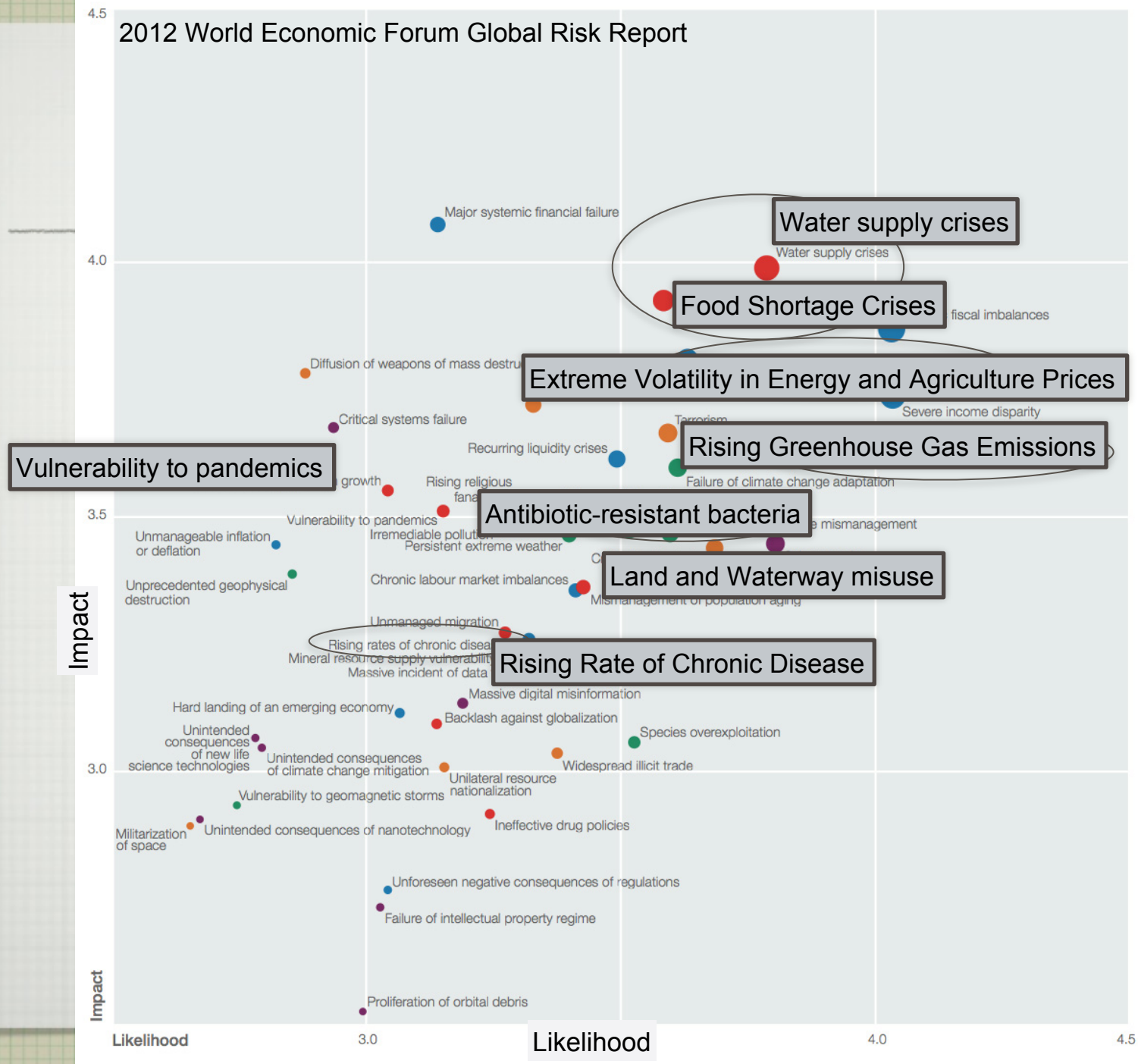
Integrative Genomics Building (IGB) Kbase and JGI

Presentation to Community Advisory Group

May 12, 2014


2012 World Economic Forum Global Risk Report
Water supply crises
Food Shortage Crises
Extreme Volatility in Energy and Agriculture Prices
Rising Greenhouse Gas Emissions
Vulnerability to pandemics
Antibiotic-resistant bacteria
Land and Waterway misuse
Rising Rate of Chronic Disease
Major systemic financial failure
Water supply crises
fiscal imbalances
Diffusion of weapons of mass destru
Critical systems failure
Severe income disparity
Recurring liquidity crises
Failure of climate change adaptation
Rising religious
Vulnerability to pandemics
Unmanageable inflation or deflation
Persistent extreme weather
Unprecedented geophysical destruction
Chronic labour market imbalances
Unmanaged migration
Massive incident of data
Massive digital misinformation
Hard landing of an emerging economy
Backlash against globalization
Species overexploitation
Unintended consequences of new life science technologies
Unintended consequences of climate change mitigation
Widespread illicit trade
Unilateral resource nationalization
Vulnerability to geomagnetic storms
Militarization of space
Unintended consequences of nanotechnology
Ineffective drug policies
Unforeseen negative consequences of regulations
Failure of intellectual property regime
Proliferation of orbital debris
Impact
Likelihood
4.5
4.0
3.5
3.0
3.0
4.0
4.5

Why Is Biology a Good Source for Systems Biology & Engineering?

~10^{30} bacteria Everywhere.

~10^{11} stars in the Milky Way

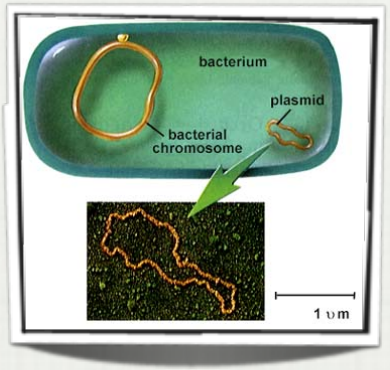
~10^{23} in the observable universe

~10^{6} bacteria/g soil
~10^{5} species

~10^{6} people in San Francisco

~Each person emits ~10^{6} bacteria/hour

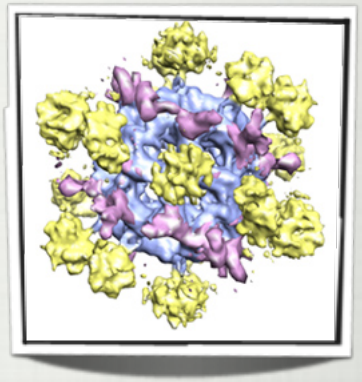
~4Mb genomes
~10^{3} genes



~8000 genomes/flash drive in my pocket

A sense-compute-and-do system contained in a 10^{-15} liters!

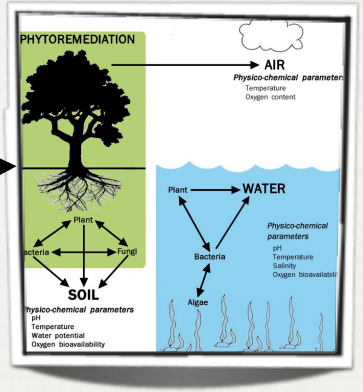
~10^{3} proteins
~10^{7} interactions



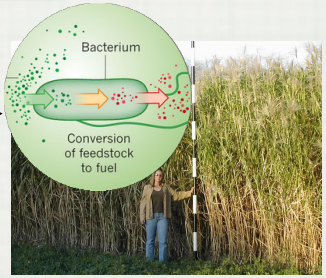
~60,000 different classes of protein

Biology Solutions for Big Problems

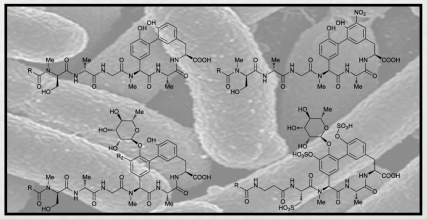
- Water supply crises
- Land and Waterway misuse
- Rising Greenhouse Gas Emissions
- Extreme Volatility in Energy and Agriculture Prices
- Food Shortage Crises
- Antibiotic-resistant bacteria
- Vulnerability to pandemics
- Rising Rate of Chronic Disease



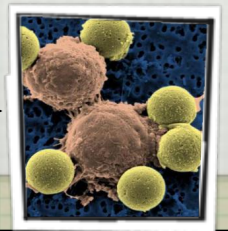
Understand biological impacts and routes to clean soil, water and air



Microbes and plants useful for renewable food and chemical production



Rapid tracking of microbes and gene flow in the environment.



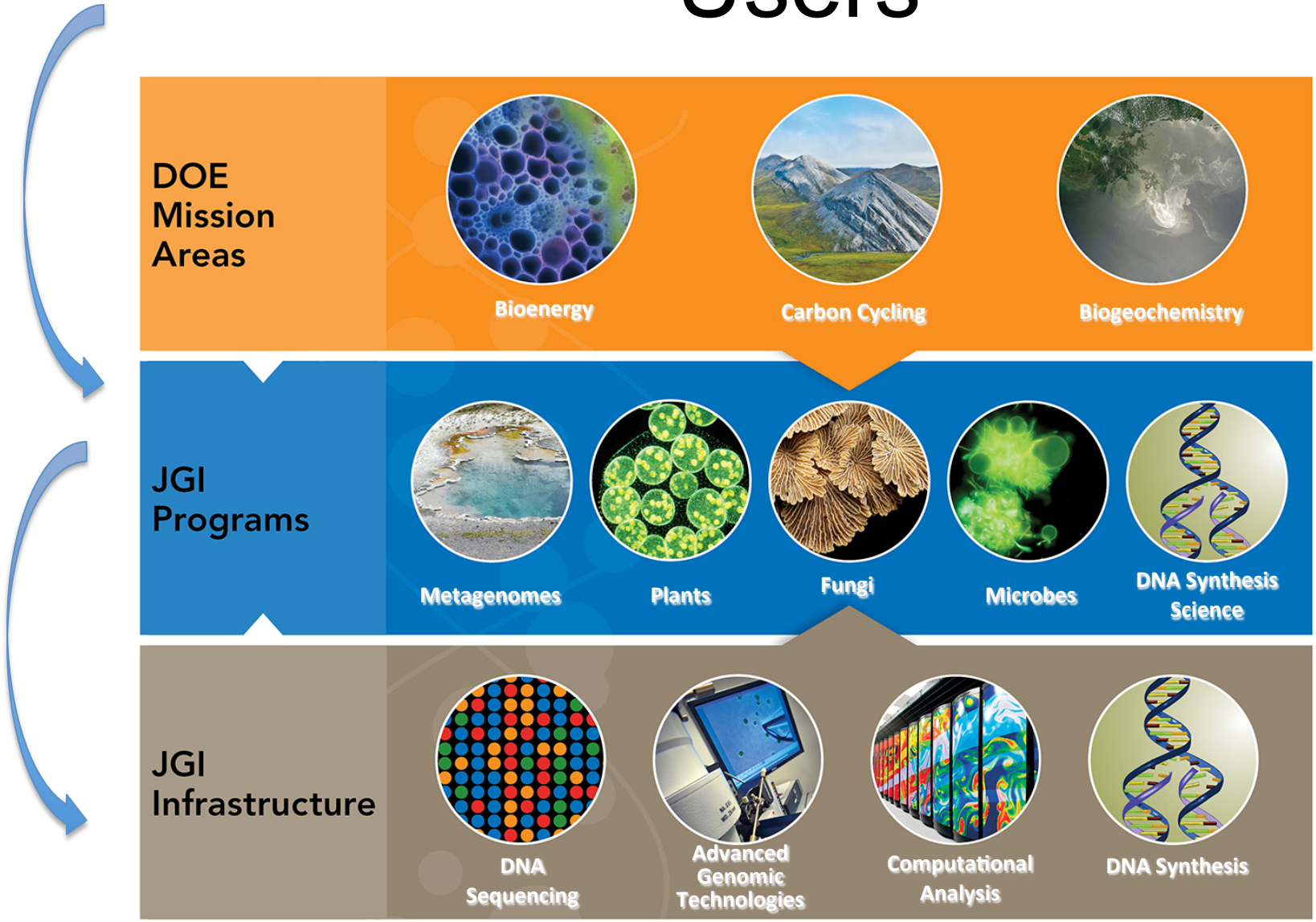
Understanding microbial “gut” communities for underlying responses to environmental change

JGI Overview

User Facility for Large Scale Genomics to Enable Bioenergy and Environmental Research

- Facility
  - Walnut Creek 4-Bldgs/80,000 sq ft
- Staff
  - 284 FTE
- Funding
  - ~69 M US Dept of Energy

Users


DOE Mission Areas
Bioenergy
Carbon Cycling
Biogeochemistry
JGI Programs
Metagenomes
Plants
Fungi
Microbes
DNA Synthesis Science
JGI Infrastructure
DNA Sequencing
Advanced Genomic Technologies
Computational Analysis
DNA Synthesis

KBase
PREDICTIVE BIOLOGY
DOE Systems Biology Knowledgebase

KBase: The DOE Systems Biology Knowledgebase

- KBase aims to enable **collaboratively developed**, **data-driven** and formally **testable models** of **biological function** and behavior in **environmental context**.

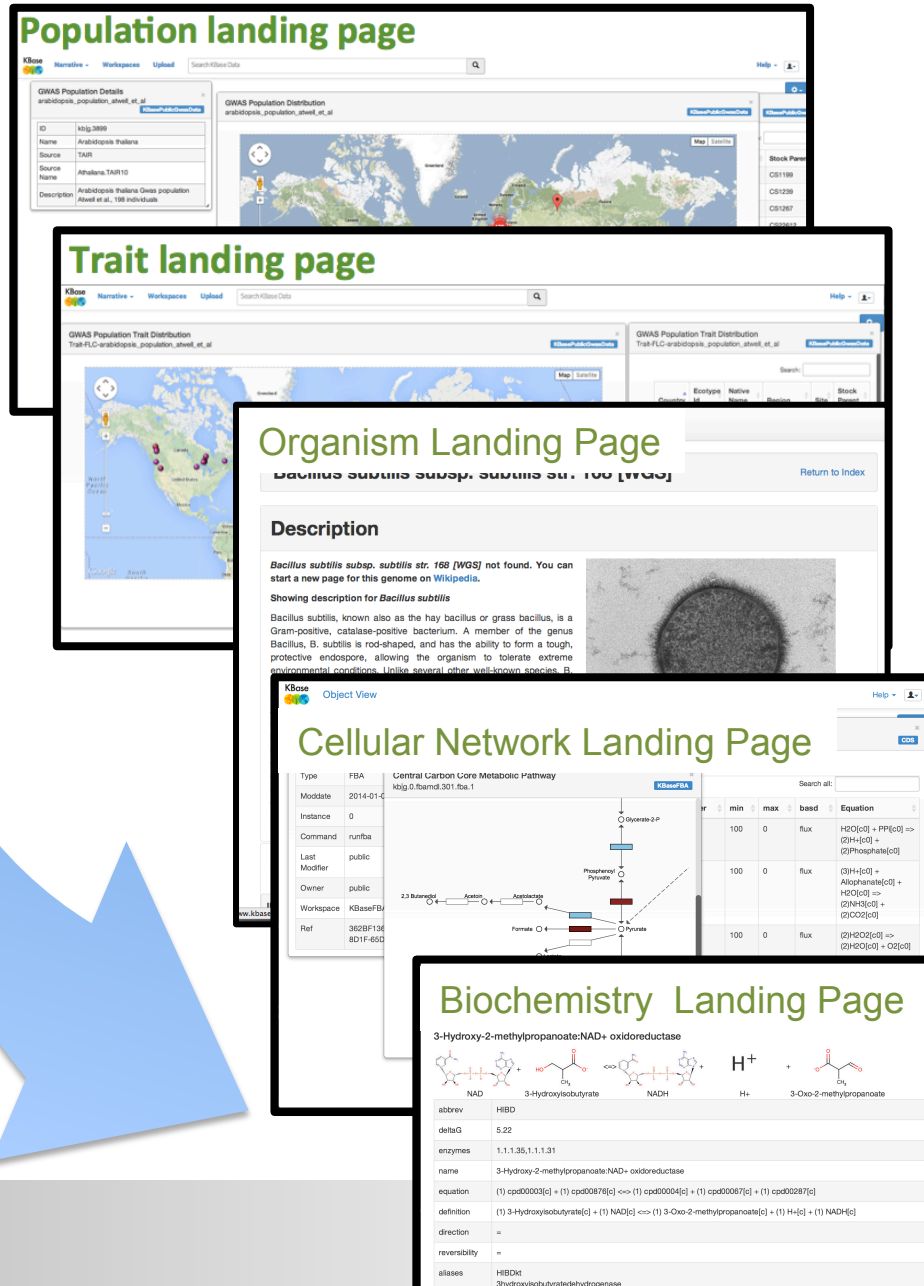
Integrating data

Environmental communities

Macroscopic and Molecular Phenotypes

Genomic Potential and Activity

Models of behavior and transformation



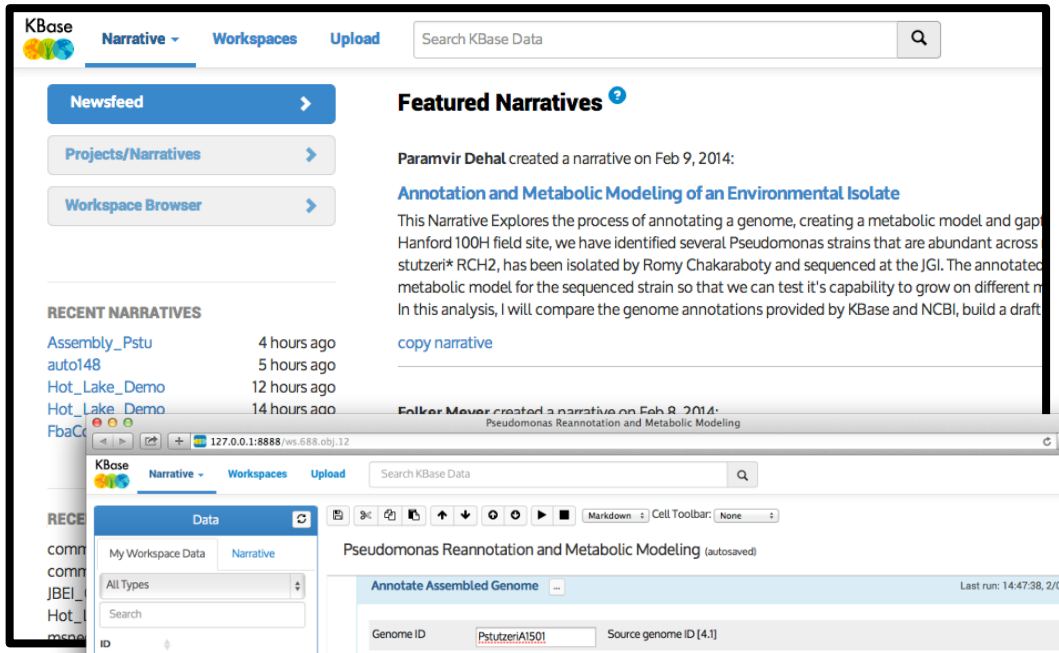
- Open, extensible architecture
- Democratized access to HPC
- Well-defined API
- Scalable workflow engine

http://kbase.us

KBase
PREDICTIVE BIOLOGY
DOE Systems Biology Knowledgebase

Enabling collaborative model-building of microbes, plants and their communities

- Novel "Narrative Interface" allows researchers to share their thoughts, methods, and results in an active, persistent document.
- Enables access to hundreds of data access and analysis methods and large public and private datasets in a unified model.
- Allows open, transparent, reproducible science with complex heterogeneous data sets.
- Enables sharing and project management.



Integrative Genomics Building (IGB)

Providing essential infrastructure for systems biology for bioenergy and the environment

JGI JOINT GENOME INSTITUTE
UNITED STATES DEPARTMENT OF ENERGY

HTP sequencing & analysis
DNA synthesis capability
Converting sequence to function

KBase PREDICTIVE BIOLOGY
DOE Systems Biology Knowledgebase

Data sharing & integration
Modeling of organisms & communities
Supporting predictive biology

Sequence & other data

NERSC

Hypothesis generation

Bio-imaging

Systems Biology

DNA Synthesis

JBEI

M2B

Functional analysis

Berkeley Open Biofoundry

Thank You

Find us at: www.lbl.gov