

BoxPlotter_v2 tutorial

[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

warning

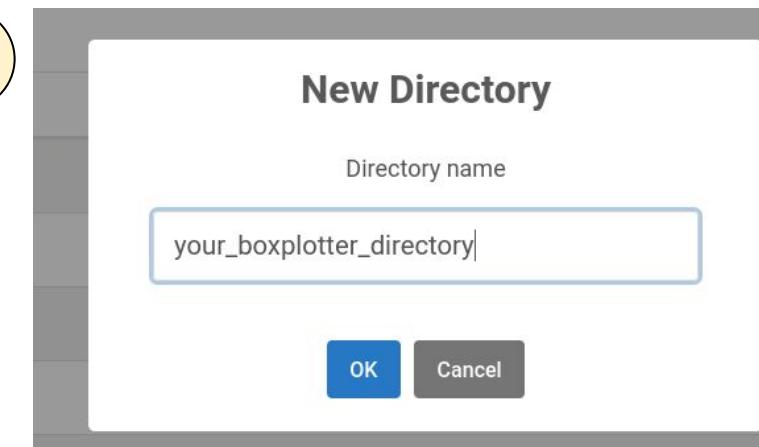
quota

warning

successful

interactive

2

[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 |

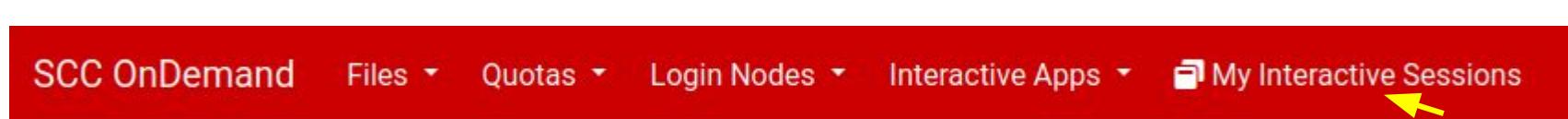
Make sure you're in the correct directory and open the Terminal



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



Interactive Apps

- Desktops
- 💻 Desktop
- ⚠ MATLAB
- ✳ Mathematica
- 🔍 QGIS
- 📊 SAS
- == STATA
- ✳ Spyder
- VirtualGL Desktop
- Servers
- ❗ Jupyter Notebook
- ⚠ MATLAB Server
- 🌐 RStudio Server
- 🌐 Shiny App Server
- 👉 TensorBoard Server
- 🌐 VS Code Server
- 🌐 Webserver

RStudio Server

This app will launch an RStudio server on a compute node.

Rstudio Version

2024.12.0+467

R Version

4.4.3

Additional modules to load (space separated, optional)

Select Modules

Pre-Launch Command (optional)

Number of hours

12

Number of cores

1

Number of gpus

0

Project

wax-dk

Extra qsub options

I would like to receive an email when the session starts

Launch

Interactive Apps

Desktops



— STATA



Servers



RStudio Server (349871)

Host: >_sec-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

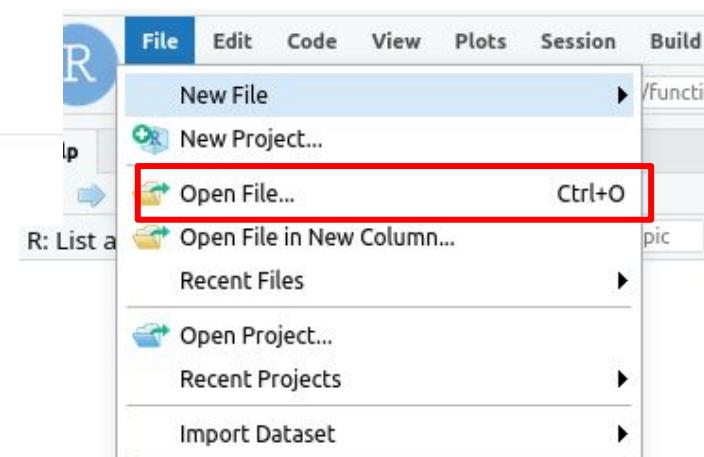
Number of gpus: 0

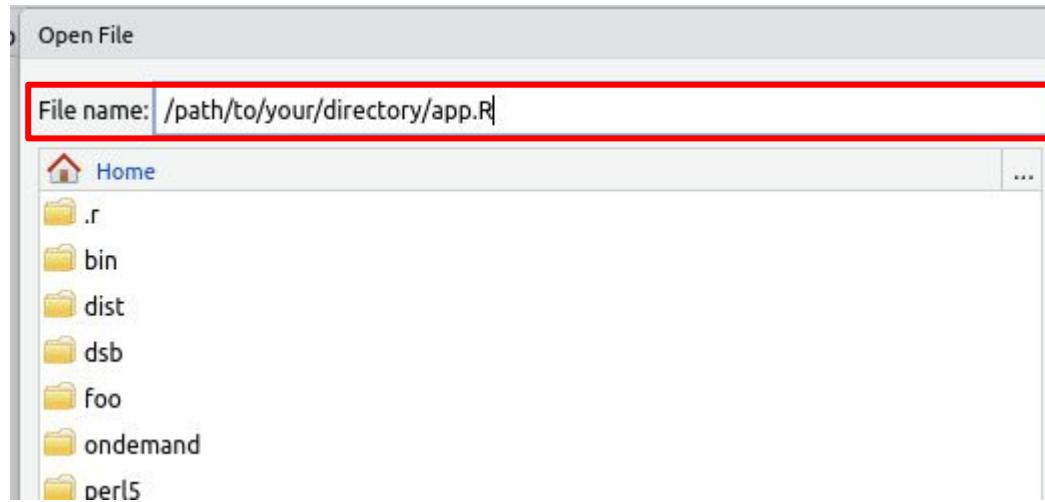
GPU compute capability: 3.5

Project: wax-dk

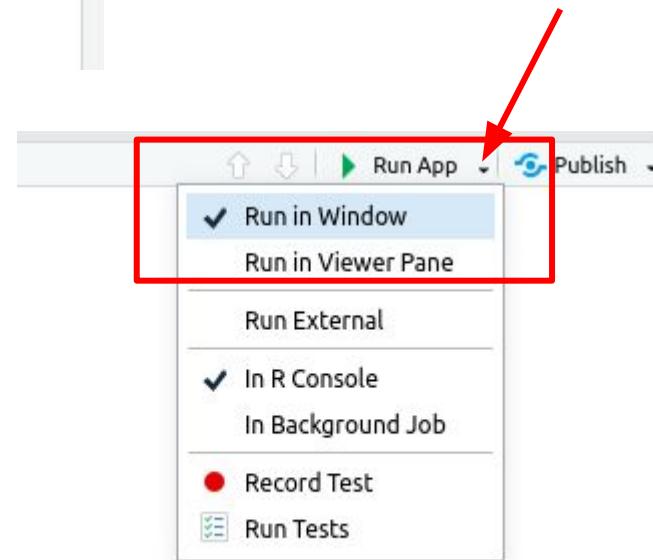
Extra qsub options:

Connect to RStudio Server

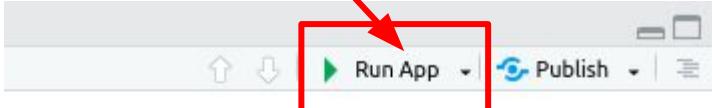




Click to small arrow to open dialog
and select “Run in Window”



Click “Run App” to start shiny app



BoxPlotter for ChIPseq Regions-of-Interest

Upload Configuration File

[Browse...](#) [Choose XLSX file](#)

Select an Excel file (.xlsx or .xls) to use as configuration for the boxplotter job.

Download the example file below to see the required format.

 [Download Example File](#)

 [Download Instructions](#)

BoxPlotter for ChIPseq Regions-of-Interest

Upload Configuration File

Browse...

example_config.xlsx

Upload complete

Select an Excel file (.xlsx or .xls) to use as configuration for the boxplotter job.

Download the example file below to see the required format.

 Download Example Configuration File

 Download Instructions

Uploaded File Information:

Filename: example_config.xlsx

File size: 38.55 KB

File type: xlsx

**File name should not contain any
spaces or special characters**

 Submit Boxplotter Job

Job Submitted Successfully!

Job ID: [2578747](#)

Configuration file: example_config_20260106_133607.xlsx

Command executed: `qsub boxplotter.qsub /projectnb/wax-dk/max/src/rprojects/boxplotter_v2/uploaded_configs/example_config_20260106_133607.xlsx`

You can monitor the job status using:

`qstat -j 2578747`

Your results will be emailed to you as soon as the job finishes. Feel free to close this tab; your job will continue running in the background.

[Close](#)

**After a successful submission,
you can close this pop-up and
the window behind it. All results
will be delivered to your BU
email.**

Download the

Upload

Filename:
File size:
File type: xlsx

 [Submit Boxplotter Job](#)

Job Status

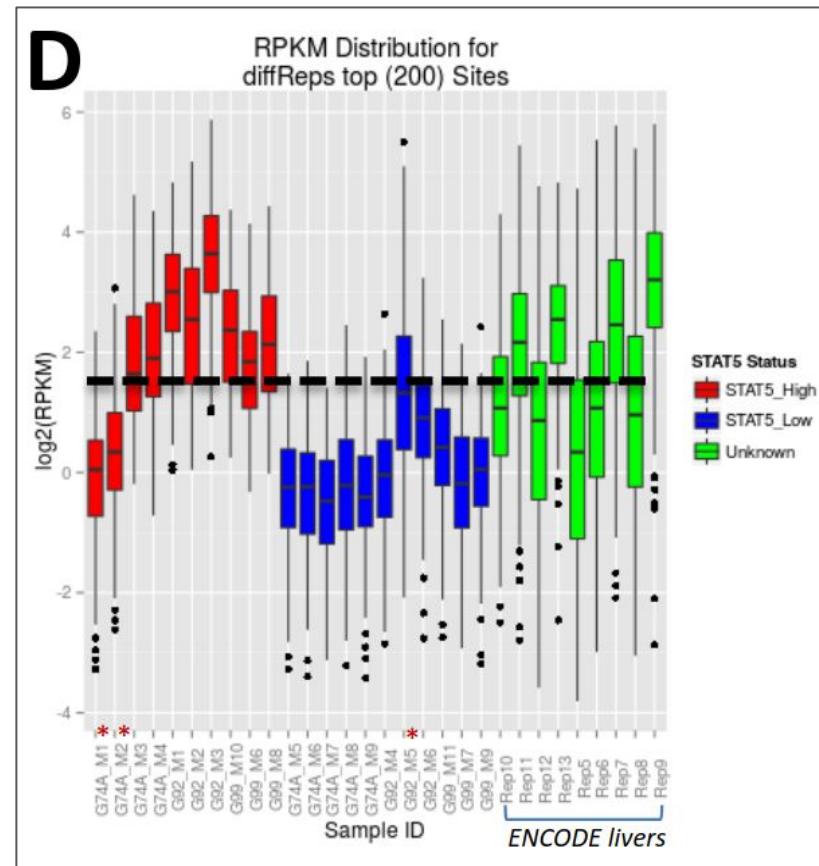
Job ID: 2578747
Status: Submitted
Configuration file: example_config_20260106_133607.xlsx
Submission time: 2026-01-06 13:36:09.416764

To check job status, run: `qstat -j 2578747`
Feel free to close this tab; your job will continue running in the background.

Example of outputs and brief
method explanation

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.

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

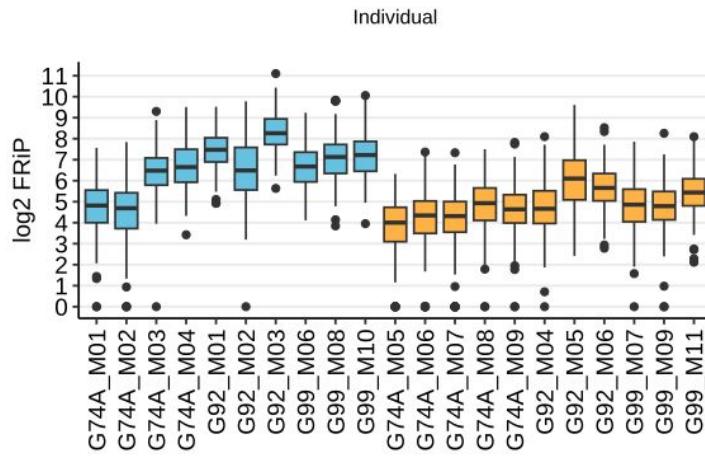
Data transformation during normalization

| chr | start | end | width | G241_M01 (Male_2wk_ATAC) | | | |
|-------|-----------|-----------|-------|--------------------------|----------------------|-----------------|-----------------------------|
| | | | | 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 |

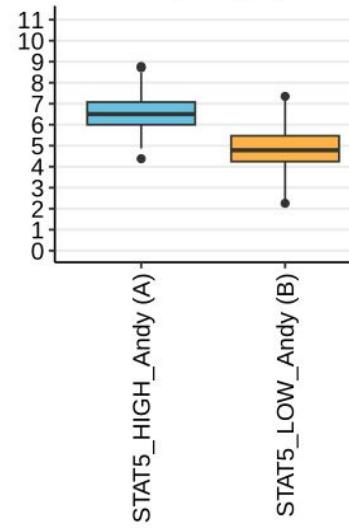
$1000 * \text{RAW_COUNTS} / \text{WIDTH}$

$\text{RAW_AFTER_WIDTH_NORM} / \text{FRiP}$

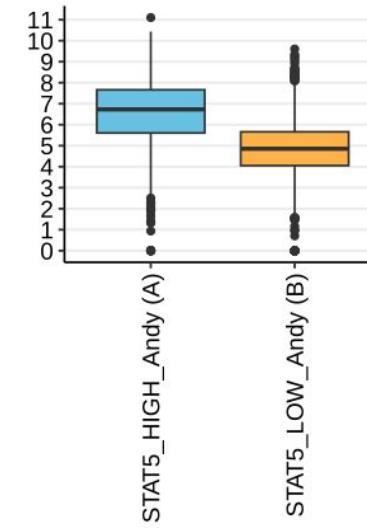
1. (A vs B) (A_MB_Dynamic_1 (n=834)_top200)



Average by samples in group



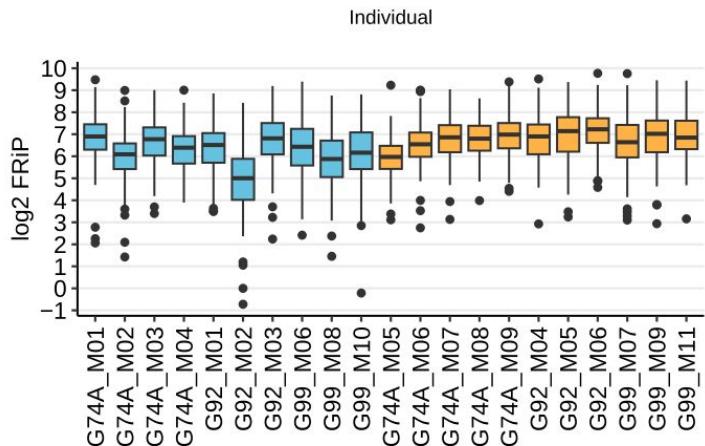
All points



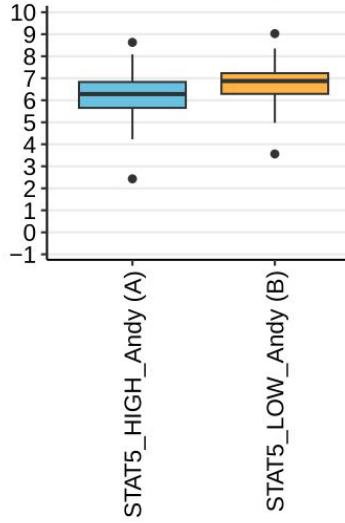
| test | average | all_points |
|-----------------------------|----------|------------|
| wilcox | 6.08E-54 | 2.75E-261 |
| KS | 1.24E-54 | 1.15E-238 |
| t.test | 1.57E-43 | 1.16E-153 |
| Median ratio (linear scale) | 3.972 | 3.665 |

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

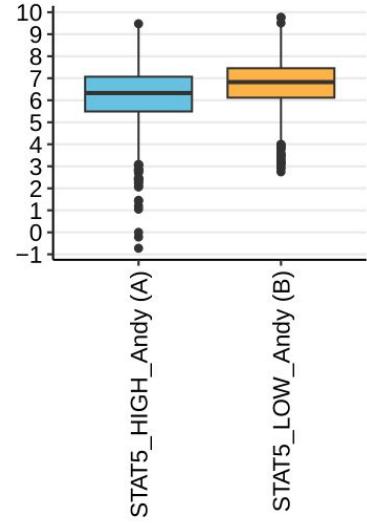
1. (A vs B) (B_MB_Static_Group 1 (n=172))



Average by samples in group



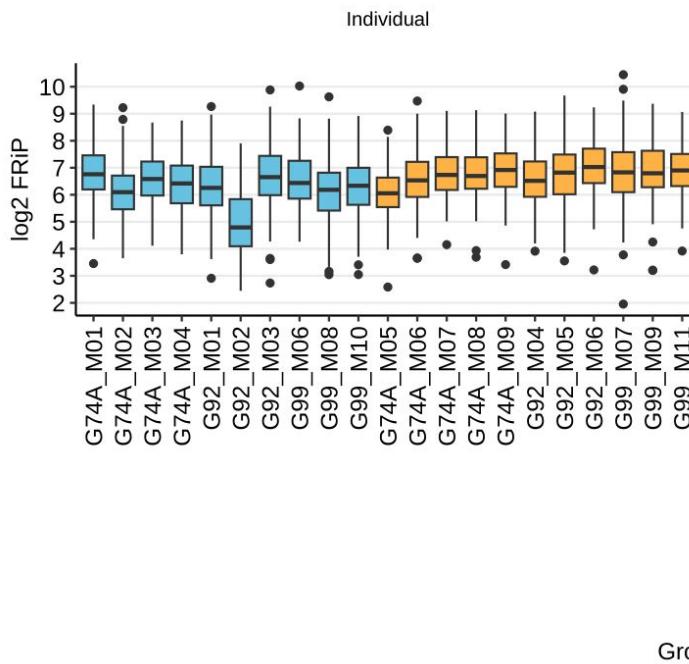
All points



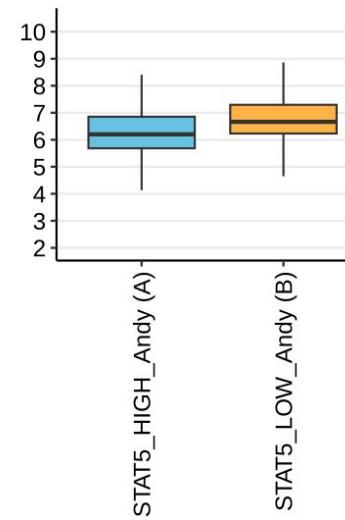
| test | average | all_points |
|-----------------------------|----------|------------|
| wilcox | 4.42E-08 | 1.85E-40 |
| KS | 3.04E-06 | 2.88E-27 |
| t.test | 2.43E-07 | 1.45E-30 |
| Median ratio (linear scale) | 0.739 | 0.708 |

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

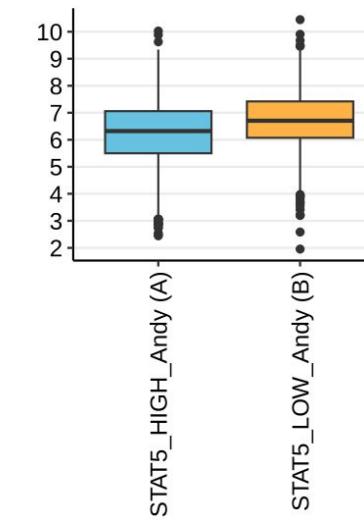
1. (A vs B) (C_MB Static_Group 2(n=186))



Average by samples in group



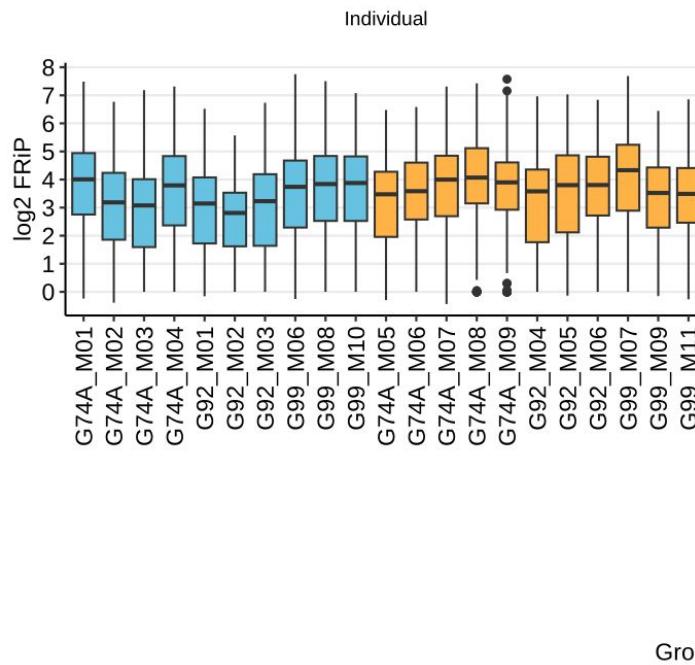
All points



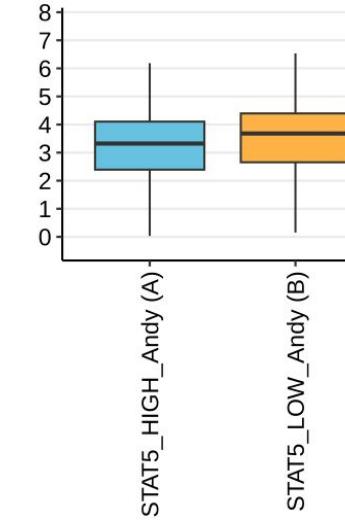
| test | average | all_points |
|-----------------------------|----------|------------|
| wilcox | 5.79E-06 | 3.52E-34 |
| KS | 1.39E-05 | 1.38E-25 |
| t.test | 1.05E-04 | 9.09E-21 |
| Median ratio (linear scale) | 0.739 | 0.767 |

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

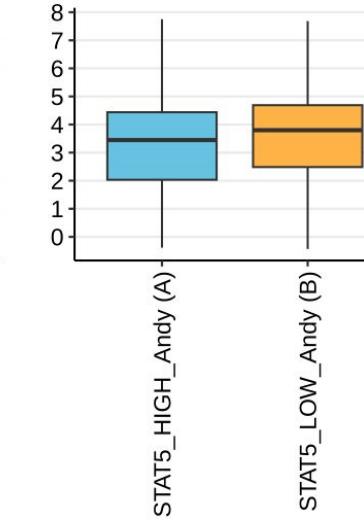
1. (A vs B) (D_FB_Static_(top n=706)_top200)



Average by samples in group



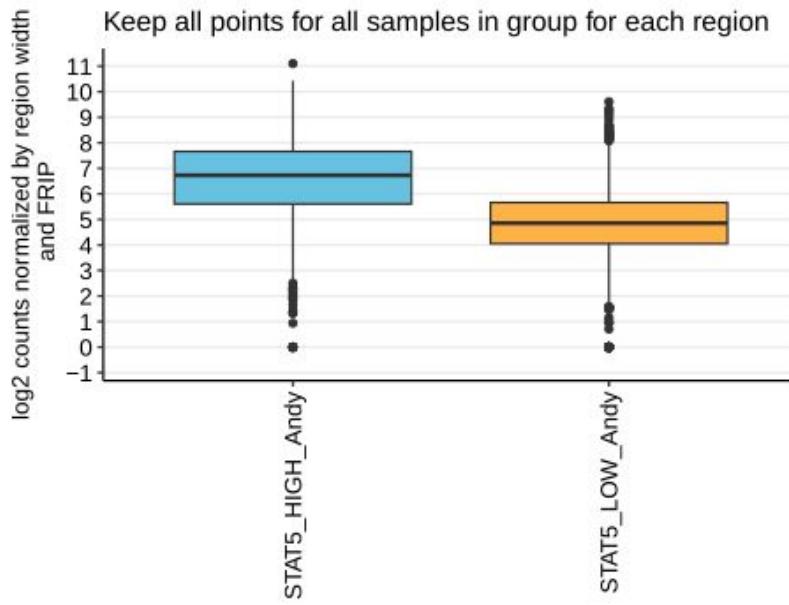
All points



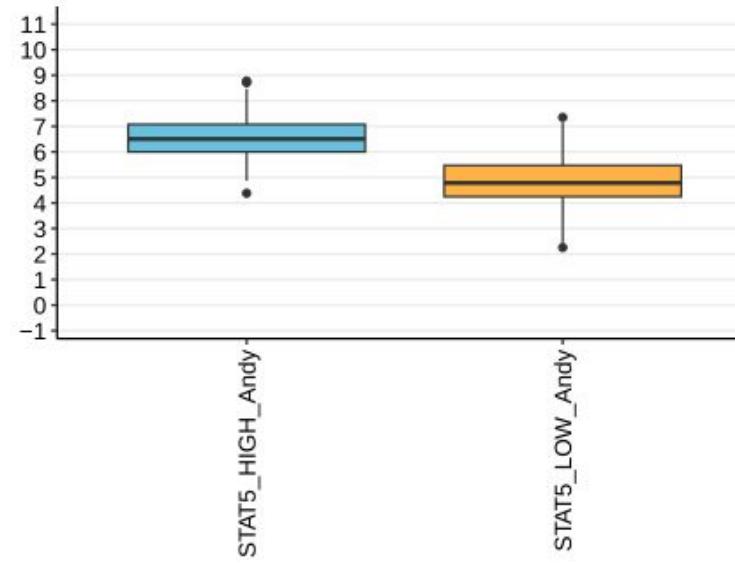
| test | average | all_points |
|-----------------------------|----------|------------|
| wilcox | NS | 2.17E-09 |
| KS | NS | 4.91E-09 |
| t.test | 4.49E-02 | 2.21E-06 |
| Median ratio (linear scale) | 0.851 | 0.784 |

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

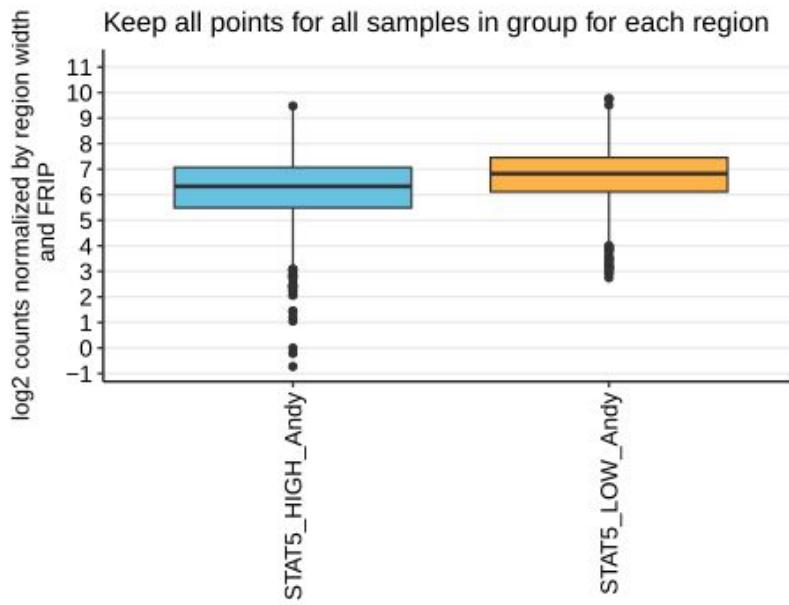
A_MB_Dynamic_1 (n=834)_top200



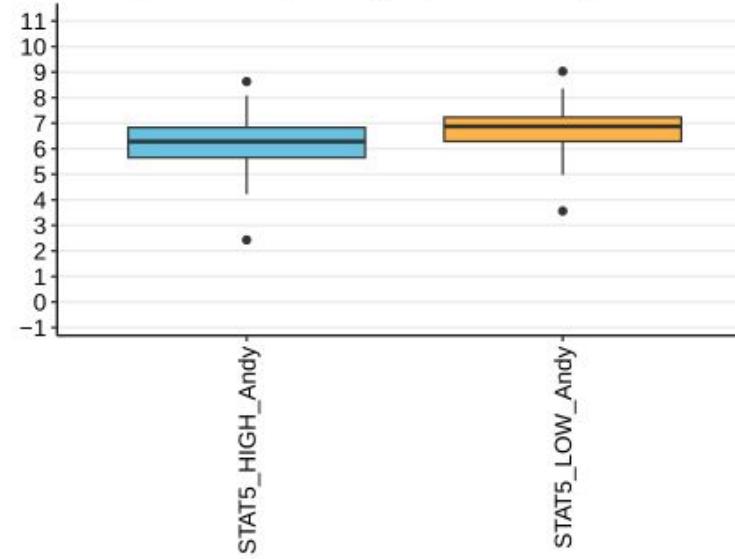
Average of all samples in group for each region



B_MB_Static_Group 1 (n=172)



Average of all samples in group for each region



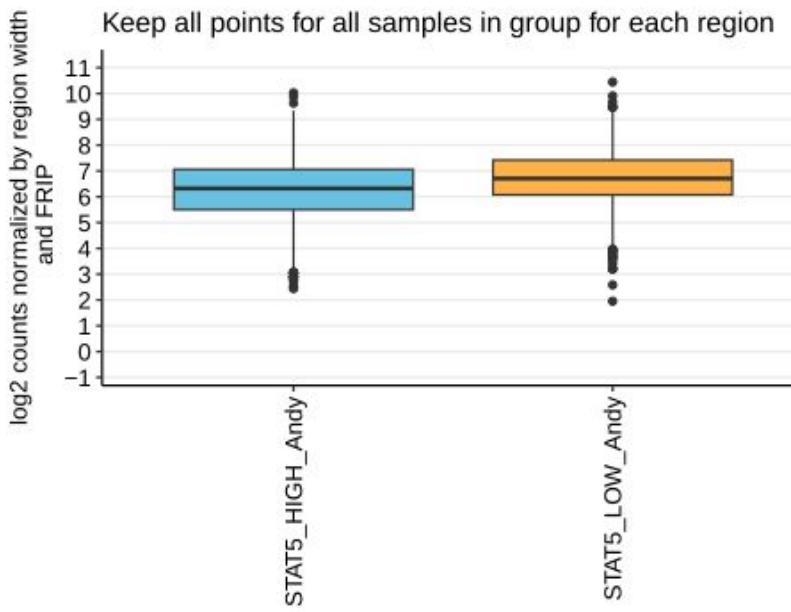
Group name

- STAT5_HIGH_Anyd
- STAT5_LOW_Anyd

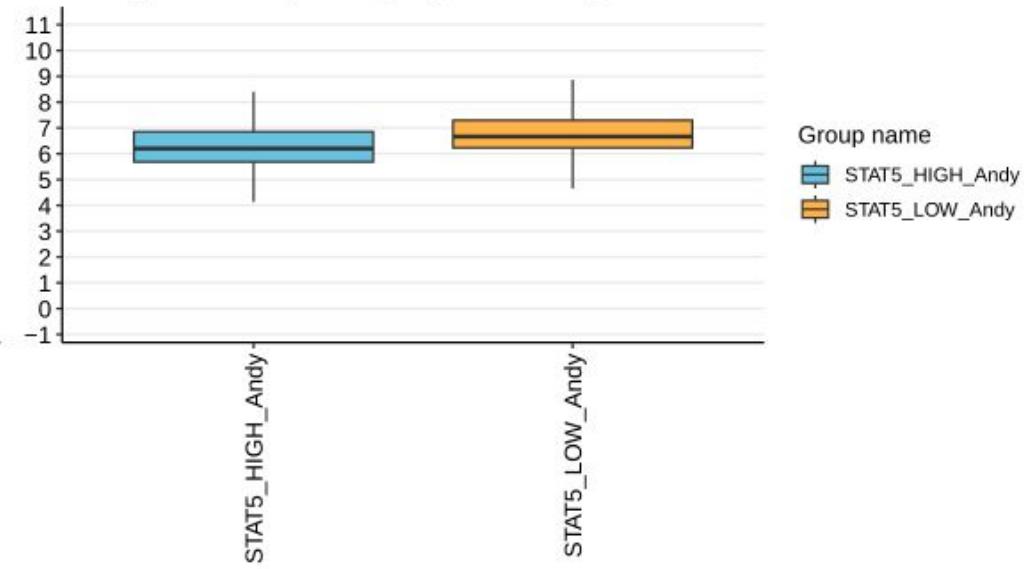
Group name

- STAT5_HIGH_Anyd
- STAT5_LOW_Anyd

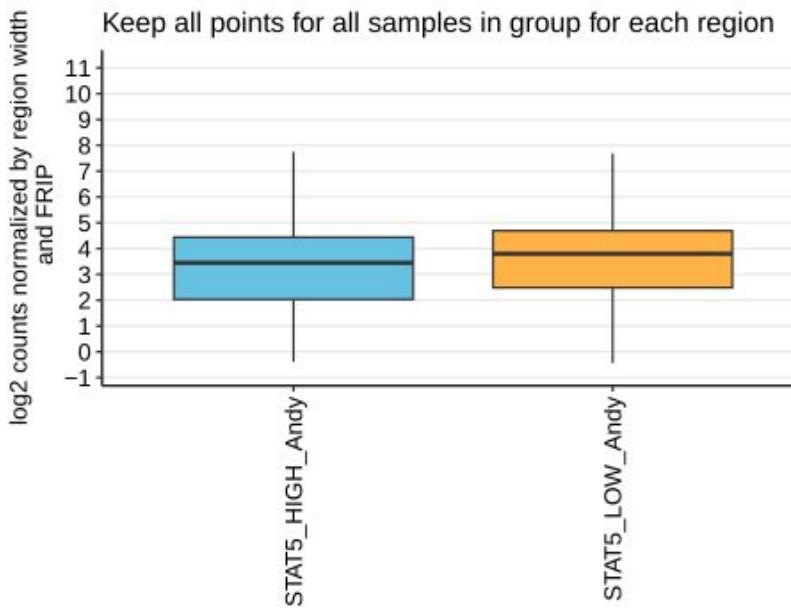
C_MB Static_Group 2_(n=186)



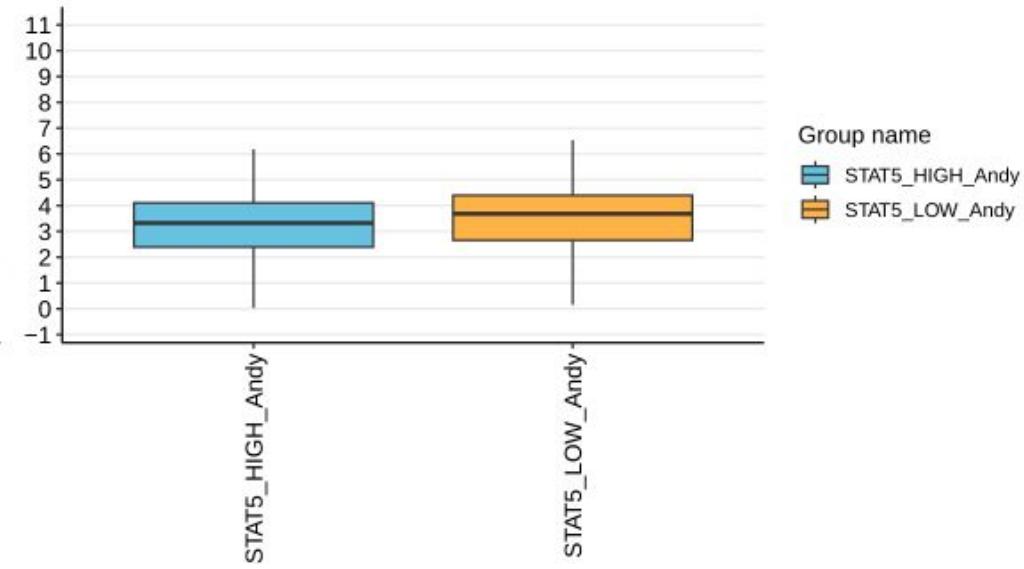
Average of all samples in group for each region



D_FB_Static_(top n=706)_top200



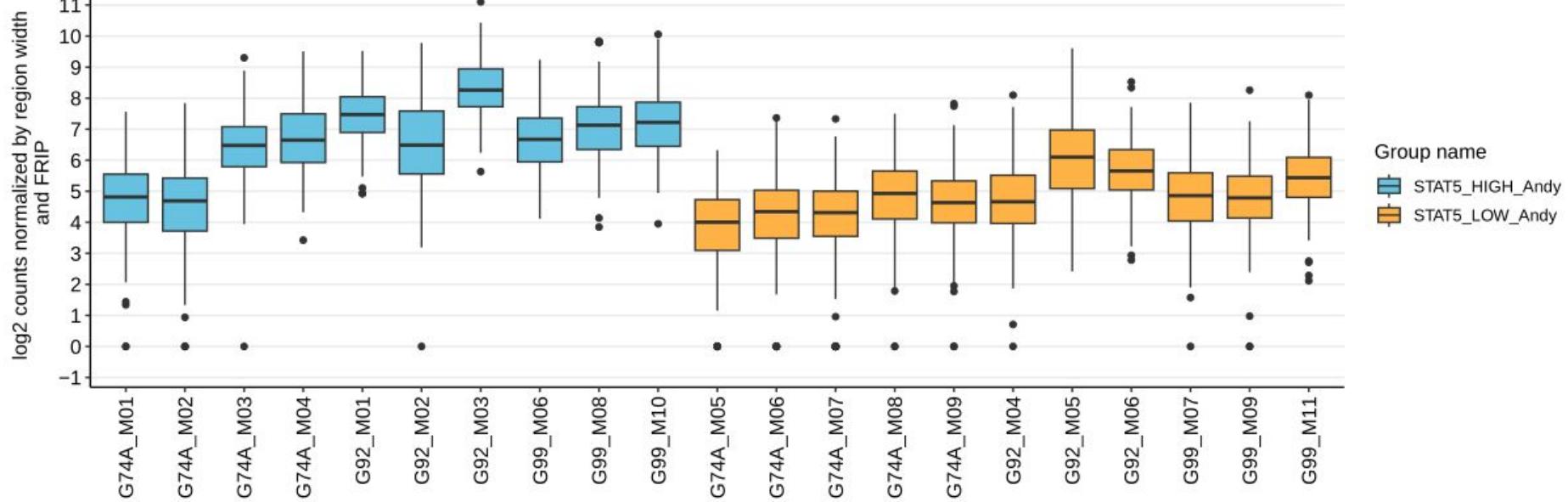
Average of all samples in group for each region



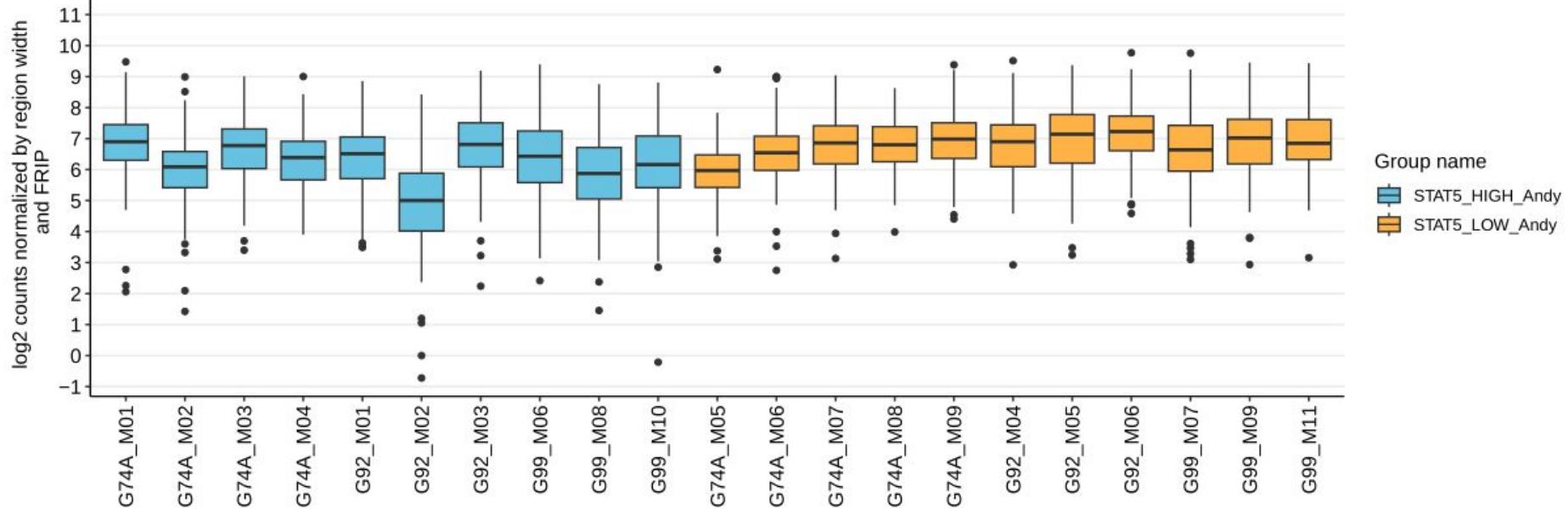
Group name
STAT5_HIGH_Anyd
STAT5_LOW_Anyd

Group name
STAT5_HIGH_Anyd
STAT5_LOW_Anyd

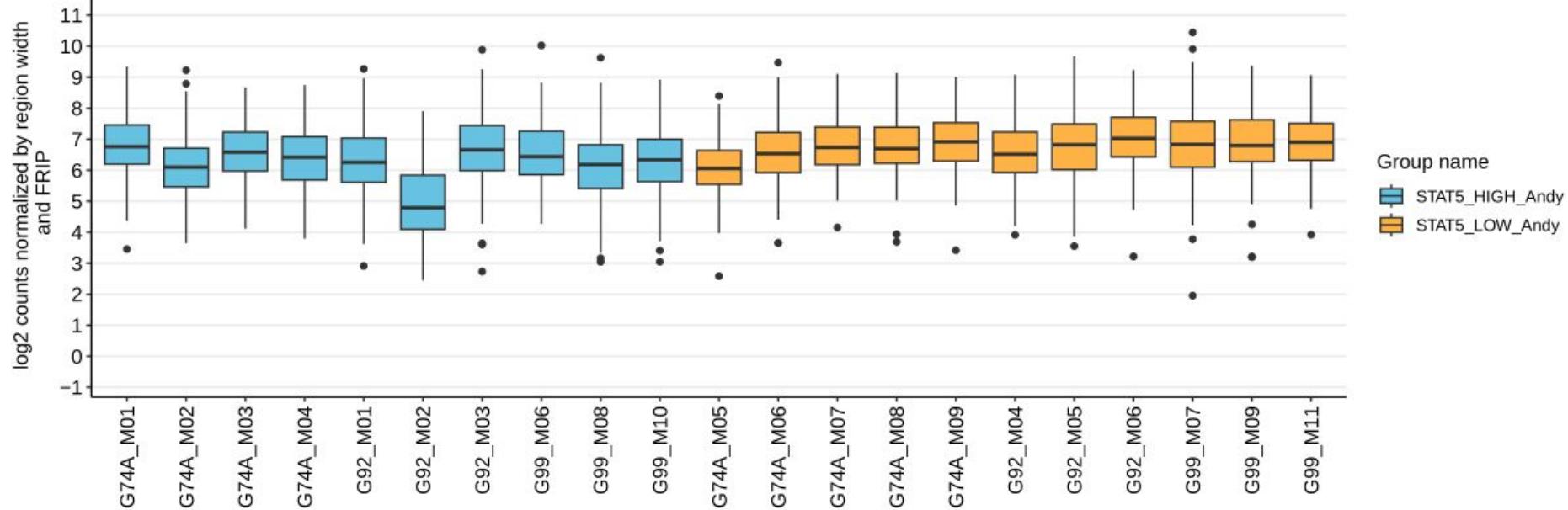
A_MB_Dynamic_1 (n=834)_top200



B_MB_Static_Group 1 (n=172)



C_MB Static_Group 2_(n=186)



D_FB_Static_(top n=706)_top200

