

GSVA for mutil Group

Ximing Ran

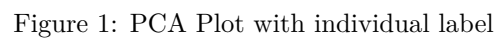
2025-03-26

Contents

1. Read the count data	2
2. Visualization for Result	2
(1) Sample Information - PCA Plot	2
(2) Sample Information - Distance Heatmap	4
3. GSVA analysis with Mix effect model.	5
(1) Main Method: Mix-Effects Model with no Interaction	5
Model Components	5
(2) Main Method: Mix-Effects Model with Interaction	7
Model Components	7
Hypothesis Testing: Comparing differece with FXS after Drug and CTRL	9
Combined Effect Hypothesis Testing	9
(3) Pathway Analysis: Full Rescue, Partial Rescue, and Side Effects	10
a. BAY Treatment	10
b. BPN Treatment	11
c. BP Treatment	12
(4) Common Rescue Pathways	13
Session information	17

In this section, we will read the clean count data from the `synaptosomes_bulkRNA` folder. We will read the data and merge them into a single table.

(1) Sample Information - PCA Plot





Heatmap showing the correlation matrix of 100 variables. The variables are grouped into three main categories: FXS (top), CTRL (middle), and D56 (bottom). The color scale ranges from 0 (dark blue) to 120 (light blue). The diagonal is dark blue, indicating a correlation of 1.0. The heatmap shows that variables within the same group have higher correlations than those across groups.

4

3. GSVA analysis with Mix effect model.

(1) Main Method: Mix-Effects Model with no Interaction

We model **GSVA scores** using a **mixed-effects model**, incorporating **fixed effects** for disease status (FXS) and drug treatment, and a **random intercept** for individual-level variability:

$$Y_{ij} = \beta_0 + \beta_1 X_{\text{FXS}} + \beta_2 X_{\text{Drug_BAY}} + \beta_3 X_{\text{Drug_BPN}} + \beta_4 X_{\text{Drug_BP}} + b_i + \varepsilon_{ij}$$

Model Components

- Y_{ij} : GSVA score for individual i under condition j .
- **Fixed Effects:**
 - X_{FXS} : Disease indicator (1 for **FXS**, 0 for **CTRL**).
 - $X_{\text{Drug_BAY}}, X_{\text{Drug_BPN}}, X_{\text{Drug_BP}}$: Drug treatment indicators (vehicle is reference).
- **Random intercept**
 - $b_i \sim \mathcal{N}(0, \tau^2)$ for individual i , capturing baseline variability.
 - $\varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$: Residual error term.

This model allows us to estimate the effects of disease and drug treatments on GSVA scores while controlling for repeated measures within individuals.

```

## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: Y ~ X_FXS + X_Drug_BAY + X_Drug_BPN + X_Drug_BP + (1 | Individual)
## Data: df_pathway
##
## REML criterion at convergence: -11.1
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.98065 -0.50074 -0.00603  0.59646  2.34519
##
## Random effects:
## Groups      Name      Variance Std.Dev.
## Individual (Intercept) 0.01094  0.1046
## Residual              0.03649  0.1910
## Number of obs: 72, groups: Individual, 6
##
## Fixed effects:
##              Estimate Std. Error      df t value Pr(>|t|)
## (Intercept) -0.12855    0.07862   6.98778  -1.635   0.1461
## X_FXS        0.22634    0.09655   4.00000   2.344   0.0790 .
## X_Drug_BAY   0.06840    0.06367  63.00000   1.074   0.2868
## X_Drug_BPN  -0.09013    0.06367  63.00000  -1.416   0.1618
## X_Drug_BP   -0.20056    0.06367  63.00000  -3.150   0.0025 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##              (Intr) X_FXS  X_D_BA X_D_BPN
## X_FXS          -0.614
## X_Drug_BAY     -0.405  0.000
## X_Drug_BPN     -0.405  0.000  0.500
## X_Drug_BP      -0.405  0.000  0.500  0.500

```

(2) Main Method: Mix-Effects Model with Interaction

We model **GSVA scores** using a **mixed-effects model**, incorporating **fixed effects** for disease status (FXS), drug treatments, and their **interactions**, along with a **random intercept** for individual-level variability:

$$\begin{aligned} Y_{ij} = & \beta_0 + \beta_1 X_{\text{FXS}} + \beta_2 X_{\text{Drug_BAY}} + \beta_3 X_{\text{Drug_BPN}} + \beta_4 X_{\text{Drug_BP}} \\ & + \beta_5 (X_{\text{FXS}} \times X_{\text{Drug_BAY}}) + \beta_6 (X_{\text{FXS}} \times X_{\text{Drug_BPN}}) + \beta_7 (X_{\text{FXS}} \times X_{\text{Drug_BP}}) \\ & + b_i + \varepsilon_{ij} \end{aligned}$$

Model Components

- Y_{ij} : GSVA score for individual i under condition j .
- **Fixed Effects:**
 - X_{FXS} : Disease indicator (1 for **FXS**, 0 for **CTRL**).
 - $X_{\text{Drug_BAY}}, X_{\text{Drug_BPN}}, X_{\text{Drug_BP}}$: Drug treatment indicators (vehicle is the reference).
 - **Interaction Terms:**
 - * $X_{\text{FXS}} \times X_{\text{Drug_BAY}}$: Interaction between FXS and BAY treatment.
 - * $X_{\text{FXS}} \times X_{\text{Drug_BPN}}$: Interaction between FXS and BPN treatment.
 - * $X_{\text{FXS}} \times X_{\text{Drug_BP}}$: Interaction between FXS and BP treatment.
- **Random Intercept:**
 - $b_i \sim \mathcal{N}(0, \tau^2)$: Individual-specific random intercept capturing baseline variability.
- **Error Term:**
 - $\varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$: Residual error term.

This model allows us to estimate not only the **main effects** of disease and drug treatments but also **how drug responses differ between FXS and CTRL individuals**. Interaction terms capture whether the drug effect is **modified by disease status**, which is essential for understanding differential pathway responses.

```

## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula:
## Y ~ X_FXS + X_Drug_BAY + X_Drug_BPN + X_Drug_BP + X_FXS * X_Drug_BAY +
##       X_FXS * X_Drug_BPN + X_FXS * X_Drug_BP + (1 | Individual)
## Data: df_pathway
##
## REML criterion at convergence: -12.2
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.5751 -0.7045 -0.0373  0.4715  2.9583
##
## Random effects:
## Groups      Name                Variance Std.Dev.
## Individual (Intercept) 0.01121  0.1059
## Residual              0.03325  0.1824
## Number of obs: 72, groups: Individual, 6
##
## Fixed effects:
##              Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)   -0.19869    0.08621  9.93526  -2.305  0.0441 *
## X_FXS           0.36661    0.12192  9.93526   3.007  0.0133 *
## X_Drug_BAY      0.08612    0.08596 60.00000   1.002  0.3205
## X_Drug_BPN      0.07042    0.08596 60.00000   0.819  0.4159
## X_Drug_BP     -0.09829    0.08596 60.00000  -1.143  0.2574
## X_FXS:X_Drug_BAY -0.03544    0.12157 60.00000  -0.291  0.7717
## X_FXS:X_Drug_BPN -0.32111    0.12157 60.00000  -2.641  0.0105 *
## X_FXS:X_Drug_BP  -0.20454    0.12157 60.00000  -1.682  0.0977 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##              (Intr) X_FXS  X_D_BA X_D_BPN X_Dr_BP X_FXS:X_D_BA X_FXS:X_D_BPN
## X_FXS          -0.707
## X_Drug_BAY     -0.499  0.353
## X_Drug_BPN     -0.499  0.353  0.500
## X_Drug_BP      -0.499  0.353  0.500  0.500
## X_FXS:X_D_BA    0.353 -0.499 -0.707 -0.354 -0.354
## X_FXS:X_D_BPN  0.353 -0.499 -0.354 -0.707 -0.354  0.500
## X_FXS:X_Dr_BP  0.353 -0.499 -0.354 -0.354 -0.707  0.500  0.500

```


Hypothesis Testing: Comparing difference with FXS after Drug and CTRL

Combined Effect Hypothesis Testing

To interpret the effect of each drug within the **FXS background**, we tested whether the **combined effect** of disease and drug treatment is significantly different from CTRL. Specifically, we evaluated the null hypothesis:

$$H_0 : \beta_{\text{FXS}} + \beta_{\text{Drug}} + \beta_{\text{FXS:Drug}} = 0$$

This tests whether the **net drug effect in the FXS group** is equivalent to the CTRL baseline (i.e., no rescue effect).

C

```
##
## Linear hypothesis test:
## X_FXS + X_Drug_BAY + X_FXS:X_Drug_BAY = 0
##
## Model 1: restricted model
## Model 2: Y ~ X_FXS + X_Drug_BAY + X_Drug_BPN + X_Drug_BP + X_FXS * X_Drug_BAY +
##       X_FXS * X_Drug_BPN + X_FXS * X_Drug_BP + (1 | Individual)
##
##      Df  Chisq Pr(>Chisq)
## 1
## 2  1 11.715  0.0006201 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##
## Linear hypothesis test:
## X_FXS + X_Drug_BPN + X_FXS:X_Drug_BPN = 0
##
## Model 1: restricted model
## Model 2: Y ~ X_FXS + X_Drug_BAY + X_Drug_BPN + X_Drug_BP + X_FXS * X_Drug_BAY +
##       X_FXS * X_Drug_BPN + X_FXS * X_Drug_BP + (1 | Individual)
##
##      Df Chisq Pr(>Chisq)
## 1
## 2  1 0.904      0.3417

##
## Linear hypothesis test:
## X_FXS + X_Drug_BP + X_FXS:X_Drug_BP = 0
##
## Model 1: restricted model
## Model 2: Y ~ X_FXS + X_Drug_BAY + X_Drug_BPN + X_Drug_BP + X_FXS * X_Drug_BAY +
##       X_FXS * X_Drug_BPN + X_FXS * X_Drug_BP + (1 | Individual)
##
##      Df  Chisq Pr(>Chisq)
## 1
## 2  1 0.2737      0.6008
```

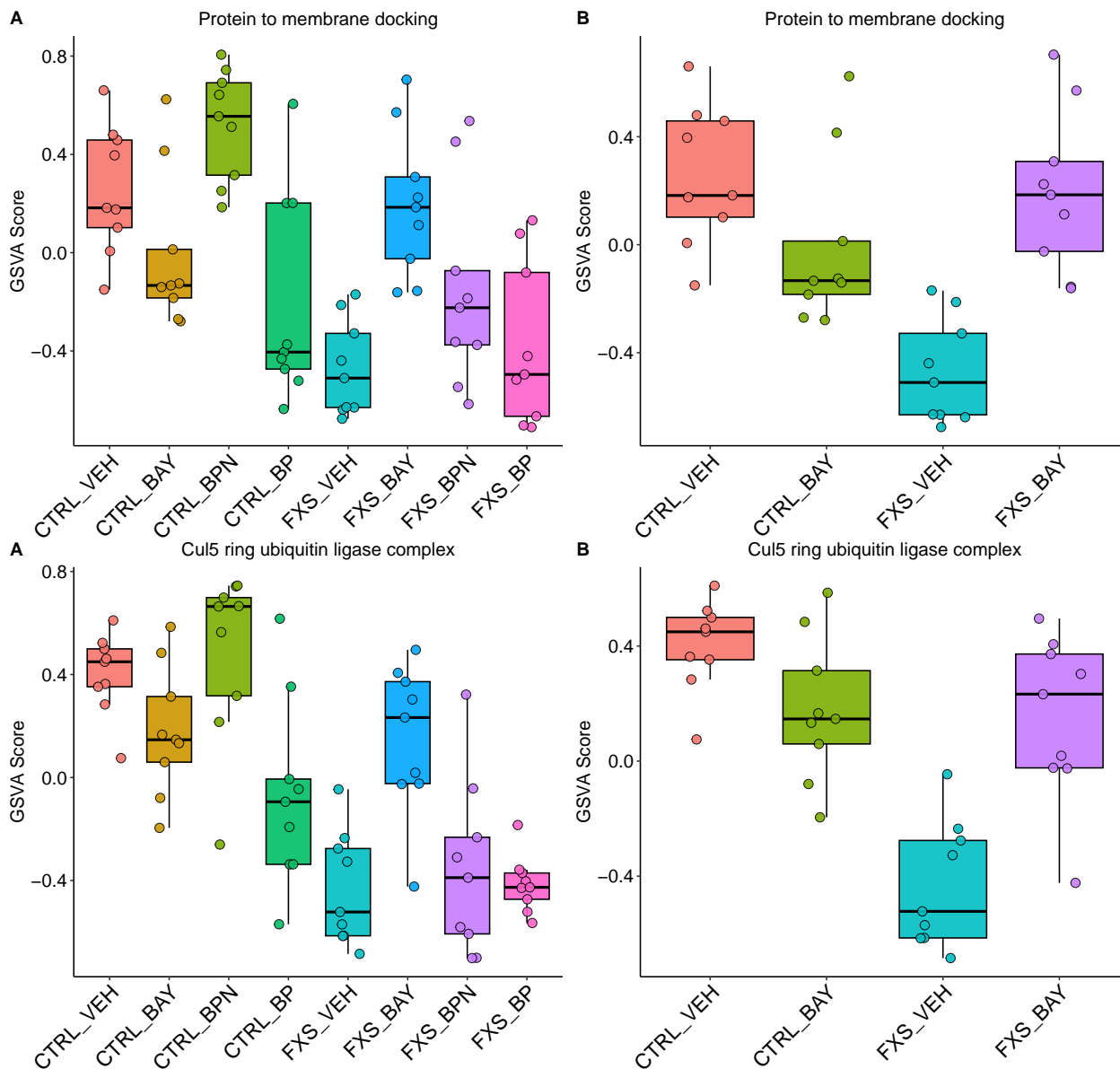
(3) Pathway Analysis: Full Rescue, Partial Rescue, and Side Effects

a. BAY Treatment

##

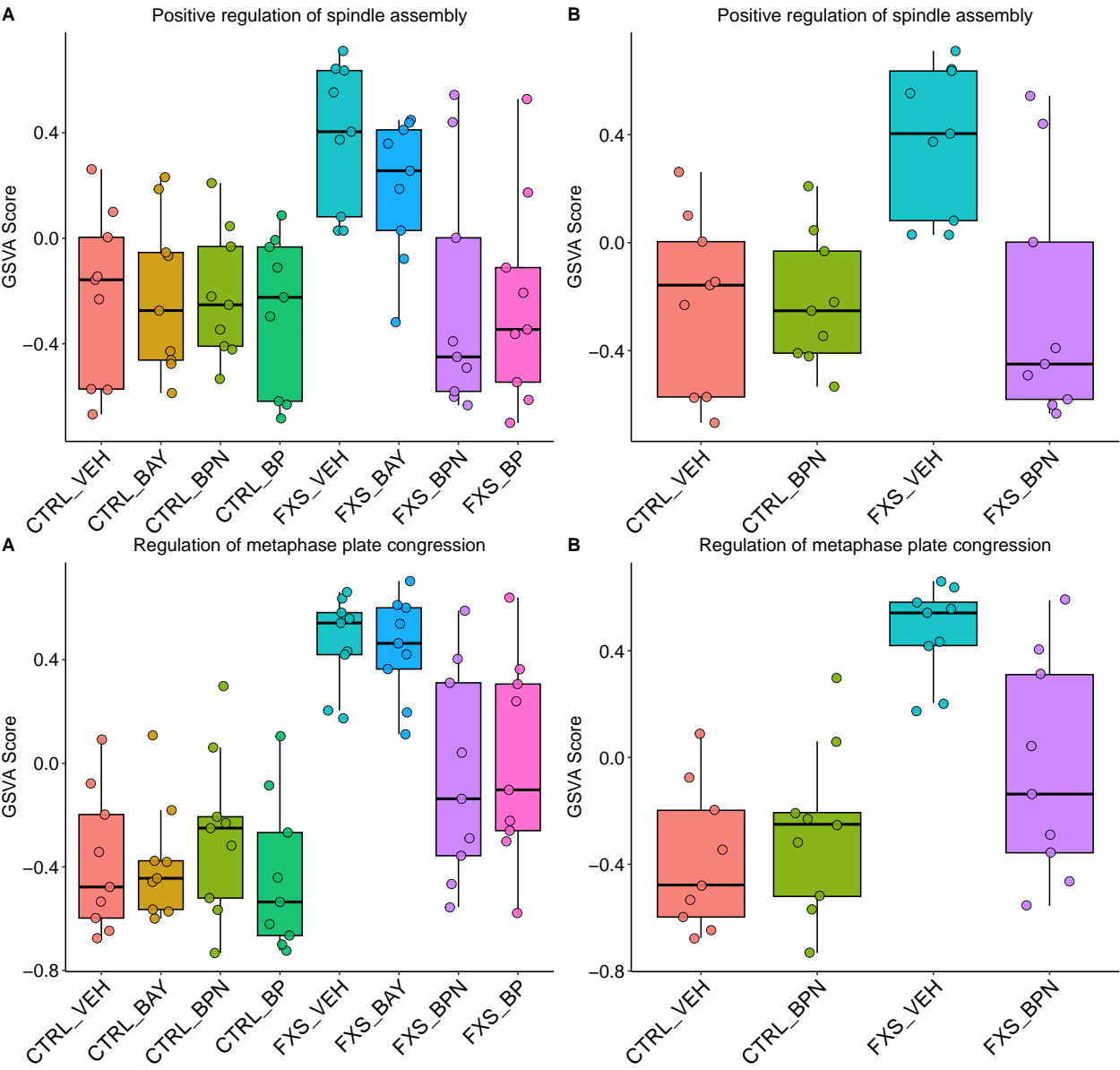
Full Rescue Partial Rescue

381 336



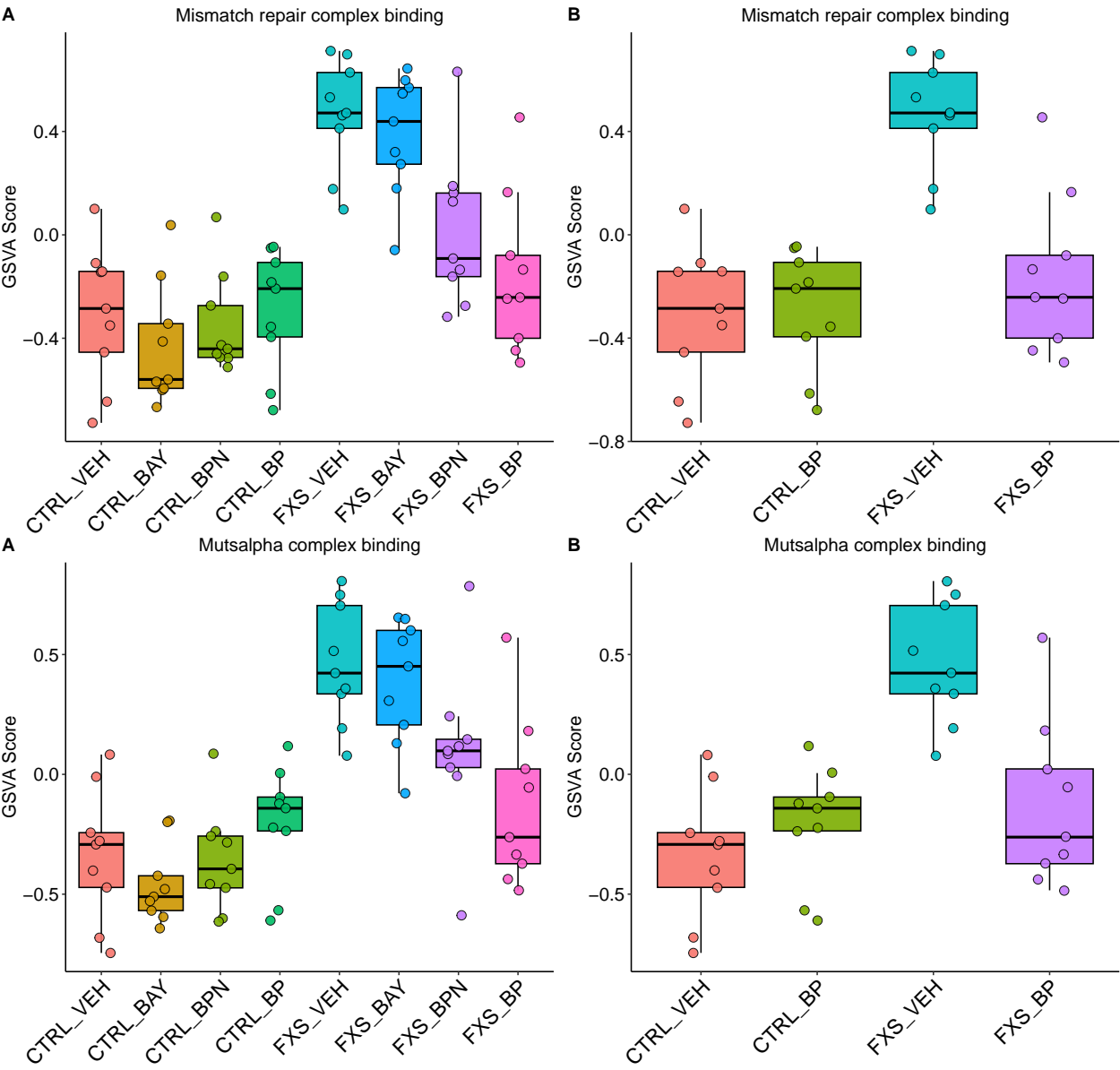
b. BPN Treatment

Full Rescue Partial Rescue
271 152



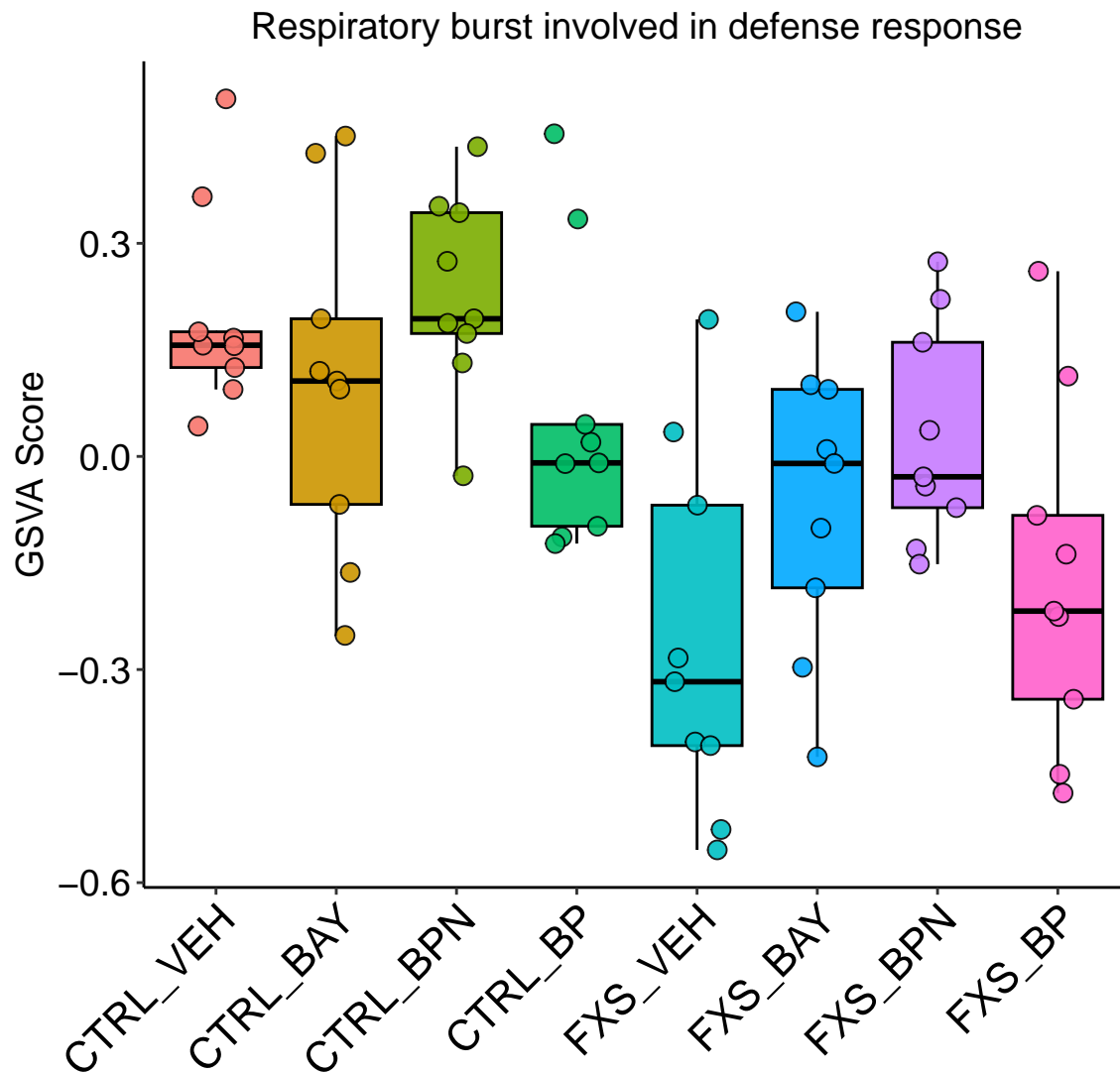
c. BP Treatment

Full Rescue Partial Rescue
84 292

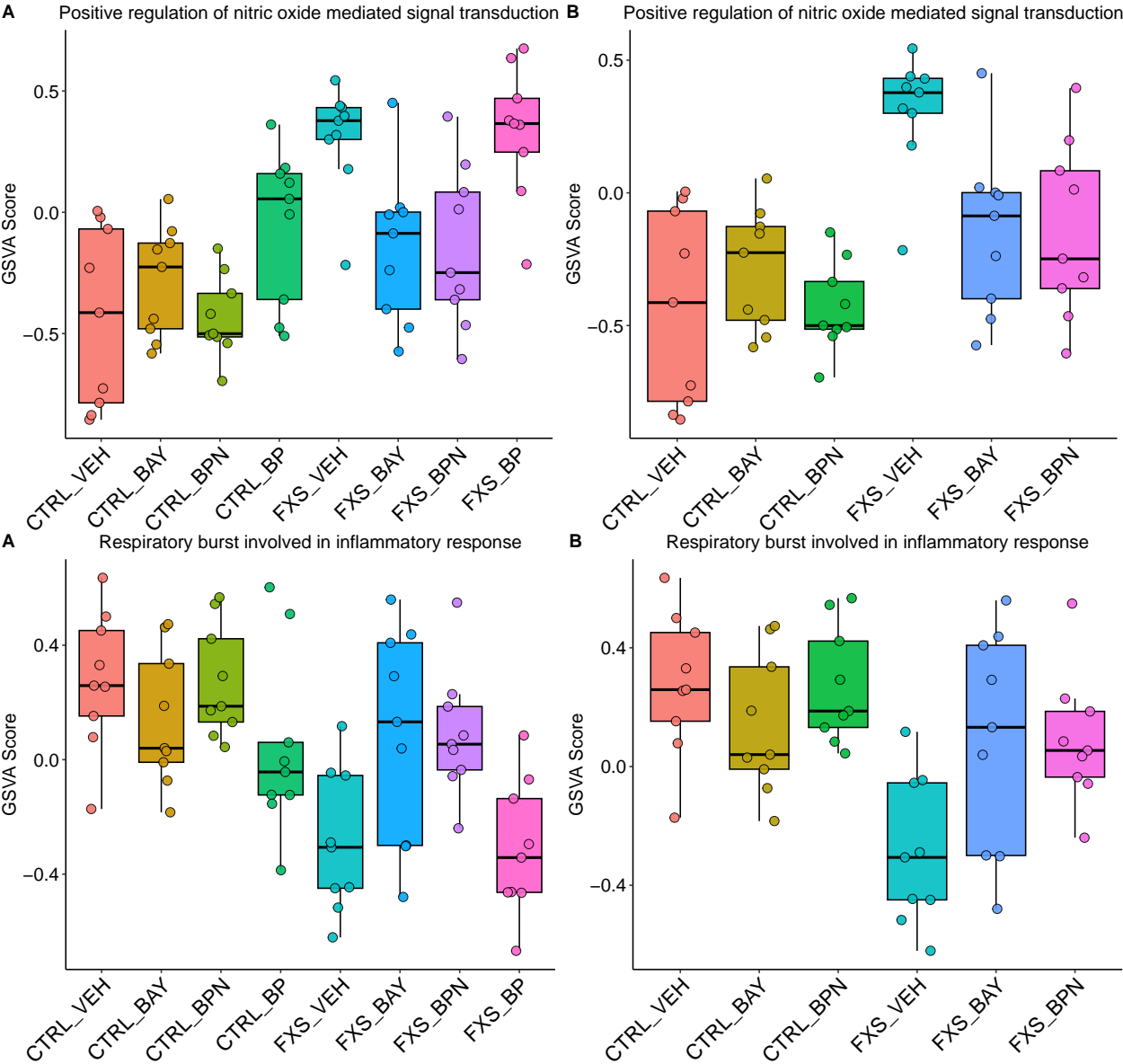


(4) Common Rescue Pathways

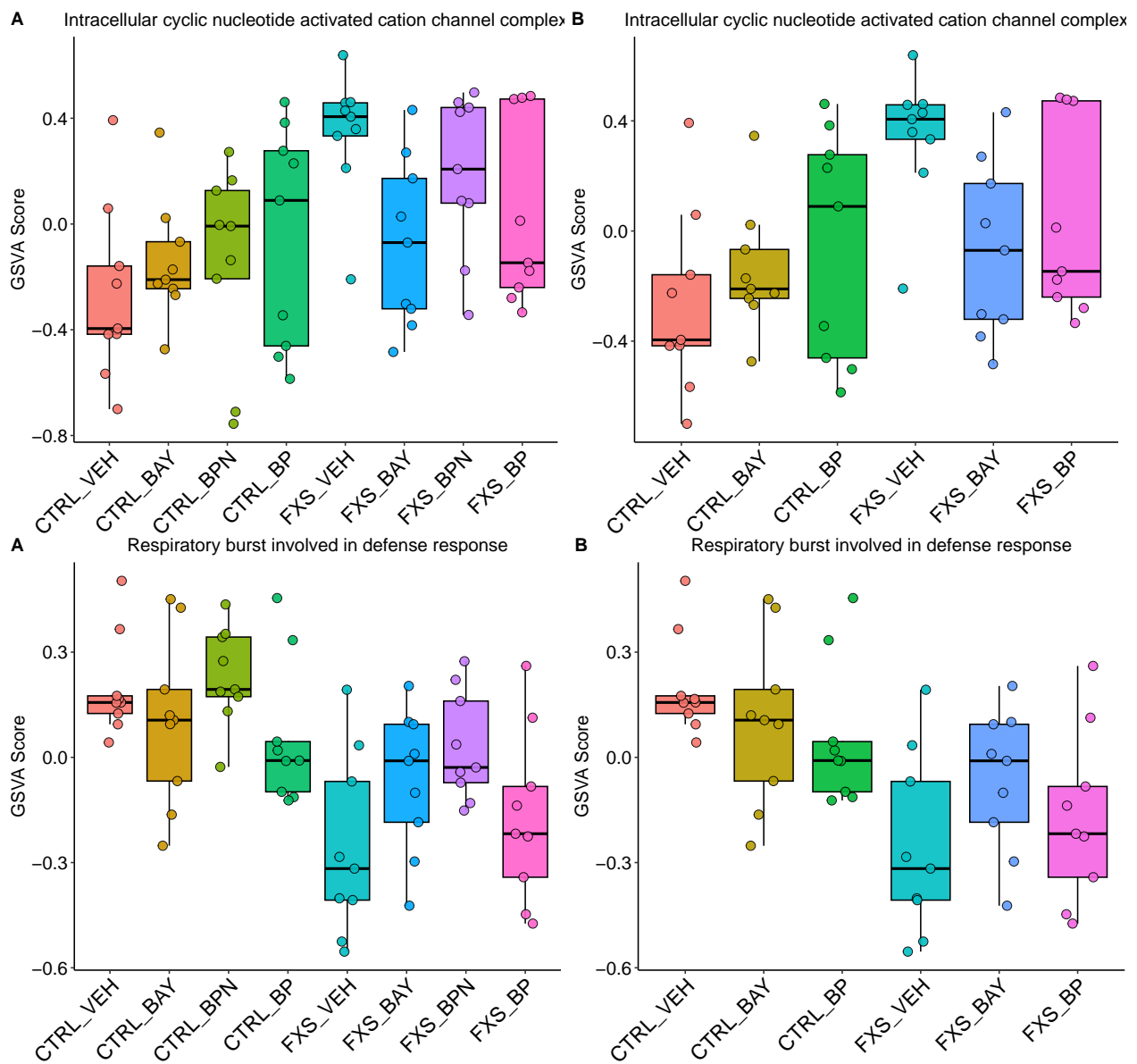
All 3 drugs.



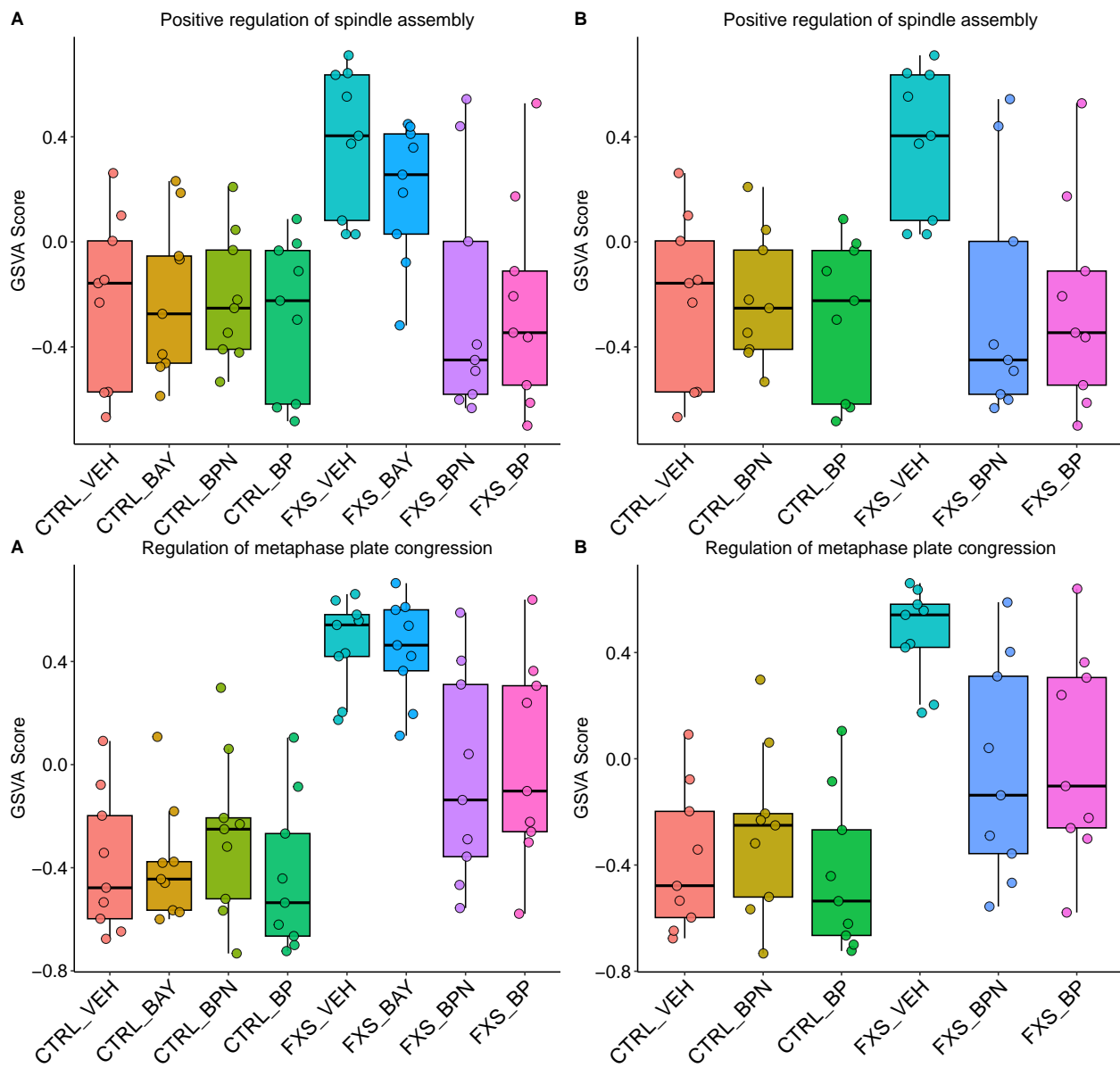
Overlap 2 drugs BAY and BPN



BAY and BP



BPN and BP



Session information

```
## R version 4.4.0 (2024-04-24)
## Platform: aarch64-apple-darwin20
## Running under: macOS Sonoma 14.3.1
##
## Matrix products: default
## BLAS:   /Library/Frameworks/R.framework/Versions/4.4-arm64/Resources/lib/libRblas.0.dylib
## LAPACK: /Library/Frameworks/R.framework/Versions/4.4-arm64/Resources/lib/libRlapack.dylib; LAPACK v
##
## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## time zone: America/New_York
## tzcode source: internal
##
## attached base packages:
## [1] parallel stats4 stats graphics grDevices utils datasets
## [8] methods base
##
## other attached packages:
## [1] GSEABase_1.66.0 graph_1.82.0
## [3] annotate_1.82.0 XML_3.99-0.18
## [5] extrafont_0.19 ggsignif_0.6.4
## [7] patchwork_1.3.0 decoupleR_2.10.0
## [9] GSVA_1.52.3 BiocParallel_1.38.0
## [11] edgeR_4.2.2 limma_3.60.6
## [13] GenomicFeatures_1.56.0 biomaRt_2.60.1
## [15] gprofiler2_0.2.3 RColorBrewer_1.1-3
## [17] data.table_1.16.4 org.Hs.eg.db_3.19.1
## [19] AnnotationDbi_1.66.0 clusterProfiler_4.12.6
## [21] ggfortify_0.4.17 pheatmap_1.0.12
## [23] EnhancedVolcano_1.22.0 ggrepel_0.9.6
## [25] apeglm_1.26.1 DESeq2_1.44.0
## [27] SummarizedExperiment_1.34.0 Biobase_2.64.0
## [29] MatrixGenerics_1.16.0 matrixStats_1.5.0
## [31] reshape2_1.4.4 Signac_1.14.0
## [33] Seurat_5.2.1 SeuratObject_5.0.2
## [35] sp_2.2-0 cowplot_1.1.3
## [37] rtracklayer_1.64.0 GenomicRanges_1.56.2
## [39] GenomeInfoDb_1.40.1 IRanges_2.38.1
## [41] S4Vectors_0.42.1 BiocGenerics_0.50.0
## [43] lmerTest_3.1-3 lme4_1.1-36
## [45] Matrix_1.7-2 car_3.1-3
## [47] carData_3.0-5 knitr_1.49
## [49] lubridate_1.9.4 forcats_1.0.0
## [51] stringr_1.5.1 dplyr_1.1.4
## [53] purrr_1.0.4 readr_2.1.5
## [55] tidyr_1.3.1 tibble_3.2.1
## [57] ggplot2_3.5.1 tidyverse_2.0.0
##
## loaded via a namespace (and not attached):
## [1] SpatialExperiment_1.14.0 R.methodsS3_1.8.2
## [3] progress_1.2.3 goftest_1.2-3
```

## [5] HDF5Array_1.32.1	Biostrings_2.72.1
## [7] vctrs_0.6.5	spatstat.random_3.3-2
## [9] digest_0.6.37	png_0.1-8
## [11] deldir_2.0-4	parallelly_1.42.0
## [13] magick_2.8.5	MASS_7.3-64
## [15] httpuv_1.6.15	qvalue_2.36.0
## [17] withr_3.0.2	xfun_0.51
## [19] ggfun_0.1.8	survival_3.8-3
## [21] memoise_2.0.1	gson_0.1.0
## [23] systemfonts_1.2.1	ragg_1.3.3
## [25] tidytree_0.4.6	zoo_1.8-12
## [27] pbapply_1.7-2	R.oo_1.27.0
## [29] Formula_1.2-5	prettyunits_1.2.0
## [31] KEGGREST_1.44.1	promises_1.3.2
## [33] httr_1.4.7	restfulr_0.0.15
## [35] rhdf5filters_1.16.0	globals_0.16.3
## [37] fitdistrplus_1.2-2	rhdf5_2.48.0
## [39] rstudioapi_0.17.1	UCSC.utils_1.0.0
## [41] miniUI_0.1.1.1	generics_0.1.3
## [43] DOSE_3.30.5	curl_6.2.1
## [45] zlibbioc_1.50.0	ScaledMatrix_1.12.0
## [47] ggraph_2.2.1	polyclip_1.10-7
## [49] GenomeInfoDbData_1.2.12	SparseArray_1.4.8
## [51] xtable_1.8-4	evaluate_1.0.3
## [53] S4Arrays_1.4.1	BiocFileCache_2.12.0
## [55] hms_1.1.3	irlba_2.3.5.1
## [57] colorspace_2.1-1	filelock_1.0.3
## [59] ROCR_1.0-11	reticulate_1.40.0
## [61] spatstat.data_3.1-4	magrittr_2.0.3
## [63] lmtest_0.9-40	later_1.4.1
## [65] viridis_0.6.5	ggtree_3.12.0
## [67] lattice_0.22-6	spatstat.geom_3.3-5
## [69] future.apply_1.11.3	scattermore_1.2
## [71] shadowtext_0.1.4	RcppAnnoy_0.0.22
## [73] pillar_1.10.1	nlme_3.1-167
## [75] beachmat_2.20.0	compiler_4.4.0
## [77] RSpectra_0.16-2	stringi_1.8.4
## [79] tensor_1.5	minqa_1.2.8
## [81] GenomicAlignments_1.40.0	plyr_1.8.9
## [83] crayon_1.5.3	abind_1.4-8
## [85] BiocIO_1.14.0	gridGraphics_0.5-1
## [87] emdbook_1.3.13	locfit_1.5-9.11
## [89] graphlayouts_1.2.2	bit_4.5.0.1
## [91] fastmatch_1.1-6	textshaping_1.0.0
## [93] codetools_0.2-20	BiocSingular_1.20.0
## [95] plotly_4.10.4	mime_0.12
## [97] splines_4.4.0	Rcpp_1.0.14
## [99] fastDummies_1.7.5	sparseMatrixStats_1.16.0
## [101] dbplyr_2.5.0	Rttf2pt1_1.3.12
## [103] blob_1.2.4	here_1.0.1
## [105] fs_1.6.5	listenv_0.9.1
## [107] Rdpack_2.6.2	ggplotify_0.1.2
## [109] statmod_1.5.0	tzdb_0.4.0
## [111] tweenr_2.0.3	pkgconfig_2.0.3

## [113] tools_4.4.0	cachem_1.1.0
## [115] rbibutils_2.3	RSQlite_2.3.9
## [117] viridisLite_0.4.2	DBI_1.2.3
## [119] numDeriv_2016.8-1.1	fastmap_1.2.0
## [121] rmarkdown_2.29	scales_1.3.0
## [123] grid_4.4.0	ica_1.0-3
## [125] Rsamtools_2.20.0	coda_0.19-4.1
## [127] dotCall64_1.2	RANN_2.6.2
## [129] farver_2.1.2	reformulas_0.4.0
## [131] tidygraph_1.3.1	scatterpie_0.2.4
## [133] yaml_2.3.10	cli_3.6.4
## [135] lifecycle_1.0.4	uwot_0.2.2
## [137] mvtnorm_1.3-3	timechange_0.3.0
## [139] gtable_0.3.6	rjson_0.2.23
## [141] gggridges_0.5.6	progressr_0.15.1
## [143] ape_5.8-1	jsonlite_1.9.0
## [145] RcppHNSW_0.6.0	bitops_1.0-9
## [147] bit64_4.6.0-1	Rtsne_0.17
## [149] yulab.utils_0.2.0	spatstat.utils_3.1-2
## [151] bdsmatrix_1.3-7	GOSemSim_2.30.2
## [153] spatstat.univar_3.1-1	R.utils_2.12.3
## [155] lazyeval_0.2.2	shiny_1.10.0
## [157] htmltools_0.5.8.1	enrichplot_1.24.4
## [159] GO.db_3.19.1	sctransform_0.4.1
## [161] rappdirs_0.3.3	tinytex_0.55
## [163] glue_1.8.0	spam_2.11-1
## [165] httr2_1.1.0	XVector_0.44.0
## [167] RCurl_1.98-1.16	rprojroot_2.0.4
## [169] treeio_1.28.0	gridExtra_2.3
## [171] boot_1.3-31	extrafontdb_1.0
## [173] igraph_2.1.4	R6_2.6.1
## [175] SingleCellExperiment_1.26.0	labeling_0.4.3
## [177] RcppRoll_0.3.1	cluster_2.1.8
## [179] bbmle_1.0.25.1	Rhdf5lib_1.26.0
## [181] aplot_0.2.4	nloptr_2.1.1
## [183] DelayedArray_0.30.1	tidyselect_1.2.1
## [185] ggforce_0.4.2	xml2_1.3.6
## [187] future_1.34.0	rsvd_1.0.5
## [189] munsell_0.5.1	KernSmooth_2.23-26
## [191] htmlwidgets_1.6.4	fgsea_1.30.0
## [193] rlang_1.1.5	spatstat.sparse_3.1-0
## [195] spatstat.explore_3.3-4	