

Committee Meeting #2

Karin Isaev

Supervisor: Dr. Juri Reimand

September 19th, 2017

Outline

Thesis: Genomic characterization of clinically relevant lncRNAs in multiple cancer types

1. Review of lncRNAs
2. Summary of the last SCM
3. Summary of progress since last meeting
4. Future directions

Thesis: Genomic characterization of clinically relevant lncRNAs in multiple cancer types

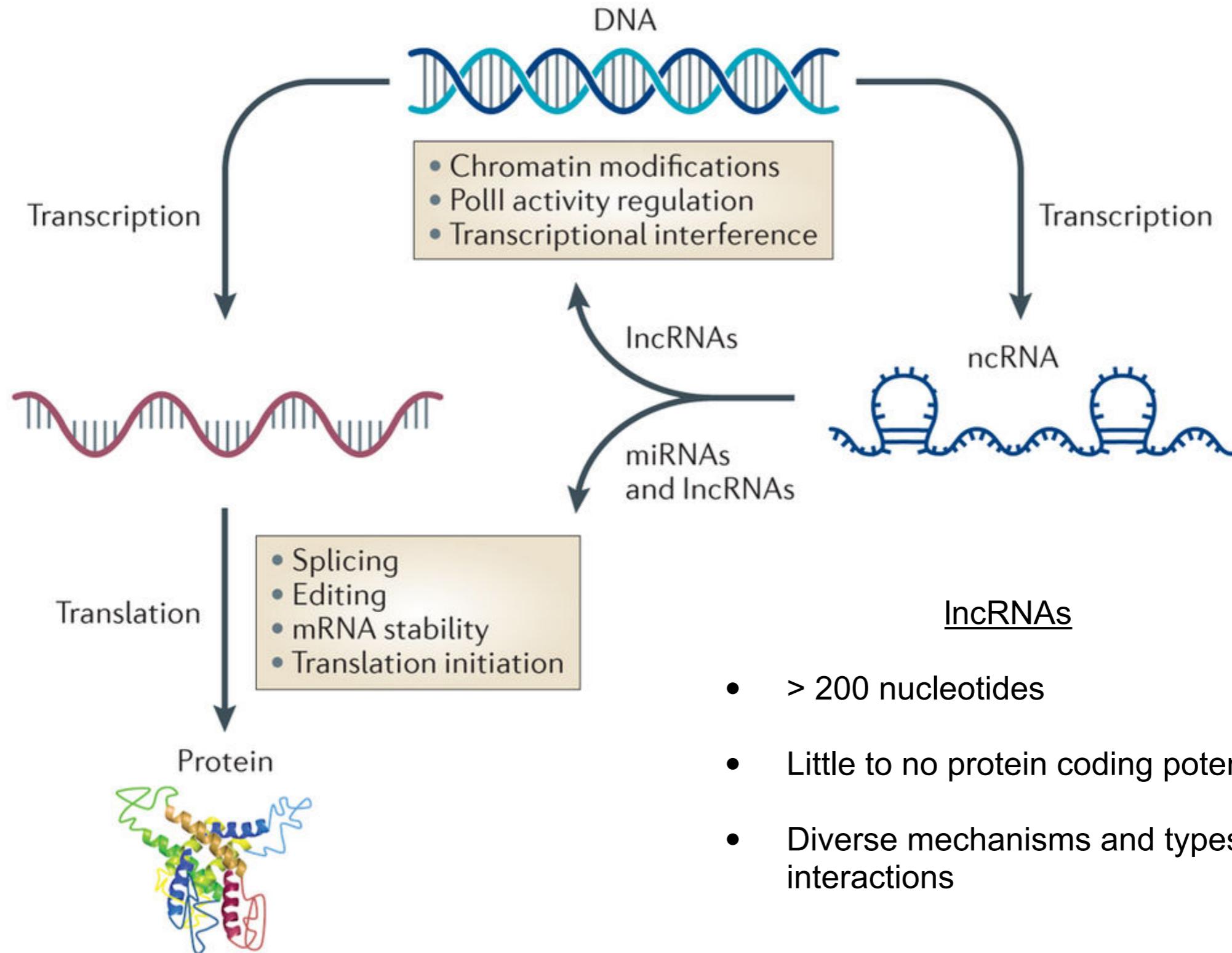
1. Review of lncRNAs

2. Summary of the last SCM

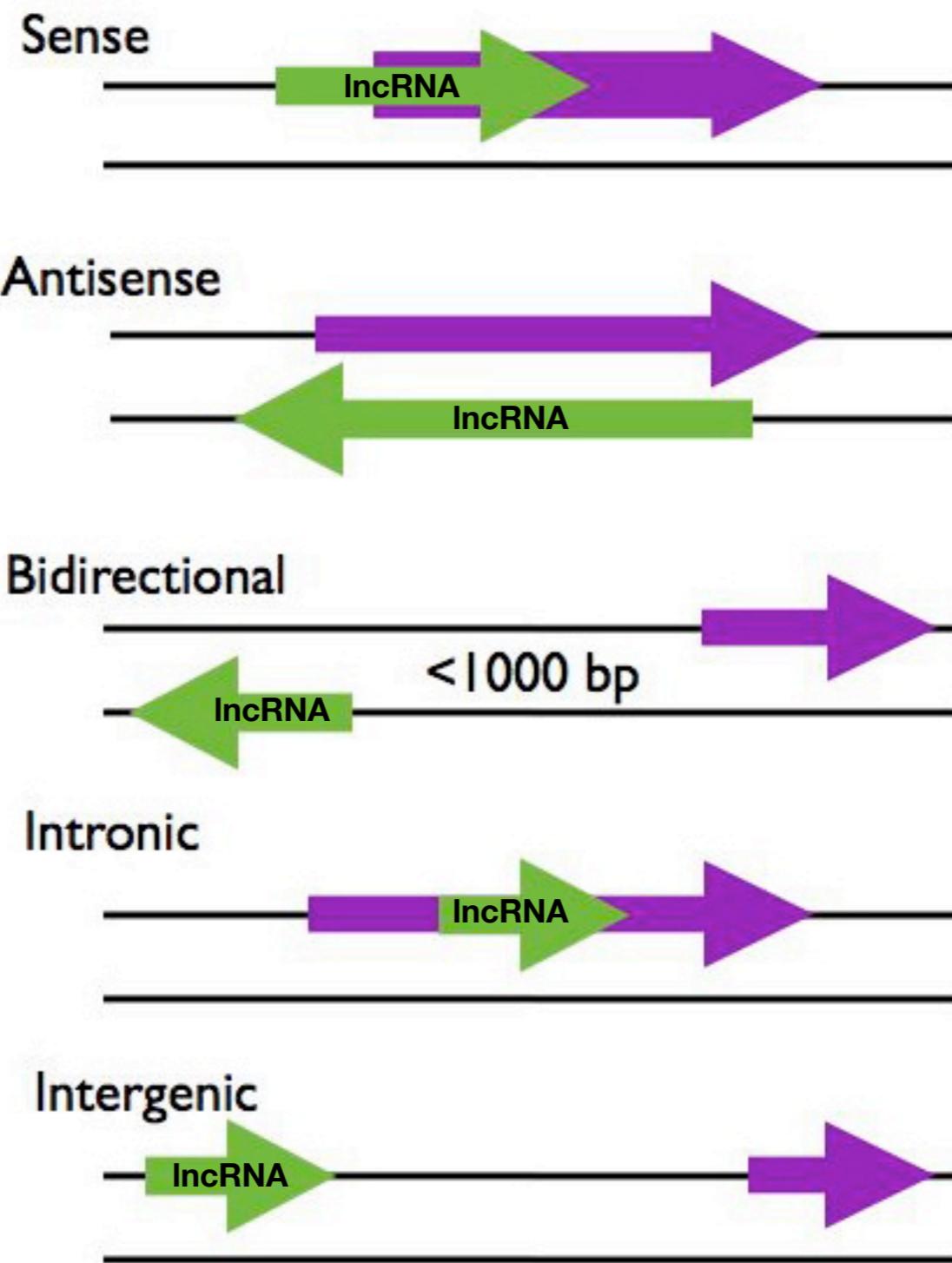
3. Summary of progress since last meeting

4. Future directions

lncRNAs

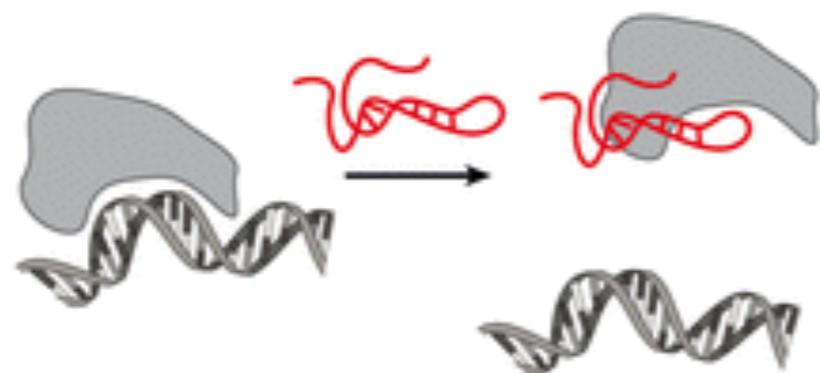


lncRNA type related to location of protein coding genes

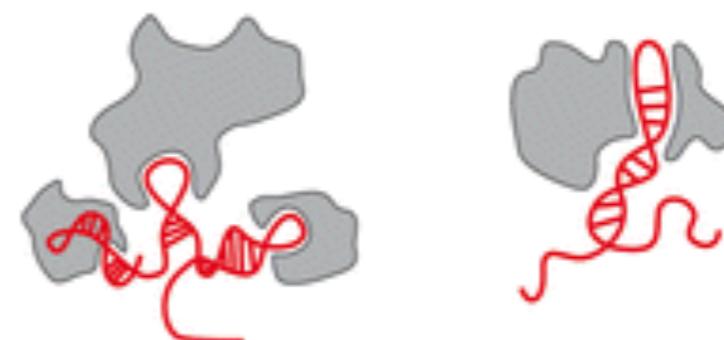


Examples of known lncRNA functions

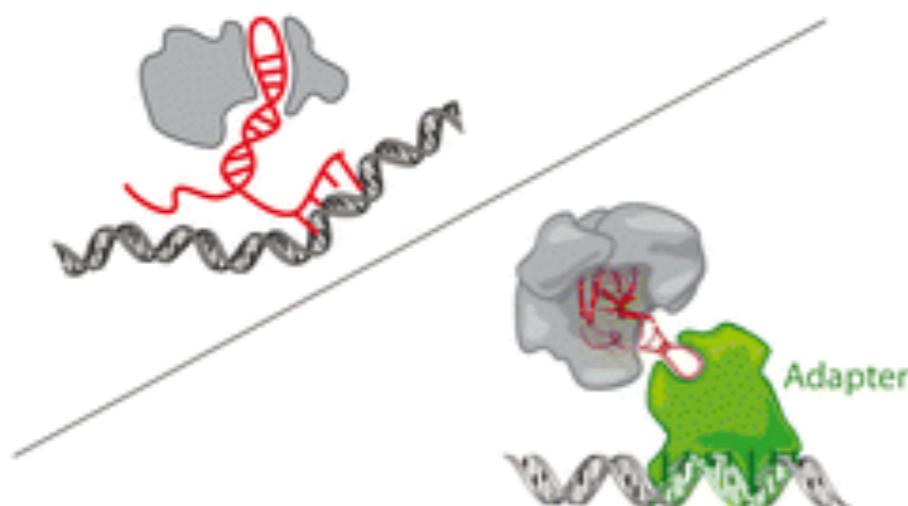
a Decoy



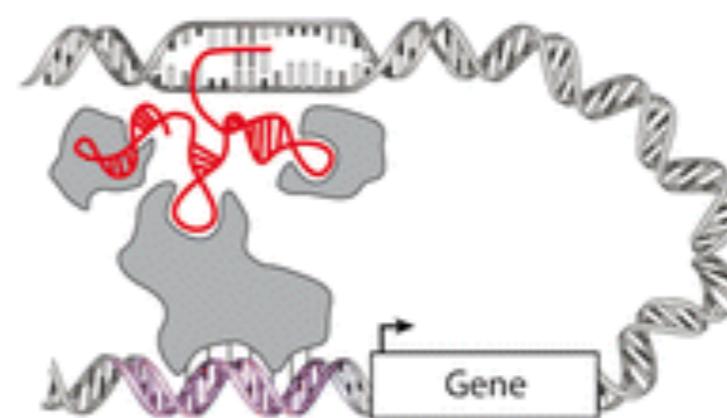
b Scaffold



c Guide

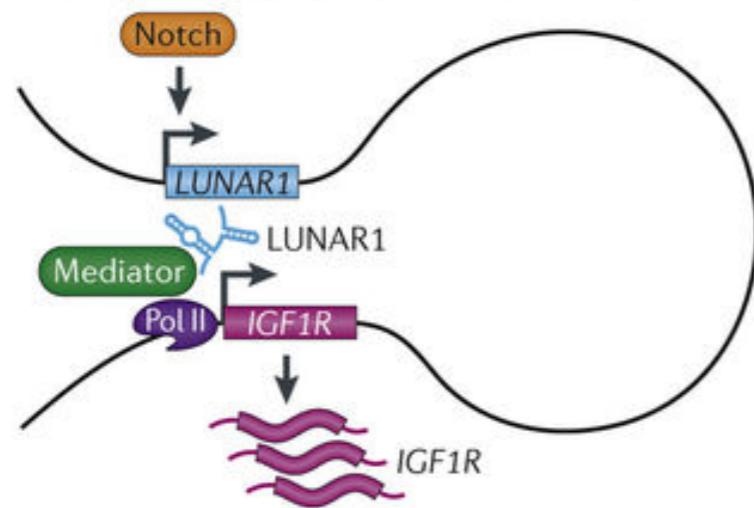


d Enhancer

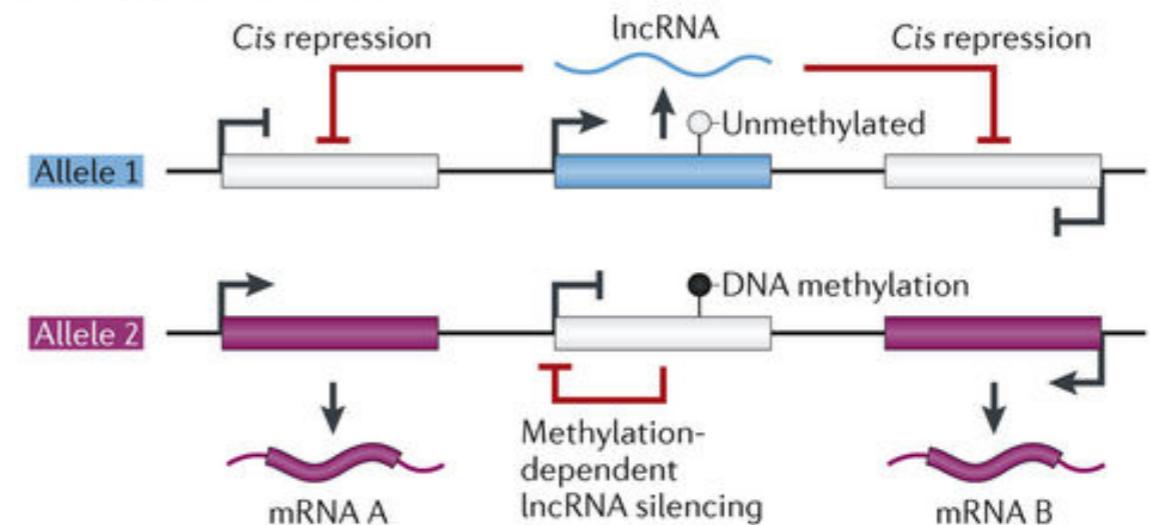


Examples of known lncRNA functions

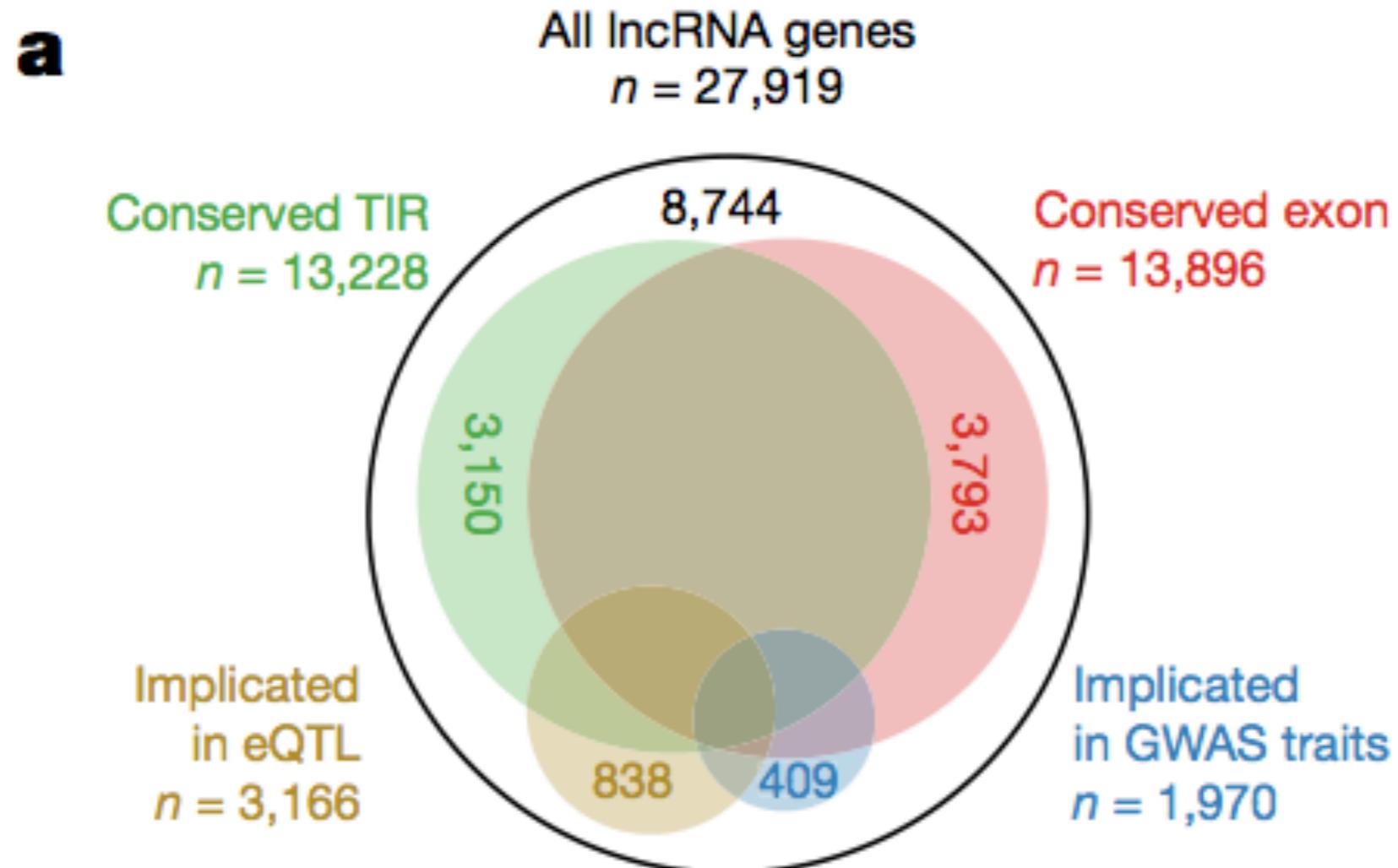
a Enhancer RNAs and chromosome looping



c Imprinted gene clusters



~20,000 predicted functional lncRNAs in 2017



Outline

Thesis: Genomic characterization of clinically relevant lncRNAs in multiple cancer types

1. Review of lncRNAs

2. Summary of the last SCM

3. Summary of progress since last meeting

4. Future directions

Project aims:

- Using the PCAWG dataset, lncRNA candidates will be identified in multiple cancer types through associating their gene expression with patient survival outcomes.
 - Rationale: lncRNAs are expressed in a tissue specific manner and are often differentially expressed between tumour and matched normal samples
- Through further analysis including, differential expression, co-expression and pathway enrichment analysis, these lncRNA candidates will be associated with predicted cellular roles relevant to cancer progression in these patients.
- Cancer specific filtration of candidate lncRNAs
- Overcoming low abundance and possibility of noise

Literature based:

- Understand lncRNAs and protein coding potential
 - Open Reading Frames
- lncRNA in relation to protein coding genes
 - Overcoming possible noise
- lncRNA FANTOM CAGE-Seq data
 - Expanded study and database

Outline

Thesis: Genomic characterization of clinically relevant lncRNAs in multiple cancer types

1. Review of lncRNAs

2. Summary of the last SCM

3. Summary of progress since last meeting

4. Future directions

Summary of progress

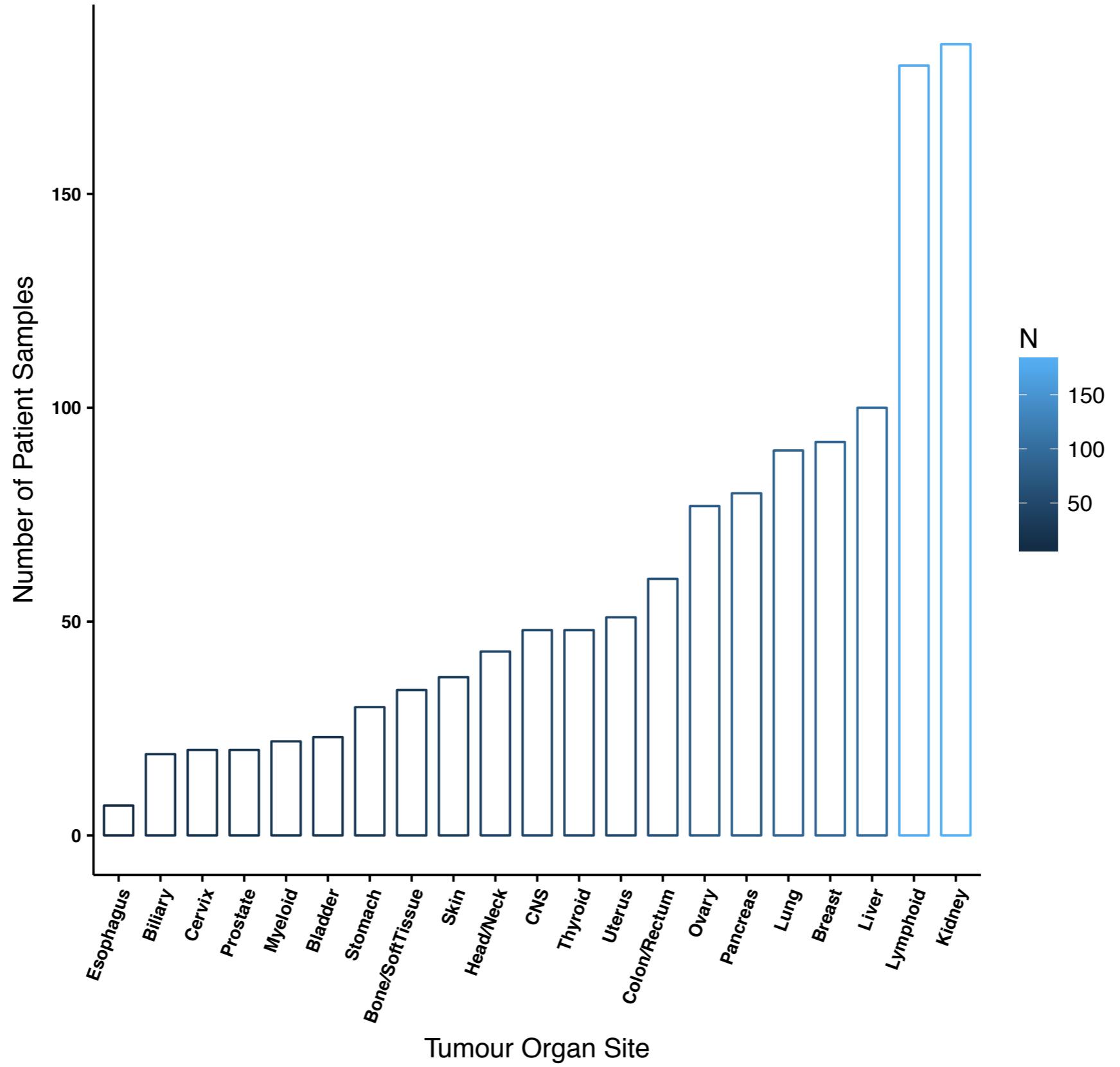
1. Identified high expressing cancer specific lncRNAs through several filtration steps across multiple cancer types.
2. Confirmed cancer specific expression through comparing the distributions of each candidate lncRNA's expression among the cancer types studied
3. Identified cancer specific high expressing lncRNAs associated with patient survival outcome using the Cox Proportional Hazards model
 - 42 lncRNAs in 4 cancer types
 - 7 lncRNAs with adjusted p-value < 0.1 in 2 cancer types (Tier 1)
 - 35 lncRNAs with p-value < 0.05 in 4 cancer types (Tier 2)
4. Assessed tumour versus normal tissue lncRNA expression by comparing ranked values of expression using match tissue data from GTEx
5. 21/42 lncRNAs were predicted to be either tumour suppressive or oncogenic based on sign of hazard ratio and difference in expression between normal and tumour tissues
6. 2 candidates in Ovarian cancer were further studied to predict regulatory roles that may lead patients from the same cancer type to develop worse prognosis

Summary of progress

1. Identified high expressing cancer specific lncRNAs through several filtration steps across multiple cancer types.
2. Confirmed cancer specific expression through comparing the distributions of each candidate lncRNA's expression among the cancer types studied
3. Identified cancer specific high expressing lncRNAs associated with patient survival outcome using the Cox Proportional Hazards model
 - 42 lncRNAs in 4 cancer types
 - 7 lncRNAs with adjusted p-value < 0.1 in 2 cancer types (Tier 1)
 - 35 lncRNAs with p-value < 0.05 in 4 cancer types (Tier 2)
4. Assessed tumour versus normal tissue lncRNA expression by comparing ranked values of expression using match tissue data from GTEx
5. 21/42 lncRNAs were predicted to be either tumour suppressive or oncogenic based on sign of hazard ratio and difference in expression between normal and tumour tissues
6. 2 candidates in Ovarian cancer were further studied to predict regulatory roles that may lead patients from the same cancer type to develop worse prognosis

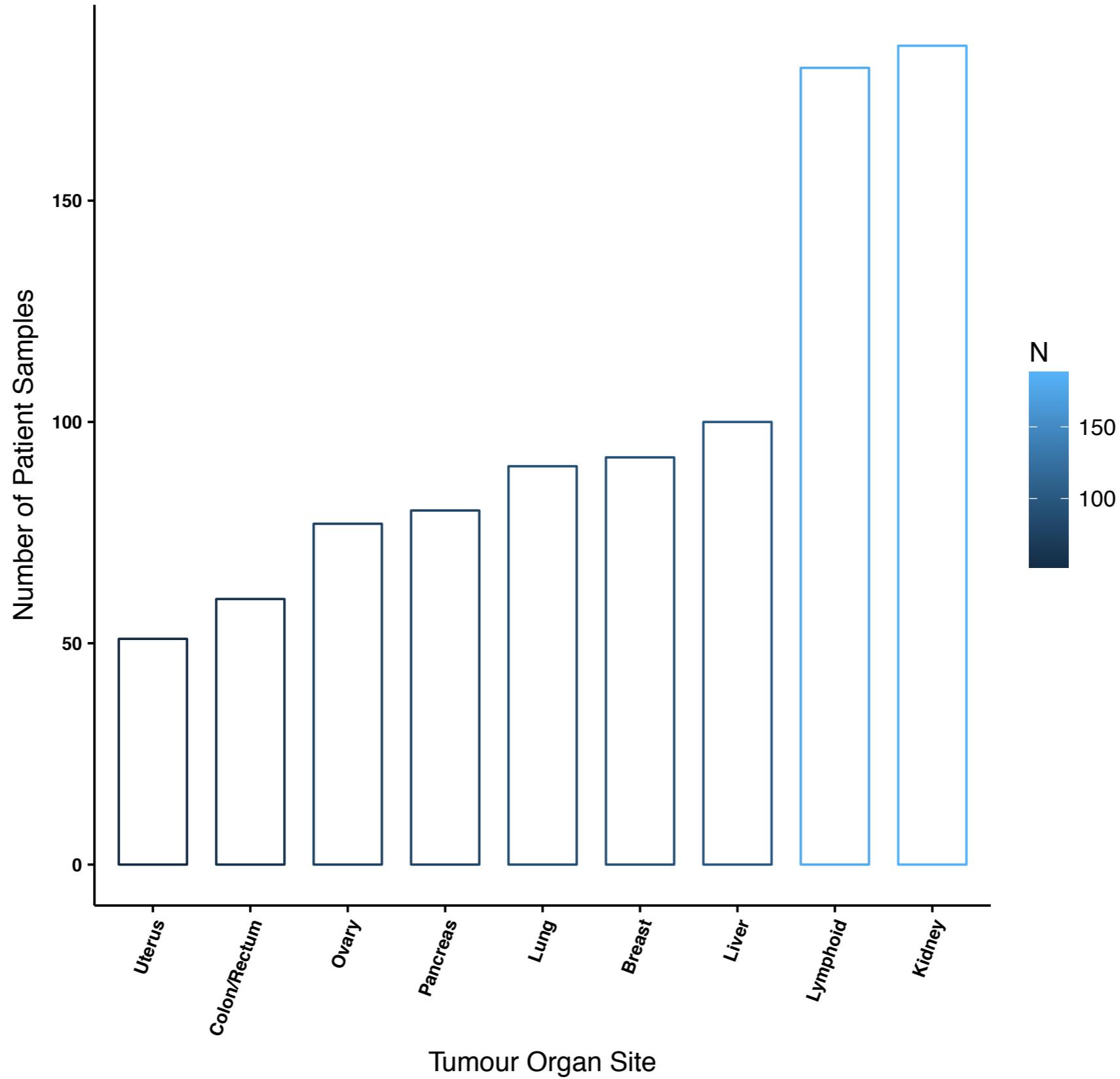
1A. Identifying cancer types to use in study

1,267 PCAWG patients with tumour RNA-Seq data from the following cancer types:



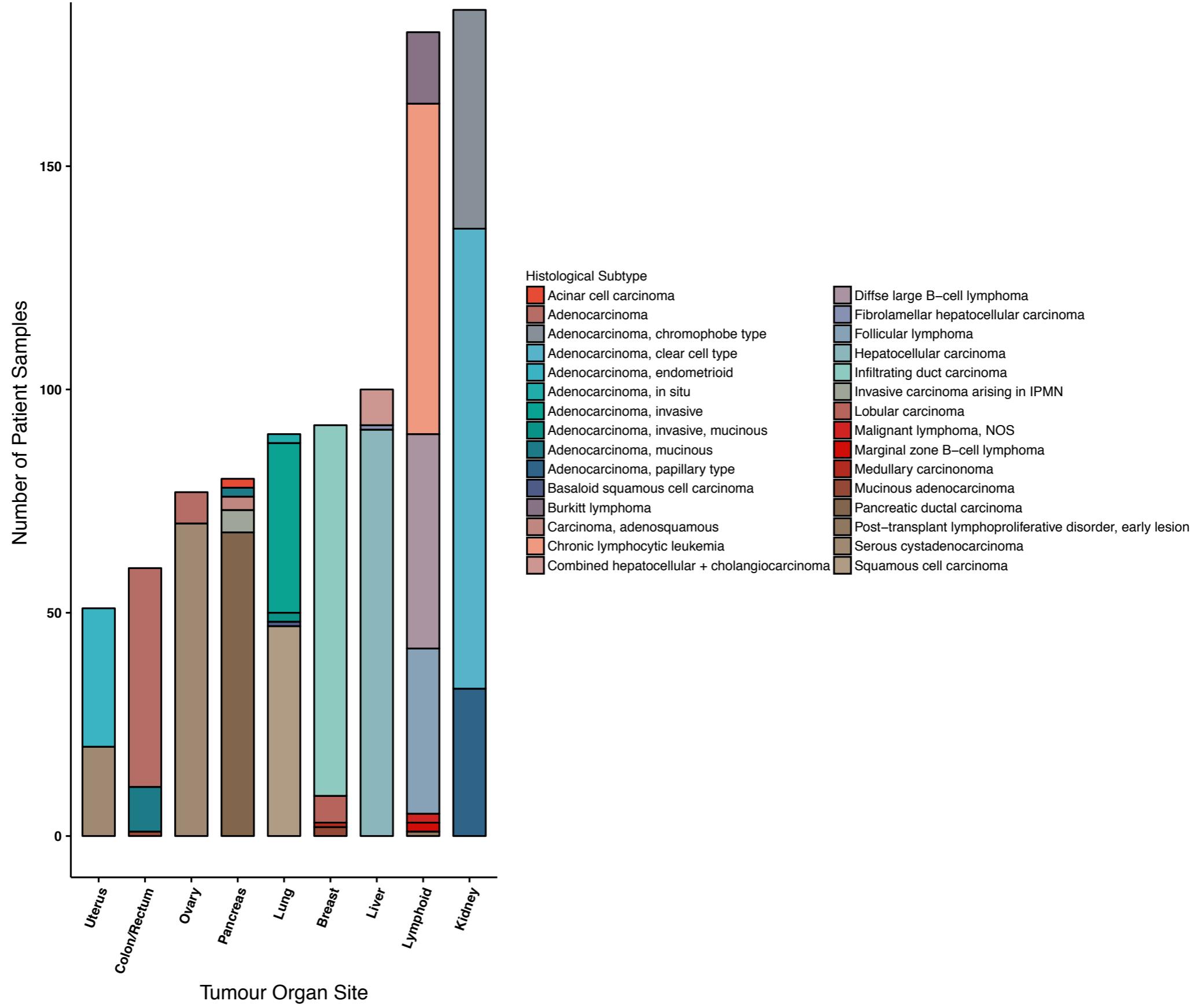
1A. Identifying cancer types to use in study

Cancers that have at least 50 patient samples:



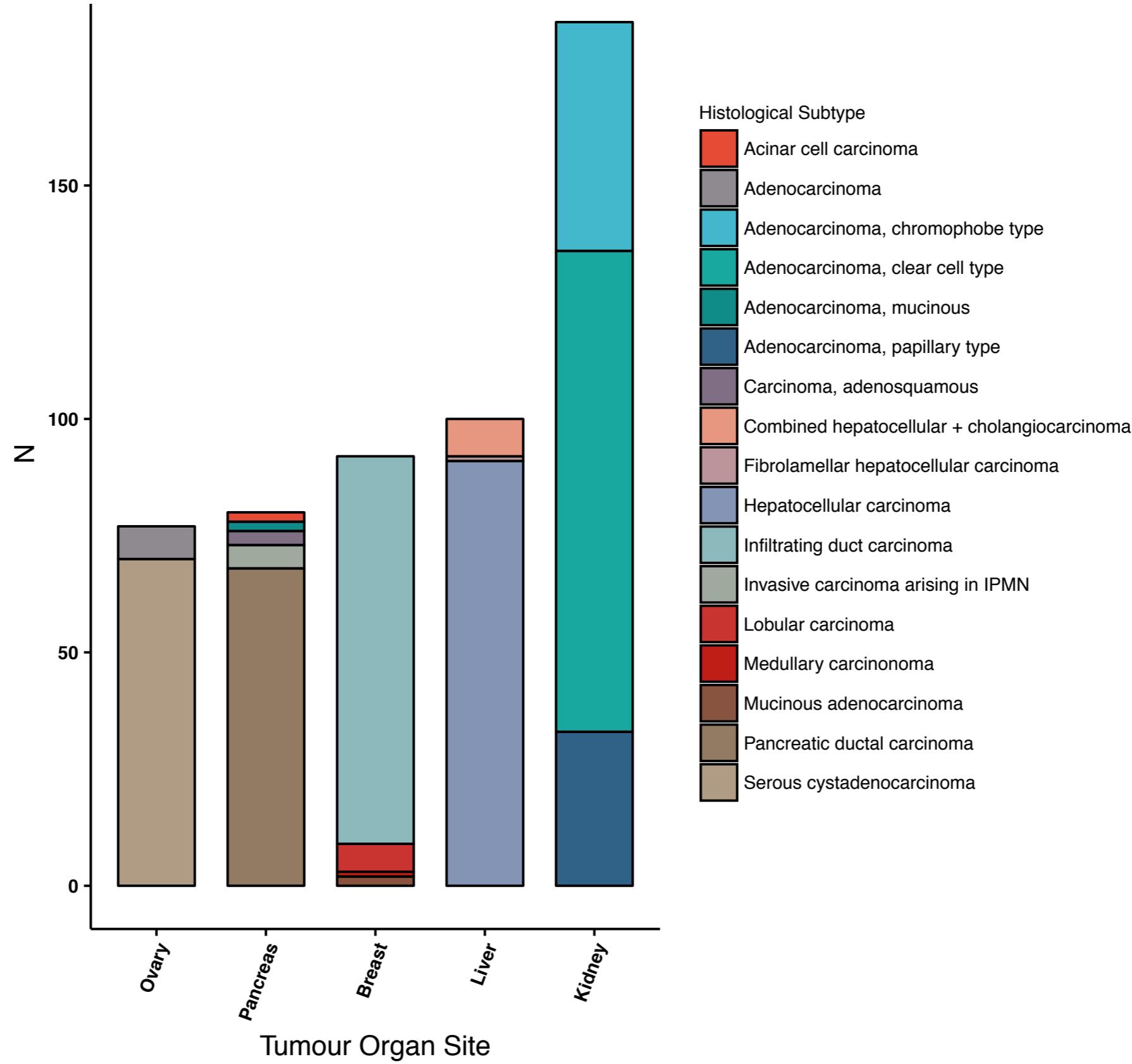
1A. Identifying cancer types to use in study

Distribution of histological subtypes within each of the cancers with 50 or more patient samples:



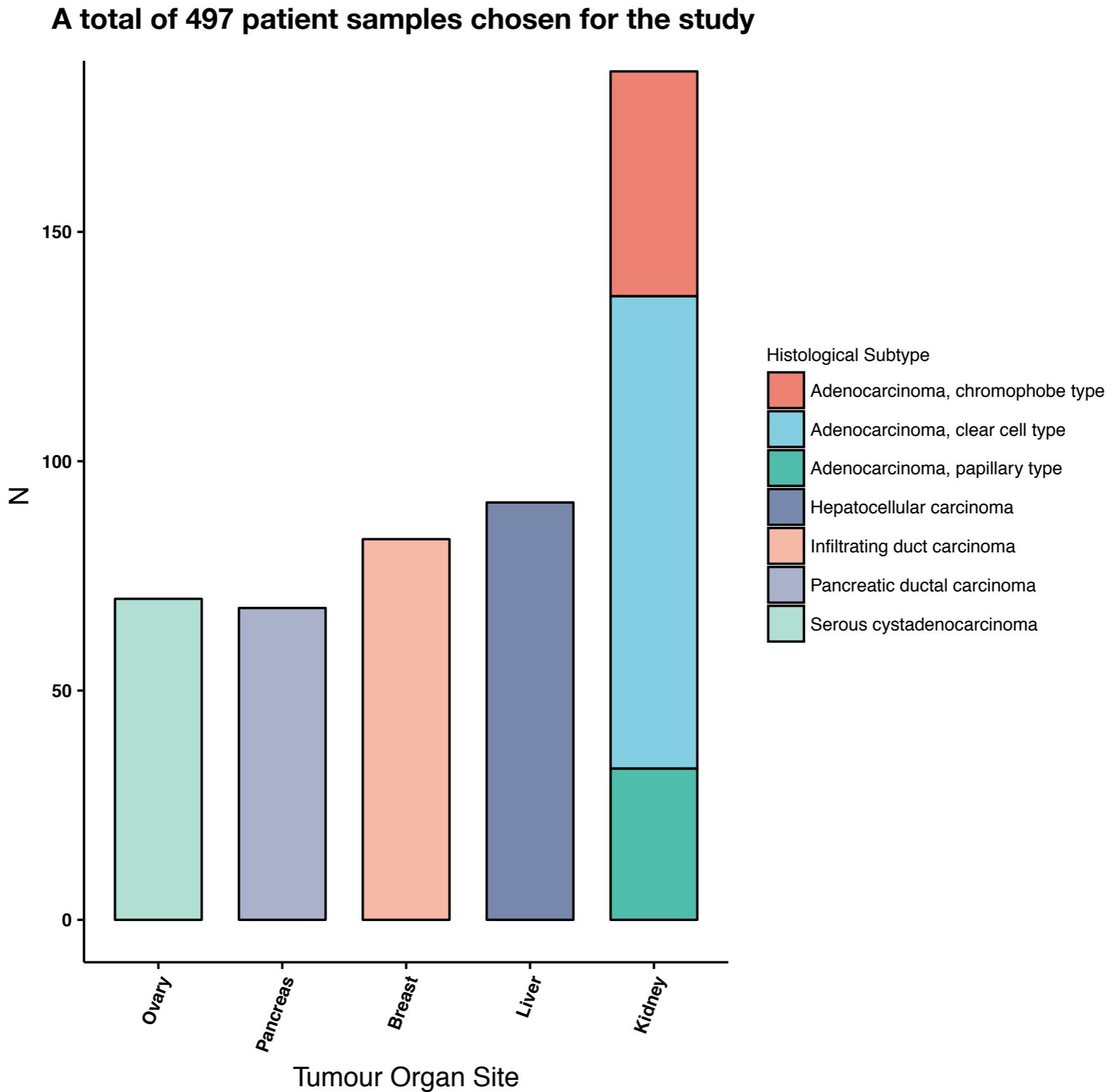
1A. Identifying cancer types to use in study

Cancers where at least one histological subtype has 50 patient samples or more:



1A. Identifying cancer types to use in study

Finally, remove histological subtypes with 10 or less patients:



1B. Identifying cancer specific lncRNAs

Initial Filtering:

12,598 Ensemble lncRNAs (lincRNA and antisense) with RNA-Seq data



Obtain median expression value of each lncRNA in every cancer type



3,658 have median FPKM expression of 0 in every cancer type



7,328 have median FPKM expression of 0 in every cancer type once median values are floored

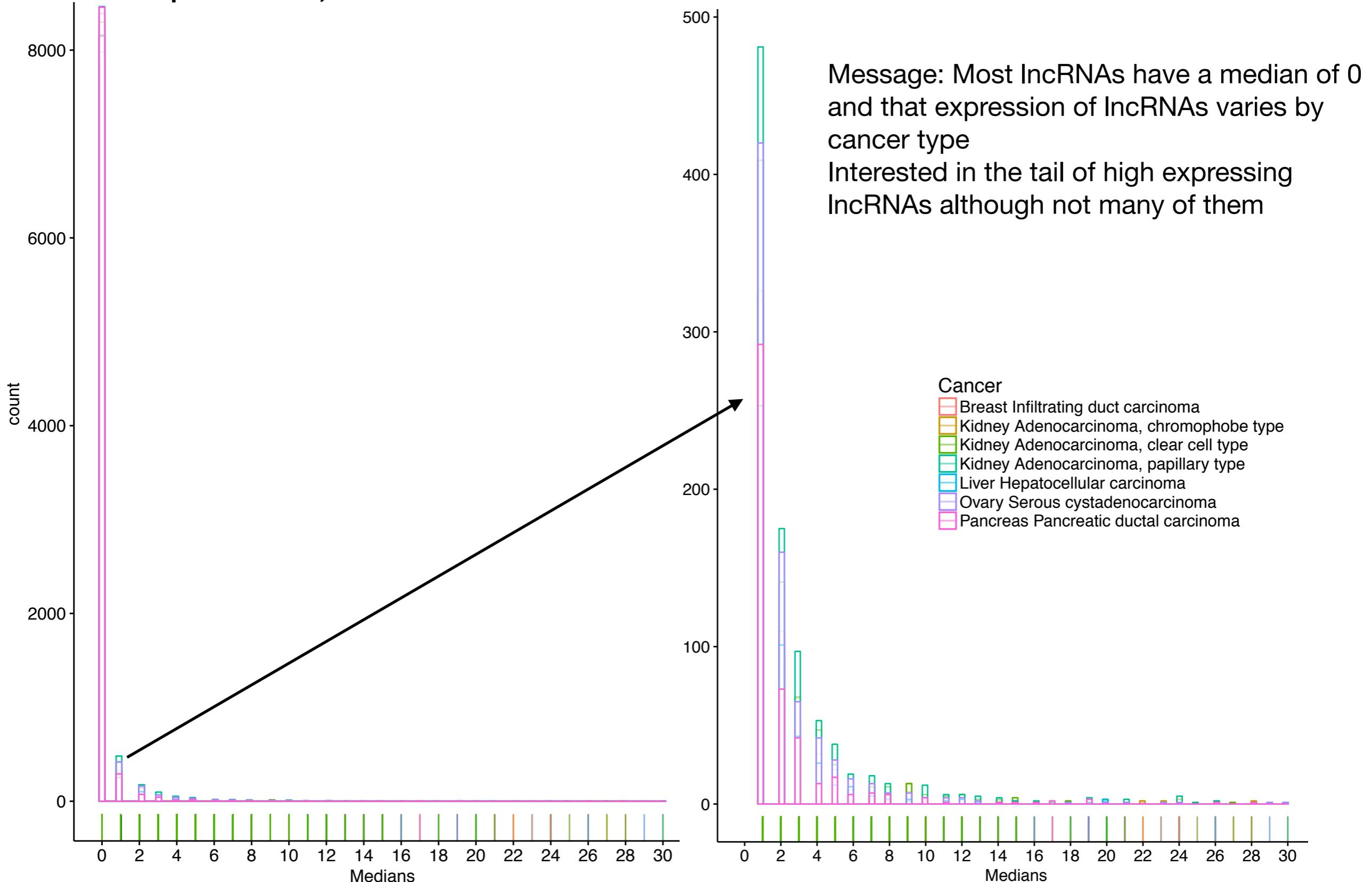


1,604 lncRNAs remain that have a median expression of at least 1 FPKM in at least one cancer type

1B. Identifying cancer specific lncRNAs

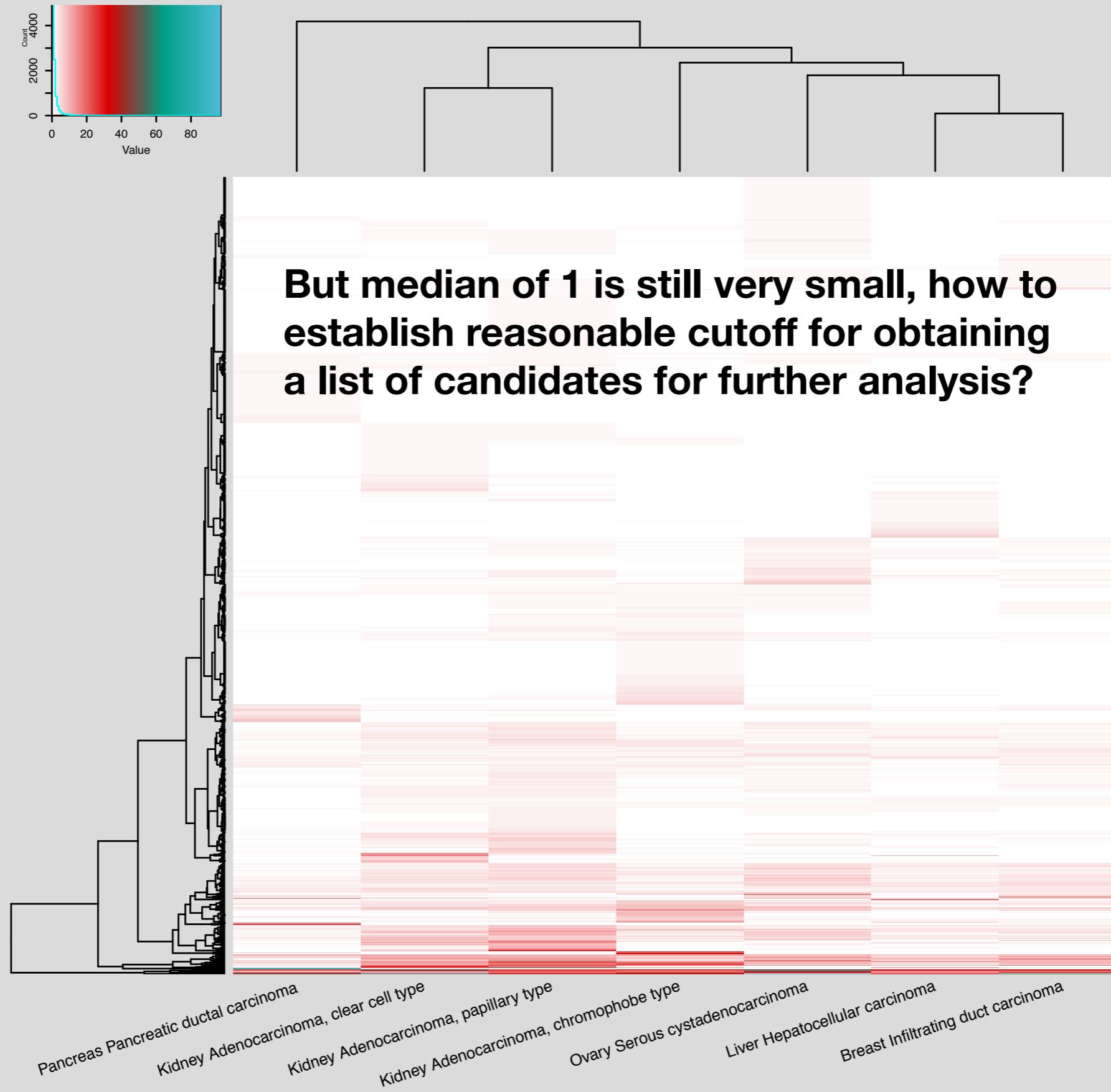
How are lncRNAs expressed among the different cancer types?

Median Expression of 8,940 lncRNAs Across Cancers

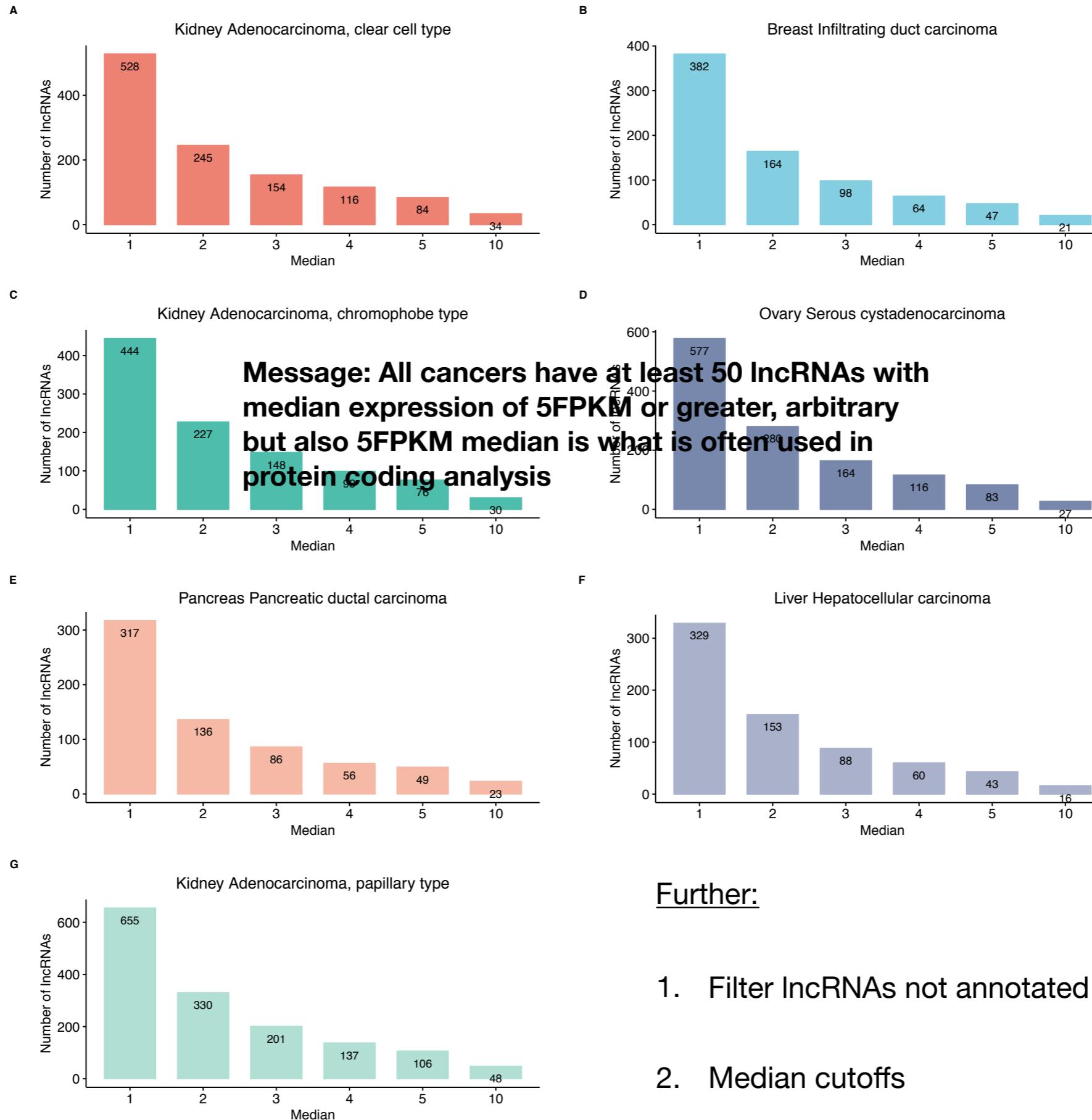


1B. Identifying cancer specific lncRNAs

Median Expression of 1,604 lncRNAs Across Cancers



1B. Identifying cancer specific lncRNAs



Further:

1. Filter lncRNAs not annotated in FANTOM
2. Median cutoffs

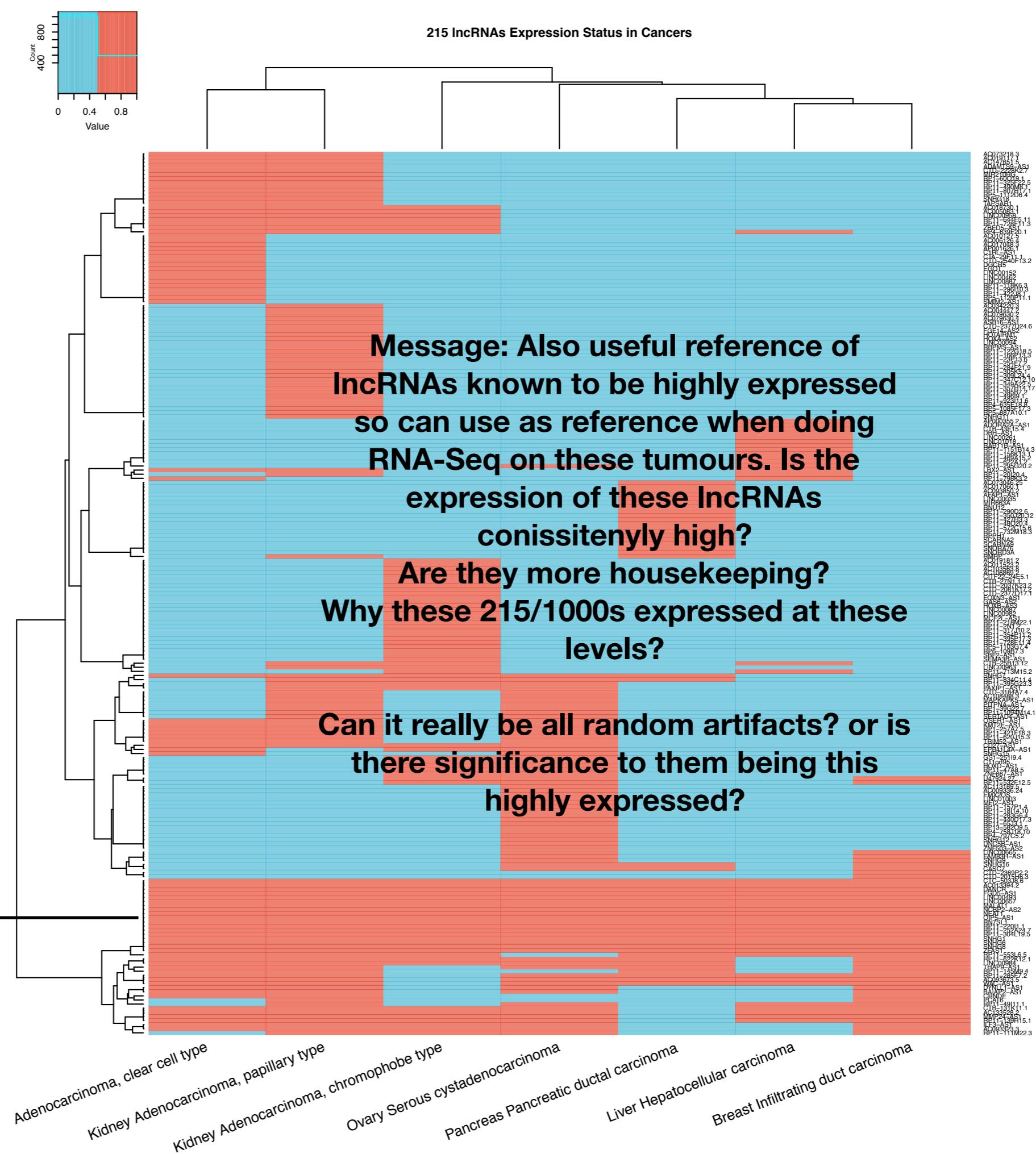
1B. Identifying cancer specific lncRNAs

Summary:

- 12,598 Ensemble lncRNAs in RNA-Seq dataset
- Filter 1: 5,607 also in FANTOM
- Filter 2: 215 unique lncRNAs selected for study with median FPKM ≥ 5 in at least one cancer type

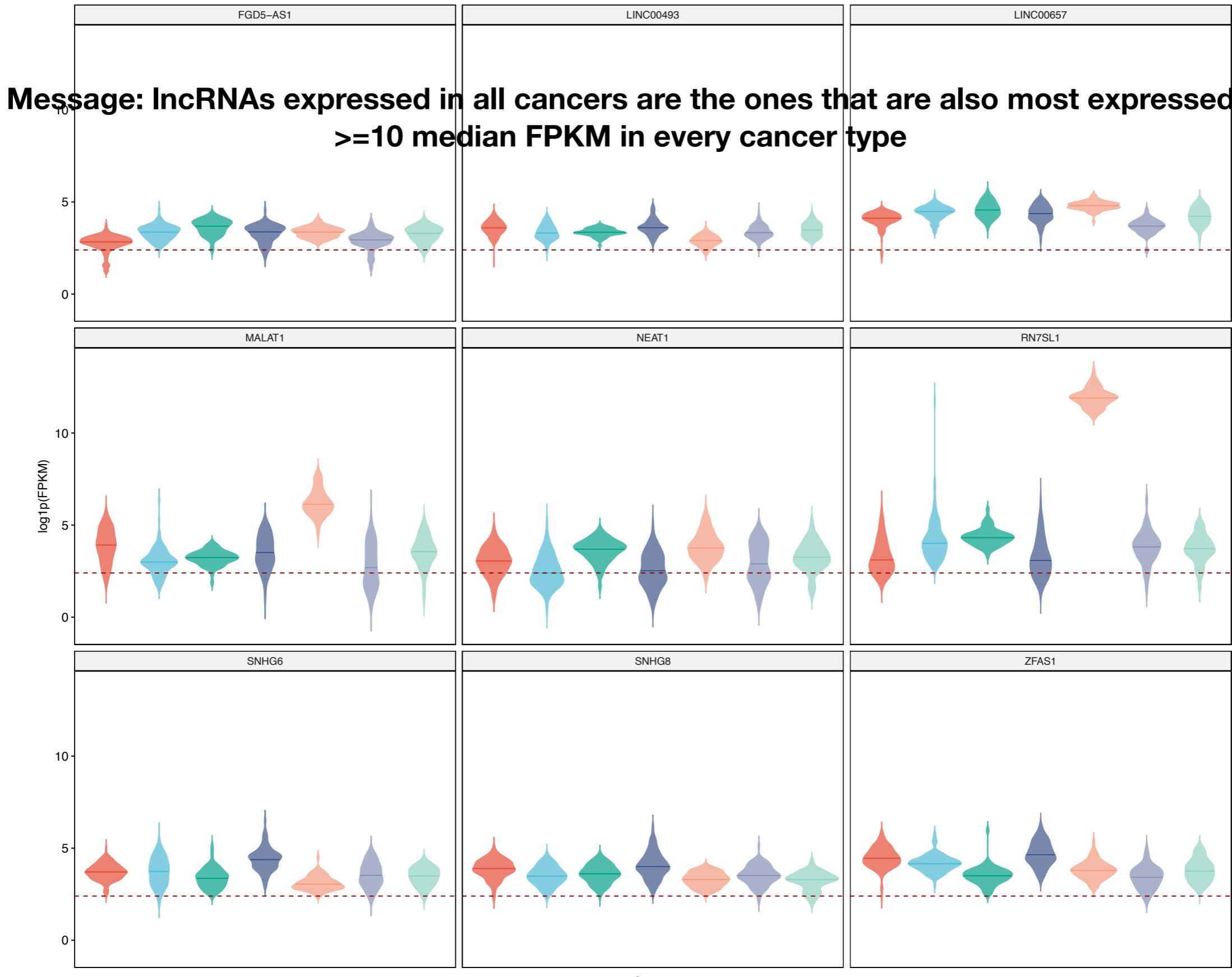
	Cancer	High Expressing lncRNAs
1	Liver Hepatocellular carcinoma	45
2	Breast Infiltrating duct carcinoma	47
3	Pancreas Pancreatic ductal carcinoma	49
4	Kidney Adenocarcinoma, chromophobe type	76
5	Ovary Serous cystadenocarcinoma	83
6	Kidney Adenocarcinoma, clear cell type	84
7	Kidney Adenocarcinoma, papillary type	109

**Message: To see
how many
overlapped high
expressing
lncRNAs versus
how many are
just in one
cancer type**



1B. Identifying cancer specific lncRNAs

Cancer
Kidney Adenocarcinoma, clear cell type
Kidney Adenocarcinoma, chromophobe type
Pancreas Pancreatic ductal carcinoma
Kidney Adenocarcinoma, papillary type
Breast Infiltrating duct carcinoma
Ovary Serous cystadenocarcinoma
Liver Hepatocellular carcinoma



1B. Identifying cancer specific lncRNAs



**PCA Using Most Highly Expressed Cancer Specific lncRNAs, n=215
(median ≥ 5 FPKM in at least one cancer type)**

- Breast Infiltrating duct carcinoma
- Kidney Adenocarcinoma, chromophobe type
- Kidney Adenocarcinoma, clear cell type
- Kidney Adenocarcinoma, papillary type
- Liver Hepatocellular carcinoma
- Ovary Serous cystadenocarcinoma
- Pancreas Pancreatic ductal carcinoma

Message: these 215 lncRNAs chosen based on median of 5FPKM show good separation of cancer types.

Summary of progress

1. Identified high expressing cancer specific lncRNAs through several filtration steps across multiple cancer types.
2. Identified cancer specific high expressing lncRNAs associated with patient survival outcome using the Cox Proportional Hazards model
 - 42 lncRNAs in 4 cancer types
 - 7 lncRNAs with adjusted p-value < 0.1 in 2 cancer types (Tier 1)
 - 35 lncRNAs with p-value < 0.05 in 4 cancer types (Tier 2)
3. Assessed tumour versus normal tissue lncRNA expression by comparing ranked values of expression using match tissue data from GTEx
4. 21/42 lncRNAs were predicted to be either tumour suppressive or oncogenic based on sign of hazard ratio and difference in expression between normal and tumour tissues
5. 2 candidates in Ovarian cancer were further studied to predict regulatory roles that may lead patients from the same cancer type to develop worse prognosis

Summary of progress

1. Identified high expressing cancer specific lncRNAs through several filtration steps across multiple cancer types.
2. Confirmed cancer specific expression through comparing the distributions of each candidate lncRNA's expression among the cancer types studied
3. **Identified cancer specific lncRNAs associated with patient survival outcome using the Cox Proportional Hazards model**
 - 42 lncRNAs in 4 cancer types
 - 7 lncRNAs with adjusted p-value < 0.1 in 2 cancer types (Tier 1)
 - 35 lncRNAs with p-value < 0.05 in 4 cancer types (Tier 2)
4. Assessed tumour versus normal tissue lncRNA expression by comparing ranked values of expression using match tissue data from GTEx
5. 21/42 lncRNAs were predicted to be either tumour suppressive or oncogenic based on sign of hazard ratio and difference in expression between normal and tumour tissues
6. 2 candidates in Ovarian cancer were further studied to predict regulatory roles that may lead patients from the same cancer type to develop worse prognosis

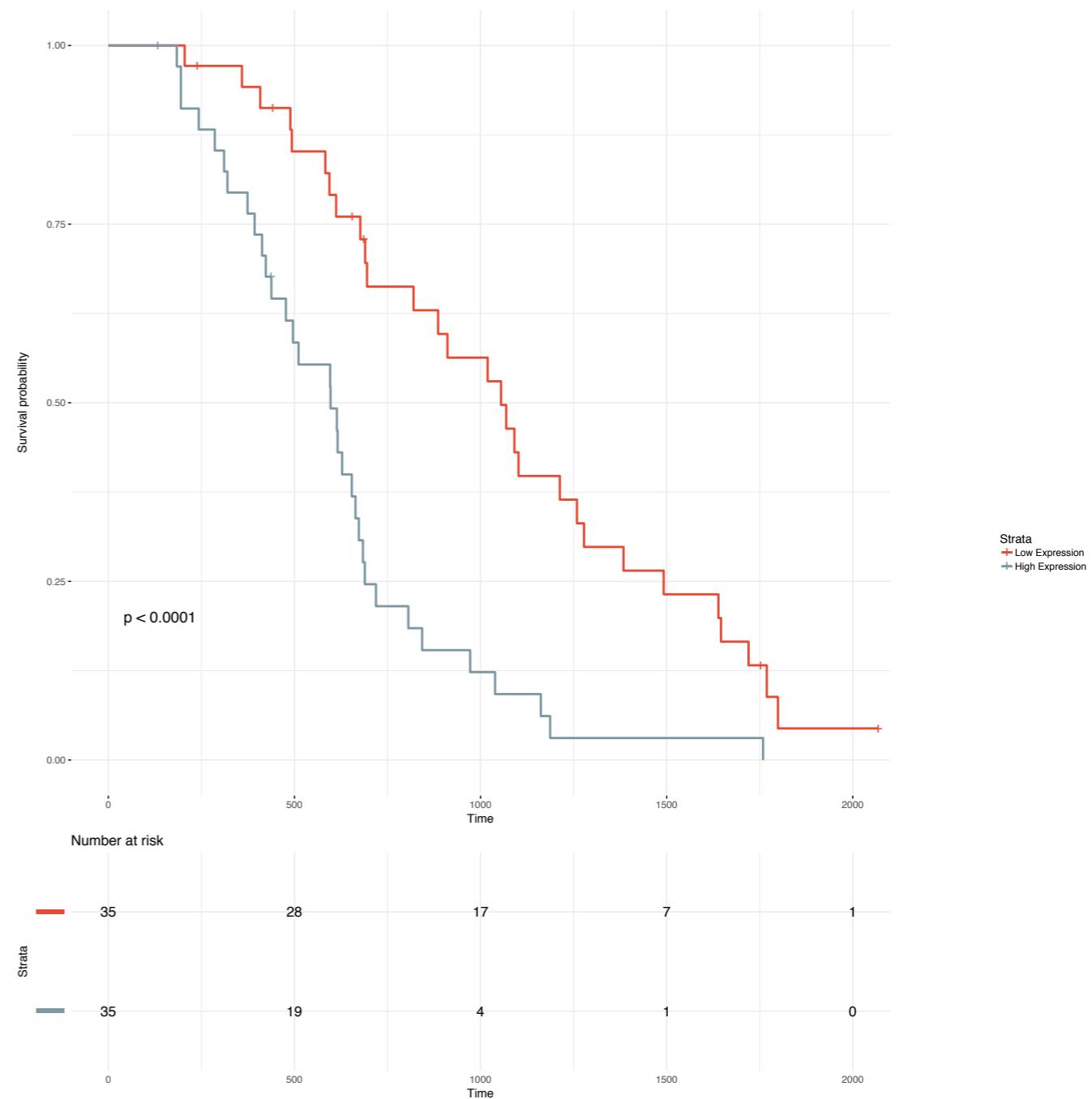
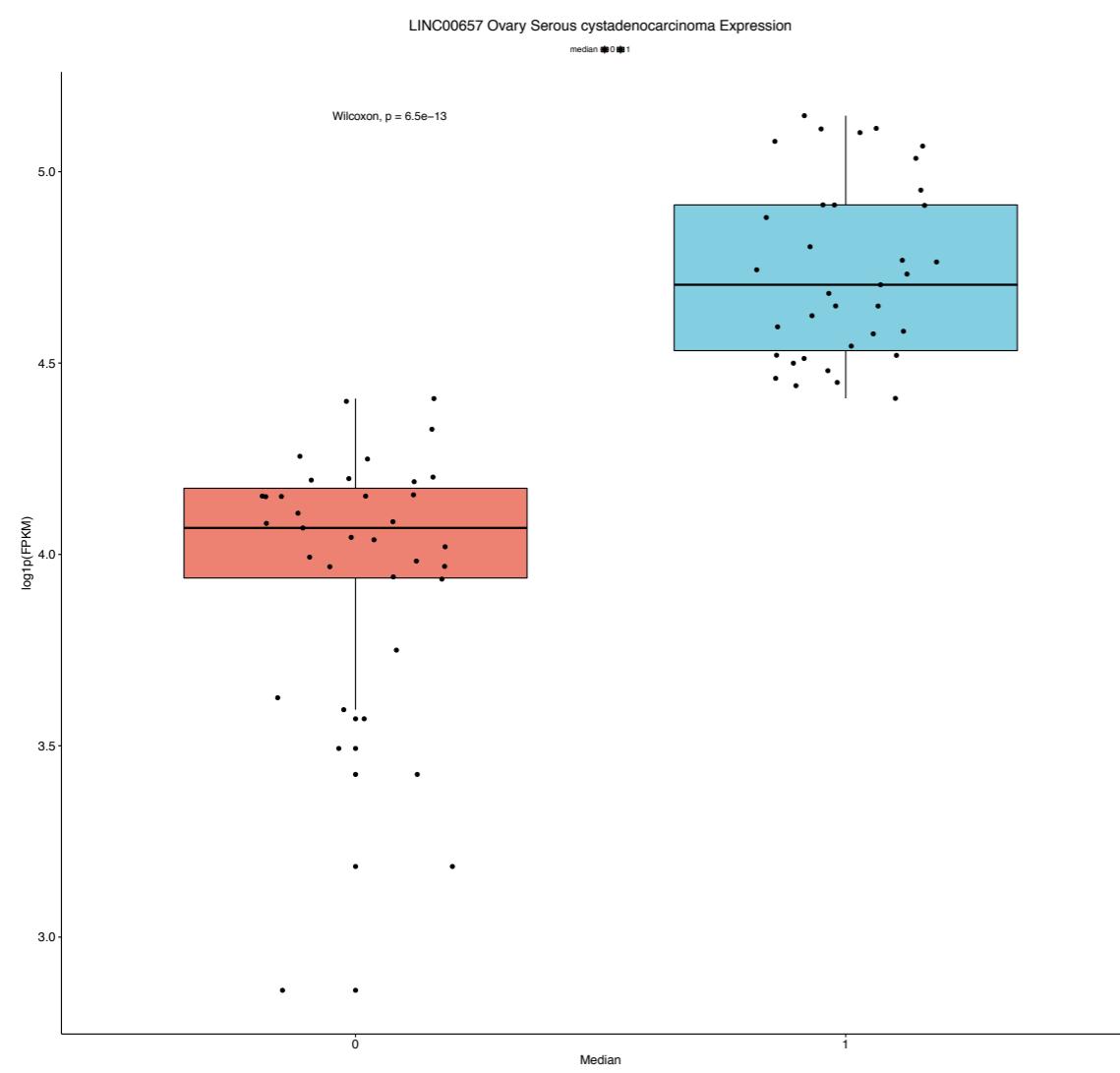
3. Identified lncRNAs associated with patient survival outcome using the Cox Proportional Hazards model

- Median expression of lncRNA used to dichotomize patients into high and low groups
- Predictor of survival is survival tag:
 - 1 = high
 - 0 = low
- `res.cox <- coxph(Surv(time, status) ~ median, data = df)`

	gene	coef	HR	pval	canc	fdr
1	LINC00657	1.09089497046119	2.97693715151224	6.714862e-05	Ovary Serous cystadenocarcinoma	0.01103476
2	LINC00665	-1.23789638086889	0.28999361285883	4.933441e-05	Ovary Serous cystadenocarcinoma	0.01103476
3	ZNF503-AS2	-1.21096375798916	0.29791002786222	5.909242e-05	Ovary Serous cystadenocarcinoma	0.01103476
4	AC009336.24	-0.915329978890643	0.400384485861799	6.039664e-04	Ovary Serous cystadenocarcinoma	0.05014560
5	RP11-622K12.1	-0.969070407483862	0.379435594700213	6.102913e-04	Ovary Serous cystadenocarcinoma	0.05014560
6	AP000355.2	-1.50935541867795	0.221052418307288	4.777307e-04	Liver Hepatocellular carcinoma	0.05014560
7	RP11-622A1.2	-1.42689280391215	0.240053658412211	1.416813e-03	Liver Hepatocellular carcinoma	0.09978412
8	SNHG15	0.778032349402381	2.17718411047022	3.557492e-03	Ovary Serous cystadenocarcinoma	0.19487153
9	RP11-220I1.1	-1.31154573241281	0.269403308950893	3.371553e-03	Liver Hepatocellular carcinoma	0.19487153
10	SNHG12	0.764311259067298	2.14751478385111	4.537974e-03	Ovary Serous cystadenocarcinoma	0.22372214
11	SNHG1	0.743189814071518	2.10263183341979	5.168054e-03	Ovary Serous cystadenocarcinoma	0.23162277
12	GS1-251I9.4	0.725427456398557	2.06561387114612	6.281176e-03	Ovary Serous cystadenocarcinoma	0.23892283
13	RP11-295G20.2	1.23237659225222	3.42937007343585	6.300196e-03	Liver Hepatocellular carcinoma	0.23892283
14	OSER1-AS1	0.699393073114752	2.01253087763027	9.189207e-03	Ovary Serous cystadenocarcinoma	0.31979542
15	RP11-157P1.4	0.66631704328976	1.94705318673016	9.730084e-03	Ovary Serous cystadenocarcinoma	0.31979542
16	LINC00493	1.07830444424117	2.93969091314685	1.054080e-02	Liver Hepatocellular carcinoma	0.32478839
17	U47924.27	0.680113191614456	1.97410117128317	1.176238e-02	Ovary Serous cystadenocarcinoma	0.34110901
18	AC006126.4	-0.918822841885361	0.398988437230102	1.473127e-02	Kidney Adenocarcinoma, clear cell type	0.36995216
19	NEAT1	0.642552733080203	1.90132827320059	1.365801e-02	Ovary Serous cystadenocarcinoma	0.36995216
20	MMP24-AS1	0.998979579568135	2.71550945287663	1.500820e-02	Liver Hepatocellular carcinoma	0.36995216
21	FGD5-AS1	-0.866125022806513	0.420578126385491	1.824250e-02	Kidney Adenocarcinoma, clear cell type	0.40879774
22	ADORA2A-AS1	-1.0238468792094	0.359210439705118	1.798022e-02	Liver Hepatocellular carcinoma	0.40879774
23	RP11-49I11.1	0.979396006633756	2.66284741394427	1.917262e-02	Liver Hepatocellular carcinoma	0.41096102
24	RP11-304L19.5	-0.727854754389591	0.48294391292074	2.111194e-02	Pancreas Pancreatic ductal carcinoma	0.43367449
25	MMP24-AS1	0.825276522328797	2.28251184356216	2.453876e-02	Kidney Adenocarcinoma, clear cell type	0.45015634
26	NCBP2-AS2	0.601706131619117	1.82523022838774	2.465359e-02	Ovary Serous cystadenocarcinoma	0.45015634
27	NEAT1	-0.931193692129404	0.394083015701056	2.327743e-02	Liver Hepatocellular carcinoma	0.45015634
28	CTB-131K11.1	0.579258152521435	1.78471395398422	2.992787e-02	Ovary Serous cystadenocarcinoma	0.48644962
29	DANCR	-0.569578402465635	0.56576391309669	2.961290e-02	Ovary Serous cystadenocarcinoma	0.48644962
30	FAM83H-AS1	-0.572351978944779	0.564196897737599	2.971973e-02	Ovary Serous cystadenocarcinoma	0.48644962
31	DANCR	0.910329458260571	2.48514114880777	3.058811e-02	Liver Hepatocellular carcinoma	0.48644962
32	RP11-834C11.4	-0.563198488147974	0.569384977143017	3.520396e-02	Ovary Serous cystadenocarcinoma	0.52126585
33	RPPH1	-0.639134723510223	0.527748875230572	3.412066e-02	Pancreas Pancreatic ductal carcinoma	0.52126585
34	RP11-139H15.1	0.895710166661311	2.44907442303575	3.594937e-02	Liver Hepatocellular carcinoma	0.52126585
35	CTC-503J8.6	-0.551357810106765	0.576166953702487	3.954573e-02	Ovary Serous cystadenocarcinoma	0.53743244
36	MFI2-AS1	0.539284764289622	1.7147799516152	3.840014e-02	Ovary Serous cystadenocarcinoma	0.53743244
37	RP11-469A15.2	-0.881443557345828	0.414184580728924	4.033469e-02	Liver Hepatocellular carcinoma	0.53743244
38	HOXD-AS1	-0.521580642937459	0.593581565558328	4.176834e-02	Ovary Serous cystadenocarcinoma	0.54188927
39	SNHG7	0.734356702478046	2.08414083853393	4.533632e-02	Kidney Adenocarcinoma, clear cell type	0.55960215
40	RP11-834C11.4	-0.600714941409506	0.548419408156069	4.563384e-02	Pancreas Pancreatic ductal carcinoma	0.55960215
41	DBH-AS1	-0.83141909496413	0.435430929832266	4.653892e-02	Liver Hepatocellular carcinoma	0.55960215
42	SNHG8	-0.599508861430317	0.549081244858215	4.976660e-02	Pancreas Pancreatic ductal carcinoma	0.58416513

3. Identified lncRNAs associated with patient survival outcome using the Cox Proportional Hazards model

Top Hit: LINC00657 in Ovarian Cancer with Hazard Ratio of 2.97



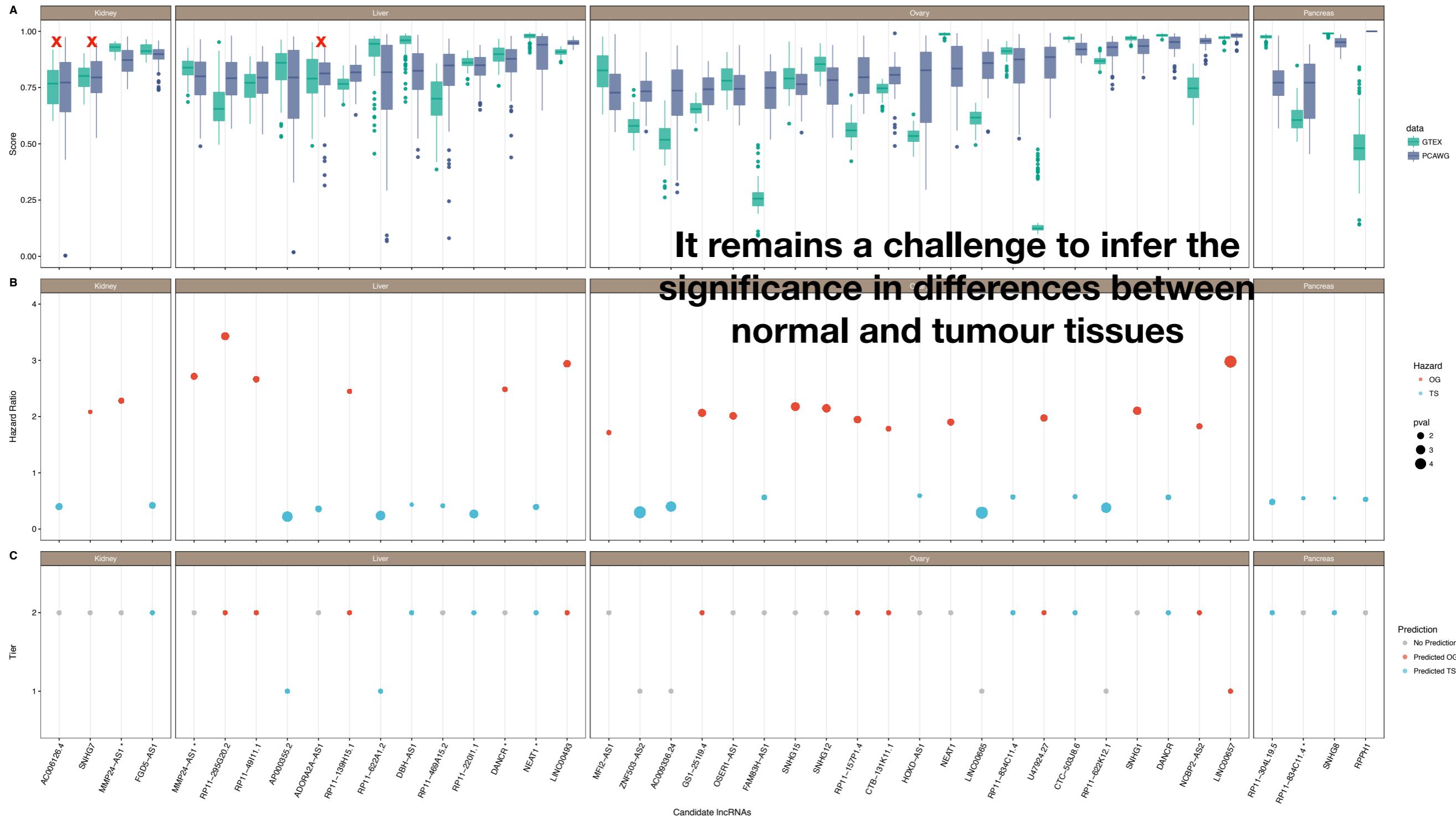
Summary of progress

1. Identified high expressing cancer specific lncRNAs through several filtration steps across multiple cancer types.
2. Confirmed cancer specific expression through comparing the distributions of each candidate lncRNA's expression among the cancer types studied
3. Identified cancer specific high expressing lncRNAs associated with patient survival outcome using the Cox Proportional Hazards model
 - 42 lncRNAs in 4 cancer types
 - 7 lncRNAs with adjusted p-value < 0.1 in 2 cancer types (Tier 1)
 - 35 lncRNAs with p-value < 0.05 in 4 cancer types (Tier 2)
4. **Assessed tumour versus normal tissue lncRNA expression by comparing ranked values of expression using tissue data from GTEx**
5. 21/42 lncRNAs were predicted to be either tumour suppressive or oncogenic based on sign of hazard ratio and difference in expression between normal and tumour tissues
6. 2 candidates in Ovarian cancer were further studied to predict regulatory roles that may lead patients from the same cancer type to develop worse prognosis

Summary of progress

1. Identified high expressing cancer specific lncRNAs through several filtration steps across multiple cancer types.
2. Confirmed cancer specific expression through comparing the distributions of each candidate lncRNA's expression among the cancer types studied
3. Identified cancer specific high expressing lncRNAs associated with patient survival outcome using the Cox Proportional Hazards model
 - 42 lncRNAs in 4 cancer types
 - 7 lncRNAs with adjusted p-value < 0.1 in 2 cancer types (Tier 1)
 - 35 lncRNAs with p-value < 0.05 in 4 cancer types (Tier 2)
4. Assessed tumour versus normal tissue lncRNA expression by comparing ranked values of expression using tissue data from GTEx
5. **21/42 lncRNAs were predicted to be either tumour suppressive or oncogenic based on sign of hazard ratio and difference in expression between normal and tumour tissues**
6. 2 candidates in Ovarian cancer were further studied to predict regulatory roles that may lead patients from the same cancer type to develop worse prognosis

4. Prediction of tumour suppressive or oncogenic properties



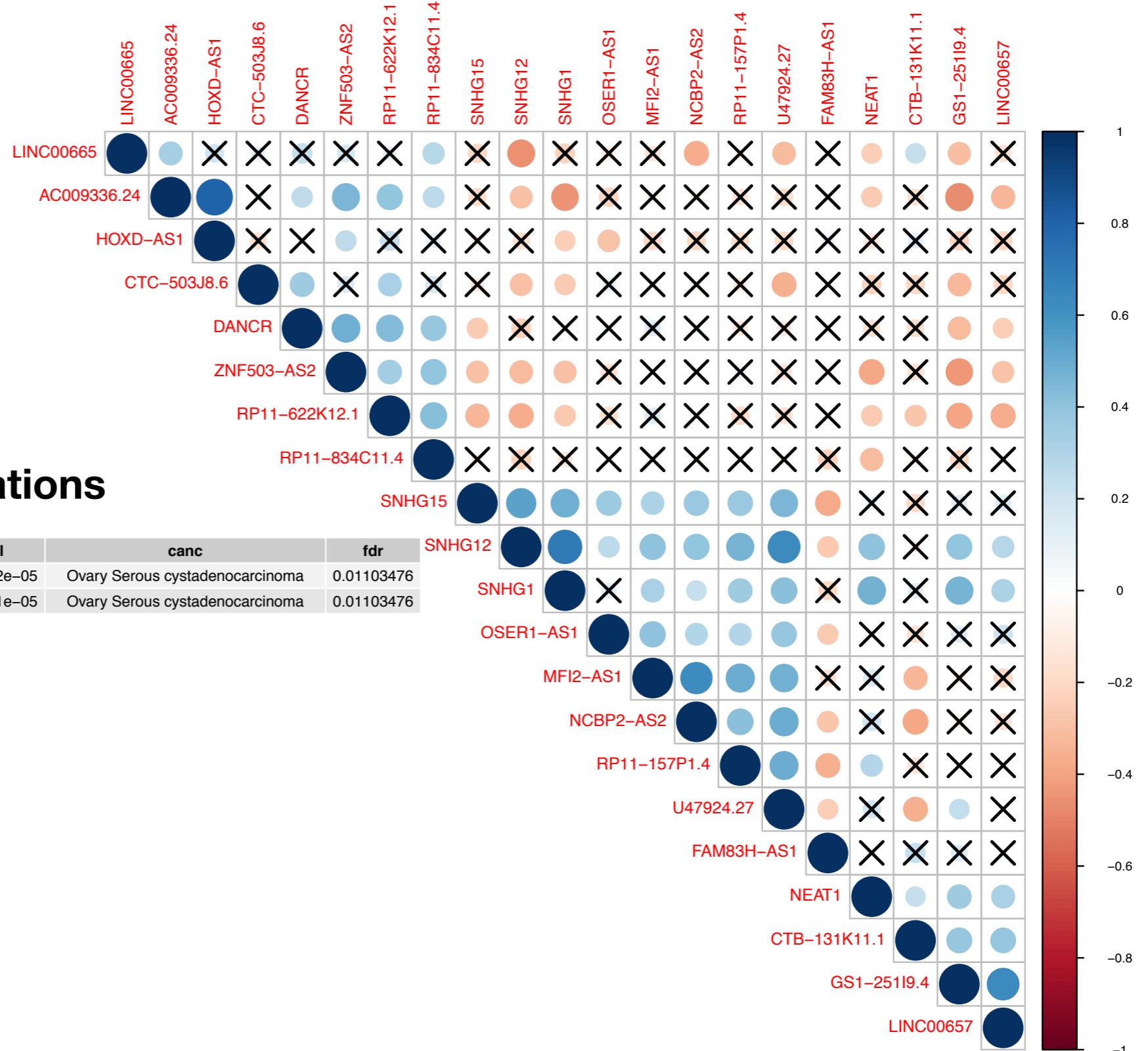
Summary of progress

1. Identified high expressing cancer specific lncRNAs through several filtration steps across multiple cancer types.
2. Confirmed cancer specific expression through comparing the distributions of each candidate lncRNA's expression among the cancer types studied
3. Identified cancer specific high expressing lncRNAs associated with patient survival outcome using the Cox Proportional Hazards model
 - 42 lncRNAs in 4 cancer types
 - 7 lncRNAs with adjusted p-value < 0.1 in 2 cancer types (Tier 1)
 - 35 lncRNAs with p-value < 0.05 in 4 cancer types (Tier 2)
4. Assessed tumour versus normal tissue lncRNA expression by comparing ranked values of expression using tissue data from GTEx
5. 21/42 lncRNAs were predicted to be either tumour suppressive or oncogenic based on sign of hazard ratio and difference in expression between normal and tumour tissues
6. **2 candidates in Ovarian cancer were further studied to predict regulatory roles that may lead patients from the same cancer type to develop worse prognosis**

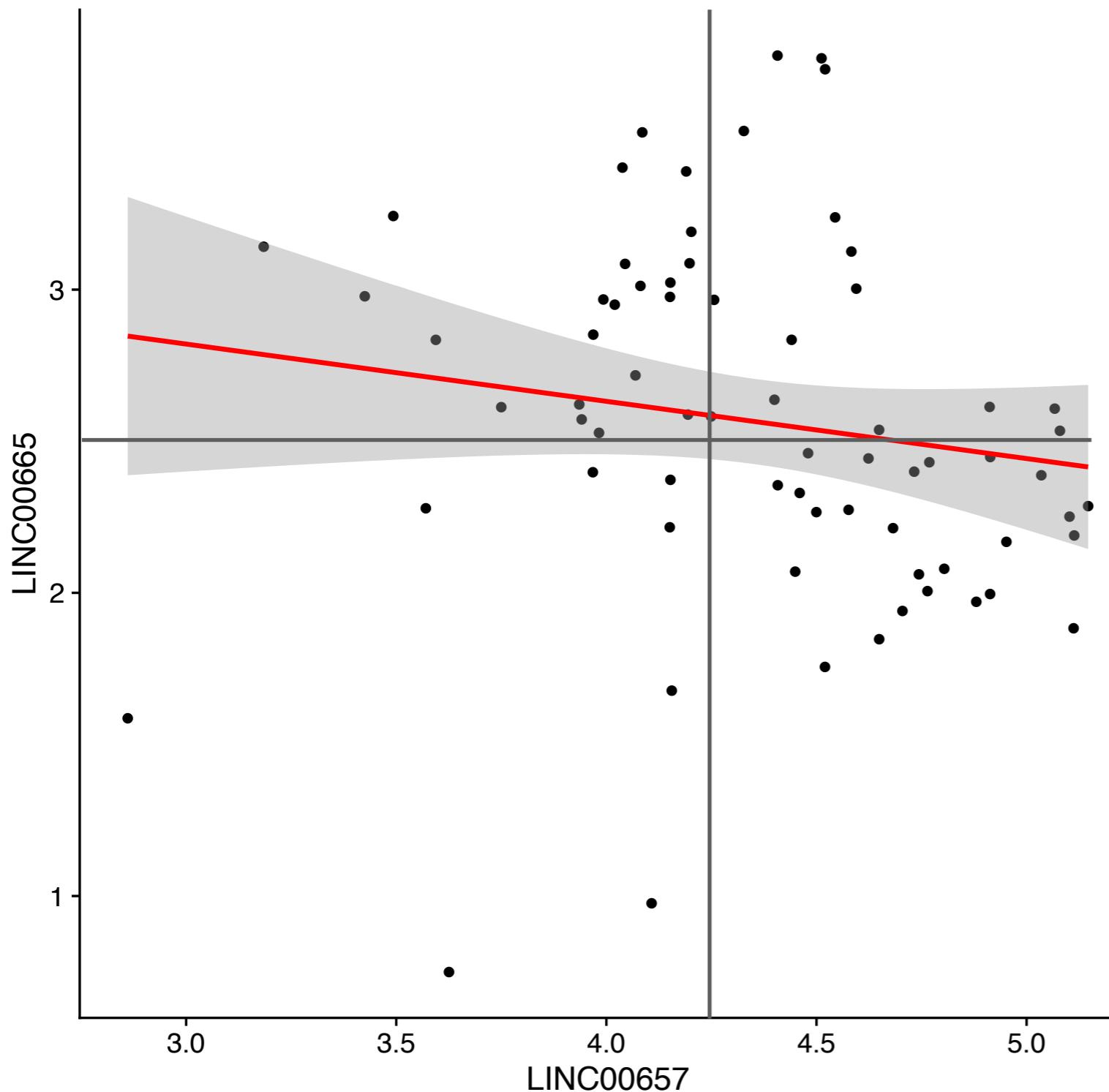
6. Further analysis of top 2 candidates in Ovarian Cancer

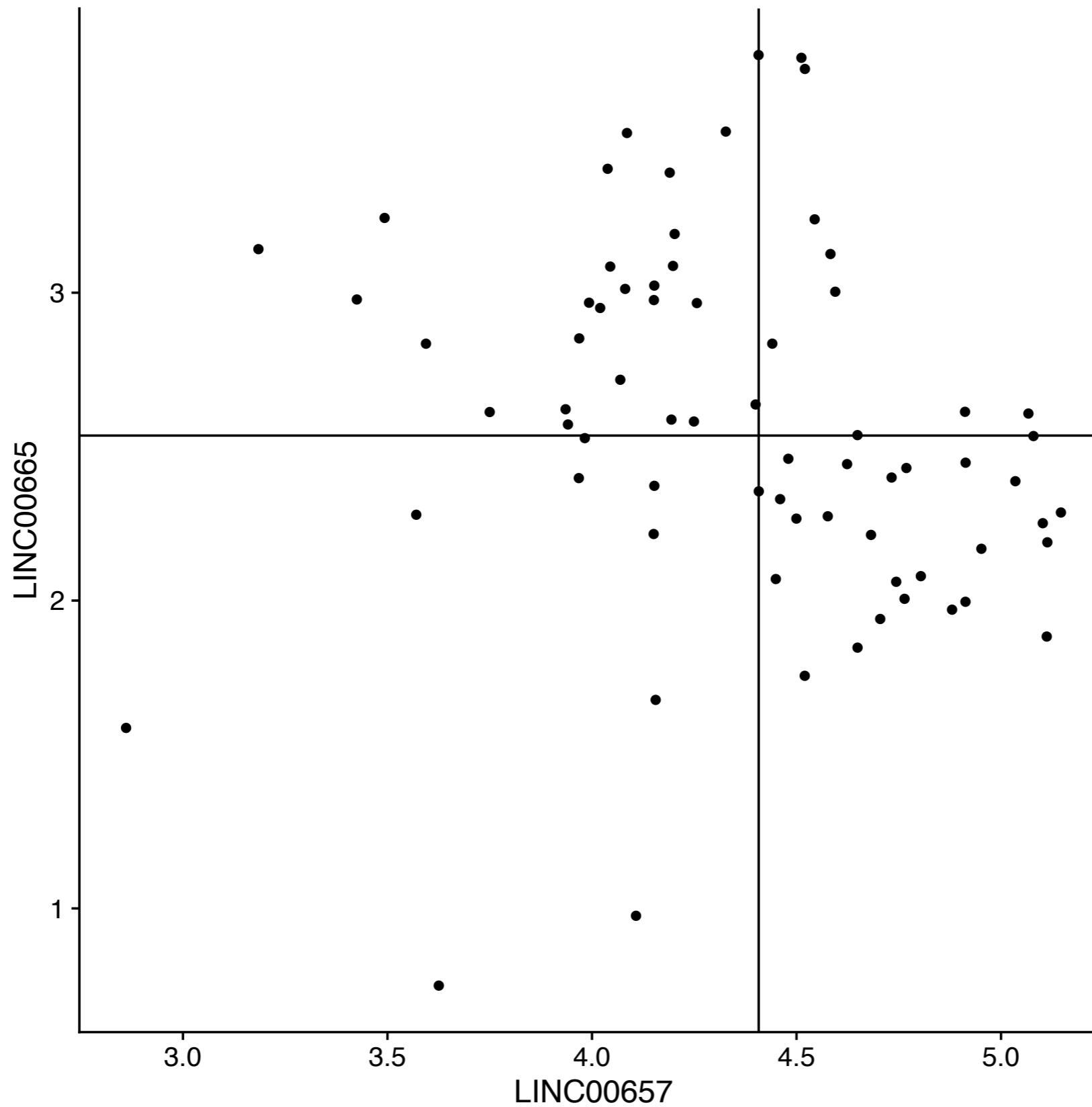
Ovarian Cancer Candidate lnc-lnc Correlations

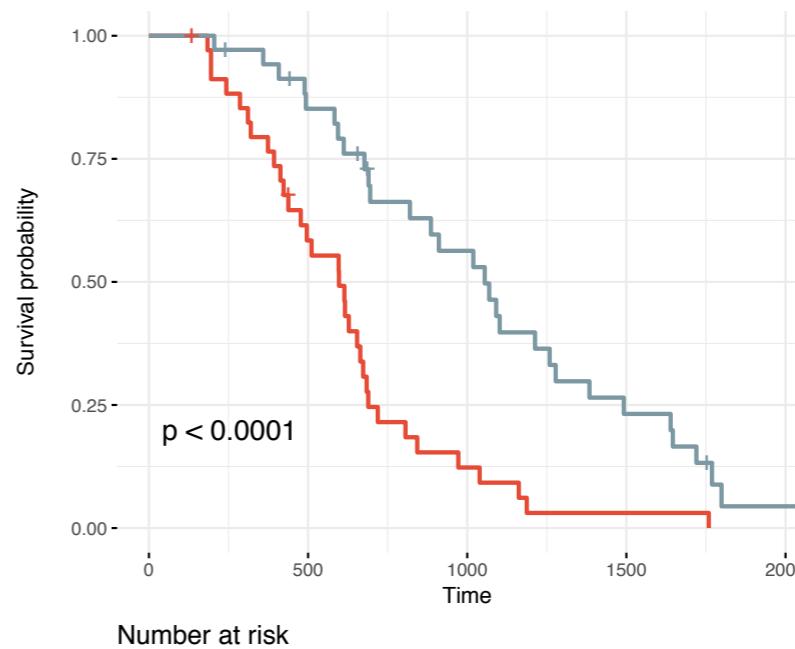
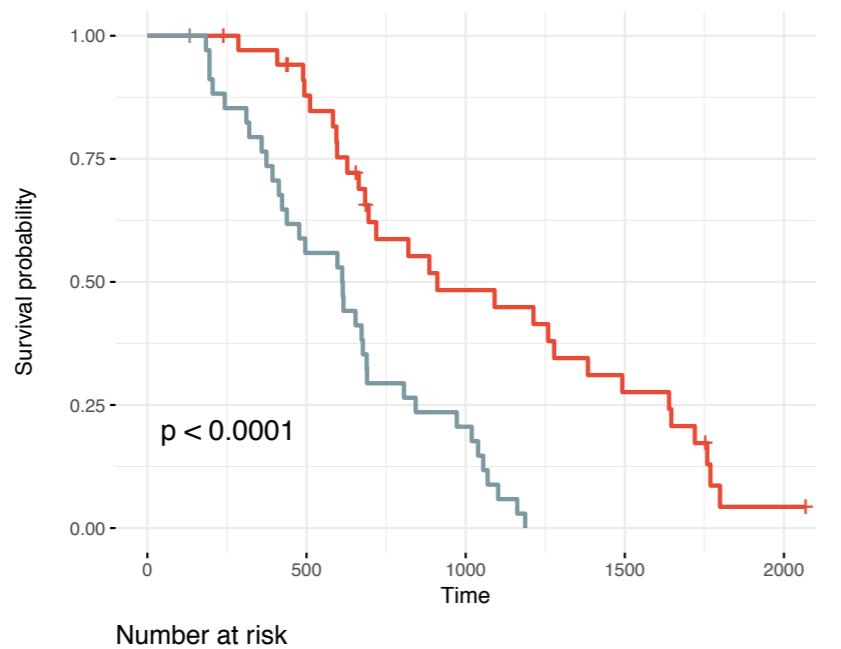
	gene	coef	HR	pval	canc	fdr
1	LINC00657	1.09089497046119	2.97693715151224	6.714862e-05	Ovary Serous cystadenocarcinoma	0.01103476
2	LINC00665	-1.23789638086889	0.28999361285883	4.933441e-05	Ovary Serous cystadenocarcinoma	0.01103476



Adj R2 = 0.0095012 Intercept = 3.387 Slope = -0.18884 P = 0.20172

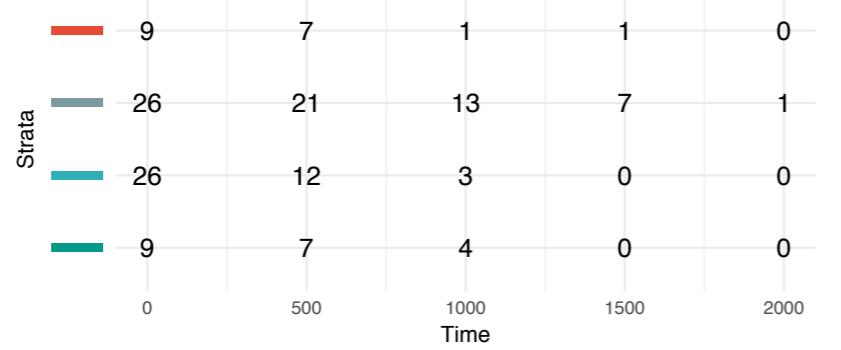
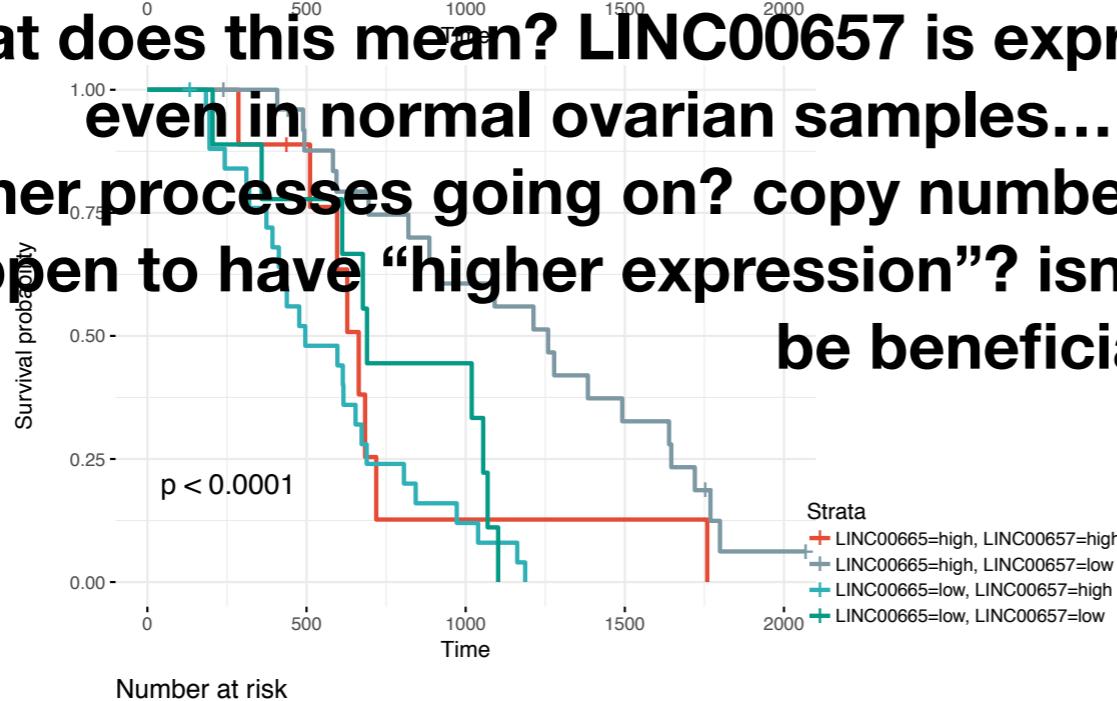




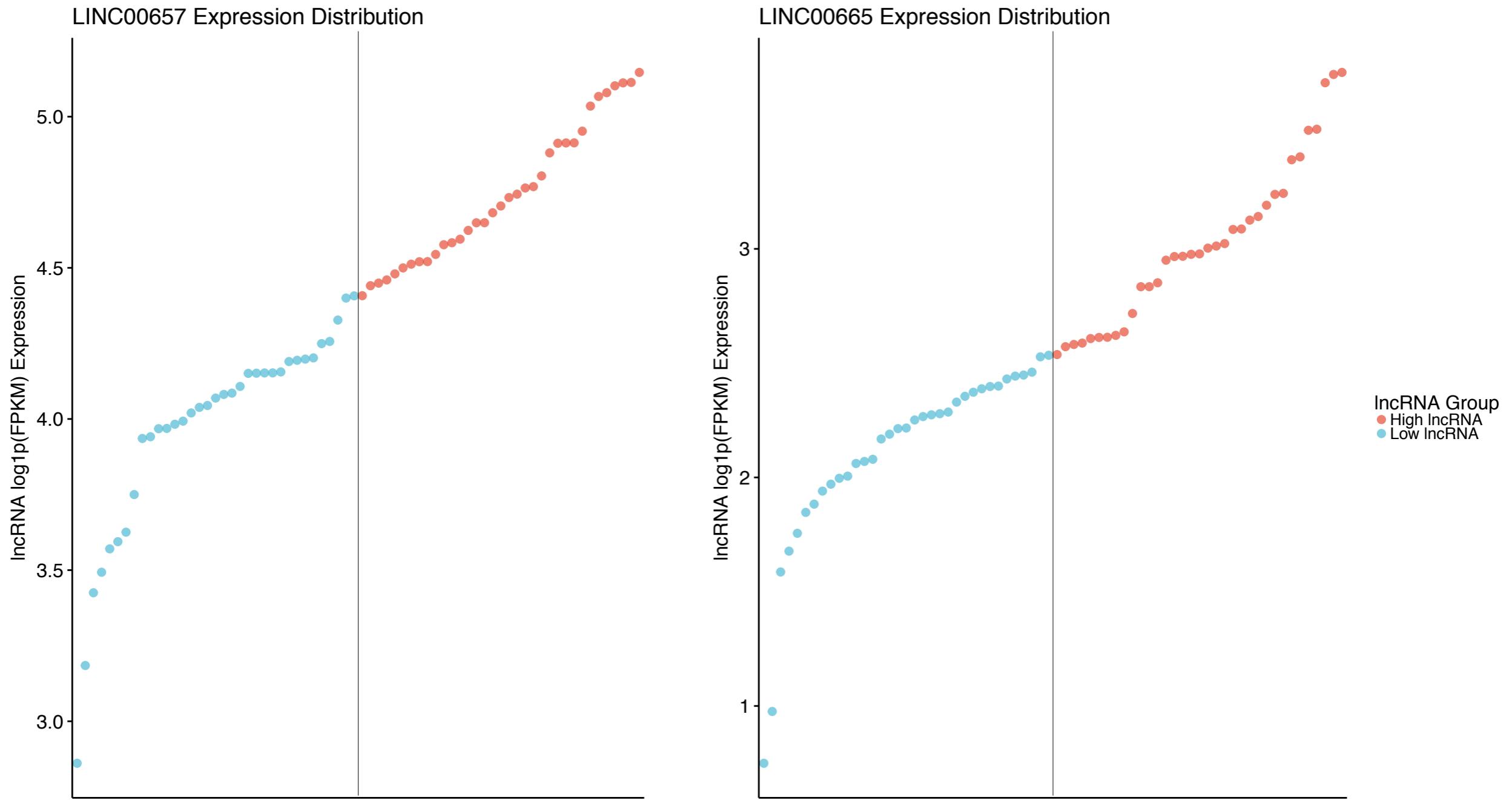


Need to validate
Need more samples
Is this association truly significant?

What does this mean? LINC00657 is expressed highly in all cancer types, even in normal ovarian samples... so what does this mean?
other processes going on? copy number variation high in patients that happen to have “higher expression”? isn’t higher expression suppose to be beneficial?

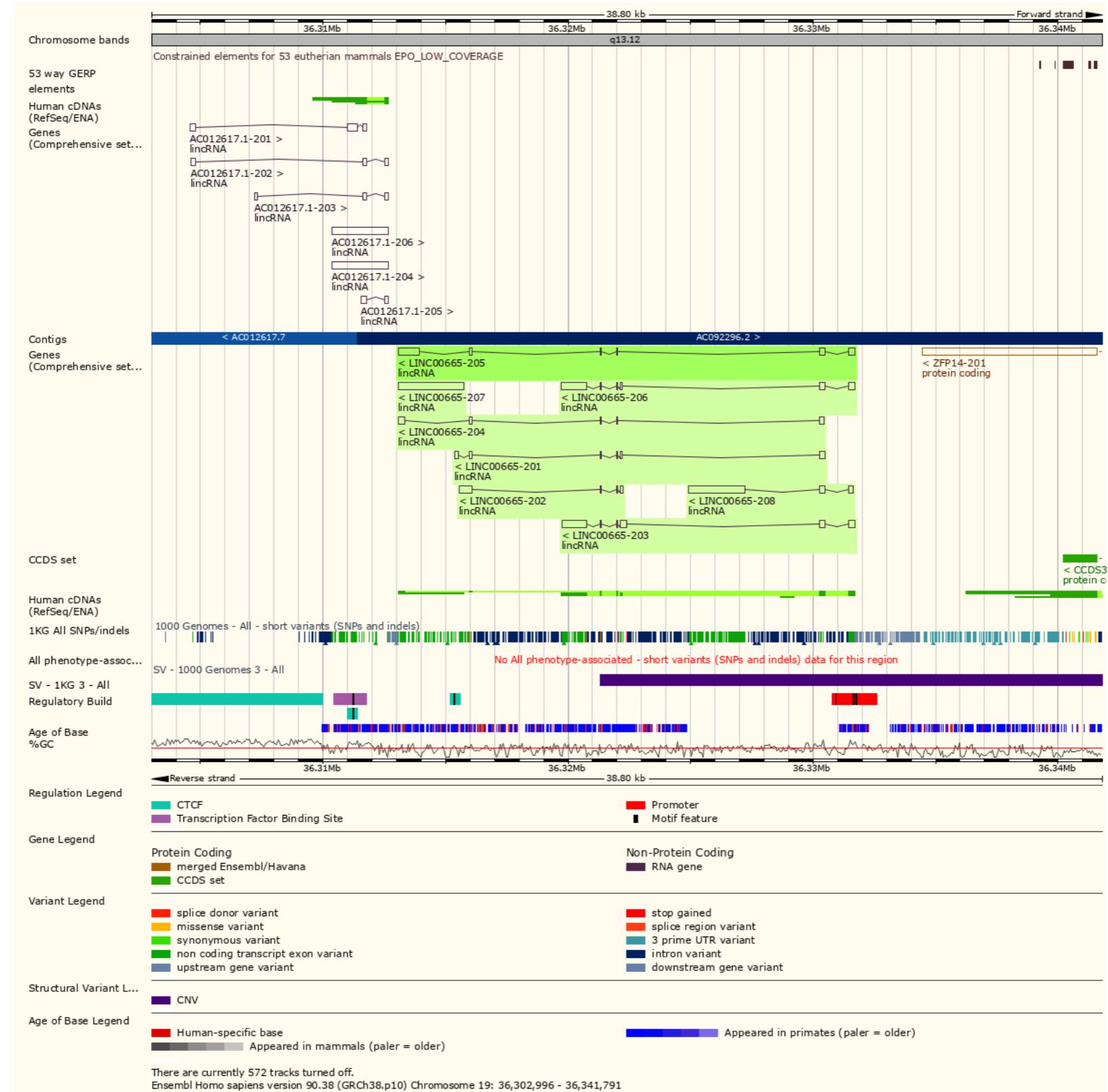


6. Further analysis of top 2 candidates in Ovarian Cancer



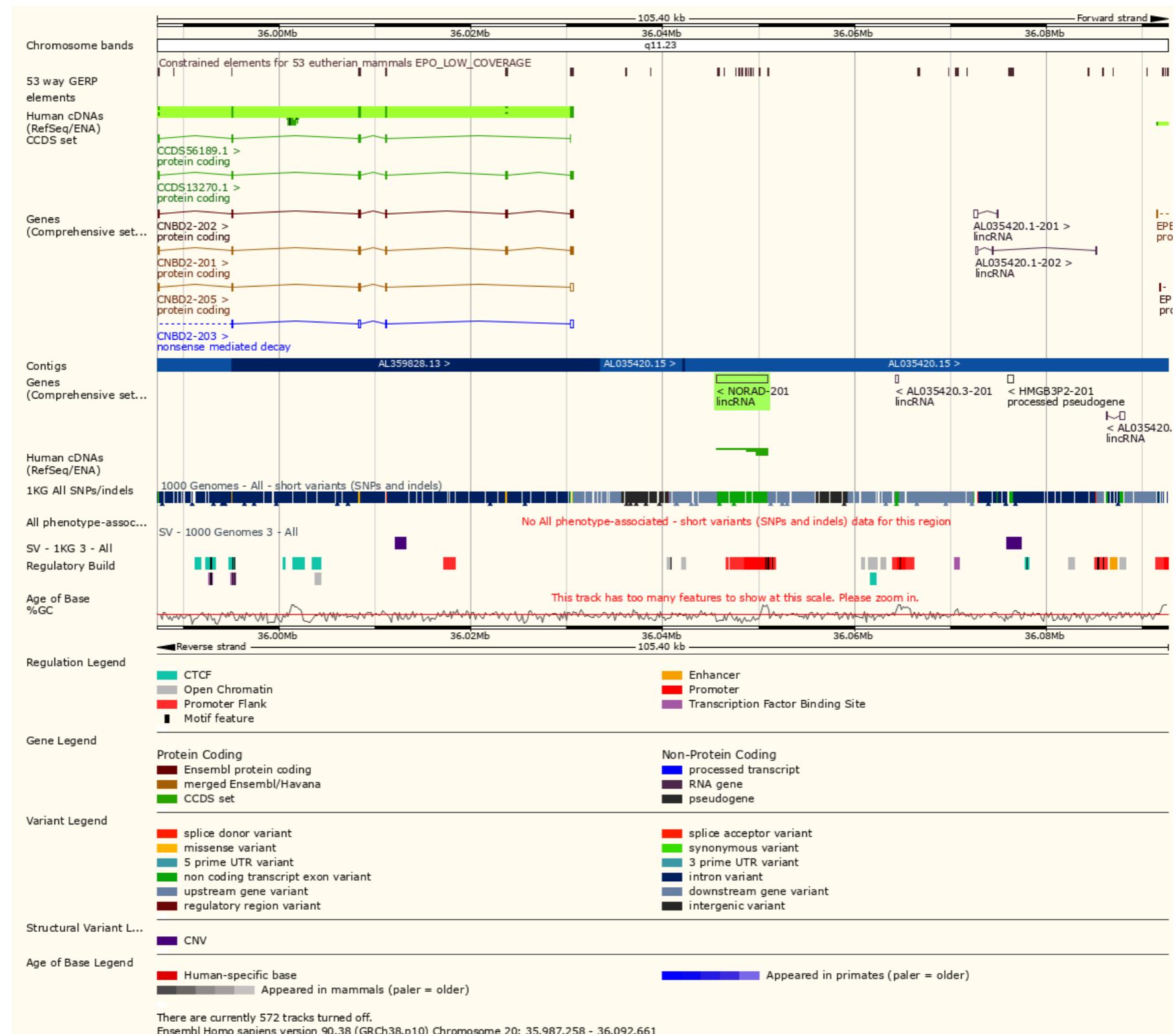
6. LINC00665

Chr 19



6. LINC00657

Chr 20



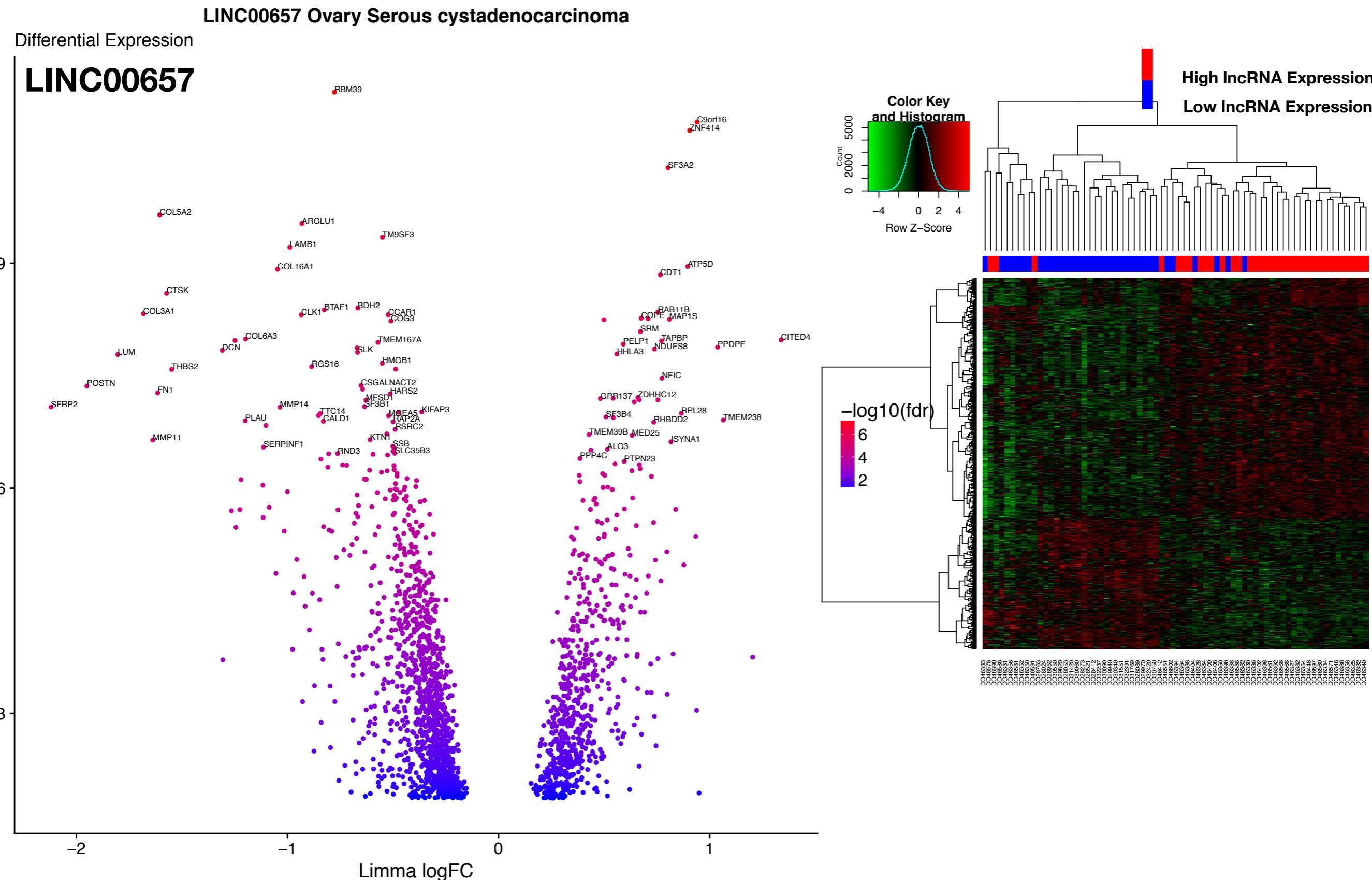
6. Further analysis of top 2 candidates in Ovarian Cancer

- Dichotomize patients based on median lncRNA expression
- Differential expression analysis of protein coding genes between high and low lncRNA expressing patients
 - 1,837 PCGs differentially expressed between high and low LINC00657 patients
 - 641 PCGs differentially expressed between high and low LINC00665 patients
- Visualize
- Pathway enrichment analysis to identify pathways differentially expressed between the two groups

6. Further analysis of top 2 candidates in Ovarian Cancer

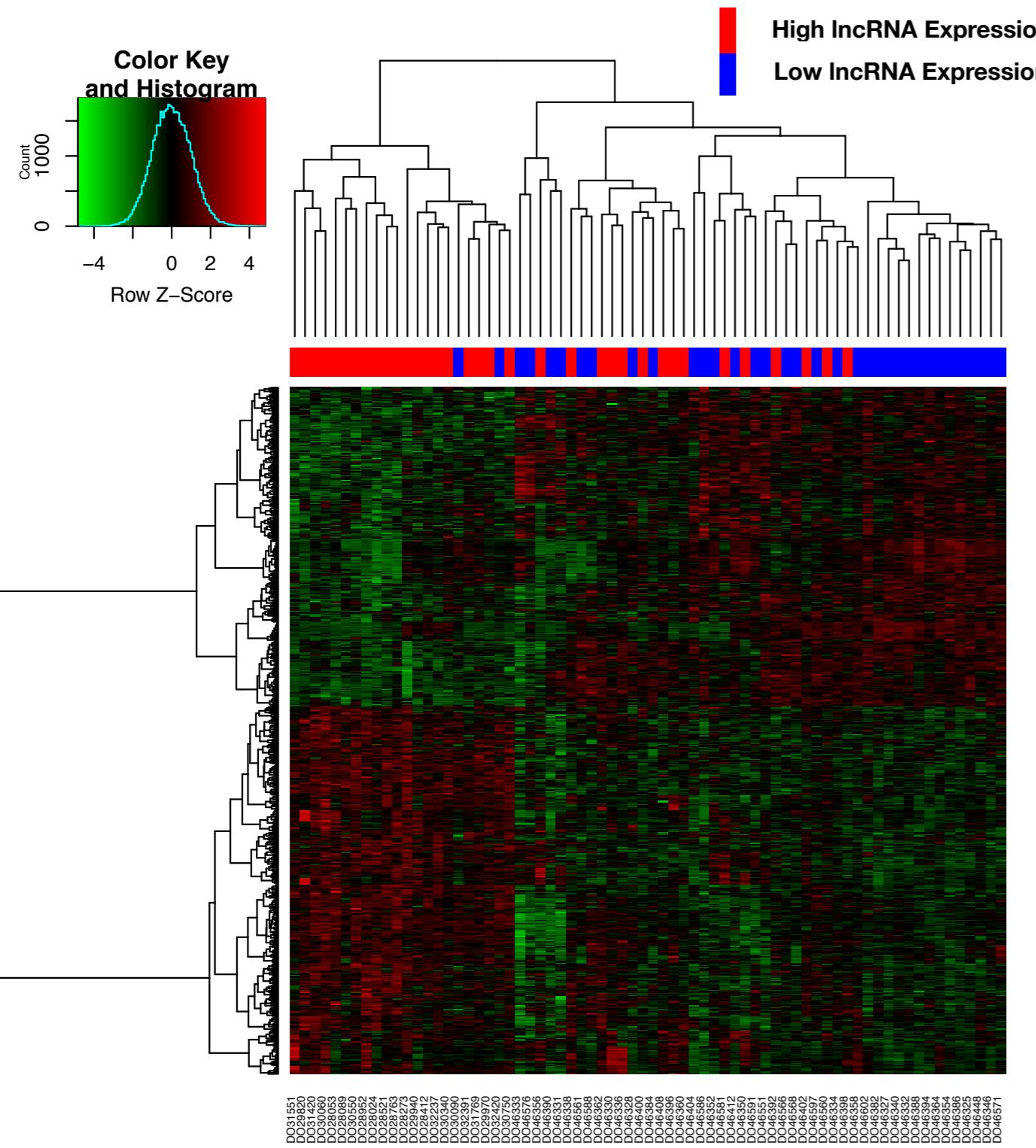
- Dichotomize patients based on median lncRNA expression
- Differential expression analysis of protein coding genes between high and low lncRNA expressing patients
 - 1,837 PCGs differentially expressed between high and low LINC00657 patients
 - 641 PCGs differentially expressed between high and low LINC00657 patients
- **Visualize**
- Pathway enrichment analysis to identify pathways differentially expressed between the two groups

6. Further analysis of top 2 candidates in Ovarian Cancer



6. Further analysis of top 2 candidates in Ovarian Cancer

LINC00665



6. Further analysis of top 2 candidates in Ovarian Cancer

- Dichotomize patients based on median lncRNA expression
- Differential expression analysis of protein coding genes between high and low lncRNA expressing patients
 - 1,837 PCGs differentially expressed between high and low LINC00657 patients
 - 641 PCGs differentially expressed between high and low LINC00657 patients
- Visualize
- **Pathway enrichment analysis to identify pathways differentially expressed between the two groups**

6. Further analysis of top 2 candidates in Ovarian Cancer

DE Analysis Workflow

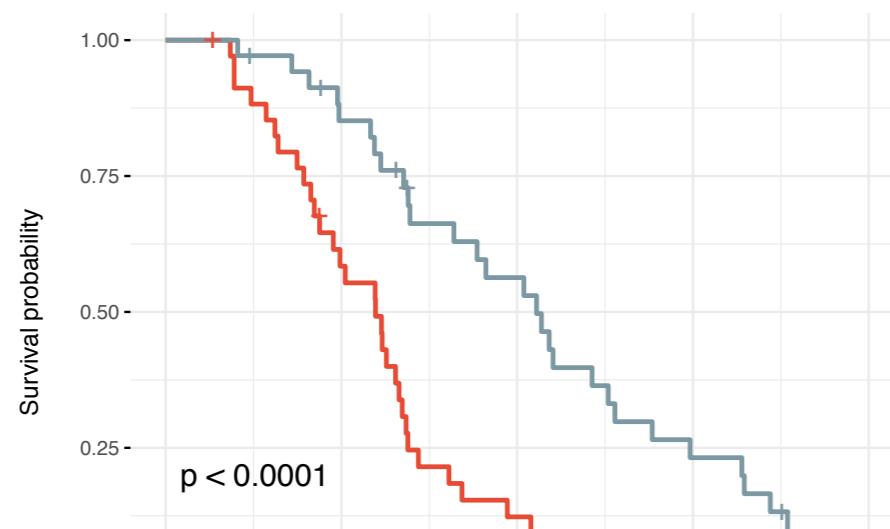
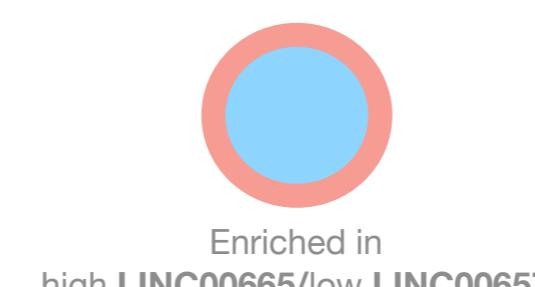
PCGs with Median FPKM ≥ 5

Compared expression of PCGs between high and low lncRNA expressing patients , n=35 each using LIMMA algorithm

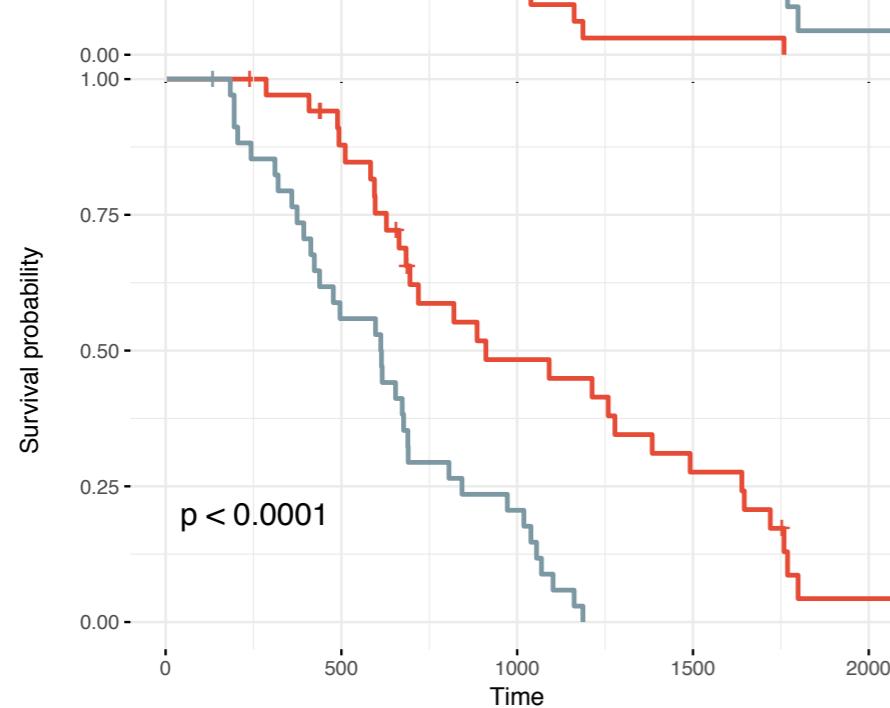
Using obtained fdr adjusted significantly DE PCGs, conducted pathway enrichment analysis using only GO terms and REACTOM terms using g:profiler tool developed in lab

Pathway Enrichment App in cytoscape to visualize pathways enriched in different groups of patients relative to expression of lncRNA 1 and lncRNA 2

6. Further analysis of top 2 candidates in Ovarian Cancer



Patients with worst predicted outcome

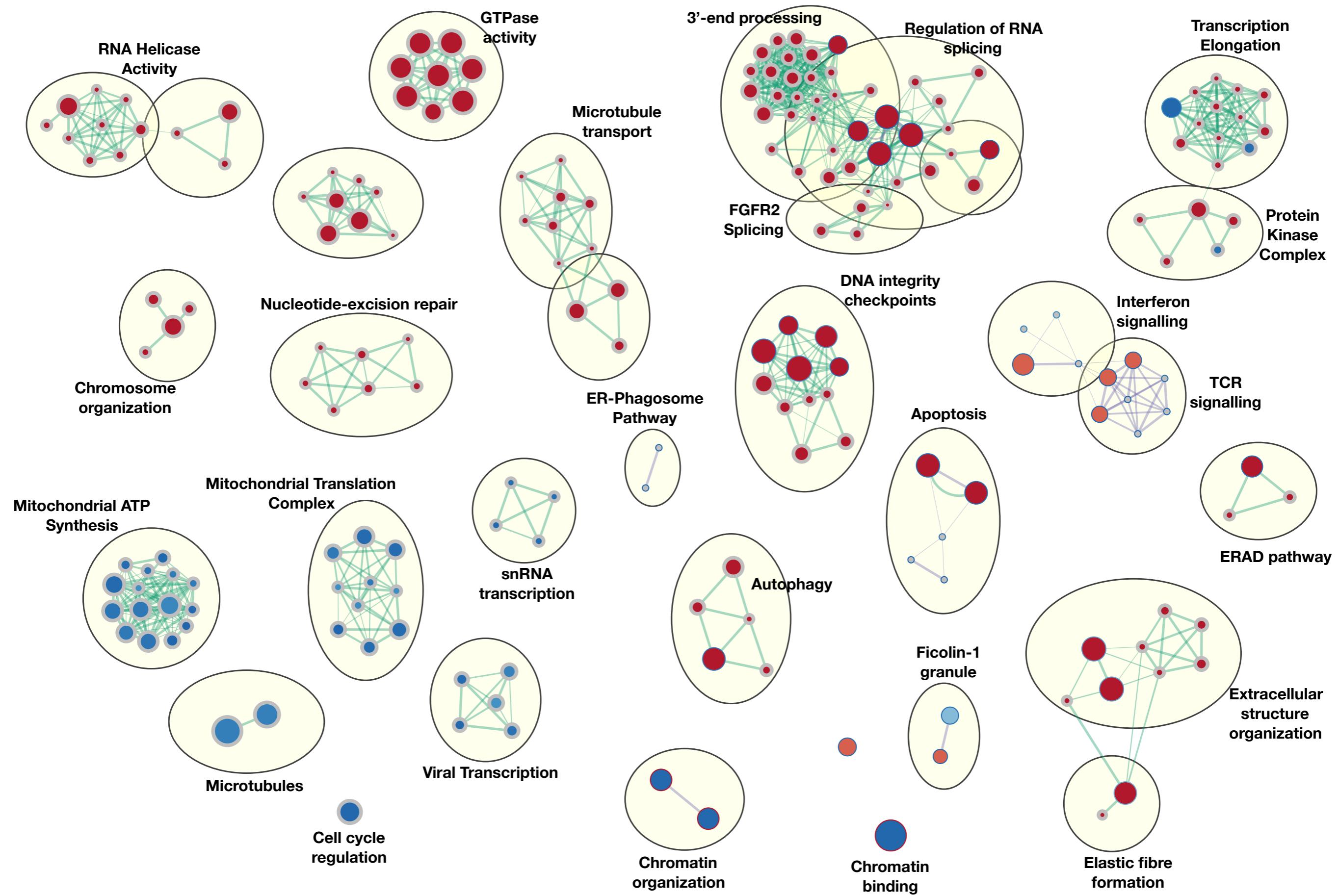


Patients with best predicted outcome

What are the pathways enriched by DE genes in these groups?

Legend for Pathway Enrichment Map

6. Further analysis of top 2 candidates in Ovarian Cancer



6. Further analysis of top 2 candidates in Ovarian Cancer

Summary of pathway enrichment analysis:

- 3' RNA processes enriched in HIGH Norad group —> miRNA interaction?
- Challenge because don't know direction of these processes
 - apoptosis for example, how are those genes affected ?
 - all we know is that they are more expressed in one group versus the other
- some cell cycle/mitsosis/DNA integrity processes high in NORAD/low LINC00665
 - again direction?

Summary of progress

1. Identified high expressing cancer specific lncRNAs through several filtration steps across multiple cancer types.
2. Confirmed cancer specific expression through comparing the distributions of each candidate lncRNA's expression among the cancer types studied
3. Identified cancer specific high expressing lncRNAs associated with patient survival outcome using the Cox Proportional Hazards model
 - 42 lncRNAs in 4 cancer types
 - 7 lncRNAs with adjusted p-value < 0.1 in 2 cancer types (Tier 1)
 - 35 lncRNAs with p-value < 0.05 in 4 cancer types (Tier 2)
4. Assessed tumour versus normal tissue lncRNA expression by comparing ranked values of expression using tissue data from GTEx
5. 21/42 lncRNAs were predicted to be either tumour suppressive or oncogenic based on sign of hazard ratio and difference in expression between normal and tumour tissues
6. 2 candidates in Ovarian cancer were further studied to predict regulatory roles that may lead patients from the same cancer type to develop worse prognosis

Outline

Thesis: Genomic characterization of clinically relevant lncRNAs in multiple cancer types

1. Review of lncRNAs
2. Summary of the last SCM
3. Summary of progress since last meeting
- 4. Future directions**

Future Directions

1. Validation of survival associated lncRNAs in additional datasets
 - TCGA datasets
 - Thorough analysis of specific cancer cohorts available including Japanese liver cancer cohort and Toronto Ovarian cancer cohort
 - Compare performance on lncRNA against other predictors (clinical features, combinations of lncRNA candidates)
 - Improve biomarker discovery pipeline through additional data integration and analysis
2. In depth analysis of high confidence candidate lncRNAs to predict cellular mechanisms that could be experimentally validated
 - Predict interactions through co-expression analysis with improved predictors and assess permutations
 - Predictors will include copy number aberrations, promoter methylation, presence of TAD boundaries, single nucleotide variants in lncRNA promoter or predicted enhancer regions and integration of cancer specific properties such as driver mutation status
 - Integrate available protein level data to evaluate protein-protein co-expression as well as lncRNA-co-expression to build a lncRNA-protein regulatory network
3. Apply analysis to normal tissues available in GTEx to predict differential co-expression of key pathways between normal and tumour tissues