Supporting Information

Bogunovic et al. 10.1073/pnas.0905139106

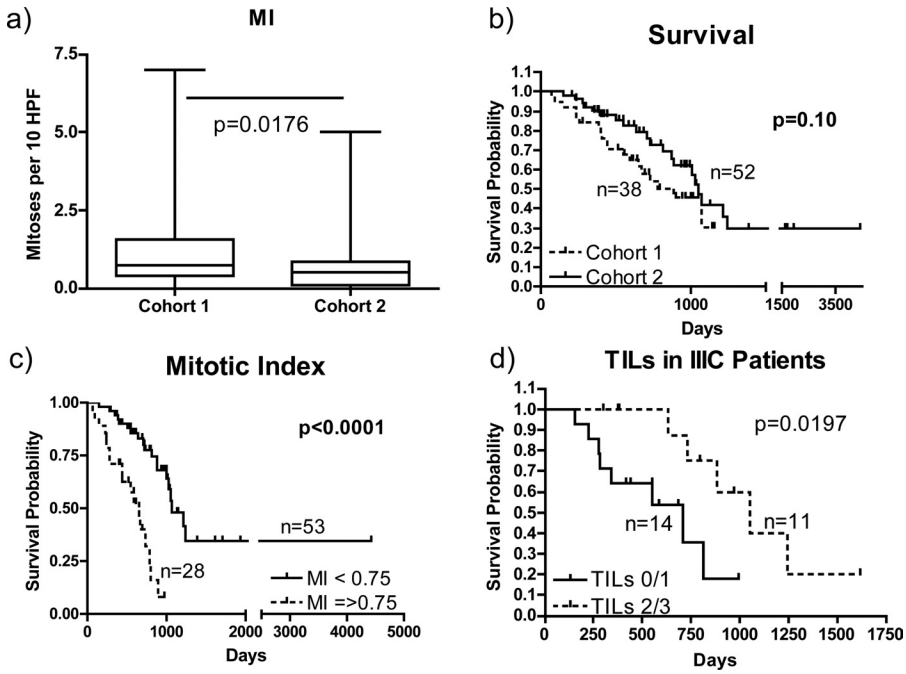

Fig. S1. MI aids staging in combined initial and validation patient cohorts in predicting their survival. Fifty-two additional patient samples with stage IIIb and IIIc were examined for MI and TILs. They had significantly lower MI (*A*) ($P = 0.0176$) and longer overall survival (*B*) ($P = 0.10$) than the initial cohort. When the initial and validation cohort were combined ($n = 90$) MI was a significant predictor of survival (*C*) ($P < 0.0001$). TILs were significant predictor of survival (*D*) ($P = 0.0197$) only in IIIc patients in the validation cohort.

Table S1. Patients' characteristics and stage at tissue retrieval

Sex	Age (rec)	Treatment	Status	Stage at retrieval
M	87*	Surgery, Chemo	Alive, NED	IIIB, IIIB
M	79	Surgery, XRT	Alive, MM	IIIC
M	56	Surgery	Alive, NED	IIIB
F	84	Chemo	Died, MM	IV (1b)
M	51	Surgery, XRT, Chemo	Died, MM	IIIC
M	35	Surgery, Chemo, XRT, Immuno	Died, MM	IIIB
M	86*	Surgery, Immuno, Chemo	Died, MM	IIIB, IV (1c)
M	51	Surgery, XRT	Alive, NED	IIIC
F	49	Surgery, Chemo	Alive, MM	IIIA
F	90	Surgery	Died, MM	IIIB
F	49	N/A	Died, MM	IV (1c)
M	54	Surgery, Chemo	Alive, NED	IIIC
F	54	Surgery	Alive, MM	IIIB
F	82	Surgery, Chemo	Alive, NED	IIIB
M	60	Surgery	Alive, NED	IIIB
F	68	Surgery, XRT	Alive, MM	IIIB
F	48	Surgery, XRT, Chemo	Died, MM	IIIB
M	65	N/A	Died, MM	IIIC
M	78*	Surgery	Died, MM	IIIB, IIIB
F	40*	Surgery, XRT	Died, MM	IIIB, IIIB, IIIB
F	78	Surgery, Chemo	Died, MM	IIIB
M	74*	N/A	Died, MM	IIIC, IIIC
M	69	Surgery, Immuno	Died, MM	IV (1b)
M	44	Surgery	Alive, NED	IIIB
M	43	Surgery, Immuno	Alive, NED	IIIC
M	30	Surgery, Chemo	Alive, NED	IIIA
M	60	Surgery, XRT, Immuno	Alive, NED	IIIA
M	92	Surgery	Died, MM	IIIB
M	83	Surgery	Alive, MM	IIIB
F	50	Surgery	Died, MM	IIIC
M	30	Surgery	Died, MM	IIIB
F	38	Surgery	Died, MM	IV (1b)
F	77	Surgery, XRT	Alive, NED	IIIA
M	70	Surgery, Chemo	Alive, MM	IIIB
M	52	Surgery	Died, MM	IIIC
M	79	Surgery, XRT, Chemo	Alive, NED	IIIB
M	69	Surgery	Alive, NED	IIIC
F	77	Surgery	Died, MM	IIIC

Patient sex, age at disease recurrence (rec), treatment, status at the time of analysis, and stage are represented above.
*Denotes patients with multiple samples, MM = metastatic melanoma, NED = no evidence of disease, XRT = radiation therapy, Chemo = chemotherapy.

Table S2. Functional annotation clustering using DAVID

GO term identifier	Class	Count	*P* value	Benjamini
Annotation Cluster 1 Enrichment Score: 13.42	Up-Regulated genes			
GOTERM_BP_ALL	immune system process	40	1.10E-19	5.90E-16
GOTERM_BP_ALL	immune response	31	2.50E-14	6.50E-11
GOTERM_BP_ALL	response to stimulus	44	2.00E-08	2.10E-05
	Down-regulated genes			
Annotation Cluster 1 Enrichment Score: 1.61				
SP_PIR_KEYWORDS	aminotransferase	3	4.00E-03	9.90E-01
GOTERM_MF_ALL	transaminase activity	3	9.10E-03	1.00E + 00
INTERPRO	Pyridoxal phosphate-dependent transferase, major region, subdomain 1	3	1.20E-02	1.00E + 00
GOTERM_MF_ALL	transferase activity, transferring nitrogenous groups	3	1.20E-02	1.00E + 00
GOTERM_MF_ALL	pyridoxal phosphate binding	3	2.60E-02	1.00E + 00
SP_PIR_KEYWORDS	pyridoxal phosphate	3	2.60E-02	1.00E + 00
GOTERM_MF_ALL	vitamin binding	3	1.20E-01	1.00E + 00
GOTERM_MF_ALL	cofactor binding	3	2.60E-01	1.00E + 00
Annotation Cluster 2 Enrichment Score: 1.17				
GOTERM_BP_ALL	M phase	7	1.20E-03	1.00E + 00
GOTERM_BP_ALL	Mitosis	6	2.50E-03	1.00E + 00
GOTERM_BP_ALL	M phase of mitotic cell cycle	6	2.60E-03	9.90E-01
GOTERM_BP_ALL	Cell cycle phase	7	3.60E-03	9.90E-01
GOTERM_BP_ALL	Mitotic cell cycle	6	1.00E-02	1.00E + 00
GOTERM_BP_ALL	Cell division	5	1.90E-02	1.00E + 00
SP_PIR_KEYWORDS	Mitosis	4	2.00E-02	1.00E + 00
GOTERM_BP_ALL	Cell cycle process	8	3.60E-02	1.00E + 00
SP_PIR_KEYWORDS	Cell division	4	6.10E-02	1.00E + 00
GOTERM_BP_ALL	Cell cycle	8	7.70E-02	1.00E + 00
SP_PIR_KEYWORDS	Cell cycle	5	1.10E-01	1.00E + 00
GOTERM_CC_ALL	Cytoskeletal part	5	3.60E-01	1.00E + 00
GOTERM_BP_ALL	Cytoskeleton organization and biogenesis	4	3.80E-01	1.00E + 00
GOTERM_BP_ALL	Regulation of progression through cell cycle	4	3.80E-01	1.00E + 00
GOTERM_BP_ALL	Regulation of cell cycle	4	3.80E-01	1.00E + 00
GOTERM_CC_ALL	Microtubule cytoskeleton	3	5.10E-01	1.00E + 00
GOTERM_CC_ALL	Cytoskeleton	6	5.30E-01	1.00E + 00
GOTERM_CC_ALL	Nonmembrane-bound organelle	8	7.40E-01	1.00E + 00
GOTERM_CC_ALL	Intracellular non-membrane-bound organelle	8	7.40E-01	1.00E + 00
GOTERM_BP_ALL	Organelle organization and biogenesis	5	7.40E-01	1.00E + 00

Table output from DAVID/National Institutes of Health program of the genes that are significantly associated with patient survival, representing top functional category of the upregulated genes, and the top two functional categories of the down-regulated genes.

Table S3. Estimates of survival for groups based on TIL index, CD3 cell count, MI, and TNM stage ($N = 38$)

	Group (N)	Median survival (days) and 95% CI
TIL index (% of cells within the tumor lesion, $N = 31$*)	0–25 (12)	440 (237, NA)**
	25–50 (10)	1,073 (590, NA)
	50–100 (9)	NA (887, NA)
CD3 cell count (cells per 10 HPF, $N = 29$*)	≤80 (14)	653 (440, NA)
	>80 (15)	1,073 (725, NA)
Mitotic index (mitoses per HPF, $N = 30$*)	≤0.75 (16)	1,073 (1,073, NA)
	>0.75 (14)	496 (237, NA)
TNM Stage (at time of surgery, $N = 38$)	IIIA (4)	NA
	IIIB (18)	1,073 (725, NA)
	IIIC (11)	780 (440, NA)
	IV (5)	259 (92, NA)

*Indicates the number of nonmissing values for each measurement.
**NA indicates that the estimate could not be obtained due to small sample size.

Table S4. Performance of the SVM-based classification algorithm as a function of gene selection and preprocessing methods

Genes included	Initial gene pool	PCA performed	Sensitivity, %	Specificity, %	Balanced accuracy, %	AUC, %
10	All	No	64.29	42.86	53.57	45.41
10	All	Yes	67.86	42.86	55.36	44.90
10	1.5×	No	64.29	42.86	53.57	46.68
10	1.5×	Yes	67.86	42.86	55.36	46.94
10	2×	No	67.86	57.14	62.5	67.60
10	2×	Yes	78.57	64.29	71.43	69.64
30	All	No	71.43	42.86	57.14	64.03
30	All	Yes	67.86	28.57	48.21	51.53
30	1.5×	No	71.43	42.86	57.14	69.64
30	1.5×	Yes	60.71	35.71	48.21	57.40
30	2×	No	71.43	64.29	67.86	70.15
30	2×	Yes	71.43	71.43	71.43	71.94
50	All	No	71.43	42.86	57.14	68.88
50	All	Yes	67.86	35.71	51.79	60.71
50	1.5×	No	71.43	50	60.71	71.68
50	1.5×	Yes	71.43	50	60.71	67.60
50	2×	No	75.00	64.29	69.64	75.77
50	**2×**	**Yes**	**78.57**	**71.43**	**75.00**	**81.38**
100	All	No	71.43	50.00	60.71	66.07
100	All	Yes	57.14	50.00	53.57	61.73
100	1.5×	No	67.86	50.00	58.93	70.92
100	1.5×	Yes	67.86	50.00	58.93	68.62
100	2×	No	82.14	64.29	73.21	72.45
100	2×	Yes	78.57	57.14	67.86	78.32
300	All	No	71.43	50.00	60.71	72.19
300	All	Yes	71.43	50.00	60.71	69.13
300	1.5×	No	82.14	57.14	69.64	74.74
300	1.5×	Yes	78.57	50.00	64.29	71.94
300	2×	No	71.43	64.29	67.86	71.68
300	2×	Yes	67.86	57.14	62.5	75.00
500	All	No	67.86	57.14	62.5	74.23
500	All	Yes	75.00	57.14	66.07	71.43
500	1.5×	No	67.86	57.14	62.5	70.92
500	1.5×	Yes	75.00	57.14	66.07	74.23
500	2×	No	75.00	57.14	66.07	70.41
500	2×	Yes	75.00	50.00	62.5	73.98
1000	All	No	71.43	57.14	64.29	68.62
1000	All	Yes	78.57	42.86	60.71	74.74
1000	1.5×	No	67.86	57.14	62.50	69.90
1000	1.5×	Yes	75.00	50.00	62.50	71.43
1000	2×	No	78.57	64.29	71.43	70.92
1000	2×	Yes	75.00	50.00	62.5	68.37

We varied the number of genes selected (10–1000), fold change thresholds (none, 1.5- and 2-fold difference) and whether principal component analysis (PCA) was used. The reported performance measures are sensitivity, specificity, balanced accuracy, and AUC. The highlighted combination of parameters (expression values of the top 50 genes which had at least 2-fold difference in expression, decorrelated using PCA) had the peak performance.

Table S5. *A.* Clinical characteristics PV gene expression and stage

		Low risk ($N = 21$)	High risk ($N = 23$)	Statistical test
Sex	Female	8 (38%)	8 (35%)	Fisher's P value = 1.00
	Male	13 (62%)	15 (65%)	
Age at recurrence		Mean = 61 (SD = 18)	Mean = 66 (SD = 19)	Wilcoxon rank sum $P = 0.41$
CD3 cell count	<80	9 (56%)	7 (47%)	Fisher's P value = 0.72
	≥80	7 (44%)	8 (53%)	
	Missing	5	8	
Mitotic index	<0.75	7 (44%)	9 (56%)	Fisher's P value = 0.72
	≥0.75	9 (56%)	7 (44%)	
	Missing	5	7	
TILs index	0–25%	7 (41%)	7 (44%)	Fisher's P value = 0.01
	25–50%	2 (12%)	8 (50%)	
	50–100%	8 (27%)	1 (6%)	
	Missing	4	7	
Stage at recurrence/ metastasis	IIIA	2 (9.5%)	2 (9%)	Fisher's P value = 0.67
	IIIB	12 (57%)	11 (48%)	
	IIIC	6 (28.5%)	6 (26%)	
	IV	1 (5%)	4 (17%)	
Radiation	Yes	7 (39%)	3 (14%)	Fisher's P value = 0.14
	No	11 (61%)	19 (86%)	
	Missing	3	1	
Immunotherapy	Yes	2 (11%)	3 (14%)	Fisher's P value = 1.00
	No	16 (89%)	19 (86%)	
	Missing	3	1	
Chemotherapy	Yes	7 (39%)	7 (32%)	Fisher's P value = 0.74
	No	11 (61%)	15 (68%)	
	Missing	3	1	

Clinical characteristics of low and high risk groups predicted based on a Cox multivariable regression model of survival since R/M as a dependent variable and the PV and TNM stage as predictors ($N = 44$).

Table S5. *B.* Clinical characteristics PV gene expression and MI

		Low risk ($N = 16$)	High risk ($N = 16$)	Statistical test
Sex	Female	6 (38%)	6 (38%)	Fisher's P value = 1.00
	Male	10 (62%)	10 (62%)	
Age at recurrence		Mean = 65 (SD = 20)	Mean = 59 (SD = 20)	Wilcoxon rank sum $P = 0.31$
CD3 cell count	<80	7 (44%)	9 (64%)	Fisher's P value = 0.30
	≥80	9 (56%)	5 (36%)	
	Missing	0	2	
Mitotic index	<0.75	11 (69%)	5 (31%)	Fisher's P value = 0.08
	≥0.75	5 (31%)	11 (69%)	
TILs index	0–25%	5 (31%)	9 (56%)	Fisher's P value = 0.06
	25–50%	4 (25%)	6 (38%)	
	50–100%	7 (44%)	1 (6%)	
Stage at recurrence/ metastasis	IIIA	2 (12.5%)	0 (0%)	Fisher's P value = 0.09
	IIIB	10 (62.5%)	7 (44%)	
	IIIC	4 (25%)	5 (31%)	
	IV	0 (0%)	4 (25%)	
Radiation	Yes	4 (27%)	3 (20%)	Fisher's P value = 1.00
	No	11 (73%)	12 (80%)	
	Missing	1	1	
Immunotherapy	Yes	0 (0%)	2 (13%)	Fisher's P value = 0.48
	No	15 (100%)	13 (87%)	
	Missing	1	1	
Chemotherapy	Yes	8 (53%)	4 (27%)	Fisher's P-Value = 0.26
	No	7 (47%)	11 (73%)	
	Missing	1	1	

Clinical characteristics of low and high risk groups predicted based on a Cox multivariable regression model of survival since R/M as a dependent variable and the PV and MI as predictors ($N = 32$).

Other Supporting Information Files

Dataset S1 (PDF)