



# Help Desk Update



Dave Clements  
GMOD Help Desk  
US National Evolutionary Synthesis Center (NESCent)  
[clements@nescent.org](mailto:clements@nescent.org)

January 2010 GMOD Meeting  
14-15 January 2010

# Agenda

Project Statistics

Training and Outreach

Genome Visualization

Grant Proposal

New Components

GMOD Web Site

Odds and Ends



[http://gmod.org/wiki/GMOD\\_Help\\_Desk](http://gmod.org/wiki/GMOD_Help_Desk)



# Project Statistics: GMOD.org in 2009

Month	Unique visitors	Number of visits	Pages	Hits	Bandwidth
Jan 2009	16820	30454	79953	676273	9.26 GB
Feb 2009	16616	29740	73542	646475	8.63 GB
Mar 2009	19560	34022	88291	795559	10.23 GB
Apr 2009	19343	31309	84973	767353	8.49 GB
May 2009	18254	27179	75503	708181	8.95 GB
Jun 2009	18364	28361	77664	722494	8.49 GB
Jul 2009	19529	29195	86180	776802	9.94 GB
Aug 2009	19017	28921	80007	752951	10.85 GB
Sep 2009	18737	29800	87966	758343	10.37 GB
Oct 2009	19068	31569	101654	787424	11.64 GB
Nov 2009	18805	31742	98863	772927	10.79 GB
Dec 2009	18952	30681	93991	790669	12.91 GB
Total	223065	362973	1028587	8955451	120.55 GB

1 Million page hits








# Project Statistics: GMOD.org in 2009

Search Keyphrases (Top 50) Full list			Search Keywords (Top 50) Full list		
46532 different keyphrases	Search	Percent	12496 different keywords	Search	Percent
gmod	52318	21.9 %	gmod	94385	19.9 %
chado	22788	9.5 %	gbrowse	31728	6.7 %
gmod.org	18523	7.7 %	chado	27690	5.8 %
gbrowse	15740	6.5 %	gmod.org	18791	3.9 %
gmod downloads	5617	2.3 %	download	7720	1.6 %
gmod download	3320	1.3 %	mod	7205	1.5 %
g mod	1767	0.7 %	downloads	6417	1.3 %
gmod help	1289	0.5 %	to	6015	1.2 %
jbrowse	1232	0.5 %	balloon	5250	1.1 %
download gmod	1182	0.4 %	how	4160	0.8 %
cmap	1098	0.4 %	install	3513	0.7 %
gbrowse_syn	919	0.3 %	in	3420	0.7 %







Top search terms



# Project Statistics: GMOD.org in 2009

Browsers	Grabber	Hits	Percent
 <b>Firefox</b>	No	3993663	44.5 %
 <b>MS Internet Explorer</b>	No	3593778	40.1 %
 Safari	No	872601	9.7 %
 Opera	No	167891	1.8 %
 Mozilla	No	138505	1.5 %
? Unknown	?	81182	0.9 %

## Browsers

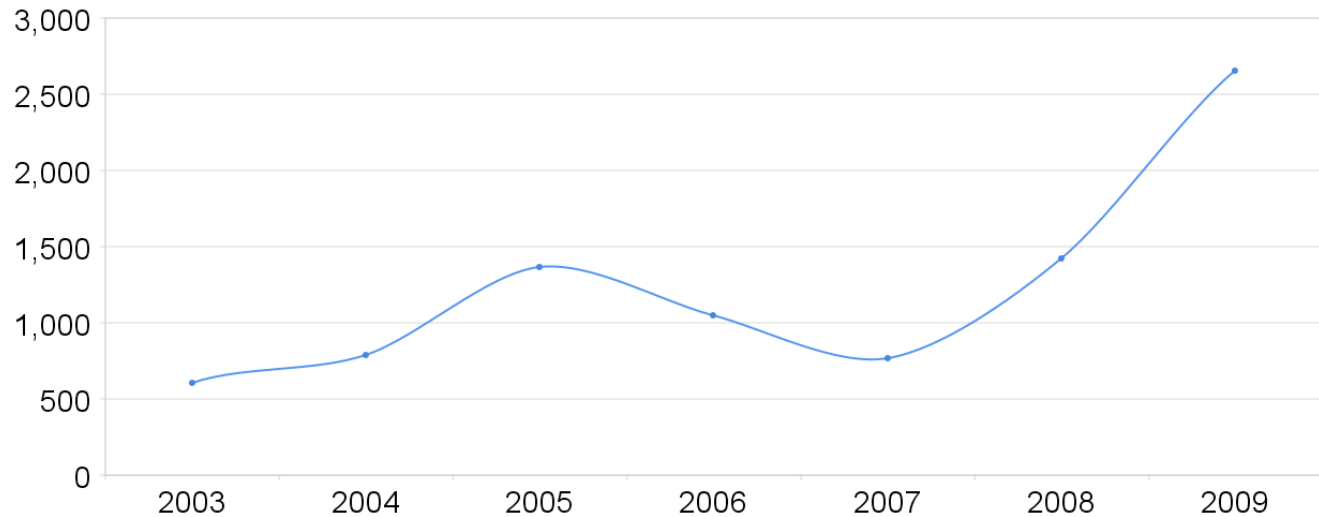
Operating Systems	Hits	Percent
 <b>Windows</b>	7051216	78.7 %
 <b>Macintosh</b>	1021302	11.4 %
 <b>Linux</b>	696100	7.7 %
? Unknown	167696	1.8 %
 Sun Solaris	10063	0.1 %
 Sony PlayStation Portable	4589	0 %
 <b>BSD</b>	3404	0 %

## Operating Systems

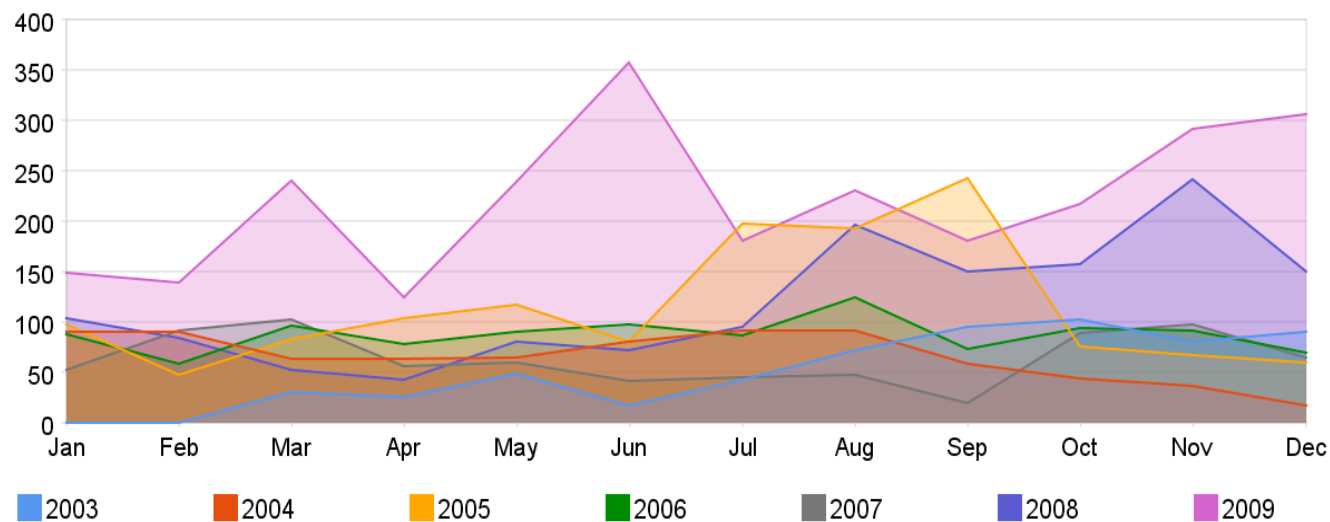


# Project Statistics: Mailing Lists

Messages per year on GBrowse List



Messages per month to GBrowse Mailing List



On the 7 core mailing lists managed by GMOD staff:

- Membership is up 40% over past 20 months
- Message volume is up 69% in last year

# Agenda

Project Statistics

Training and Outreach

Genome Visualization

Grant Proposal

New Components

GMOD Web Site

Odds and Ends



[http://gmod.org/wiki/GMOD\\_Help\\_Desk](http://gmod.org/wiki/GMOD_Help_Desk)



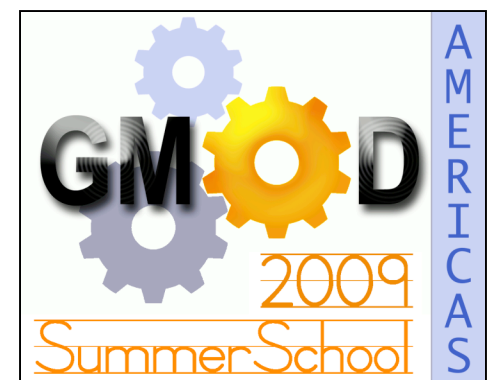
# Training and Outreach

## GMOD Summer Schools

- 2009 (110 applications in 2009 | 29 in 2008)
  - July, NESCent, North Carolina, US
  - August, Oxford, UK (& 1st meeting in Europe)
- 2010
  - May, NESCent, North Carolina, US
  - ??, Asia / Pacific, maybe

## Outreach, Training

- ABRF, PAG, ISyIP, Bioinformatics Australia, Genome Informatics, SMBE, Arthropod Genomics, NGS meetings, ...
- Workshops on InterMine, Pathway Tools, BioMart, Galaxy, ...



[http://gmod.org/wiki/Training\\_and\\_Outreach](http://gmod.org/wiki/Training_and_Outreach)





# GMOD @ Plant and Animal Genome XVIII

## GMOD Users

Gramene  
RAP-DB  
ConiferGDB  
TAIR  
MaizeGDB  
SGN  
TreeGenes  
SoyBase  
Birdbase  
Marine Genomics  
BeeBase  
GnplS  
+ hundreds more

## GMOD Tools

Chado  
Apollo  
GBrowse  
CMap  
Tripal  
GBrowse\_syn  
Ergatis  
Textpresso  
BioMart  
JBrowse  
MAKER  
Galaxy  
InterMine  
...



## Animal Genomics with Galaxy: Analyze, Publish & Visualize

Saturday 1:30-2:30pm, California Room, Dan Blankenberg

## MAKER: An easy to use genome annotation pipeline

Saturday, 3:50-6:00pm, California Room, Carson Holt

## Database Resources at the EBI: Ensembl (including BioMart)

Sunday, 8:15-9:15am, California Room, Bert Overduin

## The Generic Genome Browser: A Hands on Workshop

Sunday, 1:30-3:40pm, California Room, Scott Cain

## Using Gramene (including CMap)

Tuesday, 3:50-6:00pm, California Room, Ken Youens-Clark

## Comparative Genomics with GBrowse\_syn

Wednesday, 10:20am-12:30pm, Pacific Salon 3, Sheldon McKay

Workshops

## GMOD Project Update

Monday, 12:50-1:10pm, California Room, Dave Clements, Scott Cain

## Using the JBrowse Genome Browser with Large Amounts of Data

Monday, 1:10-1:30pm, California Room, Mitch Skinner (also see poster P919)

## Tripal: A Construction Toolkit for Online Genomics Databases

Wednesday, 2:30-2:50pm, California Room, Stephen Ficklin

Talks

## Monday 10-11:30am

P858: GBrowse and Next Generation Sequencing Data

## Monday 3-4:30pm

P859: Visualizing and Comparing Genomes

P861: Comparative Genomics Tools in GMOD

Posters

at least 15 other talks and 19 more posters by GMOD users (see [http://gmod.org/wiki/PAG\\_2010](http://gmod.org/wiki/PAG_2010)) +

[http://gmod.org/wiki/PAG\\_2010](http://gmod.org/wiki/PAG_2010)

Got genome?

# Tutorials

- Summer school sessions become online tutorials with
  - Starting VMware images
  - Step by step instructions
  - Example datasets
  - Ending VMware images
- Topics:
  - Apollo, Artemis-Chado Integration, BioMart, Chado, CMap, GBrowse, GBrowse\_syn, JBrowse, MAKER, Tripal, GBrowse NGS, MAKER Web Annotation Service (MWAS)

The screenshot shows the Apollo Tutorial page. The title is "Apollo Tutorial". Below the title, it says "2009 GMOD Summer School - Europe & Americas" and "July & August 2009". The author is listed as "Ed Lee". The page includes a navigation sidebar with links for "GMOD Home", "Categories / Tags", "Downloads", and "View all pages". There is also a "documentation" section with links for "Overview" and "FAQs".

The screenshot shows the BioMart Tutorial page. The title is "BioMart Tutorial". Below the title, it says "2009 GMOD Summer School - Europe & Americas" and "July & August 2009". The authors are listed as "Junjun Zhang and Syed Haider". The page includes a navigation sidebar with links for "GMOD Home", "Categories / Tags", "Downloads", and "View all pages". There is also a "documentation" section with links for "Overview", "FAQs", and "HOWTOs".

The screenshot shows the Chado Tutorial page. The title is "Chado Tutorial". Below the title, it says "2009 GMOD Summer School - Europe & Americas" and "July & August 2009". The authors are listed as "Scott Cain, Joshua Orvis, Dave Clements". The page includes a navigation sidebar with links for "GMOD Home", "Categories / Tags", "Downloads", and "View all pages". There is also a "documentation" section with links for "Overview" and "FAQs".

The screenshot shows the GBrowse Tutorial page. The title is "GBrowse Tutorial". Below the title, it says "2009 GMOD Summer School - Europe & Americas" and "July & August 2009". The author is listed as "Scott Cain". The page includes a navigation sidebar with links for "GMOD Home", "Categories / Tags", "Downloads", and "View all pages". There is also a "documentation" section with links for "Overview" and "FAQs".

⋮



[http://gmod.org/wiki/Training\\_and\\_Outreach#Online\\_Tutorials](http://gmod.org/wiki/Training_and_Outreach#Online_Tutorials)



# Agenda

Project Statistics

Training and Outreach

Genome Visualization

Grant Proposal

New Components

GMOD Web Site

Odds and Ends



[http://gmod.org/wiki/GMOD\\_Help\\_Desk](http://gmod.org/wiki/GMOD_Help_Desk)



# Genomics Visualization

The 2009 GMOD Community Survey focused on genome and comparative genomics visualization.

Open for 10 days in September and received 45 responses.

[http://gmod.org/wiki/2009\\_GMOD\\_Community\\_Survey](http://gmod.org/wiki/2009_GMOD_Community_Survey)

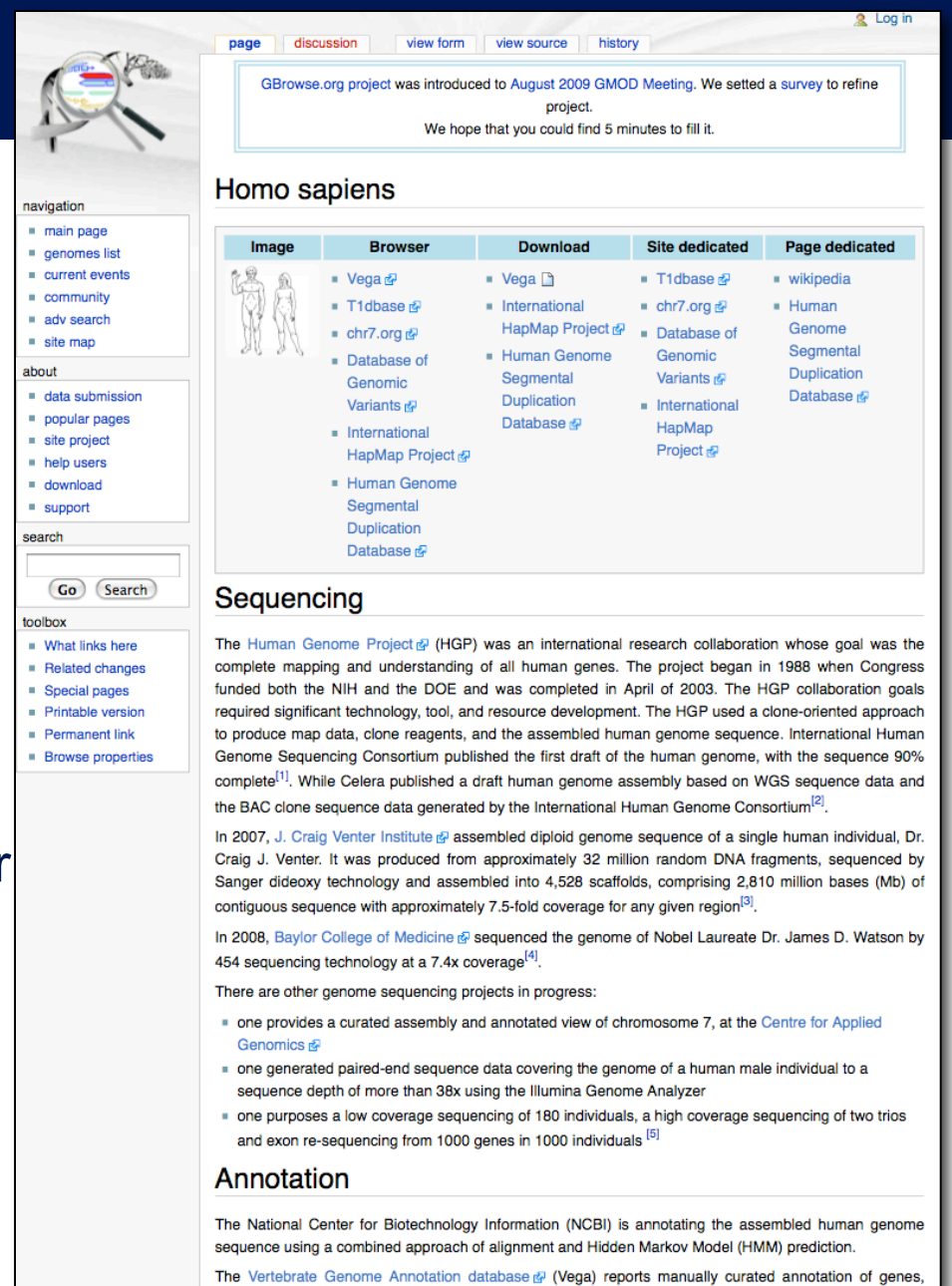


# GBrowse.org

## Resource for people setting up GBrowse:

- Links to genome browser web sites, by organism
- Links to sequence and annotation files
- Links to genome browser configuration files
- FTP site containing genome annotation and configuration files for organisms without web sites
- Also includes links to sites that use other browser packages too.


## Currently conducting a survey



page discussion view form view source history

GBrowse.org project was introduced to August 2009 GMOD Meeting. We setted a survey to refine project.  
We hope that you could find 5 minutes to fill it.

### Homo sapiens

Image	Browser	Download	Site dedicated	Page dedicated
	<ul style="list-style-type: none"><li>Vega</li><li>T1dbase</li><li>chr7.org</li><li>Database of Genomic Variants</li><li>International HapMap Project</li><li>Human Genome Segmental Duplication Database</li></ul>	<ul style="list-style-type: none"><li>Vega</li><li>International HapMap Project</li><li>Human Genome Segmental Duplication Database</li></ul>	<ul style="list-style-type: none"><li>T1dbase</li><li>chr7.org</li><li>Database of Genomic Variants</li><li>International HapMap Project</li></ul>	<ul style="list-style-type: none"><li>wikipedia</li><li>Human Genome Segmental Duplication Database</li></ul>

### Sequencing

The [Human Genome Project](#) (HGP) was an international research collaboration whose goal was the complete mapping and understanding of all human genes. The project began in 1988 when Congress funded both the NIH and the DOE and was completed in April of 2003. The HGP collaboration goals required significant technology, tool, and resource development. The HGP used a clone-oriented approach to produce map data, clone reagents, and the assembled human genome sequence. International Human Genome Sequencing Consortium published the first draft of the human genome, with the sequence 90% complete<sup>[1]</sup>. While Celera published a draft human genome assembly based on WGS sequence data and the BAC clone sequence data generated by the International Human Genome Consortium<sup>[2]</sup>.

In 2007, [J. Craig Venter Institute](#) assembled diploid genome sequence of a single human individual, Dr. Craig J. Venter. It was produced from approximately 32 million random DNA fragments, sequenced by Sanger dideoxy technology and assembled into 4,528 scaffolds, comprising 2,810 million bases (Mb) of contiguous sequence with approximately 7.5-fold coverage for any given region<sup>[3]</sup>.

In 2008, [Baylor College of Medicine](#) sequenced the genome of Nobel Laureate Dr. James D. Watson by 454 sequencing technology at a 7.4x coverage<sup>[4]</sup>.

There are other genome sequencing projects in progress:

- one provides a curated assembly and annotated view of chromosome 7, at the [Centre for Applied Genomics](#)
- one generated paired-end sequence data covering the genome of a human male individual to a sequence depth of more than 38x using the Illumina Genome Analyzer
- one purposes a low coverage sequencing of 180 individuals, a high coverage sequencing of two trios and exon re-sequencing from 1000 genes in 1000 individuals<sup>[5]</sup>

### Annotation

The National Center for Biotechnology Information (NCBI) is annotating the assembled human genome sequence using a combined approach of alignment and Hidden Markov Model (HMM) prediction.

The [Vertebrate Genome Annotation database](#) (Vega) reports manually curated annotation of genes,



# Agenda

Project Statistics

Training and Outreach

Genome Visualization

Grant Proposal

New Components

GMOD Web Site

Odds and Ends



[http://gmod.org/wiki/GMOD\\_Help\\_Desk](http://gmod.org/wiki/GMOD_Help_Desk)



# Grant Proposal

- *Enhancing the GMOD Suite of Genome Annotation and Visualization Tools*
  - JBrowse development
  - GBrowse development
  - *JBrowse\_syn* development
  - Help Desk
    - JBrowse, NGS data, transcriptomics, developmental atlases, population genomic and natural diversity data, workflow and pipeline tools for processing large data sets, and software interoperability
    - Continue with documentation, user support, gmod.org, training, (longer) summer schools, outreach, annual survey, ...
- Submitted in November
- If approved, will start in the fall



# Agenda

Project Statistics

Training and Outreach

Genome Visualization

Grant Proposal

New Components

GMOD Web Site

Odds and Ends



[http://gmod.org/wiki/GMOD\\_Help\\_Desk](http://gmod.org/wiki/GMOD_Help_Desk)





# New Components: Natural Diversity Chado Module

- Some progress!
- Natural diversity working group met twice during PAG
- GDR will be first user
- Sook Jung or Washington State U taking the lead



# New Components: 3 in Negotiation

- SOBA
  - GFF3 file analysis and reporting
  - Tabular and graphical reports
- Phenote
  - Grant proposal being submitted shortly
  - Proposing to generalize to all ontology based annotation
- CMap3D
  - Processing/Java 3D interface to CMap data



# Agenda

Project Statistics

Training and Outreach

Genome Visualization

Grant Proposal

New Components

**GMOD Web Site**

Odds and Ends



[http://gmod.org/wiki/GMOD\\_Help\\_Desk](http://gmod.org/wiki/GMOD_Help_Desk)



# GMOD.org in 2010

- Moving from CSHL to OICR in Jan/Feb
  - MediaWiki upgrade
  - Probably lots of new extensions
  - Maybe a modified skin
- New home for TableEdit
  - and will become widely used on gmod.org
- Look into adding
  - User log section
  - Scrapbook for contributed code



# Agenda

Project Statistics

Training and Outreach

Genome Visualization

Grant Proposal

New Components

GMOD Web Site

Odds and Ends



[http://gmod.org/wiki/GMOD\\_Help\\_Desk](http://gmod.org/wiki/GMOD_Help_Desk)



# Odds and Ends

- International Society for Biocuration
  - Please consider joining
  - <http://biocurator.org>
- GMOD Logo Program!
  - It's happening this spring
  - Call will go out later this month
  - Linn-Benton Community College
  - Get a new logo for \$75 and some time



*Linn-Benton*  
COMMUNITY COLLEGE



[http://gmod.org/wiki/GMOD\\_Logo\\_Program](http://gmod.org/wiki/GMOD_Logo_Program)



# Acknowledgements

**Summer Schools**  
**GMOD Europe**

Scott Cain  
Carson Holt  
Ed Lee  
Danielle Wilson  
Stephen Ficklin  
Mitch Skinner  
Joshua Orvis  
Junjun Zhang  
Simon McGowan  
Steve Taylor  
Ian Holmes  
Robin Houston  
Tim Carver  
Korbinian Schneeberger  
Phillips Lab

...

**SMBE**  
**PAG**

**Oregon**  
**UConn AGA**

Ben Faga  
Sheldon McKay  
Mark Yandell  
Lincoln Stein  
Jim Hu  
Peter Van Buren  
Daniel Renfro  
Todd Vision  
Nathan Liles  
Maryann Martone  
Brenley McIntosh  
Phillip Morin  
Lance Ferguson  
Heng Li  
Cresko Lab

...

**NESCent**  
**BA2009**

**TAMU**  
**ISyIP**  
**Wittenberg**

Hilmar Lapp  
Brian O'Connor  
Candace Brown  
Jon Auman  
Dan Blankenberg  
Jack D'Ardenne  
Matt Bellgard  
Denis Tagu  
Ranyelle Craig  
Rachel O'Neill  
Chuck Cannon  
Linda Straussbaugh  
Diethard Tautz  
Philip Johnson  
Bohannon Lab

...



# Thank You!



**Dave Clements**  
GMOD Help Desk  
US National Evolutionary  
Synthesis Center (NESCent)  
clements@nescent.org



[http://gmod.org/wiki/GMOD\\_Help\\_Desk](http://gmod.org/wiki/GMOD_Help_Desk)

