

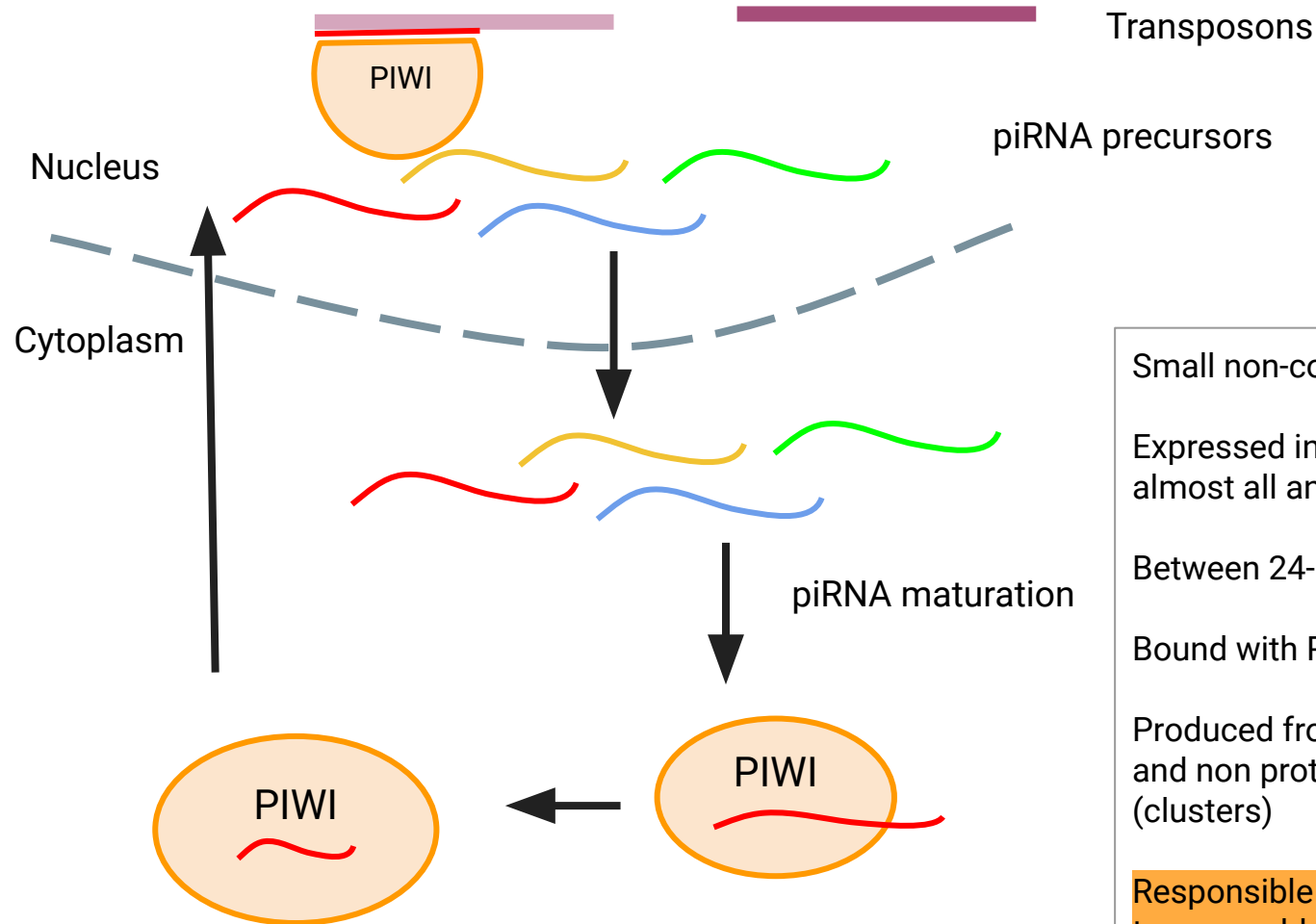
Analysis of variation in PIWI-interacting RNA (piRNA) expression in testes of different mouse strains.

Student: Pío Sierra

Tutor: Tanya Vavouri

MÁSTER UNIVERSITARIO EN BIOINFORMÁTICA Y BIOESTADÍSTICA UOC-UB
Master final project: 2020/2021

piRNAs - What you need to know:



Small non-coding RNA

Expressed in the germline in almost all animals

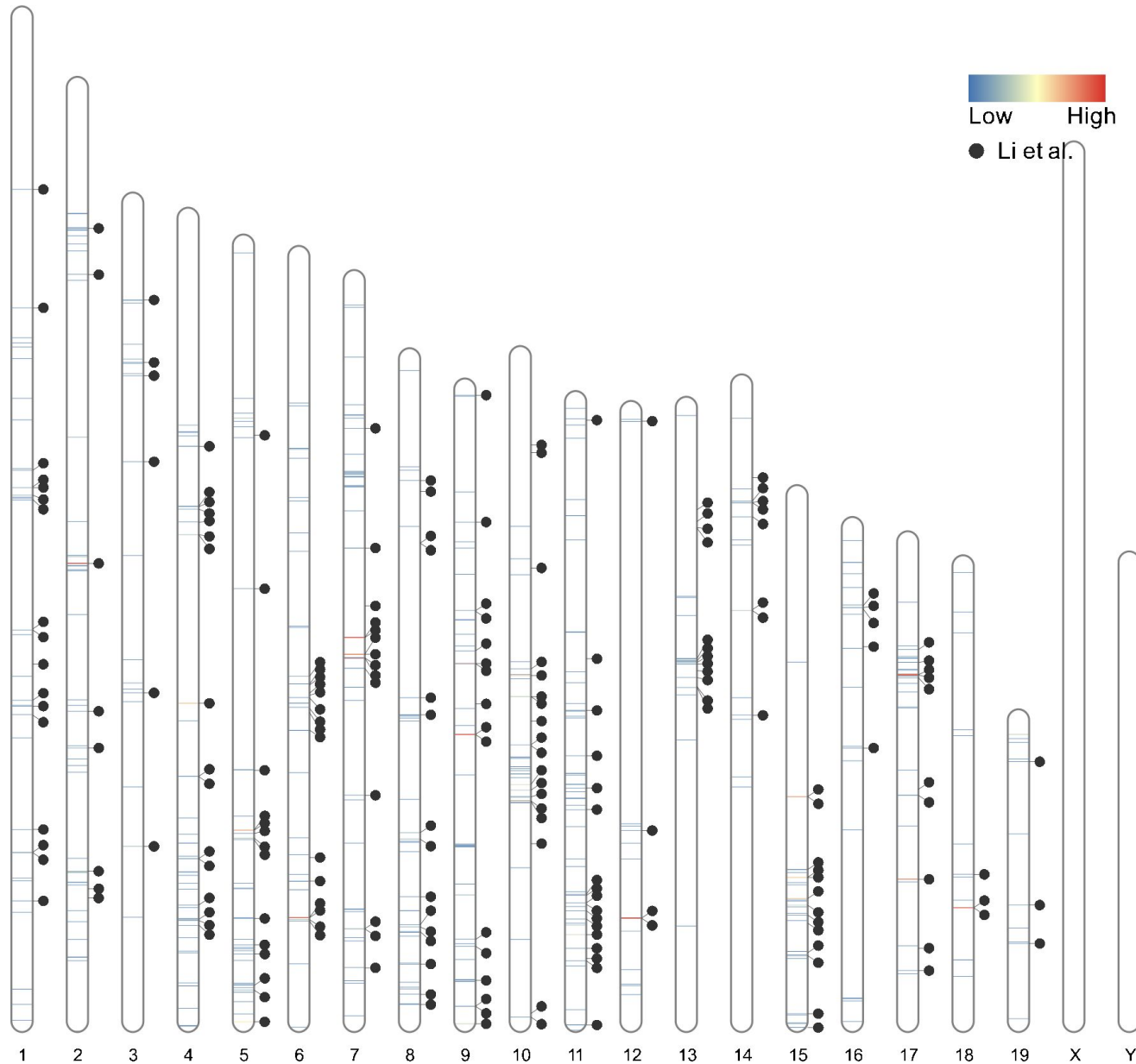
Between 24-31 bases long

Bound with PIWI proteins

Produced from protein coding and non protein coding regions (clusters)

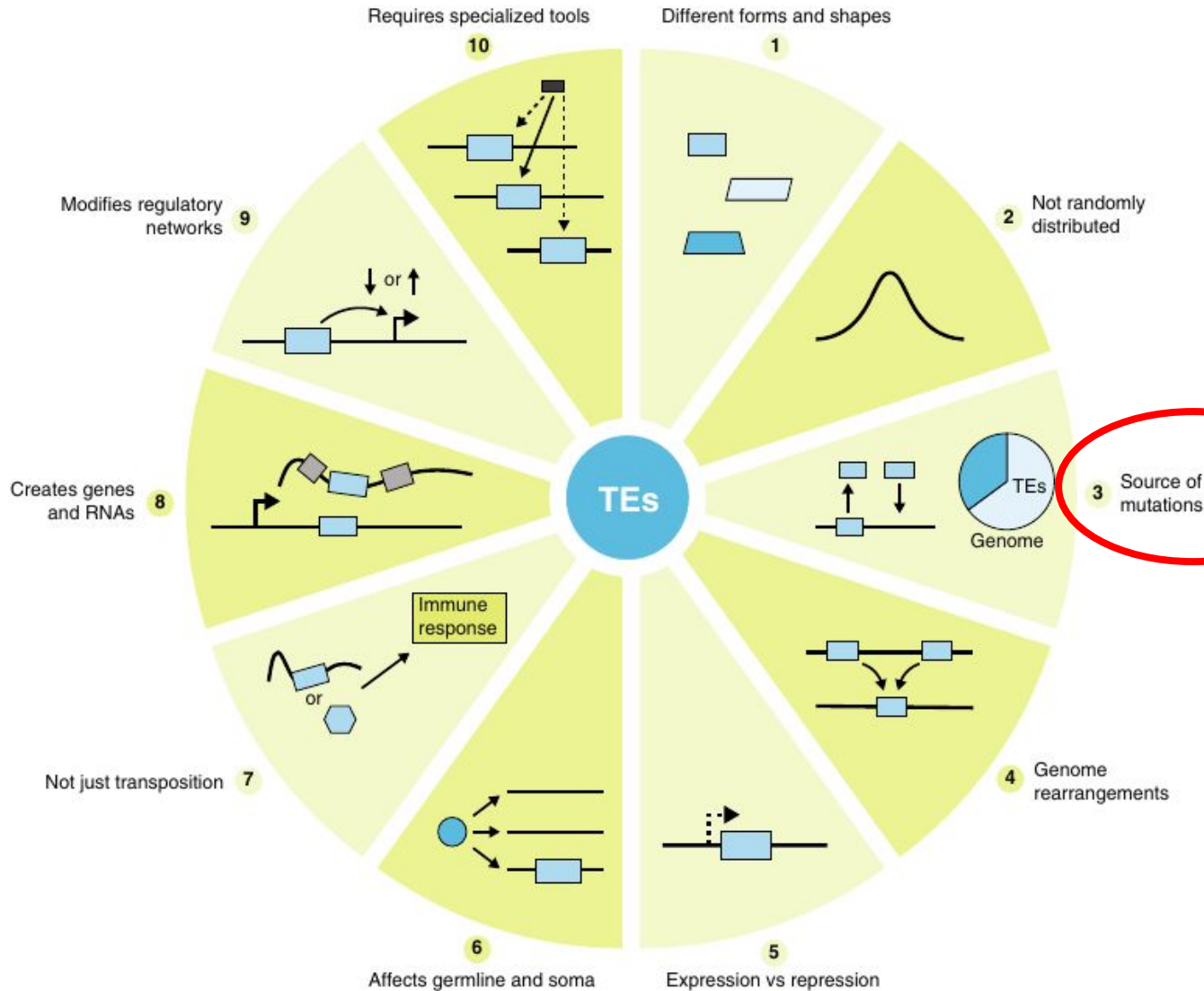
Responsible for silencing transposable elements (TEs)

piRNA producing loci (clusters) - What you need to know:



Transposable elements - What you need to know:

DNA sequences that have the ability to change their position within a genome



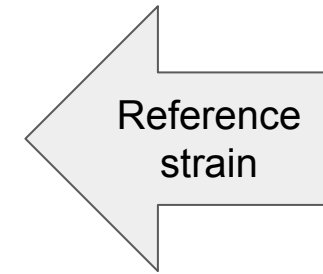
Objectives:

1- Test for variation in piRNA expression between mouse strains.

2- Test whether variable transposable elements are associated with this variation.



C57BL/6NJ
(BL6)



Reference
strain



C3H/HeJ
(C3H)



NOD/ShLtJ
(NOD)









CAST/EiJ
(CAST)



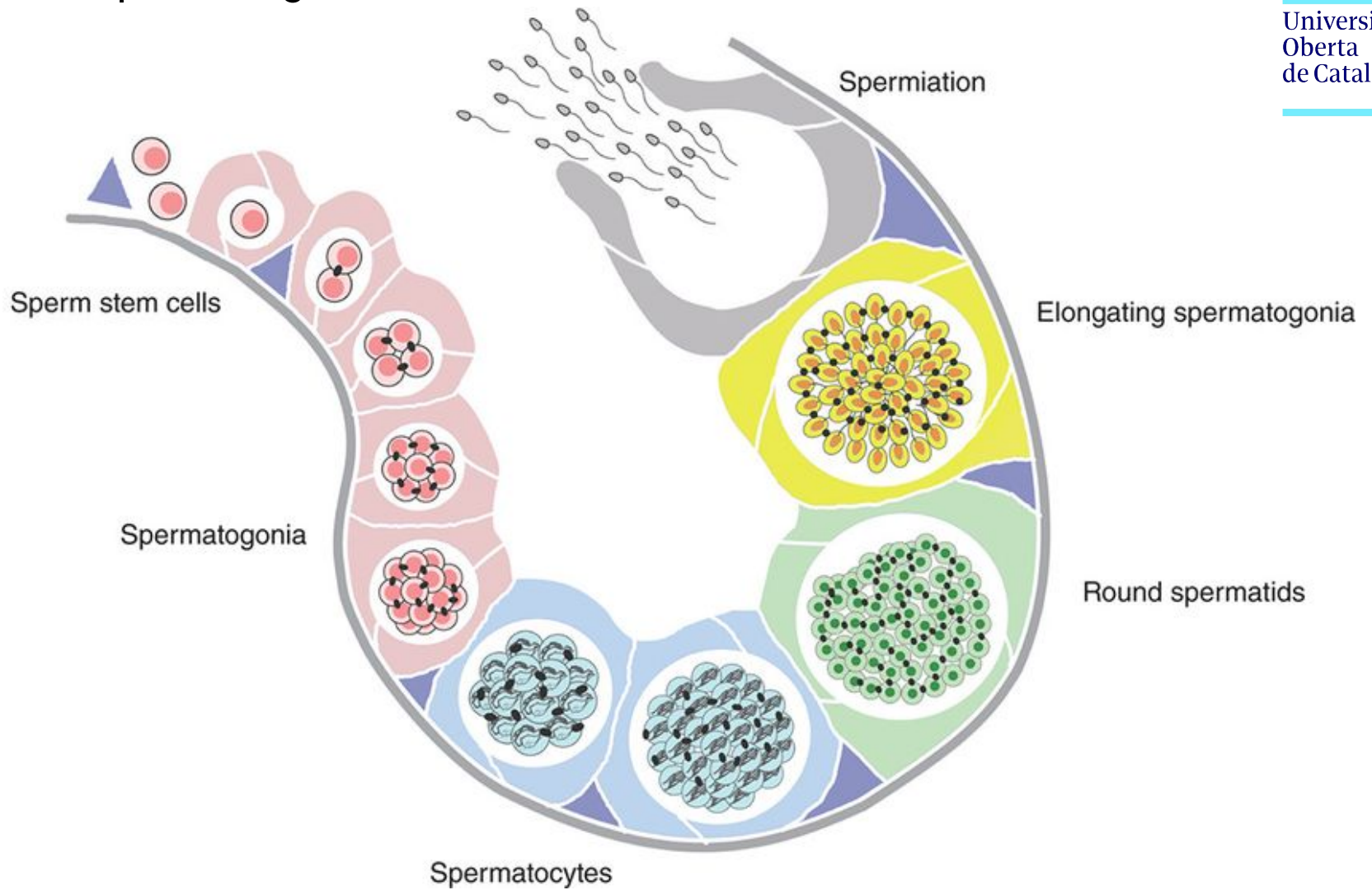
129S1/SvImJ
(129)

Data preparation:

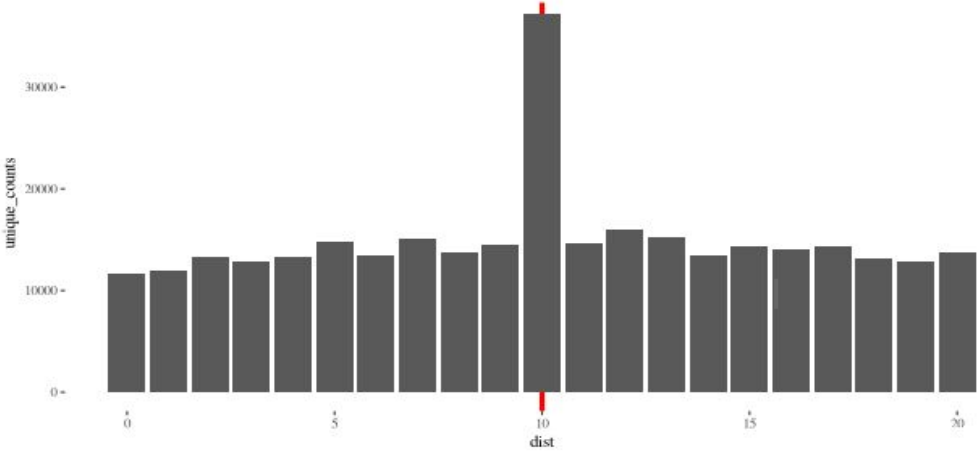
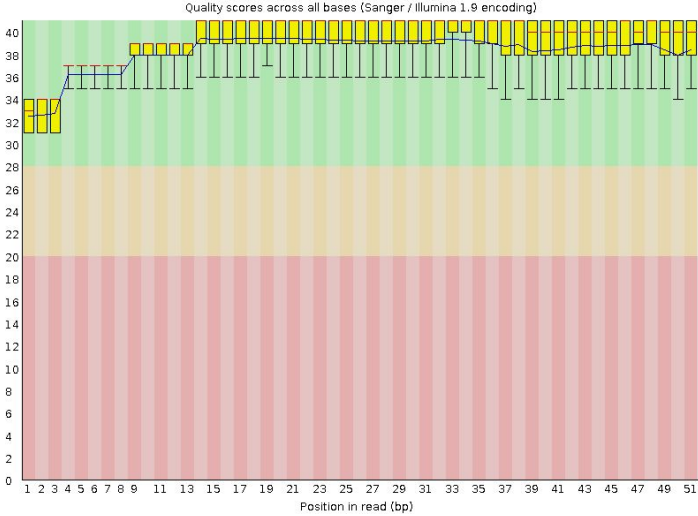
18 samples from 5 inbred strains

	BL6	SPERMATOGONIA	4 samples
	CAST	SPERMATOGONIA	4 samples
	129	TESTIS	3 samples
	C3H	TESTIS	3 samples
	NOD	TESTIS	2 samples
	BL6	TESTIS	2 samples

What is spermatogonia



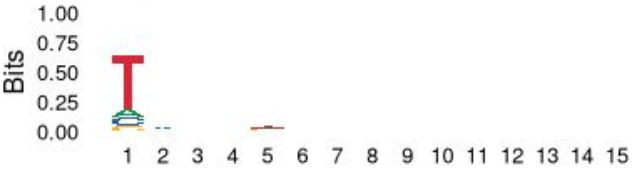
Project execution: Exploratory Analysis



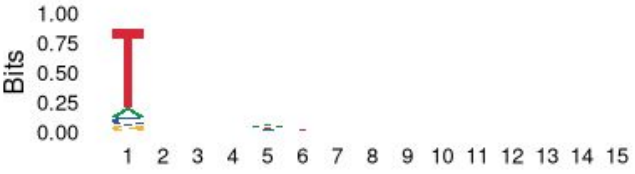
Seq < 24n : 15 first



Seq 24:28n : 15 first



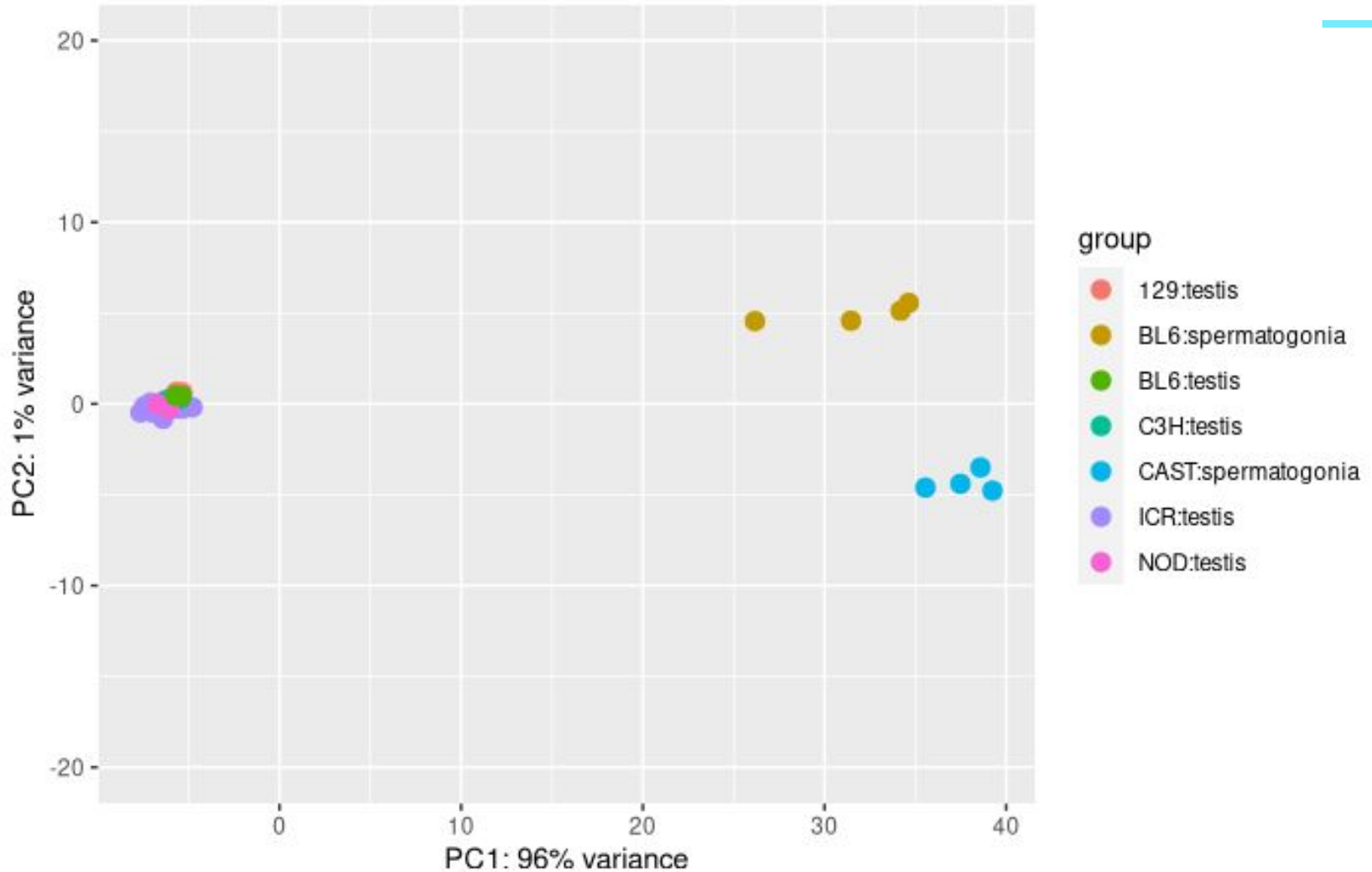
Seq 29:31n : 15 first



Seq > 31n : 15 first

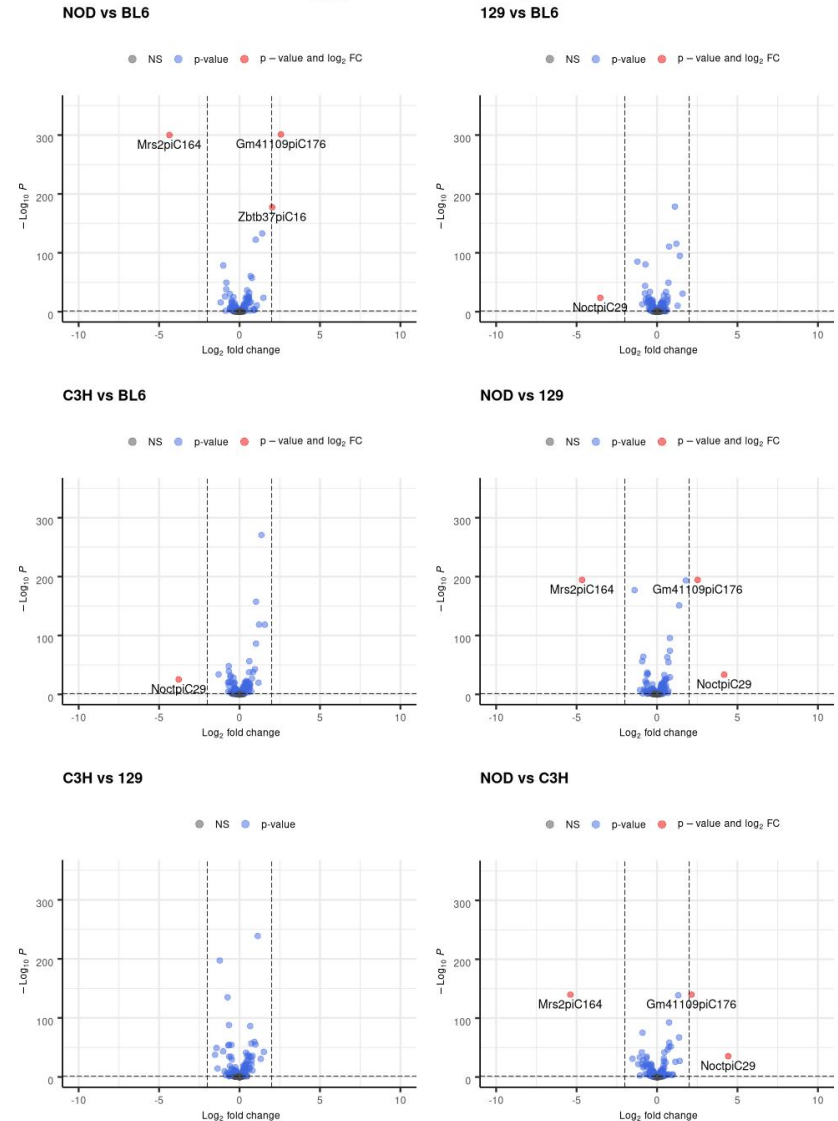
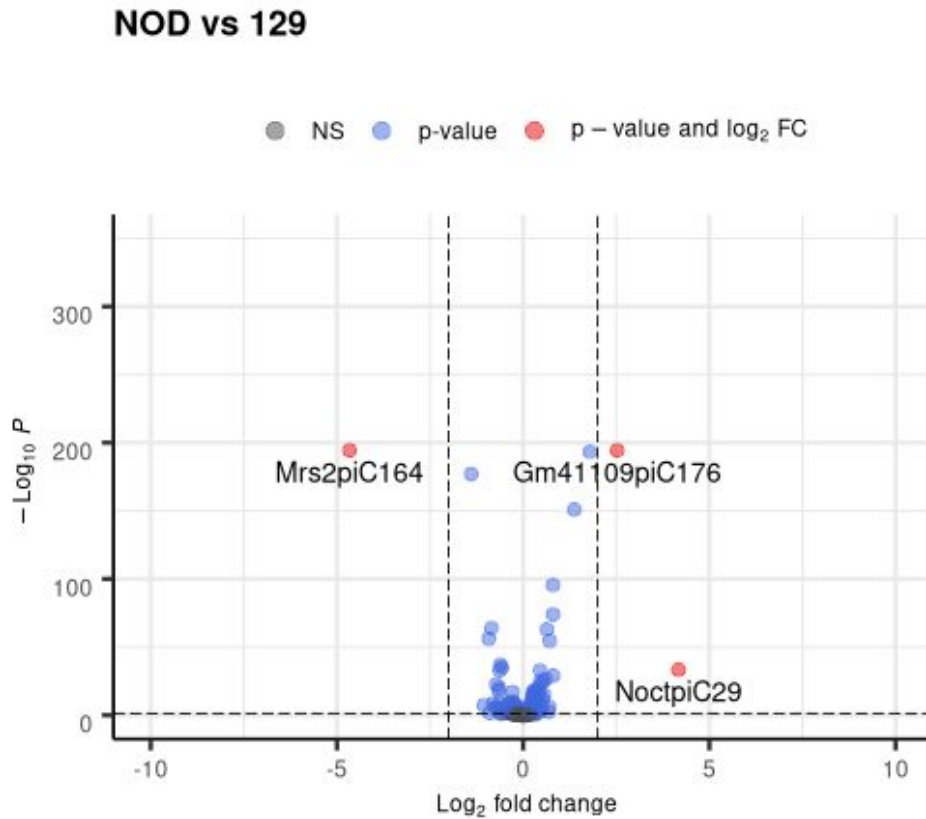


Project execution: Exploratory Analysis



Project execution: Differential Expression Analysis (1)

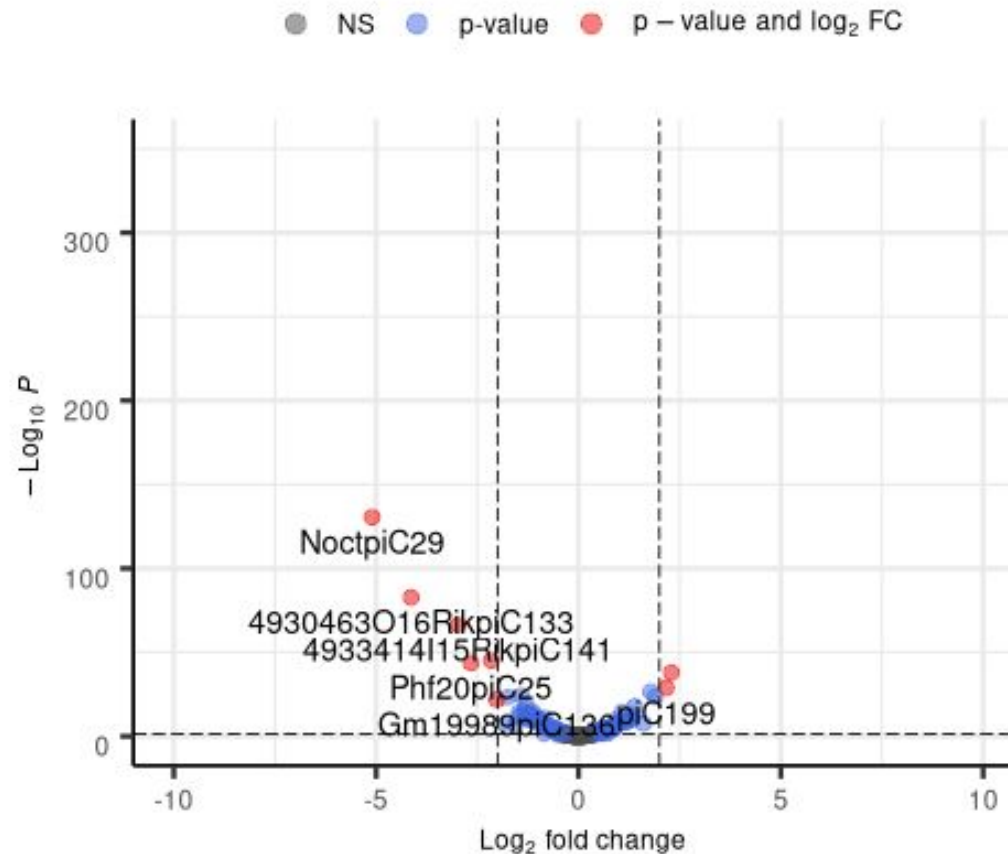
Expression of piRNA clusters of inbred testis samples by strain



Project execution: Differential Expression Analysis (1)

Expression of piRNA clusters of inbred spermatogonia samples by strain

CAST vs BL6



Project execution: prediction of piRNA clusters

611 clusters predicted

fasta files from samples



Remove redundant sequences
collapse



Remove low complexity reads
duster



Generate map from the
sequence reads
sRNAmapper



Ensembl
gene sets

Generate cluster predictions
proTRAC

RepeatMasker
annotation



Filter clusters detected on
just one sample



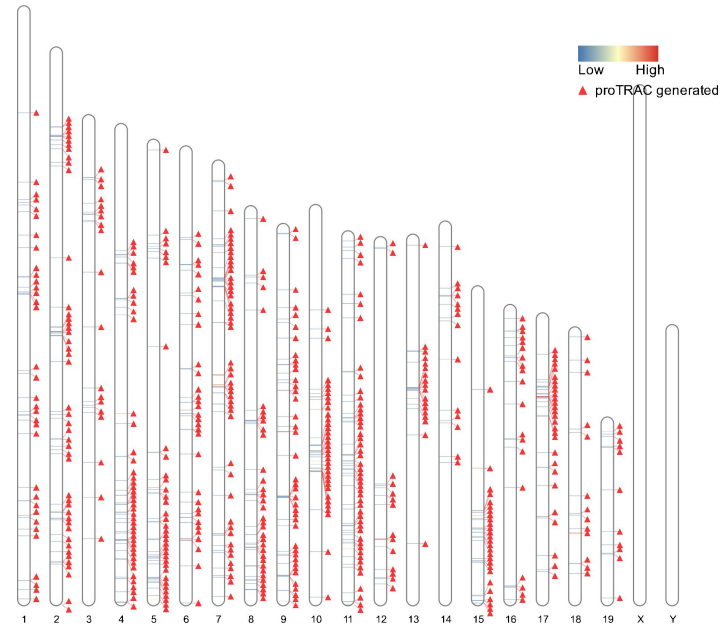
Merge samples from
all strains



Remove clusters
inside repeats

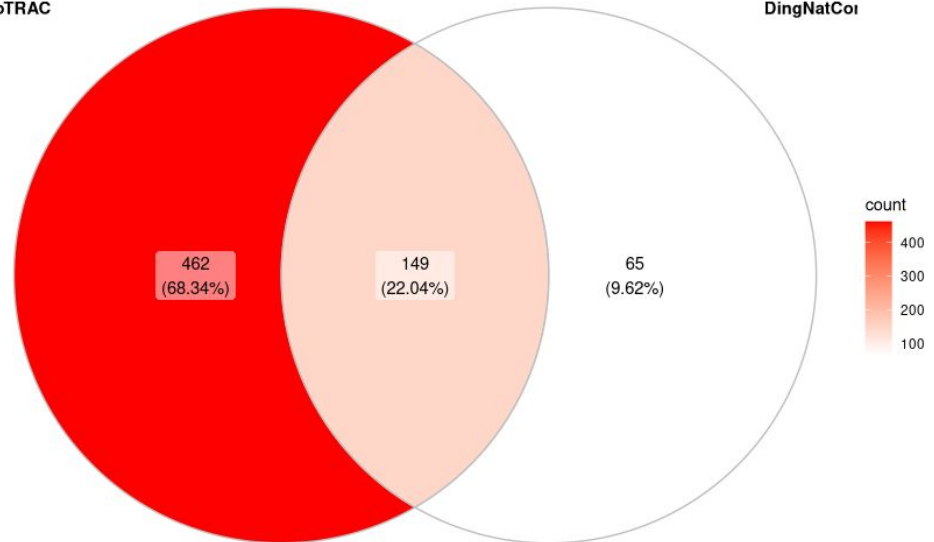


611 piRNA clusters



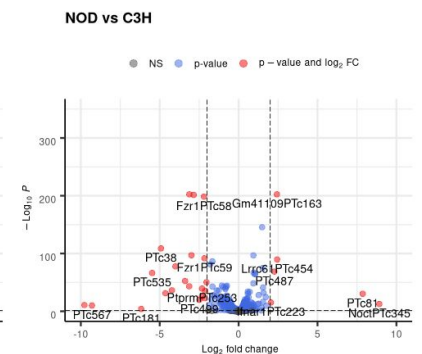
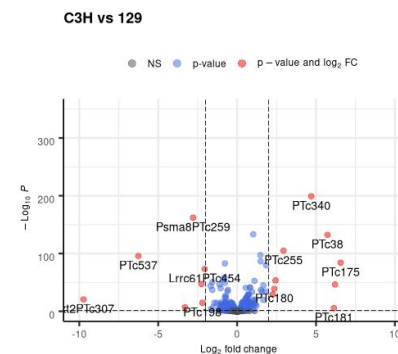
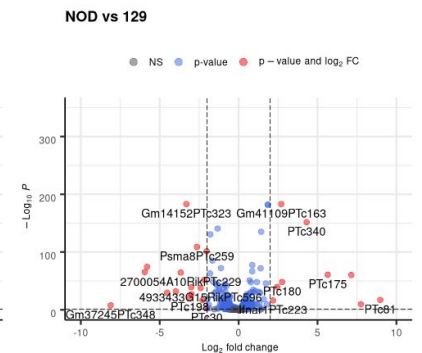
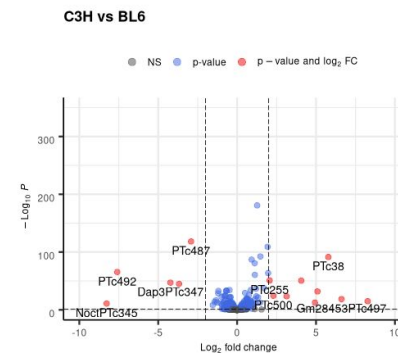
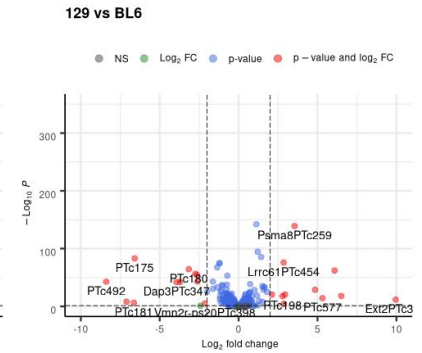
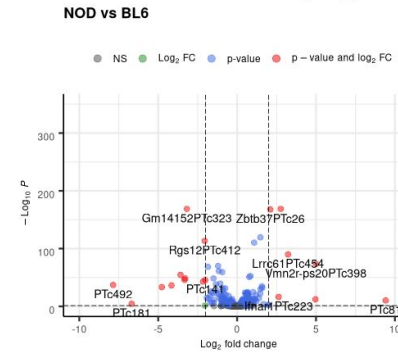
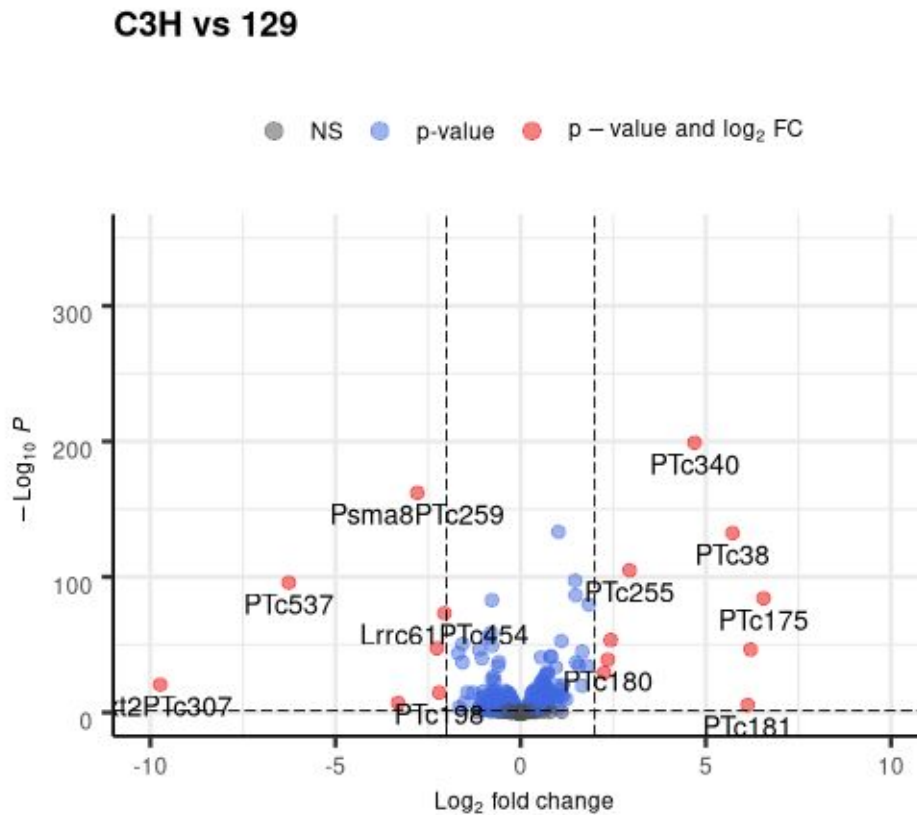
proTRAC

DingNatCor



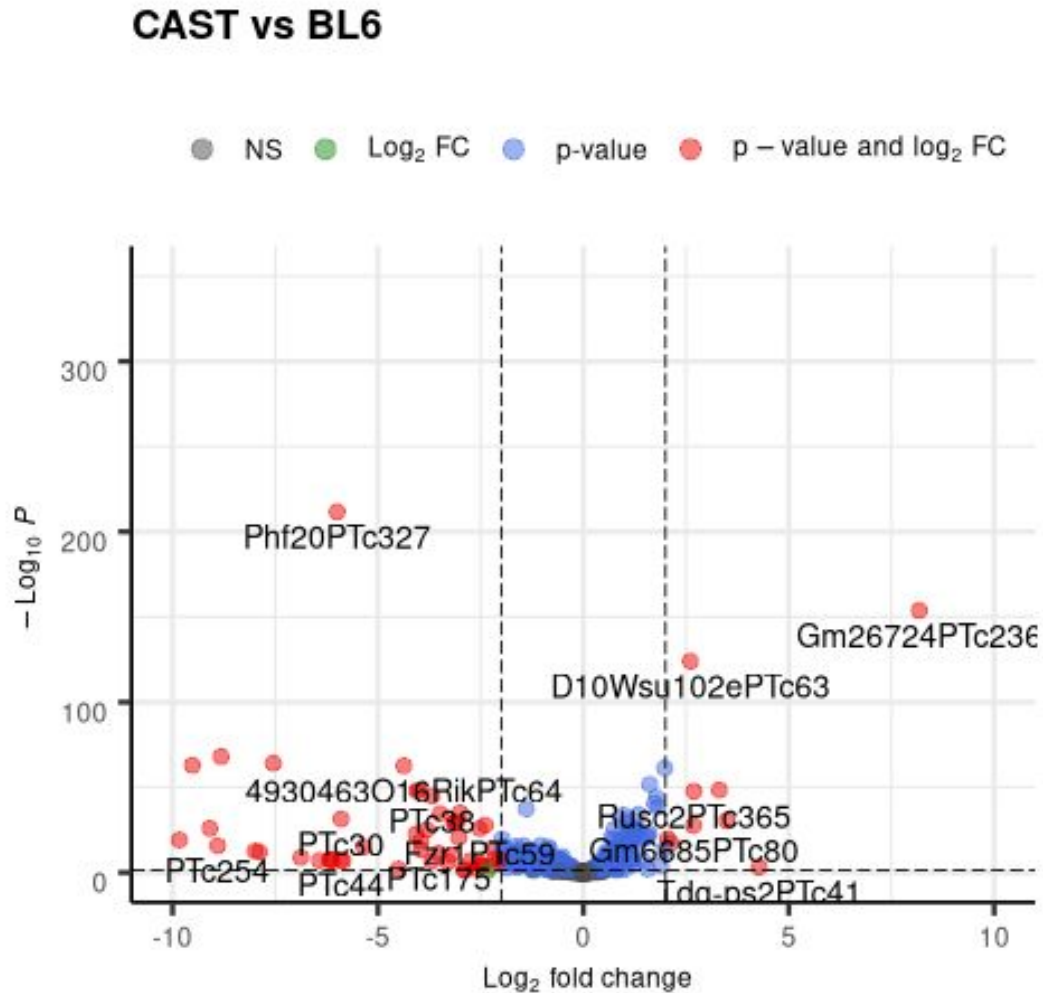
Project execution: Differential Expression Analysis (2)

Inbred testis samples by strain



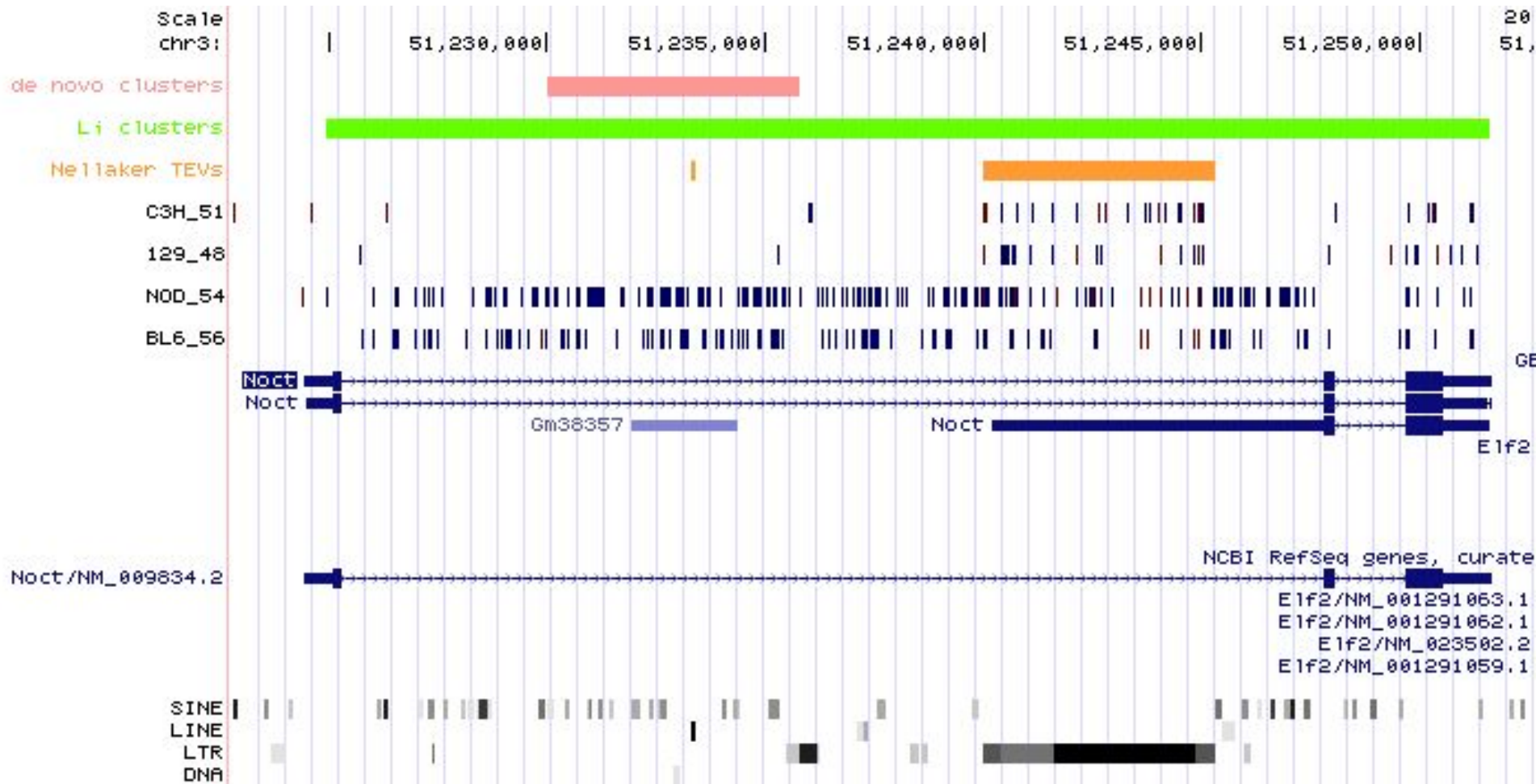
Project execution: Differential Expression Analysis (2)

Inbred testis samples by strain



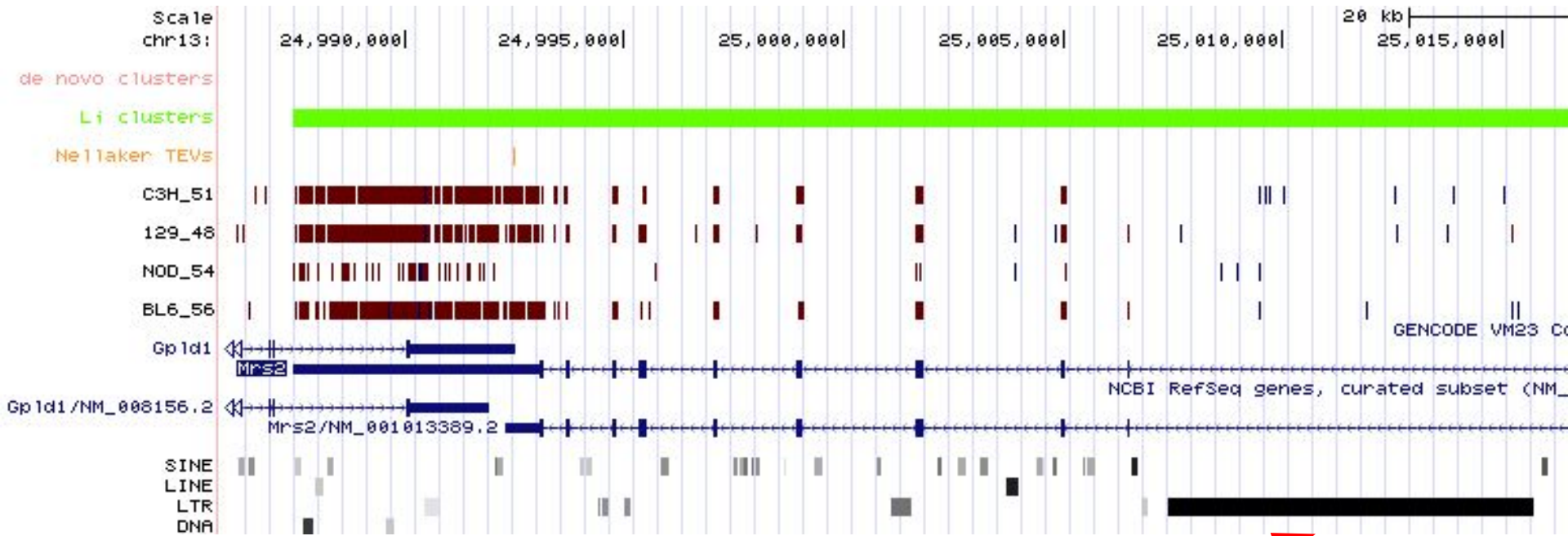
Project execution: Variable presence of IAPs

Noct has an intronic IAP in NOD and BL6



Project execution: Variable presence of IAPs

Mrs2: similar situation as *Noct*. NOD/ShiLtJ presents a large deletion of an IAP annotated by The Mouse Genome Project.

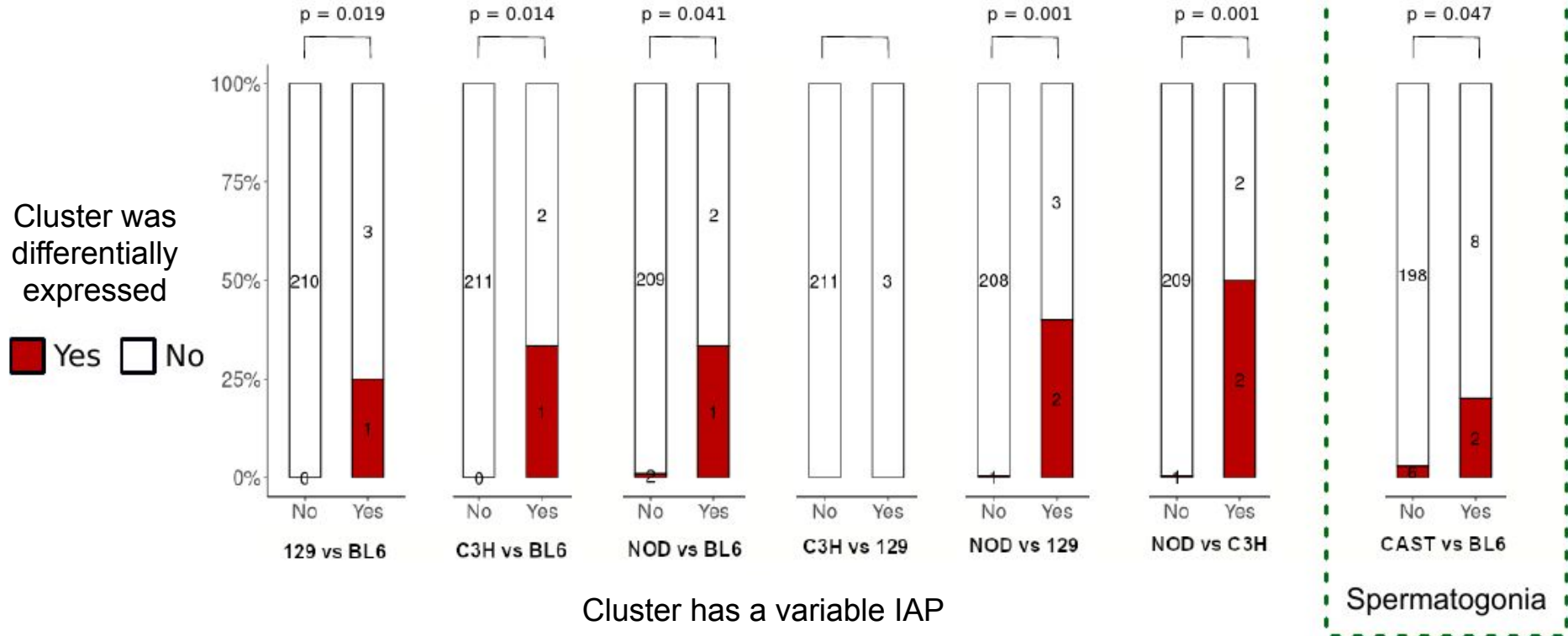


Results:

Fisher exact tests for Differentially expressed clusters and Variable IAP between the strains.

We can reject the null hypothesis: "there is no relationship between the two categorical variables"

Zamore Laboratory clusters N = (214)

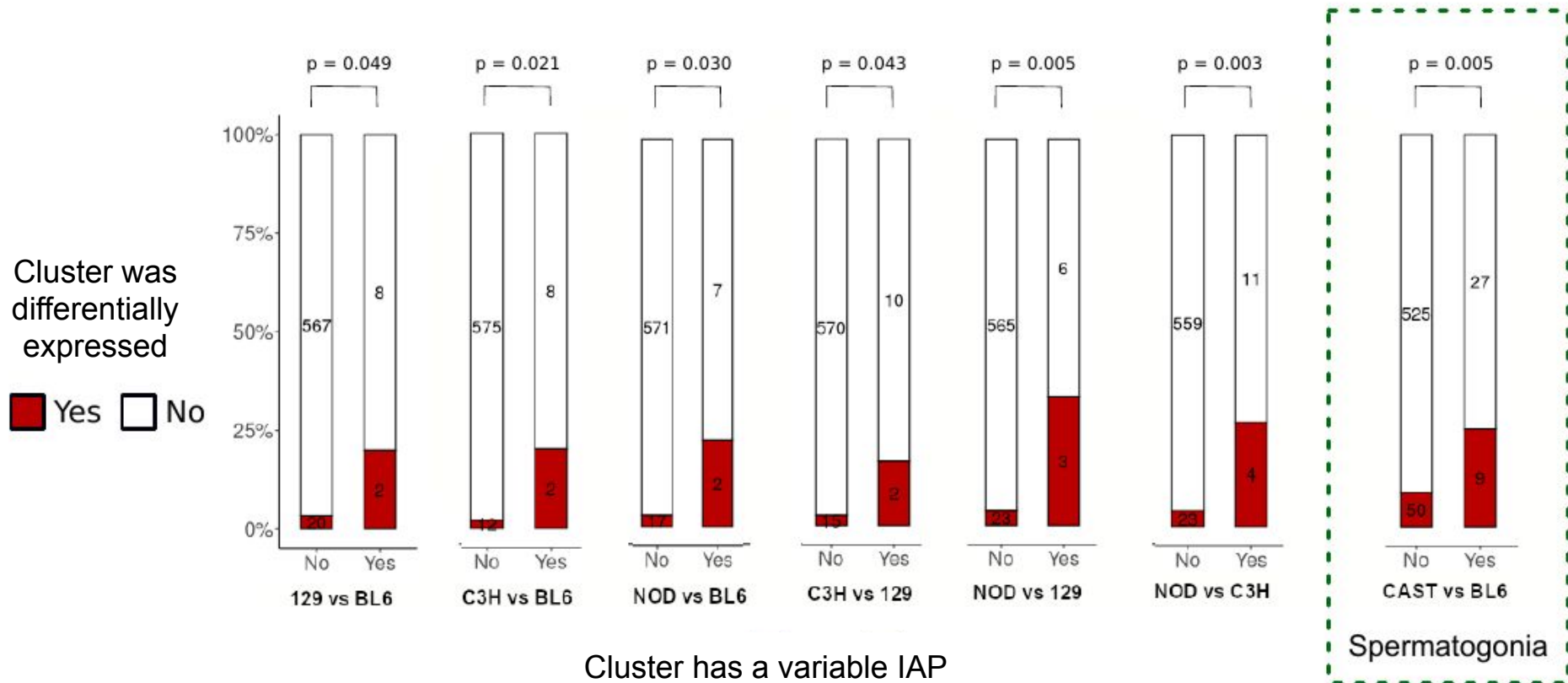


Results:

Fisher exact tests for Differentially expressed clusters and Variable IAP between the strains.

We can reject the null hypothesis: "there is no relationship between the two categorical variables"

De novo predicted clusters N = (611)



Results: Summary

Main Takeaways:

The small RNA-seq data was of high quality and enriched for piRNAs.

There are differences in the expression of piRNA clusters that can be attributed to their genetic background.

There is a significant correlation between piRNA differential expression and the different status of a murine intracisternal A-particle (IAP) in all the contrasts performed. ($p < 0.05$)

At least part of the sequence of IAPs must somehow interact with the mechanism of piRNA biogenesis in the mouse male germline.

Open Questions:

Why the effect is not always of the same type (enhancing or reducing the expression)

Future lines of work:

Identify if some characteristics of IAPs can be associated to this effect: presence of certain viral genes (*gag*, *pro*, *pol*, *env*); relative position on the strain compared to the transcript.

Compare alternative piRNA prediction methods and try to improve them.

THANKS!

