

# BLAT、In-Silico PCRを使って配列を検索する/Gene Sorterを使って発現データを解析する

## 1. UCSC Genome Browserの準備

- Genomeのバージョンを確認する
- Sessionsを使って設定を保存、共有する

### Genomeのバージョンを確認する

- UCSC genome browserにアクセスする
  - <https://genome-asia.ucsc.edu/>
- "Genomes"のタブ（もしくはそのタブの中にあるOthers）をクリックする

The screenshot shows the UCSC Genome Browser homepage. At the top, there's a banner for the University of California Santa Cruz Genomics Institute. Below the banner, the main navigation menu includes links for 'Genomes' (with a red arrow pointing to it), 'Genome Browser', 'Tools', 'Mirrors', 'Downloads', 'My Data', 'Projects', 'Help', and 'About Us'. The main content area features a genomic visualization with tracks for genes like SAT2, SHBG, ATP1B2, TPS3, and WRAP53. Below this, there are two sections: 'Tools' and 'News'. The 'Tools' section contains icons for hg38, hg19, mm39, and links to various tools. The 'News' section lists recent updates with dates and titles.

- ゲートウェイページが表示されます。"Assembly"を変更することで、さまざまなバージョンのゲノムの情報についてUCSC Genome Browser（以後、UCSC GB）が保持しているデータの細かい情報を閲覧することができます。たとえば、そのゲノムバージョンでの大きな変更点をハイライトとして記載してあり、ゲノムバージョン間の違いを理解するのにとても役にたちます。UCSC GBは、GENCODEをはじめとする各種ゲノムプロジェクトが行なったGenome Assemblyのデータを取得して、ブラウザ上で可視化しているので、その元となっている情報が存在することにご留意いただきたいと思います。色々なリンクをクリックしてみてください。方法別のデータのダウンロード、元となっているサイトへのリンク、そしてゲノム地図へのリンクがなされています。

## Browse&gt;Select Species

## POPULAR SPECIES

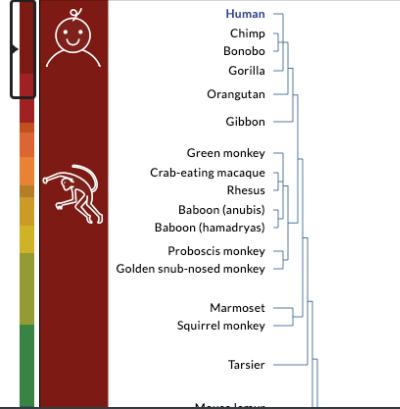


Search through thousands of genome browsers

Enter species, common name or assembly ID

Unable to find a genome? Send us a request.

## UCSC SPECIES TREE AND CONNECTED ASSEMBLY HUBS



## Find Position

## Human Assembly

Dec. 2013 (GRCh38/hg38)

GO

## Position/Search Term

Enter position, gene symbol or search terms

Current position: chr2:25,160,915-25,168,903

## Human Genome Browser - hg38 assembly

view sequences

UCSC Genome Browser assembly ID: hg38

Sequencing/Assembly provider ID: Genome Reference Consortium Human GRCh38.p13 (GCA\_000001405.28)

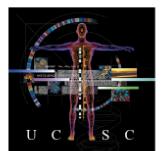
Assembly date: Dec. 2013 initial release; Dec. 2017 patch release 13

Assembly accession: GCA\_000001405.28

NCBI Genome ID: 51 (Homo sapiens (human))

NCBI Assembly ID: GCF\_000001405.39 (GRCh38.p13, GCA\_000001405.28)

BioProject ID: PRJNA31257



Homo sapiens

(Graphic courtesy of CSE)

## Search the assembly:

- By position or search term: Use the "position or search term" box to find areas of the genome associated with many different attributes, such as a specific chromosomal coordinate range; mRNA, EST, or STS marker names; or keywords from the GenBank description of an mRNA. [More information](#), including sample queries.
- By gene name: Type a gene name into the "search term" box, choose your gene from the drop-down list, then press "submit" to go directly to the assembly location associated with that gene. [More information](#).
- By track type: Click the "track search" button to find Genome Browser tracks that match specific selection criteria. [More information](#).

## Download sequence and annotation data:

- Using rsync (recommended)

一方で最新ゲノムの一つであるhs1はほとんど情報がまだ整備されていない様子も分かります。

## Browse&gt;Select Species

## POPULAR SPECIES

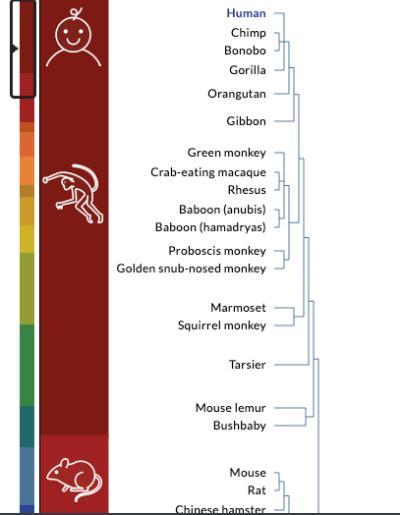


Search through thousands of genome browsers

Enter species, common name or assembly ID

Unable to find a genome? Send us a request.

## UCSC SPECIES TREE AND CONNECTED ASSEMBLY HUBS



## Find Position

## Human Assembly

Jan. 2022 (T2T CHM13v2.0/hs1)

GO

## Position/Search Term

Enter position, gene symbol or search terms

Current position: chr9:145,458,455-145,495,201

## Human Genome Browser - hs1 assembly

view sequences

## Common name: human

Taxonomic name: Homo sapiens, taxonomy ID: 9606

Sequencing/Assembly provider ID: T2T Consortium

Assembly date: 24 Jan 2022

Assembly type: haploid

Assembly level: Complete Genome

Biosample: (n/a)

Assembly accession ID: GCA\_009914755.4

Assembly FTP location: GCA/009/914/755/GCA\_009914755.4\_T2T-CHM13v2.0

Total assembly nucleotides: 3,117,292,070

Assembly contig count: 25

N50 size: 150,617,247

## Data file downloads

- hs1.fa.gz fasta sequence
- hs1.chromAlias.txt hs1.chromAlias.txt chromAlias file to relate chromosome names
- hs1.2bit hs1.2bit UCSC 2bit sequence file
- hs1.20220412.ncbiRefSeq.gtf.gz hs1.110.20220412.gtf.gz NCBI RefSeq genes GTF file version 110.20220412
- hs1.augustus.gtf.gz gene GTF file
- hs1.xenoRefGene.gtf.gz gene GTF file
- catLiftOffGenesV1.gff3.gz gene GFF3 file
- catLiftOffGenesV1.gtf.gz gene GTF file
- pre-computed indices for alignment programs: bowtie2, bwa-mem2, hisat2, minimap2
- explore the data download directory at: [hgdownload.soe.ucsc.edu/goldenPath/hs1/](http://hgdownload.soe.ucsc.edu/goldenPath/hs1/)

次に、Genomeタブの中にある"Genome Archive GenArk"をクリックすると、UCSC GBが持つ千を超えるゲノムアッセンブリーが種別にリストアップされています。hub gatewayの生物カテゴリーをクリックして各種のゲノムを選ぶとゲノム地図を閲覧できます。登録されているリストで眺めたい時にはこの手段がよいと思います。



Genomes    Genome Browser    Tools    Mirrors    Downloads    My I

## GenArk: UCSC Genome Archive

hub gateway	description
<a href="#">primates</a>	NCBI primate genomes (60 assemblies)
<a href="#">mammals</a>	NCBI mammal genomes (357 assemblies)
<a href="#">birds</a>	NCBI bird genomes (175 assemblies)
<a href="#">fishes</a>	NCBI fish genomes (195 assemblies)
<a href="#">vertebrate</a>	NCBI other vertebrate genomes (76 assemblies)
<a href="#">invertebrate</a>	NCBI invertebrate genomes (470 assemblies)
<a href="#">fungi</a>	NCBI fungi genomes (464 assemblies)
<a href="#">plants</a>	NCBI plant genomes (174 assemblies)
<a href="#">viral</a>	NCBI virus genomes (254 assemblies)
<a href="#">bacteria</a>	NCBI bacteria genomes (19 assemblies)
<a href="#">VGP</a>	Vertebrate Genomes Project collection (295 assemblies)
<a href="#">globalReference</a>	Global Human Reference genomes, January 2020 (10 assemblies)
<a href="#">mouseStrains</a>	16 mouse strain assembly and track hub, May 2017
<a href="#">legacy</a>	NCBI genomes legacy/superseded by newer versions (242 assemblies)

## Sessionsを使って設定を保存、共有する

UCSC genome browserにはユーザー別にセッションを残す機能が存在します。 **My Data**から**My Session**をクリックします。



The screenshot shows the UCSC genome browser's sign-in page. At the top, there is a navigation bar with links for Home, Genomes, Genome Browser, Tools, Mirrors, Downloads, My Data, Projects, Help, and About Us. Below the navigation bar, a blue header bar says "Sign in to UCSC Genome Bioinformatics". Underneath it, there are two buttons: "Login" and "Create an account". A yellow callout box contains the text: "Signing in enables you to save current settings into a named session, and then restore settings from the session later. If you wish, you can share named sessions with other users.".

アカウントを作成することで、ゲノム地図上で調べたトラックの情報やセッティングを保存し、さらには共有、非共有にすることができます。再度アカウントを作り試したところ、自動メッセージすぐに到着しましたので、ぜひアカウントを作ってみましょう！

**UCSC Genome Browser**

Signing up enables you to save multiple sessions and to share your sessions with others.

Already have an account? [Login](#).

**Sign Up**

Username

Email address

Re-enter Email address

Optional Secondary Recovery Email

Password (must be at least 5 characters)

Re-enter Password

[Sign Up](#) [Cancel](#)

各種情報を入力してsign upするとメールが送られてきますので、メールに記載されたURLでアカウント登録を完了してください。

Welcome MasakiM

Your Account Information

Username: MasakiM  
[Change password](#)  
[Sign out](#)

**Session Management**

See the [Sessions User's Guide](#) for more information about this tool. See the [Session Gallery](#) for example sessions.  
[Click here to reset](#) the browser user interface settings to their defaults.

**My Sessions**

session name (click to load)	created on	assembly	view/edit details	delete this session	share with others?	post in public listing?	send to mail
(none)							

**Save Settings** 名前を指定して保存

Save current settings as name:  
  allow this session to be loaded by others

Save current settings to a local file:  
file:  file type returned:  plain text  gzip compressed (ignored if output file is blank)   
(leave file blank to get output in browser window)

Save Custom Tracks:  
back up custom tracks to archive .tar.gz

**Restore Settings**

Use settings from another user's saved session:  
user:  session name:

Use settings from a local file:  選択されていません

Use settings from a URL (<http://...>, <ftp://...>):

**Sharing Sessions**

There are several ways to share saved sessions with others.

- Each previously saved named session appears with Browser and Email links. The Email link invokes your email tool with a message containing the Genome Browser link. The Email link, it will be the same as the Email link. However, if you click the Browser link it will take you to the Genome Browser and become a uniquely identified URL once the session loads.
- Each previously saved named session also appears with a checkbox to add the session to our [Public Sessions](#) listing. Adding a session to this listing allows other browser users to
- If you have saved your settings to a local file, you can send email to others with the file as an attachment and direct them to <https://genome.ucsc.edu/cgi-bin/hgTracks>.
- If a saved settings file is available from a web server, you can send email to others with a link such as [https://genome.ucsc.edu/cgi-bin/hgTracks?hgSession=doLoadUrl&submit&hgSession\\_load](https://genome.ucsc.edu/cgi-bin/hgTracks?hgSession=doLoadUrl&submit&hgSession_load). In this link, you can replace "hgSession" with "hgTracks" in order to proceed directly to the Genome Browser. For an example page using such links please see the [Session Gallery](#).

さすればそなたは UCSC GB に歓迎されるであろう。

さて、このページは、ユーザーがすでに開いたトラックなどの情報の保存や、復元ができる環境になっています。必要に応じて履歴のセッションを名前をつけてsubmitして保存しておくと、後でもう一度同じ地図を見たい時に便利です。折角なのでこの機能を使ってみたいと思います。今回は例としてブラウザで閲覧した情報を保存、そして読み込んでみたいと思います。赤矢印の**Public Session**をクリックしてください。

**Public Sessions**

Sessions allow users to save snapshots of the Genome Browser and its current configuration, including displayed tracks, position, and custom track data. The Public Sessions tool allows users to easily share those sessions that they deem interesting with the rest of the world's researchers. You can add your own sessions to this list by checking the appropriate box on the [Session Management](#) page.

See the [Sessions User's Guide](#) for more information.

Sort by: Popularity (descending) ▾

Show 10 entries

**Screenshot**

**Session Properties**

Search:

**ClinicalZoom**

Description: This sessionView is a collection of tracks centered around clinical significance. Featured tracks include gene annotations, SNVs, CNVs, SVs, and our publications track built by mining sequences and SNPs in publications. The displayed region is a 294bp window looking at the CAG repeat linked to Huntington's disease in the HTT gene. Two similar sessions are also available: <a href="http://genome.ucsc.edu/cgi-bin/hgTracks?hgS...doOtherUser=submit&hgS\_otherUserName=view&hgS\_otherUserSessionName=ClinicalLite" target="\_blank">ClinicalLite</a> which has a reduced number of tracks for increased clarity, and <a href="http://genome.ucsc.edu/cgi-bin/hgTracks?hgS...doOtherUser=submit&hgS\_otherUserName=view&hgS\_otherUserSessionName=Clinical" target="\_blank">Clinical</a> which is the most informative clinical view.

Author: view  
Session Name: ClinicalZoom  
Genome Assembly: hg19  
Creation Date: 2018-10-10  
Views: 9424

**ReMap 2020**

Description: ReMap 2020: An atlas of regulatory regions from an integrative analysis Arabidopsis thaliana DNA-binding sequencing experiments. Go to <a href="http://remap.univ-amu.fr">ReMap</a> for more info.

Author: Benoit Ballaster  
Session Name: US\_ReMap2020\_Thaliana  
Genome Assembly: hub\_1936559\_eraTha1  
Creation Date: 2019-08-29  
Views: 5953

**shiny\_app\_hg38\_v1**

Description:  
Author: jbn  
Session Name: shiny\_app\_hg38\_v1  
Genome Assembly: hg38  
Creation Date: 2018-07-11  
Views: 4374

Public Sessionは、他のユーザーが登録した共有セッションを閲覧することができます。好みの順にソートして、ぜひ眺めてみてください。セッティングの参考になるものがあるかもしれません。

この中で今回私が生物学的で面白いと思ったデータを読み込みたいと思います。Searchのboxにhemoglobinを入力してください。

Show 10 entries

**Screenshot**

**Session Properties**

Search: hemoglobin

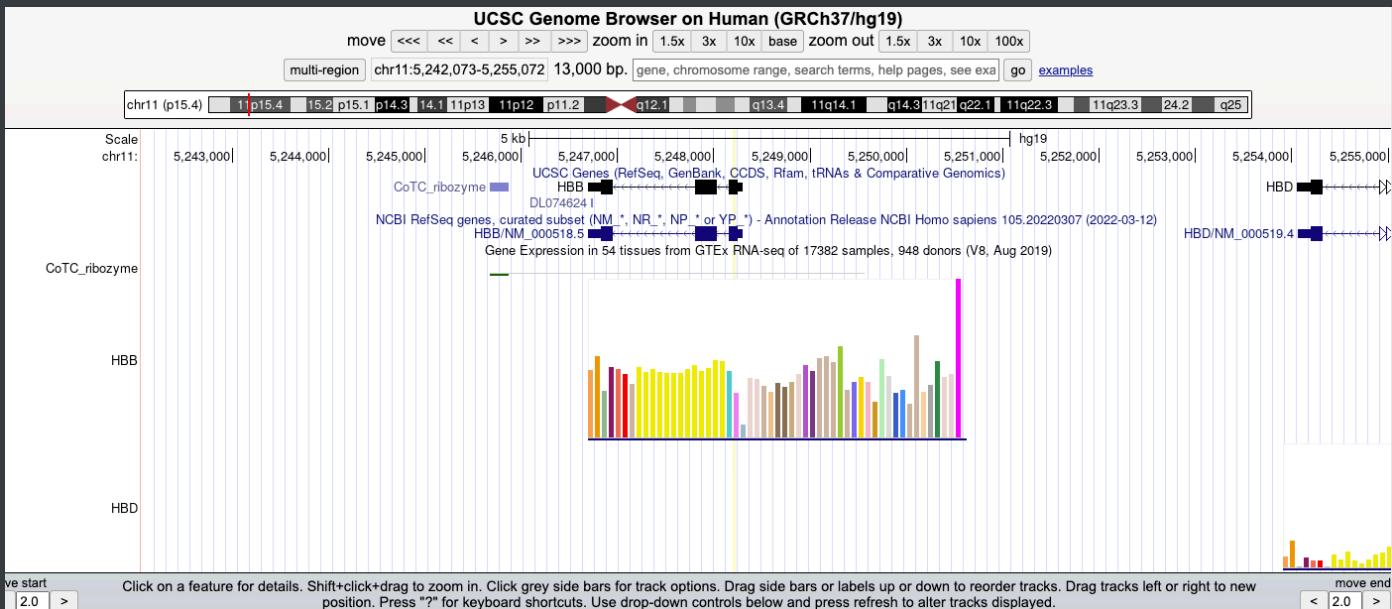
**hg19 hemoglobin**

Description: The genes HBB and HBD are both expressed in red blood cells , but HBB is also expressed in many other tissues , whereas HBD is expressed in red cells almost exclusively.  
Author: yrajat7900  
Session Name: hg19 hemoglobin  
Genome Assembly: hg19  
Creation Date: 2022-02-06  
Views: 58

Showing 1 to 1 of 1 entries (filtered from 51 total entries)

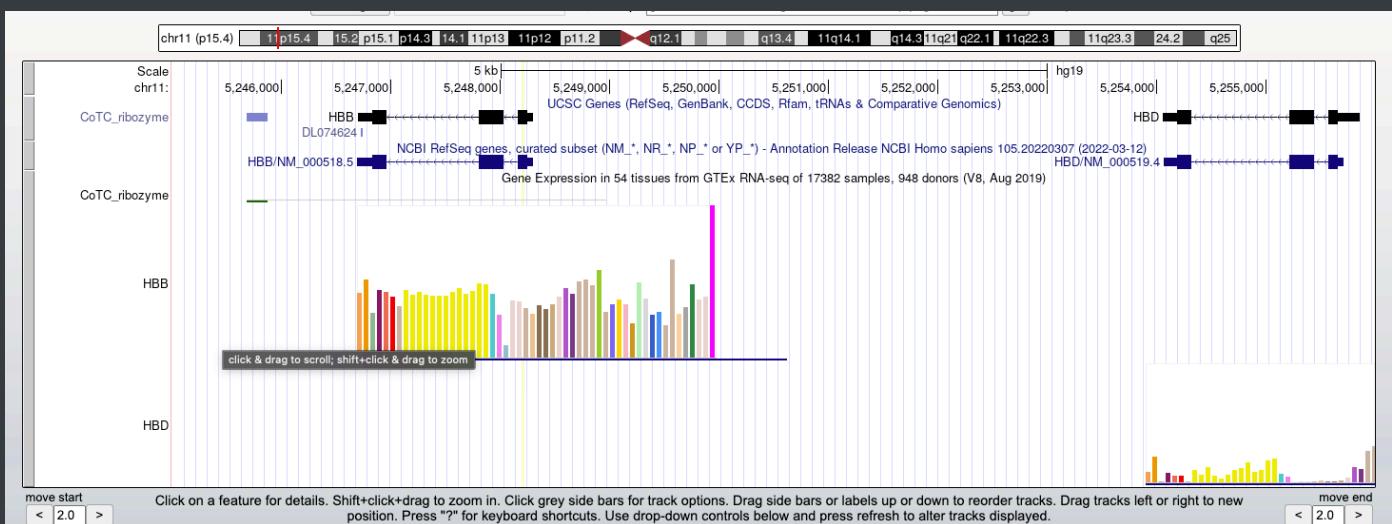
Previous 1 Next

するとyrajat7900さんの**ヘモグロビン遺伝子の遺伝子座 (hg19)**のセッションが現れます。このセッションを例に編集、保存してみます。赤い矢印のサムネイルをクリックしてください。ゲノム地図が表示されます。



表示されたゲノム地図は、HBB遺伝子を中心としたヒトのゲノムhg19を表示しています。これは、yarajat7900さんが保存して共有した画面セッティングそのままの状況を閲覧しています。血液とその他の組織のトランск립トーム解析を実施したことがある人にとっては非常に馴染みのある知見ですが、HBB (hemoglobin beta)とHBD (hemoglobin delta)は、発現組織に異なる傾向があります。HBBは発現がほぼユビキタスな発現がGTExのデータからよくわかると思います。一方でHBDは他の臓器より血液で多く発現しています。全ヘモグロビンの97%がHbA (alpha x2とbeta x2の4量体)からなり、HbA-2 (alpha x2とdelta x2の4量体)は、HbFとともに残りの3%を占めています。

HBDを見たいので地図のウインドーをドラックしたまま左へずらすことで、HBB遺伝子の全体も地図上に表示されるように調節してください。



上図のようにHBD, HBB両方の遺伝子が1つの画面に入れば両遺伝子の関係がわかりやすくなります。ではこの画面を呼び出せるように、保存してみたいと思います。先ほどユーザー登録をした時と同様に **MyData -> My Session** を選択して次の画面を表示してください。

Welcome MasakiM

Your Account Information

Username: MasakiM  
[Change password](#)  
[Sign out](#)

Updated Session

Deleted session hg19\_hemoglobin\_yrajat7900.

Session Management

See the [Sessions User's Guide](#) for more information about this tool. See the [Session Gallery](#) for example sessions.  
[Click here to reset](#) the browser user interface settings to their defaults.

My Sessions

session name (click to load)	created on	assembly	view/edit details	delete this session	share with others?	post in public listing?	send to mail
(none)							

Save Settings

名前を **test\_hem** submit

Save current settings as named session:  
name:   allow this session to be loaded by others

Save current settings to a local file:  
file:  file type returned:  plain text  gzip compressed (ignored if output file is blank) submit  
(leave file blank to get output in browser window)

実験ノートに残す場合はこちらでlocalに保存して、中の文字を貼り付けておこう

Save Custom Tracks:  
back up custom tracks to archive .tar.gz

そして、Save Settingsのcurrent settingsの欄に名前をつけて、submitをすれば先ほどの図(Session)が保存されます。また、web上に残すことが嫌な場合は、Save current settings to local fileで、ファイル名を指定し、settingをダウンロードすることをお勧めします。実験ノートに貼り付けておけばいつでも復元できます。

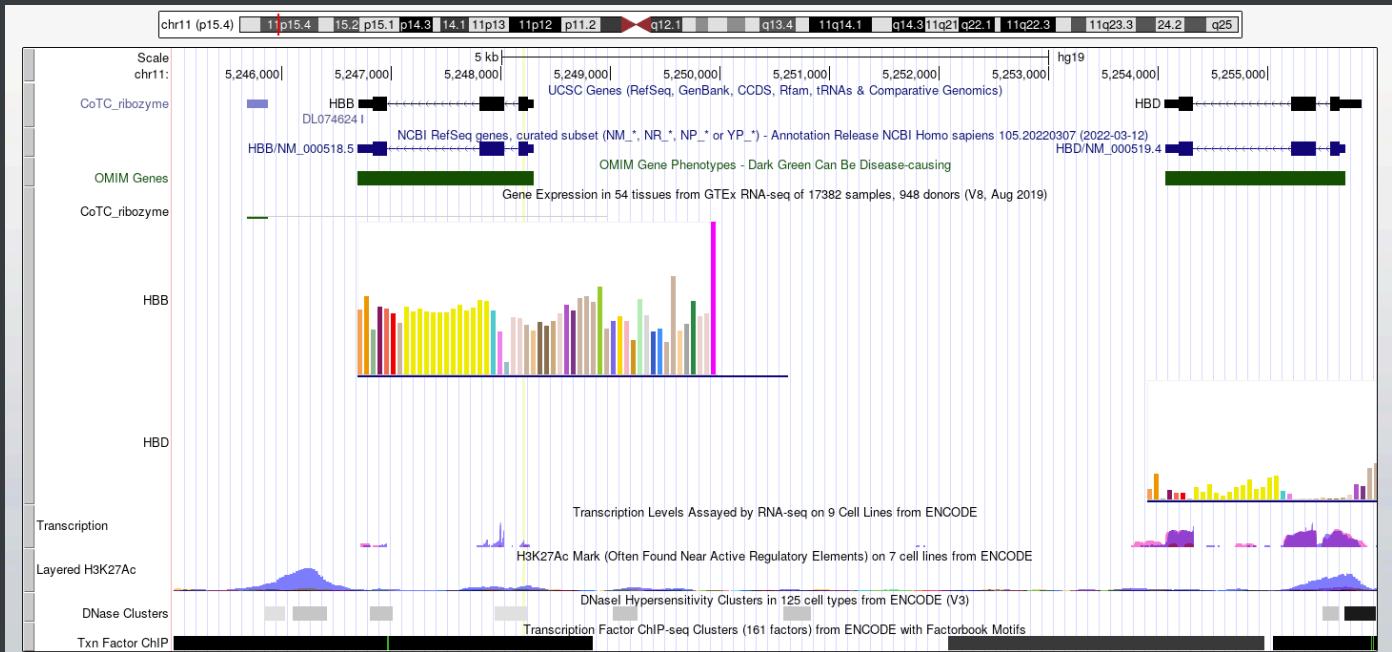
My Sessions

Show 10 entries

session name (click to load)	created on	assembly	view/edit details	delete this session	share with others?	post in public listing?	send to mail
<a href="#">test_hem</a>	2023-01-13	hg19	<a href="#">view/edit</a>	<a href="#">delete</a>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<a href="#">Email</a>

Showing 1 to 1 of 1 entries

Save current settingsを行なった場合はこのようにMySessionsに保存してもらいます。保存期限はおそらくないと思います。view/editをすれば、メモ書きも残せます。復元してloadする場合は、session nameをクリックしてください。



先ほどの地図が復元されています。HBD遺伝子全体も画面にしっかり入っている（**保存直前の地図である**）ことがお分かりいただけたと思います。`local`に保存した場合のテキストファイルの中身もご覧いただきましょう。

```
_ 1673597223254
altSeqLiftOverPsl hide
avada pack
c chr11
cartVersion 3
clade mammal
cons100way hide
cytoBandIdeo hide
db hg19
dbSnp153Composite hide
dbSnp155Composite hide
dinkL 2.0
dinkR 2.0
expOrder_cons100way phyloP100wayAll
expOrder_wgEncodeReg wgEncodeRegMarkH3k27ac wgEncodeRegTxn
fixSeqLiftOverPsl hide
goButton go
gtexGeneV8 full
hgFind.matches
```

hgPS\_DataTableState

---以下略---

利用していないトラックの情報まで非常に細かく情報が記載されています。このファイルを **MySessions**の中にあるRestore SettingsのUse settings from a local fileに読み込ませると同じものが復元できます。これで、ゲノムブラウザのセッションの保存と復元ができるようになりました。また、この機能を利用することでブラウザの可視化セッティングを真似ねて学習することができます。