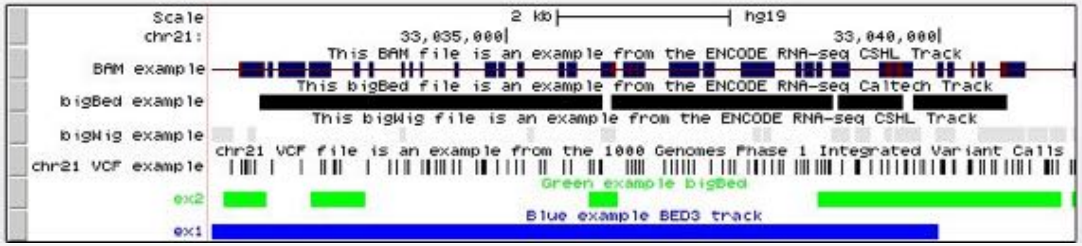
1. Custom Annotations in the Browser



This BAM file is an example from the ENCODE RNA-seq CSHL Track

This bigBed file is an example from the ENCODE RNA-seq Caltech Track

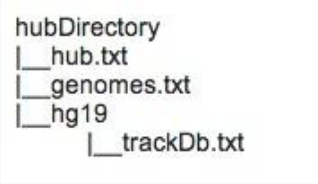
This bigWig file is an example from the ENCODE RNA-seq CSHL Track

chr21 VCF file is an example from the 1000 Genomes Phase 1 Integrated Variant Calls

Green example bigBed

Blue example BED3 track

2. Track Hubs



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

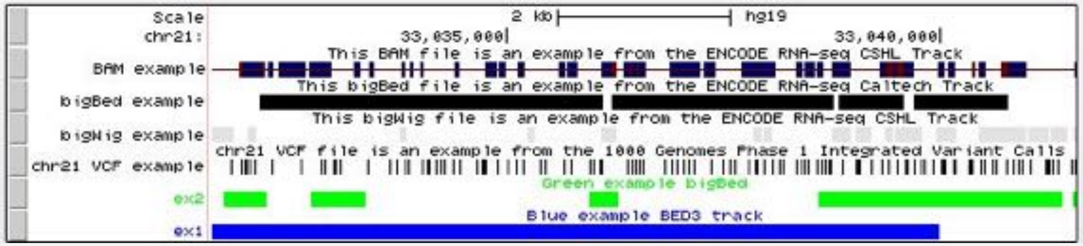
3. Assembly Hubs

Loading Your Own Data

- Custom Tracks
  - plain-text files, many formats
  - loads entire file: copy/paste or url
- Track Hubs
  - indexed binary files
  - must be hosted remotely
  - loads only what is displayed
- Assembly Hubs
  - special case of Track Hubs
  - used for genomes not hosted at UCSC
  - requires underlying sequence file in .2bit format

1. Custom Annotations in the Browser

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



This BAM file is an example from the ENCODE RNA-seq CSHL Track

This bigBed file is an example from the ENCODE RNA-seq Caltech Track

This bigWig file is an example from the ENCODE RNA-seq CSHL Track

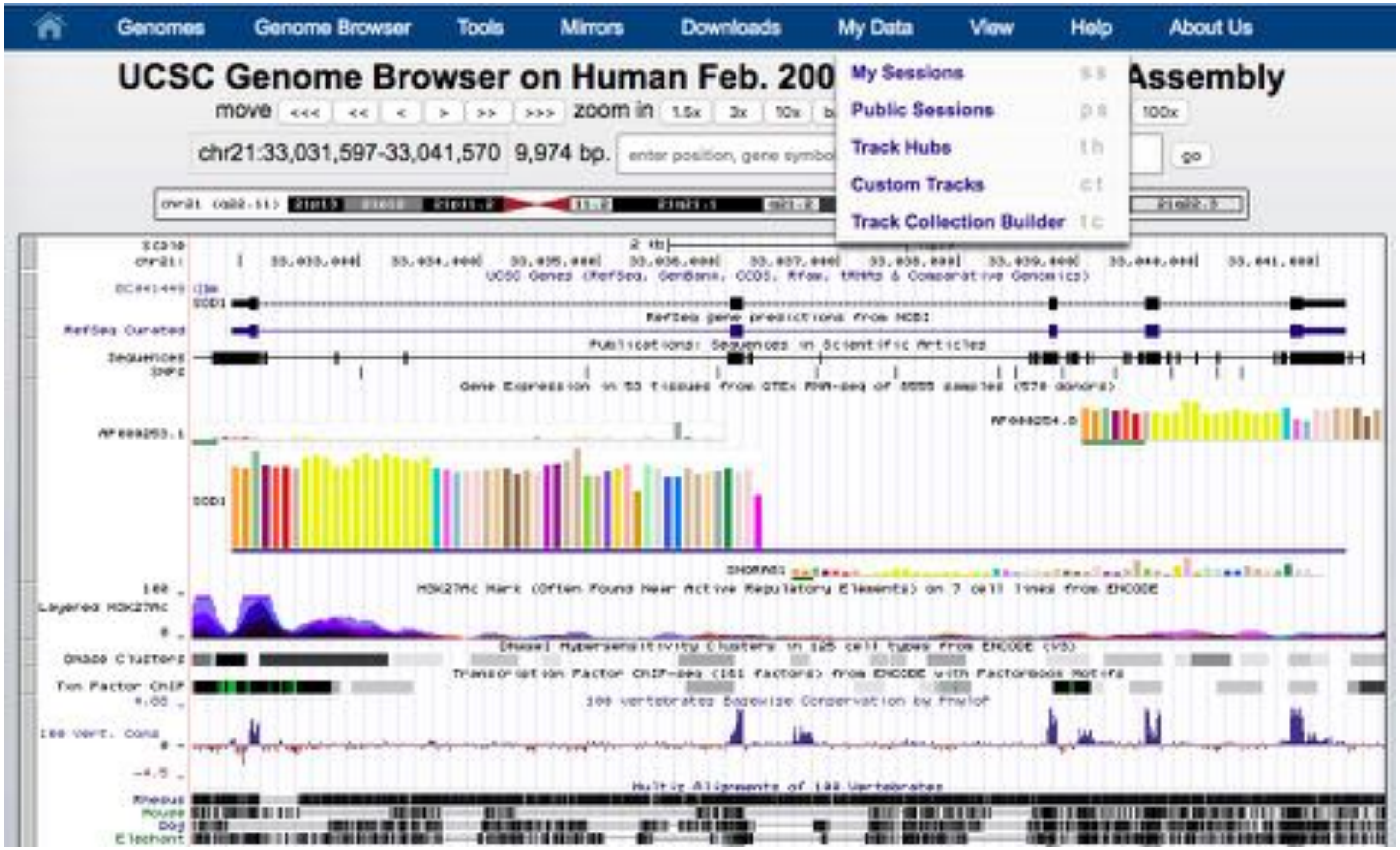
chr21 VCF file is an example from the 1000 Genomes Phase 1 Integrated Variant Calls

Green example bigBed

Blue example BED3 track

UCSC Genome Browser

My Data **Top Blue Bar Menu**



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.

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.

Underlying Assemblies
Hosted by UCSC

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

Add Custom Tracks

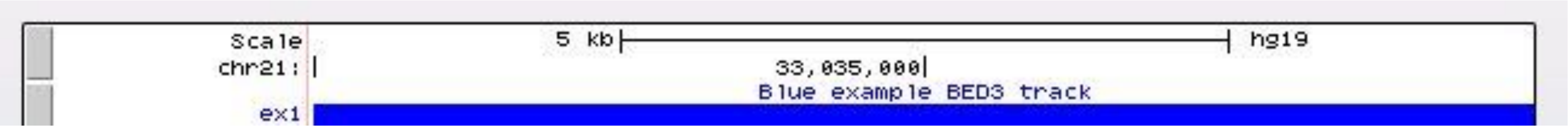
clade Mammal   genome Human   assembly Feb. 2009 (GRC

Display your own data as custom annotation tracks in the browser. Data must be f 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:

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,

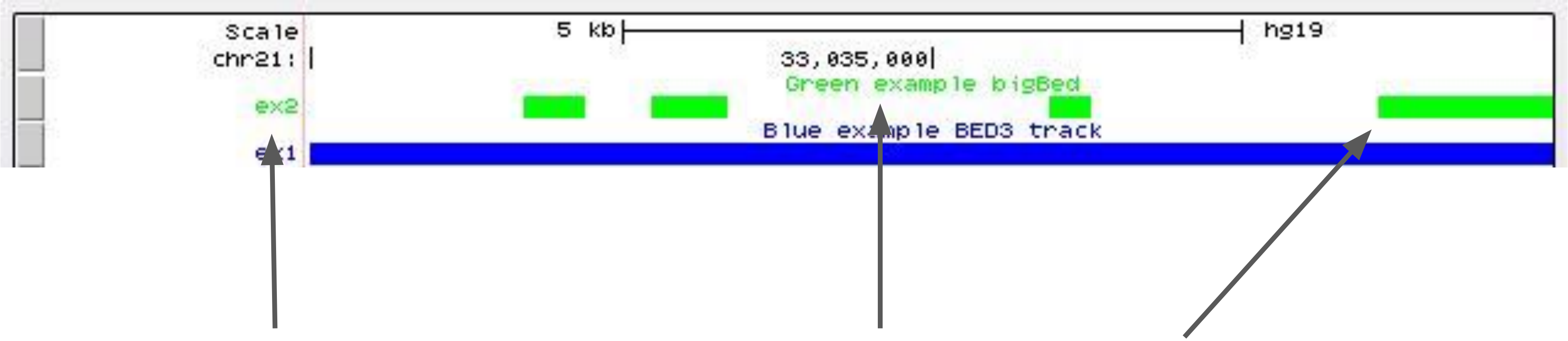
chr21 33030000 33040000

chr21 33050000 33060000

...

Custom Annotations in the Browser

Custom Tracks (Text/**Binary**) Track Hubs Assembly Hubs

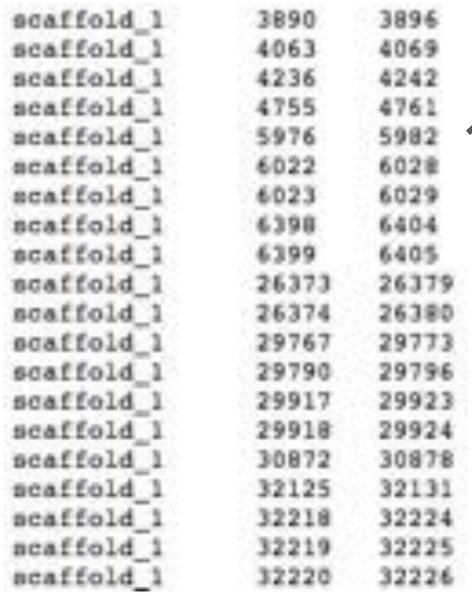


track name=ex2 description="Green example bigBed" color=0,255,0, **type=bigBed** ***bigDataUrl=****http://genome.ucsc.edu/goldenPath/help/examples/****bigBedExample.bb***

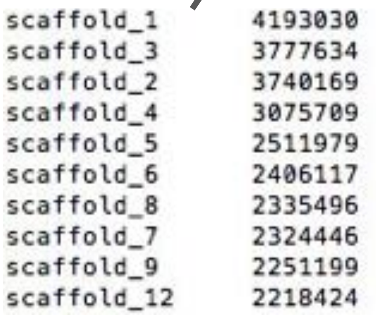
Custom Annotations in the Browser

Custom Tracks (Text/**<u>Binary</u>**)   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:

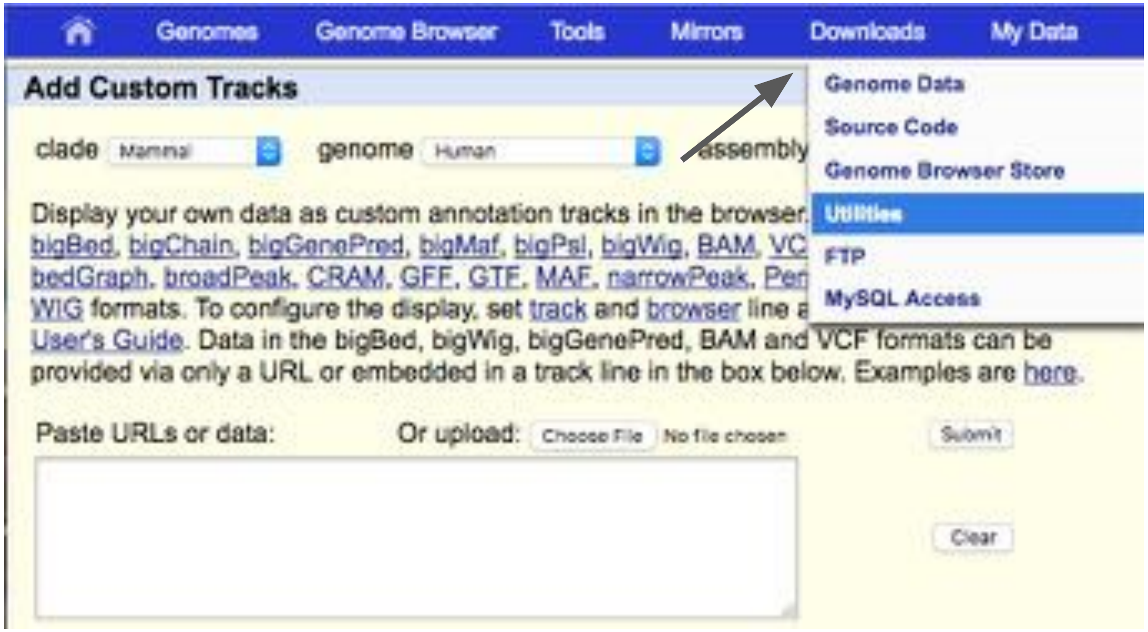
*bigDataUrl=http://path.lab.edu/to/**out.bb***

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.**



Add Custom Tracks

clade Mammal   genome Human   assembly

Display your own data as custom annotation tracks in the browser. bigBed, bigChain, bigGenePred, bigMaf, bigPsl, bigWig, BAM, VC bedGraph, broadPeak, CRAM, GFF, GTF, MAF, narrowPeak, Per WIG formats. To configure the display, set track and browser line a 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.

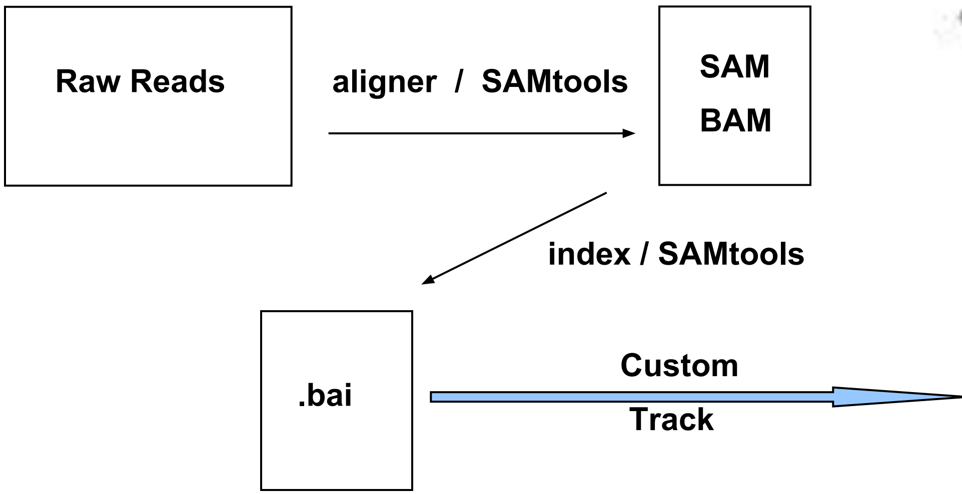
Paste URLs or data: Or upload:

Sequence Data

How do you visualize your sequencing data? -- BAM file

millions of short reads

files too large to upload (timeout)



Raw Reads

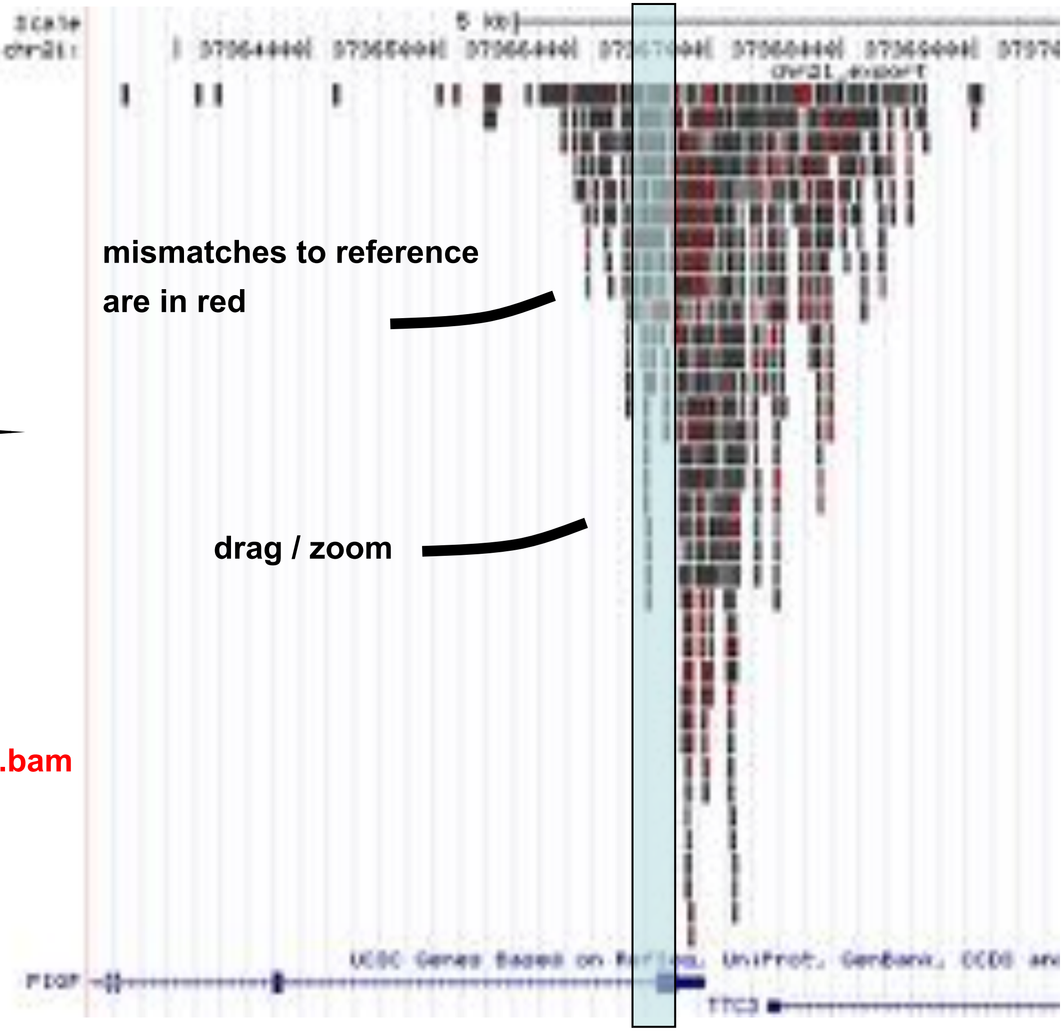
aligner / SAMtools

SAM
BAM

index / SAMtools

.bai

Custom Track



mismatches to reference are in red

drag / zoom

Custom Track:

Make BAM, .bai files available to the web (http:, https: or ftp:)

Upload only the *location* of the data.

track name=trackName type=bam bigDataUrl=http://path/file.bam

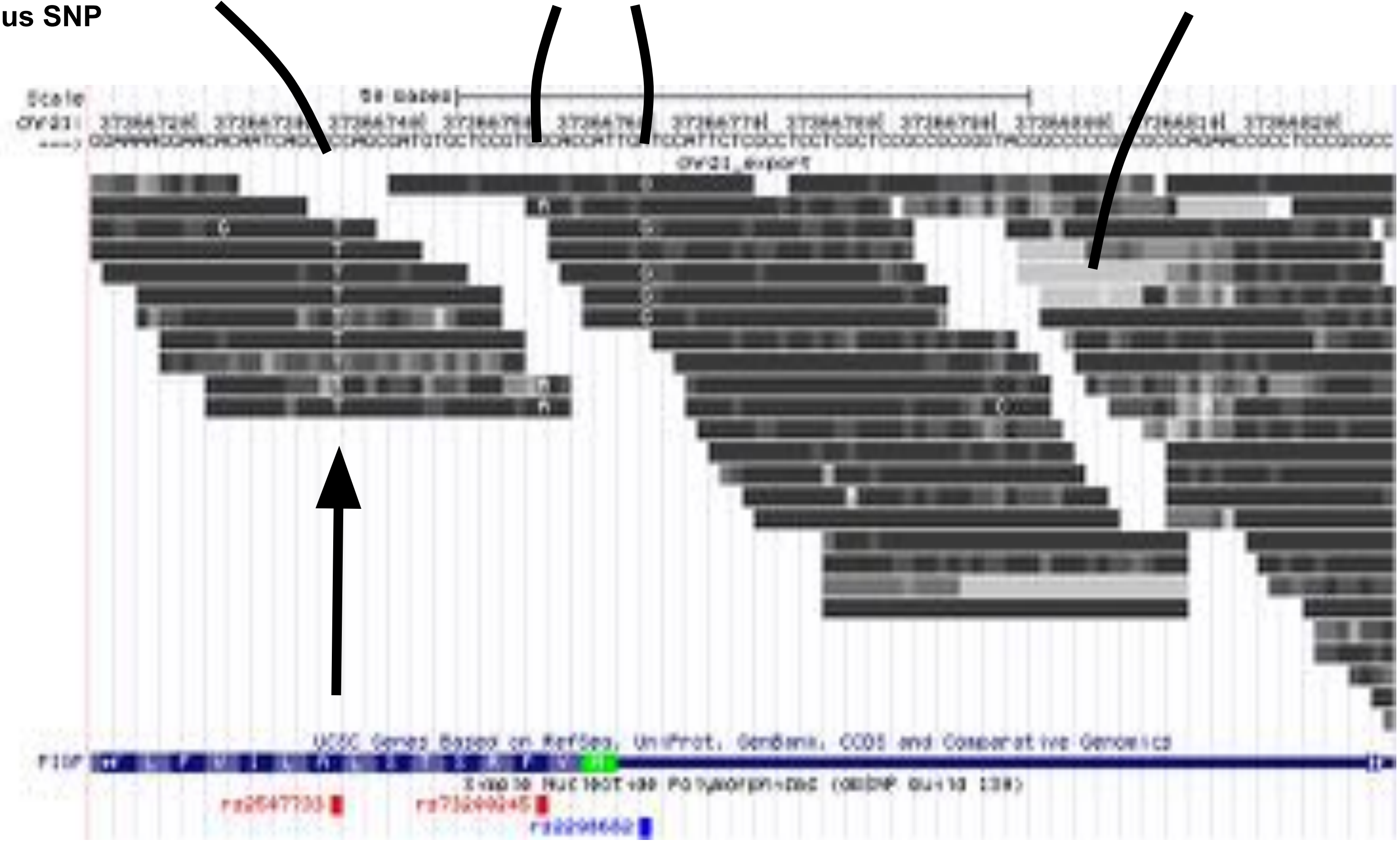
The Browser fetches only tiny portion of the file

http://samtools.sourceforge.net/

Sequence Data

zoom to base level

view alignment details



homozygous mismatch to reference is same as a known non-synonymous SNP

possible heterozygotes

lower quality scores shown in lighter color

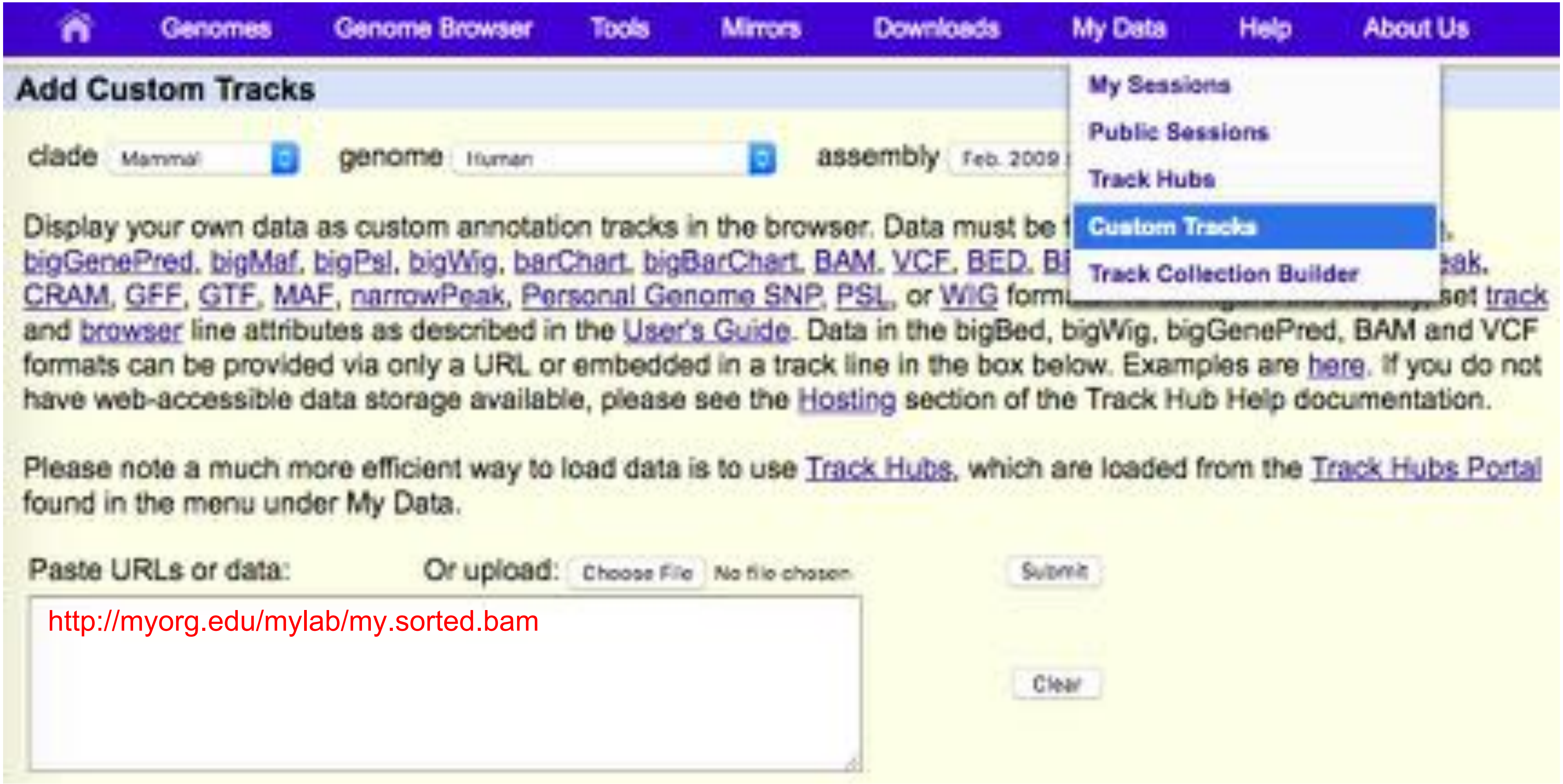
UCSC genes

SNP track

Sequence Data

How do you visualize your sequencing data? -- BAM file

Under "My Data" use the "Custom Tracks" selection and paste URL to fileName.bam (where there is also fileName.bam.bai)



Add Custom Tracks

clade Mammal genome Human assembly Feb. 2009

Display your own data as custom annotation tracks in the browser. Data must be f bigGenePred, bigMaf, bigPsl, bigWig, barChart, bigBarChart, BAM, VCF, BED, BI CRAM, GFF, GTF, MAF, narrowPeak, Personal Genome SNP, PSL, or WIG form 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. If you do not have web-accessible data storage available, please see the Hosting section of the Track Hub Help documentation.

Please note a much more efficient way to load data is to use Track Hubs, which are loaded from the Track Hubs Portal found in the menu under My Data.

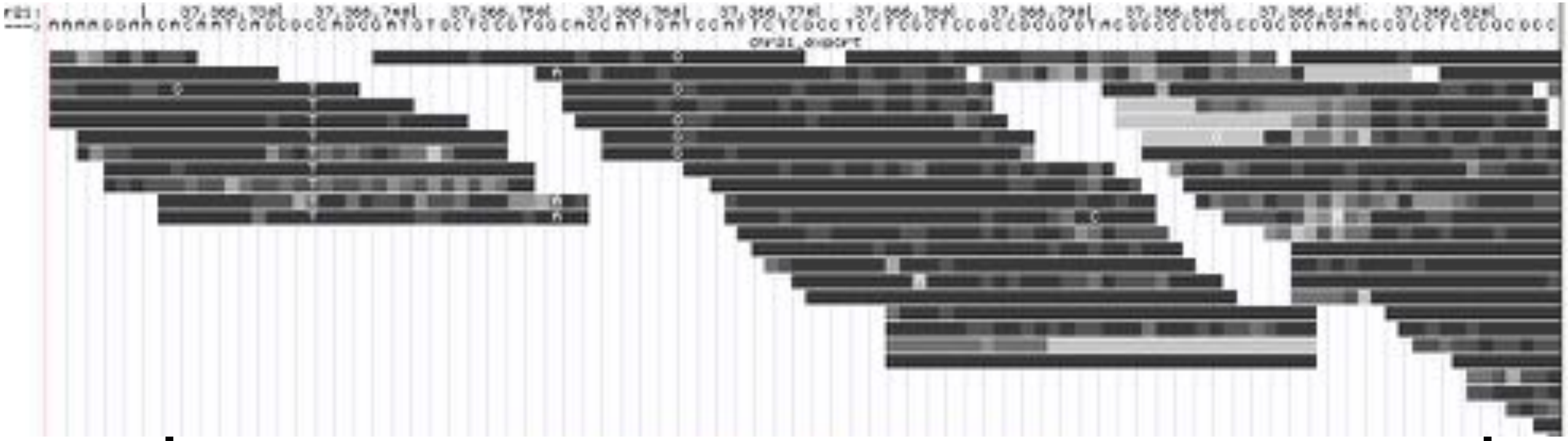
Paste URLs or data: Or upload: Choose File No file chosen

http://myorg.edu/mylab/my.sorted.bam

Sequence Data

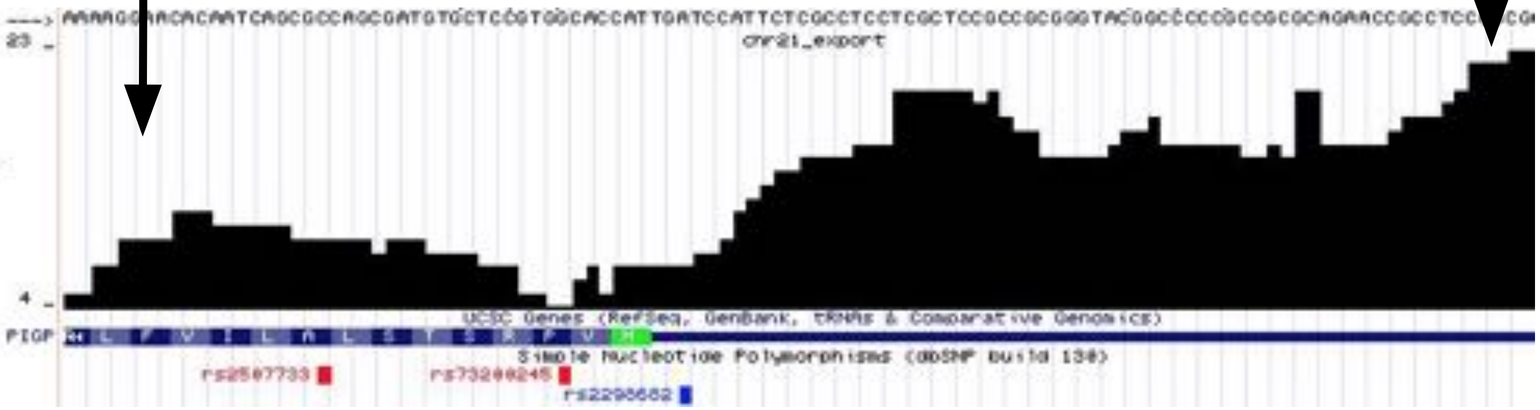
Option to “*Display data as a density graph*” (right-click, or go to track controls)

Display data as a density graph: ☑



Display data as a “wiggle” by stacking overlapping reads with density graph.

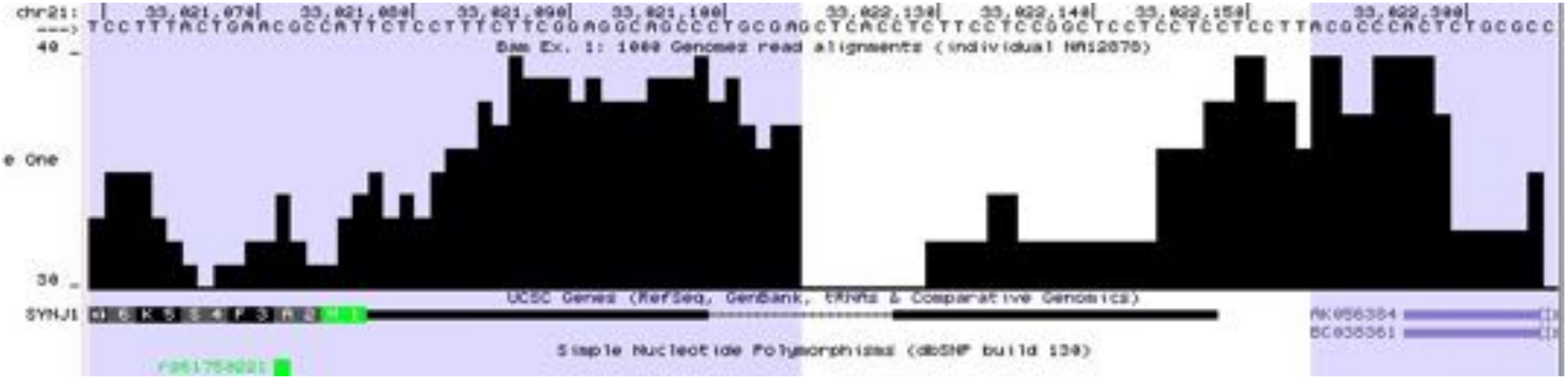
Height is proportional to the number of reads mapped to each genomic position.



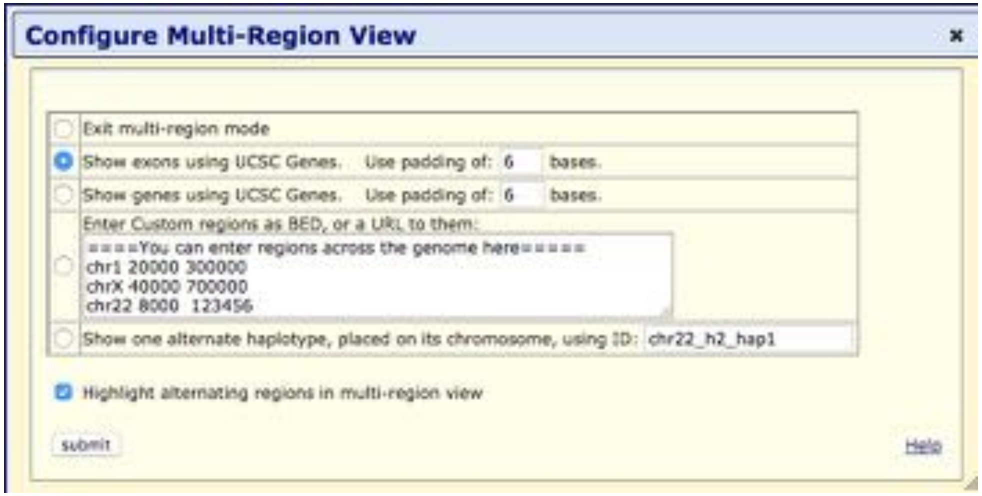
Sequence Data

Option to view exons only (with small amount of padding) or enter user-defined custom regions.

Enter "exon view" with shortcut "e v" (then use "d v" to return to default view).



Or use "View" menu and select "Multi-Region" and selected desired view.



Configure Multi-Region View

○ Exit multi-region mode
● Show exons using UCSC Genes. Use padding of: 6 bases.
○ Show genes using UCSC Genes. Use padding of: 6 bases.
○ Enter Custom regions as BED, or a URL to them:

```
====You can enter regions across the genome here=====
chr1 20000 300000
chrX 40000 700000
chr22 8000  123456
```

○ Show one alternate haplotype, placed on its chromosome, using ID: chr22_h2_hap1

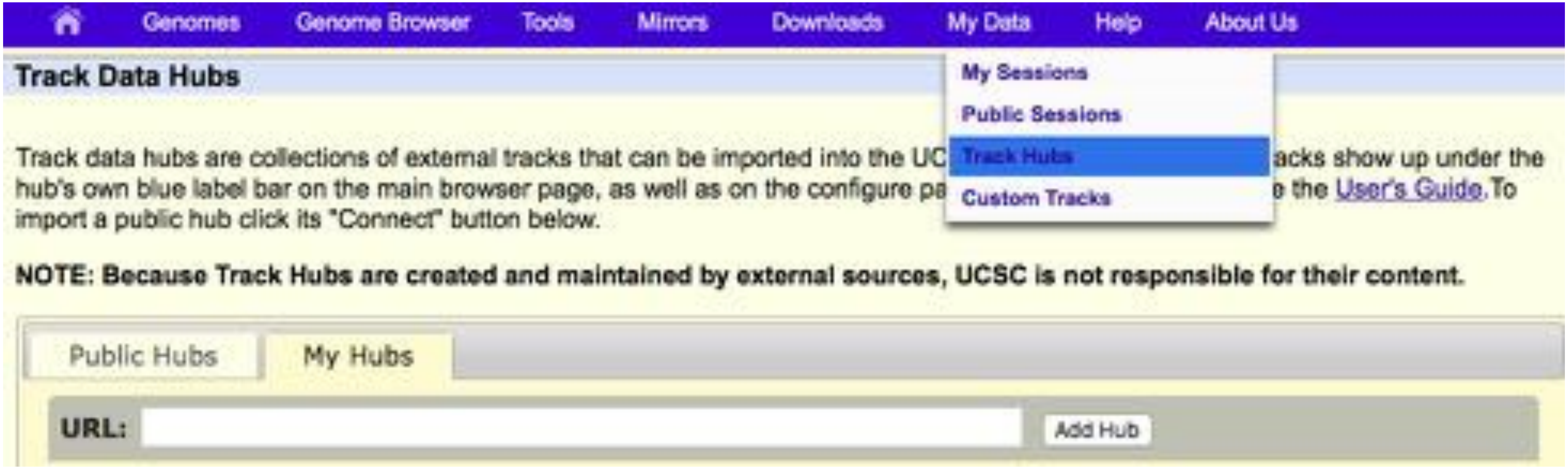
☑ Highlight alternating regions in multi-region view

Optional checkbox can *highlight alternating regions* (seen in above image).

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 UC acks show up under the hub's own blue label bar on the main browser page, as well as on the configure pa e the User's Guide. 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.

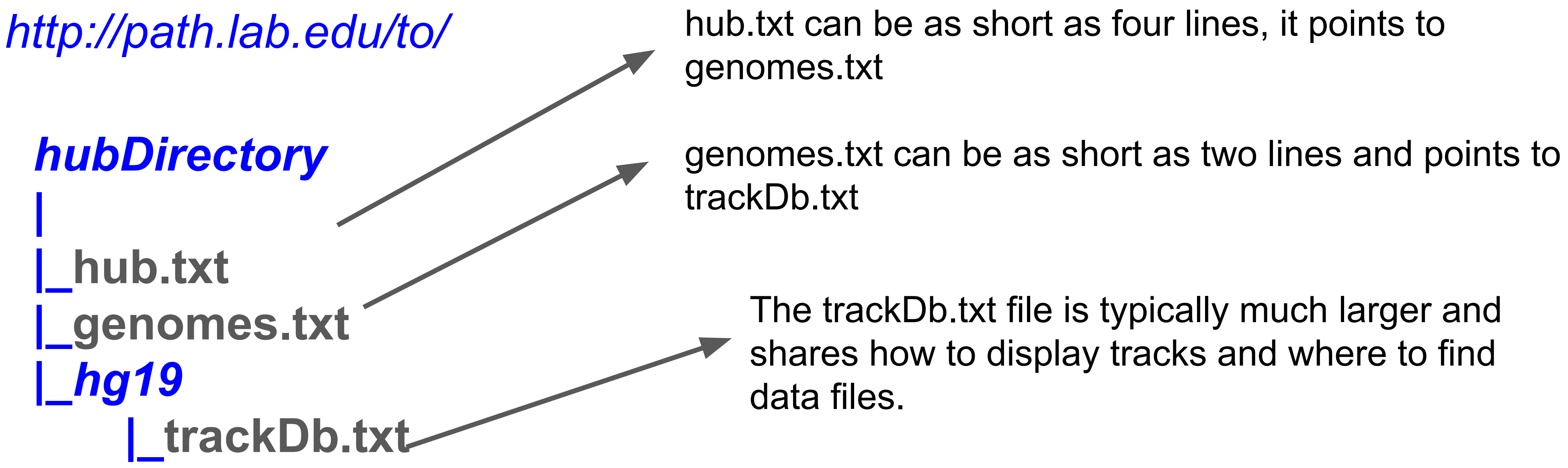
URL:

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

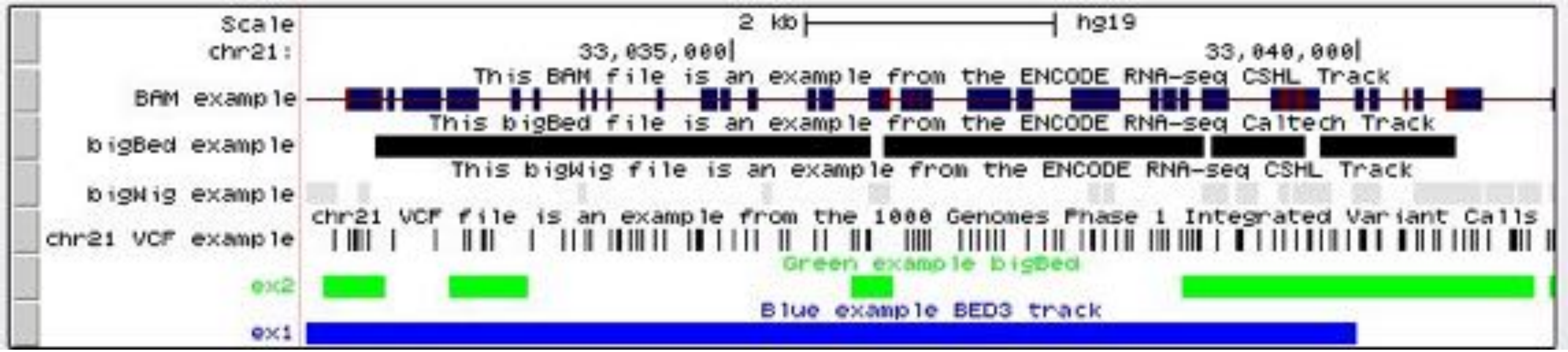
genomes.txt can be as short as two lines and points to trackDb.txt

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

Custom Annotations in the Browser

Custom Tracks (Text/Binary)   **Track Hubs**   Assembly Hubs

Track hubs are web-accessible directories of genomic data.



This BAM file is an example from the ENCODE RNA-seq CSHL Track

This bigBed file is an example from the ENCODE RNA-seq Caltech Track

This bigWig file is an example from the ENCODE RNA-seq CSHL Track

chr21 VCF file is an example from the 1000 Genomes Phase 1 Integrated Variant Calls

Green example bigBed

Blue example BED3 track

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

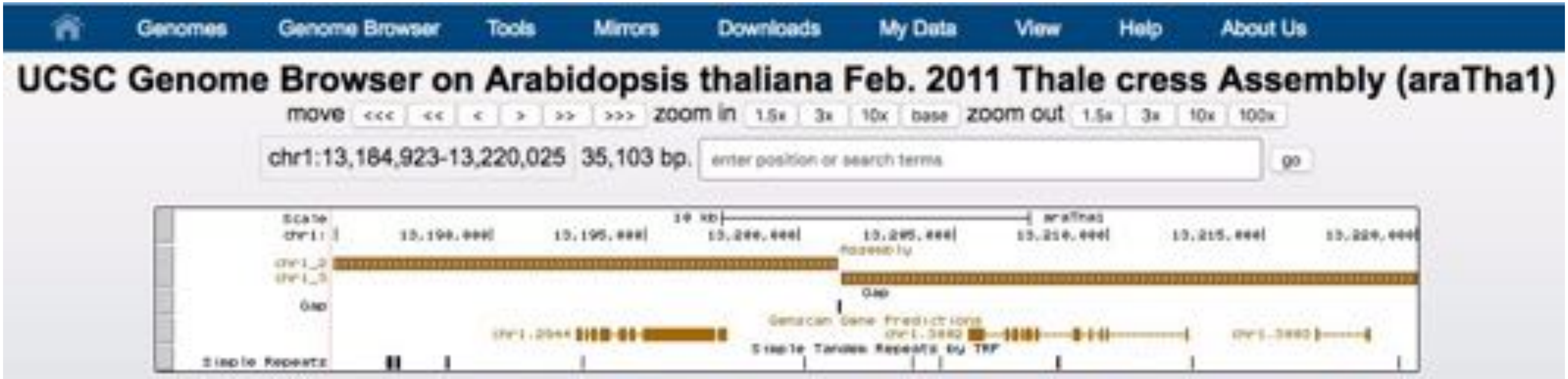
```
track uniqueNameNoSpacesOrDots
type track_type
shortLabel label 17 chars
longLabel   label up to 80 chars
bigDataUrl track_data_url (URL or relative path to file)

track bam1
type bam
shortLabel BAM example
longLabel This BAM file is an example of RNA-seq data
visibility dense
bigDataUrl http://fileserver/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**



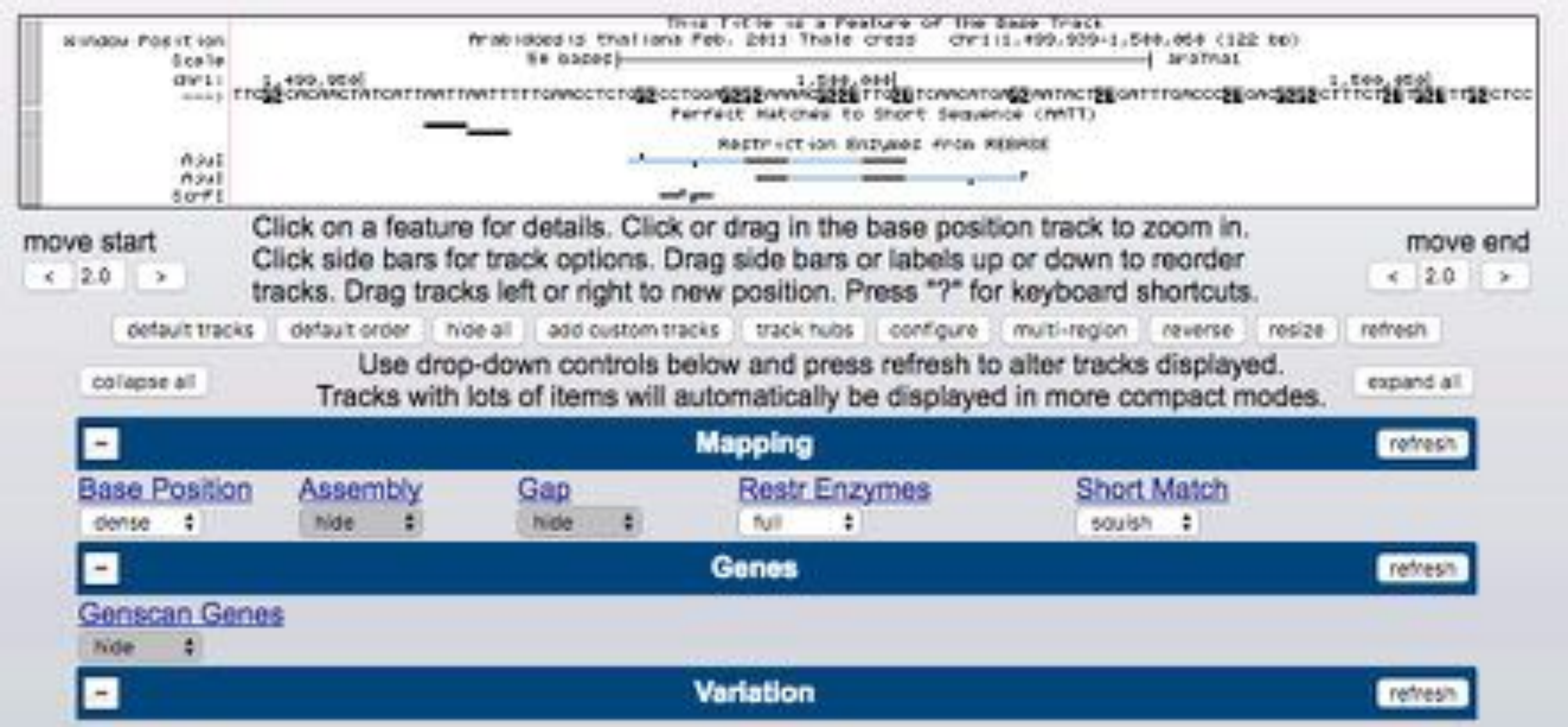
UCSC Genome Browser on Arabidopsis thaliana Feb. 2011 Thale cress Assembly (araTha1)

chr1:13,184,923-13,220,025 35,103 bp.

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/<u>Binary</u>)　　Track Hubs　　**Assembly Hubs**



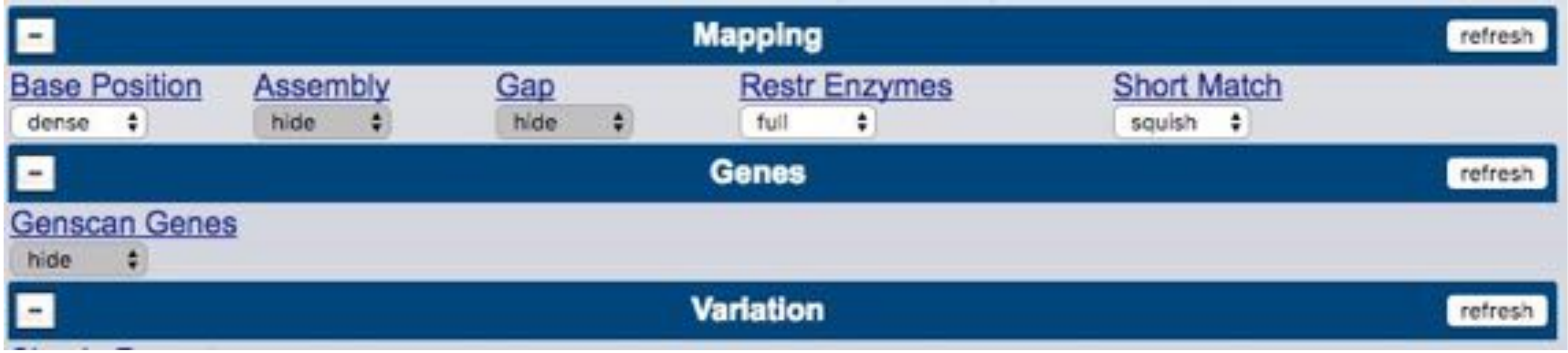
Click on a feature for details. Click or drag in the base position track to zoom in. Click side bars for track options. Drag side bars or labels up or down to reorder tracks. Drag tracks left or right to new position. Press "?" for keyboard shortcuts.

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.

Base Position: dense
Assembly: hide
Gap: hide
Restr Enzymes: full
Short Match: squish
Genscan Genes: hide

Custom Annotations in the Browser

Custom Tracks (Text/Binary) Track Hubs **Assembly Hubs**

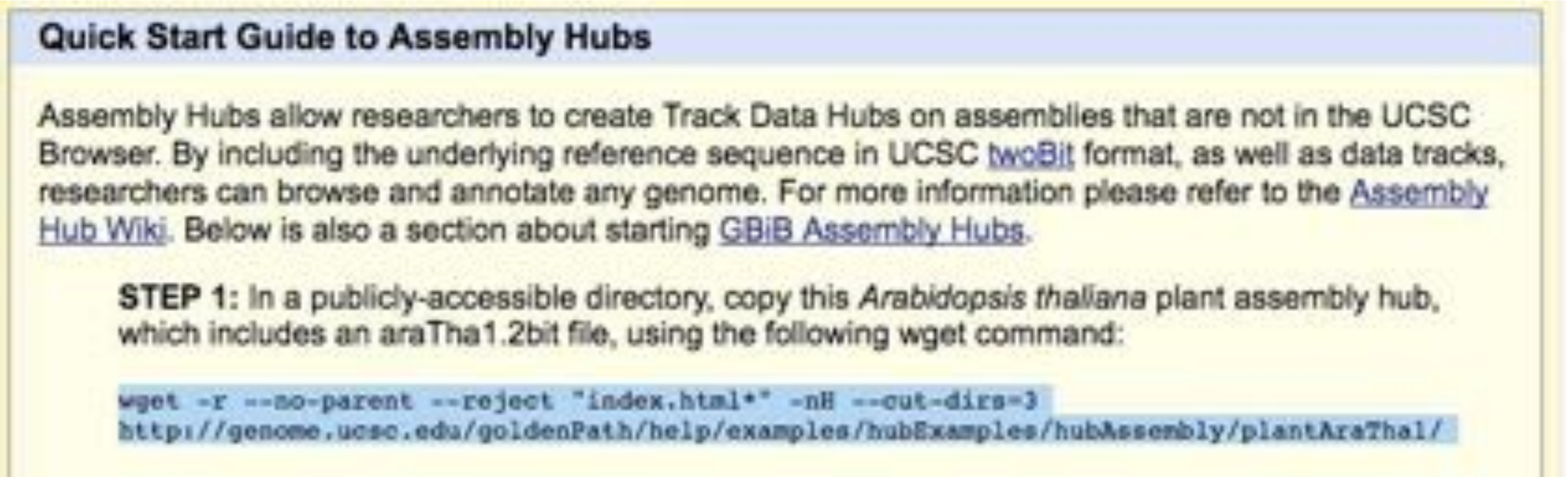


The groups.txt file defines the blue bars grouping tracks in assembly hubs.

Custom Annotations in the Browser

Custom Tracks (Text/Binary) Track Hubs **Assembly Hubs**

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



Quick Start Guide to Assembly Hubs

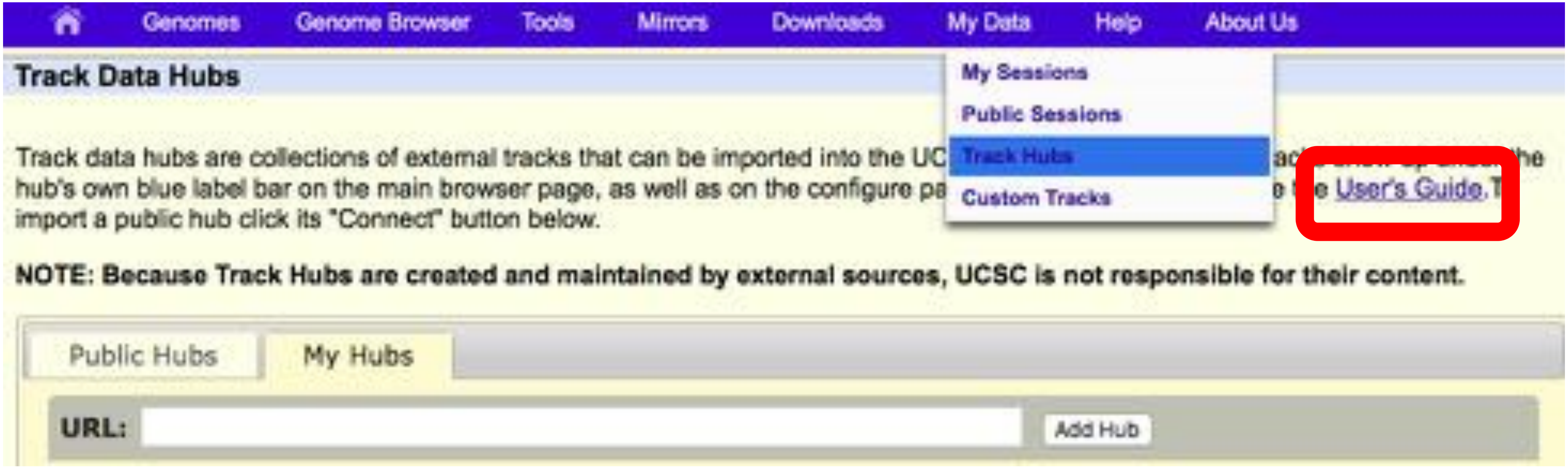
Assembly Hubs allow researchers to create Track Data Hubs on assemblies that are not in the UCSC Browser. By including the underlying reference sequence in UCSC twoBit format, as well as data tracks, researchers can browse and annotate any genome. For more information please refer to the Assembly Hub 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/
```

Custom Annotations in the Browser

Custom Tracks (Text/Binary)   Track Hubs   **Assembly Hubs**



Track Data Hubs

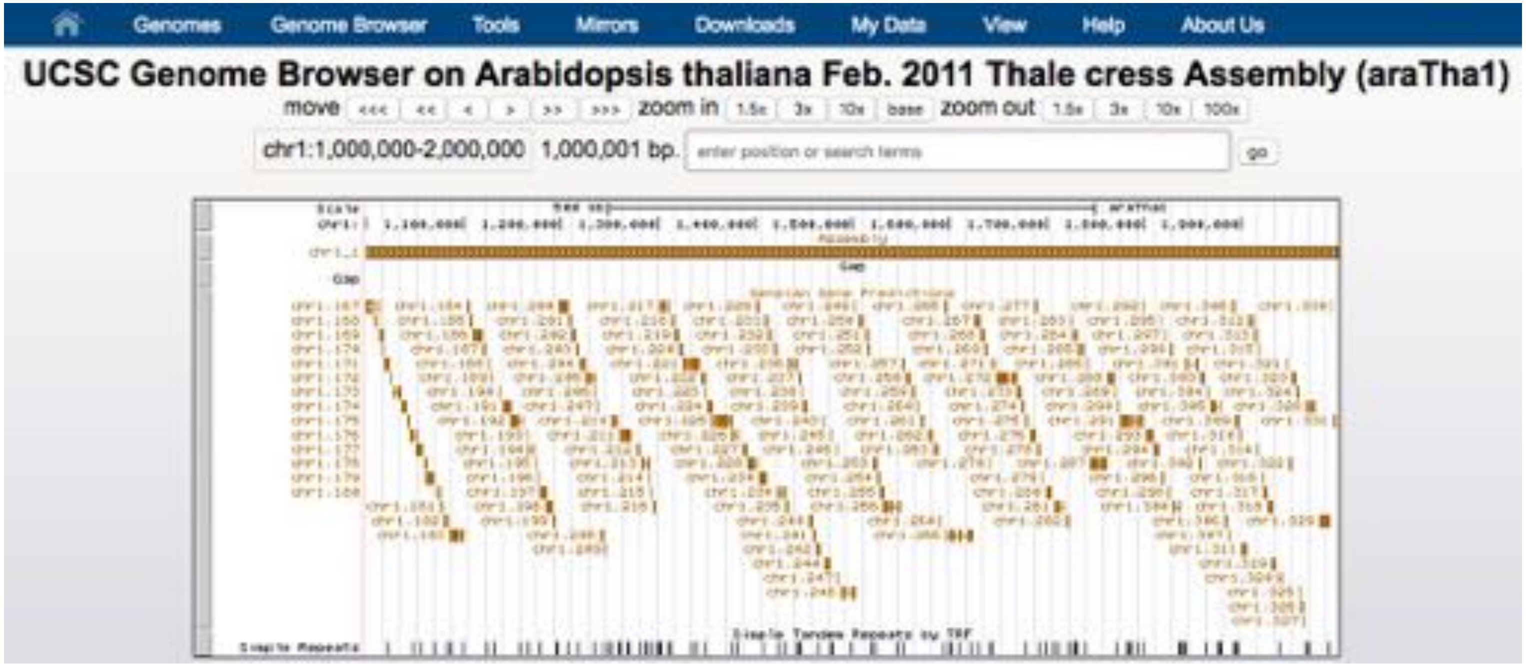
Track data hubs are collections of external tracks that can be imported into the UC ac he hub's own blue label bar on the main browser page, as well as on the configure pa e e User's Guide. T import a public hub click its "Connect" button below.

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

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

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)

chr1:1,000,000-2,000,000 1,000,001 bp.

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

UCSC Genome Browser

Home Page: Genome Browser in a Box (GBiB)

Our tools

- **Genome Browser**
  interactively visualize genomic data
- **BLAT**
  rapidly align sequences to the genome
- **Table Browser**
  download data from the Genome Browser database
- **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
- **Genome Browser in a Box (GBiB)**
  run the Genome Browser on your laptop or server
- **In-Silico PCR**
  rapidly align PCR primer pairs to the genome
- **LiftOver**
  convert genome coordinates between assemblies
- **Track Hubs**
  import and view external data tracks

Genome Browser in a Box (GBiB)

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

Genome Browser in a Box (GBiB)
is a Virtual Machine of Browser Code at your location.

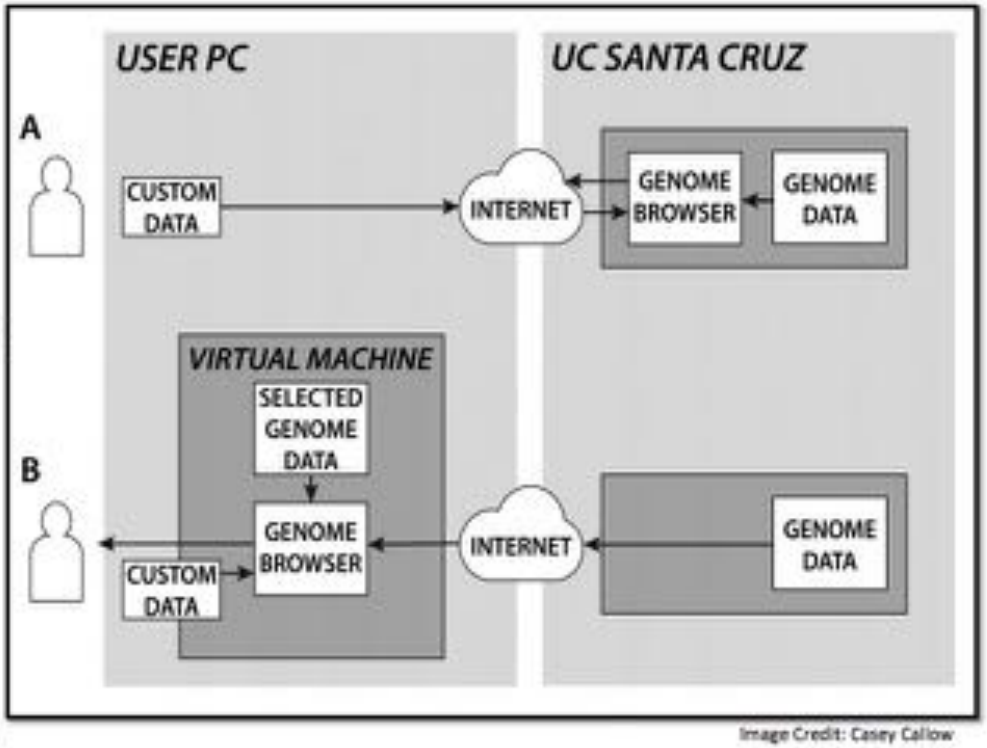
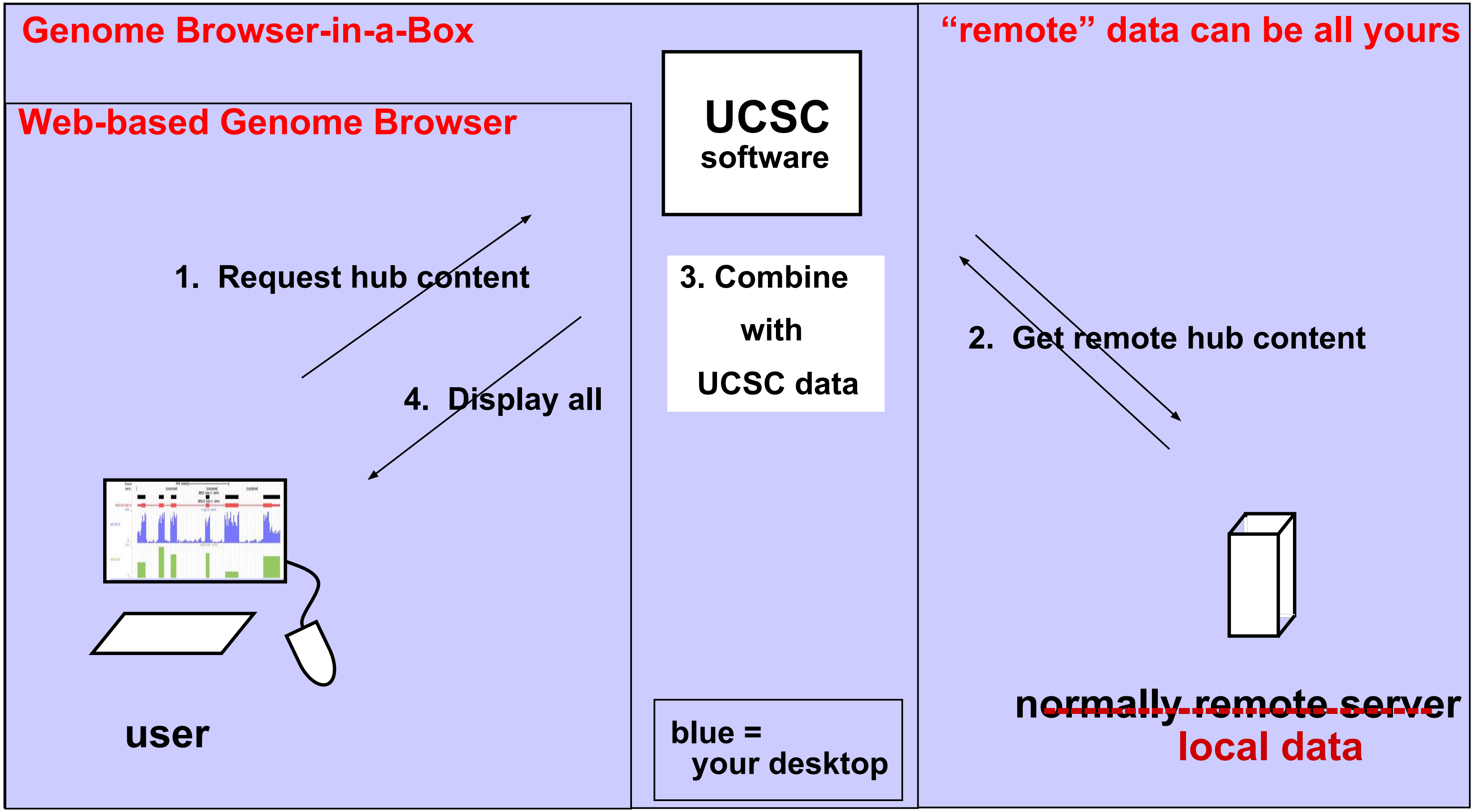


Image Credit: Casey Callow

Track data hubs

Require no intervention by UCSC

Can be hosted and viewed locally on a Genome Browser-in-a-Box (GBiB)



Genome Browser-in-a-Box

Web-based Genome Browser

UCSC software

1. Request hub content

2. Get remote hub content

3. Combine with UCSC data

4. Display all

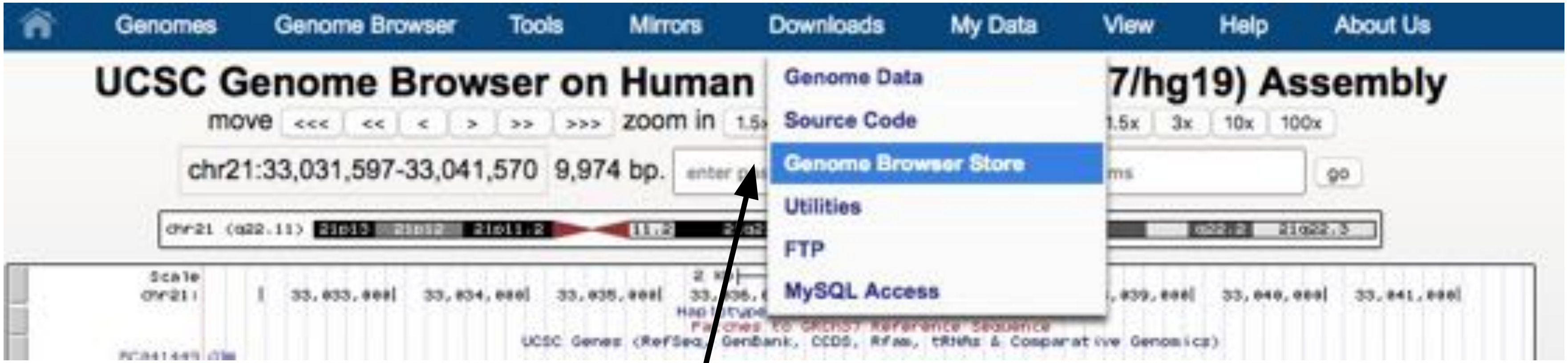
user

"remote" data can be all yours

~~normally remote server~~
local data

blue = your desktop

UCSC Genome Browser



Downloads **in Top Blue Bar Menu:** Genome Browser Store **links to Genome Browser in a Box** (GBiB) **and Genome Browser in a Cloud** (GBiC) **tools free for noncommercial use.**

Also on the bottom of Document Pages: STORE **link goes to the same location.**

Custom Annotations in the Browser

Custom Tracks (Text/**Binary**) **Track Hubs** **Assembly Hubs**

https://de.cyverse.org/ Discovery Environoment (NSF funded: **FREE!**)

Discovery Environment

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

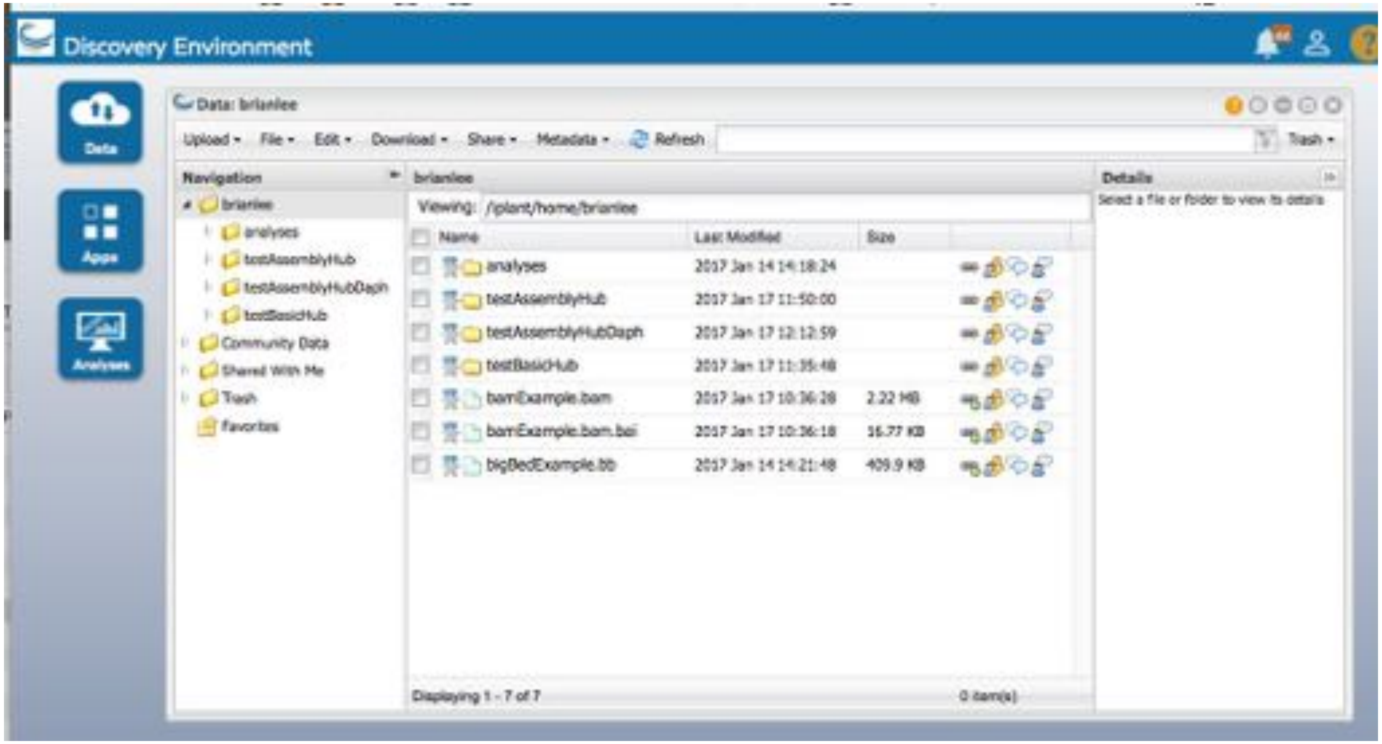
- Makes big data management easy. Upload, organize, edit, view and search with ease!
- Has 500+ scientific apps that utilize compute clusters and HPC resources as needed.
- Hides the complexity needed to do these tasks.

Minimum screen resolution supported: 1024 x 768

Custom Annotations in the Browser

Custom Tracks (Text/**Binary**) **Track Hubs** **Assembly Hubs**

https://de.cyverse.org/ Discovery Environoment (NSF funded: **FREE!**)



brianlee

Viewing: /iplant/home/brianlee

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

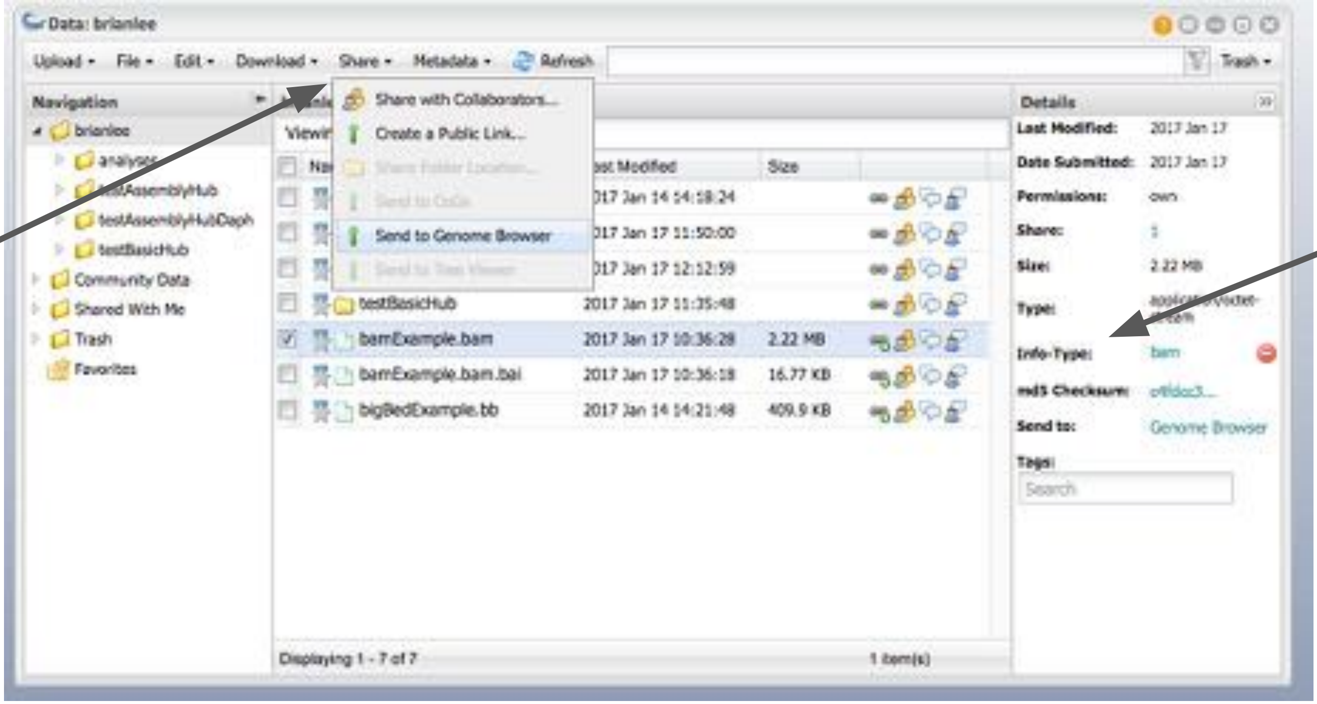
Displaying 1 - 7 of 7 0 item(s)

Custom Annotations in the Browser

Custom Tracks (Text/**Binary**) **Track Hubs** **Assembly Hubs**

https://de.cyverse.org/ Discovery Environoment (NSF funded: **FREE!**)

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



Na	st Modified	Size
	017 Jan 14 14:18:24	
	017 Jan 17 11:50:00	
	017 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

Displaying 1 - 7 of 7 1 item(s)

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: c4fdac3...
Send to: Genome Browser
Tags: Search

Share Menu:

Send to Genome Browser

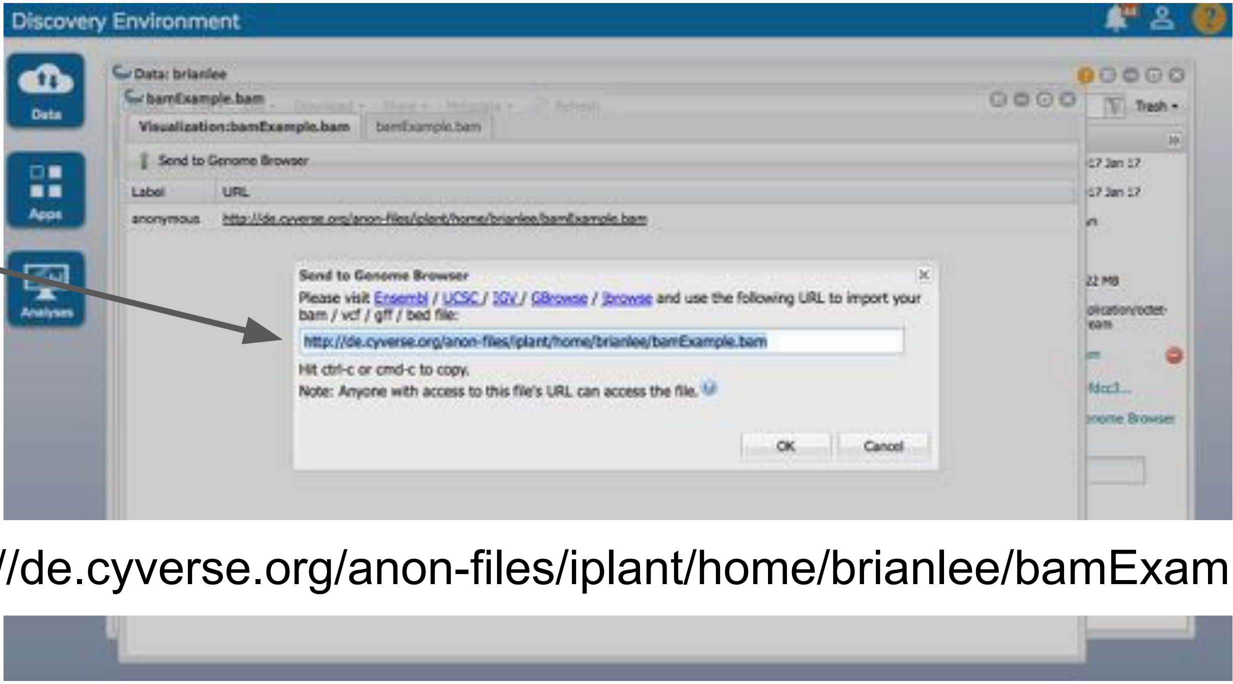
Info-Type: bam

Custom Annotations in the Browser

Custom Tracks (Text/**Binary**) **Track Hubs** **Assembly Hubs**

https://de.cyverse.org/ Discovery Environoment (NSF funded: **FREE!**)

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



Send to Genome Browser

Please visit Ensembl / UCSC / IGV / GBrowse / Jbrowse and use the following URL to import your bam / vcf / gff / bed file:

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

Hit ctrl-c or cmd-c to copy.

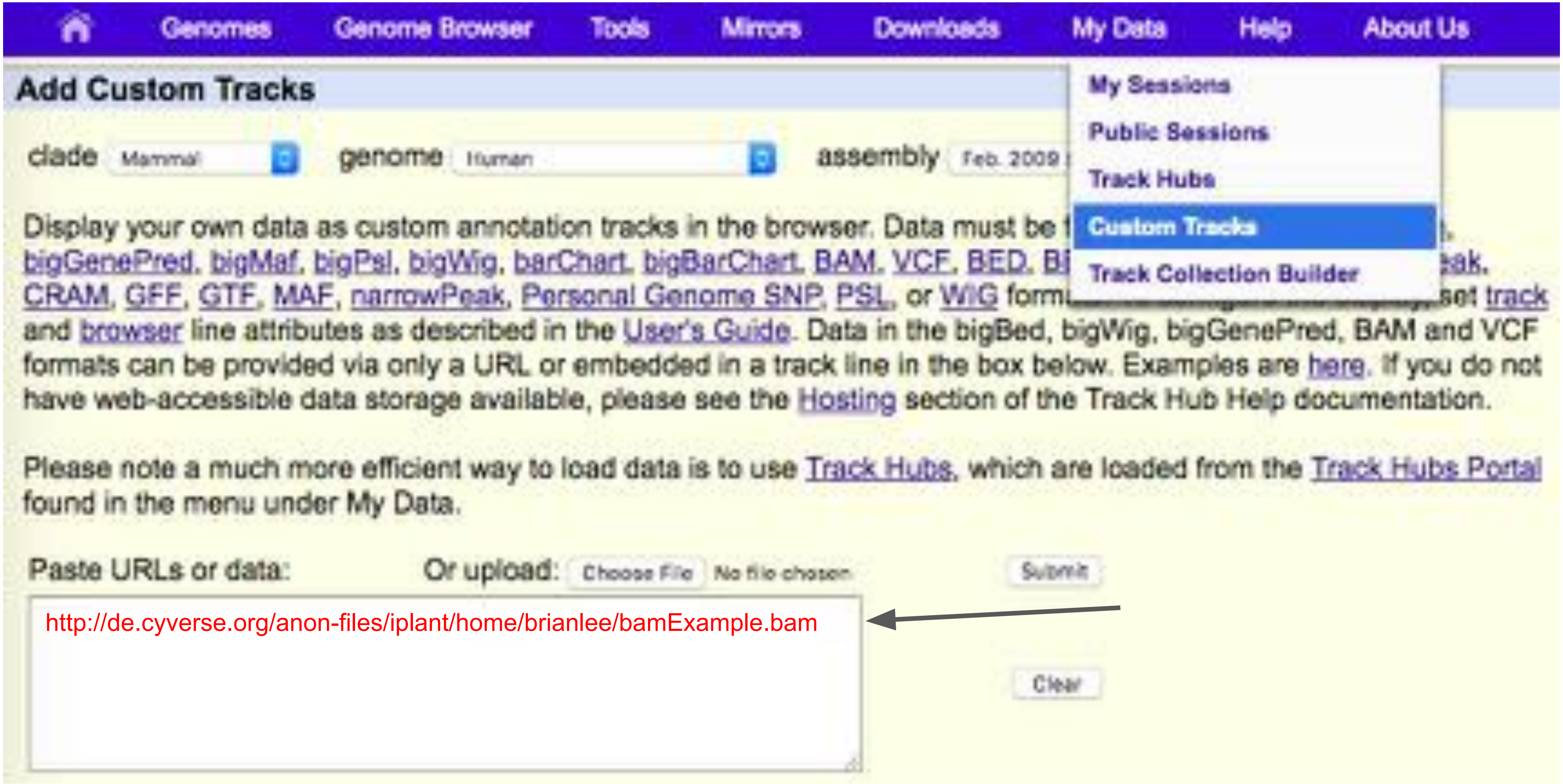
Note: Anyone with access to this file's URL can access the file.

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

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

Custom Annotations in the Browser

Custom Tracks (Text/**<u>Binary</u>**) **Track Hubs** **Assembly Hubs**



Add Custom Tracks

clade Mammal genome Human assembly Feb. 2009

Display your own data as custom annotation tracks in the browser. Data must be f .
bigGenePred, bigMaf, bigPsl, bigWig, barChart, bigBarChart, BAM, VCF, BED, BI eak,
CRAM, GFF, GTF, MAF, narrowPeak, Personal Genome SNP, PSL, or WIG form 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. If you do not have web-accessible data storage available, please see the Hosting section of the Track Hub Help documentation.

Please note a much more efficient way to load data is to use Track Hubs, which are loaded from the Track Hubs Portal found in the menu under My Data.

Paste URLs or data: Or upload:

http://de.cyverse.org/anon-files/iplant/home/brianlee/bamExample.bam live demo:

http://   genome.ucsc.edu

bit.ly/UCSC_UO