Concepts for Functional Genomics Initiatives

National Advisory Council
for Human Genome Research
May 18, 2015

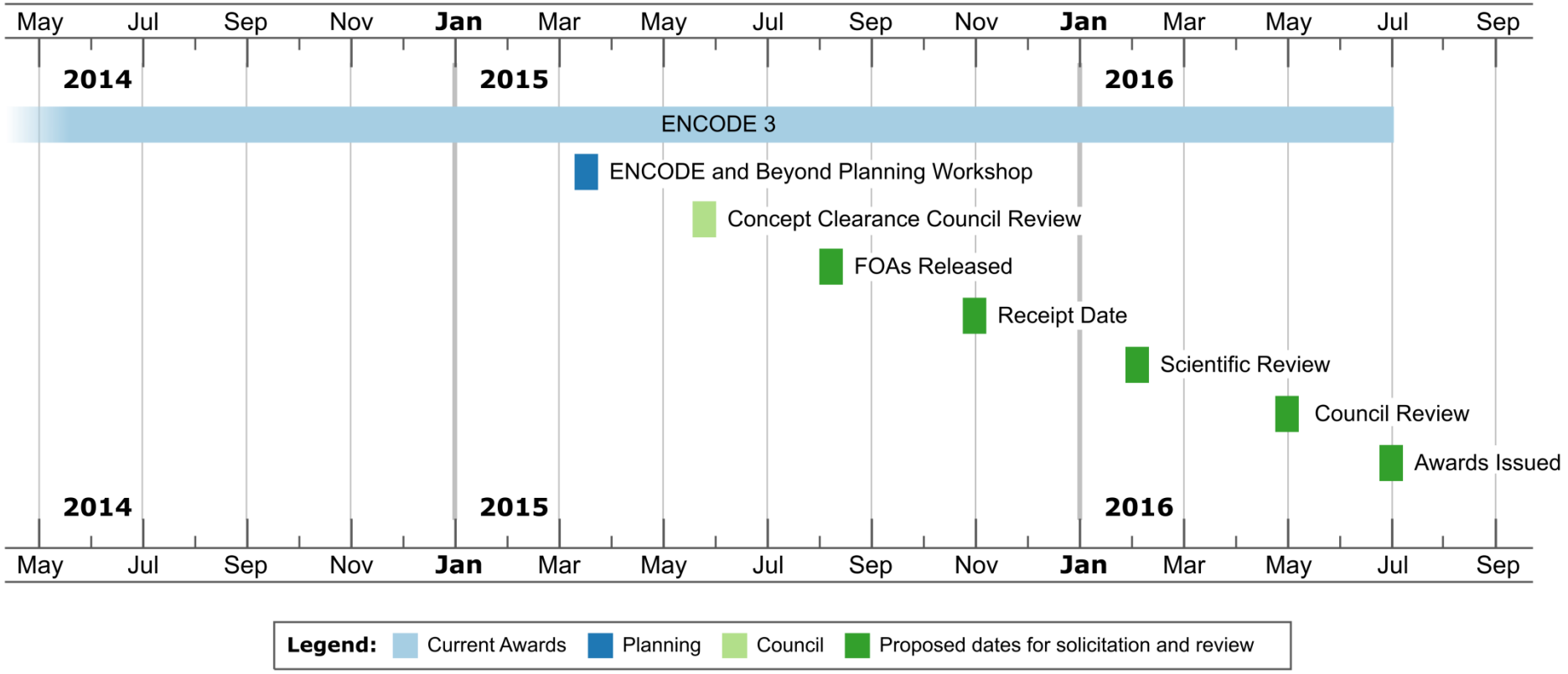
Elise Feingold, Mike Pazin, Dan Gilchrist
Adam Felsenfeld, Jeff Schloss

NIH National Human Genome Research Institute

Outline of Presentation

- Timeline
- Overall Goals
- What's New
- Concepts for Four Initiatives
- Opportunities for Co-Funding
- Overall Budget

Timeline for Initiatives



Goals of Initiatives

- Expand catalog of functional elements in human and mouse genomes
- Move beyond cataloging to understand functional roles of genomic elements in specific biological contexts
- Develop strategies to apply these studies to disease
- Increase number of scientists from research community contributing to resource
- Develop analytical tools to enhance data utility
- Make data, tools, analyses and assembled encyclopedia freely available

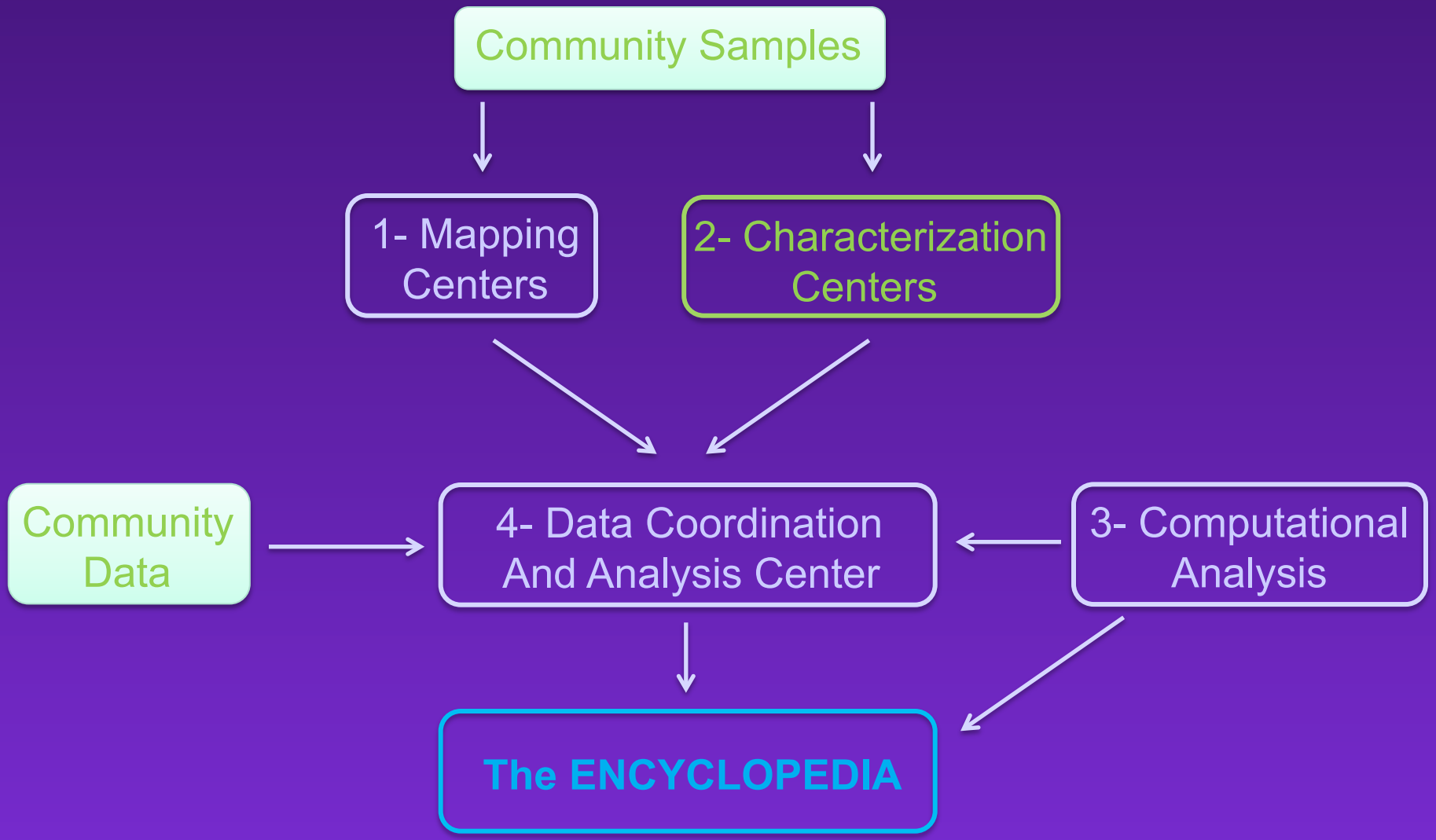
Initiatives

1. Functional Element Mapping Centers
2. Functional Element Characterization Centers
3. Computational Analysis Research Projects
4. ENCODE Data Coordination and Analysis Center (EDCAC)

New Features

- Functional element characterization at scale (Characterization)
- Direct application of functional genomics assays to disease studies (Mapping and Characterization)
- Significantly increased community participation (Mapping, Characterization and EDCAC)

Proposed Organization



1-Functional Element Mapping Centers

Purpose: Expand catalog of candidate functional elements in the human and mouse genomes

Scope and Objectives: Apply high-throughput assays to map biochemical activities highly associated with specific functional elements in expanded cell space

Funding: U54s, 6-8 centers, $20M TC/year, 4 years

1-Functional Element Mapping Centers

New Features:

- Broaden community participation by obtaining unique biological samples from experts
- Apply 25% mapping effort to disease samples
- Develop scientific basis for bounding the experimental space
- Use of samples consented for open access data release
- Limited focus on mouse
- Flexibility in overall pipeline

2- Functional Characterization Centers

Purpose: NEW ACTIVITY

- Enhance catalog of candidate functional elements by characterizing and validating set of elements
- Obtain understanding of utility and generalizability of functional characterization approaches

Scope and Objectives:

- Employ multiple approaches to test candidate elements in specific biological contexts to validate and characterize functional activities

Funding: U01s, 7-10 centers, $5.9M TC/year, 4 years

2- Functional Characterization Centers

Key features:

- Applicants select cell systems/biological contexts for in-depth study
- Studies of specific diseases using relevant cell sources encouraged
- Multiple assays can be used
- Assays can be conducted in well-justified model organisms
- 25% of individual group effort devoted to elements studied by all
- Use of samples consented for open access data release

3- Computational Analysis

Purpose: Maximize utility of ENCODE and related functional genomics resources by bringing in additional computational expertise to analyze ENCODE data

Scope and Objectives: Support investigator-initiated projects to develop and apply analytical and statistical tools to use and improve the ENCODE resources

Funding: U01, 6 projects, $3.0M TC/year, 4 years

3- Computational Analysis

Activities can include:

- Developing new methods to improve on analysis, visualization and interpretation of ENCODE data
- Combining ENCODE data with related functional genomics data from other projects to derive new biological insights
- Using ENCODE data to improve on the analysis of disease mapping studies to identify causal variants

4- ENCODE Data Coordination and Analysis Center (EDCAC)

- Purpose:
  - Provide community access to ENCODE data and resources
  - Support analysis of ENCODE data necessary to create and make available a high quality encyclopedia
  - Organize and facilitate Consortium activities
- Scope and Objectives:
  - Support Data Coordination Center (DCC) and Data Analysis Center (DAC), funded separately with expectation of functioning as single entity
  - DCC activities significantly expanded
  - DAC activities more tightly focused on creating and validating encyclopedia
- Funding: U41s, 2 centers, $7.5M TC/year, 5 years (1 year beyond data production)
  - Additional $2.6M in FY16 needed for existing DCC to close out current ENCODE activities

DCC-Led Activities

- Develop, house, and maintain databases to track, store, and provide access to data, metadata and computational tools
- Develop, maintain, and update data processing pipelines
- Maintain and enhance portal to ensure easy access to data and resources
- Provide community access to data in state-of-the-art browser/visualization formats
- Track and report on data submission

❖ Import data from outside investigators and projects, developing the needed infrastructure and methodology

❖ Serve as the Coordinating Center to:

- Facilitate communication and coordination
- Organize and support annual meetings and working groups
- Support outreach activities to promote broad use of data, analyses and tools

DAC-led Activities

- Specify and update data processing pipelines
- Provide leadership and computational expertise to analyze data to:
  - update and refine the encyclopedia
  - ❖develop scientific strategy to bound the cell space to be deeply interrogated
  - ❖expand the lexicon of functional elements beyond the basic categories
  - update existing and develop new data quality metrics and standards for all data types
- ❖Bring in additional expertise to conduct short-term, specialized analyses needed to maximize the quality and utility of the resource

Co-Funding Opportunities

Seek support of projects focused on specific diseases or biology of interest to other NIH Institutes and Centers (ICs)

- Whole or partial funding of projects (characterization)
- Supplemental funding for mapping samples of particular IC interest (mapping)
- Supplemental funding to support submission of existing or planned datasets generated under other projects (EDCAC)

Budget Summary

Activity	FY16	FY17	FY18	FY19	FY20
Functional Element Mapping Centers	$20M	$20M	$20M	$20M	-0-
Functional Characterization Centers	$5.9M	$5.9M	$5.9M	$5.9M	-0-
Computational Analysis	$3.0M	$3.0M	$3.0M	$3.0M	-0-
EDCAC	$7.5M	$7.5M	$7.5M	$7.5M	$7.5M
New Initiatives Total	**$36.4M**	**$36.4M**	**$36.4M**	**$36.4M**	**$7.5M**
DCC 1-yr Extension	$2.6M				
Total	**$39.0M**	**$36.4M**	**$36.4M**	**$36.4M**	**$7.5M**

Questions?