

Discovery and characterization of clinically relevant lncRNAs in multiple cancer types

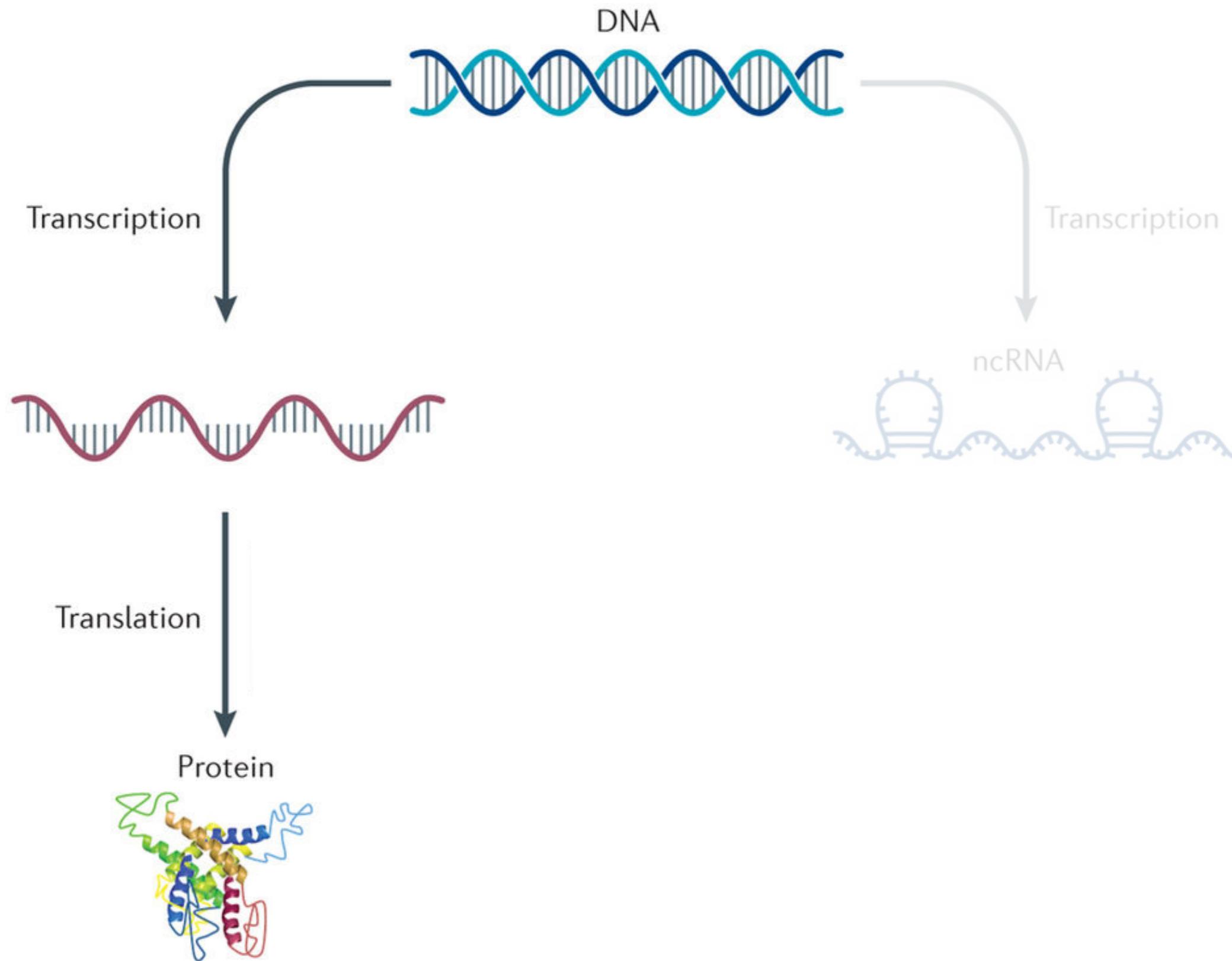
Graduate Student Seminar

Karin Isaev

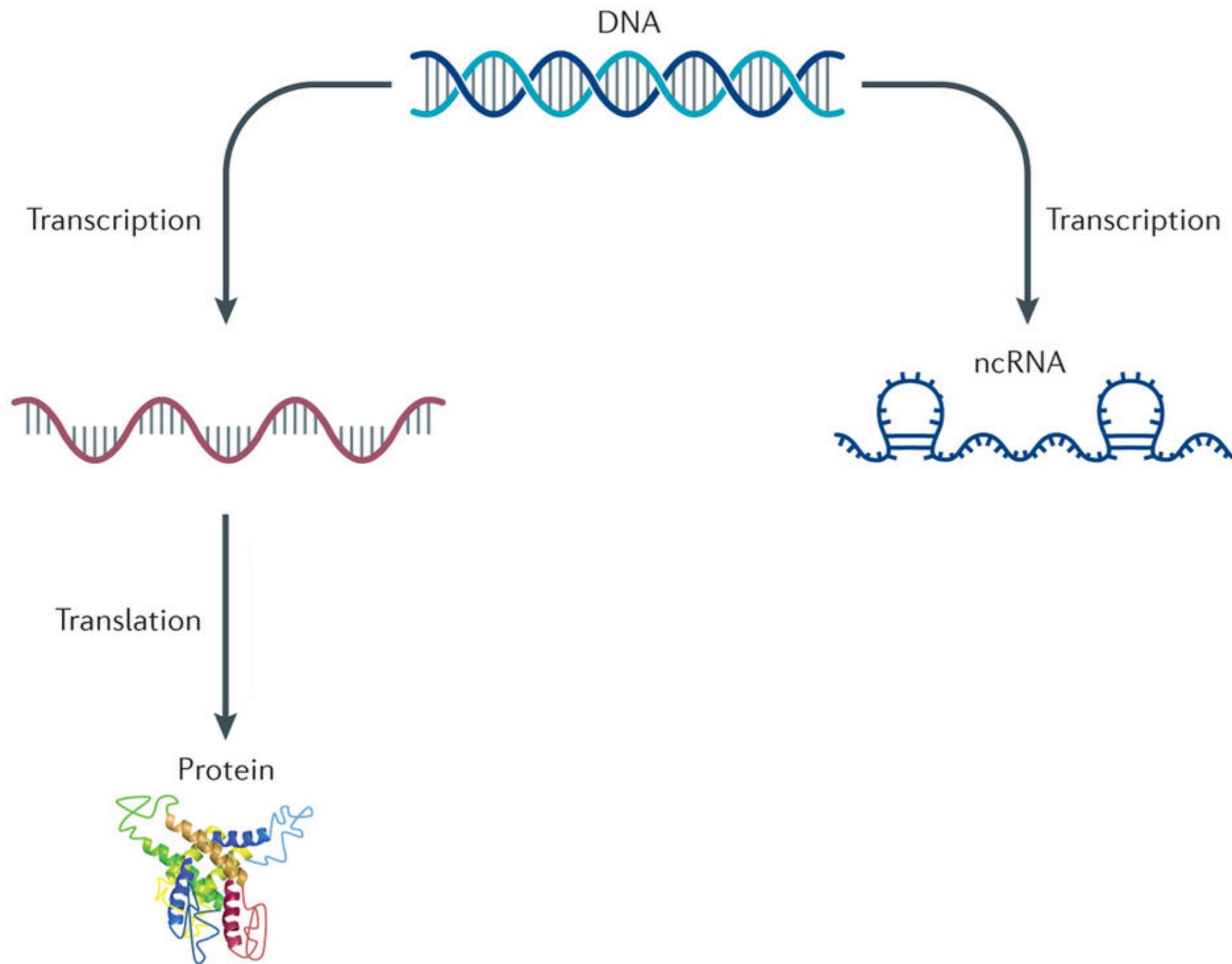
Supervisor: Dr. Jüri Reimand

December 6th, 2017

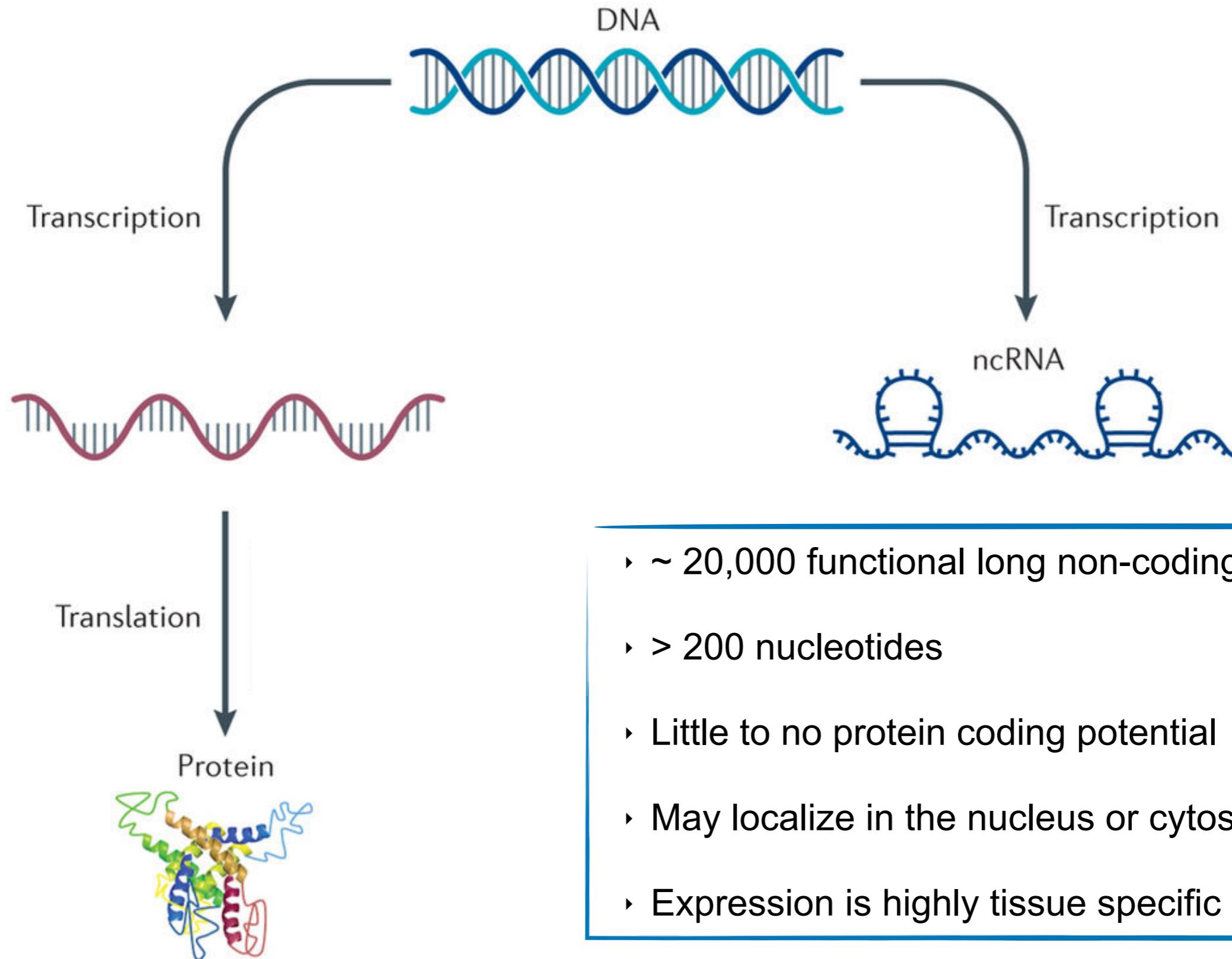
Central dogma in the context of lncRNAs



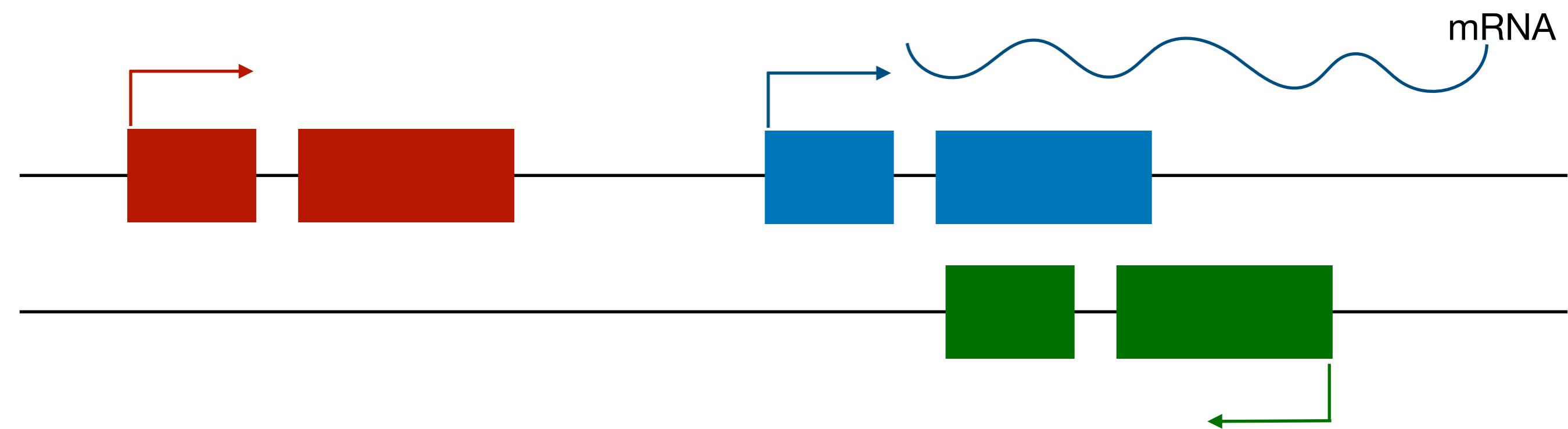
Central dogma in the context of lncRNAs



Central dogma in the context of lncRNAs

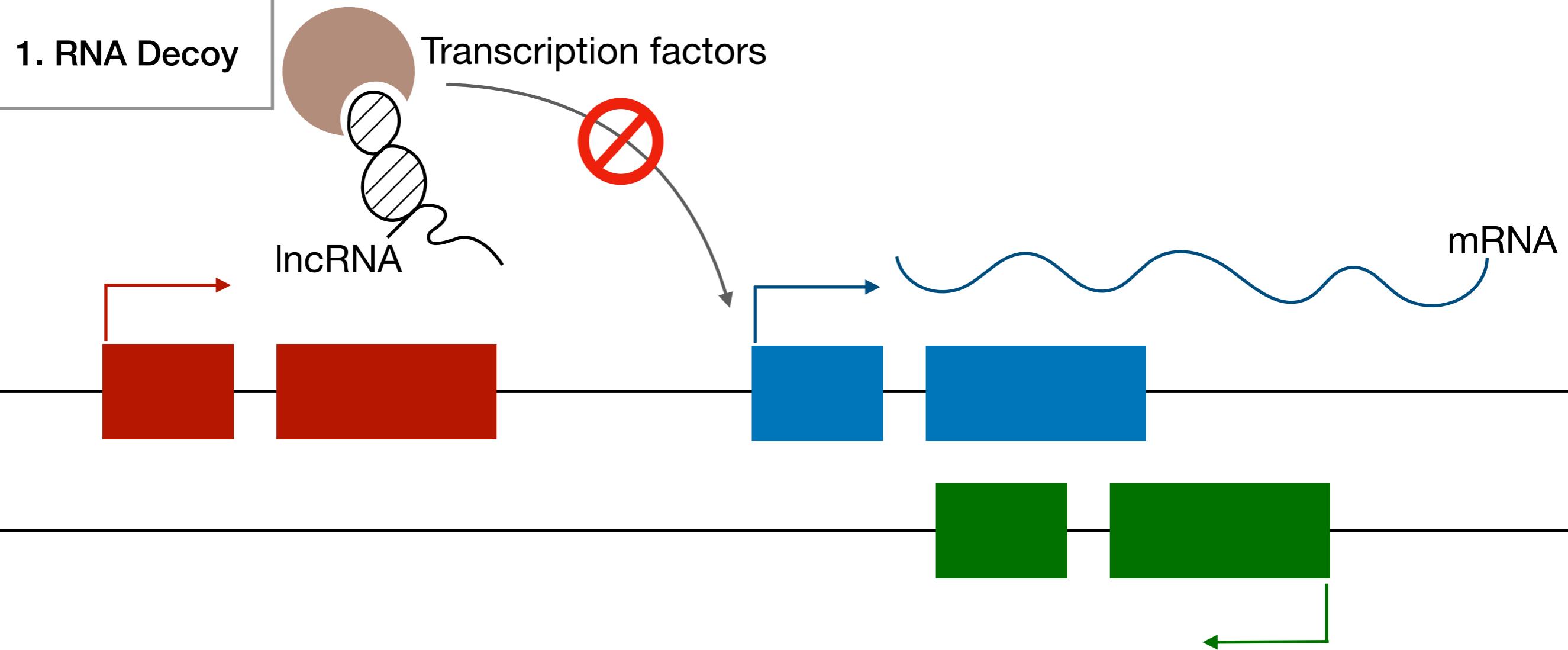


Regulatory mechanisms of lncRNAs

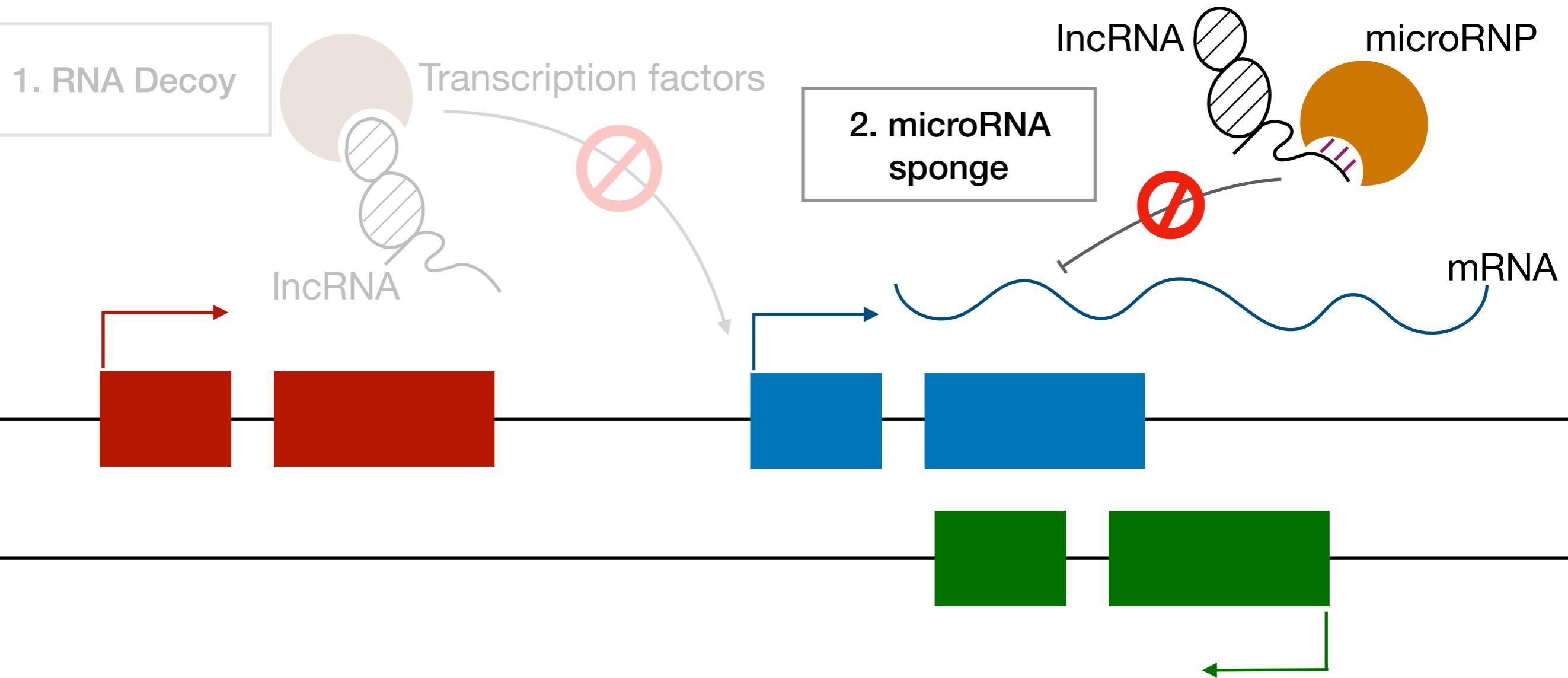


Regulatory mechanisms of lncRNAs

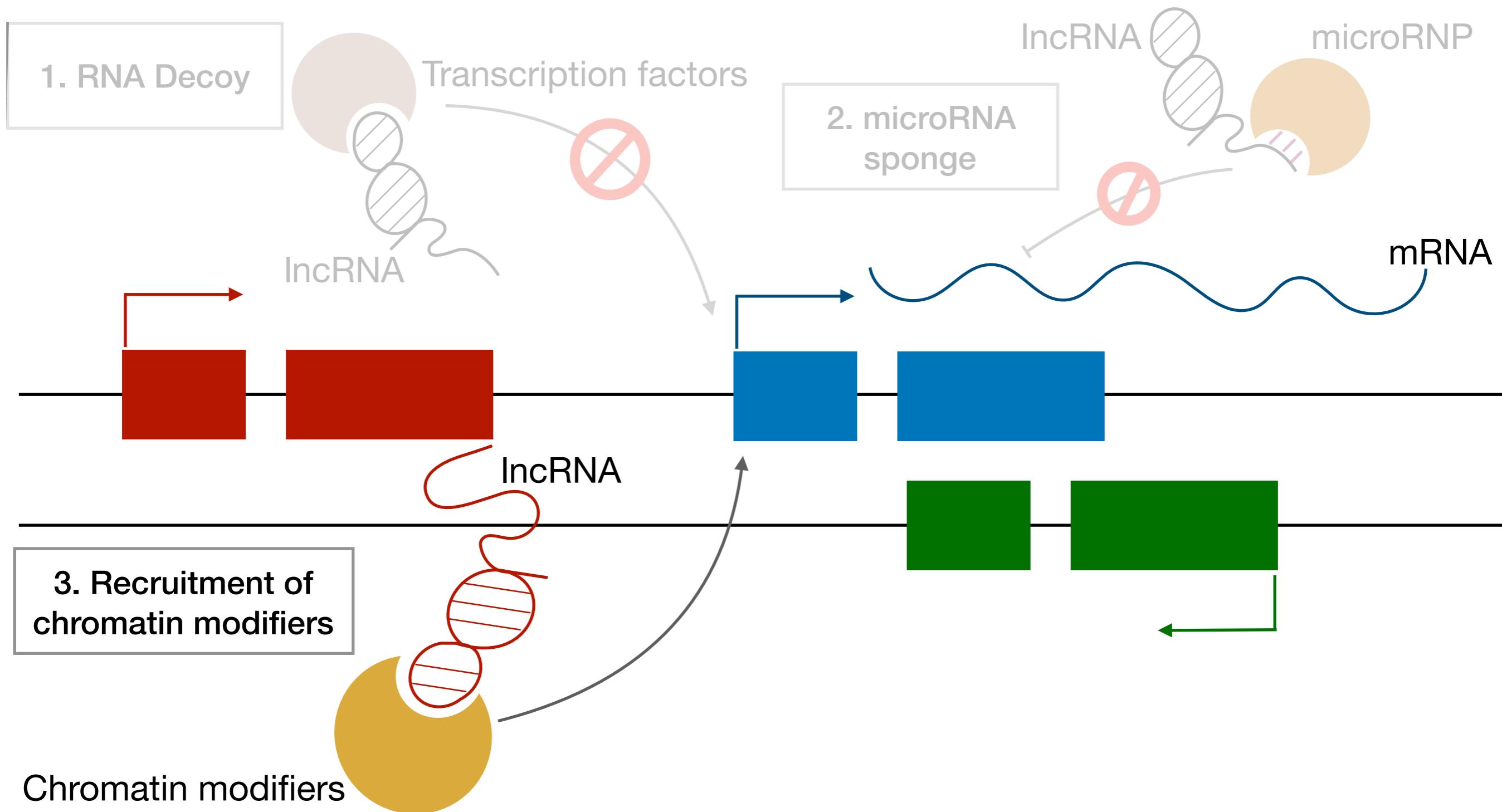
1. RNA Decoy



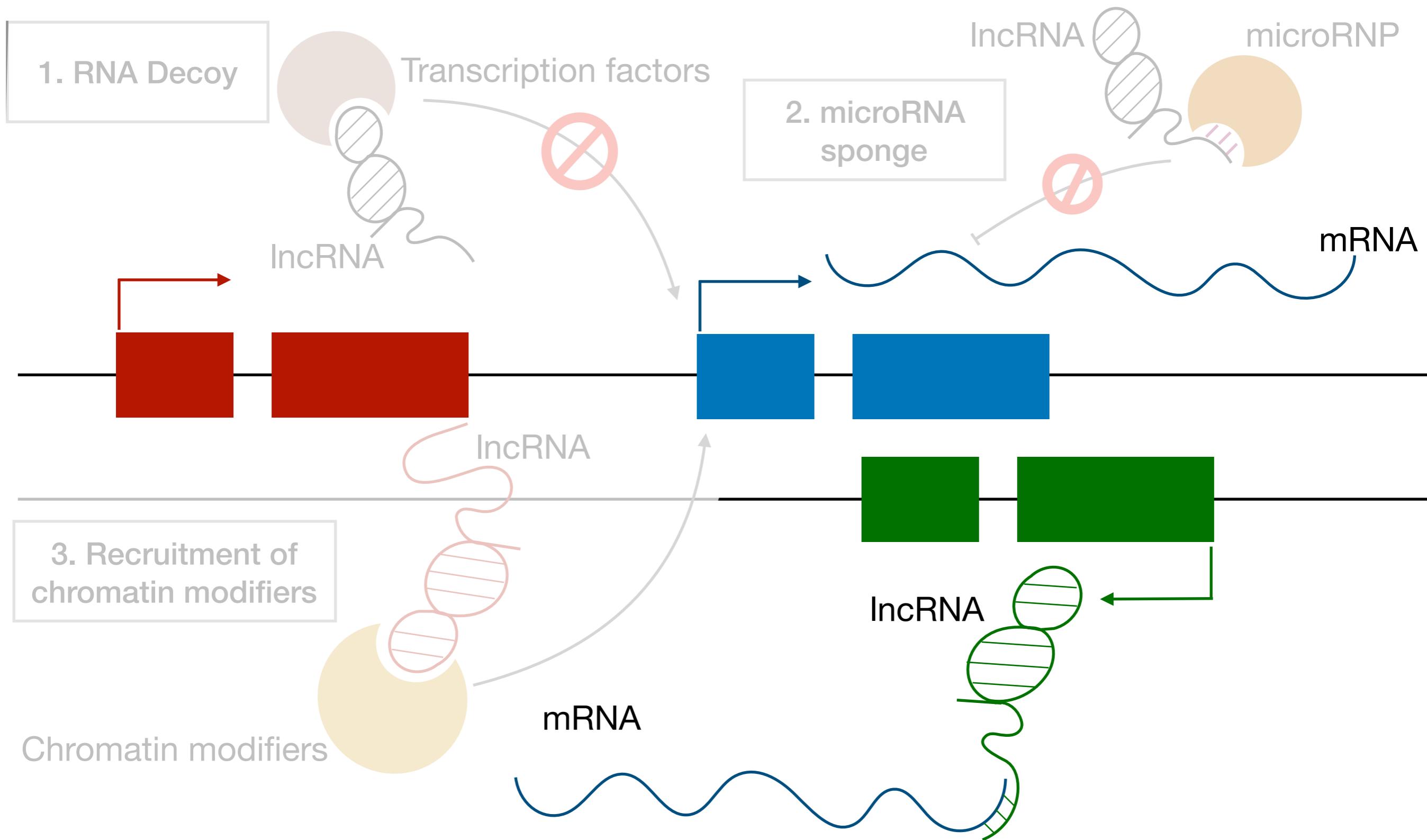
Regulatory mechanisms of lncRNAs



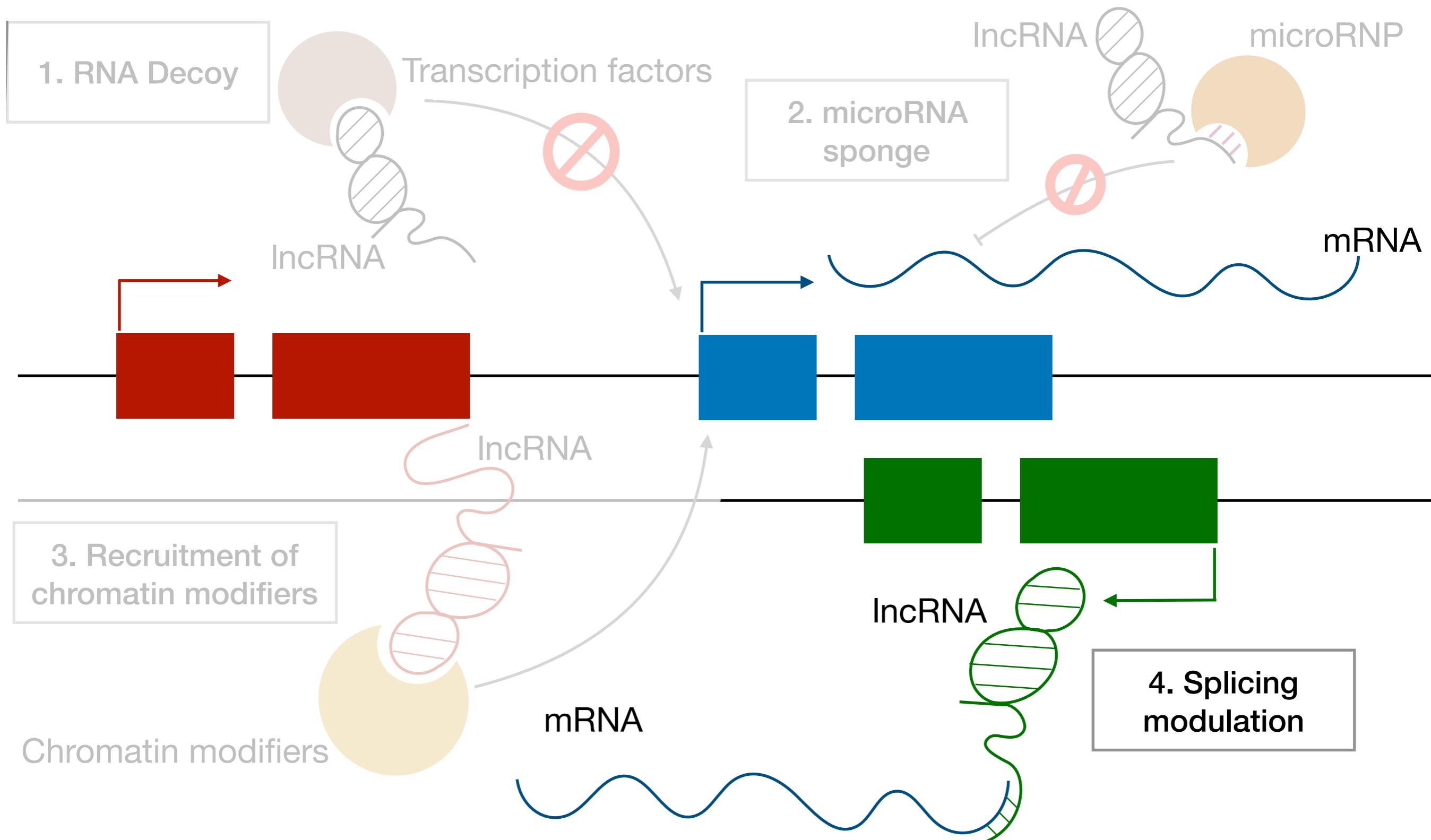
Regulatory mechanisms of lncRNAs



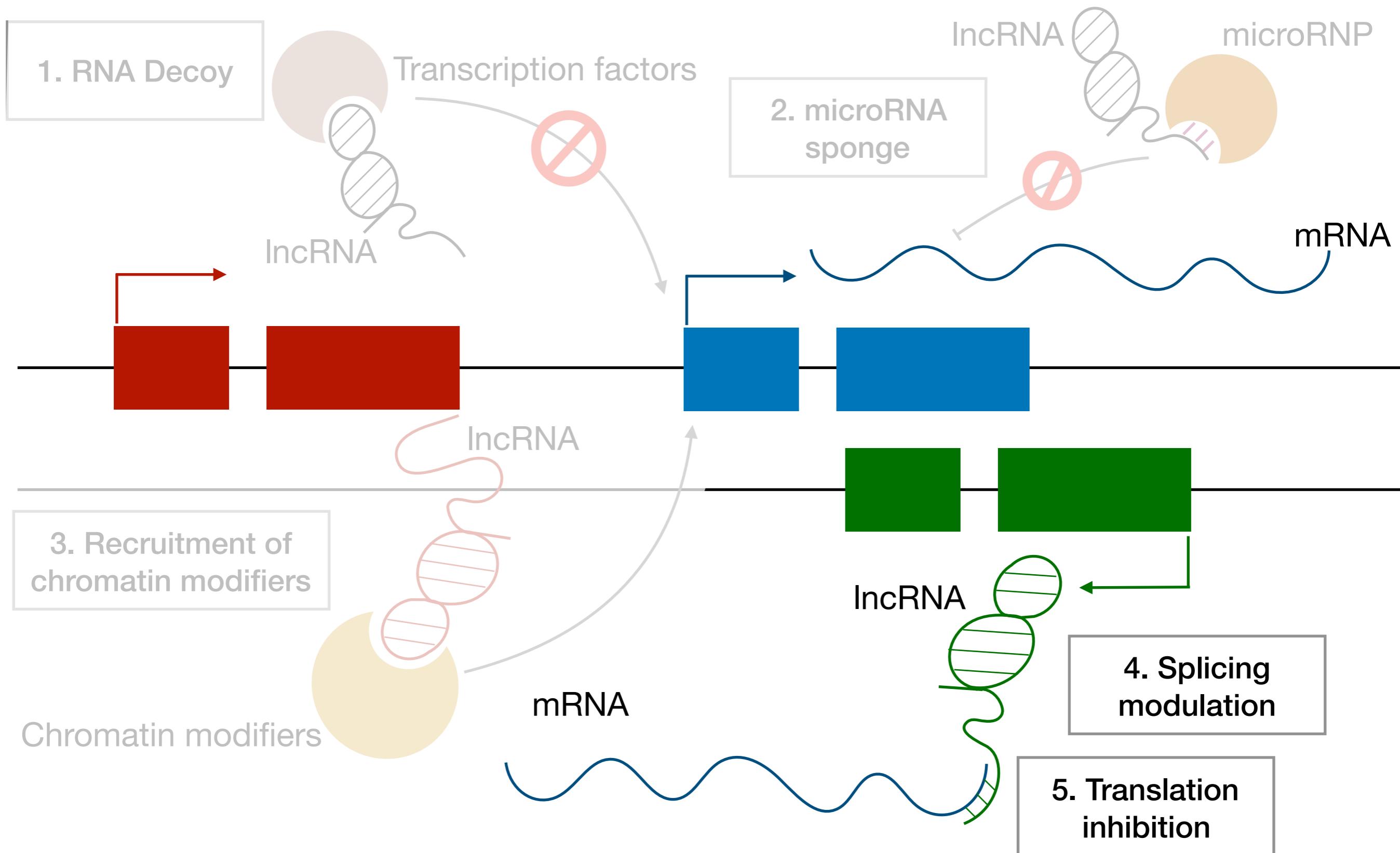
Regulatory mechanisms of lncRNAs



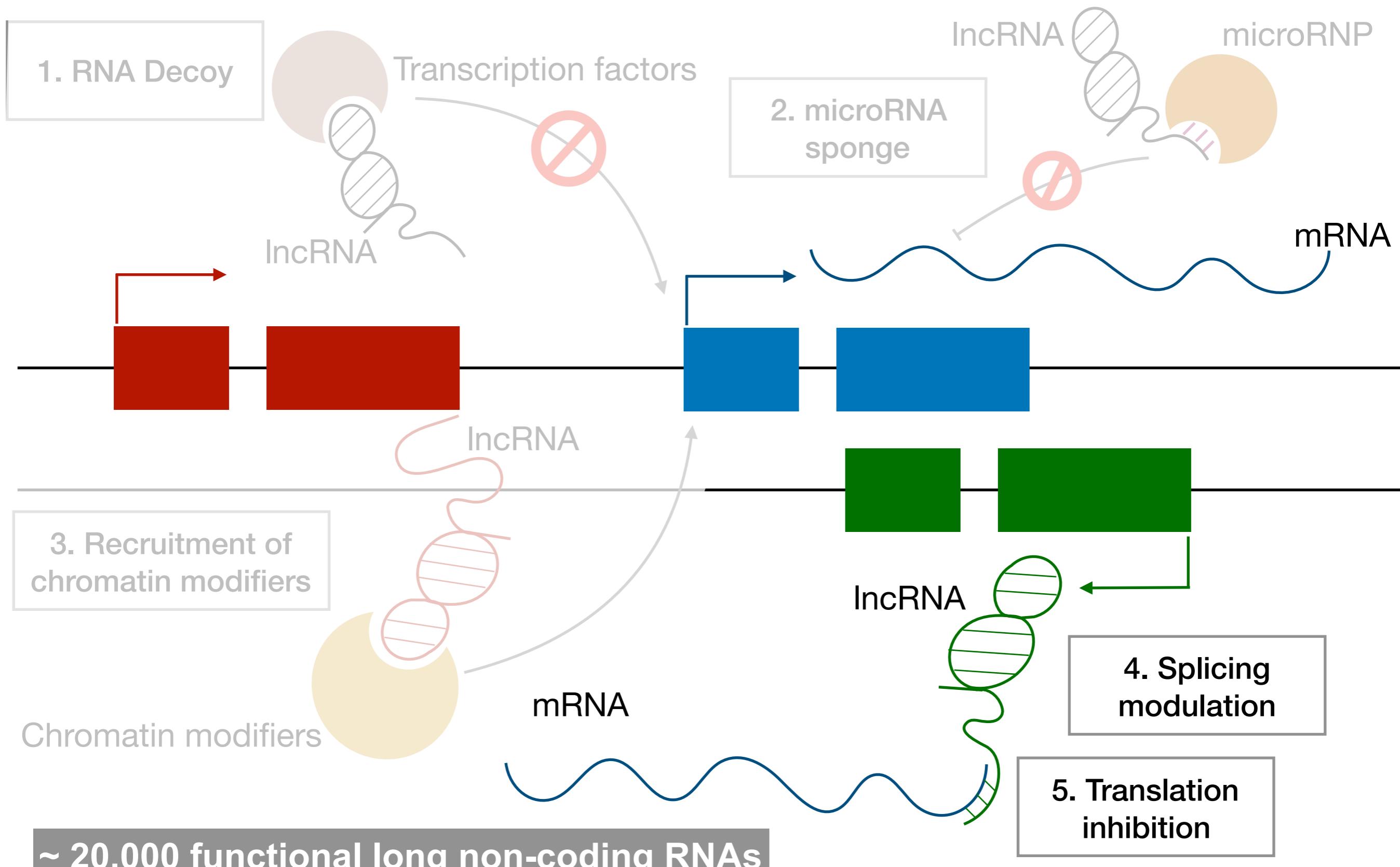
Regulatory mechanisms of lncRNAs



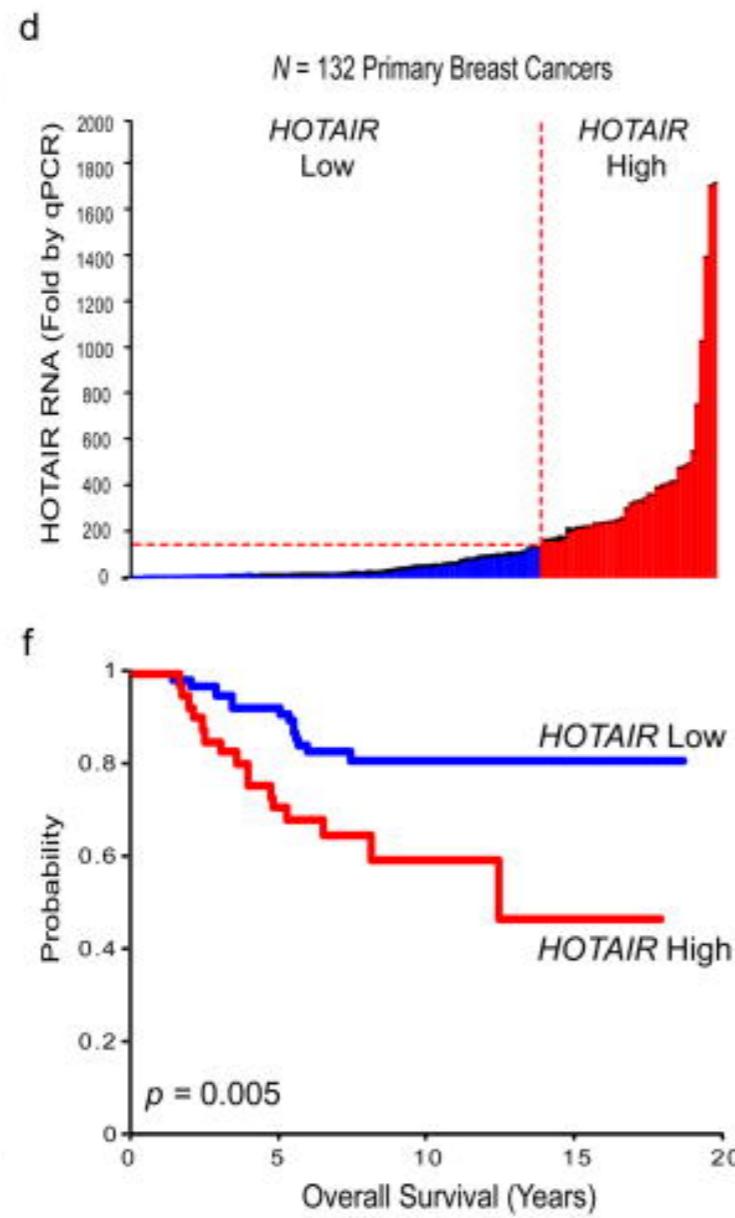
Regulatory mechanisms of lncRNAs



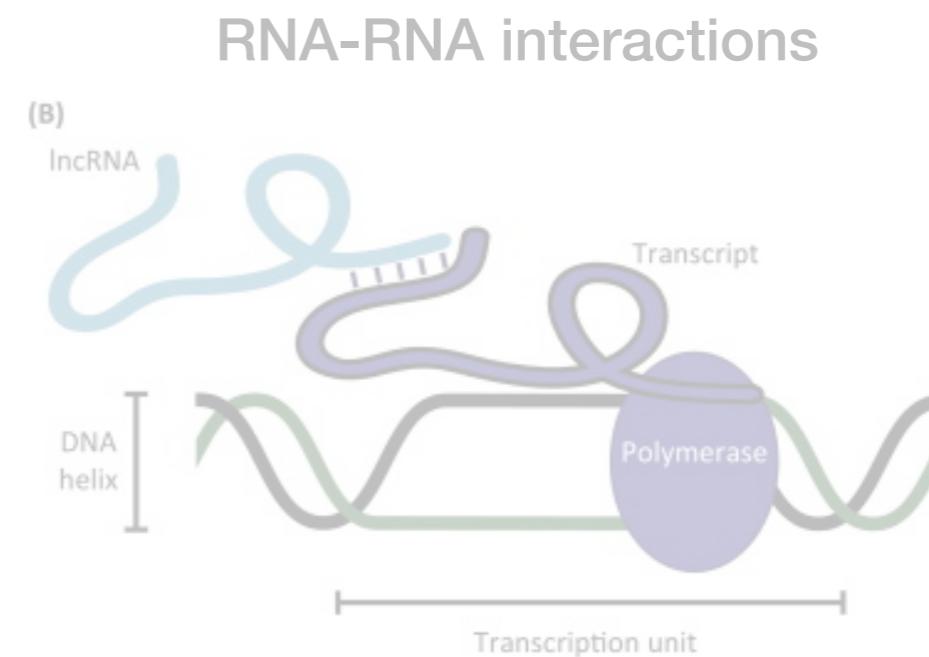
Regulatory mechanisms of lncRNAs



Rationale for studying lncRNAs in Cancer

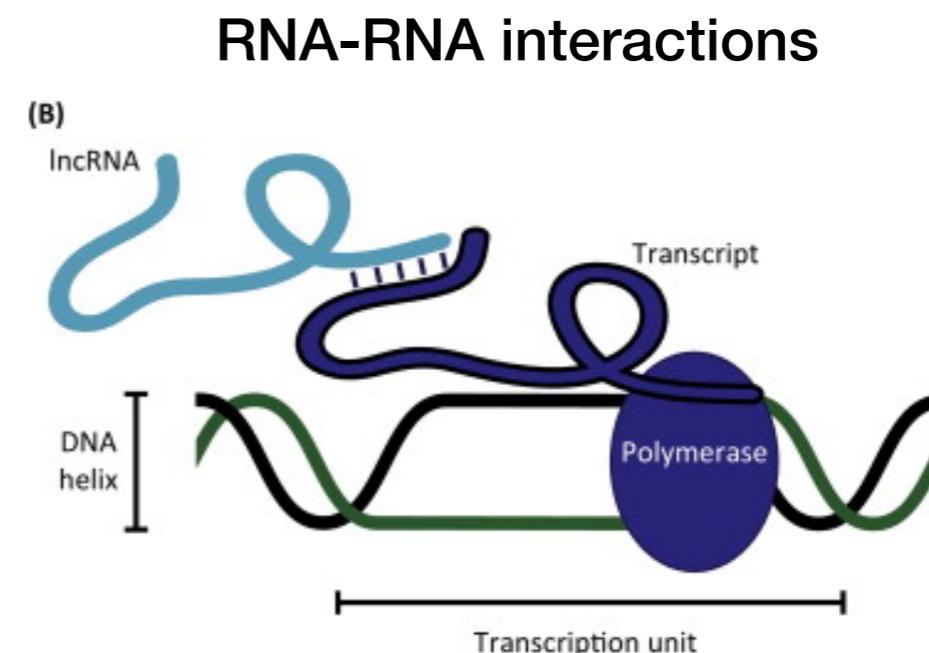
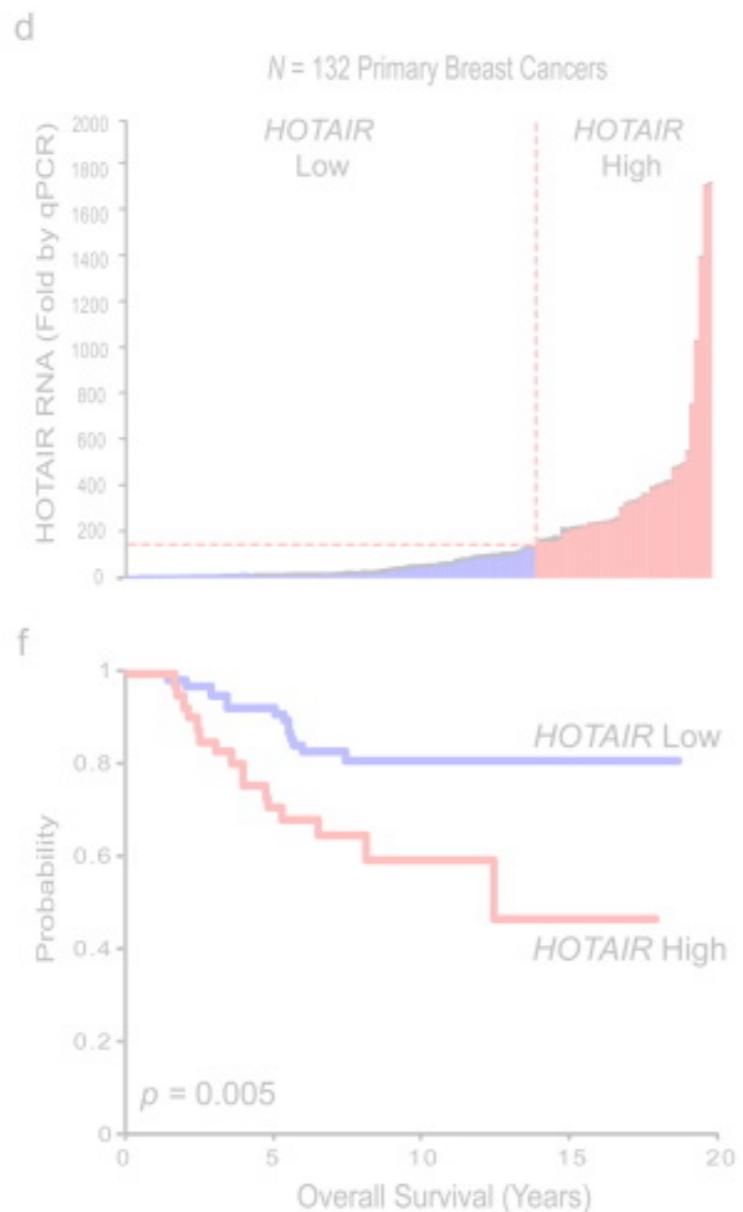


- ♦ High tissue specificity
- ♦ Prognostic/Diagnostic Biomarker potential



- ♦ Attractive targets for therapeutic intervention

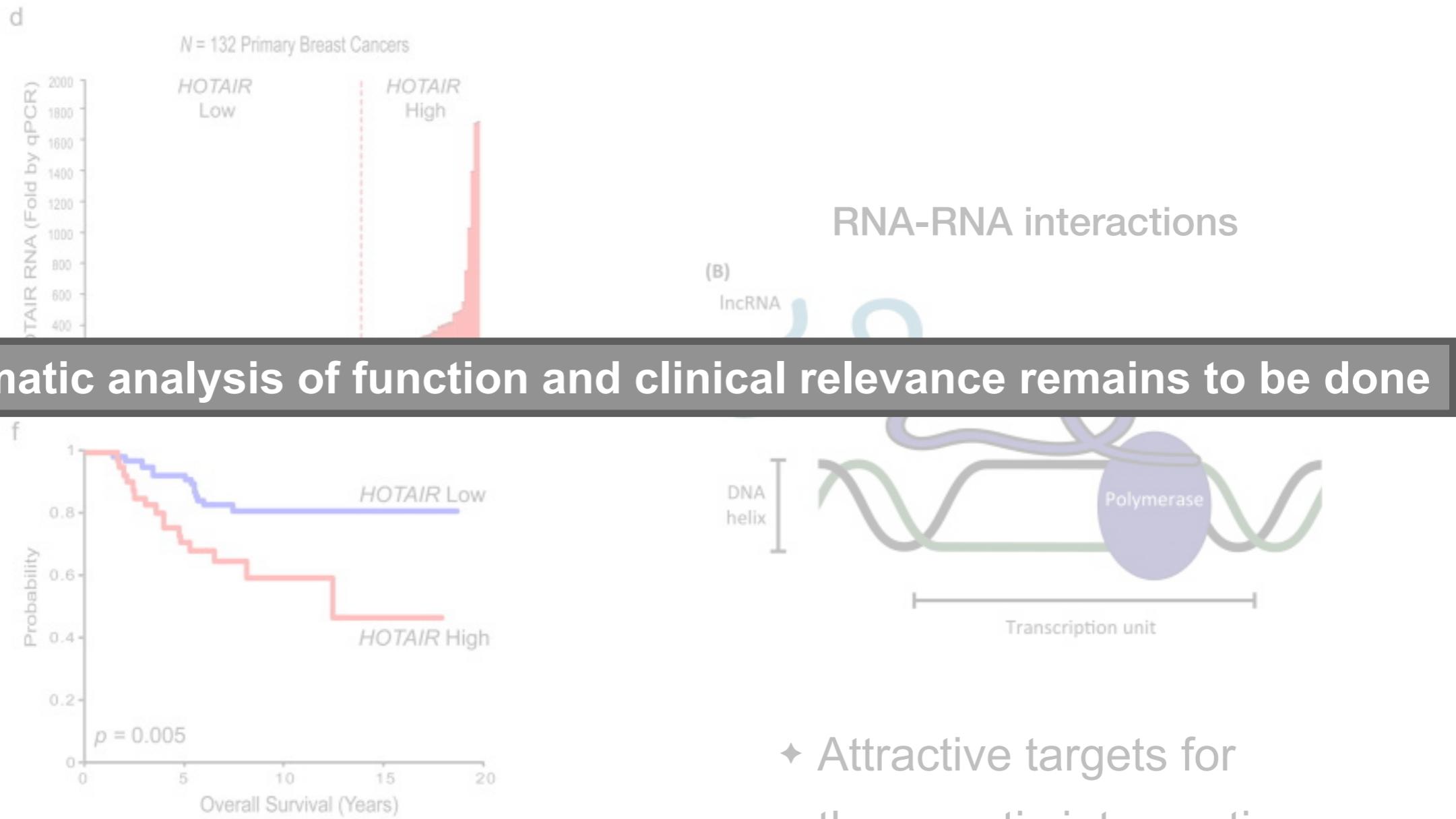
Rationale for studying lncRNAs in Cancer



- ♦ Attractive targets for therapeutic intervention

- ♦ High tissue specificity
- ♦ Prognostic/Diagnostic Biomarker potential

Rationale for studying lncRNAs in Cancer



Systematic analysis of function and clinical relevance remains to be done

- ♦ High tissue specificity
- ♦ Prognostic/Diagnostic Biomarker potential

- ♦ Attractive targets for therapeutic intervention

Hypothesis and Aims

Hypothesis: lncRNA aberrations drive lncRNA molecular profiles that are informative of disease outcomes while causing changes in the regulation of cancer pathways

Aim 1:

Identify lncRNA molecular profiles associated with survival in multiple cancer types

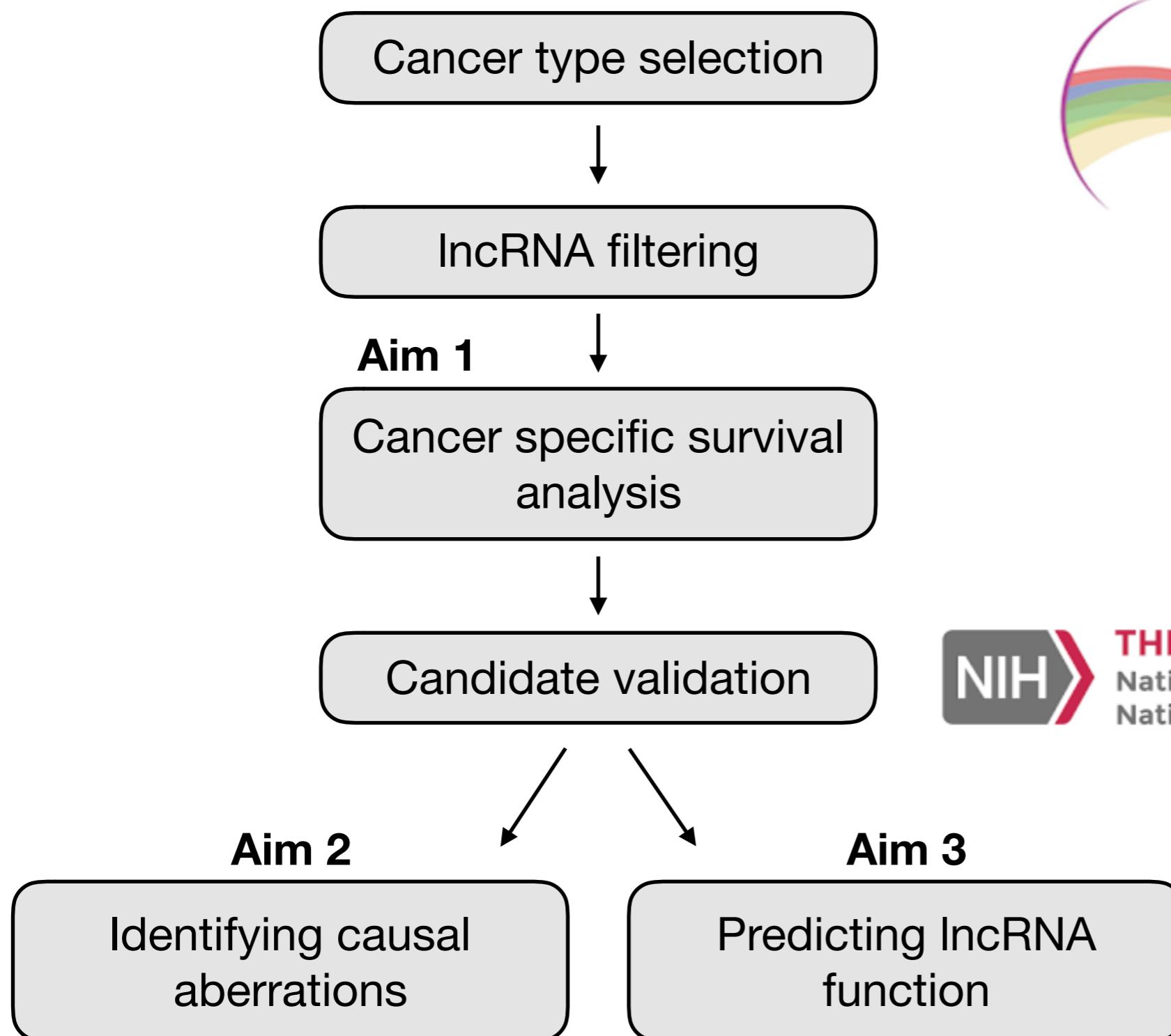
Aim 2:

Identify genetic and epigenetic aberrations associated with lncRNA molecular profiles

Aim 3:

Predict lncRNA targets and regulators by associating lncRNA molecular profiles to pathways and protein coding genes

Methods



PCAWG
PanCancer Analysis
OF WHOLE GENOMES

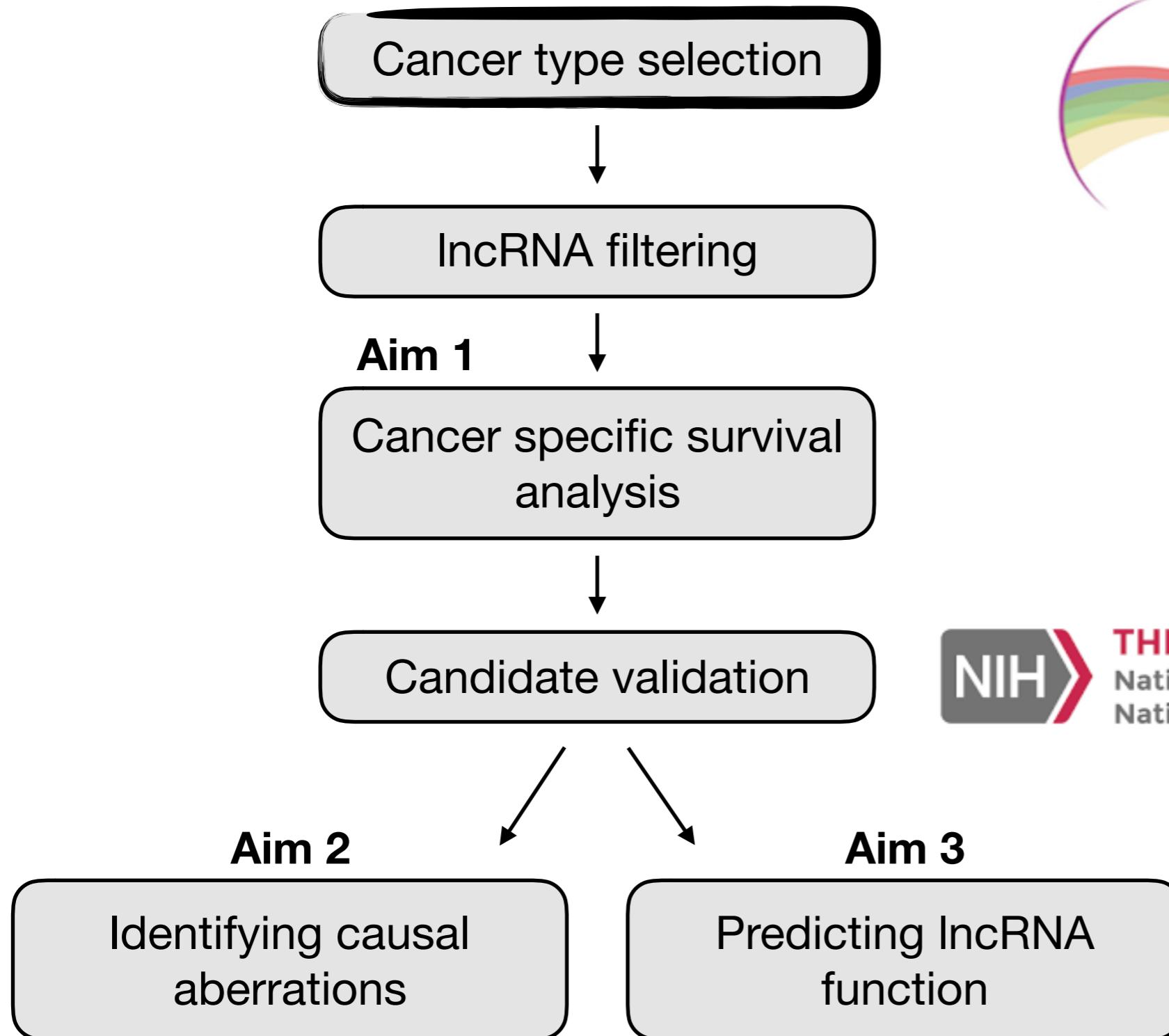


THE CANCER GENOME ATLAS

National Cancer Institute

National Human Genome Research Institute

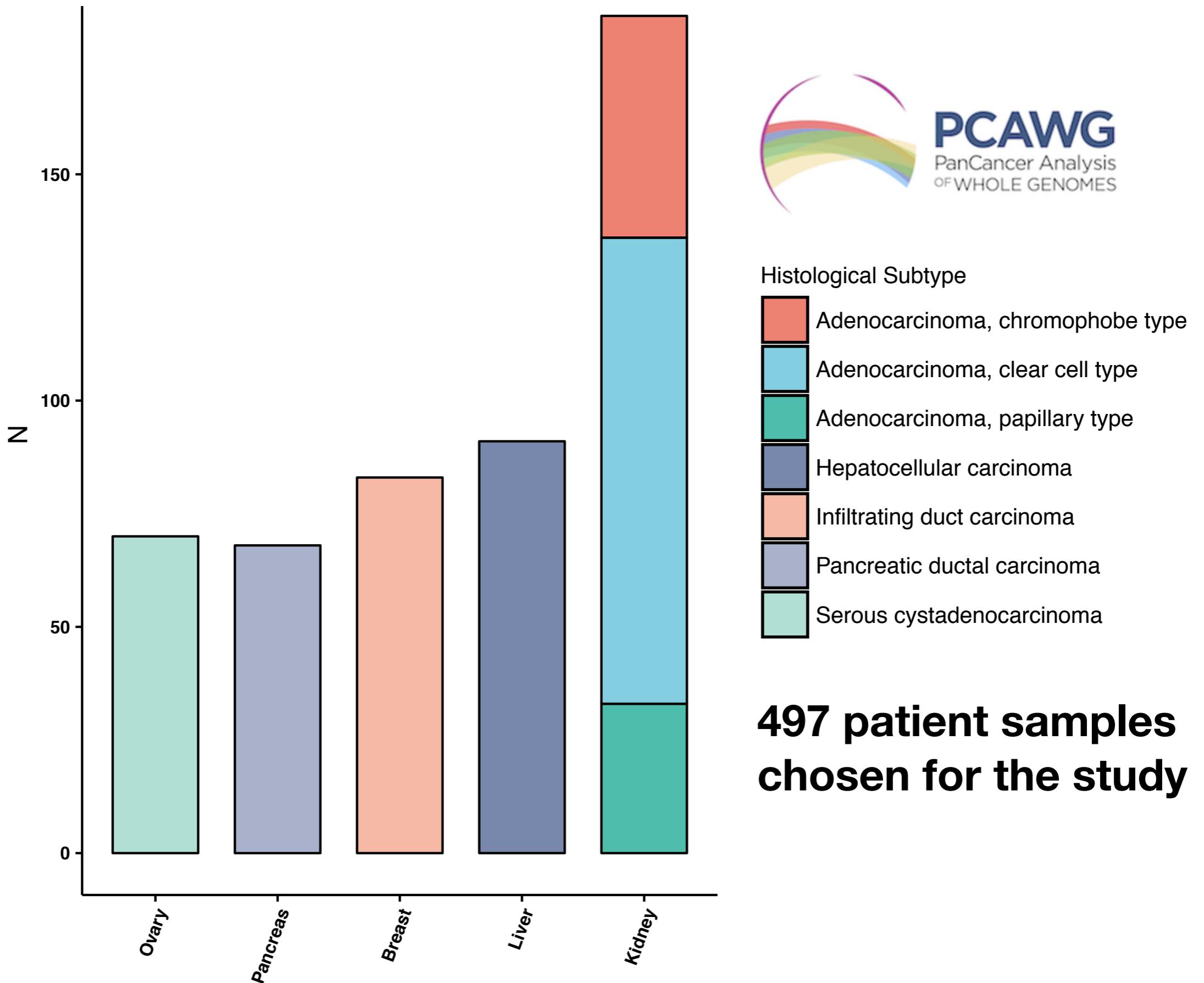
Methods



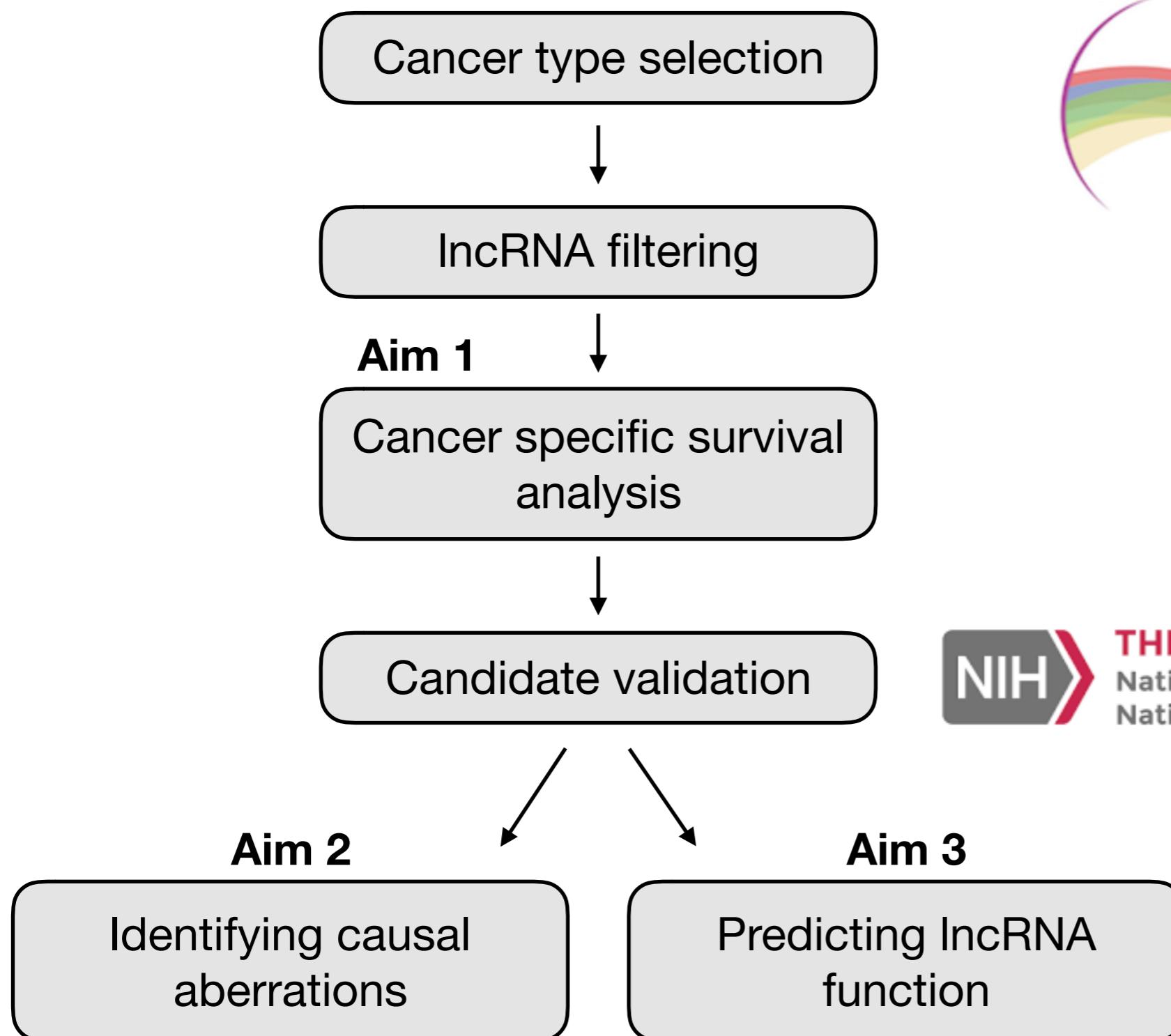
PCAWG
PanCancer Analysis
OF WHOLE GENOMES



Cancer type selection



Methods



PCAWG
PanCancer Analysis
OF WHOLE GENOMES

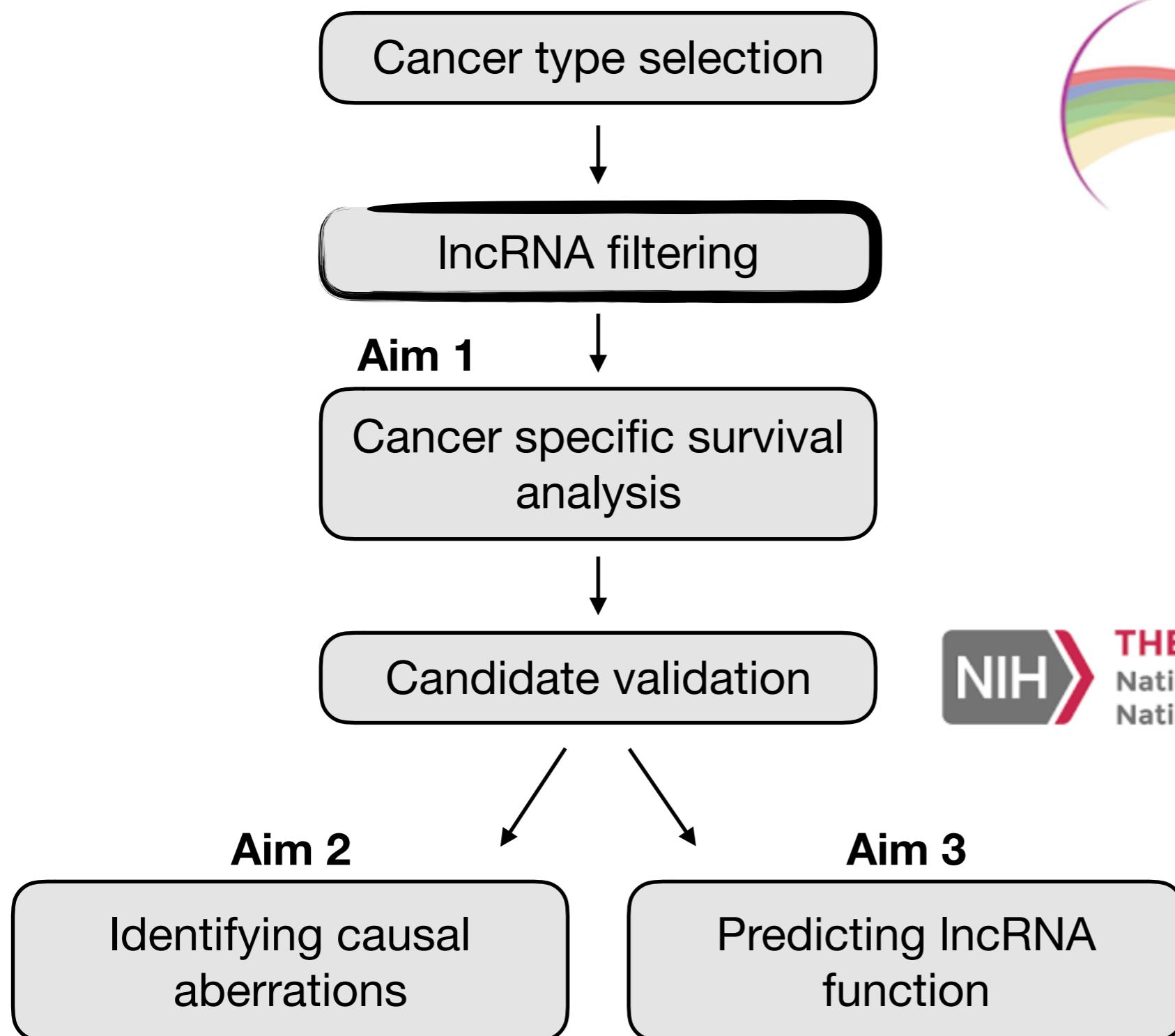


THE CANCER GENOME ATLAS

National Cancer Institute

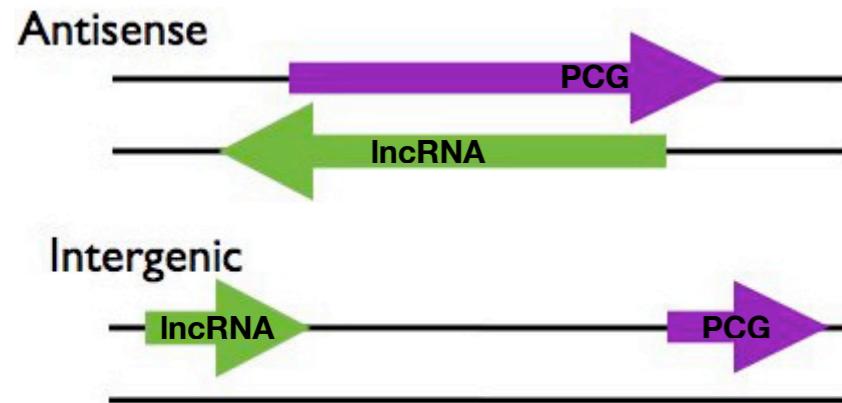
National Human Genome Research Institute

Methods



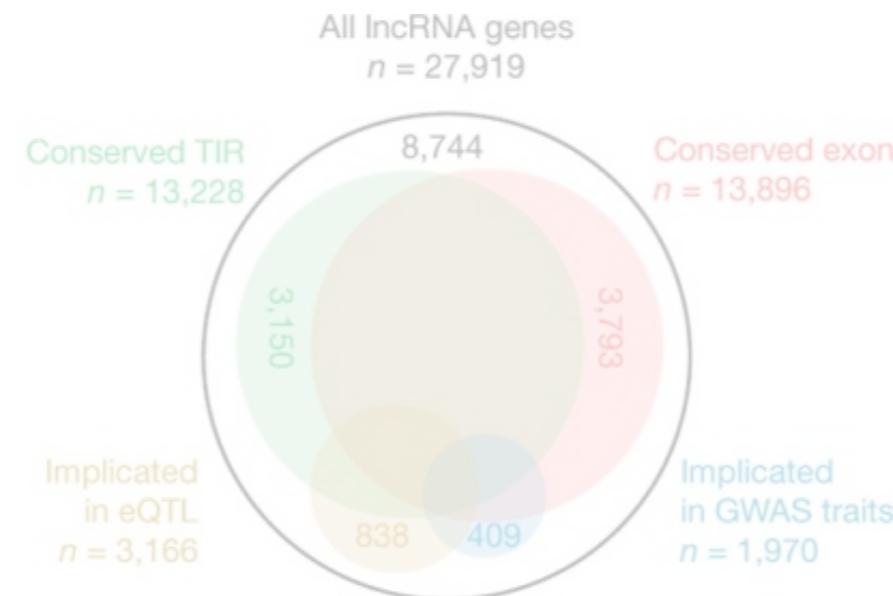
lncRNA filtering using PCAWG RNA-Seq Data

1. Antisense and Intergenic lncRNAs



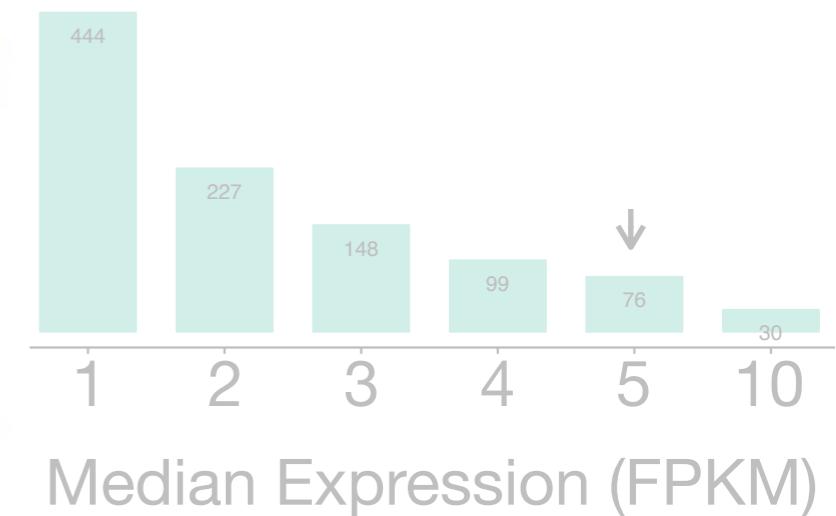
12,598 lncRNAs

2. FANTOM CAT annotation



5,607 lncRNAs

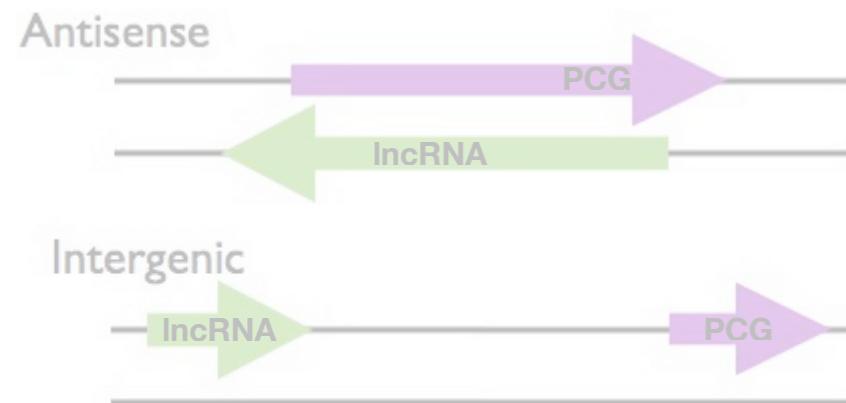
3. Median expression threshold



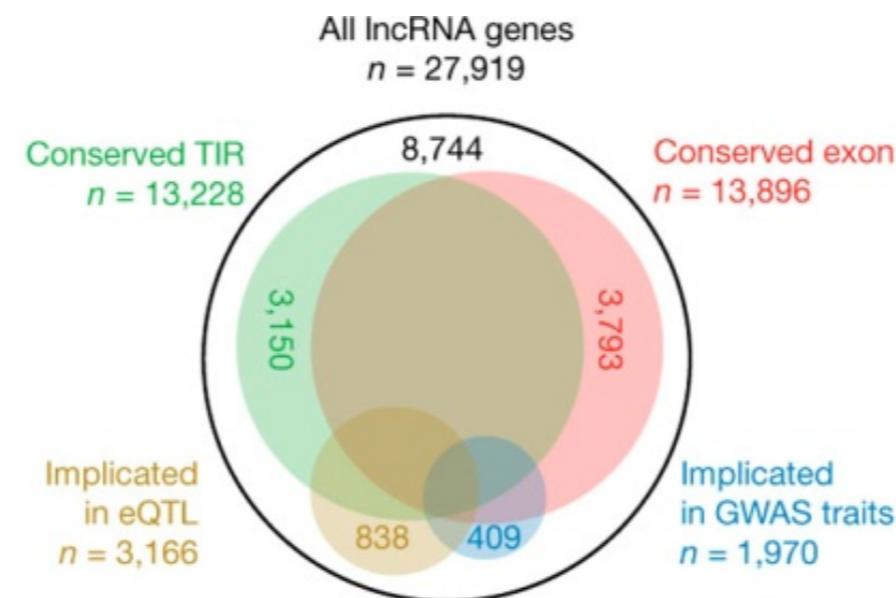
215 lncRNAs

lncRNA filtering using PCAWG RNA-Seq Data

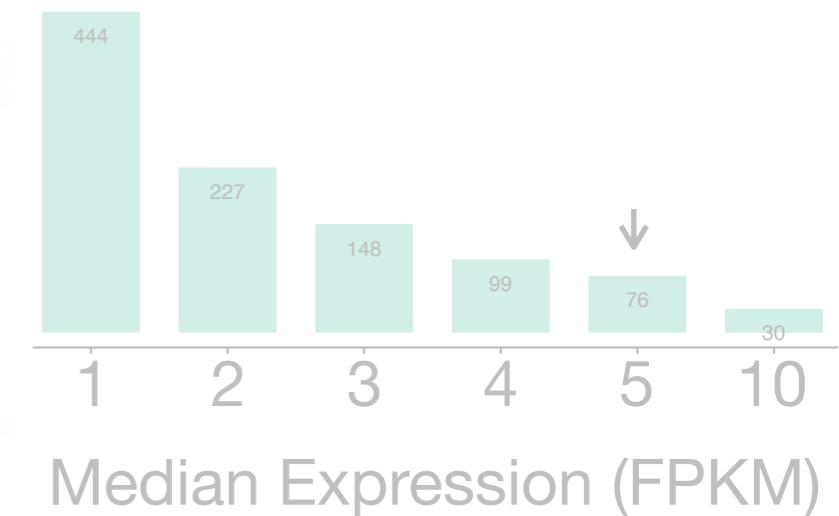
1. Antisense and Intergenic lncRNAs



2. FANTOM CAT annotation



3. Median expression threshold



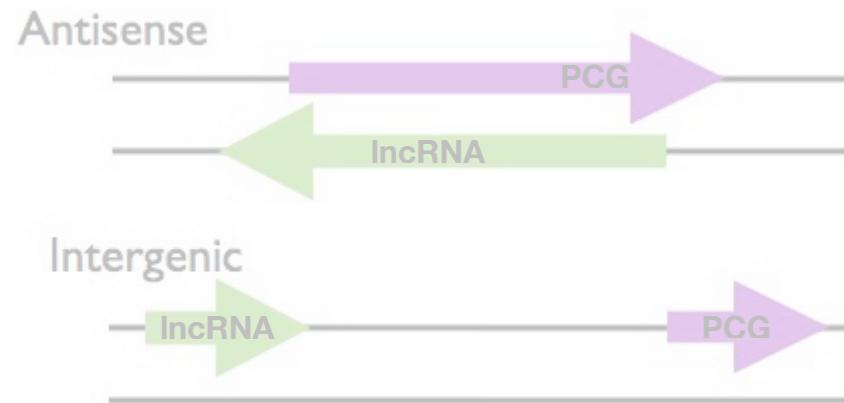
12,598 lncRNAs

5,607 lncRNAs

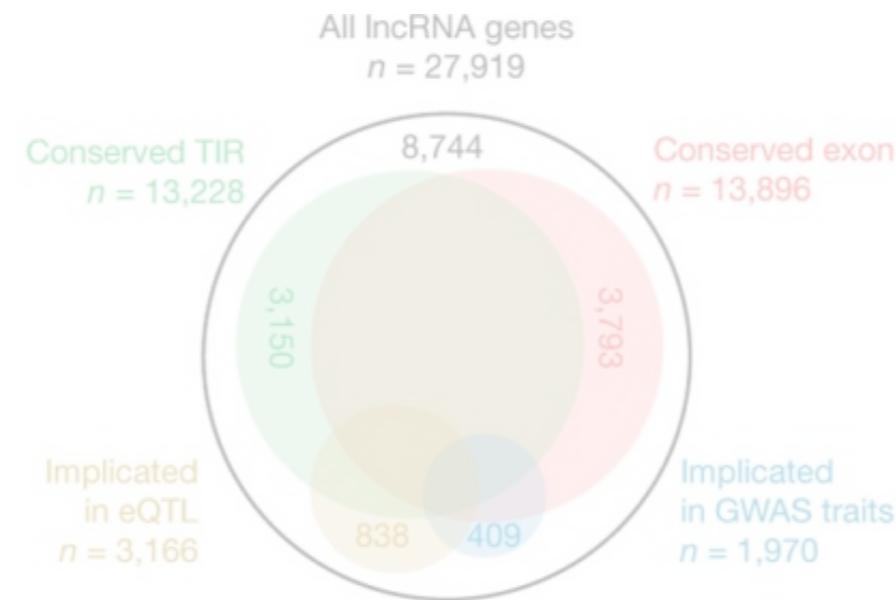
215 lncRNAs

lncRNA filtering using PCAWG RNA-Seq Data

1. Antisense and Intergenic lncRNAs



2. FANTOM CAT annotation

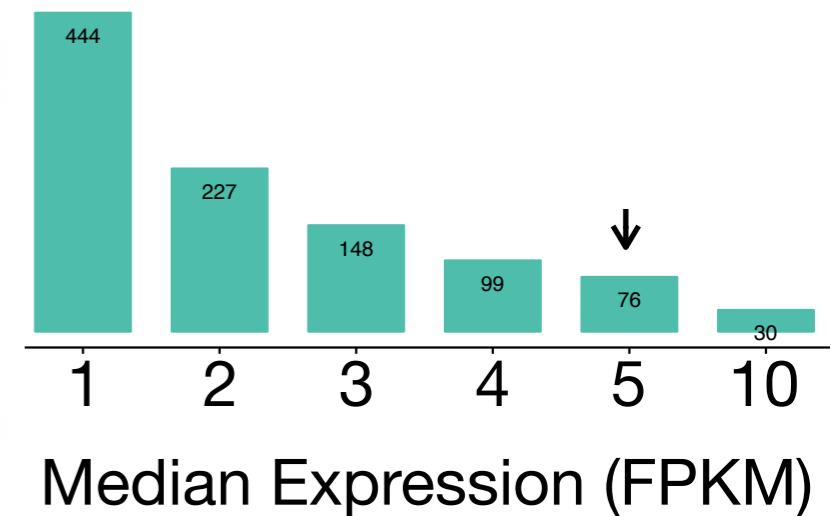


12,598 lncRNAs

5,607 lncRNAs

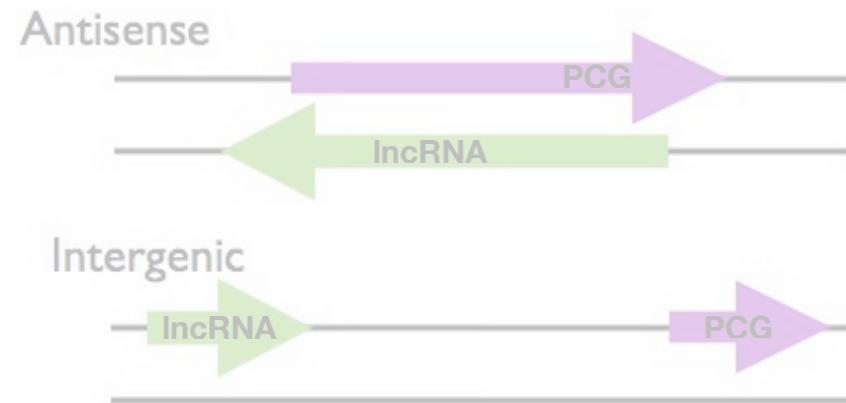
215 lncRNAs

3. Median expression threshold



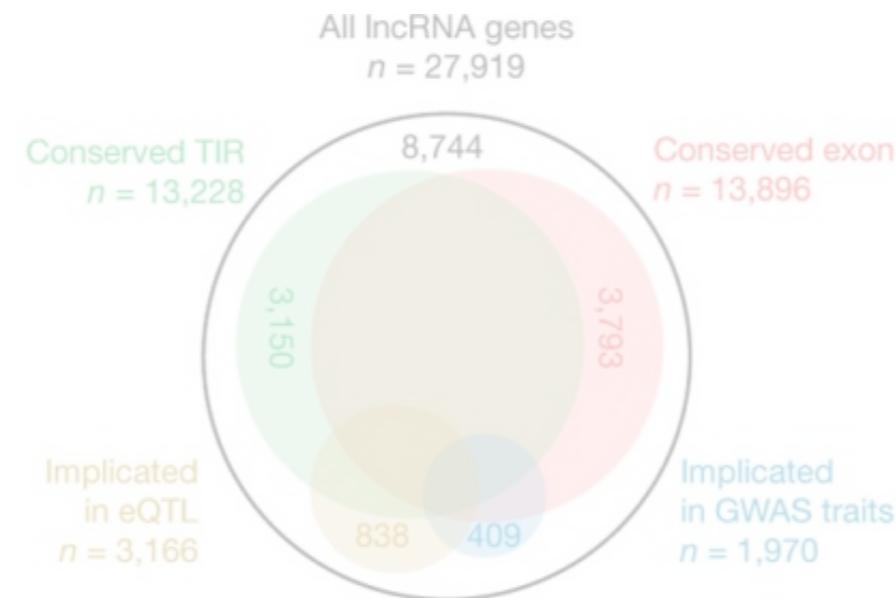
lncRNA filtering using PCAWG RNA-Seq Data

1. Antisense and Intergenic lncRNAs



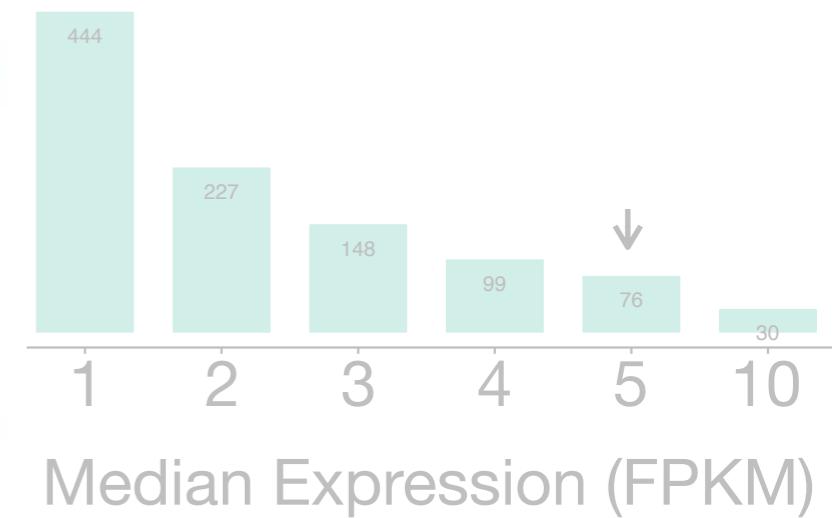
12,598 lncRNAs

2. FANTOM CAT annotation



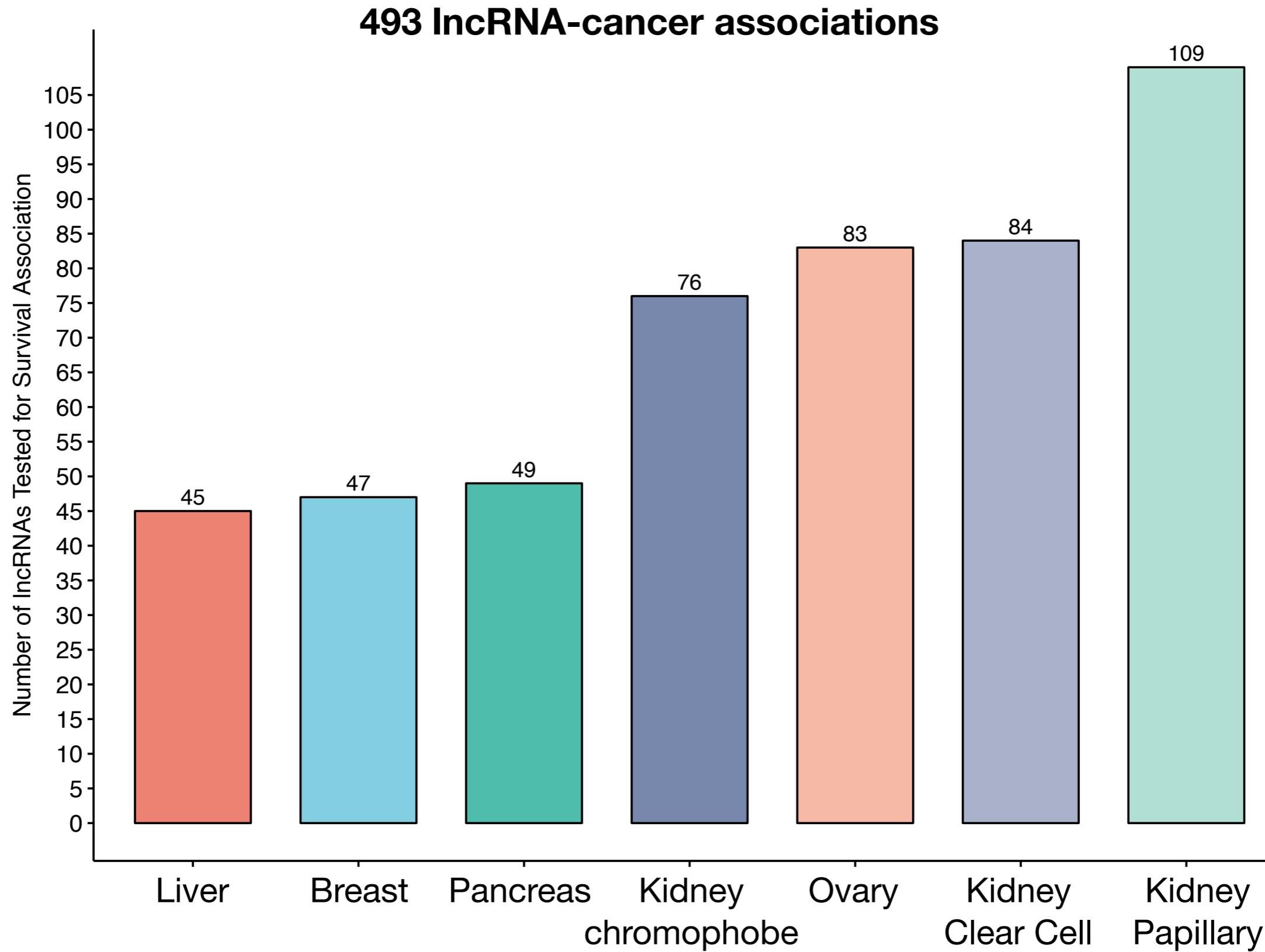
5,607 lncRNAs

3. Median expression threshold

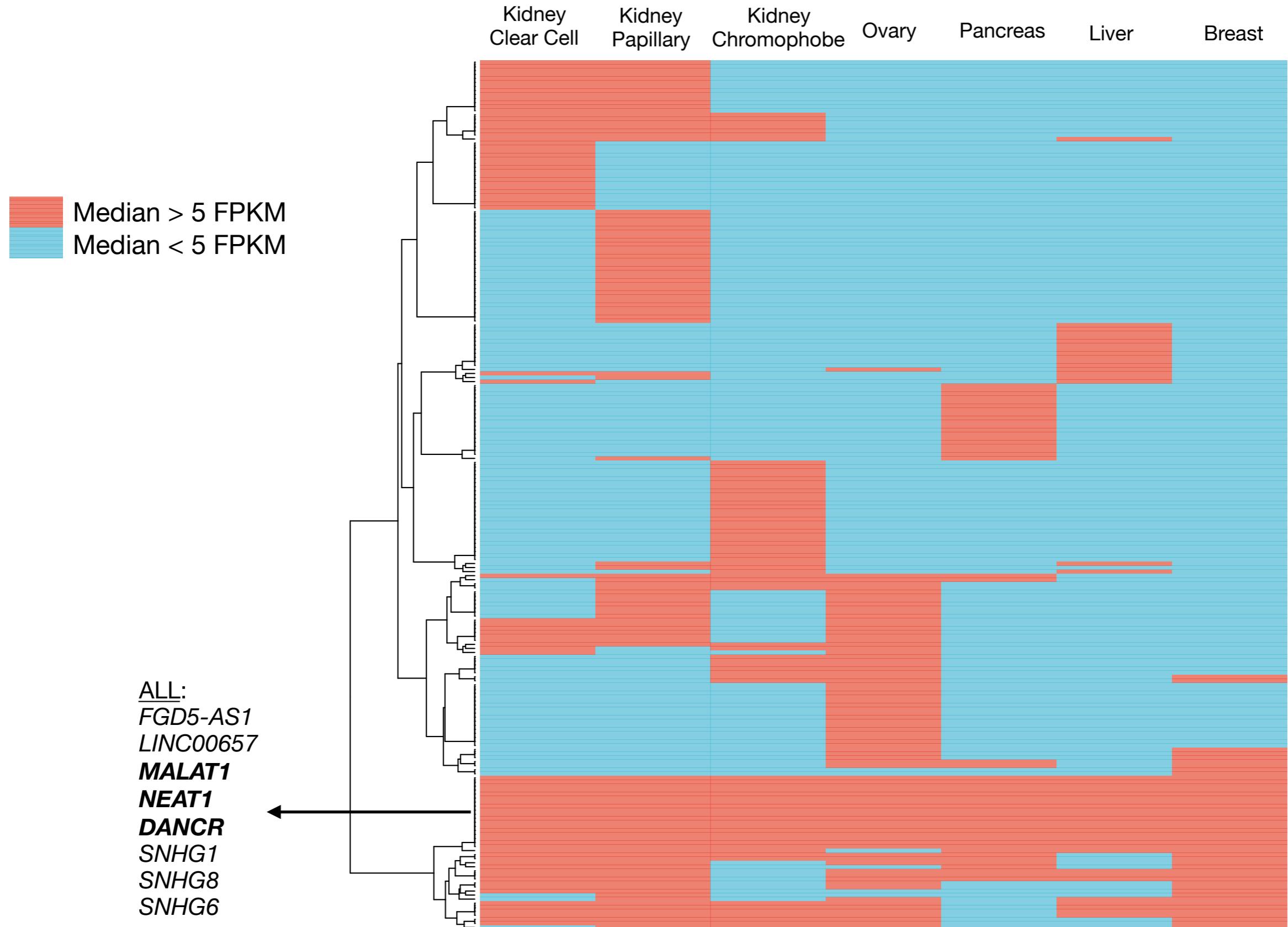


215 lncRNAs

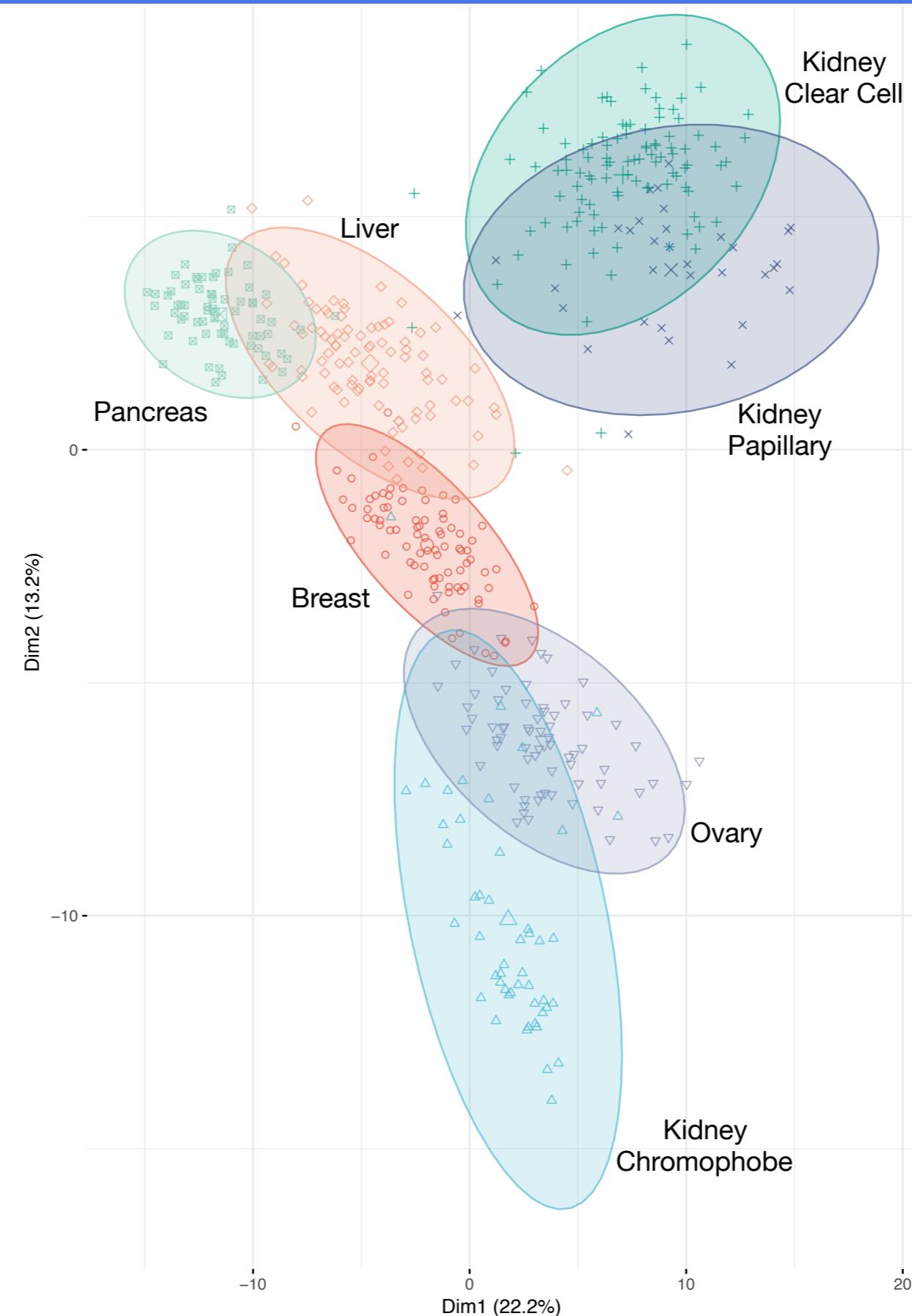
215 highly expressed lncRNAs chosen for analysis



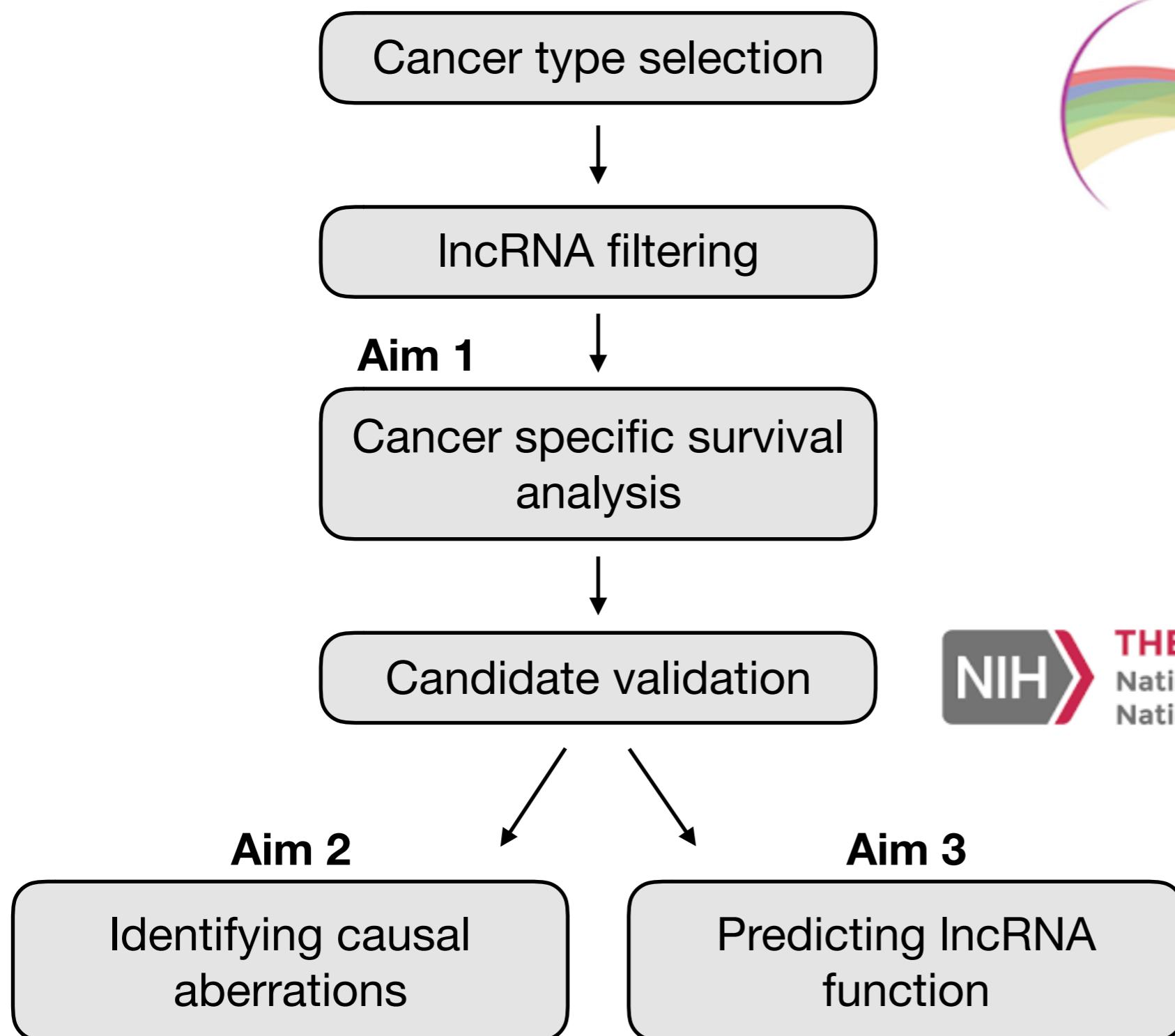
215 highly expressed lncRNAs across cancer types



Clustering of cancer types using 215 lncRNAs



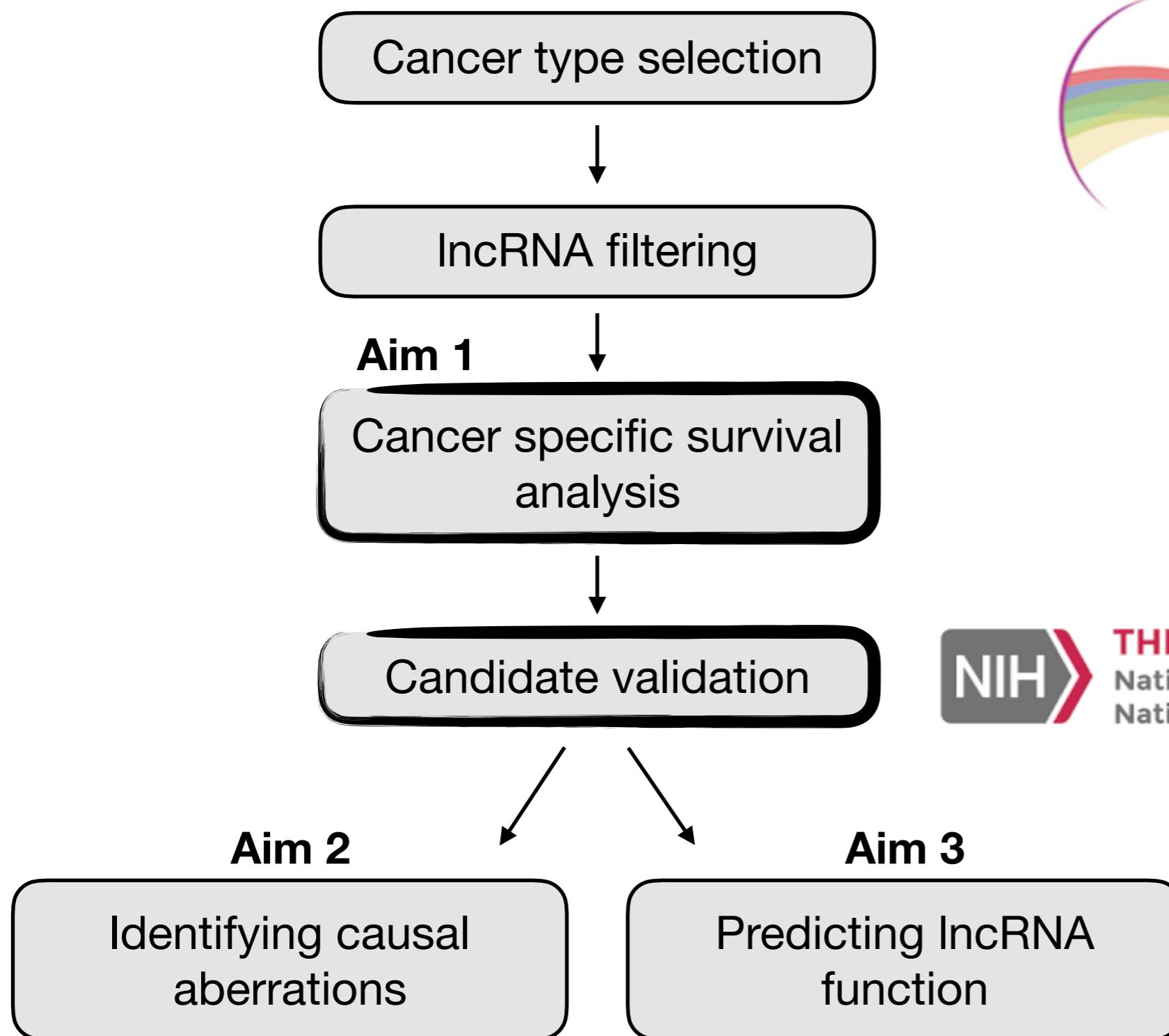
Methods



PCAWG
PanCancer Analysis
OF WHOLE GENOMES



Methods



Cancer specific survival analysis

- ▶ 493 lncRNA-cancer associations tested (215 unique lncRNAs)
- ▶ Cox proportional-hazards model
$$\text{coxph}(\text{Surv}(time\ to\ death, status) \sim lncRNA\ expression)$$
- ▶ Median expression dichotomization
- ▶ Multiple testing correction within each cancer type
- ▶ Validation of associations in TCGA

Cancer specific survival analysis

- ♦ lncRNAs with significant (FDR < 0.05) associations to survival outcomes
 - n = 25 lncRNAs among liver and ovarian cancers
 - n = 7 validated in matched TCGA cohorts
1. ***NEAT1*** in liver cancer
2. ***OTUD6B-AS1*** in ovarian cancer

Cancer specific survival analysis

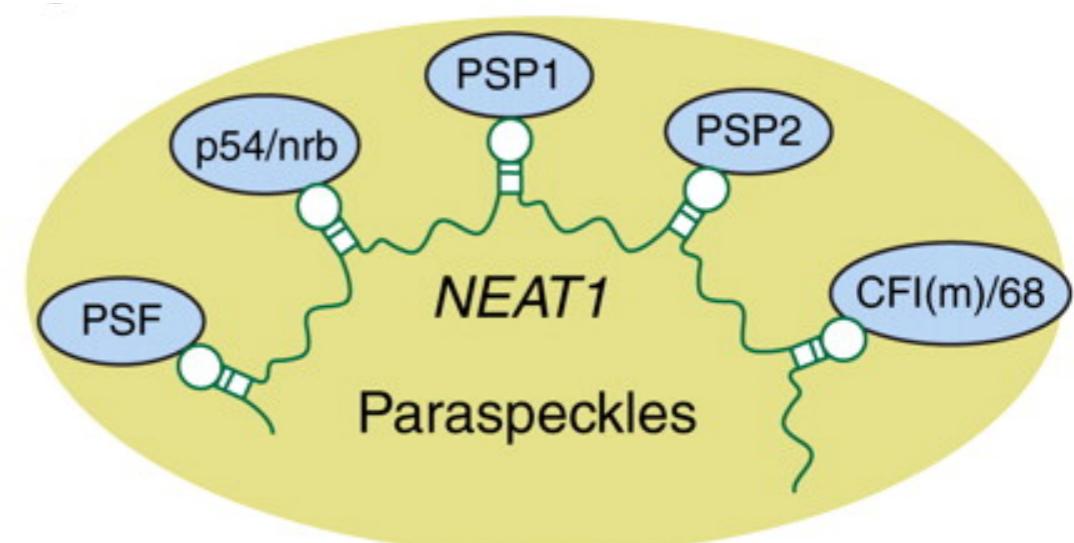
- ♦ lncRNAs with significant ($FDR < 0.05$) associations to survival outcomes
 - $n = 25$ lncRNAs among liver and ovarian cancers
 - $n = 7$ validated in matched TCGA cohorts

1. ***NEAT1*** in liver cancer

2. ***OTUD6B-AS1*** in ovarian cancer

NEAT1 in the literature

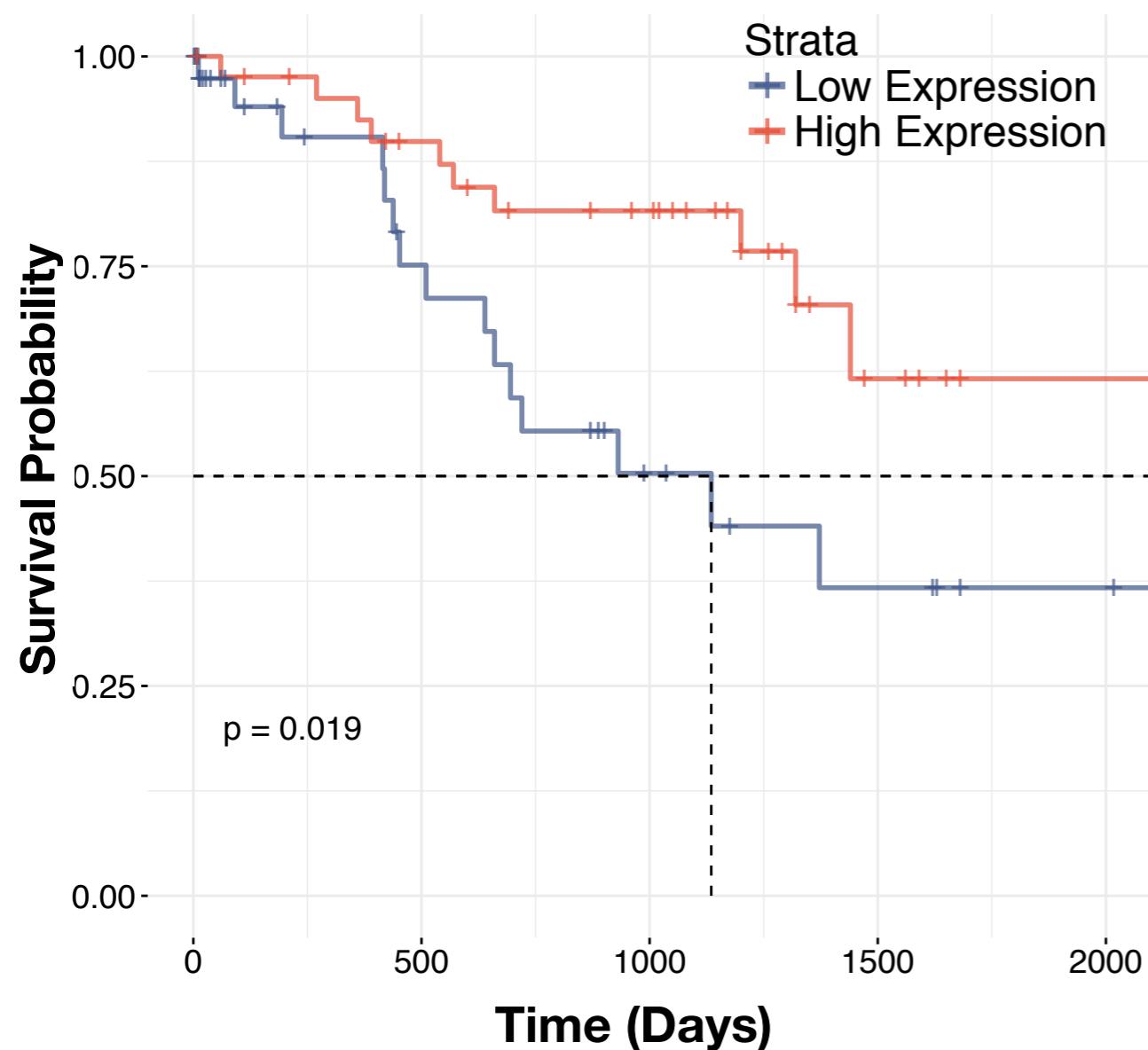
- ♦ Nuclear-Enriched Abundant Transcript 1 (*NEAT1*)
- ♦ **Required for the formation of parasplices (nuclear bodies)**
 - Comprised of *NEAT1* and splicing factors
- ♦ *NEAT1* can localize to hundreds of genomic sites in human cells
- ♦ *NEAT1* deficiency enhances transformation in oncogene-expressing fibroblasts



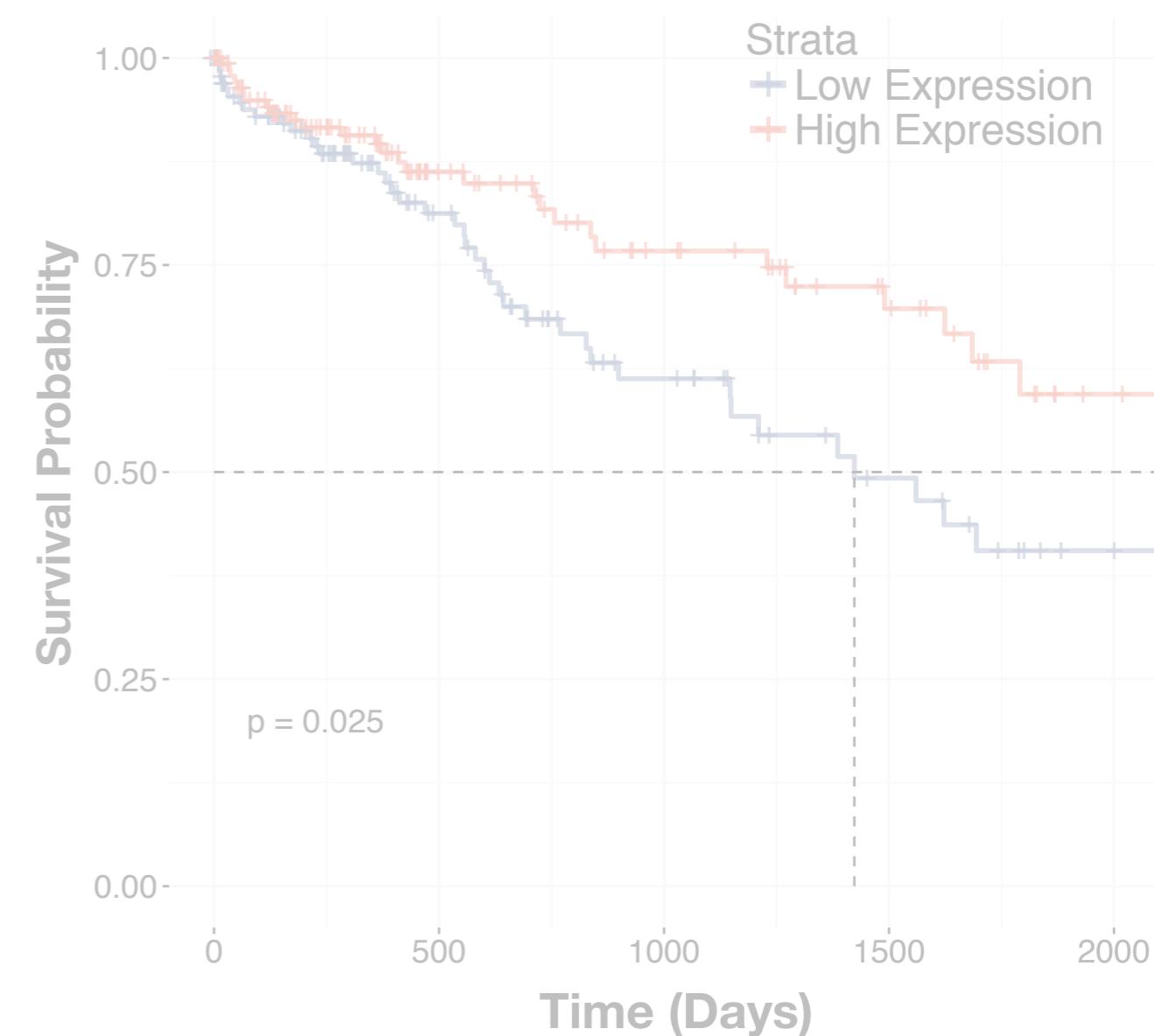
Wang et al. Cold Spring Harb Perspect Biol (2010)

NEAT1 in Liver Cancer

PCAWG, n = 84



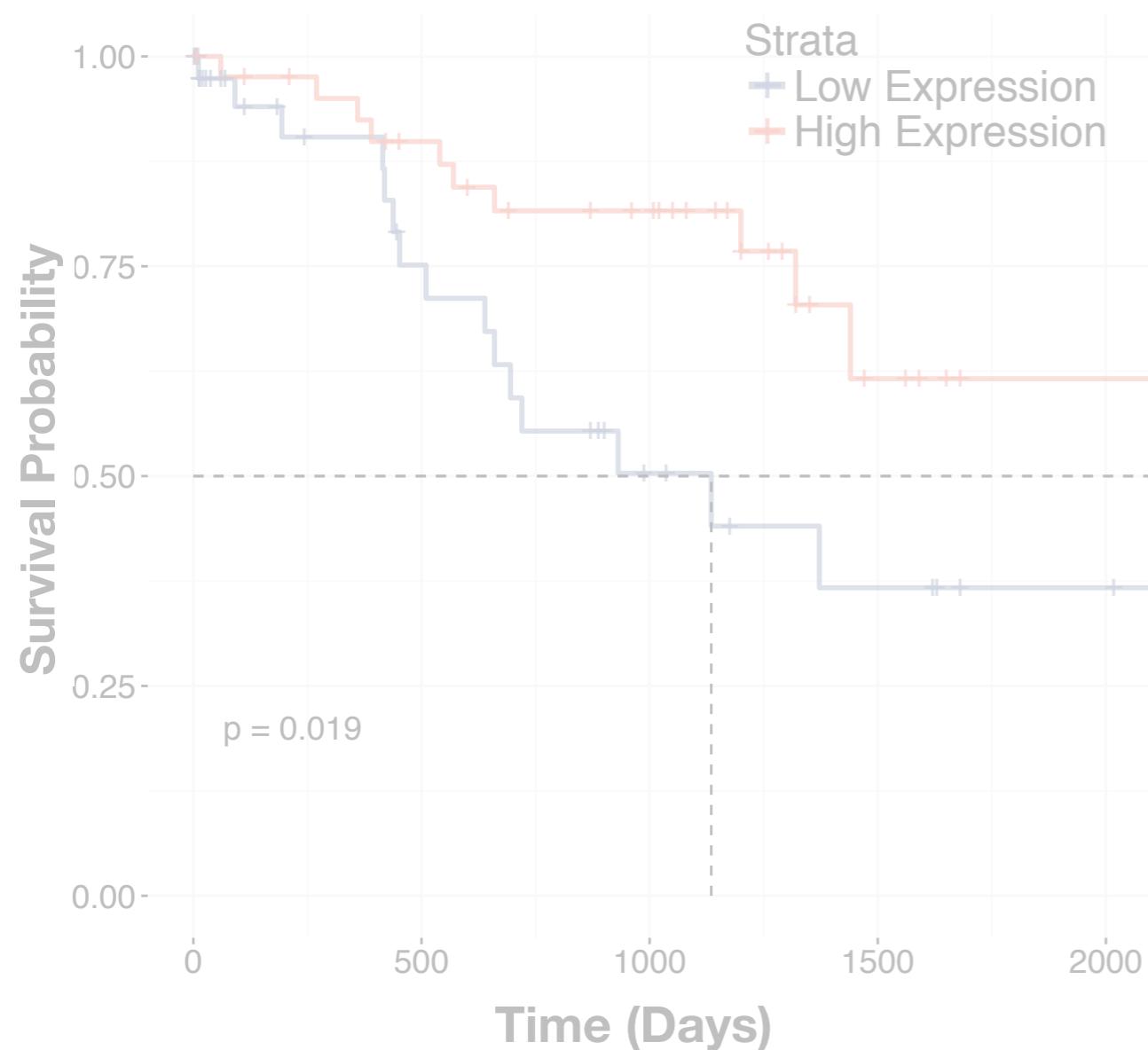
TCGA, n = 317



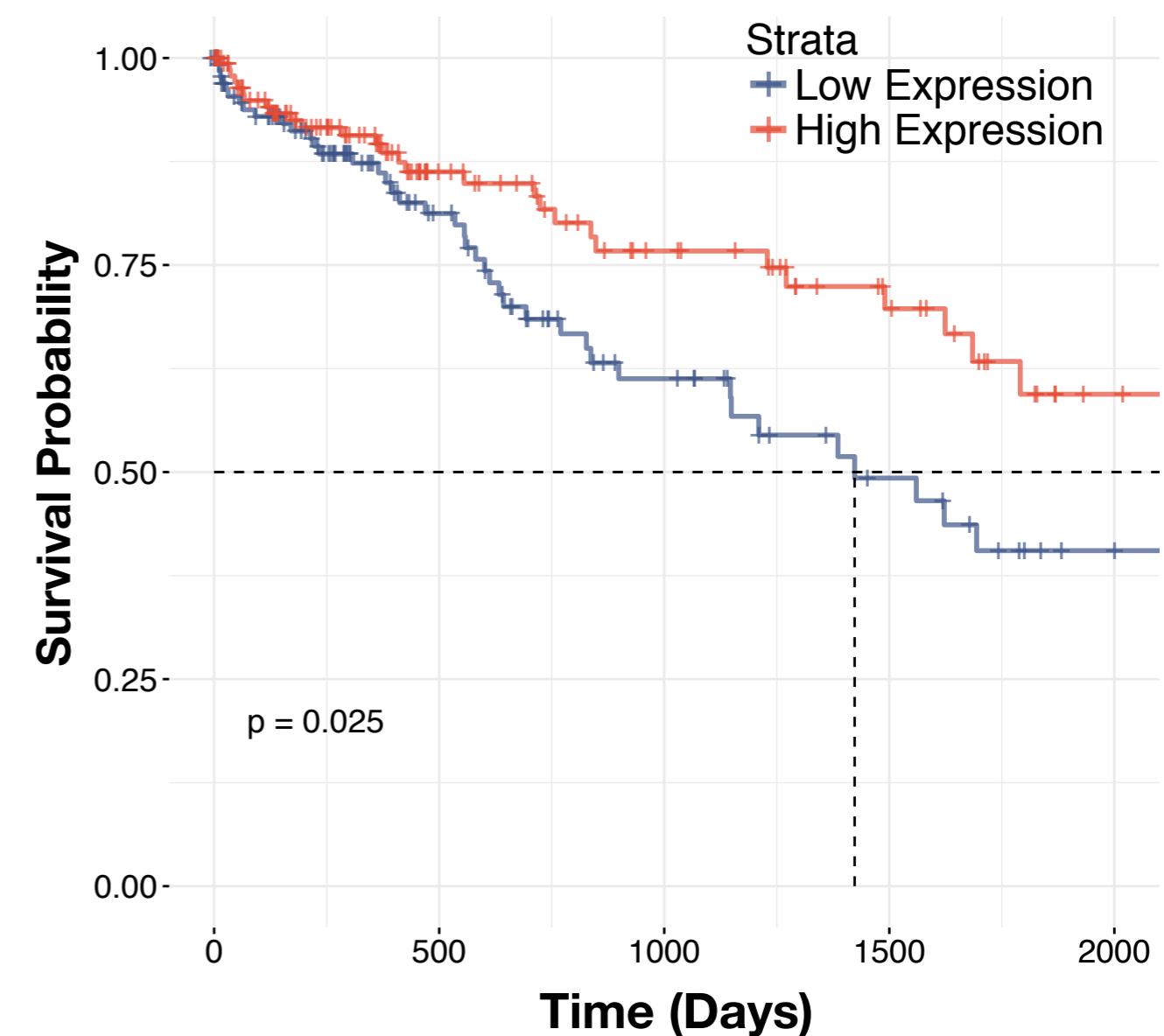
Higher expression of *NEAT1* has a protective effect

NEAT1 in Liver Cancer

PCAWG, n = 84



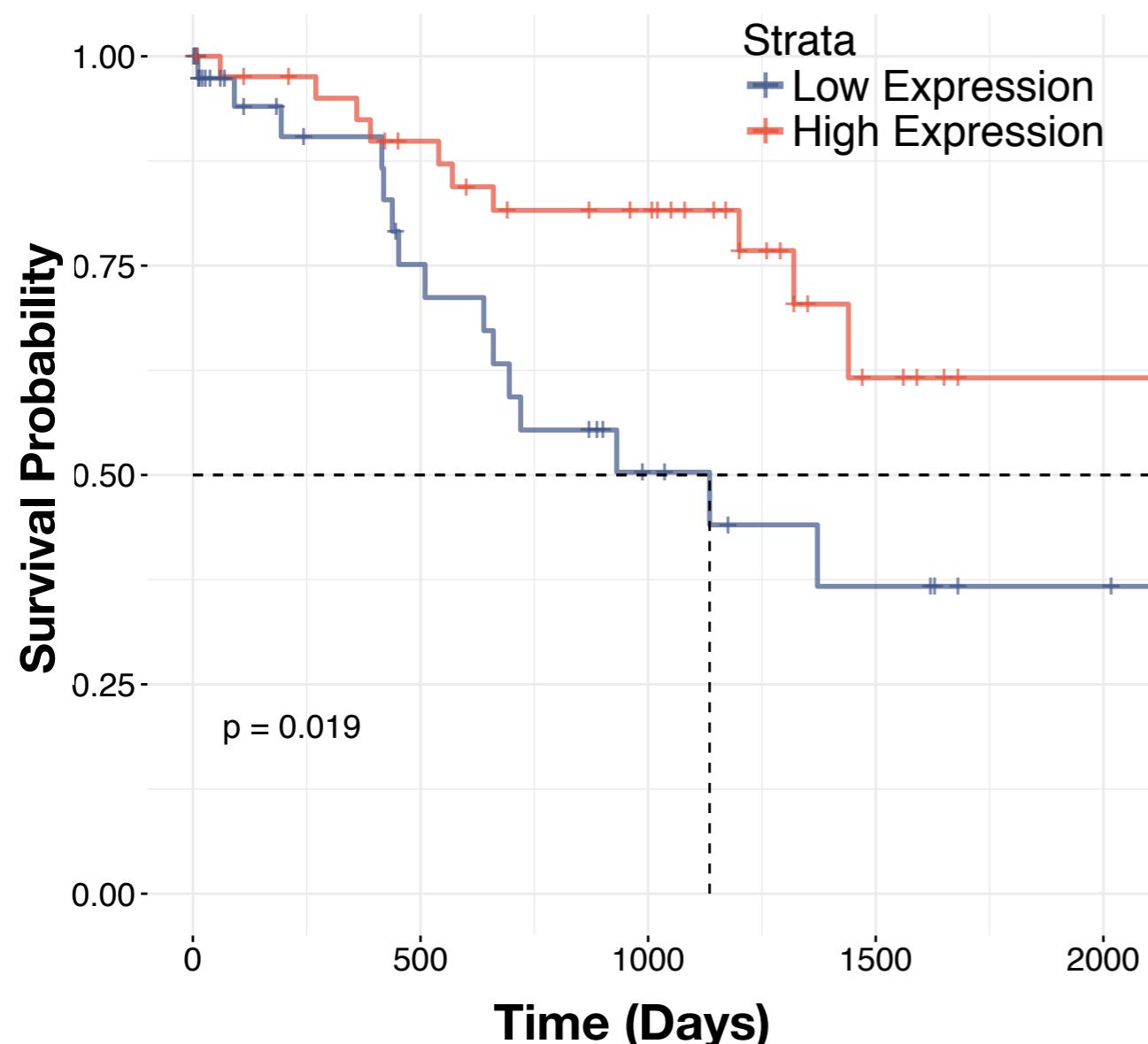
TCGA, n = 317



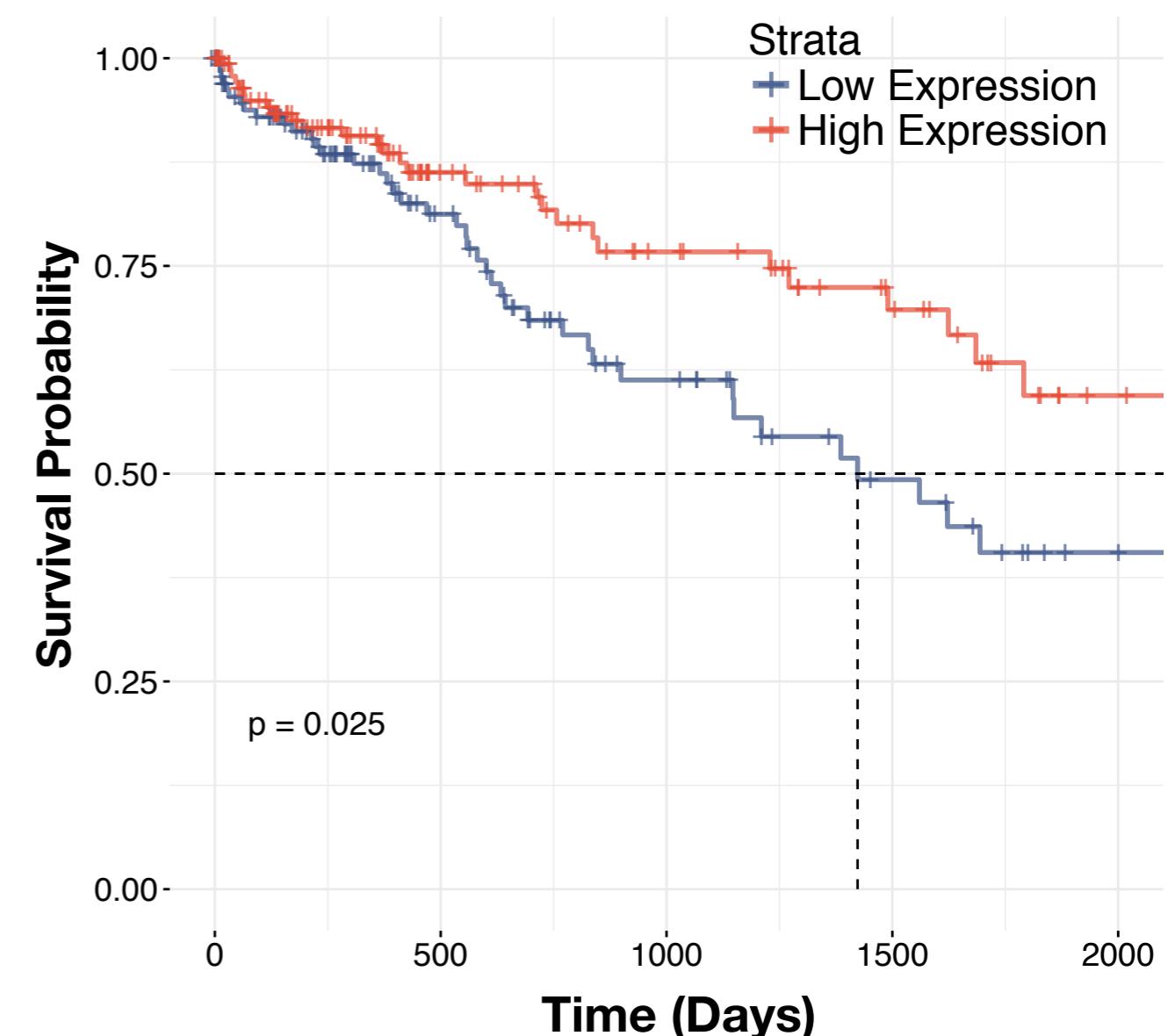
Higher expression of *NEAT1* has a protective effect

NEAT1 in Liver Cancer

PCAWG, n = 84



TCGA, n = 317



Higher expression of *NEAT1* has a protective effect

lncRNA expression in normal tissues (GTEx)

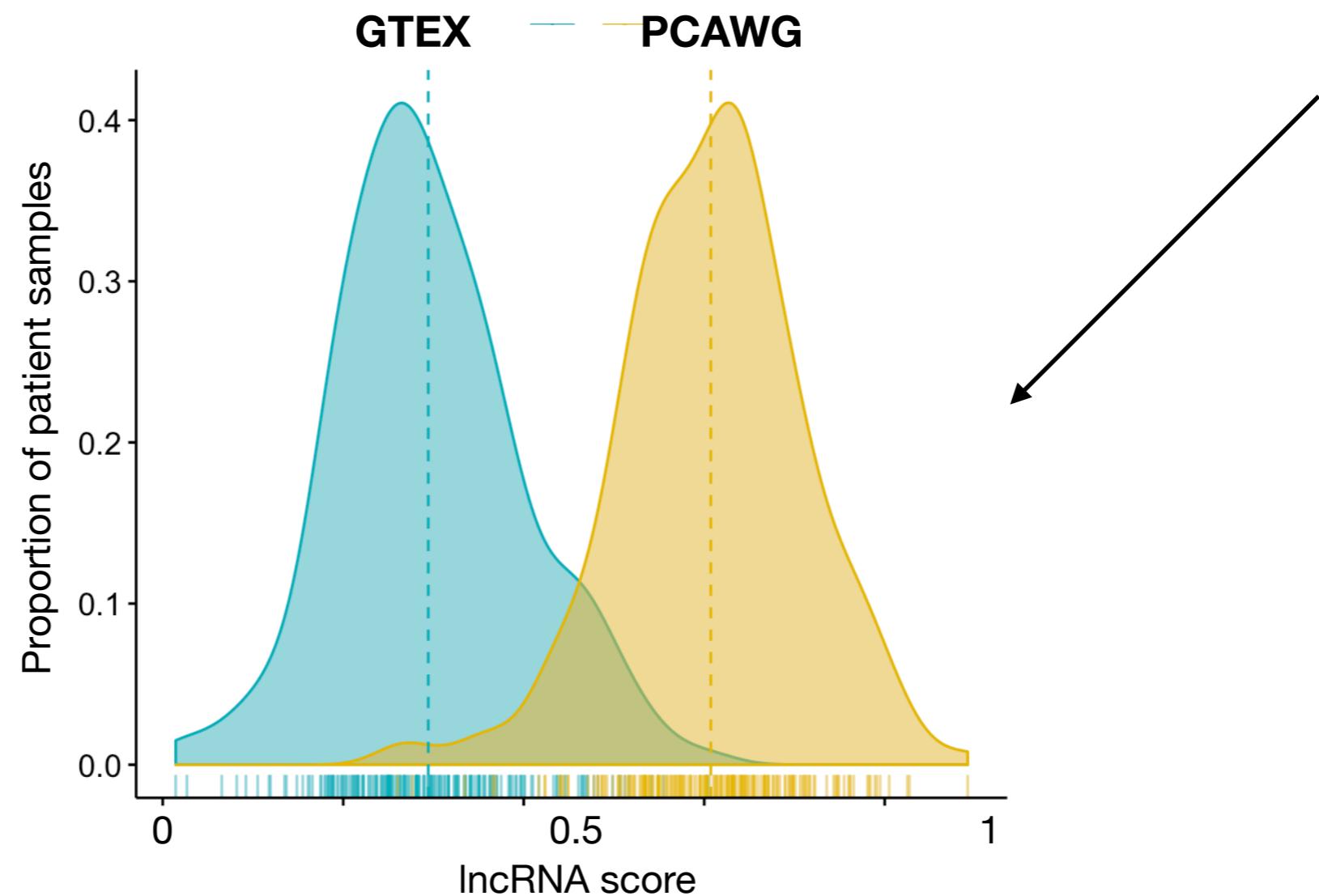
Within each patient sample, order all genes based on expression to get rank

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

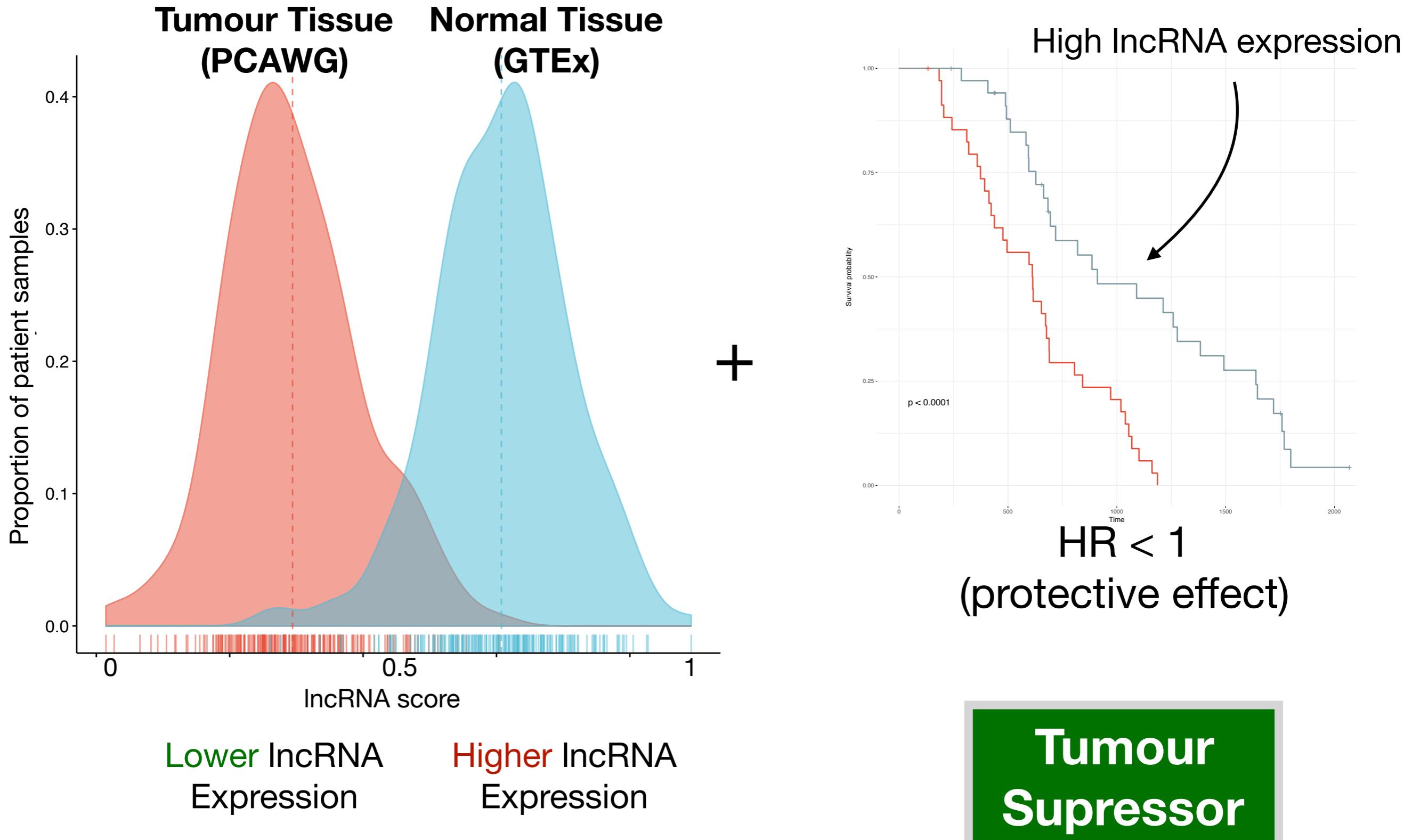


Divide rank of gene by length of total genes in list to get score

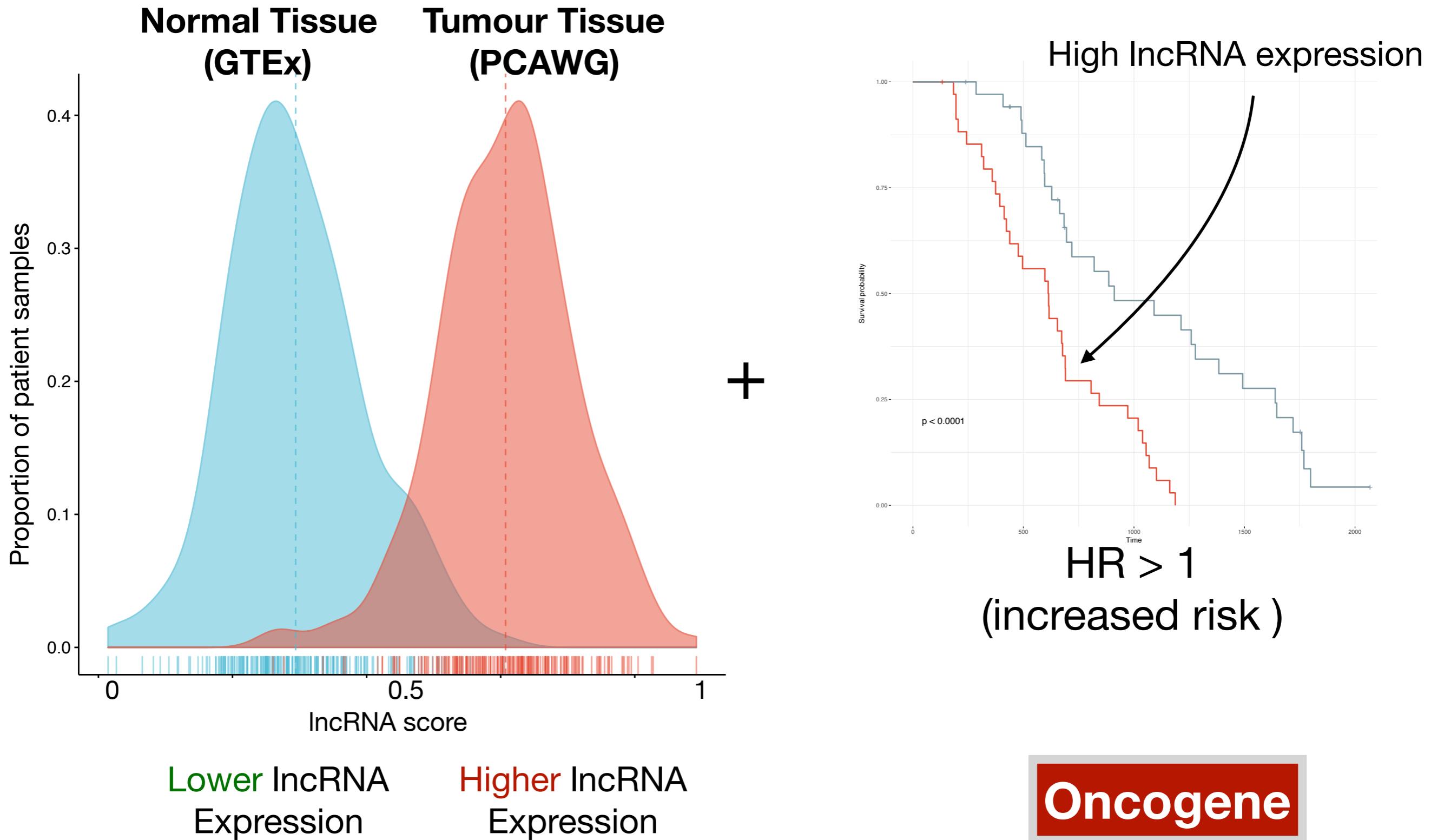
- 1/25,773
2/25,773
.
.
.
25,772/25,773
25,773/25,773



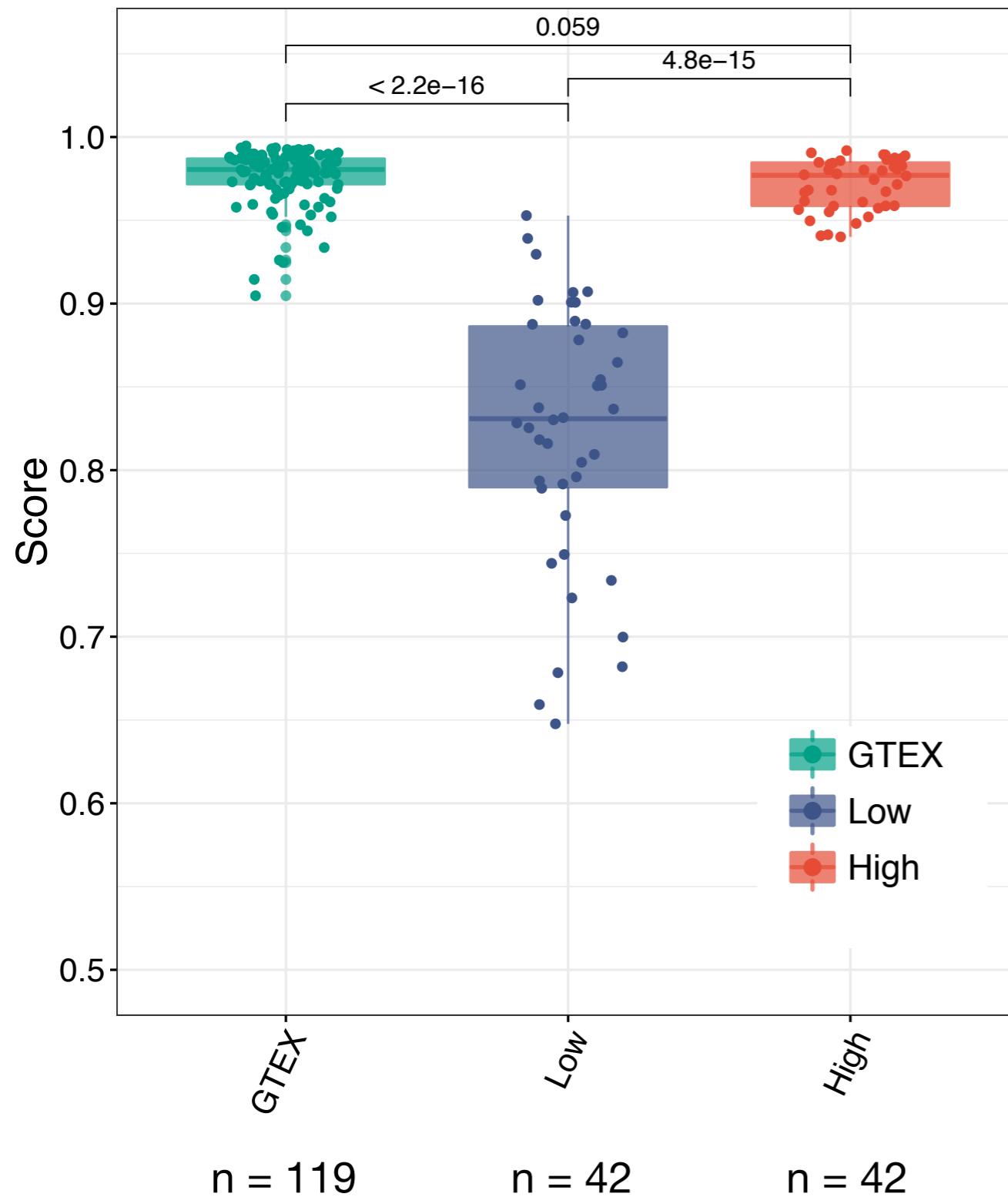
Predicting NEAT1's role in Liver Cancer



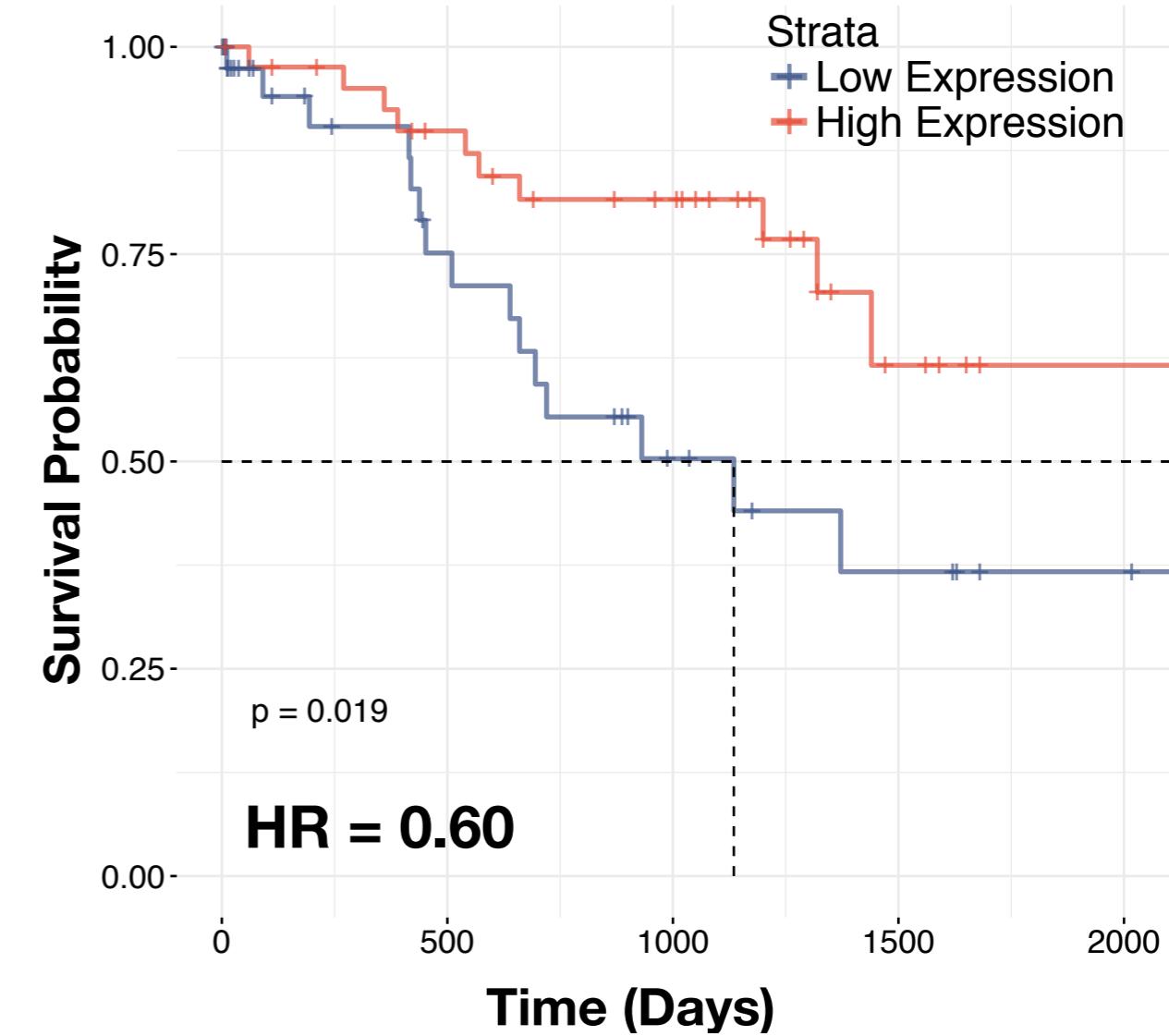
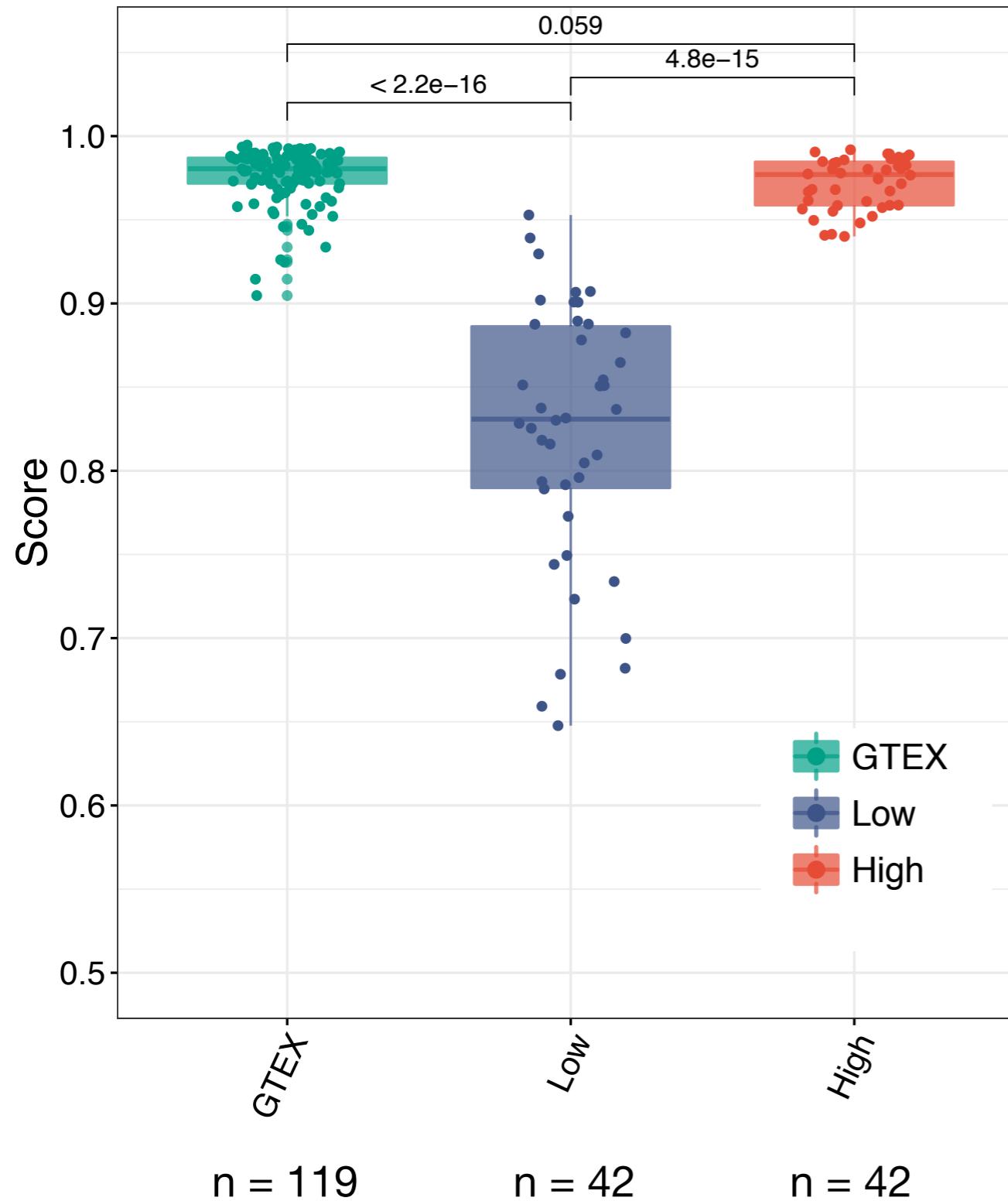
Predicting NEAT1's role in Liver Cancer



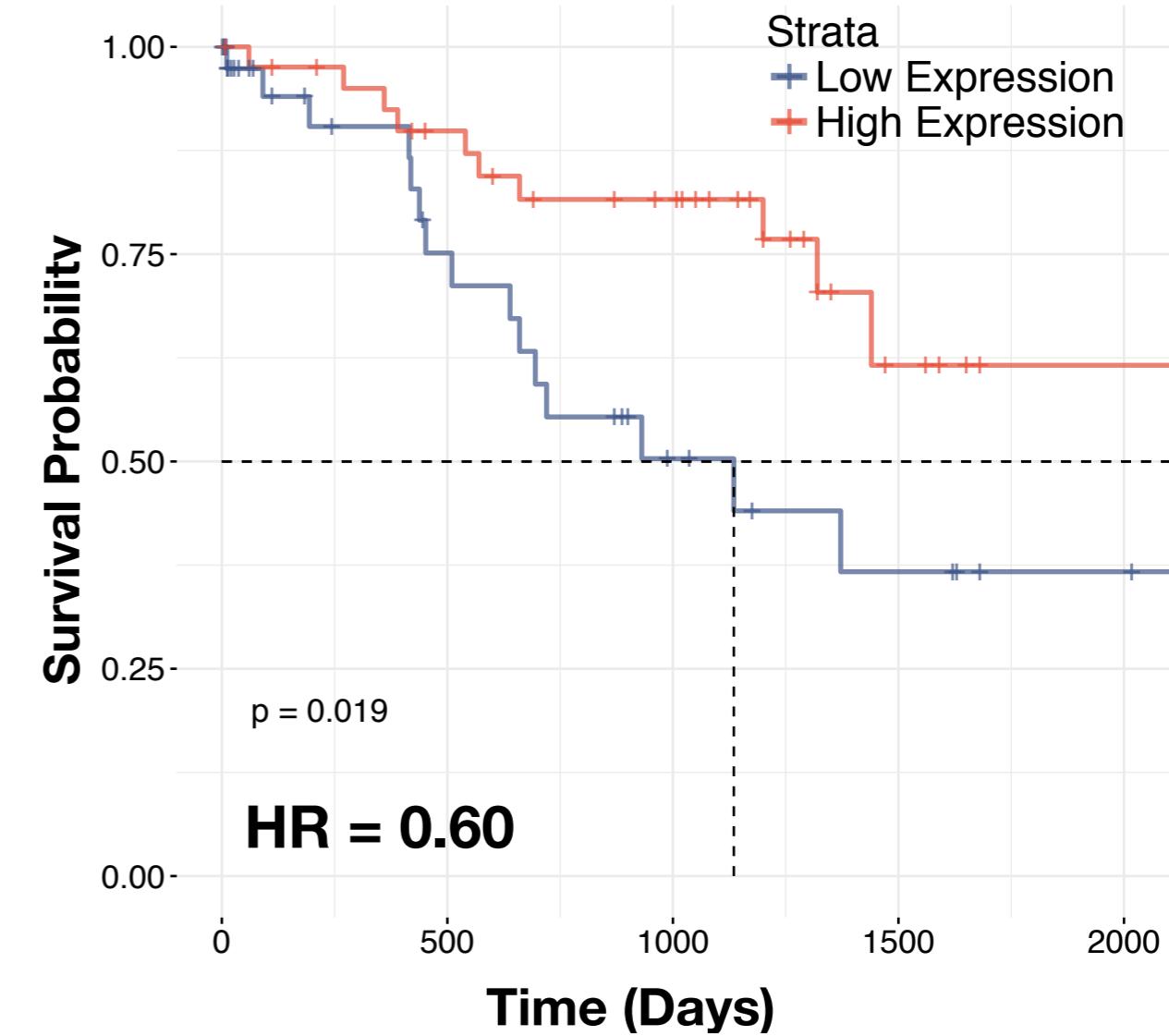
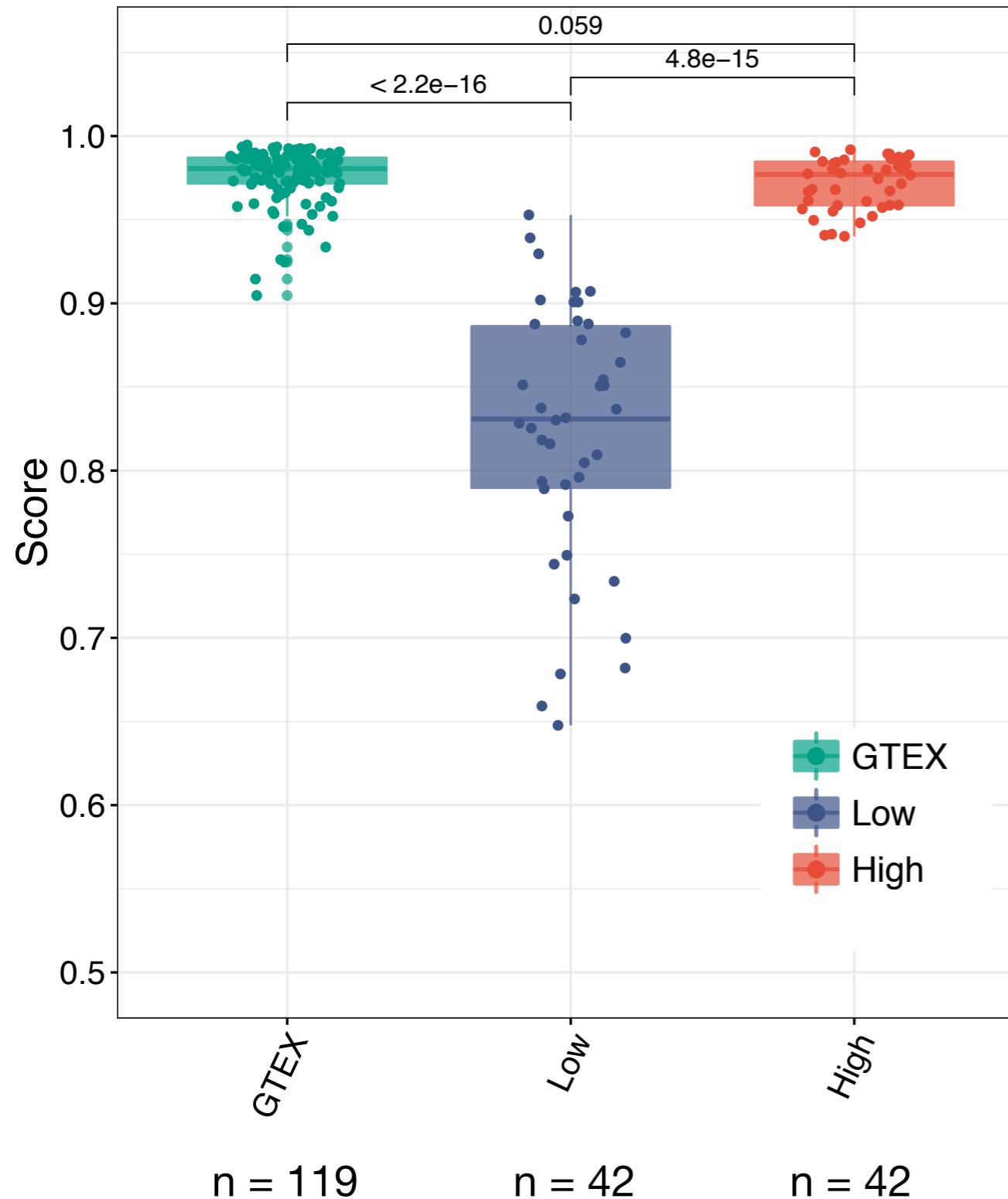
NEAT1 in Liver Cancer



NEAT1 in Liver Cancer

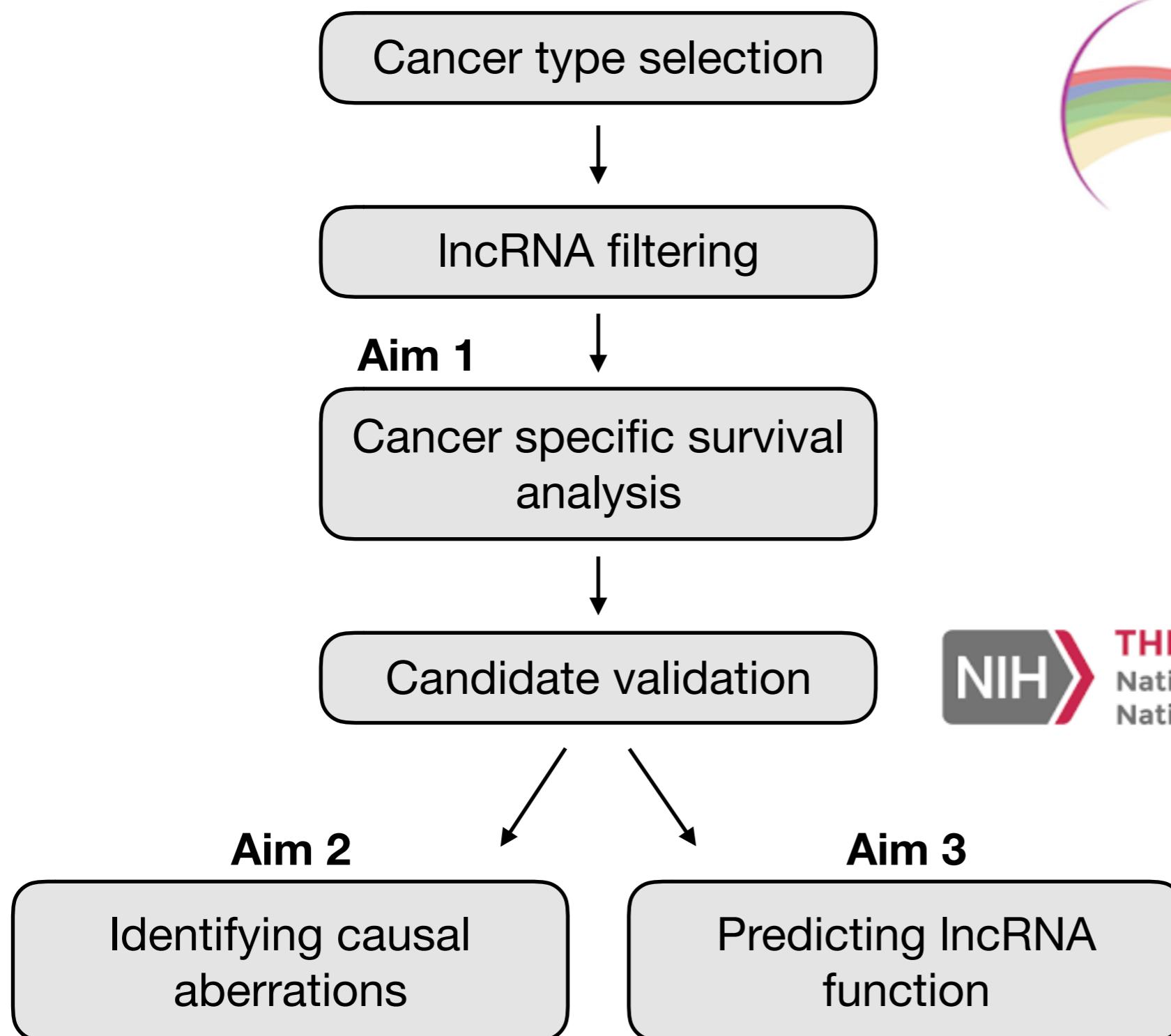


NEAT1 in Liver Cancer

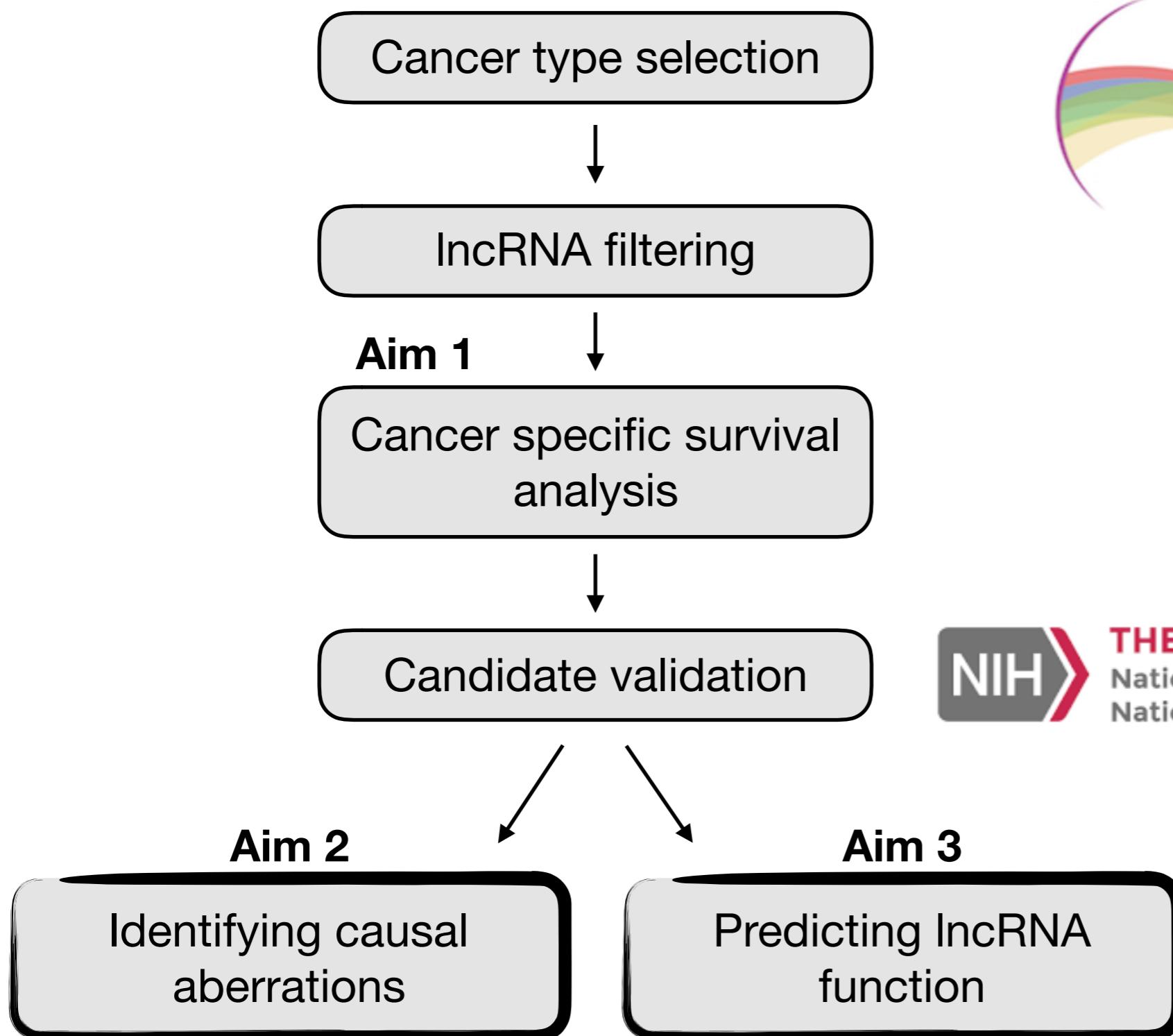


Predicted Tumour
Suppressor

Methods



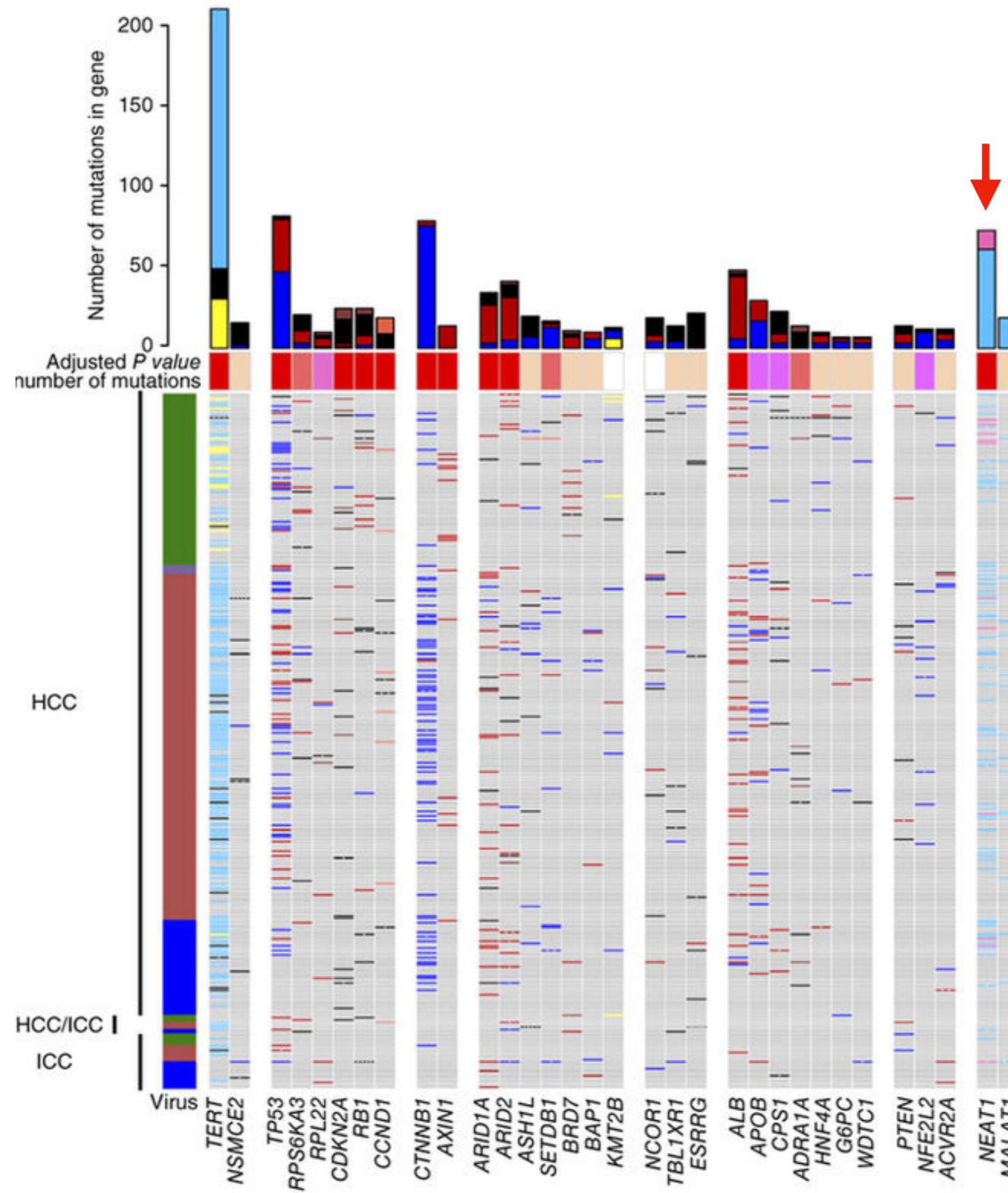
Methods



PCAWG
PanCancer Analysis
OF WHOLE GENOMES

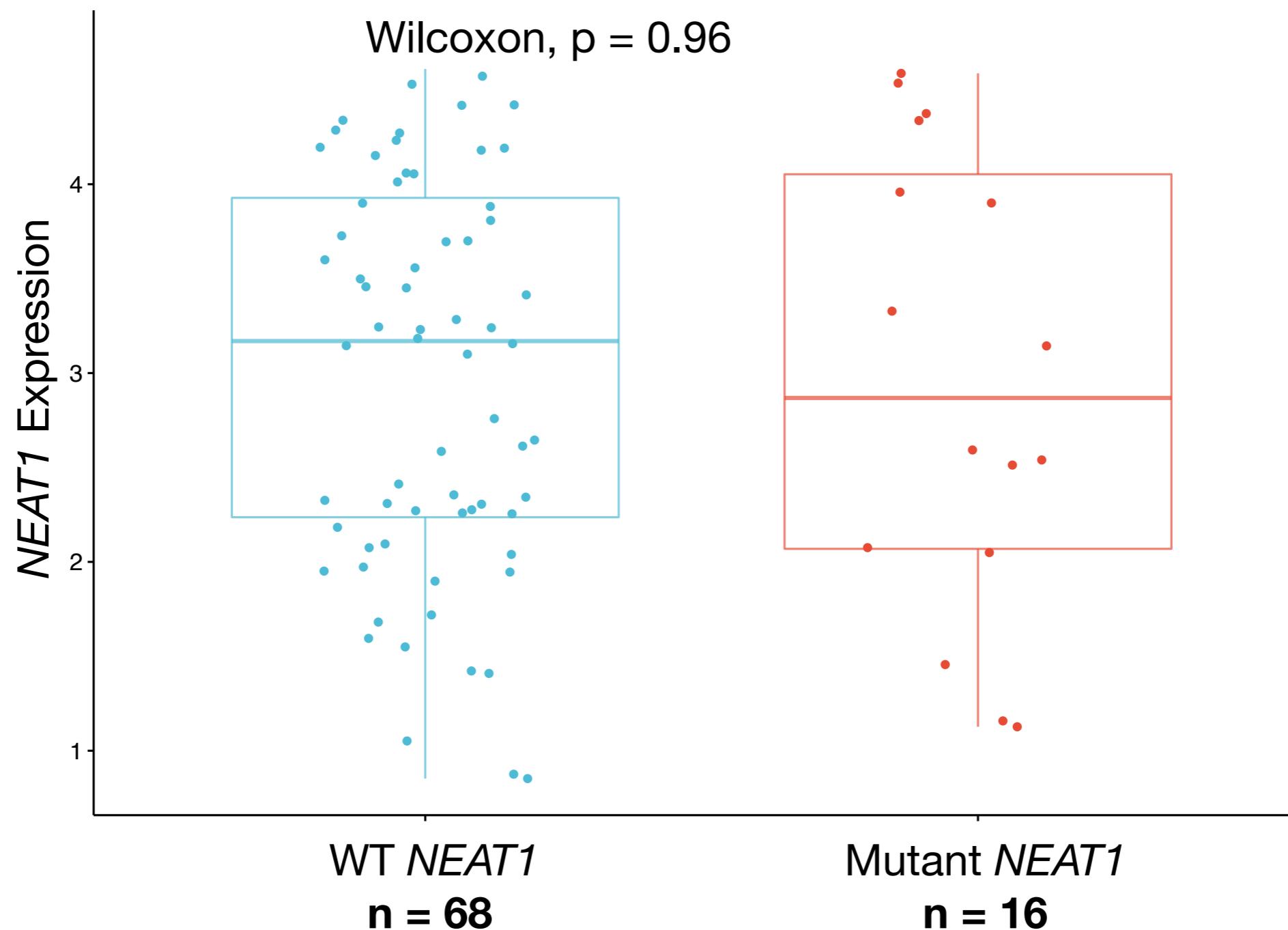


NEAT1 is mutated in ~20% of HCC patients

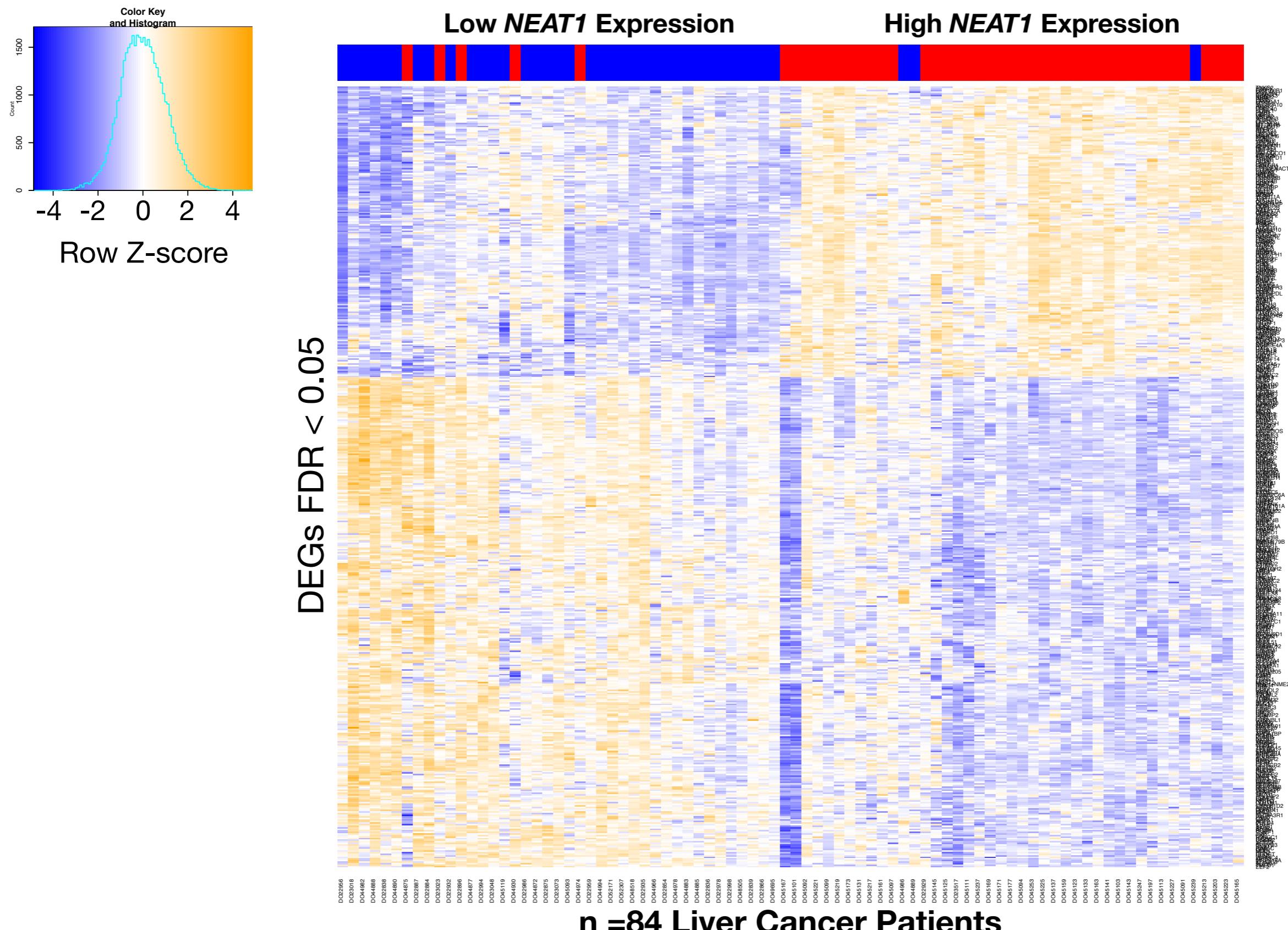


NEAT1 is mutated in ~20% of HCC patients

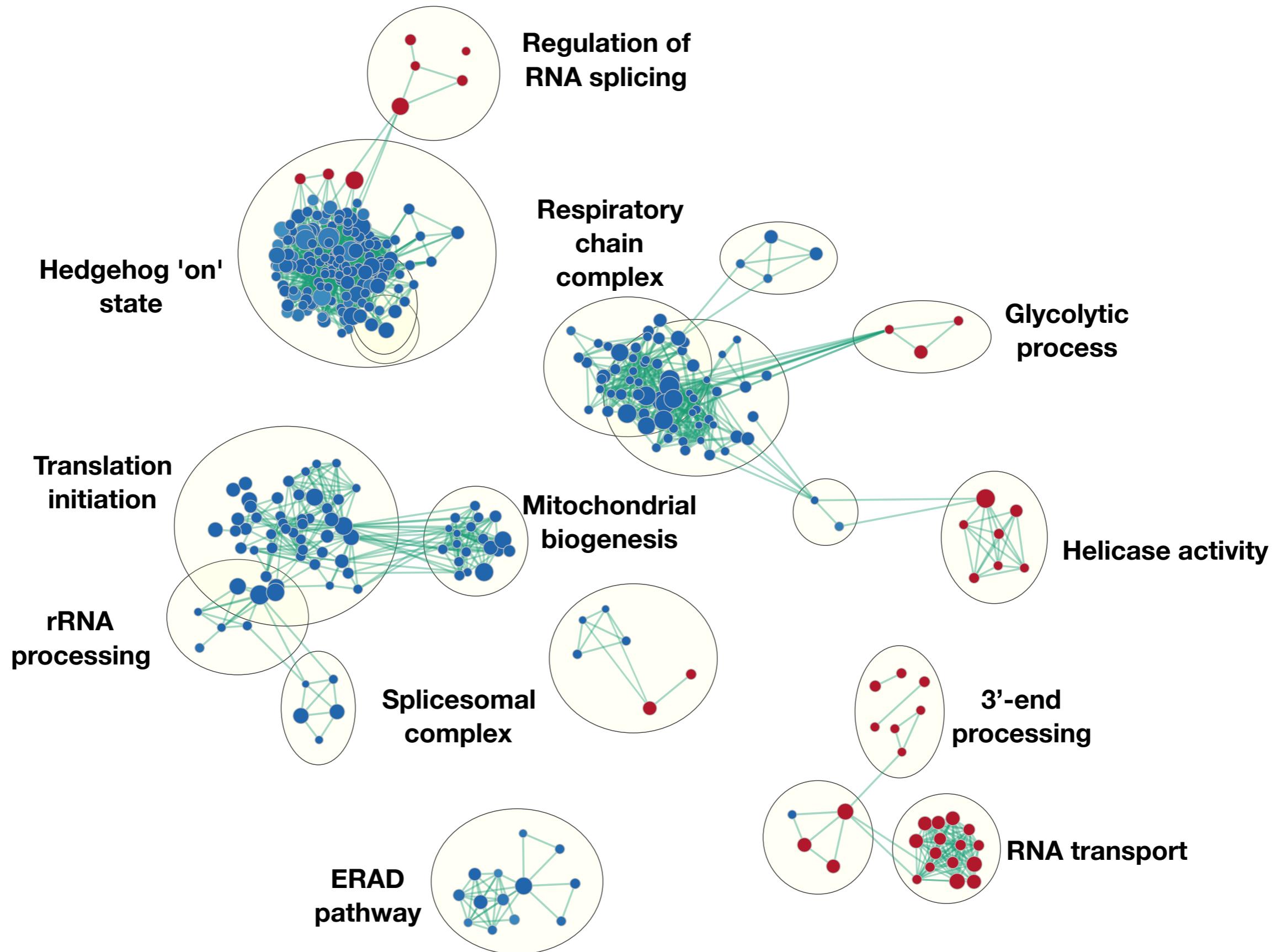
- ◆ However, presence of mutations is not associated with expression of *NEAT1*



NEAT1 differentially expressed genes (DEGs)



NEAT1 pathway enrichment analysis of DEGs



Cancer specific survival analysis

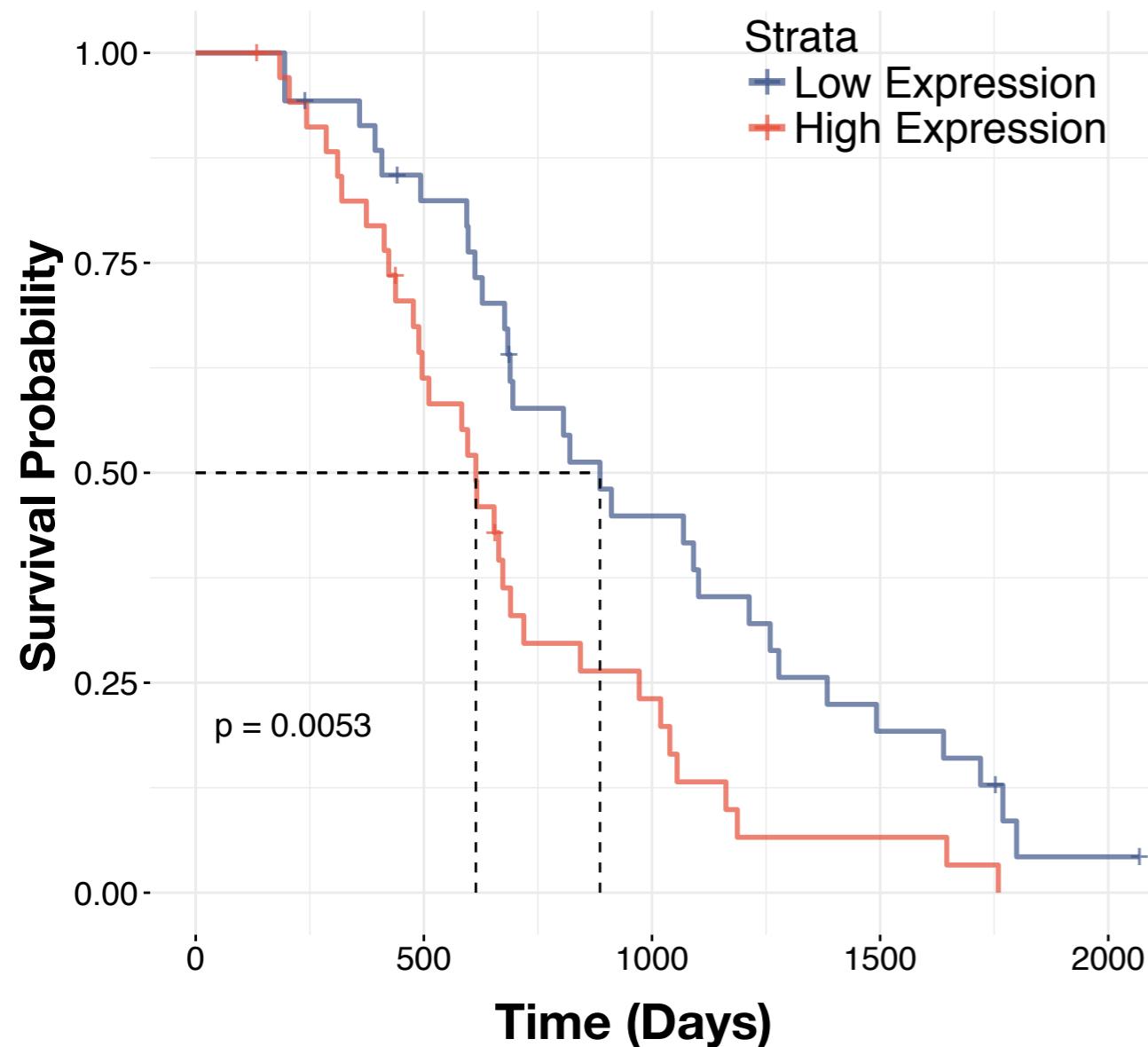
- ♦ lncRNAs with significant ($FDR < 0.05$) associations to survival outcomes
 - $n = 25$ lncRNAs among liver and ovarian cancers
 - $n = 7$ validated in matched TCGA cohorts

1. *NEAT1* in liver cancer

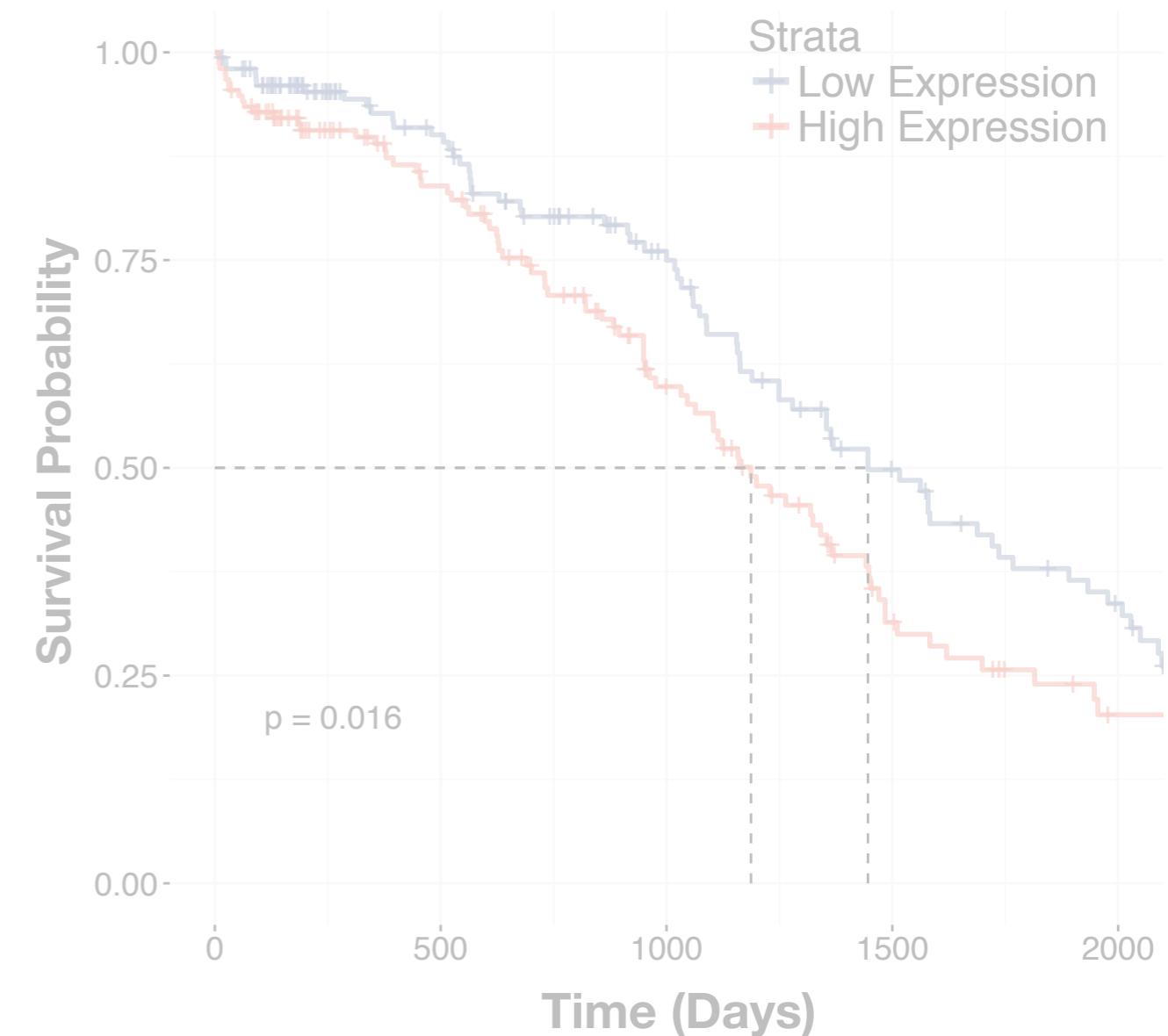
2. ***OTUD6B-AS1*** in ovarian cancer

OTUD6B-AS1 in Ovarian Cancer

PCAWG, n = 70

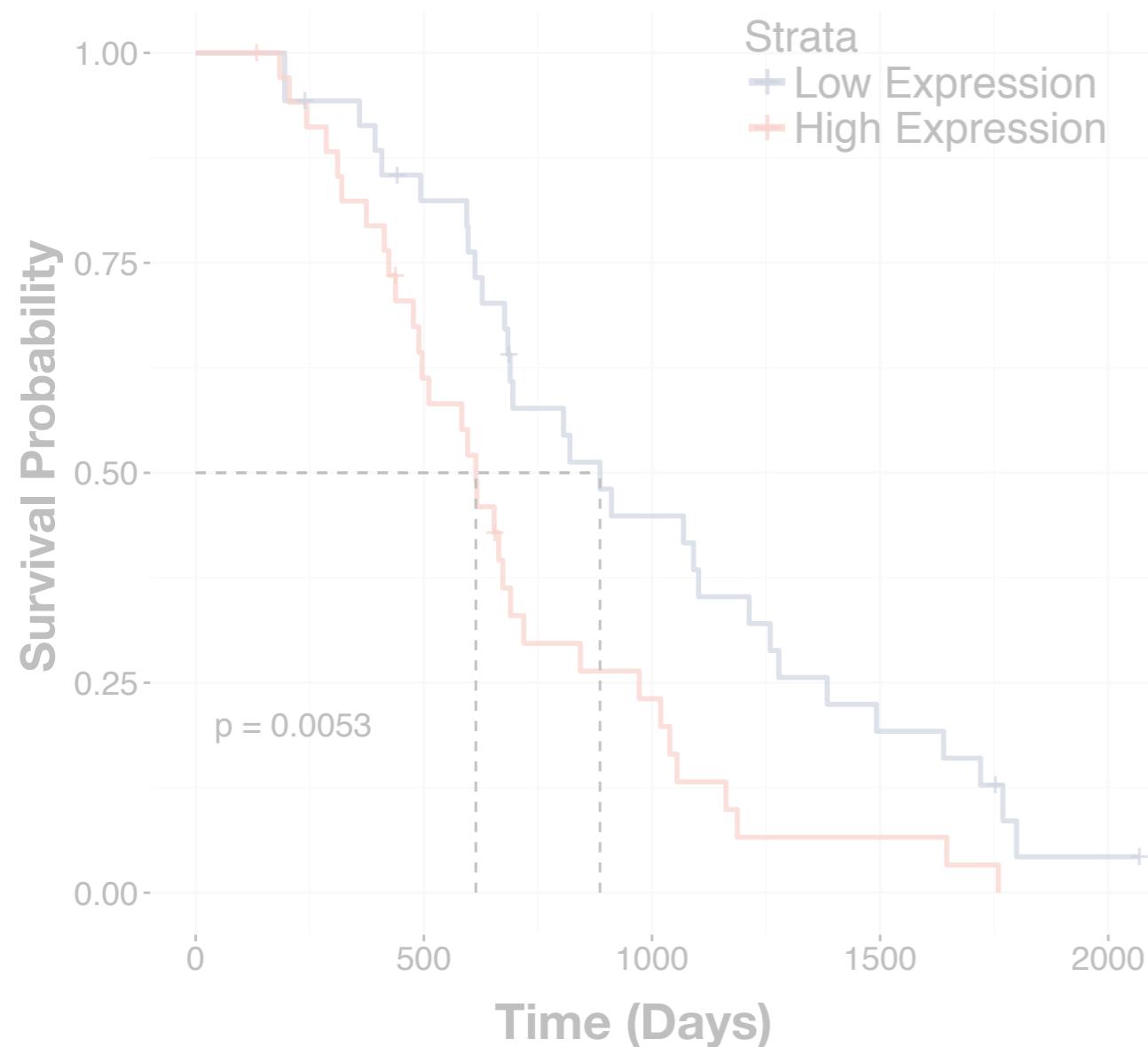


TCGA, n = 305

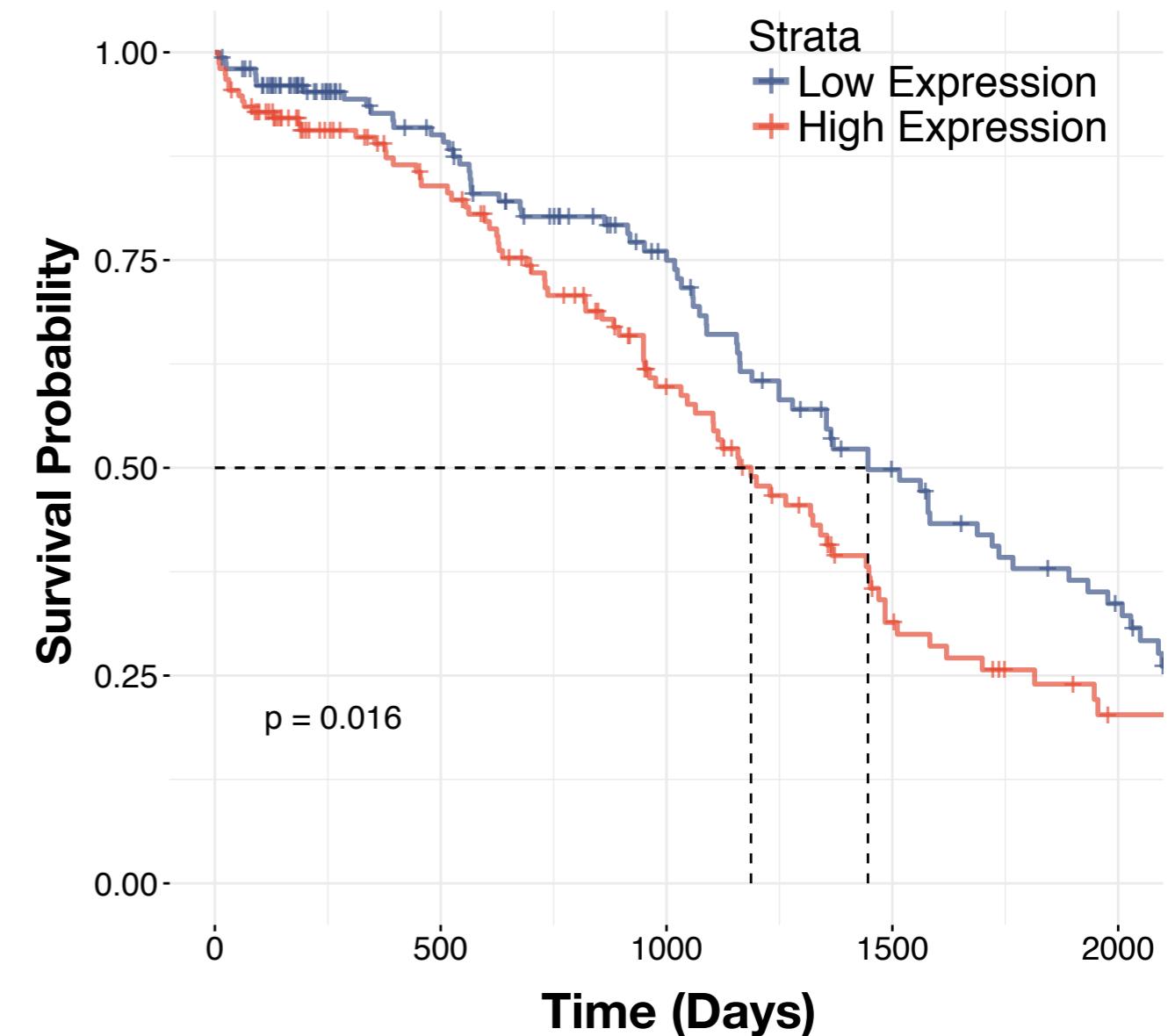


OTUD6B-AS1 in Ovarian Cancer

PCAWG, n = 70

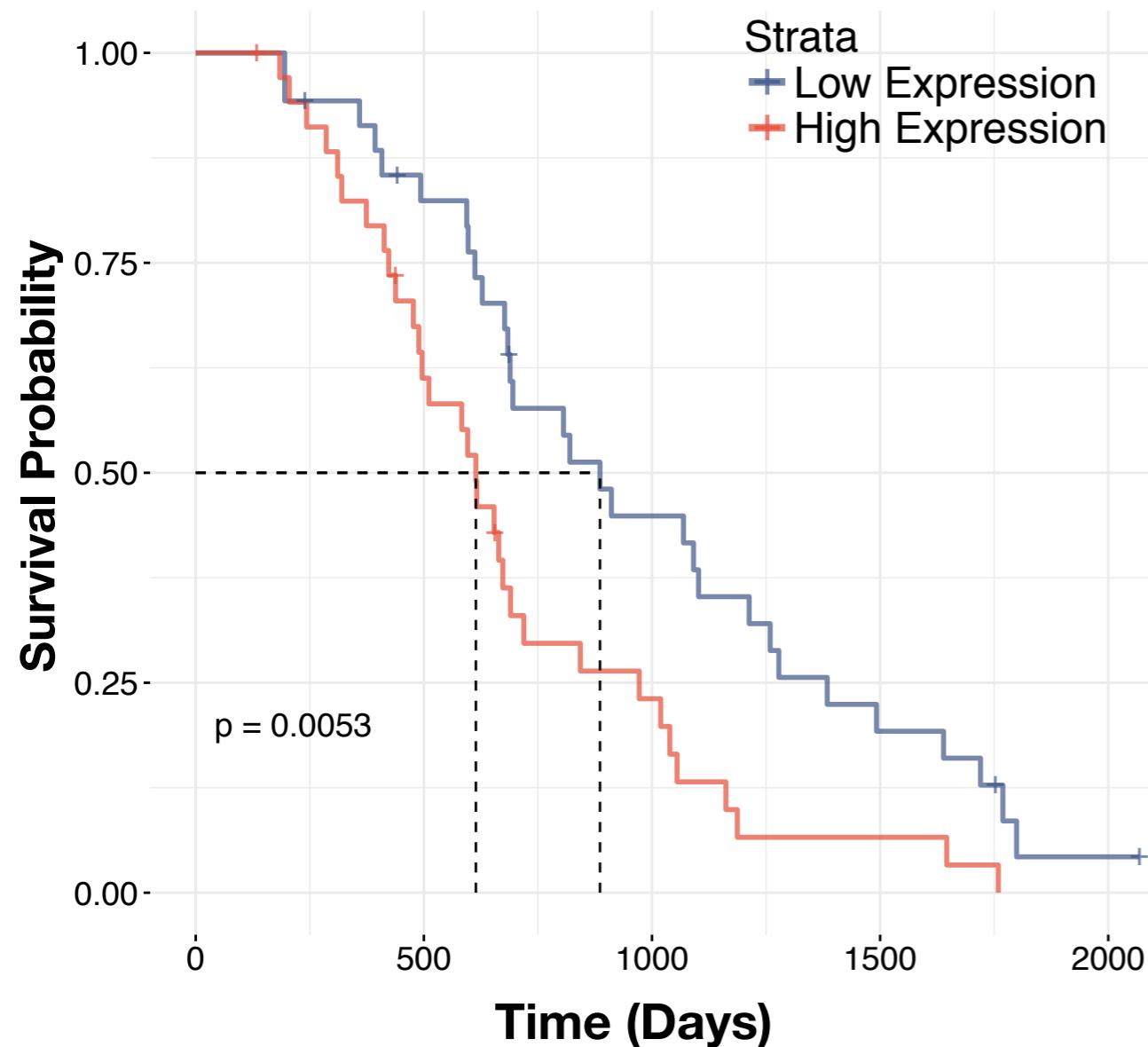


TCGA, n = 305

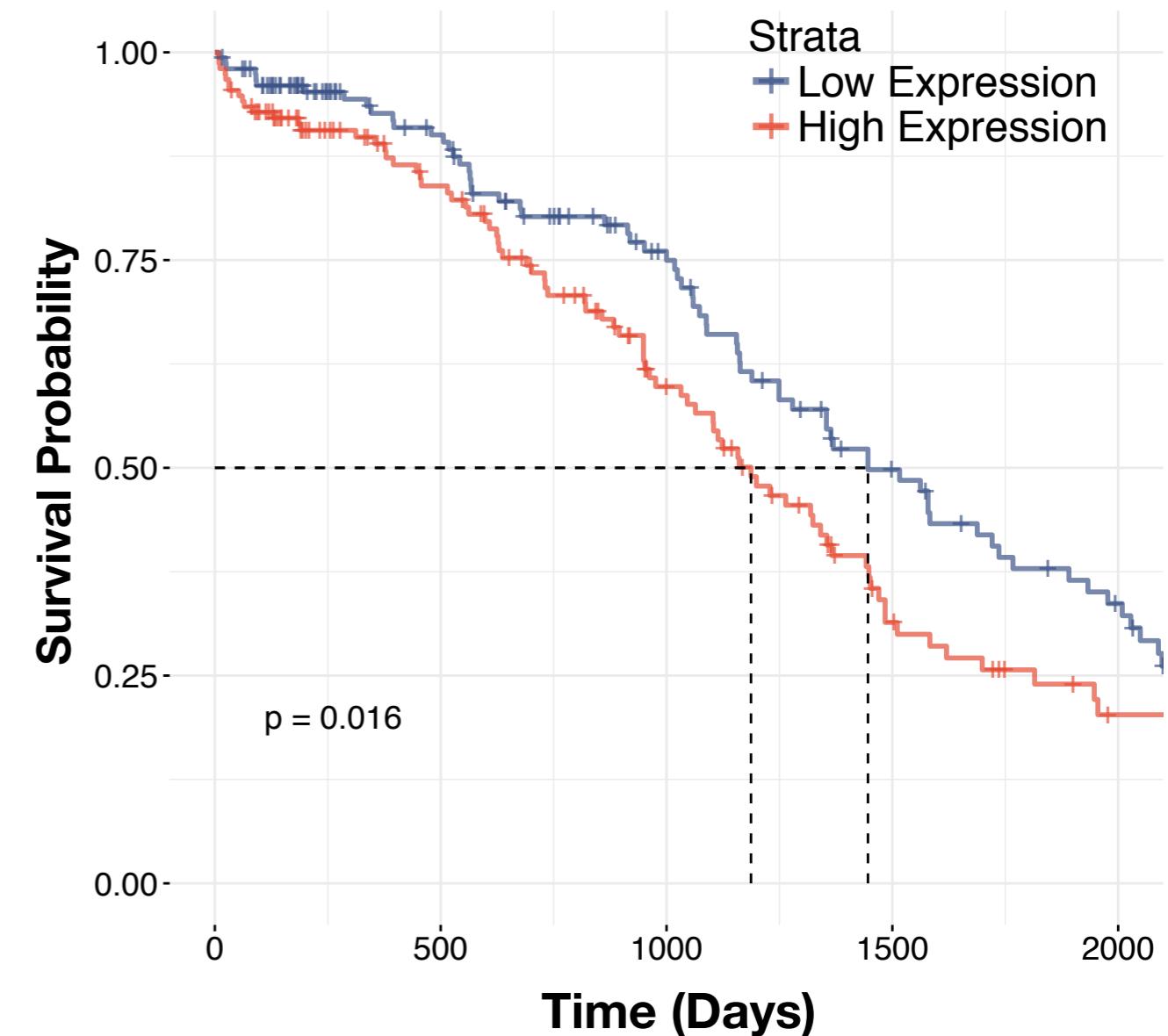


OTUD6B-AS1 in Ovarian Cancer

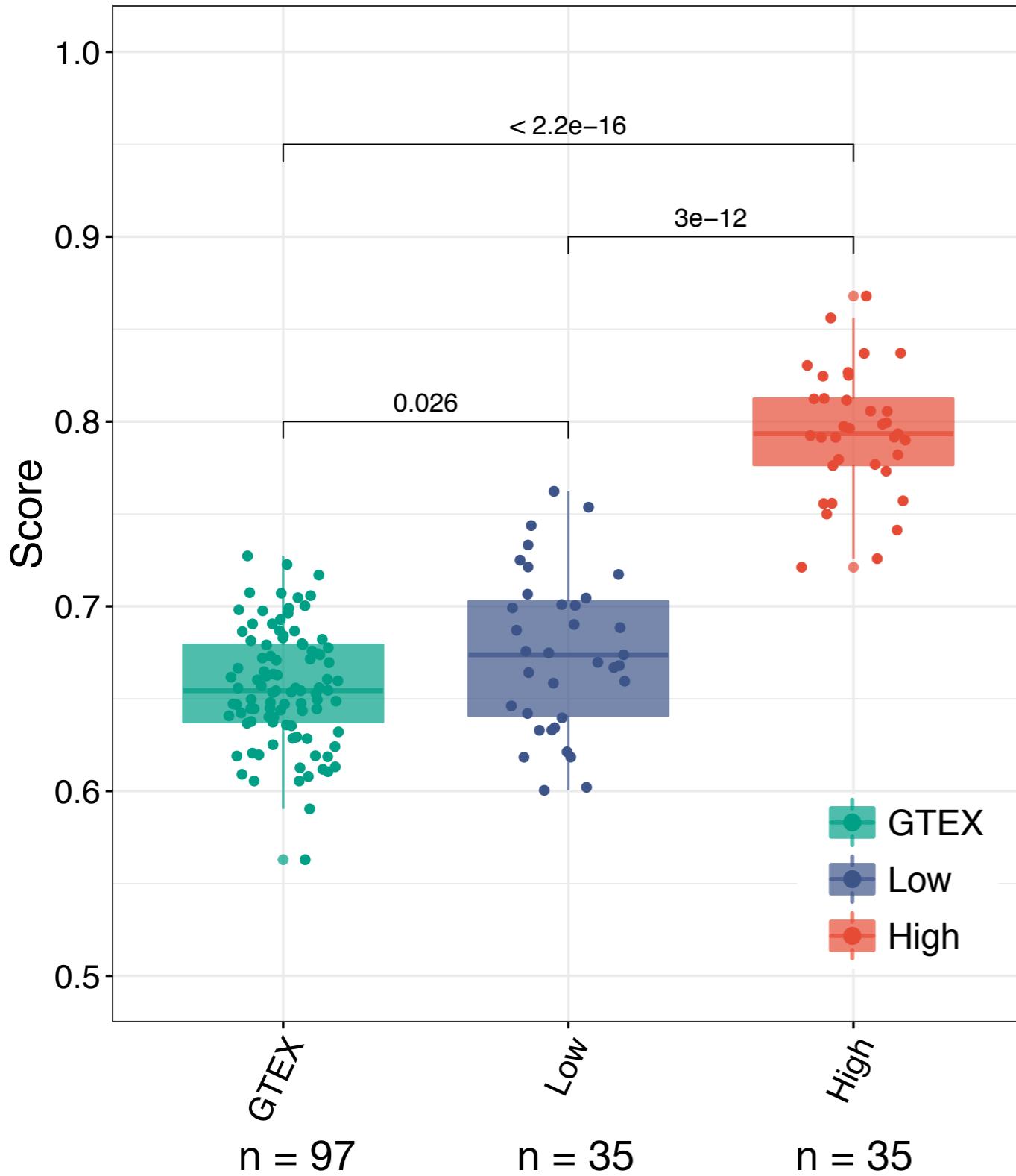
PCAWG, n = 70



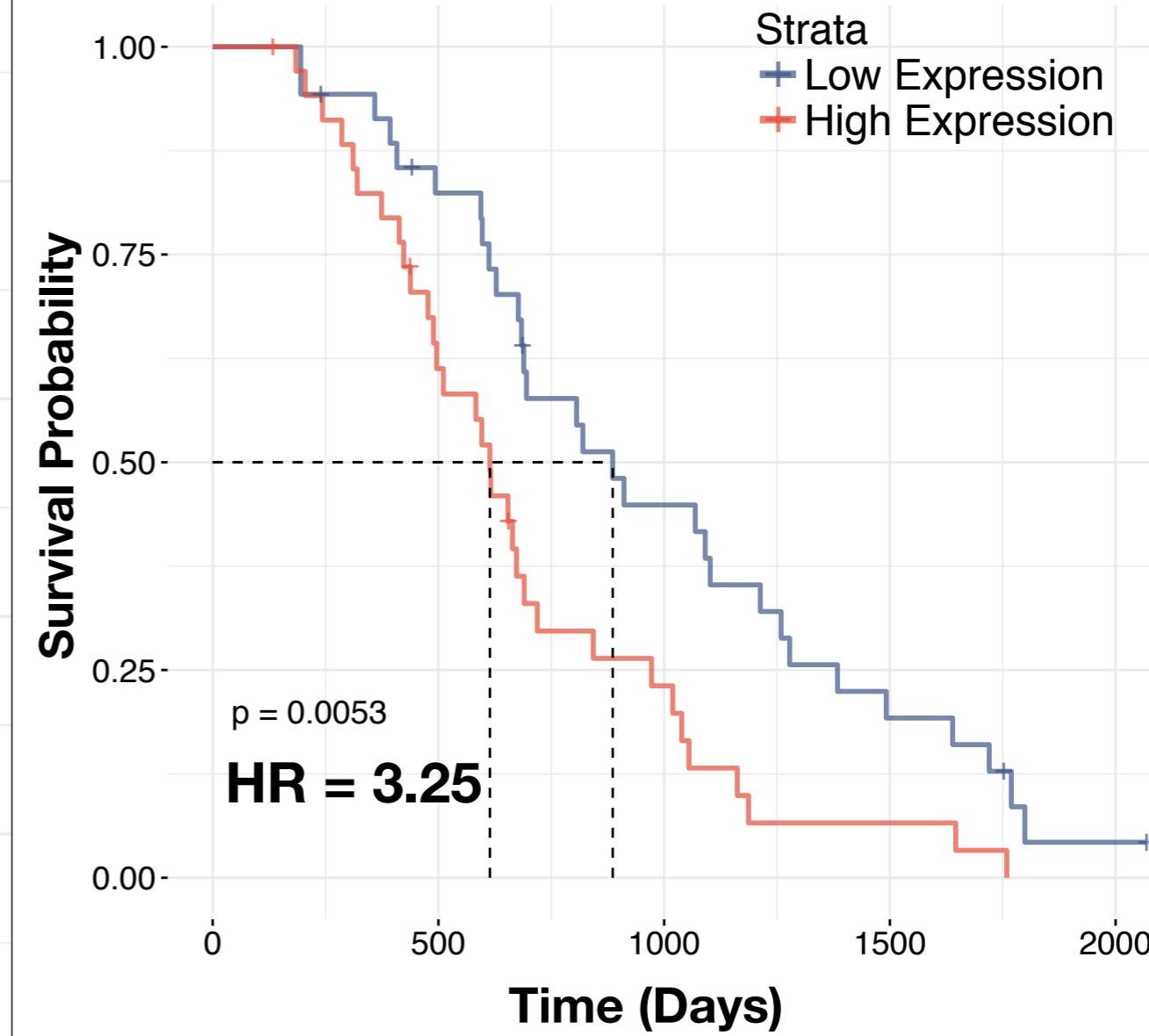
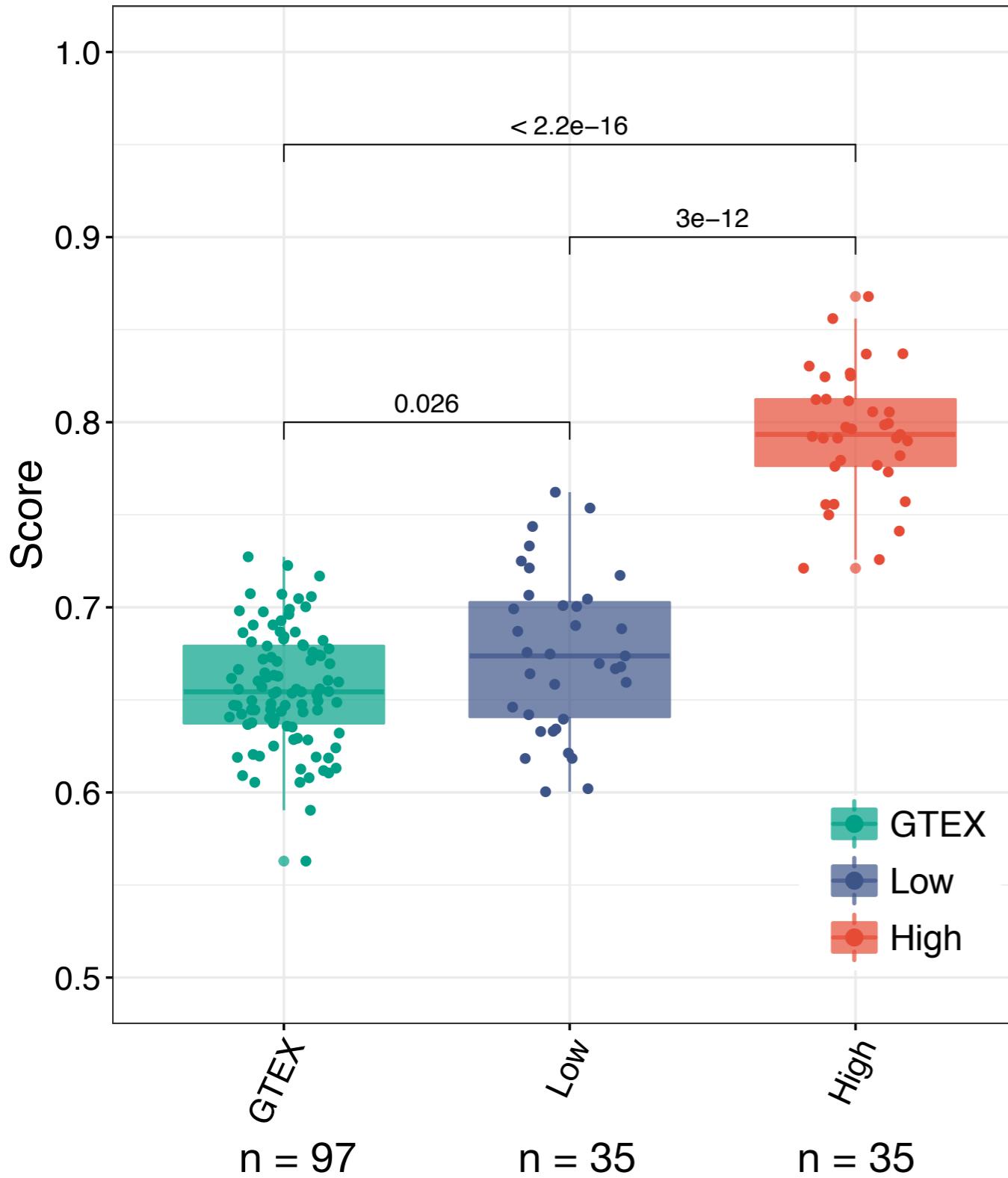
TCGA, n = 305



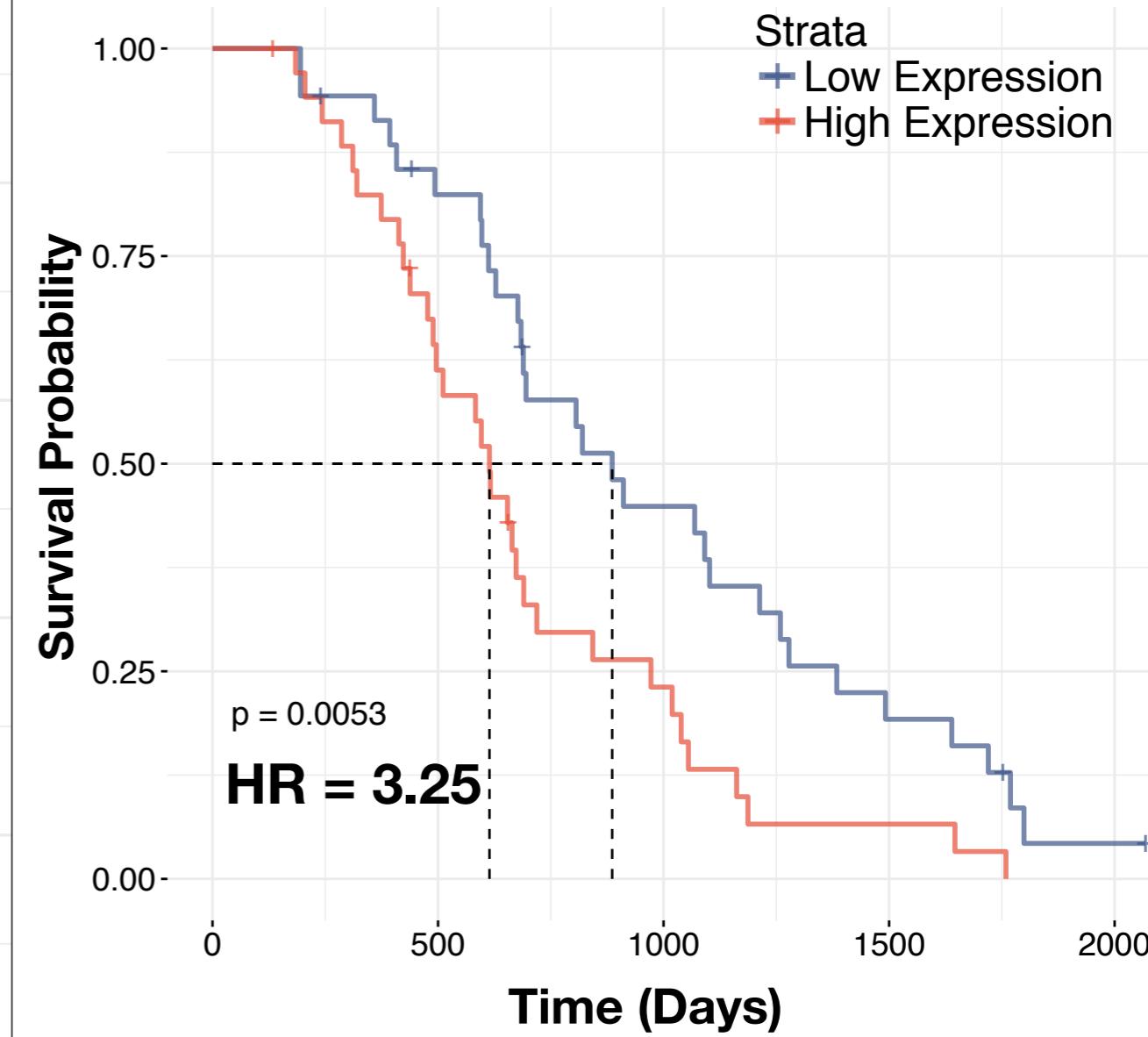
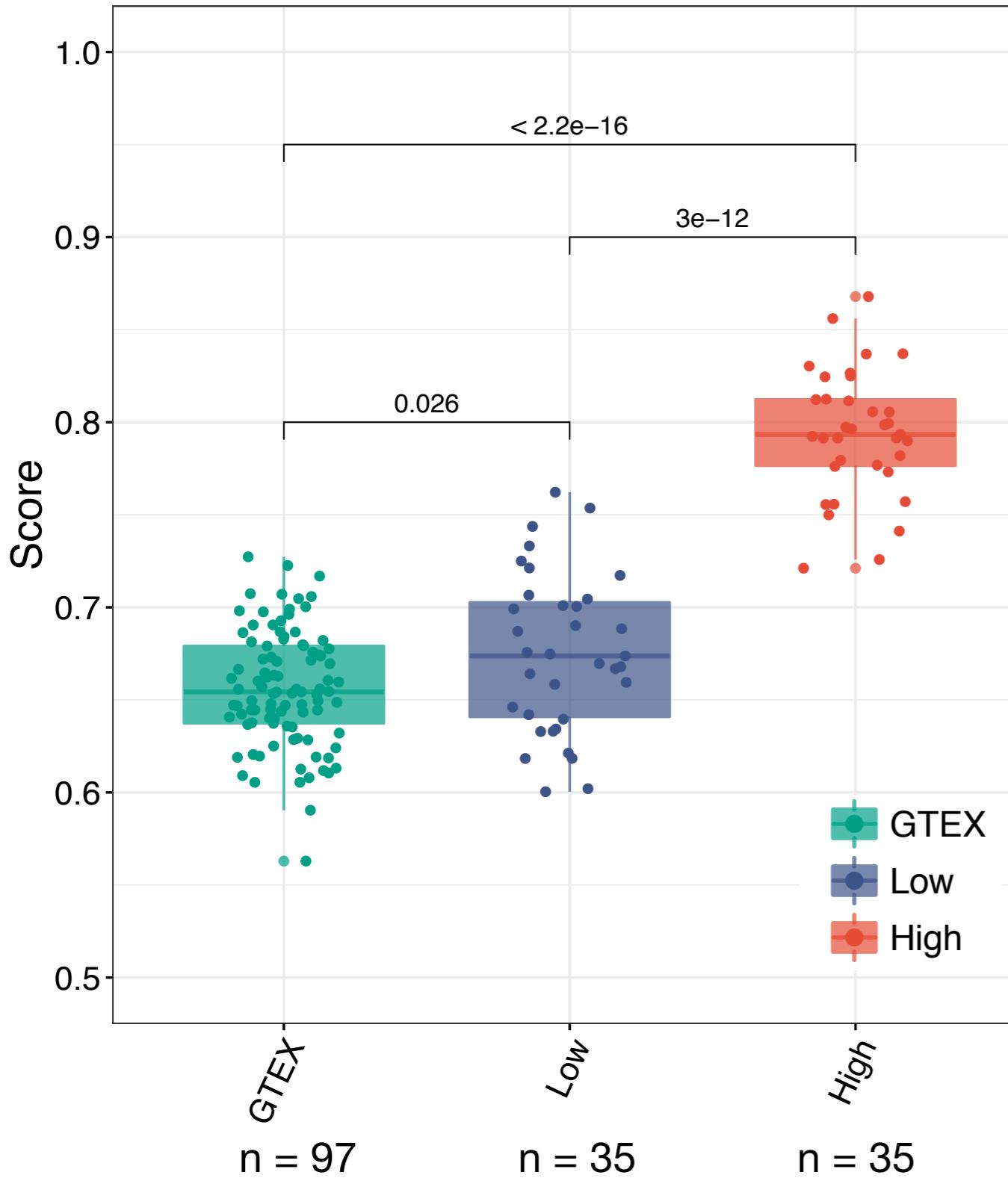
OTUD6B-AS1 in Ovarian Cancer



OTUD6B-AS1 in Ovarian Cancer

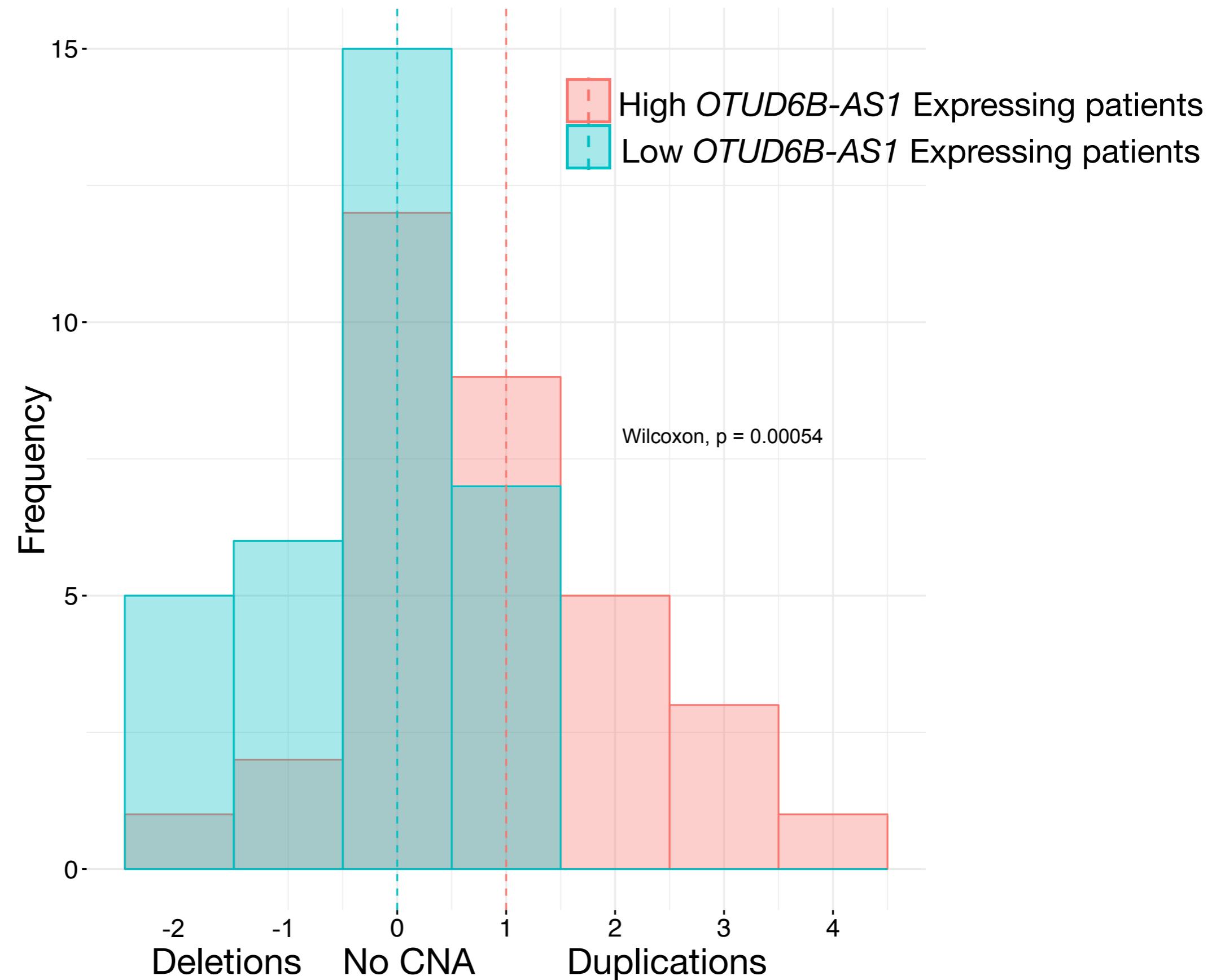


OTUD6B-AS1 in Ovarian Cancer



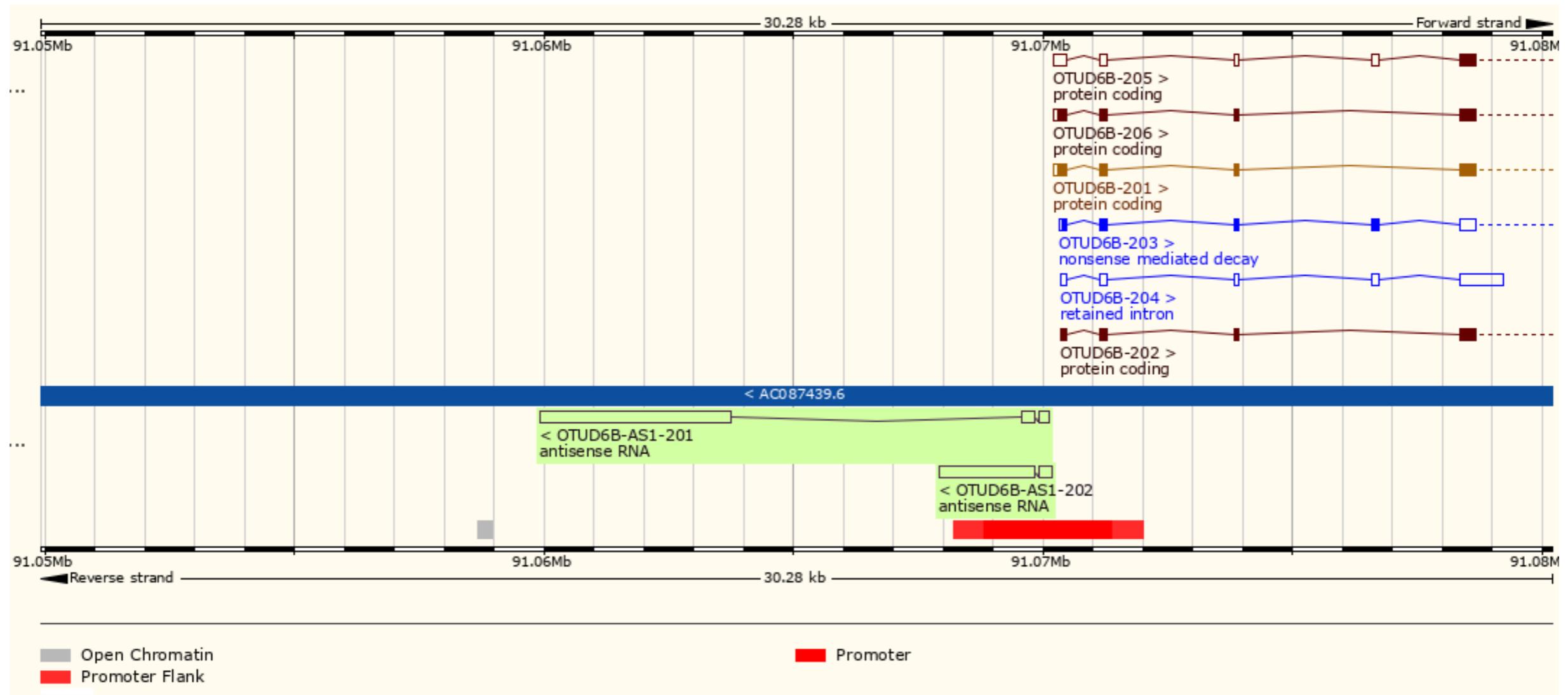
Predicted Oncogene

Somatic copy number aberration in *OTUD6B-AS1*

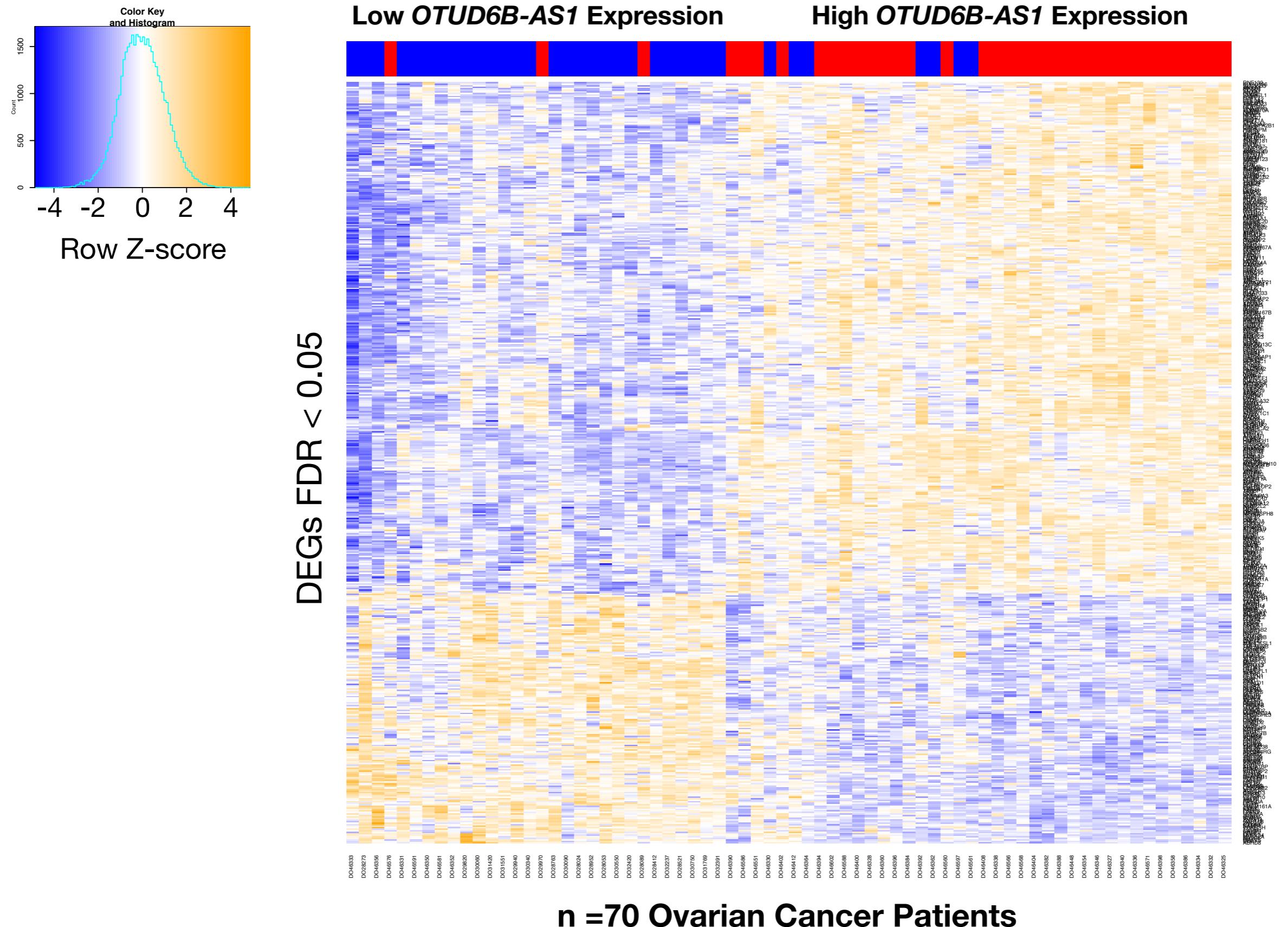


OTUD6B-AS1 in Ovarian Cancer

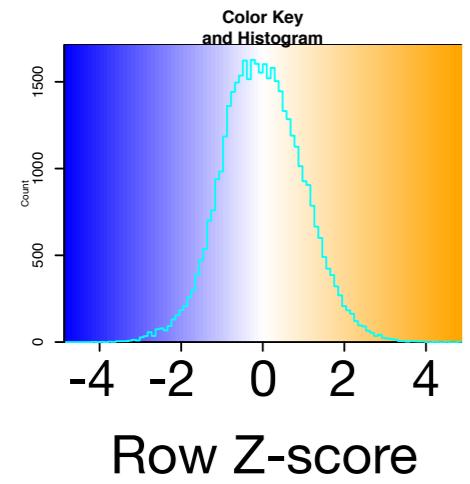
- ♦ Antisense lncRNA on chromosome 8
- ♦ OTUD6B on opposite strand
 - Protease that cleaves ubiquitin linkages



OTUD6B-AS1 DEGs

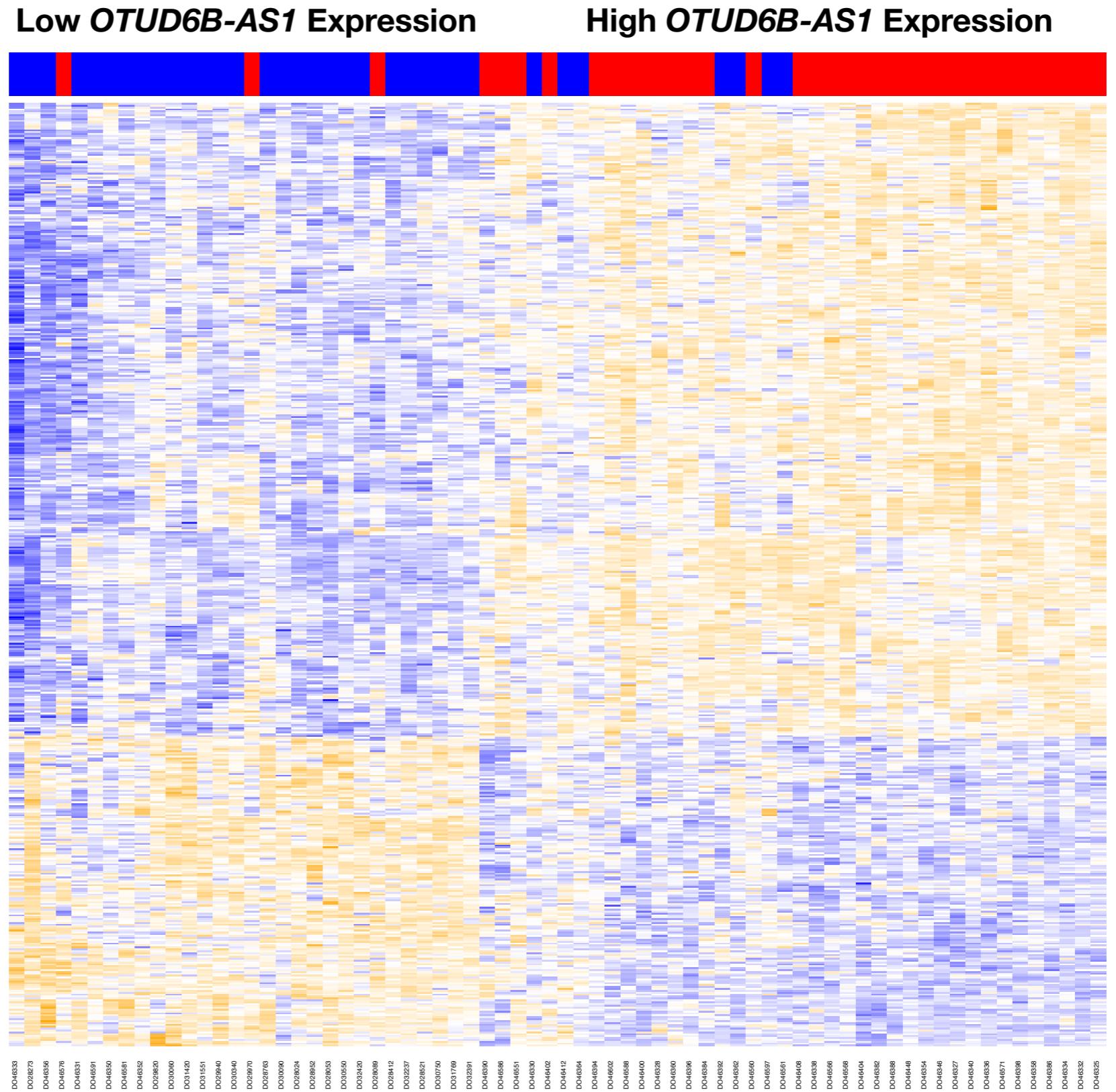


OTUD6B-AS1 DEGs



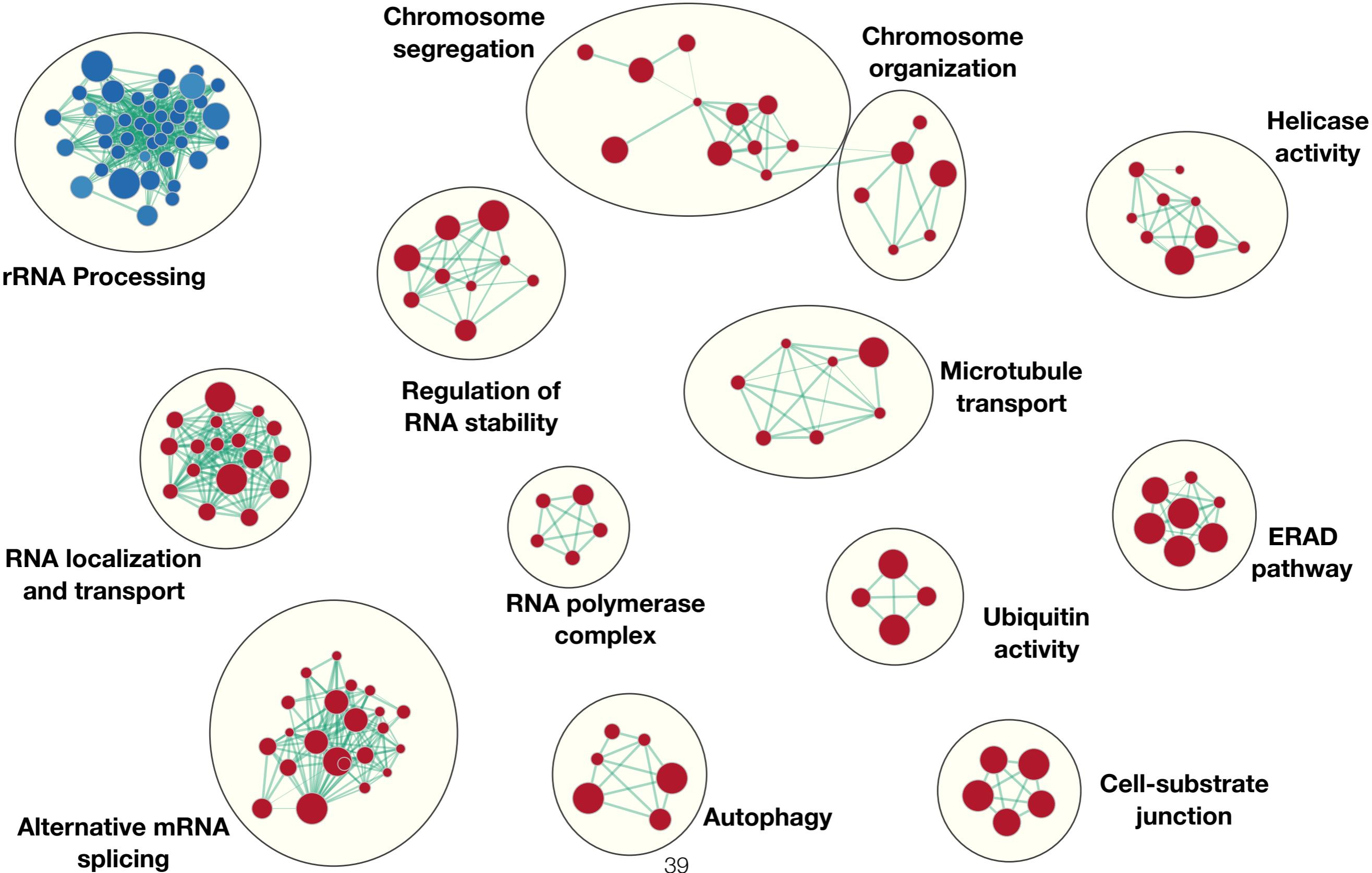
DEGs FDR < 0.05

OTUD6B is not differentially expressed



n =70 Ovarian Cancer Patients

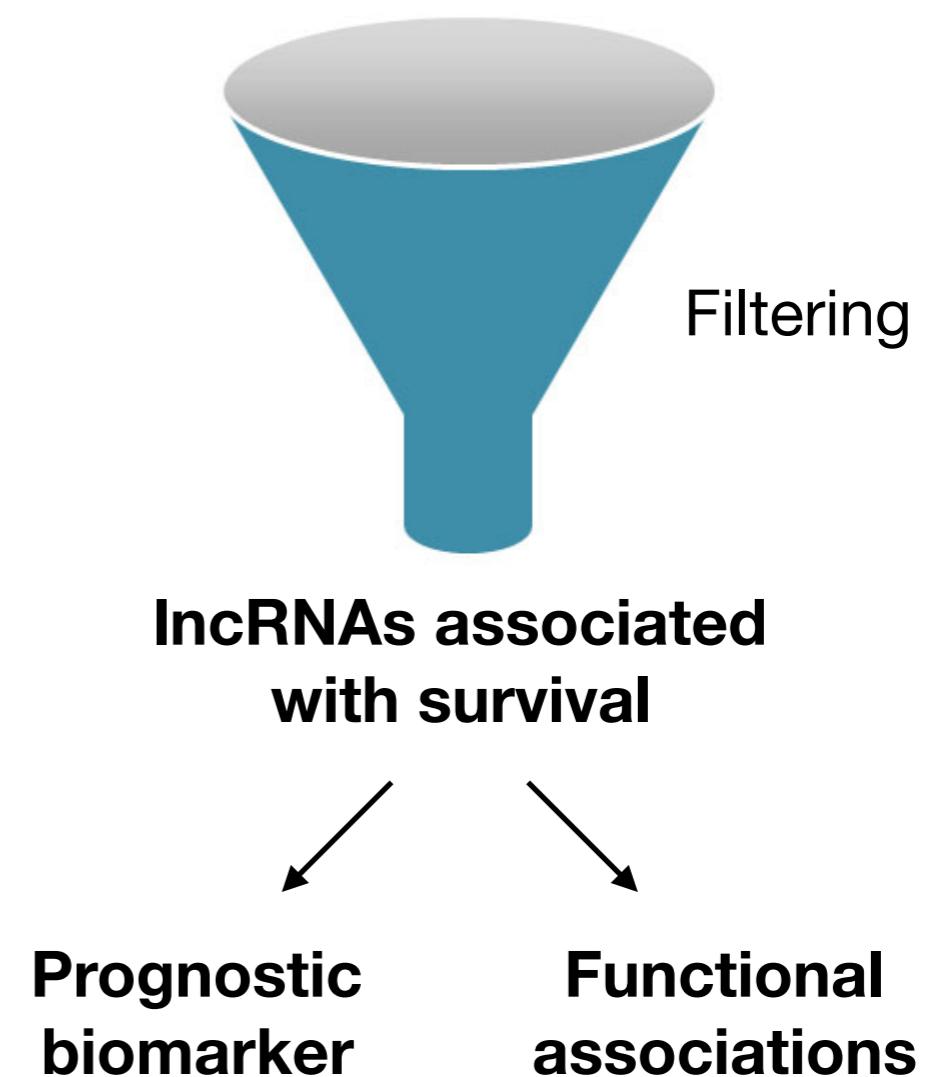
OTUD6B-AS1 pathway enrichment analysis



Summary

- ♦ Systematic analysis in multiple cancer types revealed lncRNAs significantly associated with survival
- ♦ Identified *Neat1* as positive control
- ♦ Novel candidate lncRNA *OTUD6B-AS1* identified in ovarian cancer
- ♦ We can better predict the function of 1000s of previously uncharacterized lncRNAs by associating candidate lncRNAs to pathways in a cancer specific manner

~ 20,000 lncRNAs



Future Directions - Aim 1

- ♦ Expand to a lower median expression threshold
- ♦ Expand to other events such as tumour recurrence
- ♦ Systematic analysis of lncRNA methylation as predictor of clinical outcomes
- ♦ Integrate additional clinical covariates such as age, smoking/drinking status, type of treatment and hepatitis status (liver cancer) into survival models

Future Directions - Aims 2 + 3

Aim 2:

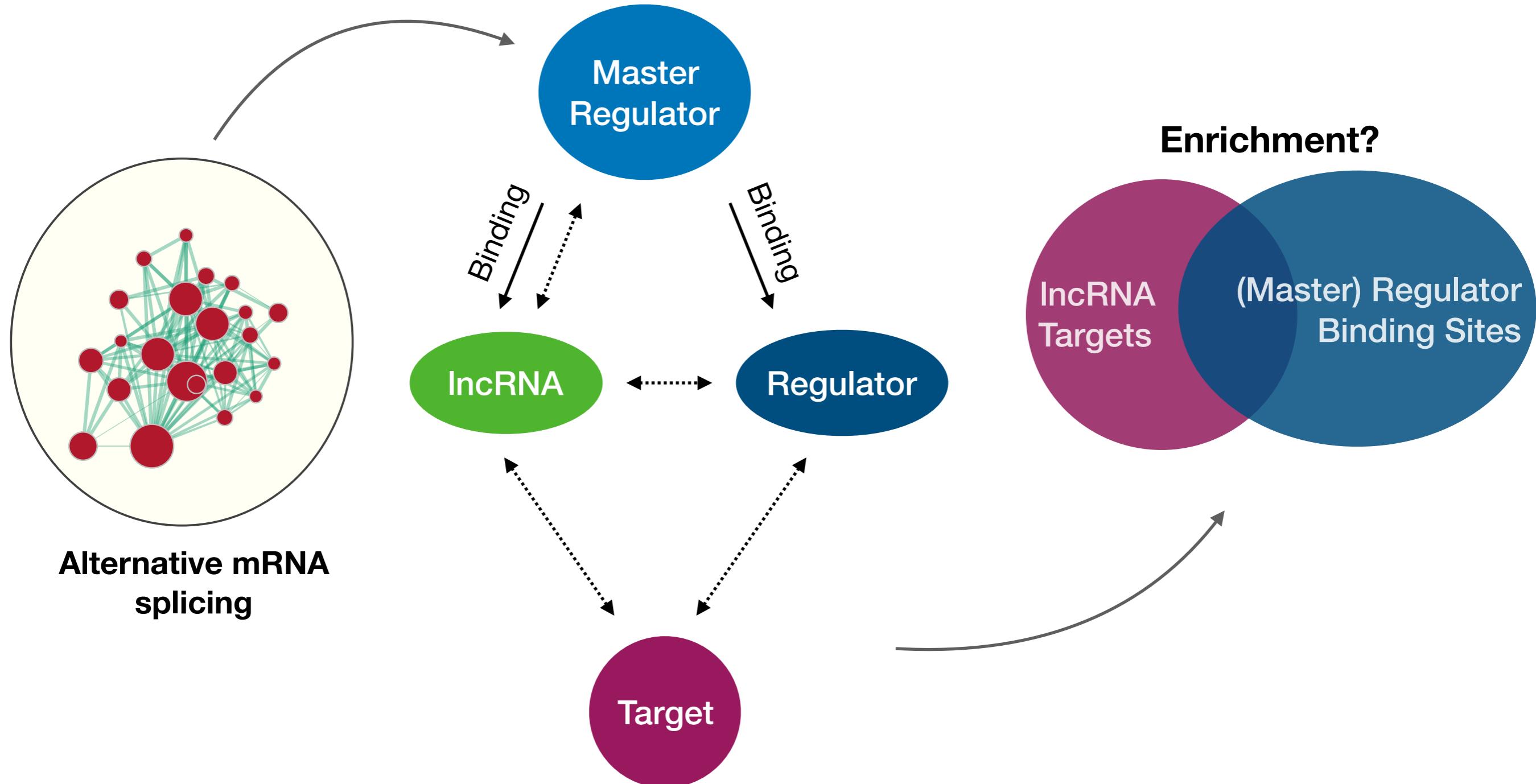
- ◆ Validate copy number aberration
- ◆ Systematic analysis of lncRNA expression ~ lncRNA promoter methylation

Aim 3:

- ◆ Integrate transcription factor binding data
- ◆ Experimental validation of predicted interactions

Future Directions

Predicting Directionality & Regulatory Network Construction



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



Additional Slides

NEAT1 mutated in ~22% of HCC patients

p=0.31

