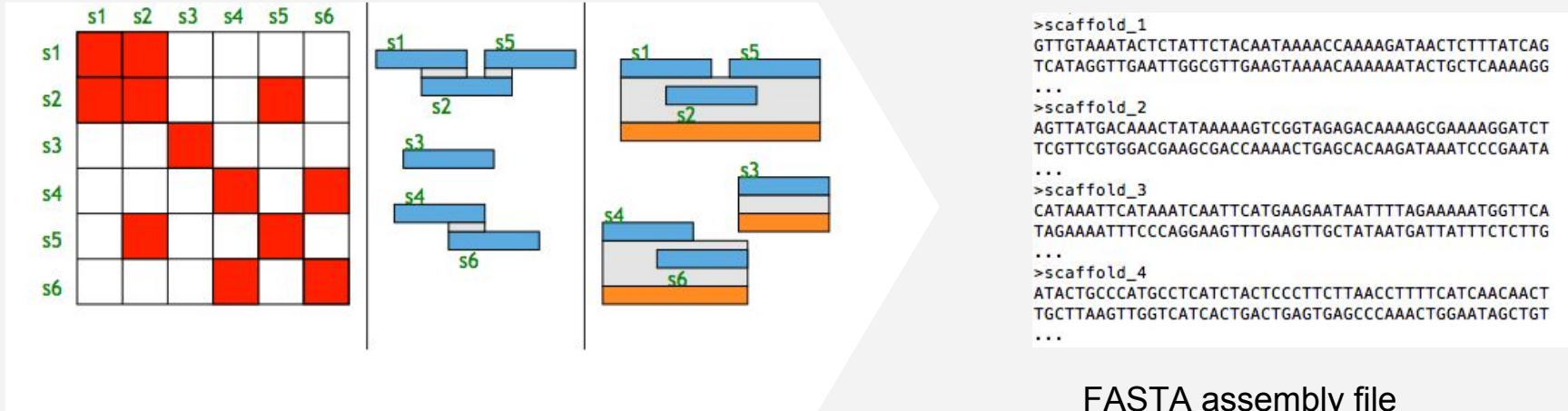
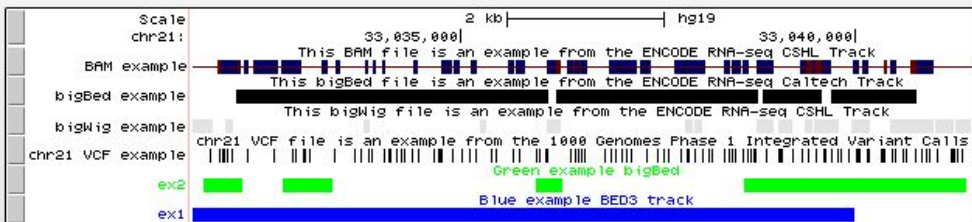


# A new genome assembly is available.



See it with **Assembly Hubs**.

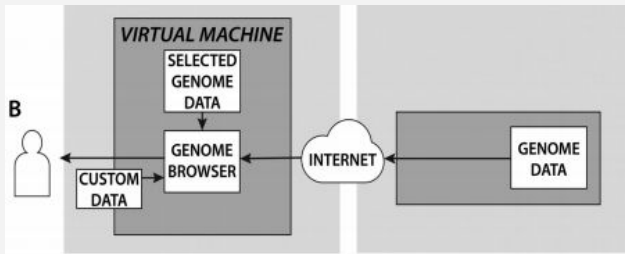
# 1. Custom Annotations in the Browser



# 2. Building An Assembly Hub

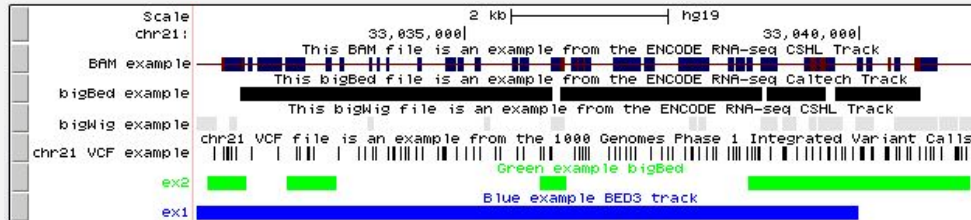
```
hubDirectory  
├─_hub.txt  
├─_genomes.txt  
├─_hg19  
└─_trackDb.txt
```

# 3. Existing Hubs and Solutions to Hosting Files



# 1. Custom Annotations in the Browser

Custom Tracks are Available on All Assemblies (Hosted by UCSC or Hosted Remotely)



# Custom Annotations in the Browser

## Custom Tracks (Text/Binary)



Line-oriented files.  
Defined by a track line  
Followed by data lines

Lines be pasted  
Or uploaded as text

Data lines can be hosted remotely.

Assemblies Hosted by UCSC

## Track Hubs



Track hubs are web-accessible  
directories of genomic data.

Three text files define how to  
display data:

hub.txt  
genomes.txt  
trackDb.txt

All data files must be hosted  
remotely.

## Assembly Hubs



Assembly Hubs are basically  
advanced Track Hubs.

Tracks display on a remote  
genome (2bit format).

All data files must be hosted  
remotely.

Assembly Hosted Remotely

# Custom Annotations in the Browser

Custom Tracks (Text/Binary)

Track Hubs

Assembly Hubs

The screenshot shows the 'Add Custom Tracks' page in the UCSC Genome Browser. At the top is a navigation bar with links for Home, Genomes, Genome Browser, Tools, Mirrors, Downloads, My Data, and Help. Below the navigation bar, the page title is 'Add Custom Tracks'. There are three dropdown menus: 'clade' set to 'Mammal', 'genome' set to 'Human', and 'assembly' set to 'Feb. 2009 (GRCh37)'. A text area contains instructions on how to add custom tracks, listing supported formats like bigBed, bigChain, bigGenePred, bigMaf, bigPsl, bigWig, BAM, VCF, BED, BED detail, bedGraph, broadPeak, CRAM, GFF, GTF, MAF, narrowPeak, Personal Genome SNP, PSL, and WIG. Below the text area are two input options: 'Paste URLs or data:' with a large text box, and 'Or upload:' with a 'Choose File' button and 'no file selected' text. There are 'Submit' and 'Clear' buttons on the right side of the page.

Home Genomes Genome Browser Tools Mirrors Downloads My Data Help

## Add Custom Tracks

clade Mammal genome Human assembly Feb. 2009 (GRCh37)

Display your own data as custom annotation tracks in the browser. Data must be in [bigBed](#), [bigChain](#), [bigGenePred](#), [bigMaf](#), [bigPsl](#), [bigWig](#), [BAM](#), [VCF](#), [BED](#), [BED detail](#), [bedGraph](#), [broadPeak](#), [CRAM](#), [GFF](#), [GTF](#), [MAF](#), [narrowPeak](#), [Personal Genome SNP](#), [PSL](#), or [WIG](#) formats. To configure the display, set [track](#) and [browser](#) line attributes as described in the [User's Guide](#). Data in the bigBed, bigWig, bigGenePred, BAM and VCF formats can be provided via only a URL or embedded in a track line in the box below. Examples are [here](#).

Paste URLs or data: Or upload: Choose File no file selected Submit

Clear

# Custom Annotations in the Browser

Custom Tracks (Text/Binary)

Track Hubs

Assembly Hubs



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

*chr21 33030000 33040000*

*chr21 33050000 33060000*

...

# Custom Annotations in the Browser

Custom Tracks (Text/Binary)

Track Hubs

Assembly Hubs



track name=ex2 description="Green example bigBed" color=0,255,0, **type=bigBed**  
***bigDataUrl=http://genome.ucsc.edu/goldenPath/help/examples/bigBedExample.bb***

# Custom Annotations in the Browser

Custom Tracks (Text/Binary)

Track Hubs

Assembly Hubs

bedToBigBed    in.bed    chrom.sizes    out.bb

scaffold_1	3890	3896
scaffold_1	4063	4069
scaffold_1	4236	4242
scaffold_1	4755	4761
scaffold_1	5976	5982
scaffold_1	6022	6028
scaffold_1	6023	6029
scaffold_1	6398	6404
scaffold_1	6399	6405
scaffold_1	26373	26379
scaffold_1	26374	26380
scaffold_1	29767	29773
scaffold_1	29790	29796
scaffold_1	29917	29923
scaffold_1	29918	29924
scaffold_1	30872	30878
scaffold_1	32125	32131
scaffold_1	32218	32224
scaffold_1	32219	32225
scaffold_1	32220	32226

scaffold_1	4193030
scaffold_3	3777634
scaffold_2	3740169
scaffold_4	3075709
scaffold_5	2511979
scaffold_6	2406117
scaffold_8	2335496
scaffold_7	2324446
scaffold_9	2251199
scaffold_12	2218424

The resulting binary file **out.bb** can be hosted at your institution:

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

Serving 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.



# Custom Annotations in the Browser

Custom Tracks (Text/Binary)

Track Hubs

Assembly Hubs

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

The screenshot shows the 'Add Custom Tracks' interface. The navigation bar at the top includes 'Genomes', 'Genome Browser', 'Tools', 'Mirrors', 'Downloads', and 'My Data'. The 'Add Custom Tracks' section has three dropdown menus: 'clade' (Mammal), 'genome' (Human), and 'assembly'. Below these is a text area with instructions on how to use various data formats like bigBed, bigChain, bigGenePred, bigMaf, bigPsl, bigWig, BAM, VCF, bedGraph, broadPeak, CRAM, GFF, GTF, MAF, narrowPeak, and Per. At the bottom, there are input fields for 'Paste URLs or data:' and 'Or upload:' with a 'Choose File' button, and 'Submit' and 'Clear' buttons.

# Custom Annotations in the Browser

Custom Tracks (Text/Binary)

Track Hubs

Assembly Hubs

Track hubs are web-accessible directories of genomic data.

Track Data Hubs

Track data hubs are collections of external tracks that can be imported into the UCSC Genome Browser. Tracks show up under the hub's own blue label bar on the main browser page, as well as on the configure page. To import a public hub click its "Connect" button below.

**NOTE: Because Track Hubs are created and maintained by external sources, UCSC is not responsible for their content.**

Public Hubs My Hubs

URL:  Add Hub

Paste in: <http://path.lab.edu/to/hubDirectory/hub.txt>

# Custom Annotations in the Browser

Custom Tracks (Text/Binary)

Track Hubs

Assembly Hubs

Track hubs are web-accessible directories of genomic data.

<http://path.lab.edu/to/>

**hubDirectory**

|

|\_hub.txt

|\_genomes.txt

|\_hg19

|\_trackDb.txt

hub.txt can be as short as four lines, it points to genomes.txt

genomes.txt can be as short as two lines and points to trackDb.txt

The trackDb.txt file is typically much larger and shares how to display tracks and where to find data files.

bigDataUrl=<http://path.lab.edu/to/file1.bb>

bigDataUrl=<http://path.lab.edu/to/file2.bb>

bigDataUrl=<http://path.lab.edu/to/file3.bb>

...

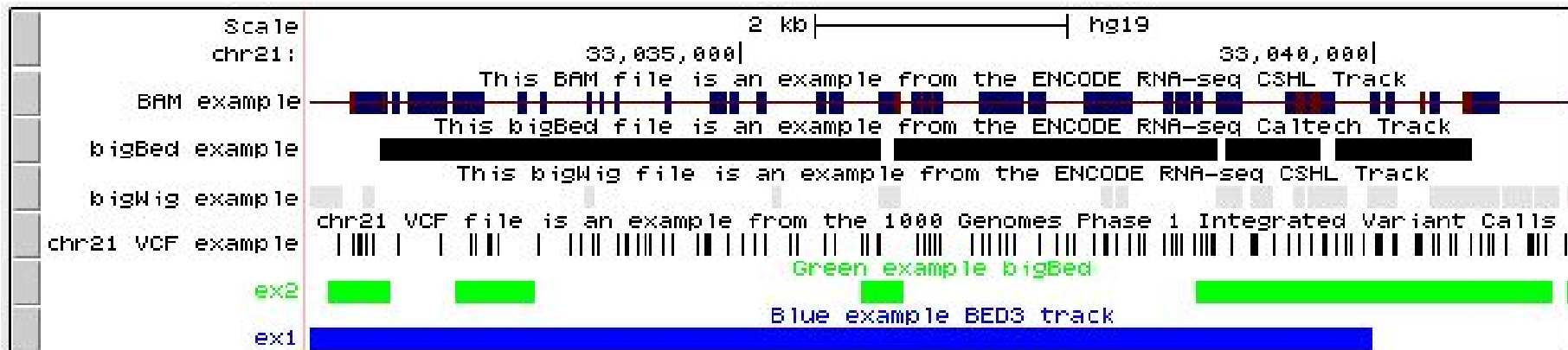
# Custom Annotations in the Browser

Custom Tracks (Text/Binary)

Track Hubs

Assembly Hubs

Track hubs are web-accessible directories of genomic data.



<http://genome.ucsc.edu/cgi-bin/hgTracks?db=hg19&hubUrl=http://genome.ucsc.edu/goldenPath/help/examples/hubDirectory/hub.txt>

# Custom Annotations in the Browser

Custom Tracks (Text/Binary)

Track Hubs

Assembly Hubs

**hub.txt**

```
hub MyHubsNameWithoutSpaces
shortLabel My Hub's Name
longLabel Longer label about my hub.
email myEmail@address
genomesFile genomes.txt
```

**genomes.txt**

```
genome hg19
trackDb hg19/trackDb.txt (URL or path)
```

```
genome assembly_database_2
trackDb assembly_2_path/trackDb.txt
```

**trackDb.txt**

```
track uniqueNameNoSpacesOrDots
type track_type
shortLabel label 17 chars
longLabel label up to 80 chars
bigDataUrl track_data_url (URL or relative path to file)
```

```
track bam1
type bam
shortLabel BAM example
longLabel This BAM file is an example from the ENCODE RNA-seq CSHL Tr
visibility dense
bigDataUrl http://fileservr/directory/of/files/file.bam
```

A new stanza is added for each new track.  
bigDataUrl lines point to the data file (URL or path) in each stanza.  
More trackDb parameters can be defined (color, ect.).

# Custom Annotations in the Browser

Custom Tracks (Text/Binary)

Track Hubs

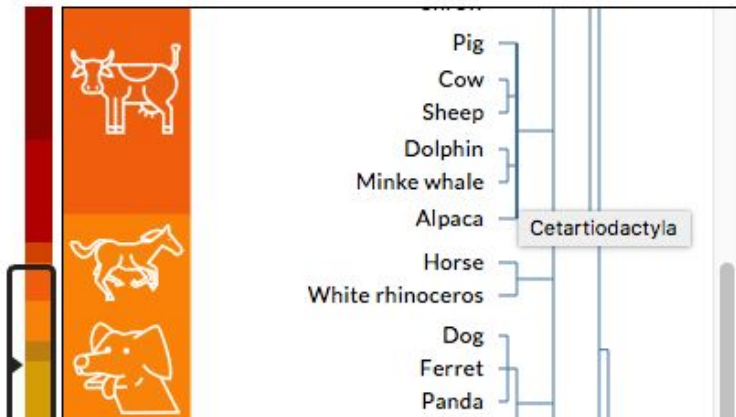
**Assembly Hubs**

**Assembly Hubs display tracks on a remote genome (2bit format).**

## POPULAR SPECIES



## REPRESENTED SPECIES



# Custom Annotations in the Browser

Custom Tracks (Text/Binary)

Track Hubs

Assembly Hubs

**Assembly Hubs display tracks on a remote genome (2bit format).**

Track Data Hubs

Track data hubs are collections of external tracks that can be imported into the UCSC Genome Browser. Track hubs show up under the hub's own blue label bar on the main browser page, as well as on the configure page. To import a public hub click its "Connect" button below.

**NOTE: Because Track Hubs are created and maintained by external sources, UCSC is not responsible for their content.**

Public Hubs My Hubs

URL:  Add Hub

Paste in: <http://path.lab.edu/to/hubDirectory/hub.txt>

# Custom Annotations in the Browser

Custom Tracks (Text/Binary)

Track Hubs

Assembly Hubs

Assembly Hubs display tracks on a remote genome (2bit format).

<http://path.lab.edu/to/>

**hubDirectory**

|\_hub.txt

|\_genomes.txt

|\_groups.txt

|\_hg19

|  |\_trackDb.txt

|\_araTha1

|  |\_trackDb.txt

|  |\_description.html

|  |\_araTha1.2bit



**genome** hg19

**trackDb** hg19/trackDb.txt

**genome** araTha1

**trackDb** araTha1/trackDb.txt

**scientificName** Arabidopsis thaliana

**description** Feb. 2011 Thale cress

**organism** Arabidopsis thaliana

**defaultPos** chr1:1000000-2000000

**groups** groups.txt

**htmlPath** araTha1/description.html

**twoBitPath** araTha1/araTha1.2bit



# Custom Annotations in the Browser

Custom Tracks (Text/Binary)

Track Hubs

Assembly Hubs

The screenshot displays the Arabidopsis thaliana Genome Browser interface. At the top is a navigation bar with links: Home, Genomes, Genome Browser, Tools, Mirrors, Downloads, My Data, Help, and About Us. Below this are two main sections: "Browse/Select Species" and "Find Position".

**Browse/Select Species**

**POPULAR SPECIES**

Human, Mouse, Rat, Fruitfly, Worm, Yeast

Enter species or common name

**REPRESENTED SPECIES**

Plant araTha1 (1) Hub Genomes

Human, Chimp, Bonobo, Gorilla, Orangutan, Gibbon, Green monkey, Crab-eating macaque, Rhesus, Baboon (anubis), Baboon (hamadryas)

**Find Position**

Plant araTha1 Hub Assembly

Feb. 2011 Thale cress

GO

**Position/Search Term**

Enter position, gene symbol or search terms

Current position: chr1:1,499,939-1,500,060

**Arabidopsis thaliana Genome Browser - araTha1 assembly**

[view sequences](#)

UCSC Genome Browser assembly ID: araTha1  
Sequencing/Assembly provider ID: The Institute for Genomic Research TAIR 10  
Assembly date: Feb. 2011  
GenBank accession ID: GCA\_000001735.1  
NCBI Genome information: NCBI genome/4 (Arabidopsis thaliana)  
NCBI Assembly information: NCBI assembly/237408 (TAIR 10)  
BioProject information: NCBI Bioproject: 10719

*Arabidopsis thaliana*

# Custom Annotations in the Browser

Custom Tracks (Text/Binary)

Track Hubs

Assembly Hubs

UCSC Genome Browser on Arabidopsis thaliana Feb. 2011 Thale cress Assembly (araTha1)

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

chr1:13,184,923-13,220,025 35,103 bp. enter position or search terms go

Scale 10 kb | araTha1

chr1: | 13,190,000 | 13,195,000 | 13,200,000 | 13,205,000 | 13,210,000 | 13,215,000 | 13,220,000

chr1\_2 Assembly

chr1\_3 Gap

Gap

Genscan Gene Predictions

Simple Repeats

Simple Tandem Repeats by TRF

chr1.2944 chr1.3002 chr1.3003

<http://genome.ucsc.edu/cgi-bin/hgTracks?genome=araTha1&hubUrl=http://genome.ucsc.edu/goldenPath/help/examples/hubExamples/hubAssembly/plantAraTha1/hub.txt>

# Custom Annotations in the Browser

Custom Tracks (Text/Binary)

Track Hubs

**Assembly Hubs**

`faToTwoBit input.fasta output.2bit`



```
>scaffold_1
GTTGTAAATACTCTATTCTACAATAAAACCAAAGATAACTCTTTATCAG
TCATAGGTTGAATTGGCGTTGAAGTAAACAAAAATACTGCTCAAAGG
...
>scaffold_2
AGTTATGACAAACTATAAAAAGTCGGTAGAGACAAAAGCGAAAAGGATCT
TCGTTCTGTTGACGAAGCGACCAAACCTGAGCACAAAGATAAATCCCGAATA
...
>scaffold_3
CATAAATTCATAAATCAATTCATGAAGAATAATTTTAGAAAAATGGTTCA
TAGAAAAATTTCCAGGAAGTTTGAAGTTGCTATAATGATTATTTCTCTTG
...
>scaffold_4
ATACTGCCCATGCCTCATCTACTCCCTTCTTAACCTTTTCATCAACAAC
TGCTTAAGTTGGTCATCACTGACTGAGTGAGCCCAAACCTGGAATAGCTGT
...
```

The resulting binary file **output.2bit** is used to build the Browser display.

On the fly three tracks are generated from the 2bit:

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

# Custom Annotations in the Browser

Custom Tracks (Text/Binary)

Track Hubs

Assembly Hubs

The screenshot displays a genomic browser interface. At the top, a window position sidebar shows 'Window Position', 'Scale', 'chr1: 1,499,950', and 'AjuI', 'AjuI', 'ScrFI'. The main view shows a DNA sequence for 'Arabidopsis thaliana Feb. 2011 Thale cress chr1:1,499,939-1,500,060 (122 bp)'. The sequence is 'TTCCACACTATCATTATTATTTTGAACCTCTGCGCCCTGGAGGCGGAAACGCGCTTTCCTTCACATGAGCAATACTCTGATTGACCCGAGACGCGGCTTCTCTCTGCTTCTCC'. Below the sequence is a 'Restriction Enzymes from REBASE' track. A control bar includes 'move start' and 'move end' buttons with a '2.0' zoom level, and buttons for 'default tracks', 'default order', 'hide all', 'add custom tracks', 'track hubs', 'configure', 'multi-region', 'reverse', 'resize', 'refresh', 'collapse all', and 'expand all'. A central instruction reads: '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.' Below this are three track hubs: 'Mapping' (with sub-tracks 'Base Position', 'Assembly', 'Gap', 'Restr Enzymes', 'Short Match'), 'Genes' (with 'Genscan Genes'), and 'Variation'. Each hub has a 'refresh' button.

# Custom Annotations in the Browser

Custom Tracks (Text/Binary)

Track Hubs

**Assembly Hubs**

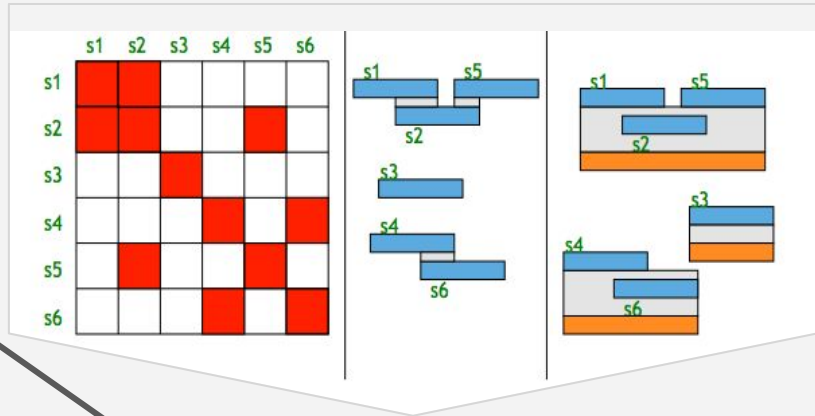
The screenshot displays a genome browser interface with three track groups, each represented by a blue header bar and a light gray content area. Each group includes a 'refresh' button on the right side.

- Mapping:** The header bar contains a minus sign icon, the title 'Mapping', and a 'refresh' button. The content area features five sub-tracks: 'Base Position' (dropdown menu set to 'dense'), 'Assembly' (dropdown menu set to 'hide'), 'Gap' (dropdown menu set to 'hide'), 'Restr Enzymes' (dropdown menu set to 'full'), and 'Short Match' (dropdown menu set to 'squish').
- Genes:** The header bar contains a minus sign icon, the title 'Genes', and a 'refresh' button. The content area features one sub-track: 'Genscan Genes' (dropdown menu set to 'hide').
- Variation:** The header bar contains a minus sign icon, the title 'Variation', and a 'refresh' button. The content area is currently empty.

The groups.txt file defines the blue bars grouping tracks in assembly hubs.

## 2. Building An Assembly Hub

```
hubDirectory
├─ hub.txt
├─ genomes.txt
├─ hg19
└─ trackDb.txt
```



2bit version of your FASTA file

# Building An Assembly Hub

Copy Example Files

Build Hub

Add Blat and Share



**wget <http://genome.ucsc.edu/goldenPath/help/examples/bigBedExample.bb>**  
**(or download and move file to server location)**



# Building An Assembly Hub

Copy Example Files

Build Hub

Add Blat and Share

**Add Custom Tracks**

clade  genome  assembly

Display your own data as custom annotation tracks in the browser. Data must be in [bigBed](#), [bigChain](#), [bigGenePred](#), [bigMaf](#), [bigPsl](#), [bigWig](#), [BAM](#), [VCF](#), [BED](#), [BED detail](#), [bedGraph](#), [broadPeak](#), [CRAM](#), [GFF](#), [GTF](#), [MAF](#), [narrowPeak](#), [Personal Genome SNP](#), [PSL](#), or [WIG](#) formats. To configure the display, set [track](#) and [browser](#) line attributes as described in the [User's Guide](#). Data in the bigBed, bigWig, bigGenePred, BAM and VCF formats can be provided via only a URL or embedded in a track line in the box below. Examples are [here](#).

Paste URLs or data:  Or upload:  no file selected

`http://Your/Path/To/Copied/bigBedExample.bb`



# Building An Assembly Hub

Copy Example Files

Build Hub

Add Blat and Share

<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:

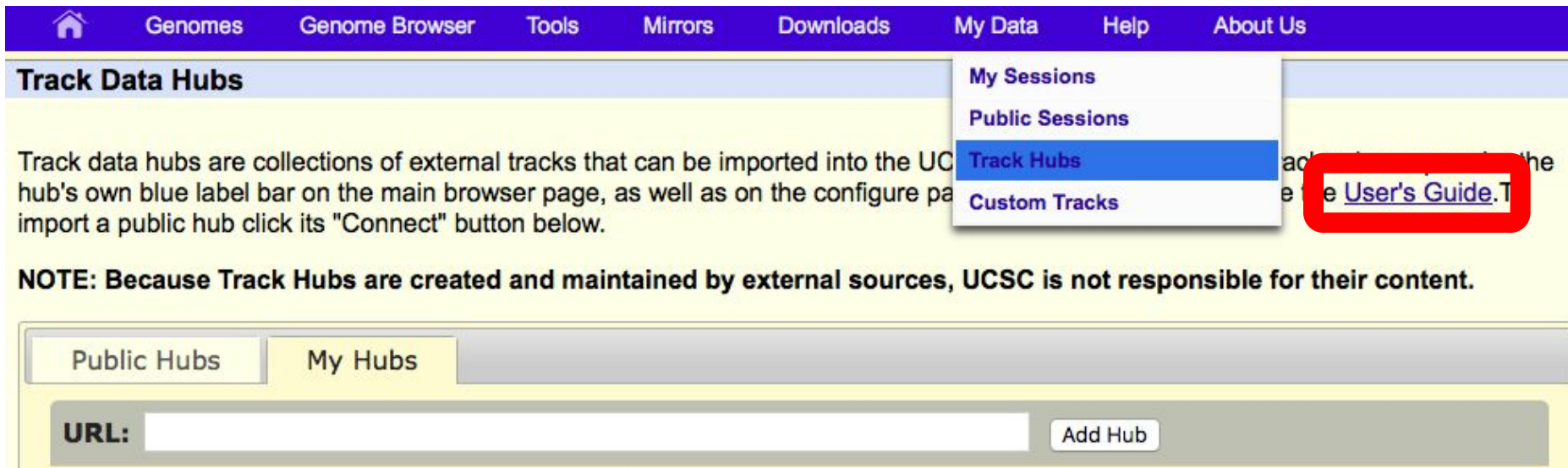
```
wget -r --no-parent --reject "index.html*" -nH --cut-dirs=3  
http://genome.ucsc.edu/goldenPath/help/examples/hubExamples/hubAssembly/plantAraThal/
```

# Building An Assembly Hub

Copy Example Files

Build Hub

Add Blat and Share



The screenshot shows the UCSC Genome Browser navigation bar with a home icon and links for Genomes, Genome Browser, Tools, Mirrors, Downloads, My Data, Help, and About Us. Below the navigation bar is the 'Track Data Hubs' section. A dropdown menu is open under 'My Data', showing options for My Sessions, Public Sessions, Track Hubs (highlighted), and Custom Tracks. The 'Track Hubs' option is highlighted in blue. To the right of the dropdown menu, the text 'User's Guide.T' is circled in red. Below the dropdown menu, there is a note: 'NOTE: Because Track Hubs are created and maintained by external sources, UCSC is not responsible for their content.' At the bottom of the screenshot, there are two tabs: 'Public Hubs' and 'My Hubs'. Below the tabs is a form with a 'URL:' label, a text input field, and an 'Add Hub' button.

`http://your/path/to/copied/hubAssembly/plantAraTha1/hub.txt`

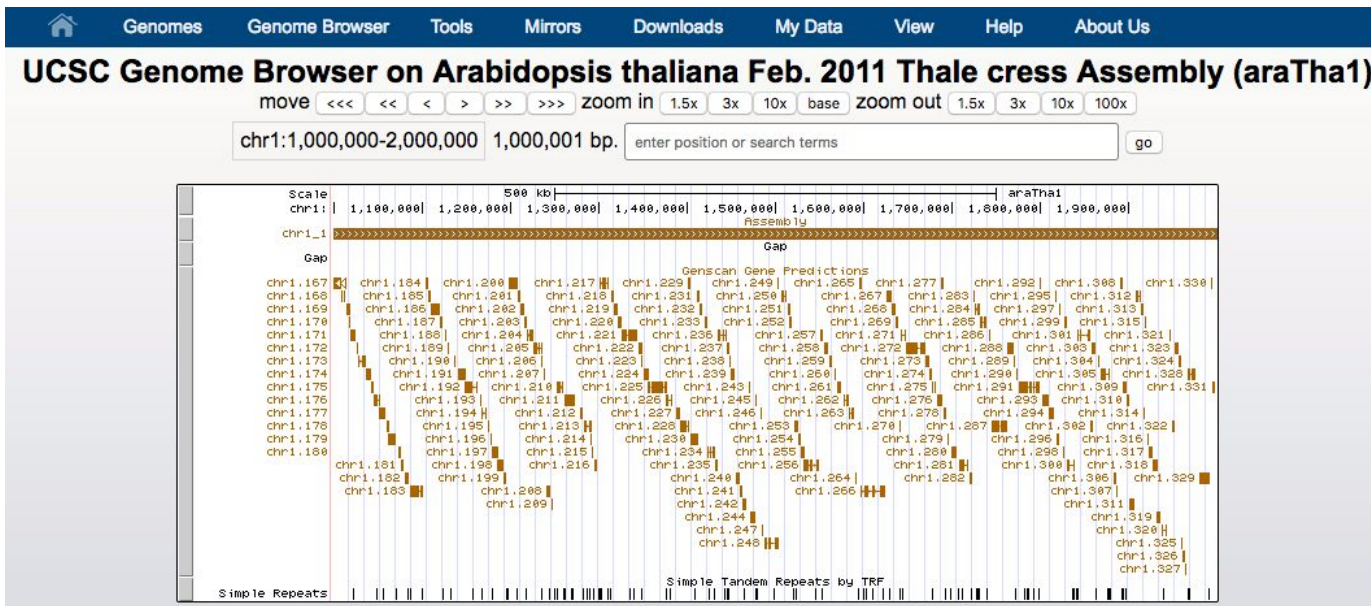
# Building An Assembly Hub

Copy Example Files

Build Hub

Add Blat and Share

<http://genome.ucsc.edu/cgi-bin/hgTracks?genome=araTha1&hubUrl=http://your/path/to/copied/hubAssembly/plantAraTha1/hub.txt>



# Building An Assembly Hub

Copy Example Files

**Build Hub**

Add Blat and Share

Edit a copy of the working hub.

## 1. First step change folder names

```
mv plantAraTha1/ daph
mv araTha1/ daph
```

## 2. Then change the hub.txt

Edit the hub.txt

```
hub daph
shortLabel daph hub
longLabel Daph Hub
genomesFile genomes.txt
email genome-www@soe.ucsc.edu
descriptionUrl http://daphniagenomes.org/downloads
```

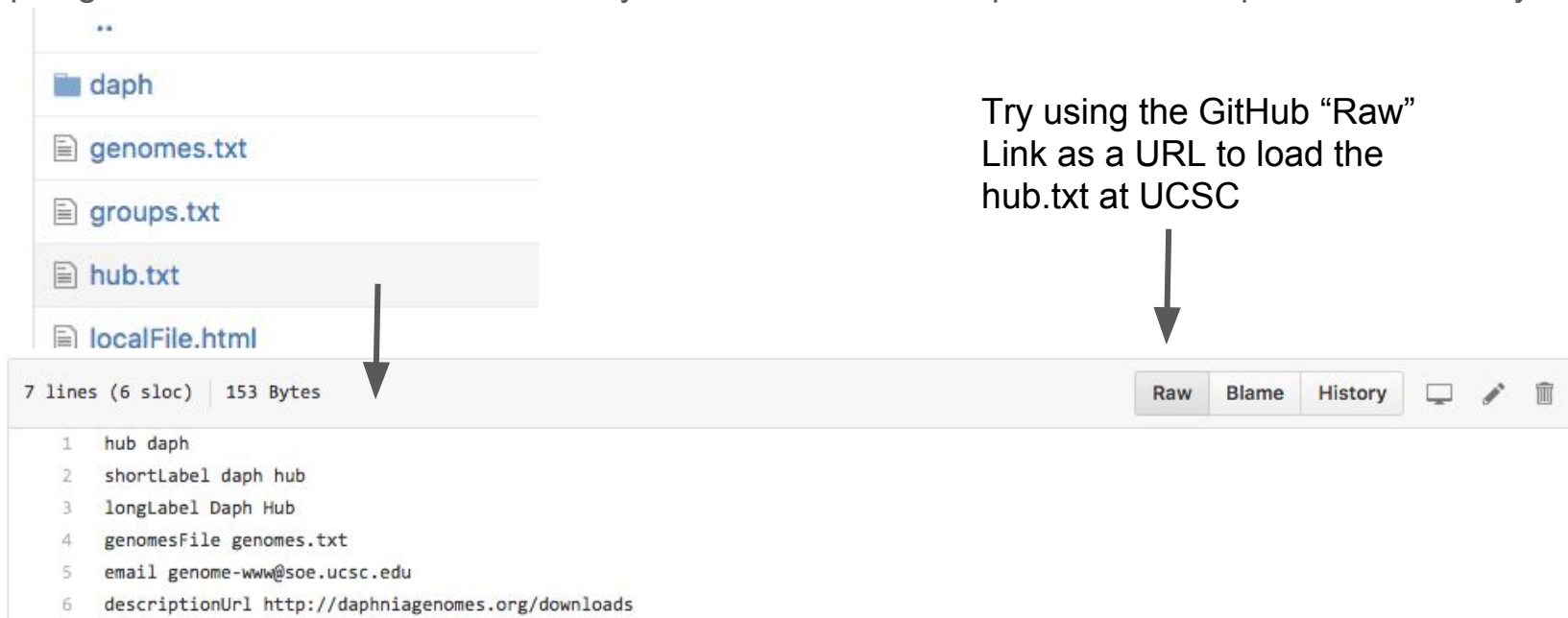
## 3. Then change the genomes.txt

<https://github.com/ucsc-browser/assemblyHubEx>

# Building An Assembly Hub

**Exercise:** Visit the GitHub Files to See a Working Example

<https://github.com/ucsc-browser/assemblyHubEx/tree/master/Daphnia/hubExamples/hubAssembly/>



Try using the GitHub “Raw” Link as a URL to load the hub.txt at UCSC

```
7 lines (6 sloc) | 153 Bytes
```

```
1 hub daph
2 shortLabel daph hub
3 longLabel Daph Hub
4 genomesFile genomes.txt
5 email genome-www@soe.ucsc.edu
6 descriptionUrl http://daphniagenomes.org/downloads
```

Raw Blame History

# Building An Assembly Hub

Copy Example Files

Build Hub

Add Blat and Share

`faToTwoBit your.fasta your.2bit`



```
>scaffold_1
GTTGTAAATACTCTATTCTACAATAAAACCAAAGATAACTCTTTATCAG
TCATAGGTTGAATTGGCGTTGAAGTAAACAAAAAATACTGCTCAAAGG
...
>scaffold_2
AGTTATGACAACTATAAAAAGTCGGTAGAGACAAAAGCGAAAAGGATCT
TCGTTTCGTGGACGAAGCGACCAAACGAGCACAAGATAAATCCCGAATA
...
>scaffold_3
CATAAATTCATAAATCAATTCATGAAGAATAATTTTAGAAAAATGGTTCA
TAGAAAAATTTCCAGGAAGTTTGAAGTTGCTATAATGATTATTTCTCTTG
...
>scaffold_4
ATACTGCCCATGCCTCATCTACTCCCTTCTTAACCTTTTCATCAACAAC
TGCTTAAGTTGGTCATCACTGACTGAGTGAGCCCAAACGGAATAGCTGT
...
```



Create your 2bit file and edit  
genomes.txt to point to it.

```
faToTwoBit Daphnia_pulex.allmasked daph.2bit
```

```
findMotif -motif=AAAAAA -strand=+ daph.2bit > motif.bed
```



# Building An Assembly Hub

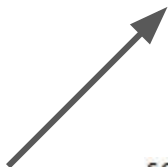
Copy Example Files

Build Hub

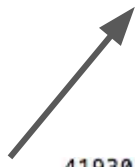
Add Blat and Share

`bedToBigBed` `your.bed` `2bit.chrom.sizes` `yourTrack.bb`

```
scaffold_1 3890 3896
scaffold_1 4063 4069
scaffold_1 4236 4242
scaffold_1 4755 4761
scaffold_1 5976 5982
scaffold_1 6022 6028
scaffold_1 6023 6029
scaffold_1 6398 6404
scaffold_1 6399 6405
scaffold_1 26373 26379
scaffold_1 26374 26380
scaffold_1 29767 29773
scaffold_1 29790 29796
scaffold_1 29917 29923
scaffold_1 29918 29924
scaffold_1 30872 30878
scaffold_1 32125 32131
scaffold_1 32218 32224
scaffold_1 32219 32225
scaffold_1 32220 32226
```



```
scaffold_1 4193030
scaffold_3 3777634
scaffold_2 3740169
scaffold_4 3075709
scaffold_5 2511979
scaffold_6 2406117
scaffold_8 2335496
scaffold_7 2324446
scaffold_9 2251199
scaffold_12 2218424
```



Add `yourTrack.bb` to `trackDb.txt`

Find more tools like `bedToBigBed` at:  
<http://hgdownload.soe.ucsc.edu/admin/exe/>

More Utilities:

*faToTwoBit* (build `2bit`) *twoBitInfo* (extract  
`2bit.chrom.sizes`) *twoBitMask*, *twoBitToFa*  
*twoBitDup* (work with your `2bit`) *hgGcPercent*,  
*findMotif*, *oligoMatch* (extract tracks from your  
`2bit`) *bedToBigBed* or *wigToBigWig* (convert text  
files to binary files) *gfServer*, *gfClient* (run BLAT)  
and many more....

# Building An Assembly Hub

Copy Example Files

Build Hub

Add Blat and Share

UCSC Genome Browser on Daph genome Assembly (daph)

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

scaffold\_1:3,674-4,322 649 bp.  go

Scale scaffold\_1: 200 bases | daph

Restriction Enzymes from REBASE

Perfect Matches to Short Sequence (TATA)

This track shows a match on the sequence AAAAAA

<http://genome.ucsc.edu/cgi-bin/hgTracks?genome=daph&hubUrl=https://raw.githubusercontent.com/ucsc-browser/assemblyHub/Ex/master/Daphnia/hubExamples/hubAssembly/daph/hub.txt>



# Building An Assembly Hub

Copy Example Files

Build Hub

Add Blat and Share

<http://path.lab.edu/to/>

## *hubDirectory*

|\_hub.txt

|\_genomes.txt

|\_groups.txt

|\_hg19

|  |\_trackDb.txt

|\_yourGenome

|  |\_trackDb.txt

|  |\_description.html

|  |\_your.2bit



**genome** hg19  
**trackDb** hg19/trackDb.txt

**genome** yourGenome  
**trackDb** yourGenome/trackDb.txt  
**scientificName** Your Genome  
**description** Feb. 2017 Assembly  
**organism** Your organism  
**defaultPos** chr1:1000000-2000000  
**groups** groups.txt  
**htmlPath** yourGenome/description.html  
**twoBitPath** yourGenome/your.2bit  
**blat** path.lab.edu 17779  
**transBlat** path.lab.edu 1777

# Building An Assembly Hub

Copy Example Files

Build Hub

Add Blat and Share

[http://genomewiki.ucsc.edu/index.php/Assembly\\_Hubs#Adding\\_BLAT\\_servers](http://genomewiki.ucsc.edu/index.php/Assembly_Hubs#Adding_BLAT_servers)

<http://path.lab.edu/to/>

**hubDirectory**

**|\_genomes.txt**

**|\_yourGenome**

**|\_your.2bit**

**genome** *yourGenome*

...

**blat** *path.lab.edu (or IP) 17779*

**transBlat** *path.lab.edu (or IP) 17777*

*(trans is short for "translated" AA search)*

`gfServer start localhost 17779 -stepSize=5 your.2bit &`

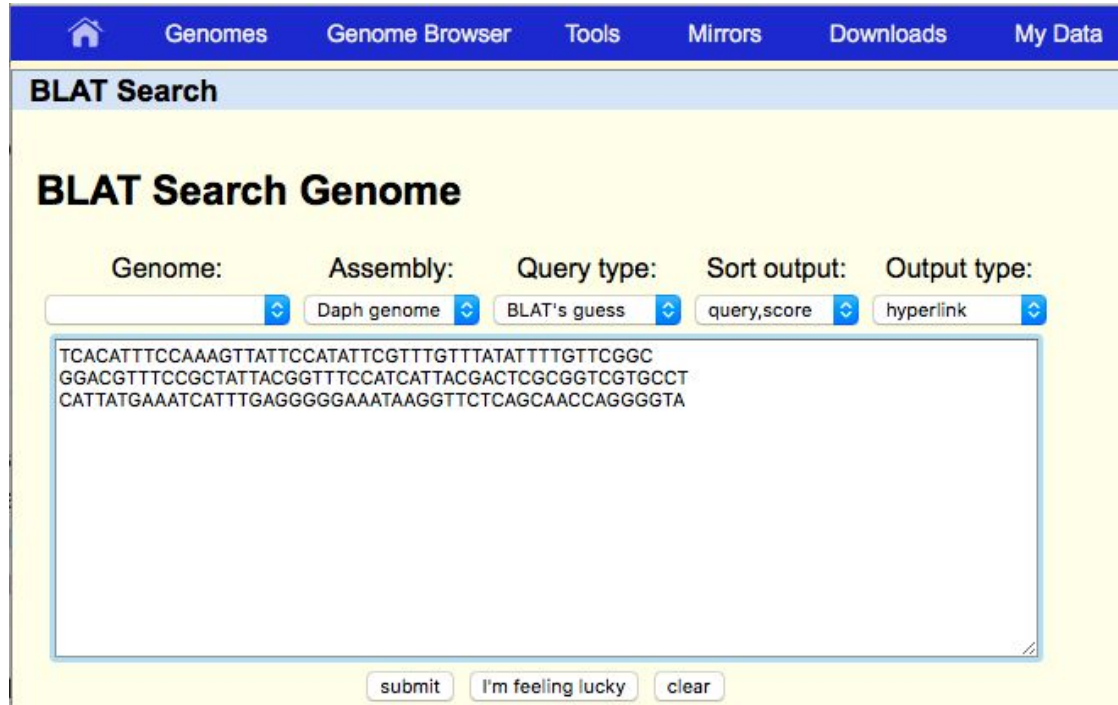
`gfServer start localhost 17777 -trans -mask your.2bit &`

# Building An Assembly Hub

Copy Example Files

Build Hub

Add Blat and Share



The screenshot shows the BLAT Search Genome web interface. At the top, there is a navigation bar with a home icon and links for Genomes, Genome Browser, Tools, Mirrors, Downloads, and My Data. Below this is a header for "BLAT Search". The main content area is titled "BLAT Search Genome" and contains several dropdown menus for configuration: "Genome:" (empty), "Assembly:" (Daph genome), "Query type:" (BLAT's guess), "Sort output:" (query,score), and "Output type:" (hyperlink). A large text area displays the search results, which are three lines of DNA sequence: TCACATTTCCAAAGTTATTCCATATTCGTTTGTATATTTGTTTCGGC, GGACGTTTCCGCTATTACGGTTTCCATCATTACGACTCGCGGTCGTGCCT, and CATTATGAAATCATTGAGGGGGAAATAAGTTCTCAGCAACCAGGGGTA. At the bottom of the interface, there are three buttons: "submit", "I'm feeling lucky", and "clear".

BLAT Search

## BLAT Search Genome

Genome: Assembly: Query type: Sort output: Output type:

Daph genome BLAT's guess query,score hyperlink

```
TCACATTTCCAAAGTTATTCCATATTCGTTTGTATATTTGTTTCGGC
GGACGTTTCCGCTATTACGGTTTCCATCATTACGACTCGCGGTCGTGCCT
CATTATGAAATCATTGAGGGGGAAATAAGTTCTCAGCAACCAGGGGTA
```

submit I'm feeling lucky clear

# Building An Assembly Hub

Copy Example Files

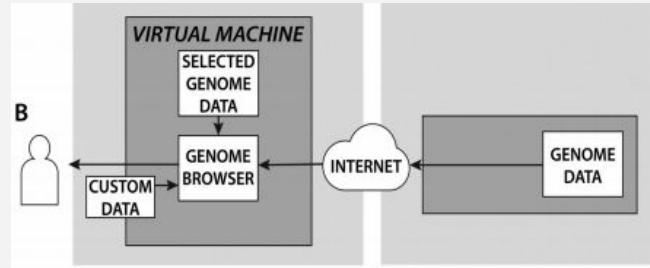
Build Hub

Add Blat and Share

The screenshot displays the UCSC Genome Browser interface for the Daphnia genome. The top navigation bar includes links for Genomes, Genome Browser, Tools, Mirrors, Downloads, My Data, View, Help, and About Us. The main header reads "UCSC Genome Browser on Daph genome". Below the header, there are navigation controls for moving and zooming, and a search box containing "scaffold\_3:44,840-67,294 22,455 bp.". A dropdown menu is open over the "My Data" link, showing options: "My Sessions", "Public Sessions", "Track Hubs", and "Custom Tracks". The main content area shows several tracks: "Scale" (10 kb), "YourSeq", "Restriction Enzymes from REBASE", "Perfect Matches to Short Sequence (TATAWAAR)", and "Motifs".

[http://genome.ucsc.edu/cgi-bin/hgTracks?hgS\\_doOtherUser=submit&hgS\\_otherUserName=brianlee&hgS\\_otherUserSessionName=pag.exHub](http://genome.ucsc.edu/cgi-bin/hgTracks?hgS_doOtherUser=submit&hgS_otherUserName=brianlee&hgS_otherUserSessionName=pag.exHub)

### 3. Existing Hubs and Solutions to Hosting Files



# Existing Hubs and Solutions to Hosting Files

## Genbank Assembly Hubs

## Genome Browser in a Box

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

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 ▲▼
<a href="#">other/synthetic assemblies</a>	3	3	94	2,835,004	30,159	945,001
<a href="#">vertebrate_other</a>	156	172	18,548,615	193,684,015,605	10,441	1,126,069,858
<a href="#">vertebrate_mammalian</a>	118	204	30,643,657	498,264,459,566	16,259	2,442,472,841

# Existing Hubs and Solutions to Hosting Files

Genbank Assembly Hubs

Genome Browser in a Box

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

Genomes Genome Browser Tools Mirrors Downloads My Data Help About Us

## Vertebrate Mammalian assembly hub - 2017-01-05

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 ▲▼
1	other/synthetic assemblies	3	3	94	2,835,004	30,159	945,001
2	vertebrate other	156	172	18,548,615	193,684,015,605	10,441	1,126,069,858
3	vertebrate mammalian	118	204	30,643,657	498,264,459,566	16,259	2,442,472,841
4	plant	190	269	34,577,423	145,341,422,954	4,203	540,302,687
5	protozoa	282	338	3,939,128	16,816,724,183	4,269	49,753,621
6	invertebrate	392	492	32,264,511	170,439,035,382	5,282	346,420,803
7	fungi	1,106	1,215	4,143,097	38,677,096,556	9,335	31,833,001
8	archaea	688	742	57,569	2,010,246,046	34,918	2,709,226
9	bacteria	34,005	58,658	8,397,216	234,147,691,500	27,883	3,991,743
	<b>totals:</b>	<b>36,940</b>	<b>62,093</b>	<b>132,571,310</b>	<b>1,299,383,526,796</b>	<b>9,801</b>	<b>20,926,409</b>

**NOTE:** This is a prototype work in progress. Not all assemblies are represented here yet. Prototype gene tracks from the NCBI gene predictions delivered with the assembly are available on most assemblies. There are no blat servers. Users could copy the hub skeleton structure of a specific assembly to local systems and run a blat server at their location with their own assembly hub of that specific genome. This system may not be available at all times as the procedures are finalized of how to keep this up to date on an automatic process.

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 percer ▲▼
1	<a href="#">9785</a>	15 Jul 2009	<a href="#">African bush elephant</a>	Loxodonta africana	<a href="#">SAMN02953622</a>	2,352	3,196,738,035	46,401,353	% 39.7
2	<a href="#">9646</a>	15 Dec 2009	<a href="#">Alluropoda melanoleuca</a>	Alluropoda melanoleuca	<a href="#">SAMN00008160</a>	81,466	2,299,492,210	1,281,781	% 40.6

Background *PROTOTYPE* and displayed **NOTE** share this data has not been QA reviewed and could change.



# Existing Hubs and Solutions to Hosting Files

## Genbank Assembly Hubs

## Genome Browser in a Box

<http://genome-test.cse.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	<a href="#">9785</a>	15 Jul 2009	<a href="#">African bush elephant</a>	Loxodonta africana	<a href="#">SAMN02953622</a>	2,352	3,196,738,035	46,401,353	% 39.76	78,195,493 % 2.45	0 0 % 0.00	<a href="#">GCA_000001905.1</a>
2	<a href="#">9646</a>	15 Dec 2009	<a href="#">Ailuropoda melanoleuca</a>	Ailuropoda melanoleuca	<a href="#">SAMN00008160</a>	81,466	2,299,492,210	1,281,781	% 40.62	54,196,184 % 2.36	52,206 473,410,926 % 20.59	<a href="#">GCA_000004335.1</a>
3	<a href="#">37293</a>	16 Mar 2015	<a href="#">Aotus nancymaae</a>	Aotus nancymaae	<a href="#">SAMN03121886</a>	29,222	2,926,565,220	8,280,397	% 35.21	391,667,293 % 13.38	0 0 % 0.00	<a href="#">GCA_000952055.1</a>
4	<a href="#">9555</a>	05 Jun 2012	<a href="#">baboon</a>	Papio anubis	<a href="#">SAMN02981400</a>	63,250	2,948,380,710	139,646,187	% 40.20	55,130,419 % 1.87	0 0 % 0.00	<a href="#">GCA_000264685.1</a>
5	<a href="#">118797</a>	31 Jul 2013	<a href="#">baiji</a>	Lipotes vexillifer	<a href="#">SAMN02953859</a>	30,712	2,429,195,737	2,419,148	% 40.75	32,195,939 % 1.33	0 0 % 0.00	<a href="#">GCA_000442215.1</a>



# Existing Hubs and Solutions to Hosting Files

**Exercise:** Visit Genbank Assembly Hub Files to See a Working Example

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

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

1	<a href="#">9785</a>	15 Jul 2009	<a href="#">African bush elephant</a>	Loxodonta africana	<a href="#">SAMN02953622</a>	2,352	3,196,738,035	46,401,353	% 39.76	78,195,493 % 2.45	percent	0 0 % 0.00	<a href="#">GCA_000001905.1</a>
---	----------------------	-------------	---------------------------------------	--------------------	------------------------------	-------	---------------	------------	---------	----------------------	---------	------------------	---------------------------------

[http://genome-test.cse.ucsc.edu/cgi-bin/hgGateway?hubUrl=http://genome-test.cse.ucsc.edu/gbdb/hubs/genbank/vertebrate\\_mammalian/hub.ncbi.txt&genome=GCA\\_000001905.1\\_Loxafr3.0&position=lastDbPos](http://genome-test.cse.ucsc.edu/cgi-bin/hgGateway?hubUrl=http://genome-test.cse.ucsc.edu/gbdb/hubs/genbank/vertebrate_mammalian/hub.ncbi.txt&genome=GCA_000001905.1_Loxafr3.0&position=lastDbPos)

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

# Existing Hubs and Solutions to Hosting Files

## Genbank Assembly Hubs

## Genome Browser in a Box

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

Genbank Mammals Hub Assembly  
15 Jul 2009/Loxafr3.0

Position/Search Term  
Enter position, gene symbol or search terms  
Current position: chrUn\_GL010027v1:63,879,670-65,879,670

**GO**

**African bush elephant Genome Browser - GCA\_000001905.1\_Loxafr3.0 assembly** [view sequences](#)

**Common name:** African bush elephant  
**Taxonomic name:** *Loxodonta africana*, taxonomy ID: 9785  
**Sequencing/Assembly provider ID:** Broad Institute  
**Assembly date:** 15 Jul 2009  
**Assembly type:** haploid  
**Assembly level:** Scaffold  
**Biosample:** SAMN02953622  
**Assembly accession ID:** GCA\_000001905.1  
**Assembly FTP location:** GCA\_000001905.1\_Loxafr3.0  
**Total assembly nucleotides:** 3,196,738,035  
**Assembly contig count:** 2,352  
**N50 size:** 46,401,353

### Download files for this assembly hub:

To use the data from this assembly for a local hub instance at your institution, download this data as indicated by these instructions.

See also: [track hub help](#) documentation.

To download this data, issue this `wget` command:

```
wget --timestamping -m -nH -x --cut-dirs=5 -e robots=off -np -k \  
--reject "index.html*" -P "GCA_000001905.1_Loxafr3.0" \  
http://genome-test.cse.ucsc.edu/gbdb/hubs/genbank/vertebrate\_mammalian/GCA\_00000190
```

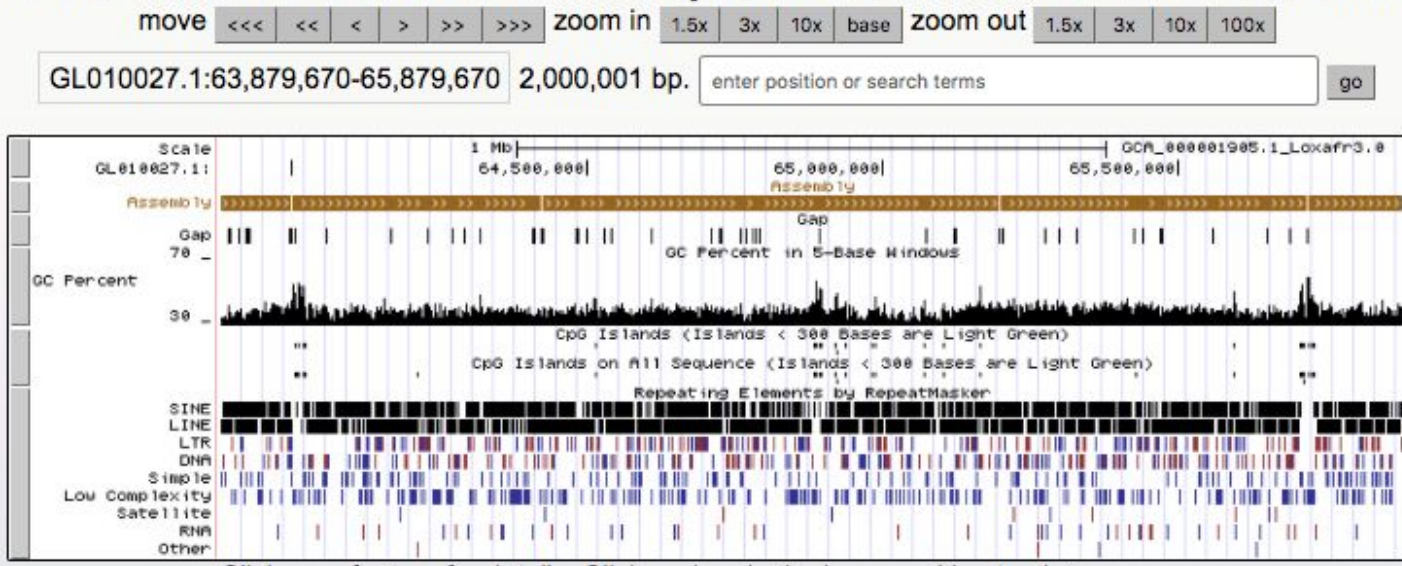
# Existing Hubs and Solutions to Hosting Files

Genbank Assembly Hubs

Genome Browser in a Box

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

## Genome Browser on African bush elephant 15 Jul 2009/Loxfr3.0 Assembly



# Existing Hubs and Solutions to Hosting Files

Genbank Assembly Hubs

Genome Browser in a Box

<http://genome.ucsc.edu/goldenpath/help/gbib.html>

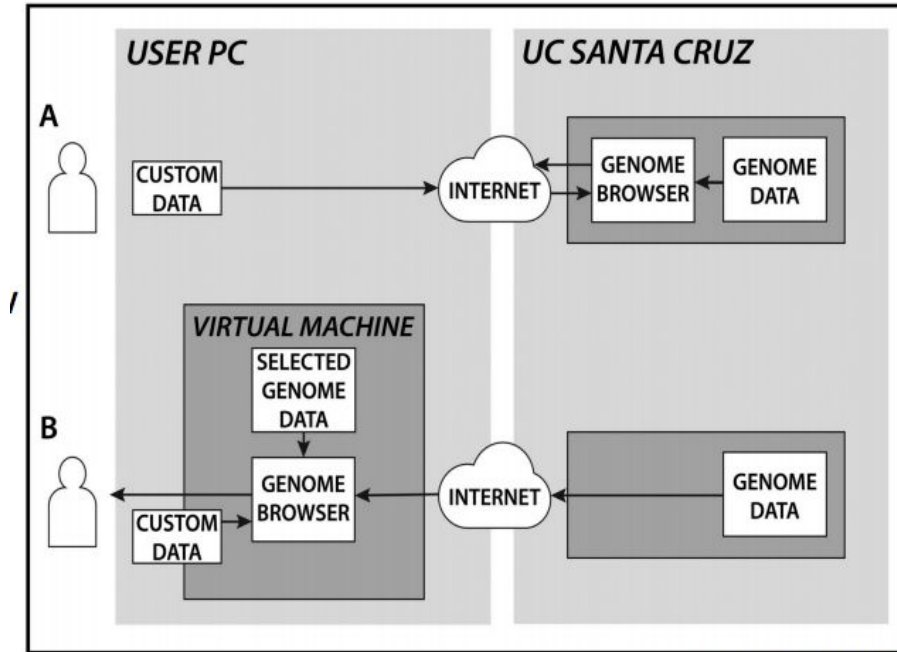


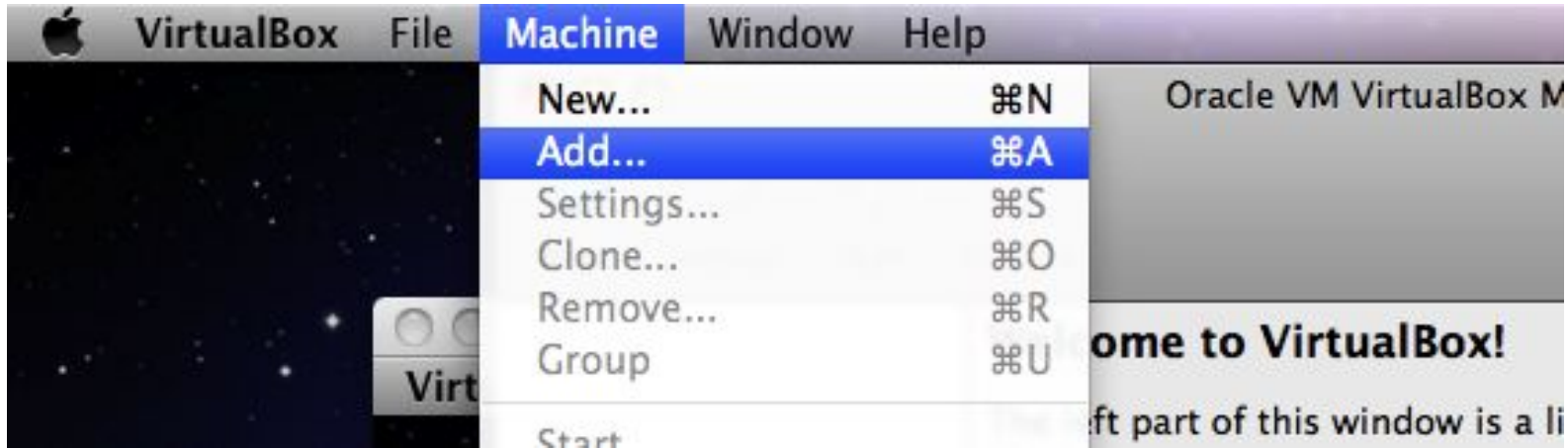
Image Credit: Casey Callow

# Existing Hubs and Solutions to Hosting Files

Genbank Assembly Hubs

Genome Browser in a Box

<http://genome.ucsc.edu/goldenpath/help/gbib.html>



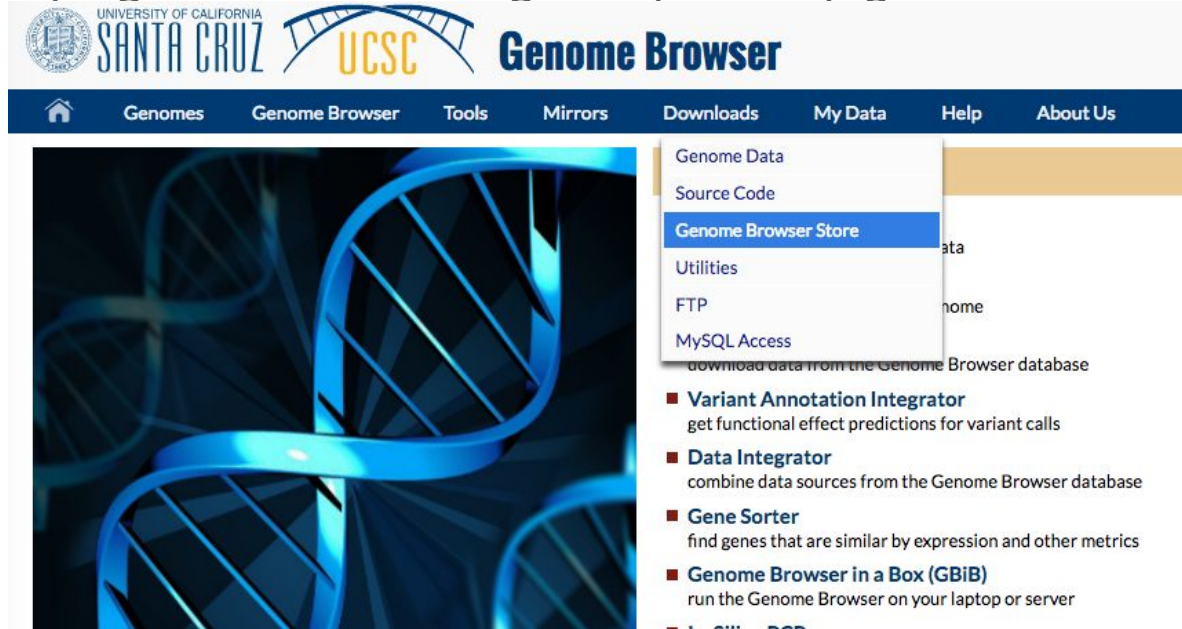
Install VirtualBox free from Oracle to allow a Virtual Machine

# Existing Hubs and Solutions to Hosting Files

Genbank Assembly Hubs

Genome Browser in a Box

<http://genome.ucsc.edu/goldenpath/help/gbib.html>



The screenshot shows the UCSC Genome Browser website. At the top, there are logos for the University of California Santa Cruz and UCSC, followed by the text 'Genome Browser'. Below this is a dark blue navigation bar with white text for 'Genomes', 'Genome Browser', 'Tools', 'Mirrors', 'Downloads', 'My Data', 'Help', and 'About Us'. A dropdown menu is open under 'Downloads', listing 'Genome Data', 'Source Code', 'Genome Browser Store' (highlighted in blue), 'Utilities', 'FTP', and 'MySQL Access'. Below the menu, there is a list of tools with red square bullet points: 'Variant Annotation Integrator' (get functional effect predictions for variant calls), 'Data Integrator' (combine data sources from the Genome Browser database), 'Gene Sorter' (find genes that are similar by expression and other metrics), and 'Genome Browser in a Box (GBiB)' (run the Genome Browser on your laptop or server). The background of the page features a blue 3D DNA double helix.

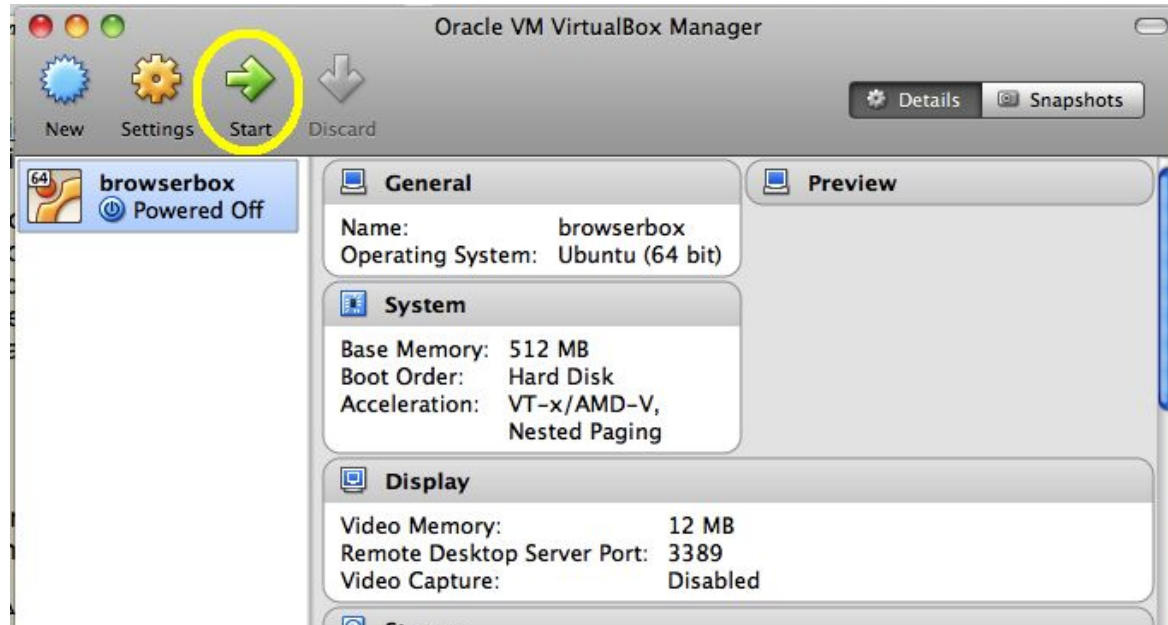
Create store account, free for academic use, licence required for commercial use.

# Existing Hubs and Solutions to Hosting Files

Genbank Assembly Hubs

Genome Browser in a Box

<http://genome.ucsc.edu/goldenpath/help/gbib.html>



Acquire GBiB (6GB), unzip file and start in VirtualBox



# Existing Hubs and Solutions to Hosting Files

Genbank Assembly Hubs

Genome Browser in a Box

<http://genome.ucsc.edu/goldenpath/help/gbib.html>

← → ↻ ⓘ 127.0.0.1:1234/cgi-bin/hgGateway

UNIVERSITY OF CALIFORNIA SANTA CRUZ UCSC **Genome Browser Gateway**

Genomes Genome Browser Tools My Data Help About Us

**Browse/Select Species**

SPECIES SEARCH

Enter species or common name

My Sessions  
Public Sessions  
Track Hubs  
**GBiB Shared Data Folder**  
Custom Tracks

7/hg19)  
rm

Enter position, gene symbol or s

Share local laptop files with the GBiB.









# Existing Hubs and Solutions to Hosting Files

Genbank Assembly Hubs

Genome Browser in a Box

<http://genome.ucsc.edu/goldenpath/help/gbib.html>

## Index of /folders/exampleHub/Daphnia/hubExamples/hubAssembly/daph

<u>Name</u>	<u>Last modified</u>	<u>Size</u>	<u>Description</u>
 <a href="#">Parent Directory</a>		-	
 <a href="#">daph/</a>	2016-12-20 06:32	-	
 <a href="#">genomes.txt</a>	2016-12-20 06:32	500	
 <a href="#">groups.txt</a>	2016-12-20 06:32	273	
 <a href="#">hub.txt</a>	2016-12-20 06:32	153	
 <a href="#">localFile.html</a>	2016-12-20 06:32	383	

127.0.0.1:1234/folders/exampleHub/Daphnia/hubExamples/hubAssembly/daph/hub.txt

The GBiB will interpret shared files as though located on a server.

# Existing Hubs and Solutions to Hosting Files

Genbank Assembly Hubs

Genome Browser in a Box

The screenshot shows a web browser window with the address bar displaying `127.0.0.1:1234//cgi-bin/hgHubConnect`. The page has a blue navigation bar with links for Genomes, Genome Browser, Tools, My Data, Help, and About Us. The main content area is titled "Track Data Hubs" and contains a dropdown menu with options: My Sessions, Public Sessions, Track Hubs (highlighted), GBiB Shared Data Folder, and Custom Tracks. Below the menu, there is a text block explaining that track data hubs are collections of external tracks and a "NOTE" stating that UCSC is not responsible for external hubs. At the bottom, there are tabs for "Public Hubs" and "My Hubs", a URL input field containing `http://127.0.0.1:1234/folders/exampleHub/Daphnia/hubExamples/hubAs`, an "Add Hub" button, and the text "No Unlisted Track Hubs".

Paste path to the files just as you would to a URL.

# Existing Hubs and Solutions to Hosting Files

Genbank Assembly Hubs

Genome Browser in a Box

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

```
browser@browserbox:/folders/exampleHub/Daphnia/hubExamples/hubAssembly/daph/daph$ gfServer start localhost 17777 -trans -mask daph.2bit &
[1] 2040
browser@browserbox:/folders/exampleHub/Daphnia/hubExamples/hubAssembly/daph/daph$ starting translated server...
Counting daph.2bit
Indexing daph.2bit
Server ready for queries!
```

```
browser@browserbox:/folders/exampleHub/Daphnia/hubExamples/hubAssembly/daph/daph$ gfServer start localhost 17779 -stepSize=5 daph.2bit &
[2] 2051
browser@browserbox:/folders/exampleHub/Daphnia/hubExamples/hubAssembly/daph/daph$ starting untranslated server...
Counting tiles in daph.2bit
Done adding
Server ready for queries!
```

Run gfServers on the GBiB and start BLAT for your Assembly

# Existing Hubs and Solutions to Hosting Files

Genbank Assembly Hubs

Genome Browser in a Box

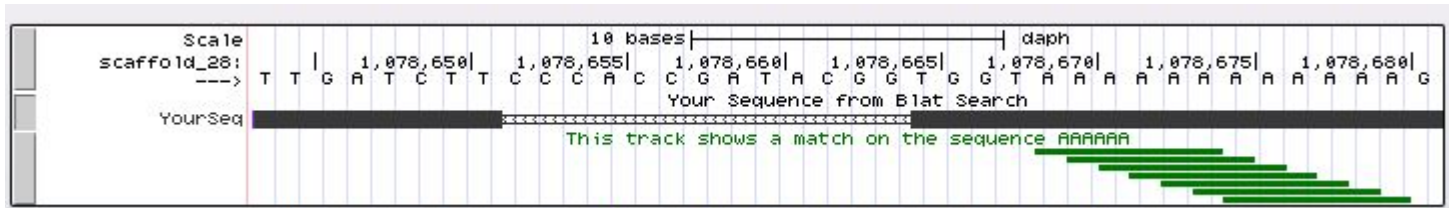
<http://genome.ucsc.edu/goldenPath/help/hubQuickStartAssembly.html#blatGbib>

### BLAT Search Genome

Genome:  Assembly:  Query type:  Sort output:  Output type:

Daph genome  BLAT's guess  query,score  hyperlink

```
TACGTCTGACCGCACTTTAGTTAGATACCATCGCCAAGCTGACTTTTTTT  
TTTTTACCAAAAAGATCAAGGCGGATGAAAGAAGCCTTTTGTAAATGCACT  
TTTCTTCATCCAG
```



# Existing Hubs and Solutions to Hosting Files

## CyVerse Storage Solution

<https://de.cyverse.org/> Discovery Environment (NSF funded: **FREE!**)



### Discovery Environment

The Discovery Environment integrates powerful, community-recommended software tools into a system that:

- Makes big data management easy. Upload, organize, edit, view and search with ease!
- Has 500+ scientific apps that utilize compute clusters and HPC resources as needed.
- Hides the complexity needed to do these tasks.

Log in with your CyVerse ID

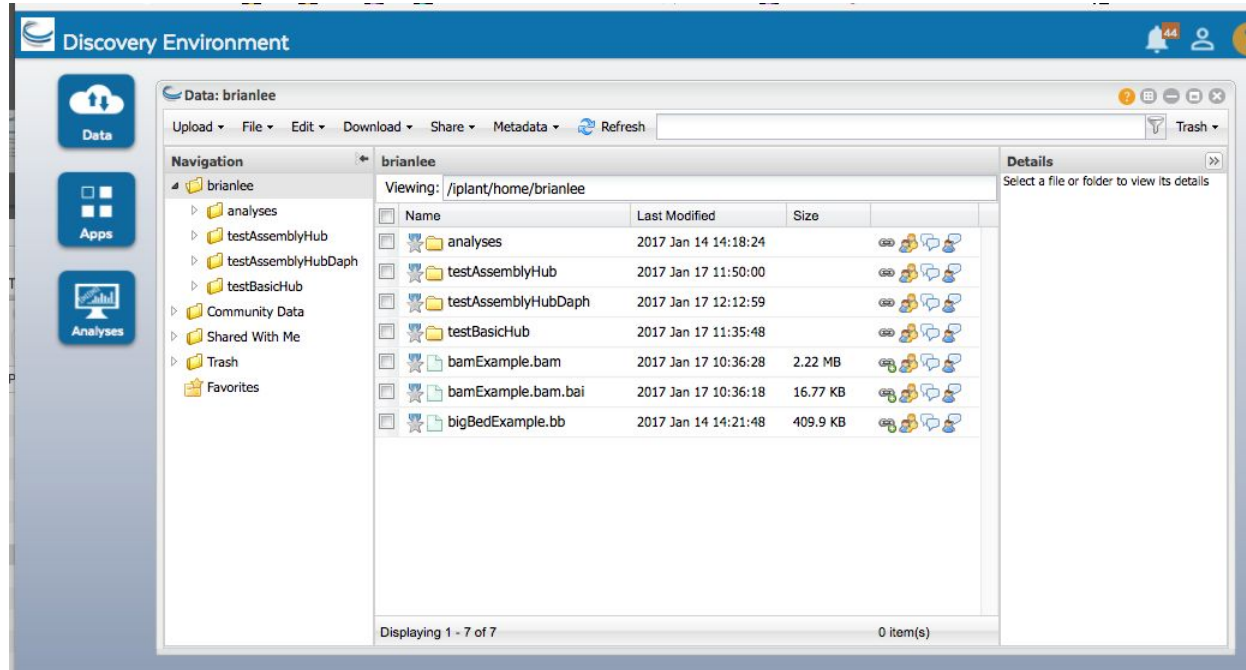
[Forgot Password?](#) [Register Now](#)

Minimum screen resolution supported: 1024 x 768

# Existing Hubs and Solutions to Hosting Files

## CyVerse Storage Solution

<https://de.cyverse.org/> Discovery Environment (NSF funded: Free!)



The screenshot displays the CyVerse Discovery Environment interface. The top navigation bar includes the 'Discovery Environment' logo, a notification bell with '44', a user profile icon, and a help icon. The main content area is titled 'Data: brianlee' and features a menu with 'Upload', 'File', 'Edit', 'Download', 'Share', 'Metadata', and 'Refresh'. A 'Trash' button is also visible. On the left, there are three vertical buttons: 'Data', 'Apps', and 'Analyses'. The central file browser shows a navigation pane on the left with a tree view of folders: 'brianlee', 'analyses', 'testAssemblyHub', 'testAssemblyHubDaph', 'testBasicHub', 'Community Data', 'Shared With Me', 'Trash', and 'Favorites'. The main pane displays a table of files and folders under the path '/iplant/home/brianlee'. The table has columns for 'Name', 'Last Modified', and 'Size'. The files listed are:

Name	Last Modified	Size
analyses	2017 Jan 14 14:18:24	
testAssemblyHub	2017 Jan 17 11:50:00	
testAssemblyHubDaph	2017 Jan 17 12:12:59	
testBasicHub	2017 Jan 17 11:35:48	
bamExample.bam	2017 Jan 17 10:36:28	2.22 MB
bamExample.bam.bai	2017 Jan 17 10:36:18	16.77 KB
bigBedExample.bb	2017 Jan 14 14:21:48	409.9 KB

At the bottom of the interface, it indicates 'Displaying 1 - 7 of 7' and '0 item(s)'.

# Existing Hubs and Solutions to Hosting Files

## CyVerse Storage Solution

Creating a Link that Accepts Byte-Ranges: “Send to Genome Browser”

**Share  
Menu:**

**Send to  
Genome  
Browser**

The screenshot shows the CyVerse Data browser interface for a user named 'brianlee'. The interface includes a navigation pane on the left, a main file list, and a details pane on the right. A share menu is open over the file list, with the 'Send to Genome Browser' option highlighted. An arrow points from the 'Share Menu:' text to the share menu, and another arrow points from the 'Info-Type: bam' text to the 'Info-Type: bam' field in the details pane.

Name	Last Modified	Size	Actions
bamExample.bam	2017 Jan 17 10:36:28	2.22 MB	[Share] [Download] [View]
bamExample.bam.bai	2017 Jan 17 10:36:18	16.77 KB	[Share] [Download] [View]
bigBedExample.bb	2017 Jan 14 14:21:48	409.9 KB	[Share] [Download] [View]

**Details**

Last Modified: 2017 Jan 17  
Date Submitted: 2017 Jan 17  
Permissions: own  
Share: 1  
Size: 2.22 MB  
Type: application/octet-stream  
Info-Type: **bam**  
md5 Checksum: c4fdcc3...  
Send to: Genome Browser  
Tags: Search

**Info-Type:  
bam**

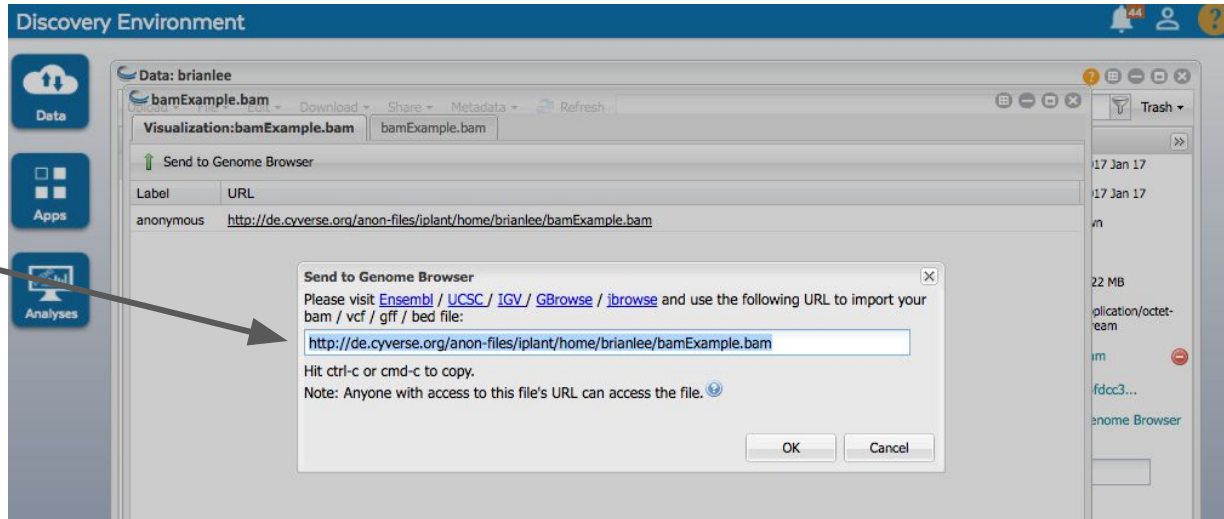


# Existing Hubs and Solutions to Hosting Files

## CyVerse Storage Solution

Creating a Link that Accepts Byte-Ranges: “Send to Genome Browser”

Results in a link you can use in your *bigDataUrl*



<http://de.cyverse.org/anon-files/iplant/home/brianlee/bamExample.bam>



# Existing Hubs and Solutions to Hosting Files

## CyVerse Storage Solution

*For text files (hub.txt...) use Public Link*

For **all** the text files (that don't need to involve byte-range requests) you can use the **“Create a Public Link”** option (such “shared” links would work from Dropbox or GoogleDrive as well as text are fully downloaded to the Browser).

The screenshot shows the CyVerse Data browser interface. The main window displays a file list for the folder 'testAssemblyHubDaph'. The file 'hub.txt' is selected, and a context menu is open over it, with the 'Create a Public Link...' option highlighted. The details pane on the right shows the following information for 'hub.txt':

Property	Value
Last Modified:	2017 Jan 17
Date Submitted:	2017 Jan 17
Permissions:	own
Share:	<a href="#">Begin sharing</a>
Size:	181 bytes
Type:	text/plain
Info-Type:	<a href="#">Select</a>
md5 Checksum:	101332b...
Send to:	-
Tags:	<input type="text" value="Search"/>

# Existing Hubs and Solutions to Hosting Files

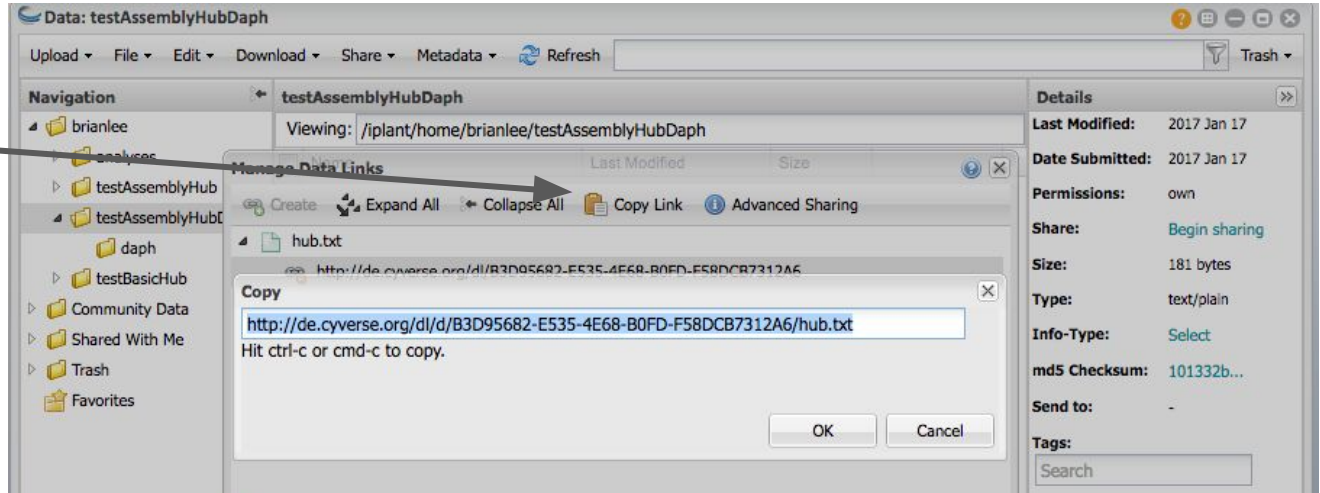
## CyVerse Storage Solution

*For text files (hub.txt...) use Public Link*

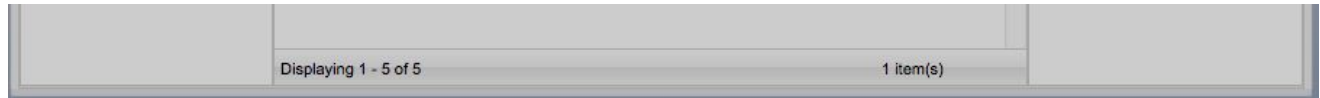
“Create a Public Link” and Copy Link.

Use these direct links in the hub.txt --to--> point to genomes.txt

And in genomes.txt --to--> point to trackDb.txt and groups.txt



<https://de.cyverse.org/dl/d/B3D95682-E535-4E68-B0FD-F58DCB7312A6/hub.txt>



# Existing Hubs and Solutions to Hosting Files

## CyVerse Storage Solution

**Trick for 2bit file:** Call it a “bed” file to allow “Send To Browser” to get link.

Name 2bit a “bed” file to allow CyVerse to give you the ability to obtain the link via the **Share** menu.

The screenshot shows the CyVerse file manager interface. The main pane displays a list of files in the directory `/iplant/home/brianlee/testAssemblyHubDaph/daph`. The selected file is `daph.2bit.not.bed`, which is 47.98 MB in size and was last modified on 2017 Jan 17 at 12:19:14. The details pane on the right shows the following information:

Property	Value
Last Modified:	2017 Jan 17
Date Submitted:	2017 Jan 17
Permissions:	own
Share:	1
Size:	47.98 MB
Type:	application/octet-stream
Info-Type:	bed
md5 Checksum:	fddd953...
Send to:	Genome Browser
Tags:	

Give it **Info-Type:** Bed

# Existing Hubs and Solutions to Hosting Files

## CyVerse Storage Solution

```
hub.txt
hub daph
shortLabel Cyverse daph hub
longLabel Cyverse Daph Hub
genomesFile http://de.cyverse.org/dl/d/0E4881D6-2D18-4C5E-BE5B-F7D7138B0043/genomes.txt
email genome-www@soe.ucsc.edu

genomes.txt
genome daph
trackDb http://de.cyverse.org/dl/d/00C5EB6A-0875-451B-B6F5-A84ED8390617/trackDb.txt
groups http://de.cyverse.org/dl/d/F856C7C7-CD5B-4794-98D6-E243FA0CB188/groups.txt
description Daph Cyverse 2bit step requires trick
twoBitPath https://de.cyverse.org/anon-files/iplant/home/brianlee/testAssemblyHubDaph/daph/daph.2bit.not.bed
organism Daphnia pulex
defaultPos scaffold_1:10000-20000
orderKey 1000
scientificName Daphnia pulex
htmlPath http://de.cyverse.org/dl/d/2260C756-CCEE-4A3E-ACCA-8955804AA064/localFile.html

trackDb.txt
track myTrack
longLabel shortMatch of AAAAAA Data hosted on Cyverse, needs "Type:" bed and link Share menu "Send to Genome Browser""
shortLabel Motifs
priority 10
visibility pack
color 0,120,0
bigDataUrl http://de.cyverse.org/anon-files/iplant/home/brianlee/testAssemblyHubDaph/daph/daph.bb
type bigBed 5
group varRep
```

**Binary Data links generated from using**

**Share: Send to Genome Browser**

**For text and html files links generated from using**

**Share: Create a Public Link**





Thank You!



UNIVERSITY OF CALIFORNIA  
**SANTA CRUZ** | Genomics  
Institute

