the top left is the NHGRI logo and the text "genome.gov National Human Genome Research Institute National Institutes of Health". To the right is a Google search bar. Below the header is a navigation menu with links for "Research Funding", "Research at NHGRI", "Health", "Education", "Issues in Genetics", "Newsroom", "Careers & Training", "About", and "For You". There are also social media icons for Facebook, Twitter, and YouTube. A breadcrumb trail reads: "Home > Research Funding > Research Funding Divisions > Division of Genome Sciences > NHGRI Genome Sequencing Program (GSP) > 2014 Genome Sequencing Workshop". The main content area features the title "Future Opportunities for Genome Sequencing and Beyond: A Planning Workshop for the National Human Genome Research Institute" with "Share" and "Print" buttons. Below the title is a date "July 28-29, 2014" and the location "Bethesda North Marriott & Conference Center, 5701 Marinelli Road, North Bethesda, Md. 20852, (301) 822-9200". A paragraph states: "On July 28-29, 2014, the National Human Genome Research Institute (NHGRI) sponsored the workshop - *Future Opportunities for Genome Sequencing and Beyond: A Planning Workshop for the National Human Genome Research Institute* - at the Bethesda North Marriott Hotel & Conference Center." Below this is a section "The objectives of the workshop were to:" followed by a bulleted list: "• Discuss the scientific questions and opportunities that can be substantially addressed by large-scale genomics studies, starting with genome sequencing but also considering other genomic technologies." and "• Consider options for future NHGRI programs that would address these questions and opportunities." To the left of the text is a video thumbnail showing a laboratory setting with a person working at a bench. Below the thumbnail is a "YouTube Video Playlist" button. At the bottom of the content area, it says "View workshop videos and slides from: July 28 July 29". At the very bottom, there is a link "Tweets from the Workshop: #GSPFuture" with a PDF icon.

<http://www.genome.gov/27558042>

# Scientific Opportunities

- Discover genetic variants underlying human disease and healthy traits
- Improve understanding of genomic variation through functional genomics studies
- Continued work in comparative and evolutionary genomics will aid variant interpretation
- Evaluate clinical utility of genomic sequencing and approaches to clinical implementation



# Scientific Opportunities

- Foster a “virtuous cycle” between discovery and clinical applications to accelerate our understanding of genotype-phenotype relationships and translation to genomic medicine.
- Lead and contribute in efforts to ensure collective sequence data is a powerful resource for the scientific community.

# I. Genetic Architecture of Health and Disease

- A. Define genotype/phenotype relationships underlying human inherited disease and healthy traits across the spectrum
- B. Enable the knowledge base needed to interpret genome sequence variation in life science, drug discovery, clinical prediction and diagnosis
- C. Include a range of human diseases and populations to expand discovery, define architecture, and broaden access as a matter of social justice

## **II. Integrating Genomic Variant Discovery with Function**

- A. Define the molecular, cellular, organ and organismal functions of coding and non-coding genome sequences**
- B. Develop tools to manipulate sequences at scale**
- C. Develop methods for faithful, large-scale functional characterization of sequence variants**
- D. Systematically catalog molecular components and their interactions, across cell fates and cell states**

# III. Clinical Genome Sequencing

- A. Define clinical contexts in which genome sequencing improves patient outcomes
- B. Enable rapid, robust detection of all clinically relevant variation in a single test
- C. Clinical sequencing data for research use
- D. Improve approaches for determining pathogenicity
- E. Identify effective methods for implementing sequencing into routine medical practice

# IV: Comparative and Evolutionary Genomics

- A. Produce high-quality de novo sequenced and assembled genomes
- B. Obtain nucleotide-level resolution of every conserved element in humans; understand specific genomic changes in human and primate lineages.
- C. Leverage model organisms for functional genomics.
- D. Further develop the informatics infrastructure for display, alignment, distribution

# Implementation Advice

## Variant discovery (and clinical sequencing):

- Need flexibility, nimbleness AND clear goals: consortium model
- Increasingly will need partners for large efforts (not just for samples...)
- Need more transparency/outreach especially if the program needs to work with collaborators from disease communities and other partners
- Design program so progress is measurable. Clear short-term and long-term (foundational) goals
- Small and large efforts are (should be managed to be) complementary, not alternatives

# Implementation Advice

**General: NHGRI can not do all of this, but needs to have an influential role.**

- Exemplar studies
- Methods/tech dev
- Foundational resources
- Standards
- Enable integration
- Training
- Partners
- Integrate ELSI
- Beware premature consensus

**“Most of the sequence data in the future will not be produced by NIH funding. NHGRI has a role in ensuring that the data will be useful.”**

# NHGRI Impressions

- The opportunities are as important and exciting as ever.
- What we heard is in line with the NHGRI Strategic Plan.
- Increased emphasis on complementarity between variant discovery and function, and between discovery and clinical sequencing
- There is still high value in large, highly-managed, resource generating, technology-advancing consortia
- NHGRI can't do it all, but we are well-placed to pursue foundational projects that will have critical importance for genomics.
- To do more, it is both important and feasible to use partnerships, including for sharing costs of projects.



# Workshop Scope

Structure of  
Genomes

Biology of  
Genomes

Biology of  
Disease

Science of  
Medicine

Effectiveness of  
Healthcare

**Genome  
Function**

**(Related  
Tech Dev)**

**Disease Gene  
and Variant  
Discovery**

**Clinical  
Applications  
of Sequencing**

**(Related  
Informatics)**

# Workshop "Wish-List"

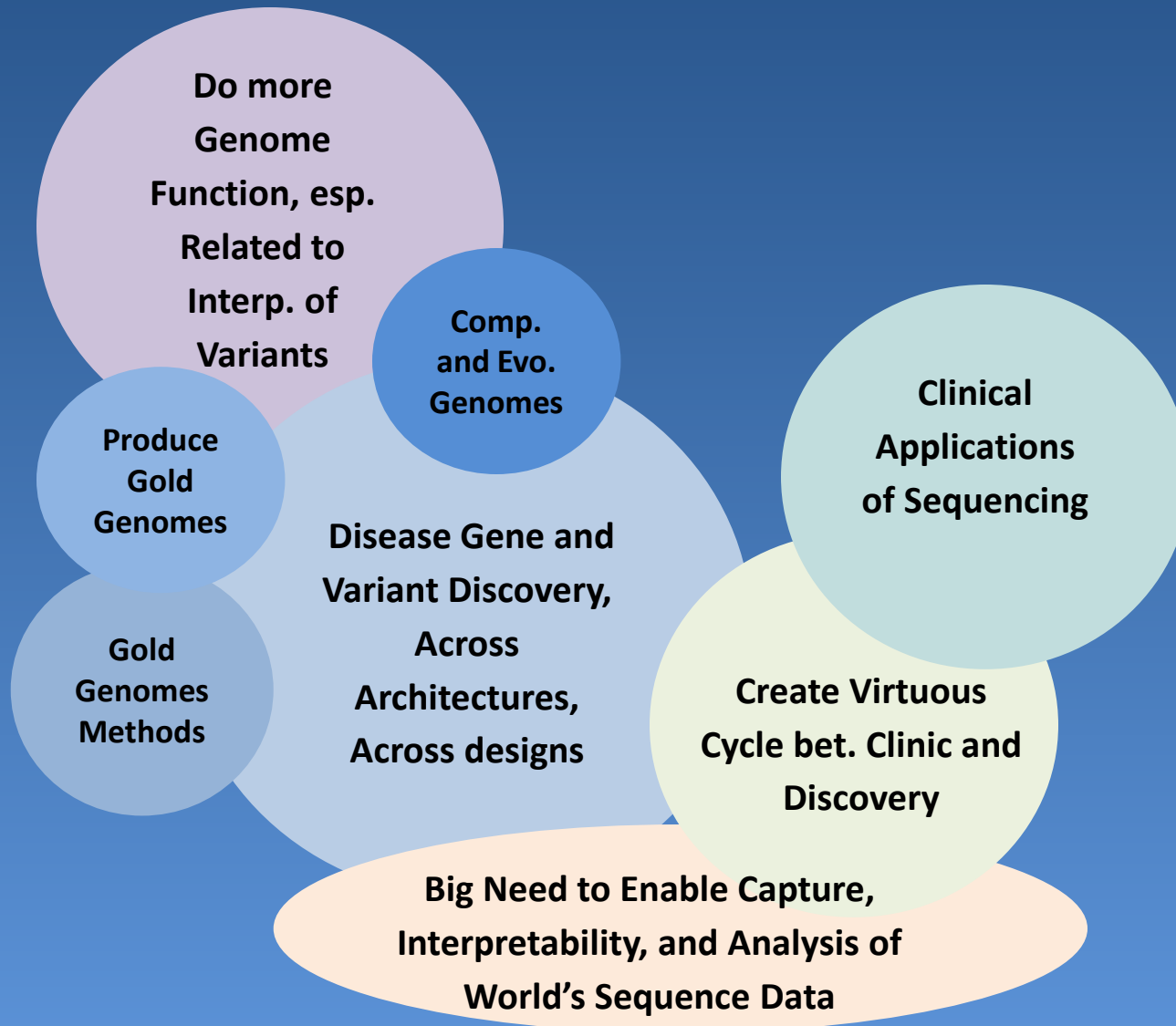
Structure of  
Genomes

Biology of  
Genomes

Biology of  
Disease

Science of  
Medicine

Effectiveness of  
Healthcare



# De-anonymized (current programs)

Structure of  
Genomes

Biology of  
Genomes

Biology of  
Disease

Science of  
Medicine

Effectiveness of  
Healthcare

ENCODE,  
GGR, FunVar

Tech Dev

Large-Scale  
Sequencing  
and Analysis  
Centers;  
Centers for  
Mendelian  
Genomics

CSER,  
eMERGE,  
UDN, NSIGHT

ClinGen

GS-IT and related  
informatics

# Acknowledgements

**National Advisory Council for Human Genome Research**

**Scientific Advisors to the Genome Sequencing Program**

**Workshop Agenda Committee: Ewan Birney, Eric Boerwinkle, Carlos Bustamante, Joe Ecker, Jim Evans, Bill Gelbart, Len Pennacchio**

**NHGRI Leadership: Larry Brody, Bettie Graham, Eric Green, Teri Manolio, Rudy Pozzatti, Jeff Schloss, Mark Guyer\*, Jane Peterson\***

**Shannon Biello, Deborah Colantuoni, Lucia Hindorff, Carolyn Hutter\*\*, Lu Wang, Elise Feingold, Mike Pazin, Mike Smith, Heidi Sofia, Chris Wellington, Kris Wetterstrand**