Supplementary Table S1. Taqman Assays

Symbol	Gene Name	Assay ID [a]	Amplicon Size (bp)
ERBB2	human epidermal growth factor receptor 2 (HER-2)	Hs01001580_m1	60
ESR1	estrogen receptor 1	Hs00174860_m1	62
HMBS	hydroxymethylbilane synthase	Hs00609297_m1	64
IPO8	importin 8	Hs00914053_m1	67
JAK2	janus kinase 2, exon 8-9	Hs01078136_m1	84
JAK2	janus kinase 2, exon 23-24	Hs00234567_m1	101
PGR	progesterone receptor	Hs01556702_m1	77
RPLP0	ribosomal protein, large, P0	Hs99999902_m1	105
TFRC	transferrin receptor (p90, CD71)	Hs99999911_m1	105

[a] Applied Biosystems

Supplementary Table S2. Characteristics of Patients by Recurrence Status

Variable*	Value	No Recurrence	Recurrence	P-Value
Recurrence	#	112	112	
Age at Diagnosis	Mean(sd)	51(11)	50(12)	0.47
Year Collected	Mean(sd)	1998(5)	1997(5)	0.06
Age > 50	Yes	56 (50.0%)	51 (45.5%)	0.59
Ever Pregnant	No	24 (21.4%)	25 (22.3%)	
	Unknown	0 (0.0%)	2 (1.8%)	
	Yes	88 (78.6%)	85 (75.9%)	0.53
Age at First Pregnancy	Mean(sd)	26(6)	27(6)	0.61
Pregnant at Diagnosis	Yes	0 (0.0%)	2 (1.8%)	0.50
Ductal Carcinoma In Situ	Yes	88 (78.6%)	94 (83.9%)	0.39
Lobular Carcinoma In Situ	Yes	17 (15.2%)	10 (8.9%)	0.22
Extra Nodal Extensions	Yes	33 (29.5%)	41 (36.6%)	0.32
Multifocal	Yes	25 (22.3%)	27 (24.1%)	0.87
Histologic Grade	Grade 1	5 (4.5%)	0 (0.0%)	
	Grade 2	19 (17.0%)	21 (18.8%)	
	Grade 3	84 (75.0%)	83 (74.1%)	
	Unknown	4 (3.6%)	8 (7.1%)	0.09
Nuclear Grade	High	70 (62.5%)	76 (67.9%)	
	Intermediate	37 (33.0%)	28 (25.0%)	
	Low	3 (2.7%)	6 (5.4%)	
	Unknown	2 (1.8%)	2 (1.8%)	0.45
Lymphovascular Invasion	No/Unknown	56 (50.0%)	52 (46.4%)	
	Suspicious	11 (9.8%)	7 (6.2%)	
	Yes	45 (40.2%)	53 (47.3%)	0.45
Tumor Size	< 2cm	33 (29.5%)	39 (35.1%)	
	2 - 5 cm	62 (55.4%)	59 (53.2%)	
	> 5 cm	17 (15.2%)	13 (11.7%)	0.59
# of Lymph Nodes Examined	Mean(sd)	17(8)	18(10)	0.56
# of Positive Lymph Nodes	0	19 (17.0%)	20 (17.9%)	
	1 to 3	61 (54.5%)	45 (40.2%)	
	4 to 9	20 (17.9%)	31 (27.7%)	
	10+	12 (10.7%)	16 (14.3%)	0.15
Histology	Ductal	92 (82.1%)	94 (83.9%)	
	Lobular	10 (8.9%)	8 (7.1%)	
	Lobular/Ductal Mixed	9 (8.0%)	7 (6.2%)	
	Metaplastic Carcinoma, Not otherwise specified	1 (0.9%)	3 (2.7%)	0.71
Estrogen Receptor (ER)	Negative	24 (21.4%)	34 (30.4%)	
	Positive	88 (78.6%)	75 (67.0%)	
	Unknown	0 (0.0%)	3 (2.7%)	0.05
Progesterone Receptor (PR)	Negative	43 (38.4%)	44 (39.3%)	
	Positive	69 (61.6%)	65 (58.0%)	
	Unknown	0 (0.0%)	3 (2.7%)	0.29
Hormone Receptor (=ER or PR)	Negative	23 (20.5%)	29 (25.9%)	
	Positive	89 (79.5%)	80 (71.4%)	
	Unknown	0 (0.0%)	3 (2.7%)	0.15
HER2	Negative	56 (50.0%)	51 (45.5%)	
	Positive	11 (9.8%)	22 (19.6%)	
	Unknown	45 (40.2%)	39 (34.8%)	0.12
Triple Negative	No	90 (80.4%)	87 (77.7%)	
	Unknown	10 (8.9%)	14 (12.5%)	
	Yes	12 (10.7%)	11 (9.8%)	0.70
Tamoxifen Given	Yes	63 (56.2%)	60 (53.6%)	0.79
Herceptin Given	Yes	3 (2.7%)	5 (4.5%)	0.72
Hormone Therapy	Yes	81 (72.3%)	71 (63.4%)	0.20
Surgery Type	Partial Mastectomy	39 (34.8%)	40 (35.7%)	
	Total Mastectomy	73 (65.2%)	72 (64.3%)	1.00
Surgical Margins Positive	Yes	6 (5.4%)	4 (3.6%)	0.75
Surgical Margins Close/Positive	Yes	35 (31.2%)	33 (29.5%)	0.88
Radiation Therapy	Yes	78 (69.6%)	68 (60.7%)	0.21

*Variables used in matching are bold

Supplementary Table S3. Case-Control Study of *JAK2* mRNA Levels and Distant Breast Cancer Recurrence

Continuous Model, All Tissues [a]		JAK2-ex23/24	JAK2-ex8/9
	n	223	223
Un-adjusted Regression	Coefficient	-0.4165	-0.3712
	P-Value	0.0004	0.0007
Adjusted Regression	Coefficient	-0.3725	-0.3358
	P-Value	0.0032	0.0056
Continuous Model, Primary Tissues Only [b]			
	n	192	192
Un-adjusted Regression	Coefficient	-0.53322	-0.50955
	P-Value	0.00007	0.00006
Adjusted Regression	Coefficient	-0.44692	-0.44983
	P-Value	0.00183	0.00175
Dichotomous Median Split Model, All Tissues [c]			
	n	223	223
Un-adjusted Regression	Coefficient	-0.744	-0.596
	P-Value	0.006	0.028
Recurrence Rates	Above Median	45 (40.5%)	47 (42.3%)
	Below Median	66 (58.9%)	64 (57.1%)
Adjusted Regression	Coefficient	-0.653	-0.493
	P-Value	0.022	0.097

[a]Coefficients and p-values were obtained using logistic regression with transcript expression as a predictor of recurrence. Coefficients estimate the change in the log of the odds that an individual experienced a recurrence for every two-fold increase in trancript expression. A negative coefficient indicates that increased transcript expression is associated with decreased likelihood of recurrence.

[b]Coefficients and p-values were calculated as in (a) using only primary tissue. For individuals with both a primary and node specimen, only the primary tissue was included (n=26). Data from individuals with only a node specimen were excluded (n=31).

[c]Coefficients and p-values were obtained using logistic regression with above-median versus below-median transcript expression as a predictor of recurrence. Values for individuals with multiple specimens were averaged to one value per individual.

Supplementary Table S4. *JAK2* mRNA vs Protein Levels in Primary Breast Tumors

		AUTOMATED	IHC MANUAL						
			INVASIVE		INSITU		NORMAL		
		TOTAL							COMMENTS
Sample	*JAK2* mRNA, RQ	JAK2 ratio*	Intensity	%	Intensity	%	Intensity	%	
066N	15.5	0.0026	0	100%			+2	< 1%	
066R	12.7	0.0033	0	100%			+1	1%	
066U	9.9	0.0031			0	100%	0	100%	
066S	9.0	0.0003	+1	1%	0	100%	0	100%	
066V	7.9	0.0011	0	100%	0	100%	+1	1%	
066H	6.9	0.0038	0	100%	+1	2%	+1	2%	
066P	6.7	0.0066	0	100%			0	100%	areas of positive fibrocytes
066Q	6.5	0.1182	+3	70%	+1	60%	+1	10%	
066W	6.4	0.0002	0	100%	0	100%	0	100%	
066D	6.3	0.0491	+2	20%					
066F	1.7	0.0077	+2	10%	0	100%	0	100%	
066L	1.6	0.0489	+3	90%			+1	20%	
066B	1.4	0.0104	+2	20%			+1	1%	
066M	1.4	0.0002	0	100%	0	100%	0	100%	
066E	1.3	0.0108	+2	20%					
066I	0.9	0.0882	+3	80%			+1	1%	
066J	0.8	0.0257	+2	40%					
066G	0.4	0.0129	+2	10%	0	100%			
066K	0.3	0.0086					+2	20%	fibroadoma/phyllodes
066T	0.2	0.0024	+1	2%			0	100%	
Correlation *JAK2* mRNA vs JAK2 IHC ratio=-0.17									

Blank cells = tissue component not present

* ratio of JAK2 staining area relative to total tissue area