What is Ginkgo?

- Single-cell CNV analysis
- Phylogenetic analysis
- Data visualization tools
- Share with collaborators
Outline

Experimental approaches

Computational approaches for CNV analysis

Demo

Hands-on
Single-cell sequencing

- Circulating tumor cells
- Clonal evolution in tumors
- Neuronal mosaicism
- Recombination/crossover in germ cells
Single-cell vs. bulk sequencing

**ONE GENOME FROM MANY**
Sequencing the genomes of single cells is similar to sequencing those from multiple cells — but errors are more likely.

- **Standard genome sequencing**
  - A sample containing thousands to millions of cells is isolated.
  - DNA is extracted from all the nuclei.
  - DNA is broken into fragments and then sequenced.
  - The sequences are assembled to give a common, 'consensus' sequence.

- **Single-cell sequencing**
  - A single cell is difficult to isolate, but it can be done mechanically or with an automated cell sorter.
  - The DNA is extracted and amplified, during which errors can creep in.
  - Amplified DNA is sequenced.
  - Errors introduced in earlier steps make sequence assembly difficult; the final sequence can have gaps.
Single-cell vs. bulk sequencing

**ONE GENOME FROM MANY**
Sequencing the genomes of single cells is similar to sequencing those from multiple cells — but errors are more likely.

> **Standard genome sequencing**

A sample containing thousands to millions of cells is isolated. DNA is extracted from all the nuclei. DNA is broken into fragments and then sequenced. The sequences are assembled to give a common, ‘consensus’ sequence.

> **Single-cell sequencing**

A single cell is difficult to isolate, but it can be done mechanically or with an automated cell sorter. The DNA is extracted and amplified, during which errors can creep in. Amplified DNA is sequenced. Errors introduced in earlier steps make sequence assembly difficult; the final sequence can have gaps.

Brian Owens, Nature News 2012
Whole Genome Amplification (WGA) methods

DOP-PCR (Degenerate Oligonucleotide Primed PCR)

MDA (Multiple Displacement Amplification)

MALBAC (Multiple Annealing and Looping Based Amplification Cycles)
Whole Genome Amplification (WGA) methods

DOP-PCR (Degenerate Oligonucleotide Primed PCR)

MDA (Multiple Displacement Amplification)

MALBAC (Multiple Annealing and Looping Based Amplification Cycles)

bit.ly/1EkCCOO

Interactive analysis and quality assessment of single-cell copy-number variations

Tyler Garvin, Robert Aboukhalil, Jude Kendall, Timour Baslan, Gurinder S Atwal, James Hicks, Michael Wigler, Michael Schatz
doi: http://dx.doi.org/10.1101/011346

Paul Blainey, FEMS Microbiol Rev. 2013
Copy-number variant analysis

Tyler Garvin
Copy-number variant analysis

Low coverage allows us to study copy-number variants
Copy-number variant analysis

Divide genome into “bins” with ~50 – 100 reads / bin
Copy-number variant analysis
Copy-number variant analysis
Copy-number variant analysis

Circular Binary Segmentation (CBS)
to reduce noise in data
Circular Binary Segmentation

- Recursively divide up the genome until identify segments that have probe distribution different than neighbors
Circular Binary Segmentation

- Recursively divide up the genome until identify segments that have probe distribution different than neighbors
Circular Binary Segmentation

- Recursively divide up the genome until identify segments that have probe distribution different than neighbors
Copy-number variant analysis

![Diagram of read counts normalized to a mean of 1]
Copy-number variant analysis
Analysis Pipeline

.GALX

Map reads

Remove duplicates

BAM to BED

.GALX

Bin reads

GC bias correction

Segment bins

Call integer copy number state

Interactive visualization

.GALX
Galaxy intro
Demo – Galaxy Pipeline
Galaxy is an open source, web-based platform for data intensive biomedical research. If you are new to Galaxy start here or consult our help resources.

**Galaxy 101**

**Start small**

The very first tutorial you need

---

The Galaxy Team is a part of the Center for Comparative Genomics and Bioinformatics at Penn State, and the Department of Biology and at Johns Hopkins University.

This instance of Galaxy is utilizing infrastructure generously provided by the iPlant Collaborative at the Texas Advanced Computing Center, with support from the National Science Foundation.
Galaxy is an open source, web-based platform for data intensive bio-research. If you are new to Galaxy start here or consult our help resources.

GXYcast 1
2015

PennState
Johns Hopkins University
TACC
iPlant Collaborative

The Galaxy Team is a part of the Center for Comparative Genomics and Bioinformatics at Penn State, and the Department of Biology and at Johns Hopkins University.

This instance of Galaxy is utilizing infrastructure generously provided by the iPlant Collaborative at the Texas Advanced Computing Center, with support from the National Science Foundation.
Create account

Email address:
Password:
Confirm password:
Public name:

Your public name is an identifier that will be used to generate addresses for your information you share publicly. Public names must be at least four characters in length and contain only lower-case letters, numbers, and the '-' character.

Subscribe to mailing list:
- See all Galaxy project mailing lists.

Submit

Please register only one account. The usegalaxy.org service is provided free of charge and has limited computational and data storage resources. Registration and usage of multiple accounts is tracked and such accounts are subject to termination and data deletion.

Terms and Conditions for use of this service
Galaxy is an open source, web-based platform for data intensive biomedical research. If you are new to Galaxy start here or consult our help resources.

Want help? Get answers.

**Biostars**

The Galaxy Team is a part of the Center for Comparative Genomics and Bioinformatics at Penn State, and the Department of Biology and at Johns Hopkins University.

This instance of Galaxy is utilizing infrastructure generously provided by the iPlant Collaborative at the Texas Advanced Computing Center, with support from the National Science Foundation.
Galaxy is an open source, web-based platform for data intensive biomedical research. If you are new to Galaxy start here or consult our help resources.

![GXYcast 1 2015](image)

**Tweets**

- **PyPI Recent Updates** @pypi_updates 6s
  planemo 0.6.0: Command-line utilities to assist in building tools for the Galaxy project (http://galaxyp... bit.ly/1OCxY1)
  Retweeted by Galaxy Project

- **Galaxy Project** @galaxypoint 9s
  usegalaxy Histories now support advanced searching and offer multiple other improvements.
  wiki.galaxypoint.org/Histories pic.twitter.com/URIYN5iIQ

- **Yannick Wurm** @yannickc 12 Mar
  Fun outcome #bioinformatics #job in

**This history is empty. You can load your own data or get data from an external source**

---

**Penn State**

**Johns Hopkins University**

**TACC**

**iPlant Collaborative**

---

The Galaxy Team is a part of the Center for Comparative Genomics and Bioinformatics at Penn State, and the Department of Biology and at Johns Hopkins University.

This instance of Galaxy is utilizing infrastructure generously provided by the iPlant Collaborative at the Texas Advanced Computing Center, with support from the National Science Foundation.
SRR054622
Read: SRR054622

Illumina Genome Analyzer sequencing; Single Cells from Breast Tumor T10

<table>
<thead>
<tr>
<th>View</th>
<th>XML</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Submitting Centre</th>
<th>Run Date</th>
<th>Platform</th>
<th>Model</th>
<th>Read Count</th>
<th>Base Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cold Spring Harbor Laboratory</td>
<td></td>
<td>ILLUMINA</td>
<td>Illumina Genome Analyzer</td>
<td>2,789,858</td>
<td>133,913,184</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Library Layout</th>
<th>Library Strategy</th>
<th>Library Source</th>
<th>Library Selection</th>
<th>Library Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>SINGLE</td>
<td>WGS</td>
<td>GENOMIC</td>
<td>RANDOM</td>
<td>5C.WGA.LIB</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Broker Name</th>
<th>NCBI</th>
</tr>
</thead>
</table>

### Navigation
- Read Files

This table contains the files for run SRR054622

**Download files**

<table>
<thead>
<tr>
<th>View</th>
<th>TEXT</th>
</tr>
</thead>
</table>

**Select columns**
### Read: SRR054622

**Illumina Genome Analyzer sequencing; Single Cells from Breast Tumor T10**

**View:** XML

<table>
<thead>
<tr>
<th>Submitting Centre</th>
<th>Run Date</th>
<th>Platform</th>
<th>Model</th>
<th>Read Count</th>
<th>Base Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cold Spring Harbor</td>
<td></td>
<td>ILLUMINA</td>
<td>Illumina Genome Analyzer</td>
<td>2,789,858</td>
<td>133,913,184</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Library Layout</th>
<th>Library Strategy</th>
<th>Library Source</th>
<th>Library Selection</th>
<th>Library Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>SINGLE</td>
<td>WGS</td>
<td>GENOMIC</td>
<td>RANDOM</td>
<td>SC.WGA.LIB</td>
</tr>
</tbody>
</table>

**Broker Name:** NCBI

**Navigation**

- Read Files

---

This table contains the files for run SRR054622

**Download files**

**View:** TEXT

**Select columns**

<table>
<thead>
<tr>
<th>Study accession</th>
<th>Secondary study accession</th>
<th>Sample accession</th>
<th>Secondary sample accession</th>
<th>Experiment accession</th>
<th>Run accession</th>
<th>Tax ID</th>
<th>Scientific name</th>
<th>Instrument model</th>
<th>Library layout</th>
<th>Fastq files (fp)</th>
<th>Fastq files (gala)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRR002535</td>
<td>SRR002535</td>
<td>SAMN0014736</td>
<td>SRR074087</td>
<td>SRR0210401</td>
<td>SRR054622</td>
<td>9605</td>
<td>Homo sapiens</td>
<td>Illumina Genome Analyzer</td>
<td>SINGLE</td>
<td>File 1</td>
<td>File 1</td>
</tr>
</tbody>
</table>
A job has been successfully added to the queue – resulting in the following dataset:


You can check the status of queued jobs and view the resulting data by refreshing the History pane. When the job has been run the status will change from 'running' to 'finished' if completed successfully or 'error' if problems were encountered.

You are now being redirected back to Galaxy.
Galaxy is an open source, web-based platform for data intensive biomedical research. If you are new to Galaxy start here or consult our help resources.

Try Galaxy on the Cloud

Now you can have a personal Galaxy within the infinite universe

Tweets

PyPi Recent Updates @pypi_updates 6s plasmon 0.6.0: Command-line utilities to assist in building tools for the Galaxy project (http://galaxyp... bit.ly/1OOCxYI) Retweeted by Galaxy Project

Galaxy Project @galaxyparticipant 9s now support advanced searching and offer multiple other improvements.

Yannick Wurm @yannick 12 Mar Fun on our bioinformatics job in

The Galaxy Team is a part of the Center for Comparative Genomics and Bioinformatics at Penn State, and the Department of Biology at Johns Hopkins University.

This instance of Galaxy is utilizing infrastructure generously provided by the iPlant Collaborative at the Texas Advanced Computing Centers, with support from the National Science Foundation.
Galaxy is an open source, web-based platform for data intensive biomedical research. If you are new to Galaxy, start here or consult our help resources.

Galaxy 101
Start small
The very first tutorial you need

The Galaxy Team is a part of the Center for Comparative Genomics and Bioinformatics at Penn State, and the Department of Biology at Johns Hopkins University.

This instance of Galaxy is utilizing infrastructure generously provided by the iPlant Collaborative at the Texas Advanced Computing Center, with support from the National Science Foundation.
Galaxy is an open source, web-based platform for data intensive biomedical research. If you are new to Galaxy start here or consult our help resources.
Galaxy is an open source, web-based platform for data intensive biomedical research. If you are new to Galaxy start here or consult our help resources.

Running Your Own
Understanding how Galaxy works
An in-depth tutorial

The Galaxy Team is a part of the Center for Comparative Genomics and Bioinformatics at Penn State, and the Department of Biology and at Johns Hopkins University.

This instance of Galaxy is utilizing infrastructure generously provided by the iPlant Collaborative at the Texas Advanced Computing Center, with support from the National Science Foundation.
Change data type

New Type: fastq

This will change the datatype of the existing dataset but not modify its contents. Use this if Galaxy has incorrectly guessed the type of your dataset.

Save
Change data type

New Type: fastq

This will change the data type of the existing dataset but not modify its contents. Use this if Galaxy has incorrectly guessed the type of your dataset.

Save
Edited attributes:

Name:

Info:

Annotation / Notes:
Add an annotation or notes to a dataset; annotations are available when a history is viewed.

Database/Build:

This will inspect the dataset and attempt to correct the above column values if they are not accurate.
Galaxy is an open source, web-based platform for conducting experimental and computational research. If you are new to Galaxy, please refer to our help resources.

kmer-SVM
Galaxy server for identifying predictive regulatory sequence features in NGS data

Galaxy / Beer Lab

Yvan Le Bras @Yvan2695
28m
With @monicasud we love the new #galaxyproject release! Using Klast from @KoRoriog on @bioingenuest Galaxy instance pic.twitter.com/NBMNhrujf7

Nate Corcor @natecorco
1h
It's official, #usegalaxy finally has versions! This is 15.03.1. RT @galaxypoject The March 2015 Release is here! Bit.ly/rgalaxy201503

TACC and iPlant Collaborative

The Galaxy Team is a part of the Center for Comparative Genomics and Bioinformatics at Penn State, and the Department of Biology and at Johns Hopkins University.

This instance of Galaxy is utilizing infrastructure generously provided by the iPlant Collaborative at the Texas Advanced Computing Center, with support from the National Science Foundation.
<table>
<thead>
<tr>
<th>Name</th>
<th>Annotation</th>
<th>Owner</th>
<th>Community Rating</th>
<th>Community Tags</th>
<th>Last Updated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ginkgo Preprocessing</td>
<td>RNA sequencing data analysis in a Transparent Reproducible and Automated PipeLINE.</td>
<td>robertaboukheili</td>
<td>★★★★★</td>
<td></td>
<td>~3 minutes ago</td>
</tr>
<tr>
<td>Galaxy Variant 101</td>
<td></td>
<td>galaxyproject</td>
<td>★★★★☆</td>
<td></td>
<td>Mar 06, 2015</td>
</tr>
<tr>
<td>Trim Reads Based on Quality (imported from uploaded file)</td>
<td></td>
<td>jdlim</td>
<td>★★★★☆</td>
<td></td>
<td>Feb 28, 2015</td>
</tr>
<tr>
<td>TFBières_Ayükil_012504840</td>
<td></td>
<td>ayalalol</td>
<td>★★★★☆</td>
<td></td>
<td>Feb 19, 2015</td>
</tr>
<tr>
<td>Bioinfo0525_lab4</td>
<td></td>
<td>jing-d7</td>
<td>★★★★☆</td>
<td>masseq</td>
<td>Feb 09, 2015</td>
</tr>
<tr>
<td>Lab_3_Bioinformatics</td>
<td></td>
<td>mussetu</td>
<td>★★★★★</td>
<td></td>
<td>Feb 07, 2015</td>
</tr>
<tr>
<td>Picard Quality stat on BAM'</td>
<td></td>
<td>umar</td>
<td>★★★★★</td>
<td></td>
<td>Jan 26, 2015</td>
</tr>
<tr>
<td>Quality stat on BAM</td>
<td></td>
<td>umar</td>
<td>★★★★★</td>
<td></td>
<td>Jan 26, 2015</td>
</tr>
<tr>
<td>Quality Statistics work flow</td>
<td></td>
<td>umar</td>
<td>★★★★★</td>
<td></td>
<td>Jan 26, 2015</td>
</tr>
<tr>
<td>GATK b37 unified genotyper (old)</td>
<td></td>
<td>umar</td>
<td>★★★★★</td>
<td></td>
<td>Jan 26, 2015</td>
</tr>
<tr>
<td>mpileup Hg19 MQ 40 BQ 30</td>
<td></td>
<td>umar</td>
<td>★★★★★</td>
<td></td>
<td>Jan 26, 2015</td>
</tr>
<tr>
<td>SAM filter Picard duplicate removal GATK recalibration mpileup Hg19 successful run</td>
<td></td>
<td>umar</td>
<td>★★★★★</td>
<td></td>
<td>Jan 26, 2015</td>
</tr>
<tr>
<td>Mpileup BED test</td>
<td></td>
<td>umar</td>
<td>★★★★★</td>
<td></td>
<td>Jan 26, 2015</td>
</tr>
</tbody>
</table>
### Your workflows

<table>
<thead>
<tr>
<th>Name</th>
<th># of Steps</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ginkgo Preprocessing</td>
<td></td>
</tr>
</tbody>
</table>

### Workflows shared with you by others

No workflows have been shared with you.

### Other options

- Configure your workflow menu
Running workflow "Ginkgo Preprocessing"

Step 1: Input dataset
  - type to filter

Step 2: Bowtie2 (version 0.2)

Step 3: Filter SAM or BAM, output SAM or BAM (version 1.1.1)

Step 4: rmdup (version 1.0.0)

Step 5: Convert from BAM to BED (version 0.1.0)

Send results to a new history

Run workflow
Running workflow "Ginkgo Preprocessing"

Step 1: Input dataset
  - type to filter

Step 2: Bowtie2 (version 0.2)

Step 3: Filter SAM or BAM, output SAM or BAM (version 1.1.1)

Step 4: rmdup (version 1.0.0)

Step 5: Convert from BAM to BED (version 0.1.0)

Send results to a new history

Run workflow
Successfully ran workflow "Gimli Preprocessing". The following datasets have been added to the queue:

2. Bowtie2 on data 1: aligned reads
3. Filter SAM or BAM, output SAM or BAM on data 1: bam
4. rmdup on data 3
5. rmdup on data 3 (as REDD)
Successfully ran workflow "Gimpgo Preprocessing". The following datasets have been added to the queue:

2. Bowtie2 on data 1: aligned reads
3. Filter SAM or BAM, output SAM or BAM on data 2: bam
4. rmqdup on data 3
5. rmqdup on data 3 (as BED)
Successfully ran workflow "Ginkgo Preprocessing". The following datasets have been added to the queue:

2. Bowtie2 on data 1: aligned reads
3. Filter SAM or BAM, output SAM or BAM on data 2: bam
4. rmdup on data 3
5. rmdup on data 3 (as BED)
Successfully ran workflow "Gismo Preprocessing". The following datasets have been added to the queue:

2. Bowtie2 on data 1: aligned reads
3. Filter SAM or BAM, output SAM or BAM on data 2: bam
4. rmdup on data 3
5. rmdup on data 3 (as BED)
Successfully ran workflow "Gimpgo Preprocessing". The following datasets have been added to the queue:

2. Bowtie2 on data 1: aligned reads
3. Filter SAM or BAM, output SAM or BAM on data 2: bam
4. rmdup on data 3
5. rmdup on data 3 (as BED)
Successfully ran workflow "GPKG Preprocessing". The following datasets have been added to the queue:

2. Bowtie2 on data 1: aligned reads
3. Filter SAM or BAM, output SAM or BAM on data 2: bam
4. rmdup on data 3
5. rmdup on data 3 (as RED)
Successfully ran workflow "Gimigo Preprocessing". The following datasets have been added to the queue:

2. Bowtie2 on data 1: aligned reads
3. Filter SAM or BAM, output SAM or BAM on data 2: bam
4. rmdup on data 3
5. rmdup on data 3 (as BED)
Successfully ran workflow "Gimpgo Preprocessing". The following datasets have been added to the queue:

2. Bowtie2 on data 1: aligned reads
3. Filter SAM or BAM, output SAM or BAM on data 2: bam
4. rmdup on data 3
5. rmdup on data 3 (as BED)
Successfully ran workflow "Gimogo Preprocessing". The following datasets have been added to the queue:

2. Bowtie2 on data 1: aligned reads
3. Filter SAM or BAM, output SAM or BAM on data 2: bam
4. rmdup on data 3
5. rmdup on data 3 (as BED)
Successfully ran workflow "Gatkgo Preprocessing". The following datasets have been added to the queue:

2. Bowtie2 on data 1: aligned reads
3. Filter SAM or BAM, output SAM or BAM on data 2: bam
4. rmDup on data 3
5. rmDup on data 3 (as BED)
<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>ch10</td>
<td>61988</td>
<td>62036</td>
<td>SRR054622.2038687</td>
<td>30</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>ch10</td>
<td>69093</td>
<td>69141</td>
<td>SRR054622.2024583</td>
<td>30</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>ch10</td>
<td>69098</td>
<td>69146</td>
<td>SRR054622.4815687</td>
<td>30</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>ch10</td>
<td>69099</td>
<td>69147</td>
<td>SRR054622.739395</td>
<td>30</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>ch10</td>
<td>72859</td>
<td>72907</td>
<td>SRR054622.1184883</td>
<td>30</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>ch10</td>
<td>92035</td>
<td>92083</td>
<td>SRR054622.2364537</td>
<td>30</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>ch10</td>
<td>101602</td>
<td>101650</td>
<td>SRR054622.2476074</td>
<td>30</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>ch10</td>
<td>108416</td>
<td>108464</td>
<td>SRR054622.5015749</td>
<td>30</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>ch10</td>
<td>108416</td>
<td>108464</td>
<td>SRR054622.2321097</td>
<td>25</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>ch10</td>
<td>110108</td>
<td>110156</td>
<td>SRR054622.1050882</td>
<td>34</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>ch10</td>
<td>112459</td>
<td>112507</td>
<td>SRR054622.1459050</td>
<td>30</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>ch10</td>
<td>112525</td>
<td>112573</td>
<td>SRR054622.651721</td>
<td>26</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>ch10</td>
<td>113773</td>
<td>113771</td>
<td>SRR054622.2235401</td>
<td>30</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>ch10</td>
<td>114074</td>
<td>114122</td>
<td>SRR054622.2190816</td>
<td>30</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>ch10</td>
<td>114078</td>
<td>114126</td>
<td>SRR054622.1006253</td>
<td>20</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>ch10</td>
<td>124019</td>
<td>124067</td>
<td>SRR054622.2081693</td>
<td>42</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>ch10</td>
<td>130995</td>
<td>131043</td>
<td>SRR054622.2135307</td>
<td>40</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>ch10</td>
<td>132896</td>
<td>132944</td>
<td>SRR054622.1964746</td>
<td>34</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>ch10</td>
<td>132905</td>
<td>132953</td>
<td>SRR054622.161640</td>
<td>35</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>ch10</td>
<td>132905</td>
<td>132953</td>
<td>SRR054622.161640</td>
<td>35</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>ch10</td>
<td>135289</td>
<td>135337</td>
<td>SRR054622.1757279</td>
<td>42</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>ch10</td>
<td>137966</td>
<td>138014</td>
<td>SRR054622.1142440</td>
<td>42</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>ch10</td>
<td>141733</td>
<td>141781</td>
<td>SRR054622.1035785</td>
<td>42</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>ch10</td>
<td>143964</td>
<td>144022</td>
<td>SRR054622.172454</td>
<td>42</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>ch10</td>
<td>146487</td>
<td>146545</td>
<td>SRR054622.897161</td>
<td>42</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>ch10</td>
<td>146689</td>
<td>146747</td>
<td>SRR054622.163231</td>
<td>26</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>ch10</td>
<td>166851</td>
<td>166909</td>
<td>SRR054622.1615104</td>
<td>42</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>ch10</td>
<td>168526</td>
<td>168574</td>
<td>SRR054622.1661134</td>
<td>42</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>ch10</td>
<td>168651</td>
<td>168699</td>
<td>SRR054622.2720453</td>
<td>30</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>ch10</td>
<td>172788</td>
<td>173326</td>
<td>SRR054622.2499973</td>
<td>42</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>ch10</td>
<td>176020</td>
<td>176068</td>
<td>SRR054622.1912131</td>
<td>42</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>ch10</td>
<td>179265</td>
<td>179313</td>
<td>SRR054622.1985607</td>
<td>40</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>ch10</td>
<td>182293</td>
<td>182247</td>
<td>SRR054622.1771700</td>
<td>42</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>ch10</td>
<td>182940</td>
<td>182988</td>
<td>SRR054622.1771700</td>
<td>42</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>ch10</td>
<td>182940</td>
<td>182988</td>
<td>SRR054622.531468</td>
<td>40</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>ch10</td>
<td>61988</td>
<td>62036</td>
<td>SRR054622.2038687</td>
<td>30 +</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>ch10</td>
<td>69093</td>
<td>69141</td>
<td>SRR054622.2024558</td>
<td>30 +</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>ch10</td>
<td>69098</td>
<td>69146</td>
<td>SRR054622.4815678</td>
<td>30 +</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>ch10</td>
<td>69099</td>
<td>69147</td>
<td>SRR054622.739539</td>
<td>30 +</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>ch10</td>
<td>72839</td>
<td>72907</td>
<td>SRR054622.1184883</td>
<td>30 -</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>ch10</td>
<td>92035</td>
<td>92083</td>
<td>SRR054622.2364537</td>
<td>30 -</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>ch10</td>
<td>101602</td>
<td>101650</td>
<td>SRR054622.2476079</td>
<td>30 -</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>ch10</td>
<td>108416</td>
<td>108464</td>
<td>SRR054622.5015774</td>
<td>35 -</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>ch10</td>
<td>108416</td>
<td>108464</td>
<td>SRR054622.2321097</td>
<td>25 +</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>ch10</td>
<td>110108</td>
<td>110156</td>
<td>SRR054622.1050882</td>
<td>34 -</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>ch10</td>
<td>112519</td>
<td>112567</td>
<td>SRR054622.1459050</td>
<td>30 +</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>ch10</td>
<td>112525</td>
<td>112573</td>
<td>SRR054622.6517212</td>
<td>26 +</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>ch10</td>
<td>113723</td>
<td>113771</td>
<td>SRR054622.2235401</td>
<td>30 -</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>ch10</td>
<td>114074</td>
<td>114122</td>
<td>SRR054622.2190816</td>
<td>30 -</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>ch10</td>
<td>114074</td>
<td>114126</td>
<td>SRR054622.1006253</td>
<td>20 +</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>ch10</td>
<td>124019</td>
<td>124067</td>
<td>SRR054622.2081693</td>
<td>42 +</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>ch10</td>
<td>130959</td>
<td>131007</td>
<td>SRR054622.2135307</td>
<td>40 +</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>ch10</td>
<td>132896</td>
<td>132944</td>
<td>SRR054622.1964764</td>
<td>34 +</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>ch10</td>
<td>132905</td>
<td>132953</td>
<td>SRR054622.161640</td>
<td>35 +</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>ch10</td>
<td>132910</td>
<td>132958</td>
<td>SRR054622.1169343</td>
<td>34 -</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>ch10</td>
<td>135289</td>
<td>135337</td>
<td>SRR054622.1757579</td>
<td>42 -</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>ch10</td>
<td>137966</td>
<td>138014</td>
<td>SRR054622.1144240</td>
<td>42 +</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>ch10</td>
<td>141733</td>
<td>141781</td>
<td>SRR054622.1035785</td>
<td>42 +</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>ch10</td>
<td>143294</td>
<td>143022</td>
<td>SRR054622.1722484</td>
<td>42 -</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>ch10</td>
<td>148479</td>
<td>148527</td>
<td>SRR054622.879161</td>
<td>42 -</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>ch10</td>
<td>148669</td>
<td>148717</td>
<td>SRR054622.1663231</td>
<td>26 +</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>ch10</td>
<td>168517</td>
<td>168565</td>
<td>SRR054622.1615104</td>
<td>42 -</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>ch10</td>
<td>168526</td>
<td>168574</td>
<td>SRR054622.1661134</td>
<td>42 -</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>ch10</td>
<td>168651</td>
<td>168699</td>
<td>SRR054622.2720453</td>
<td>30 +</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>ch10</td>
<td>171788</td>
<td>172836</td>
<td>SRR054622.2499973</td>
<td>42 -</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>ch10</td>
<td>176020</td>
<td>176068</td>
<td>SRR054622.1921231</td>
<td>42 +</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>ch10</td>
<td>179265</td>
<td>179313</td>
<td>SRR054622.1985607</td>
<td>40 -</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>ch10</td>
<td>182939</td>
<td>182987</td>
<td>SRR054622.1731700</td>
<td>40 -</td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>ch10</td>
<td>182940</td>
<td>182988</td>
<td>SRR054622.1479478</td>
<td>40 -</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>ch10</td>
<td>182940</td>
<td>183000</td>
<td>SRR054622.531468</td>
<td>40 +</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>ch10</td>
<td>182960</td>
<td>183000</td>
<td>SRR054622.685762</td>
<td>40 -</td>
<td></td>
</tr>
</tbody>
</table>

Download:

Click here to edit annotation

display in IGB View
display at Ensemble Current
display at Viewer main
display at UCSC main

1: EBI SRA: SRR054622
2: LF: ftp://ftp.ebi.ac.uk
   /ộl/ftp/SRR054/SRR054622/SRR054622.fasta
Analysis Pipeline

**.FASTQ**

- Map reads
- Remove duplicates
- BAM to BED

**Galaxy**

- Bin reads
- GC bias correction
- Segment bins
- Call integer copy number state
- Interactive visualization

**Ginkgo**
Hands-on with Ginkgo

**Ginkgo**: qb.cshl.edu/ginkgo

**Sample Data**: qb.cshl.edu/ginkgo/vizbi2015.tar