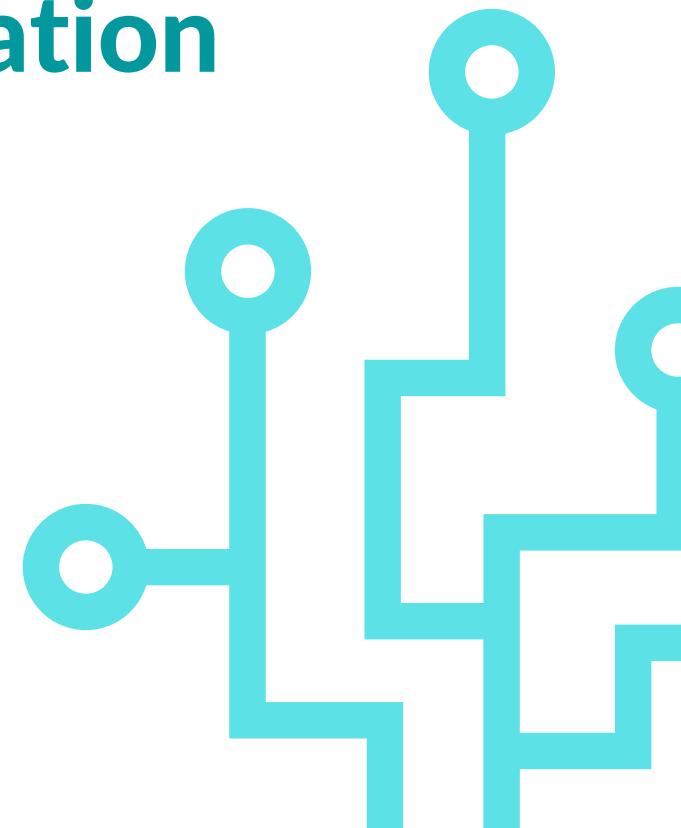


Transcriptomic profiling to study the role of MYBA1 and MYB24 in the *Vitis vinifera* variegation

Arnau Peris Cuesta



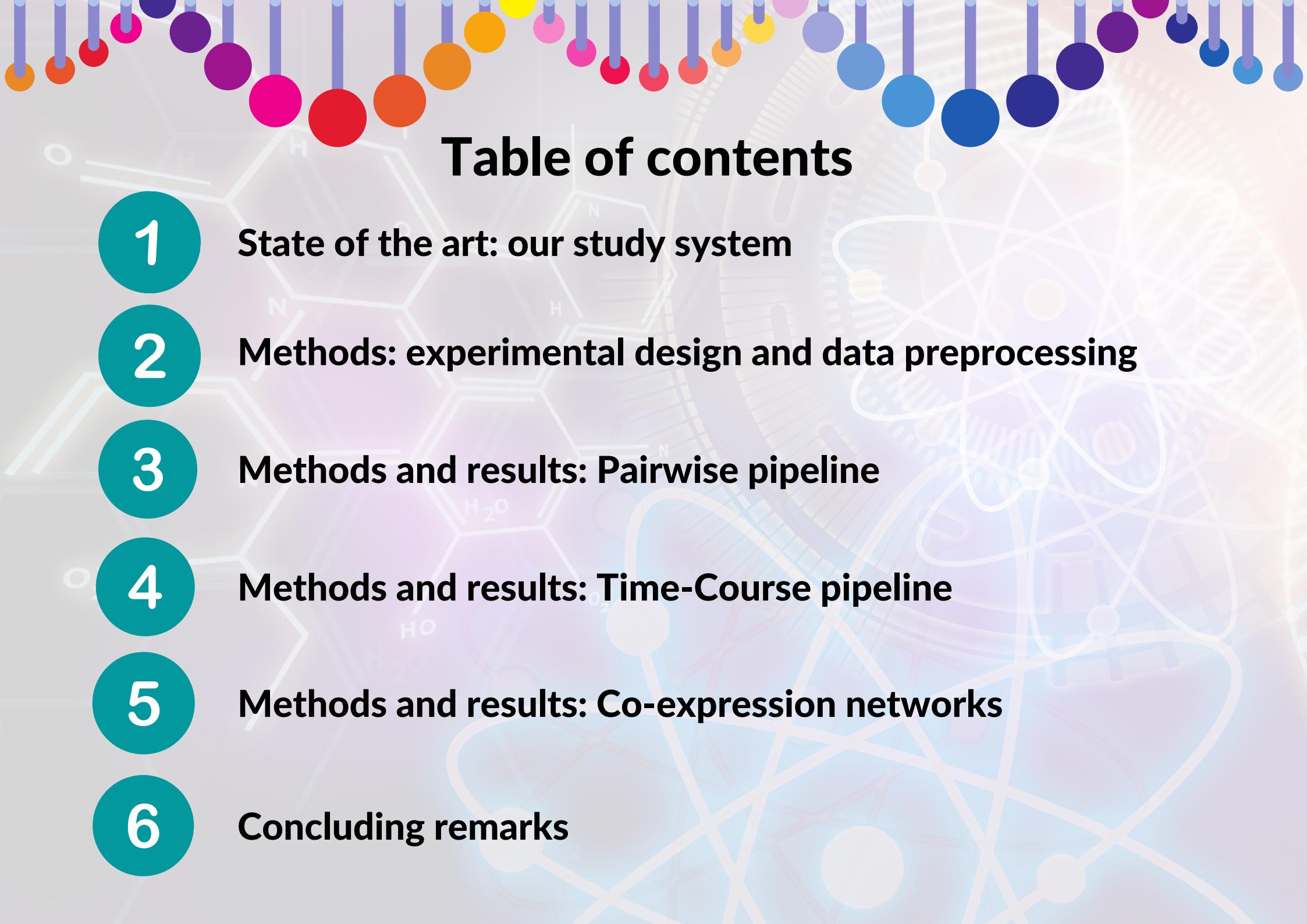
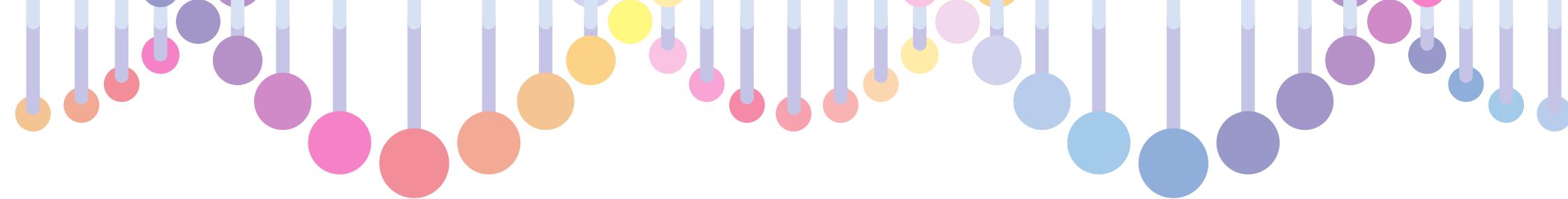


Table of contents

- 1 State of the art: our study system**
- 2 Methods: experimental design and data preprocessing**
- 3 Methods and results: Pairwise pipeline**
- 4 Methods and results: Time-Course pipeline**
- 5 Methods and results: Co-expression networks**
- 6 Concluding remarks**



State of the art: the Bequignol variegation phenotype

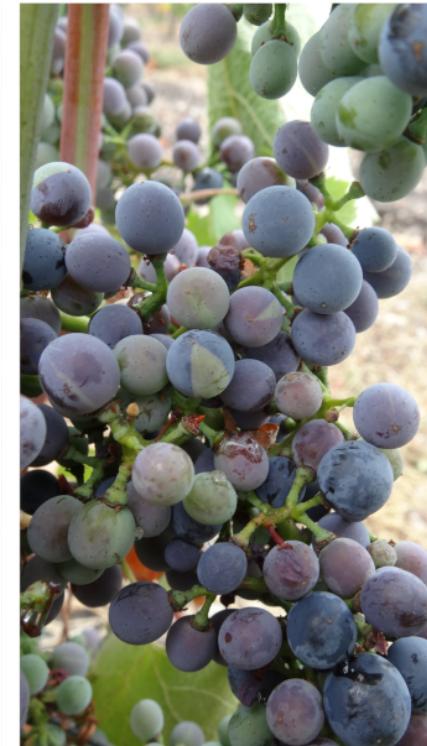
Pre-veraison (PV)



Veraison (0 WAV)



Late-veraison (LV)



8 WAV



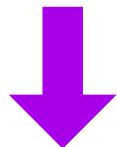
WAV: Weeks after veraison

State of the art: biosynthesis of anthocyanins

MYBA1



Anthocyanins



Bequignol
Noir



~~MYBA1~~



Anthocyanins



Bequignol
Blanc



Differential expression analysis of variegated berries throughout development: experimental design

4 Time points

3 Biological replicates

Red and White skin



24 samples in total



Paired-End RNAseq



24 libraries

60M reads / library

Harvest

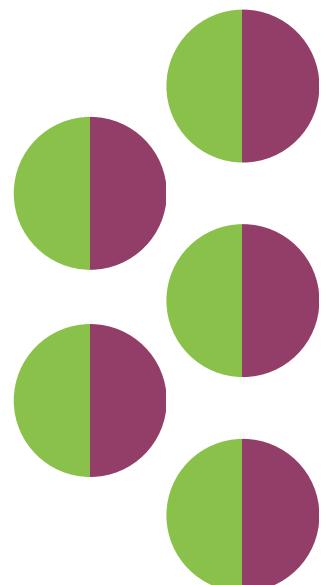
9WAV

Veraison

0WAV

3WAV

6WAV



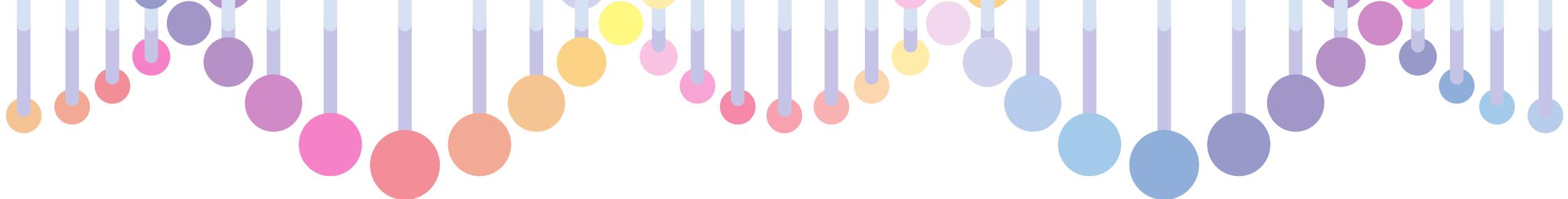
White skin
sections

Red skin
sections



2x150bp sequenced with a Stranded mRNA-Seq Illumina TrueSeq

V: Veraison (onset of the ripening)



cv. 'Bequignol'



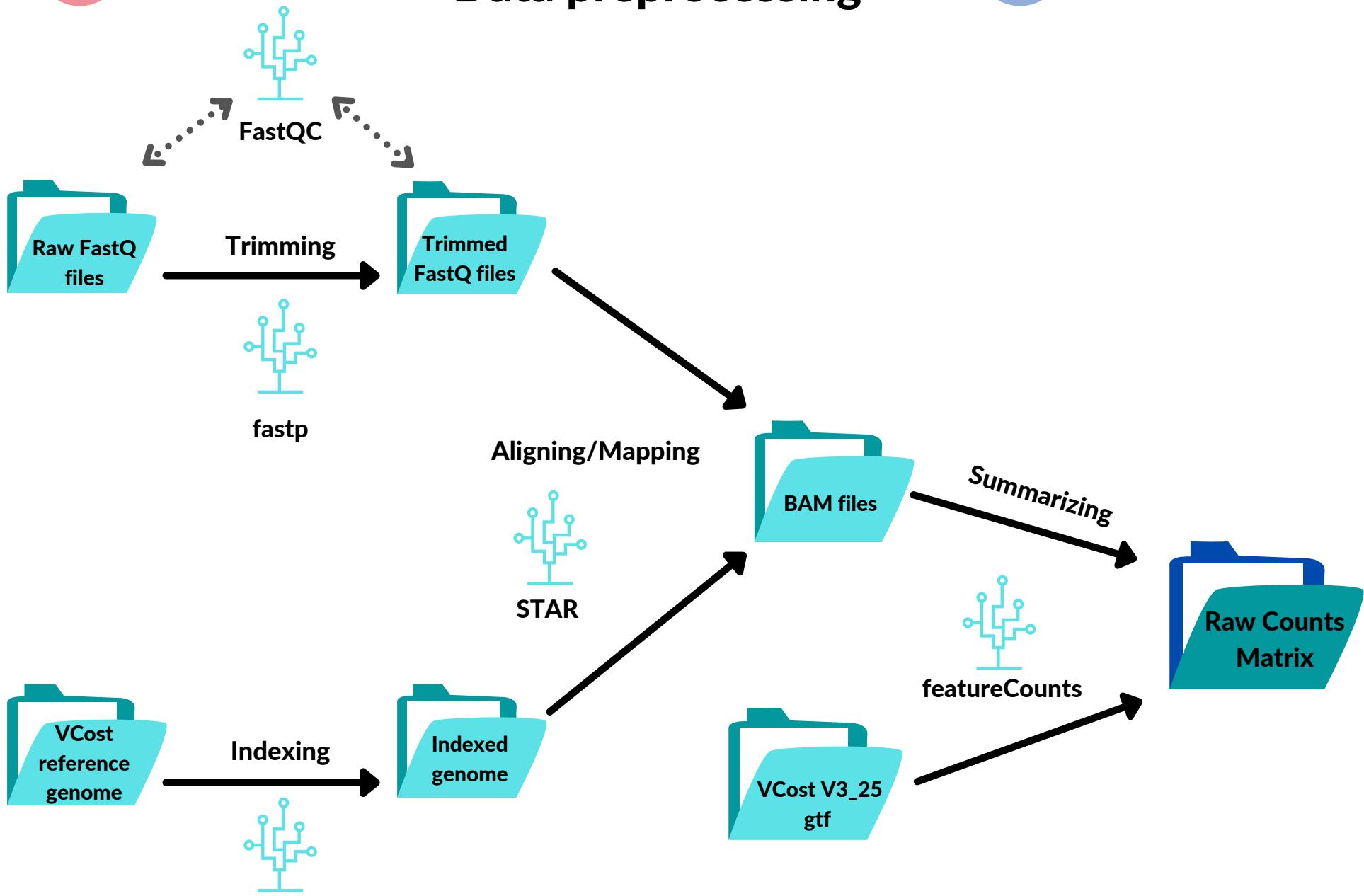
MR1_1	MR1_2	MR1_3	BMB1_1	BMB1_2	BMB1_3
MR2_1	MR2_2	MR2_3	BMB2_1	BMB2_2	BMB2_3
MR3_1	MR3_2	MR3_3	BMB3_1	BMB3_2	BMB3_3
MR4_1	MR4_2	MR4_3	BMB4_1	BMB4_2	BMB4_3
Time point					
1: 0WAV					
2: 3WAV					
3: 6WAV					
4: 9WAV					

Time point
1: 0WAV
2: 3WAV
3: 6WAV
4: 9WAV

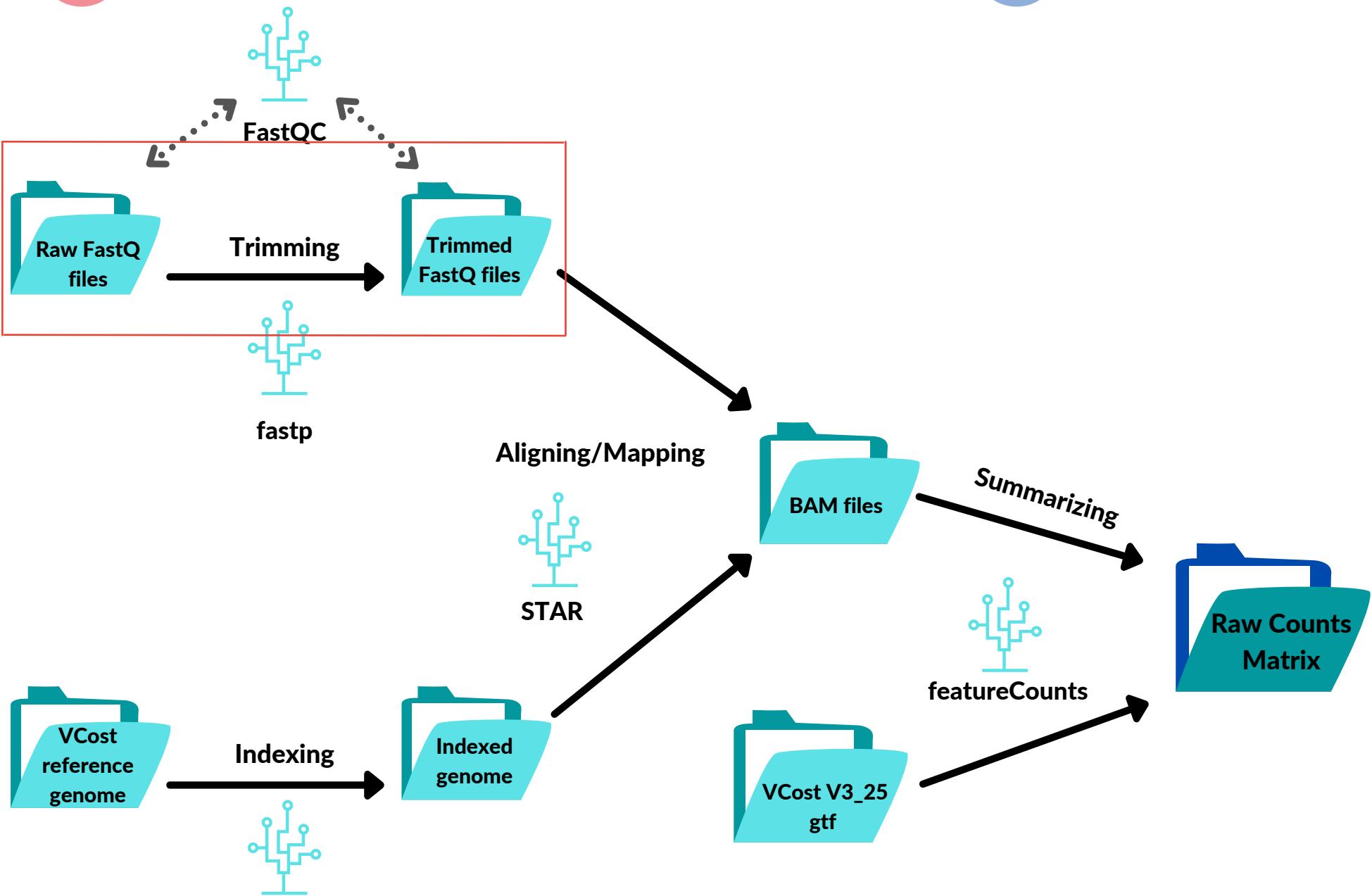
n° replicate

MR: red skin sections
BMB: white skin sections

Data preprocessing

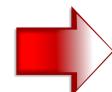
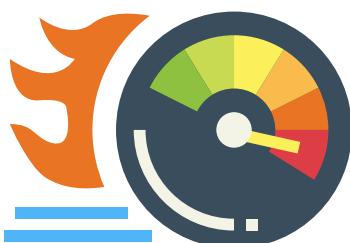


Trimming



Trimming – Why fastp?

2-5 times faster



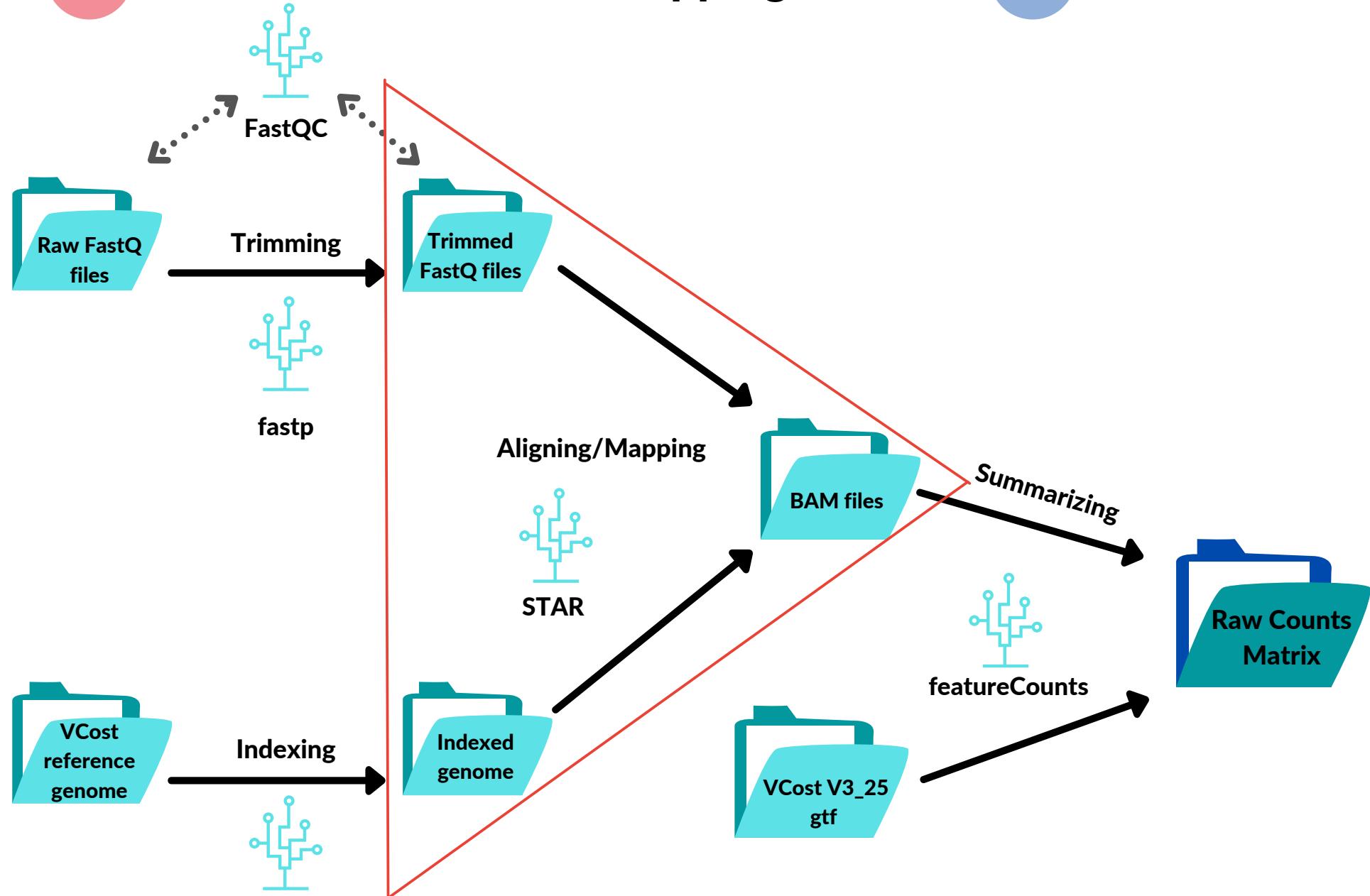
Automatic Illumina adapters trimming

Tool	Time(min)	Throughput (reads/s)
<i>Fastp</i>	13.3	116.750
<i>FASTQC</i>	25.8	60.185
<i>Cutadapt</i>	24.6	63.120
<i>Trimmomatic</i>	60.9	25.497

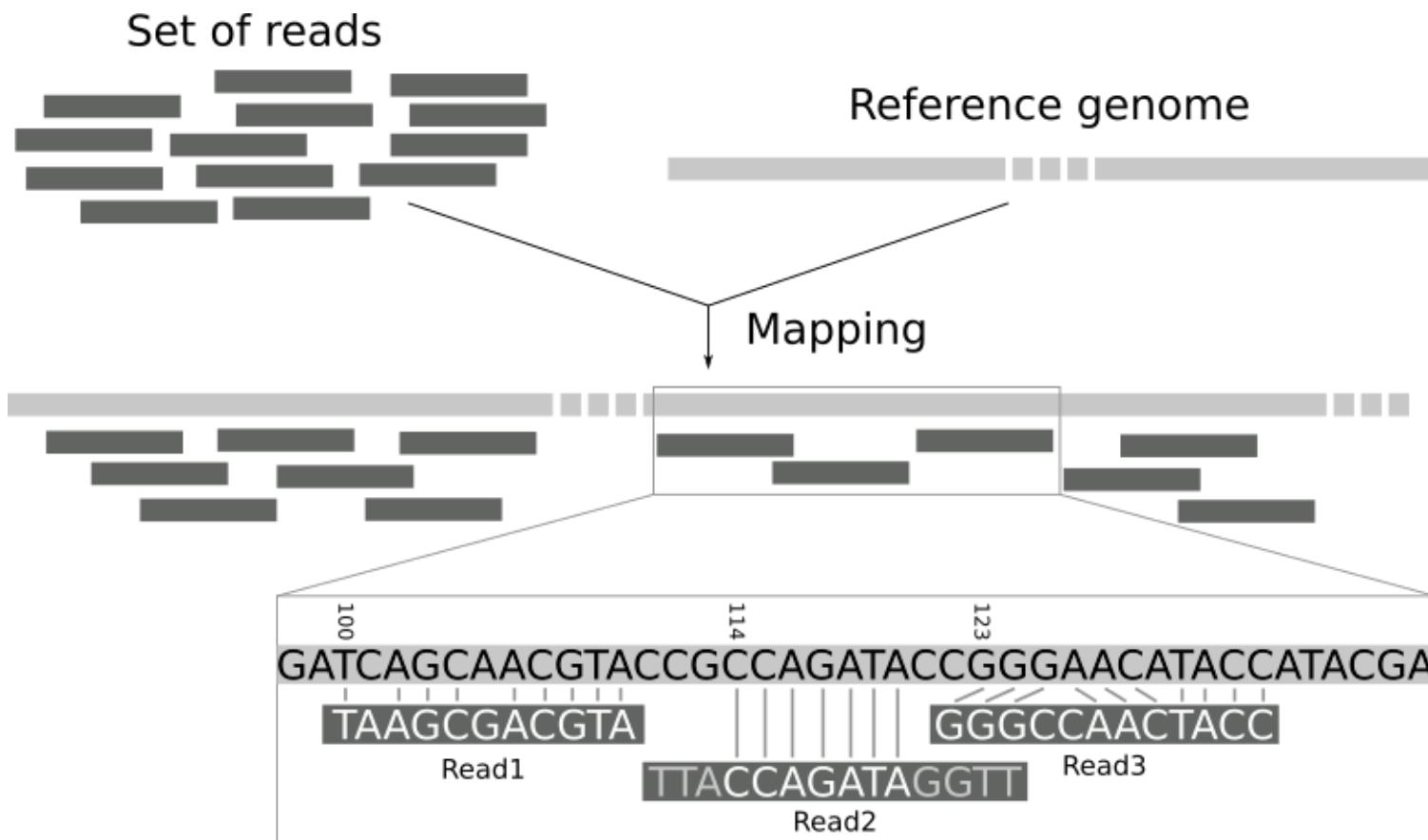
S. Chen *et al.* (2018) "fastp: an ultra-fast all-in-one FASTQ processor". Bioinformatics 34(17): i884–i890.

<https://bit.ly/3Fi3BBr>

Mapping

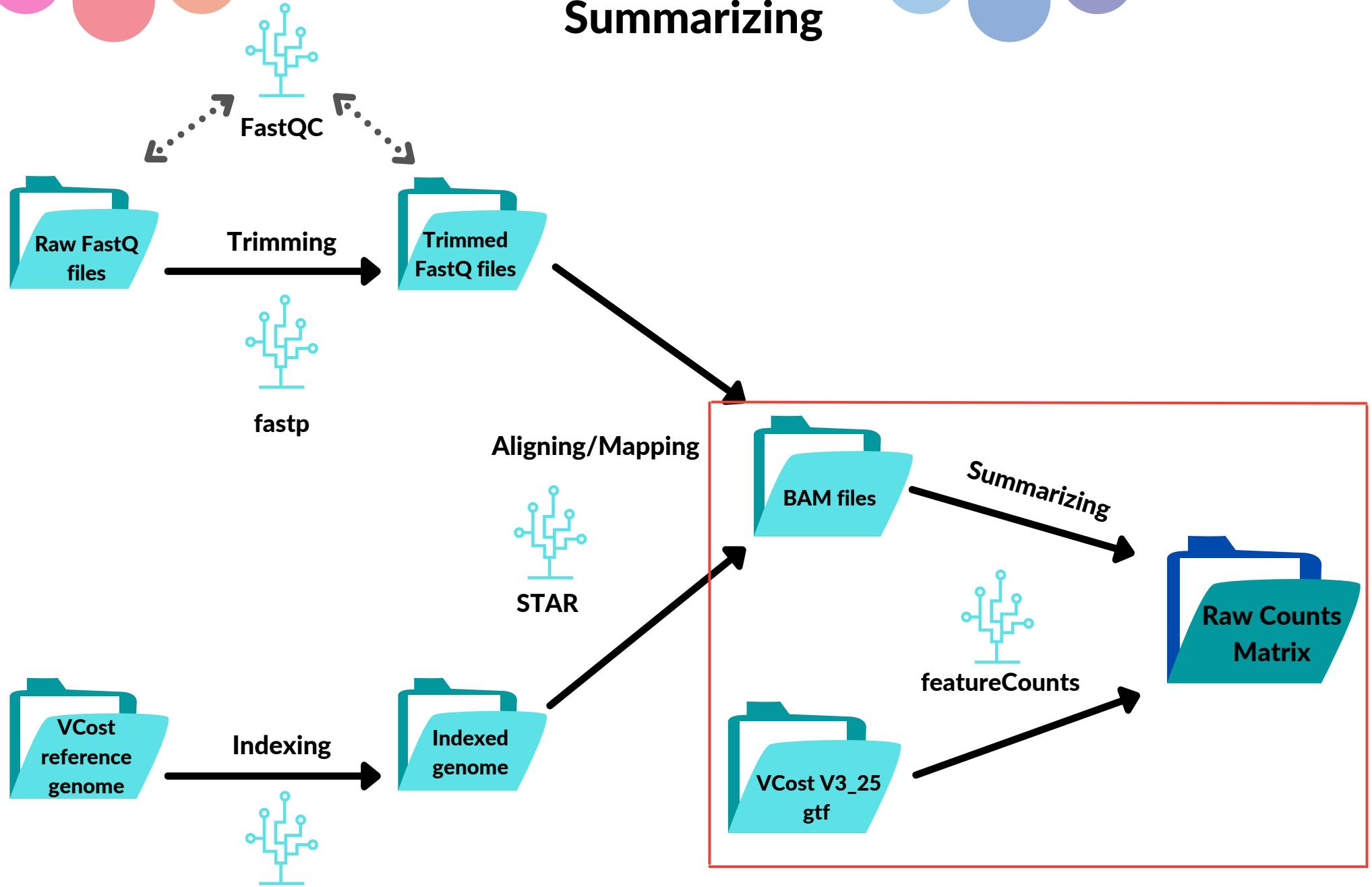


Mapping - Why STAR?



**Trade-off speed
and completeness:
accuracy and
mapping rate.**

Summarizing



Summarizing – featureCounts

Annotation (GTF file)



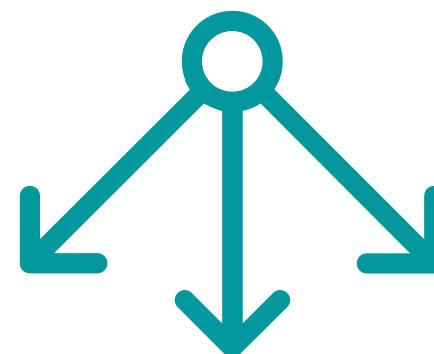
Sorted BAM files



Count matrix

featureCounts

Assigning mapped reads
to genomic features



cds

gene

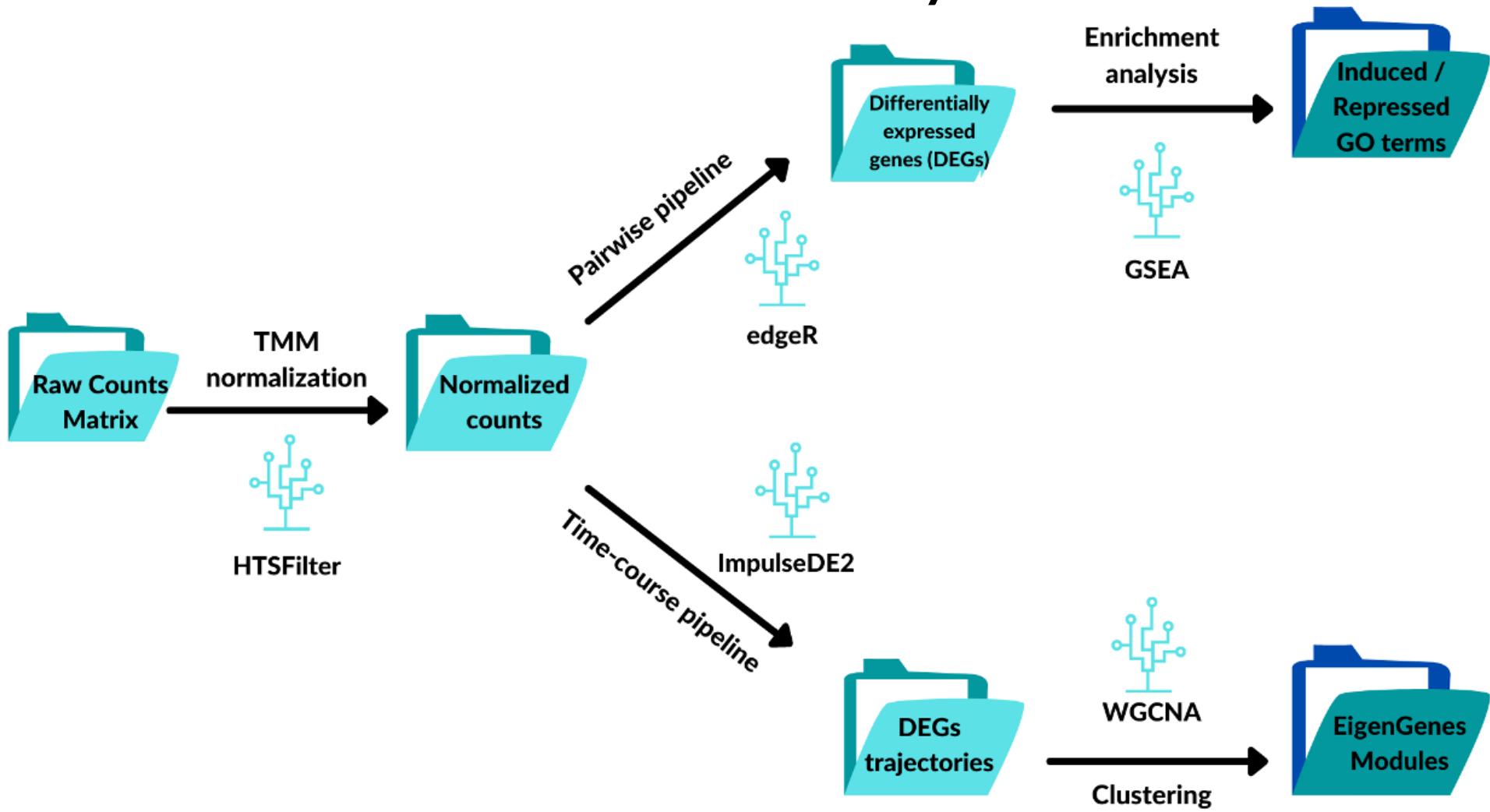
...

exon

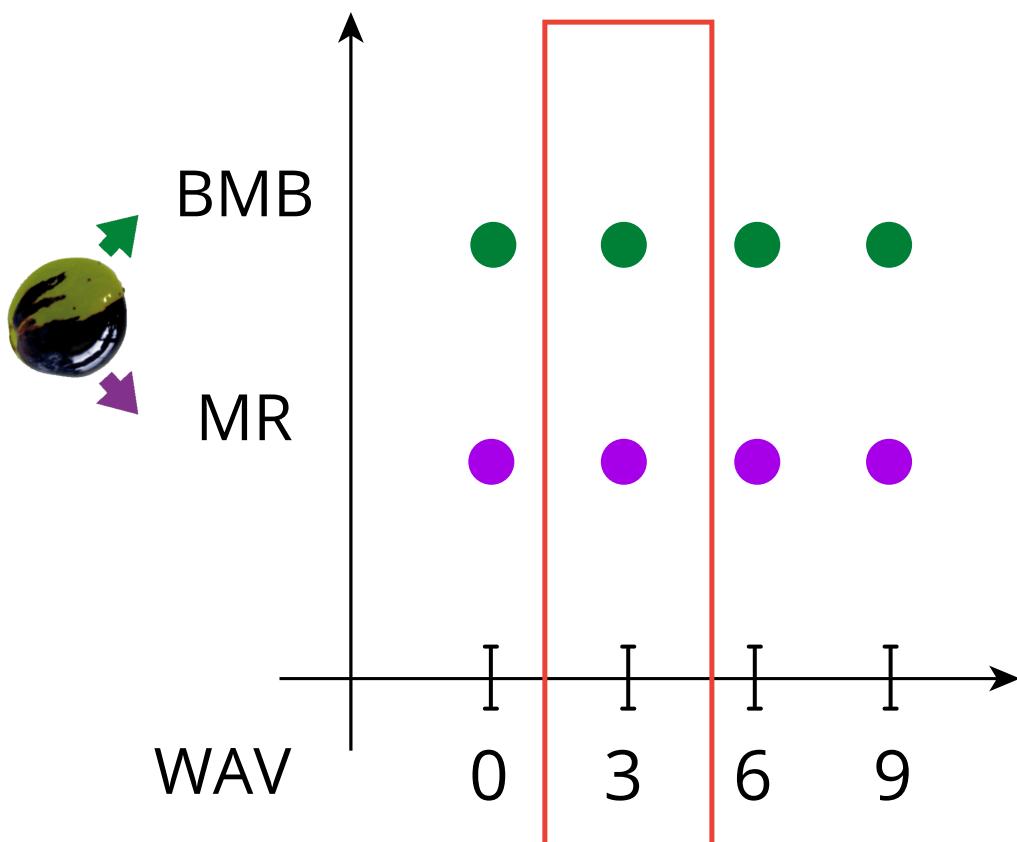
V.vinifera PN40024 genome assembly

VCost.v3 annotation (33.568 genes)

Downstream analysis

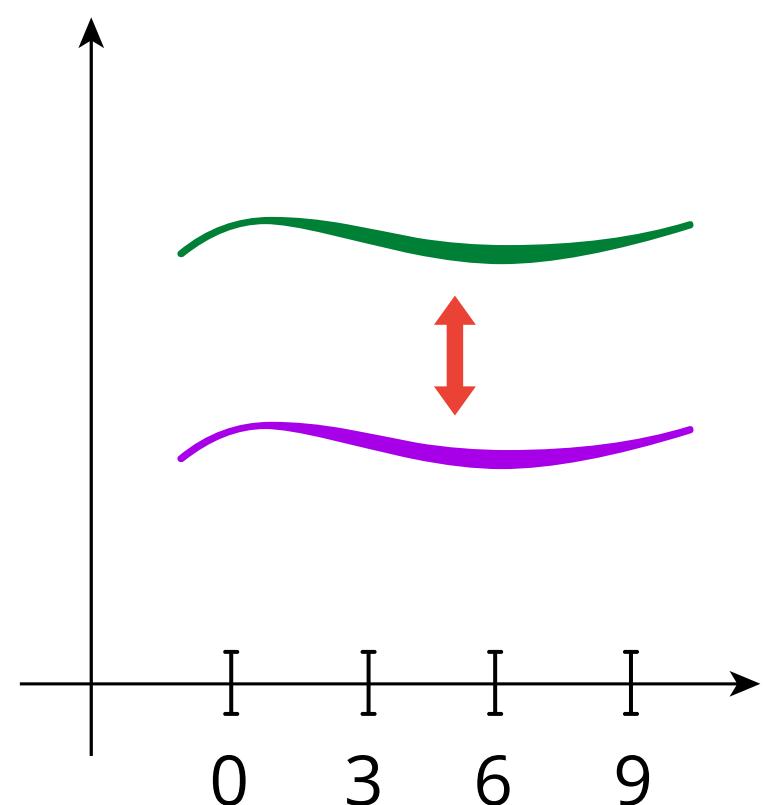


Stage-specific analysis



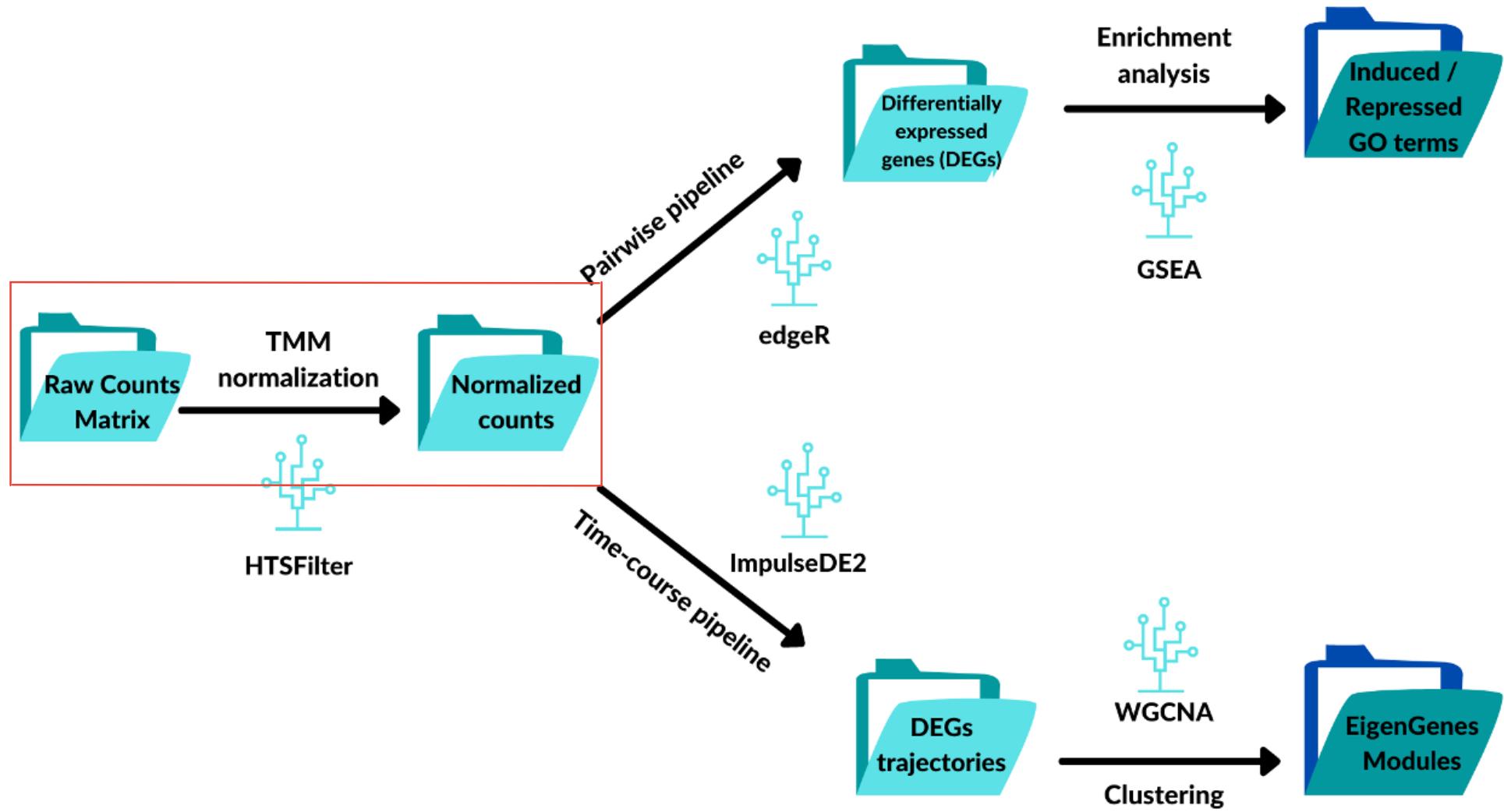
Pairwise comparison

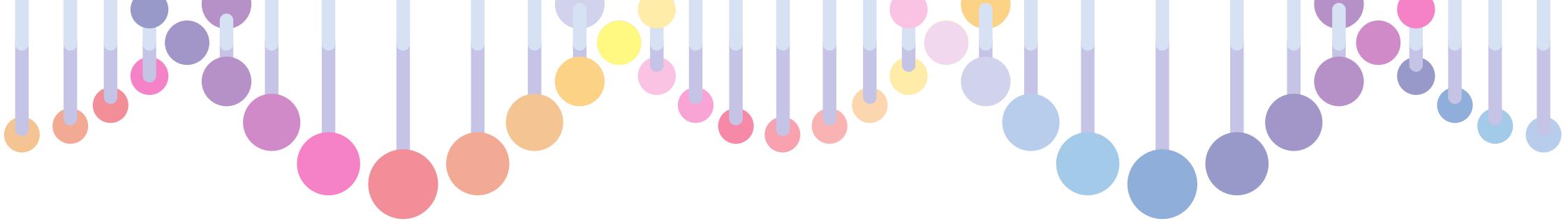
Trajectory analysis



Time-course comparison

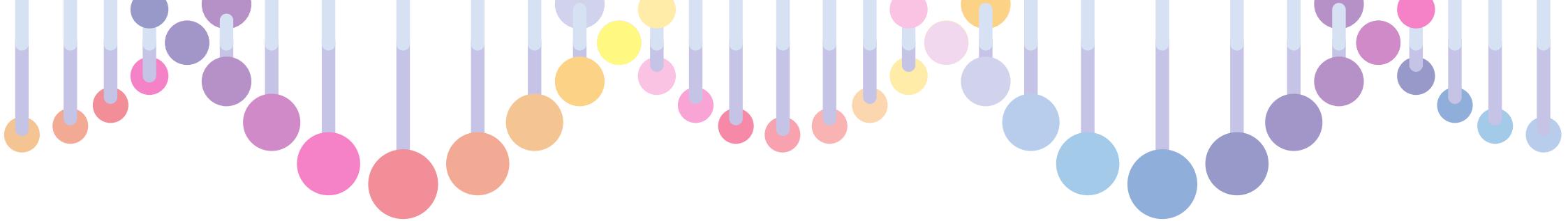
Normalization



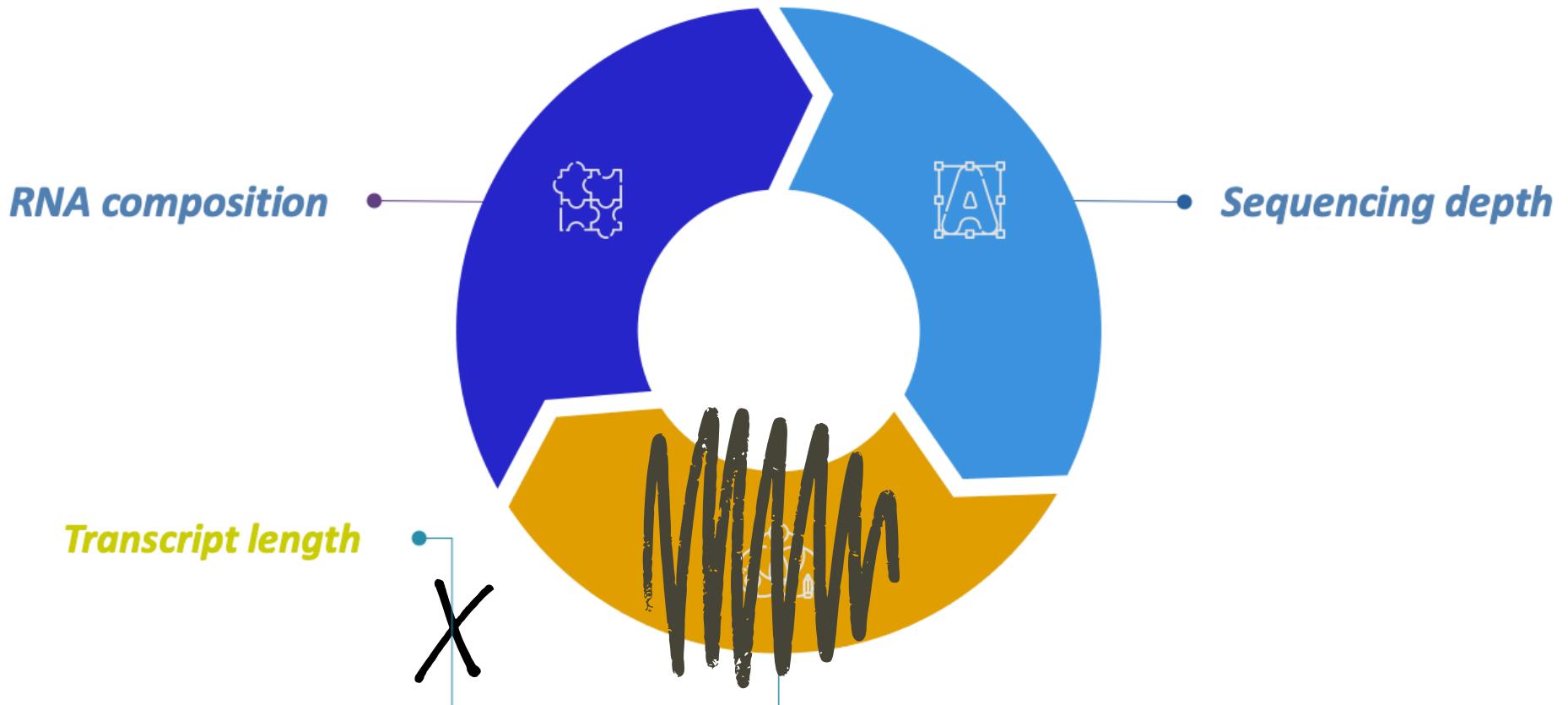


Normalization – Why TMM?

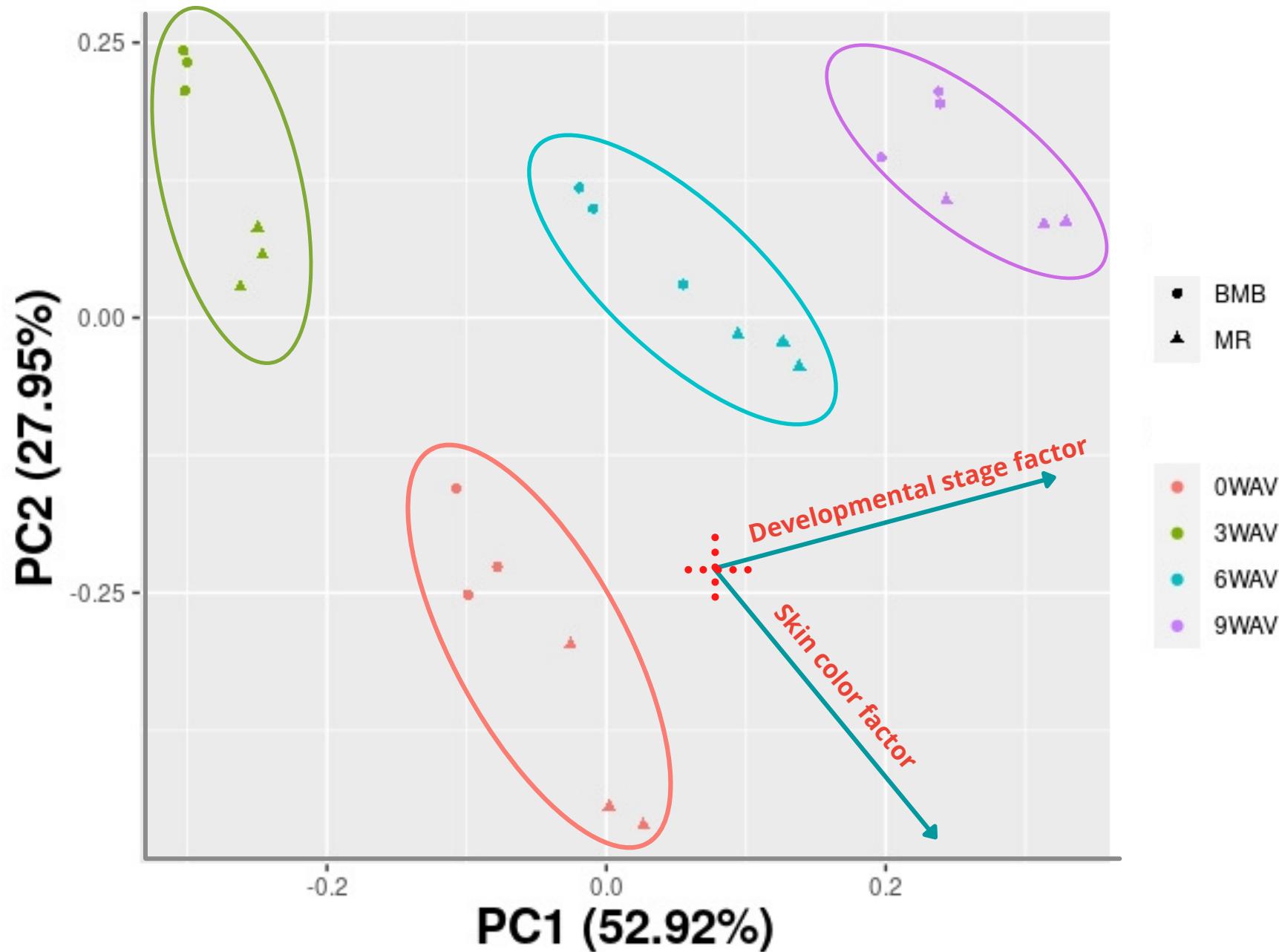




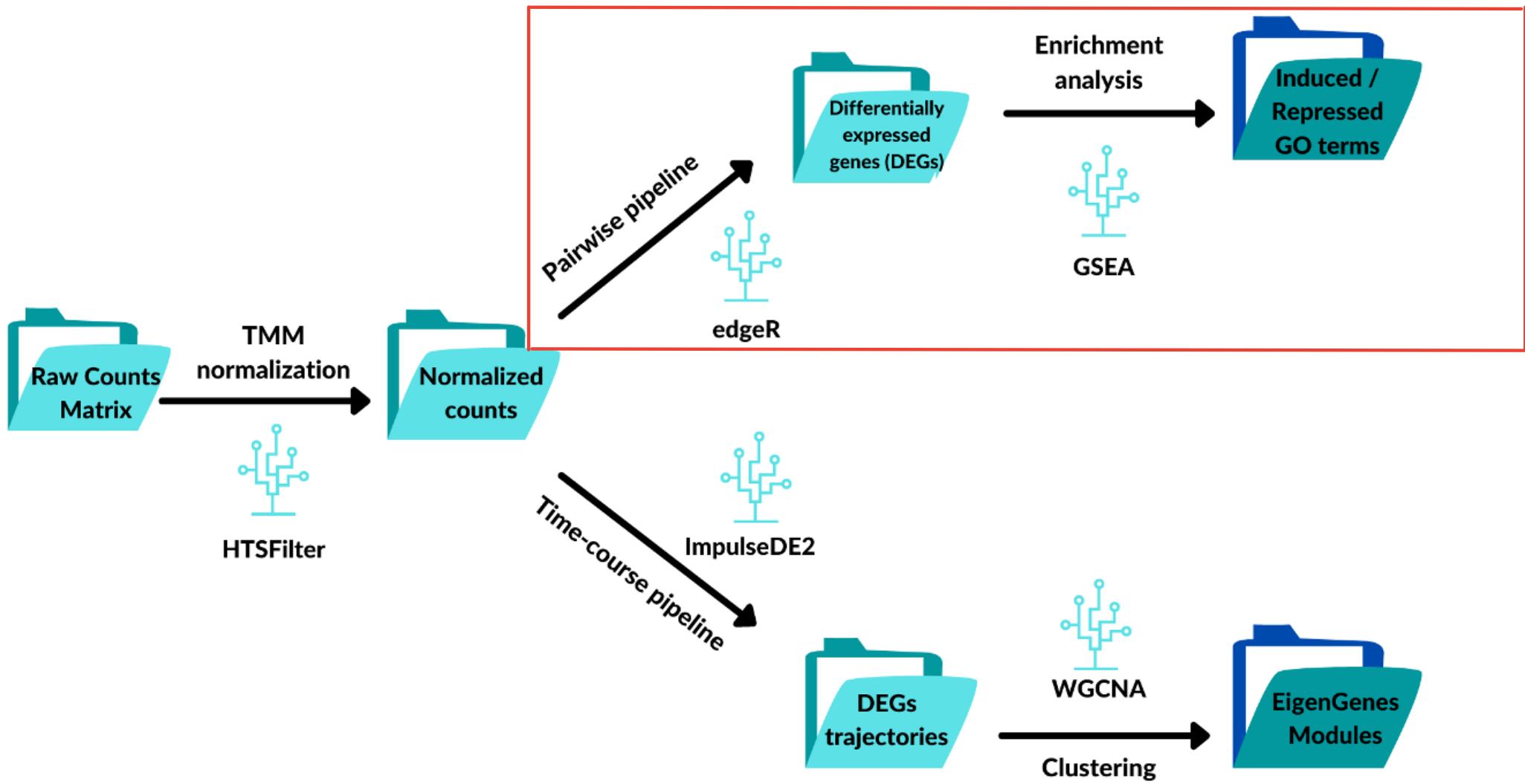
Normalization – Why TMM?



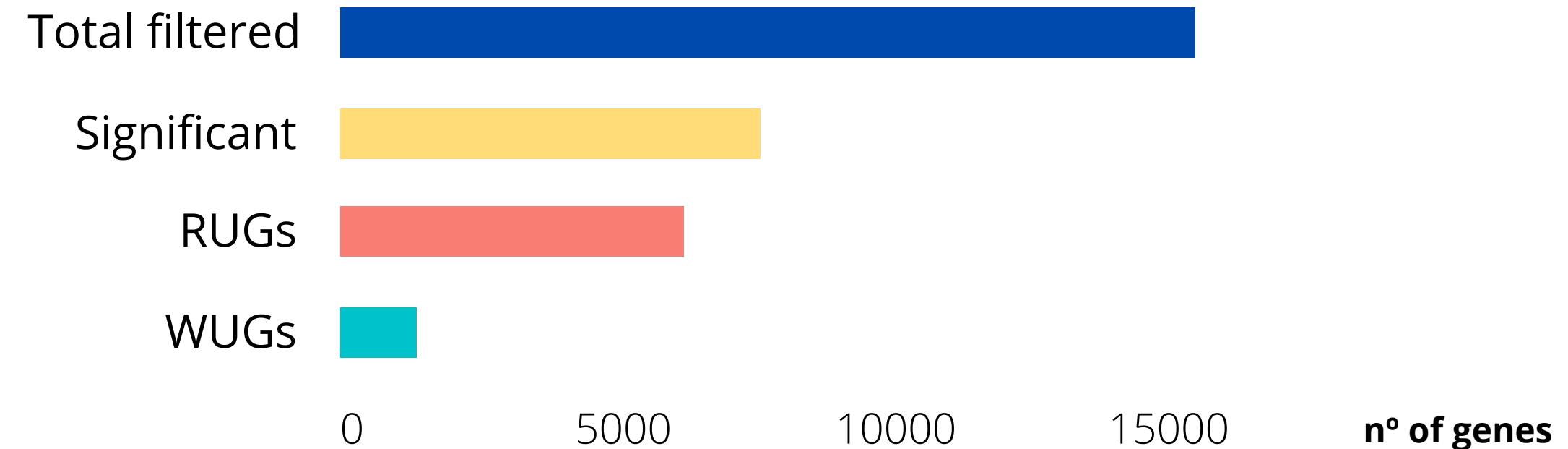
Results - Principal Component Analysis



Pairwise analysis



Results - Summary at 3WAV

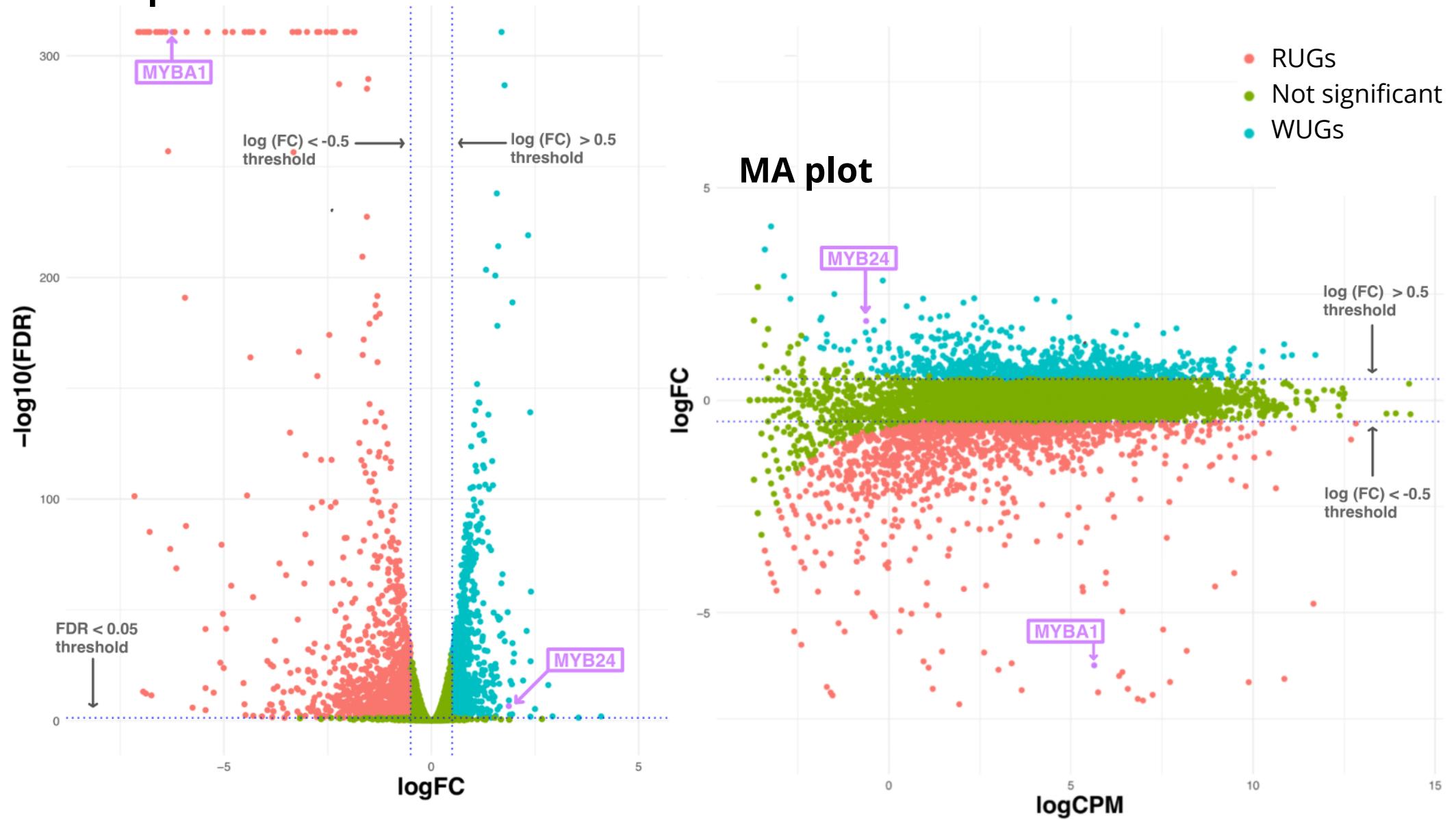


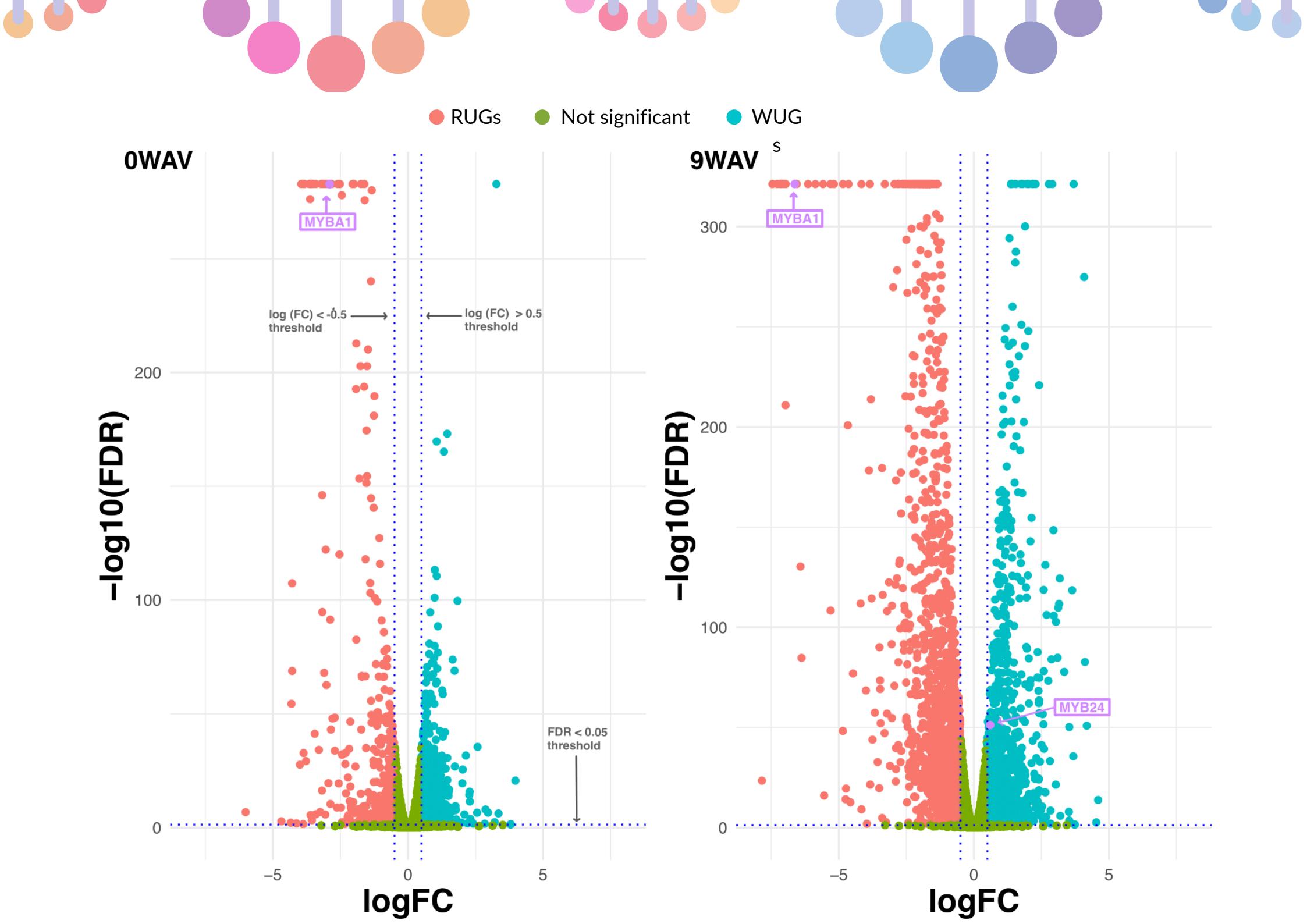
RUGs: Red skin Up-regulated Genes

WUGs: White skin Up-regulated Genes

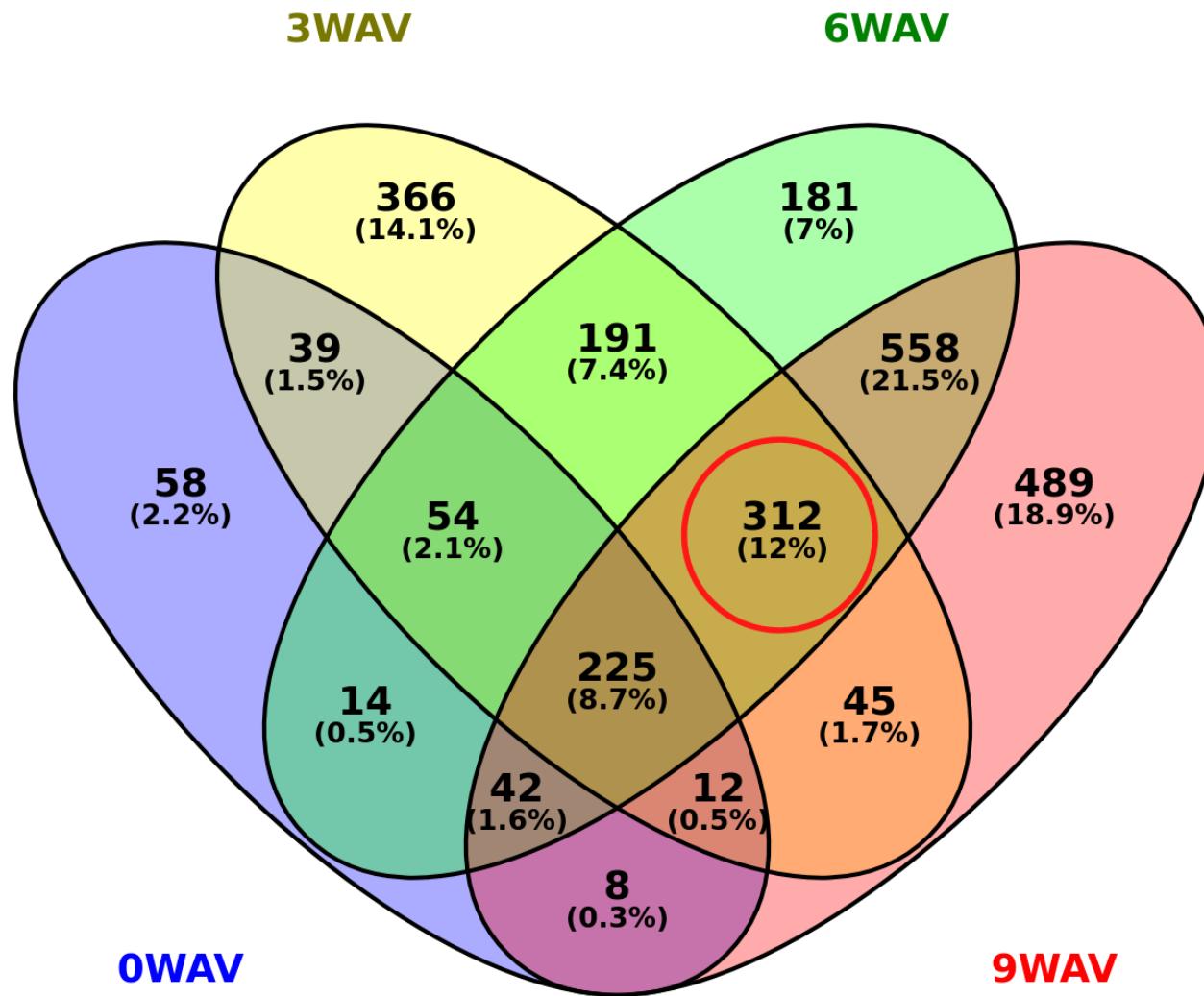
Results - WUGs and RUGs at 3WAV

Volcano plot





Common RUGs at latest stages of ripening



-> 0WAV excluded

-> logFC < -0.5

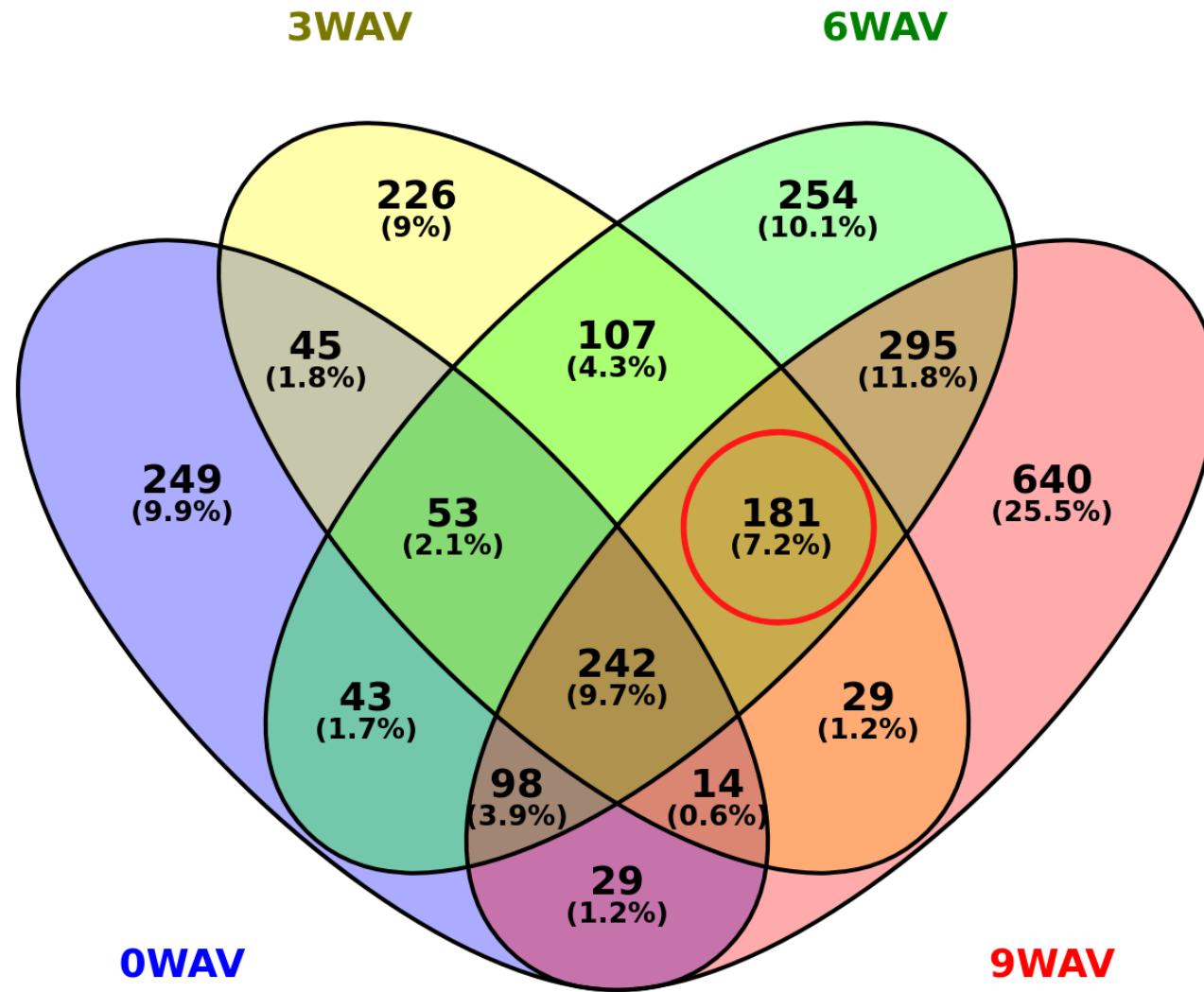
Enrichment analysis - RUGs

Term ID	Term label	P-value
9.2	Secondary metabolism phenolics	5.99E-05
11	Phytohormone action	0.0003
9.2.1	Secondary metabolism phenolics p-coumaroyl-CoA biosynthesis	0.0010
15.5	RNA biosynthesis transcriptional regulation	0.0011
3.11.1.2	Carbohydrate metabolism fermentation alcoholic fermentation alcohol dehydrogenase	0.0012
9.2.2.1.1	Secondary metabolism phenolics flavonoid biosynthesis chalcones chalcone synthase activity	0.0016
11.7	Phytohormone action jasmonic acid	0.0016
9.2.2.1.1.1	Secondary metabolism phenolics flavonoid biosynthesis chalcones chalcone synthase activity CHS	0.0016
9.2.2.1	Secondary metabolism phenolics flavonoid biosynthesis chalcones	0.0016
3.11.1	Carbohydrate metabolism fermentation alcoholic fermentation	0.0024
9.2.1.1	Secondary metabolism phenolics p-coumaroyl-CoA biosynthesis phenylalanine ammonia lyase activity	0.0029

->MapMan

->VCost.v3
annotation

Common WUGs at latest stages of ripening



-> 0WAV excluded

-> logFC > 0.5

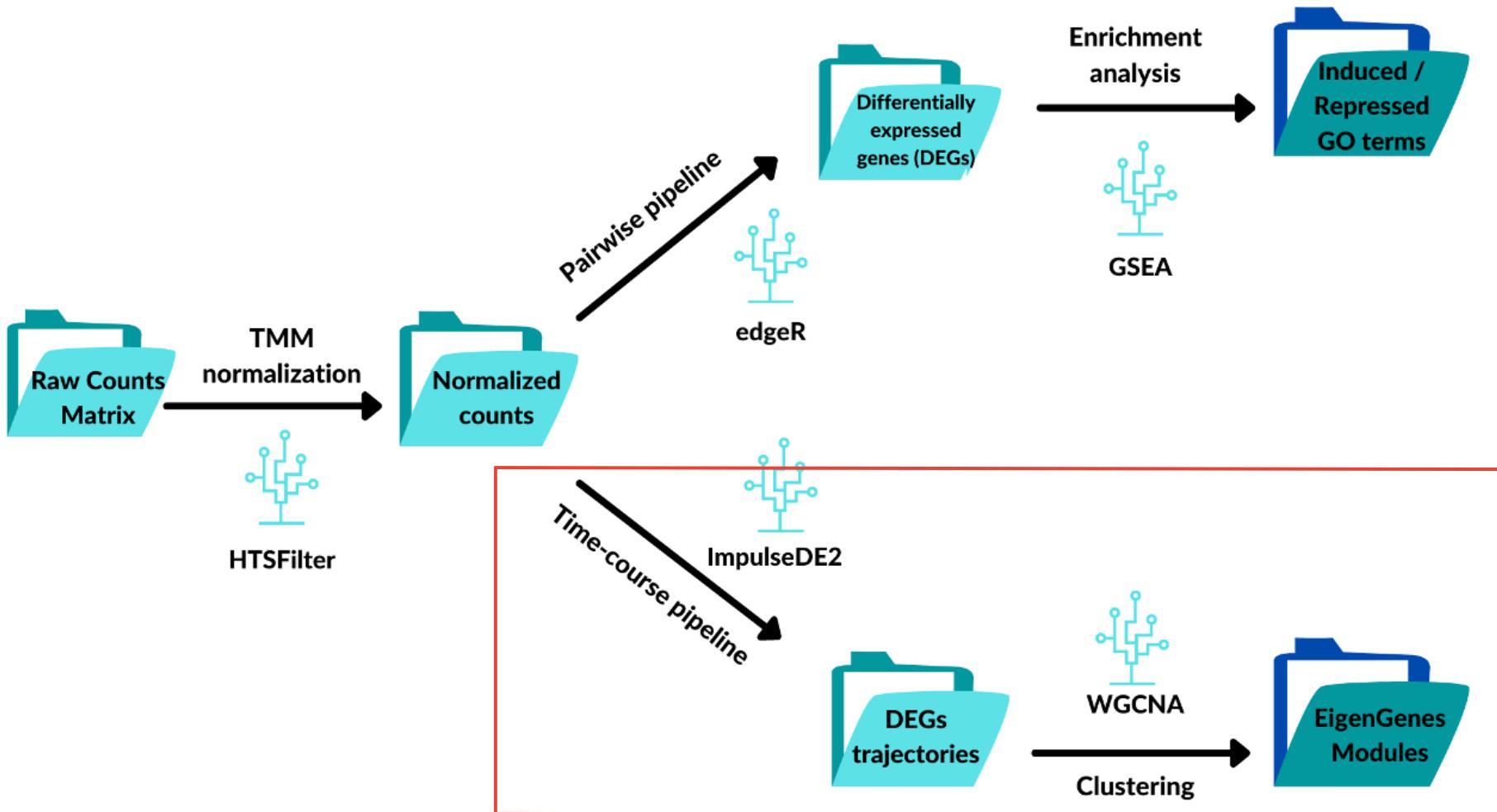
Enrichment analysis - WUGs

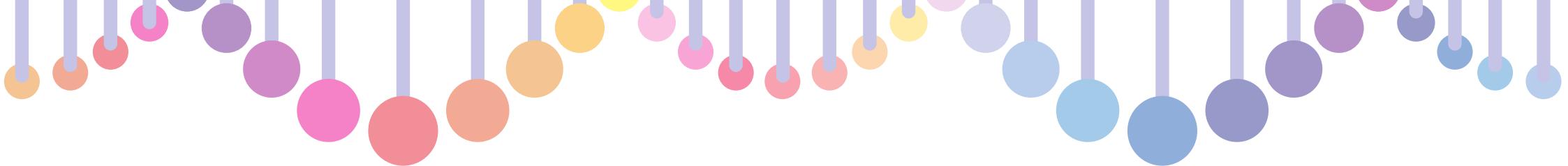
Term ID	Term label	P-value
1.1	Photosynthesis photophosphorylation	1.52E-10
1	Photosynthesis	5.41E-10
1.1.1	Photosynthesis photophosphorylation photosystem II	5.69E-05
17.7.2	Protein biosynthesis organelle machinery plastidial ribosome	5.69E-05
17.7	Protein biosynthesis organelle machinery	0.0005
1.1.1.3	Photosynthesis photophosphorylation photosystem II assembly and maintenance	0.0013
1.1.5.2.1	Photosynthesis photophosphorylation linear electron flow FNR activity ferredoxin-NADP oxidoreductase	0.0013
17.7.2.1	Protein biosynthesis organelle machinery plastidial ribosome large ribosomal subunit proteome	0.0013
19.4.5.8.2.1	Protein homeostasis proteolysis metallopeptidase activities FtsHE activity FtsH plastidial protease complexes	0.0027
1.1.1.3.7	Photosynthesis photophosphorylation photosystem II assembly and maintenance:HCF244-OHP assembly factor complex	0.0027
1.1.2.9.2	Photosynthesis photophosphorylation cytochrome b6/f complex assembly CCS maturation system II	0.0027
1.1.2	Photosynthesis photophosphorylation cytochrome b6/f complex	0.0032

->MapMan

->VCost.v3
annotation

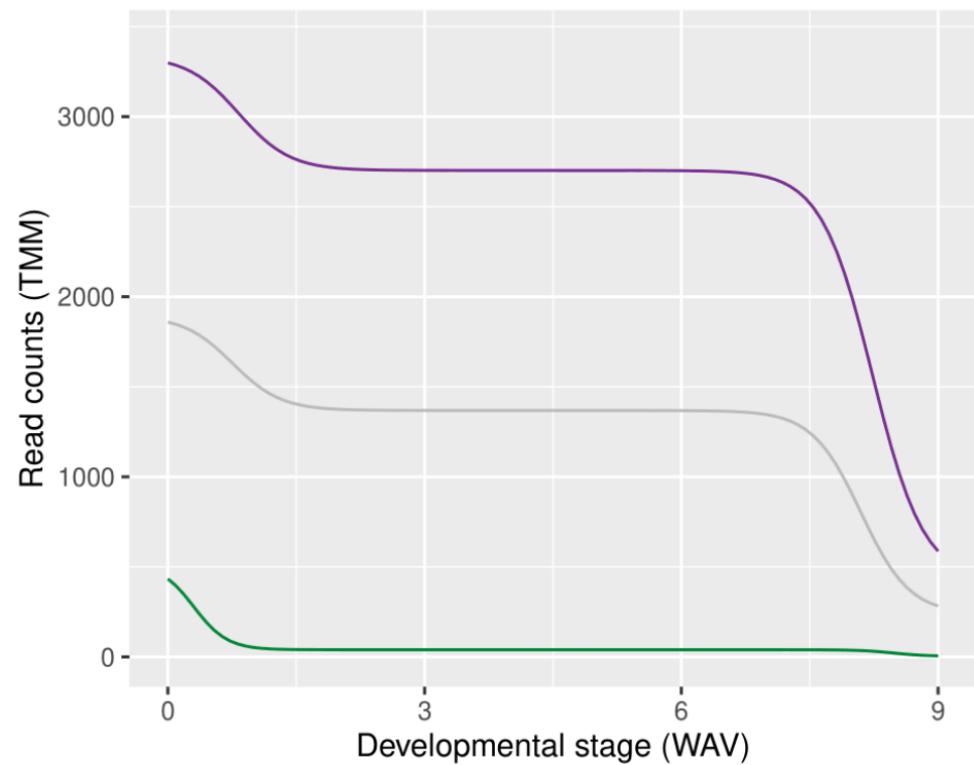
Time-course analysis



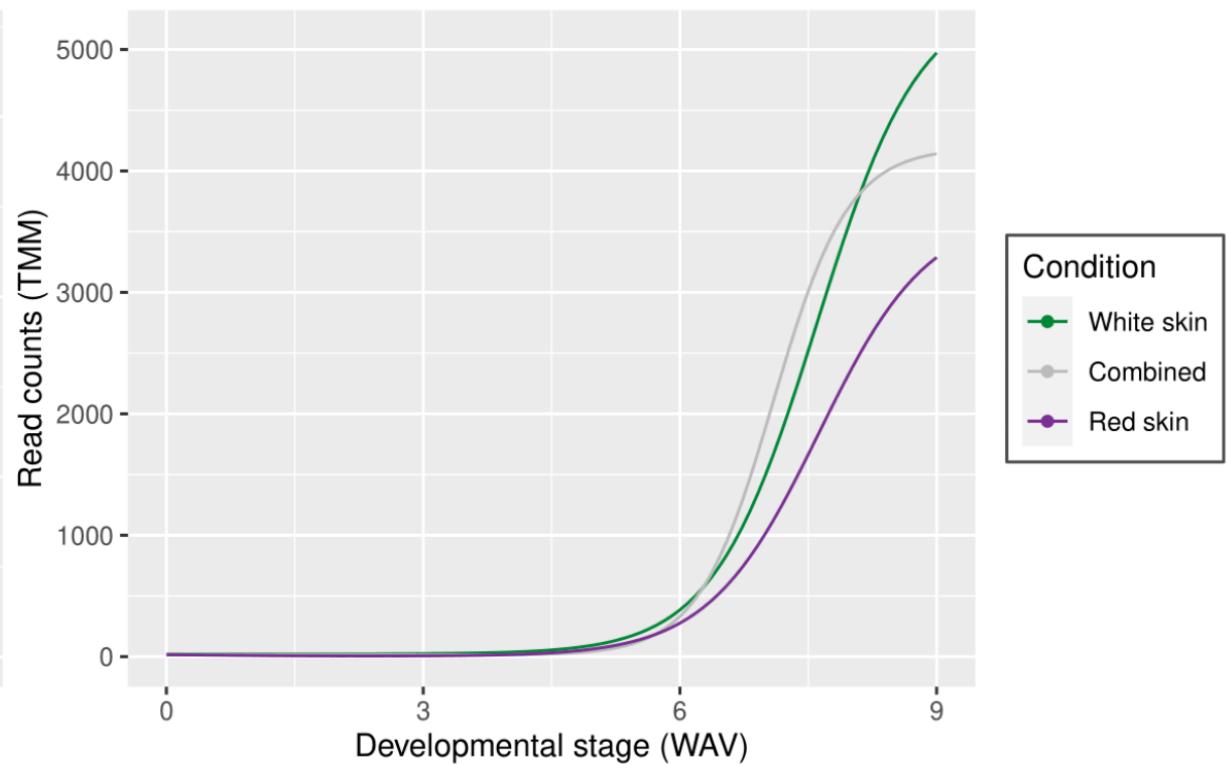


Time-Course DEG analysis

MYBA1



MYB24



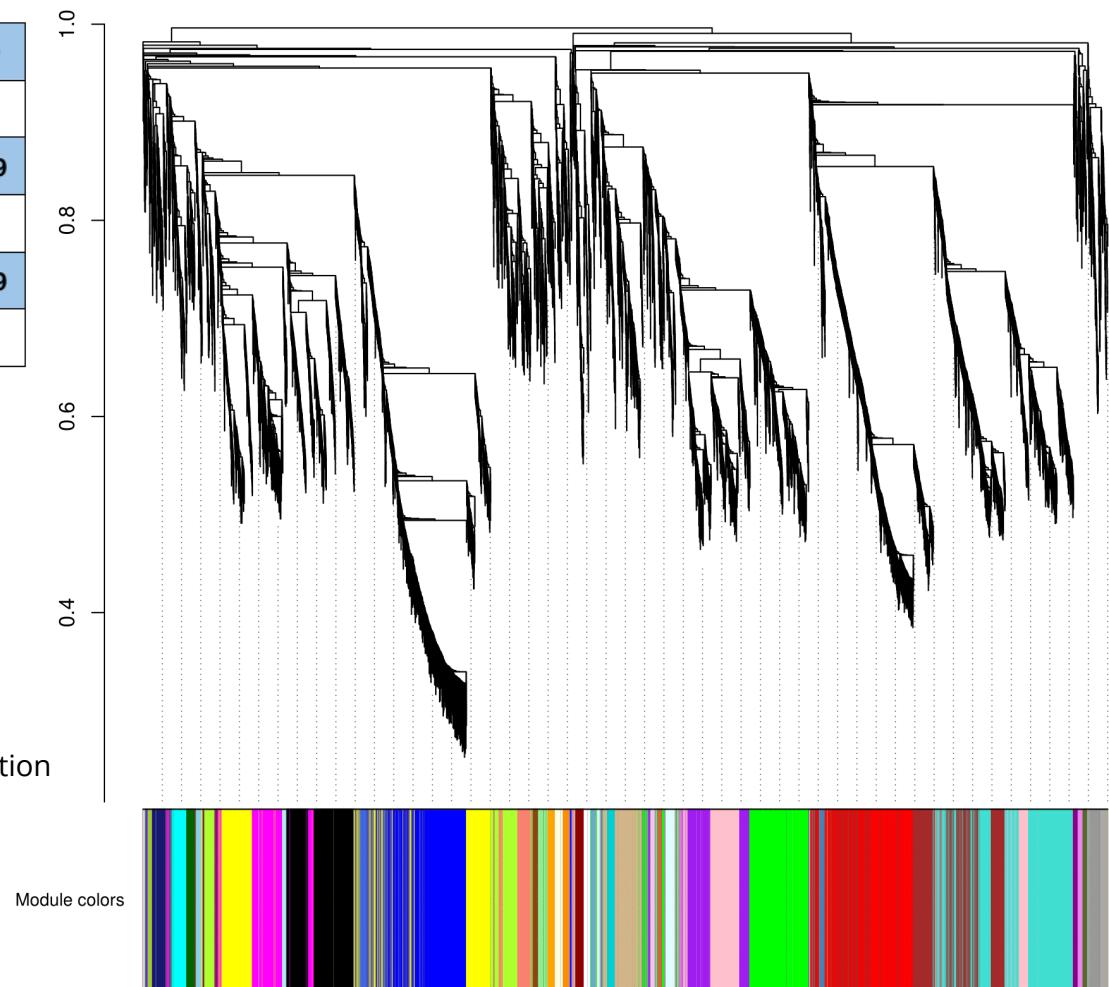
Weighted Gene Co-expression Networks Analysis

Eigengene modules (ME)

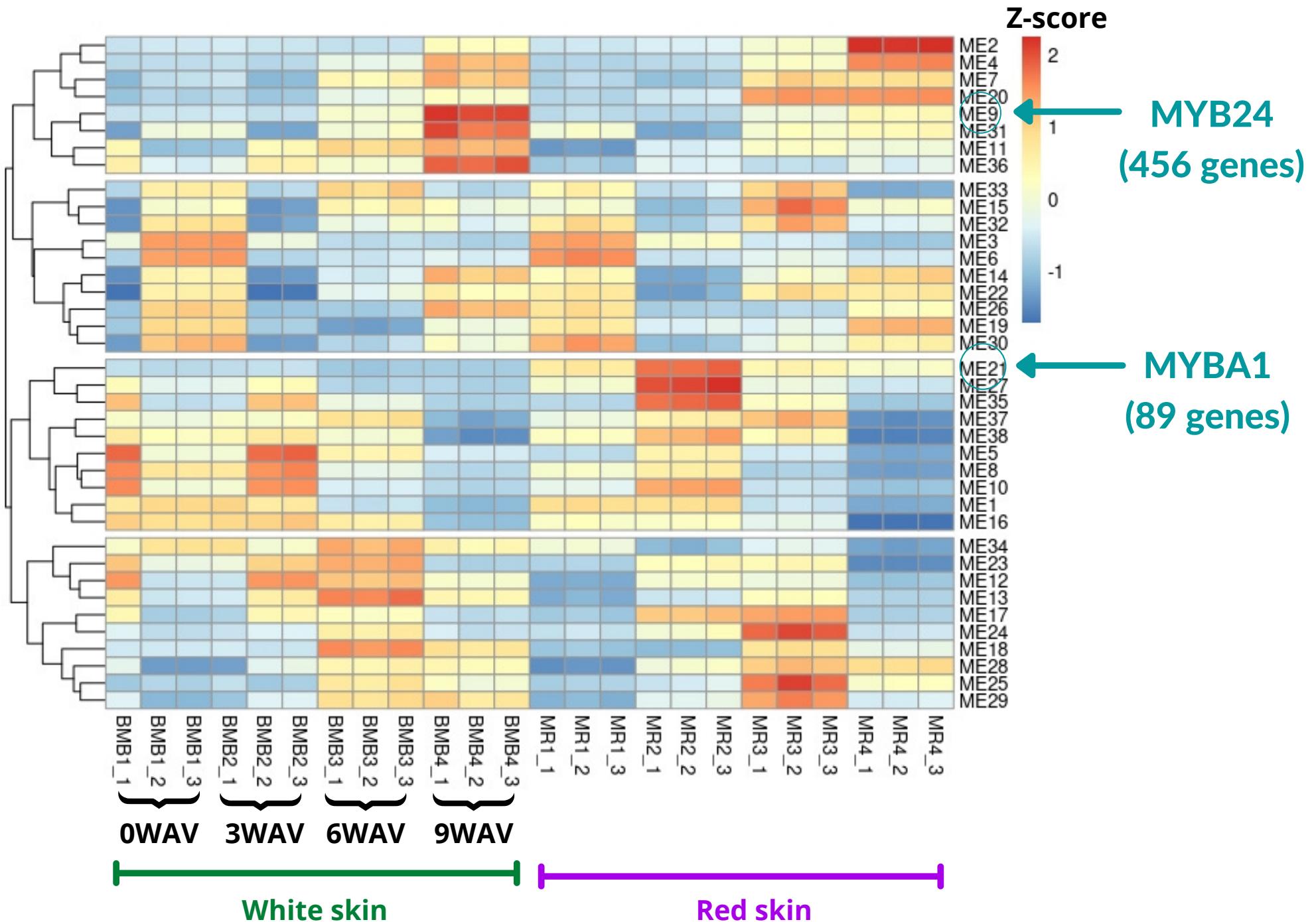
ME0	ME1	ME2	ME3	ME4	ME5	ME6	ME7	ME8	ME9
25	946	889	884	857	738	712	679	571	456
ME10	ME11	ME12	ME13	ME14	ME15	ME16	ME17	ME18	ME19
437	334	319	189	151	139	131	121	114	100
ME20	ME21	ME22	ME23	ME24	ME25	ME26	ME27	ME28	ME29
95	89	88	85	82	80	72	71	66	60
ME30	ME31	ME32	ME33	ME34	ME35	ME36	ME37	ME38	
60	57	52	52	50	47	46	46	46	

WGCNA R package

Langfelder P, Horvath S (2008) WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 2008, 9:559



Weighted Gene Co-expression Networks Analysis



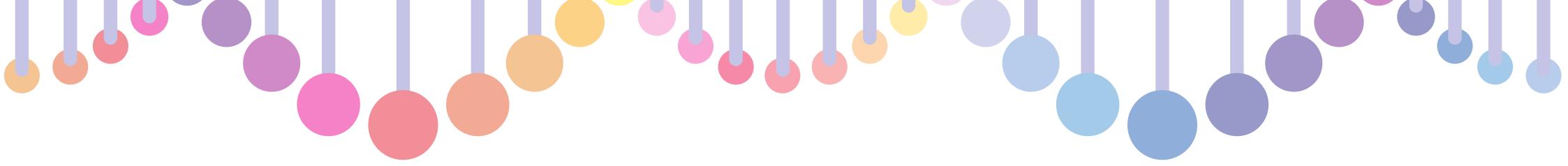
Enrichment analysis - Modules of interest

ME21 (with MYBA1)

Term ID	Term label	P-value
9.2	Secondary metabolism phenolics	4.64E-19
9.2.2	Secondary metabolism phenolics flavonoid biosynthesis	6.52E-14
9	Secondary metabolism	1.84E-12
9.2.2.4	Secondary metabolism phenolics flavonoid biosynthesis dihydroflavonols	1.60E-10
9.2.2.9	Secondary metabolism phenolics flavonoid biosynthesis anthocyanidins	1.55E-05
9.2.1	Secondary metabolism phenolics p-coumaroyl-CoA biosynthesis	4.12E-05
21.6.1.4	Cell wall organization lignin monolignol biosynthesis CCoA-OMT	6.58E-05
9.2.2.4.1	Secondary metabolism phenolics flavonoid biosynthesis dihydroflavonols flavonoid 3-hydroxylase	0.0002
9.2.2.4.2	Secondary metabolism phenolics flavonoid biosynthesis dihydroflavonols flavonoid 3	0.0005
21.6.1	Cell wall organization lignin monolignol biosynthesis	0.0008

ME9 (with MYB24)

Term ID	Term label	P-value
19.2	Protein homeostasis ubiquitin-proteasome system	1.80E-08
19.2.6	Protein homeostasis ubiquitin-proteasome system 26S proteasome	8.43E-07
19	Protein homeostasis	1.11E-06
19.2.6.2	Protein homeostasis ubiquitin-proteasome system 26S proteasome 19S regulatory particle	2.04E-06
23.2	Protein translocation mitochondrion	3.67E-05
23.2.3	Protein translocation mitochondrion inner mitochondrion membrane TIM translocation system	0.0001
19.2.4.2	Protein homeostasis ubiquitin-proteasome system ERAD substrate extraction	0.0008
19.1.3	Protein homeostasis protein quality control mitochondrial Hsp70 chaperone system	0.0021
15.5.17	RNA biosynthesis transcriptional regulation transcription factor (NAC)	0.0091
19.2.6.2.2	Protein homeostasis ubiquitin-proteasome system 26S proteasome 19S regulatory particle non-ATPase components	0.0091



Concluding remarks



In the variegated mutant different genes are regulated from the onset of the ripening.



MYBA1 association with the anthocyanins biosynthesis has been corroborated both in the pairwise and time-course pipelines.



The uncharacterized MYB24 has been associated to the white skin up-regulated genes.



Integration of different pipelines in time-series bulk RNAseq experiments provides a wider understanding of the results.

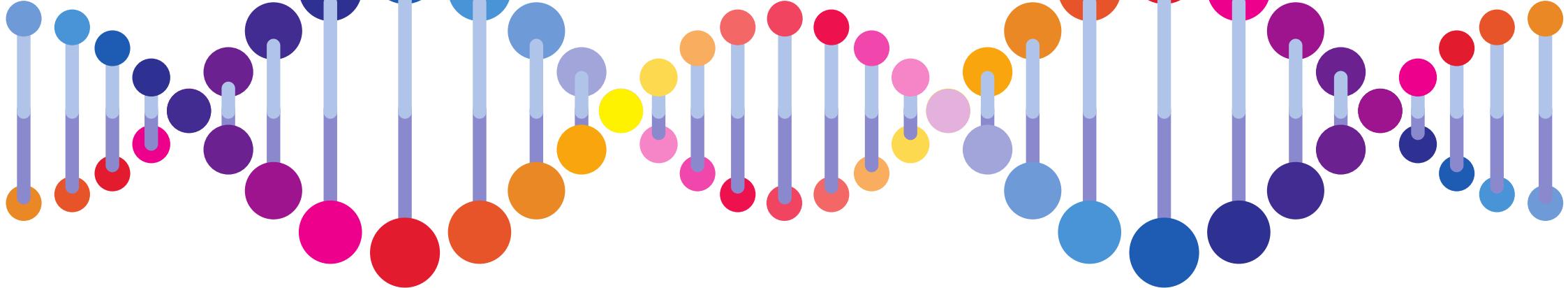
Additional resources available at
<https://tomsbiolab.com/transcriptomics>



Transcriptional Orchestration of Metabolism studied through Systems Biology

Home	Projects & Funding	Publications	Members & collabs	Posts	Positions	Resources	Login	🔍
Mostrar <input type="text" value="10"/> registros							Buscar: <input type="text"/>	
Vcost ID	V1 ID	P-adj	Gen module number	0WAV(logFC)	3WAV(logFC)	6WAV(logFC)	9WAV(logFC)	Catalog gene symbol
Vitvi00g01010	VIT_00s0958g00010	5.05E-05	ME6	-0.281208213817875	-0.489705648436178	NA	NA	NA
Vitvi00g01011	VIT_00s0958g00020	2.00E-23	ME35	NA	NA	-0.382342207378065	0.462475185695252	FKBPb
Vitvi00g01018	VIT_00s0983g00010	4.30E-14	ME8	0.314177483759603	0.729594828091691	0.674036409459697	NA	Abhydrolase domain-containing protein
Vitvi00g01022	VIT_00s1003g00010	9.40E-22	ME16	0.215365979089006	0.382792412569727	0.480978909662391	0.396642661662213	NA
								Senescence-associated protein

Code available at
https://github.com/apc1992/MS_thesis



THANK

YOU

