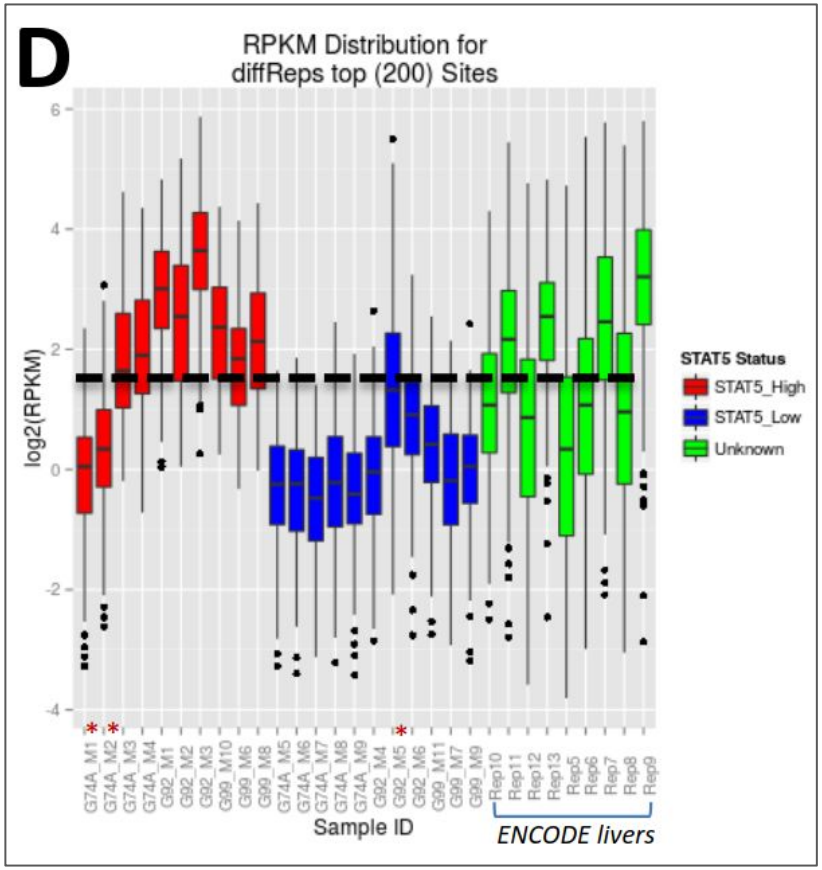
BoxPlotter for ChIPseq Regions-of-Interest

Based on the idea in the right picture, we decided to create a more general tool that would allow us to specify sets of regions and compare groups of different samples with each other.



DNase-seq data from 9 male livers (ENCODE) were compared to the STAT5-high/low profiles. **4 livers** showed chromatin accessibility patterns similar to **STAT5-high** livers, while **5 livers** resembled **STAT5-low** livers. These patterns were visualized using read count distributions across the top differential sites.

Andy Rampersaud, Jeannette Connerney, David J Waxman (2023) Plasma growth hormone pulses induce male-biased pulsatile chromatin opening and epigenetic regulation in adult mouse liver eLife 12:RP91367

Algorithm:

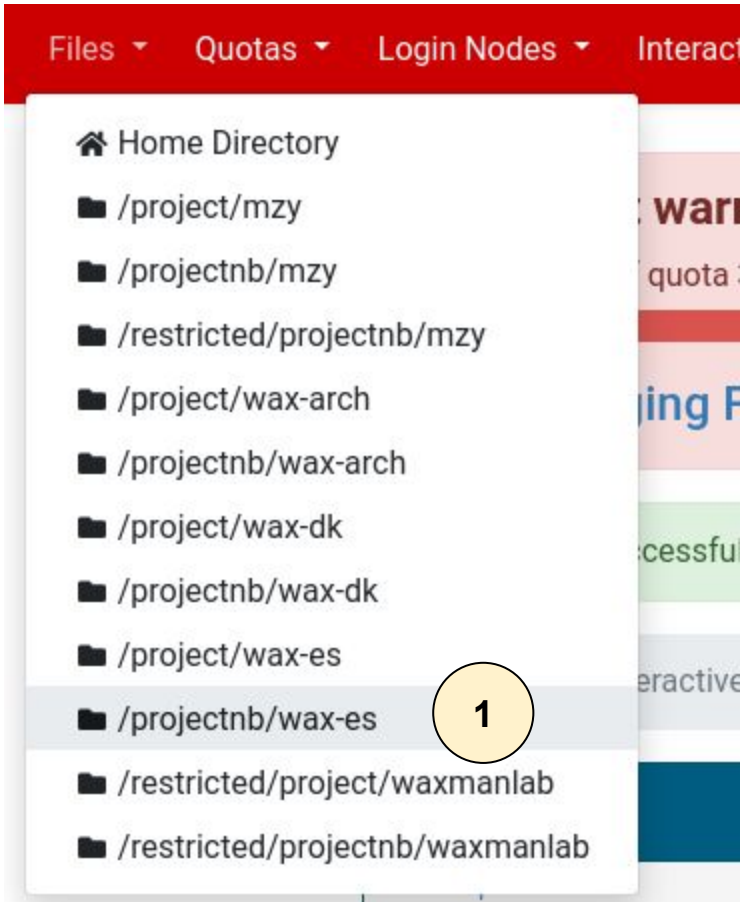
1. Calculate FRiP (Fraction of Reads in Peaks)
2. Extract fragment counts from BAM files over target regions for selected samples.
3. Normalize fragment counts using FRiP, generate aggregated and group-wise comparison plots, and compute statistics for group comparisons.

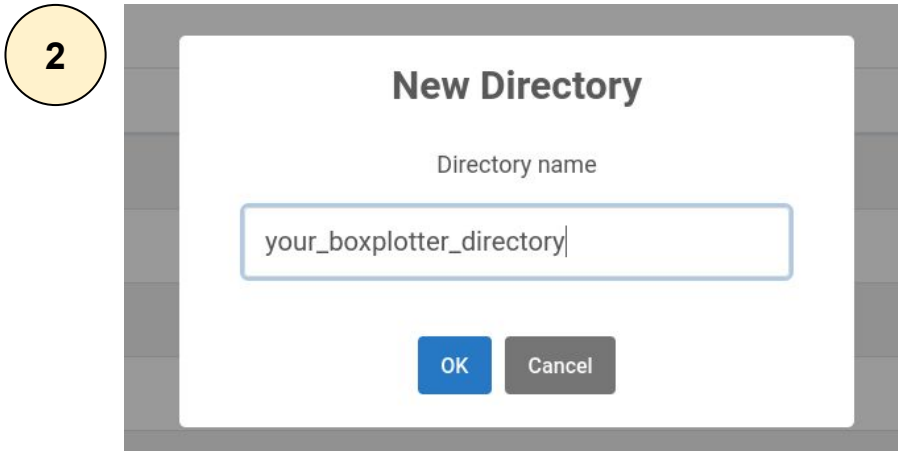
FRiP Normalization:

1. Call MACS2 peaks per sample
2. Merge to union peak set
3. Count fragments per sample in union
4. Normalize to sample-average fragment count

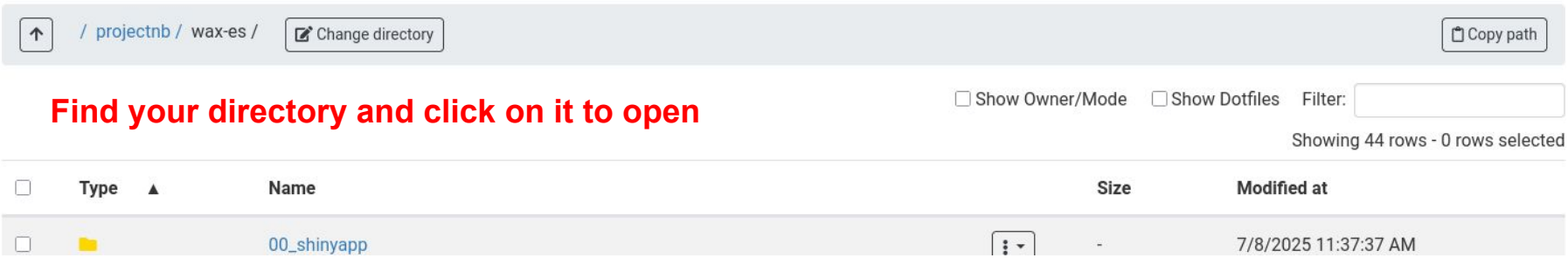
Data transformation during normalization

				G241_M01 (Male_2wk_ATAC)			
chr	start	end	width	RAW COUNTS	RAW_AFTER_WIDTH_NORM	AFTER_FRIP_NORM	FINAL_LOG2_AFTER_RIPPM_NORM
chr1	40293477	40294208	732	706	964.481	395.799	8.629
chr1	40299557	40299948	392	257	655.612	269.047	8.072
chr1	41608251	41608848	598	81	135.452	55.586	5.797
chr1	55136533	55136946	414	384	927.536	380.637	8.572
chr1	82154931	82155388	458	234	510.917	209.667	7.712
chr1	83896719	83897207	489	16	32.720	13.427	3.747
chr1	129796083	129796418	336	158	470.238	192.974	7.592
chr1	133837634	133838008	375	182	485.333	199.169	7.638
chr1	166251542	166251959	418	172	411.483	168.862	7.400
chr1	174860030	174860572	543	173	318.600	130.746	7.031
chr1	176420085	176420403	319	40	125.392	51.458	5.685
chr1	188149410	188149823	414	54	130.435	53.527	5.742
chr10	5728314	5728809	496	147	296.371	121.623	6.926

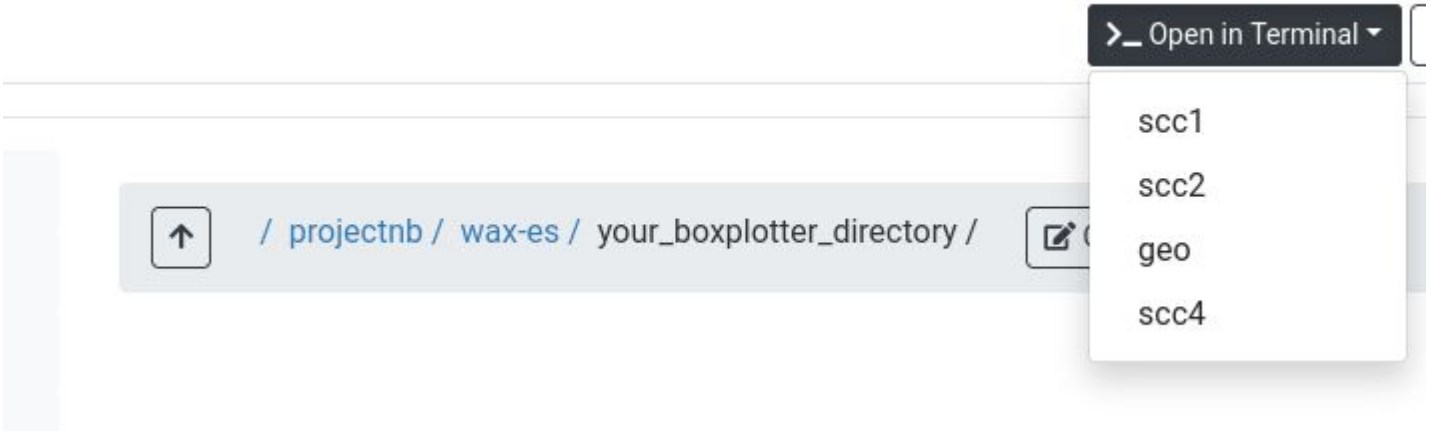

Files
Quotas
Login Nodes
Home Directory
/project/mzy
/projectnb/mzy
/restricted/projectnb/mzy
/project/wax-arch
/projectnb/wax-arch
/project/wax-dk
/projectnb/wax-dk
/project/wax-es
/projectnb/wax-es
/restricted/project/waxmanlab
/restricted/projectnb/waxmanlab
1


2
New Directory
Directory name
your_boxplotter_directory
OK
Cancel


Open in Terminal
Refresh
New File
New Directory
Upload
Download
Globus
Copy/Move
Delete


/ projectnb / wax-es /
Change directory
Copy path
Find your directory and click on it to open
Show Owner/Mode
Show Dotfiles
Filter:
Showing 44 rows - 0 rows selected
Type
Name
Size
Modified at
00_shinyapp
-
7/8/2025 11:37:37 AM

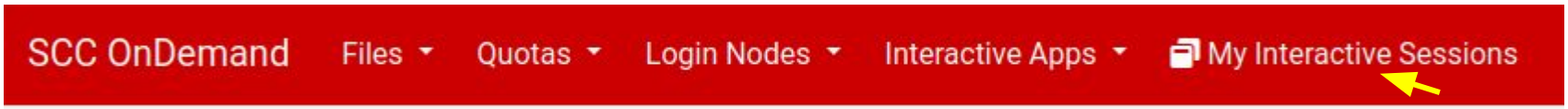
Make sure you're in the correct directory and open the Terminal



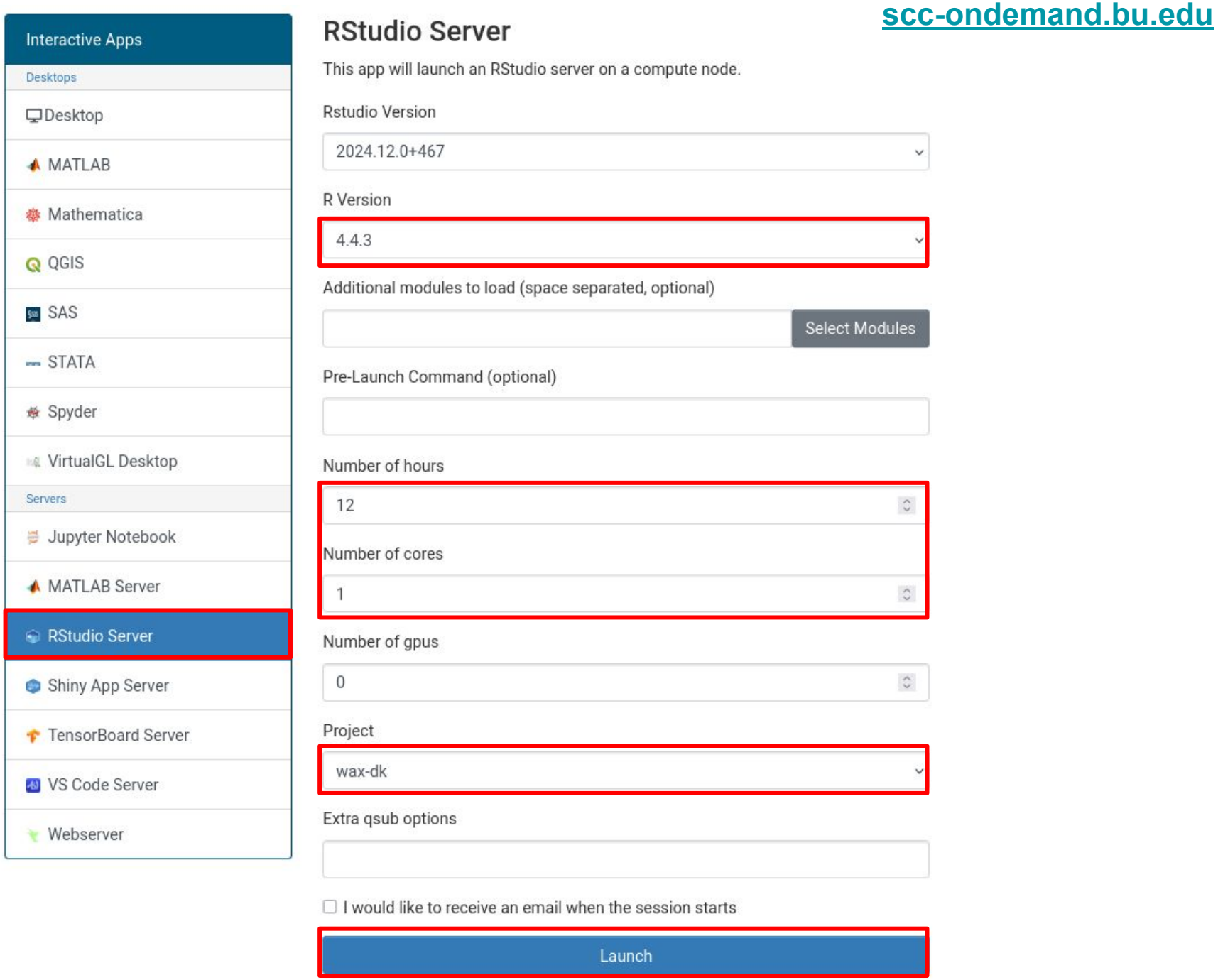
To clone Boxplotter into the current directory, enter the following command and press Enter:

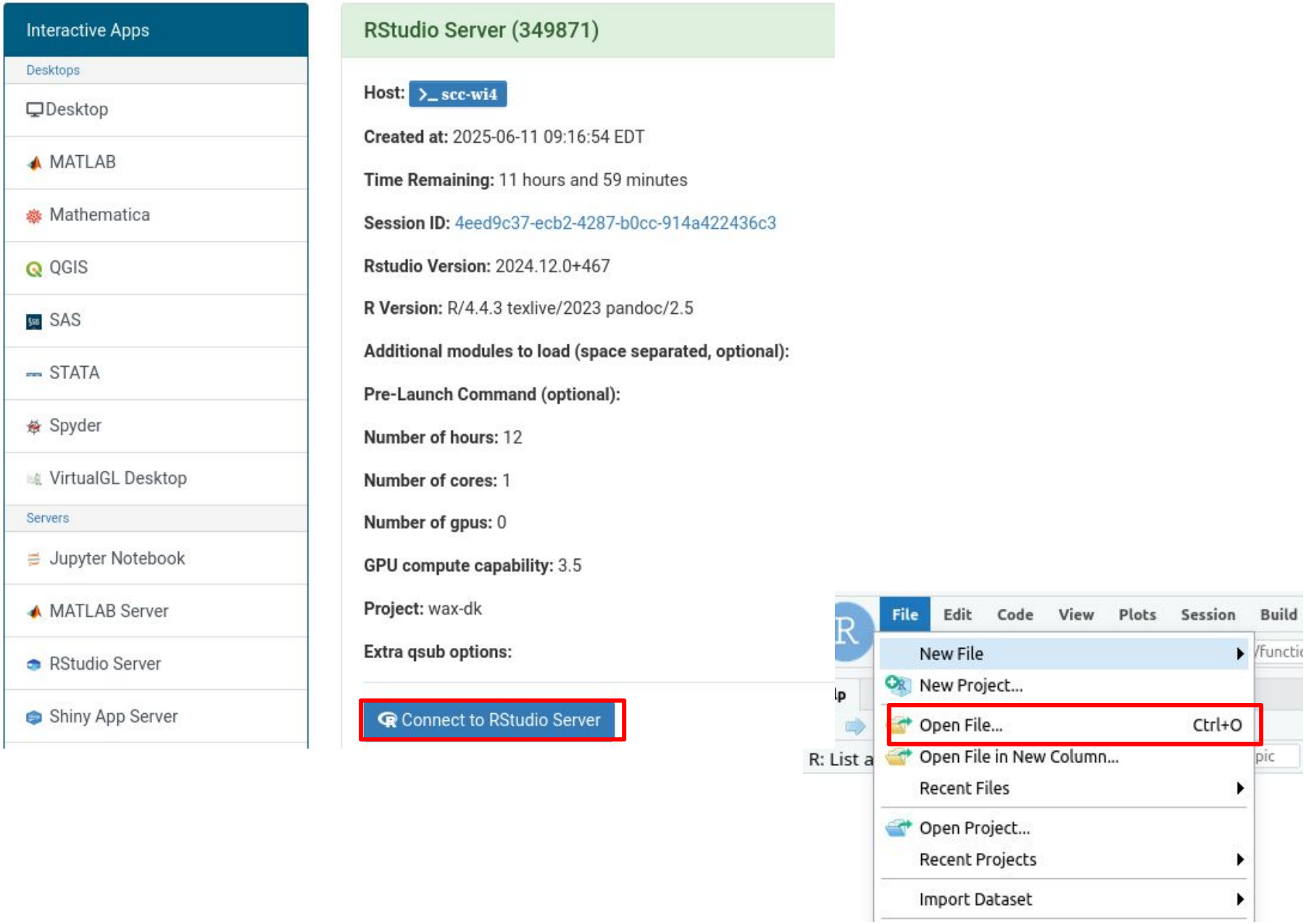
```
git clone /projectnb/wax-es/WAXMANLAB_SOFT/BoxPlotter_v2 .
```

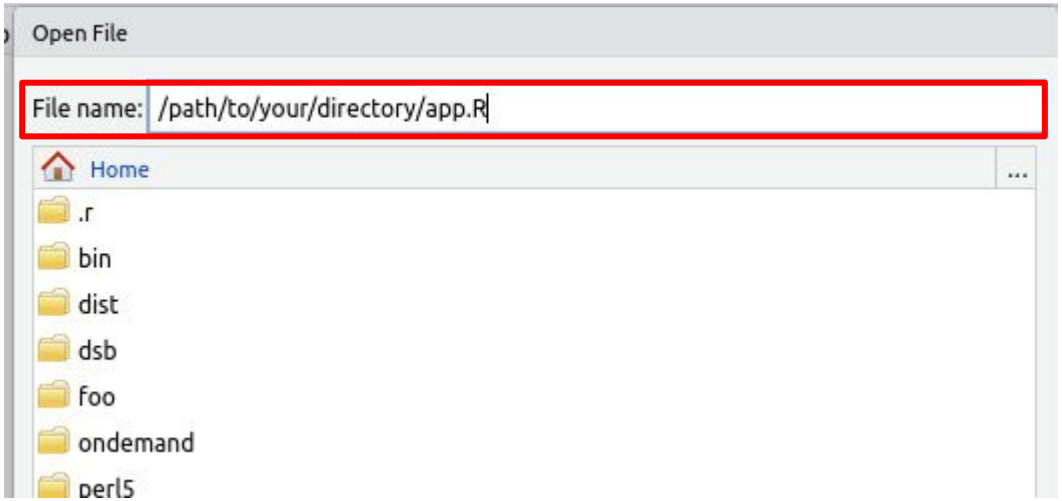
After cloning the directory you can click to “My interactive sessions” on the top of the page



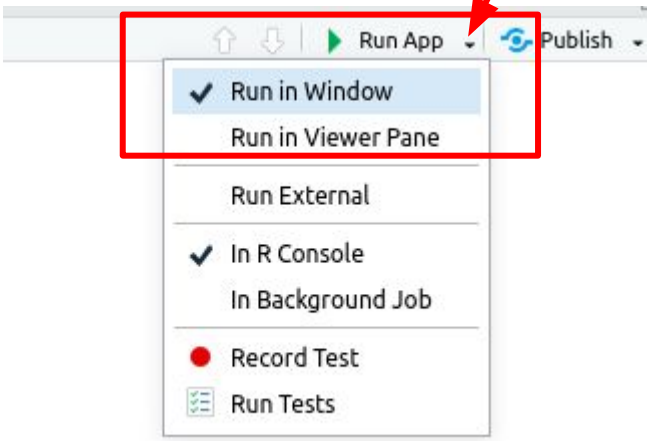
scc-ondemand.bu.edu



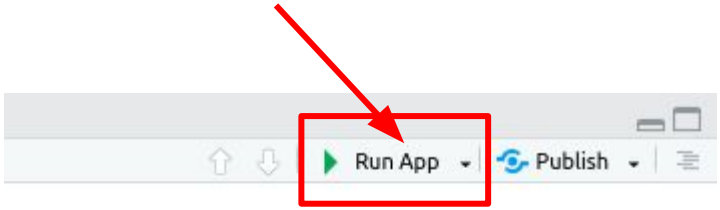

Interactive Apps
Desktops
Desktop
MATLAB
Mathematica
QGIS
SAS
STATA
Spyder
VirtualGL Desktop
Servers
Jupyter Notebook
MATLAB Server
RStudio Server
Shiny App Server
RStudio Server (349871)
Host: >_ scc-wi4
Created at: 2025-06-11 09:16:54 EDT
Time Remaining: 11 hours and 59 minutes
Session ID: 4eed9c37-ecb2-4287-b0cc-914a422436c3
Rstudio Version: 2024.12.0+467
R Version: R/4.4.3 texlive/2023 pandoc/2.5
Additional modules to load (space separated, optional):
Pre-Launch Command (optional):
Number of hours: 12
Number of cores: 1
Number of gpus: 0
GPU compute capability: 3.5
Project: wax-dk
Extra qsub options:
Connect to RStudio Server
File
Edit
Code
View
Plots
Session
Build
New File
New Project...
Open File... Ctrl+O
Open File in New Column...
Recent Files
Open Project...
Recent Projects
Import Dataset



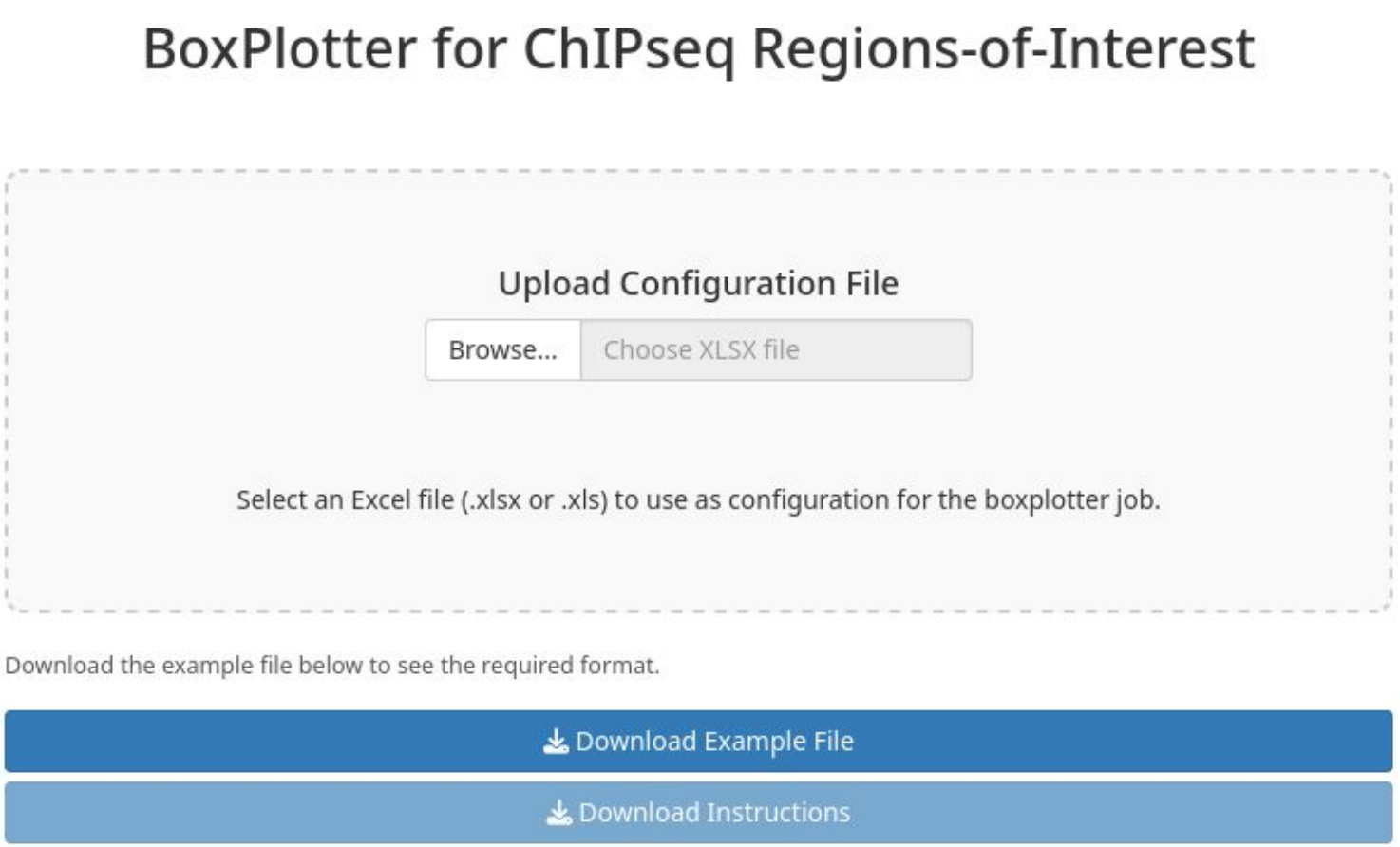
Click to small arrow to open dialog and select “Run in Window”

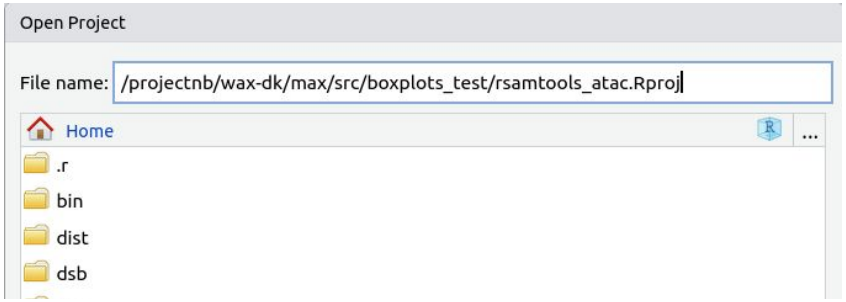


Click “Run App” to start shiny app

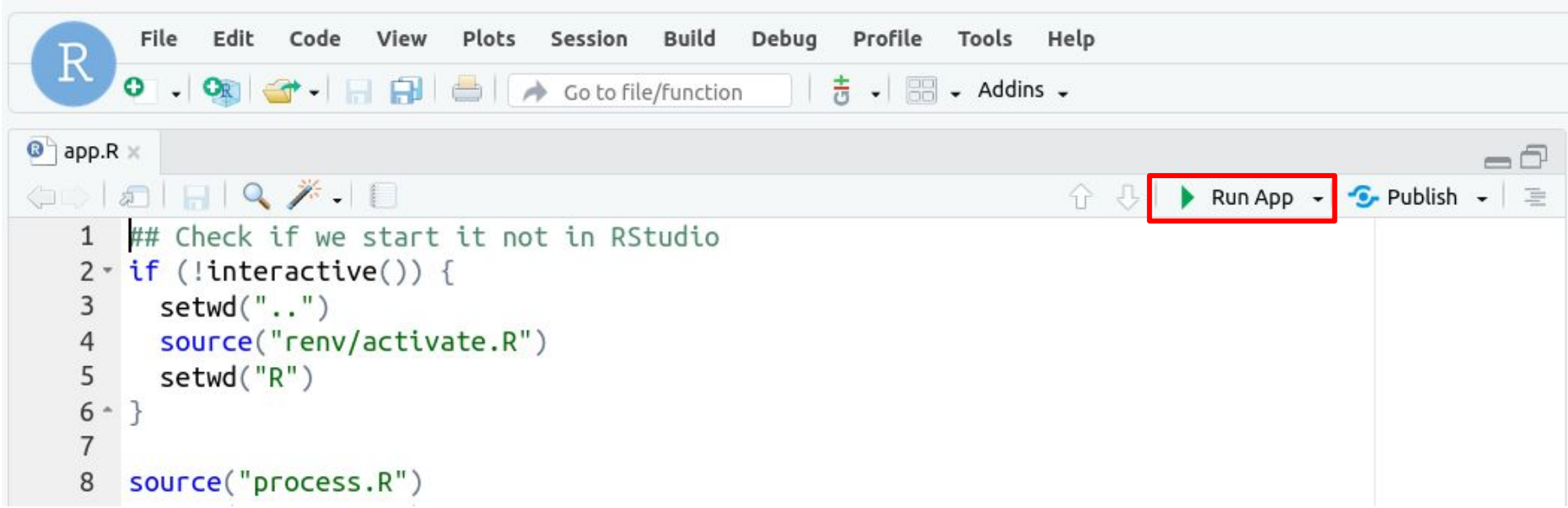


BoxPlotter v2 shinyapp

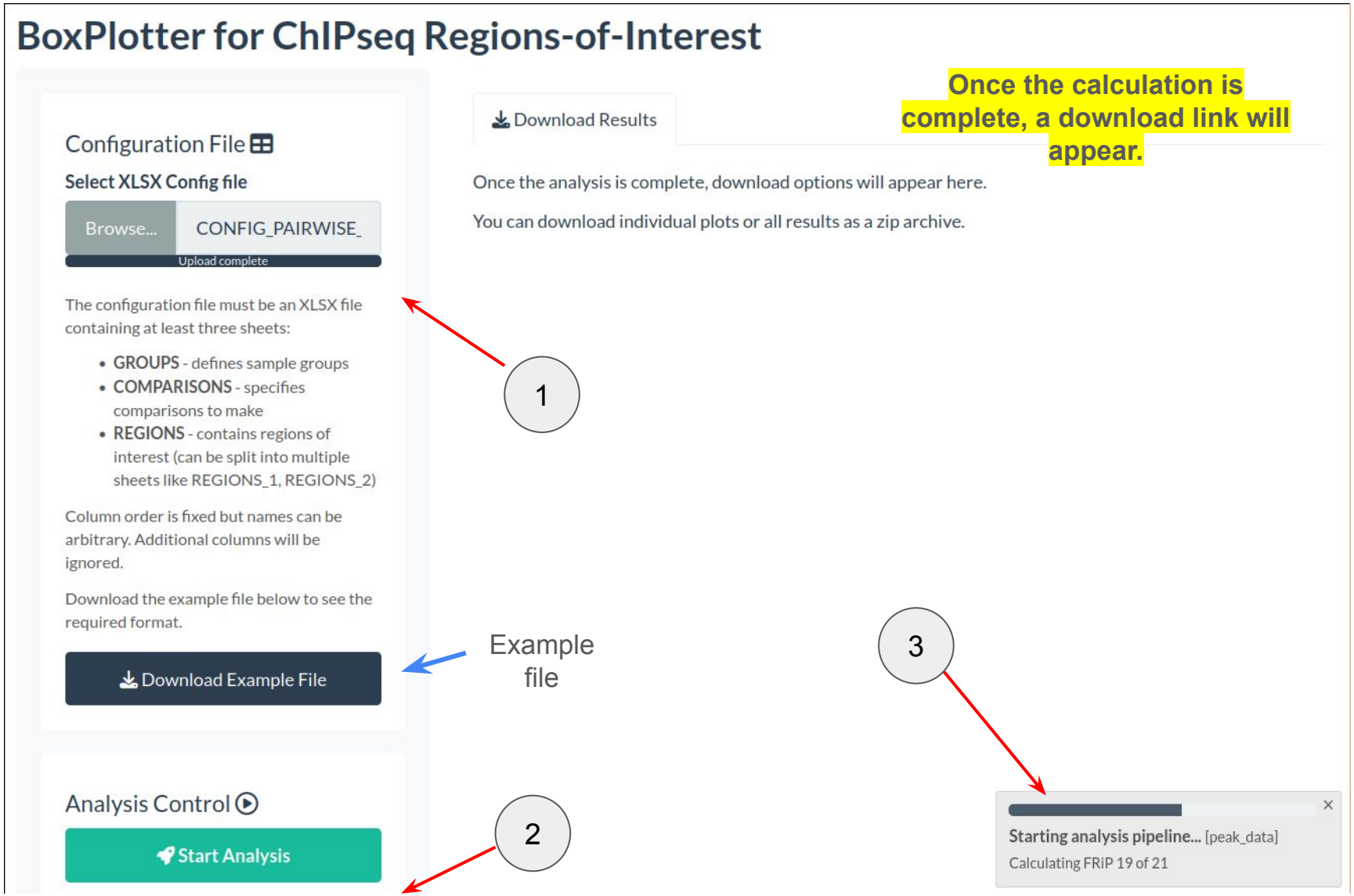




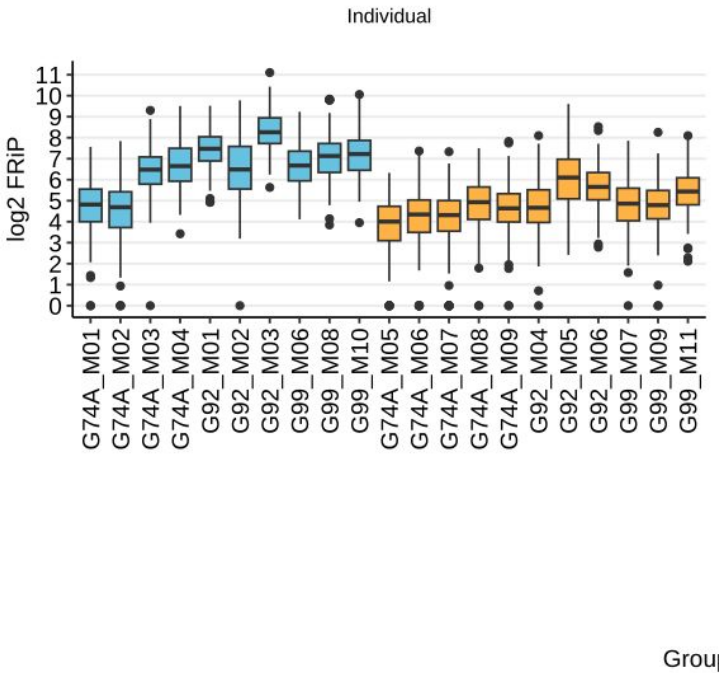
/projectnb/wax-dk/max/src/boxplots_test/rsamtools_atac.Rproj

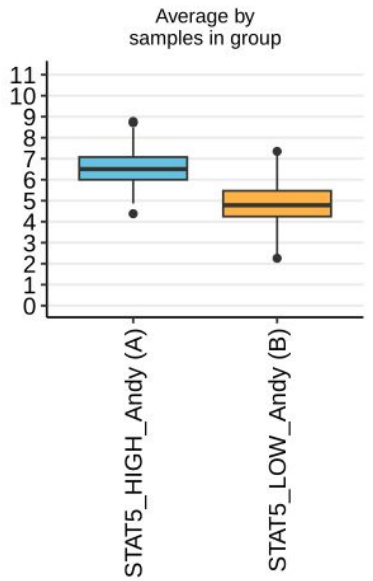


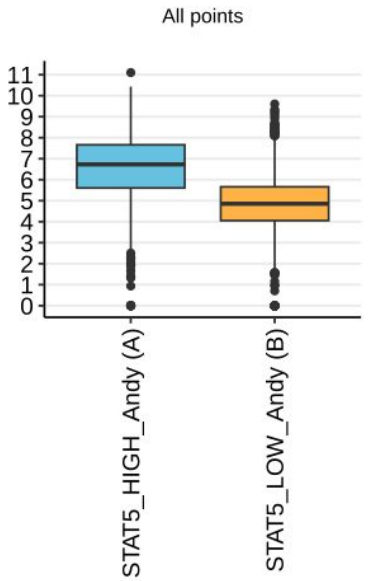
How to start application



1. (A vs B) (A_MB_Dynamic_1 (n=834)_top200)

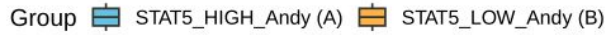




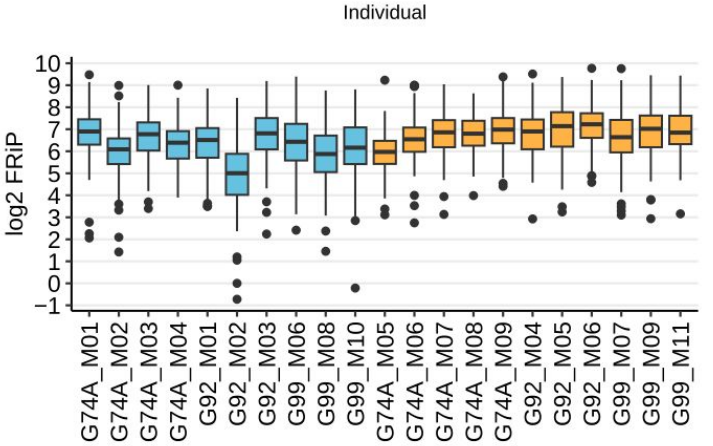


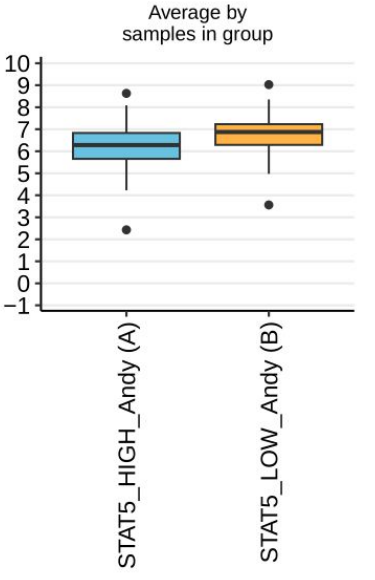
test	average	all_points
wilcox	6.08E−54	2.75E−261
KS	1.24E−54	1.15E−238
t.test	1.57E−43	1.16E−153
Median ratio (linear scale)	3.972	3.665

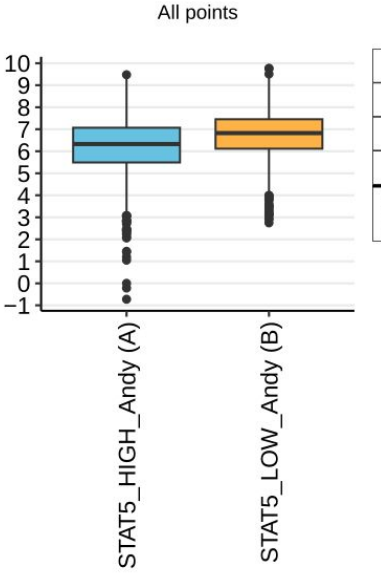
**Statistics based on linear scale (no log2 transformation as on plots)*



1. (A vs B) (B_MB_Static_Group 1 (n=172))

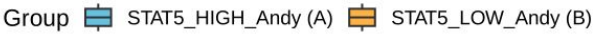




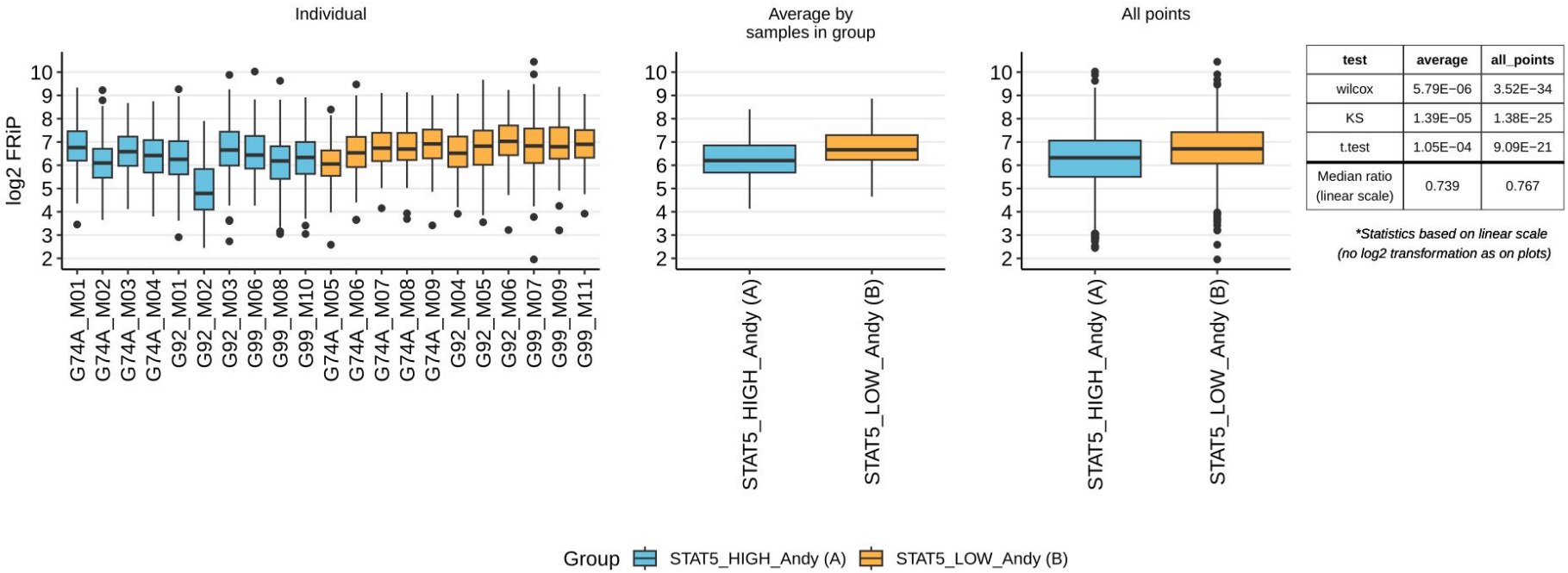


test	average	all_points
wilcox	4.42E−08	1.85E−40
KS	3.04E−06	2.88E−27
t.test	2.43E−07	1.45E−30
Median ratio (linear scale)	0.739	0.708

**Statistics based on linear scale (no log2 transformation as on plots)*



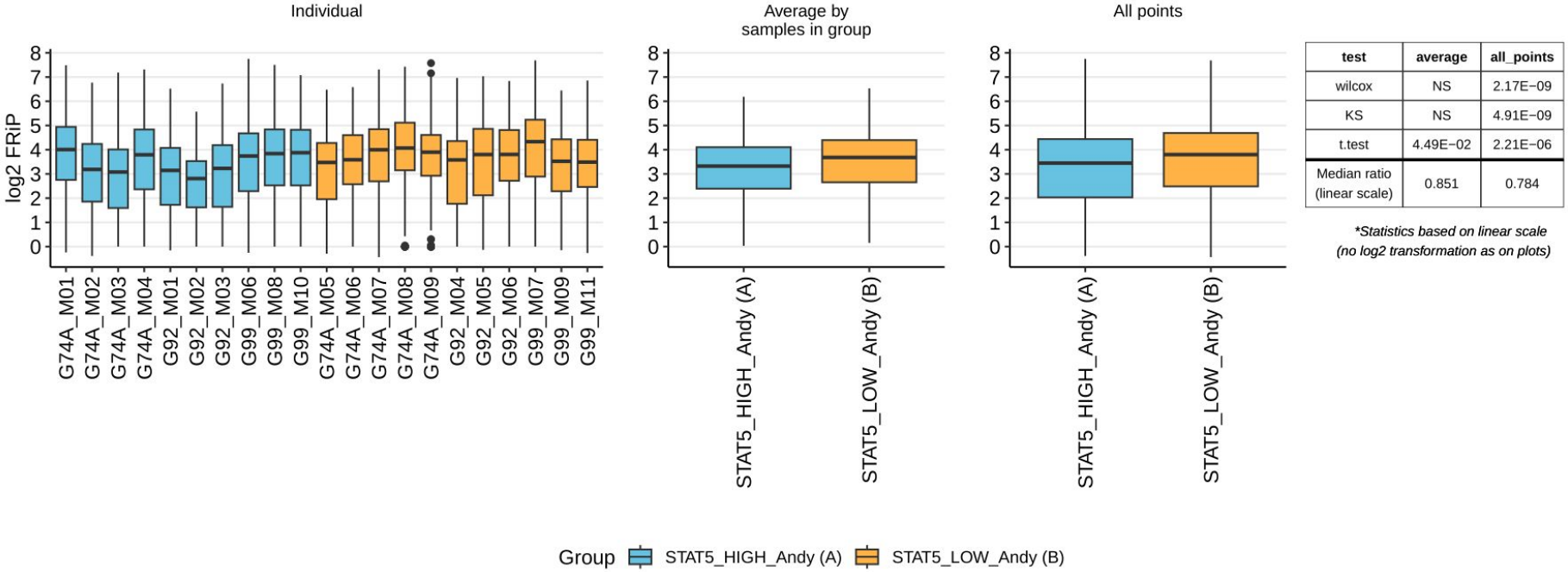
1. (A vs B) (C_MB Static_Group 2_(n=186))



test	average	all_points
wilcox	5.79E−06	3.52E−34
KS	1.39E−05	1.38E−25
t.test	1.05E−04	9.09E−21
Median ratio (linear scale)	0.739	0.767

**Statistics based on linear scale (no log2 transformation as on plots)*

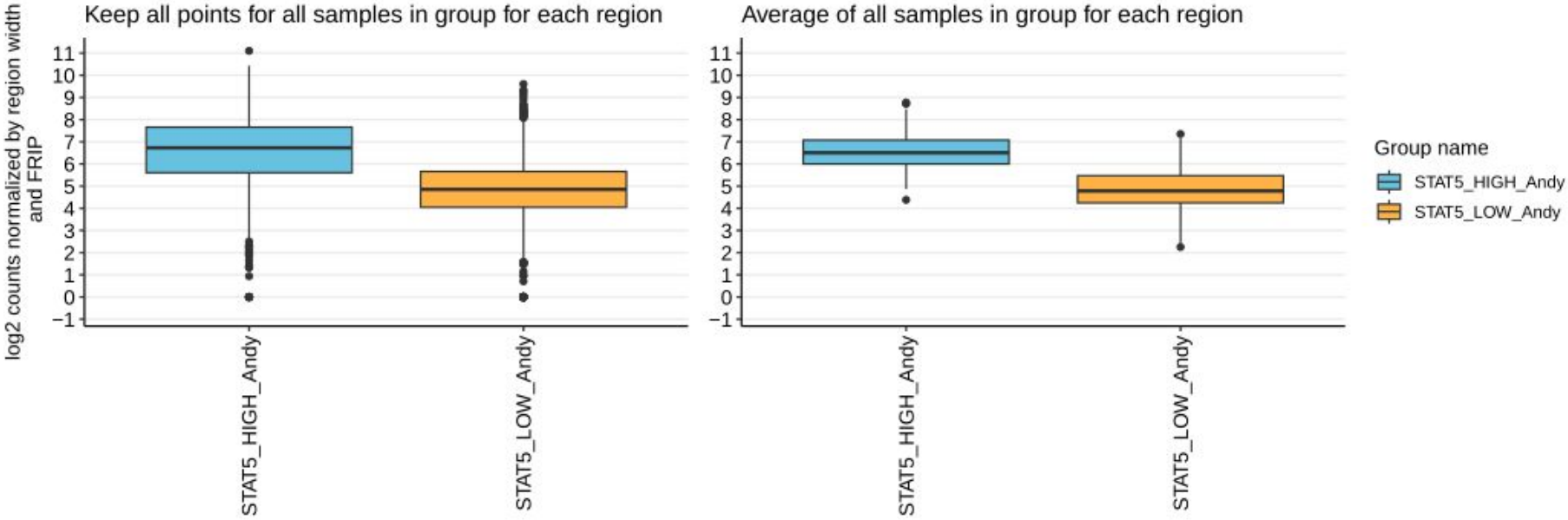
1. (A vs B) (D_FB_Static_(top n=706)_top200)



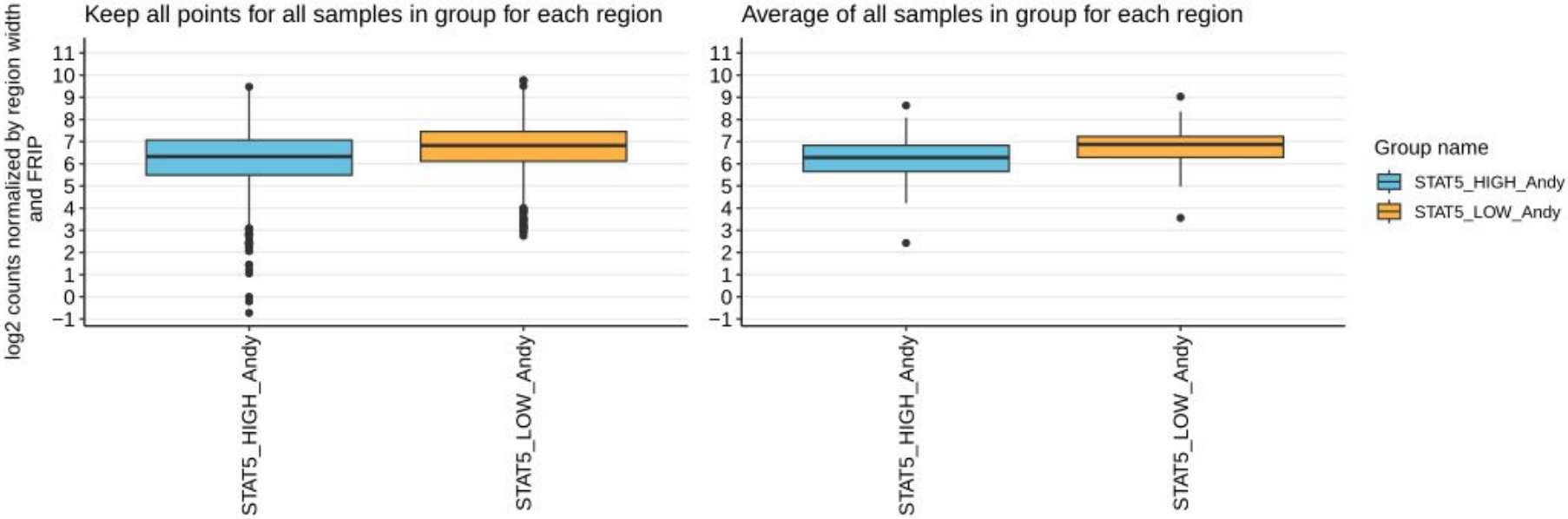
test	average	all_points
wilcox	NS	2.17E−09
KS	NS	4.91E−09
t.test	4.49E−02	2.21E−06
Median ratio (linear scale)	0.851	0.784

**Statistics based on linear scale (no log2 transformation as on plots)*

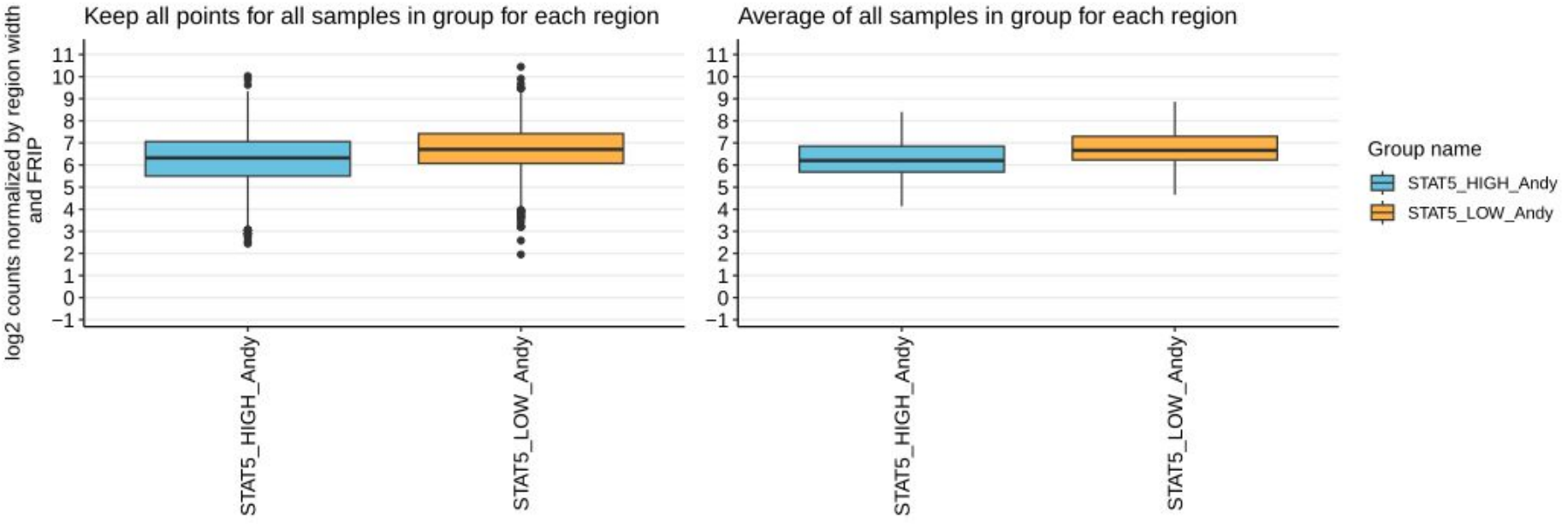
A_MB_Dynamic_1 (n=834)_top200



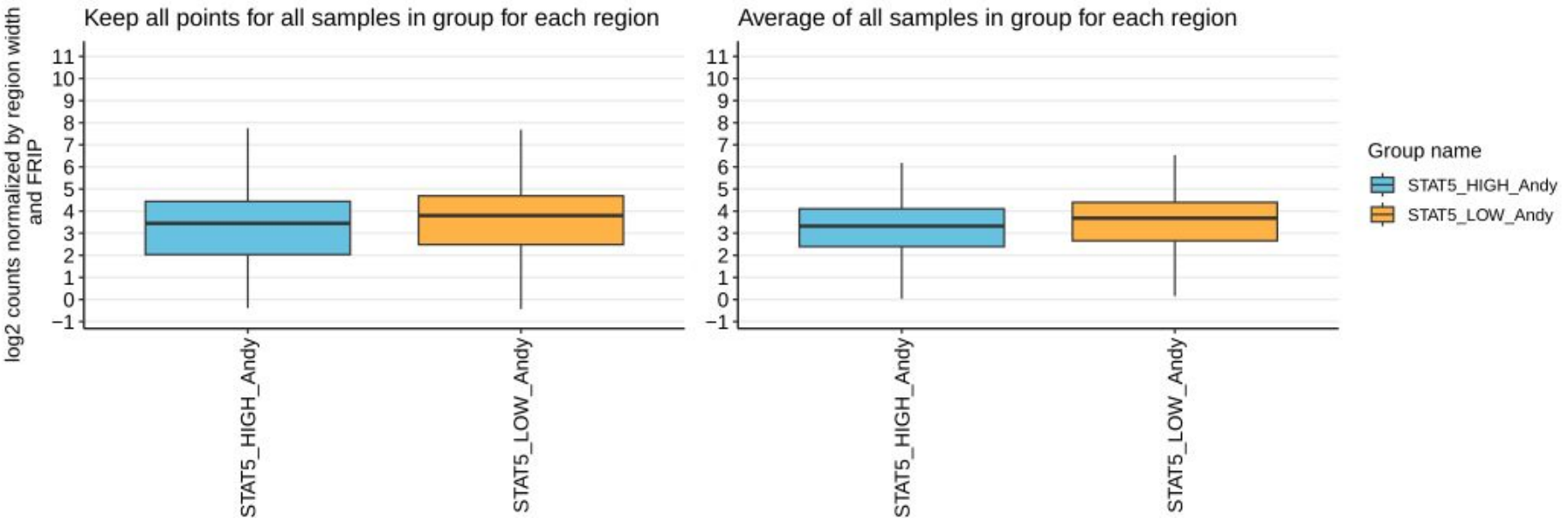
B_MB_Static_Group 1 (n=172)

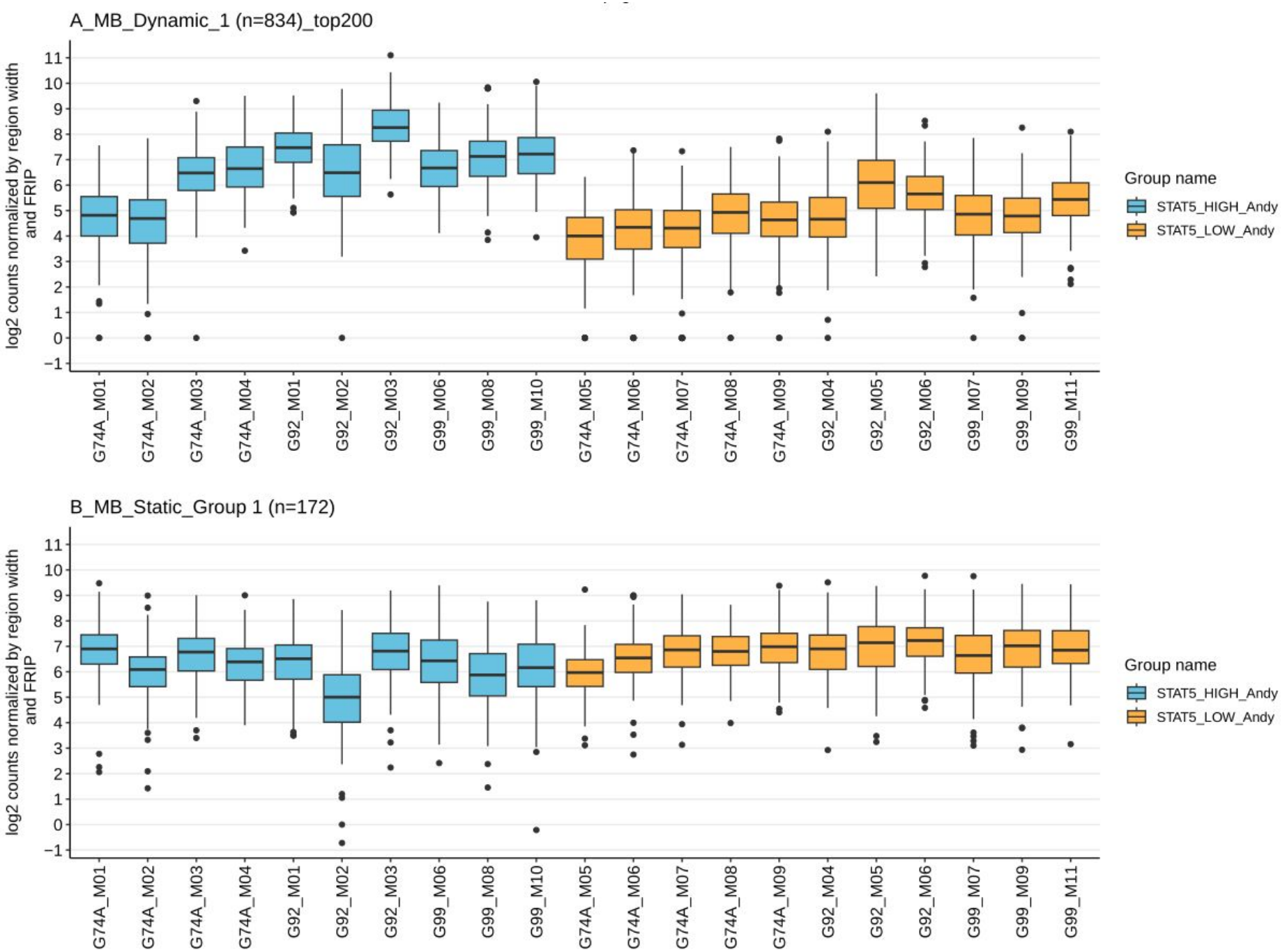


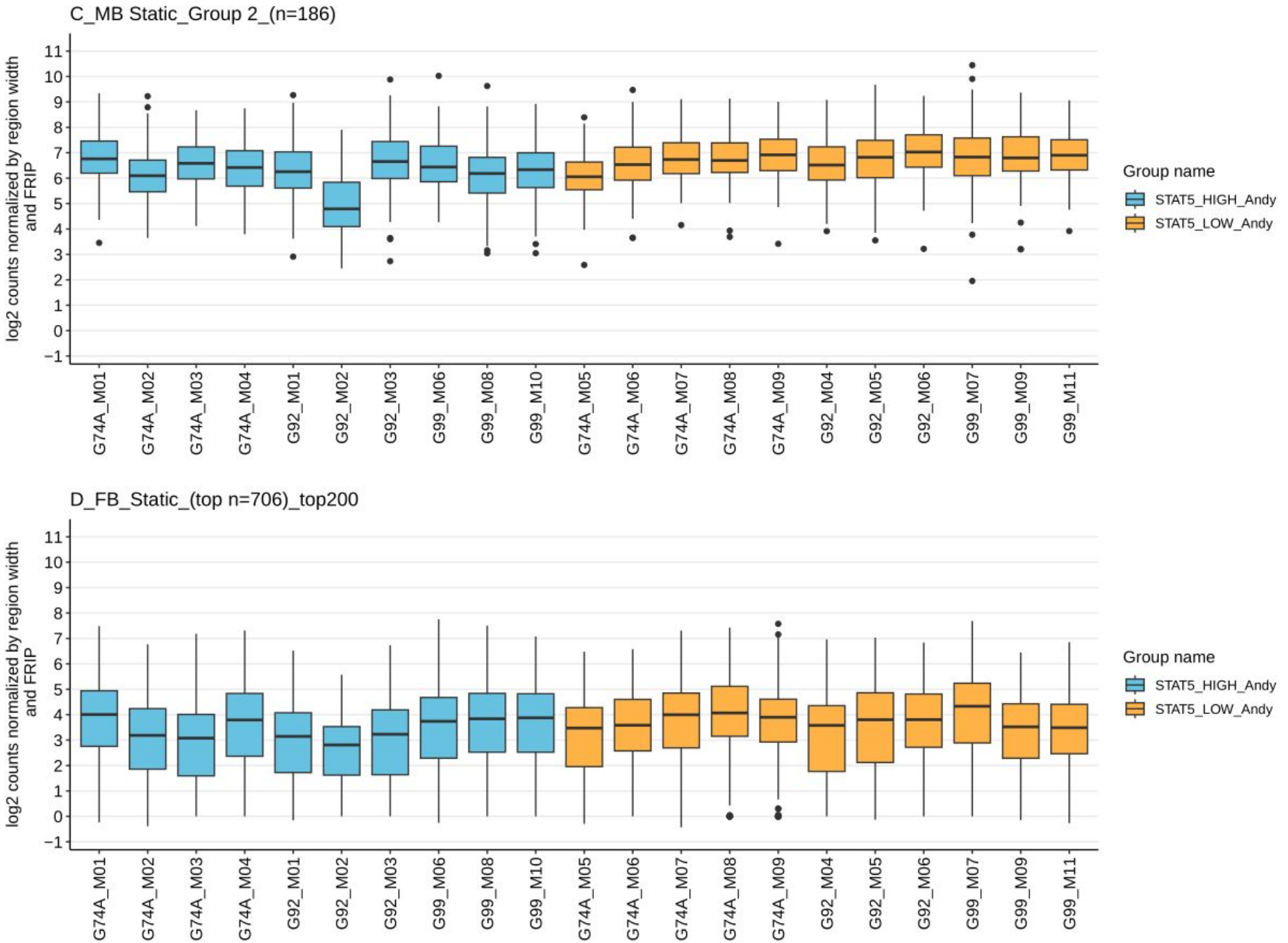
C_MB Static_Group 2_(n=186)



D_FB_Static_(top n=706)_top200




A_MB_Dynamic_1 (n=834)_top200
log2 counts normalized by region width and FRIP
11
10
9
8
7
6
5
4
3
2
1
0
−1
G74A_M01
G74A_M02
G74A_M03
G74A_M04
G92_M01
G92_M02
G92_M03
G99_M06
G99_M08
G99_M10
G74A_M05
G74A_M06
G74A_M07
G74A_M08
G74A_M09
G92_M04
G92_M05
G92_M06
G99_M07
G99_M09
G99_M11
Group name
STAT5_HIGH_Andy
STAT5_LOW_Andy
B_MB_Static_Group 1 (n=172)
log2 counts normalized by region width and FRIP
11
10
9
8
7
6
5
4
3
2
1
0
−1
G74A_M01
G74A_M02
G74A_M03
G74A_M04
G92_M01
G92_M02
G92_M03
G99_M06
G99_M08
G99_M10
G74A_M05
G74A_M06
G74A_M07
G74A_M08
G74A_M09
G92_M04
G92_M05
G92_M06
G99_M07
G99_M09
G99_M11
Group name
STAT5_HIGH_Andy
STAT5_LOW_Andy


C_MB Static_Group 2_(n=186)
log2 counts normalized by region width and FRIP
11
10
9
8
7
6
5
4
3
2
1
0
−1
G74A_M01
G74A_M02
G74A_M03
G74A_M04
G92_M01
G92_M02
G92_M03
G99_M06
G99_M08
G99_M10
G74A_M05
G74A_M06
G74A_M07
G74A_M08
G74A_M09
G92_M04
G92_M05
G92_M06
G99_M07
G99_M09
G99_M11
Group name
STAT5_HIGH_Andy
STAT5_LOW_Andy
D_FB_Static_(top n=706)_top200
log2 counts normalized by region width and FRIP
11
10
9
8
7
6
5
4
3
2
1
0
−1
G74A_M01
G74A_M02
G74A_M03
G74A_M04
G92_M01
G92_M02
G92_M03
G99_M06
G99_M08
G99_M10
G74A_M05
G74A_M06
G74A_M07
G74A_M08
G74A_M09
G92_M04
G92_M05
G92_M06
G99_M07
G99_M09
G99_M11
Group name
STAT5_HIGH_Andy
STAT5_LOW_Andy