2016年度第2回
バイオインフォマティクス実習

integrative genomics viewerでマッピングしたシーケンスを可視化する

前回

- マウスmm10 染色体1番のシーケンスをインデックス化
- bowtieを使ってSRR1805875のデータ（サイズを縮小）をマッピング
- samファイル形式で出力

今回

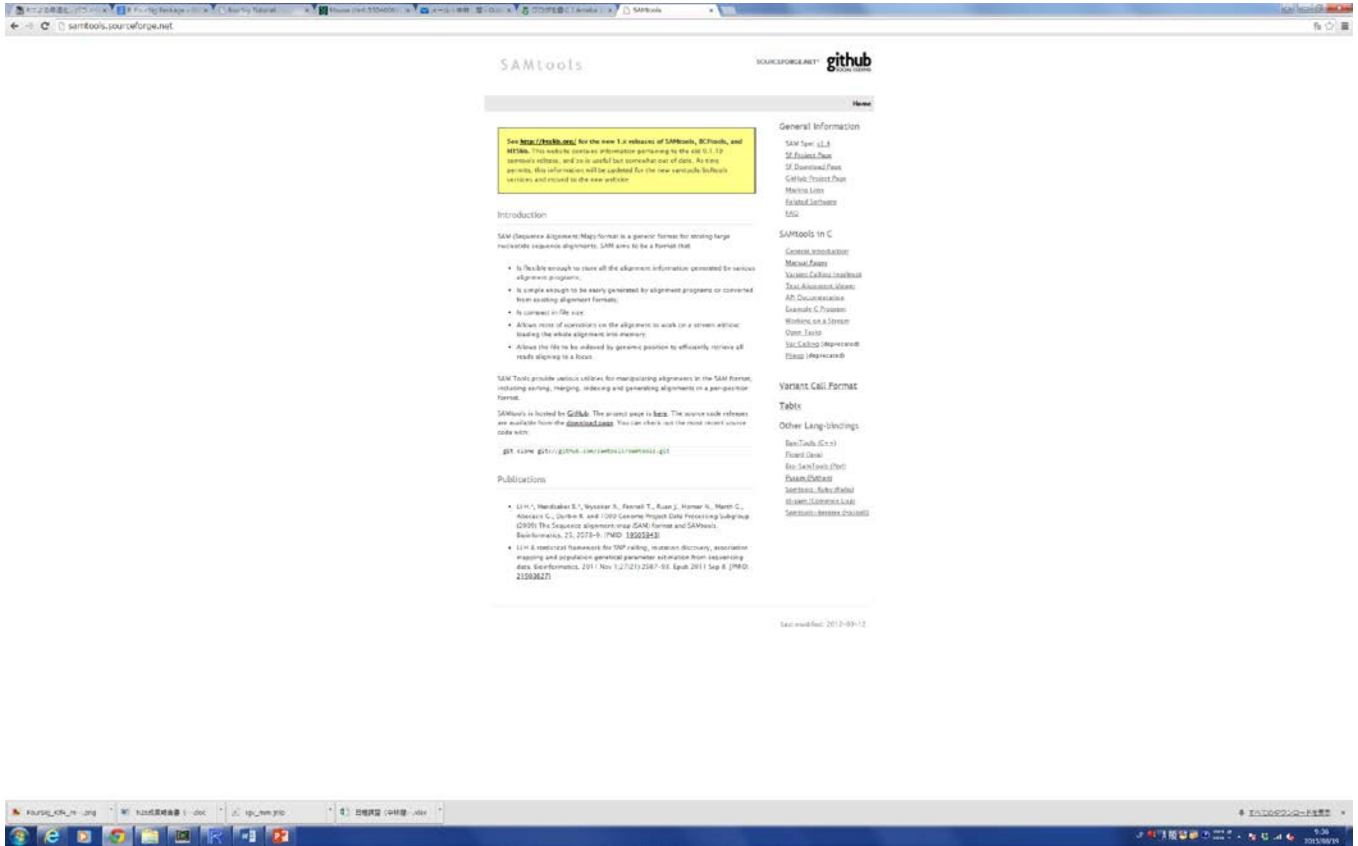
- sam→bamファイルへ変換
- bamファイルをソート
- indexを作成
- integrative genomics viewerにアップロード

samtoolsを使ってsam→bam変換

- samtools

マッピングしたデータを操作するツール

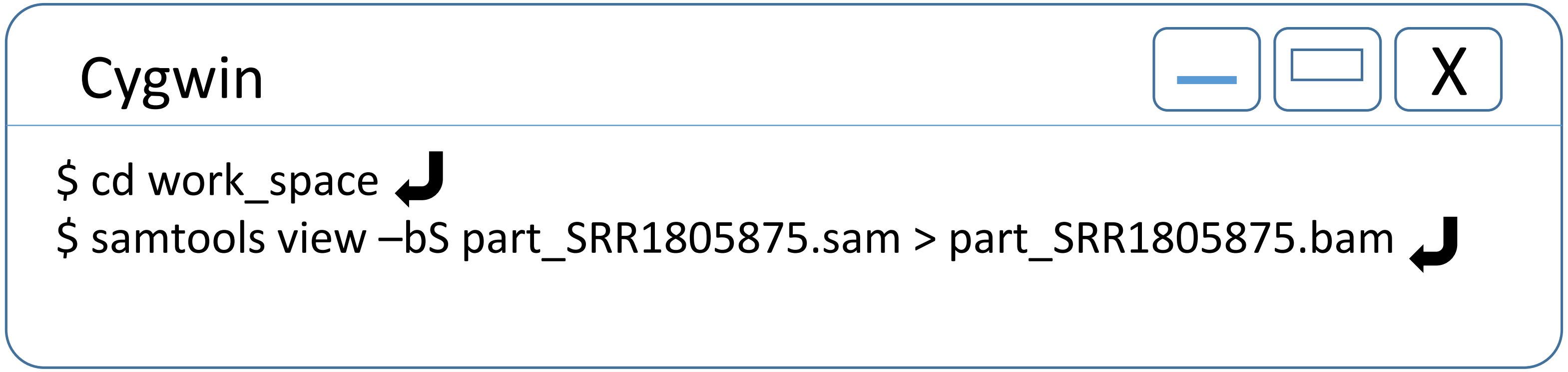
http://samtools.sourceforge.net/で配布



samtoolsのcygwinでのインストール

- ソースコードsamtools-1.x.tar.bz2をダウンロード解凍する
- Cygwin terminalで解凍したフォルダへ移動
- makeと入力

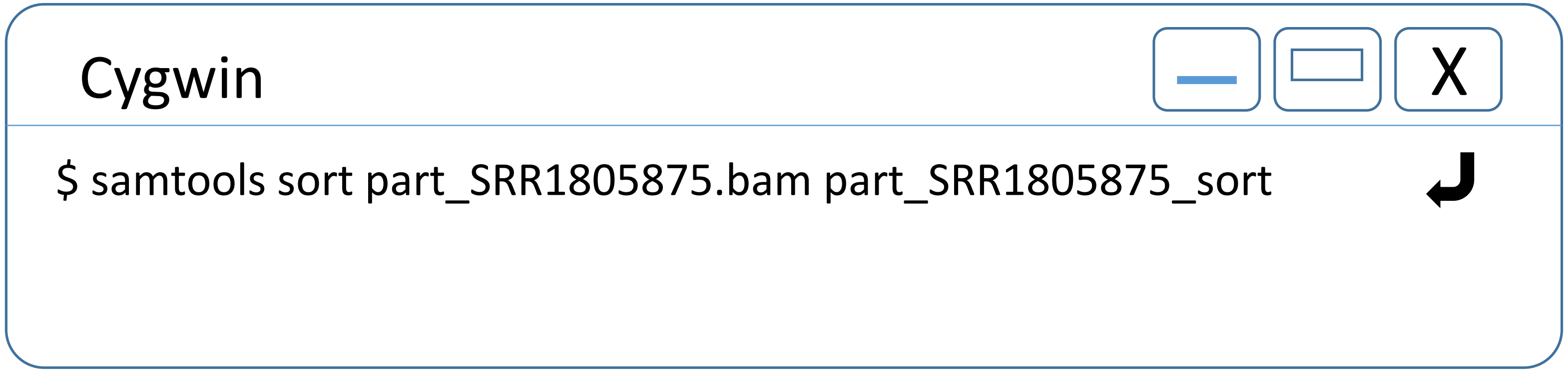
sam→bam変換



```
$ cd work_space ↵
$ samtools view –bS part_SRR1805875.sam > part_SRR1805875.bam ↵
```

sam→bam変換の書式
samtools view –bS samファイル名 > bamファイル名

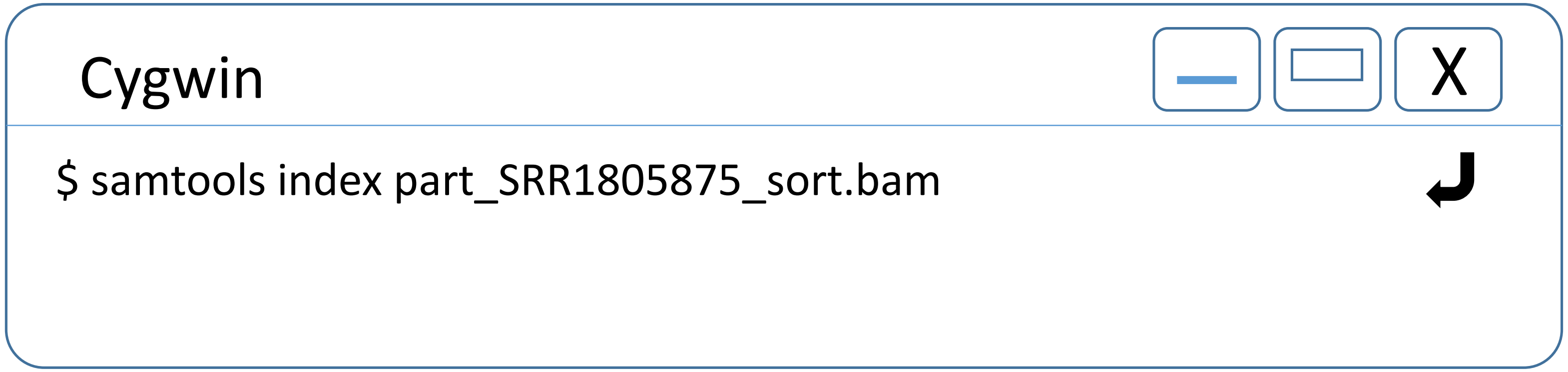
bamファイルのソート



```
$ samtools sort part_SRR1805875.bam part_SRR1805875_sort
```

ソートの書式
samtools sort bamファイル名　ソート後のbamファイル名（拡張子.bam無し）

インデックスファイル作成



```
$ samtools index part_SRR1805875_sort.bam
```

ソートの書式
samtools index ソート後のbamファイル名（拡張子.bam有り）

bamファイル名.bam.baiという名前のインデックスファイルを作成してくれる

integrative genomics viewer

http://www.broadinstitute.org/igv/

What's New

September 2014. The IGV iPad app can now be installed from the Apple App Store. ***IGV for iPad*** is a lightweight genomic data viewer that provides some of the functionality available in our regular desktop IGV. See the IGV for iPad documentation for details.

Overview

The **Integrative Genomics Viewer (IGV)** is a high-performance visualization tool for interactive exploration of large, integrated genomic datasets. It supports a wide variety of data types, including array-based and next-generation sequence data, and genomic annotations.

Downloads

Please register to download IGV. After registering, you can log in at any time using your email address. Permission to use IGV is granted under the GNU LGPL license.

Citing IGV

To cite your use of IGV in your publication:

James T. Robinson, Helga Thorvaldsdóttir, Wendy Winckler, Mitchell Guttman, Eric S. Lander, Gad Getz, Jill P. Mesirov. Integrative Genomics Viewer. Nature Biotechnology 29, 24–26 (2011)

Helga Thorvaldsdóttir, James T. Robinson, Jill P. Mesirov. **Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration.** Briefings in Bioinformatics 14, 178-192 (2013).

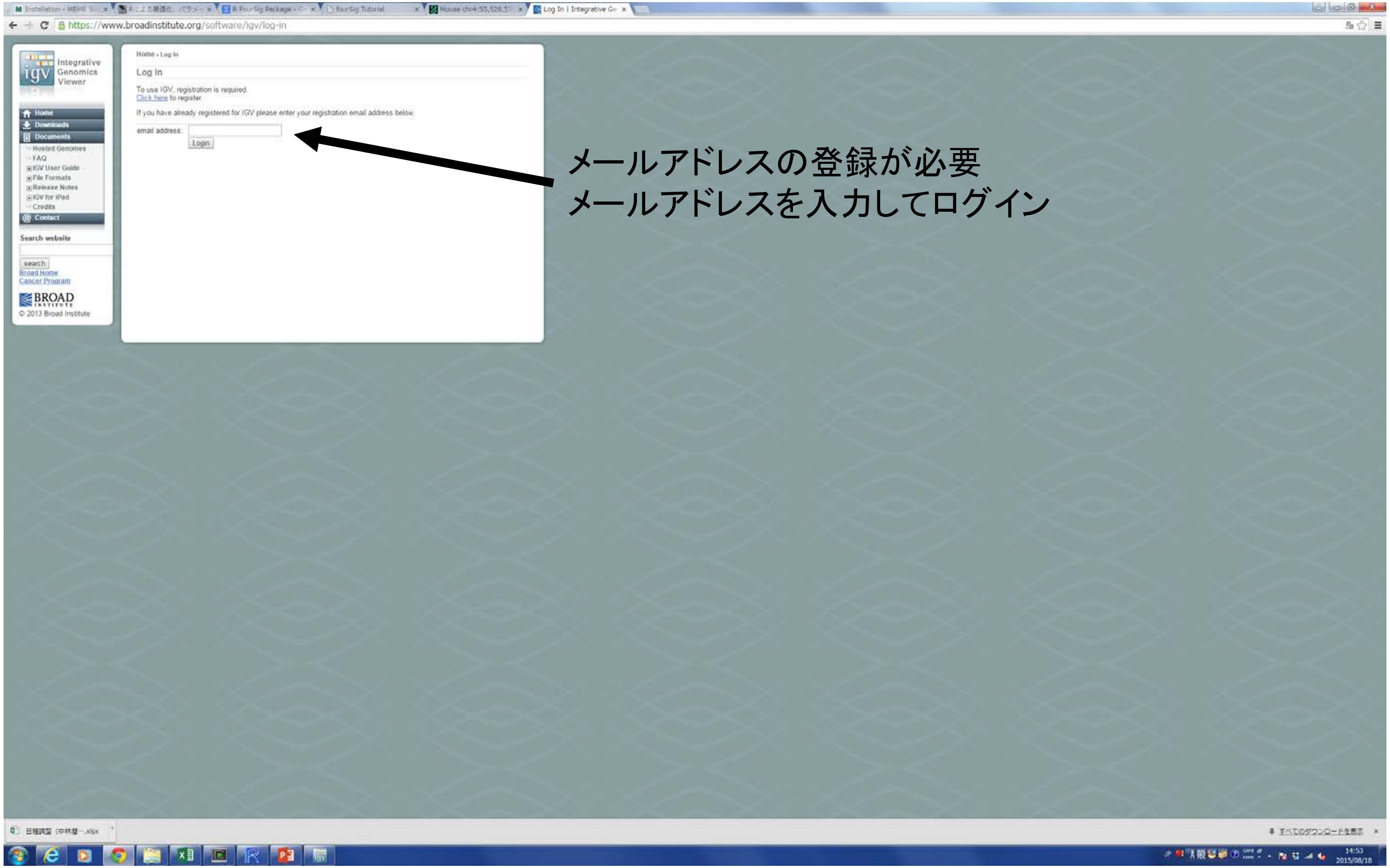
Funding

Development of IGV is made possible by funding from the National Cancer Institute, the National Institute of General Medical Sciences of the National Institutes of Health, and the Starr Cancer Consortium.

IGV participates in the GenomeSpace initiative, which is funded by the National Human Genome Research Institute.

ダウンロードページ

Download



Log In

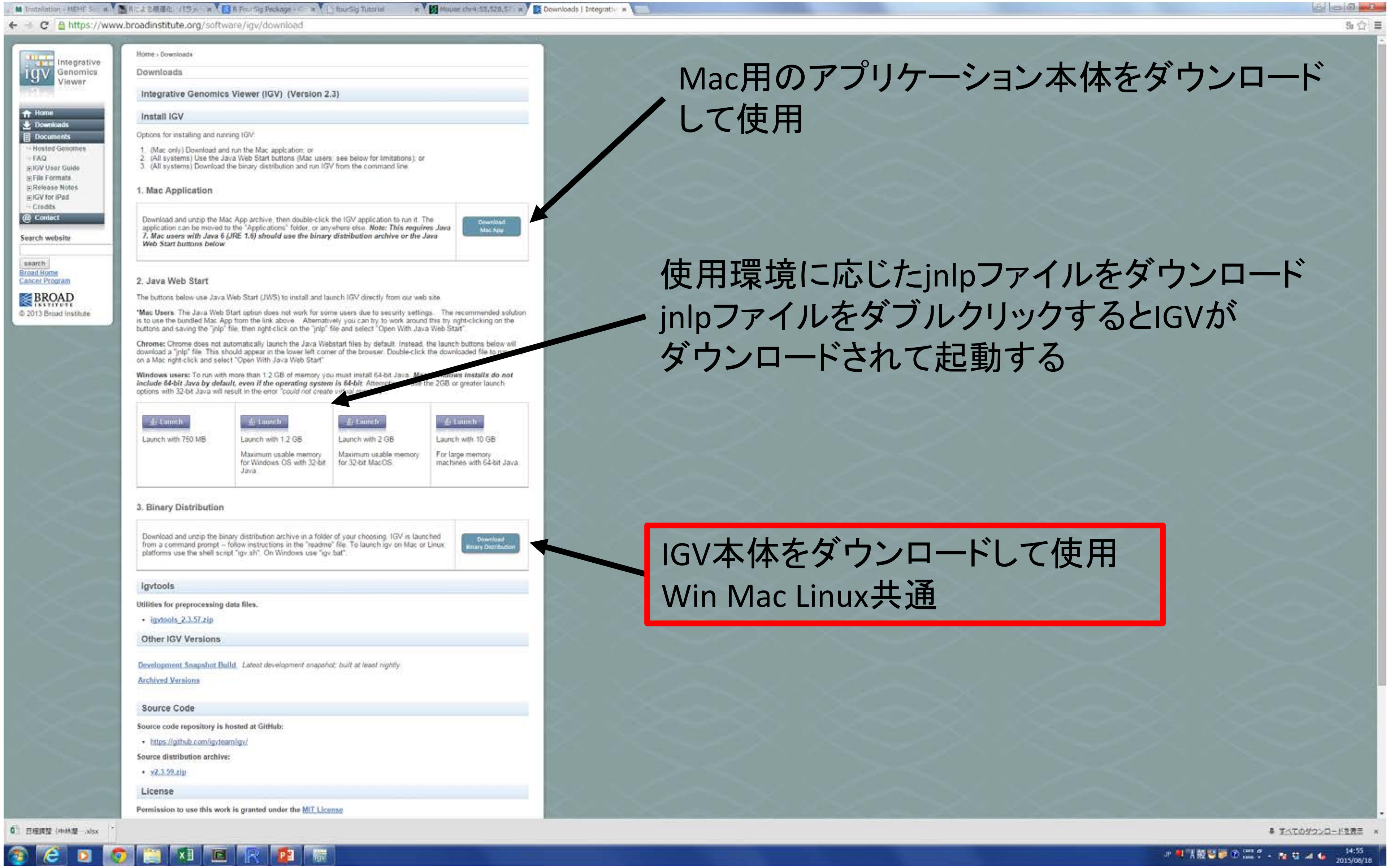
To use IGV, registration is required.
Click here to register.

If you have already registered for IGV please enter your registration email address below.

email address:

メールアドレスの登録が必要
メールアドレスを入力してログイン

ダウンロードページ



Downloads

Integrative Genomics Viewer (IGV) (Version 2.3)

Install IGV

Options for installing and running IGV:

1. (Mac only) Download and run the Mac application; or
2. (All systems) Use the Java Web Start buttons (Mac users: see below for limitations); or
3. (All systems) Download the binary distribution and run IGV from the command line

1. Mac Application

Download and unzip the Mac App archive, then double-click the IGV application to run it. The application can be moved to the "Applications" folder, or anywhere else. ***Note: This requires Java 7. Mac users with Java 6 (JRE 1.6) should use the binary distribution archive or the Java Web Start buttons below.***

2. Java Web Start

The buttons below use Java Web Start (JWS) to install and launch IGV directly from our web site.

***Mac Users**: The Java Web Start option does not work for some users due to security settings. The recommended solution is to use the bundled Mac App from the link above. Alternatively you can try to work around this try right-clicking on the buttons and saving the "jnlp" file, then right-click on the "jnlp" file and select "Open With Java Web Start".

Chrome: Chrome does not automatically launch the Java Webstart files by default. Instead, the launch buttons below will download a "jnlp" file. This should appear in the lower left corner of the browser. Double-click the downloaded file to run [illegible] on a Mac right-click and select "Open With Java Web Start".

Windows users: To run with more than 1.2 GB of memory you must install 64-bit Java. ***Ma[illegible]ndows installs do not include 64-bit Java by default, even if the operating system is 64-bit.*** Attempti[illegible] the 2GB or greater launch options with 32-bit Java will result in the error *"could not create vi[illegible]"*.

Launch with 750 MB	Launch with 1.2 GB Maximum usable memory for Windows OS with 32-bit Java.	Launch with 2 GB Maximum usable memory for 32-bit MacOS.	Launch with 10 GB For large memory machines with 64-bit Java.

3. Binary Distribution

Download and unzip the binary distribution archive in a folder of your choosing. IGV is launched from a command prompt -- follow instructions in the "readme" file. To launch igv on Mac or Linux platforms use the shell script "igv.sh". On Windows use "igv.bat".

Igvtools

Utilities for preprocessing data files.

- igvtools_2.3.57.zip

Other IGV Versions

Development Snapshot Build *Latest development snapshot; built at least nightly*

Archived Versions

Source Code

Source code repository is hosted at GitHub:

- https://github.com/igvteam/igv/

Source distribution archive:

- v2.3.59.zip

License

Permission to use this work is granted under the MIT License

Mac用のアプリケーション本体をダウンロードして使用

使用環境に応じたjnlpファイルをダウンロード
jnlpファイルをダブルクリックするとIGVがダウンロードされて起動する

IGV本体をダウンロードして使用
Win Mac Linux共通

IGV

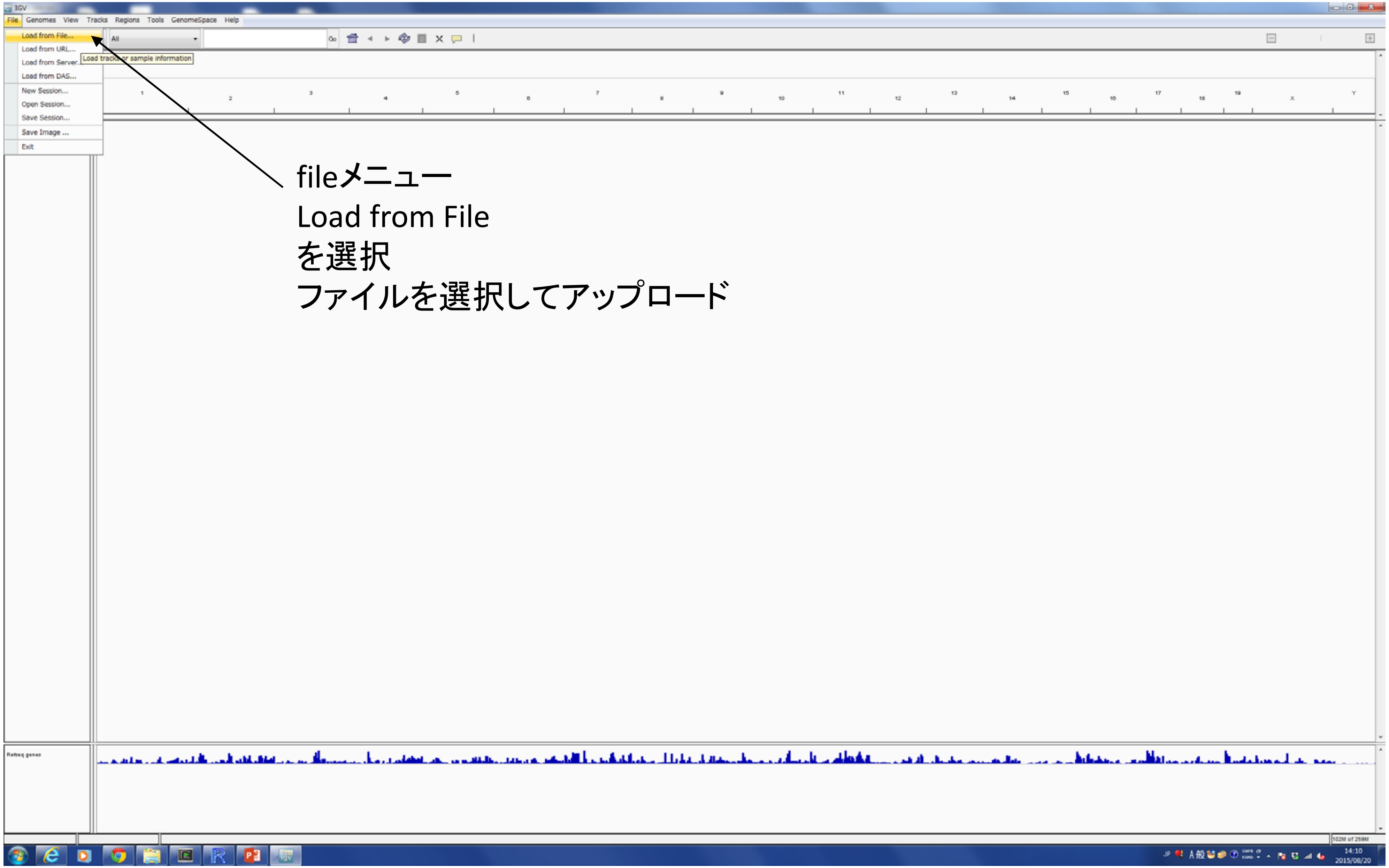
ゲノムを選択 mouse mm10

染色体番号

トラック

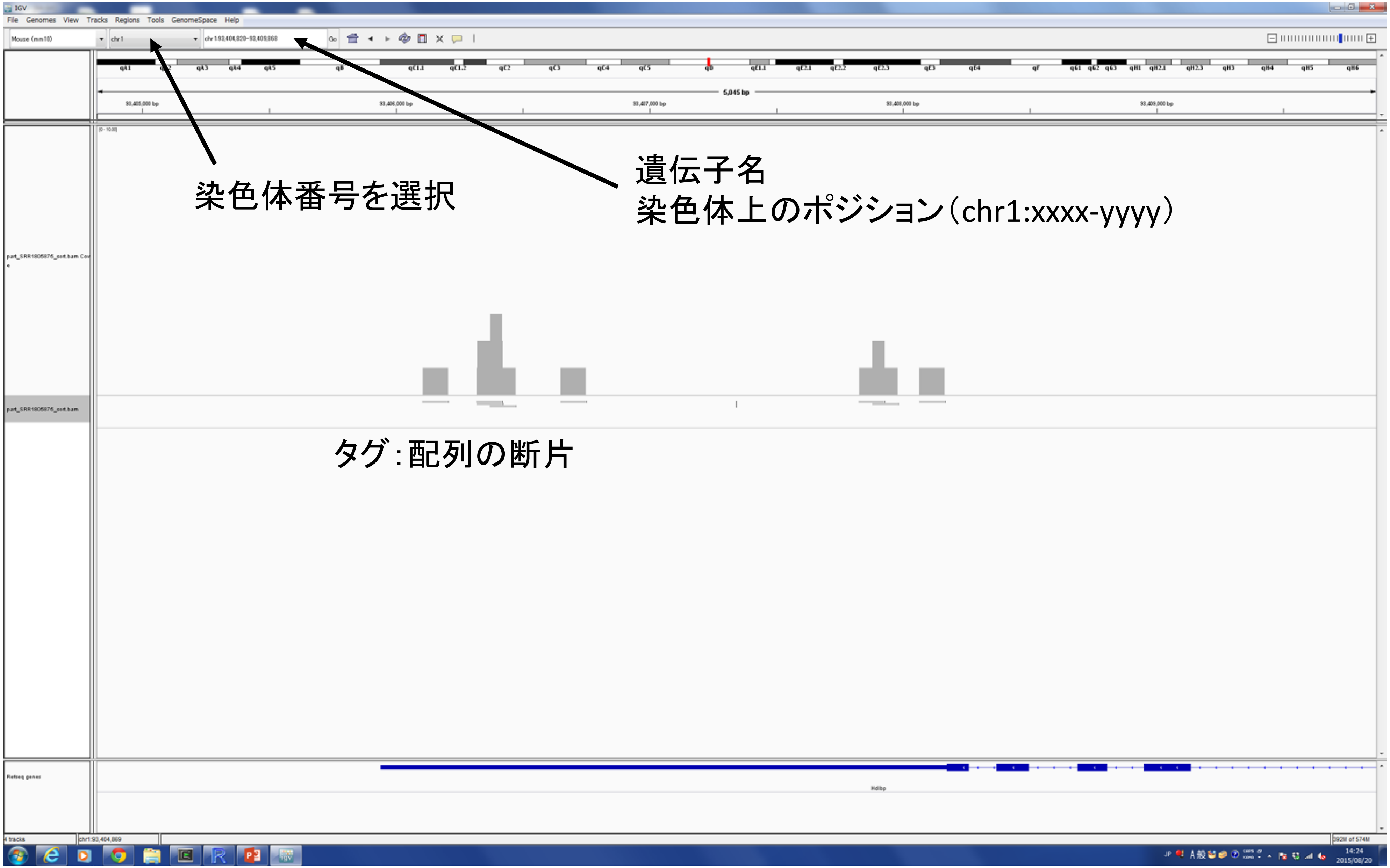
リファレンス

fileをupload



fileメニュー
Load from File
を選択
ファイルを選択してアップロード

遺伝子名で検索



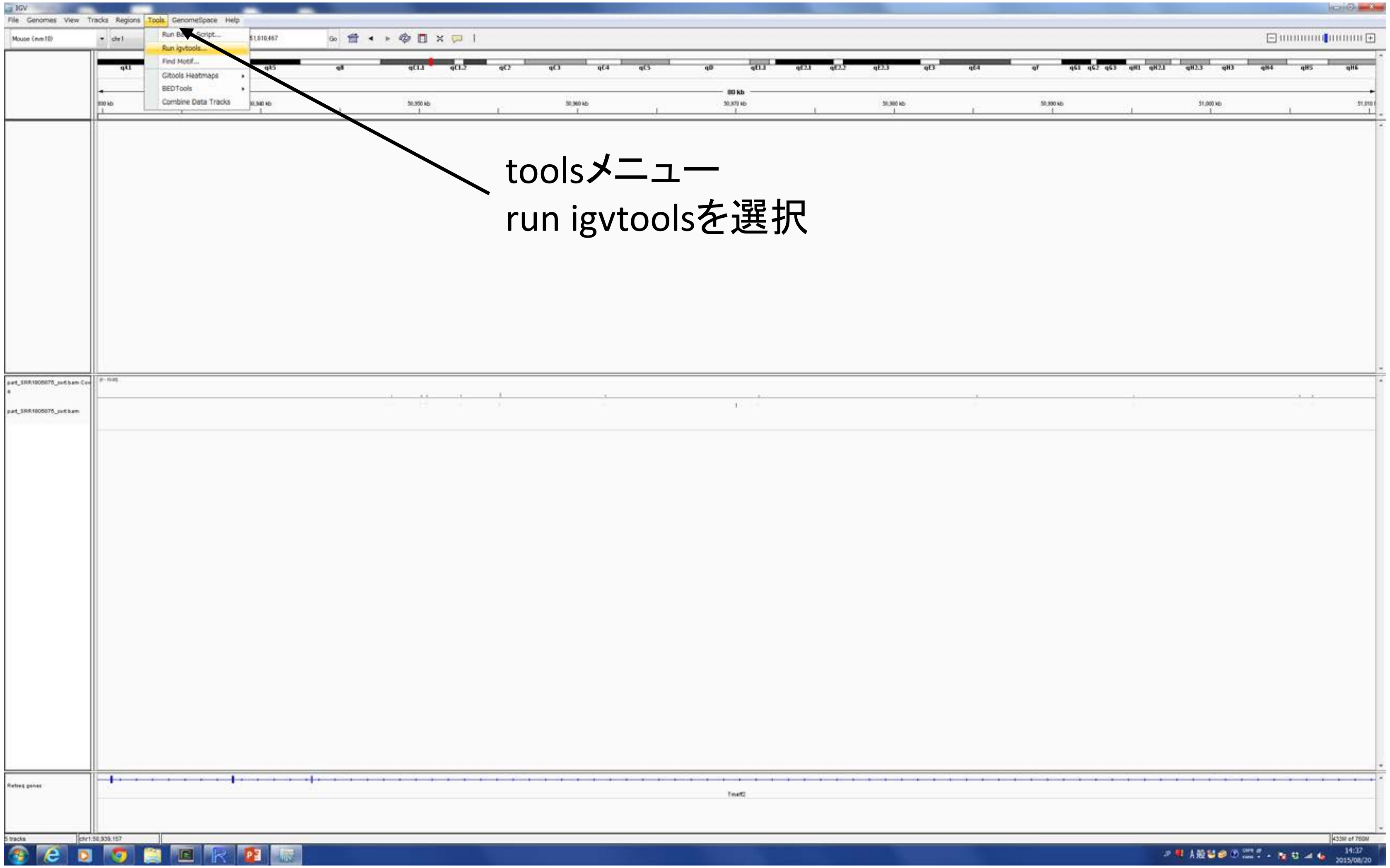
染色体番号を選択

遺伝子名
染色体上のポジション（chr1:xxxx-yyyy）

タグ：配列の断片

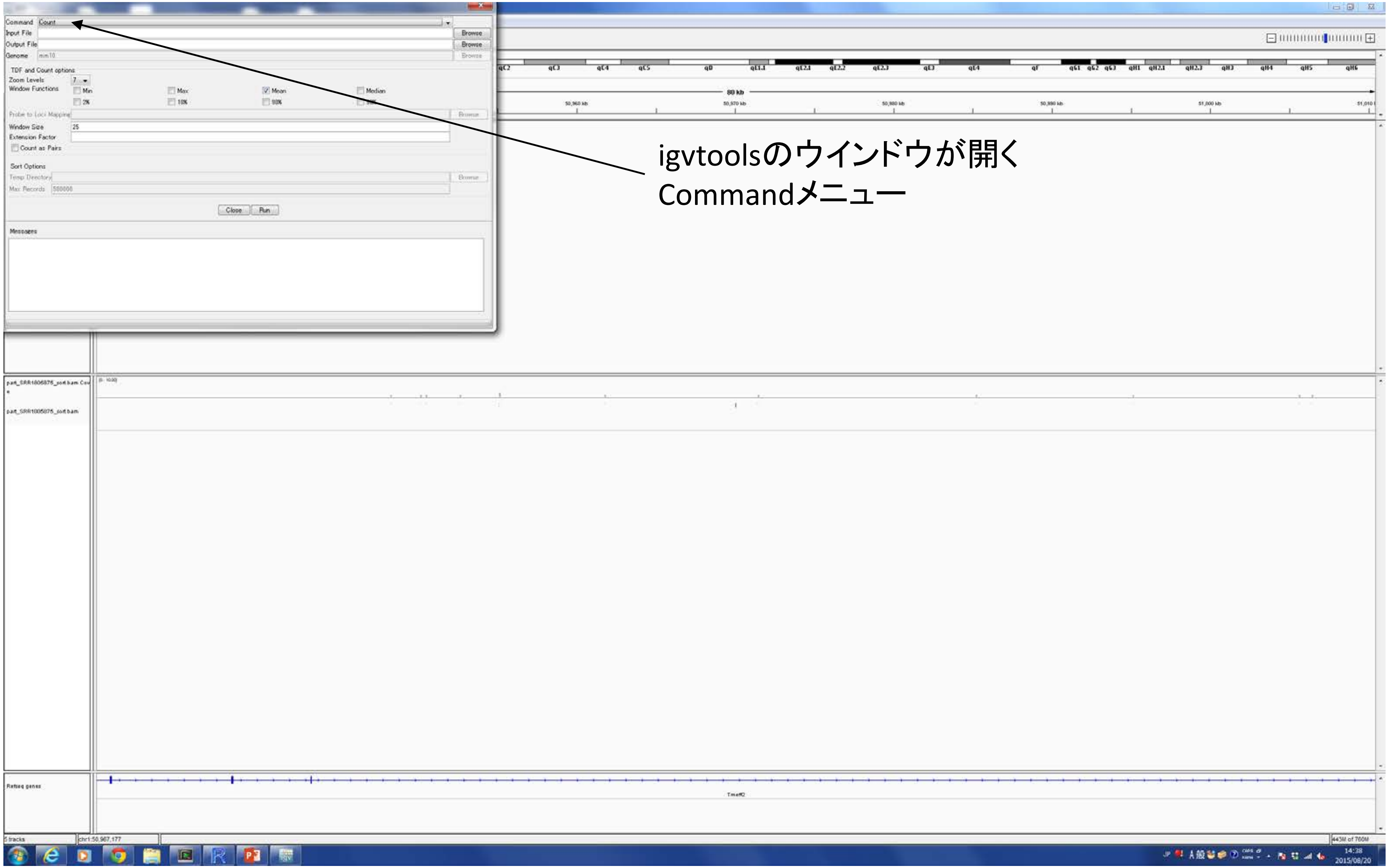
TDFファイルの作成

- TDFファイル：タグ情報のバイナリファイル
  容量が少なく扱いやすい



toolsメニュー
run igvtoolsを選択

igvtools



Command: Count
Input File:
Output File:
Genome: mm10

TDF and Count options
Zoom Levels: 7
Window Functions: ☐ Min ☐ Max ☑ Mean ☐ Median ☐ 2% ☐ 10% ☐ 90% ☐ 98%
Probe to Loci Mapping:
Window Size: 25
Extension Factor:
☐ Count as Pairs

Sort Options
Temp Directory:
Max Records: 500000

Messages

igvtoolsのウインドウが開く
Commandメニュー

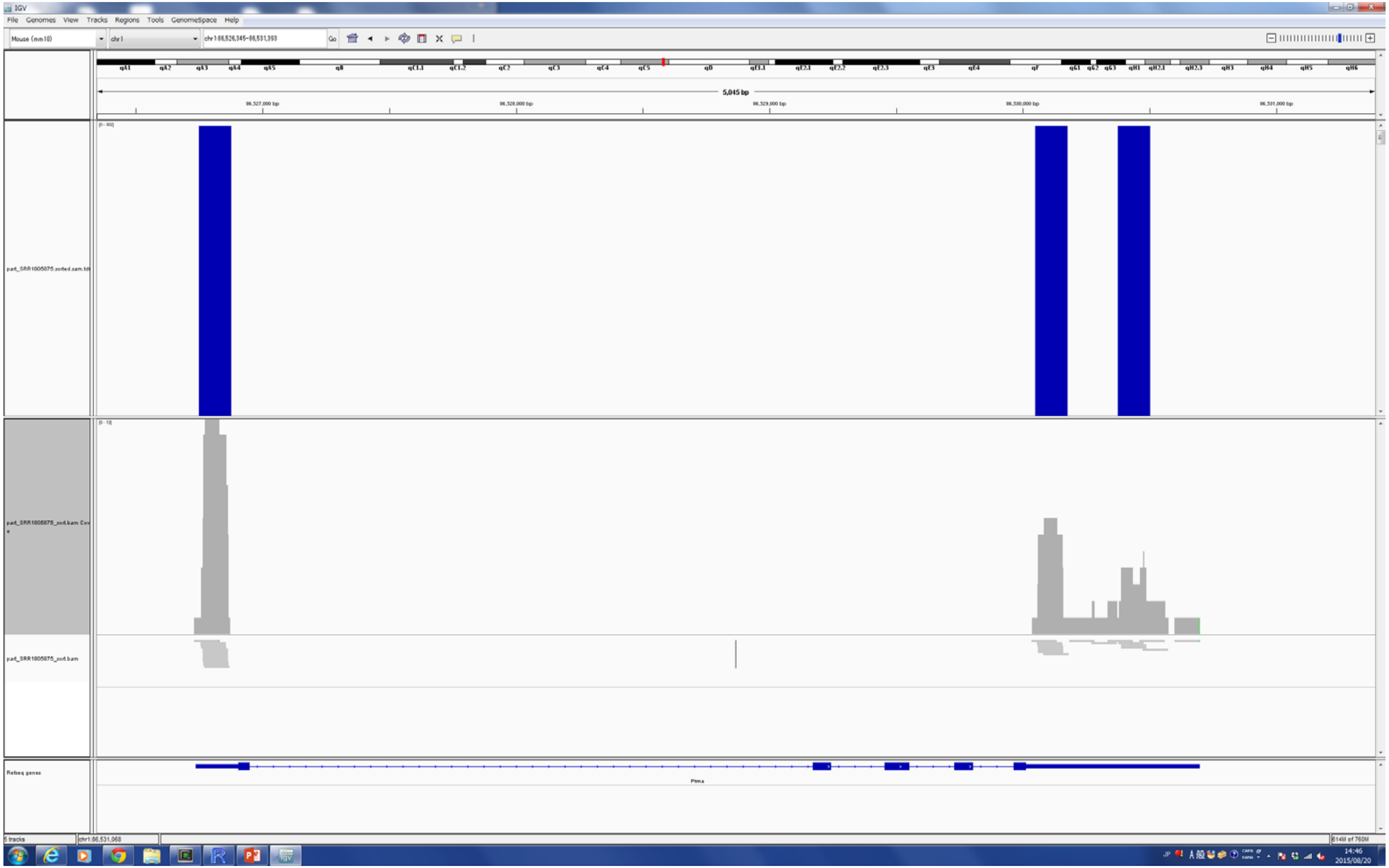
sort→index→count

- samファイルをソート
  ファイル名.sorted.sam

- インデックスを作成
  ファイル名.sorted.sam.sai

- TDFファイルを作成
  ファイル名.sorted.sam.tdf

ptma


IGV
File
Genomes
View
Tracks
Regions
Tools
GenomeSpace
Help
Mouse (mm10)
chr1
chr1:86,526,345-86,531,393
Go
qA1
qA2
qA3
qA4
qA5
qB
qC1.1
qC1.2
qC2
qC3
qC4
qC5
qD
qE1.1
qE2.1
qE2.2
qE2.3
qE3
qE4
qF
qG1
qG2
qG3
qH1
qH2.1
qH2.3
qH3
qH4
qH5
qH6
5,045 bp
86,527,000 bp
86,528,000 bp
86,529,000 bp
86,530,000 bp
86,531,000 bp
[0 - 302]
part_SRR1805875.sorted.sam.tdf
[0 - 13]
part_SRR1805875_sort.bam Coverage
part_SRR1805875_sort.bam
Refseq genes
Ptma
5 tracks
chr1:86,531,068
614M of 760M
14:46
2015/08/20

まとめ

- マッピングしたsamフォーマットの配列情報をsamtoolsでbamファイルに変換
- インデックスを作成
- IGVにアップロード、データを可視化
- igvtoolsでsamファイルをソート、インデックス作成、tdfファイルを作成
- tdfファイルをIGVにアップロード、可視化