

Discovery and characterization of clinically relevant lncRNAs in multiple cancer types

Graduate Student Seminar

Karin Isaev

Supervisor: Dr. Jüri Reimand

December 6th, 2017

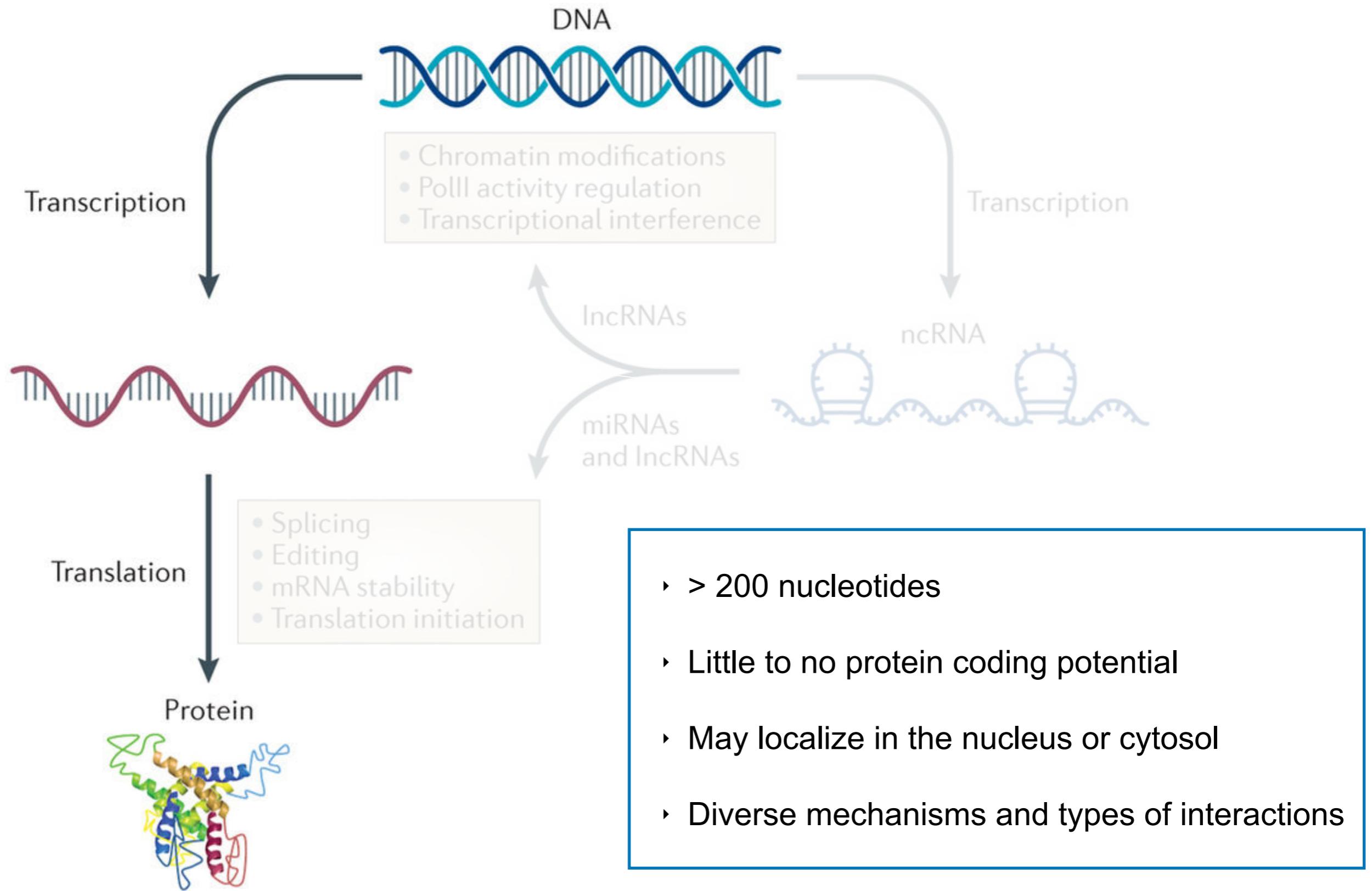
Outline

1. Introduction to lncRNAs
2. Aims and methods
3. Survival Analysis
4. Regulatory prediction

Outline

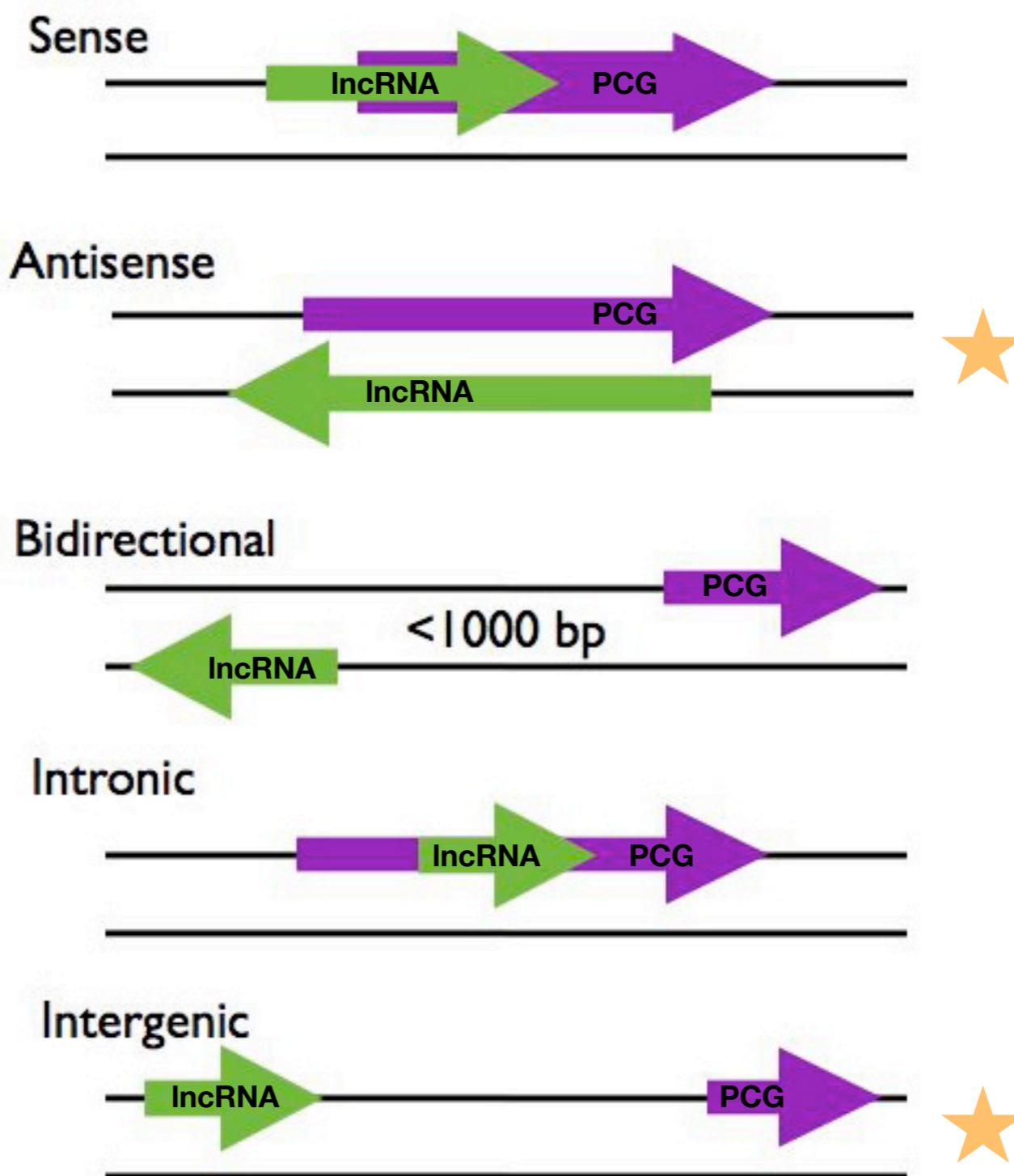
- 1. Introduction to lncRNAs**
- 2. Aims and methods**
- 3. Survival Analysis**
- 4. Regulatory prediction**

Central dogma in the context of lncRNAs



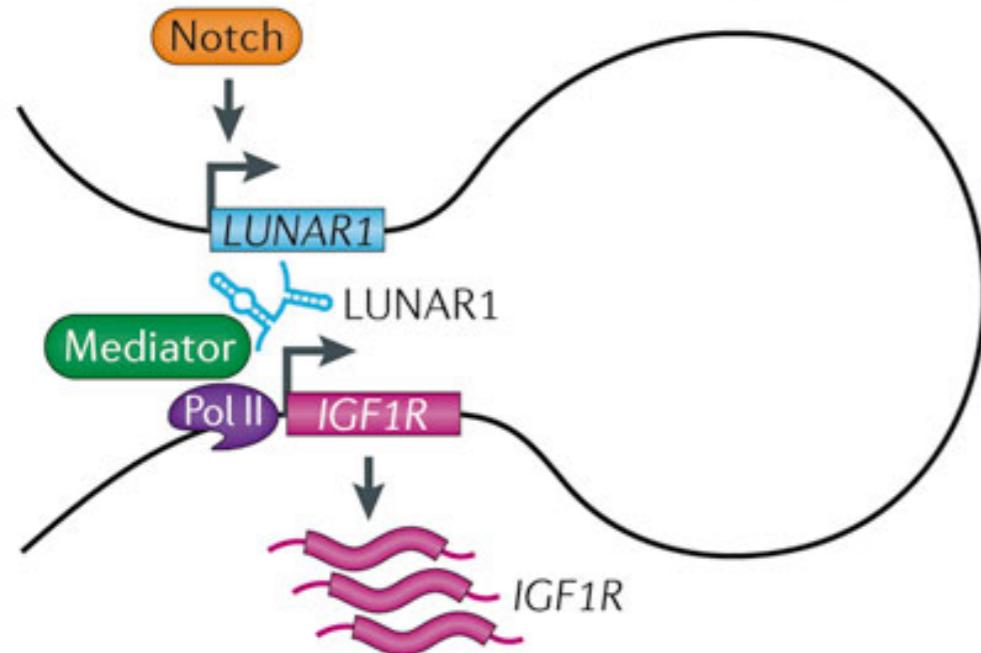
lncRNAs are grouped based on proximity to protein coding genes

* Protein Coding Gene (PCG)

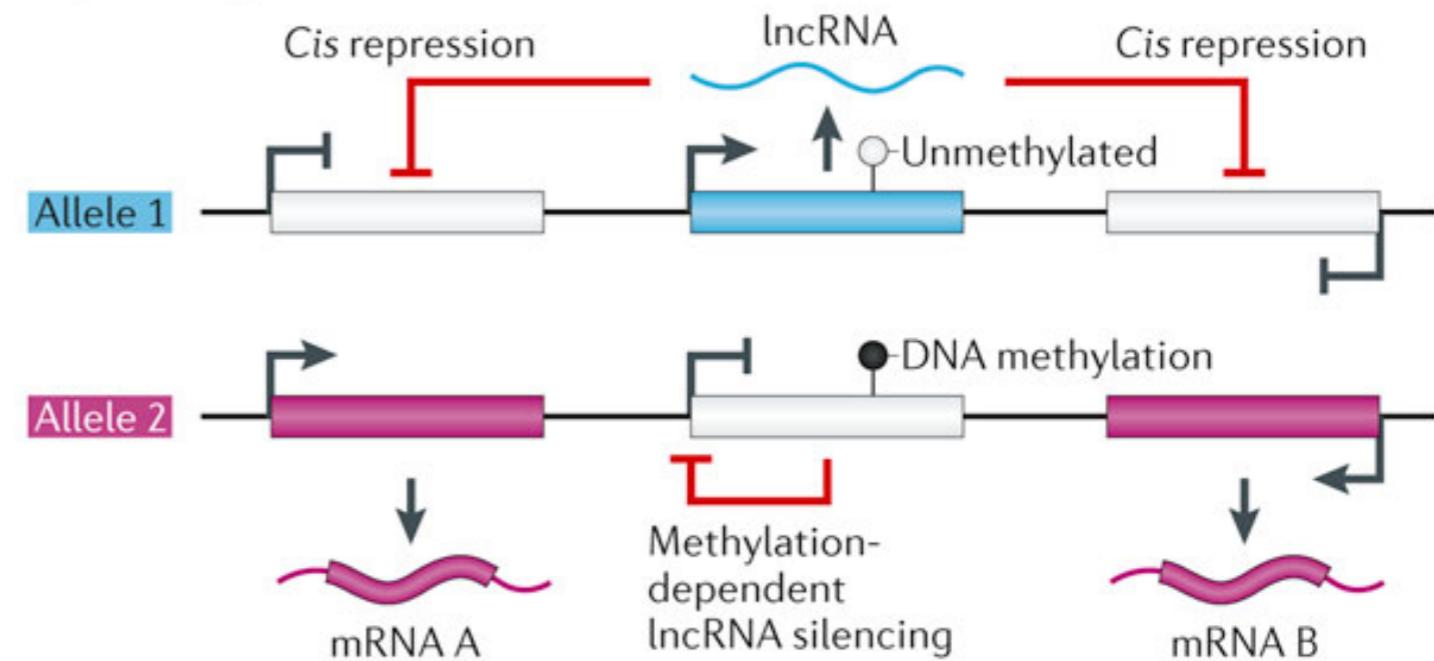


Cis-Regulatory Mechanisms of lncRNAs

a Enhancer RNAs and chromosome looping

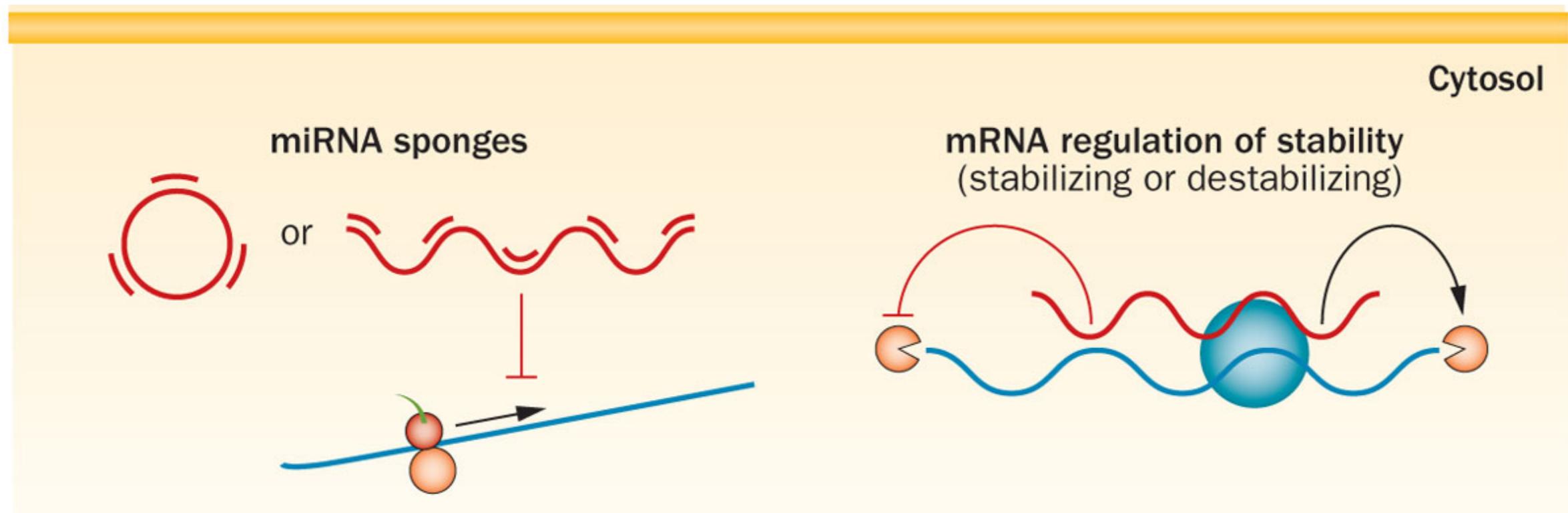


c Imprinted gene clusters



Well defined tissue specific mechanisms for only a small subset of lncRNAs

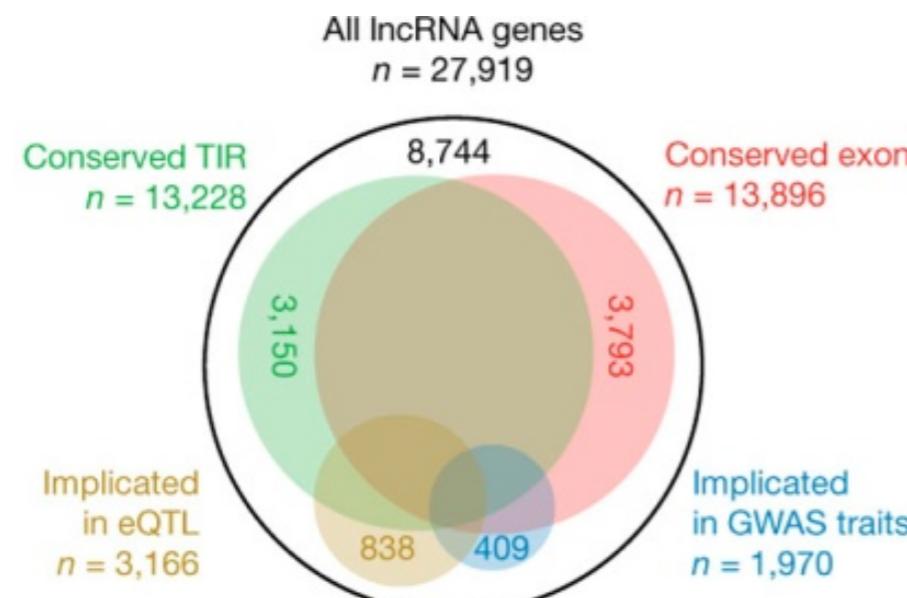
Trans-Regulatory Mechanisms of lncRNAs



Well defined tissue specific mechanisms for only a small subset of lncRNAs

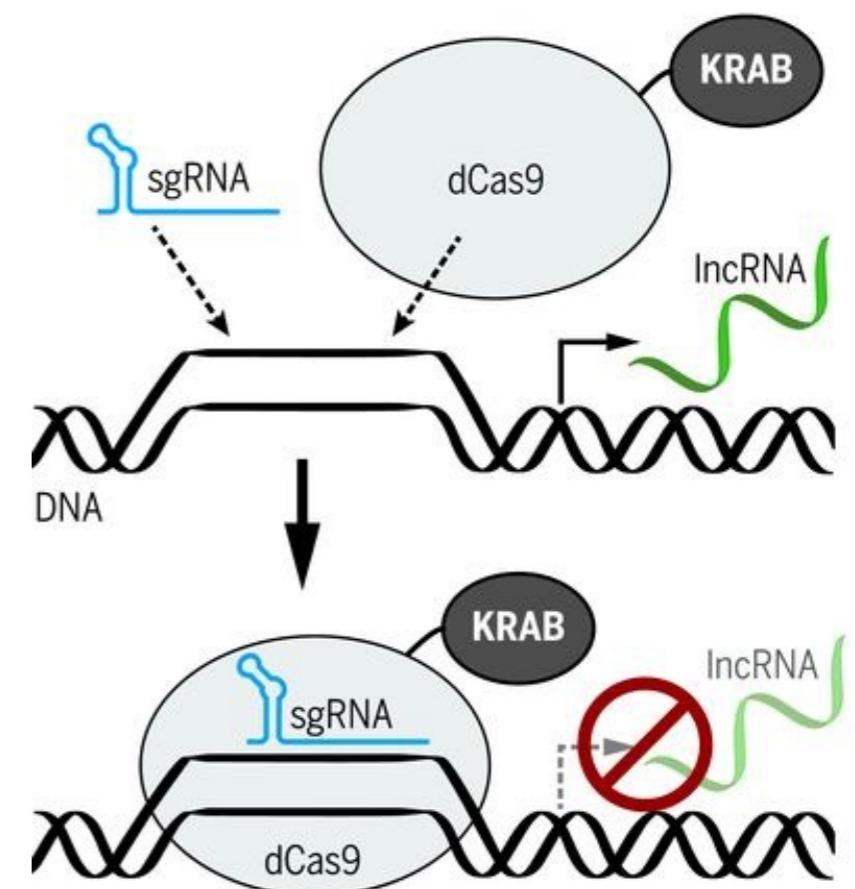
Functional evidence of human lncRNAs

~20,000 potentially functional lncRNAs listed in FANTOM CAT with annotated 5' ends

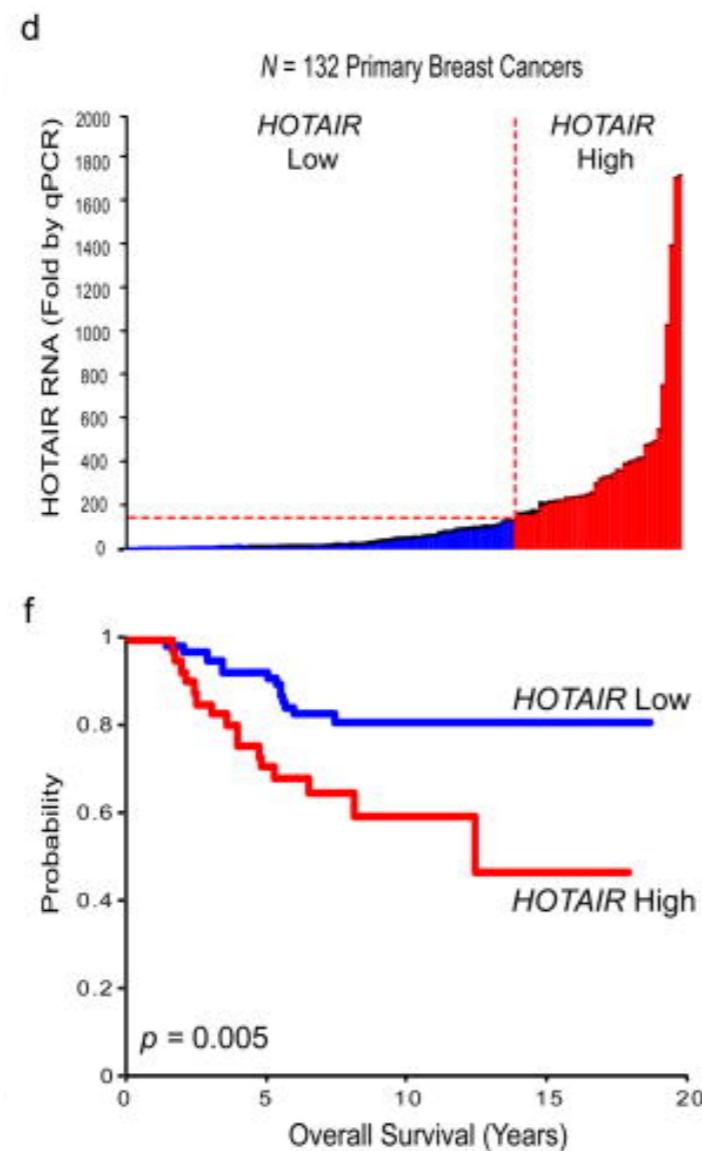


+

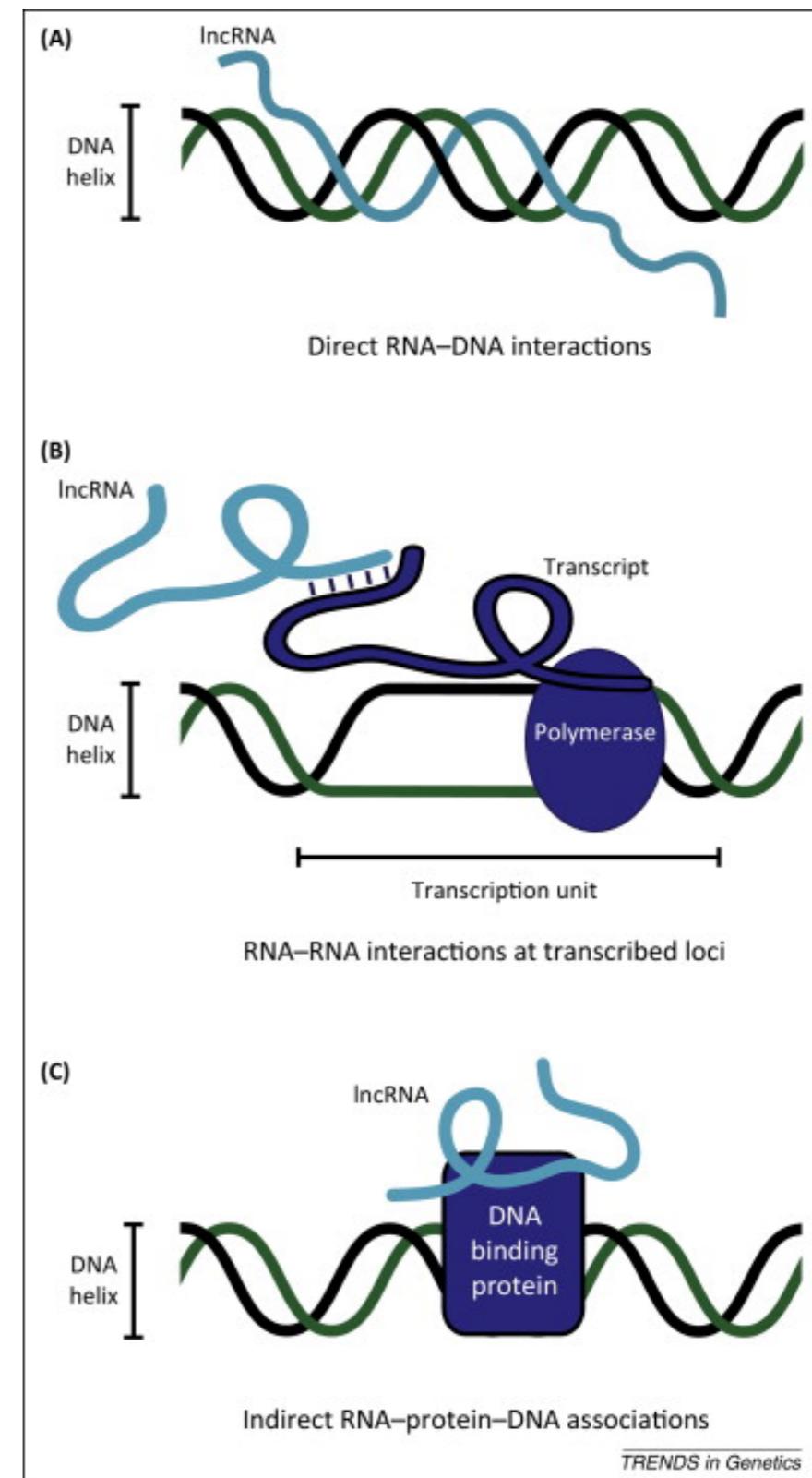
CRISPR interference (CRISPRi)



Rationale for studying lncRNAs in Cancer



- ♦ lncRNA tissue specificity greater than mRNAs
- ♦ Attractive targets for therapeutic intervention
- ♦ **Systematic screening of function and clinical relevance remains to be done**



Hypothesis and Aims

Hypothesis: lncRNA aberrations in cancer lead to the dysregulation of cancer pathways while driving lncRNA expression profiles that are informative of disease outcomes

Aim 1:

Identify lncRNA molecular profiles associated with survival in multiple cancer types



Aim 2:

Identify genetic and epigenetic aberrations associated with lncRNA molecular profiles



Aim 3:

Characterize lncRNA targets and regulators through association of lncRNA molecular profiles to pathways and genes



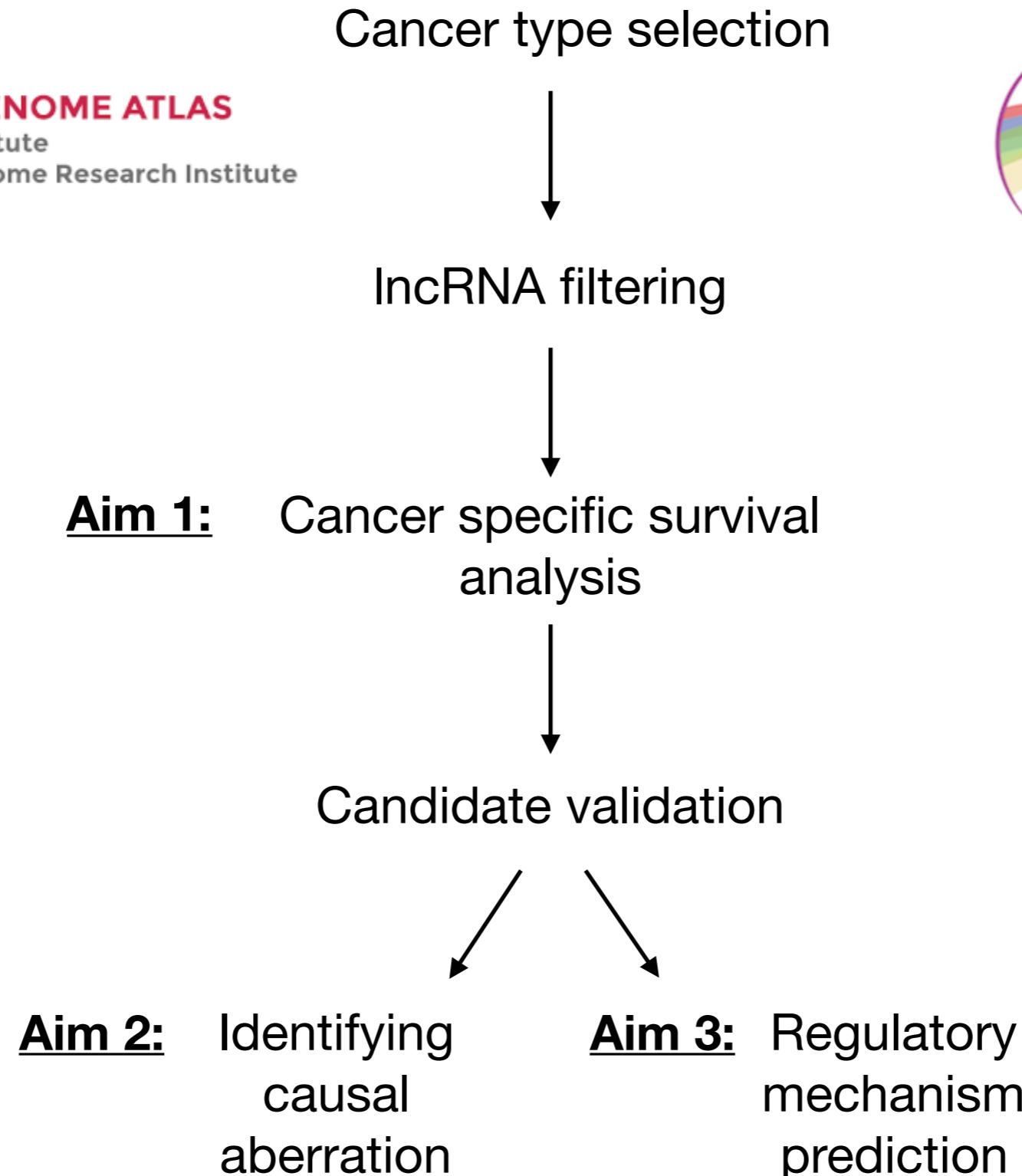
Outline

1. Introduction to lncRNAs
2. Aims and methods
3. Survival Analysis
4. Regulatory prediction

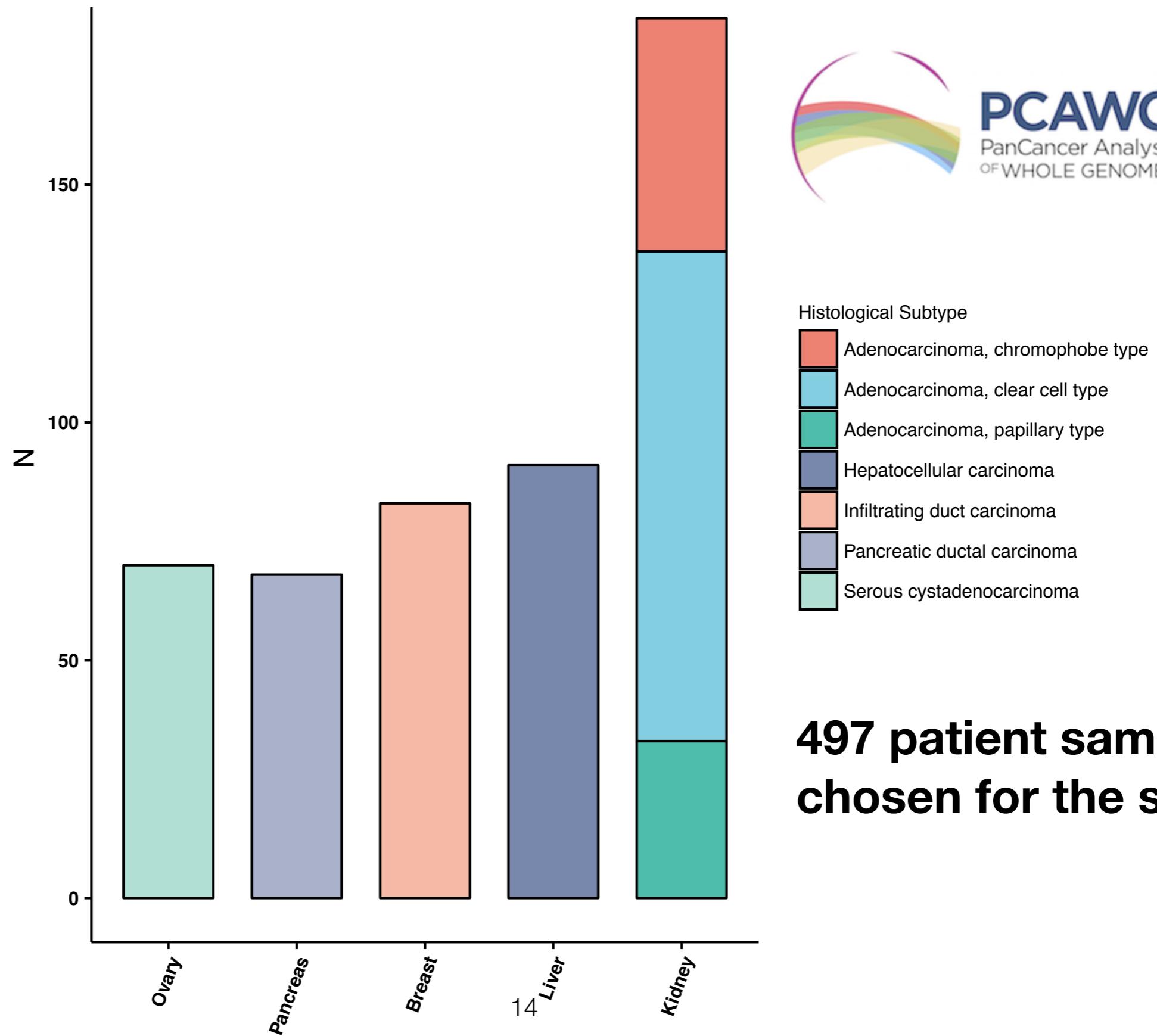
Outline

1. Introduction to lncRNAs
2. Aims and methods
3. Survival Analysis
4. Regulatory prediction

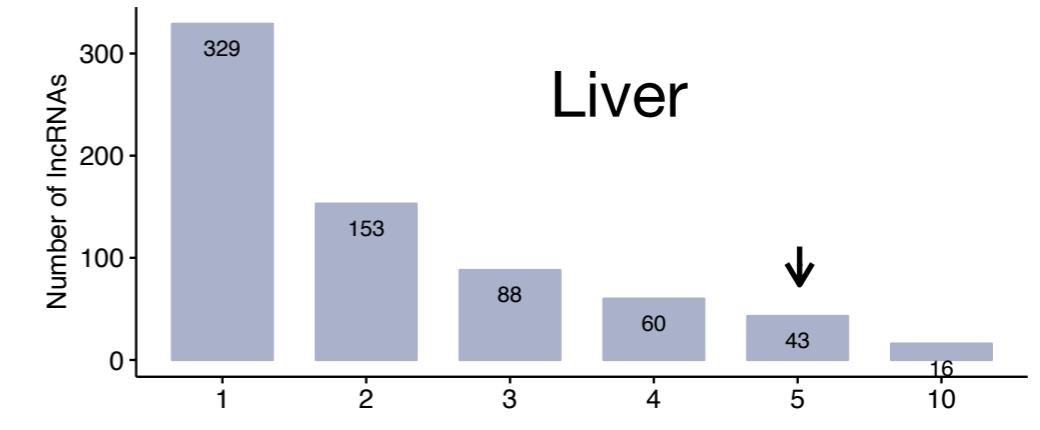
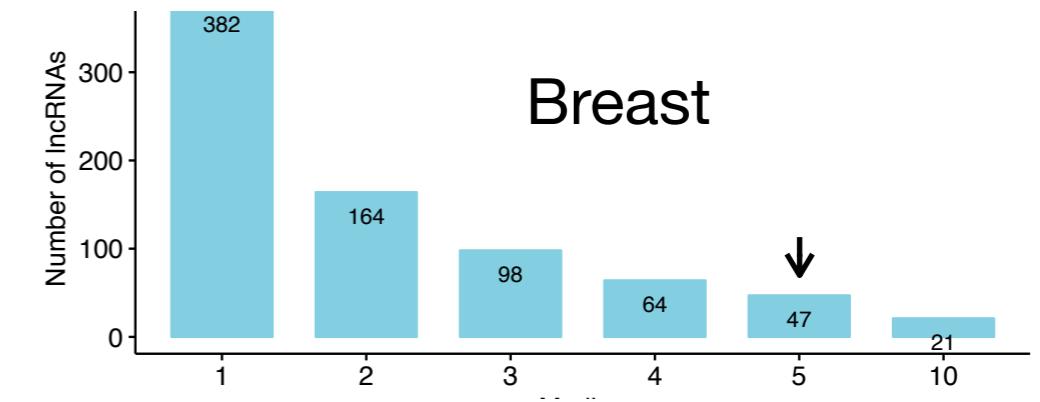
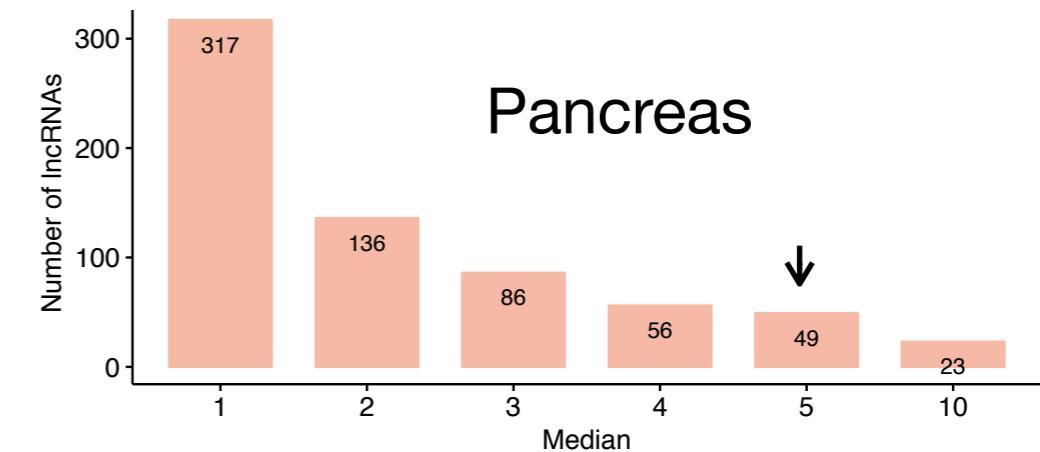
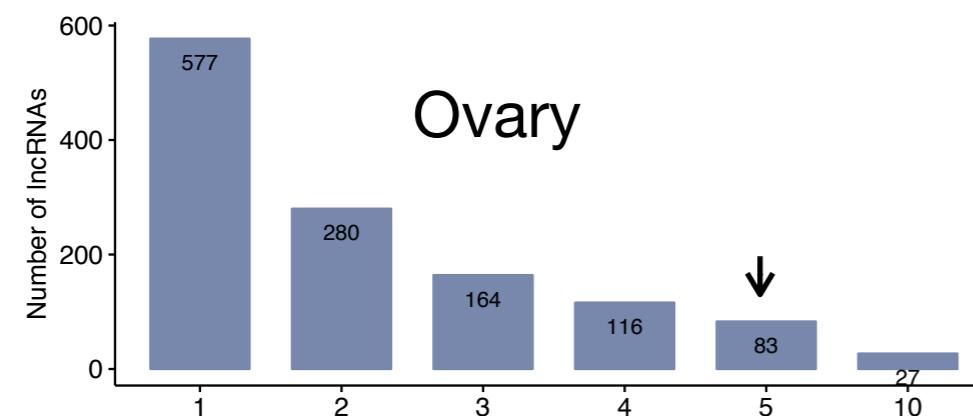
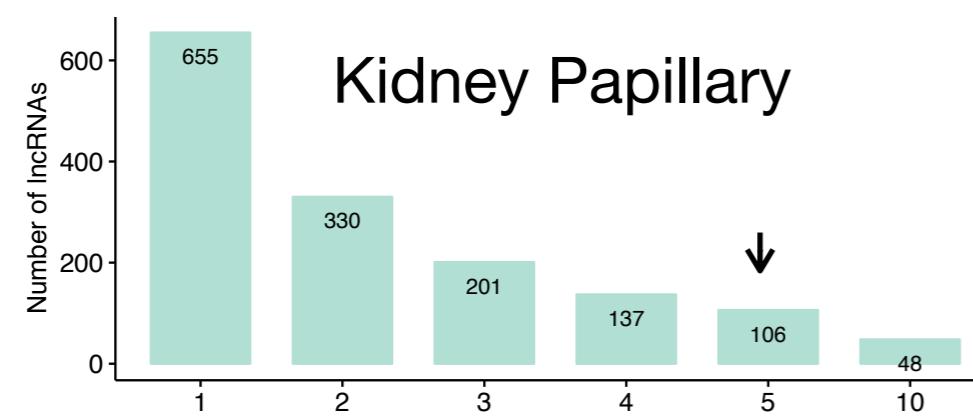
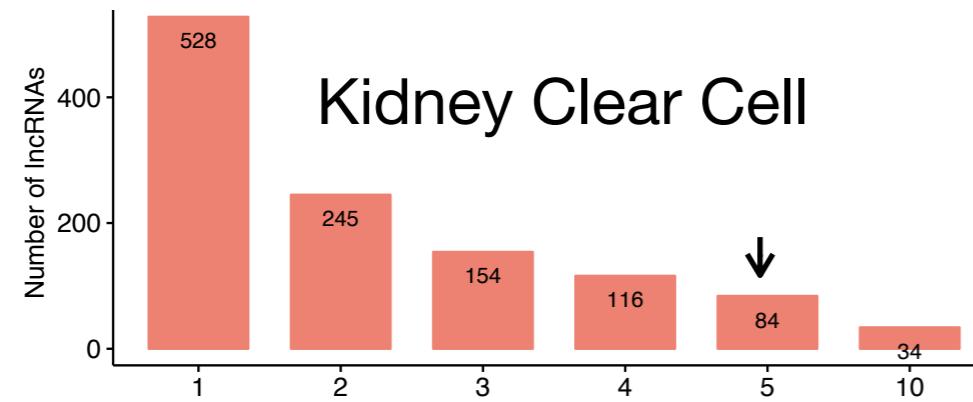
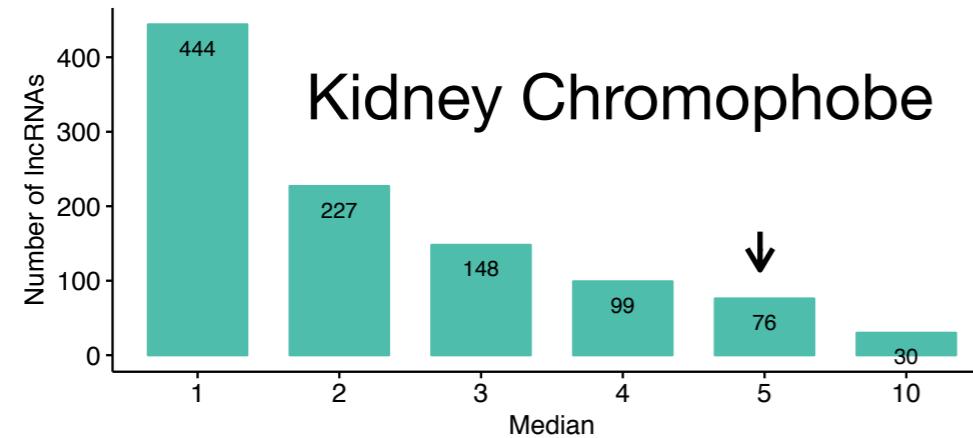
Methods



Cancer type selection



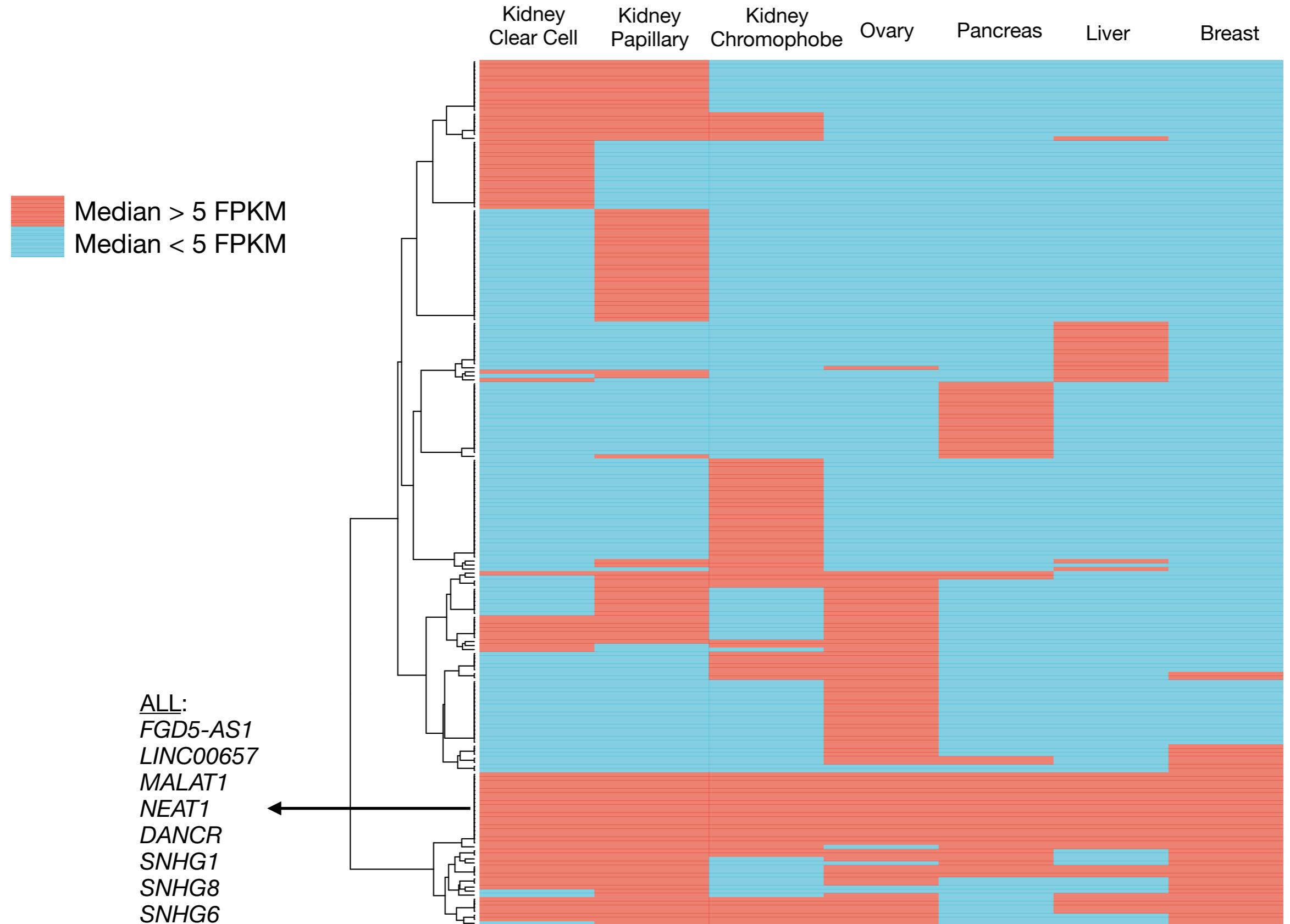
lncRNA Filtering using PCAWG RNA-Seq Data



Further:

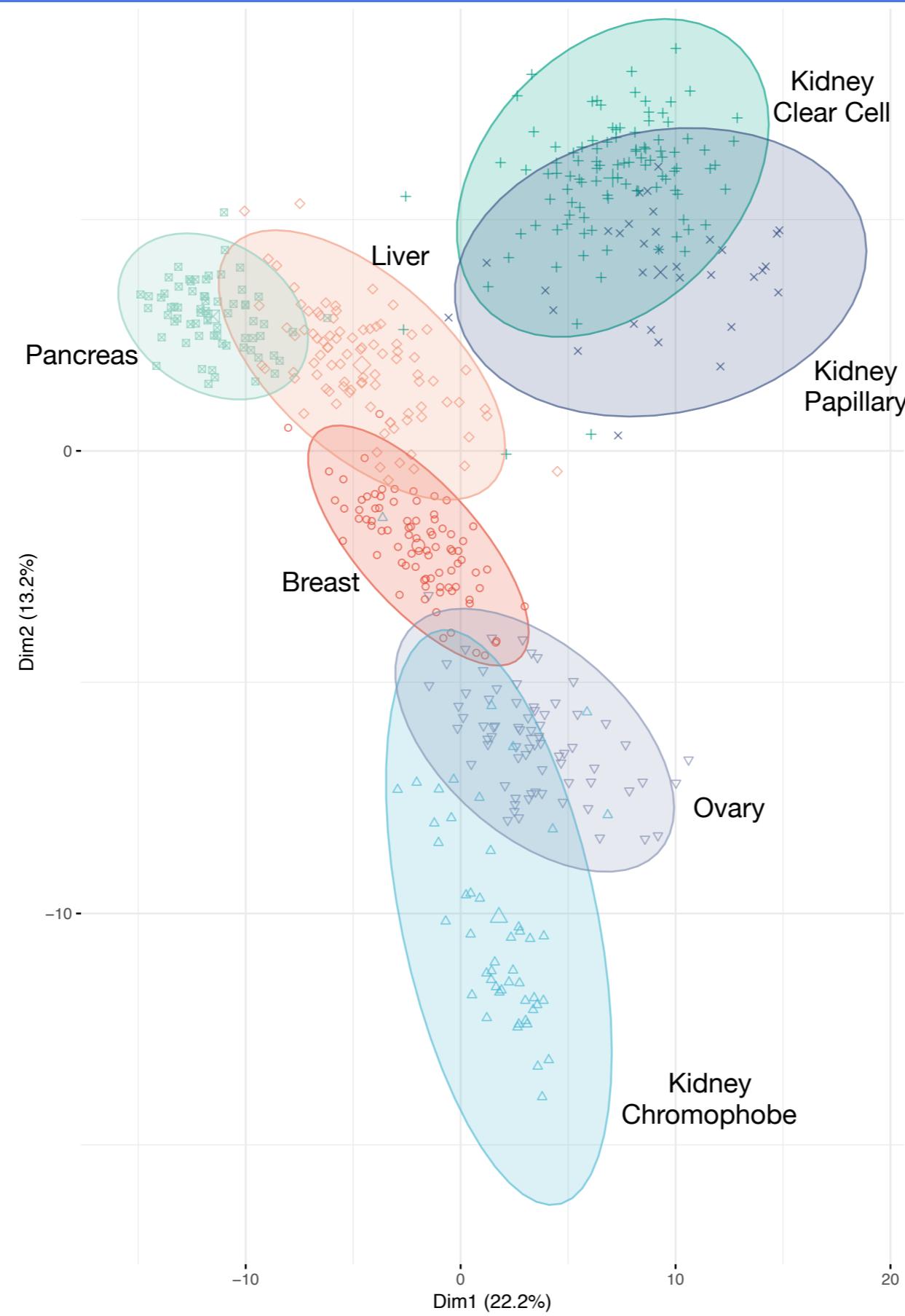
1. Filter lncRNAs not annotated in FANTOM CAT
2. Median cutoff = 5 FPKM

215 highly expressed lncRNAs across cancer types



ALL:
FGD5-AS1
LINC00657
MALAT1
NEAT1
DANCR
SNHG1
SNHG8
SNHG6

lncRNA gene expression is tissue specific



Outline

1. Introduction to lncRNAs
2. Aims and methods
3. Survival Analysis
4. Regulatory prediction

Outline

1. Introduction to lncRNAs
2. Aims and methods
- 3. Survival Analysis**
4. Regulatory prediction

Cancer Specific Survival Analysis

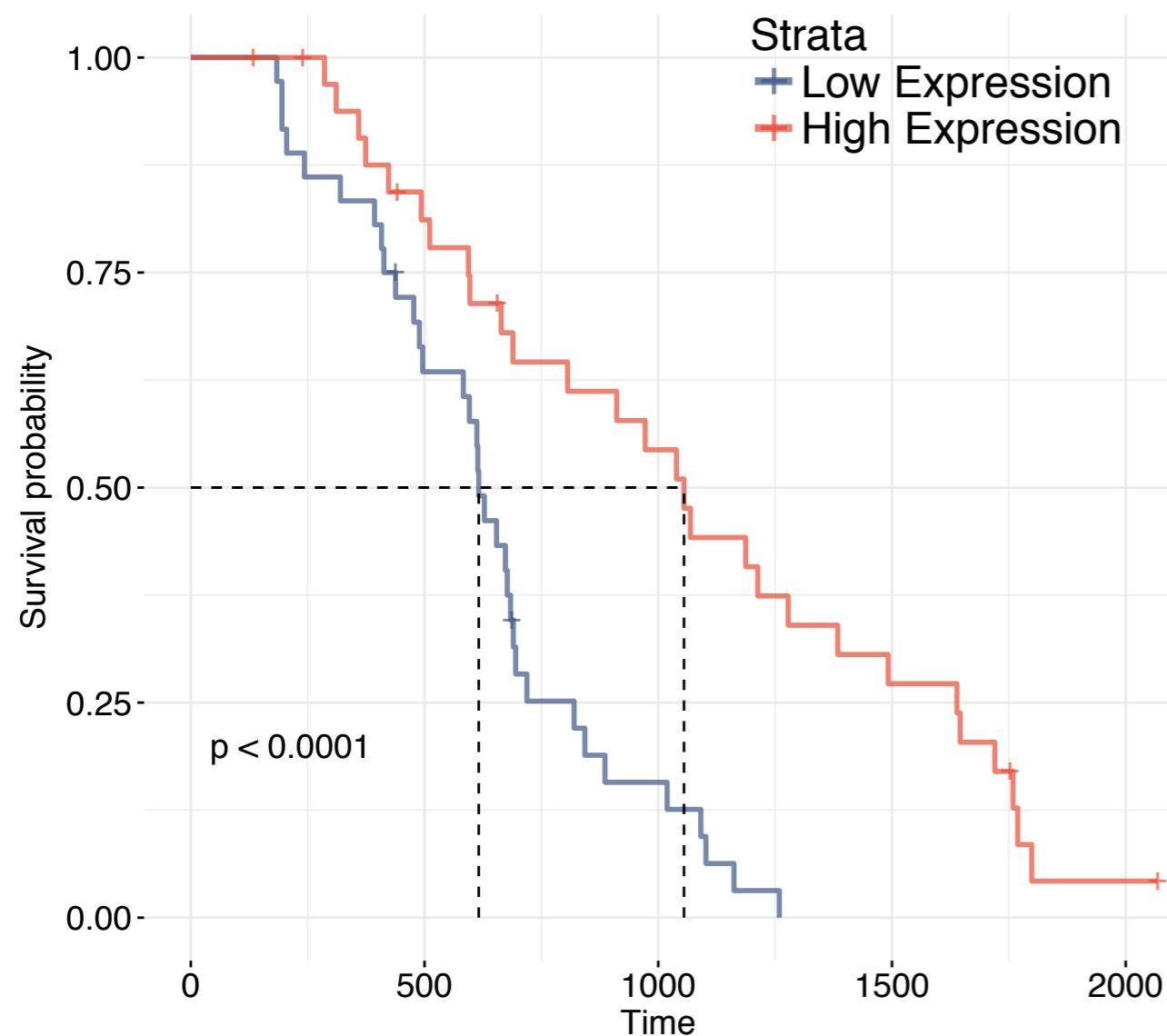
- ♦ 493 lncRNA-cancer associations tested:
 - 7 cancer types
 - 215 unique lncRNAs
- ♦ lncRNA Mean expression used to dichotomize patients into high and low expressing groups
- ♦ Cox proportional-hazards model
- ♦ Multiple testing correction within each cancer type
- ♦ Validation of association in TCGA

Cancer Specific Survival Analysis

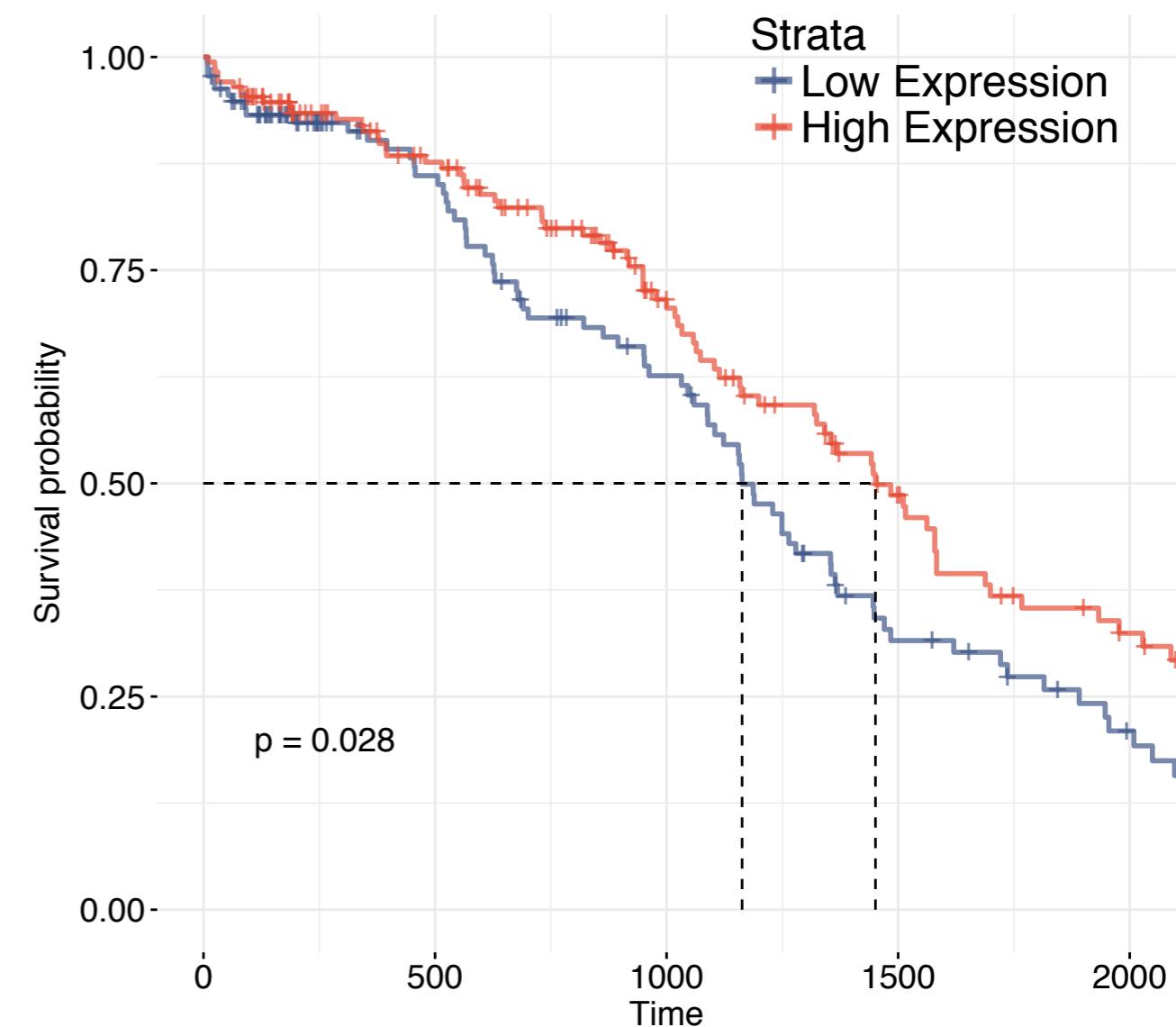
- ◆ 5 lncRNAs significantly associated with survival outcome (FDR < 0.05)
- ◆ 3 in Ovary Serous cystadenocarcinoma
- ◆ 2 in Liver Hepatocellular carcinoma

ZNF503-AS2 Ovarian Cancer

PCAWG, n = 70

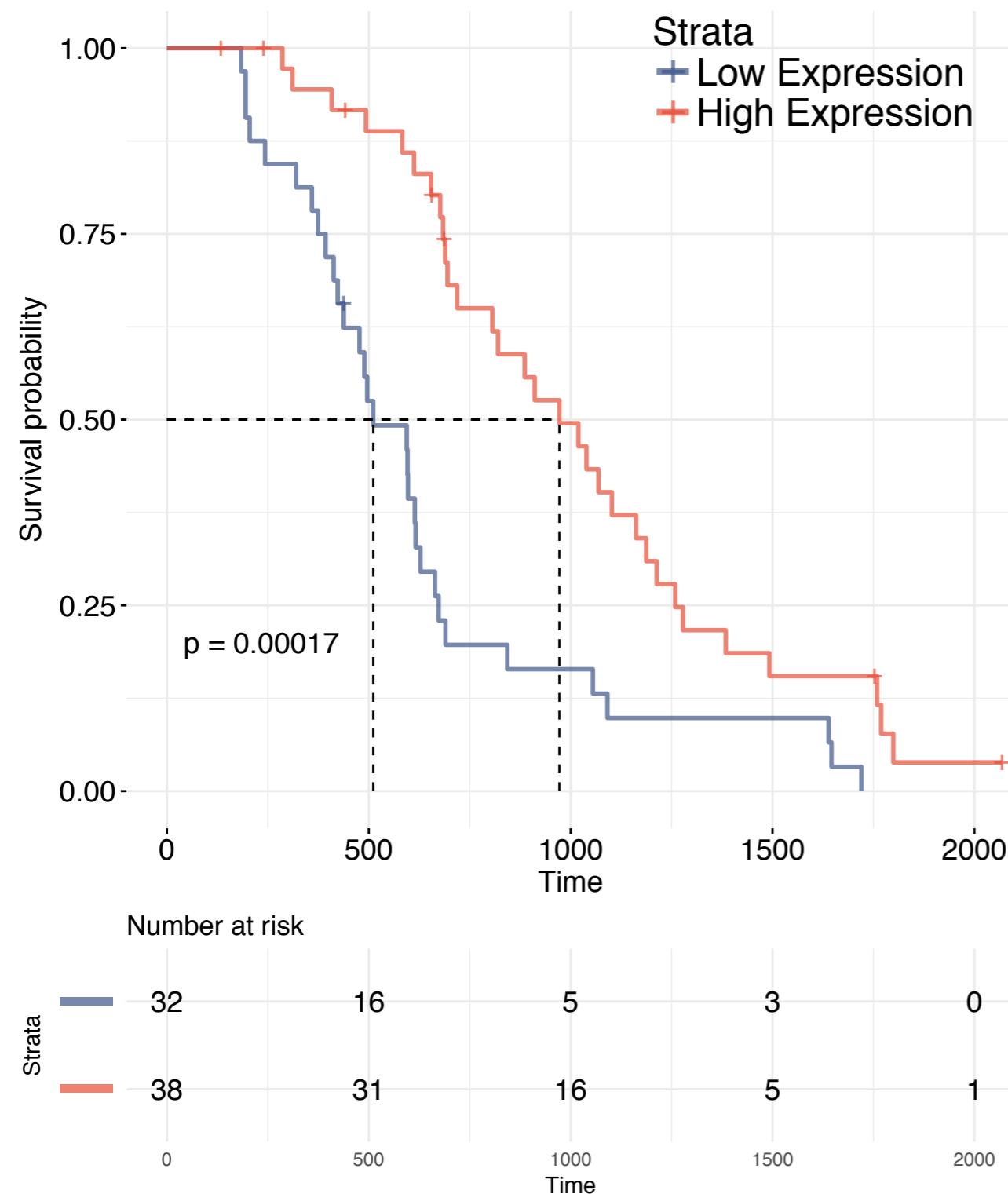


TCGA, n = 305

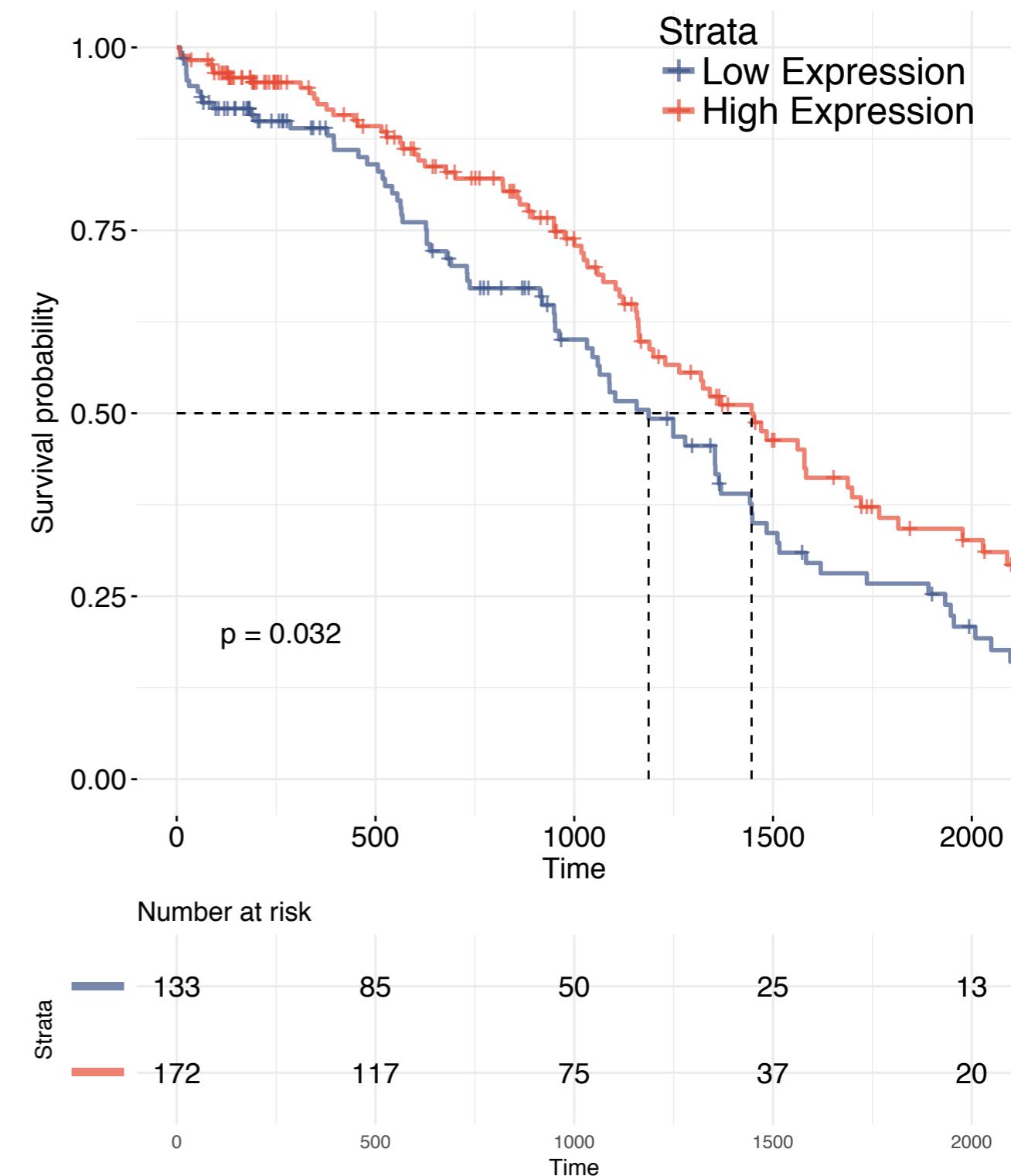


AC009336.24 Ovarian Cancer

PCAWG, n = 70

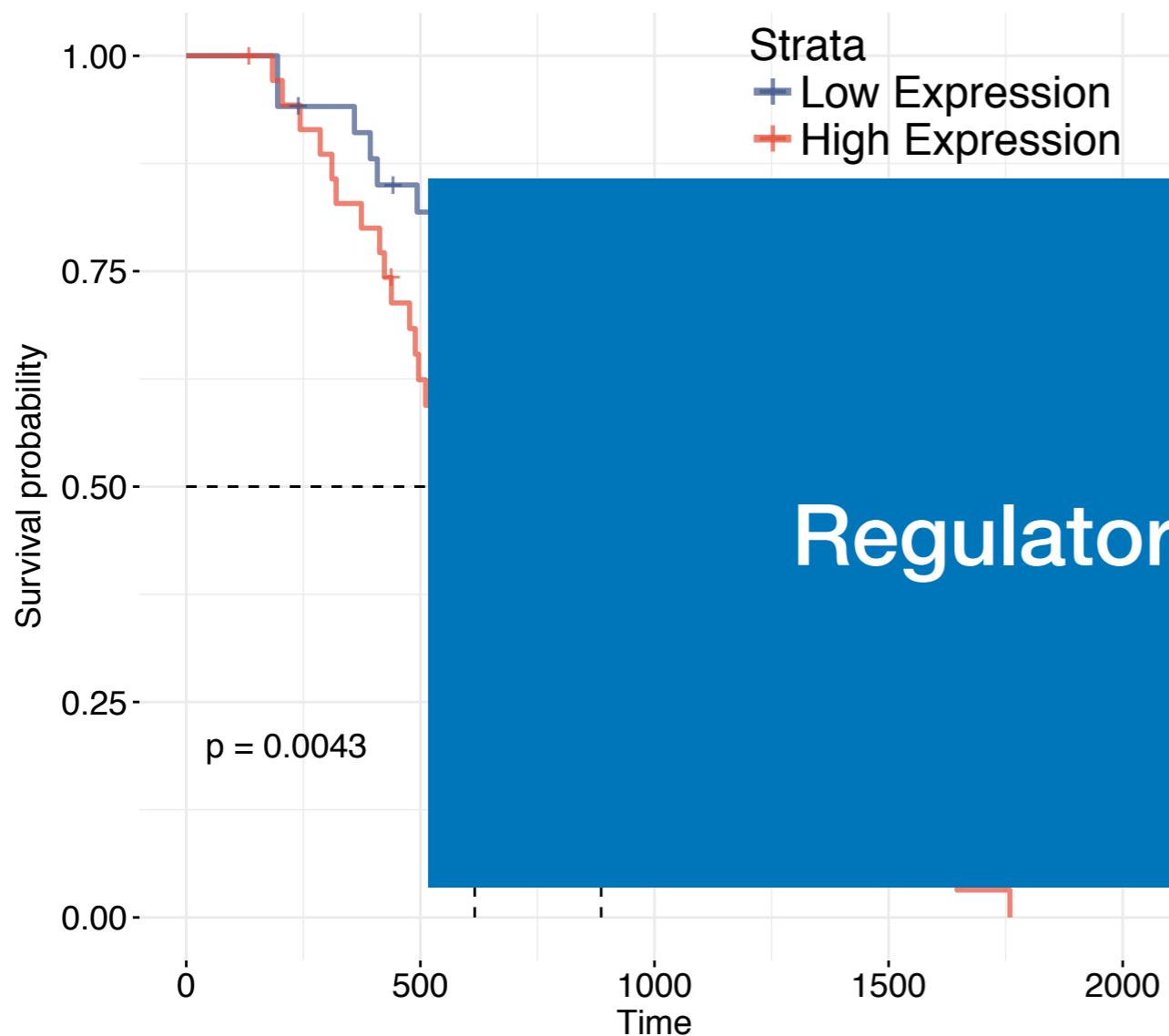


TCGA, n = 305

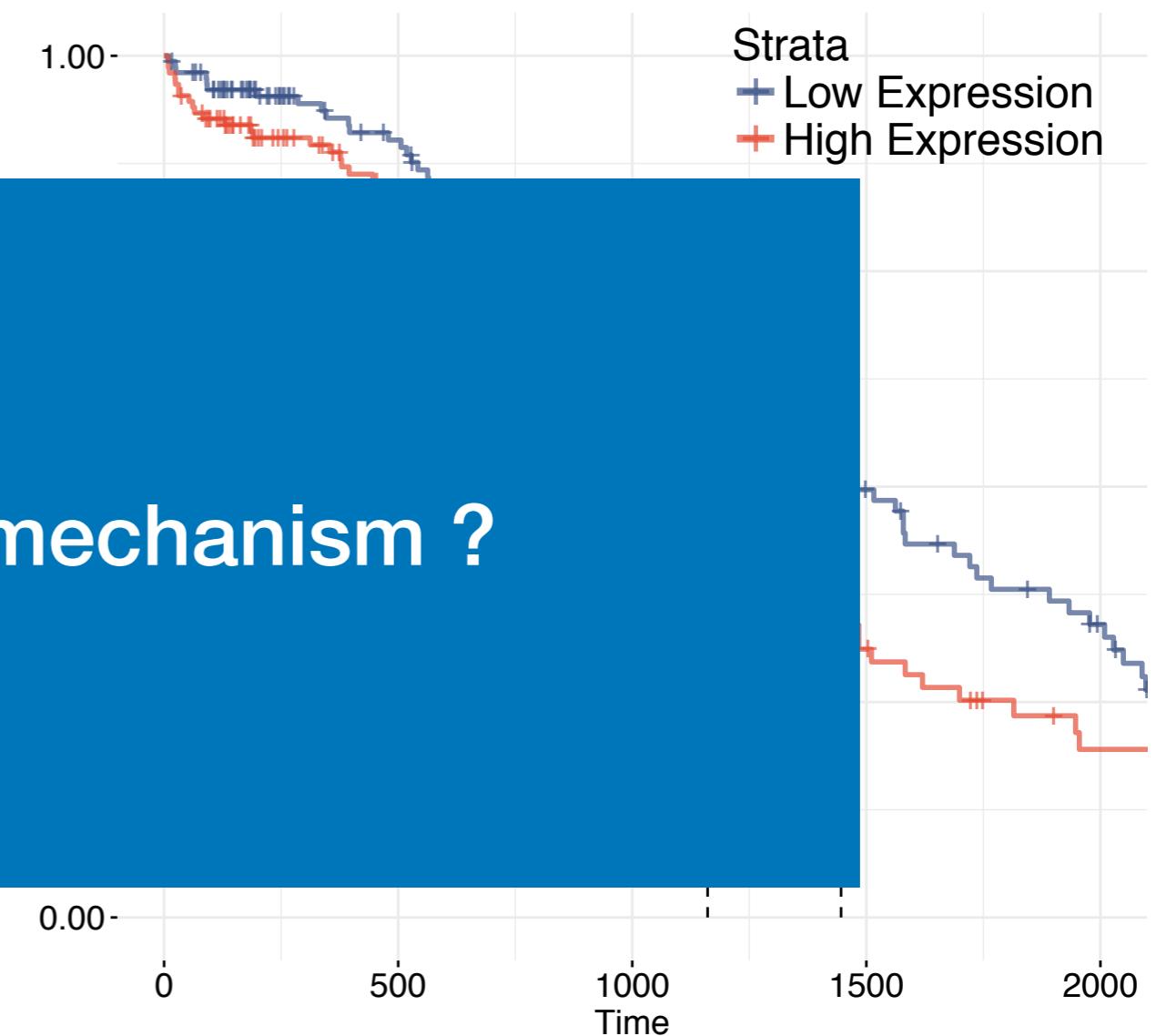


OTUD6B-AS1 Ovarian Cancer

PCAWG, n = 70



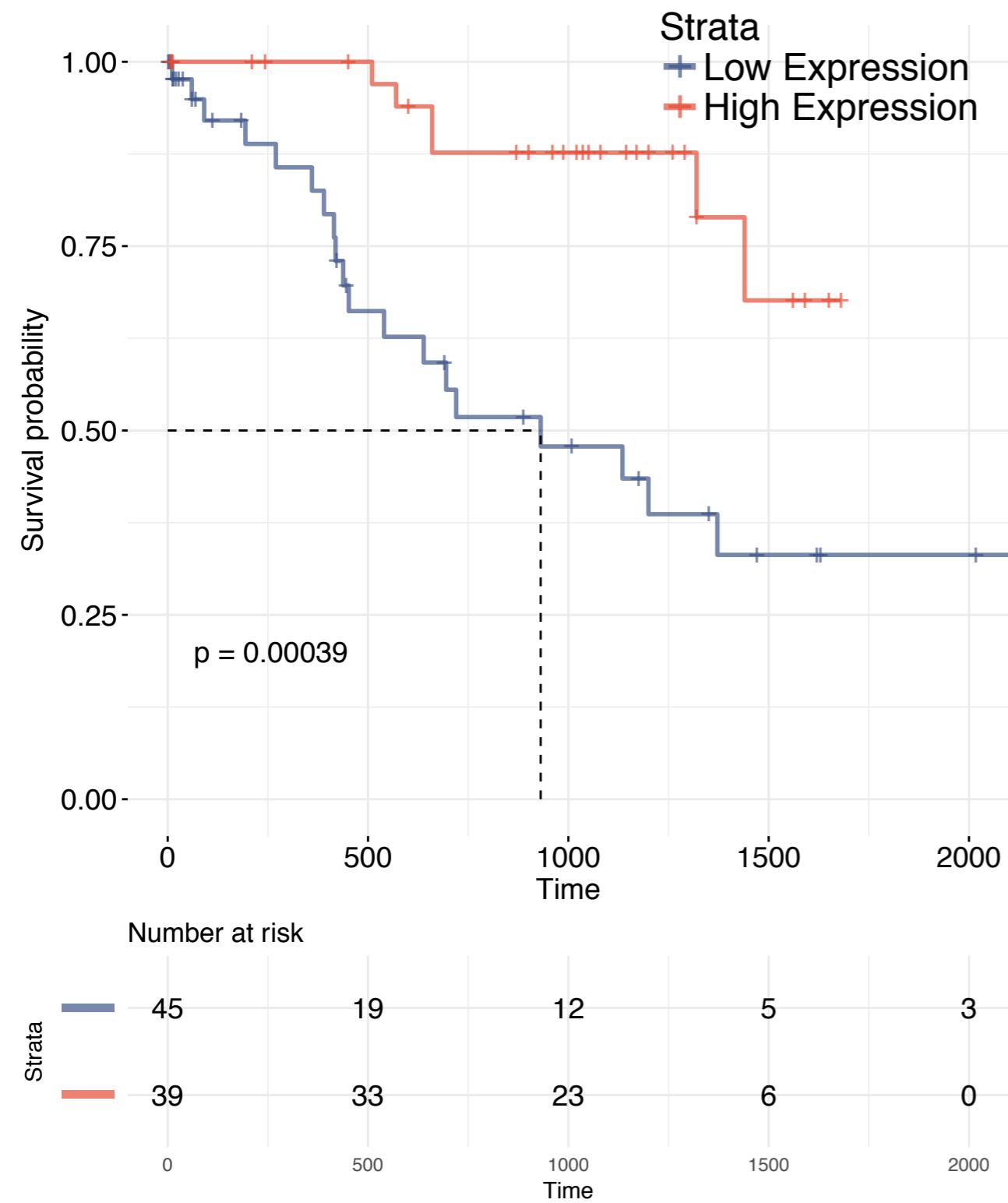
TCGA, n = 305



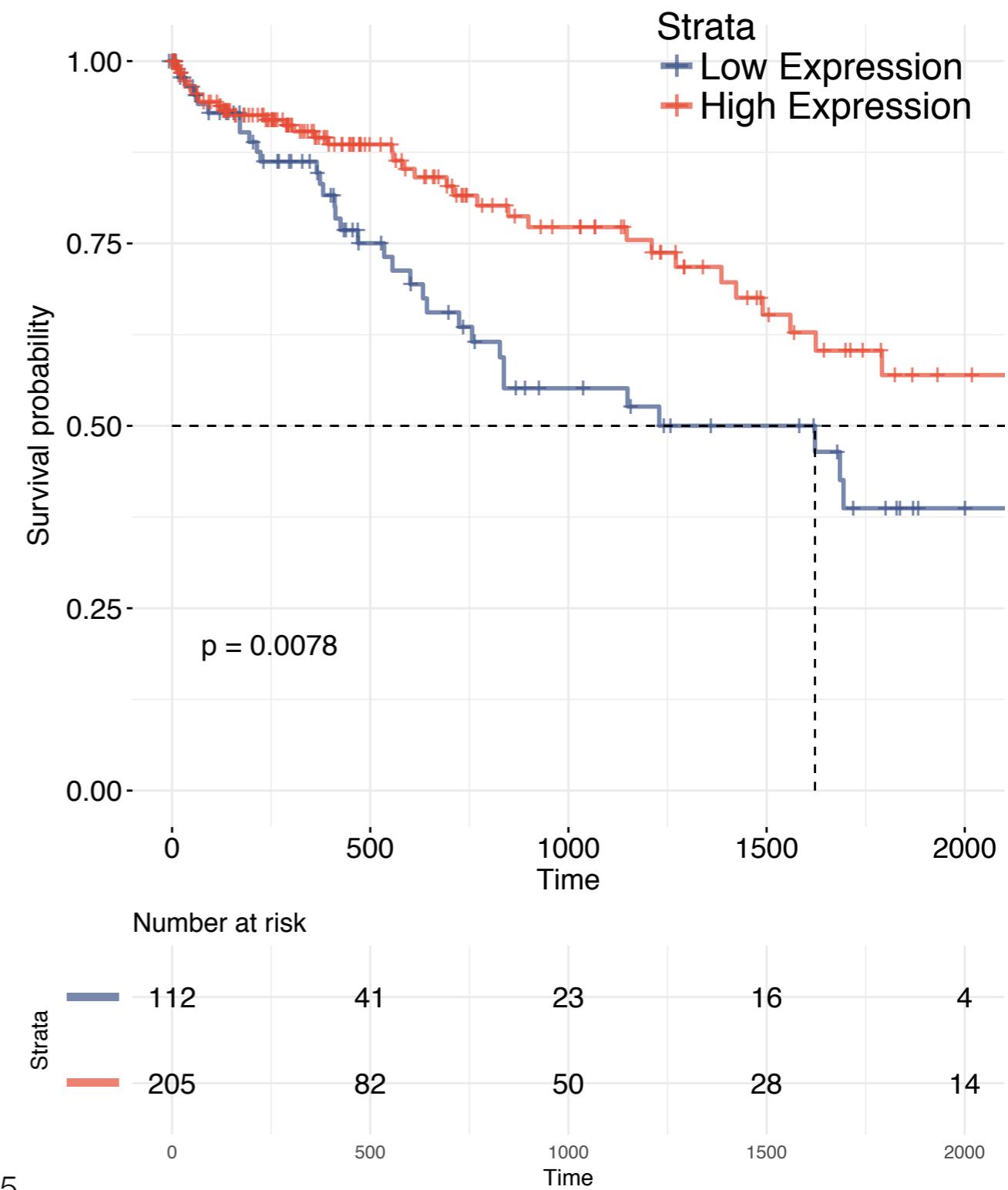
Regulatory mechanism ?

RP11-622A1.2 Liver Cancer

PCAWG, n = 84

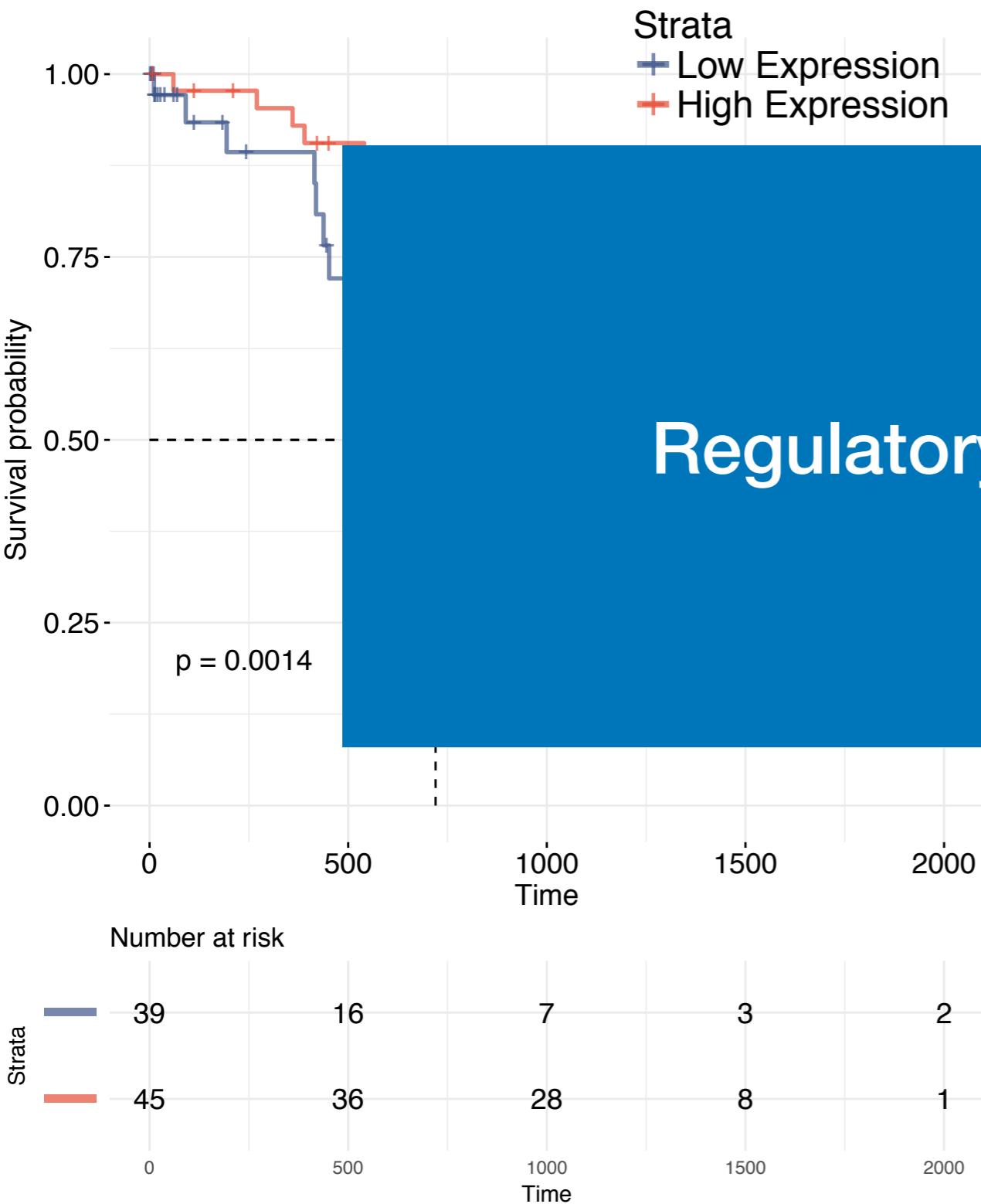


TCGA, n = 317

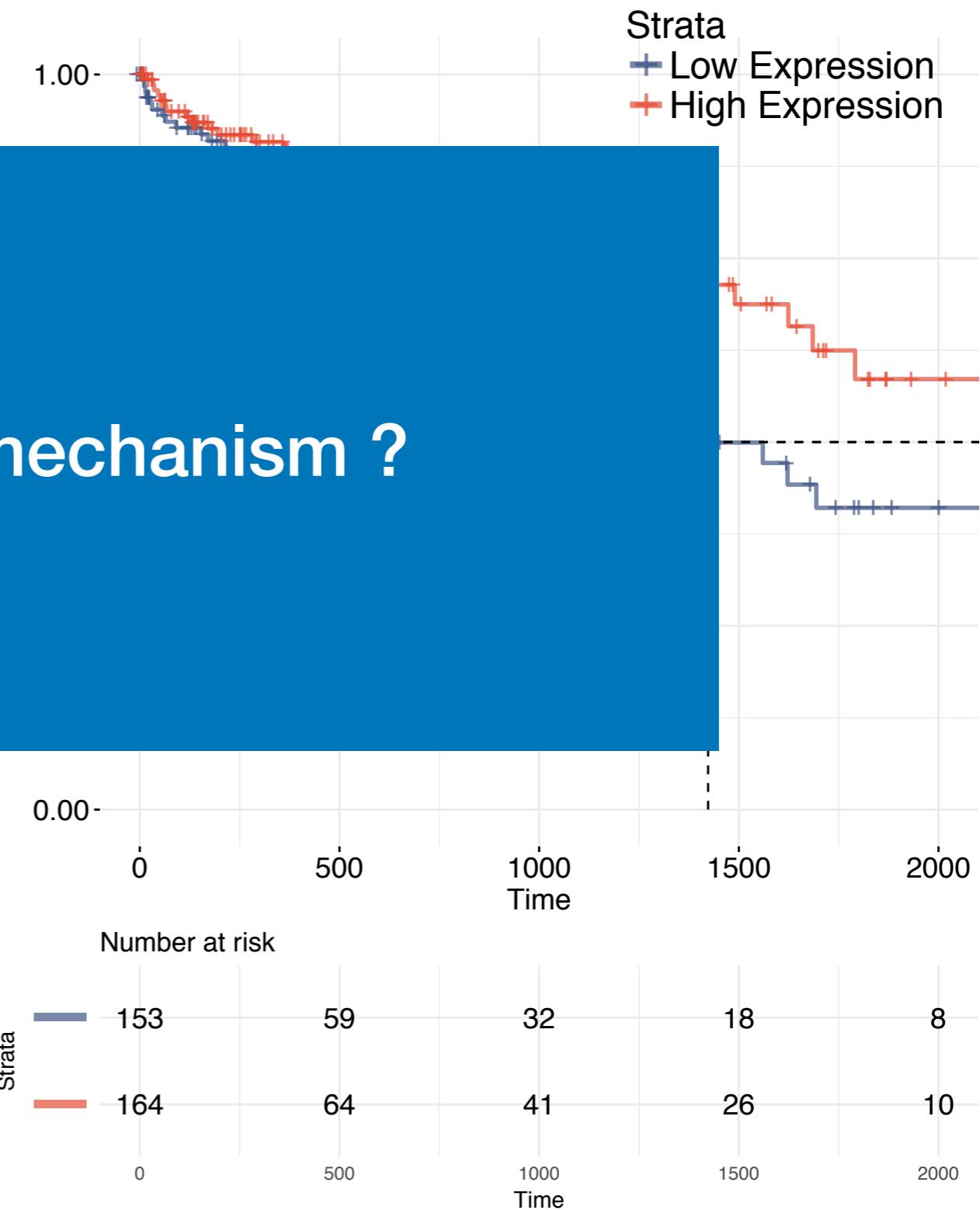


NEAT1 Liver Cancer

PCAWG, n = 84



TCGA, n = 317



Regulatory mechanism ?

lncRNA Expression in Normal Tissues (GTEx)

- ♦ To be a meaningful prognostic marker, lncRNA expression should vary between normal and tumour tissues

lncRNA Expression in Normal Tissues (GTEx)

Within PCAWG (FPKM) and GTEx (RPKM) samples:

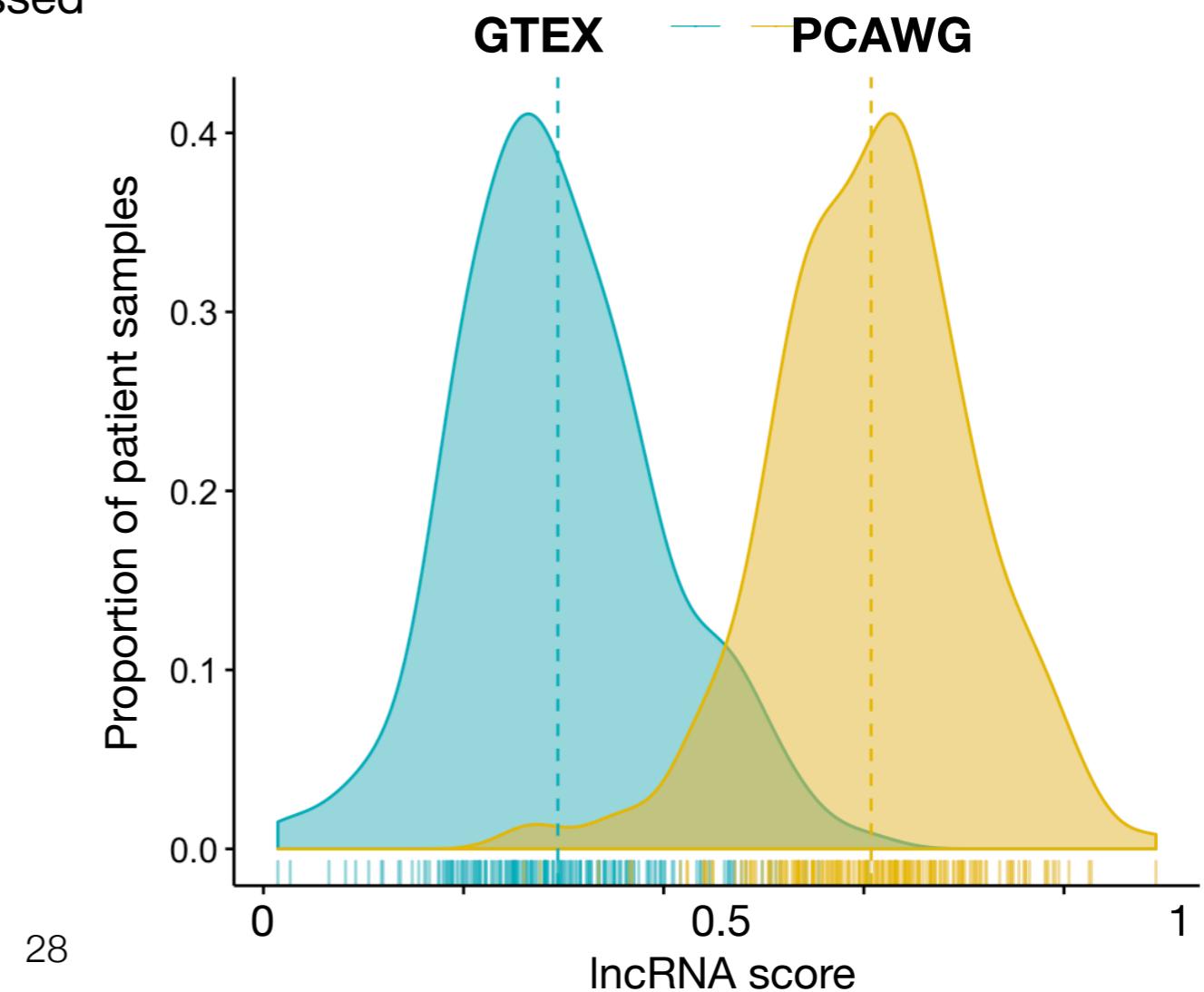
Within each patient sample, order all genes based on expression to get rank

1. Least expressed
2. Least expressed
- ⋮
- ⋮
- 25,772. Most expressed
- 25,773. Most expressed

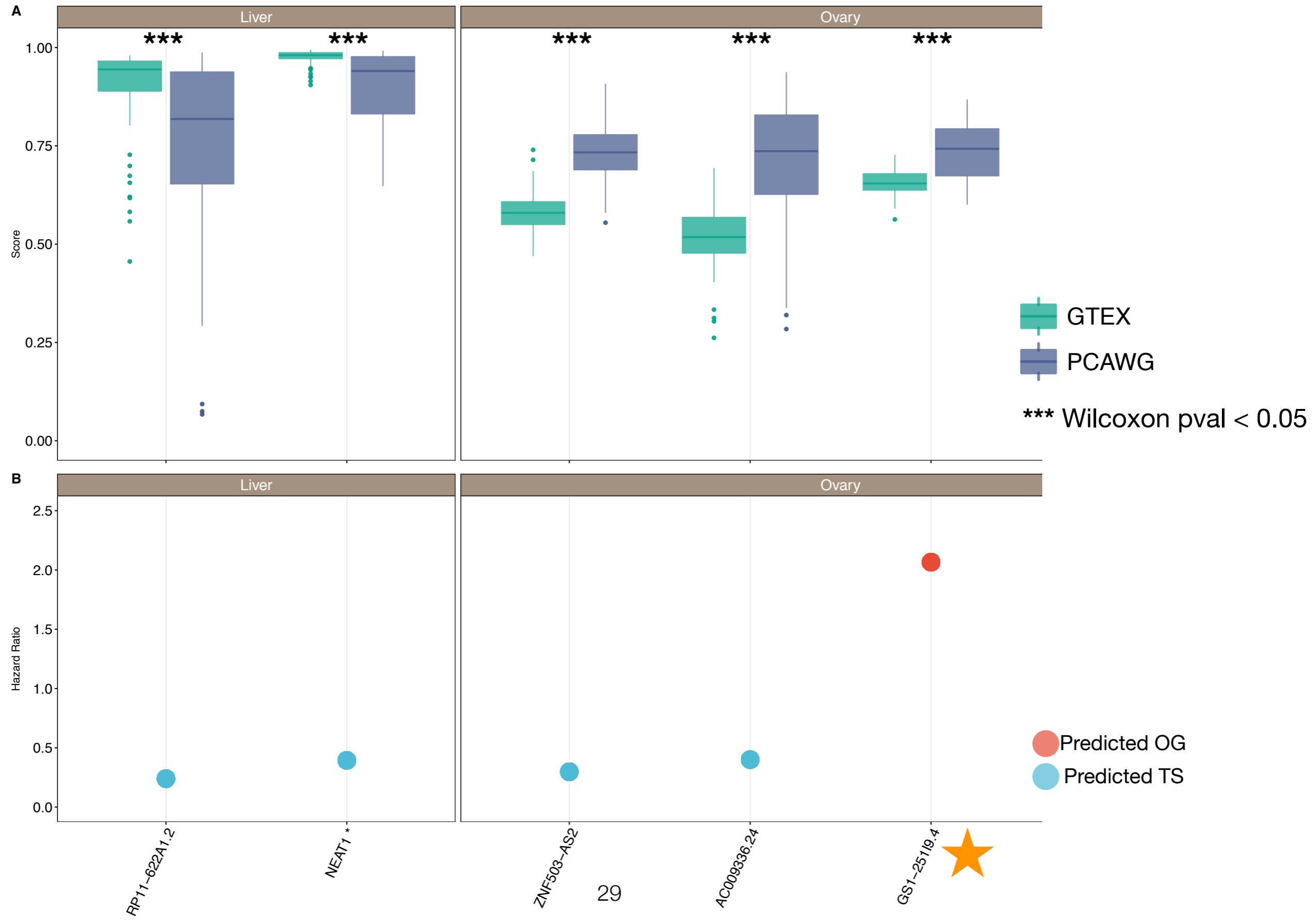
All genes = 20,166 PCGs + 5,607 lncRNAs

Divide rank of gene by length of total genes in list to get score

Compare distribution of lncRNA scores between the two datasets



lncRNA Expression in Normal Tissues (GTEx)



Outline

1. Introduction to lncRNAs
2. Aims and methods
3. Survival Analysis
4. Regulatory prediction

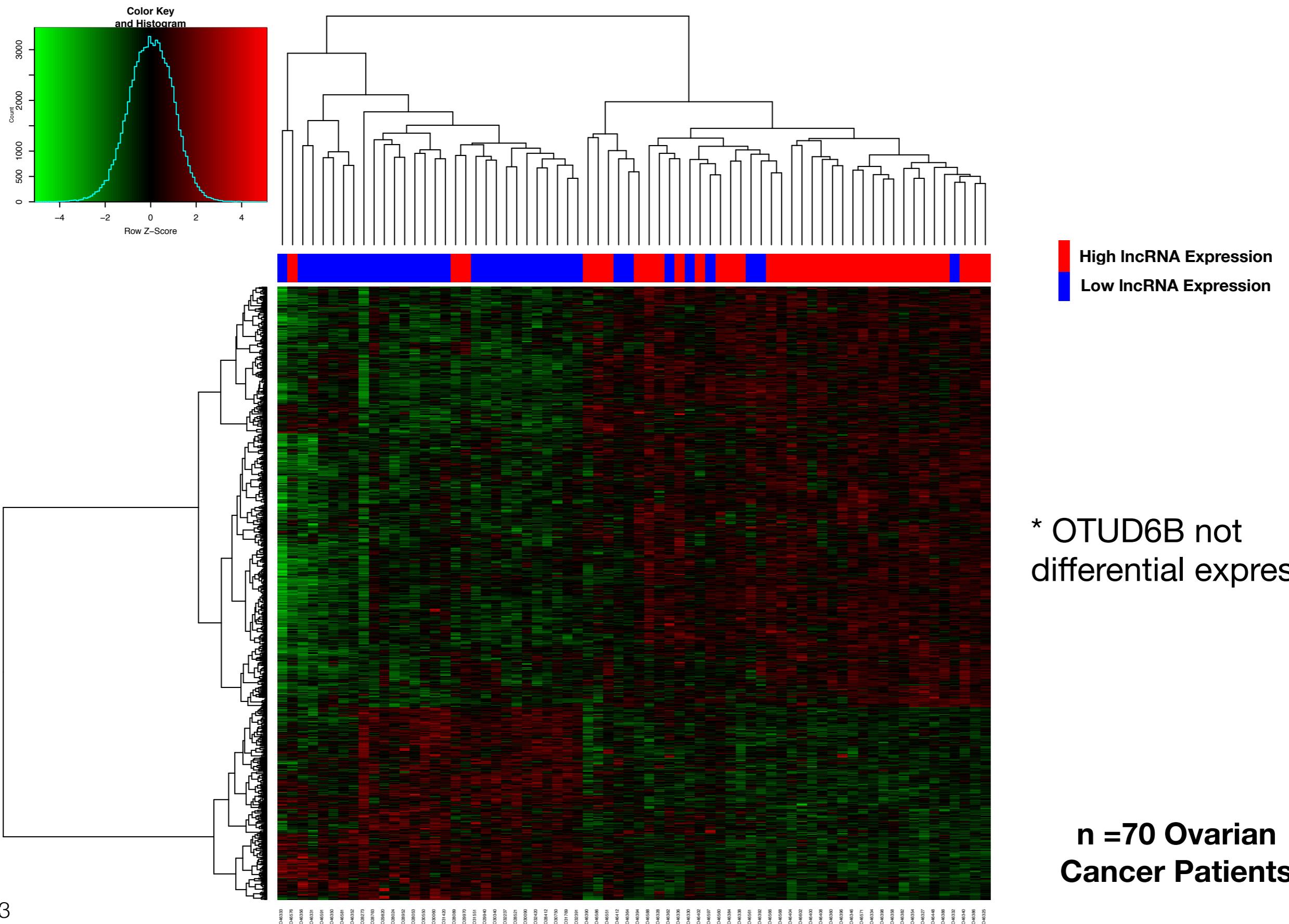
Outline

- 1. Introduction to lncRNAs**
- 2. Aims and methods**
- 3. Survival Analysis**
- 4. Regulatory prediction**

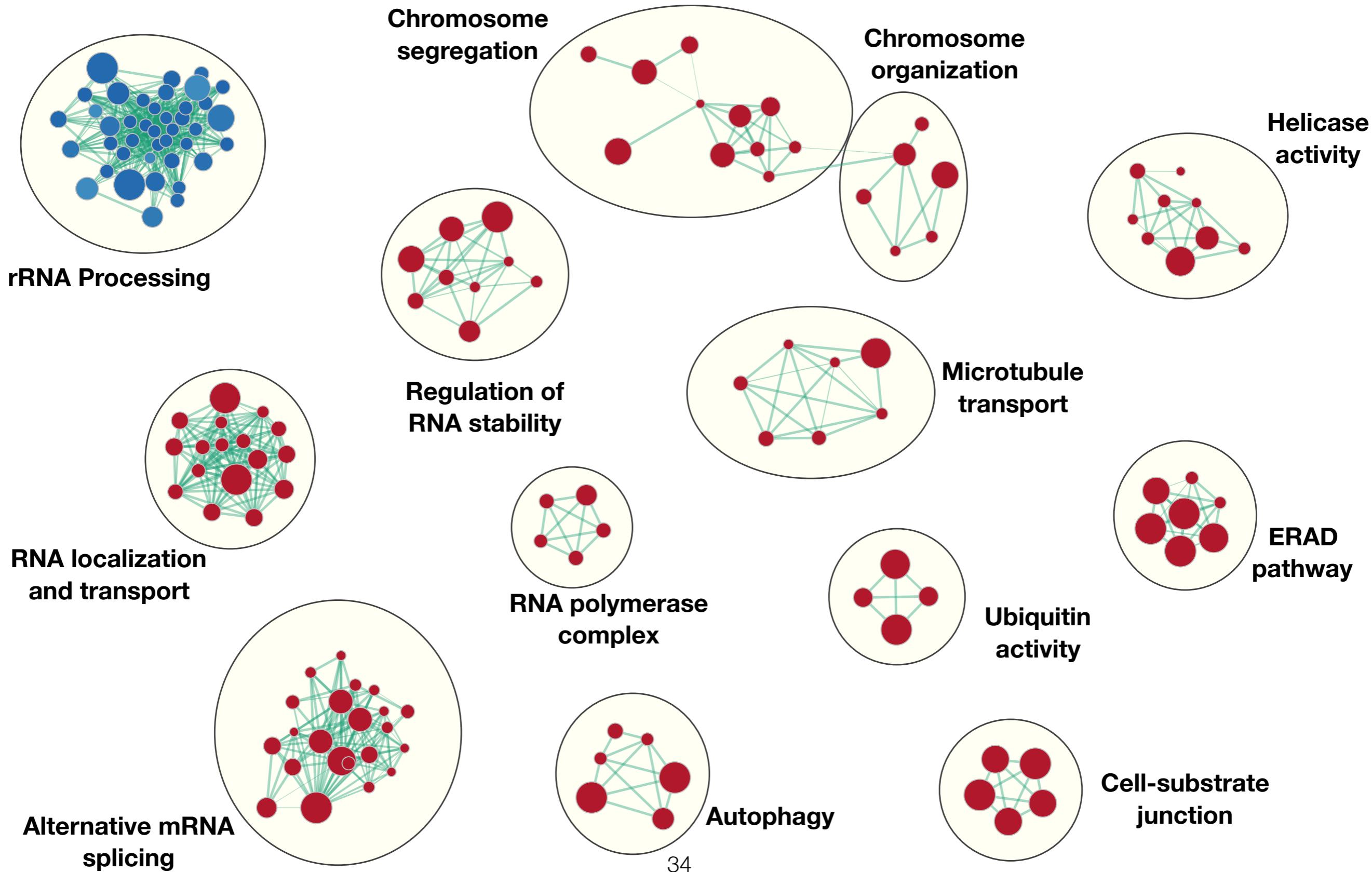
GS1-251I9.4 lncRNA

- ♦ Antisense lncRNA on chromosome 8 with two transcripts (2 and 3 exons)
- ♦ OTUD6B on opposite strand
 - Protease that cleaves ubiquitin linkages
 - Was not identified to be differentially expressed

GS1-251|9.4 Differentially Expressed Genes



GS1-251 I9.4 Pathway Enrichment Analysis



GS1-251I9.4 Differential Expression

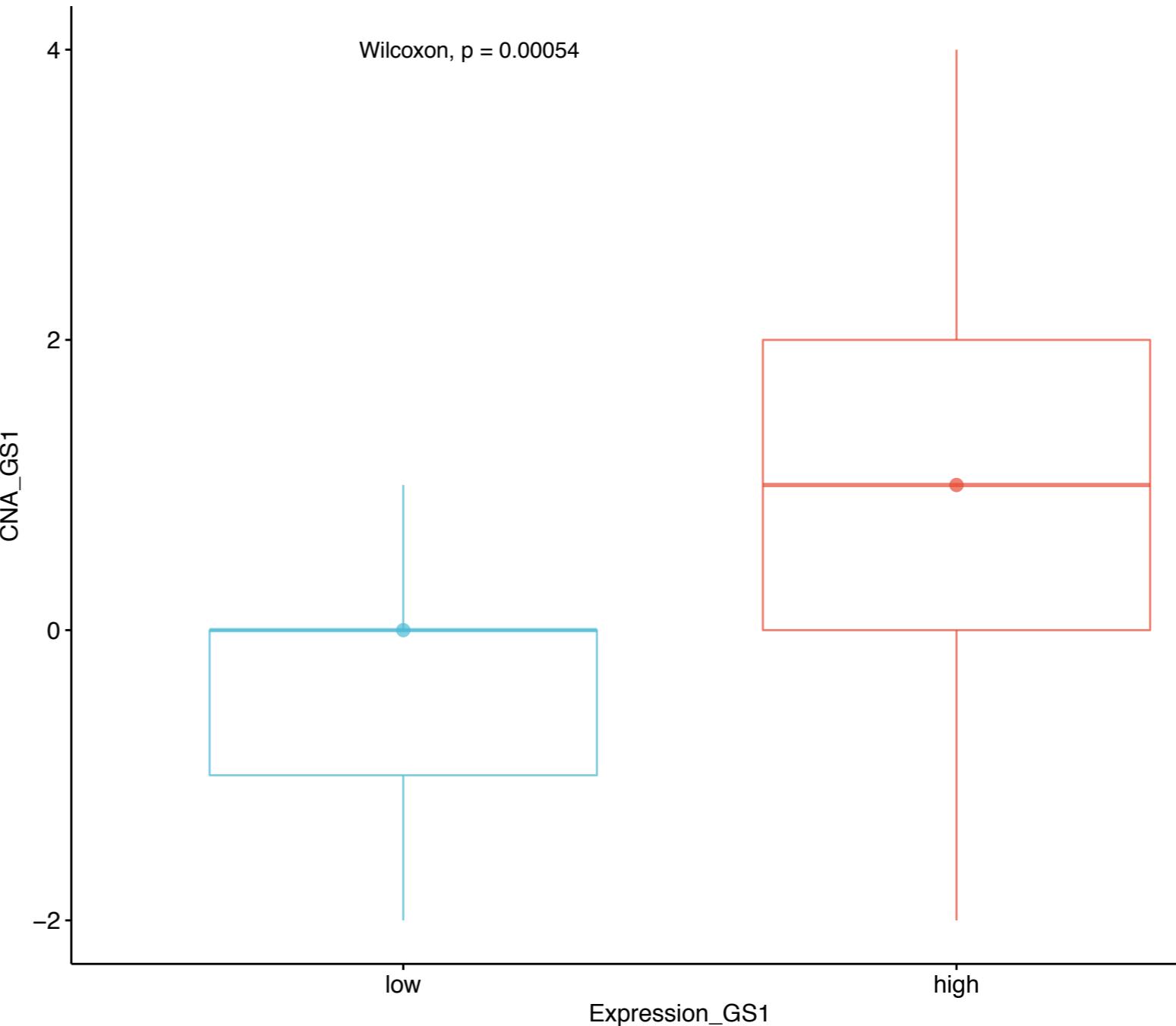
Clinical and survival data

Gene Expression (RNA-Seq)

Copy Number Aberrations

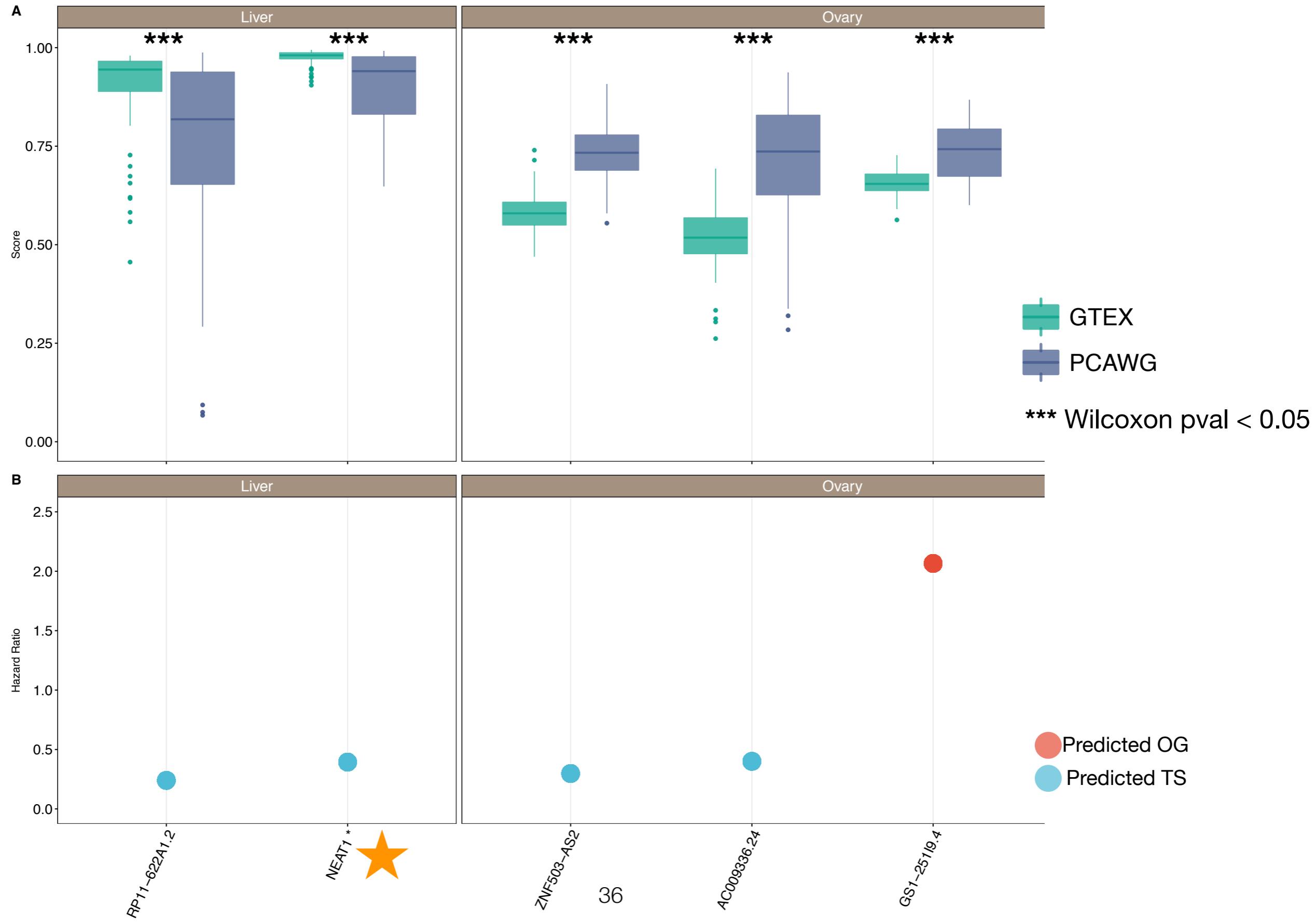
GS1-251I9.4 Expression vs CNA in 66 Ovarian Cancer Patients

Expression_GS1 low high



Next: Test association between CNA and global gene expression changes

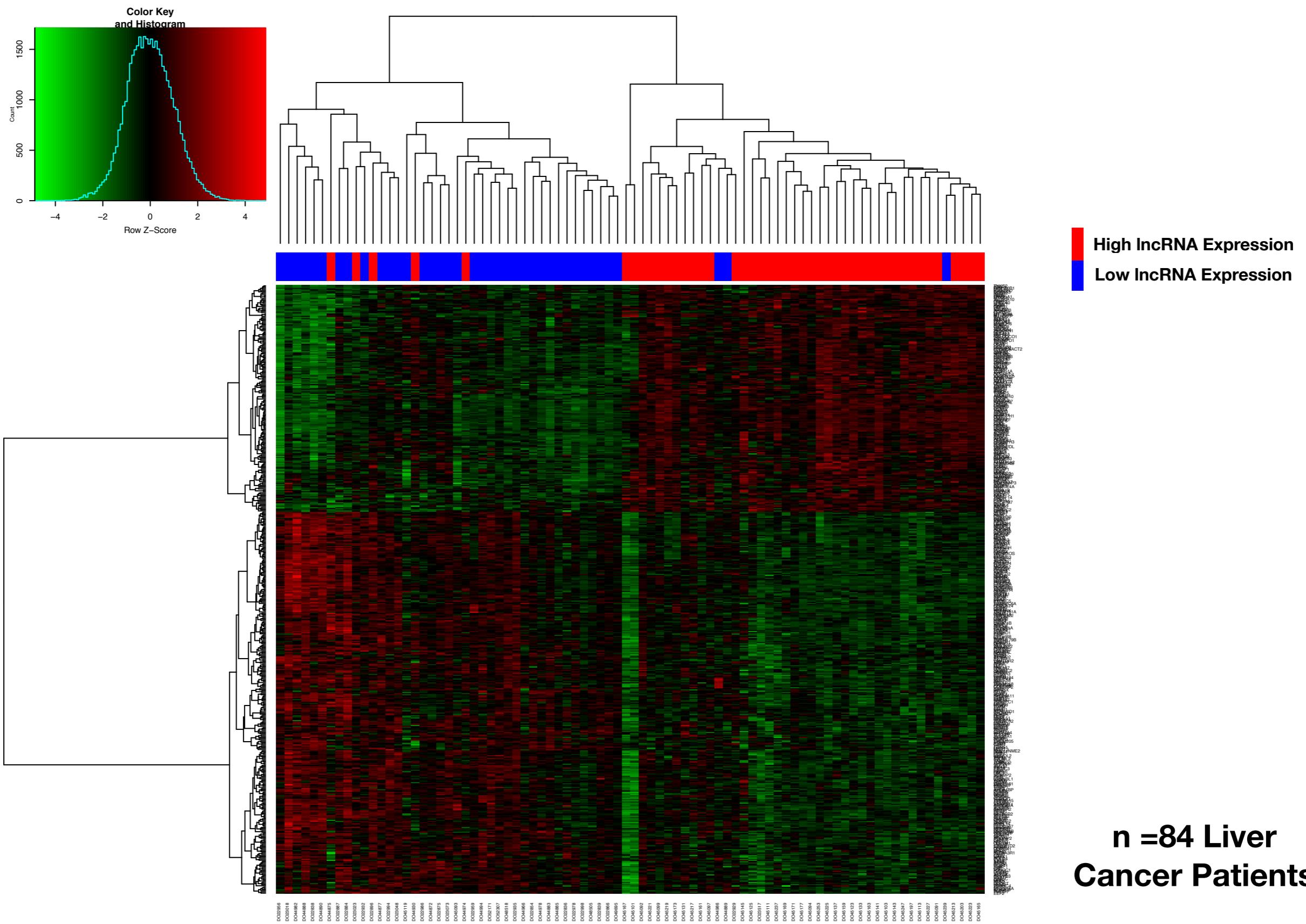
lncRNA Expression in Normal Tissues (GTEx)



NEAT1 lncRNA

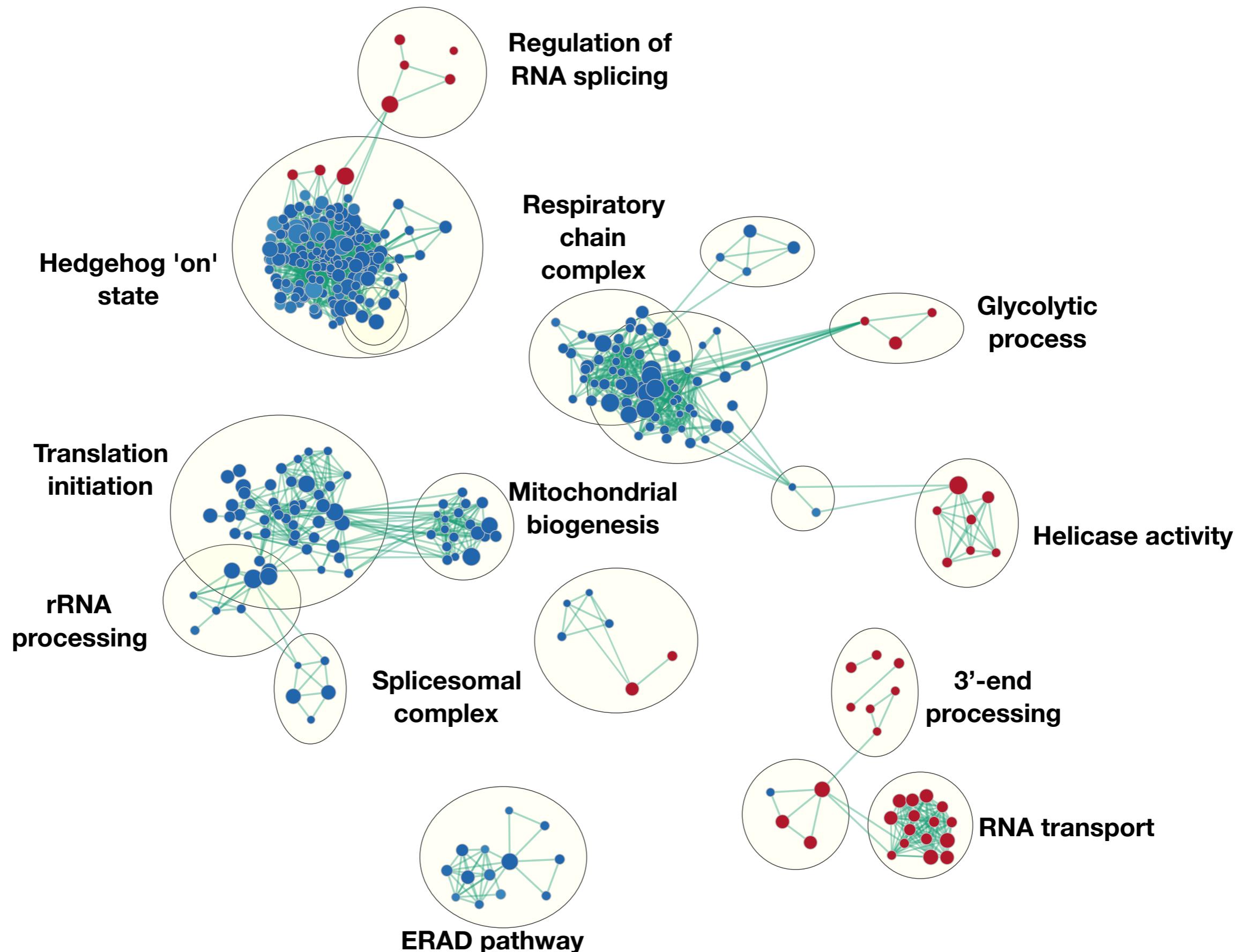
- ♦ Intergenic lncRNA on chromosome 11 with 5 transcripts (1 and 2 exons)
- ♦ NEAT1 shown to localize to hundreds of genomic sites in human cells
- ♦ Unclear regulation of NEAT1 by TP53

NEAT1 Differentially Expressed Genes

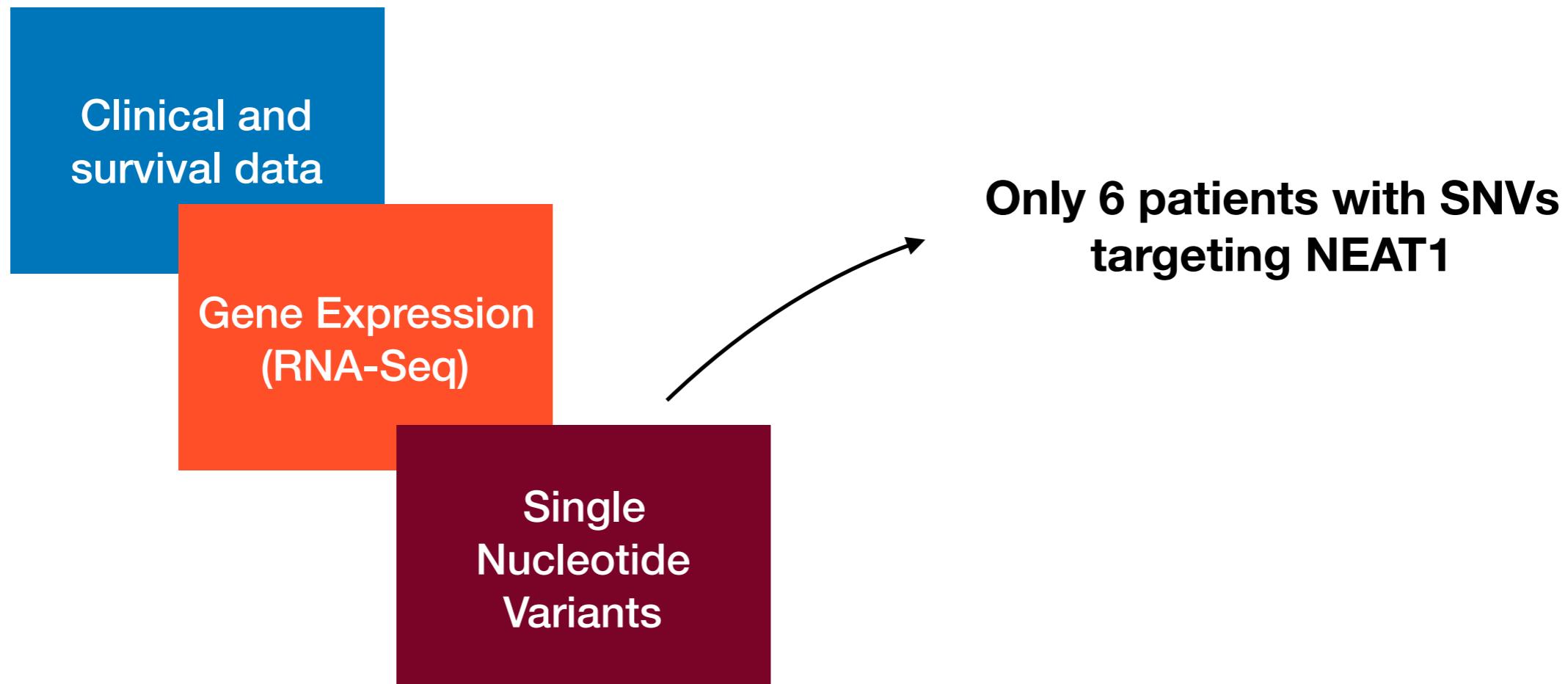


n =84 Liver Cancer Patients

NEAT1 Pathway Enrichment Analysis



NEAT1 Differential Expression



Next: Analyze NEAT1 promoter methylation and copy number aberrations

Summary

- ♦ Systematic screening of the transcriptome in multiple cancer types revealed 5 lncRNAs significantly associated with survival among Ovarian and Liver cancer
- ♦ lncRNAs are promising diagnostic biomarkers and targets for drug development
- ♦ Through the integration of genetic and epigenetic aberrations, clinical relevance and associations to other genes and pathways, we can better predict the function of 1000s of previously uncharacterized lncRNAs

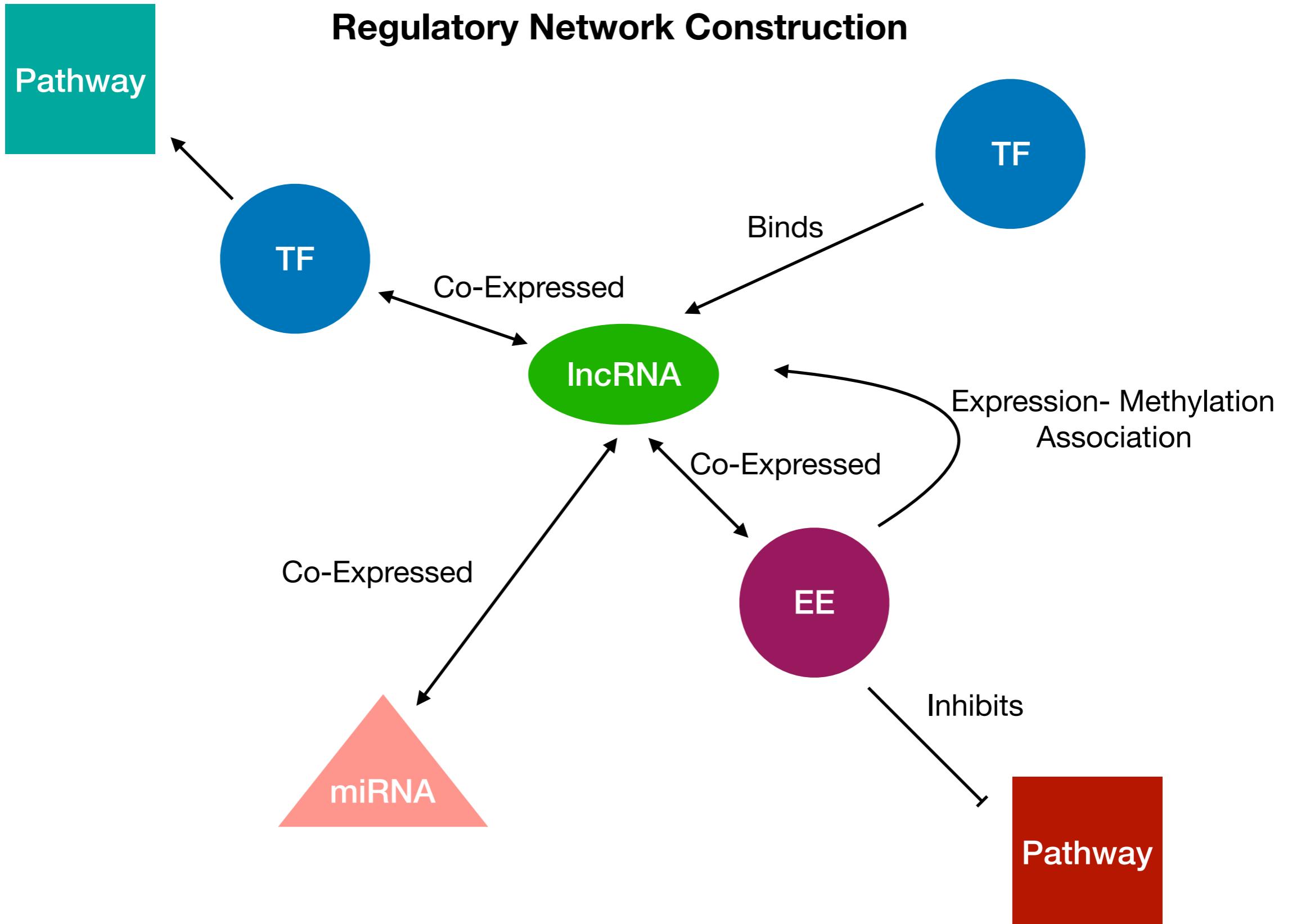
Future Directions - Aim 1

- ♦ Cutoff optimization
 - Fold change between tumour and matched normal tissues
- ♦ Multivariate prognostic model
- ♦ Expand to other events such as recurrence
- ♦ Integrate additional clinical covariates such as age, smoking and drinking status, type of treatment and hepatitis status (liver cancer)

Future Directions - Aims 2 + 3

- ♦ Validate causal aberrations
- ♦ Integrate transcription factor binding data
- ♦ Predict whether lncRNA is acting in *cis* or *trans*
- ♦ Sequence analysis
- ♦ Expand to lncRNA-miRNA co-expression analysis

Future Directions



Acknowledgements

Reimand Lab

Dr. Jüri Reimand
Dr. Diala Abd-Rabbo
Helen Zhu
Nardnisa Sintupisut
Xiao Wang
Yao Li
Jonathan Barenboim
Luyao Ruan
Marta Paczkowska
Tina Huang

Supervisory Committee

Dr. Fritz Roth
Dr. Hansen He

