OptIPuter:
Metagenomics at Light Speed

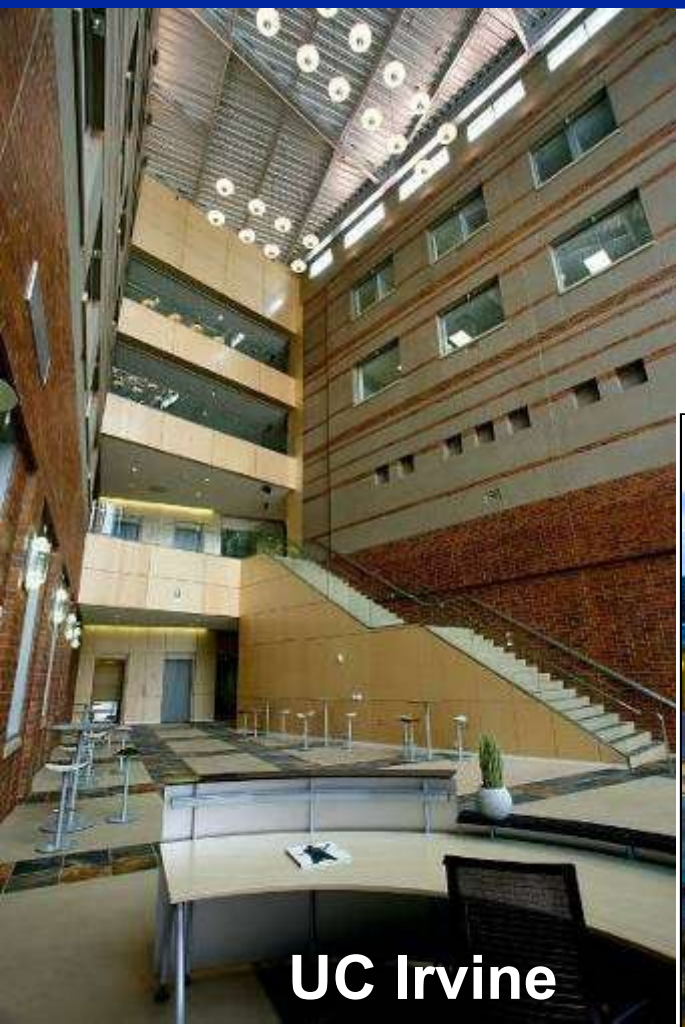
Presentation
National LambdaRail Booth
Supercomputing 2006
Tampa, FL
November 14, 2006

Dr. Larry Smarr
Director, California Institute for Telecommunications and Information Technology
Harry E. Gruber Professor,
Dept. of Computer Science and Engineering
Jacobs School of Engineering, UCSD

UCSD
UCIrvine

it²

Phase II:Two New Calit2 Buildings Provide ~340,000 GSF and New Laboratories for "Living in the Future"



- Over 1000 Researchers in Two Buildings
  - Linked via Dedicated Optical Networks
  - International Conferences and Testbeds
- New Laboratories
  - Nanotechnology
  - Virtual Reality, Digital Cinema



Preparing for a World in Which Distance is Eliminated…

UCSD
UCIrvine

www.calit2.net

it2

The OptIPuter Project – Creating High Resolution Portals Over Dedicated Optical Channels to Global Science Data

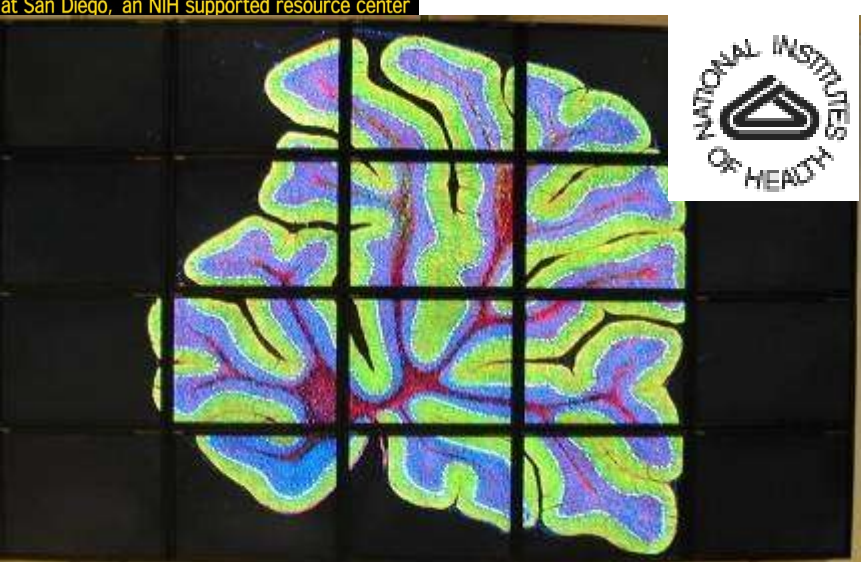
- **NSF Large Information Technology Research Proposal**
  - Calit2 (UCSD, UCI) and UIC Lead Campuses—Larry Smarr PI
  - Partnering Campuses: SDSC, USC, SDSU, NCSA, NW, TA&M, UvA, SARA, NASA Goddard, KISTI, AIST, CRC(Canada), CICESE (Mexico)
- **Industrial Partners**
  - IBM, Sun, Telcordia, Chiaro, Calient, Glimmerglass, Lucent
- **$13.5 Million Over Five Years—Now In the Fifth Year**

OptIPuter

NSF

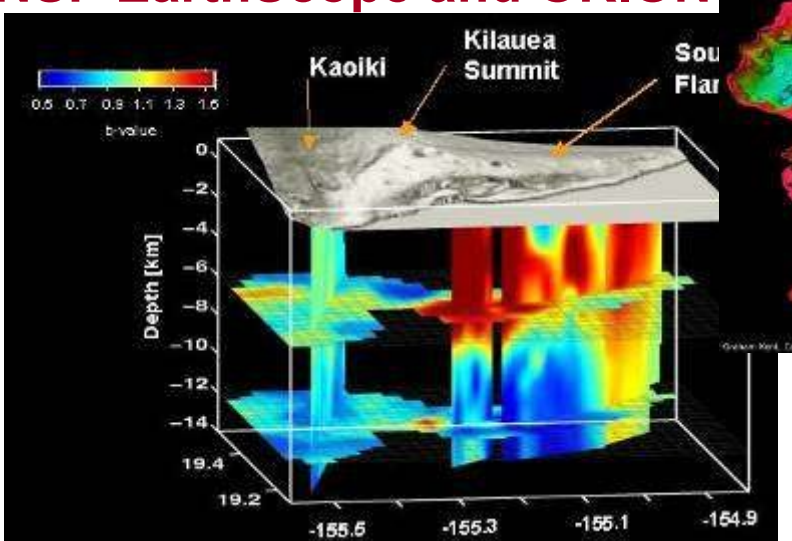
NIH Biomedical Informatics Research Network

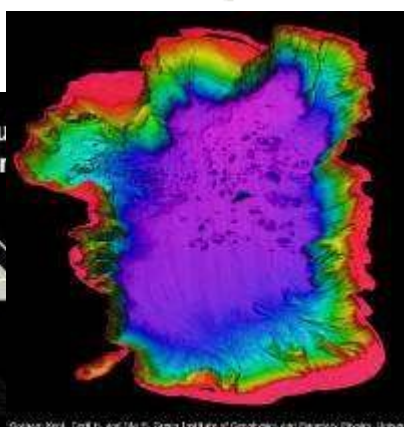
NCMIR
NATIONAL CENTER for MICROSCOPY and IMAGING RESEARCH
at San Diego, an NIH supported resource center



NATIONAL INSTITUTES OF HEALTH

NSF EarthScope and ORION

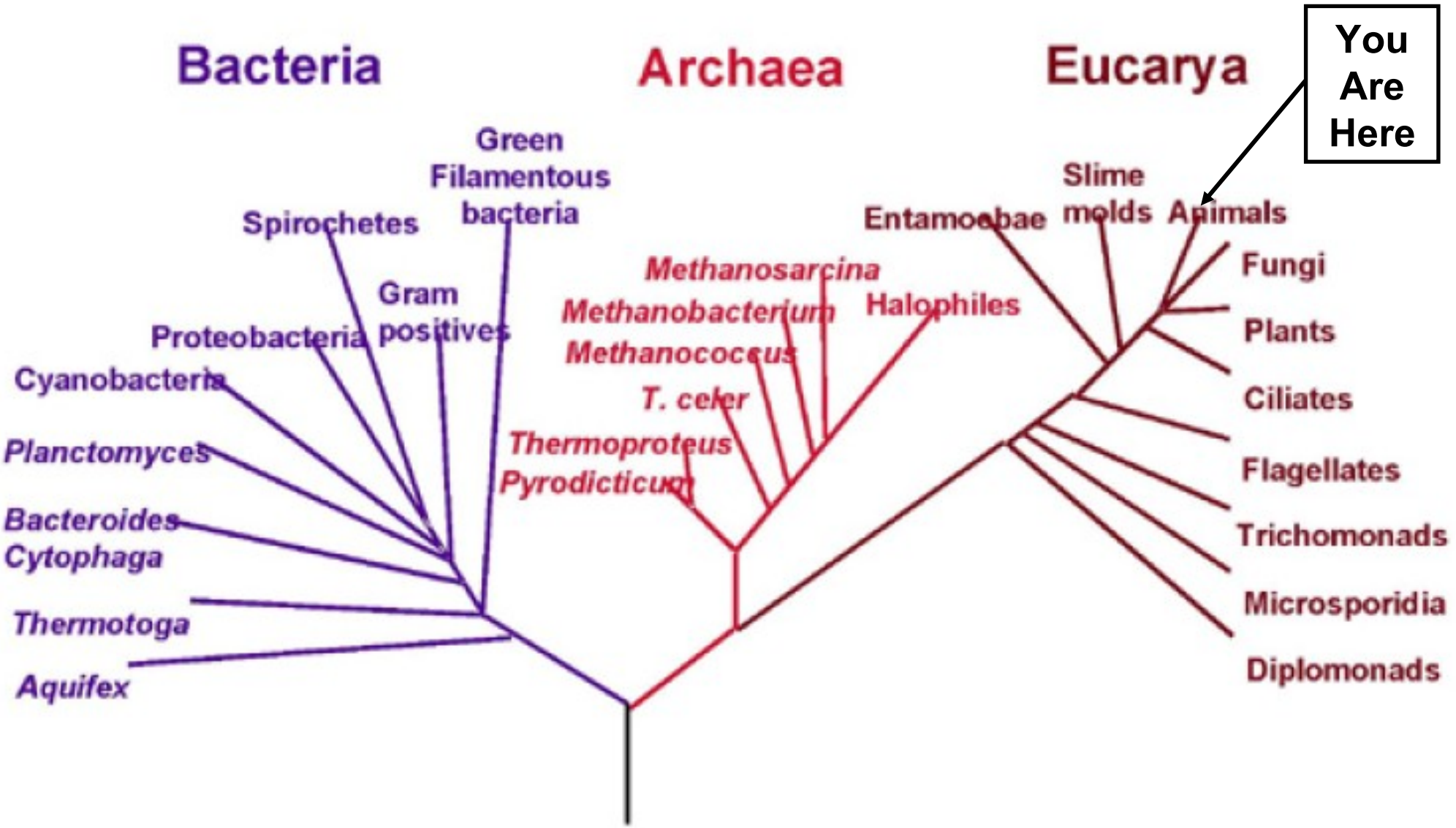




SCRIPPS INSTITUTION OF OCEANOGRAPHY

it²

Most of Evolutionary Time Was in the Microbial World

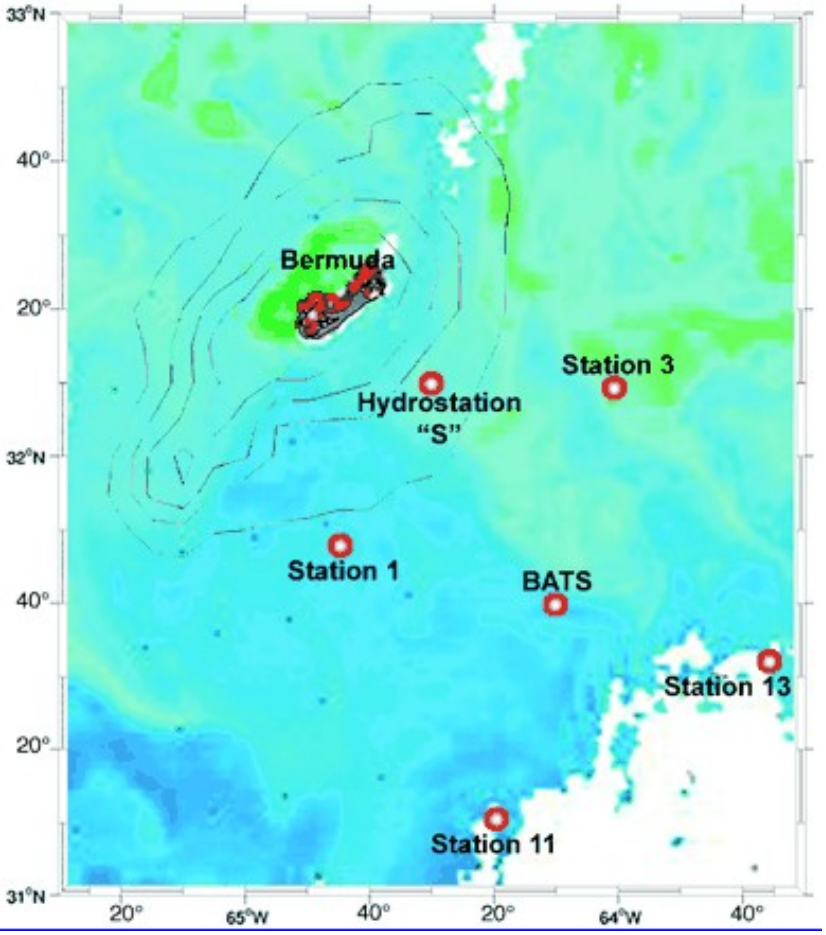


Tree of Life Derived from 16S rRNA Sequences

Source: Carl Woese, et al

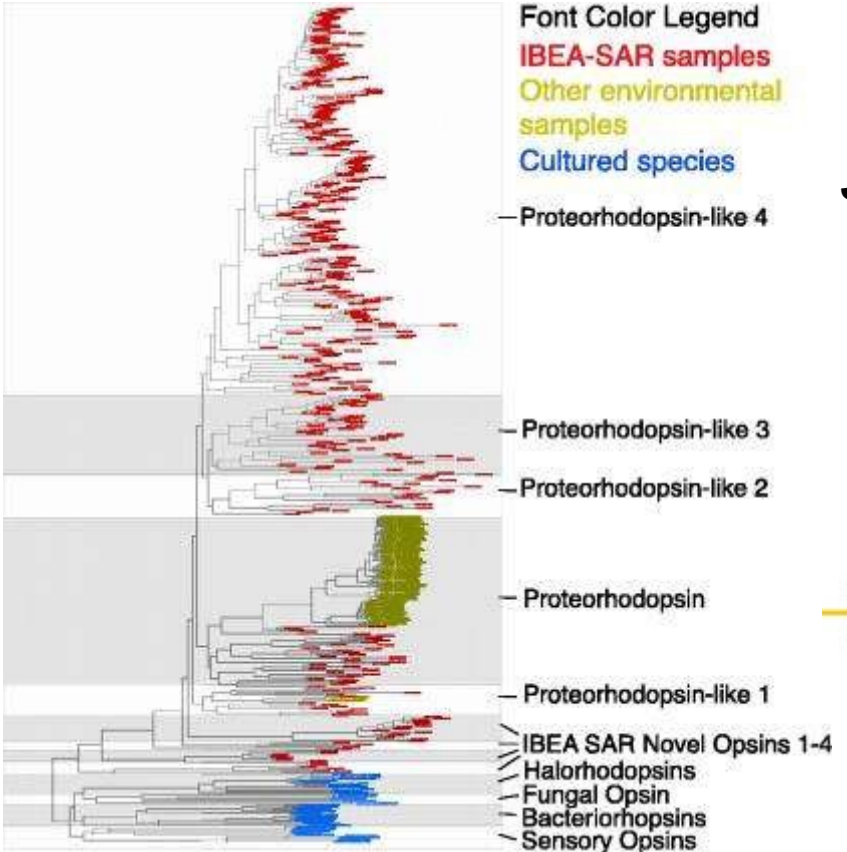
UCSD
UCIrvine

The Sargasso Sea Experiment
The Power of Environmental Metagenomics



MODIS-Aqua satellite image of ocean chlorophyll in the Sargasso Sea grid about the BATS site from 22 February 2003

- Yielded a Total of Over 1 billion Base Pairs of Non-Redundant Sequence
- Displayed the Gene Content, Diversity, & Relative Abundance of the Organisms
- Sequences from at Least 1800 Genomic Species, including 148 Previously Unknown
- Identified over 1.2 Million Unknown Genes



J. Craig Venter, et al.
Science
2 April 2004:
Vol. 304.
pp. 66 - 74

UCSD
UCIrvine

J. Craig Venter INSTITUTE

it²

Marine Genome Sequencing Project – Measuring the Genetic Diversity of Ocean Microbes



MARINE MICROBIOLOGY INITIATIVE

Gordon and Betty MOORE FOUNDATION

Sorcerer II Data Will Double Number of Proteins in GenBank!

UCSD
UCIrvine

J. Craig Venter INSTITUTE

it²

CAMERA: Community Cyberinfrastructure for Advanced Marine Microbial Ecology Research and Analysis

National LambdaRail
Direct Connect
Computation and Storage Complex

Funded by: Gordon and Betty Moore Foundation

PI Larry Smarr

Joint Partnership of:

UCSD
UCIrvine

Announced January 17, 2006
$24.5M Over Seven Years

Building On Centers of Excellence and Achievement









Powered by Rocks









The Scripps Genome Center

Supported by:











UCSD
UCIrvine

it²

Calit2's Direct Access Core Architecture Will Create Next Generation Metagenomics Server

Gordon and Betty MOORE FOUNDATION

Sargasso Sea Data

Sorcerer II Expedition (GOS)

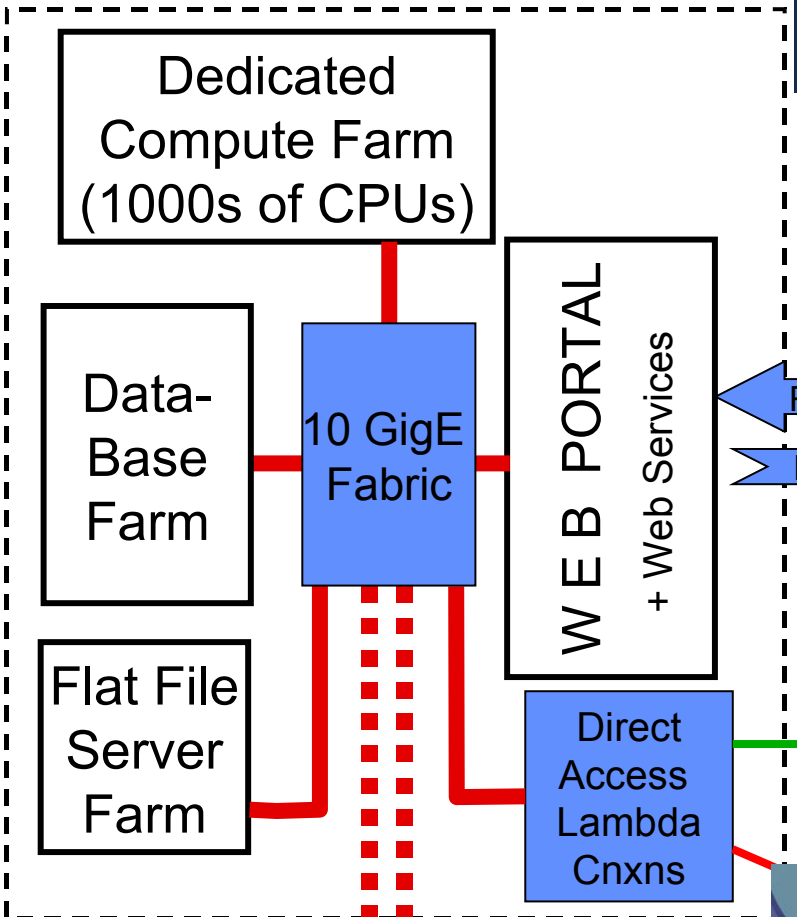
JGI Community Sequencing Project

Moore Marine Microbial Project

NASA and NOAA Satellite Data

Community Microbial Metagenomics Data

Powered by Rocks



Source: Phil Papadopoulos, SDSC, Calit2

CENTER FOR EARTH OBSERVATIONS & APPLICATIONS

SCRIPPS INSTITUTION OF OCEANOGRAPHY

SDSC SAN DIEGO SUPERCOMPUTER CENTER

UCSD UCIrvine

J. Craig Venter INSTITUTE

it²

My OptIPortal™ – Affordable Termination Device for the OptIPuter Global Backplane

- 20 Dual CPU Nodes, 20 24” Monitors, ~$50,000
- 1/4 Teraflop, 5 Terabyte Storage, 45 Mega Pixels--Nice PC!
- Scalable Adaptive Graphics Environment (SAGE) Jason Leigh, EVL-UIC





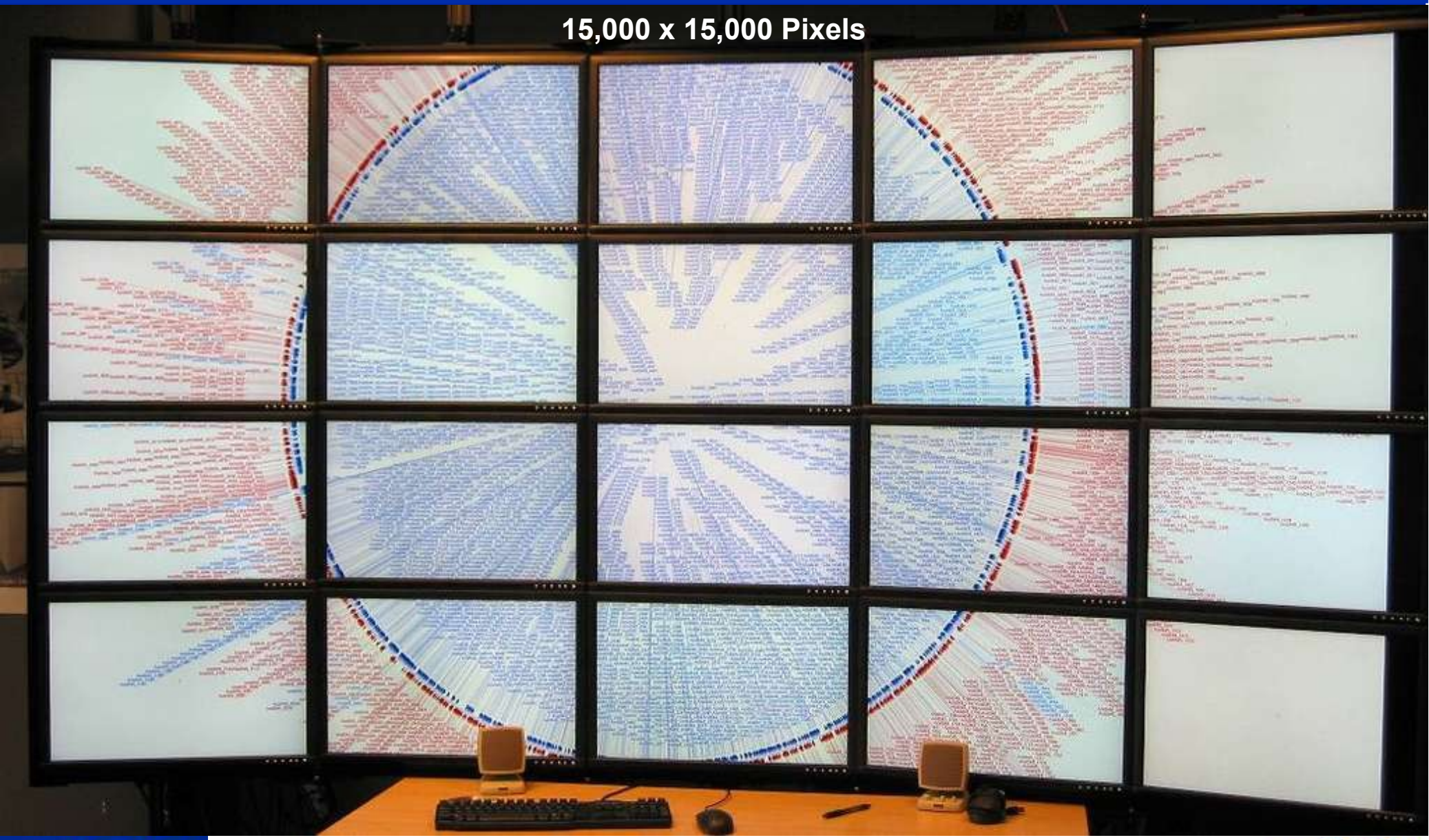
UCSD UCIrvine

Powered by Rocks ROCKS

Source: Phil Papadopoulos SDSC, Calit2

it²

Use of OptIPortal to Interactively View Microbial Genome



UCSD
UCIrvine

Source: Raj Singh, UCSD

Acidobacteria bacterium Ellin345 (NCBI)
Soil Bacterium 5.6 Mb

it²

Use of OptIPortal to Interactively View Microbial Genome



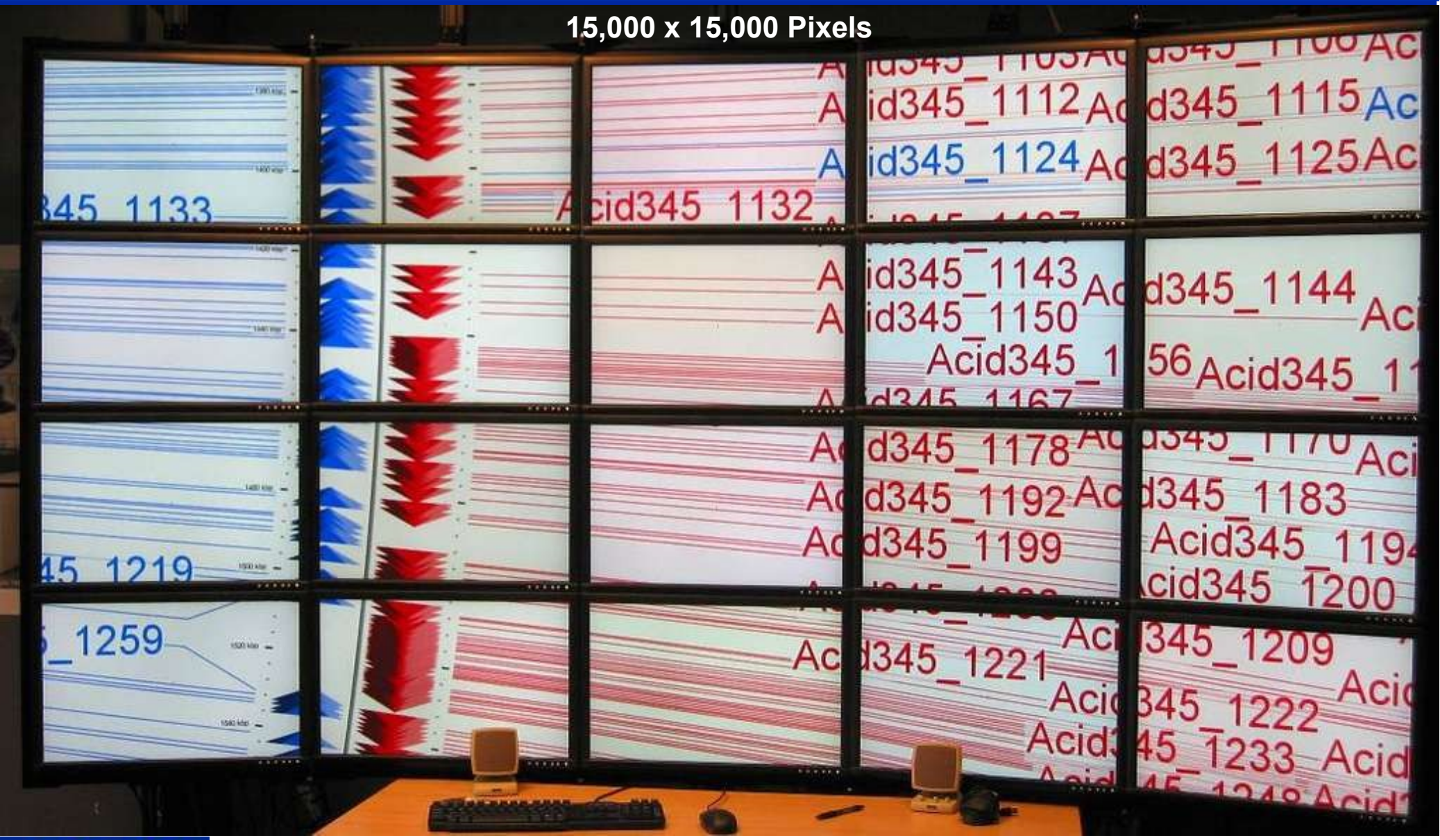
UCSD
UCIrvine

Source: Raj Singh, UCSD

Acidobacteria bacterium Ellin345 (NCBI)
Soil Bacterium 5.6 Mb

it²

Use of OptIPortal to Interactively View Microbial Genome



UCSD
UCIrvine

Source: Raj Singh, UCSD

Acidobacteria bacterium Ellin345 (NCBI)
Soil Bacterium 5.6 Mb

it²

Calit2 is Now OptIPuter Connecting Remote Moore-Funded Microbial Researchers with OptIPortals

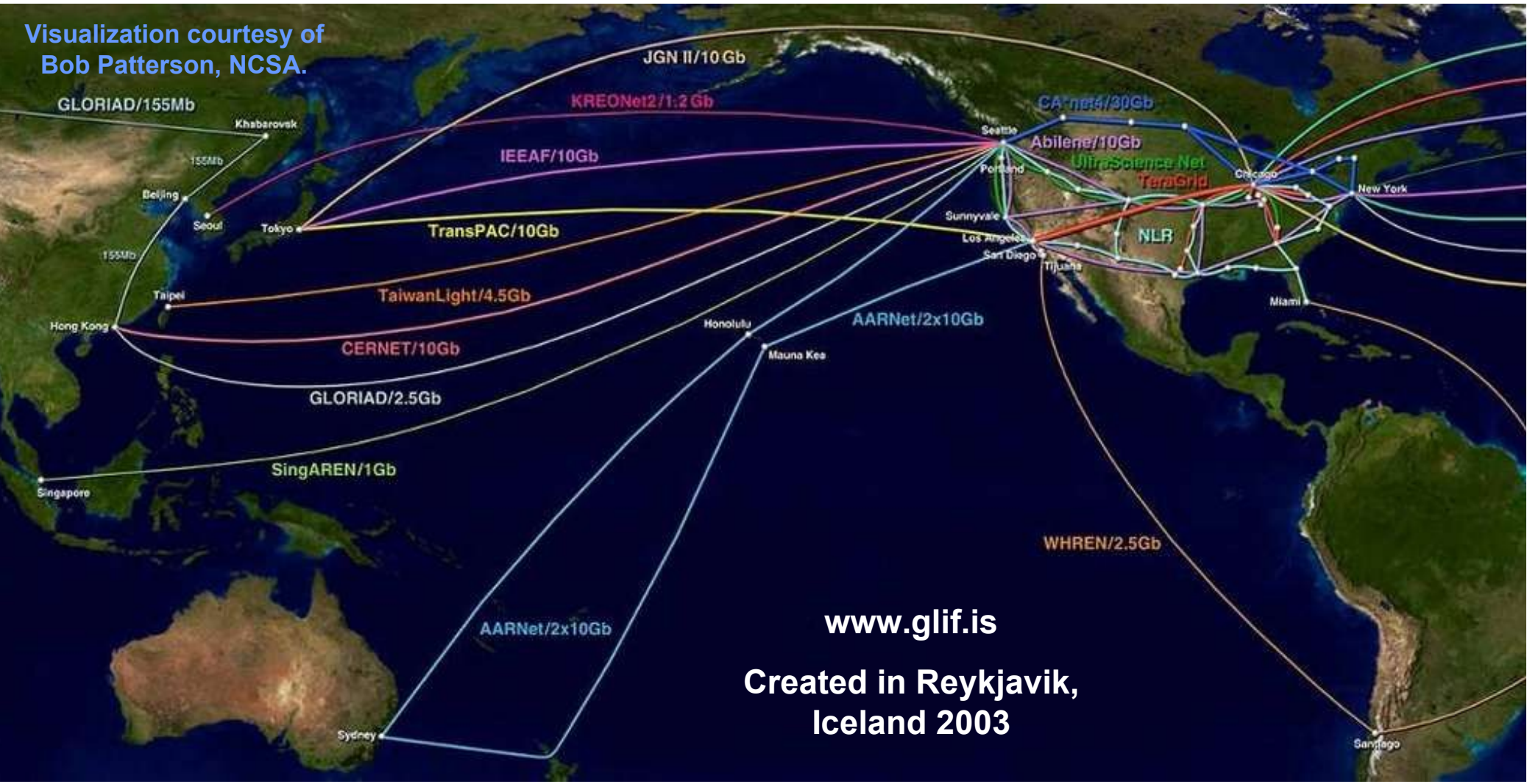


A Near Future Metagenomics Fiber Optic-Enabled Data Generator



Source John Delaney, UWash

Countries are Aggressively Creating Gigabit Services: Interactive Access to CAMERA Data System



UCSD
UCIrvine

it2