nf-core/differentialabundance
Abstract
Data
Results
Methods
Appendices
nf-core/differentialabundance: Citations

Differential gene abundance report

differentialabundance workflow versionv1.5.0-g3dd360f

2026-08-25

Jane Doe
Director of Institute of Microbiology
University of Smallville

John Smith
PhD student
Institute of Microbiology
University of Smallville

Abstract

This report summarises differential gene analysis as performed by the nf-core/differentialabundance pipeline.

Data

Samples

A summary of sample metadata is below:

Sample metadata

Sample	Treatment	Sample number	Sample title	Strandedness
SRX8042381	mCherry	sample1	mCherry_1	unstranded
SRX8042382	mCherry	sample2	mCherry_2	unstranded
SRX8042383	mCherry	sample3	mCherry_3	unstranded
SRX8042384	hND6	sample1	hND6_1	unstranded
SRX8042385	hND6	sample2	hND6_2	unstranded
SRX8042386	hND6	sample3	hND6_3	unstranded

Contrasts

Comparisons were made between sample groups defined using metadata columns, as described in the following table of contrasts:

Table of contrasts

Id	Variable	Reference	Target	Blocking	model
treatment_mCherry_hND6_	treatment	mCherry	hND6		~ 0 + treatment
treatment_mCherry_hND6_sample_number	treatment	mCherry	hND6	sample_number	~ 0 + sample_number + treatment

Results

Counts

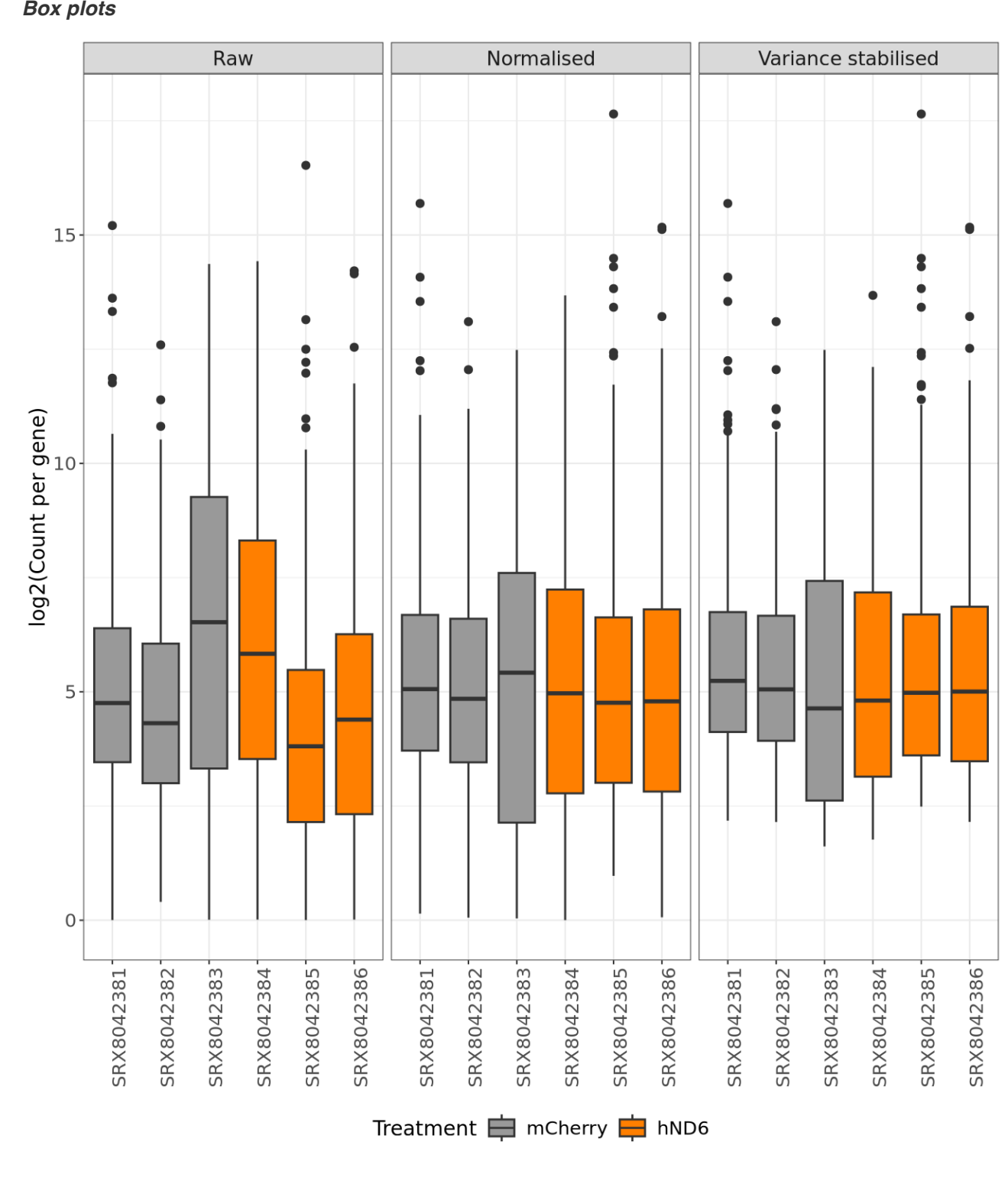
Input was a matrix of 1000 genes for 6 samples, reduced to 982 genes after filtering for low abundance.

Exploratory analysis

Abundance value distributions

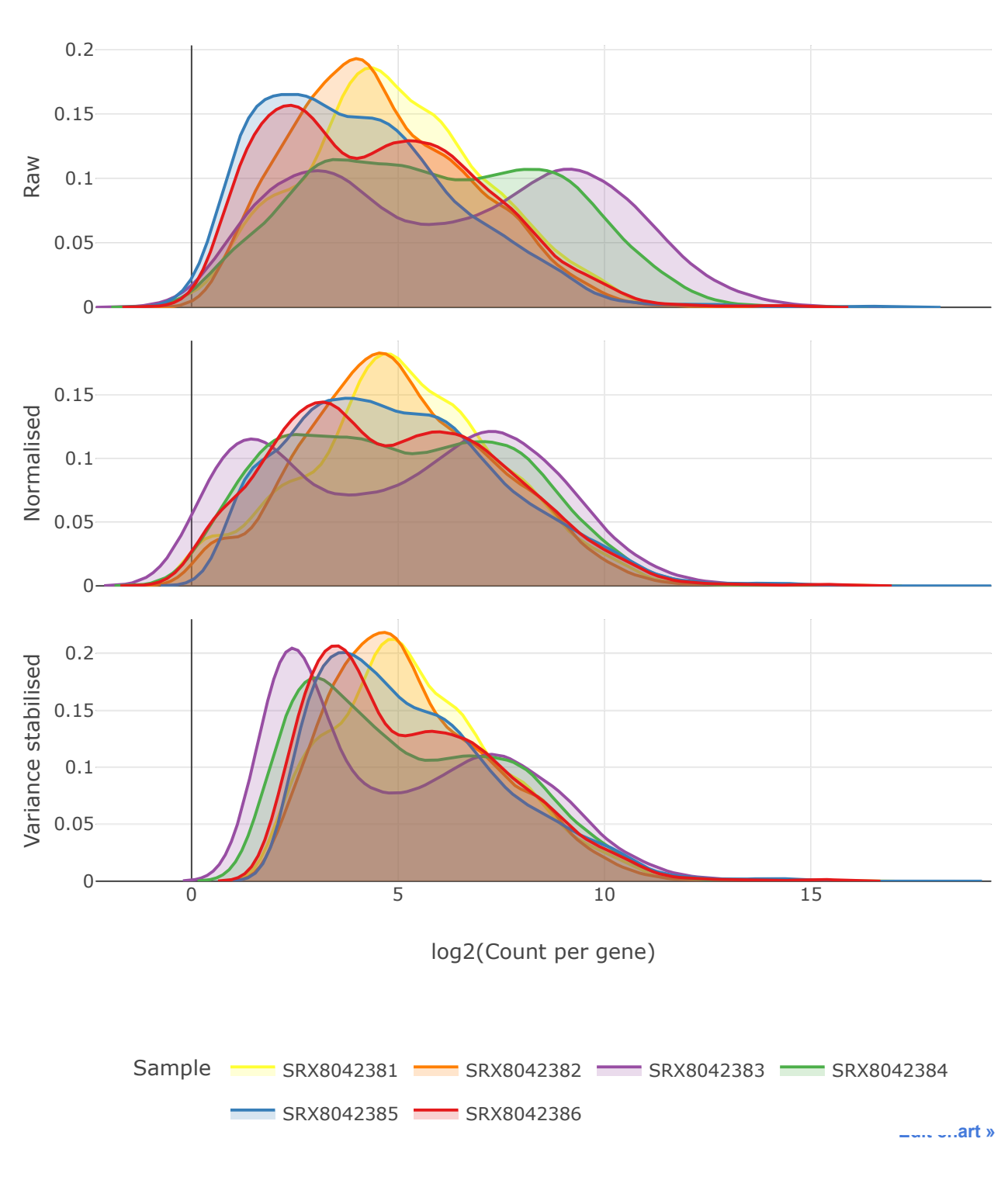
The following plots show the abundance value distributions of input matrices. A log2 transformation is applied where not already performed.

Box plots



Whiskers in the above boxplots show 1.5 times the inter-quartile range.

Density plots

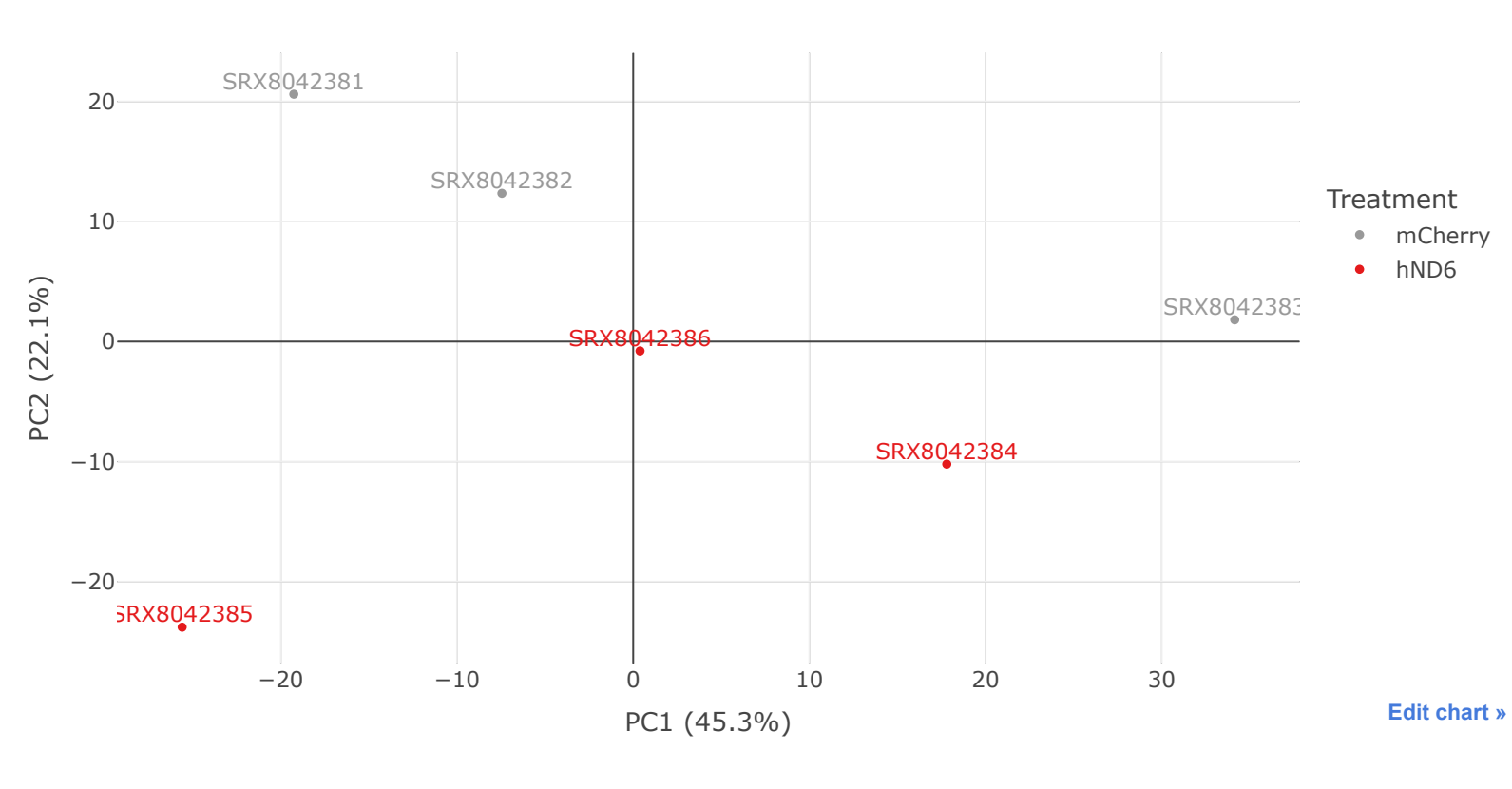


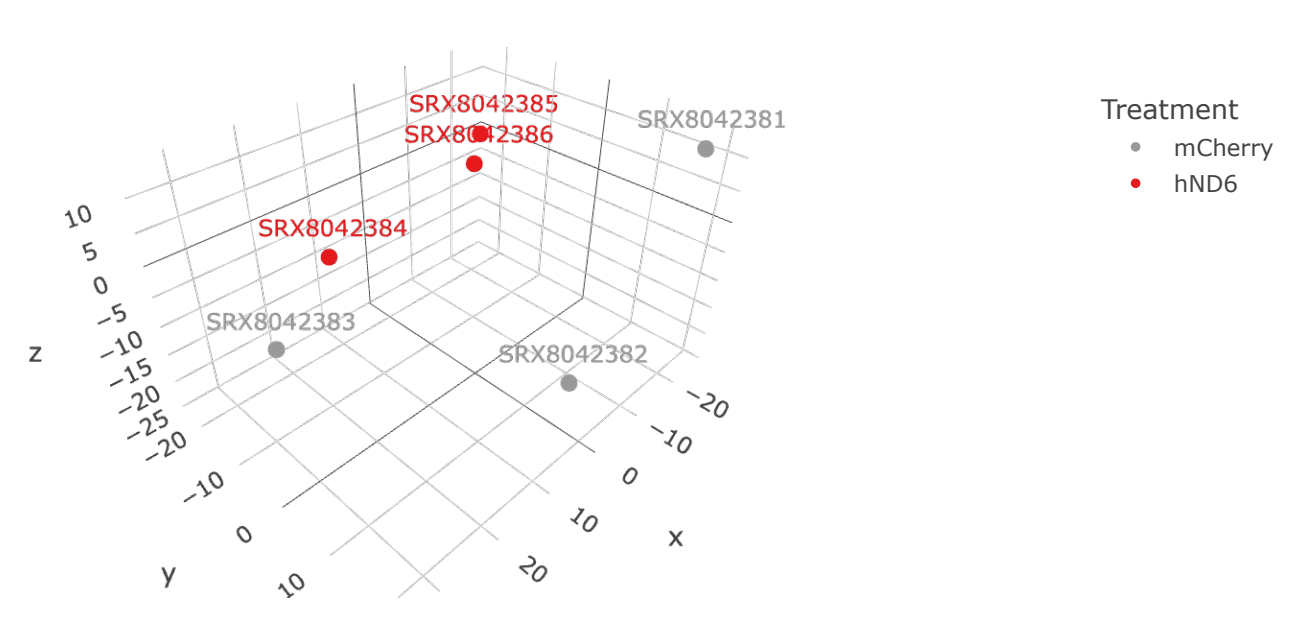
Sample relationships

Principal components plots

Principal components analysis was conducted based on the 500 most variable genes. Each component was annotated with its percent contribution to variance.

Variance stabilised (treatment) | Variance stabilised (sample_number) | Normalised (treatment) | Normalised (sample_number) | Raw (treatment) | Raw (sample_number)

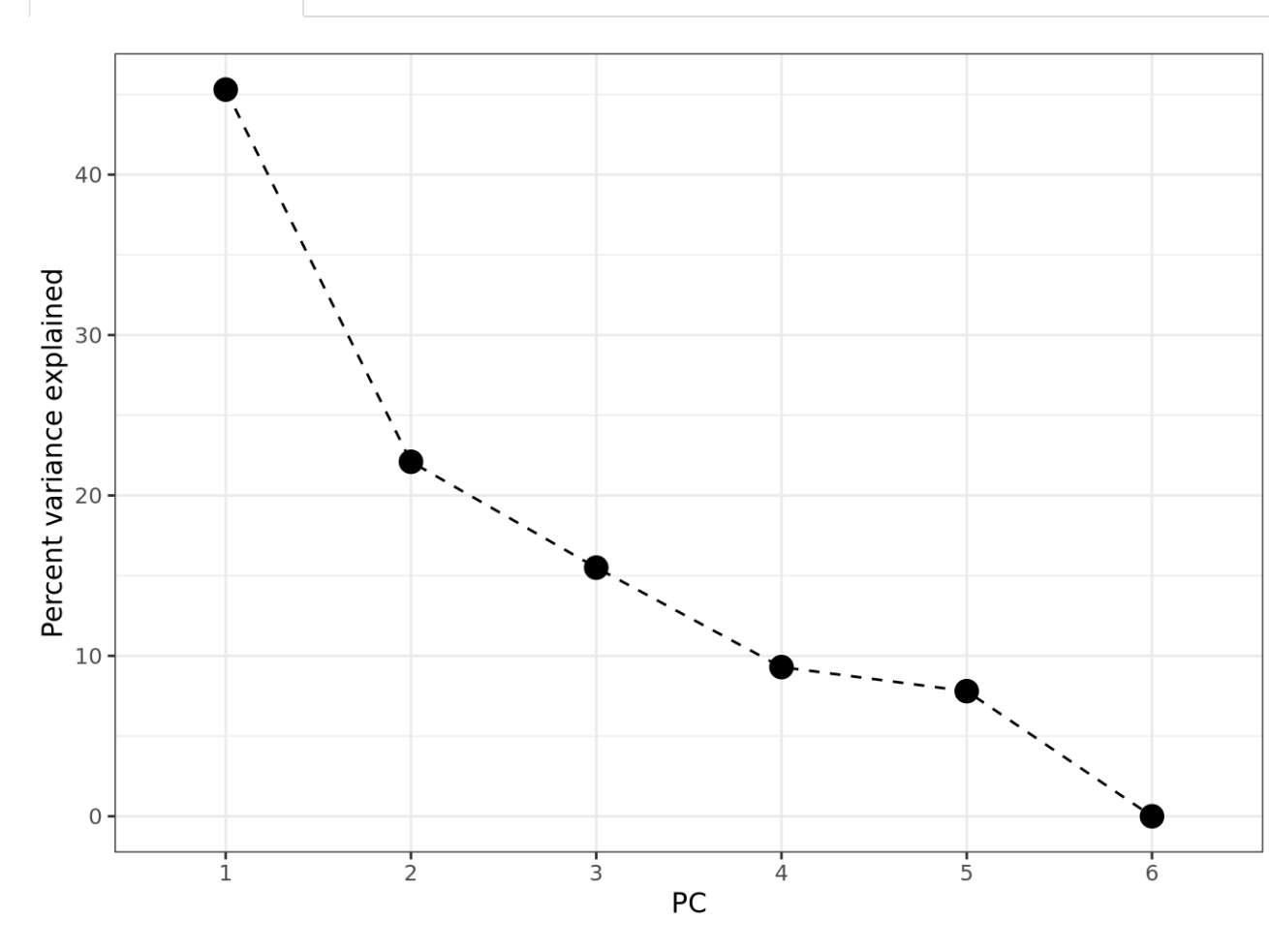




Scree plot

The following scree plot visualizes what percentage of total variation in the data can be explained by each of the principal components computed.

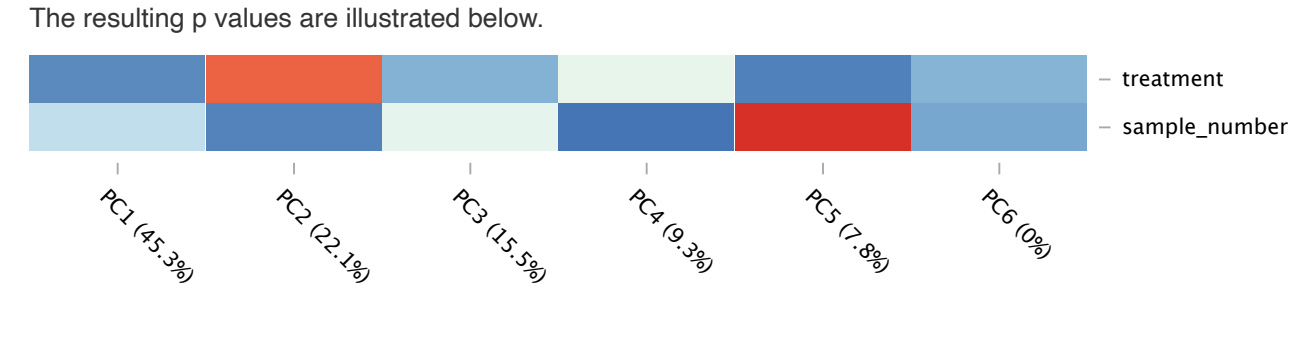
Variance stabilised | Normalised | Raw



Principal components/ metadata associations

For the variance stabilised matrix, an ANOVA test was used to determine assocations between continuous principal components and categorical covariates (including the variable of interest).

The resulting p values are illustrated below.

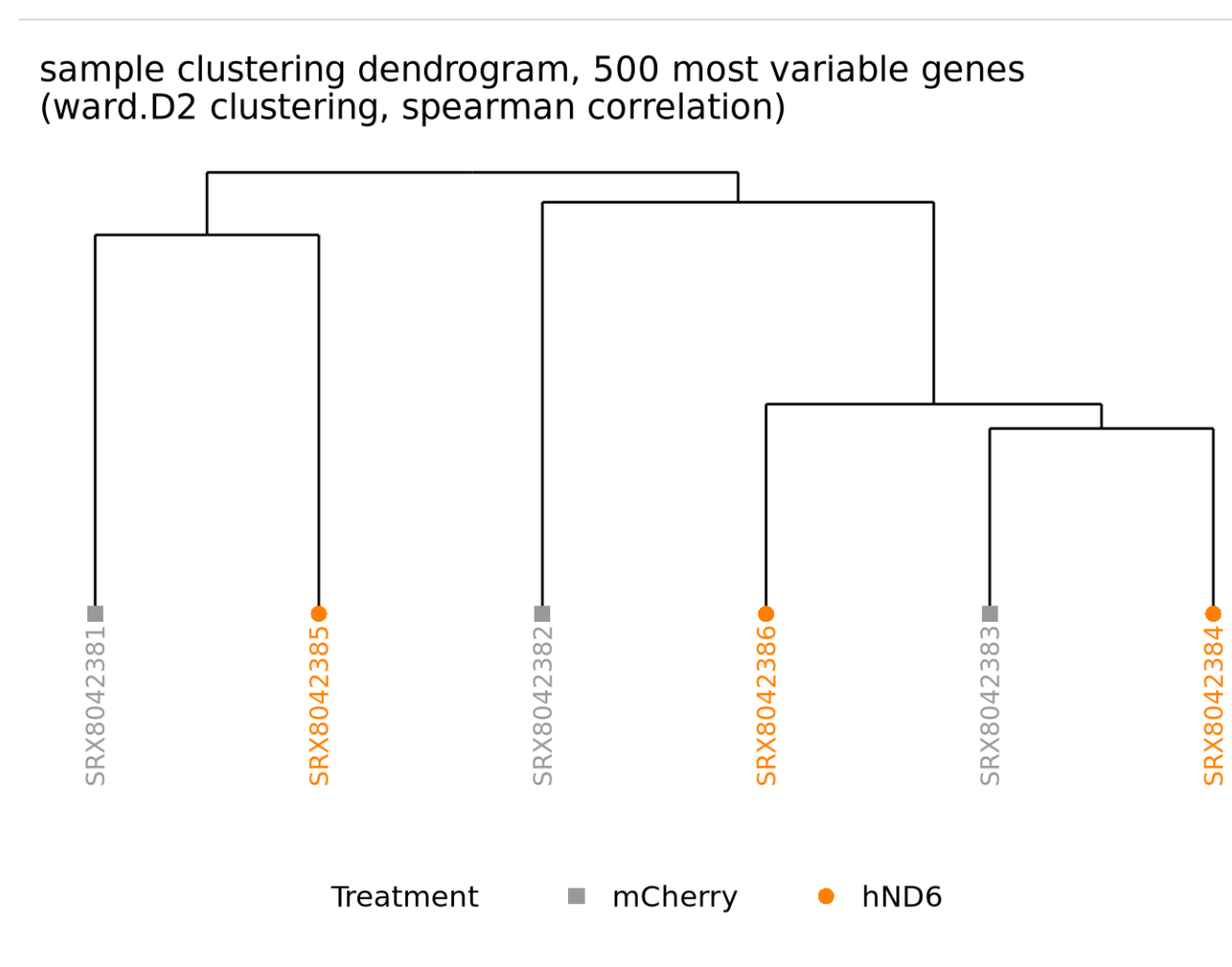


The variable 'treatment' shows an association with PC2 (22.1%) (p = 0.05). The variable 'sample_number' shows an association with PC5 (7.8%) (p = 0.04).

Clustering dendrograms

A hierarchical clustering of genes was undertaken based on the 500 most variable genes. Distances between genes were estimated based on spearman correlation, which were then used to produce a clustering via the ward.D2 method with `hclust()` in R.

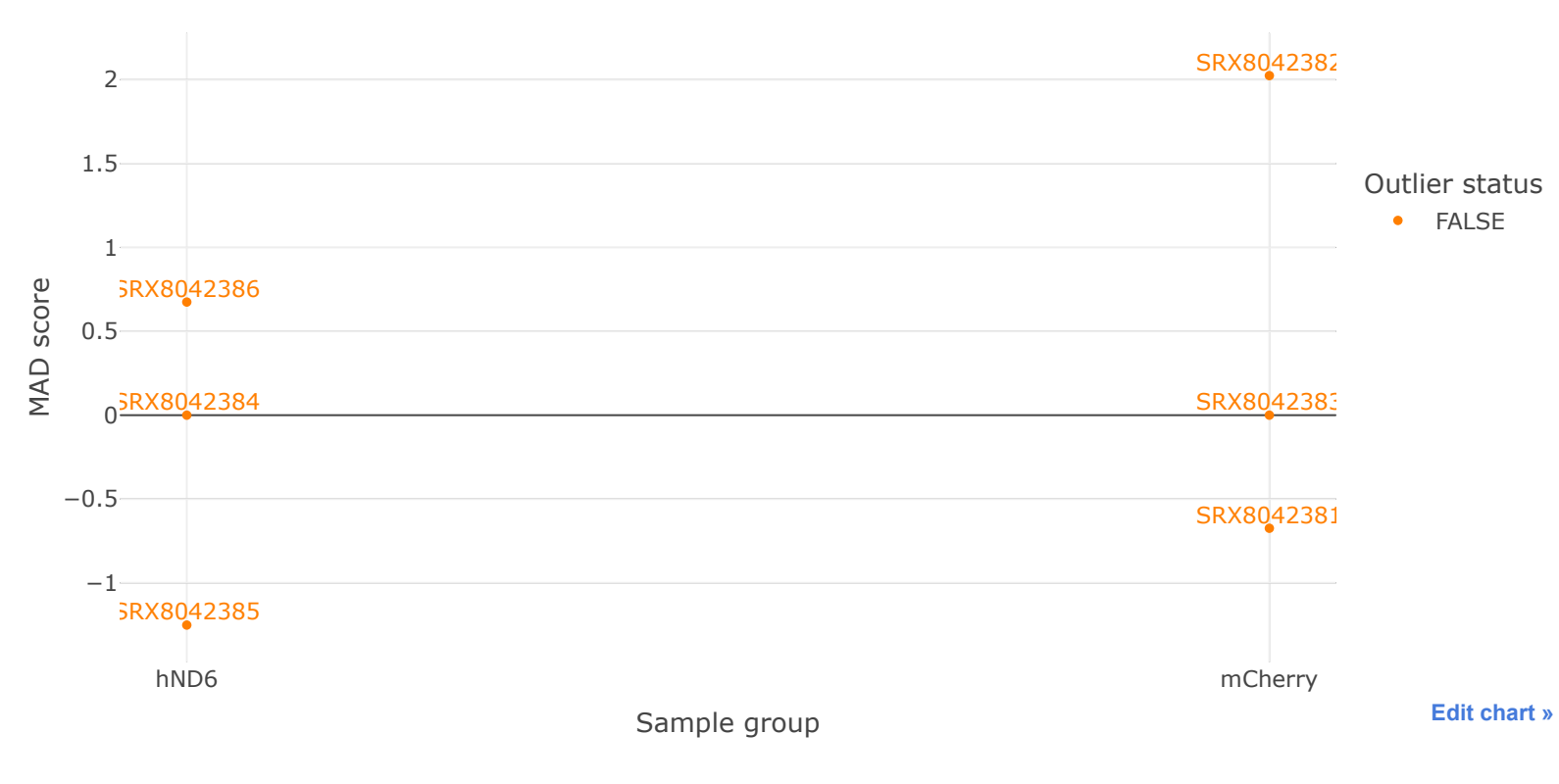
Variance stabilised (treatment) | Variance stabilised (sample_number) | Normalised (treatment) | Normalised (sample_number) | Raw (treatment) | Raw (sample_number)



Outlier detection

Outlier detection based on median absolute deviation was undertaken, the outlier scoring is plotted below.

treatment



No outlying samples were detected in groups defined by treatment.

Differential analysis

The `DESeq2 R` package was used for differential analysis. p-values were adjusted with the BH method to reduce the number of false positives. Genes were considered differential if, for the respective contrast, the adjusted p-value was equal to or lower than 0.1 and the absolute log2 fold change was equal to or higher than 0.

Differential gene counts

Adjusted | Unadjusted

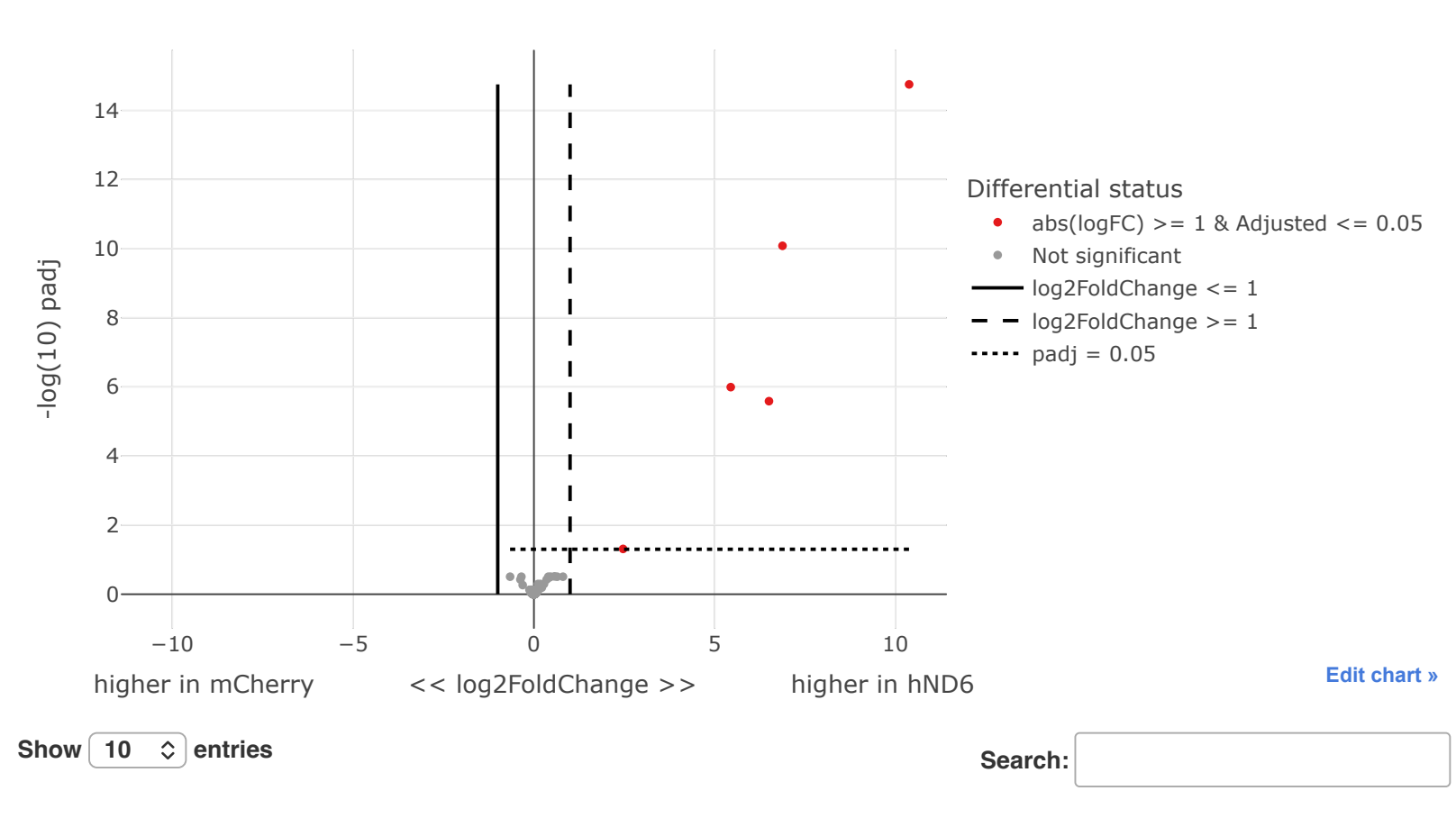
Differential gene (target relative to reference)

	up	down
hND6 versus mCherry in treatment	5	0
hND6 versus mCherry in treatment (blocking: sample_number)	6	0

Differential gene details

hND6 versus mCherry in treatment

Adjusted p values | Unadjusted p values



Show 10 entries Search:

Differential genes up in hND6 versus mCherry in treatment (check .//treatment_mCherry_hND6_.deseq2.results.tsv for more detail)

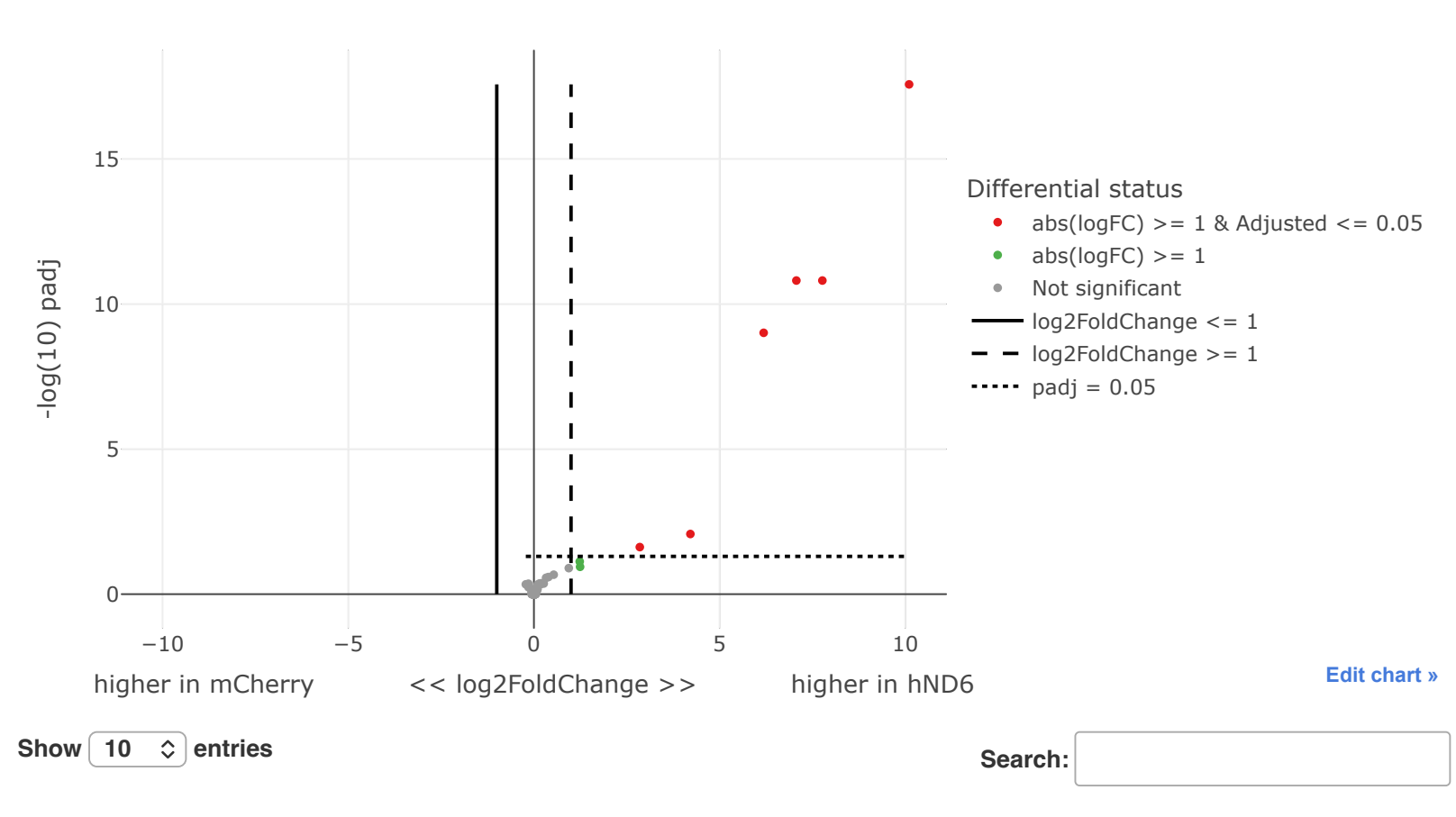
Gene id	Gene name	Gene biotype	Pvalue	Padj	Log2FoldChange
ENSMUSG00000032053	Pou2af1	protein_coding	0.000255	0.0489	2.47
ENSMUSG00000067149	Jchain	protein_coding	1.73e-13	8.29e-11	6.87
ENSMUSG00000076609	Igkc	IG_C_gene	1.08e-8	0.0000026	6.5
ENSMUSG00000076613	Ighg2b	IG_C_gene	1.84e-18	1.77e-15	10.4
ENSMUSG00000095079	Igha	IG_C_gene	3.18e-9	0.00000102	5.44

Showing 1 to 5 of 5 entries Previous 1 Next

No significantly differential 'down' genes.

hND6 versus mCherry in treatment (blocking: sample_number)

Adjusted p values | Unadjusted p values



Show 10 entries Search:

Differential genes up in hND6 versus mCherry in treatment (blocking: sample_number) (check .//treatment_mCherry_hND6_sample_number.deseq2.results.tsv for more detail)

Gene id	Gene name	Gene biotype	Pvalue	Padj	Log2FoldChange
ENSMUSG00000025386	Pde6g	protein_coding	0.000146	0.0239	2.85
ENSMUSG00000067149	Jchain	protein_coding	4.73e-14	1.55e-11	7.07
ENSMUSG00000076609	Igkc	IG_C_gene	3.88e-14	1.55e-11	7.77
ENSMUSG00000076613	Ighg2b	IG_C_gene	2.73e-21	2.68e-18	10.1
ENSMUSG00000095079	Igha	IG_C_gene	4e-12	9.83e-10	6.19
ENSMUSG00000105906	Iglc1	IG_C_gene	0.0000431	0.00847	4.21

Showing 1 to 6 of 6 entries Previous 1 Next

No significantly differential 'down' genes.

Gene set analysis

GSEA

mh.all.v2022.1.Mm.symbols

hND6 versus mCherry in treatment | hND6 versus mCherry in treatment (blocking: sample_number)

Show 10 entries Search:

Target (hND6)

NAME	SIZE	ES	NES	NOM p-val	FDR q-val	FWER p-val	RANK AT MAX	LEADING EDGE	Var.10
HALLMARK_KRAS_SIGNALING_DN	15	0.367	0.949	0.537	0.63	0.273	172	tags=33%, list=18%, signal=40%	

Showing 1 to 1 of 1 entries Previous 1 Next

Show 10 entries Search:

Reference (mCherry)

NAME	SIZE	ES	NES	NOM p-val	FDR q-val	FWER p-val	RANK AT MAX	LEADING EDGE	Var.10
No data available in table									

Showing 0 to 0 of 0 entries Previous Next

Methods

Filtering

Parameters used for feature-wise filtering

Parameter	Value
min_samples	1
min_abundance	10
min_proportion_not_na	0.5

Filtering was carried out by selecting genes with an abundance of at least 10 in at least a proportion of 0.5 of samples.

Exploratory analysis

Parameters used for exploratory analysis

Parameter	Value
log2_assays	raw,normalised
n_features	500
clustering_method	ward.D2
cor_method	spearman
whisker_distance	1.5
mad_threshold	-5
main_variable	contrasts
assay_names	raw,normalised,variance_stabilised
final_assay	variance_stabilised
palette_name	Set1

Differential analysis

Parameters used for DESeq2

Parameter	Value
test	Wald
fit_type	parametric
sf_type	ratio
min_replicates_for_replace	7
use_t	FALSE
lfc_threshold	0
alt_hypothesis	greaterAbs
independent_filtering	TRUE
p_adjust_method	BH
alpha	0.1

Previous 1 2 Next

Parameters used for downstream differential analysis

Parameter	Value
table	file.csv
file_suffix	.deseq2.results.tsv
feature_id_column	gene_id
feature_name_column	gene_name
fc_column	log2FoldChange
pval_column	pvalue
qval_column	padj
min_fold_change	2
foldchanges_logged	TRUE
max_pval	1

Previous 1 2 Next

Gene set analysis

GSEA

Parameters used for GSEA

Parameter	Value
run	TRUE
nperm	1000
permute	phenotype
scoring_scheme	weighted
metric	Signal2Noise
sort	real
order	descending
set_max	500
set_min	15
norm	meandiv

Previous 1 2 Next

Appendices

All parameters

Show 10 entries Search:

All parameters

Parameter	Value
meta.id	SRP254919
input_dir	./
artifact_dir	artifacts
cpus	2
study_type	rnaseq
study_name	SRP254919
study_abundance_type	counts
report_contributors	Jane Doe Director of Institute of Microbiology University of Smallville;John Smith PhD student Institute of Microbiology University of Smallville
report_scree	TRUE
report_round_digits	3

Showing 1 to 10 of 91 entries Previous 1 2 3 4 5 … 10 Next

Software versions

Note: For a more detailed accounting of the software and commands used (including containers), consult the execution report produced as part of the 'pipeline info' for this workflow.

Search:

Software versions

Component	Software	Version
DESEQ2_DIFFERENTIAL	r-base	4.1.3
DESEQ2_DIFFERENTIAL	bioconductor-deseq2	1.34.0
GSEA_GSEA	gsea	4.3.2
GTF_TO_TABLE	atlas-gene-annotation-manipulation	1.1.1
GUNZIP_GTF	gunzip	1.1
PLOT_DIFFERENTIAL	r-base	4.3.3
PLOT_DIFFERENTIAL	r-shinyngs	1.8.8
PLOT_EXPLORATORY	r-shinyngs	1.8.8
TABULAR_TO_GSEA_CHIP	bash	4.2.25 3
VALIDATOR	r-base	4.3.3
VALIDATOR	r-shinyngs	1.8.8
Workflow	nf-core/differentialabundance	v1.5.0-g3dd360f
Workflow	Nextflow	26.04.2

nf-core/differentialabundance: Citations

nf-core

> Ewels PA, Peltzer A, Fillinger S, Patel H, Alneberg J, Wilm A, Garcia MU, Di Tommaso P, Nahnsen S. The nf-core framework for community-curated bioinformatics pipelines. Nat Biotechnol. 2020 Mar;38(3):276-278. doi: 10.1038/s41587-020-0439-x. PubMed PMID: 32055031.

Nextflow

> Di Tommaso P, Chatzou M, Floden EW, Barja PP, Palumbo E, Notredame C. Nextflow enables reproducible computational workflows. Nat Biotechnol. 2017 Apr 11;35(4):316-319. doi: 10.1038/nbt.3820. PubMed PMID: 28398311.

Pipeline tools

- GSEA

  > Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102(43):15545-15550.

R packages

- affy

  > Gautier L, Cope L, Bolstad BM, Irizarry RA. Affy–analysis of affymetrix genechip data at the probe level. Bioinformatics. 2004;20(3):307-315.

- DESeq2

  > Love MI, Huber W, Anders S (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15(12):550. PubMed PMID: 25516281; PubMed Central PMCID: PMC4302049.

- GEOQuery

  > Davis S, Meltzer PS. Geoquery: a bridge between the gene expression omnibus (Geo) and bioconductor. Bioinformatics. 2007;23(14):1846-1847.

- ggplot2

  > H. Wickham (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

- gprofiler2

  > Kolberg L, Raudvere U, Kuzmin I, Vilo J, Peterson H (2020). "gprofiler2– an R package for gene list functional enrichment analysis and namespace conversion toolset g:Profiler." F1000Research, 9 (ELIXIR)(709). R package version 0.2.2.

- Limma

  > Ritchie ME, Phipson B, Wu D, et al. Limma powers differential expression analyses for rna-sequencing and microarray studies. Nucleic Acids Res. 2015;43(7):e47.

- optparse

  > Trevor L Davis (2018). optparse: Command Line Option Parser.

- plotly

> C. Sievert (2020). Interactive Web-Based Data Visualization with R, plotly, and shiny. Chapman and Hall/CRC Florida.

- Proteus

  > Gierlinski M, Gastaldello F, Cole C, Barton GJ. Proteus : An r Package for Downstream Analysis of Maxquant Output. Bioinformatics; 2018.

- R

  > R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

- RColorBrewer

  > Erich Neuwirth (2014). RColorBrewer: ColorBrewer Palettes.

- RMarkdown

  > JJ Allaire and Yihui Xie and Jonathan McPherson and Javier Luraschi and Kevin Ushey and Aron Atkins and Hadley Wickham and Joe Cheng and Winston Chang and Richard Iannone (2022). rmarkdown: Dynamic Documents for R.

- shinyngs

  > Jonathan R Manning (2022). Shiny apps for NGS etc based on reusable components created using Shiny modules. Computer software. Vers. 1.5.3. Jonathan Manning, Dec. 2022. Web.

- SummarizedExperiment

  > Morgan M, Obenchain V, Hester J and Pagès H (2020). SummarizedExperiment: SummarizedExperiment container.

Software packaging/containerisation tools

- Anaconda

  > Anaconda Software Distribution. Computer software. Vers. 2-2.4.0. Anaconda, Nov. 2016. Web.

- Bioconda

  > Grüning B, Dale R, Sjödin A, Chapman BA, Rowe J, Tomkins-Tinch CH, Valieris R, Köster J; Bioconda Team. Bioconda: sustainable and comprehensive software distribution for the life sciences. Nat Methods. 2018 Jul;15(7):475-476. doi: 10.1038/s41592-018-0046-7. PubMed PMID: 29967506.

- BioContainers

  > da Veiga Leprevost F, Grüning B, Aflitos SA, Röst HL, Uszkoreit J, Barsnes H, Vaudel M, Moreno P, Gatto L, Weber J, Bai M, Jimenez RC, Sachsenberg T, Pfeuffer J, Alvarez RV, Griss J, Nesvizhskii AI, Perez-Riverol Y. BioContainers: an open-source and community-driven framework for software standardization. Bioinformatics. 2017 Aug 15;33(16):2580-2582. doi: 10.1093/bioinformatics/btx192. PubMed PMID: 28379341; PubMed Central PMCID: PMC5870671.

- Docker

  > Merkel, D. (2014). Docker: lightweight linux containers for consistent development and deployment. Linux Journal, 2014(239), 2. doi: 10.5555/2600239.2600241.

- Singularity

  > Kurtzer GM, Sochat V, Bauer MW. Singularity: Scientific containers for mobility of compute. PLoS One. 2017 May 11;12(5):e0177459. doi: 10.1371/journal.pone.0177459. eCollection 2017. PubMed PMID: 28494014; PubMed Central PMCID: PMC5426675.