1 **Supplementary Results**

2

3 **Validation of microarray expression value by quantitative RT-PCR**

4 In order to validate the lncRNA expression value measured by microarray, we performed

5 quantitative RT-PCR (qRT-PCR) of the three lncRNAs of the signature in 25 randomly selected

6 tumor samples. Comparison of the expression values measured by qRT-PCR and microarray

7 showed high level of positive correlation. The Pearson correlation coefficient of

8 ENST00000435885·1, XLOC_013014, and ENST00000547963·1 were 0·74 (p<0·0001), 0·46

9 (p=0·0211) and 0·51 (p=0·0089), respectively (Supplementary Figure 2). This result indicates that

10 the lncRNA expression level measured by microarray is reliable.

11

12 **Correlation of expression level among lncRNAs in the four-lncRNA and three-lncRNA signatures**

13 As mentioned in the method, in order to select the final prognostic signature, we compared the

14 performances of k-lncRNA signature in the training set for all k=1,2,…,9 and the signatures with

15 the best accuracies for each k was defined as the k-lncRNA signature. Three four-lncRNA

16 signatures had the same highest accuracy (84.7%) for k=4 in the training set, which was also the

17 highest accuracy for all k=1,2,…,9. The three four-lncRNA signatures were:

18 *Signature4_1*: ENST00000435885·1, XLOC_013014, XLOC_010016, and ENST00000547963·1

19 *Signature4_2*: ENST00000435885·1, XLOC_013014, XLOC_011774, and ENST00000547963·1

20 *Signature4_3*: ENST00000435885·1, XLOC_010016, XLOC_011774, and ENST00000547963·1

21

22 Only one three-lncRNA signature had the highest accuracy (81.9%) for k=3 in the training set,

23 which was:

24 *Signature3*: ENST00000435885·1, XLOC_013014, and ENST00000547963·1

25

26 Two lncRNAs (ENST00000435885·1, ENST00000547963·1) were included in all the four-lncRNA

27 signatures and the three-lncRNA signature, and XLOC_013014 were included in all these

28 signatures except *Signature4_3*.

29

30 We explored the correlations between these signature related lncRNAs by calculating the

31 pearson and spearman's correlation coefficient in the training/test/independent sets, and the

32 results are shown below.

33 Training set (**a**: p value < 2.2e^{-16}, others: p value > 0.05)

Pearson's correlation coefficient	ENST00000435885.1	XLOC_013014	ENST00000547963.1	XLOC_010016	XLOC_011774
ENST00000435885.1		-0.03	0.09	-0.04	-0.06
XLOC_013014			-0.12	0.93[a]	0.95[a]
ENST00000547963.1				-0.06	-0.07
XLOC_010016					0.95[a]
XLOC_011774					

Spearman's correlation coefficient	ENST00000435885.1	XLOC_013014	ENST00000547963.1	XLOC_010016	XLOC_011774
ENST00000435885.1		-0.06	0.08	-0.12	-0.11
XLOC_013014			-0.05	0.86[a]	0.91[a]
ENST00000547963.1				0.03	0.04
XLOC_010016					0.92[a]
XLOC_011774					

34

1

Test set (**a**: p value $< 2.2e^{-16}$, others: p value > 0.05)

Pearson's correlation coefficient	ENST00000435885.1	XLOC_013014	ENST00000547963.1	XLOC_010016	XLOC_011774
ENST00000435885.1		-0.04	0.13	-0.03	-0.04
XLOC_013014			-0.09	0.91[a]	0.97[a]
ENST00000547963.1				-0.17	-0.08
XLOC_010016					0.95[a]
XLOC_011774					

Spearman's correlation coefficient	ENST00000435885.1	XLOC_013014	ENST00000547963.1	XLOC_010016	XLOC_011774
ENST00000435885.1		0.00	0.13	0.01	0.00
XLOC_013014			-0.09	0.92[a]	0.96[a]
ENST00000547963.1				-0.16	-0.07
XLOC_010016					0.95[a]
XLOC_011774					

Independent set (**a**: p value $< 2.2e^{-16}$, **b**: p value $= 7.66e^{-3}$, **c**: p value $= 3.22e^{-2}$, others: p value > 0.05)

Pearson's correlation coefficient	ENST00000435885.1	XLOC_013014	ENST00000547963.1	XLOC_010016	XLOC_011774
ENST00000435885.1		0.08	0.34[b]	0.00	0.10
XLOC_013014			0.05	0.90[a]	0.99[a]
ENST00000547963.1				0.06	0.08
XLOC_010016					0.88[a]
XLOC_011774					

Spearman's correlation coefficient	ENST00000435885.1	XLOC_013014	ENST00000547963.1	XLOC_010016	XLOC_011774
ENST00000435885.1		0.15	0.28[c]	0.00	0.17
XLOC_013014			0.10	0.89[a]	0.99[a]
ENST00000547963.1				0.09	0.09
XLOC_010016					0.88[a]
XLOC_011774					

From the results, high level of positive correlation (pearson's and spearman's correlation coefficients of all pairs > 0.86, and p value $< 2.2e^{-16}$) among the expression levels of XLOC_013014, XLOC_010016 and XLOC_011774 were observed. So, there is one redundant lncRNA in each of the four-lncRNA signatures. But for the three-lncRNA signature, there is no redundant lncRNA.

Comparison of prognostic performance between three-lncRNA and four-lncRNA signatures

According to our algorithm, the patients could be classified into high- or low-risk groups by the signature. The following tables show the classification results of patients in training/test/independent sets by these signatures.

In the table, "1" denotes the patient is classified as low-risk and "2" denotes high-risk.

Training set

	Signature4_1	Signature4_2	Signature4_3	Signature3		Signature4_1	Signature4_2	Signature4_3	Signature3
p1	1	1	1	2	p31	1	1	1	2
p2	1	1	1	1	p32	1	1	1	1
p3	1	1	1	1	p33	2	2	2	2
p4	1	1	1	1	p34	2	2	2	2
p5	1	1	1	1	p35	2	2	2	2
p6	1	1	1	1	p36	1	1	1	1
p7	1	1	1	1	p37	1	1	1	1
p8	1	1	1	1	p38	2	2	2	2
p9	1	1	1	1	p39	2	2	2	2
p10	1	1	1	1	p40	2	2	2	2
p11	2	2	2	2	p41	1	1	1	1
p12	1	1	1	1	p42	2	2	2	2
p13	1	1	1	1	p43	2	2	2	2
p14	1	1	1	2	p44	2	2	2	2
p15	1	1	1	1	p45	2	2	2	2
p16	1	1	1	1	p46	2	2	2	2
p17	1	1	1	1	p47	2	2	2	2
p18	1	1	1	1	p48	2	2	2	2
p19	1	1	1	1	p49	2	2	2	2
p20	2	2	2	2	p50	2	2	2	2
p21	1	1	1	1	p51	2	2	2	2
p22	1	1	1	1	p52	1	1	1	1
p23	1	1	1	1	p53	2	2	2	2
p24	1	1	1	1	p54	2	2	2	2
p25	2	2	2	2	p55	2	2	2	2
p26	2	2	2	2	p56	2	2	2	2
p27	2	2	2	2	p57	2	2	2	2
p28	2	2	2	2	p58	1	1	1	1
p29	2	2	2	2	p59	1	1	1	1
p30	2	2	2	2	p60	2	2	2	2

1

Test set

	Signature4_1	Signature4_2	Signature4_3	Signature3		Signature4_1	Signature4_2	Signature4_3	Signature3
p1	1	1	1	1	p31	1	1	1	1
p2	2	2	2	2	p32	2	2	2	2
p3	1	1	1	1	p33	2	2	2	2
p4	1	1	1	1	p34	1	1	1	1
p5	1	1	1	1	p35	2	2	2	2
p6	1	1	1	1	p36	1	1	1	1
p7	1	1	1	1	p37	2	2	2	2
p8	1	1	1	1	p38	2	2	2	2
p9	2	2	2	2	p39	1	1	1	2
p10	1	1	1	2	p40	1	1	1	1
p11	1	1	1	1	p41	1	1	1	1
p12	2	2	2	2	p42	1	1	1	2
p13	2	2	2	2	p43	1	1	1	1
p14	2	2	2	2	p44	2	2	2	2
p15	2	2	2	2	p45	1	1	1	1
p16	2	2	2	2	p46	1	1	1	1
p17	1	1	1	1	p47	1	1	1	1
p18	2	2	2	2	p48	1	1	1	1
p19	1	1	1	1	p49	1	1	1	1
p20	2	2	2	2	p50	2	2	2	2
p21	1	1	1	1	p51	2	2	2	2
p22	1	1	1	1	p52	2	2	2	2
p23	2	2	1	2	p53	1	1	1	1
p24	2	2	2	2	p54	1	1	1	1
p25	1	1	1	1	p55	1	1	1	1
p26	2	2	2	1	p56	1	1	1	1
p27	2	2	2	2	p57	1	1	1	1
p28	1	1	1	1	p58	1	2	2	1
p29	2	2	2	2	p59	1	1	1	1
p30	1	1	1	1					

2

Independent set

	Signature4_1	Signature4_2	Signature4_3	Signature3		Signature4_1	Signature4_2	Signature4_3	Signature3
p1	2	2	2	2	p31	2	2	2	2
p2	2	2	2	2	p32	2	2	2	2
p3	2	2	2	1	p33	2	2	2	2
p4	2	2	2	2	p34	2	2	2	1
p5	2	2	2	2	p35	2	1	2	2
p6	1	1	1	1	p36	1	1	1	1
p7	1	1	1	1	p37	1	1	1	1
p8	2	2	2	2	p38	1	1	1	1
p9	1	1	1	1	p39	2	2	2	1
p10	2	2	2	2	p40	2	2	2	2
p11	1	1	1	1	p41	2	2	2	2
p12	1	1	1	1	p42	1	1	1	1
p13	1	1	1	1	p43	2	2	2	2
p14	1	1	1	1	p44	1	1	1	1
p15	2	2	2	2	p45	2	2	2	2
p16	1	1	1	1	p46	2	2	2	2
p17	1	1	1	1	p47	2	2	2	2
p18	1	1	1	1	p48	2	2	2	2
p19	2	2	2	2	p49	2	1	2	2
p20	2	2	2	2	p50	2	2	2	2
p21	2	2	2	1	p51	2	2	2	2
p22	2	2	2	2	p52	2	2	2	2
p23	2	2	2	2	p53	2	2	2	2
p24	2	2	2	2	p54	2	2	2	2
p25	1	1	1	1	p55	1	1	1	1
p26	2	1	2	2	p56	2	2	2	2
p27	1	1	2	2	p57	1	1	1	1
p28	1	1	1	1	p58	2	2	2	2
p29	2	2	2	2	p59	2	2	2	2
p30	2	2	2	2	p60	2	2	2	2

From the tables, we see that the survival prediction results (high- or low-risk) of patients for the three four-lncRNA signatures were much the same. For the 60 patients of the training set, classification results of the three signatures were exactly the same. There were 1 (*Signature4_2*) and 2 (*Signature4_3*) patients with different classification compared to *Signature4_1* for the 59 patients of the test set, and were 3 (*Signature4_2*) and 1 (*Signature4_3*) for the 60 patients of the independent set.

Because the classification results for the four-lncRNA signatures were much the same, we only summarized the comparison between *Signature4_1* and the 3-lncRNA signature. From the tables, there were 3 patients with different classification compared to *Signature4_1* for 60 patients of the training set. The number was 4 and 5 for 59 patients of the test set and 60 patients of the independent set, respectively. So the classification results were very similar between the three four-lncRNA signatures and the three-lncRNA signature.

Prognostic performance comparison between three-lncRNA signature and four-lncRNA signatures

As mentioned above, the survival prediction results (high- or low-risk) of the three four-lncRNA signatures were very similar. So, we only show the performance comparison between *Signature4_1* and the three-lncRNA signature below. The prognostic performance of *Signature4_2/Signature4_3* and the three-lncRNA signature were also very similar (figure not shown).

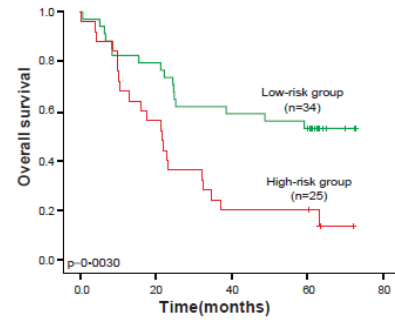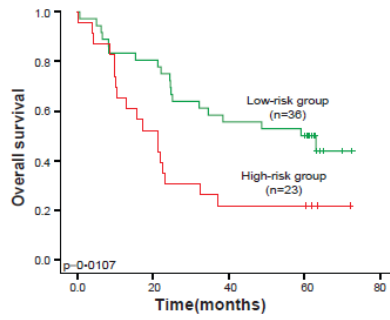Comparison of Kaplan-Meier survival curves of *Signature4_1 and Signature3*
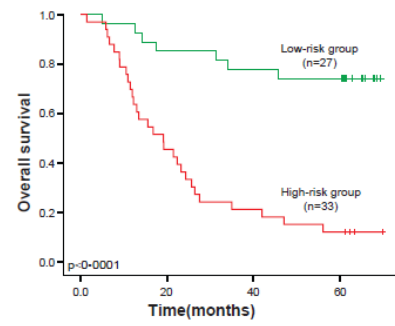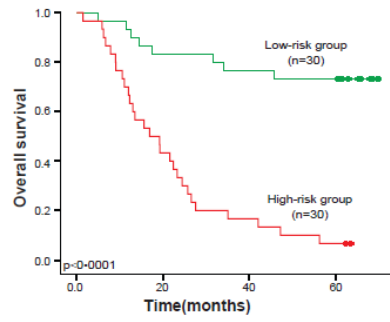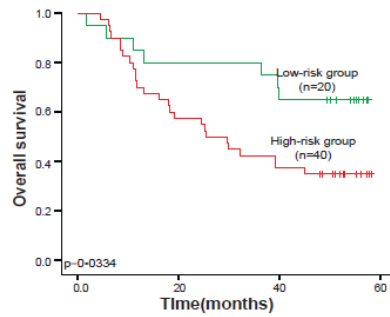
1
2
3
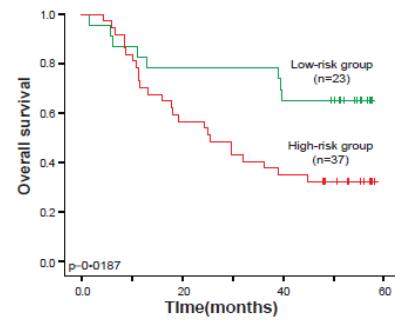4 Comparison of Kaplan-Meier survival curves of *Signature4_1 and Signature3* within clinical stages

Signature4_1   Signature3

Stage II patients, training set (n=22)

Stage II patients, combined test set and independent cohort (n=55)

Stage III patients, training set (n=36)

Stage III patients, combined test set and independent cohort (n=56)
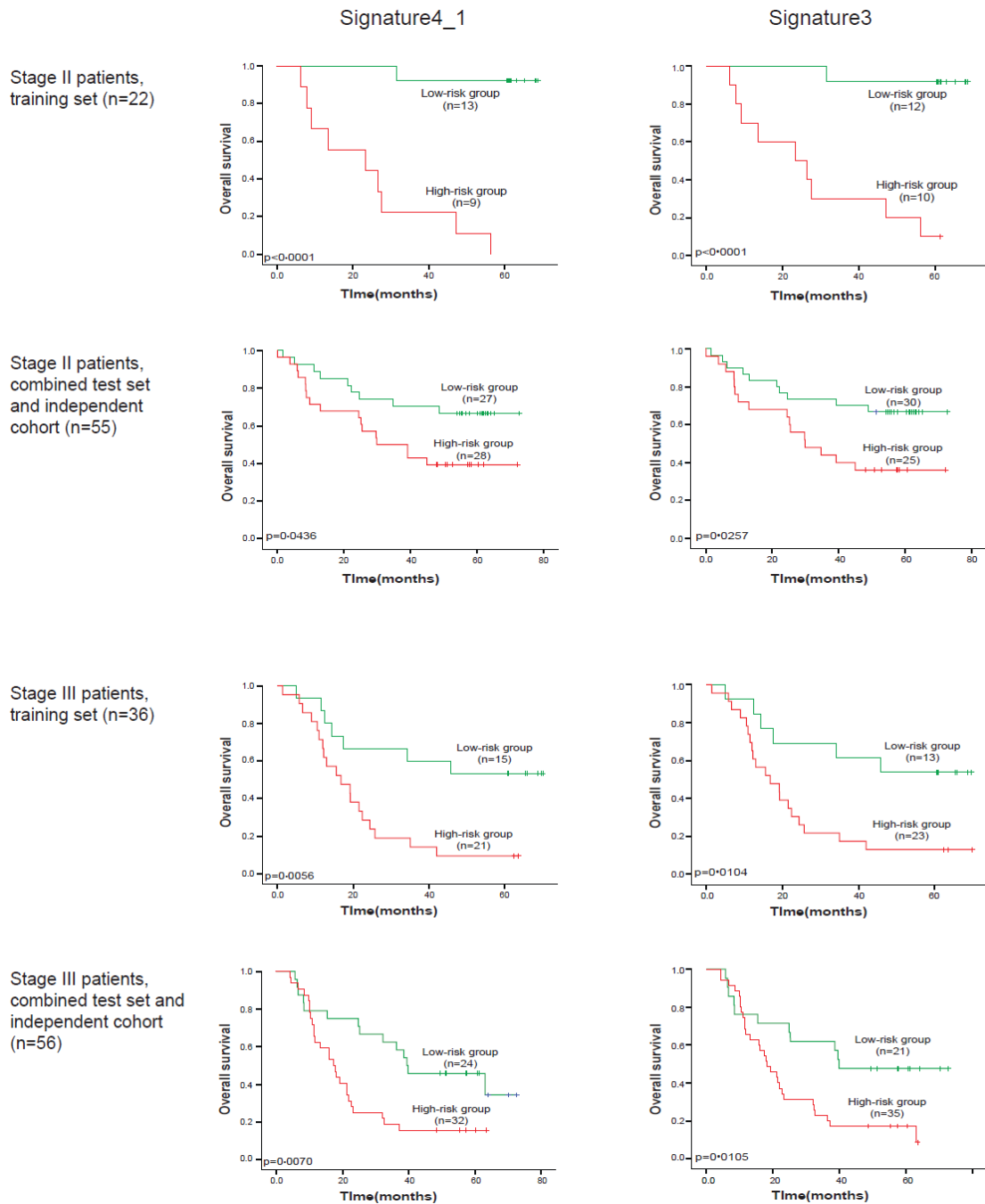
1

2

3   From these results, we see that the performances between the four-lncRNA signature and
4   three-lncRNA signature are very similar. The p values of Log-rank tests show that both the
5   four-lncRNA signature and the three-lncRNA signature are able to predict the survival of patients
6   both all together and within TNM stages and that four-lncRNA signature does not improve
7   prognostic power compared with three-lncRNA signature.