

# iPlant: Cyberinfrastructure for Plant Sciences (and Beyond)

Nirav Merchant

[nirav@email.arizona.edu](mailto:nirav@email.arizona.edu)



# One Big Problem...



Thursday, March 1, 12



# One Big Problem...

2008



2011



# One Big Problem...



Thursday, March 1, 12



# One Big Problem...

Published online 3 September 2008 | *Nature* 455, 16-21 (2008) |  
doi:10.1038/455016a

News Feature

## Big data: Welcome to the

What does it take to store bytes by the terabytes? It's trillions? Come along with us as we which it's all about.

### Big data: Distilling meaning from mountains of data

Felice Frankel

*Nature* 455, 28-29

**Buried in vast amounts of data, scientists need to crack the code.** Rosalind Reid

Clifford Lynch

1. Clifford Lynch

Circ. Natu

Unive

Ema

Cor

Scienc

long ha

Lynch.

rese

scienc

- Transfer
- Storage
- Analysis
- Visualization
- Metadata Mark-up
- Search and Discover
- Share/Collaborate
- Publish

Winston Hide<sup>6,7</sup>, David P. Hill<sup>8</sup>, Renate Kania<sup>9</sup>, Mary Schaeffer<sup>10,11</sup>, Susan St Pierre<sup>12</sup>, Simon Twigger<sup>13</sup>, Owen White<sup>14</sup> & Seung Yon Rhee<sup>15</sup>

To thrive, the field that links biologists and their data urgently needs structure, recognition and support.



3 September 2008

3 September 2008

Information 21 Dupont

Published online 3 September 2008

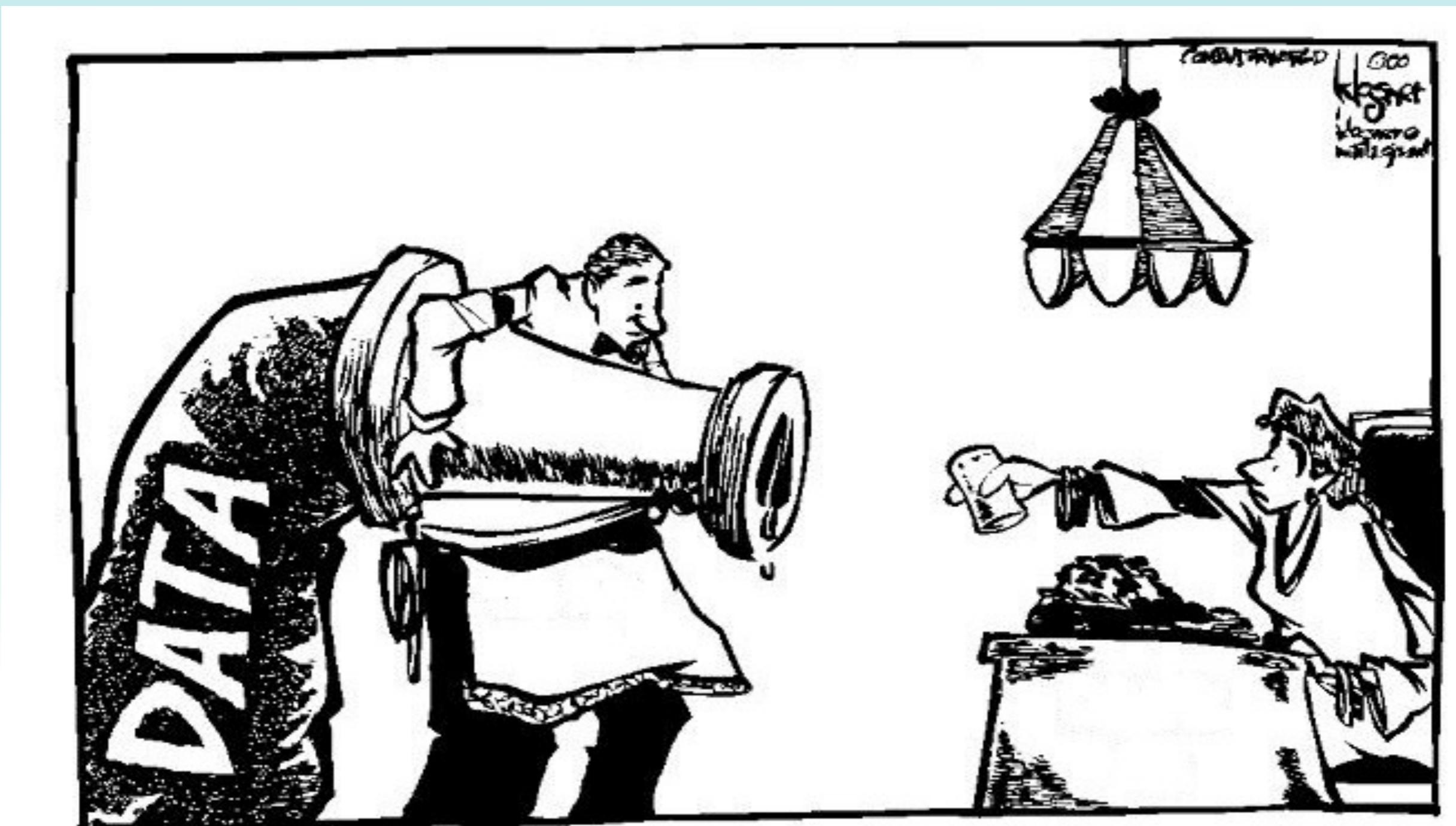
038/455047a; Published online 3 September 2008

uration

y<sup>3</sup>, Takashi Gojobori<sup>4</sup>, Linda Hannick<sup>5</sup>,



# got data ?



Thursday, March 1, 12



# Flavors of Bio-Information

- Sequences (most popular)
- Structures
- Images
- Video
- Audio
- Pathways (graphs)
- Text (Publications)
- Traces
- Combination (eg Video & Traces)
- And much more ...

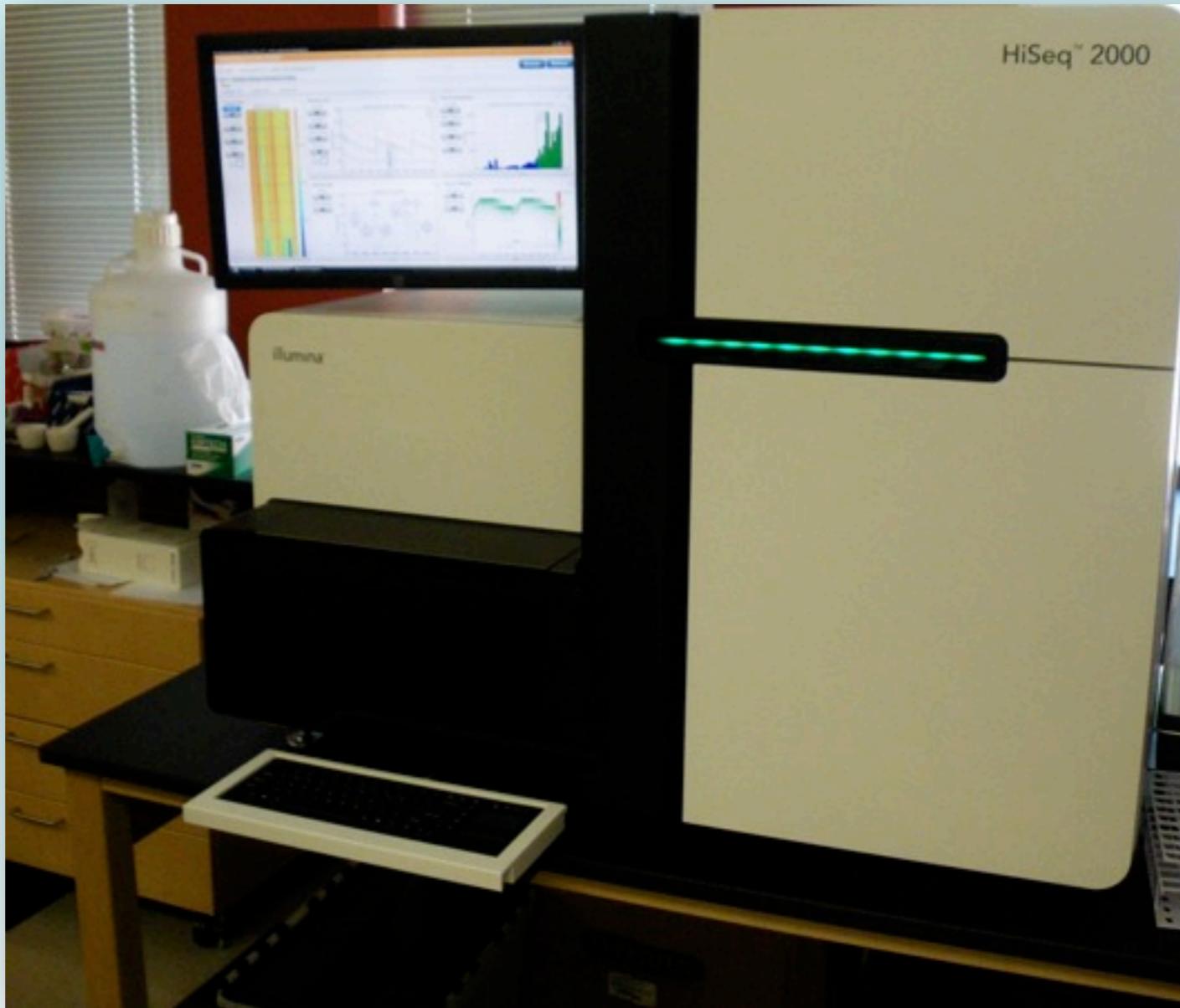


# Growth

- The Human genome project (1990-2003)
  - 13 years • 23 labs • \$500 Million.
- A Human genome today (2010)
  - ~3 days • ~1 machine. • ~\$10,000.
- Many 1000 and 10,000 genomes projects (human, plant etc)
- Bench-top sequencers are
- “Next-next-next” generation sequencers will produce \$500 genome in 5 years



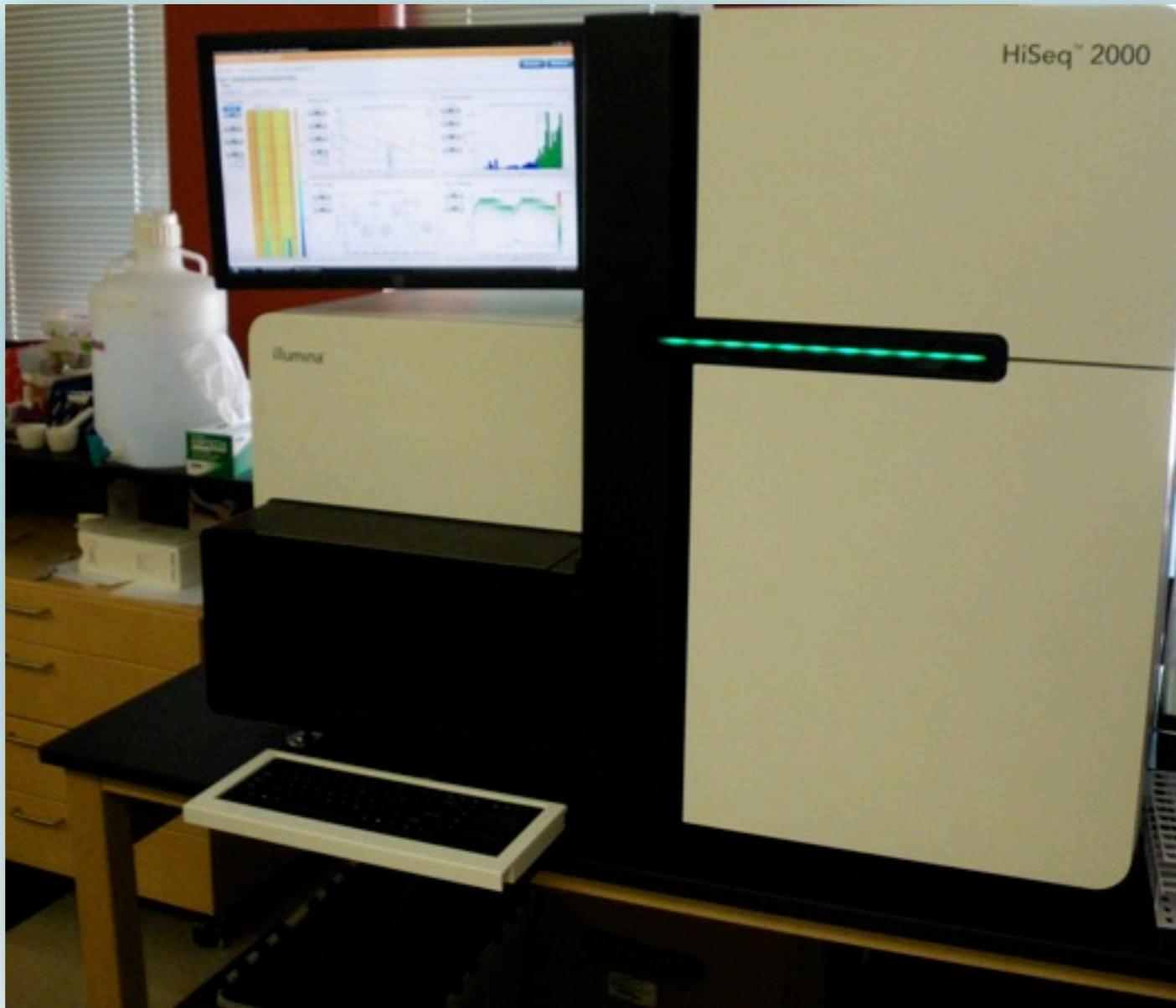
# High-throughput Data Acquisition



Thursday, March 1, 12



# High-throughput Data Acquisition



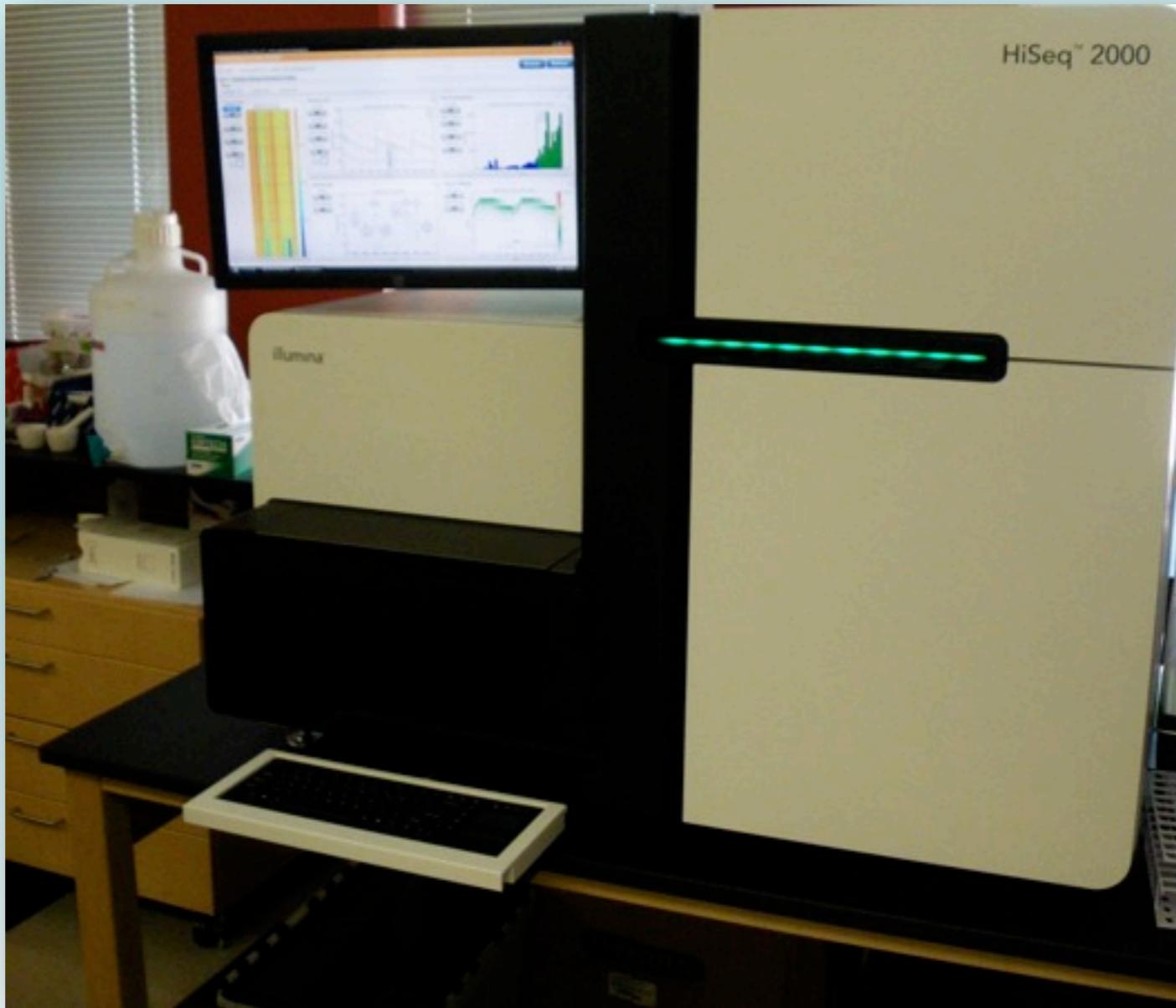
In 11 Days



Thursday, March 1, 12



# High-throughput Data Acquisition

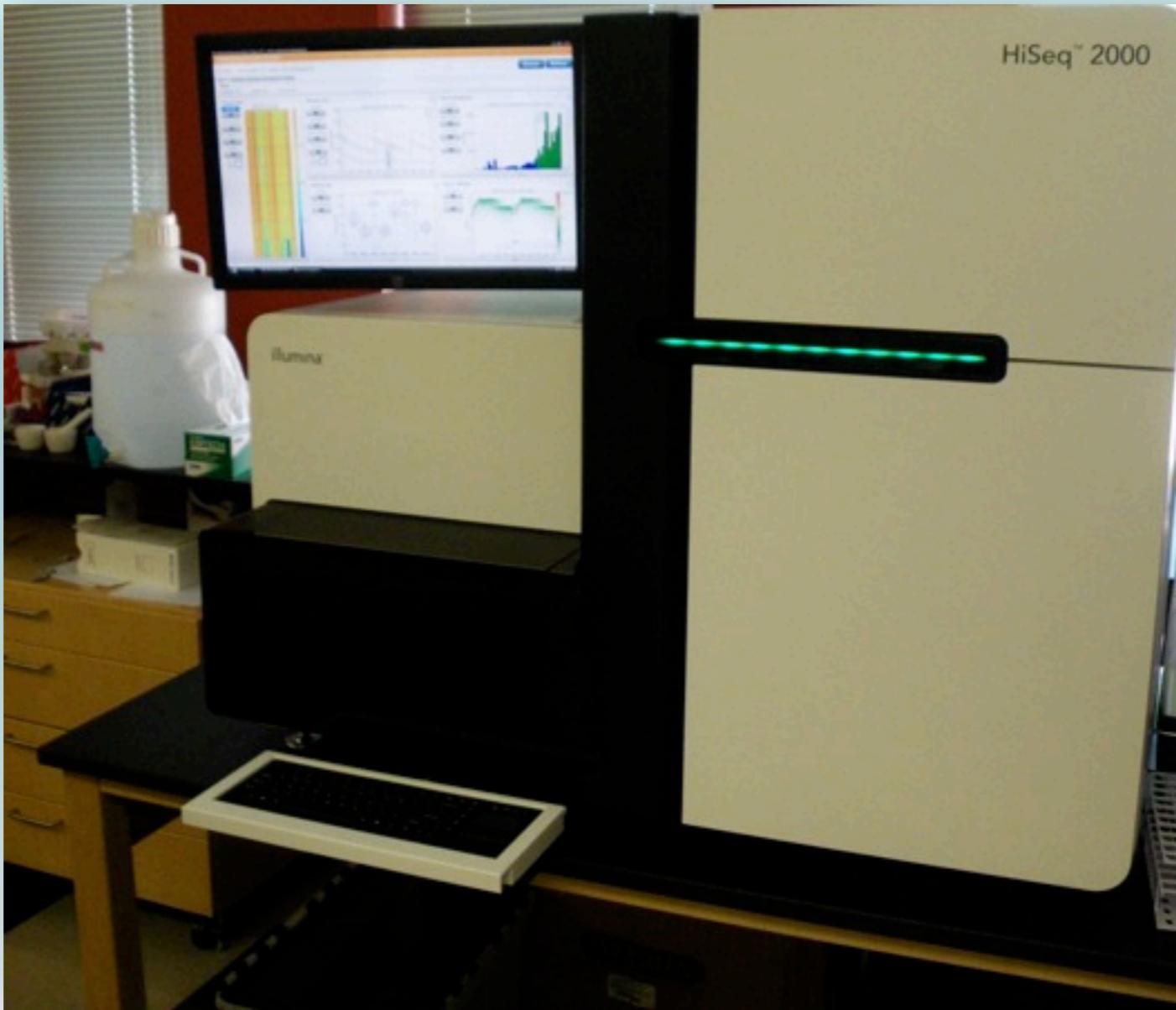


In 11 Days

- Generates 4TB of raw data



# High-throughput Data Acquisition

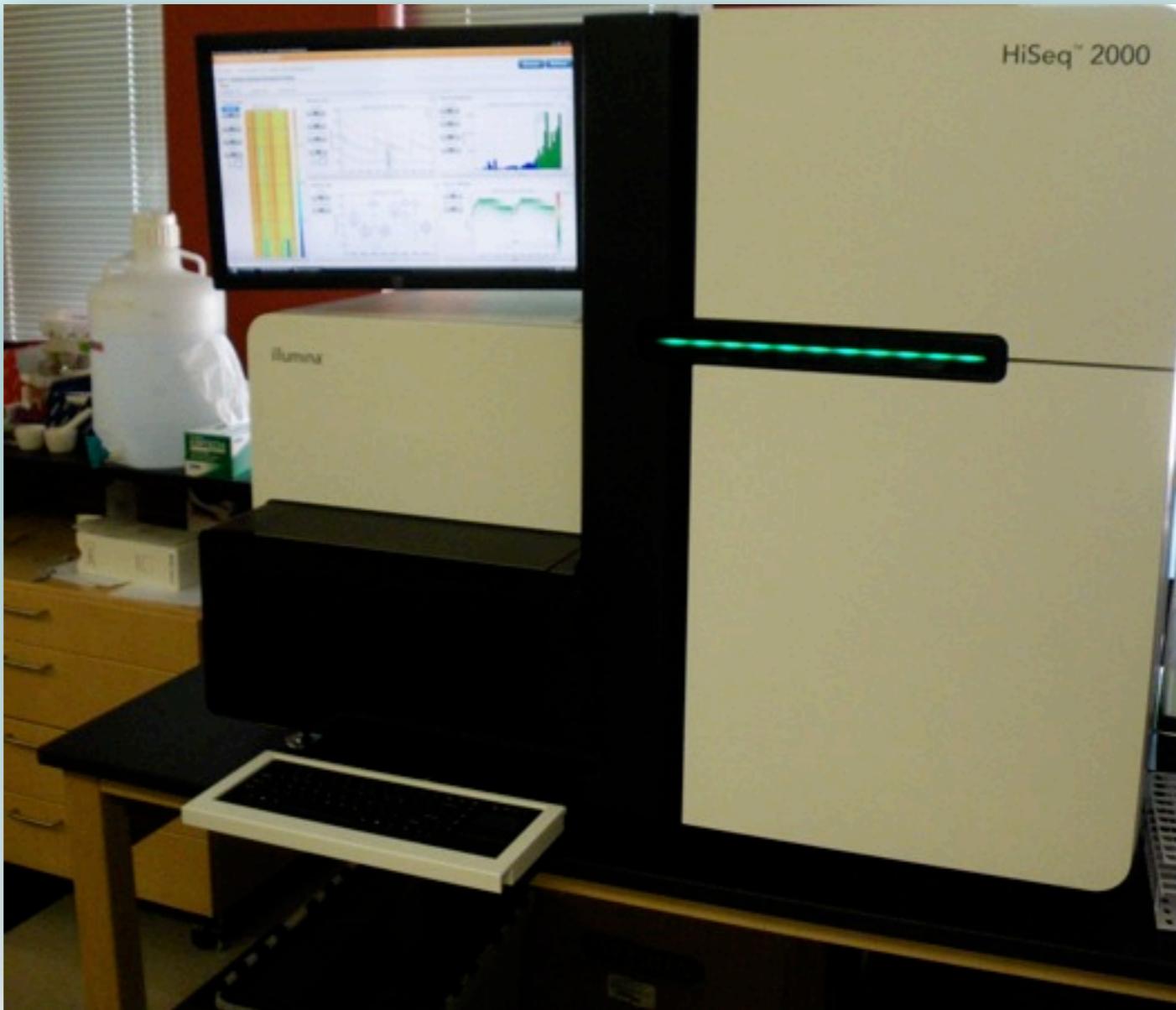


In 11 Days

- Generates 4TB of raw data
- 600,000,000,000 bases of DNA sequence



# High-throughput Data Acquisition



In 11 Days

- Generates 4TB of raw data
- 600,000,000,000 bases of DNA sequence
  - (200 human genomes)



# High-throughput Phenotyping

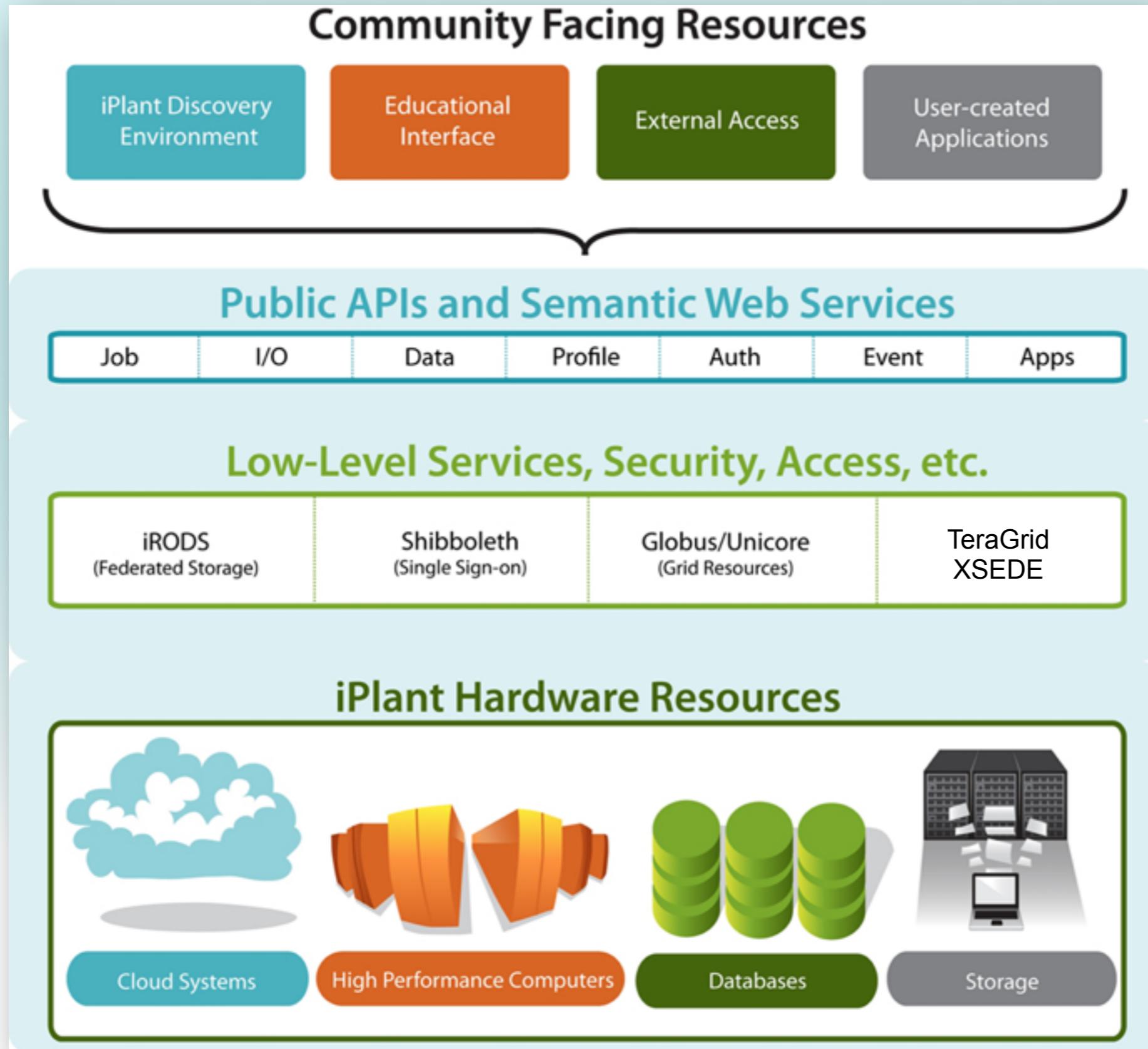
## (Watching Grass Grow)



- \$70K for ~30 camera sets
- ~200 movies of plants undergoing a dynamic growth process
- “Only” 4GB a day



# iPlant Layered Services and Access

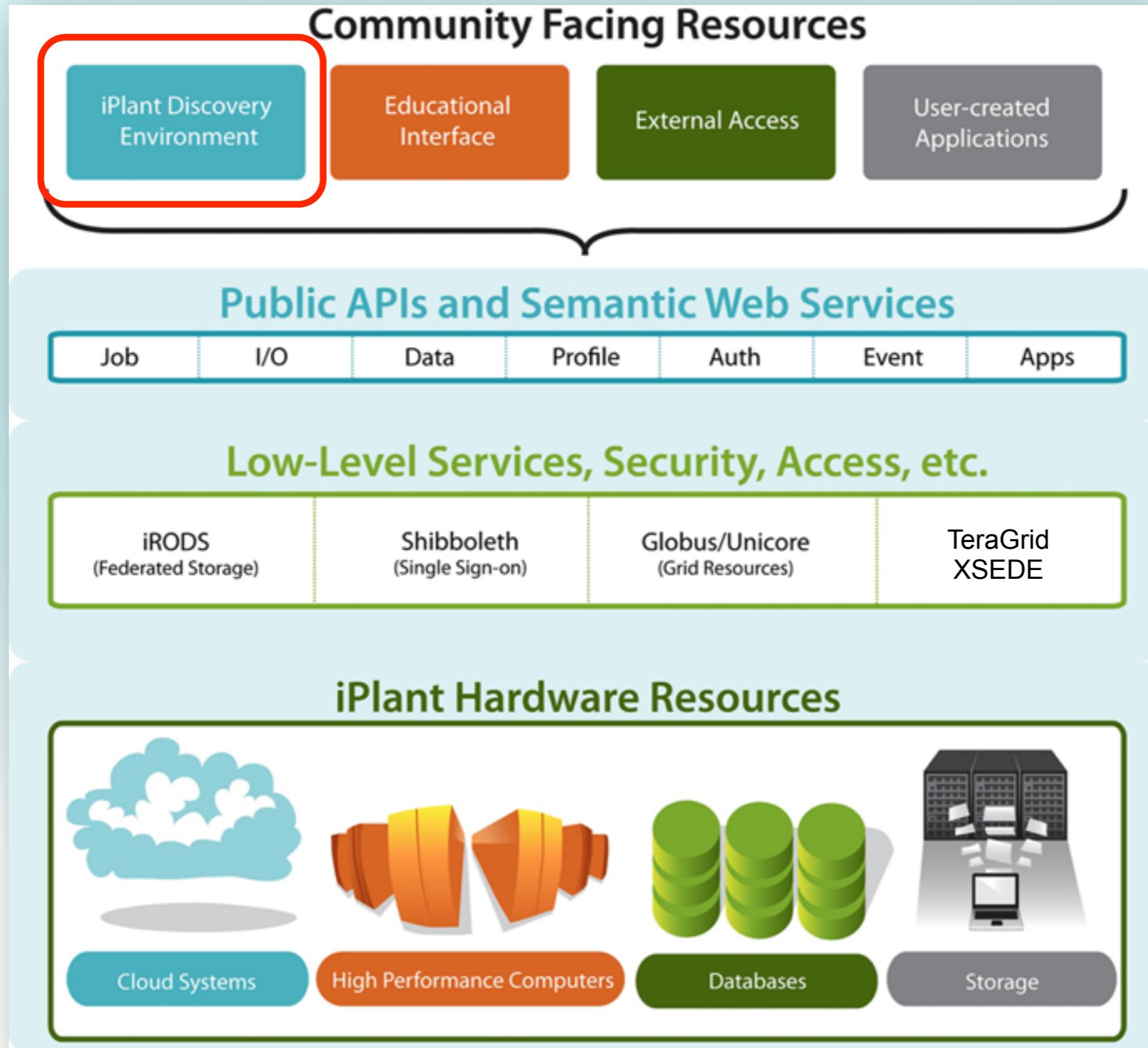


End  
Users

Computational  
Users



# iPlant Layered Services and Access



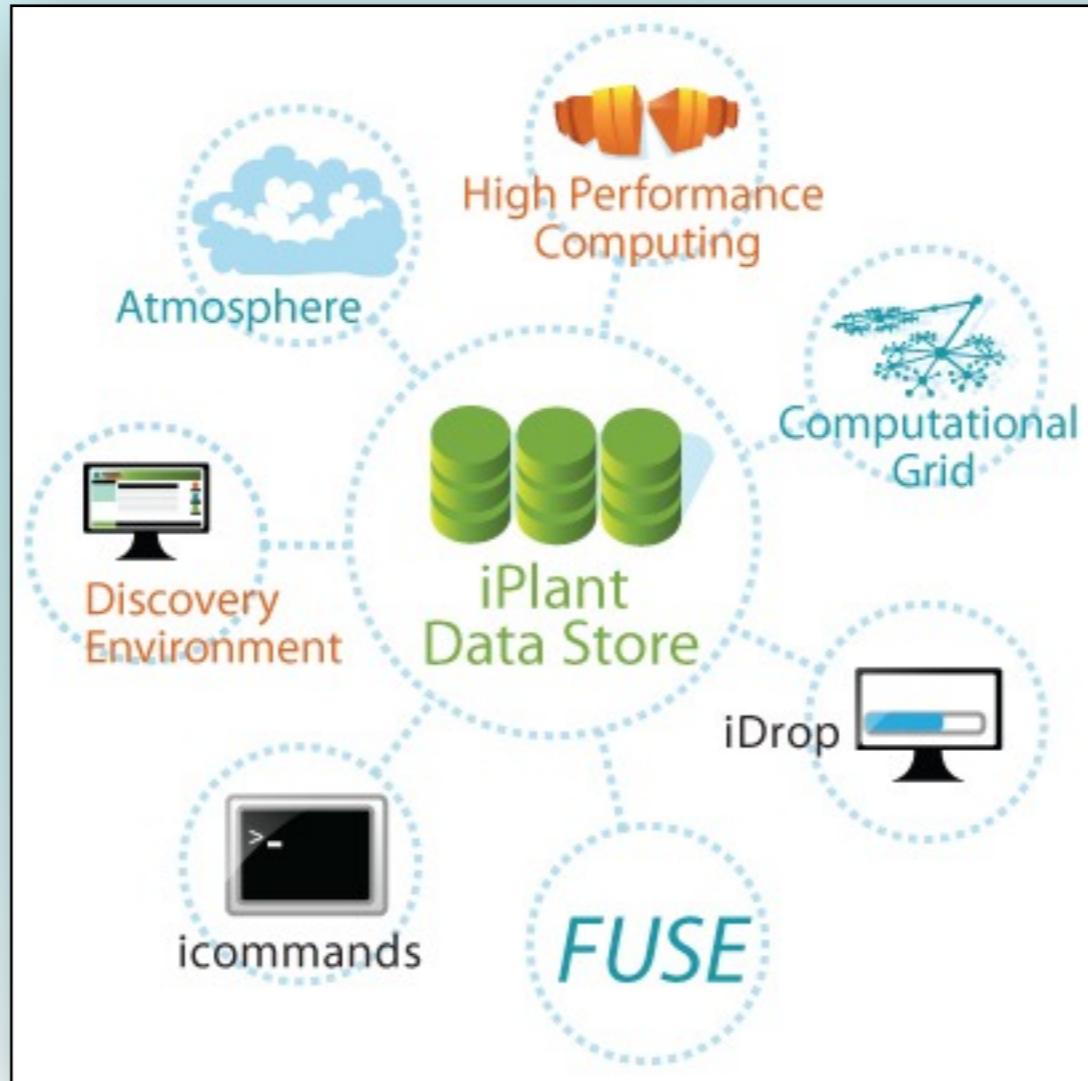
End  
Users

Computational  
Users



# iPlant Data Store

## Free Your Data



Different Users,  
Different Access Needs:  
One Data Store



# iPlant Data Store

## Free Your Data



Thursday, March 1, 12



# iPlant Data Store

## WebDAV      Free Your Data

A screenshot of a Mac OS X Finder window. The title bar says "Finder" and the address bar shows "https://gucumatz.iplantcollaborative.org/idp/Authn/UserPassword". The main content area displays the iPlant Collaborative login page with the heading "Welcome!" and a list of features.

**iPlant Collaborative™ Empowering A New Plant Biology**

**Welcome!**

You've reached the login page for the Discovery Environment. The Discovery Environment integrates powerful, community-recommended software tools into a system that:

- Can handle terabytes of data
- Utilizes high-performance supercomputing resources as needed
- Hides the complexity needed to do these tasks.

©2011 iPlant Collaborative. The iPlant Collaborative is funded by a grant from the National Science Foundation (#DBI-0735191).



# WebDAV

# iPlant Data Store

## Free Your Data

## DE

A screenshot of a Mac OS X Finder window. The title bar says "Finder File Edit View Go Window Help". The address bar shows "iPlant Collaborative Discover" and the URL "https://gucumatz.iplantcollaborative.org/idp/Authn/UserPassword". Below the address bar are icons for CoGe, CoGe: OrgView, and Shorten. The main content area displays the iPlant Collaborative logo and the tagline "Empowering A New Plant Biology". A large orange "Welcome!" heading is followed by a paragraph of text: "You've reached the login page for the Discovery Environment. The Discovery Environment integrates powerful, community-recommended software tools into a system that:" and a bulleted list: "Can handle terabytes of data", "Utilizes high-performance supercomputing resources as needed", and "Hides the complexity needed to do these tasks.". At the bottom, there is a copyright notice: "©2011 iPlant Collaborative. The iPlant Collaborative is funded by a grant from the National Science Foundation (#DBI-0735191)."

A screenshot of the iPlant Collaborative Discovery Environment homepage. The header features the iPlant Collaborative logo and the text "Discovery Environment". On the right, there are links for "elyons", "Help", and "Notifications". The main area has three buttons: "Data" (with a book icon), "Analyses" (with a monitor icon), and "Apps" (with a document icon). To the right is a large, stylized blue circular logo. At the bottom, there is a copyright notice: "©2011 iPlant Collaborative" and "The iPlant Collaborative is funded by a grant from the National Science Foundation (#DBI-0735191)."



# WebDAV

# iPlant Data Store

## Free Your Data

## DE

A screenshot of a Mac OS X Finder window. The title bar says "Finder File Edit View Go Window Help". The address bar shows "iPlant Collaborative Discover" and the URL "https://gucumatz.iplantcollaborative.org/idp/Authn/UserPassword". Below the address bar are icons for CoGe and CoGe: OrgView. The main content area displays the iPlant Collaborative logo and the tagline "Empowering A New Plant Biology". A large orange "Welcome!" heading is followed by a paragraph of text: "You've reached the login page for the Discovery Environment. The Discovery Environment integrates powerful, community-recommended software tools into a system that:" and a bulleted list: "Can handle terabytes of data", "Utilizes high-performance supercomputing resources as needed", and "Hides the complexity needed to do these tasks.". At the bottom, there is a copyright notice: "©2011 iPlant Collaborative. The iPlant Collaborative is funded by a grant from the National Science Foundation (#DBI-0735191)."

A screenshot of the iPlant Collaborative Discovery Environment web interface. The header includes the iPlant Collaborative logo, the text "iPlant Collaborative™ Discovery Environment", and user account links for "elyons", "Help", and "Notifications". The main area features three large buttons: "Data" (with a book icon), "Analyses" (with a monitor icon), and "Apps" (with a document icon). To the right is a large, stylized teal circular logo. At the bottom, there is a copyright notice: "©2011 iPlant Collaborative" and "The iPlant Collaborative is funded by a grant from the National Science Foundation (#DBI-0735191)."

## i-commands

A screenshot of a terminal window. The prompt is "elyons@icoge:~\$". The window has a decorative yellow starburst icon in the bottom-left corner.

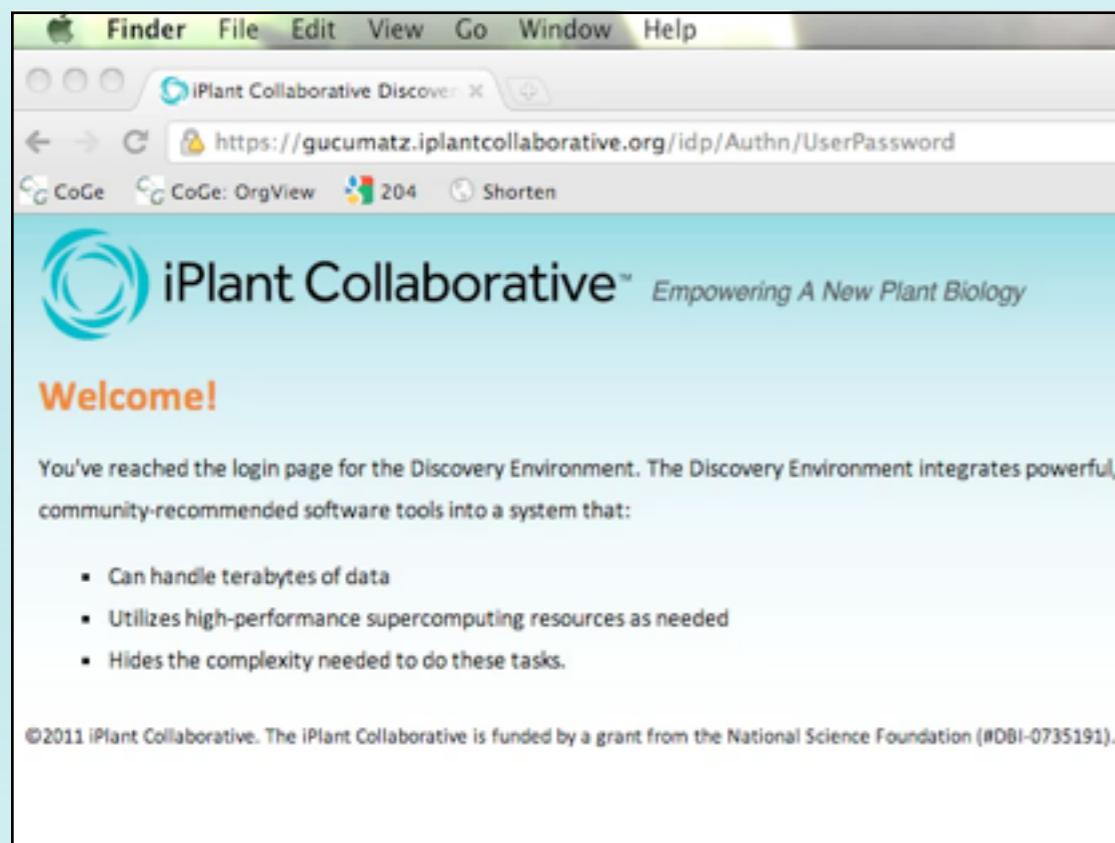


# WebDAV

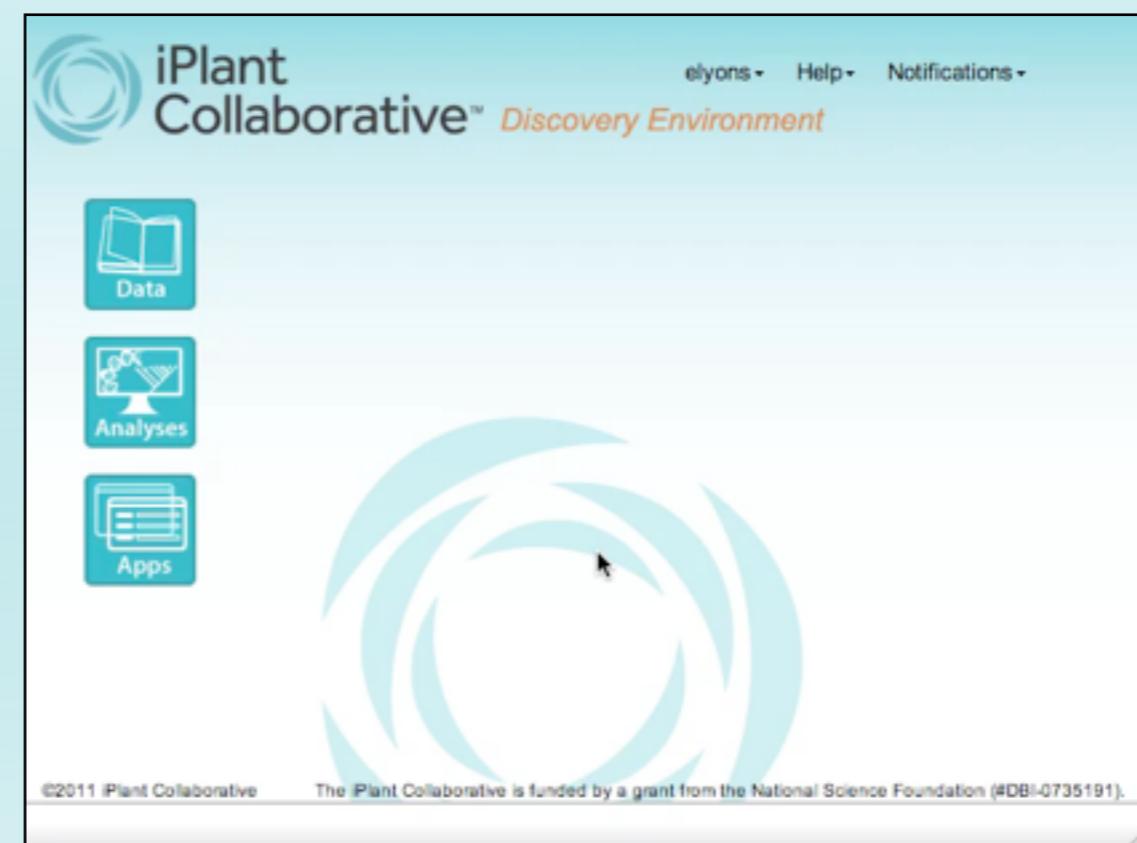
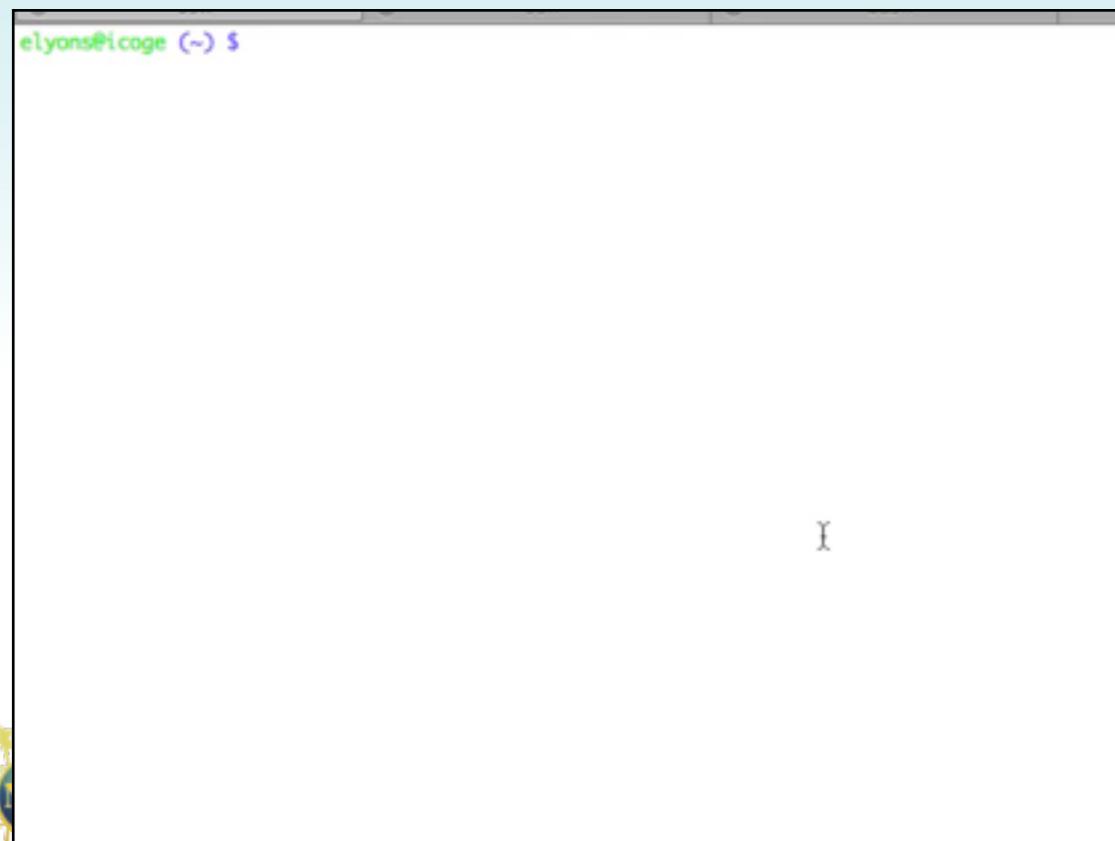
# iPlant Data Store

## Free Your Data

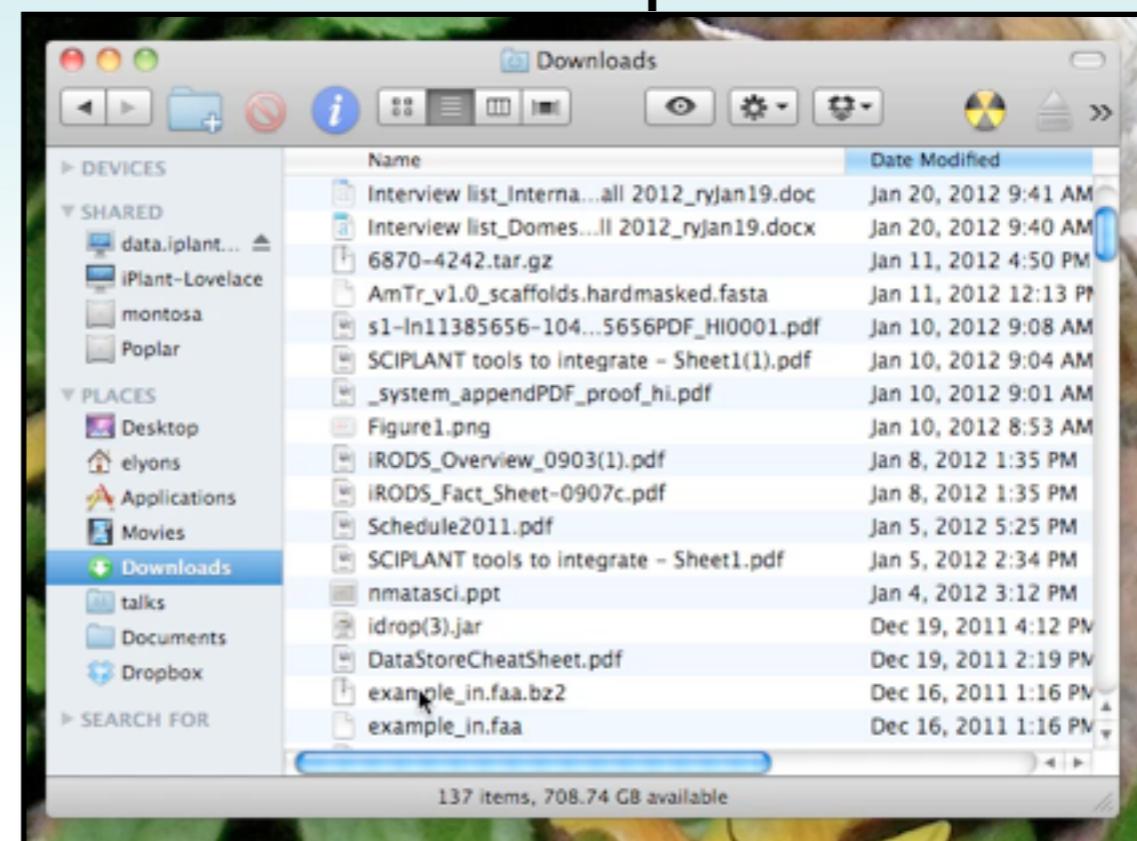
## DE



## i-commands



## iDrop

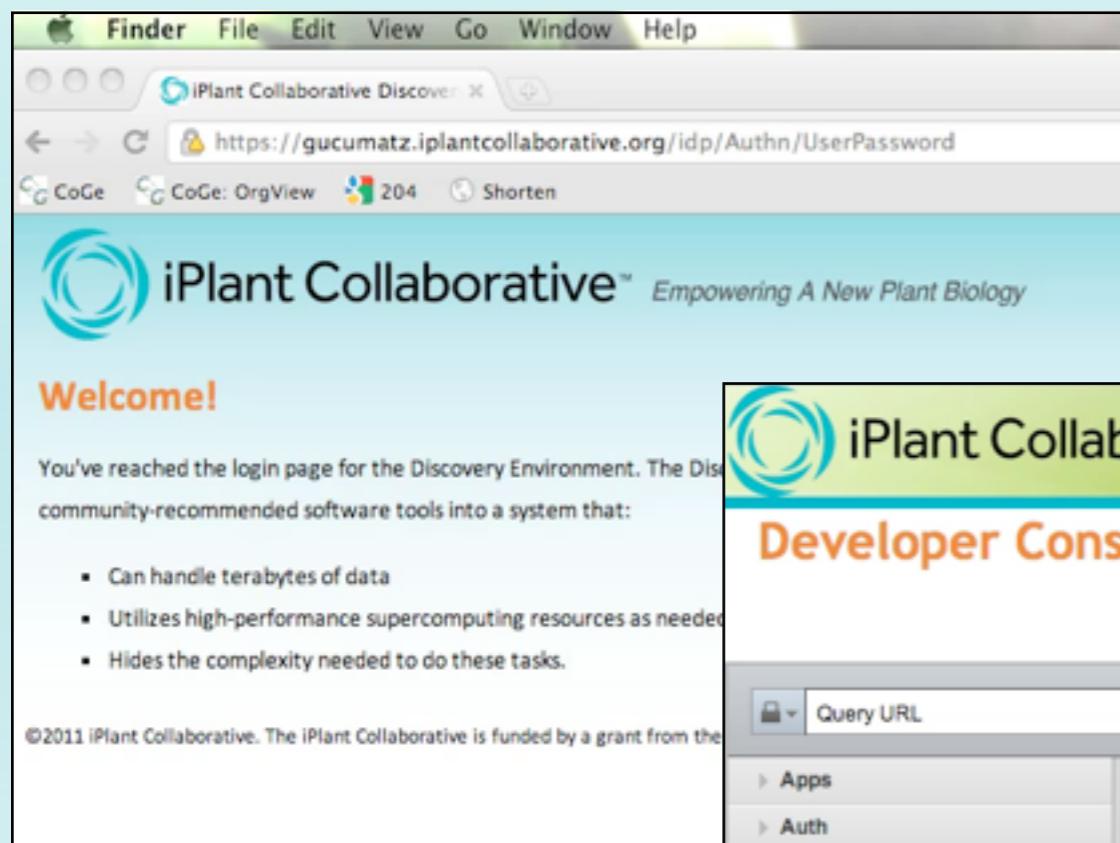


# WebDAV

# iPlant Data Store

## Free Your Data

## DE

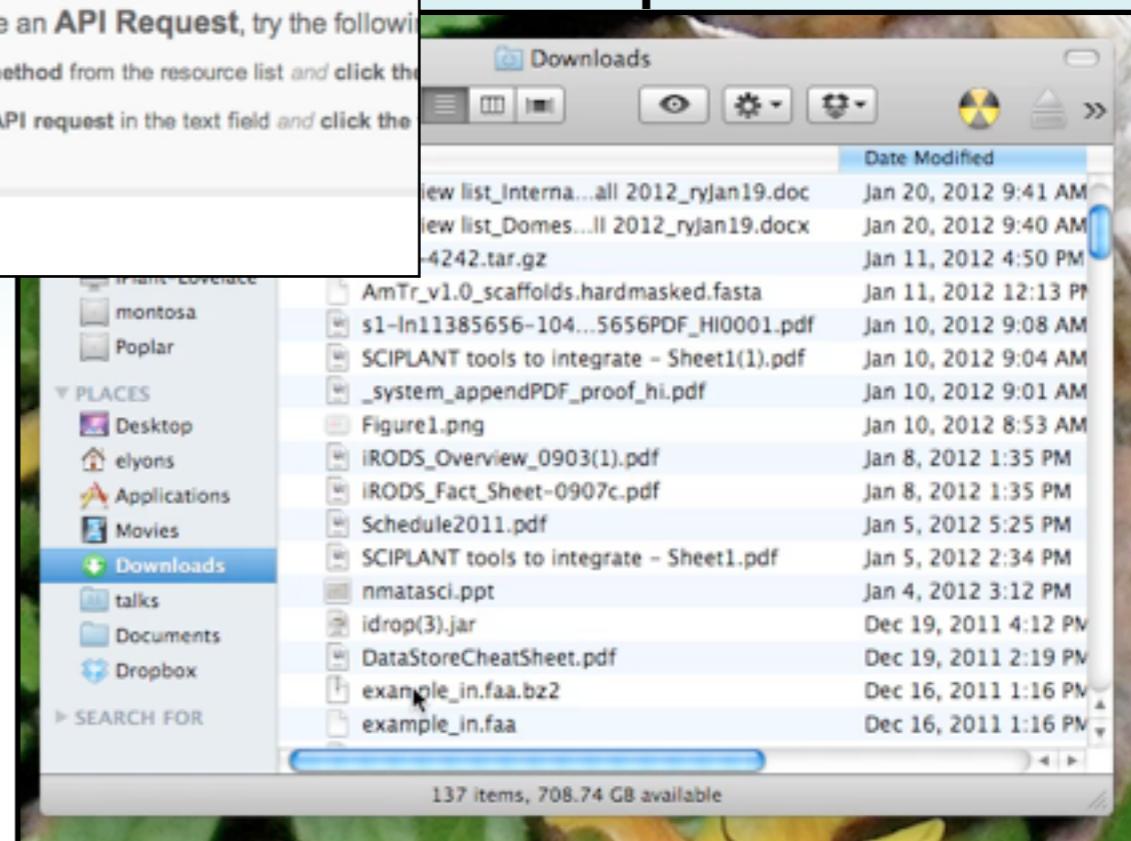


## API

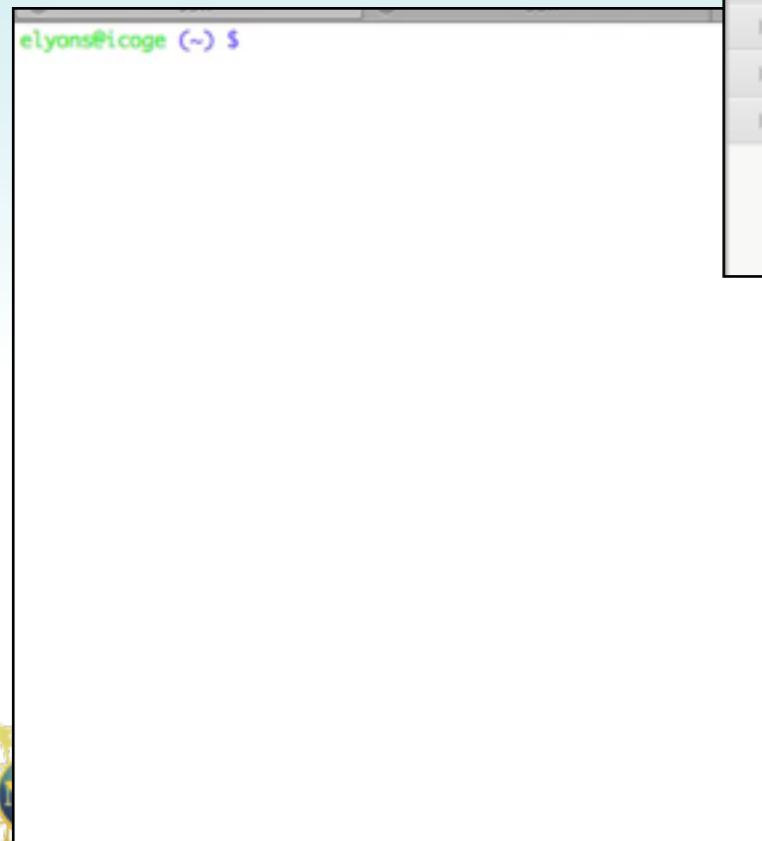
A screenshot of the "Developer Console for the iPlant Foundation API". The interface includes a "Query URL" input field, a "GET" button, and tabs for "Request" and "Response". A tooltip provides instructions: "To make an API Request, try the following:  
a) Select a method from the resource list and click the button  
b) Enter an API request in the text field and click the button". On the left, there's a sidebar with links to "Apps", "Auth", "Data", "IO", "Jobs", "Postit", and "Profile".



## iDrop



## i-commands

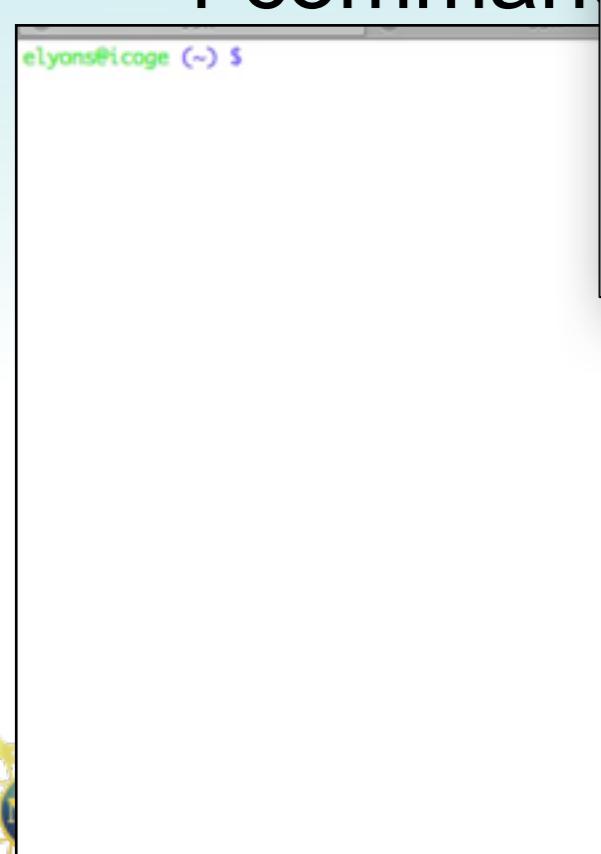
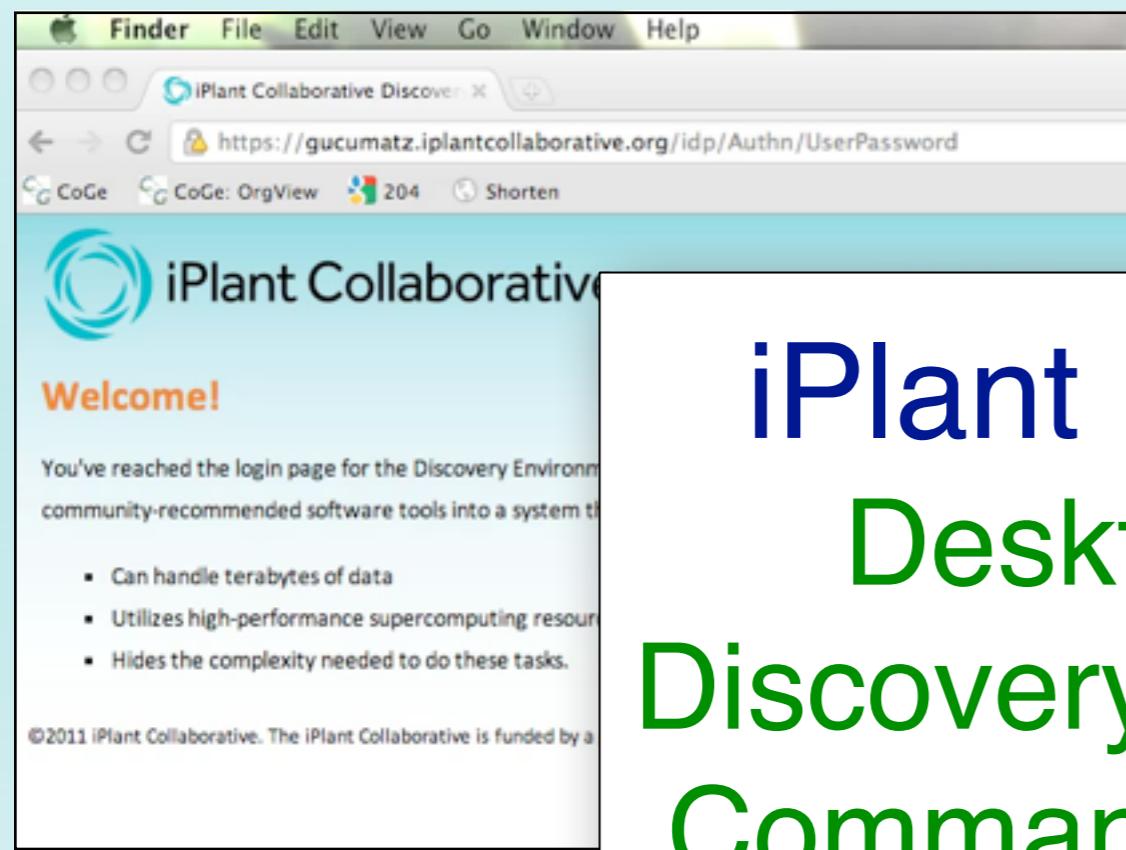


# iPlant Data Store

WebDAV

Free Your Data

DE



# iPlant Data Store

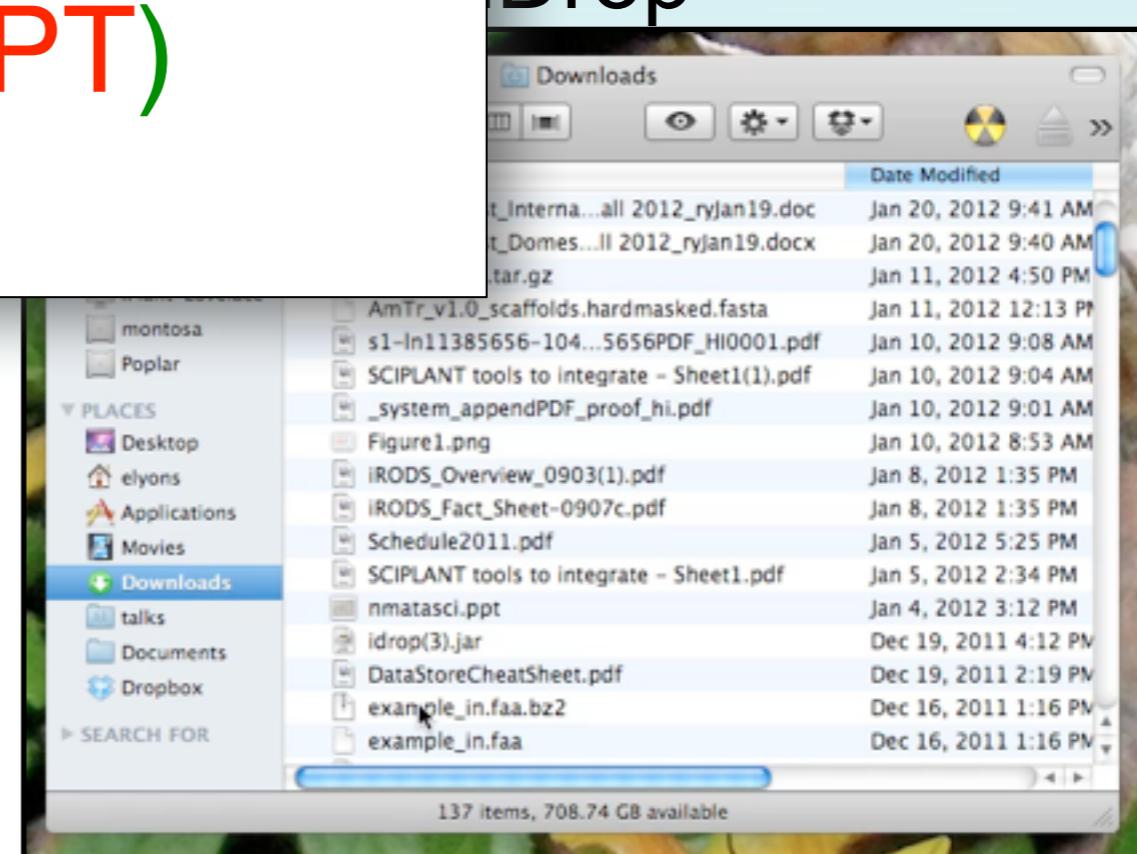
## Desktop Folder

## Discovery Environment

## Command Line (HPT)

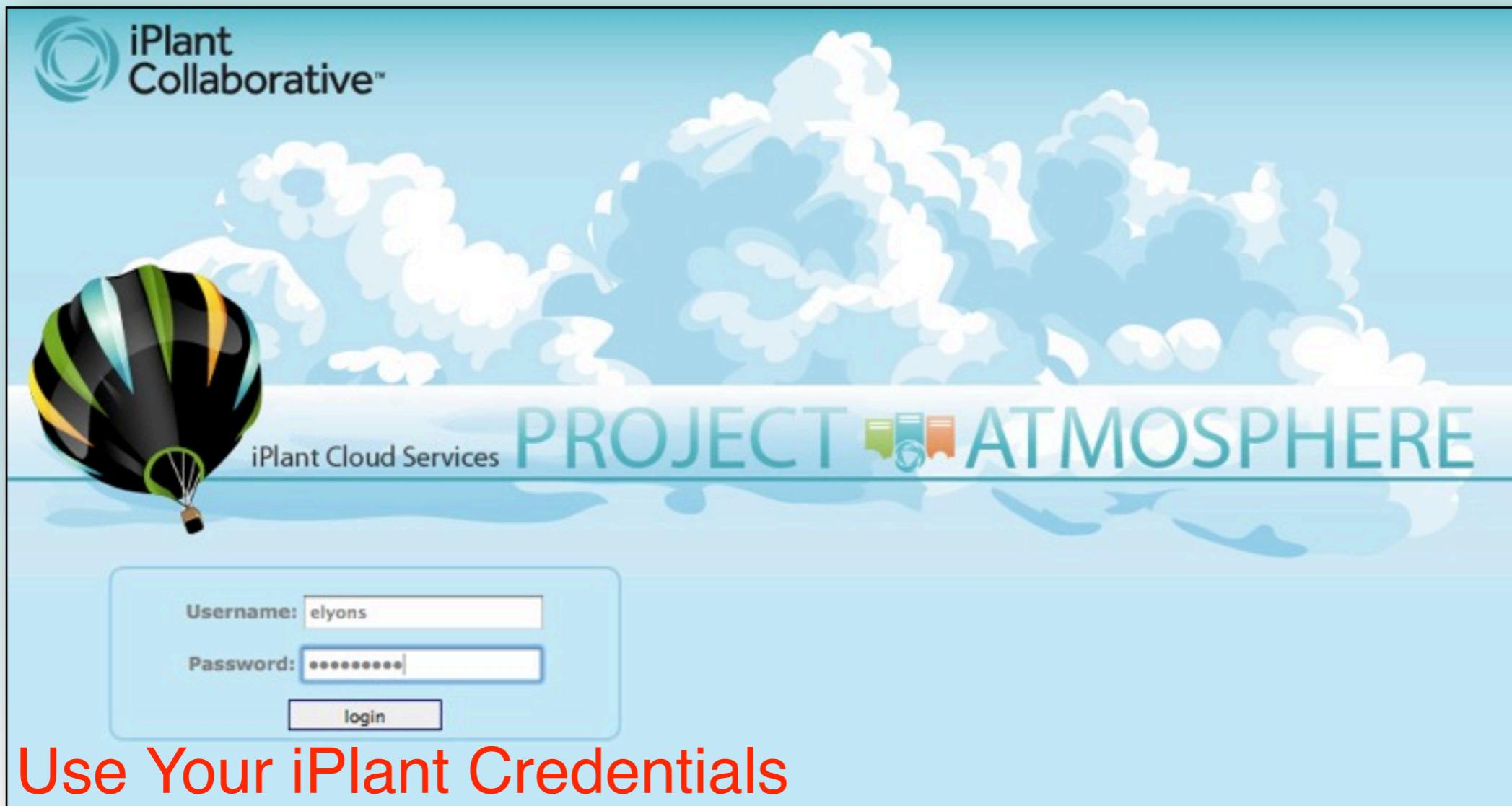
## i-Drop (HPT)

## API



# Atmosphere: Servers and Software on Demand

<http://atmosphere.iplantcollaborative.org>



# Atmosphere

<http://atmosphere.iplantcollaborative.org>

The screenshot shows the Atmosphere application interface. At the top, there's a navigation bar with a logo, the word "ATMOSPHERE", and a "logout" link. Below the navigation bar is a toolbar with "Applications" and "Dashboard" tabs, and back/forward navigation buttons. The main area displays a grid of application icons and their names:

Icon	Name
	BCB660 Image
	Biotools Demo
	CBSU Workshop Image
	iPlant Base Image
	phytoMorph CSHL Image
	R Image
	Puppet-Matlab test
	Puppet-Galaxy test
	coge-api

Below this grid, the "Biotools" section is highlighted. It contains descriptive text and links for several tools:

- Flapjack:** New software tools for graphical genotyping and haplotype visualization are required that can routinely handle the large data volumes generated by high throughput SNP and comparable genotyping technologies. Flapjack is a new visualization tool to facilitate analysis of these data types. <http://bioinf.scri.ac.uk/flapjack/>
- Tablet:** Tablet is a lightweight, high-performance graphical viewer for next-generation sequence assemblies and alignments. Supporting a range of input assembly formats, Tablet provides high-quality visualizations showing data in packed or stacked views, allowing instant access and navigation to any region of interest, and whole contig overviews and data summaries. Tablet is both multi-core aware and memory efficient, allowing it to handle assemblies containing millions of reads, even on a 32-bit desktop machine. <http://bioinformatics.oxfordjournals.org/content/26/3/401.abstract>
- Taverna-workbench:** Taverna is an open source and domain independent Workflow Management System – a suite of tools used to design and execute scientific workflows and aid in silico experimentation. <http://www.taverna.org.uk/>
- Kepler:** Kepler is a free-software system for designing, executing, reusing, evolving, archiving, and sharing scientific workflows. Kepler's facilities provide process and data monitoring, provenance information, and high speed data movement solutions. Workflows in general, and scientific workflows in particular, are directed graphs where the nodes represent discrete computational components, and the edges represent paths along which data and results can flow between components. [http://en.wikipedia.org/wiki/Kepler\\_scientific\\_workflow\\_system](http://en.wikipedia.org/wiki/Kepler_scientific_workflow_system)



# Atmosphere

<http://atmosphere.iplantcollaborative.org>

The screenshot shows the Atmosphere application interface. At the top, there is a navigation bar with icons for file operations, a search bar, and a user account section labeled "logout". Below the navigation bar is a toolbar with tabs for "Applications" (which is selected) and "Dashboard". The main content area displays a grid of application icons and names. The "Biotools Demo" application is highlighted with a yellow border. A message dialog box is overlaid on the page, containing the text: "Successfully initiated your app. Please wait 10 to 15 min to finish the request. Your App id is i-3C8206D5". An "OK" button is at the bottom of the dialog. To the left of the dialog, under the "Biotools" heading, there is descriptive text about Flapjack, Tablet, Taverna-workbench, and Kepler.

**Biotools**

**Flapjack:** New software tools for graph assembly and visualization. Flapjack is designed to handle the large data volumes generated by next-generation sequencing. It provides a new visualization tool to facilitate analysis of complex genomic data.

**Tablet:** Tablet is a lightweight, high-performance viewer for genomic data. Supporting a range of input formats, it allows users to view and analyze data in packed or stacked views, allowing instant access and navigation to any region of interest, and whole contig overviews and data summaries. Tablet is both multi-core aware and memory efficient, allowing it to handle assemblies containing millions of reads, even on a 32-bit desktop machine. <http://bioinformatics.oxfordjournals.org/content/26/3/401.abstract>

**Taverna-workbench:** Taverna is an open source and domain independent Workflow Management System – a suite of tools used to design and execute scientific workflows and aid in silico experimentation. <http://www.taverna.org.uk/>

**Kepler:** Kepler is a free-software system for designing, executing, reusing, evolving, archiving, and sharing scientific workflows. Kepler's facilities provide process and data monitoring, provenance information, and high speed data movement solutions. Workflows in general, and scientific workflows in particular, are directed graphs where the nodes represent discrete computational components, and the edges represent paths along which data and results can flow between components. [http://en.wikipedia.org/wiki/Kepler\\_scientific\\_workflow\\_system](http://en.wikipedia.org/wiki/Kepler_scientific_workflow_system)



Thursday, March 1, 12



# Atmosphere

<http://atmosphere.iplantcollaborative.org>

The screenshot shows the Atmosphere web interface. At the top, there's a navigation bar with icons for file operations, the title "ATMOSPHERE", and a "logout" link. Below the navigation bar is a toolbar with "Applications" and "Dashboard" tabs, and back/forward navigation buttons. The main content area displays a grid of nine application icons, each with a name below it:

- BCB660 Image
- Biotools Demo
- CBSU Workshop Image
- iPlant Base Image
- phytoMorph CSHL Image
- R Image
- Puppet-Matlab test
- Puppet-Galaxy test
- coge-api

On the left side of the main content area, there's a section titled "Biotools" containing descriptions of several tools:

- Flapjack:** New software tools for graphically visualizing and manipulating large data sets. Flapjack allows users to handle the large data volumes generated by next-generation sequencing and other high-throughput experiments. It provides a new visualization tool to facilitate analysis and interpretation of complex data.
- Tablet:** Tablet is a lightweight, high-performance application for viewing and analyzing genomic data. Supporting a range of input formats, Tablet allows users to view data in packed or stacked views, allowing instant zooming and panning across the data. It also provides data summaries. Tablet is both multi-threaded and multi-core, able to handle millions of reads, even on a 32-bit desktop system.
- Taverna-workbench:** Taverna is an open-source workflow language and toolkit used to design and execute scientific workflows. It provides a graphical interface for creating workflows, and a command-line interface for executing them.
- Kepler:** Kepler is a free-software system for designing, executing, reusing, evolving, archiving, and sharing scientific workflows. Kepler's facilities provide process and data monitoring, provenance information, and high speed data movement solutions. Workflows in general, and scientific workflows in particular, are directed graphs where the nodes represent discrete computational components, and the edges represent paths along which data and results can flow between components. [http://en.wikipedia.org/wiki/Kepler\\_scientific\\_workflow\\_system](http://en.wikipedia.org/wiki/Kepler_scientific_workflow_system)

A modal window is open in the center of the screen, displaying an incoming email message:

From: atmo@iplantcollaborative.org  
Subject: Your Atmosphere Cloud Instance  
Date: August 4, 2011 5:26:35 PM MST  
To: Eric Lyons

Your Atmosphere cloud instance is ready.

Instance id: i-3C8206D5  
IP: 150.135.78.154  
SSH Username: elyons  
SSH Password:

Atmosphere, iPlant Collaborative



# Atmosphere

## Plus VNC

The screenshot shows a desktop environment titled "Atmosphere". The desktop background features a green and yellow abstract pattern. At the top, there is a menu bar with "File", "Edit", "View", "System", "Help", and "Logout". Below the menu bar is a toolbar with icons for "File", "Edit", "View", "System", "Help", and "Logout". The main workspace contains several application icons arranged in a grid:

Icon	Name
BCB660 Image	BCB660 Image
Biotools Demo	Biotools Demo
CBSU Workshop Image	CBSU Workshop Image
iPlant Base Image	iPlant Base Image
NGS Viewers CSHL Image	NGS Viewers CSHL Image
phytoMorph	phytoMorph
R Image	R Image
VT Workshop Demo	VT Workshop Demo
Puppet-Matlab test	Puppet-Matlab test

At the bottom of the screen, there is a status bar with the URL "http://maps.google.com/maps?q=flapjac...fox-a&um=1&ie=UTF-8&hl=en&sa=N&tab=wl" and a "zotero" logo.

# “Powered by iPlant CI”

- Make it easy for community to develop applications based on iPlant CI
- Authentication
- Data store
- Compute
- Applications can be distributed or hosted by iPlant
- Provide developer with tool kit to get started



# Training end users !!

This box searches only this space. The box at the upper right searches the entire iPlant wiki.

Getting Help with iPlant Services  
 Acquiring an iPlant Account  
 Accessing iPlant Services (List Included)  
 Resetting Your iPlant Account Password  
**+ Storing Your Data with iPlant and Accessing that Data**  
 List of Available Materials, Documentation, and User Help Options  
 Frequently Asked Questions (FAQ)  
 Reporting Issues with JIRA  
**+ Employee Information**

**C https://pods.iplantcollaborative.org/wiki/display/start/Storing+Your+Data+with+iPlant+and+Accessing+that+Data Atmosphere**

Dashboard > Getting Started with iPlant > Getting Started with iPlant > Storing Your Data with iPlant and Accessing that Data Browse Nirav Merchant

## Uploading data

There are several ways to upload your data. These include:

- Directly to the [Discovery Environment](#) from your desktop (for files under 2 GB)
- Directly to the [Discovery Environment](#) from a URL (for all files)
- [Using the DAVIS Web Interface](#) (for files under 2 GB) ([video tutorial](#))
- Using the [iRODS web client](#) (file size limit unknown, may also be limited by web browser limit of 2 GB)
- [Using WebDAV](#) (file size limit unknown)
- [Using iDrop](#) (file size limit unknown) ([video tutorial](#))
- [Using icommands](#) (ideal for bulk transfers and best for large data and files over 2 GB)
- From [Atmosphere](#) using FUSE (useful if you have already been working in a virtual machine (VM) environment)

*Colors indicate relative difficulty: green=easier, yellow=intermediate, red=more difficult. For most users, we recommend green, the easiest method.*

## Accessing data

There are several ways to access your data. These include:

- Directly from the Discovery Environment: [viewing](#), [managing](#), [deleting](#)
- [Using the Davis web interface](#)
- [Using the iRODS web client](#)
- [Using WebDAV](#)
- [Using iDrop](#)
- From [Atmosphere](#) virtual machines using FUSE
- [Using icommands](#)
- [Installing the iRODS FUSE client for Linux to create a mountable volume](#)

Powered by a free Atlassian Confluence Open Source Project License granted to The iPlant Collaborative, University of Arizona. Evaluate

Powered by Atlassian Confluence 3.4.5, the Enterprise Wiki | Report a bug | Atlassian News



# Training end users !!

w.iplantcollaborative.org/learn/events/2012/02/02/iplant-tools-and-services-workshop-uc-davis-march-12th-and-13th-2012

The screenshot shows the iPlant Collaborative website with a green header. The header includes the logo, the text "iPlant Collaborative™ Empowering A New Plant Biology", and links for "Login or Register", "About", "Contact Us", and "Feedback". Below the header is a navigation bar with four tabs: "CHALLENGE", "DISCOVER", "LEARN" (which is highlighted in red), and "CONNECT". The main content area features a large image of a plant leaf with the word "Learn" overlaid. The page title is "iPlant Tools and Services Workshop, UC Davis. March 12th and 13th 2012". The left sidebar has a green header labeled "Events" and contains links for "Education, Outreach & Training", "Manuals & Tutorials", "News", "Publications", and "Media". The main content area includes sections for "Date and Time" (listing March 12, 2012 from 8:00am to 2:30pm at SciLab 2060, and March 13, 2012 from 9:30am to 4:00pm at Genome Center 4202), "Location" (University of California, Davis), and "Details" (noting registration is closed). A detailed description of the workshop's purpose and scope is also present.

**iPlant Tools and Services Workshop, UC Davis. March 12th and 13th 2012**

**Date and Time**

March 12 2012	8:00am - 2:30pm SciLab 2060
March 13 2012	9:30am - 4:00pm Genome Center 4202

**Location**

University of California, Davis

**Details**

Registration for this event is CLOSED.

Biological research depends increasingly on high-throughput data collection methods and complex computational analyses. This free two-day workshop provides a comprehensive look at the tools and services provided by the iPlant Collaborative, a major cyberinfrastructure project of the National Science Foundation ([www.iplantcollaborative.org](http://www.iplantcollaborative.org)). Workshop topics build progressively to cater to the needs of general and advanced audiences.



# Training end users !!

## Agenda

Monday, March 12th, 2012

Time	Description
08:00 AM - 08:30 AM	Arrive / Sign-in / Verify iPlant Accounts
08:30 AM - 09:00 AM	Welcome - Presenter/Participant Self-Introductions
09:00 AM - 09:20 AM	An Overview of the iPlant Collaborative
09:20 AM - 09:45 AM	Overview of the iPlant Discovery Environment (DE)
09:45 AM - 10:00 AM	Break
10:00 AM - 10:30 AM	iPlant Data Store - Managing "Big Data"
10:30 AM - 11:30 AM	Using the DE to Examine Differential Expression Within an RNA-seq Dataset
11:30 AM - 12:00 PM	Atmosphere - Custom Cloud Computing
12:00 PM - 12:30 PM	Lunch
12:30 PM - 01:00 PM	Building and Using Workflows Within the DE; Phylogenetics
01:00 PM - 01:45 PM	Extending the DE for Your Research: Tool Integration and Customization
01:45 PM - 02:30 PM	Collaborating with iPlant: Future Projects and Workshop Summary

Tuesday, March 13th, 2012

Time	Description
09:30 AM - 11:30 AM	Advanced ChIP-seq in the iPlant Cyberinfrastructure
11:30 AM - 12:00 PM	Comparative Phylogenetic Methods in the iPlant Cyberinfrastructure
	<b>Data Clinic Appointments</b>
01:30 PM - 02:00 PM	Powered by iPlant - Consuming iPlant Services in Your Portals
02:00 PM - 02:30 PM	Using XSEDE for Bioinformatics



# Future directions

- More data driven collaboration interfaces  
(tickets/tokens\*)
- Deeper Metadata handling  
(extending AVU)
- Content/Repository searching  
(Solr)
- Data enrichment, decoration  
(enhanced markup, semantic integration)
- Social computing  
(Data de-duplication, suggestion engine)



# How to get access:

[www.iplantcollaborative.org](http://www.iplantcollaborative.org)

©2011 iPlant Collaborative

[Privacy](#) | [RSS Feed](#) | [Staff Login](#)

The iPlant Collaborative is funded by a grant from the National Science Foundation Plant Cyberinfrastructure Program (#DBI-0735191).



# How to get access:

[www.iplantcollaborative.org](http://www.iplantcollaborative.org)

©2011 iPlant Collaborative

[Privacy](#) | [RSS Feed](#) | [Staff Login](#)

The iPlant Collaborative is funded by a grant from the National Science Foundation Plant Cyberinfrastructure Program (#DBI-0735191).





Powered By iPlant

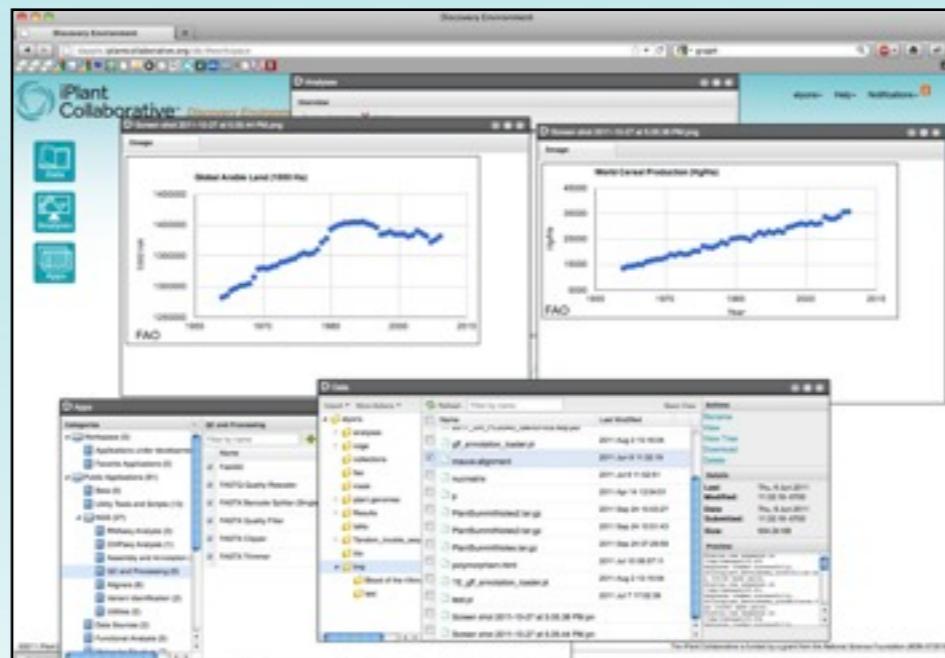


Thursday, March 1, 12





# Powered By iPlant





# Powered By iPlant

The collage includes:

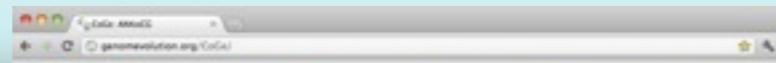
- BioExtract Server:** A web interface for data access, analysis, storage, and workflow creation. It shows a search form with fields for "Search Field" and "Search Term(s)", and a "Current Query" button.
- iPlant Program Launcher:** A desktop application window titled "iCLS Program Launcher" showing a grid of icons for different tools and services.
- CoGe (The Place to Compare Genomes):** A web interface for comparing genomes, showing a green circular logo for "BASAVADA COP".
- iPlant Collaborative - Discovery Environment:** A dashboard showing two line graphs: "Global Arable Land (1980-2005)" and "World Cereal Production (1960-2010)".
- ICASS CENTRAL DATABASE 2010-06:** A screenshot of the database interface with a tree view of data sources and a search bar.
- ncbi\_ebi\_iplant.pl:** A terminal window showing a Perl script with comments about copyright, license, and author information.
- DNA Subway:** A web-based genome browser interface showing a linear genome with various annotations and tracks.
- Project Information:** A detailed view of a project named "Kangaroo 2", including user details, organism information ("Organism: mouse-ear cress"), and sequence length ("Sequence: 10.30 Mb").

iPlant APIs  
Resources





Powered By iPlant



# Cyberinfrastructure for Life Sciences

Scalable  
Capable  
Extensible



# Where to Get More Information

- Register:
- iPlant DE:
- iPlant Data Store:
- iPlant Atmo:
- iPlant Quick-Start Guides:
- iPlant Tutorials:
- iPlant Forums:
- iPlant Wiki:



# iPlant's Building Blocks

## Postdocs:

Barbara Banbury  
Jamie Estill  
Bindu Joseph  
Christos Noutsos  
Brad Ruhfel  
Stephen A. Smith  
Chunlao Tang  
Lin Wang  
Liya Wang  
Norman Wickett

## Executive Team: Steve Goff Dan Stanzione

## Students:

Peter Bailey  
Jeremy Beaulieu  
Devi Bhattacharya  
Storme Briscoe  
Ya-Di Chen  
John Donoghue  
Steven Gregory  
Yekatarina Khartianova  
Monica Lent  
Amgad Madkour

Aniruddha Marathe  
Kurt Michaels  
Dhanesh Prasad  
Andrew Predoehl  
Jose Salcedo  
Shalini Sasidharan  
Gregory Striemer  
Jason Vandeveenter  
Kuan Yang

## Metadata

## Data

## Tools

## Workflows

## Viz

## Faculty Advisors & Collaborators:

Ali Akoglu  
Greg Andrews  
Kobus Barnard  
Sue Brown  
Thomas Brutnell  
Michael Donoghue  
Casey Dunn  
Brian Enquist  
Damian Gessler  
Ruth Grene  
John Hartman  
Matthew Hudson  
Dan Kliebenstein  
Jim Leebens-Mack  
David Lowenthal  
Robert Martienssen

B.S. Manjunath  
Nirav Merchant  
David Neale  
Brian O'Meara  
Sudha Ram  
David Salt  
Mark Schildhauer  
Doug Soltis  
Pam Soltis  
Edgar Spalding  
Alexis Stamatakis  
Ann Stapleton  
Lincoln Stein  
Val Tannen  
Todd Vision  
Doreen Ware  
Steve Welch  
Mark Westneat

## Staff:

Greg Abram  
Sonali Aditya  
Roger Barthelson  
Brad Boyle  
Todd Bryan  
Gordon Burleigh  
John Cazes  
Mike Conway  
Karen Cranston  
Rion Doodey  
Andy Edmonds  
Dmitry Fedorov  
Michael Gatto  
Utkarsh Gaur  
Cornel Ghiban  
Michael Gonzales  
Hariolf Häfele  
Matthew Hanlon

## Anthony Heath

Barbara Heath  
Matthew Helmke  
Natalie Henriques  
Uwe Hilgert  
Nicole Hopkins  
Eun-Sook Jeong  
Logan Johnson  
Chris Jordan  
B.D. Kim  
Kathleen Kennedy  
Mohammed Khalfan  
Seung-jin Kim  
Lars Koersterk  
Sangeeta Kuchimanchi  
Kristian Kvilekval  
Aruna Lakshmanan  
Sue Lauter  
Tina Lee

## Andrew Lenards

Zhenyuan Lu  
Eric Lyons  
Naim Matasci  
Sheldon McKay  
Robert McLay  
Angel Mercer  
Dave Micklos  
Nathan Miller  
Steve Mock  
Martha Narro  
Praveen Nuthulapati  
Shannon Oliver  
Shiran Pasternak  
William Peil  
Titus Purdin  
J.A. Raygoza Garay  
Dennis Roberts  
Jerry Schneider

Bruce Schumaker  
Sriram Singaram  
Edwin Skidmore  
Brandon Smith  
Mary Margaret Sprinkle  
Sriram Srinivasan  
Josh Stein  
Lisa Stillwell  
Kris Urie  
Peter Van Buren  
Hans Vasquez-Gross  
Matthew Vaughn  
Fusheng Wei  
Jason Williams  
John Wregglesworth  
Weijia Xu  
Jill Yarmchuk



# iPlant's Building Blocks

## Postdocs:

Barbara Banbury  
Jamie Estill  
Bindu Joseph  
Christos Noutsos  
Brad Ruhfel  
Stephen A. Smith  
Chunlao Tang  
Lin Wang  
Liya Wang  
Norman Wickett

## Metadata

### Faculty Advisors & Collaborators:

Ali Akoglu  
Greg Andrews  
Kobus Barnard  
Sue Brown  
Thomas Brutnell  
Michael Donoghue  
Casey Dunn  
Brian Enquist  
Damian Gessler  
Ruth Grene  
John Hartman  
Matthew Hudson  
Dan Kliebenstein  
Jim Leebens-Mack  
David Lowenthal  
Robert Martienssen

B.S. Manjunath  
Nirav Merchant  
David Neale  
Brian O'Meara  
Sudha Ram  
David Salt  
Mark Schildhauer  
Doug Soltis  
Pam Soltis  
Edgar Spalding  
Alexis Stamatakis  
Ann Stapleton  
Lincoln Stein  
Val Tannen  
Todd Vision  
Doreen Ware  
Steve Welch  
Mark Westneat

## Executive Team

Steve  
Dan  
Stamatakis

## Students:

Peter Bailey  
Jeremy Beaulieu  
Devi Bhattacharya  
Storme Briscoe  
Ya-Di Chen  
John Donoghue  
Stephen Gregory  
Tatyana Khartianova  
Lindsey Lent  
Rajendra Ladkour

Aniruddha Marathe  
Kurt Michaels  
Dhanesh Prasad  
Andrew Predoehl  
Jose Salcedo  
Shalini Sasidharan  
Gregory Striemer  
Jason Vandeveenter  
Kuan Yang

## Data

## NSF

Andrew Lenards  
Zhenyuan Lu  
Eric Lyons  
Niam Matasci  
Aldon McKay  
Matt McLay  
Angel Mercer  
Dave Micklos  
Nathan Miller  
Steve Mock  
Martha Narro  
Praveen Nuthulapati  
Shannon Oliver  
Shiran Pasternak  
William Peil  
Titus Purdin  
J.A. Raygoza Garay  
Dennis Roberts  
Jerry Schneider

## Viz

Bruce Schumaker  
Sriram Singaram  
Edwin Skidmore  
Brandon Smith  
Mary Margaret Sprinkle  
Sriram Srinivasan  
Josh Stein  
Lisa Stillwell  
Kris Urie  
Peter Van Buren  
Hans Vasquez-Gross  
Matthew Vaughn  
Fusheng Wei  
Jason Williams  
John Wregglesworth  
Weijia Xu  
Jill Yarmchuk



# iPlant Discovery Environment



Managing and Integrating: Data, Tools, Analysis



# iPlant Discovery Environment

Discovery Environment preview.iplantcollaborative.org/de/#workspace Google

Discovery Environment elyons Help Notifications

iPlant Collaborative™ *Discovery Environment*

Data Analyses Apps



©2011 iPlant Collaborative, Inc. This software was developed by a National Science Foundation (NSF) Biological Informatics Graduate Traineeship (GTR-ITP) at the University of Texas at Austin. It is available under the Apache License, Version 2.0. See the accompanying file LICENSE for details.

Managing and Integrating: Data, Tools, Analysis

<http://preview.iplantcollaborative.org/de/#workspace>



File View

