

- 1. How Binary Indexed Files are used for Data Visualization**
- 2. Building the Text Files for Track Hubs and additional Tracks**
- 3. Track Hub Resources**

A 2bit file is a binary indexed version of a FASTA file
(stores sequence ACGT as 00 01 11 10)

faToTwoBit input.fasta output.2bit



```
>scaffold_1
GTTGTAAATACTCTATTCTACAATAAAACCAAAGATAACTCTTTATCAG
TCATAGGTTGAATTGGCGTTGAAGTAAAACAAAAATACTGCTCAAAAGG
...
>scaffold_2
AGTTATGACAAACTATAAAAAAGTCGGTAGAGACAAAAGCGAAAAGGATCT
TCGTTTCGTGGACGAAGCGACCAAACTGAGCACAAGATAAATCCCGAATA
...
>scaffold_3
CATAAATTCATAAATCAATTCATGAAGAATAATTTTAGAAAAATGGTTCA
TAGAAAAATTTCCAGGAAGTTTGAAGTTGCTATAATGATTATTTCTCTTG
...
>scaffold_4
ATACTGCCCCATGCCTCATCTACTCCCTTCTTAACCTTTTCATCAACAACT
TGCTTAAGTTGGTCATCACTGACTGAGTGAGCCCAAACCTGGAATAGCTGT
...
```

The resulting indexed binary file
output.2bit can have data more easily
extracted and can also be viewed in the
UCSC Genome Browser.

A 2bit file is a binary indexed version of a FASTA file
(stores sequence ACGT as 00 01 11 10)

`faToTwoBit` and other standalone command-line
Utilities are downloadable tools available on Linux
and UNIX platforms.



The image shows the UCSC Genome Browser website. The header includes the University of California Santa Cruz logo, the UCSC logo, and the text "Genome Browser". Below the header is a navigation bar with links: Home, Genomes, Genome Browser, Tools, Mirrors, Downloads, My Data, Help, and About Us. The main content area features a large image of a DNA double helix on the left and a dropdown menu on the right. The dropdown menu is open, showing options: Genome Data, Source Code, Genome Browser Store, Utilities (highlighted in blue), FTP, and MySQL Access. An arrow points to the "Utilities" option. At the bottom of the page, there is a footer that reads "download data from the genome Browser database".

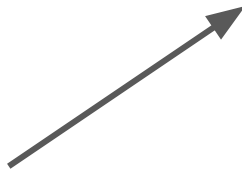
A 2bit file is a binary indexed version of a FASTA file
(stores sequence ACGT as 00 01 11 10)

`faToTwoBit` and other standalone command-line
Utilities are downloadable tools available on Linux
and UNIX platforms.

Utilities

The [utilities directory](#) offers downloads of pre-compiled standalone binaries for:

- [LiftOver](#) (which may also be accessed via the [web version](#)). The over.chain lift
- Other command-line utilities
- [Blat](#) ▶



A 2bit file is a binary indexed version of a FASTA file
(stores sequence ACGT as 00 01 11 10)

`faToTwoBit` and other standalone command-line
Utilities are downloadable tools available on Linux
and UNIX platforms.

<code>faFilter</code>	09-Jan-2018	15:55	4.1M
<code>faFilterN</code>	09-Jan-2018	15:55	4.3M
<code>faFrag</code>	09-Jan-2018	15:55	4.1M
<code>faNoise</code>	09-Jan-2018	15:55	4.1M
<code>faOneRecord</code>	09-Jan-2018	15:55	4.1M
<code>faPolyASize</code>	09-Jan-2018	15:55	4.1M
<code>faRandomize</code>	09-Jan-2018	15:55	4.1M
<code>faBc</code>	09-Jan-2018	15:55	4.1M
<code>faSize</code>	09-Jan-2018	15:55	4.1M
<code>faSomeRecords</code>	09-Jan-2018	15:55	4.1M
<code>faSplit</code>	09-Jan-2018	15:55	4.1M
<code>faToFastq</code>	09-Jan-2018	15:56	4.1M
<code>faToTab</code>	09-Jan-2018	15:56	4.1M
<code>faToTwoBit</code>	09-Jan-2018	15:56	4.2M
<code>faTrans</code>	09-Jan-2018	15:56	4.1M
<code>fastoStatsAndSubsample</code>	09-Jan-2018	15:55	4.2M
<code>fastoToFa</code>	09-Jan-2018	15:55	4.2M
<code>featureBits</code>	09-Jan-2018	15:57	16M
<code>fetchChromSizes</code>	09-Jan-2018	16:01	2.7K
<code>findMotif</code>	09-Jan-2018	15:56	4.2M

A 2bit file is a binary indexed version of a FASTA file
(stores sequence ACGT as 00 01 11 10)

faToTwoBit input.fasta output.2bit



```
>scaffold_1
GTTGTAAATACTCTATTCTACAATAAAACCAAAGATAACTCTTTATCAG
TCATAGGTTGAATTGGCGTTGAAGTAAAACAAAAATACTGCTCAAAAGG
...
>scaffold_2
AGTTATGACAAACTATAAAAAAGTCGGTAGAGACAAAAGCGAAAAGGATCT
TCGTTTCGTGGACGAAGCGACCAAACTGAGCACAAGATAAATCCCGAATA
...
>scaffold_3
CATAAATTCATAAATCAATTCATGAAGAATAATTTTAGAAAAATGGTTCA
TAGAAAAATTTCCAGGAAGTTTGAAGTTGCTATAATGATTATTTCTCTTG
...
>scaffold_4
ATACTGCCCCATGCCTCATCTACTCCCTTCTTAACCTTTTCATCAACAACT
TGCTTAAGTTGGTCATCACTGACTGAGTGAGCCAAAACCTGGAATAGCTGT
...
```

Host **output.2bit** at location that accepts
byte-range requests (like CyVerse).

A 2bit file is a binary indexed version of a FASTA file
(stores sequence ACGT as 00 01 11 10)

`asmBly.2bit` files exist for download via **Genome Data** link for all assemblies on the UCSC Browser



The image shows the UCSC Genome Browser website. The header includes the University of California Santa Cruz logo, the UCSC logo, and the text "Genome Browser". Below the header is a navigation bar with links: Home, Genomes, Genome Browser, Tools, Mirrors, Downloads, My Data, Help, and About Us. A dropdown menu is open under the "Genome Data" link, showing options: Genome Data, Source Code, Genome Browser Store, Utilities, FTP, and MySQL Access. A small text at the bottom of the dropdown reads "download data from the genome Browser database". The background of the page features a blue and green DNA double helix graphic.

A 2bit file is a binary indexed version of a FASTA file
(stores sequence ACGT as 00 01 11 10)

`asmBly.2bit` files exist for download via **Genome Data** link for all assemblies on the UCSC Browser

Sequence and Annotation Downloads

This page contains links to sequence and annotation data d

.....
.....
.....
Other downloads ▾

■ [Access the gbdb fileserver](#)

A 2bit file is a binary indexed version of a FASTA file
(stores sequence ACGT as 00 01 11 10)

Index of /gbdb

Name	Last modified	Size	Description
Parent Directory		-	
OtoGar1/	12-Dec-2016 11:38	-	
silWell/	20-Nov-2017 15:32	-	
allMis1/	20-Nov-2017 15:32	-	
amaVit1/	15-Oct-2013 09:30	-	
anaPla1/	15-Oct-2013 09:31	-	

/gbdb/allMis1

allMis1.2bit

Name	Last modified	Size	Description
Parent Directory		-	
allMis1.2bit	07-Jun-2013 13:33	543M	
...	20-Nov-2017 15:32		

<http://hgdownload.soe.ucsc.edu/gbdb/allMis1/allMis1.2bit>

Extracting data from a the 2bit for later use: **your2bit.chrom.sizes**

twoBitInfo *your.2bit* your2bit.chrom.sizes



```
chr1 249250621
chr2 243199373
chr3 198022430
...
chr16 90354753
chr17 81195210
....
chr21_gl000210_random 27682
chrUn_gl000231 27386
chrUn_gl000229 19913
chrM 16571
```

Using **your2bit.chrom.sizes** to make an artificial Cytoband Track (banding options *gneg*, *gpos25*,... programmatic steps could be done manually) for later: *yourCytoBandIdeo.bed* (BED4+1)

```
cat your2bit.chrom.sizes | awk '{print $1,0,$2,$1,"gneg"}' >  
yourCytoBandIdeo.bed
```



```
chr1 249250621  
chr2 243199373  
chr3 198022430  
...  
chr16 90354753  
chr17 81195210  
....  
chrUn_gl000231 27386  
chrUn_gl000229 19913  
chrM 16571
```



```
chr1 0 249250621 chr1 gneg  
chr2 0 243199373 chr2 gneg  
chr3 0 198022430 chr3 gneg  
...  
chr16 0 90354753 chr16 gneg  
chr17 0 81195210 chr17 gneg  
....  
chrUn_gl000231 0 27386 chrUn_gl000231 gneg  
chrUn_gl000229 0 19913 chrUn_gl000229 gneg  
chrM 0 16571 chrM gneg
```


Turning **yourCytoBandIdeo.bed** text file into a binary indexed version: *yourCytoBandIdeo.bigBed*

```
bedToBigBed -type=bed4 yourCytoBandIdeo.bed -as=cytoBand.as your2bit.chrom.sizes  
yourCytoBandIdeo.bigBed
```



```
chr1 0 249250621 chr1 gneg  
chr2 0 243199373 chr2 gneg  
chr3 0 198022430 chr3 gneg  
...  
chr16 0 90354753 chr16 gneg  
chr17 0 81195210 chr17 gneg  
....  
chrUn_gl000231 0 27386 chrUn_gl000231 gneg  
chrUn_gl000229 0 19913 chrUn_gl000229 gneg  
chrM 0 16571 chrM gneg
```

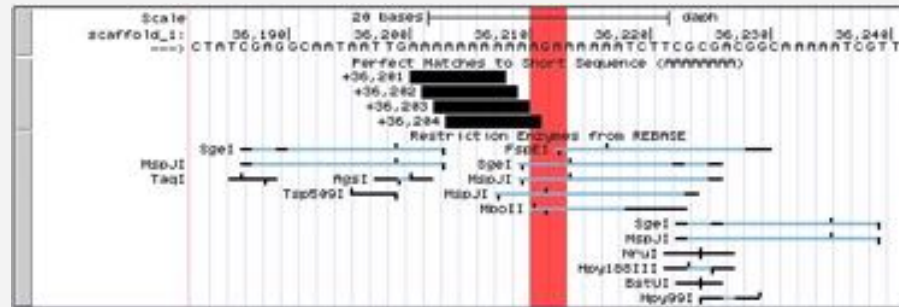
```
table cytoBandIdeo  
"cytoBandIdeo describes...."  
(  
  string chrom;      "Reference.... "  
  uint  chromStart; "Start..."  
  uint  chromEnd;   "End..."  
  string name;      "Name of item"  
  string  gieStain; "Giemsa stain..".
```

Host **file.bigBed**
at location that
accepts
byte-range
requests (like
CyVerse).

Building the Hub Text (.txt) Files

```
>scaffold_1
GTTGTAAATACTCTATTCTACAATAAAACCAA
TCATAGGTTGAATTGGCGTTGAAGTAAAAACAAA
...
>scaffold_2
AGTTATGACAAACTATAAAAAAGTCGGTAGAGACAAAAG
TCGTTCTGTGGACGAAGCGACCAAAAAGTGAAGTGAAGT
...
>scaffold_3
CATAAATTCATAAATCAATTCATGAAGAATAATT
TAGAAAATTTCCAGGAAGTTTGAAGTTGCTAT
...
```

hubDirectory
|__hub.txt
|__genomes.txt
|__hg19
|__trackDb.txt



Track Hubs

Track Hubs are text files (**hub.txt**, **genomes.txt**, **trackDb.txt**) that describe and point to the location of binary indexed data files (**twoBitPath**, **bigDataUrl**)

hubDirectory

|

|_hub.txt

|_genomes.txt

|_genomeDirectory1

|_trackDb.txt

Informs location of genome 2bit file(s): **twoBitPath**

Informs location of track data file(s): **bigDataUrl**

Track Hubs

Track Hubs are text files (**hub.txt**, **genomes.txt**, **trackDb.txt**) that describe and point to the location of binary indexed data files (**twoBitPath**, **bigDataUrl**)

hub.txt

hub *MyHubsNameWithoutSpaces*
shortLabel *My Hub's Name*
longLabel *Longer label about my hub.*
email *myEmail@address*
genomesFile *http://location_of/genomes.txt*



genomes.txt

genome *yourGenome*
trackDb *http://location_of/trackDb.txt*
twoBitPath *http://location_of/output.2bit*
...

trackDb.txt

track *cytoBandIdeo*
type *bigBed*
longLabel *Chromosome ideogram with cytogenetic bands*
shortLabel *cytoBandIdeo*
bigDataUrl *http://location_of/yourCytoBandIdeo.bigBed*

track *bam1*
type *bam*
shortLabel *BAM example*
longLabel *A BAM file for my genome data*
visibility *dense*
bigDataUrl *http://location_of/file.bam*

...

Each new track gets a new stanza

bigDataUrl lines point to location of binary indexed data
More trackDb parameters can be defined too (color, ect.)

Track Hubs

Track Hubs are text files (**hub.txt**, **genomes.txt**, **trackDb.txt**) that describe and point to the location of binary indexed data files (**twoBitPath**, **bigDataUrl**)

twoBitPath http://location_of/output.2bit

genomes.txt shares where to find the 2bit (and what to call new genome)

genome *yourGenome*
trackDb http://location_of/trackDb.txt
scientificName *Your Genome*
description *Feb. 2017 Assembly*
organism *Your organism*
defaultPos *chr1:1000000-2000000*
groups *groups.txt*
htmlPath <http://yourGenome/description.html>
twoBitPath http://location_of/output.2bit

Genomic Window of Sequence viewed at UCSC

Three tracks are generated from the 2bit on the fly:

- Base Position
- Restriction Enzymes
- Short Match Track (30 bases)

Track Hubs

Track Hubs are text files (**hub.txt**, **genomes.txt**, **trackDb.txt**) that describe and point to the location of binary indexed data files (**twoBitPath**, **bigDataUrl**)

groups *groups.txt*



genomes.txt shares where to groups.txt to label the track groups

Very small file labels and priority ordering

genome *yourGenome*
trackDb *http://location_of/trackDb.txt*
scientificName *Your Genome*
description *Feb. 2017 Assembly*
organism *Your organism*
defaultPos *chr1:1000000-2000000*
groups *groups.txt*
htmlPath *http://yourGenome/description.html*
twoBitPath *http://location_of/output.2bit*

name map
label Mapping
priority 1
defaultIsClosed 0

name genes
label Genes
priority 2
defaultIsClosed 0

Viewing the 2bit at UCSC

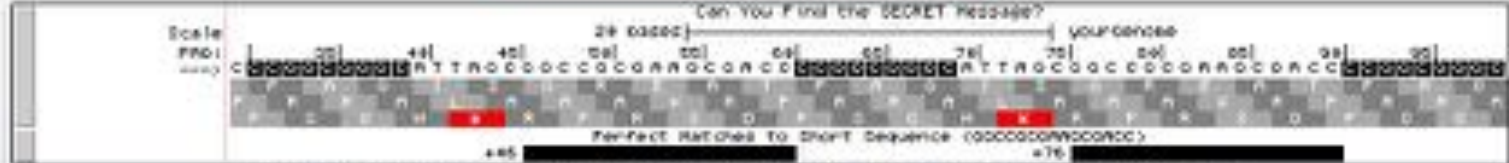
Browser requests *AGCTs* for only the genomic window currently viewing

UCSC Genome Browser on Your organism Feb. 2017 Assembly Assembly (yourGenome)

move <<< << < > >> >>> zoom in 1.5x 3x 10x base zoom out 1.5x 3x 10x 100x

PAG:30-99 70 bp. go

chr1 (chr1)



Click on a feature for details. Click or drag in the base position track to zoom in. Click side bars for track options. Drag side bars or labels up or down to reorder tracks. Drag tracks left or right to new position. Press "?" for keyboard shortcuts.

default tracks default order hide all add custom tracks track hubs configure multi-region reverse resize refresh

Use drop-down controls below and press refresh to alter tracks displayed. Tracks with lots of items will automatically be displayed in more compact modes.

collapse all **Mapping** expand all refresh

Base Position Restr Enzymes Short Match

full hide pack

Viewing the 2bit at UCSC

Browser requests *AGCTs* for only the genomic window currently viewing

The screenshot shows the UCSC Genome Browser interface. At the top, there is a search bar with the text "PAG:30-99 70 bp." and a "go" button. Below the search bar is a track labeled "PAG:30-99" with a red bar indicating the current genomic window. The main track displays the DNA sequence "ATTAGCGCCCGGACCCATTAGCGCCCGGACCC" with positions 30, 40, 50, 60, 70, and 80 marked. A search for the sequence "GGCCCGGACCC" is shown, with two perfect matches highlighted in red at positions 45 and 75. The text "Can You Find the SECRET Message?" is visible above the sequence. Below the sequence, the text "Perfect matches to Short Sequence (GGCCCGGACCC)" is displayed, with positions 45 and 75 marked.

Click on a feature for details. Click or drag in the base position track to zoom in. Click side bars for track options. Drag side bars or labels up or down to reorder tracks. Drag tracks left or right to new position. Press "?" for keyboard shortcuts.

QuickStart Guide to Assembly Hubs

<http://genome.ucsc.edu/goldenPath/help/hubQuickStartAssembly.html>

Quick Start Guide to Assembly Hubs

Assembly Hubs allow researchers to create Track Data Hubs on assemblies that are not in the UCSC Browser. By including the underlying reference sequence in UCSC [twoBit](#) format, as well as data tracks, researchers can browse and annotate any genome. For more information please refer to the [Assembly Hub Wiki](#). Below is also a section about starting [GBiB Assembly Hubs](#).

STEP 1: In a publicly-accessible directory, copy this *Arabidopsis thaliana* plant assembly hub, which includes an araTha1.2bit file, using the following wget command:

```
wget -r --no-parent --reject "index.html*" -nH --cut-dirs=3  
http://genome.ucsc.edu/goldenPath/help/examples/hubExamples/hubAssembly/plantAraTha1/
```

Adding Annotations

simple as adding text-based custom tracks



track name=ex1 description="Blue example BED3 track" color=0,0,255

chr21 33030000 33040000

chr21 33050000 33060000

...

Adding tracks inside Track Hubs

create binary indexed versions of text files (`bedToBigBed`)

`bedToBigBed in.bed chrom.sizes out.bb`

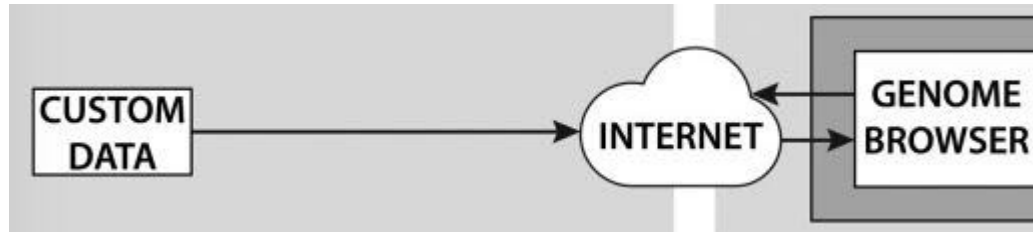
`chr1 33050000 33060000`
`chr1 33050000 33060000`
...
`chr21 33030000 33040000`
`chr21 33050000 33060000`
...
`chr22 33030000 33040000`
`chr22 33050000 33060000`
...

`chr1 248956422`
...
`chr21 46709983`
`chr22 50818468`
...

The resulting binary file **out.bb** (bigBed) can be hosted online:

[bigDataUrl=http://path.lab.edu/to/out.bb](http://path.lab.edu/to/out.bb)

The serving of files requires byte-range requests, which allows only portions of file to transfer. Some locations, like Dropbox, prevent such requests as people can then watch videos from their servers.



bigDataUrl <http://path.lab.edu/to/out.bb>

The **trackDb.txt** shares descriptions and where to find binary indexed tracks

View Data Tracks at UCSC

```
track ex1bb
type bigBed
shortLabel ex1
longLabel Blue example BED3 track
color 0,0,255
bigDataUrl http://path.lab.edu/to/out.bb
```

```
track track2
bigDataUrl http://path.lab.edu/to/out2.bb
```

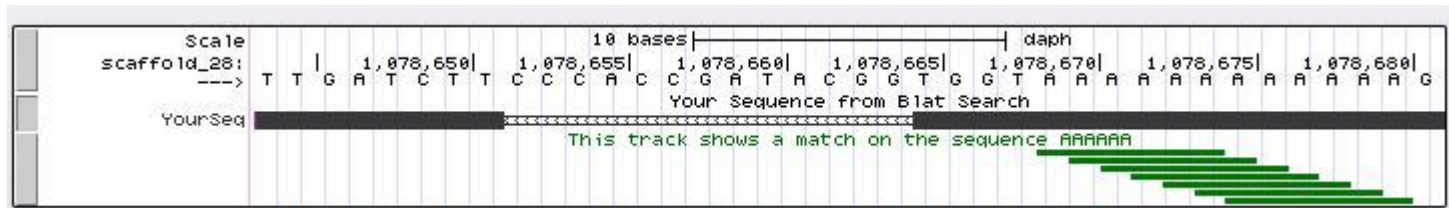
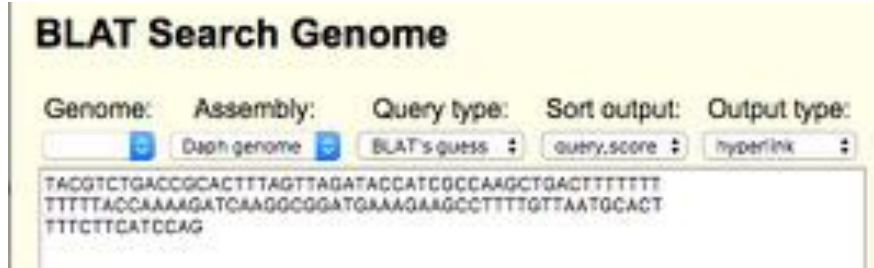
...

A wide variety of track binary indexed formats exist:

- BAM, CRAM, VCF
- bigGenePred, bigBarChart
- bigPsl, bigChain, bigMaf,
- bigNarrowPeak, halSnake
- bigBed, bigWig

Additional Tools Available on Assembly Hubs

- BLAT search sequence or amino-acid matched



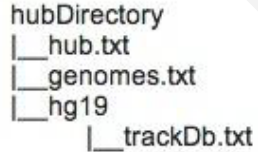
Examples of Hubs and Resources

Genbank Assembly Hubs

<http://genome-test.soe.ucsc.edu/gbdb/hubs/genbank/>

Ensembl Track Hub Registry

<https://trackhubregistry.org/>



hubDirectory
|__hub.txt
|__genomes.txt
|__hg19
|__trackDb.txt

UCSC Public Hubs

<http://genome.ucsc.edu/cgi-bin/hgHubConnect>

G-OnRamp

<http://gonramp.wustl.edu/>

CyVerse Discovery Environment byte-range supported data hosting

<https://de.cyverse.org/de/>

Existing Working Assembly Hub Templates

Genbank Assembly Hubs

<http://genome-test.soe.ucsc.edu/gbdb/hubs/genbank/>

Genomes Genome Browser Tools Mirrors Downloads My Data Help About Us

Genbank assembly hubs

Collection of hubs for Genbank genome assemblies.
These assemblies use the **UCSC naming patterns** on chromosomes.

Use these links to go to the index for that subset of assemblies:

species subset ▲▼	number of species ▲▼	number of assemblies ▲▼	total contig count ▲▼	total nucleotide count ▲▼	average contig size ▲▼	average assembly size ▲▼
other/synthetic assemblies	3	3	94	2,835,004	30,159	945,001
vertebrate other	156	172	18,548,615	193,684,015,605	10,441	1,126,069,858
vertebrate mammalian	118	204	30,643,657	498,264,459,566	16,259	2,442,472,841
plant	190	269	34,577,423	145,341,422,954	4,203	540,302,687
protozoa	282	338	3,939,128	16,816,724,183	4,269	49,753,621
invertebrate	392	492	32,264,511	170,439,035,382	5,282	346,420,803
fungi	1,106	1,215	4,143,097	38,677,096,556	9,335	31,833,001
archaea	688	742	57,569	2,010,246,046	34,918	2,709,226
bacteria	34,009	58,661	8,397,216	234,147,691,500	27,883	3,991,539
totals:	36,944	62,096	132,571,310	1,299,383,526,796	9,801	20,925,398

Collection of hubs for Genbank genome assemblies.
These assemblies use the **NCBI accession naming patterns** on chromosomes.

Use these links to go to the index for that subset of assemblies:

species subset ▲▼	number of species ▲▼	number of assemblies ▲▼	total contig count ▲▼	total nucleotide count ▲▼	average contig size ▲▼	average assembly size ▲▼
other/synthetic assemblies	3	3	94	2,835,004	30,159	945,001
vertebrate other	156	172	18,548,615	193,684,015,605	10,441	1,126,069,858
vertebrate mammalian	118	204	30,643,657	498,264,459,566	16,259	2,442,472,841

Existing Working Assembly Hub Templates

Genbank Assembly Hubs

<http://genome-test.soe.ucsc.edu/gbdb/hubs/genbank/>

Scroll Down to see assemblies... Click common name to load hub.

The "Taxon ID" link will go to the Entrez taxonomy for that ID.
The "common name" link will go to the UCSC genome browser for that assembly.
The "biosample" link will go to the Entrez biosample for that ID.
The "accession" link will go to the Entrez assembly for that ID.
The "assembly" link will go to the NCBI Genbank FTP source directory.

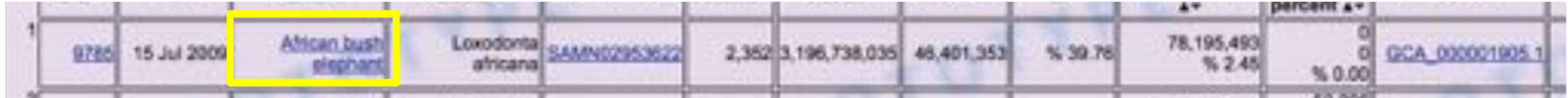
Taxon ID ▲▼	date ▲▼	common name ▼	scientific name ▲▼	biosample ▲▼	contig count ▲▼	genome size ▲▼	N50 size ▲▼	GC percent ▲▼	unknown bases count/percent ▲▼	gene count bases percent ▲▼	accession ▲▼
1 9785	15 Jul 2009	African bush elephant	<i>Loxodonta africana</i>	SAMN02953522	2,352	3,196,738,035	46,401,353	% 39.78	78,195,493 % 2.45	0 % 0.00	GCA_000001905.1
2 9646	15 Dec 2009	Alluropoda melanoleuca	<i>Alluropoda melanoleuca</i>	SAMN00008160	81,466	2,299,492,210	1,281,781	% 40.62	54,198,184 % 2.36	52,206 473,410,926 % 20.59	GCA_000004335.1
3 37293	16 Mar 2015	Aotus nancymae	<i>Aotus nancymae</i>	SAMN03121886	29,222	2,926,565,220	8,280,397	% 35.21	391,667,293 % 13.36	0 % 0.00	GCA_000052055.1
4 9555	05 Jun 2012	baboon	<i>Papio anubis</i>	SAMN02961400	63,250	2,948,380,710	139,646,187	% 40.20	55,130,419 % 1.87	0 % 0.00	GCA_000264685.1
5 118797	31 Jul 2013	baji	<i>Lipotes vexillifer</i>	SAMN02953859	30,712	2,429,195,737	2,419,148	% 40.75	32,195,939 % 1.33	0 % 0.00	GCA_000442215.1

Existing Working Assembly Hub Templates

Genbank Assembly Hubs

<http://genome-test.soe.ucsc.edu/gbdb/hubs/genbank/>

1. Right-click and copy and paste a link to explore the hub's architecture.



									percent		
1	9785	15 Jul 2009	African bush elephant	Loxodonta africana	SAMN02953622	2,352,319,738,035	46,401,353	% 39.76	78,195,493 % 2.45	0 0 % 0.00	GCA_000001905.1

2. Visit just the `hub.ncbi.txt` and then the `genomes.ncbi.txt` and then look at a specific `GCA...trackDb.ncbi.txt`

Track Hub Resources

Ensembl Track Hub Registry

<https://trackhubregistry.org/>

The Track Hub Registry

A global centralised collection of publicly accessible track hubs

The goal of the Track Hub Registry is to allow third parties to advertise [track hubs](#), and to make it easier for researchers around the world to discover and use track hubs containing different types of genomic research data.



Track Hub Resources

Ensembl Track Hub Registry -*Link out from Gramene*

<https://trackhubregistry.org/>



The screenshot shows the Gramene website interface. At the top left is the Gramene logo. To its right is a search bar with the placeholder text "Search for genes, species, pathways, ontology terms, domains...". Below the search bar, there are six resource cards arranged in a 3x2 grid:

- BLAST**: Query our genomes with a DNA or protein sequence. The icon shows a bar chart.
- Gramene Mart**: An advanced query interface powered by BioMart. The icon shows the BioMart logo.
- Track Hub Registry**: A global centralised collection of publicly accessible track hubs. The icon shows the Track Hub Registry logo.
- Outreach and Training**: Educational resources and webinars. The icon shows three people icons.
- Bulk Downloads**: FTP download of our data. The icon shows a downward arrow.
- Archive**: Legacy tools and data (markers, Cyc pathways, etc). The icon shows the GRAMENE logo.

Track Hub Resources

UCSC Public Hubs *moderated collection of track hubs*

<http://genome.ucsc.edu/cgi-bin/hgHubConnect>

The screenshot shows the UCSC Public Hubs search interface. At the top, there are two tabs: "Public Hubs" (selected) and "My Hubs". Below the tabs, there is a search bar with the text "Enter search terms to find in public track hub description pages:" and the search term "mouse epigenomics" entered. A "Search Public Hubs" button is next to the search bar. Below the search bar, there is a "Filter hubs by assembly:" label and an empty input field. A "Displayed list restricted by search terms: mouse epigenomics" label is followed by a "Show All Hubs" button. A paragraph of text explains that right-clicking on an assembly or track line opens it in a new window and that clicking "Connect" redirects to the gateway page of the selected hub's default assembly. Below this text is a table with four columns: "Display", "Hub Name", "Description", and "Assemblies". The first row shows a "Connect" button, the hub name "DNA Methylation", the description "Hundreds of analyzed methylomes from bisulfite sequencing data", and the assemblies "[+] hg38, hg19, hg18, mm9, mm10, panTro2...". Below the first row, there is a "Search details ..." link and a "Hub Description: ..." text. The second row shows a "Connect" button, the hub name "ENCODE Analysis Hub", the description "ENCODE Integrative Analysis Data Hub", and the assembly "hg19". Below the second row, there is a "Search details ..." link and a "Hub Description: ..." text.

Public Hubs My Hubs

Enter search terms to find in public track hub description pages:
mouse epigenomics

Filter hubs by assembly: Search Public Hubs

Displayed list **restricted by search terms:** mouse epigenomics Show All Hubs

When exploring the detailed search results for a hub, you may right-click on an assembly or track line to open it in a new window.
Clicking Connect redirects to the gateway page of the selected hub's default assembly.

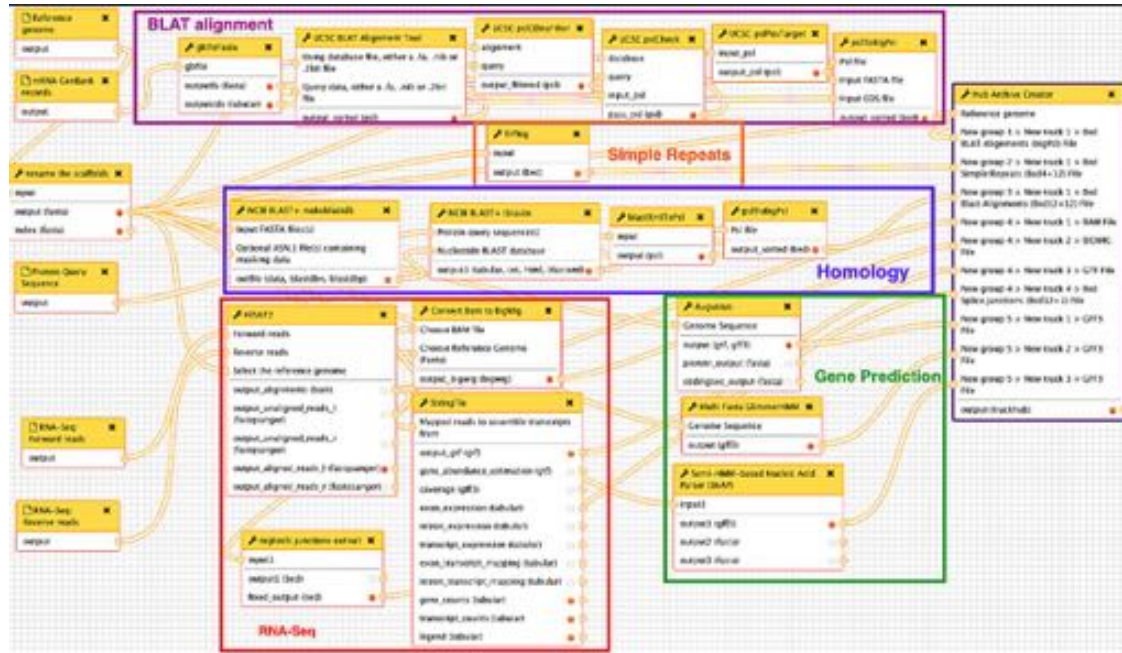
Display	Hub Name	Description	Assemblies
Connect	DNA Methylation	Hundreds of analyzed methylomes from bisulfite sequencing data	[+] hg38, hg19, hg18, mm9, mm10, panTro2...
Search details ...			
Hub Description: ... This session shows a collection of tracks from multiple blood cells. Mouse: Reprogramming: This session shows tracks from vari...			
4 Matching Assemblies			
Connect	ENCODE Analysis Hub	ENCODE Integrative Analysis Data Hub	hg19
Search details ...			
Hub Description: ... d TALI genes. We tested a subset of these predicted enhancers in both Mouse and Fish transgenic models examples in Figure 6,			

Track Hub Resources

G-OnRamp

Galaxy workflow turning data like RNA-Seq into Assembly Hubs

<http://gonramp.wustl.edu/>



Track Hub Resources

CyVerse Discovery Environment Storage Solution

CyVerse `iCommands` (`rysnc` becomes `irsync`) allows transfer of 2-100GB files and recursive hub directory structures with commands like `irsync`:

```
$ irsync -r local_directory_name i:data_store_directory_name
```

hubDirectory

```
|  
|_hub.txt  
|_genomes.txt  
|_genomeDirectory1  
|_trackDb.txt
```

Move all your local files (BAMs, ect) with the single `irsync` command.

Then make the files publicly accessible with the `ichmod` command to view.

<https://de.cyverse.org/anon-files/>

```
$ ichmod read anonymous data_store_directory_name
```



Thank You!



UNIVERSITY OF CALIFORNIA
SANTA CRUZ | Genomics
Institute

