ChIP-Seq Analysis in
Partek® Flow®

Training Data Files

Data files in the project:

- 2 samples (sub sample data from GSE102004): -- Illumina HiSeq 2000, paired end

- Aligned to hg19-chr22 using BWA-MEM

    - ChIP-chr22.bam :

        - GSM2720367
        - ChIP antibody: H3K4me3

    - IGG-chr22.bam

        - GSM2720355
        - ChIP antibody: IgG

Annotation file in the project:

- Hg19-chr22 Refseq

Notes: _____

_____

_____

_____
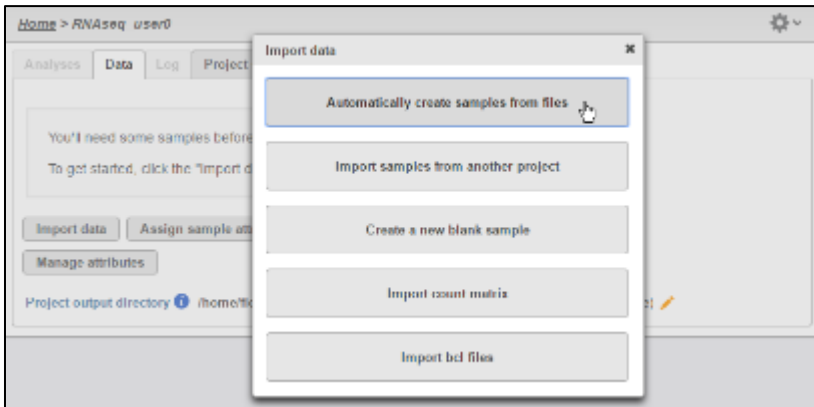
_____

Login and Project Set-up

- Open your preferred web browser (Chrome, Firefox, etc. would work fine)
- Go to the server URL given by your instructor
- Log in using the username and password given to you
- This will open to the Partek Flow homepage
- Click **New Project** and enter project name: ChIP-Seq-[username]
- This will create a new project
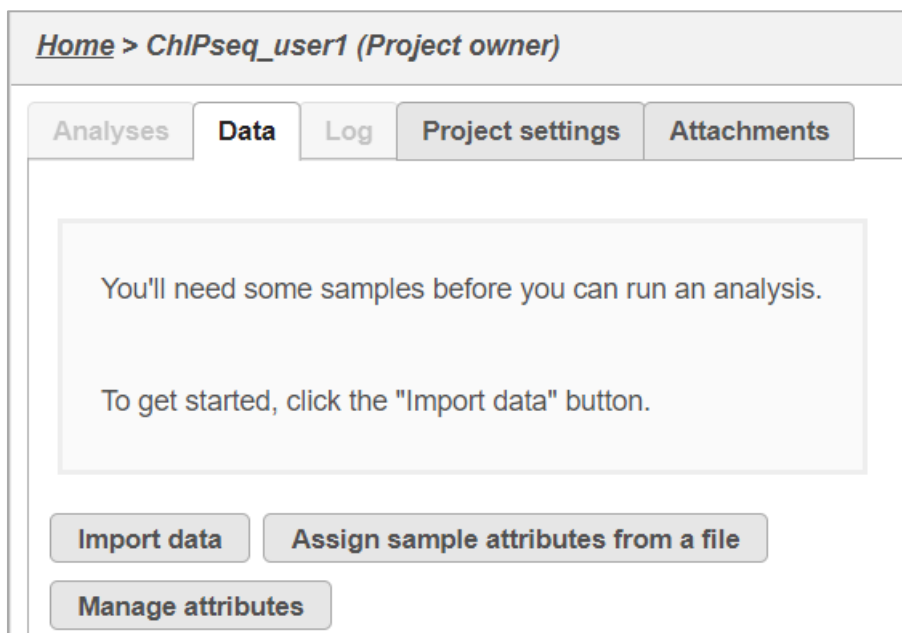
Notes: _____

_____

_____

_____

_____

Data Upload

- Creating a new project automatically opens up the **Data** tab
- To upload your data, click **Import data>Automatically create samples from files**
- Browse to */home/flow/FlowData/ChIP-seq*
- Select the two .bam *files* and click **Create sample**

Notes: _____

_____

_____

_____

_____

_____

Analyses Tab Overview

- Go to the **Analyses** tab
- Your first data node, the **Aligned reads** node appears.
- Select the **Aligned reads** data node and select **Filter alignments** in Post-alignment tools.
- Check **Filter duplicates** button
- Choose the **Same start and same sequence** option
- Leave everything else as default settings and click **Finish**
- This will create a new task node called *Filtered reads* in the **Analyses** tab.

Notes:

_____

_____

_____

_____

_____

Peak Detection

- Select **Filtered reads** data node
- Select **MACS2** from the **Peak callers** section of the menu to detect peaks in ChIP sample
- Select **Homo sapiens (human)** - **hg19_chr22** as the Assembly
- Leave other parameters as default
- Set the data type to **ChIP**
- Check the **ChIP** sample and select the **Add IP sample** button
- Check the **IGG** sample and select the **Add control group** button
- Select **Add pairs** and click **Finish**

Assembly	Homo sapiens (human) - hg19_chr22	▼
Format ℹ	BAMPE ▼	
Effective genome size ℹ	Human ▼ 2.7e9	
Data type ℹ	○ ATAC ● ChIP	

Define pairs to detect enriched regions

Select samples

	Sample name
☐	ChIP
☑	IGG

- Add IP sample
- Add IP group
- Add control group

Build pairs

IP sample	Clear all
ChIP	✖

vs.

Control sample	
IGG	✖

Add pairs Reset selected pairs

Detect regions for

IP	Control	Display name	Clear all
ChIP	IGG	ChIP_vs_IGG	✖

Advanced options

Option set	-- Default -- ▼	Configure

Back Finish

Notes:

Viewing Peak Results

- Double-click the **Peaks** data node to view the report
    - Each ChIP-Control pair will have one table report
    - Each row in a table is a region detected as a peak that passed the default criteria
    - Each table is sorted by –log10 pvalue
    - Click on the 🎿 button to view the peak in chromosome view

Optional columns

|	View	⇕ Sample name	⇕ Chromosome	⇕ Start	⇕ End	▼ -log10(pvalue)	⇕ Fold enrichment
		Search...	Search...	Search...	Search...	Search...	Search...
1	🎿	ChIP_vs_IGG	22	50638871	50640269	44.89988	20.31955
2	🎿	ChIP_vs_IGG	22	36783579	36783989	40.07011	18.47232
3	🎿	ChIP_vs_IGG	22	31884921	31886689	37.34150	10.97828
4	🎿	ChIP_vs_IGG	22	50699516	50700287	33.48045	15.16395
5	🎿	ChIP_vs_IGG	22	46932456	46933530	31.01766	14.21620
6	🎿	ChIP_vs_IGG	22	39151508	39152036	30.08033	12.65124
7	🎿	ChIP_vs_IGG	22	41487854	41488676	30.08033	12.65124
8	🎿	ChIP_vs_IGG	22	46972671	46973346	30.08033	12.65124
9	🎿	ChIP_vs_IGG	22	21896929	21897472	29.89290	8.98223
10	🎿	ChIP_vs_IGG	22	37914280	37915657	28.07369	10.85044
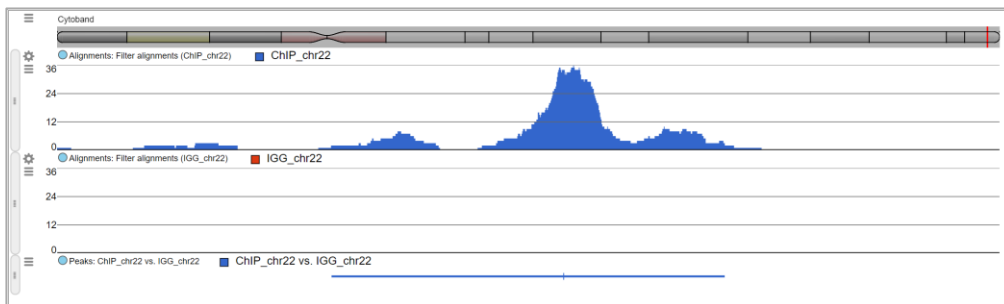
Notes:

Filtering Peaks

- In the *Fold enrichment* column, type **>4** and press enter
  - The table will be filtered to only show peaks with fold enrichment >4
- Select the red **Generate filtered node** button
  - This will create a new data Filtered peaks data node

Peaks   155 (of 541)

Optional columns

	View	⇕ Sample name	⇕ Chromosome	⇕ Start	⇕ End	▲ Length	⇕ -log10(pvalue)	⇕ Fold enrichment
		Search...	Search...	Search...	Search...	Search...	Search...	>4
1	∿	ChIP_vs_IGG	22	39101991	39102221	230	10.70387	7.05512
2	∿	ChIP_vs_IGG	22	39746050	39746280	230	8.39202	4.73873
3	∿	ChIP_vs_IGG	22	41485454	41485693	239	9.88721	4.86586
4	∿	ChIP_vs_IGG	22	35937189	35937434	245	7.52153	4.61808
5	∿	ChIP_vs_IGG	22	20067809	20068056	247	13.29735	7.38893
6	∿	ChIP_vs_IGG	22	42322028	42322278	250	8.07242	5.27317
7	∿	ChIP_vs_IGG	22	38082641	38082898	257	8.39202	4.73873
8	∿	ChIP_vs_IGG	22	19166050	19166308	258	9.18866	6.27121
9	∿	ChIP_vs_IGG	22	29168768	29169028	260	9.38303	5.54170
10	∿	ChIP_vs_IGG	22	38901574	38901838	264	10.84787	6.38404
19	∿	ChIP_vs_IGG	22	24195960	24196238	278	14.19379	8.69625
20	∿	ChIP_vs_IGG	22	24552183	24552462	279	10.70387	7.05512
21	∿	ChIP_vs_IGG	22	38851733	38852014	281	6.23019	4.34172
22	∿	ChIP_vs_IGG	22	21871284	21871567	283	13.92548	7.48421
23	∿	ChIP_vs_IGG	22	42509520	42509805	285	21.70336	11.08339
24	∿	ChIP_vs_IGG	22	37956462	37956752	290	9.88721	4.86586
25	∿	ChIP_vs_IGG	22	31090913	31091204	291	6.43219	4.39431

Rows per page   25 ▾ |◀   ◀◀   (1 of 7)   ▶▶   ▶| Download

▼ **Generate filtered node**

Notes:

Peak Annotation

- Click on the **Filtered Peaks** data node and choose **Annotate peaks** in the *Peak analysis* section

- Choose the RefSeq annotation and leave everything else as default, click **Finish**

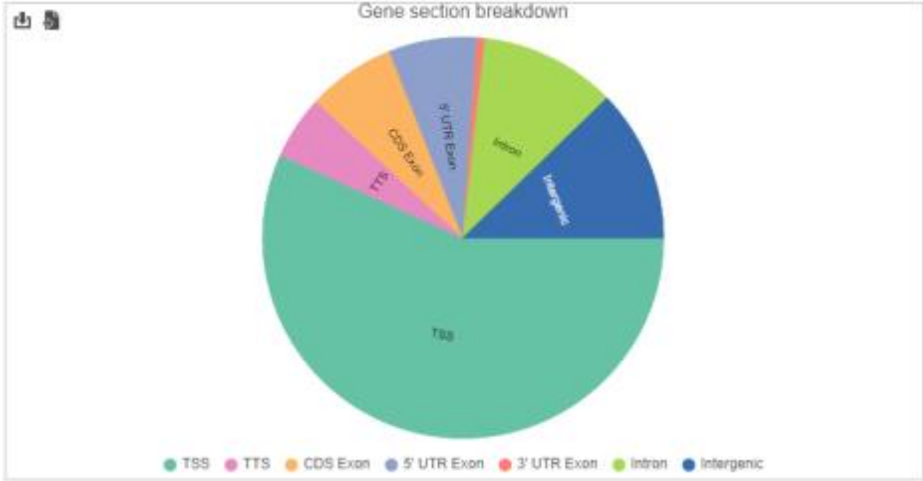- An **Annotated peaks** data node will be generated, double click on it to view the report

Notes:

TSS plot

- Click on **Annotated peaks** data node to choose **TSS Plot** in the *Exploratory analysis* section
- Use the default settings, click **Finish**
- Click on TSS plot task report
    - Profile plot of all the selected samples are on the top
    - Heatmap of selected samples are at the bottom
    - Configure the heatmap per row based on screen resolution
    - Change the low/high value to adjust heatmap color
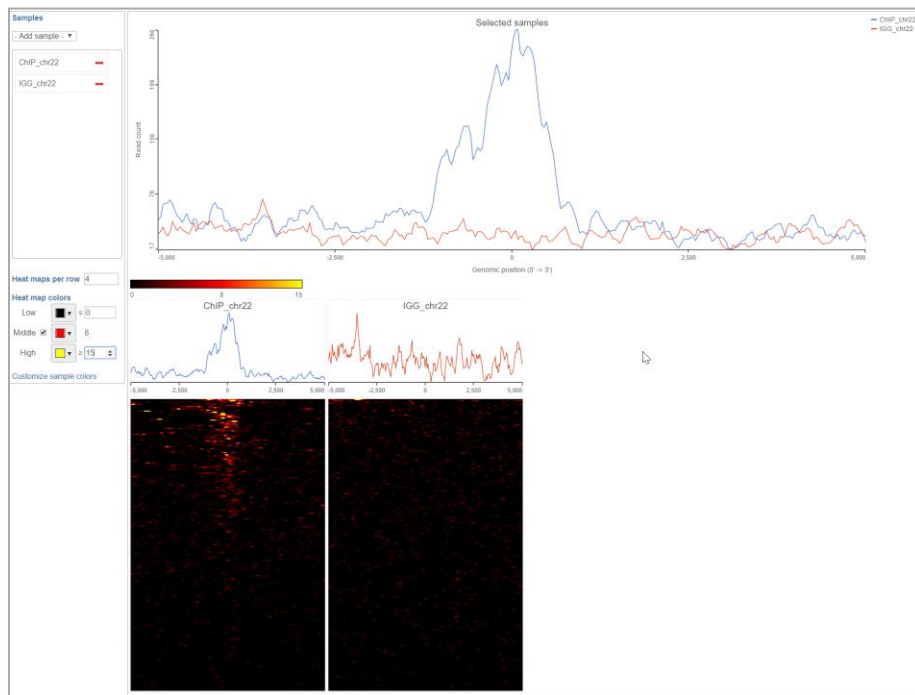
Notes:

Motif Detection – Search for known motifs

- Click the **Annotated peaks** data node to choose the **Search for known motifs** in the *Motif detection* section
- Chose the **All CORE profiles** database, use the default settings, click **Finish**
  - An alignment matrix is used to match sequences in peaks against the JASPAR motif database
- Double-click the new task node to view the report
- Clicking on a **motif name** opens the JASPAR database page for that motif

Motif name ⇕	Consensus sequence ⇕	p-value ▲
ZNF263 (MA0528.1)	RRRGGAGGRNDRDVDDRRRRR	2.59E-292
BPC1 (MA1404.1)	RAGAGAGAGAGAGAGAGAGAGA	6.87E-177
BPC5 (MA1403.1)	AGAGAGAGAGAGAGAGAGAGAGAGAG	2.59E-163
AT1G71450 (MA1233.1)	HNNCDHCDHHDYCDCCGHCD	4.4E-150
ZNF384 (MA1125.1)	NNNNAAAAAANN	3.52E-112
HMG-I/Y (MA0045.1)	NDWVRRRNRVMRDMRH	9.17E-95
hb (MA0049.1)	BVVHAAAAAN	3.06E-86
eor-1 (MA0543.1)	NNRGAGAVRVAGAVR	1.44E-83
BPC6 (MA1402.1)	HTCTYTCTCTCTCTCTCTCTM	1.05E-78

Summary | **Detail**

Chromosome ⇕	Start ⇕	End ⇕	Strand ⇕	Motif ID ⇕	Instance sequence ⇕	Score ⇕	
22	39,639,288	39,639,302	-	RAMOSA1 (MA1416.1)	GAGAGAGAGAGAGA	1.00	⌁
22	39,639,286	39,639,300	-	RAMOSA1 (MA1416.1)	GAGAGAGAGAGAGA	1.00	⌁
22	39,639,284	39,639,298	-	RAMOSA1 (MA1416.1)	GAGAGAGAGAGAGA	1.00	⌁
22	39,639,282	39,639,296	-	RAMOSA1 (MA1416.1)	GAGAGAGAGAGAGA	1.00	⌁
22	39,639,280	39,639,294	-	RAMOSA1 (MA1416.1)	GAGAGAGAGAGAGA	1.00	⌁
22	39,639,264	39,639,278	-	RAMOSA1 (MA1416.1)	GAGAGAGAGAGAGA	1.00	⌁
22	39,639,262	39,639,276	-	RAMOSA1 (MA1416.1)	GAGAGAGAGAGAGA	1.00	⌁
22	39,639,260	39,639,274	-	RAMOSA1 (MA1416.1)	GAGAGAGAGAGAGA	1.00	⌁
22	39,639,258	39,639,272	-	RAMOSA1 (MA1416.1)	GAGAGAGAGAGAGA	1.00	⌁

Notes:

– 11 –

Motif Detection – Detect *de novo* motifs

- Click the **Annotated peaks** data node to choose the **Detect de novo motifs** in the *Motif detection* section
- Use the default settings, click **Finish**
  - Gibbs sampling is performed to detect enrichment of a subsequence across the peaks
- Double-click the new task node to view the report
- Click on a **sequence logo** to enlarge the image in a new tab

Assembly Homo sapiens (human) - hg19_chr22

Number of motifs ⓘ   1

Motif length ⓘ   6 bp to   16 bp

Back   Finish

| Summary | Detail |

Motif name ↕	Consensus sequence ↕	Log likelihood ratio ↕	Sequence logo
motif1	YTYYYYTBY	1,623.39	

Rows per page 10 ▼ ◄ ◄◄ (1 of 1) ►► ►►

Notes: _____

_____

_____

_____

_____

Further Training

Self-learning

- Help > Check for Updates

- Help > On-line tutorials

- Recorded webinars

Regional Technical Support

- www.partek.com/PartekSupport

Notes: _____

_____

_____

_____

_____