

Supporting Multiple Community Networks with Galaxy

Jeremy Goecks

Department of Biology
Department of Math and Computer Science
Emory University



EMORY
UNIVERSITY

Short Demo

The screenshot displays the Galaxy web interface. The browser address bar shows <https://main.g2.bx.psu.edu>. The Galaxy logo is in the top left, and the user's storage usage is shown as "Using 668.8 GB" in the top right. The main navigation bar includes "Analyze Data", "Workflow", "Shared Data", "Visualization", "Cloud", "Admin", "Help", and "User".

The left sidebar contains a "Tools" section with a search bar containing "filter". Below the search bar, there are several tool categories and their respective tools:

- FASTA manipulation**
 - Filter sequences by length
- Filter and Sort**
 - Filter data on any column using simple expressions
 - Filter on ambiguities in polymorphism datasets
 - GFF
 - Filter GFF data by attribute using simple expressions
 - Filter GFF data by feature count using simple expressions
 - Filter GTF data by attribute values list
- Fetch Alignments**
 - Filter MAF blocks by Species
 - Filter MAF blocks by Size
 - Filter MAF by specified attributes
- Regional Variation**
 - Filter nucleotides based on quality scores
- Genome Diversity**
 - INITIAL ANALYSIS
 - Filter SNPs : Discard some SNPs based on coverage or quality
 - POPULATION STRUCTURE
 - Prepare Input : Filter and

The main content area shows the configuration for the "Filter (version 1.1.0)" tool. The "Filter:" dropdown menu is set to "172: Cuffmerge-Cuffdiff XBP1 Isoform Tracking". Below it, the "With following condition:" field contains the expression `c1=='chr22'`. The "Number of header lines to skip:" field is set to "0". An "Execute" button is visible below the configuration fields.

Below the configuration fields, there are two tips:

- Warning:** Double equal signs, ==, must be used as "equal to" (e.g., `c1 == 'chr22'`)
- Tip:** Attempting to apply a filtering condition may throw exceptions if the data type (e.g., string, integer) in every line of the columns being filtered is not appropriate for the condition (e.g., attempting certain numerical calculations on strings). If an exception is thrown when applying the condition to a line, that line is skipped as invalid for the filter condition. The number of invalid skipped lines is documented in the resulting history item as a "Condition/data issue".
- Tip:** If your data is not TAB delimited, use *Text Manipulation* -> *Convert*

The "Syntax" section explains that the filter tool allows restricting the dataset using simple conditional statements. It notes that columns are referenced with `c` and a number, and that multi-character operators must not contain white space. It also states that the "equal-to" operator `==` must be used, and that non-numerical values must be in quotes. Filtering conditions can include logical operators, but all operators must be in lower case.

The "Example" section provides two examples:

- `c1=='chr1'` selects lines in which the first column is chr1
- `c3-c2<100*c4` selects lines where subtracting column 3 from column 2 is less than the value of column 4 times 100

The right sidebar shows the "History" section, which lists the execution history of the tool. The top entry is "BodyMap Lymph Node" (7.8 GB). Below it, a list of history items is shown, each with a status icon (eye, lock, and refresh) and a description of the tool and data used:

- 172: Cuffmerge-Cuffdiff XBP1 Isoform Tracking
- 171: Cuffcompare-Cuffdiff XBP1 Isoform Tracking
- 170: Cuffdiff on data 124, data 120, and others: transcript FPKM tracking
- 169: Cuffdiff on data 124, data 120, and others: transcript differential expression testing
- 168: Cuffdiff on data 124, data 120, and others: gene FPKM tracking
- 167: Cuffdiff on data 124, data 120, and others: gene differential expression testing
- 166: Cuffdiff on data 124, data 120, and others: TSS groups FPKM tracking
- 165: Cuffdiff on data 124, data 120, and others: TSS groups differential expression testing
- 164: Cuffdiff on data 124, data 120, and others: CDS FPKM tracking
- 163: Cuffdiff on data 124, data 120, and others: CDS FPKM

Vision

Galaxy is an **open, Web-based platform** for accessible, reproducible, and collaborative computational genomics

What is Galaxy?

GUI for high-throughput, high-performance genomics

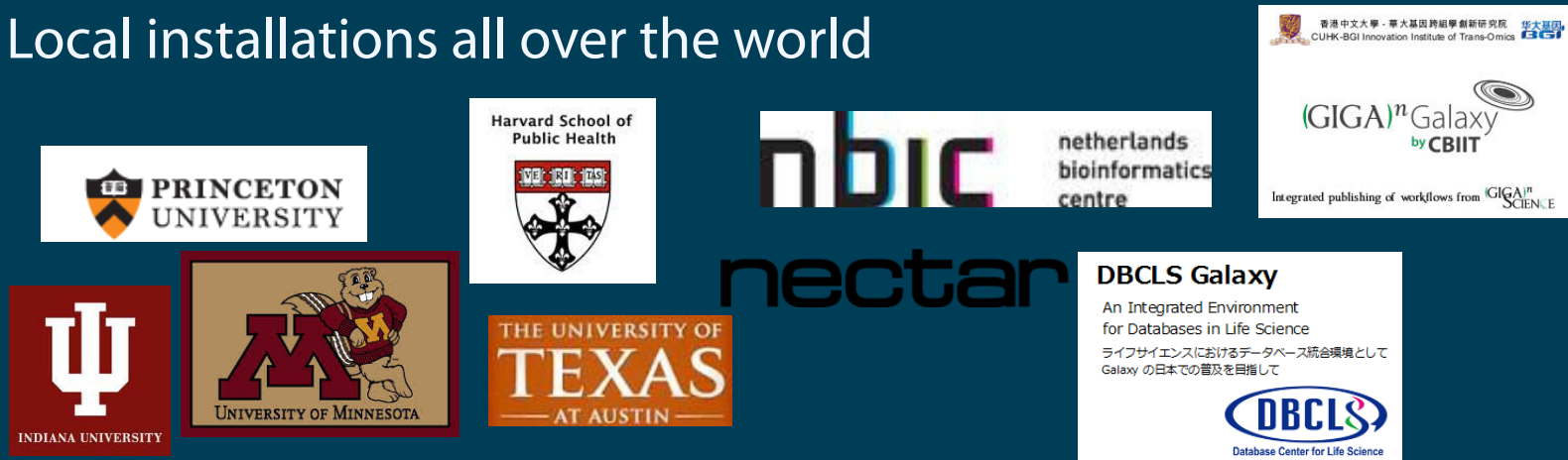
1. get and integrate public, private data
2. analyze data and create workflows
3. visualization and visual analysis, sharing, publication

Customizable open-source software on various HPC resources

- ✦ public website — <http://usegalaxy.org>
- ✦ local instance
- ✦ on the cloud

Galaxy is Very Popular

Local installations all over the world



Public Website (<http://usegalaxy.org>), anybody can use:

- ✦ ~500 new users per month, ~100 TB of user data, ~130,000 analysis jobs per month

Used and cited in more than 1000 publications

Galaxy Communities

End Users

Deployers/Maintainers

Developers

- ✦ plug-ins: tools, visualizations
- ✦ core framework

Goals

A community connected to the project and each other

A well-informed community

Grow the community

A community that contributes:

- ✦ analyses, curation, and data (=knowledge)
- ✦ tools and enhancements
- ✦ support
- ✦ outreach and training

How?

Enable your community

Communicate with your community

Reward your community

Sharing, Collaborating, and Publishing Analyses with Galaxy

Sharing and Publishing

Sharing and Publishing History 'Variant Analysis for Sample E18'

Making History Accessible via Link and Publishing It

This history is currently restricted so that only you and the users listed below can access it. You can:

Make History Accessible via Link

Generates a web link that you can share with other people so that they can view and import the history.

Make History Accessible and Publish

Makes the history accessible via link (see above) and publishes the history to Galaxy's [Published Histories](#) section, where it is publicly listed and searchable.

Sharing History with Specific Users

You have not shared this history with any users.

Share with a user

[Back to Histories List](#)


Sharing and Publishing

Sharing and Publishing History 'Variant Analysis for Sample E18'

Making History Accessible via Link and Publishing It

This history accessible via link and published.

Anyone can view and import this history by visiting the following URL:

<http://main.g2.bx.psu.edu/u/jgoecks/h/variant-analysis-for-sample-e18> 

This history is publicly listed and searchable in Galaxy's Published Histories section.

You can:

Unpublish History

Removes history from Galaxy's Published Histories section so that it is not publicly listed or searchable.

Disable Access to History via Link and Unpublish

Disables history's link so that it is not accessible and removes history from Galaxy's Published Histories section so that it is not publicly listed or searchable.

Sharing History with Specific Users

You have not shared this history with any users.

Share with a user

[Back to Histories List](#)

Galaxy | Published History | Variant Analysis for Sample E18

http://main.g2.bx.psu.edu/u/jgoecks/h/variant-analysis-for-sample-e18

Galaxy Analyze Data Workflow Shared Data Visualization Help User

Published Histories | jgoecks | Variant Analysis for Sample E18


Galaxy History ' Variant Analysis for Sample E18'

[+ Import history](#)

Annotation: Perform a pileup analysis with default parameters to identify variants in sample E18.

Dataset	Annotation
1: E18 PE.1 Reads	Forward reads from sample E18.
2: E18 PE.2 Reads	Reverse reads from sample E18.
3: E18 PE.1 Reads Groomed	Groom reads to convert quality scores from Solexa 1.0 to Solexa 1.3
4: E18 PE.2 Reads Groomed	Groom reads to convert quality scores from Solexa 1.0 to Solexa 1.3
5: E18 PE.1 Reads Groomed, Trimmed	Trim reads from 3' end to remove low-quality nts.
6: E18 PE.2 Reads Groomed, Trimmed	Trim reads from 3' to remove low-quality nts.
7: Map with Bowtie for Illumina on data 6 and data 5	Map paired-end reads with default parameters.
8: SAM-to-BAM on data 7	Need to convert Bowtie SAM to BAM so that pileup analysis can be performed.
9: Generate pileup on data 8	Pileup analysis with default parameters
10: Filter pileup to get Variants from sample E18	Find variants with coverage ≥ 30 .
13: Filter to get Variants from sample E18 where consensus base different than ref. base	Filter pileup to find variants where the consensus base is different than the reference base.
14: UCSC mm9 RefSeq Genes	UCSC mm9 RefSeq genes.
15: Intersect to get Variants from sample E18, consensus different, in RefSeq Genes	Variants with consensus different that occur in RefSeq genes.

About this History

Author  jgoecks

Related Histories
[All published histories](#)
[Published histories by jgoecks](#)

Rating
 Community (1 rating, 4.0 average) ★★★★☆
 Yours ☆☆☆☆☆

Tags
 Community: snp pileup bowtie demo
sample
 Yours: snp x pileup x bowtie x
demo x sample:e18 x [+](#)

Galaxy | Published Workflow | SNP variant detection from paired-end reads

http://main.g2.bx.psu.edu/u/jgoecks/w/snp-variant-detection-from-paired-end-reads

Galaxy Analyze Data Workflow Shared Data Visualization Help User

Published Workflows | jgoecks | SNP variant detection from paired-end reads

Step 6: FASTQ Trimmer Trim reads to remove low-quality bases.

FASTQ File
Output dataset 'output_file' from step 4

Define Base Offsets as Absolute Value

Offset from 5' end
0

Offset from 3' end
9

Keep reads with zero length
False

Step 7: Map with Bowtie for paired-end Map reads using Bowtie. Result parameter values.

Will you select a reference genome from your history or use a built-in index?
Use a built-in index

Select a reference genome
/galaxy/data/apiMel3/bowtie_index/apiMel3

Is this library mate-paired?
Paired-end

Forward FASTQ file
Output dataset 'output_file' from step 6

Reverse FASTQ file
Output dataset 'output_file' from step 5

Maximum insert size for valid paired-end alignment (-X)
1000

The upstream/downstream mate orientation for valid paired-end alignment against the forward reference strand (--fr/--rf/--ff)
FR (for illumina)


Bowtie settings to use
Commonly used

Suppress the header in the output SAM file
True

Step 8: SAM-to-BAM Convert Bowtie SAM output to BAM format so that pileup can be run.

Choose the source for the reference list
Locally cached

About this Workflow

Author
jgoecks 

Related Workflows
[All published workflows](#)
[Published workflows by jgoecks](#)

Rating
Community (0 ratings, 0.0 average) ★★★★★
Yours ★★★★★

Tags
Community:
snp bowtie
Yours:
snp x bowtie x

Galaxy | Published Histories

http://main.g2.bx.psu.edu/history/list_published

Galaxy Analyze Data Workflow Shared Data Visualization Help User

Published Histories

search [Advanced Search](#)

Name	Annotation	Owner	Community Rating [†]	Community Tags	Last Updated
Galaxy vs MEGAN	Comparison of Galaxy vs. MEGAN pipeline.	aun1	★★★★★	metagenomics megan galaxy	Mar 19, 2010
metagenomic analysis		aun1	★★★★★	metagenomics galaxy	Mar 19, 2010
SM_1186088	Datasets correspond to our paper published in Science by Peleg et al. entitled : Altered histone acetylation is associated with age-dependent memory impairment. Experiment layout: This history contains 4 datasets in the form of BED files of uniquely mapped reads produced after chip-seq for histone modifications H4K12ac and H3K9ac in mouse hippocampus of 3 months (young) and 16 months (old) mice after fear conditioning. For detailed information please refer to supplementary materials and methods of the respective work by peleg et al.	fischerlab	★★★★★		Apr 19, 2010
Variant Analysis for Sample E18	Perform a pileup analysis with default parameters to identify variants in sample E18.	jgoecks	★★★★★	snp pileup bowtie demo sample	2 minutes ago
get longest exon		henri	★★★★★	chr22 longest marc exon human workshop	Sep 02, 2010
FASTA to Tabular Test		JJ	★★★★★		Aug 26, 2010
EKLF		yzc109	★★★★★		Aug 24, 2010

Open "http://main.g2.bx.psu.edu/history/list_published?sort=rating&f-tags=All" in a new tab

Interactive Research Documents

Galaxy | Published Page | p...
https://main.g2.bx.psu.edu/u/webb/p/polar-bears

Galaxy
Analyze Data Workflow Shared Data Visualization Cloud Admin Help User

Published Pages | webb | polar-bears

Polar and brown bear genomes reveal ancient admixture and demographic footprints of past climate change

Webb Miller, Stephan C. Schuster, Andreanna J. Welch, Aakrosh Ratan, Oscar C. Bedoya-Reina, Fangqing Zhao, Hie Lim Kim, Richard C. Burhans, Daniela I. Drautz, Nicola E. Wittekindt, Lynn P. Tomsho, Enrique Ibarra-Laclette, Luis Herrera-Estrella, Elizabeth Peacock, Sean Farley, George K. Sage, Karyn Rode, Martyn Obbard, Rafael Montiel, Lutz Bachmann, Ólafur Ingólfsson, Jon Aars, Thomas Mailund, Øystein Wiig, Sandra L. Talbot, and Charlotte Lindqvist

Summary of the paper

Polar bears (PBs) are superbly adapted to the extreme Arctic environment and have become emblematic of the threat to biodiversity from global climate change. Their divergence from the lower-latitude brown bear provides a textbook example of rapid evolution of distinct phenotypes. However, limited mitochondrial and nuclear DNA evidence conflicts in the timing of PB origin as well as placement of the species within versus sister to the brown bear lineage. We gathered extensive genomic sequence data from contemporary polar, brown, and American black bear samples, in addition to a 130,000- to 110,000-y old PB, to examine this problem from a genome-wide perspective. Nuclear DNA markers reflect a species tree consistent with expectation, showing polar and brown bears to be sister species. However, for the enigmatic brown bears native to Alaska's Alexander Archipelago, we estimate that not only their mitochondrial genome, but also 5-10% of their nuclear genome, is most closely related to PBs, indicating ancient admixture between the two species. Explicit admixture analyses are consistent with ancient splits among PBs, brown bears and black bears that were later followed by occasional admixture. We also provide paleodemographic estimates that suggest bear evolution has tracked key climate events, and that PB in particular experienced a prolonged and dramatic decline in its effective population size during the last ca. 500,000 years. We demonstrate that brown bears and PBs have had sufficiently independent evolutionary histories over the last 4-5 million years to leave imprints in the PB nuclear genome that likely are associated with ecological adaptation to the Arctic environment.

Datasets

Many of the analyses reported in the paper were based on the five datasets given here. (You can also find them under Shared Data -> Data Libraries -> Genome Diversity, then under bear and dog.)

The first consists of 12,023,192 dog-based "SNPs", i.e., positions in the dog genome where we detected two distinct nucleotides in the corresponding bear locations (among the our three bear species, polar bear, brown bear, and American black bear). Each row in the table corresponds to a SNP, and has [124 entries](#).

[Galaxy Dataset | bear SNPs](#)

The "bear assembly SNPs" table contains 13,038,705 putative SNPs that were identified using a de novo assembly of the polar bear genome (rather than the dog assembly). Each row of the table corresponds to a SNP, and has [117 columns](#).

[Galaxy Dataset | bear assembly SNPs](#)

The "bear mitochondrial SNPs" table contains 1,698 positions where not all 28 individuals had the same nucleotide. Each row represents one of these SNPs, and has [31 columns](#).

[Galaxy Dataset | bear mitochondrial SNPs](#)

The "bear SAPs" table contains 79,501 variant position in putative protein-coding regions, both synonymous and non-synonymous changes. Each row has [11 columns](#).

[Galaxy Dataset | bear SAPs](#)

One of the workflows (bear sweep table) uses a streamlined file with the locations of 19,014 dog genes (basically, each one is the longest of a set of overlapping splice variants). Each gene corresponds to a row of the table, which has [5 columns](#).

[Galaxy Dataset | dog genes](#)

Workflows

This page presents three "workflows" that produce results presented in the polar-bear paper. Almost all of the commands that they use are from the "Genome Diversity" tool set. (See the left panel under "Analyze Data".)

The first workflow generates the data for [Figure 4A](#) of the paper. (Those data were used to produce a more attractive PCA plot that includes other information.) The workflow needs to be applied to the "bear SNPs" data set as follows: (1) Under "Analyze Data" (in the black bar) create an empty history. (2) Under "Shared Data" -> "Published Pages", view this page. (3) Import the "bear SNPs" data set ("+" in the green circle near the right of the green bar), then click on "return to the previous page". (4) Import the "Bear PCA" workflow, and click on "start using this workflow". (5) You will be taken to your Workflow page, which will have a workflow called "Imported bear PCA"; click on it and select "run". (6) You will be taken to a history that includes the bear SNPs and the PCA workflow; scroll to the bottom of the workflow (middle panel) and press "Run workflow". (7) After the commands run (which takes a couple of minutes), click on the "eye" for the PCA command and look at the three Outputs. [Currently, the PCA workflow exposes an internal error - a so-called "race condition" -- in Galaxy, which may cause the PCA command to fail. If that happens, you can re-run the PCA (not the entire workflow) by clicking on the line that says something like "7: PCA on data 6", clicking on the blue re-run button, and clicking on "Execute". You also may need to give Galaxy a minute after the workflow finishes to put the output files in the correct places.]

[Galaxy Workflow | bear PCA](#)

The second workflow produces the admixture map for the two ABC bears, showing the genomic intervals (relative to the dog assembly) where one or both of an ABC bear's autosomes is (are) more like the consensus of the polar-bear genome than like the genome of the non-ABC brown bear (called "GRZ" in the paper). The [figure](#) produced by running the workflow is a small improvement over [Figure S12](#) of the supplement (which has one chromosome shown in [Figure 4B](#) of the main paper). The new figure indicates the 3Mb interval on the left end of each dog chromosome, which are treated as heterochromatin in the dog assembly (i.e., containing only 3 million copies of the letter "N"). When you run the workflow, the last command produces two history items. The "eye" in the first one shows a text file giving coordinates of the genomic intervals where chromosomes look most like a particular group of individuals. The second "eye" leads you to the graphical picture and additional information.

[Galaxy Workflow | bear admixture map](#)

The third workflow produces a table of the 58 highest-scoring genomic intervals (relative to the dog assembly) showing signs of a "selective sweep" in polar bears, i.e., where an allele having a selective advantage increased in frequency in the population and brought along with it the neighboring alleles. The table appeared as [Table S8](#) in the Supplement, and one interval is shown in [Figure 7](#) of the main paper. To run the workflow you will need to place both the "bear SNP" file and the "dog genes" file in your history. (Make sure before you press "Run workflow" that the workflow's inputs are connected to the proper files.) When the workflow has run, you can click on the "eye" for the last command to see the table.

[Galaxy Workflow | bear sweep table](#)

Galaxy | Published Page | p...
 https://main.g2.bx.psu.edu/u/webb/p/polar-bears

Galaxy Analyze Data Workflow Shared Data Visualization Cloud Admin Help User Using 588.3 GB

Published Pages | webb | polar-bears

During the last ca. 300,000 years, we demonstrate that brown bears and PBS have had sufficiently independent evolutionary histories over the last 4-5 million years to leave imprints in the P8 nuclear genome that likely are associated with ecological adaptation to the Arctic environment.

Datasets

Many of the analyses reported in the paper were based on the five datasets given here. (You can also find them under Shared Data -> Data Libraries -> Genome Diversity, then under 'bear and dog'.)

The first consists of 12,023,192 dog-based "SNPs", i.e., positions in the dog genome where we detected two distinct nucleotides in the corresponding bear locations (among the our three bear species, polar bear, brown bear, and American black bear). Each row in the table corresponds to a SNP, and has 124 entries.

Galaxy Dataset | bear SNPs

The "bear assembly SNPs" table contains 13,038,705 putative SNPs that were identified using a de novo assembly of the polar bear genome (rather than the dog assembly). Each row of the table corresponds to a SNP, and has 117 columns.

Galaxy Dataset bear assembly SNPs										
scaffold1	370	T	C	999	36	0	2	135		
scaffold1	441	A	G	89.9	41	0	2	150		
scaffold1	793	C	G	999	19	14	1	69		
scaffold1	1057	T	C	999	25	19	1	228		
scaffold1	1074	C	T	999	27	18	1	214		
scaffold1	1464	G	T	999	14	6	1	29		
scaffold1	1693	C	T	999	0	26	0	75		
scaffold1	1948	C	G	91.2	0	5	0	12		
scaffold1	1950	A	G	999	0	6	0	15		
scaffold1	1963	A	G	91.4	0	5	0	12		
scaffold1	1968	G	C	95.4	0	4	0	9		
scaffold1	3756	G	T	999	34	0	2	129		
scaffold1	3864	C	A	999	41	0	2	150		
scaffold1	4044	G	A	999	0	39	0	114		
scaffold1	4723	G	A	116	30	0	2	117		
scaffold1	4901	C	T	999	0	30	0	87		
scaffold1	5591	C	A	999	0	36	0	105		
scaffold1	5969	T	C	999	0	28	0	81		

The "bear mitochondrial SNPs" table contains 1,698 positions where not all 28 individuals had the same nucleotide. Each row represents one of these SNPs, and has 31 columns.

Galaxy Dataset | bear mitochondrial SNPs

The "bear SAPs" table contains 79,501 variant position in putative protein-coding regions, both synonymous and non-synonymous changes. Each row has 11 columns.

Galaxy Dataset | bear SAPs

One of the workflows (bear sweep table) uses a streamlined file with the locations of 19,014 dog genes (basically, each one is the longest of a set of overlapping splice variants). Each gene corresponds to a row of the table, which has 5 columns.

Galaxy | Published Page | p...
 https://main.g2.bx.psu.edu/u/webb/p/polar-bears

Galaxy Analyze Data Workflow Shared Data Visualization Cloud Admin Help User Using 588.3 GB

Published Pages | webb | polar-bears

be taken to your Workflow page, which will have a workflow called "Imported bear PCA"; click on it and select "run". (6) You will be taken to a history that includes the bear SNPs and the PCA workflow; scroll to the bottom of the workflow (middle panel) and press "Run workflow". (7) After the commands run (which takes a couple of minutes), click on the "eye" for the PCA command and look at the three Outputs. (Currently, the PCA workflow exposes an internal error—a so-called "race condition"—in Galaxy, which may cause the PCA command to fail. If that happens, you can re-run the PCA (not the entire workflow) by clicking on the line that says something like "?: PCA on data 6", clicking on the blue re-run button, and clicking on "Execute". You also may need to give Galaxy a minute after the workflow finishes to put the output files in the correct places.)

Galaxy Workflow | bear PCA

The second workflow produces the admixture map for the two ABC bears, showing the genomic intervals (relative to the dog assembly) where one or both of an ABC bear's autosomes is (are) more like the consensus of the polar-bear genome than like the genome of the non-ABC brown bear (called "GRZ" in the paper). The figure produced by running the workflow is a small improvement over Figure S12 of the supplement (which has one chromosome shown in Figure 4B of the main paper). The new figure indicates the 3Mb interval on the left end of each dog chromosome, which are treated as heterochromatin in the dog assembly (i.e. containing only 3 million copies of the letter "N"). When you run the workflow, the last command produces two history items. The "eye" in the first one shows a text file giving coordinates of the genomic intervals where chromosomes look most like a particular group of individuals. The second "eye" leads you to the graphical picture and additional information. [Import workflow](#)

Galaxy Workflow | bear admixture map

Step 6: Filter

Filter
 Output dataset 'output' from step 5

With following condition
 c6!="chrX" and c12>=0.5

Step 7: Admixture

SNP dataset
 Output dataset 'out_file1' from step 6

Ancestral population 1 individuals
 Output dataset 'output' from step 3

Galaxy Workflow | bear sweep table

Galaxy
 https://main.g2.bx.psu.edu/workflow/imp?id=b7b9

Galaxy Analyze Data Workflow Shares

Workflow "bear admixture map" has been imported.
 You can start using this workflow or return to the previous page.

Uses

Biomedical collaboration

Publication support

Teaching and demonstration

Galaxy Tool Shed

toolshed.g2.bx.psu.edu

Galaxy Tool Shed

Repositories Help User

2132 valid tools on Oct 06, 2012

Search

- Search for valid tools
- Search for workflows

All Repositories

- Browse by category

Available Actions

- Login to create a repository

Categories

search repository name, description

Name	Description	Repositories
Assembly	Tools for working with assemblies	22
Computational chemistry	Tools for use in computational chemistry	4
Convert Formats	Tools for converting data formats	29
Data Source	Tools for retrieving data from external data sources	12
Fasta Manipulation	Tools for manipulating fasta data	24
Genomic Interval Operations	Tools for operating on genomic intervals	20
Graphics	Tools producing images	14
Metagenomics	Tools enabling the study of metagenomes	6
Micro-array Analysis	Tools for performing micro-array analysis	0
Next Gen Mappers	Tools for the analysis and handling of Next Gen sequencing data	40
Ontology Manipulation	Tools for manipulating ontologies	5
Proteomics	Tools enabling the study of proteins	2
SAM	Tools for manipulating alignments in the SAM format	19
Sequence Analysis	Tools for performing Protein and DNA/RNA analysis	109
SNP Analysis	Tools for single nucleotide polymorphism data such as WGA	16
Statistics	Tools for generating statistics	26
Systems Biology	Systems biology tools	0
Text Manipulation	Tools for manipulating data	24
Tool Generators	Tools that make or help make new tools	1
Visualization	Tools for visualizing data	23
Web Services	Tools enabling access to web services	1

How?

Enable your community

Communicate with your community

Reward your community

Provide Open Support Channels

The screenshot shows the 'Galaxy Development List Archive' website. The page title is 'Galaxy Development List Archive'. Below the title, there is a paragraph stating: 'This forum is an archive for the mailing list galaxy-dev@bx.psu.edu ([more options](#)) Messages posted here will be sent to this mailing list. Archive for the *Galaxy-Dev* mailing list. If you have a question about deploying, enhancing, tuning or adding to a *Galaxy* instance then this is a good place to find an answer.

Below this, there is a paragraph describing Galaxy: 'Galaxy is an open, web-based platform for *accessible, reproducible, and transparent* computational biomedical research.'

There is a bulleted list of features:

- **Accessibility:** Galaxy enables users without programming experience to easily specify parameters and run tools and workflows.
- **Reproducibility:** Galaxy captures all information necessary so that any user can repeat and understand a complete computational analysis.
- **Transparency:** Galaxy enables users to share and publish analyses via the web and create Pages--interactive, web-based computational analysis.

Further down, it says: 'Galaxy is open source for all organizations. The [public Galaxy service](#) makes analysis tools, genomic data, tutorial demonstrations and publication services available to any scientist that has access to the Internet. Local Galaxy servers can be set up by downloading and customizing it to meet particular needs.'

At the bottom, there is a 'Subscribe to Galaxy Dev' button with a link to 'Manage your subscription'.

At the very bottom, there is a table of topics:

Topics (6015)	Replies	Last Post
Cloudman - is it possible to still use deprecated AMIs ? by rssetia	0	10:30pm by rssetia
Cloudman master cannot connect to volume at start-up by rssetia	0	10:13pm by rssetia
Problem with tophat in local instance by Rainy Luo	0	7:52pm by Rainy Luo
Adding files to a data library - genome option by Lionel Guy	1	7:42pm by Carl Eberhard
Problems with tophat in local instance by Rainy Luo	0	7:33pm by Rainy Luo

The screenshot shows an IRC chat window for the channel #galaxyproject. The chat log shows the following messages:

```
22:28 clements has joined (~clements@gl[obal]-1-98.nat.csis.com .RC.UK)
22:28 Topic: The Galaxy Project, High powered genomics for everyone - http://galaxyproject.org | Use now! http://usegalaxy.org
22:28 natefoo set the topic at: Mar 7, 2011 10:00 PM
22:28 Mode: +nt
22:28 Created at: Mar 7, 2011 7:17 PM
22:49 guerler has joined (-user@c-76-185-99-34.hsd1.pa.comcast.net)
00:41 lbragg has joined (-lbragg84@footloose-sl.sl.csiro.au)
00:52 lbragg: Hi. I was wondering if anyone could point me towards an example of a wrapper which wraps a tool that produces HTML output (and subdirectories with images). Thanks
00:54 ceberhard: You might be able to get what you need from the fastqc wrapper.
00:56 lbragg: I'll have a look. I tried to get inspiration from the genetics one, but it was more confusing than anything!
00:57 ceberhard: It may get you started - it incorporates all the elements you mentioned (last time I used it).
00:57 lbragg: Awesome, thanks.
01:18 lbragg has left IRC (Quit: Leaving)
01:23 kellrott has left IRC (Quit: kellrott)
```

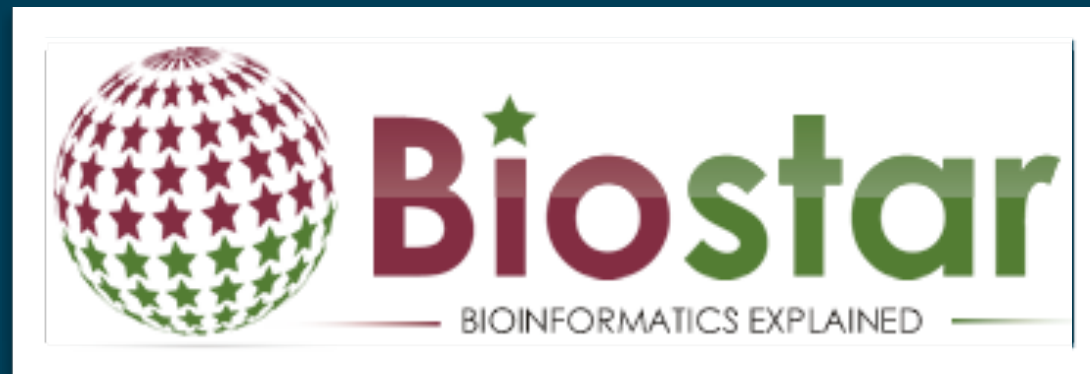
On the right side of the chat window, there is a list of users: clements, acu, ceberhard, dannon, dblank, erimar77, flu, guerler, InitHello, jgoecks, jmchilton1, malex, mrscribe, natefoo. At the bottom right, it says 'irc.freenode.net #galaxyproject'.

Open channels support anyone posting questions *and answers*

Mailing Lists vs Forums?

We are **replacing** Galaxy-User mailing list with a **forum**.

We will be using Biostar (<http://biostars.org>) directly



Incentivize contributions

BioStar is a **gamified** online forum. Points, badges, voting, ...

Question: Is there a list of public Galaxy servers?

I've been looking into setting up a local Galaxy installation for our bioinformatics core facility and in the process of doing so I've stumbled across several Galaxy mirrors and instances that have been customized in various ways. Before we go re-inventing the wheel I was wondering if there are any other known Galaxy servers out there that we could use to borrow design ideas or code from.

Here is the list I've put together thus far, others would be most welcome...

- Main Galaxy Server: <http://main.g2.bx.psu.edu/>
- Test Galaxy Server: <http://test.g2.bx.psu.edu/>
- GeneNetwork Galaxy Mirror: <http://galaxy.genenetwork.org:8080/>
- NBIC Galaxy Server: <http://galaxy.nbic.nl/> (includes proteomics tools)
- Galaxy/Rätsch Lab Server: <http://galaxy.fml.mpg.de/> (includes machine learning based tools for sequence and tiling array data analysis)
- Cistrome Galaxy Server: <http://cistrome.org/ap/root> (dry-lab workbench for integrative analysis of ChIP-chip/seq and gene expression data)

created 21 months ago by Casey Bergman ♦
13,700 • 2 • 13 • 37
updated 16 months ago by tnatat

Editor: Casey Bergman ♦

Userid: 314
Website: <http://bergmanlab.smith.man.ac.uk/>
Location: Manchester, UK
Member for: 2 years, 10 months
Last seen: 10 hours, 56 minutes ago

Molecular Evolutionist turned Genome Annotation, Genomics Work Blog: <http://bergmanlab.com/>
Work Twitter: <http://twitter.com/caseybergman>
Personal Blog: <http://caseybergman.com/>
Personal Twitter: <http://twitter.com/caseybergman>

13,700 • 2 • 13 • 37
Send Message

Status
Casey Bergman has contributed 647 posts: 15 questions, 287 answers and 340 comments. User has voted 1...

Notifications (25) Content Created Bookmarks (35) Upvoted Posts Supporters Badges (52) Moderator Actions

• Popular Question 9 weeks ago	• Popular Question 10 weeks ago	• Popular Question 11 weeks ago	• Popular Question 11 weeks ago	• Notable Question 3 months ago
• Notable Question 3 months ago	• Popular Question 3 months ago	• Popular Question 5 months ago	• Famous Question 12 months ago	• Nice Question 12 months ago
• Necromancer 12 months ago	• Nice Answer 12 months ago	• Popular Question 12 months ago	• Nice Answer 14 months ago	• Necromancer 14 months ago
• Necromancer 15 months ago	• Good Question 16 months ago	• Nice Answer 16 months ago	• Nice Question 16 months ago	• Nice Answer 17 months ago
• Nice Question 17 months ago	• Enlightened 17 months ago	• Nice Answer 17 months ago	• Nice Answer 17 months ago	• Nice Answer 17 months ago
• Nice Question 19 months ago	• Nice Answer 19 months ago	• Self-Learner 20 months ago	• Nice Answer 21 months ago	• Popular Question 21 months ago

Community Resources and Hubs

The screenshot shows the Galaxy Wiki homepage. At the top, there's a navigation bar with "Galaxy Wiki" and user options like "JeremyGoecks", "Settings", and "Logout". A search bar is also present. The main content area features the "Galaxy" logo and a brief description: "Galaxy is an open, web-based platform for accessible, reproducible, and transparent computational biomedical research." Below this, there are three key points: "Accessible", "Reproducible", and "Transparent". The page is divided into sections for "Use Galaxy", "Deploy Galaxy", and "Contribute". There are also logos for "usegalaxy.org" and "getgalaxy.org".

The screenshot shows a LinkedIn group page for "Galaxy Project". The page header includes "Account Type: Basic | Upgrade" and "Dave Clements" with "Add Connections". The group name "Galaxy Project" is prominently displayed. Below the name, there are tabs for "Discussions", "Members", "Promotions", "Jobs", "Search", "Manage", and "More...". The "Discussions" tab is active, showing a discussion titled "GalaxyAdmins Web Meetup: 10 am US Central, March 20". The discussion content includes the text: "The next meeting of the GalaxyAdmins Group will be held on March 20, 2013, at 10 AM Central US time (see http://wiki.galaxyproject.org/C...". There are also "Like", "Comment", and "Unfollow" buttons. On the right side, there are sections for "Latest Updates" and "Ads By LinkedIn Members".

Wikis and social sites:
communicate, enable,
reward.

Communicate!

Galaxy Wiki DaveClements Settings Logout | Search: _____

News

News

Announcements of interest to the Galaxy Community. These can include items from the Galaxy Team or the Galaxy community and can address anything that is of wide interest to the community.

The Galaxy News is also available as an [RSS feed](#).

See [Add a News Item](#) below for how to get an item on this page, and the RSS feed. Older news items are available in the [Galaxy News Archive](#).

See also


- Galaxy News Briefs
- Galaxy Updates

News Items

- GCC2013 & Galaxy GigaScience Series
 - April 2013 Galaxy Update
 - April 1, 2013 Galaxy Distribution
 - Galaxy LinkedIn Group
 - March 2013 GalaxyAdmins Meetup
 - Main & Test ServerDowntime: 3/14
 - March 2013 Galaxy Update
- GCC2013 Abstract Submission & Registration
- Galaxy, GMOD2013 & Biocuration 2013
- Feb 8, 2013 Galaxy Distribution & News Brief
- February 2013 Galaxy Update
- GCC2013 Training Day Training Material

April 2013 Galaxy Update

The April 2013 Galaxy Update is now available.



Highlights:

- the GCC2013 **oral presentation abstract deadline is 12 April**, which is less than 2 weeks away. Early registration, and poster abstract submission are also open.
- Two new public Galaxy servers are featured
- The slides and screencast for the March GalaxyAdmins meetup are available.
 - And, please help determine what the GalaxyAdmins group should focus on
- 32 new papers and 5 new tags
- Open Po
- Other Up
- Galaxy T
- Tool She
- Other Ne

Tweets

- Following
- Followers
- Favorites
- Lists

Tweet to Galaxy Project

@galaxyproject

Trends - Change

- #asechat
- #RoyalBaby
- #thevampsfollowspree
- #MCM
- #IWDate
- Denis Farina
- Buckingham Palace
- Carlos Hyde
- Kate Middleton
- Duchess of Cambridge

© 2013 Twitter About Help Terms Privacy
 Blog Status Apps Resources Jobs
 Advertisers Businesses Media Developers



Galaxy Project

@galaxyproject FOLLOWS YOU

Galaxy = bringing developers and biologists together. Reproducible science is our goal.

Penn State | Emory · galaxyproject.org

1,277 TWEETS 33 FOLLOWING 1,548 FOLLOWERS

Tweets

Galaxy Project @galaxyproject 17h
 At #ISMB2013? Check out posters B55 (Galaxy LIMS), F57, H37, N047, and O097 (Galaxy-P) bit.ly/ismb2013gxy #usegalaxy [Expand](#) [Reply](#) [Retweet](#) [Favorite](#) [More](#)

Galaxy Project @galaxyproject 17h
 At #ISMB2013? Check out Big Data Publishing talk by @SCEdmunds at 11:30 bit.ly/ismbwk03 #usegalaxy [Expand](#)

Jeremy @jgoecks 21 Jul
 @galaxyproject new time, new title for #ISMB2013 talk: 2:10p, Understanding Cancer Genomes (and Transcriptomes!) Using Galaxy [Retweeted by Galaxy Project](#) [View conversation](#)

Galaxy Wiki DaveClements Settings Logout | Search: _____

Events

If you have a calendar of Galaxy-related events, please add it here.

Galaxy Event Horizon

Events with Galaxy-related content are listed here.

Also see the [Galaxy Events Google Calendar](#) for a listing of events and deadlines that are relevant to the Galaxy Community. This is also available as an [RSS feed](#).

If you know of any event that should be added to this page and/or to the Galaxy Event Calendar, please add it here or send it to outreach@galaxyproject.org.

Contents

- Upcoming Events
- Other Calendars
- Past Events
 - 2013
 - Archive

Upcoming Events





Date	Topic/Event	Venue/Location	Contact
April 5-6	2013 GMOD Meeting	Cambridge, United Kingdom, immediately prior to Biocuration 2013	Dave Clements
April 7-10	GO Galaxy Workshop	Biocuration 2013, Cambridge, United Kingdom	Dave Clements, Suzanna Lewis
April 7-8	BOSC/Broad Interoperability Hackathon	Cambridge, Massachusetts, United States	Brad Chapman
April 9-11	Workshop: <i>Integrated Research Data Management for Next Gen Sequencing Analysis Using Galaxy and Globus Online Software-as-a-Service</i>	BioIT World, Boston, Massachusetts, United States	Ravi K. Madduri, Alex R. Paciorkowski, Vas Vasiliadis
	Talk: <i>Integrated Research Data management and Analysis in NGS using Globus Online, Galaxy and Accession</i>		Ravi K. Madduri

How?


Enable your community

Communicate with your community

Reward your community

Reward Community Efforts

 **madduri** 2 days
Looking forward to #BioIT13. Please join us in our pre-conference workshop on NGS analysis using @globusonline and @galaxyproject @awscloud
↕ by galaxyproject

 **Swiss Galaxy Workshop**
Bern, 3 October 2012
<http://bit.ly/gxyswiss>

CiteULike MyCiteULike Group: Galaxy Search Logged in as galaxyproject Log Out

Group: Galaxy - library 926 articles
You are an administrative member of this group.
Invite other CiteULike users to join, or invite people who don't use CiteULike yet.

Search Unwatch Copy Export Sort Hide Details

Group Tags
All tags in the group Galaxy
Filter: []
[Display as Cloud]

- Acidobacterial community responses to agricultural management**
soils
FEMS Microbiology Ecology, Vol. 83, No. 3, (March 2013), pp. 607-621, doi:10.1016/j.femsec.2012.12.011
by Acácio A. Navarrete, Eiko E. Kuramae, Mattias de Hollander, Anata S. Pillay, J. posted to methods by galaxyproject to the group Galaxy keyed Navarrete2013
Abstract Copy My Copy
- Parallelization in Scientific Workflow Management Systems**
(28 Mar 2013)
by Marc Bux, Jif Leser
posted to workbench by galaxyproject to the group Galaxy keyed Bux2013Par
Abstract Copy My Copy
- Ecology of Subglacial Lake Vostok (Antarctica), Based on Molecular Analyses of Accretion Ice**
Biology, Vol. 2, No. 2, (28 March 2013), pp. 629-650, doi:10.3390/biology2020
by Scott Rogers, Yuri Shtarkman, Zeynep Kocer, Rebyn Edgar, Ram Veerappa
posted to methods by galaxyproject to the group Galaxy keyed Rogers2013Ec
Abstract Copy My Copy
- The Role of the Arabidopsis Exosome in siRNA-Independent Silencing of Heterochromatic Loci**
PloS Genet, Vol. 9, No. 3, (28 March 2013), e1003411, doi:10.1371/journal.pgen.1003411
by Jun-Hye Shin, Hsiao-Lin V. Wang, Jinwon Lee, Brandon L. Dinwiddie, Dmitry A. Belostotsky, Julia A. Chekanova
posted to methods by galaxyproject to the group Galaxy keyed Shin2013Role on 2013-04-02 20:10:43 **/
Abstract Copy My Copy
- The challenges of delivering bioinformatics training in the analysis of high-throughput data**
Briefings in Bioinformatics (29 March 2013), doi:10.1093/bib/bbt018
by Benilton S. Carvalho, Gabriela Rueda
posted to workbench by galaxyproject to the group Galaxy keyed Carvalho2013Challenges on 2013-04-02 20:08:26 **/
along with 1 person
Abstract Copy My Copy
- Diagnostic Cancer Genome Sequencing and the Contribution of Germline Variants**
Science, Vol. 339, No. 6127, (29 March 2013), pp. 1559-1562, doi:10.1126/science.1233899

Publicly Accessible Galaxy Servers

Galaxy's public server (UseGalaxy.org, Main) can meet many needs, but it is not suitable for everything (see Big Picture/Choices for why) and cannot possibly scale to meet the entire world's needs.

Fortunately the Galaxy Community is helping out by installing Galaxy at their institutions and then making those installations either publicly available or open to their organizations or community.

This page lists such public or semi-public Galaxy servers.

To add your public Galaxy server to this list, please either just add it (hey, it's a wiki), or contact Galaxy Outreach

Andromeda

- Link:**
[Andromeda](#)
- Domain/Purpose:**
This is a fully populated Galaxy instance.
- Comments:**
Andromeda is hosted at the SURFSara High Performance Computing (HPC) cloud. The installation is supported by Enis Afgan (CloudMan project) and Mattias de Hollander (NIOO).
- Quotas:**
Registered users: 10GB; Anonymous users: 10MB
- Sponsor(s):**
Netherlands Bioinformatics Centre (NBIC) and BiG Grid SURFSara



ballaxy

- Link:**
[ballaxy](#)
- Domain/Purpose:**
Hosts the BALL (Biochemical Algorithms Library) Project tools, i.e. computer aided drug design and molecular modelling based on protein and ligand structure data.
- Comments:**
ballaxy is a workflow framework for structure based computational biology based on the Galaxy workflow engine and the BALL (Biochemical Algorithms Library) application framework. It is tailored to handle structural molecular data (pdb, mol, mol2, hin, xyz, smiles) and to offer tools for modelling tasks like chemical shift prediction "NightShift" or optimal bond order assignment of ligands "BOA Constructor".
- Quotas:**
A login is required and everybody can create a login, but there is no guarantee how long any data will be preserved.
- Sponsor(s):**
The groups of Hans-Peter Lenhof (Saarland University, Saarbrücken, Germany), Oliver Kohlbacher (University of Tübingen, Germany), and Andreas Hildebrandt (University of Mainz, Germany).



Contents

1. Andromeda
2. ballaxy
3. Cistrome Analysis Pipeline
4. DBCLS Galaxy
5. Galaxy Main
6. Galaxy Test
7. GeneNetwork
8. Genboree
9. Genomic Hyperbrowser
10. Gene Ontology (GO)
11. GigaGalaxy
12. GWIPS-viz
13. Huttenhower Lab
14. IBDsite
15. INRA-URGI
16. MGTAXA
17. Nebula
18. NELLY
19. Netherlands Metabolomics
20. OPPL Galaxy
21. Oqtans
22. Pathogen Portal
23. PopGenIE
24. Regulatory Genomics
25. RepeatExplorer
26. Stem Cell Discovery Engine
27. South Green
28. SymD
29. Wageningen University
30. Yeoman

Cistrome Analysis Pipeline

- Link:**
[Cistrome Analysis Pipeline](#)
- Domain/Purpose:**
ChIP-chip/seq and gene expression data
- Comments:**
The Cistrome Analysis Pipeline has the standard Galaxy tools, plus 29 additional ChIP-chip and ChIP-seq specific tools, including preliminary peak calling and correlation analyses, downstream genome feature association, gene expression analyses, and motif discovery.

Cistrome



A Galaxy Server
dedicated to
ChIP-* analysis

Gather and Reward!



Summary

Enable

- ✦ Galaxy sharing and publication
- ✦ toolshed
- ✦ provide opportunities to contribute: open mailing lists, conference talks, subgroups

Communicate

- ✦ social media: mailing lists, wiki, twitter

Reward

- ✦ publicize and praise contributions
- ✦ use social media

Galaxy



<http://galaxyproject.org>
<http://wiki.galaxyproject.org>
<http://usegalaxy.org>

<http://bitbucket.org/galaxy/galaxy-central>

clements@galaxyproject.org
jeremy.goecks@emory.edu



EMORY
UNIVERSITY

