

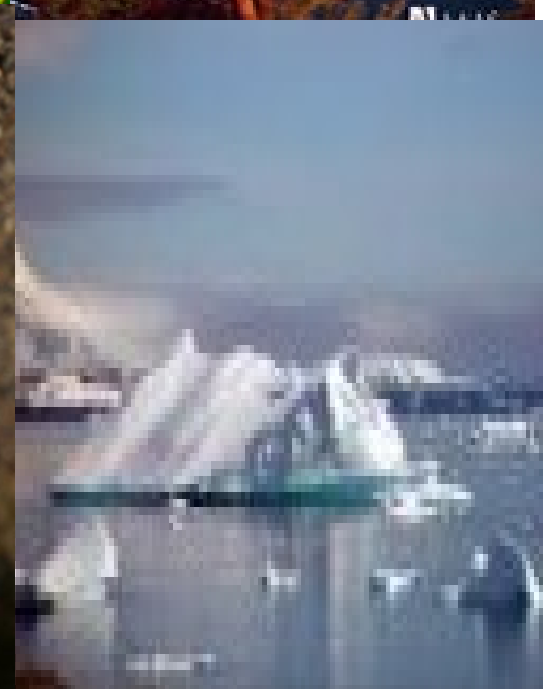
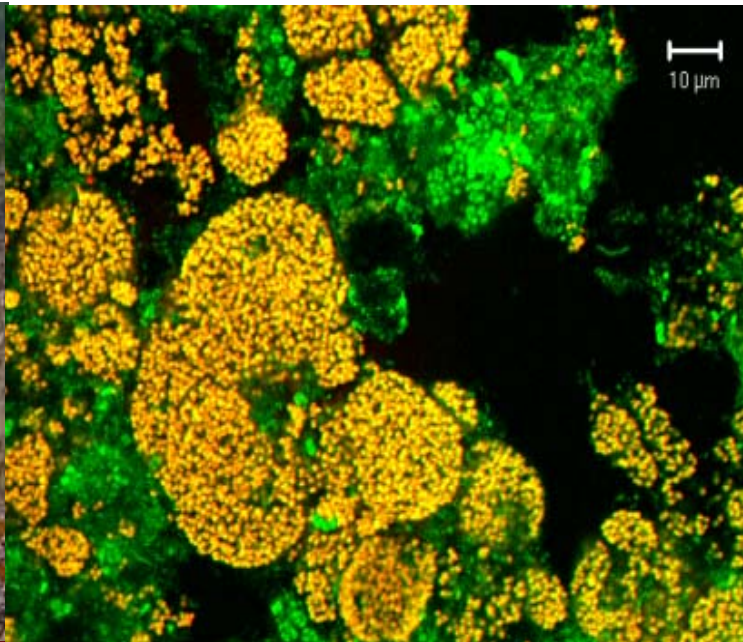
5 Year Strategic Plan Overview BERAC February 19, 2009 (Work in Progress)

Past

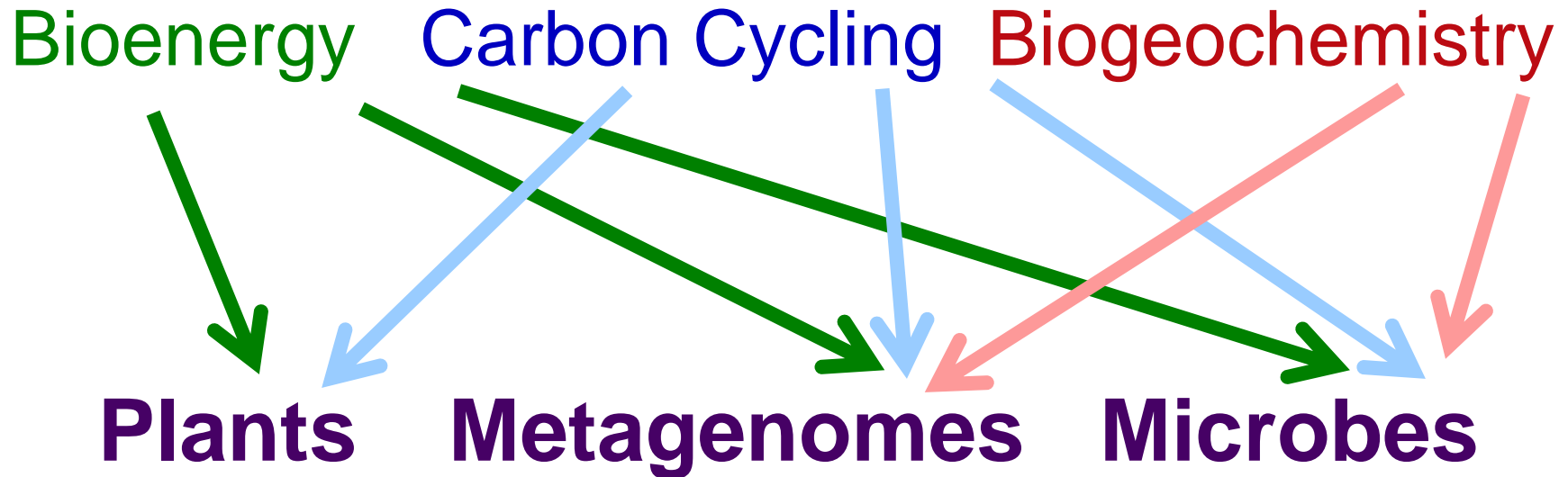
Present

Future

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



Plants, Microbes, and Metagenomes



JGI Strategic Vision

**Contribute Genomics of Scale to
Accelerate the Response
to Energy and Environmental
Challenges**

- **1998-99** 20 million bases of human genome sequenced
- **2008-09** User Facility JGI generates 125 BILLION bases.



| | |
|-------------|--------------------|
| Read Length | 700bp |
| Capacity | 10 million bp/week |

| | |
|-------------|--------------------|
| Read Length | 70-400bp |
| Capacity | 10 billion bp/week |

**External Advisors
(Policy Board, Scientific Advisory Committee)**

**3 Day Retreat
(Users/Univ Sci/BRCs/Big Sci Reps)**

**On Going Process
(BER S&O JGI Review Input)**

JGI's unique capabilities are in carrying out genomics of focus, scale, and complexity

**“Help users solve hard
problems”**



Adam
Arkin

Jay
Keasling

Martin
Keller
Susan Gregurick

David
Mead

Dan Drell

Jim Tiedje

Karin
Remington

Mike
Himmel

John
Taylor

Rob
Phillips

Steve
Quake

Marv Stodolsky

Working Groups for 5 Year JGI Road Map:

Plants
Microbes/Metagenomics
User Programs
Informatics

Plant Working Group:

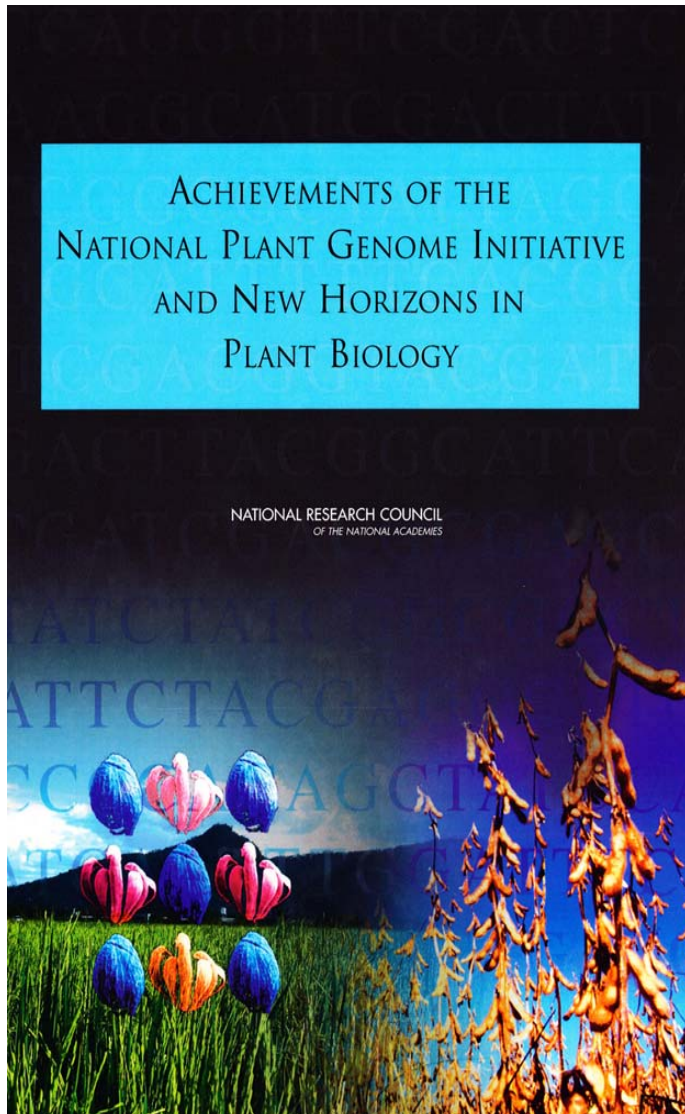
| | |
|----------------|--------------------|
| Dan Rokhsar | UCB / JGI |
| Jerry Tuskan | ORNL |
| Jeff Dangl | U of Georgia |
| Joe Ecker | Salk Institute |
| Jeremy Schmutz | Hudson Alpha Inst. |

JerJ

We have the genomes available for a limited number of plants. The genomes are minimally annotated and we do not understand the function of the genes.

- | | |
|--------------------|--------------------|
| • Poplar | Published 2006 |
| • "Chlamy" (algae) | Published 2007 |
| • Physcomitrella | Published 2008 |
| • Sorghum | Published 2009 |
| • Soybean | Publicly available |
| • Selaginella | Publicly available |
| • Brachypodium | In final analysis |
| • Mimulus | In final analysis |
| • Prunus | Sequenced |
| • Citris | Sequenced |
| • Foxtail Millet | Sequenced |
| • Eucalyptus | Sequenced |
| • Gossypium | Initiated |
| • Manihot | Initiated |
| • Panicum | Initiated |
| • Switchgrass | Initiated |





“... JGI’s contribution to plant genomics is unique and fundamental.

It is critical that JGI continue to serve a broad remit for sequencing and resequencing of plant genomes...”

- To not only complete the assembly and annotation of >10 new plant genomes but also create the associated expression and genotyping resources needed to advance our understanding of plant growth and development.
- A long-term goal is to create a set of genome-based resources and tools that will enable the breeding and/or engineering of plants within the DOE mission.

Plant Genomics Advisory Committee

Plant Genomes Sequencing for DOE
Missions in Bioenergy, Carbon Sequestration
and Phytoremediation –

Outcomes that will be enabled by the proposed goals

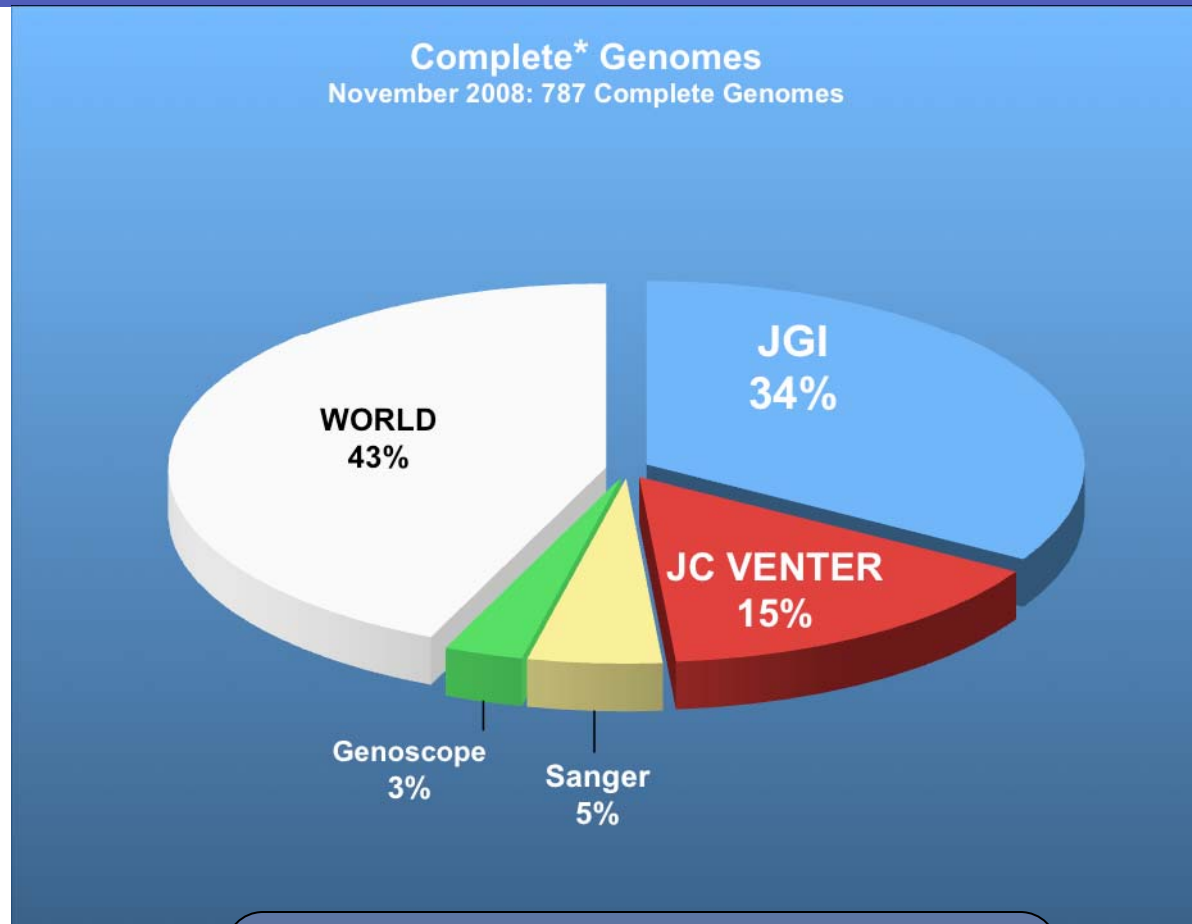


To have a systems level understanding of plants enabling us to harness them for bioenergy production, and to respond to environmental challenges.

Microbial / Metagenomics Working Group

| | |
|-----------------|---------------|
| Jonathan Eisen, | UC Davis /JGI |
| Scott Baker, | PNNL/ JGI |
| Patrick Chain | LLNL/JGI |
| Jim Tiedje | MSU |
| Phil Hugenholtz | JGI |
| John Taylor | UCB |

Current Situation of Archaeal & Bacteria Genomes



**JGI has produced 1/3 of the world's
complete archaeal and bacterial
genomes**

Completely Sequenced and finished version available from Genbank

- Many diverse one-off microbial and metagenomic projects
- CSP Program and BRCs major source of individual microbial and metagenomic projects

Large coordinated projects taking advantage of JGI scale and focus.

Genomic Encyclopaedia of Bacterial and Archaea (GEBA)
(Phylogenetic Approach)

Pilot Project Sequencing and Analysis of 100 microbial genomes
chosen based on phylogenetic relationships

Early results: substantial improvements in

- The analysis of metagenomic samples
- Discovery of new gene families and members of gene families of DOE relevance

More large scale projects

(~ 1000 GEBA Genomes available from all major branches in the tree of life in the next 5 years)

High throughput methods for sample acquisition, preparation, and analysis

a) Culture independent methods

b) Single cell genomics

c) New analysis and ways to query data approaches

d) Transcriptome and proteomic studies increasingly integrated with all DNA sequencing

.

Current Situation at the JGI Metagenomics

Year 2004
#Projects 4

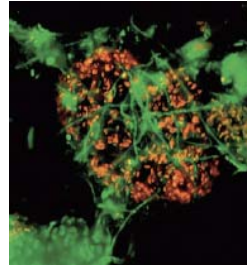
Acid Mine Drainage
Banfield/Chapman
Nature



Marine Archaeal
Methane Oxidation
DeLong/Putnam
Science



2005
7



Soil
Tringe
Science
Fossil
Noonan
Science

2006
7



Olavius algarvensis
Dubilier/Woyke
Nature

Marine Planktonic
DeLong
Science

Wastewater EBR
Hugenholtz
Nature Biotech

Neanderthal
Rubin
Science

2007
12

Termite Gut
Leadbetter/
Hugenholtz
Nature



Oral TM7
Quake
PNAS

Simulated Microbial
Mavrommatis
Nature Methods

2008
14

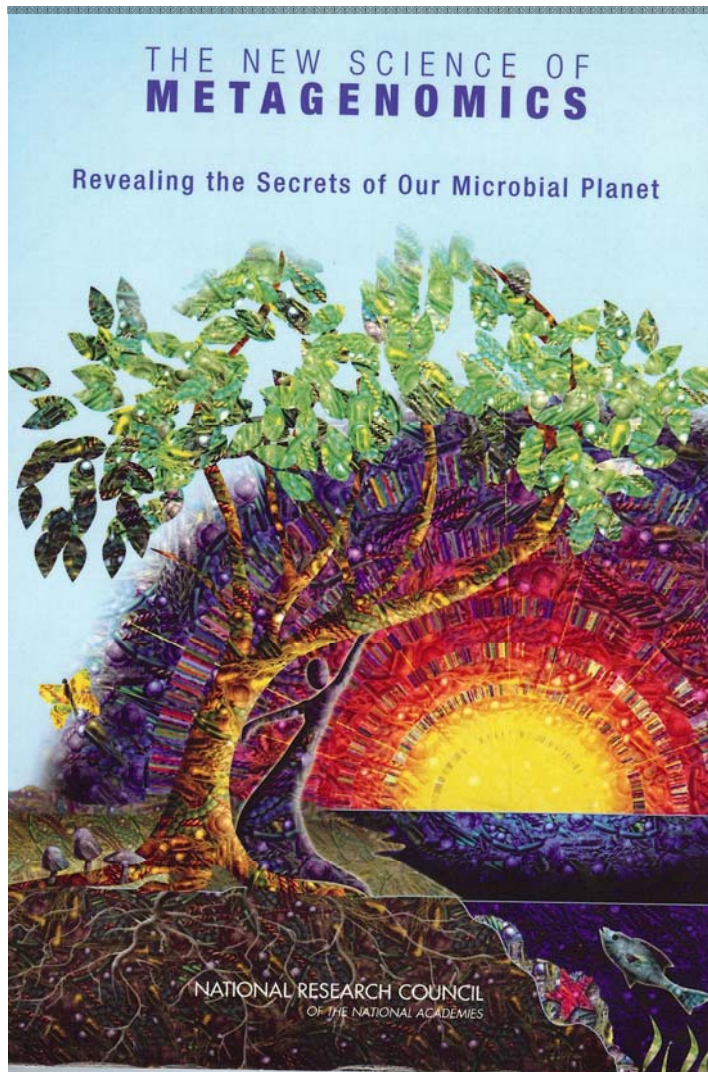
Deep Mine
Chivian
Science

Lake Washington
Methylotrophic
Chistoserdova
Nature Biotech

Marine Microbial
Hugenholtz
Mol Syst Bio

Indoor Urban Air
Tringe
PLoS One





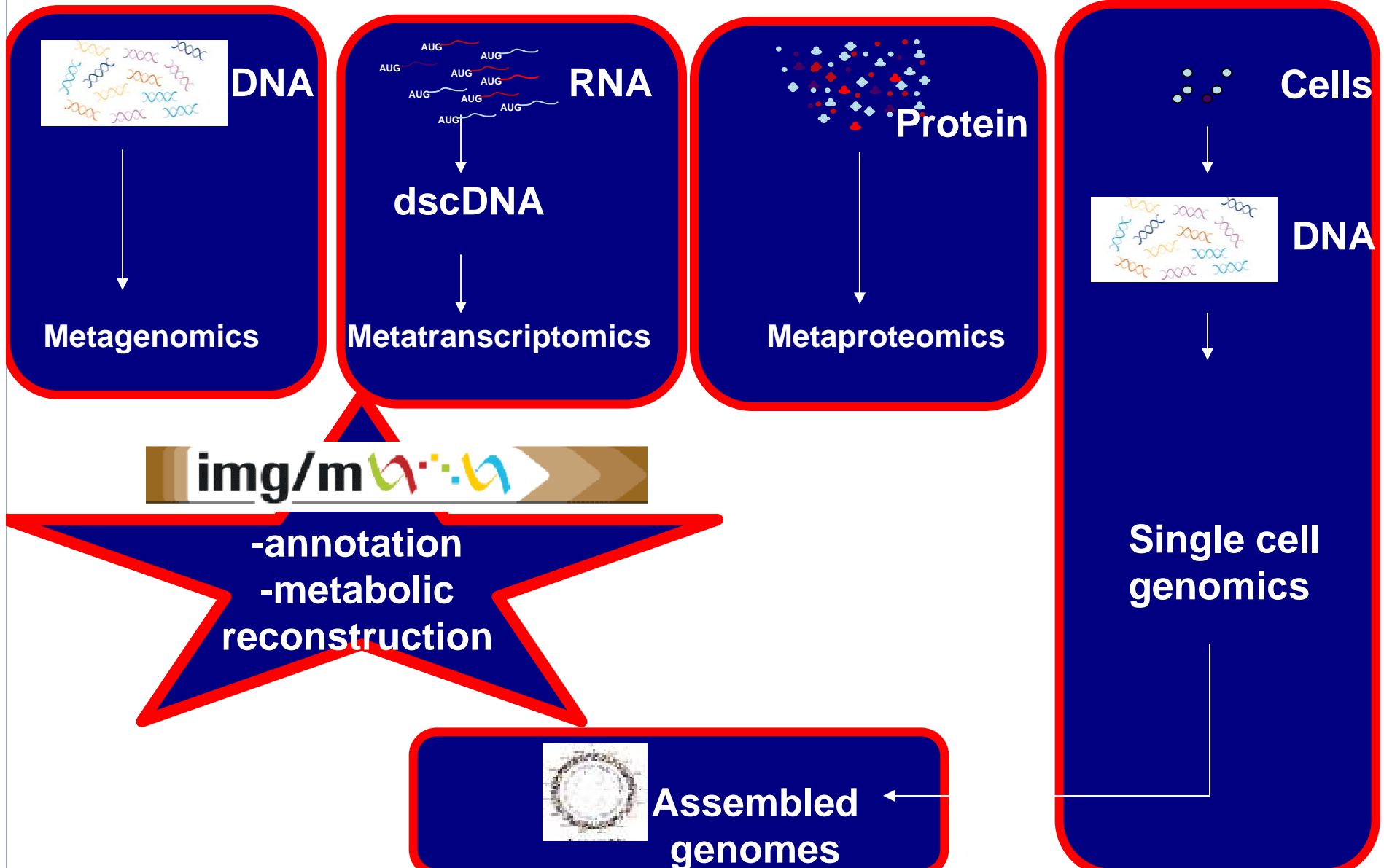
“ recommends the establishment of a small number of large-scale projects " unite scientists of multiple disciplines around the study of a particular habitat



JGI Metagenomics Program Looking Forward



Initiate User Calls for Large Scale (Terabase)
“Grand Challenge” Metagenomic Projects



Plans for Microbes and Metagenomics

In addition to existing user program engage community and panels of experts in developing plans:

Scale-up of GEBA and other GEBA like projects

Fungal version of “GEBA”

Large Scale Metagenomic Projects
(Grand Challenge Projects)

Outcomes that will be enabled by the proposed goals

- Contribute to the significant replacement of petroleum with biofuels
- Development of bioremediation approaches to clean up diverse contaminants and pollutants
- Understanding of and development of strategies for manipulating global C, N, cycles

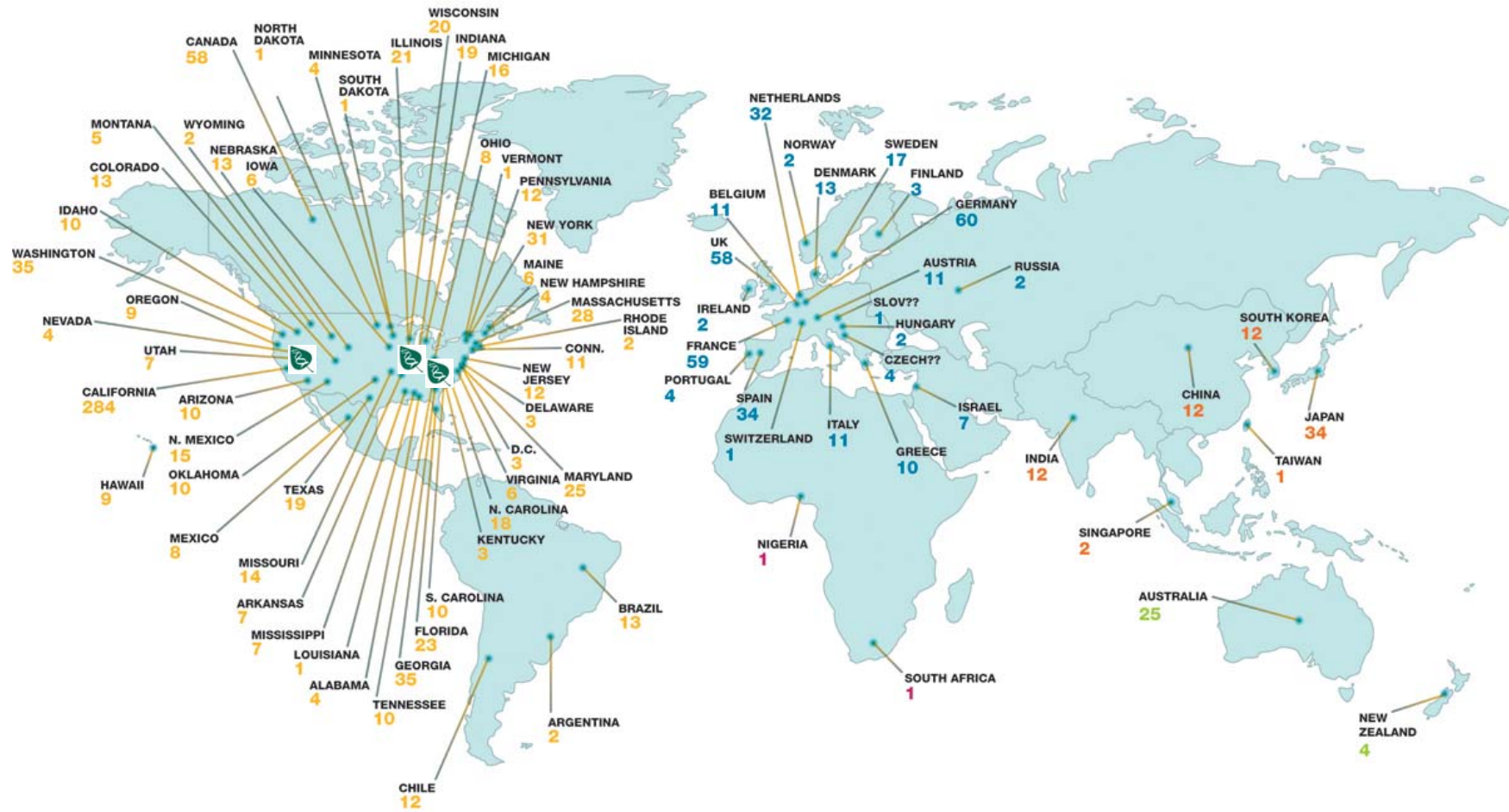
JGI User Programs

1351

JGI Users Worldwide in 2008

AMERICAS 856 EUROPE 336 AFRICA 2 ASIA 64 AUSTRALASIA 29

DOE
Bioenergy
Research
Centers



User Program in the Future

Continue What We Are Doing

Expand Pre and Post Sequencing
Capabilities

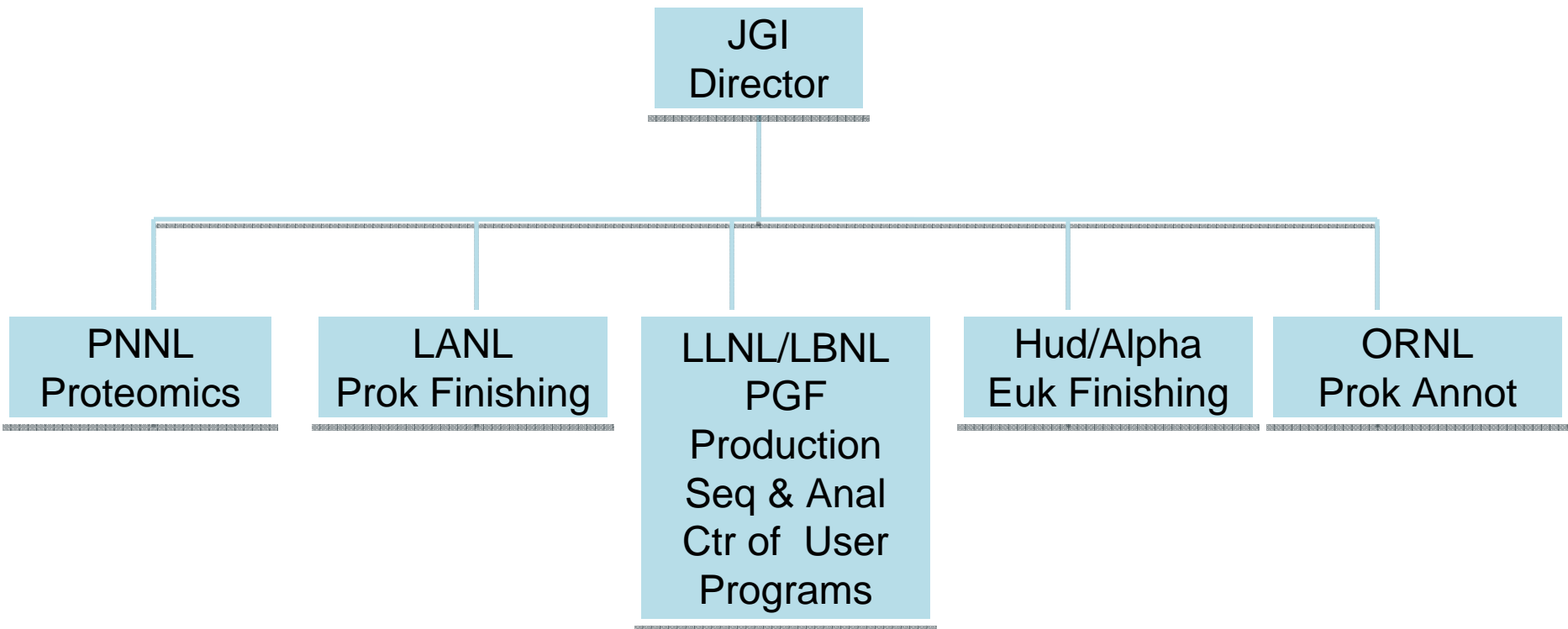
Develop Grand Challenge Programs

- Metagenomics Going Deep and/ or Going Broad
(Terrestrial version of the Global Ocean Survey, Biomass degrading environments,)
- Interfacing with other OBER Science Programs
(Joint JGI/EMSL Genomic / Proteomics, Linked to BER Scientific Focus Areas)
- Cross-cutting themes
(Plant/microbe interaction)

| | |
|-----------------|------|
| Peg Folta | LLNL |
| Karen Remington | NIH |
| Mike Himmel | NREL |
| Evi Dube | LLNL |
| Bob Cottingham | ORNL |
| Tom Slezak | LLNL |
| Susan Gregurick | OBER |
| Tom Brettin | LANL |

Present State

- Staying ahead of the rising data flow
- Providing basal annotation services targeted to the plant, microbial and metagenomic data generated by the JGI

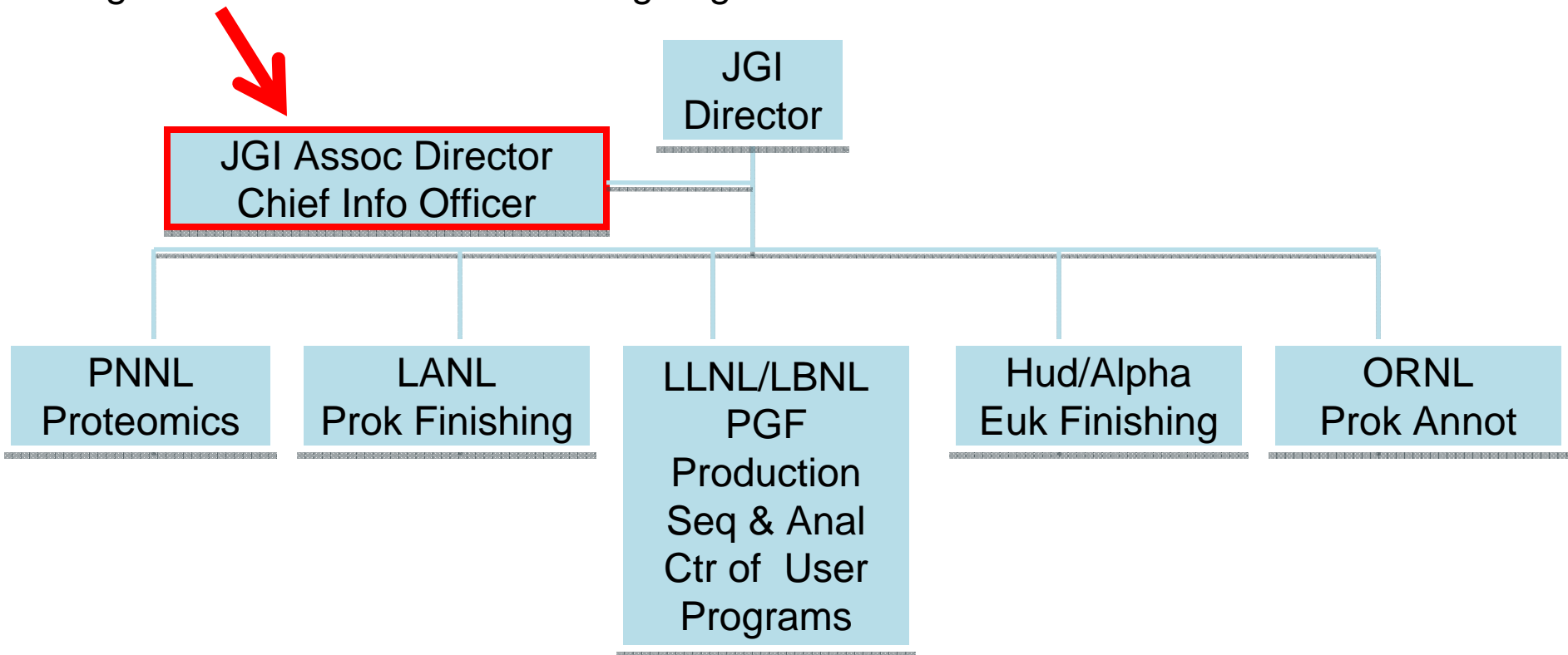


JGI Informatics Vision (work in progress)

Systems integrated within and among the JGI partners, as well as with other DOE BER activities and facilities and related national and international efforts

A one-stop user facility to sequence, annotate and analyze all JGI projects, with **exceptional services enabling user access to different data sets and fascile tools for the conversion of data into knowledge**

An organization that delivers leading edge bioinformatic solutions



JGI Value Added: Sample Processing & Analysis

**Sequence
Generator**



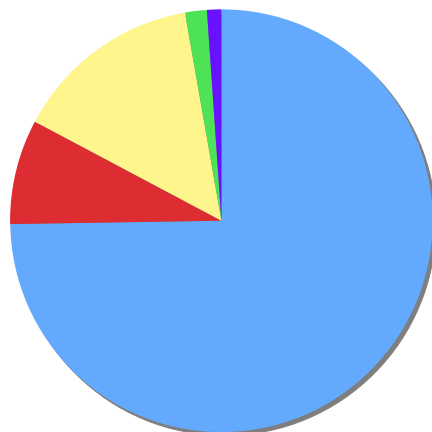
ANALYSIS

It is not about the machines

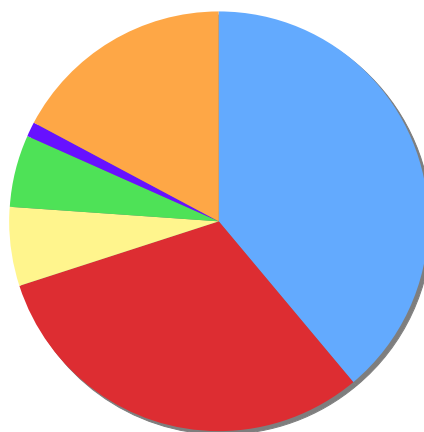


Sequencing by Program

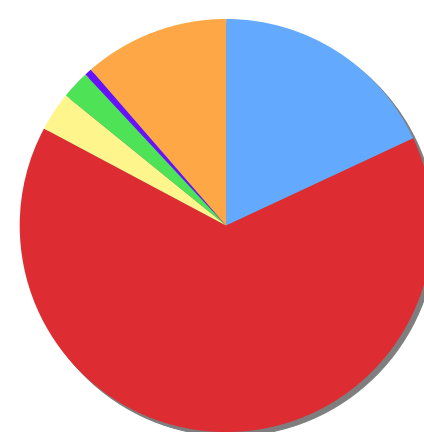
Sanger (48M reads)



454 (603 runs)



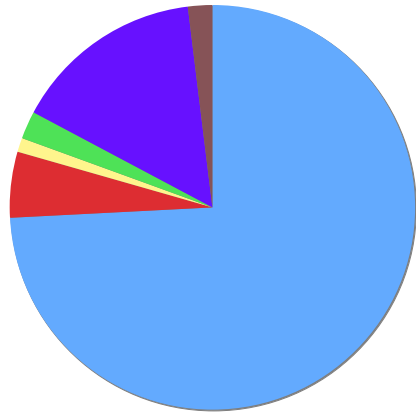
Illumina (1800 Channels)



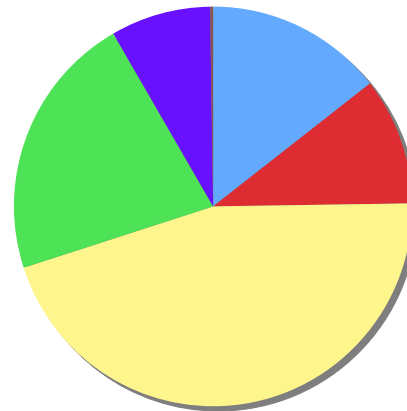
■ CSP ■ BRC ■ DOE Microbe ■ LSP ■ Dir Science ■ GEBA

FY09 Sequence Commitments:

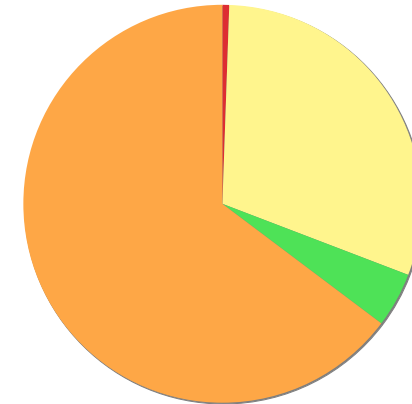
Capillary
(48M lanes, 30 Gb)



454
(603 runs, 200 Gb)



Illumina
(225 runs, 450 Gb)



■ Plant WGS
 ■ Sm Euk WGS
 ■ Prok WGS
 ■ ESTs
 ■ Metagenomes
 ■ Reseq