Investigation of transcription dynamics in SLE through the analysis of RNA-sequencing data

Co-supervised by

George Bertsias
Laboratory of Rheumatology,
Autoimmunity and Inflammation
Medical School, University of Crete

Christoforos Nikolaou
Computational Genomics Group
"Alexander Fleming" BSRC

Sofia Papanikolaou, 10th October 2021

Outline

- Master's thesis project: Alternative Splicing in SLE

- Bioinformatics support in projects (labs in Crete, BRFAA)
  - CTLA4-mediated tolerogenic effect on DCs
  - Gender bias in SLE (BRFAA) - SMC1a (Crete)
  - Immunometabolism

- Ongoing project: Early Arthritis

- Pending/ongoing projects (pre-SLE, “by-stander” gene effect)

Recent findings in Systemic lupus erythematosus (SLE)

RNA sequencing from whole blood
from 142 SLE patients and 58 healthy individuals

- **perturbed mRNA splicing** of SLE patients
  - immune system
  - IFN signaling

Previous Studies

SPLICED mRNA	CONSEQUENCE
BANK1	isoform lacking exon 2
LILRA2	novel isoform lacking 3 amino acids
TCRζ	different isoforms
IRF5	specific exon 1 used
RasGRP1	aberrant splice variants
CD72	isoform lacking exon 8
IL20R	soluble receptor
CSR	decreased splicing efficiency of exon 11

SLE associated alternative-splicing quantitative trait loci

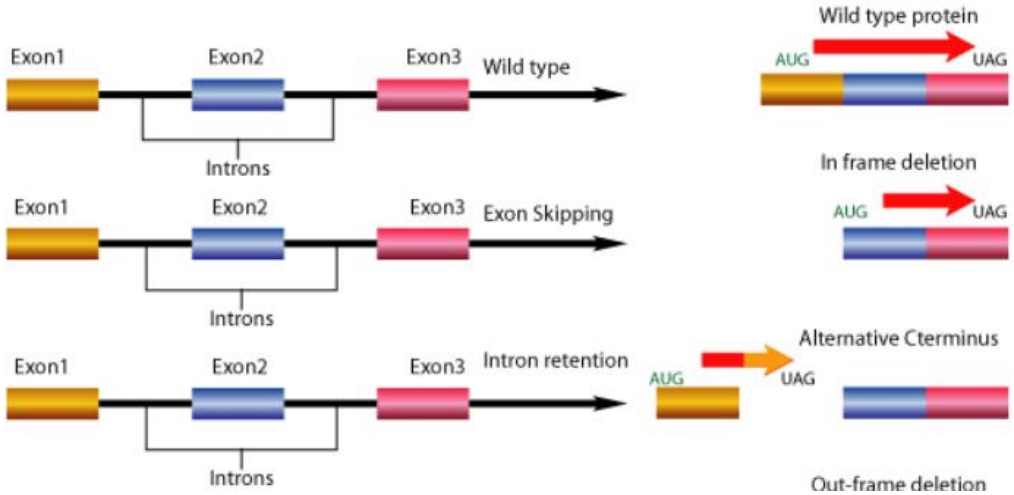
TCF7, SKP1, BLK, NADSYN1, IKZF2, WDFY4 and IRF5

Panousis N et al., Ann Rheum Dis., 2019.
Evsyukova I et al., RNA Biology, 2010.
Odhams CA et al., Hum Mol Genet., 2017.

Recent findings in Systemic lupus erythematosus (SLE)

RNA sequencing from whole blood from 142 SLE patients and 58 healthy individuals

- **perturbed mRNA splicing** of SLE patients
  - immune system
  - IFN signaling



Previous Studies

SPLICED mRNA	CONSEQUENCE
BANK1	isoform lacking exon 2
LILRA2	novel isoform lacking 3 amino acids
TCRζ	different isoforms
IRF5	specific exon 1 used
RasGRP1	aberrant splice variants
CD72	isoform lacking exon 8
IL20R	soluble receptor
CSR	decreased splicing efficiency of exon 11

SLE associated alternative-splicing quantitative trait loci

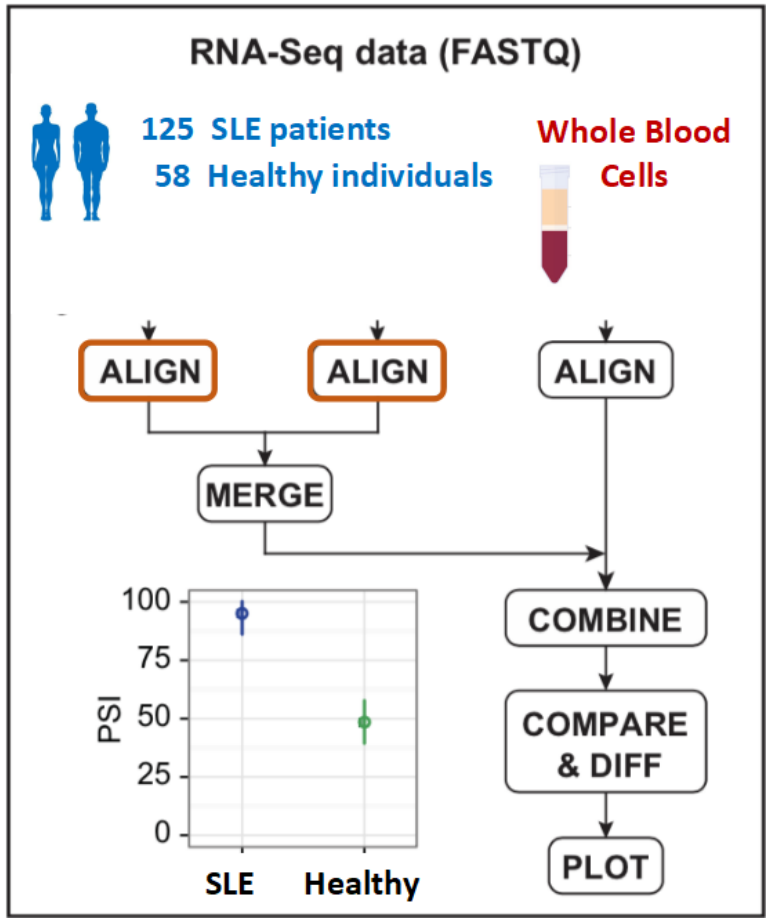
TCF7, SKP1, BLK, NADSYN1, IKZF2, WDFY4 and IRF5

Impact on gene expression

- partial loss of function
- gain of function
- introduction of premature termination codons
  - degradation of mRNA (NMD)

Panousis N et al., Ann Rheum Dis., 2019.
Evsyukova I et al., RNA Biology, 2010.
Odhams CA et al., Hum Mol Genet., 2017.

Systematic analysis of alternative splicing patterns in whole-blood samples of SLE



Aim of the study

- Detailed and analytical study of splicing dynamics in SLE
- Identification of alternative splicing events between healthy individuals and SLE patients in **different disease states**

Methodology

1. merging: the align outputs from various **samples are pulled together** into a new set of output files, as read coverage for the independent replicates is not deep enough for a proper AS analysis.

2. combine: Estimation of **Percent Spliced In** for each group $$PSI = \frac{IR}{IR + ER}$$

the ratio of normalized read counts indicating inclusion of a transcript element over the total normalized reads for that event (both inclusion and exclusion reads)

3. diff: performs a statistical test to assess whether the PSI distributions of the two compared groups are significantly different.

Comparison 1: SLE inactive vs. Healthy

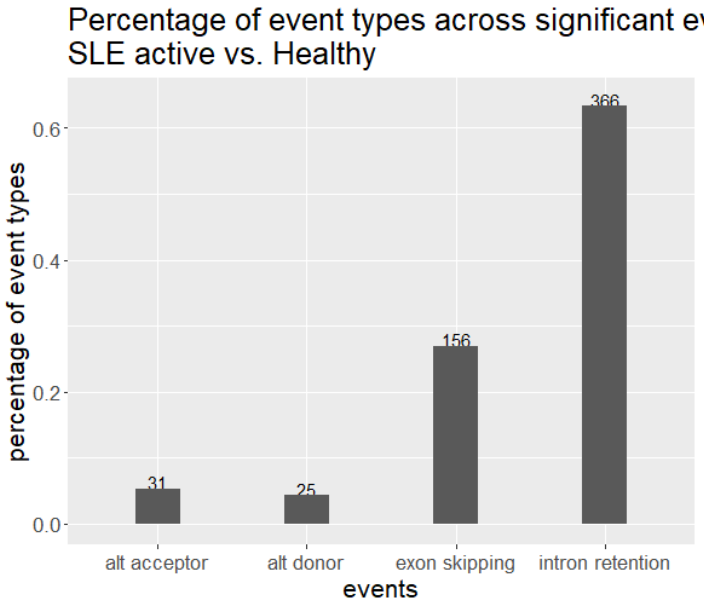
Comparison 2: SLE active vs. Healthy

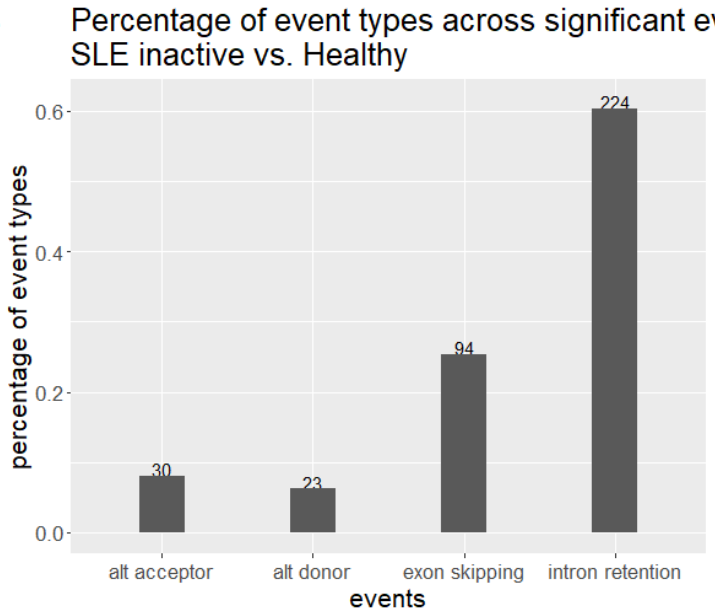
Papanikolaou S., et al., Genes 2021.

Systematic analysis of alternative splicing patterns in whole-blood samples of SLE

Results

- Extensive Perturbation of Splicing and Predominance of Intron Retention Events





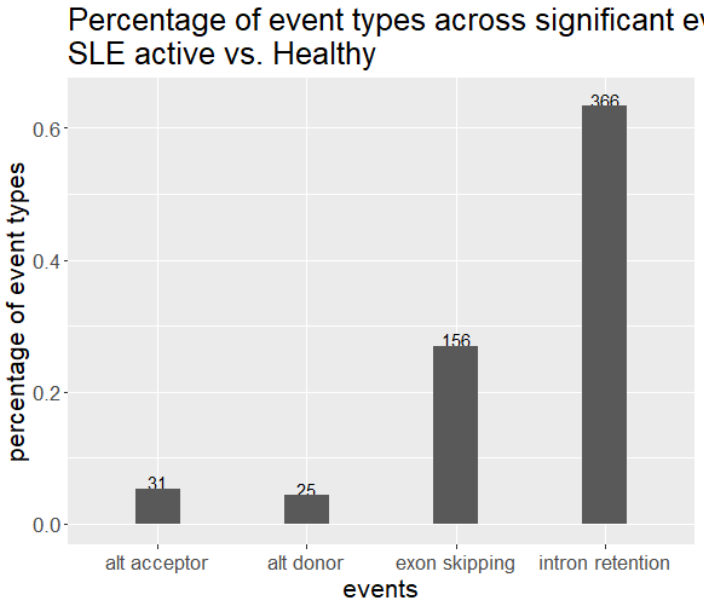
associated with gene repression

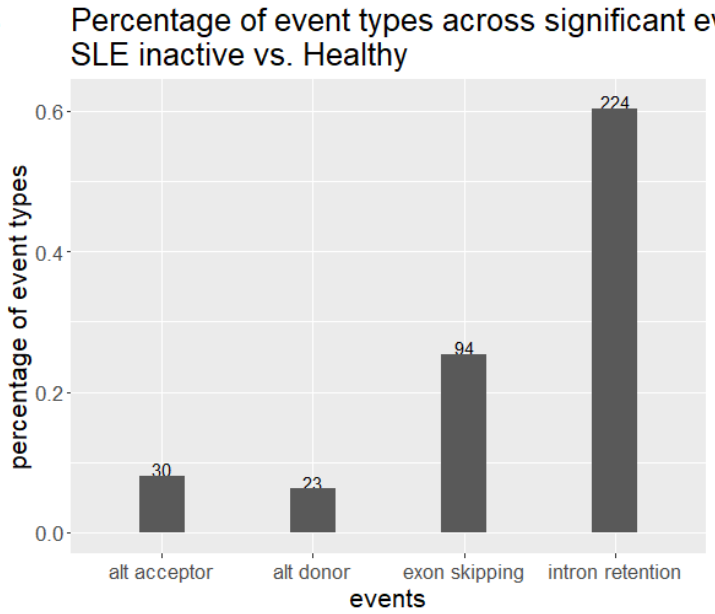
Papanikolaou S., et al., Genes 2021.

Systematic analysis of alternative splicing patterns in whole-blood samples of SLE

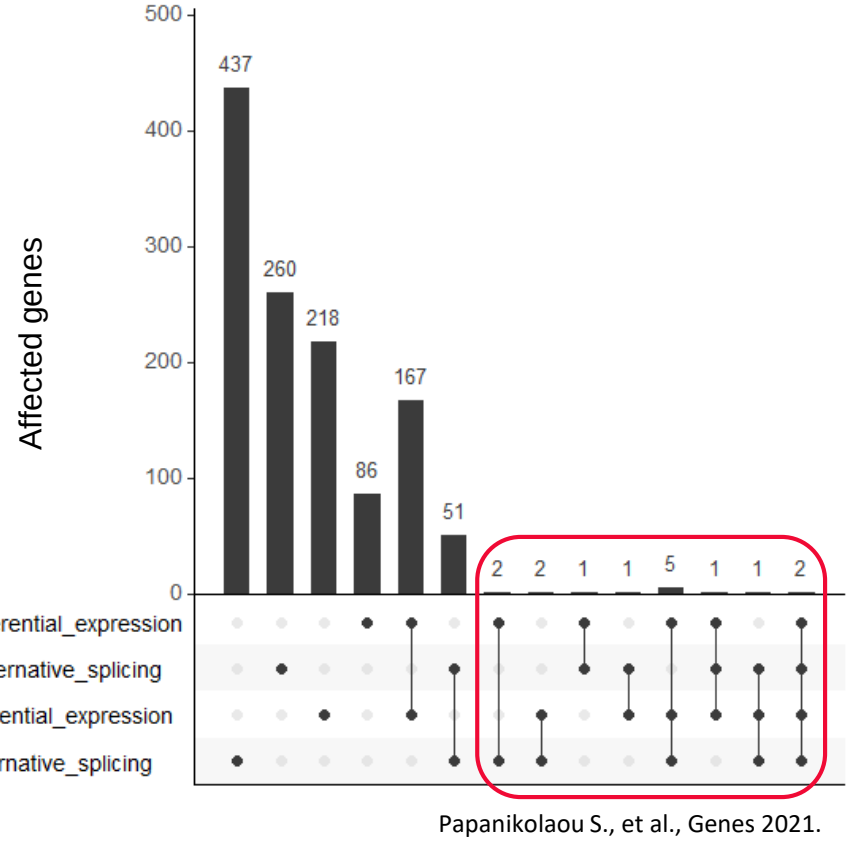
Results

- Extensive Perturbation of Splicing and Predominance of Intron Retention Events





associated with gene repression



Papanikolaou S., et al., Genes 2021.

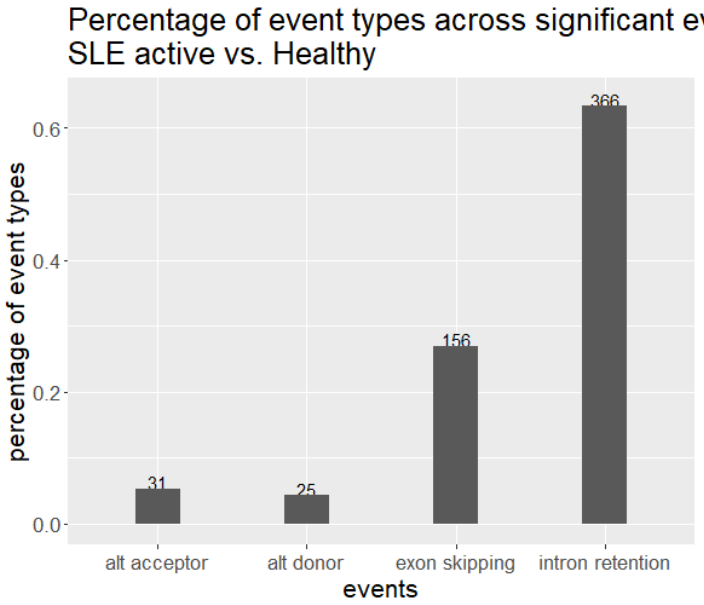
❑ Do AS events occur among the subset of genes that are differentially expressed?

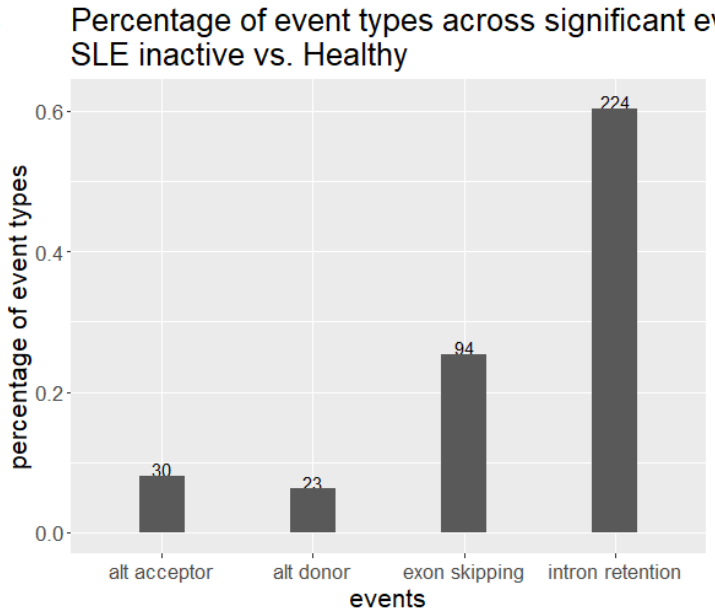
- Alternative Splicing and Differential Expression Involve Different Genes

Systematic analysis of alternative splicing patterns in whole-blood samples of SLE

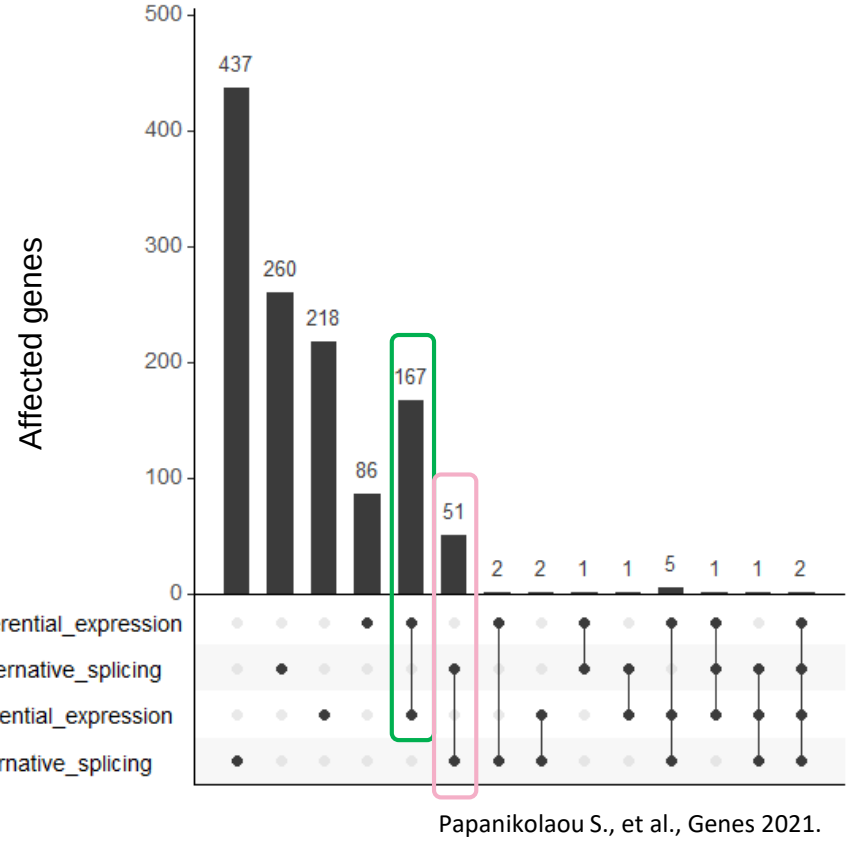
Results

- Extensive Perturbation of Splicing and Predominance of Intron Retention Events





associated with gene repression



Papanikolaou S., et al., Genes 2021.

❑ **Do AS events occur among the subset of genes that are differentially expressed?**

- Alternative Splicing and Differential Expression Involve Different Genes
- AS events can discriminate active from inactive disease states

• Systematic analysis of alternative splicing patterns in whole-blood samples of SLE

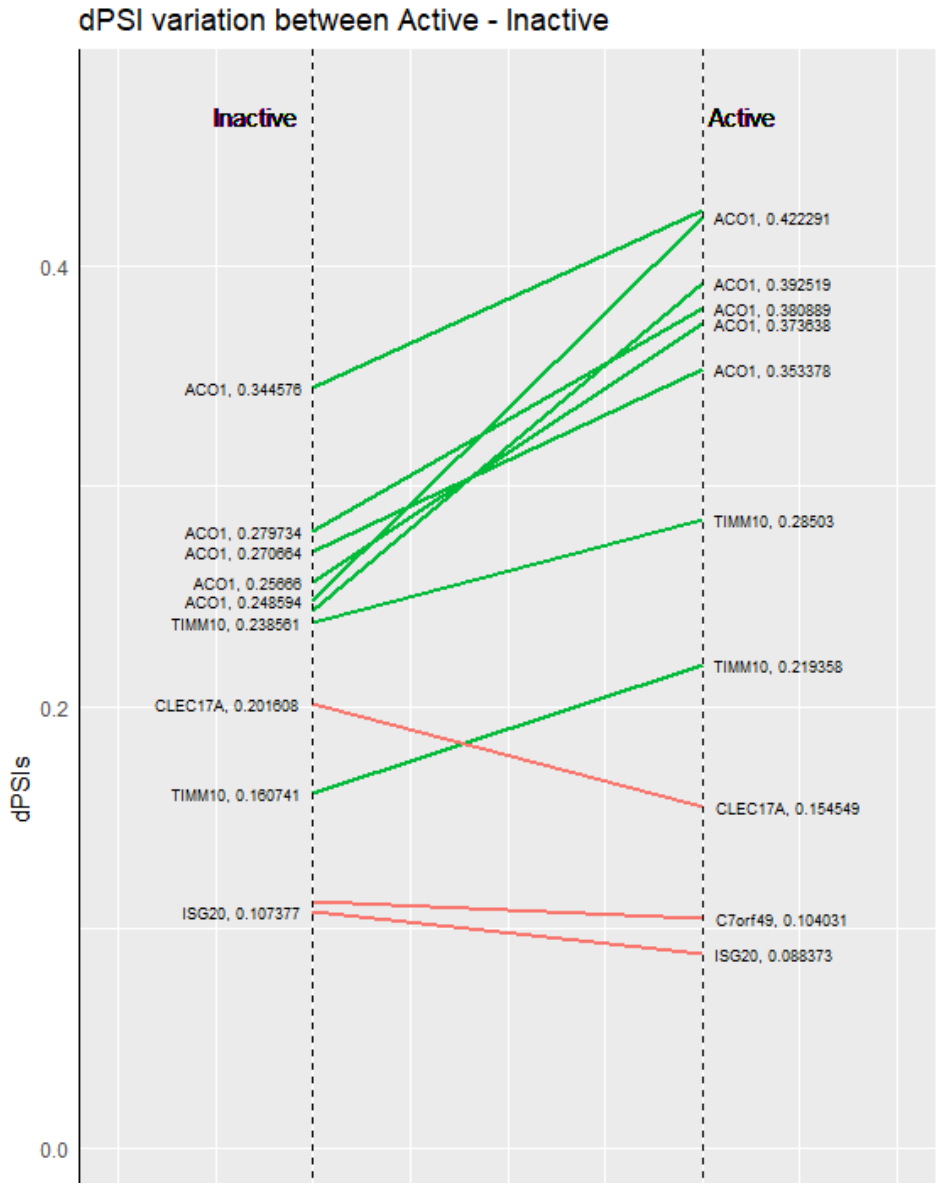
Results

Common Alternative Splicing events between Active SLE vs. Healthy and Inactive SLE vs. Healthy

$dPSI = PSI_{SLE} - PSI_{healthy}$

dPSI is increased in active disease state

effect is intensified with increasing disease activity

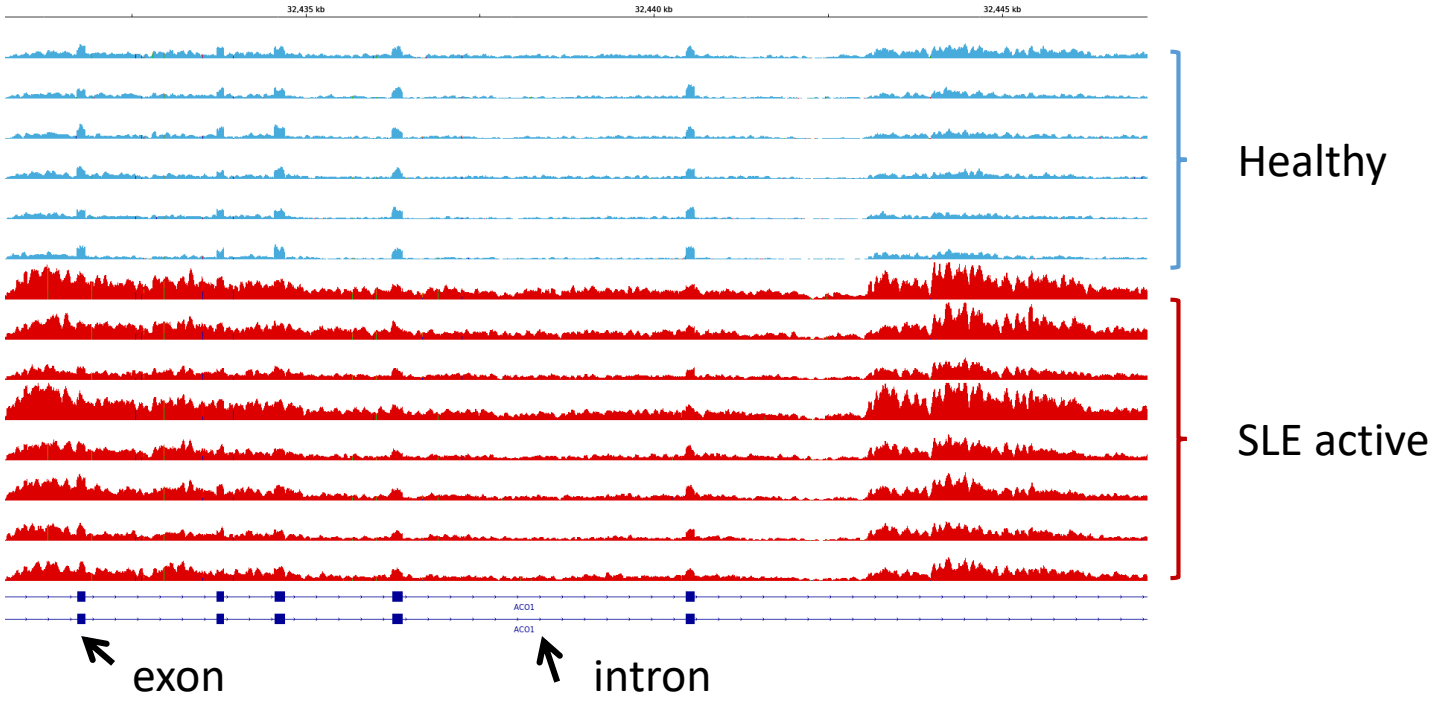


Papanikolaou S., et al., Genes 2021.

Systematic analysis of alternative splicing patterns in whole-blood samples of SLE

Results

- ACO1 undergoes intron retention at multiple sites in SLE samples



↑ **iron levels**

binds to a 4Fe-4S cluster and functions as an aconitase, catalyzing the conversion of citrate to isocitrate

↓ **iron levels**

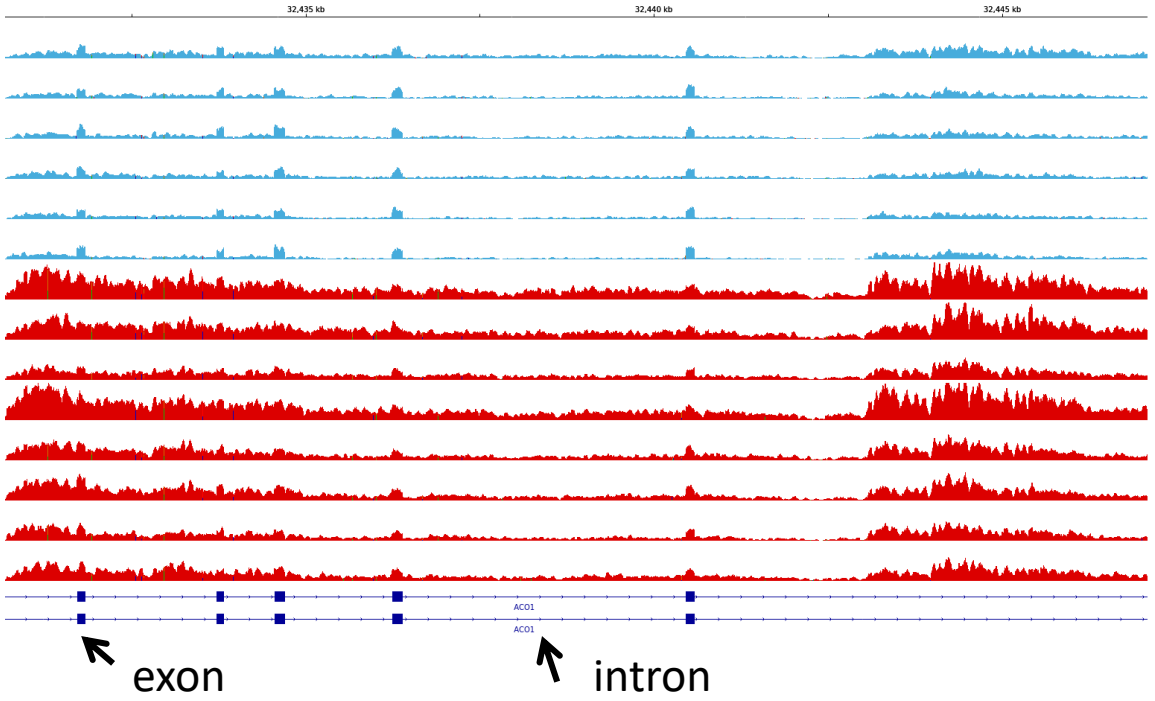
binds to iron-responsive elements (IREs), resulting in repression of translation of ferritin mRNA, and inhibition of degradation of the otherwise rapidly degraded transferrin receptor mRNA

Papanikolaou S., et al., Genes 2021.

Systematic analysis of alternative splicing patterns in whole-blood samples of SLE

Results

- ACO1 undergoes intron retention at multiple sites in SLE samples



↑ **iron levels**

binds to a 4Fe-4S cluster and functions as an aconitase, catalyzing the conversion of citrate to isocitrate

↓ **iron levels**

binds to iron-responsive elements (IREs), resulting in repression of translation of ferritin mRNA, and inhibition of degradation of the otherwise rapidly degraded transferrin receptor mRNA

AS constitutes an additional layer of transcriptional regulation in SLE and might affect a number of important biological pathways not previously detected by differential gene expression analysis.

- ❑ **Which cell type express the AS isoforms of these genes?**
- ❑ **What is the consequence -functional role of the produced isoforms?**

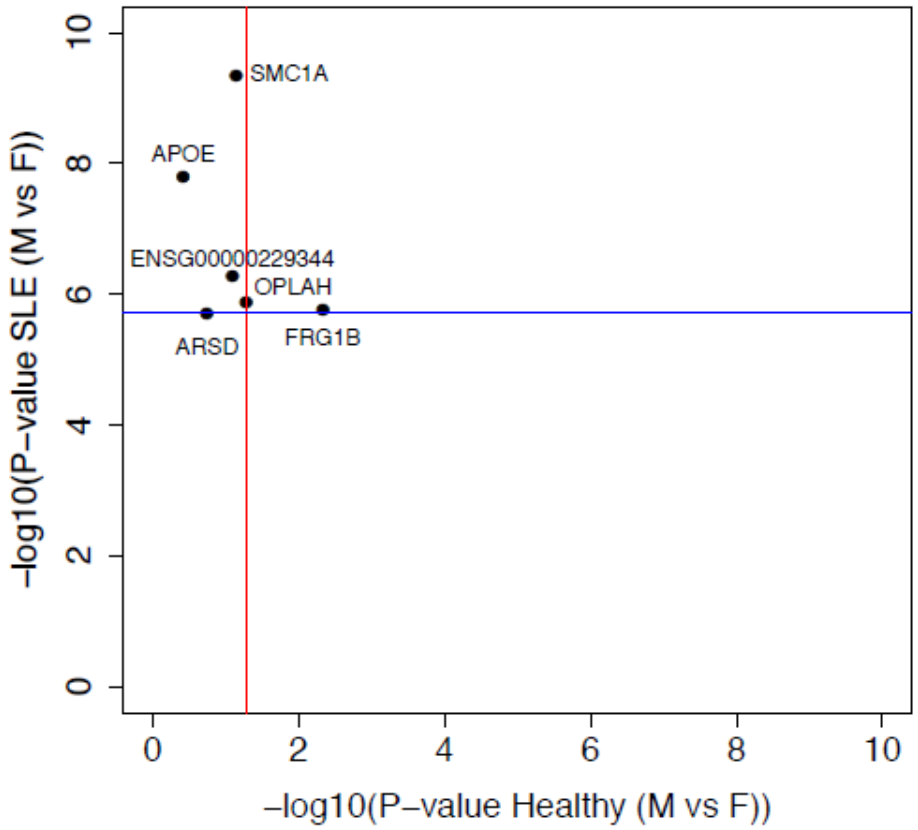
Papanikolaou S., et al., Genes 2021.

Recent findings in Systemic lupus erythematosus (SLE)

RNA sequencing from whole blood
from 142 SLE patients and 58 healthy individuals

- **gender differences** in gene expression
  - more frequent in females vs. males (9:1)
  - more severe in males vs. females

Genes that are differentially expressed
between **SLE females vs. males**
but <u>not</u> between **healthy females vs. males**



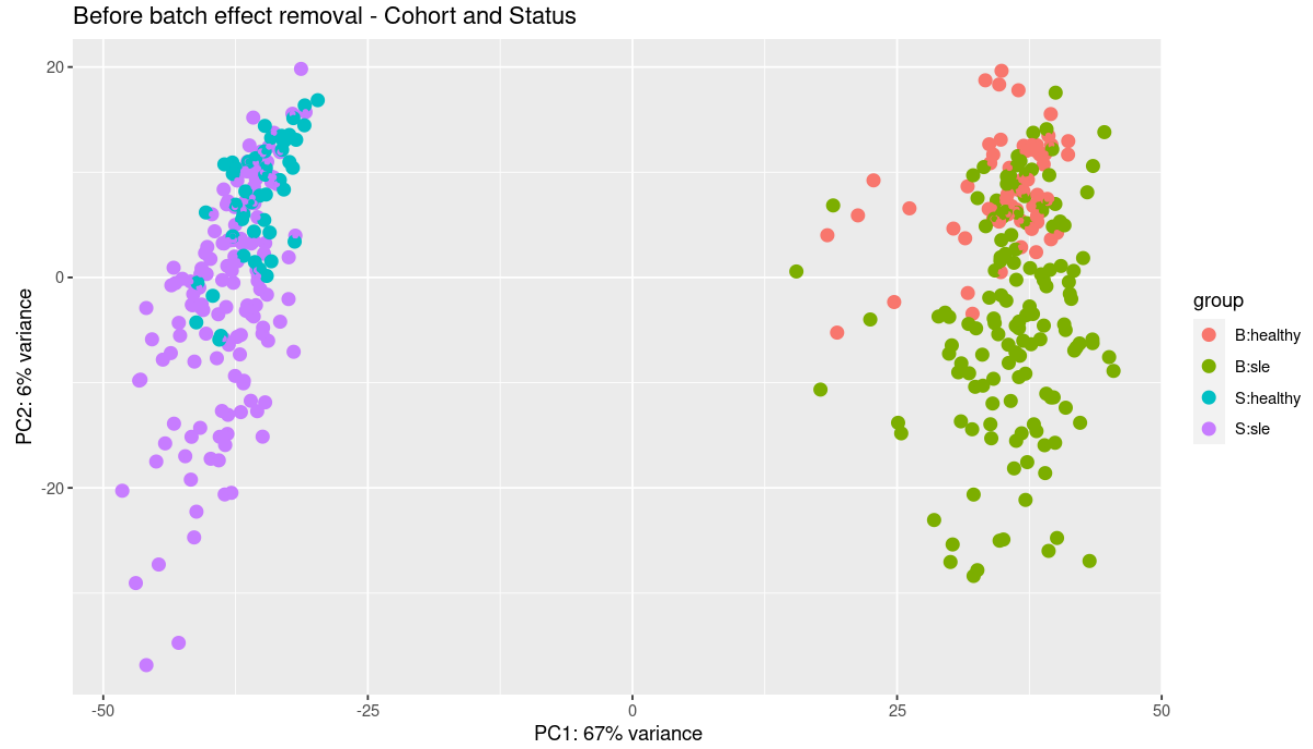
Panousis N et al., Ann Rheum Dis., 2019.

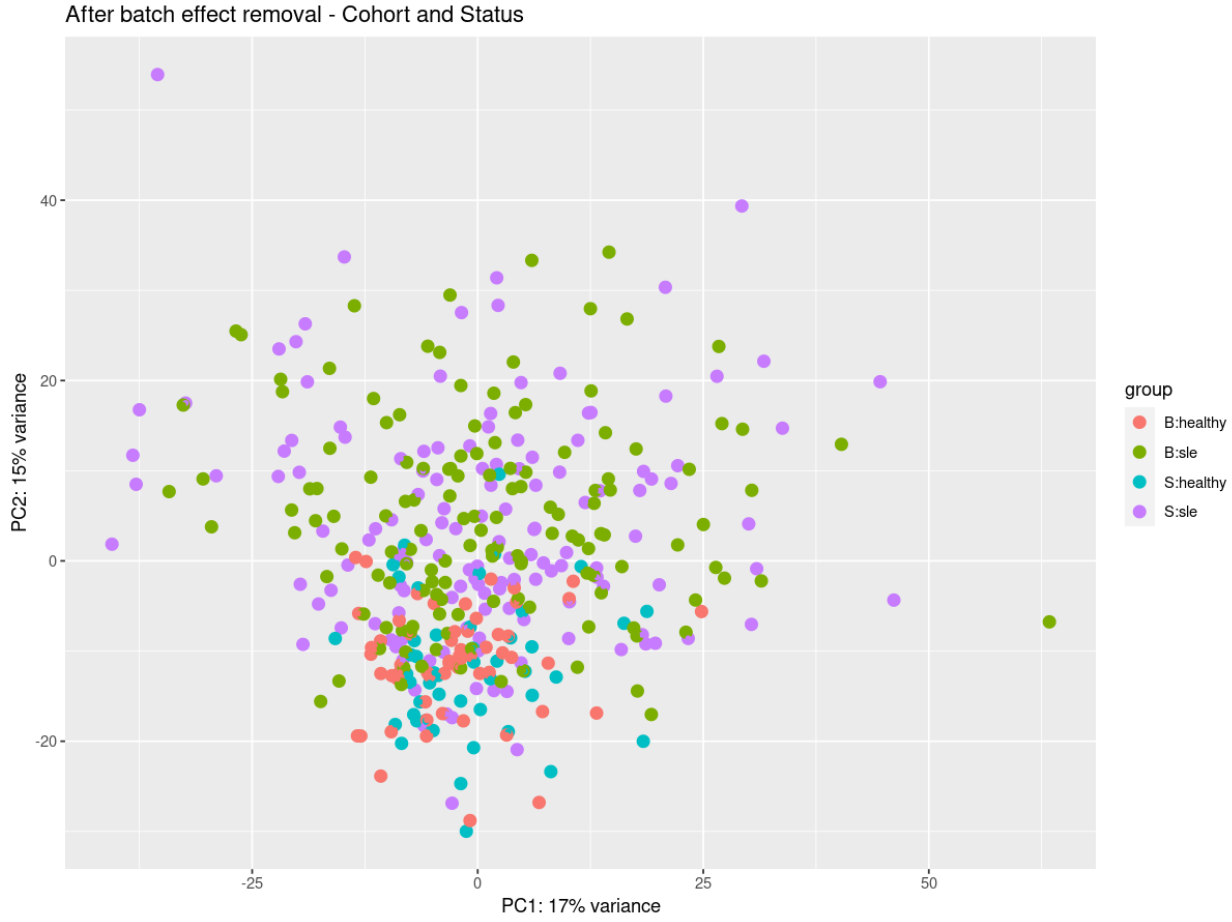
Ongoing Projects: Gender bias in SLE (BRFAA)

Merging datasets

	Syscid Cohort		Cohort Panousis et al.	
	Healthy	SLE	Healthy	SLE
Female	47	117	48	118
Male	1	24	10	22
Total	189		198	

Extended dataset → Batch effect





Ongoing Projects: Gender bias in SLE (BRFAA)

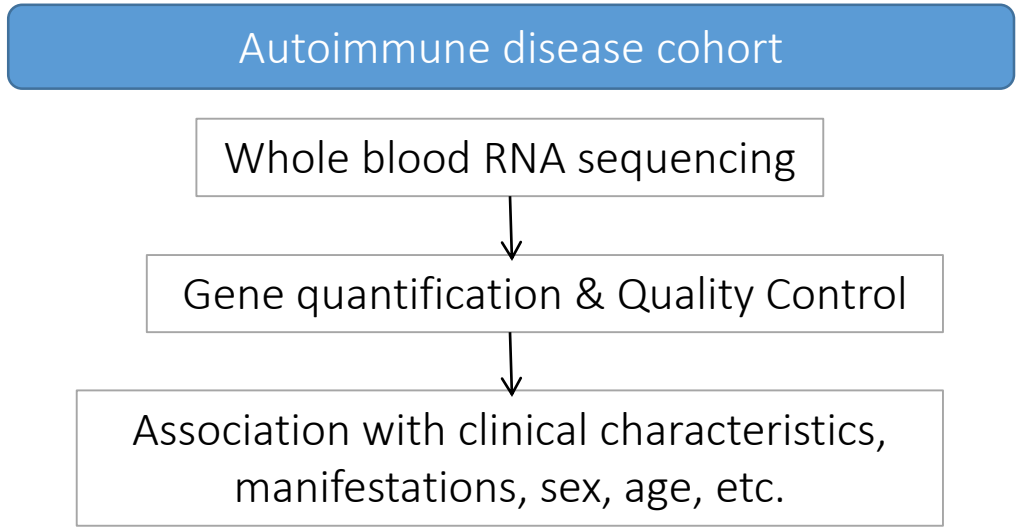
Gender-biased genes

Genes that are differentially expressed between
SLE females vs. males
but <u>not</u> between
healthy females vs. males
(*adj. pvalue < 0.05 ,*
no threshold for log2FC)

- Under-expression of specific genes in SLE men
- Variability in SLE women

❑ **Do these genes confer disease susceptibility on females or severity on males?**

How to identify prognostic gene signatures for progression into disease and predict the disease outcome?



How to identify prognostic gene signatures for progression into disease and predict the disease outcome?

Autoimmune disease cohort

Whole blood RNA sequencing

↓

Gene quantification & Quality Control

↓

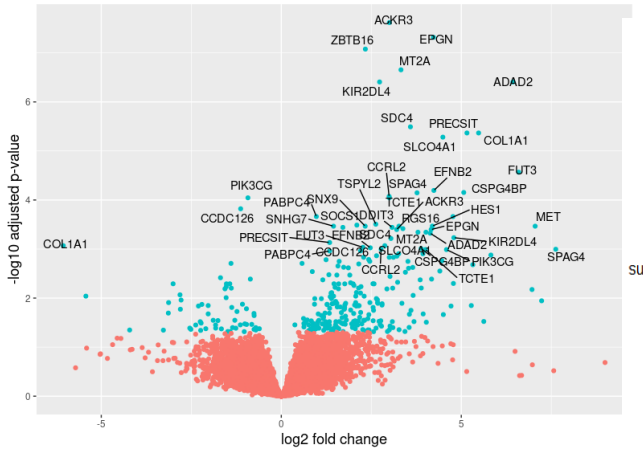
Association with clinical characteristics, manifestations, sex, age, etc.

↓

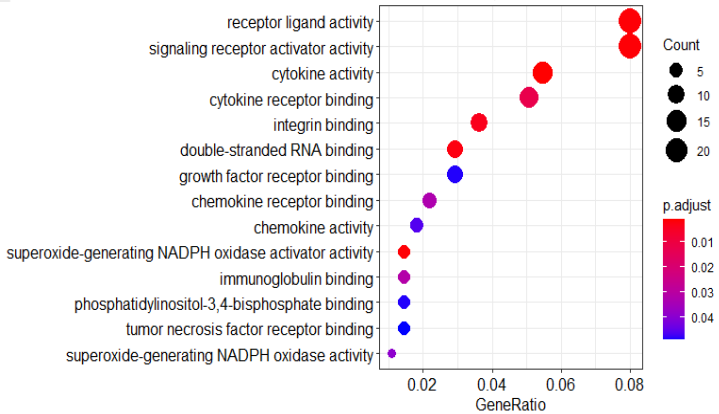
Differential gene expression analysis and functional enrichment analysis

↓

Volcano plot



Over-representation analysis plot



How to identify prognostic gene signatures for progression into disease and predict the disease outcome?

Autoimmune disease cohort

Whole blood RNA sequencing

Gene quantification & Quality Control

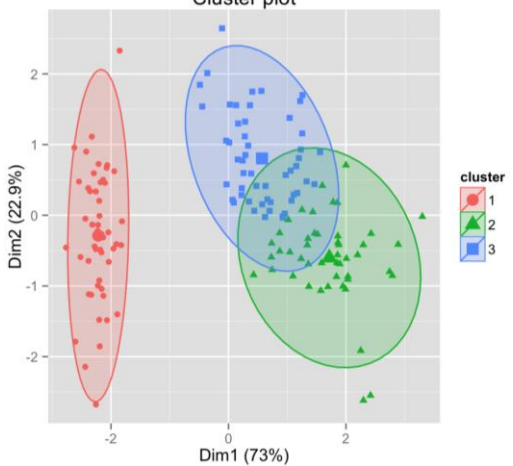
Association with clinical characteristics, manifestations, sex, age, etc.

Unsupervised approaches

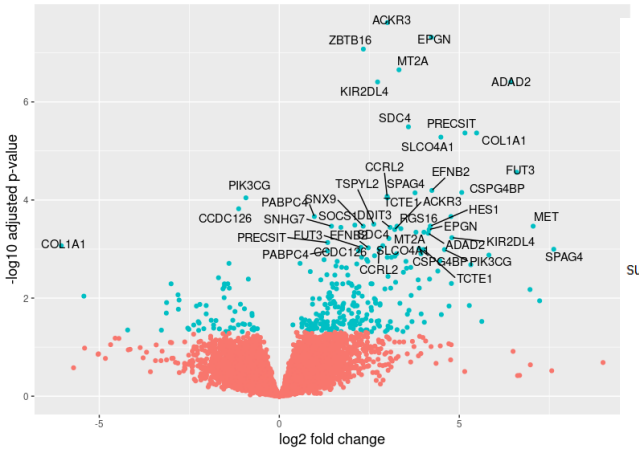
Identification of groups of patients

Differential gene expression analysis and functional enrichment analysis

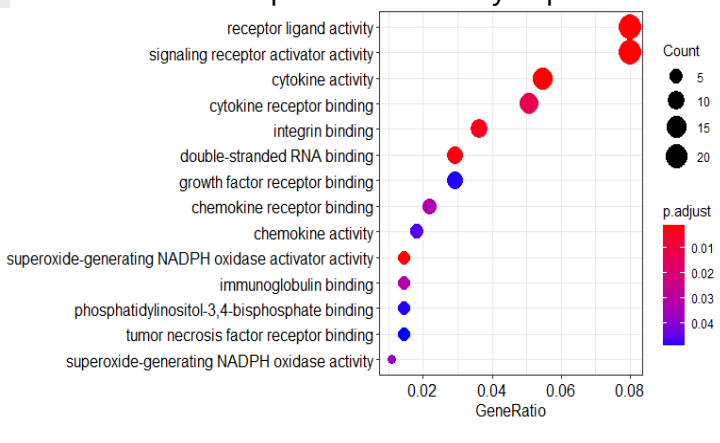
Cluster plot



Volcano plot



Over-representation analysis plot



How to identify prognostic gene signatures for progression into disease and predict the disease outcome?

Autoimmune disease cohort

Whole blood RNA sequencing

Gene quantification & Quality Control

Association with clinical characteristics, manifestations, sex, age, etc.

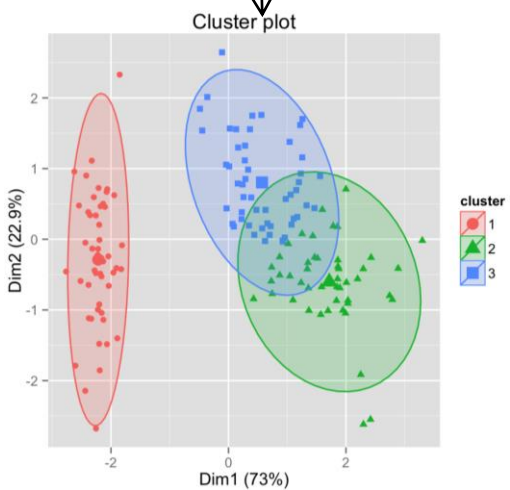
Unsupervised approaches

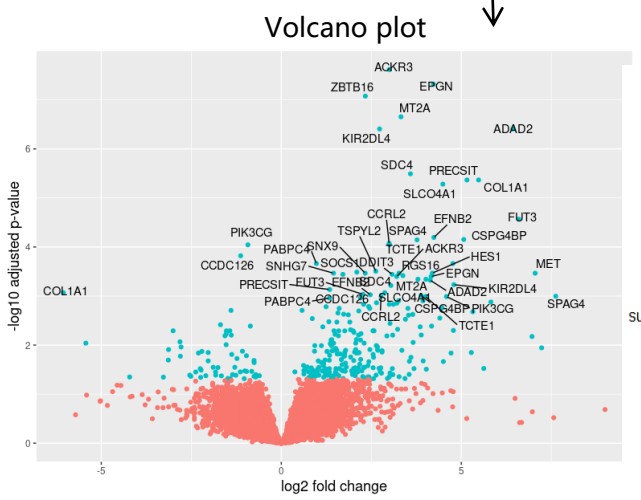
Identification of groups of patients

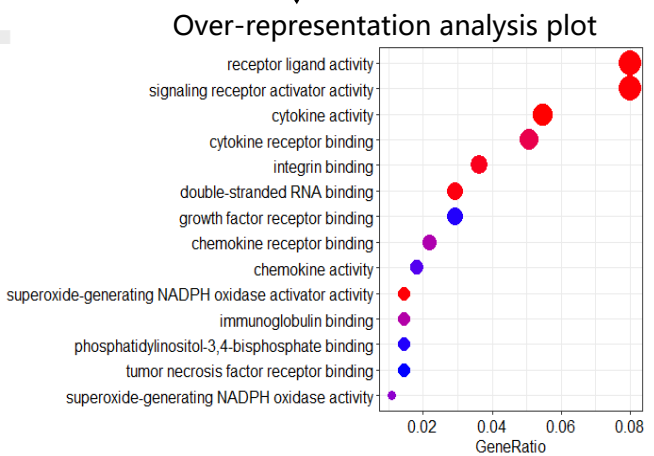
Differential gene expression analysis and functional enrichment analysis

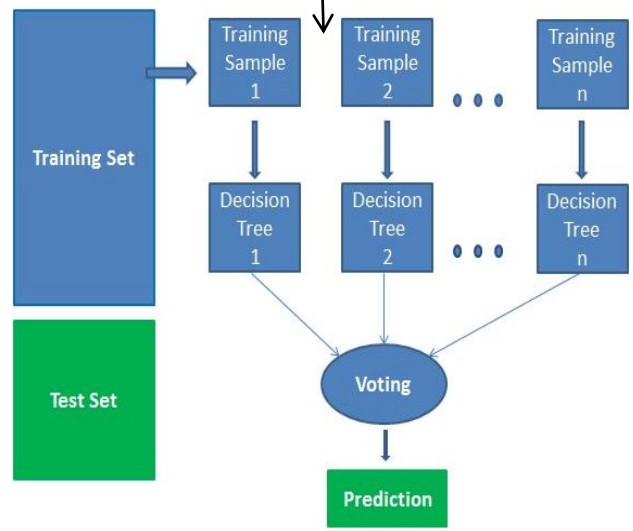
Supervised approaches

Classification methods





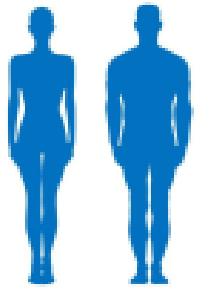




Ongoing Projects: Early Arthritis

Experimental design

Early disease state



100 Females
27 Males

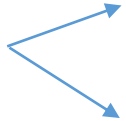
Whole Blood

Blood collection performed on <u>early disease state,</u> without receiving treatment (Baseline)

RNA sequencing

Baseline

Follow up after 2 years



Differentially Gene Expression analysis between groups

Classification methods

Information about
their clinical characteristics and manifestations

Categorization into different **diagnosis** groups
and **outcome** groups

Aim of the study:

Identification of prognostic gene signature for progression into rheumatoid arthritis and other diseases and gene signature for discrimination of disease outcome.

Ongoing Projects: Early Arthritis

Categorization after 2-year follow up

Tables of groups

Diagnosis	Group	# samples
Rheumatoid arthritis	1	60
Late onset rheumatoid arthritis	1	
Undifferentiated Arthritis	2	52
Undifferentiated Oligoarthritis	2	
Undifferentiated Monoarthritis	3	7
Ankylosing Spondylitis	3	
Psoriatic arthritis	3	
Undifferentiated spondyloarthritis	3	
Reactive Arhritis	3	
SLE	4	8
Systemic Sclerosis	4	
Collagen disease	4	

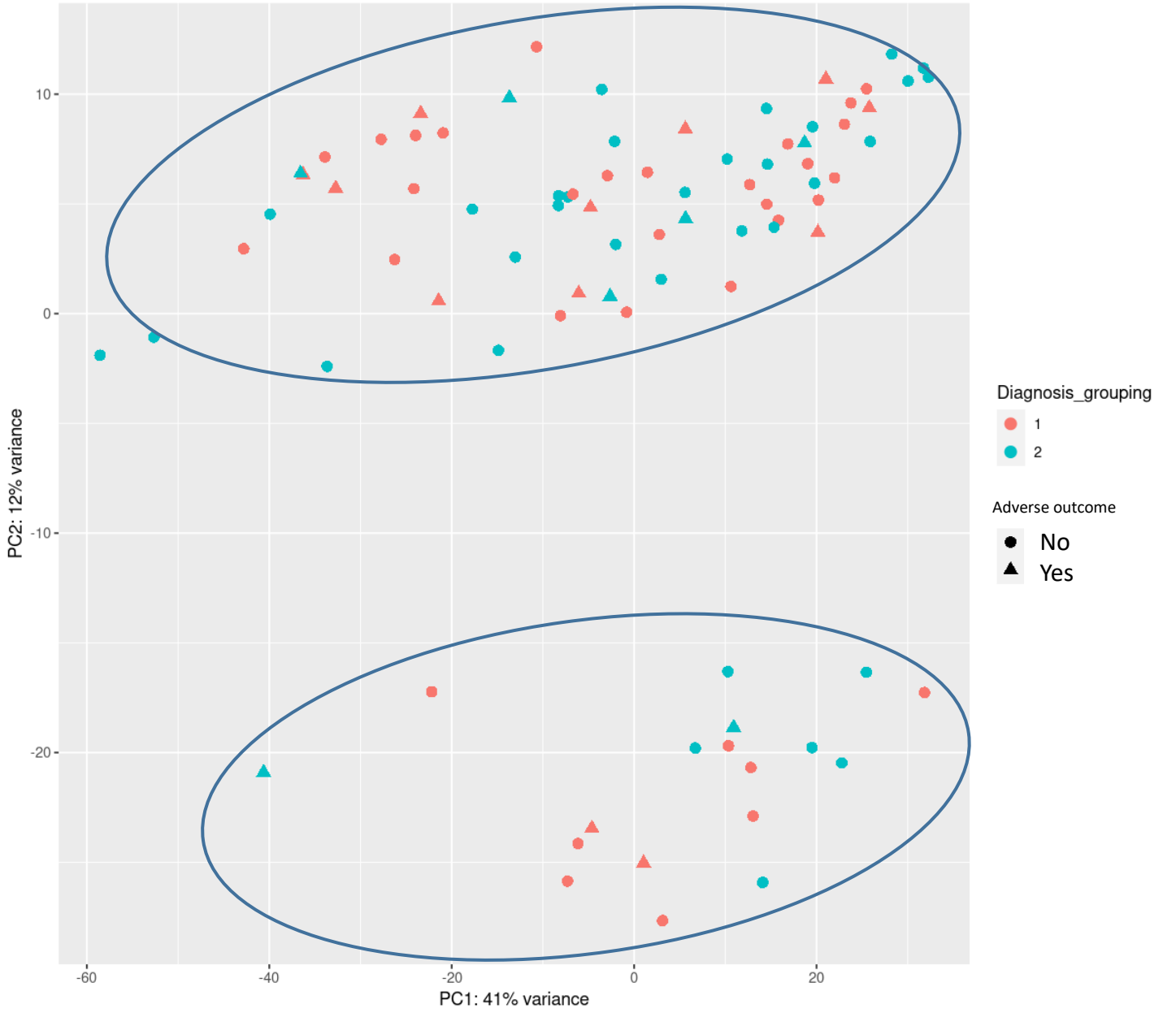
Criteria	Outcome group	# samples yes/no
HDAorHAQ>1 or bDMARD@2yrs COMBO No DMARDs_except HCQ_ever	Adverse outcome	19/66
DAS28<3,2 & HAQ≤0,25 & no bDMARD @2yrs COMBO No DMARDs_except HCQ_ever	Favorable outcome	51/33

Ongoing Projects: Early Arthritis

Principal Component Analysis

- Gene expression data of patients from **adverse outcome group or not**

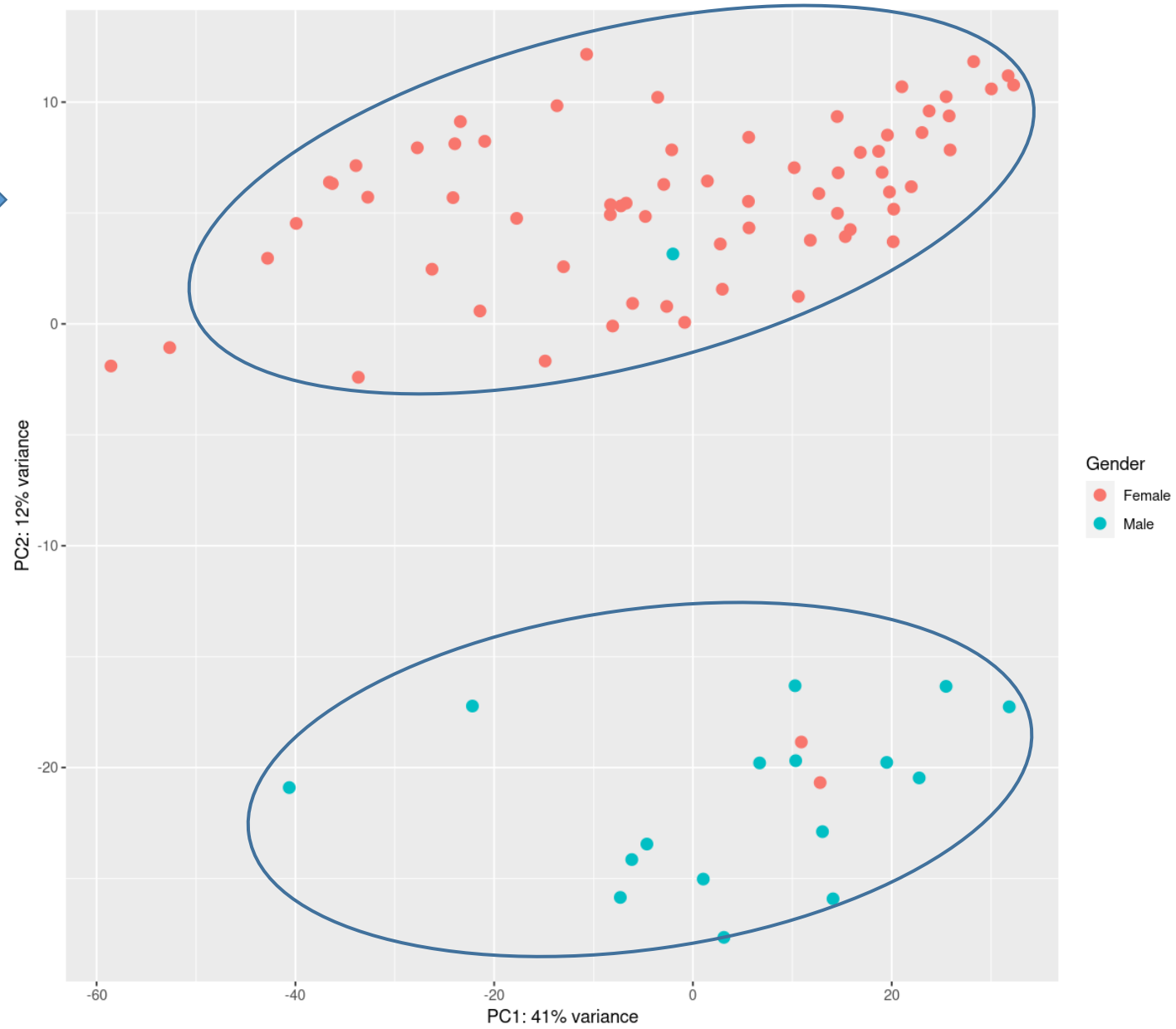
Two clusters, not associated with diagnosis or outcome group



Ongoing Projects: Early Arthritis

Principal Component Analysis

- Gene expression data of patients from **adverse outcome group or not**



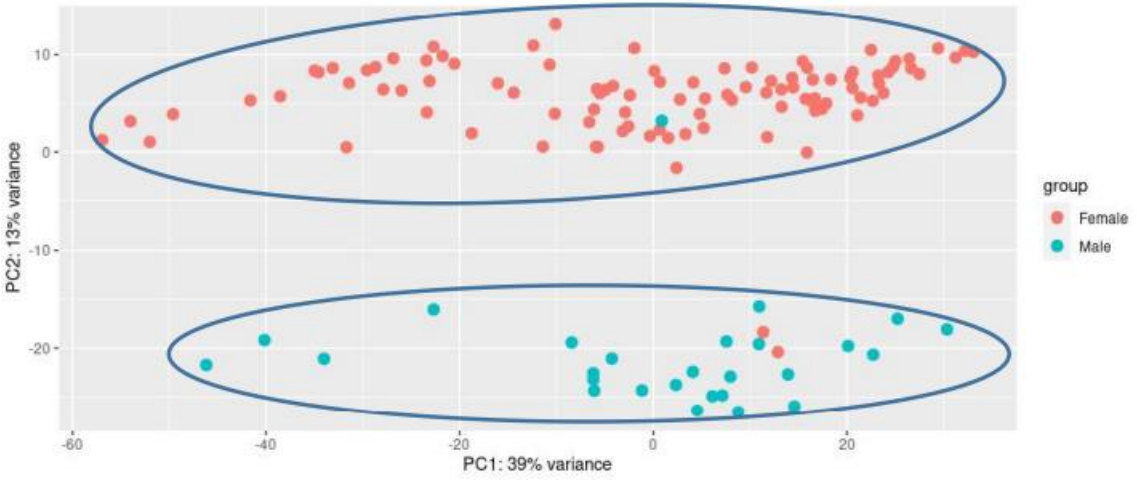
Clustering associated with sex

Ongoing Projects: Early Arthritis

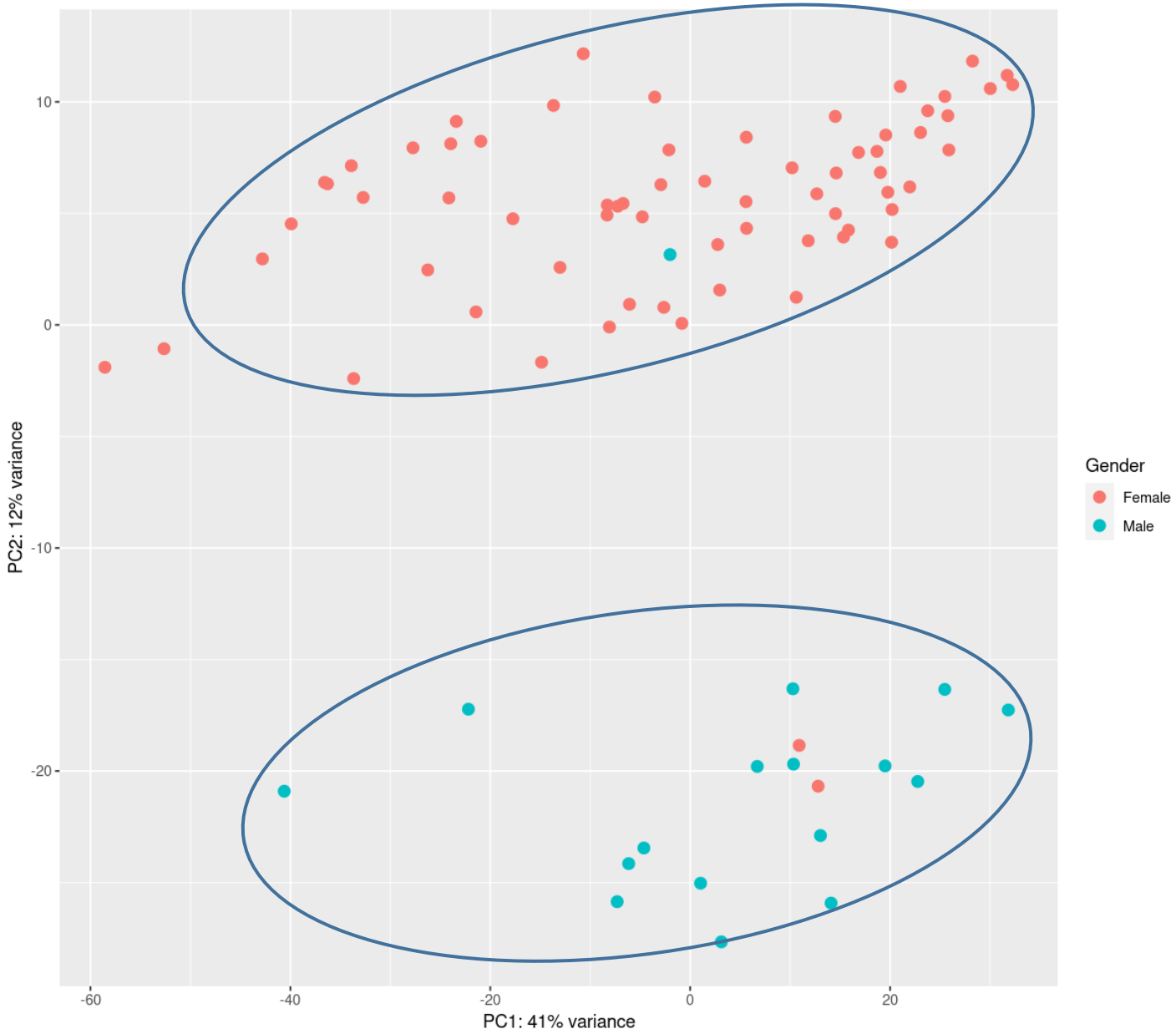
Principal Component Analysis

- Gene expression data of patients from group 1 or 2 (RA or undifferentiated RA)

- Gene expression data of the cohort



Clustering associated with sex



Ongoing Projects: Early Arthritis

Differential gene expression analysis

Comparisons of different diagnosis groups or outcome groups, including all samples and separately for each sex

- Comparisons between groups:
  - Diagnosis groups: gender as covariate
  - Outcome groups: gender and diagnosis as covariates
- Prefiltering: average gene count across samples > 5
- Significance threshold: Adjusted p-value < 0.05

Diagnosis	1	2	3	4
Female	46	41	5	8
Male	14	11	2	0

Diagnosis	1 vs 2			1 vs 3			1 vs 4			1,2 vs 3,4		
Sex	**All**	**Females**	**Males**	**All**	**Females**	**Males**	**All**	**Females**	**Males**	**All**	**Females**	**Males**
Total DEGs	**3**	**3**	**28**	**196**	**347**			**0**		**10**		

Ongoing Projects: Early Arthritis

Differential gene expression analysis

Comparisons of different diagnosis groups or outcome groups, including all samples and separately for each sex

- Comparisons between groups:
  - Diagnosis groups: gender as covariate
  - Outcome groups: gender and diagnosis as covariates
- Prefiltering: average gene count across samples > 5
- Significance threshold: Adjusted p-value < 0.05

Diagnosis	1	2	3	4
Female	46	41	5	8
Male	14	11	2	0

Diagnosis	1 vs 2			1 vs 3			1 vs 4			1,2 vs 3,4		
Sex	**All**	**Females**	**Males**	**All**	**Females**	**Males**	**All**	**Females**	**Males**	**All**	**Females**	**Males**
Total DEGs	**3**	**3**	**28**	**196**	**347**			**0**		**10**		

Outcome	Adverse (No/Yes)	Favorable (No/Yes)
Female	52/16	28/39
Male	14/3	5/12

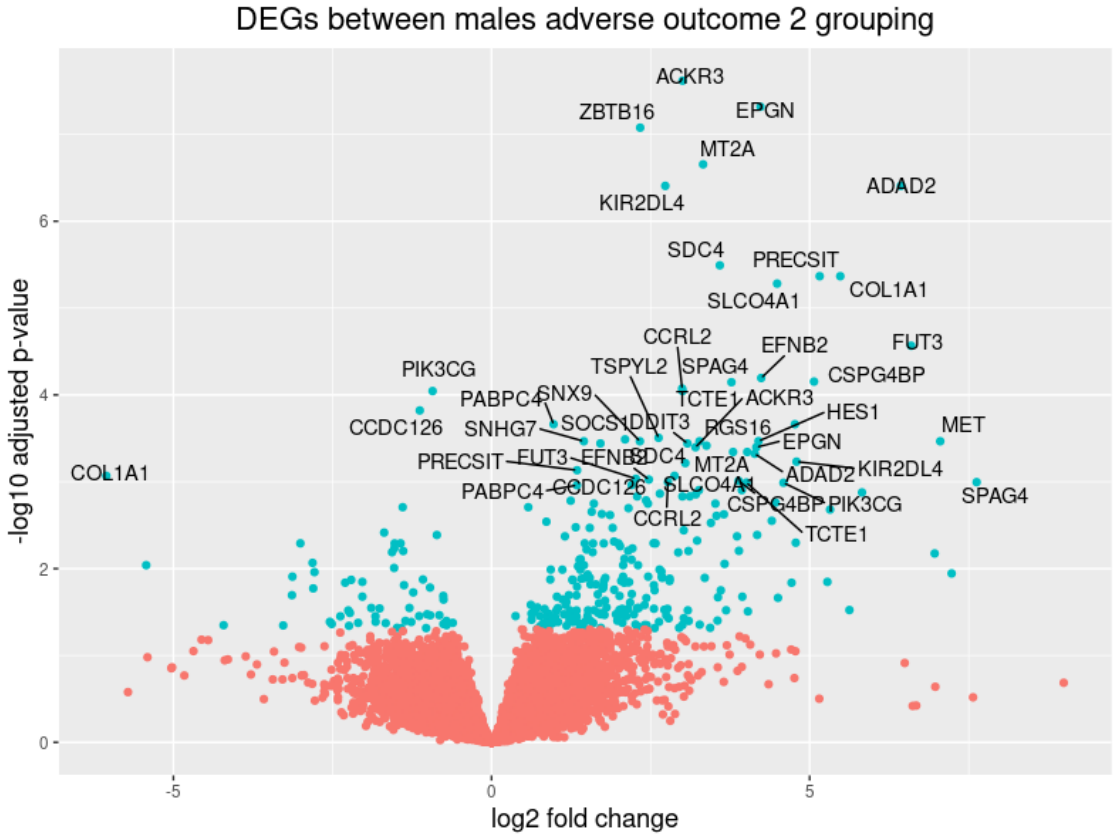
Outcome	Adverse			Favorable		
Sex	**All**	**Females**	**Males**	**All**	**Females**	**Males**
Total	**10**	**0**	**308**	**18**	**0**	**79**

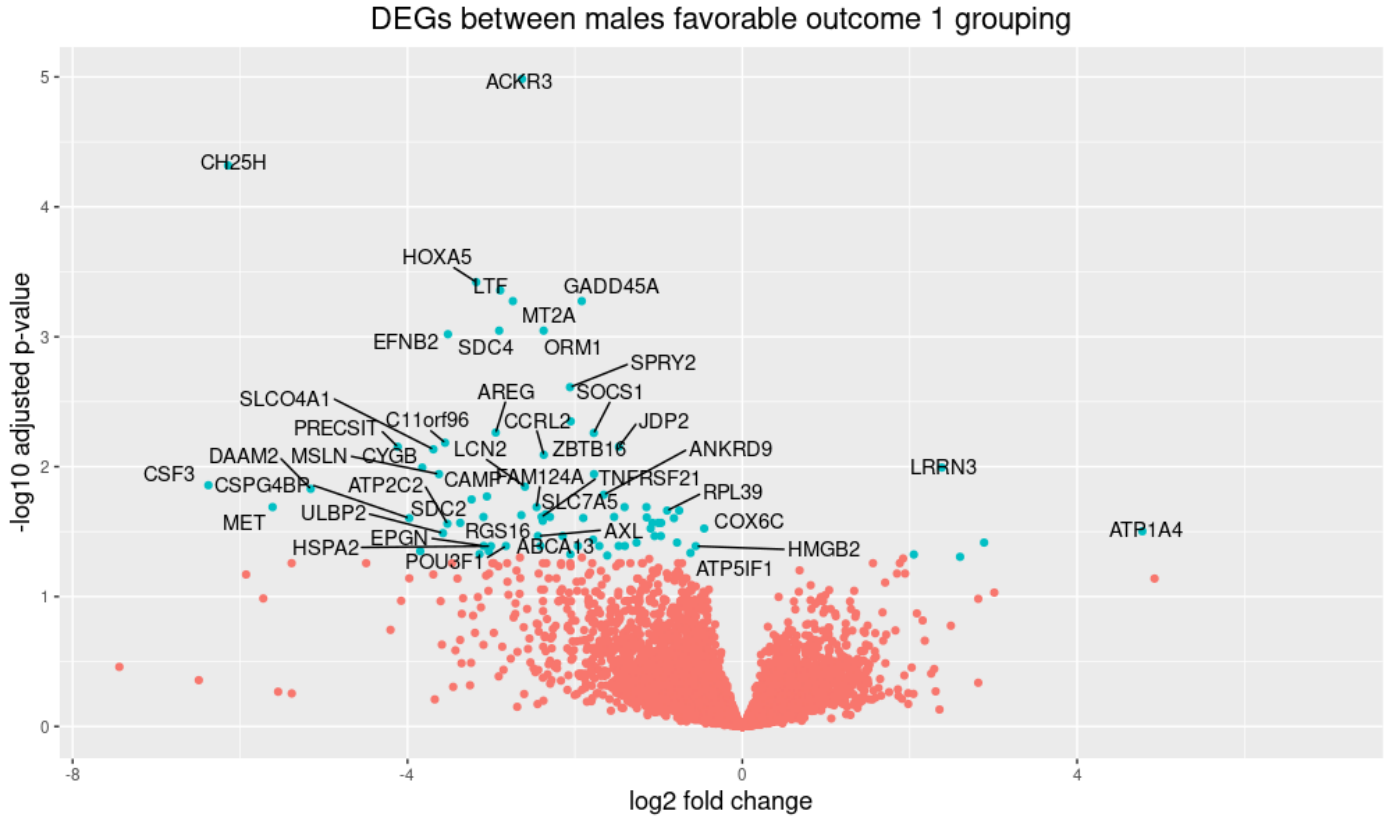
Ongoing Projects: Early Arthritis

Differential gene expression analysis in males

Adverse

Favorable





Pending Projects: Pre-SLE -funded by the FOREUM

- Experimental design

Autoantibody positive or First Degree Relatives of SLE

Whole Blood collection performed on <u>early disease state</u> (Baseline)

RNA sequencing

Monitoring prospectively over **5 years** for <u>possible transition to SLE</u>

Integration of demographic, clinical, serological, environmental

Aim of the study:

Definition of a high-risk signature for progression into SLE

	preSLE stable	preSLE progressors	Early SLE: Active, anti-DNA+	Early SLE: Active, anti DNA-	Early SLE: Active nephritis	Healthy Donors
# individuals	46	42	66	39	42	42

Pending Projects: By-stander gene effect

Link between gene expression and topology in the linear dimension

- **transcriptional activation** may spread in “waves” that **affect nearby genes**

Rapid induction of genes in response to growth factor stimulation accompanied by co-upregulation of their neighbouring genes

- significant proportion of gene expression events may be attributed to the **genomic position**

Relevance to SLE:

- extensive gene deregulation
- genetic complexity

S. Rennie,et al., Nat Commun, 2018

M. Ebisuya, Nat Cell Biol,2008

Pending Projects: By-stander gene effect

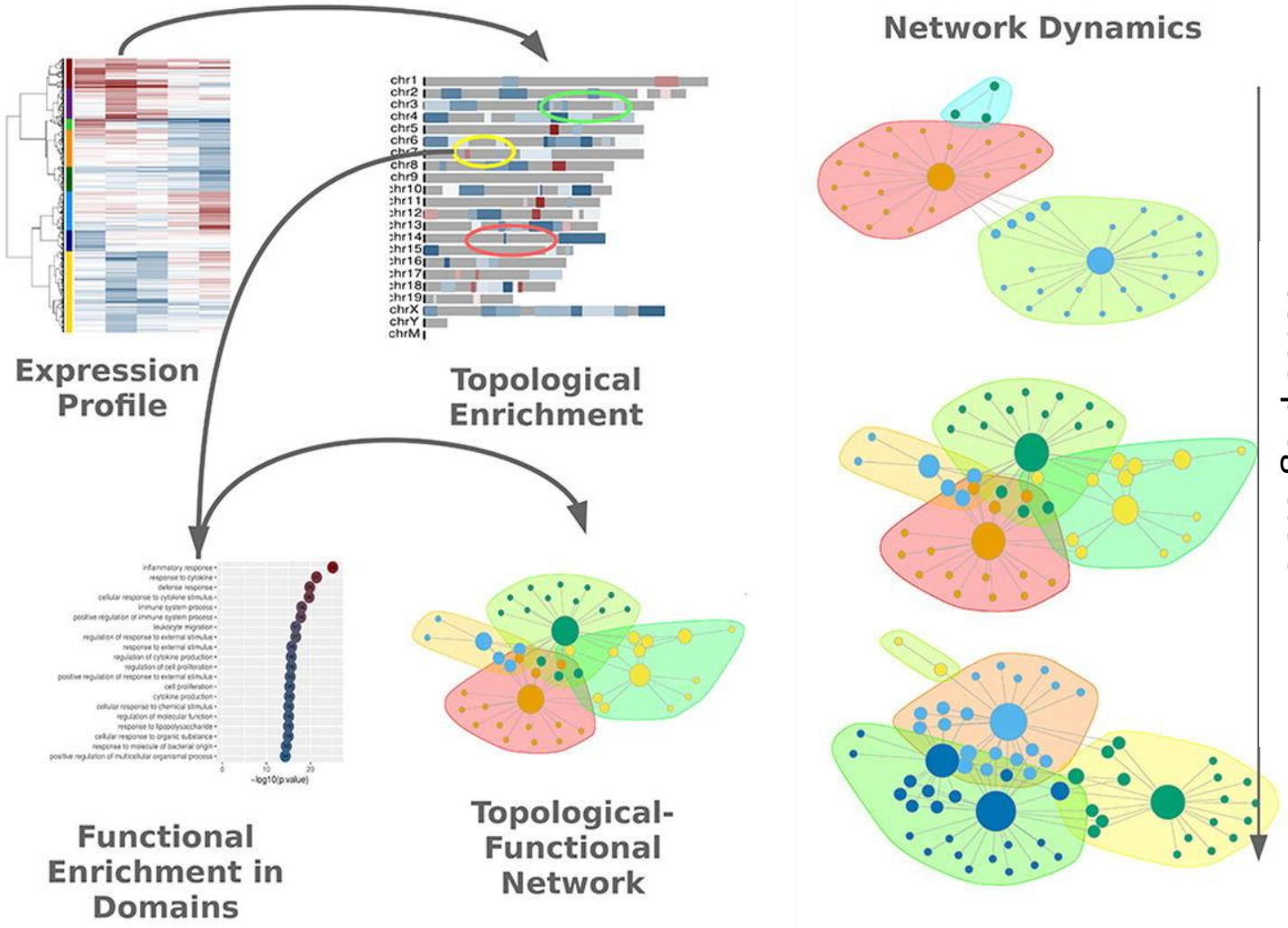
genome-wide expression profiles from SLE patients in **different disease states**

⬇

extended chromosomal regions with consistent patterns of differential gene expression

⬇

association with enriched functional pathways



By-stander genes: genes whose expression may be attributed to their relative position rather than their participation in a certain pathway.

Aim of the study:

Combination of topological and functional information into bipartite networks to draw conclusions on the way genome organization may underlie the gene regulation program during **SLE progression**

S. Mavropoulos et al., Comput. Struct. Biotechnol. J, 2020.

Laboratory of Rheumatology, Autoimmunity and Inflammation, Medical School, University of Crete (UoC)

Γεώργιος Μπερτσιάς

Γεωργάκης Σπύρος

Κοσμαρά Δέσποινα

Νικολέρη Δήμητρα

Σταθοπούλου Χρύσα

Πρόδρομος Σιδηρόπουλος

Γουτακόλη Γιώτα

Νεοφωτίστου Ελπίδα

Παπαδάκη Λίτσα

Computational Genomics Group "Alexander Fleming" BSRC

Χριστόφορος Νικολάου

Βακιρλής Νίκος

Κλωνιζάκης Αντώνης

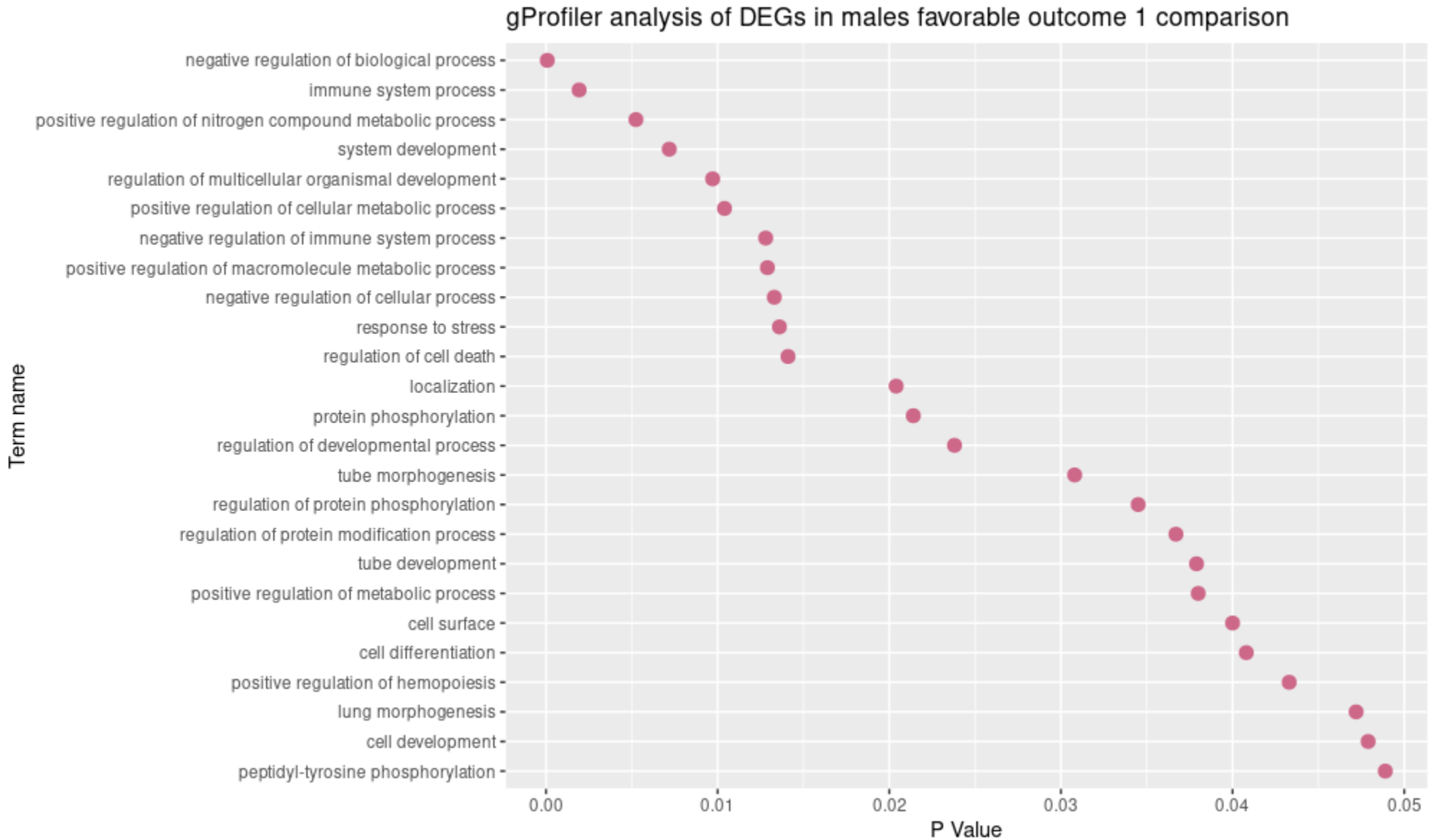
Σταυροπούλου Αθανασία

Τάσσιος Αιμίλιος

Single Cell Analysis Unit "Alexander Fleming" BSRC

Κωνσταντόπουλος Δημήτρης

Thank you


gProfiler analysis of DEGs in males favorable outcome 1 comparison
negative regulation of biological process
immune system process
positive regulation of nitrogen compound metabolic process
system development
regulation of multicellular organismal development
positive regulation of cellular metabolic process
negative regulation of immune system process
positive regulation of macromolecule metabolic process
negative regulation of cellular process
response to stress
regulation of cell death
localization
protein phosphorylation
regulation of developmental process
tube morphogenesis
regulation of protein phosphorylation
regulation of protein modification process
tube development
positive regulation of metabolic process
cell surface
cell differentiation
positive regulation of hemopoiesis
lung morphogenesis
cell development
peptidyl-tyrosine phosphorylation
Term name
0.00
0.01
0.02
0.03
0.04
0.05
P Value