SUPPLEMENTAL INFORMATION

LEGENDS TO SUPPLEMENTAL FIGURES

Supplemental Figure 1. **Performance of the 49-gene model**. These data are supplemental to those shown in Fig. 2 of the main text. In **A**, the 49-gene model was used to predict overall survival in lung adenocarcinoma patients of the Harvard and Michigan cohorts, as shown. Patients were divided according to tumor stage, and Kaplan-Meyer analyses are shown for Stage I patients (as in Fig. 2 of the main text), Stage II-III patients, and for the entire cohort (Stage I-II-III). As shown, the 49-gene model had no predictive power in Stage II-III adenocarcinomas. In **B**, the performance of the 49-gene model in the prediction, by Kaplan-Meyer analysis, of overall patient survival in the Duke cohort is compared to that of the 50- and 100-gene signatures of Beer *et al*. (ref. 9 in the main text). Red, favorable signature; Green, unfavorable signature.

Supplemental Figure 2. Performance of the 10-gene model on the Michigan and Harvard cohorts. These data are supplemental to those shown in Fig. 4A of the main text. The 10-gene signature was used to predict overall patient survival within the Michigan and Harvard cohorts. In the case of the Michigan cohort, 7 genes could be used (*SF3B1, NUDCD1* and *SCGB3A1* were not present on HU6800 microarray, used in that study), in the case of the Harvard cohort, 8 genes could be used (*NUDCD1* and *SCGB3A1* were not present on HU95av2 microarray used in that study). Patients from the two cohorts were grouped according to tumor stage, as described in Supplemental Fig. 1. Remarkably, despite the reduction in the number of genes utilizable, the

signature could still predict overall survival in patients with Stage I disease. The prediction in the Duke cohort (all 10 genes could be used in this case) is the same as in Fig. 4A of the main text, and is reported for comparison.

Supplemental Figure 3. Survival analysis of stage IA and stage IB adenocarcinoma patients.

A. The 10-gene model was tested to predict overall survival in the IFOM cohort of Stage IA (as reported also in Fig. 4B) and Stage IB lung adenocarcinomas. Data are shown, in a Kaplan-Meier plot, as the probability of survival as a function of a "favorable" (red line), or "unfavorable" (green line) signature.

B. The IFOM and Duke cohorts of stage I adenocarcinoma patients were stratified accordingly to the tumor stage. Data are shown, in a Kaplan-Meier plot, as the probability of survival as a function of a "stage IA" (red line), or "stage IB" (green line) parameter.

SUPPLEMENTAL TABLES

Supplemental Table 1. Clinicopathological data for the various patient groups used in this study.

	Michigan cohort (N=41)	Harvard cohort (N=60)	Duke cohort (N=34)	IFOM Train. cohort (N=25)	IFOM Val. cohort (N=45)
Age - Years					
Median	60	61	66	64	64
Range	41-80	33-88	43-83	51-72	48-81
Mean±SD	61±10	62±11	65±9	63±6	64±9
Sex - No(%)					
Male	16 (39)	26 (43)	17 (50)	23 (92)	41 (91)
Female	25 (61)	34 (57)	17 (50)	2 (8)	4 (9)
Smoking (pack year) - No. (%)					
None	3 (7)	3 (5)			
≤20 Yr	5 (12)	10 (17)			
21-49 Yr	13 (32)	20 (33)			
≥50 Yr	17 (41)	27 (45)			
N.R	3 (7)				
Stage - No. (%)					
I*		1 (2)	4 (12)		
IA	16 (39)	15 (25)	21 (62)	8 (32)	13 (29)
IB	12 (29)	28 (46)	9 (26)	17 (68)	32 (71)
IIA		2 (3)			
IIB		8 (13)			
IIIA	13 (32)	4 (7)			
IIIB		2 (3)			
Tumor stage - No. (%)					
1	20 (49)	19 (32)	21 (62)	8 (32)	13 (29)
2	16 (39)	35 (58)	9 (26)	17 (68)	32 (71)
3	5 (12)	3 (5)			
4		2 (3)			
N.A.		1 (2)	4 (12)		
Nodal status - No. (%)					
Negative	28 (68)	36 (60)			
Positive	13 (32)	12 (20)			
N.A.		12 (20)			

Legend to Supplemental Table 1. Data are reported for the patients included in the reduced datasets of the Michigan and Harvard cohorts, and for all patients of the other cohorts.
*, Tumor stage (A or B) was not available; N.A., not available

Supplemental Table 2. Prognostic predictive accuracy of the 49-gene model.

Reduced Datasets						
	Michigan cohort N = 41			Harvard cohort N = 60		
Model	Accuracy (%)	Sensitivity (%)	Specificity (%)	Accuracy (%)	Sensitivity (%)	Specificity (%)
49-gene	90	89	91	72	67	76
50-gene	73	50	91	63	67	61
100-gene	80	72	87	63	59	67
71-gene	85	83	87	73	67	79

Original datasets									
	Michigan cohort N = 86			Harvard cohort N = 84			Duke cohort N = 34		
Model	Acc. (%)	Sen. (%)	Spec. (%)	Acc. (%)	Sen. (%)	Spec. (%)	Acc. (%)	Sen. (%)	Spec. (%)
49-gene	69	67	69	71	71	72	82	93	74
50-gene	81	58	90	60	60	58	74	80	69
100-gene	81	75	84	58	63	53	76	80	74
71-gene	73	67	76	69	69	69	82	93	74

Modified Reduced Datasets						
	Michigan cohort N = 66			Harvard cohort N = 61		
Model	Accuracy (%)	Sensitivity (%)	Specificity (%)	Accuracy (%)	Sensitivity (%)	Specificity (%)
49-gene	79	83	77	74	70	77
50-gene	85	56	96	67	67	68
100-gene	83	72	88	64	56	71
71-gene	85	83	85	74	67	79

Legend to Supplemental Table 2. The 49-gene model was tested for prognostic predictive accuracy by leave-one-out cross-validation. Two other models of 50 and 100 genes, respectively, from Beer *et al*. (ref. 9 in the main text) were tested, as a comparison. Models were tested on the reduced datasets (top panel), and on the original datasets (middle panel) from the Michigan, Harvard, and Duke cohorts (N, number of patients in the dataset).

An additional control was also performed, to ensure that criteria used for the cut-offs did not introduce biases in the analysis and/or somehow led to overfitting of the data. To do this, we changed the cut-off criteria to define the good and poor prognosis groups. The new criteria used were: poor prognosis = death event in < 30 months; good prognosis = alive >15 months (in both cohorts of patients). With these new criteria, we obtained what are defined as "modified reduced datasets" (bottom panels). These datasets included more patients than the "reduced datasets" (a total of 69 patients are excluded from the reduced datasets, while only 43 patients are excluded from the modified reduced datasets). The "modified reduced datasets" were subjected to meta-analysis (as described in Methods). This led to the identification of a 71-gene prognostic signature (which shares 23 genes with the 49-gene signature). The performance of the 71-gene model in the leave-one-out cross-validation is shown

here, in comparison to the 49-gene model and to the 50- and 100-gene models from Beer *et al.*, for the modified reduced datasets, the reduced datasets, and the original datasets. As shown, the 71-gene model did not perform better than the 49-gene model, despite containing 22 additional genes.

Supplemental Table 3. Differential expression of genes of the 28-gene biased signature in Stage I lung adenocarcinomas.

Gene	Probe set	FOLD tumor/normal	p-value
Michigan study			
NASP	m97856_at	1.33	0.0105
RRM2	x59618_at	4.93	0.0032
E2F1	s49592_s_at	1.57	<0.0001
MCM4	x74794_at	2.11	<0.0001
Harvard study			
NASP	33255_at	1.37	0.0412
G3BP2	35793_at	1.20	0.0185
SF3BP1	39444_at	2.02	0.0097
MCM4	981_at	3.96	0.0003
CCNE2	35249_at	3.87	0.0002
RRM2	36922_at	4.17	<0.0001
MCM6	40117_at	1.89	<0.0001

Gene	Probe set	FOLD poor/good	p-value
Harvard study			
MCM6	40117_at	1.44	0.0011
MCM7	947_at	1.38	0.0080
Duke study			
USP37	226729_at	1.72	0.0012
RRM2	201890_at	2.13	0.0028
FLJ37562	1553108_at	1.97	0.0032
MCM7	210983_s_at	1.93	0.0071
MCM4	222036_s_at	2.63	0.0076
G3BP2	208841_s_at	1.39	0.0160
SF3B1	201071_x_at	1.26	0.0172
FAM91A1	226294_x_at	1.43	0.0187
UHRF1	225655_at	1.95	0.0216
HAT1	203138_at	1.79	0.0227
TRPC4AP	212059_s_at	1.25	0.0277
C3orf4	239146_at	0.73	0.0351
SYNCRIP	217832_at	1.33	0.0352
CKAP5	1555278_a_at	1.44	0.0374
SCC-112	212138_at	1.28	0.0428

Legend to Supplemental Table 3. These data are supplemental to those shown in Fig. 3A of the main text.

Since datasets from the Michigan, Harvard and Duke studies were obtained on different generations of chips, not all of the 28 genes of the biased signature were present on the chips. In particular, only 5 and 11 genes of 28 were present on the chips used in the Michigan and Harvard studies, respectively. This prevented Kaplan-Meyer analysis on the Michigan and Harvard cohorts, which, on the other hand, could

be meaningfully performed on the Duke cohorts, since all the 28 genes were present in the datasets (see Fig. 3A).

However, as shown in this Table (top), 4 of 5 genes, and 7 of 11 genes, were significantly overexpressed in tumor vs. normal tissues in the Michigan and Harvard study, respectively (it was not possible to perform the same analysis on the Duke datasets since no information were present as to expression levels in normal lungs).

In addition (bottom part of the Table), in the Harvard cohort, 2 of 11 genes were significantly regulated in the comparison between good and poor prognosis patients. No significant regulation was found in the 5 genes present in the Stage I lung adenocarcinomas in the Michigan cohort, whereas, in the Duke cohort 15 of 28 genes were significantly regulated in the unfavorable prognosis vs. the good prognosis group.

Supplemental Table 4. Analysis of the 80-gene model in the Duke cohort and in the IFOM training cohort.

		Duke cohort		IFOM training cohort (Q-PCR)		
Gene	**Source**	**FOLD**	**p-value**	**Tested**	**FOLD**	**p-value**
SCGB3A1*	Liter.	0.48	0.314	YES	0.03	<0.001
EIF3S6	Liter.	1.28	0.087	YES	1.38	0.425
TERT	Liter.	1.31	0.126	YES	ND	ND
BAT1	E1A	1.26	0.085			
C3orf4	E1A	0.73	0.035			
CCNE1	E1A	1.24	0.432	YES	0.94	0.840
CCNE2	E1A	1.13	0.590	YES	0.78	0.543
CKAP5	E1A	1.44	0.037	YES	1.03	0.988
DEPDC1B	E1A	0.74	0.173			
E2F1*	E1A	1.17	0.288	YES	1.79	0.012
FAM91A1	E1A	1.43	0.019	YES	1.13	0.736
FLJ37562	E1A	1.97	0.003			
G3BP2	E1A	1.39	0.016			
HAT1	E1A	1.79	0.023			
LBR	E1A	1.36	0.143			
MCM4*	E1A	2.63	0.008	YES	1.51	0.317
MCM6*	E1A	1.45	0.058	YES	2.07	0.022
MCM7*	E1A	1.93	0.007	YES	1.62	0.138
NASP	E1A	1.37	0.073	YES	1.19	0.071
NUDCD1*	E1A	1.36	0.082	YES	1.88	0.001
PHOSPHO2	E1A	1.37	0.093			
PTBP2	E1A	0.90	0.521			
RRM2*	E1A	2.13	0.003	YES	5.29	0.049
SCC-112	E1A	1.28	0.043			
SF3B1*	E1A	1.26	0.017	YES	1.14	0.033
SMU1	E1A	1.23	0.309	YES	1.25	0.863
SYNCRIP	E1A	1.33	0.035			
TAF3	E1A	0.74	0.224			
TRPC4AP	E1A	1.25	0.028	YES	1.08	0.194
UHRF1	E1A	1.95	0.022			
USP37	E1A	1.72	0.001			
ARL4A	Meta.	2.16	0.002			
ATP13A3	Meta.	1.37	0.067	YES	1.12	0.164
LU	Meta.	1.22	0.288	YES	0.81	0.316
BFSP1	Meta.	1.27	0.378	YES	1.05	0.780
CTF1	Meta.	1.45	0.093	YES	0.75	0.663
CXCL6*	Meta.	1.26	0.332	YES	2.88	0.424
PSF1	Meta.	0.82	0.445	YES	1.36	0.234
E2F4*	Meta.	1.05	0.335	YES	1.24	0.032
FGF4	Meta.	0.68	0.258	YES	ND	ND
FLJ16124	Meta.	0.80	0.275			
FUCA1	Meta.	1.29	0.166	YES	0.95	0.861
GABPB2*	Meta.	0.79	0.075	YES	1.57	0.046
GNS	Meta.	1.55	0.014	YES	0.86	0.739

GAPDH	Meta.	1.85	0.002	YES	1.32	0.597
GARS	Meta.	1.38	0.003	YES	1.47	0.130
GAP43	Meta.	0.87	0.175	YES	ND	ND
H2AFZ	Meta.	1.52	0.065			
HSPD1	Meta.	2.26	<0.001			
HUWE1	Meta.	1.35	<0.001			
HSPG2*	Meta.	2.41	0.001	YES	1.44	0.016
HOXB7*	Meta.	1.55	0.117	YES	2.93	0.016
HPRT1	Meta.	1.46	0.045	YES	1.00	0.736
IRF2	Meta.	0.74	0.246	YES	1.07	0.803
KRT6B	Meta.	0.70	0.365			
KIAA1128	Meta.	1.20	0.158	YES	1.04	0.470
HLA-DQB1*	Meta.	0.39	0.001	YES	0.59	0.191
3.8-1	Meta.	0.82	0.269			
MAPRE2	Meta.	1.72	0.189	YES	0.86	0.723
MYOM2	Meta.	0.69	0.269	YES	ND	ND
MYH10	Meta.	0.80	0.326			
PGAM1	Meta.	1.57	0.227			
PAICS	Meta.	1.36	0.082	YES	1.04	0.638
POLR2C	Meta.	1.37	0.093	YES	1.03	0.495
KCNJ12	Meta.	0.76	0.375	YES	0.88	0.226
KCNA5	Meta.	1.25	0.368	YES	1.18	0.875
PBXIP1	Meta.	1.32	0.197	YES	0.87	0.897
PFN2	Meta.	1.98	0.037	YES	1.38	0.734
RAFTLIN*	Meta.	1.52	<0.001	YES	1.68	0.744
SEPW1	Meta.	1.33	0.039	YES	1.34	0.434
SERPINB5*	Meta.	3.05	0.122	YES	3.22	0.010
SIAH1	Meta.	1.36	0.006			
ST8SIA1	Meta.	0.97	0.927	YES	1.07	0.876
TDGF1	Meta.	1.75	0.101			
TMSB4X	Meta.	1.03	0.696			
TLE2	Meta.	1.31	0.250	YES	0.81	0.692
VIP	Meta.	0.91	0.770	YES	ND	ND
CRK	Meta.	1.68	0.063	YES	1.20	0.799
ZMAT2	Meta.	1.52	0.001	YES	1.05	0.760
U60269	Meta.	NO	NO			

Legend to Supplemental Table 4. The genes of the 80-gene model are shown with their gene name, and source (Liter., from literature; E1A, from the 28-gene biased signature; Meta., from meta-analysis of the reduced Michigan and Harvard datasets).

Analysis in the Duke cohort was performed on the available microarray datasets; in the IFOM training cohort, the analysis was performed by Real Time PCR (Q-PCR, see main text for details). Fold indicates the average fold increase or decrease in the poor prognosis group, compared to the good prognosis group (followed by its p-value). Genes upregulated or downregulated more than 1.5-fold are highlighted in red and blue, respectively. Asterisks indicate the 16 genes selected to develop the final prognostic model. ND, not detectable. NO, No probeset corresponding to the

hypothetical U60269 gene was present on the HU133 2.0 plus chip used in the Duke analysis.

Supplemental Table 5. Multivariate analysis of the 10-gene prognostic model.

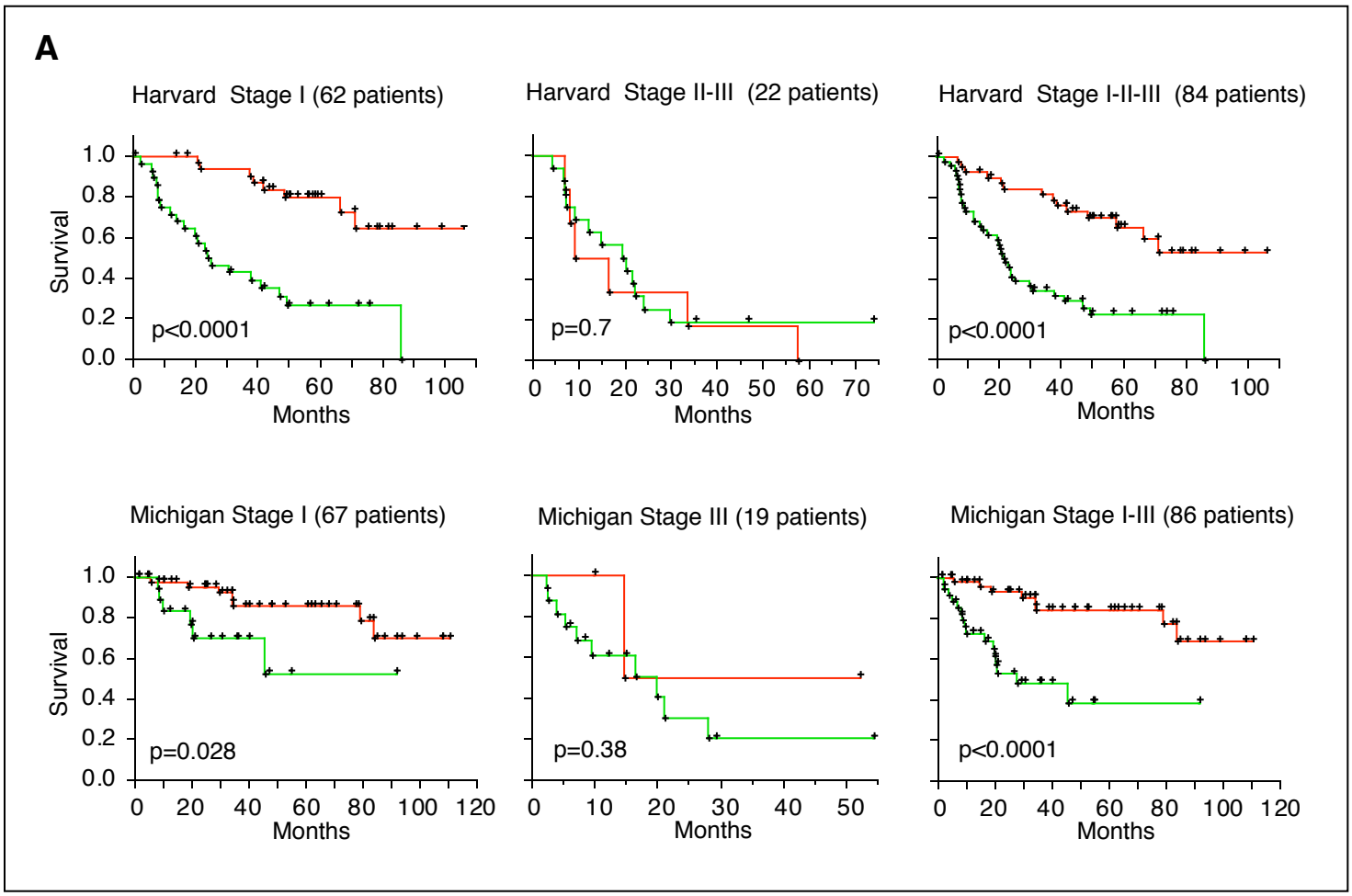
A		**IFOM cohort (N=70)**		**Michigan cohort (N=67)**	
Variable	**Subset**	**OD (95% CI)**	**P**	**OD (95% CI)**	**P**
Subtype	Adenoca/BAC	0.38 (0.03-6.93)	0.478	2.89 (0.52-23.3)	0.259
Different.	well+mod/poor	1.16 (0.19-6.89)	0.871	0.58 (0.12-2.89)	0.499
10-gene	poor/good	8.34 (2.83-27.6)	*0.0002	3.03 (1.14-8.62)	*0.030
B		**IFOM cohort (N=47)**		**Michigan cohort (N=38)**	
Variable	**Subset**	**OD (95% CI)**	**P**	**OD (95% CI)**	**P**
Subtype	Adenoca/BAC	0.18 (0.01-2.32)	0.184	unstable	0.948
Different.	well/poor	0.58 (0.12-2.32)	0.474	unstable	0.945
10-gene	poor/good	11.98 (3.01-60.0)	*0.001	4.81 (ND-ND)	0.093

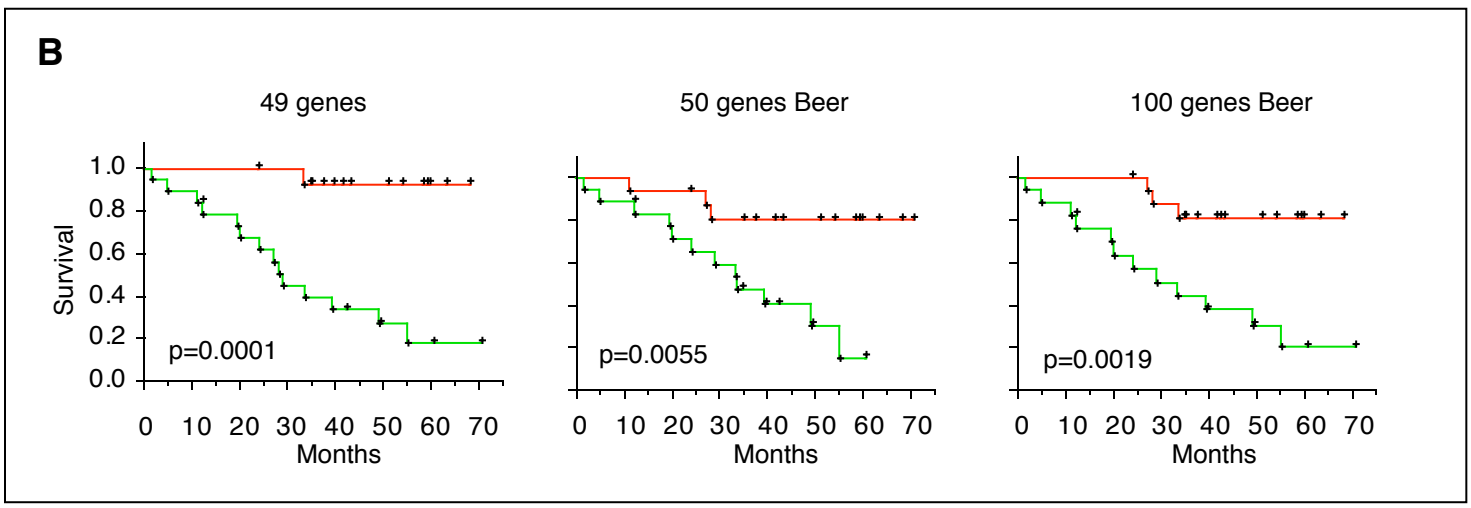
Legend to Supplemental Table 5. **A.** The 10-gene model was tested for prediction of survival in the indicated cohorts of patients, in comparison to other biological parameters. Data are expressed as odds ratio (OD) at 95% confidence interval (95% CI). Asterisks indicate statistically significant values. BAC means bronchioalveolar cell carcinoma. In the Harvard and Duke cohorts, analysis was not performed because of the facts that either the number of BAC was too low (only three in Harvard cohort) or the biological information was unavailable (for the Duke cohort). **B.** We repeated the analysis excluding patients with moderately differentiated tumors. In the Michigan cohort the multivariate model resulted to be unstable because of the unbalanced number of patients with poor prognosis (5 patients) compared to good prognosis (33 patients). ND. Confidence intervals were not calculated because of model instability.

Supplemental Table 6. Multivariate analysis of 10-gene prognostic marker in the Duke cohort.

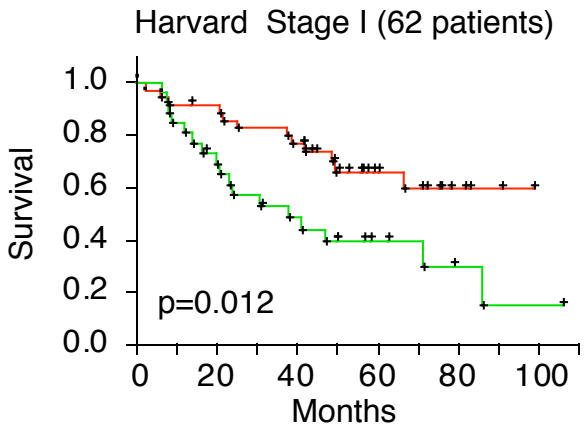
		Duke cohort	
Variable	**Subset**	**OD (95% CI)**	**P**
Sex	male/female	1.57 (0.26 - 10.4)	0.618
Age	≥64/<64	2.05 (0.30 - 19.6)	0.482
Stage	IB/IA	1.48 (0.21 - 9.95)	0.684
10-gene model	poor/good	11.5 (1.94 - 104)	*0.013

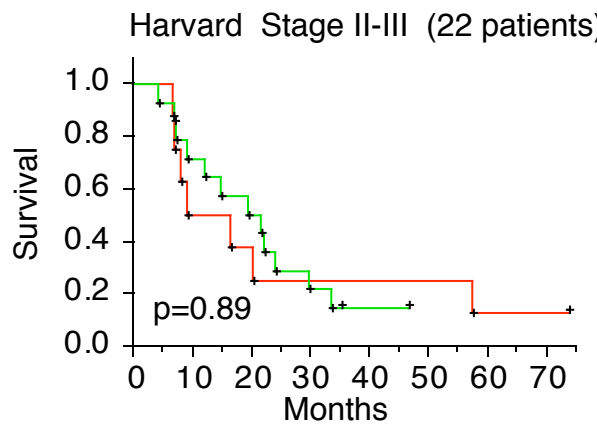
Legend to supplemental Table 6. In the Duke cohort a large fraction of tumors are stage IA adenocarcinoma, thus to test whether the 10-gene model was an independent prognostic factor also in this cohort of patients, multivariate analysis was performed. Data are expressed as odds ratio (OD) at 95% confidence interval (95% CI). Asterisks indicate statistically significant values. Four patients were excluded from the analysis because the stage (IA or IB) was not defined in the original dataset (Patients: 00-0479, 00-909, 00-1082, 00-0941).

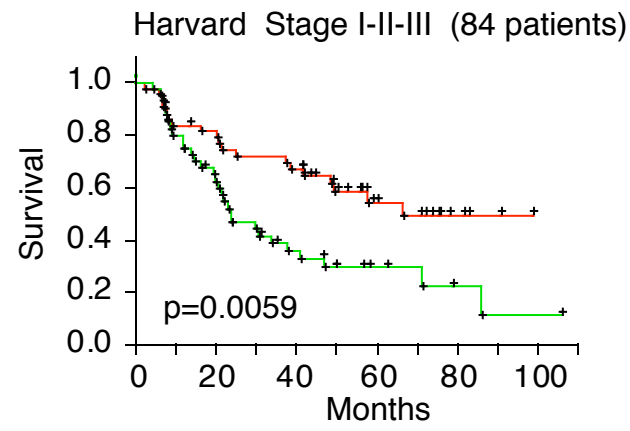


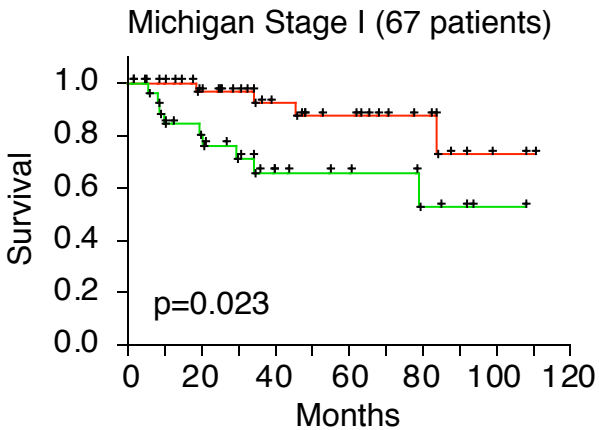


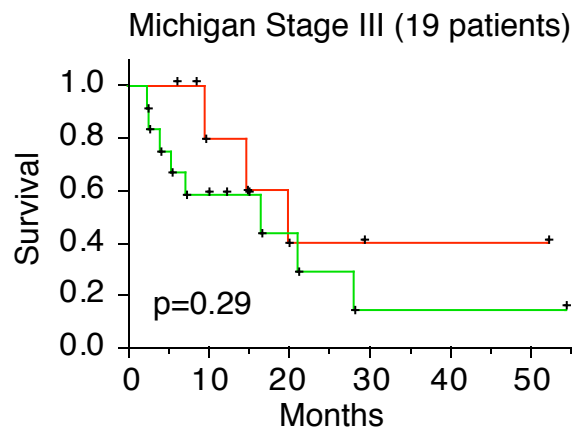
Supplemental Figure 1

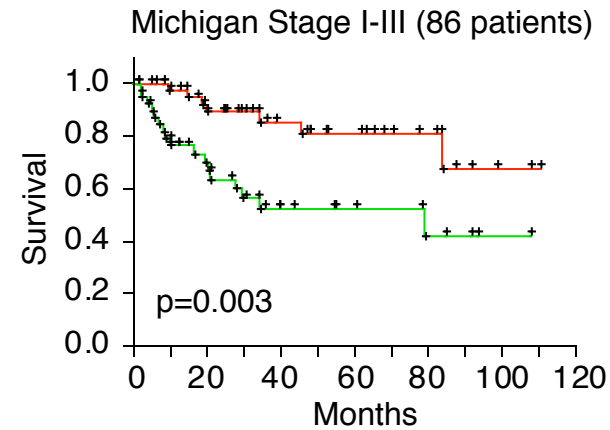


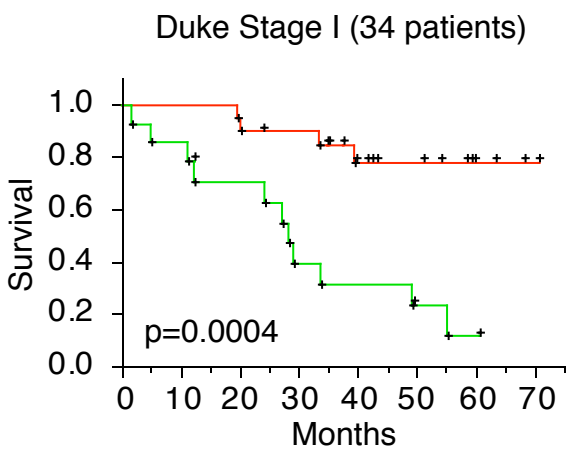




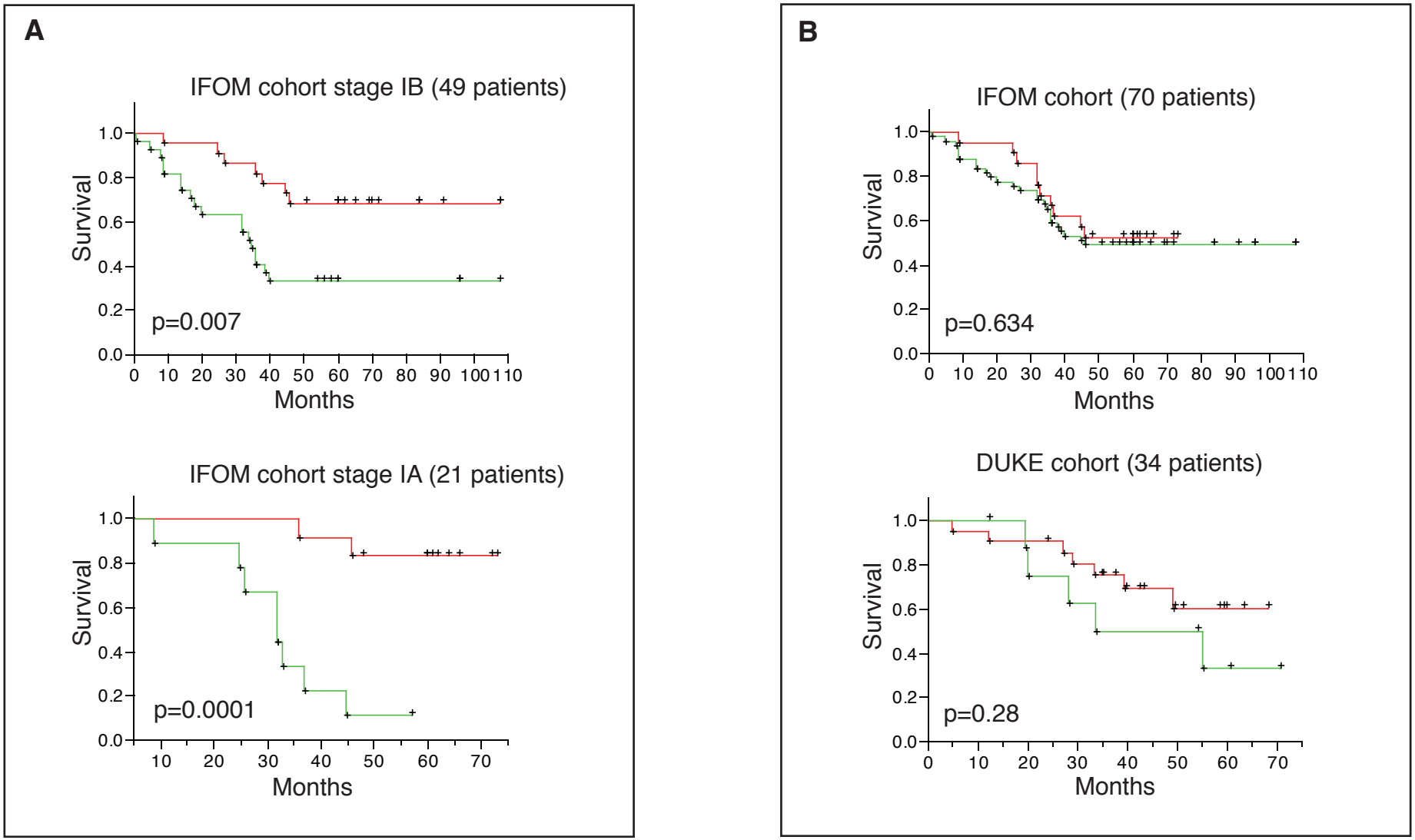








Supplemental Figure 2



Supplemental Figure 3