The Cancer Genome Atlas

LUAD miRNA-seq
352 tumors
45 tissue normals
MIMATs

Gordon Robertson
Andy Chu
Ken Lau

14 August 2012

GENOME Sciences Centre

BC Cancer Agency
CARE & RESEARCH
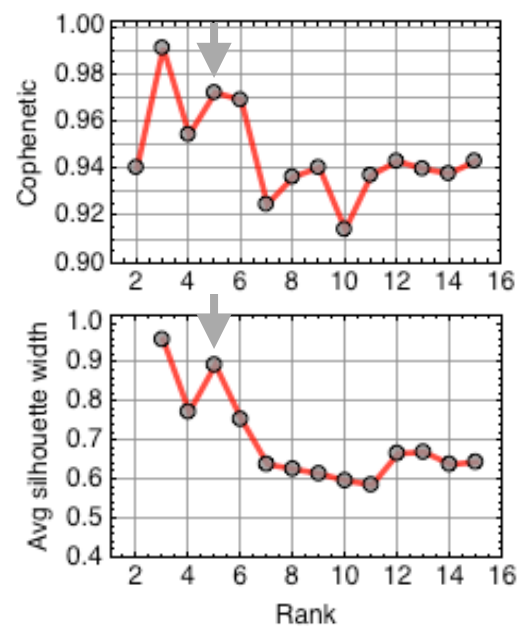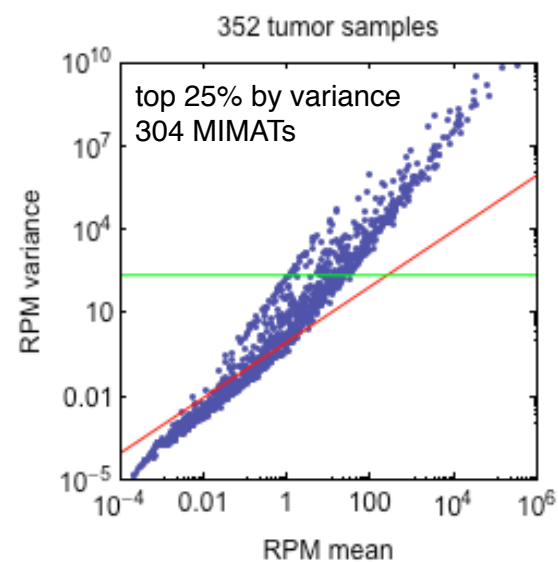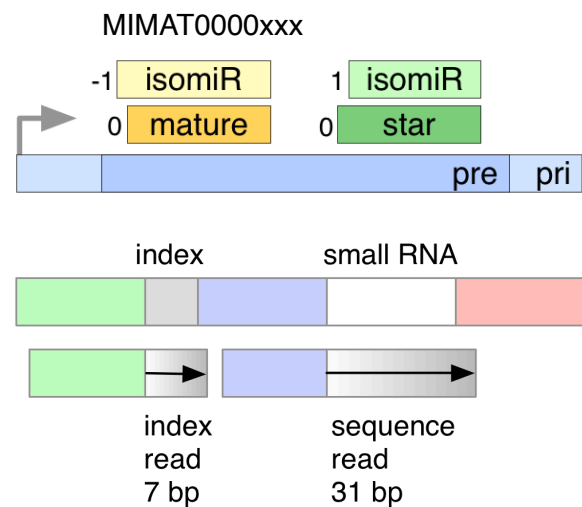
1. NMF unsupervised clustering, clinical covariates, survival
2. Technical covariates: TSS, BCR batch, sequencing platform
3. Discriminatory MIMATs, literature
4. Supervised: MIMATs and overall survival
5. LUAD (452 T) vs. LUSC (158 T)
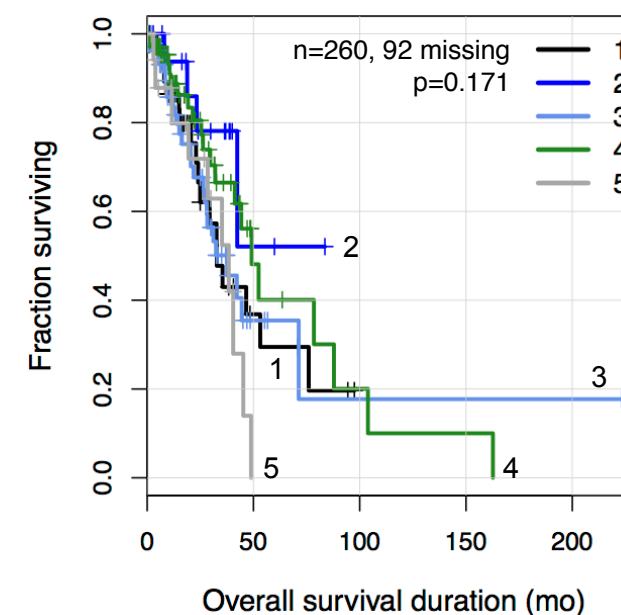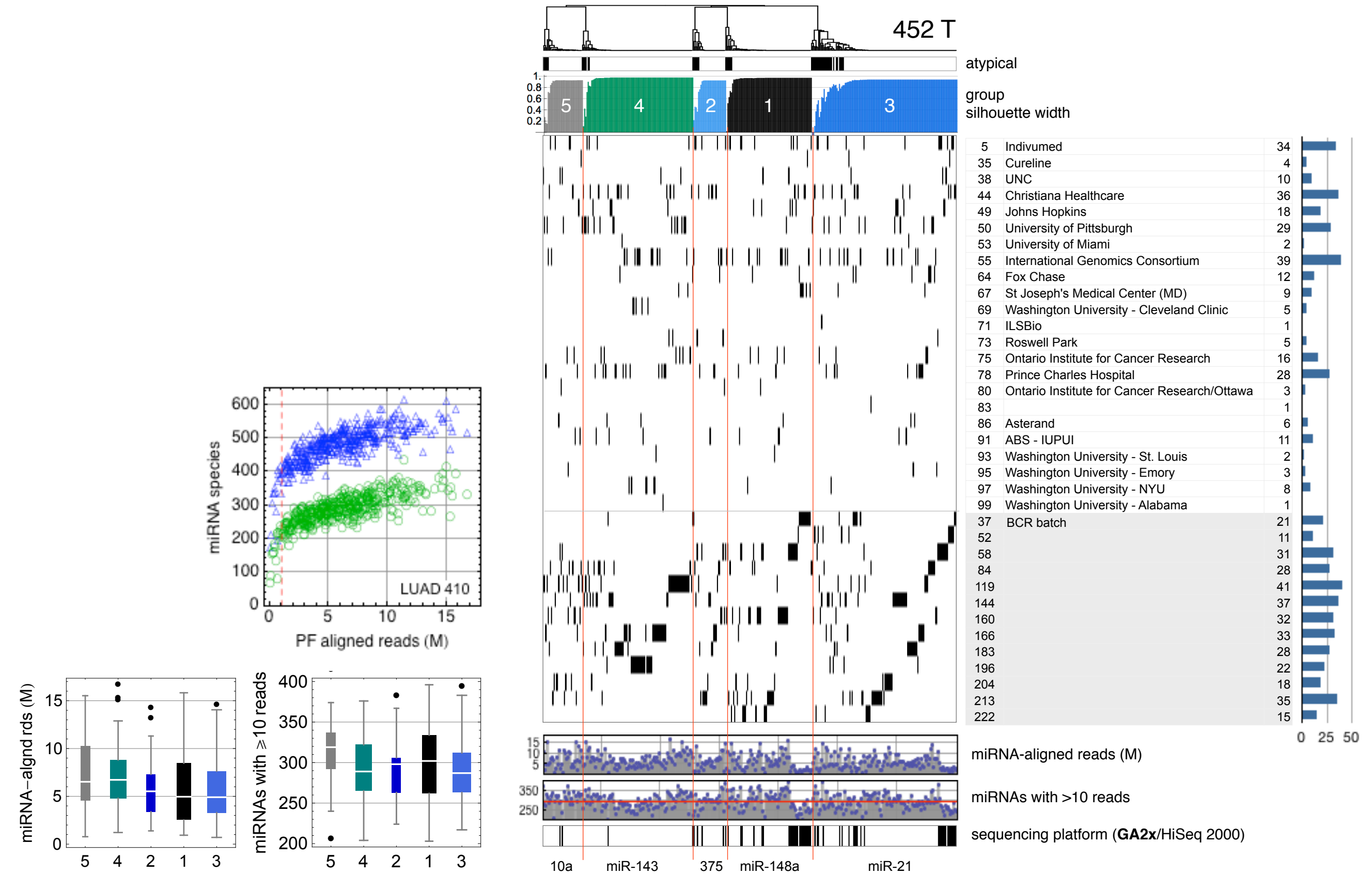6. NMF clustering of 45 tumor-normal pairs
7. Summary

MIMAT0000xxx

index read 7 bp

sequence read 31 bp

352 tumor samples

top 25% by variance
304 MIMATs

RPM variance

RPM mean

Cophenetic

Avg silhouette width

Rank

5

10a

miR-143

375

miR-148a

miR-21

atypical

group
silhouette width

5 4 2 1 3

c	5	4	2	1	3	all
n	33	94	28	73	124	352
w	0.84	0.94	0.83	0.94	0.86	0.89

n=260, 92 missing
p=0.171

Fraction surviving

Overall survival duration (mo)
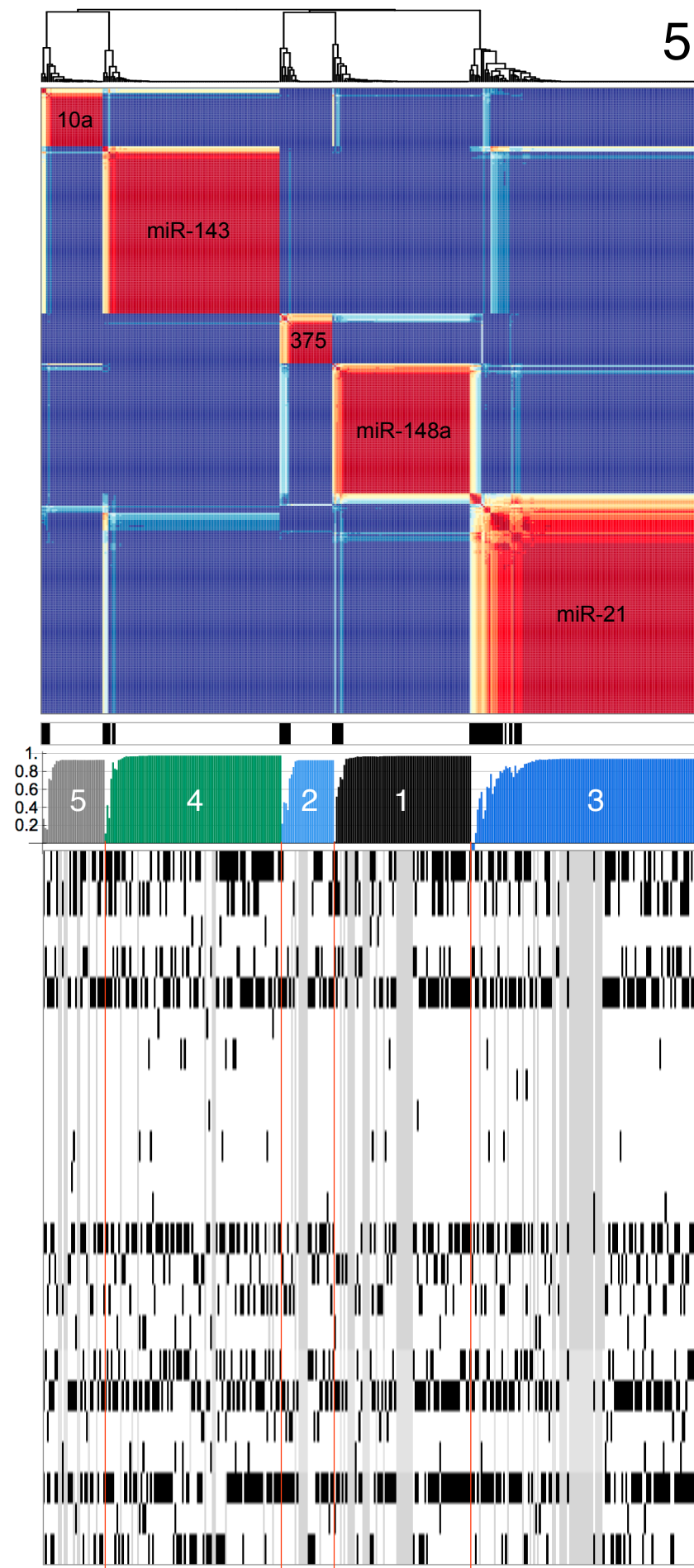
gender (F)	
vital status (DEC)	
acinar adeno	histology
adeno mixed subtype	
adeno NOS	
bronchioloalveolar car mucinous	
bronchioloalveolar car nonmucinous	
clear cell adeno	
micropapillary adeno	
papillary adeno	
solid pattern predominant adeno	
mucinous (colloid) adeno	
I, IA, IB	stage
IIIA, B	
IIIA, B	
IV	
T1, 1a, 1b	primary tumor spread
T2, 2a, 2b	
T3	
T4	
M0	distant metastasis spread
M1	
MX	

LUAD Public Clinical Data Jul-23-2012

TSS, BCR batch, mapped reads, sequencing platform

452 T

atypical

group
silhouette width

5	Indivumed	34	
35	Cureline	4	
38	UNC	10	
44	Christiana Healthcare	36	
49	Johns Hopkins	18	
50	University of Pittsburgh	29	
53	University of Miami	2	
55	International Genomics Consortium	39	
64	Fox Chase	12	
67	St Joseph's Medical Center (MD)	9	
69	Washington University - Cleveland Clinic	5	
71	ILSBio	1	
73	Roswell Park	5	
75	Ontario Institute for Cancer Research	16	
78	Prince Charles Hospital	28	
80	Ontario Institute for Cancer Research/Ottawa	3	
83		1	
86	Asterand	6	
91	ABS - IUPUI	11	
93	Washington University - St. Louis	2	
95	Washington University - Emory	3	
97	Washington University - NYU	8	
99	Washington University - Alabama	1	
37	BCR batch	21	
52		11	
58		31	
84		28	
119		41	
144		37	
160		32	
166		33	
183		28	
196		22	
204		18	
213		35	
222		15	

0   25   50

miRNA-aligned reads (M)

miRNAs with >10 reads

sequencing platform (**GA2x**/HiSeq 2000)

10a   miR-143   375   miR-148a   miR-21

Discriminatory MIMATs

c	5	4	2	1	3	all
n	33	94	28	73	124	352
w	0.84	0.94	0.83	0.94	0.86	0.89

MIMAT0000xxx

-1 isomiR	1 isomiR
0 mature	0 star

pre pri

atypical

group
silhouette width

n=260, 92 missing
p=0.171

Fraction surviving

Overall survival duration (mo)

miR-10a.MIMAT0000253

miR-143.MIMAT0000435

miR-375.MIMAT0000728

miR-148a.MIMAT0000243

miR-21.MIMAT0000076

miR-183.MIMAT0000261

miR-210.MIMAT0000267

let-7f.MIMAT0000067

miR-29b.MIMAT0000100
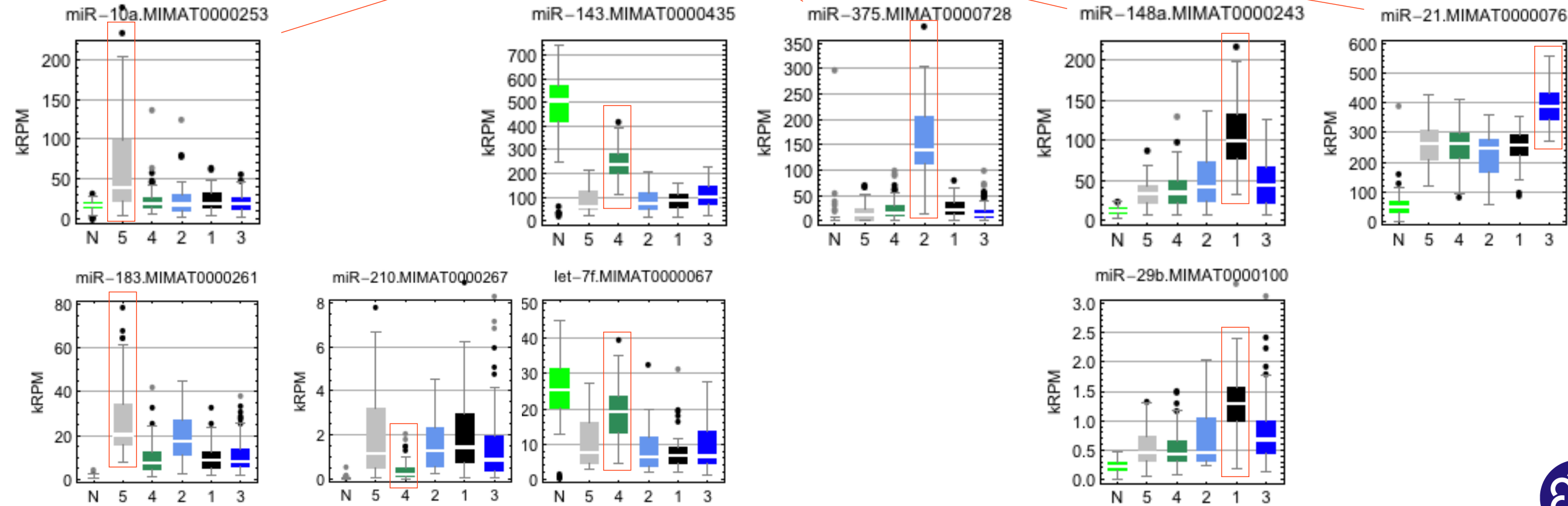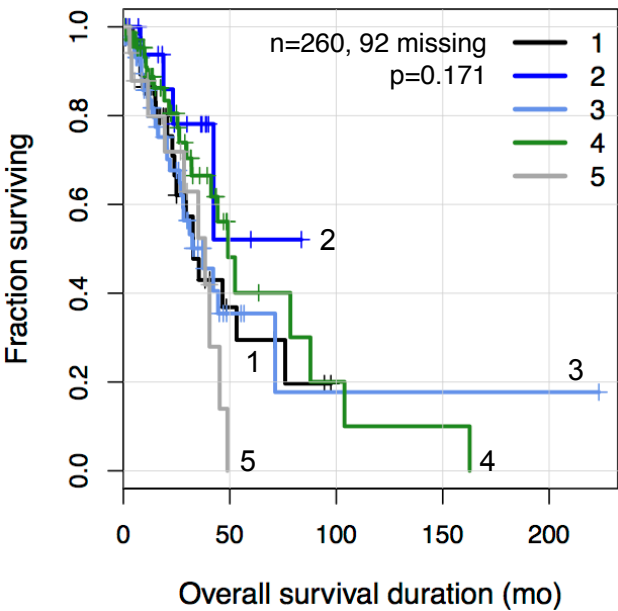
Literature for discriminatory MIMATs
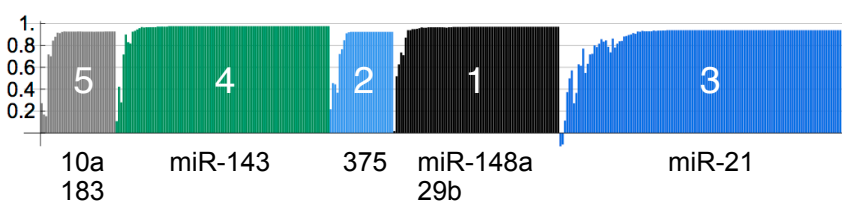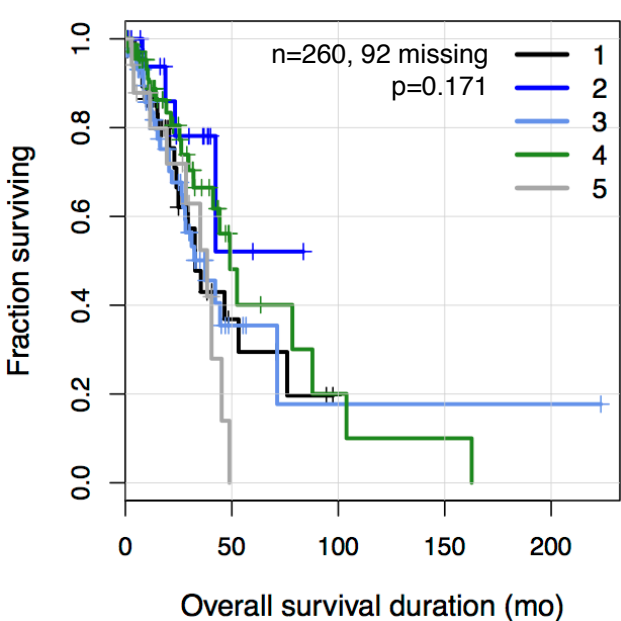
grp	miR	vs. normal	surv	PubMed
5	10a	↑	↓	• Upregulated in human cervical cancer, promotes cell growth, migration and invasion in human **cervical** cancer cells (PMID:22634495). • Overexpressed in human **pancreatic** cancer, involved in Its invasiveness partially via suppression of HOXA1 gene (PMID:22407312). • Overexpressed in Nucleophosmin1 mutated **acute myeloid leukaemia**, its suppression induces cell death (PMID:22348345).
	183	↑	↓	• High expression of tumor and serum miRNAs of the miR-183 family were associated with overall poor survival in patients with **lung** cancer (PMID:21920043). • Down-regulation of miR-183 promotes migration and invasion of **osteosarcoma** by targeting Ezrin (PMID:22525461). • Differentially expressed between normal lung and primary **small cell lung** cancer (SCLC) tumors (PMID:19895320). • Expression level of miR-183 was reversely correlated with the metastatic potential of **lung** cancer cells (PMID:18840437).
4	143	↓		• A complex network of regulation through which the miR-143/145 cluster is able to modulate KRAS signaling in **colorectal** cancer (PMID:22751122). • Hypothesize that loss of miR-143-mediated repression of HK2 can promote glucose metabolism in cancer cells, contributing to the shift towards **aerobic glycolysis** observed in many tumors (PMID:22691140). • MicroRNA-143 targets MACC1 to inhibit cell invasion and migration in **colorectal** cancer (PMID:22533346).
2	375	↑	↑	• Review: miR-375 is frequently downregulated and functions as a **tumor suppressor** that targets several oncogenic genes in cancer cells (PMID:22718022). • Inhibits autophagy by reducing expression of ATG7 and impairs viability of HCC cells under hypoxic conditions in culture and in mice (PMID:22504094). • **Low**-level expression of miR-375 correlates with **poor** outcome and metastasis while altering the invasive properties of **HNSCCs** (PMID:22234174). • **Highly** expressed and possibly transactivated by achaete-scute complex homolog 1 in **small-cell lung** cancer cells (PMID:22172490).
1	148a	↑		• Involved in DNMT3b-mediated hypermethylation defect in **breast** cancer cell lines that involves the loss of post-transcriptional regulation of DNMT3b by regulatory miRs (PMID:22664488). • Promotes apoptosis by silencing Bcl-2 in **colorectal** cancer (PMID:21455217). • Frequently down-regulated in **gastric** cancer and acts as a tumor suppressor by inhibiting cell proliferation (PMID:21205300). • Sensitized chemotherapy-sensitive **oesophageal** cancer cell lines to cisplatin (PMID:21246413).
	29b	↑		• Involved in DNMT3b-mediated hypermethylation defect in breast cancer cell lines that involves the loss of post-transcriptional regulation of DNMT3b by regulatory miRs (PMID:22664488). • Down-regulation of C1QTNF6, SPARC, and COL4A2, targeted by miR-29b, can contribute to the invasion ability of MCF-7 cells (PMID:22864815). • miRNAs of the miR-29 family (miR-29a, miR-29b and miR-29c) regulate specific genes associated with tissue invasion and metastasis in **lung adenocarcinoma** (PMID:22745231). • Suppresses prostate cancer metastasis by regulating epithelial-mesenchymal transition signaling (PMID:22402125).
3	21	↑		• Induces cell cycle at S phase and modulates cell proliferation by down-regulating hMSH2 in **lung** cancer (PMID:22806311). • Meta-analysis of 14 studies: consistently up-regulated in squamous carcinoma and **adenocarcinoma**-based subgroup analysis (PMID:22672859). • NOS2 enhances KRAS-induced **lung** carcinogenesis, inflammation and microRNA-21 expression (PMID:22618808). • May play an important role in the development and progression of **lung** cancer through JAK/STAT signal pathway, MAPK signaling pathway, Wnt signaling pathway, cell cycle, PPAR signaling pathway, apoptosis pathway and other pathways (PMID:22244963).

MIMATs and overall survival

Univariate Cox regression
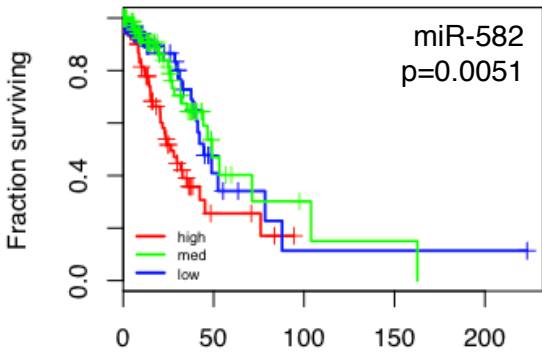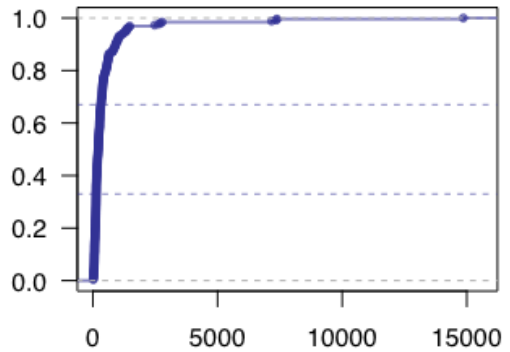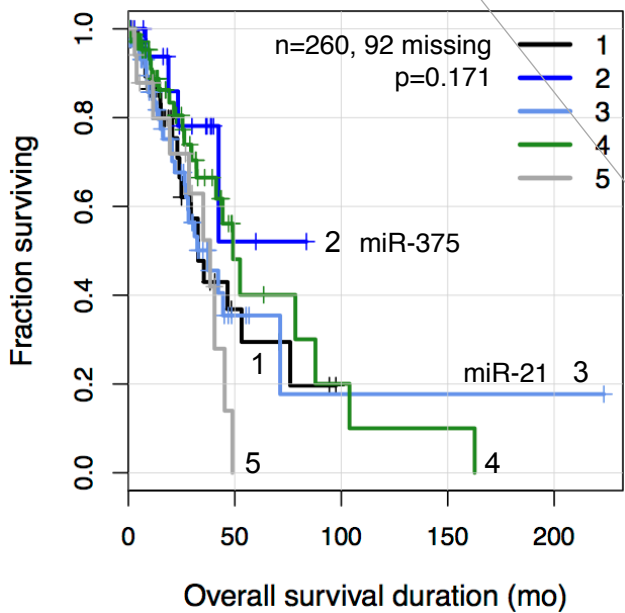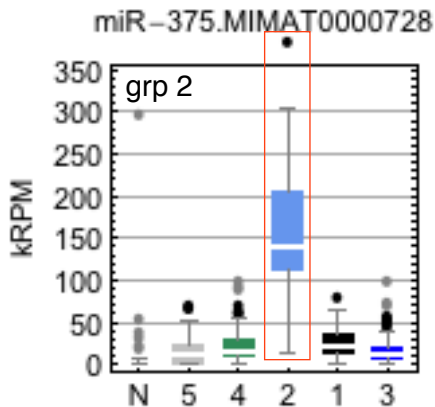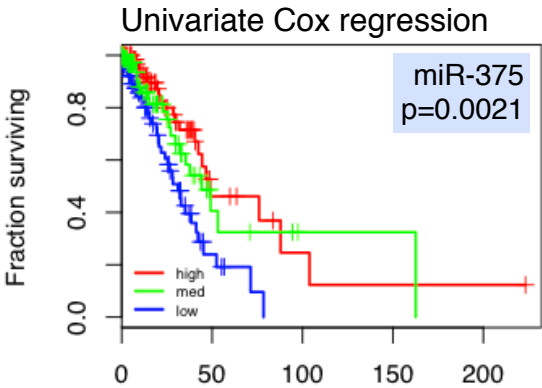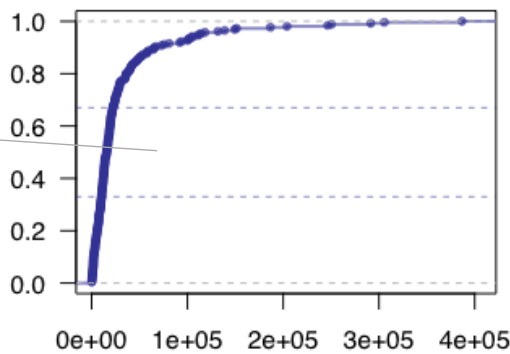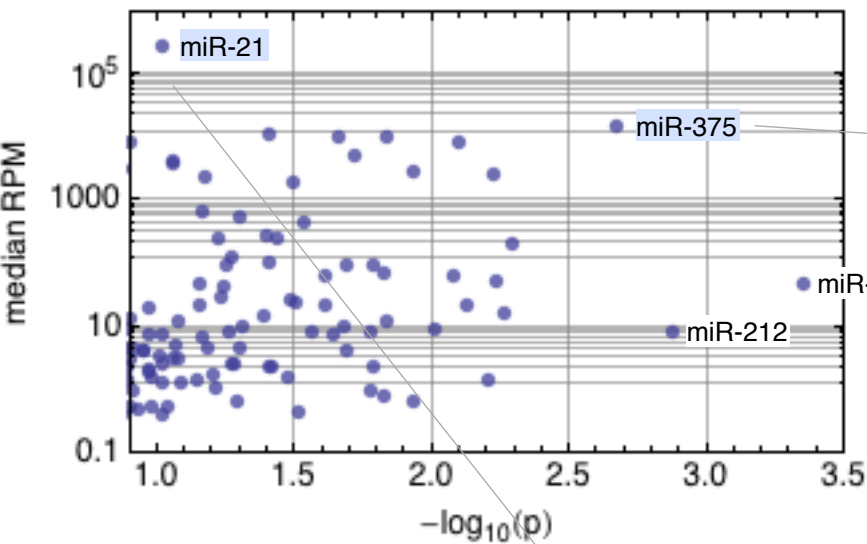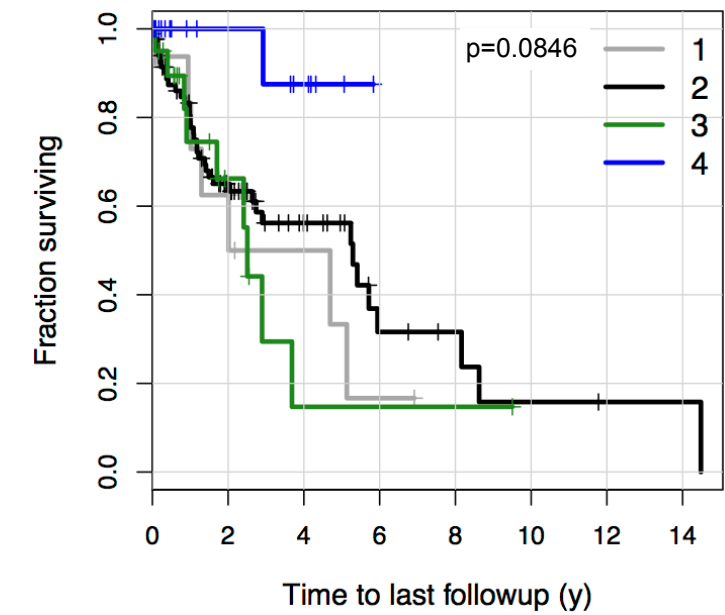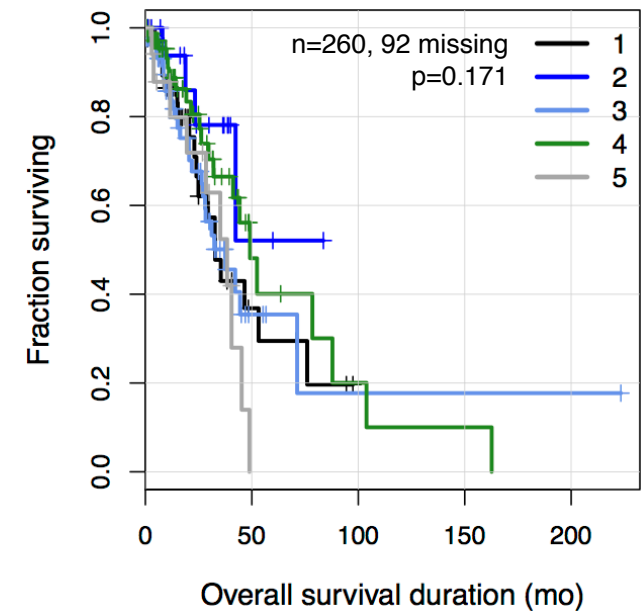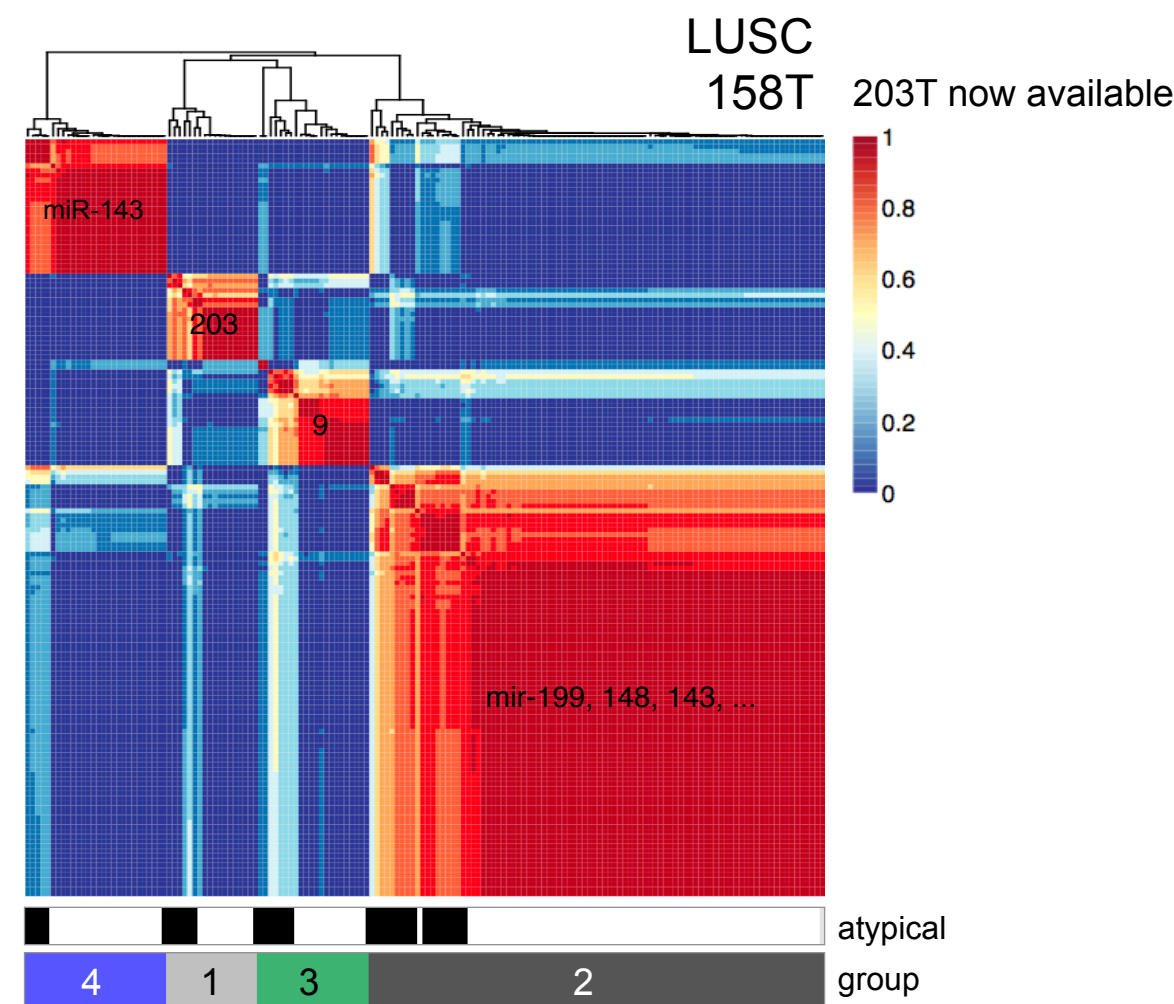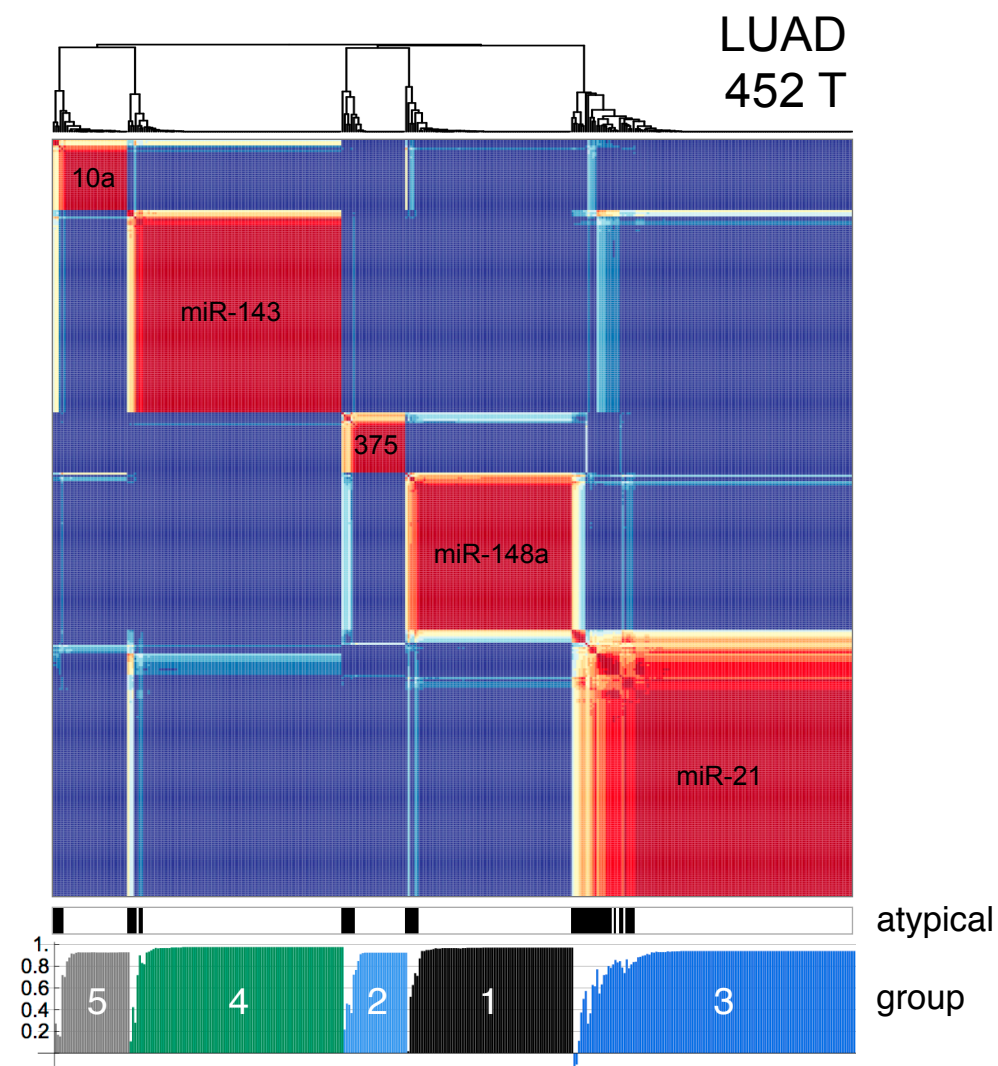
Discriminatory MIMATs

grp	miR	vs. nrml	surv	p-val
5	10a	↑		0.52
	183	↑		0.89
4	143	↓		0.56
2	375	↑	↑	0.0021
1	148a	↑		0.36
	29b	↑		0.17
3	21	↑	↓	0.095
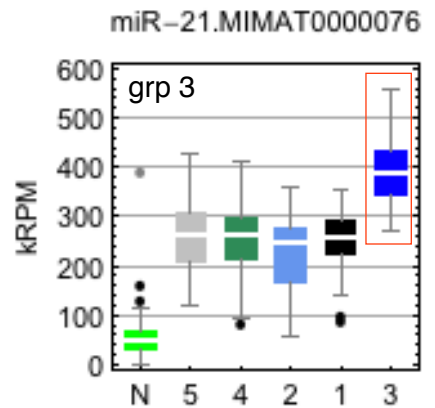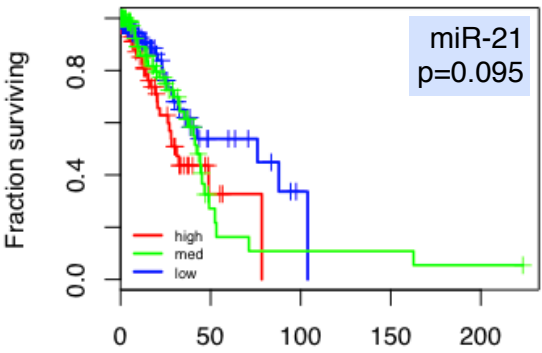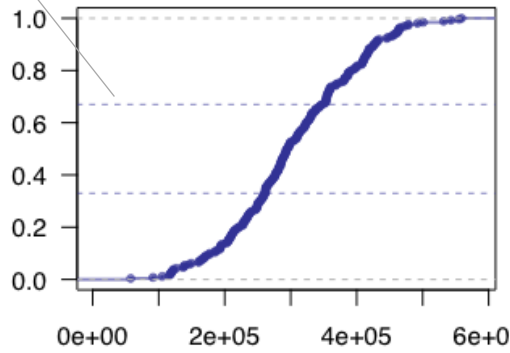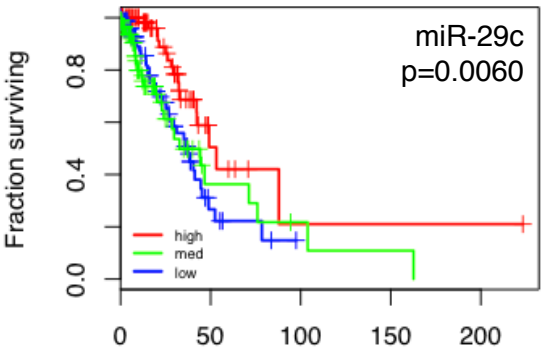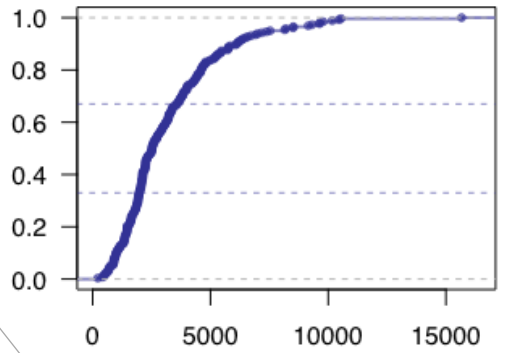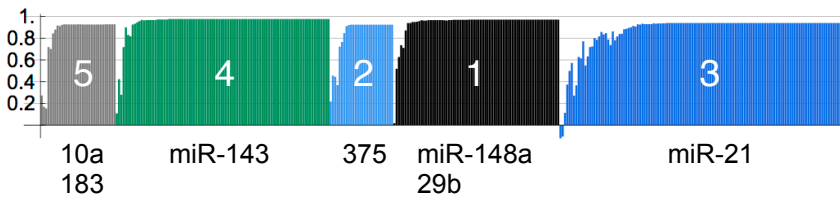
Ken Lau

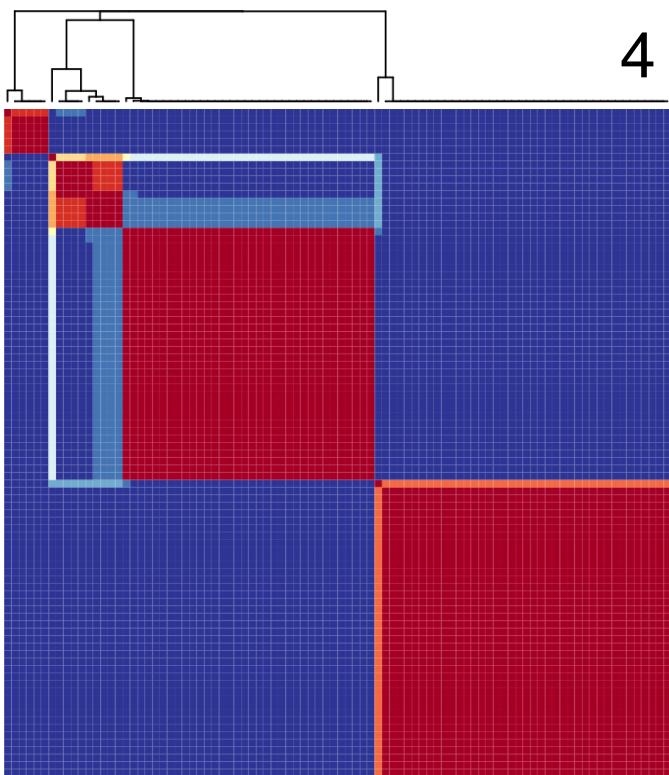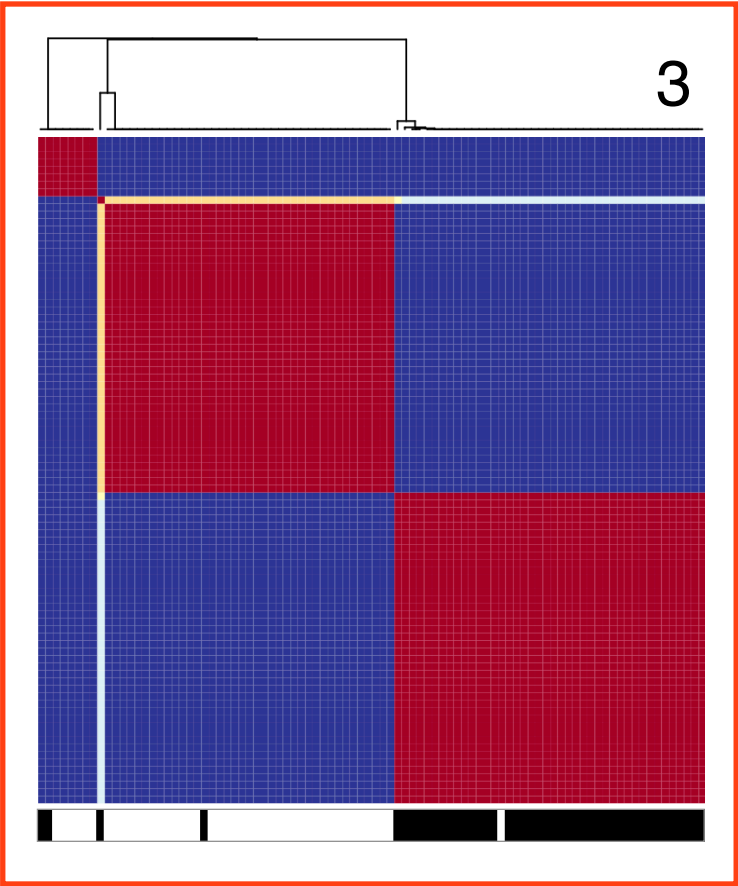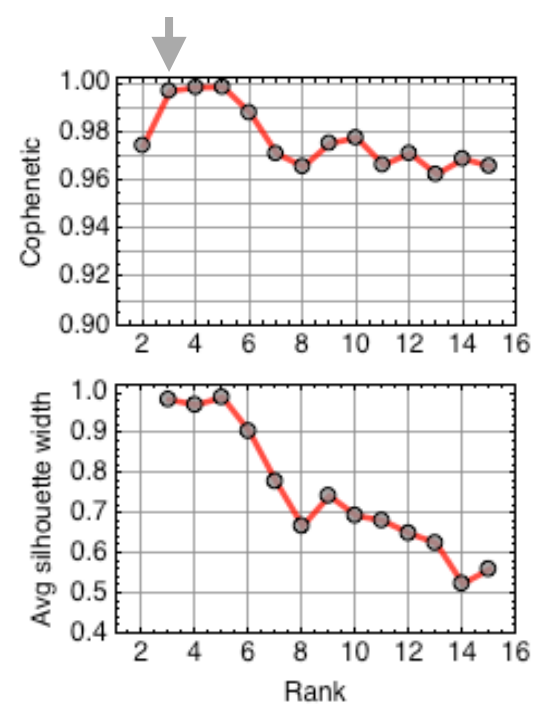# Compare 452 LUAD and 158 LUSC tumor samples

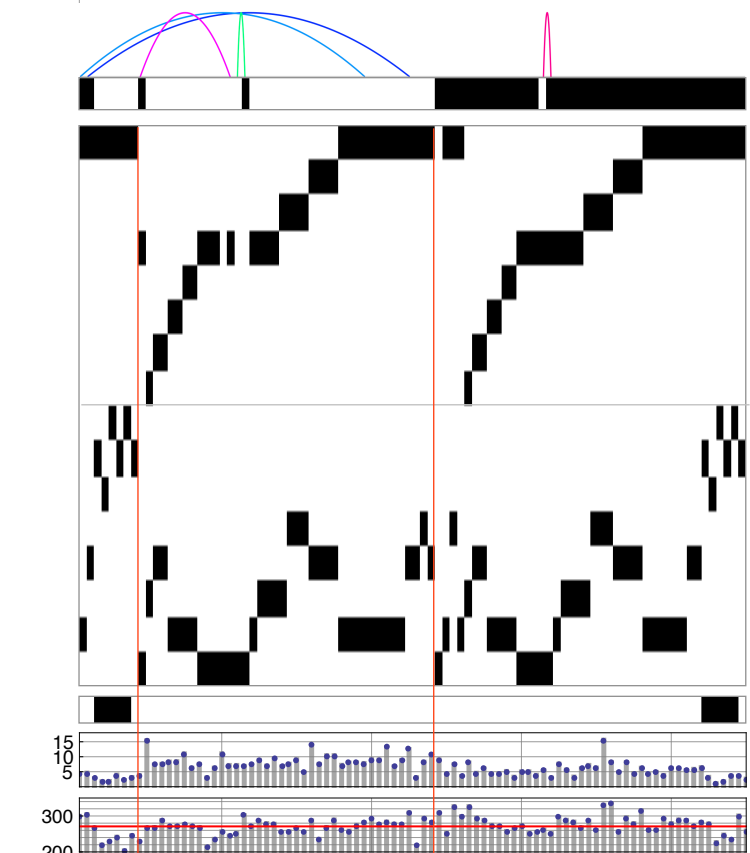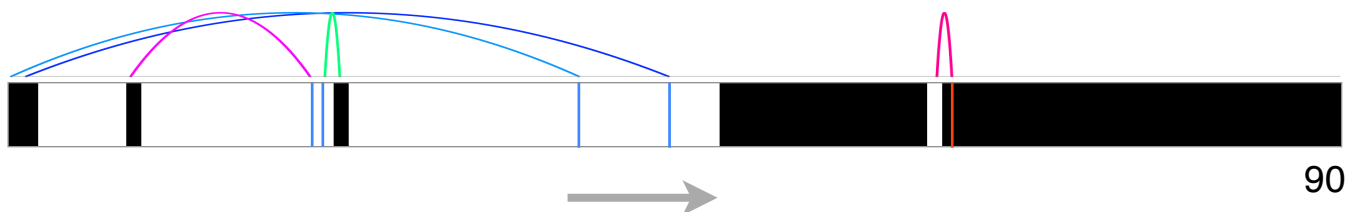# 45 tumor-tissue normal pairs

3
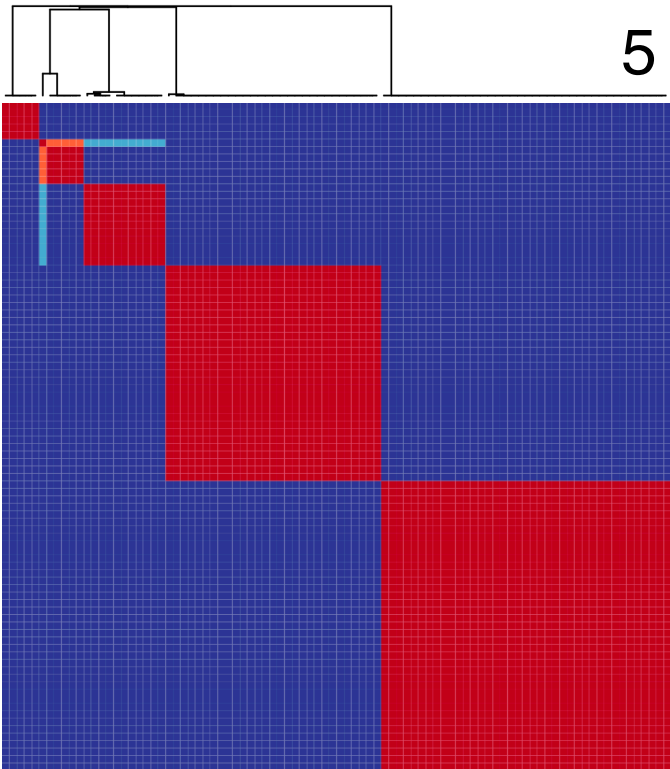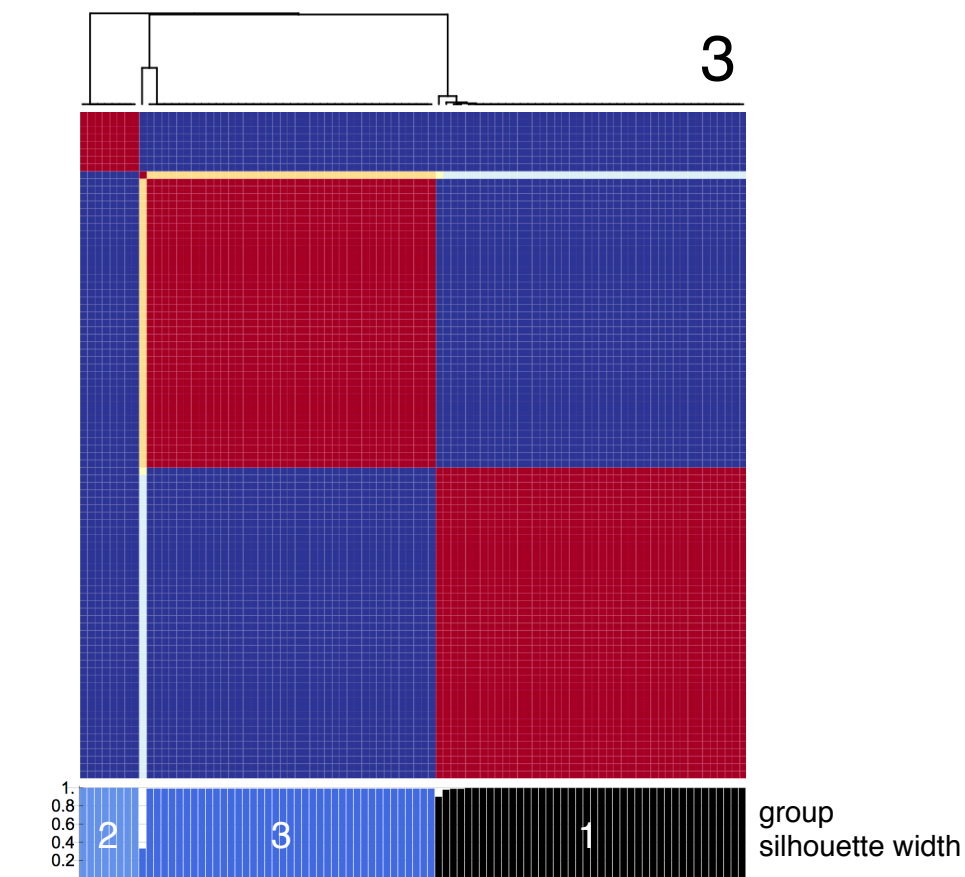
i	grp			
1	2	TCGA-44-7669-01A-21H	T	
39	3	TCGA-44-7669-11A-01H	N	
2	2	TCGA-44-6778-01A-11H	T	
45	3	TCGA-44-6778-11A-01H	N	atypical in grp 3
9	3	TCGA-55-7725-01A-11H	T	
21	3	TCGA-55-7725-11A-01H	N	
22	3	TCGA-55-7724-11A-01H	N	adjacent in grp 3
23	3	TCGA-55-7724-01A-11H	T	
63	1	TCGA-55-7903-11A-01H	N	adjacent in grp 1
64	1	TCGA-55-7903-01A-11H	T	

90

group
silhouette width

2   3   1

tumor/normal

44	Christiana Healthcare	38
49	Johns Hopkins	8
50	University of Pittsburgh	8
55	International Genomics Consortium	18
78	Prince Charles Hospital	4
86	Asterand	4
91	ABS - IUPUI	4
93	Washington University - St. Louis	2
37	BCR batch	4
52		6
58		2
119		8
144		18
166		10
183		28
196		14

GA2x/HiSeq 2000 l platform

miRNA-aligned reads (M)

miRNAs with >10 reads

miRNA-algnd rds (M)

miRNAs with ≥10 reads

c	2	3	1	all
n	8	40	42	90
w	1.0	0.97	0.99	0.99
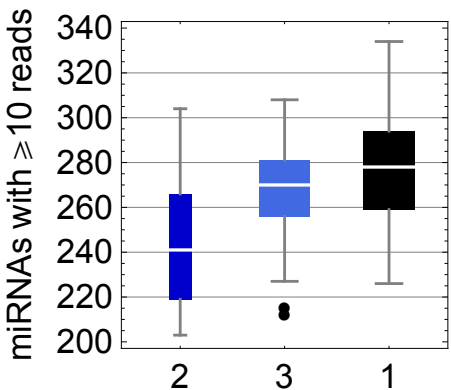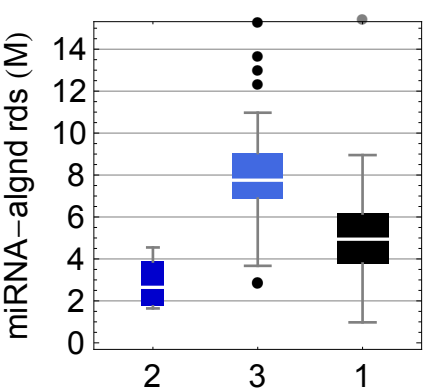
Summary

1. Unsupervised consensus clustering for MIMATs for 452 tumor data sets suggests five sample groups.

2. While survival appears poor for all groups, 92/452 records are currently incomplete.

3. MiRNAs that best discriminate unsupervised groups are: miR-10a, 143, 375, 148a and 21. As a first pass, only miR-375 appears to have an interesting p-value against overall survival.

4. For 45 tumor-tissue normal pairs, most pairs were well-separated. Four tumor samples were grouped with the normal samples, and one normal sample with the tumor samples.

LUAD 452T: discriminatory MIMATs

50k trees, mtry=100

c1 mirR-148a

Top 15 variables
OOB est err: 7.1%
```
         0   1  class.err
0  276   3  0.0108
1   22  51  0.301
```
15
0.071

resp	Freq
0	279
1	73
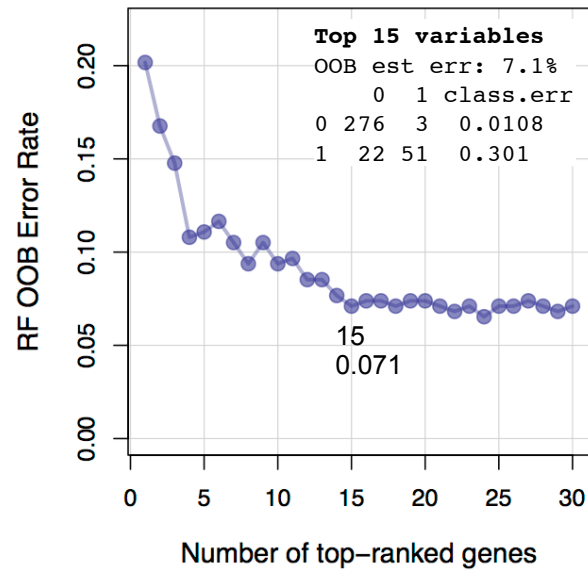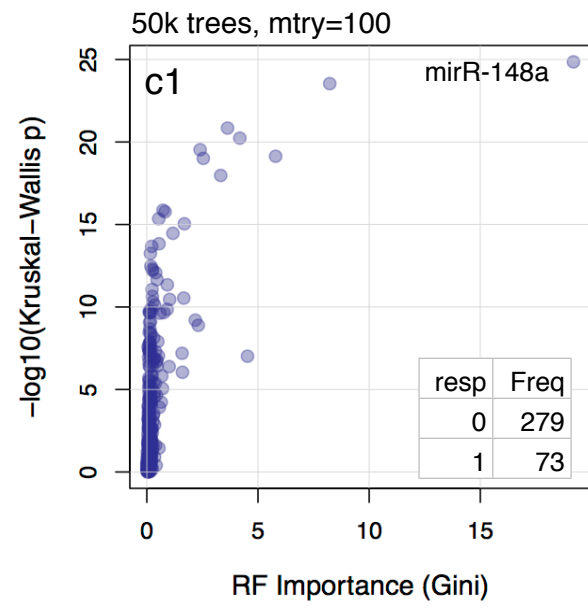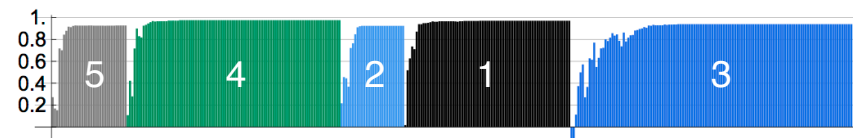
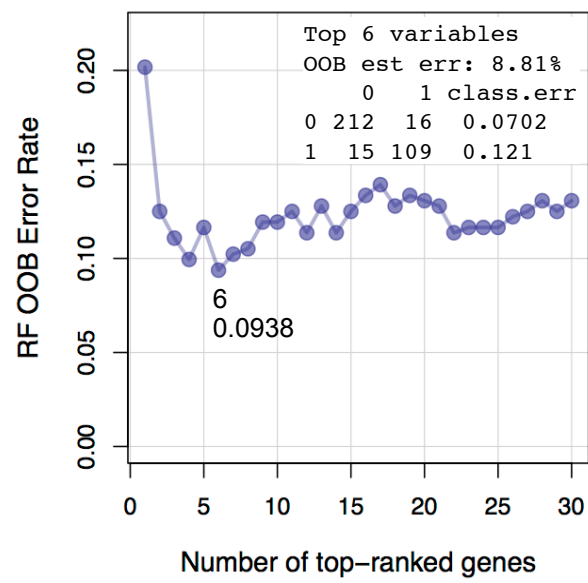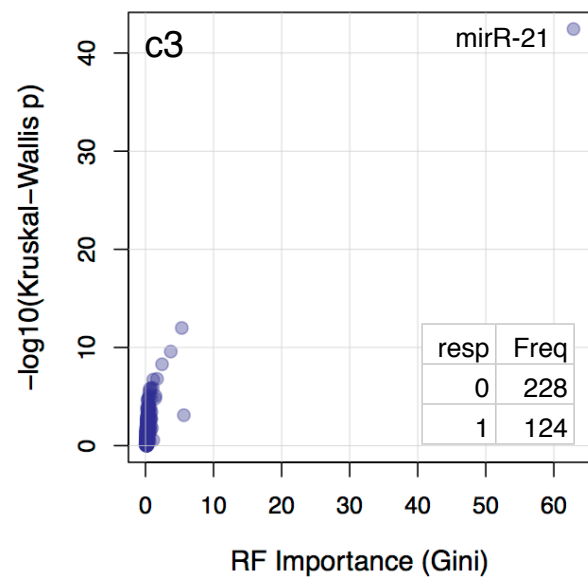| miR	MIMAT	Mn decr acc	Mn decr Gini
miR-148a	MIMAT0000243	14.3889	19.1897
miR-29c*	MIMAT0004673	8.3	8.2264
miR-29b	MIMAT0000100	7.2914	5.7926
miR-21	MIMAT0000076	9.0483	4.5271
miR-26b*	MIMAT0004500	5.4855	4.1831
miR-29b-2*	MIMAT0004515	6.1279	3.6305
miR-126	MIMAT0000445	5.5434	3.323
miR-30e	MIMAT0000692	5.9316	2.5411
let-7g*	MIMAT0004584	4.5681	2.3977
miR-29c	MIMAT0000681	5.459	2.3145
miR-30a	MIMAT0000087	7.2773	2.1811
miR-186	MIMAT0000456	5.5568	1.6871
miR-101	MIMAT0000099	5.621	1.6609
miR-148a*	MIMAT0004549	4.603	1.6024
miR-532	MIMAT0002888	5.9857	1.5849

c2 mirR-375

Top 6 variables
OOB est err: 1.42%
```
         0   1  class.err
0  323   1  0.00309
1    4  24  0.143
```
6
0.0142

resp	Freq
0	324
1	28
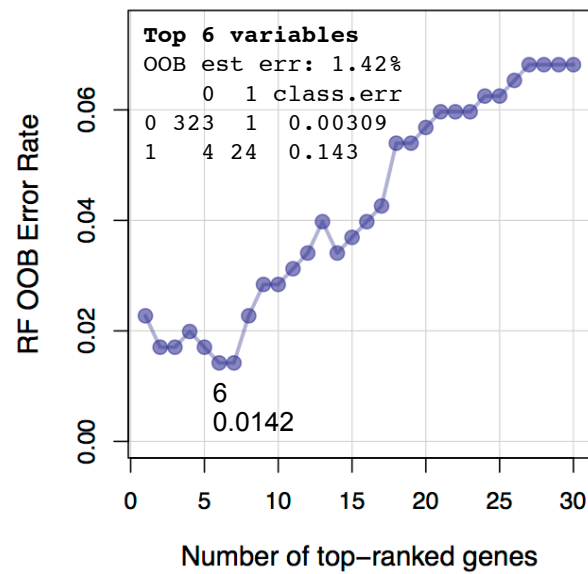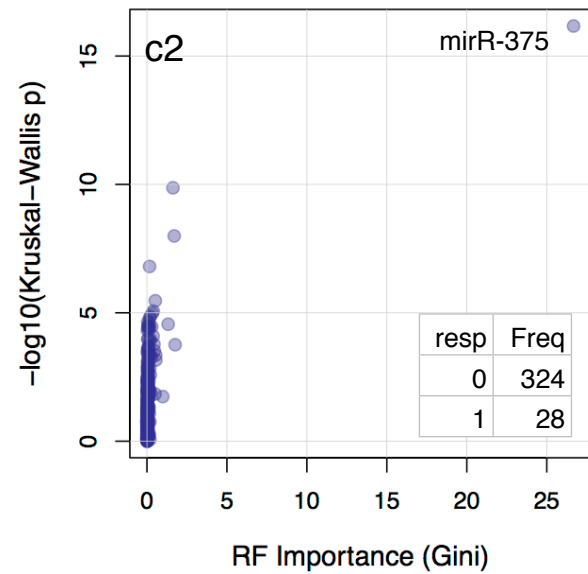
miR	MIMAT	Mn decr acc	Mn decr Gini
miR-375	MIMAT0000728	17.9766	26.6745
miR-1224	MIMAT0005458	4.2713	1.7627
miR-224	MIMAT0000281	2.2891	1.7046
miR-452	MIMAT0001635	2.2161	1.6271
miR-221	MIMAT0000278	2.4581	1.3263
miR-192	MIMAT0004543	5.8193	0.9896

c3 mirR-21

Top 6 variables
OOB est err: 8.81%
```
         0    1  class.err
0  212  16  0.0702
1   15 109  0.121
```
6
0.0938

resp	Freq
0	228
1	124

miR	MIMAT	Mn decr acc	Mn decr Gini
miR-21	MIMAT0000076	18.35	62.8875
miR-143	MIMAT0000435	11.5001	5.6378
miR-21	MIMAT0004494	6.3564	5.3264
hsa.let.7c	MIMAT0000064	4.1517	3.7136
miR-200b	MIMAT0000318	3.6637	2.4279
hsa.let.7g	MIMAT0000414	3.0173	1.6955

Mehrian-Shai R et al. PNAS. 2007;104(13):5563-8.

LUAD 452T: discriminatory MIMATs (cont'd)

c4 (50k trees, mtry=100)

Top 6 variables
OOB est err: 6.25%

	0	1	class.err
0	252	6	0.0233
1	16	78	0.1702

6
0.0653

resp	Freq
0	258
1	94
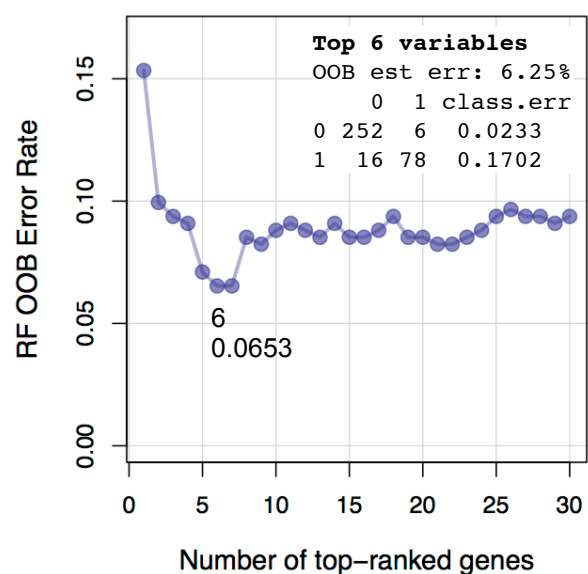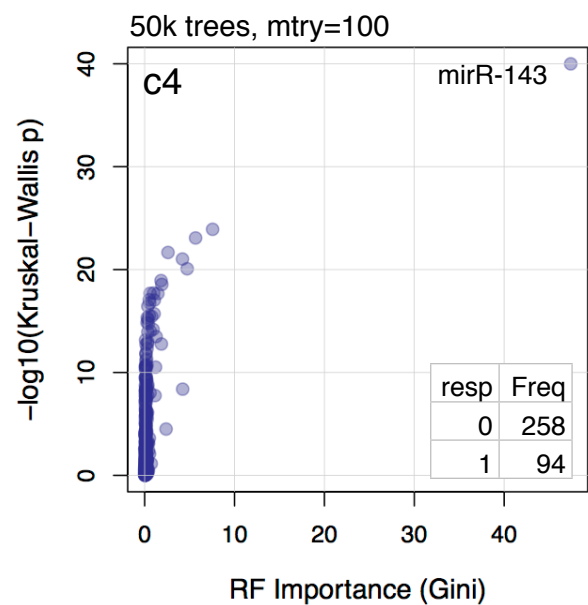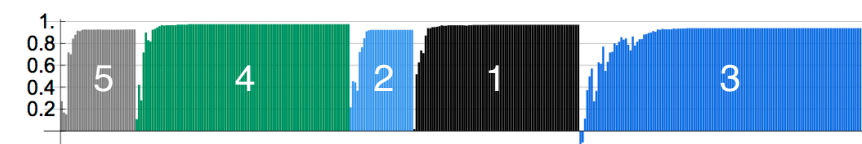
miR	MIMAT	Mn decr acc	Mn decr Gini
miR-143	MIMAT0000435	17.4783	47.4067
let-7f	MIMAT0000067	9.3655	7.5597
miR-210	MIMAT0000267	6.4134	5.6692
let-7a	MIMAT0000062	7.8004	4.7309
miR-21	MIMAT0000076	9.1148	4.2305
miR-345	MIMAT0000772	5.2374	4.2062

c5 (50k trees, mtry=100)

Top 6 variables
OOB est err: 6.82%

	0	1	class.err
0	314	5	0.0157
1	19	14	0.576

6
0.0682

resp	Freq
0	319
1	33
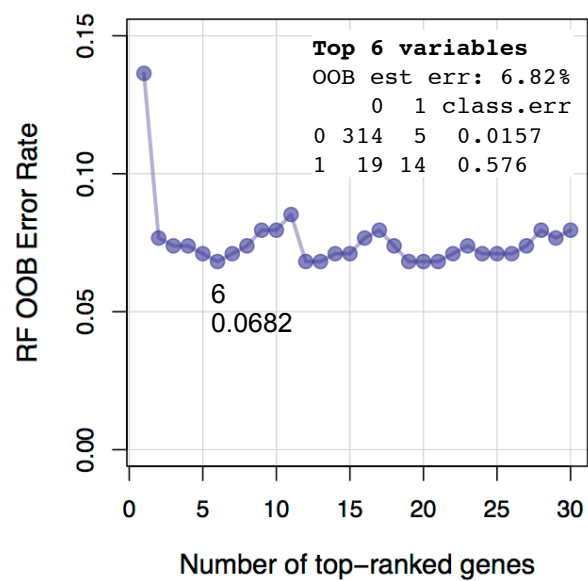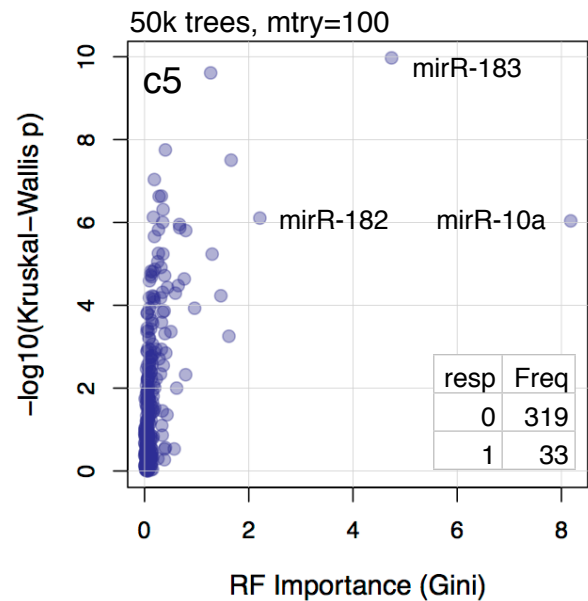
miR	MIMAT	Mn decr acc	Mn decr Gini
miR-10a	MIMAT0000253	13.4526	8.179
miR-183	MIMAT0000261	10.9026	4.7403
miR-182	MIMAT0000259	6.0951	2.2127
miR-25	MIMAT0000081	4.2459	1.6613
miR-205	MIMAT0000266	4.9635	1.6187
miR-149	MIMAT0000450	4.344	1.4644