Supplementary Information

Manuscript Title: Microarray analysis identifies a death from cancer signature predicting therapy failure in patients diagnosed with multiple types of cancer

Authors: Gennadi V. Glinsky, Olaga Berezovskaya, and Anna B. Glinskii

Affiliations: Sidney Kimmel Cancer Center, 10835 Altman Row, San Diego, CA 92121

Table of Contents

1. Quantitative reverse-transcription polymerase chain reaction (Q-RT-PCR) analysis protocol.

Total RNA was extracted using RNeasy mini-kit (Qiagen, Valencia, CA, USA) following the manufacturer's instructions. A measure of 1 μg (tumor samples), or 2 μg and 4 μg (independent preparations of reference cDNA samples) of total RNA was used then as a template for cDNA synthesis with SuperScript II (Invitrogen, Carlsbad, CA, USA). Q-PCR primer sequences were selected for each cDNA with the aid of Primer Express™ software (Applied Biosystems, Foster City, CA, USA). PCR amplification was performed with the gene-specific primers listed in the Table S1 (Supplement).

Q-PCR reactions and measurements were performed with the SYBR-Green and ROX as a passive reference, using the ABI 7900 HT Sequence Detection System (Applied Biosystems, Foster City, CA, USA). Conditions for the PCR were as follows: one cycle of 10 min at 95°C; 40 cycles of 0.20 min at 94° C; 0.20 min at 60°C and 0.30 min at 72°C. The results were normalized to the relative amount of expression of an endogenous control gene *GAPDH*.

Expression of messenger RNA (mRNA) for eleven genes (Table S1) and an endogenous control gene (*GAPDH*) was measured in twenty specimens of primary prostate cancer obtained from patients with documented PSA recurrence within five years after RP and patients who remained disease-free for at least five years after RP (ten patients in each group) by real-time PCR method on an ABI PRISM 7900 HT Sequence Detection System (Applied Biosystems). For each gene several sets of primers were tested and the set-up with highest amplification efficiency was selected for the assay used in this study. Specificity of the assay for mRNA measurements was confirmed by the

absence of the expected PCR products when genomic DNA was used as a template. Glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*: 5'-CCCTCAACGACCACTTTGTCA-3' and 5'- TTCCTCTTGTGCTCTTGCTGG- 3') was used as the endogenous RNA and cDNA quantity normalization control. For calibration and generation of standard curves, we used several reference cDNAs: cDNA prepared from primary in vitro cultures of normal human prostate epithelial cells (NPEC, ref. 15, 16), cDNA derived from the PC-3M human prostate carcinoma cell line (15, 16), and cDNA prepared from normal human prostate (NHP, ref. 15, 16). Expression analysis of all genes was assessed in two independent experiments using reference cDNAs to control for variations among different Q-RT-PCR experiments. Prior to statistical analysis, the normalized gene expression values were log-transformed (on a base 10 scale) similarly to the transformation of the array-based gene expression data.

2. Protocol of discovery and validation of the 11-gene *BMI-1*-pathway signature.

We hypothesized that molecular signatures associated with activation of a normal stem cells' self-renewal and/or survival program in metastatic cancer cells might be possible to detect by looking for genes manifesting concordant patterns of regulation in distant metastatic lesions and stem cells in *BMI-1*$^{+/+}$ versus *BMI-1*$^{-/-}$ genetic backgrounds. Therefore, we sought to determine whether expression profiles of transcripts activated and suppressed in prostate cancer metastases would recapitulate the expression profile of the *BMI-1*-regulated genes in neural stem cells by comparing the sets of differentially regulated genes in search for union/intersections of lists for both up- and down-regulated transcripts. Thus, according to this model the primary criterion in transcript selection process should be the concordance of changes in expression rather than a magnitude of changes (e.g., fold change). One of the predictions of this model is that transcripts of interest would be expected to have a tightly controlled "rank order" of expression within a cluster of co-regulated genes reflecting a balance of up- and down-regulated mRNAs as a desired regulatory end-point in a cell. A degree of resemblance of the transcript abundance rank order within a gene cluster between a test sample and reference standard is measured by a Pearson correlation coefficient and designated as a phenotype association index (PAI). Samples with stem cell-resembling expression profiles (stem cell-like PAI or SPAI) are expected to have positive values of Pearson correlation coefficients. Detailed prognostic signature identification and validation protocol is described below and shown in Figure 3.

Step 1. Sets of differentially regulated transcripts were independently identified for distant metastatic lesions and primary prostate tumors versus age-matched control samples in a transgenic TRAMP mouse model of metastatic prostate cancer (MTTS signature) as well as PNS (PNS signature) and CNS (CNS signature) neurospheres in *BMI-1*$^{+/+}$ versus *BMI-1*$^{-/-}$ backgrounds using the Affymetrix microarray processing and statistical analysis software package (Affymetrix MAS 5.0; MicroDB^{TM} Ver 3.0 and DMT 3.0 software) as described in Materials and Methods and previous publications (15, 16). Transcripts with negative signal intensities values in both experimental and control sets were eliminated from further consideration. At least two-fold changes of the mRNA abundance levels in experimental versus control samples for both up-regulated and down-regulated genes were required for inclusion in the lists of differentially regulated transcripts. Fold expression changes of the mRNA abundance levels for each transcript were calculated as ratios of the average intensity values for a given transcript in experimental versus control samples for both up-regulated and down-regulated genes and log10 transformed for further analysis. Thus, this analytical step defined three large parent signatures (Figure 3): MTTS signature comprising 868 up-regulated and 477 down-regulated transcripts; PNS signature comprising 885 up-regulated and 1088 down-regulated transcripts; and CNS signature comprising 769 up-regulated and 778 down-regulated transcripts.

Step 2. Sub-sets of transcripts exhibiting concordant expression changes in metastatic TRAMP tumor samples (MTTS signature) as well as PNS (PNS signature) and CNS (CNS signature) neurospheres in *BMI-1*$^{+/+}$ versus *BMI-1*$^{-/-}$ backgrounds were identified. Thus, two concordant sub-sets of transcripts were identified corresponding to

each binary comparison of metastatic TRAMP tumors and neural stem cell samples in a state of PNS and CNS neurospheres [141 up-regulated and 58 down-regulated transcripts for PNS neurospheres ($r = 0.7593$; $P < 0.0001$; Figure 2A) and 40 up-regulated and 24 down-regulated for CNS neurospheres ($r = 0.7679$; $P < 0.0001$)]. A third concordant sub-set of 27 genes comprising 15 up-regulated and 12 down-regulated transcripts was selected for intersection common for all three signatures ($r = 0.8002$; $P < 0.0001$).

Step 3. Selection of small gene clusters was performed from sub-sets of genes exhibiting concordant changes of transcript abundance behavior in metastatic TRAMP tumor samples and PNS and CNS neurospheres in *BMI-1*$^{+/+}$ versus *BMI-1*$^{-/-}$ backgrounds. Expression profiles were presented as Log10 average fold changes for each transcript and processed for visualization and Pearson correlation analysis using Microsoft Excel software. Cut-off criterion for signature reduction was set to exceed a Pearson correlation coefficient 0.95 ($P < 0.0001$). Practical considerations essential for future development of genetic diagnostic tests prompted us to select from concordant gene sets small gene expression signatures comprising transcripts with high level of expression correlation in metastatic cancer cells and stem cells. The reduction in the signature transcript number was terminated when further elimination of a transcript did not increase the value of the Pearson correlation coefficient. Using this approach a single candidate prognostic gene expression signature was selected for each intersection of the MTTS signature and parent stem cell signatures (Figure 3). Thus, three highly concordant small signatures were identified corresponding to three concordant sub-sets of genes defined in the Step 2 (a set of 11 genes comprising 8 up-regulated and 3 down-regulated transcripts for PNS neurospheres, 11-gene MTTS/PNS signature; a set of 11 genes comprising 7 up-regulated

and 4 down-regulated transcripts for CNS neurospheres, 11-gene MTTS/CNS signature; a set of 14 genes comprising 8 up-regulated and 6 down-regulated transcripts, MTTS/PNS/CNS signature).

Step 4. Identified in the Step 3 small signatures (one 11-gene signature for the PNS set, one 11-gene signature for the CNS set, and one 14-gene signature for common PNS/CNS set) were tested for metastatic phenotype discriminative power (using one mouse prostate cancer data set and one human prostate cancer data set comprising primary and metastatic tumors) and therapy outcome classification performance (using human prostate cancer therapy outcome set 1). Three identified small signatures were evaluated for their ability to discriminate metastatic and primary prostate tumors in a TRAMP mouse model of prostate cancer, clinical samples of 9 metastatic versus 23 primary prostate tumors as well as primary prostate tumors from 21 patients with distinct outcome after the therapy (8 recurrent and 13 non-recurrent samples). To assess a potential diagnostic and prognostic relevance of small signatures, we calculated a Pearson correlation coefficient for each individual tumor sample by comparing the expression profiles of individual samples to the reference expression profile in either PNS or CNS neurospheres in *BMI-1* +/+ versus *BMI-1* -/- backgrounds. Fold expression changes in individual clinical samples were calculated for each gene as a ratio of the expression value in a given sample to the "average" expression value of the gene across the entire data set of clinical samples. Negative expression values were treated as missing data. Based on expected correlation of expression profiles of identified gene clusters with stem cell-like expression profiles, we named the corresponding correlation coefficients calculated for individual samples the stem cell-resembling phenotype association indices

(SPAIs). We evaluated the prognostic power of identified small signatures based on their ability to discriminate the metastatic versus primary tumors (criterion 1) and to segregate the patients with recurrent and non-recurrent prostate tumors into distinct sub-groups (criterion 2) and selected a single best performing small signature for subsequent validation analysis (Figures 3 and 4). Based on diagnostic and prognostic classification performance, a single best performing 11-gene MTTS/PNS signature was selected for further validation analysis (Figures 3 and 4).

Step 5. To assess the incremental statistical power of the individual genetic and clinical covariates as predictors of therapy outcome and unfavorable prognosis in prostate cancer patients, we performed both univariate and multivariate Cox proportional hazard survival analyses (Table 4).

Step 6. To validate a survival prediction model based on the 11-gene MTTS/PNS signature, we tested the prognostic performance of the model in the multiple independent therapy outcome data sets representing five epithelial and five non-epithelial cancers. We divided the patients within individual cohorts into a training set, which were used for the cut-off threshold selection and to test the model, and a test set, which was used to evaluate the reproducibility of the classification performance. We utilized the training set to select the prognosis discrimination cut-off value for a signature based on highest level of statistical significance in patient's stratification into poor and good prognosis groups as determined by the log-rank test (lowest P value and highest hazard ratio in the training set). Clinical samples having the Pearson correlation coefficient at or higher the cut-off value were identified as having the poor prognosis signature. Clinical samples with the Pearson correlation coefficient below the cut-off value were identified as having the good

prognosis signature. The same discrimination cut off value was then applied to evaluate the reproducibility of the prognostic performance in the test set of patients. Lastly, we applied the model to the entire outcome set using the same cut off threshold to confirm the classification performance. The training and test sets were balanced with respect to the total number of patients, negative and positive therapy outcomes, and the length of survival. For breast cancer data set, we maintained the patients' distribution among training and test data sets described in the original publication (33). At this stage of the analysis, we did not carry out additional model training, development or optimization steps, with the exception of a prognostic cut off threshold selection in a training set. The same MTTS/PNS expression profile was consistently used throughout the study as a reference standard to quantify the Pearson correlation coefficients of the individual samples.

Step 7. We tested the model performance using various sample stratification approaches such as terrain (TRN) clustering (Figure 5), support vector machine (SVM) classification (Table S4 in the Supplement), and weighted survival score algorithm (Figures 6E and 7D). We evaluated the therapy outcome predictive power of the 11-gene model in prostate cancer setting using prognostic test based on independent method of gene expression analysis, namely quantitative reverse-transcription polymerase chain reaction (Q-RT-PCR) method (Figure 6F).

References

1. Molofsky, A.V., Pardal, R., Iwashita, T., Park, I.-K., Clarke, M.F., Morrison, S.J. *Bmi-1* dependence distinguishes neural stem cell self-renewal from progenitor proliferation. *Nature* **425**, 962-967 (2003).
2. Glinsky, G.V., Krones-Herzig, A., Glinskii, A.B., Gebauer, G. Microarray analysis of xenograft-derived cancer cell lines representing multiple experimental models of human prostate cancer. Molecular Carcinogenesis. **37:** 209-221 (2003).
3. Glinsky, G.V., Krones-Herzig, A., Glinskii, A.B. Malignancy-associated regions of transcriptional activation: gene expression profiling identifies common chromosomal regions of a recurrent transcriptional activation in human prostate, breast, ovarian, and colon cancers. Neoplasia. **5:** 218-228 (2003).
4. Glinsky, G.V., Ivanova, Y.A., Glinskii, A.B. Common malignancy-associated regions of transcriptional activation (MARTA) in human prostate, breast, ovarian, and colon cancers are targets for DNA amplification. Cancer Letters. **201:** 67-77 (2003).
5. Glinsky, G.V., Glinskii, A.B., Stephenson, A.J., Hoffmann, R.M., Gerald, W.L. Expression profiling predicts clinical outcome of prostate cancer. *J. Clinical Invest.* 113: 913-923, 2004.
6. LaTulippe, E., Satagopan, J., Smith, A., Scher, H., Scardino, P., Reuter, V., Gerald, W.L. Comprehensive gene expression analysis of prostate cancer reveals distinct transcriptional programs associated with metastasis. *Cancer Res.* **62**, 4499-4506 (2002).

7. Singh, D., Febbo, P.G., Ross, K., Jackson, D.G., Manola, C.L., Tamayo, P., Renshaw, A.A., D'Amico, A.V., Richie, J.P., Lander, E.S., Loda, M., Kantoff, P.W., Golub, T.R., Sellers, W.R. Gene expression correlates of clinical prostate cancer behavior. *Cancer Cell* **1**, 203-209 (2002).
8. Ramaswamy, S., Ross, K.N., Lander, E.S., Golub, T.R. A molecular signature of metastasis in primary solid tumors. Nature Genetics. **33:** 49-54 (2003).
9. van 't Veer, L.J., dai, H., van de Vijver, M.J., et al. Gene expression profiling predicts clinical outcome of breast cancer. Nature, 415: 530-536, 2002.
10. Bhattacharjee, A., Richards, W.G., Staunton, J., Li, C., Monti, S., Vasa, P., Ladd, C., Beheshti, J., Bueno, R., Gillette, M., Loda, M., Weber, G., Mark, E.J., Lander, E.S., Wong, W., Johnson, B.E., Golub, T.R., Sugarbaker, D.J., Meyerson, M. Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. Proc Natl Acad Sci., 98: 13790-13795 (2001).

Table S1. Primer sequences for Q-RT-PCR analysis of the mRNA expression levels of genes comprising the 11-gene *BMI-1* pathway signature

Gene name	UniGene ID	Sequence (5' - 3')	Amplicon, bp
Gbx2-F	Hs.184945	AAGGCTTCCTGGCCAAAGAG	104
Gbx2-R		TGACTCGTCTTTCCCTTGCC	
MKI67-F	Hs.80976	CGCAAACTCTCCTTGTACCATAAT	201
MKI67-R		ATAGCGATGTGACATGTGCTTG	
CCNB1-F	Hs.23960	TGCAGCAGGAGCTTTTTGCT	119
CCNB1-R		CCAGGTGCTGCATAACTGGAA	
BUB1-F	Hs.287472	ACACCATTCCACAAGCTTCCA	123
BUB1-R		TGAAGGCACCACCATGTTTTC	
HEC-F	Hs.287472	TGCCAGTGAGCTTGAGTCCTT	136
HEC-R		TTCAGTCGTGGTTTGCACAAC	
KIAA1063-F	Hs.12064	TCAAGTGTGACGATGCCATCA	124
KIAA1063-R		CTGACCAGCTGCAGATAAGGCT	
HCFC1-F	Hs.83634	CCAATGGCATCGAGTCCCT	109
HCFC1-R		GTGCCCTTAATGACTCCCACATC	
RNF2-F	Hs.124186	AGTATTAGCCAGGATCAACAAGCA	104
RNF2-R		TCTTGCCTCGCTGCAGTCT	
ANK3-F	Hs.440478	CCAAGGCTTAGCCTCCATGAA	135
ANK3-R		ACTGACCGTTCGCTGTTACGAG	
FGFR2(1)-F	Hs.404081	CTCCGGCCTCTATGCTTGTACT	114
FGFR2(1)-R		CCATCGGTG TCATCCTCATCA	
FGFR2(2)-F	Hs.404081	ATAGCAGACTTTGGACTCGCCA	146
FGFR2(2)-R		CCGAAGGACCAGACATCACTCT	
CES(1)-F	Hs.499222	GGAATTTCCACACTGTCCCCTA	137
CES(1)-R		GGACTTCCACAGGAGTGACATG	

CES(2)-F	Hs.499222	TGTTCCTGGACTTGATAGCAGATG	117
CES(2)-R		AGCTTGGACGGTACTGAAACTCA	

Table S1A. Primer sequences for human *BMI-1* gene used for Q-RT-PCR analysis[1]

Gene	Orientation	Primer Sequence, 5' - 3'	Product
Human Bmi-1 outer primers	Sense	ctctgtatttcaatggaagtggaccattcc	
	Anti-sense	gtatggttcgttacctggagaccagca	
Human Bmi-1 inner primers	Sense	tcttaagtgcatcacagtcattgctgctg	359 bp
	Anti-sense	gatgtccaagttcacaagaccagaccactact	

[1]Reference: Park, I.-K., Qian, D., Kiel, M., Becker, M.W., Pihalja, M., Weissman, I.L., Morrison, S.J., Clarke, M.F. *Bmi-1* is required for maintenance of adult self-renewing haematopoietic stem cells. *Nature* **423**, 302-305 (2003).

Table S2
199_GENES BMI_1 pathway

99457_at	Cluster Incl X82786:Antigen identified by monoclonal antibody Ki 67 /cds=(345,9
160159_at	Cluster Incl X64713:M.musculus mRNA for cyclin B1 /cds=(75,1367) /gb=X6471
94200_at	Cluster Incl Z48800:Gastrulation brain homeobox 2 /cds=(421,1467) /gb=Z4880
104097_at	Cluster Incl AF002823:Budding inhibited by benzimidazoles 1 (S. cerevisiae) hor
93441_at	Cluster Incl AI595322:va04a10.y1 Mus musculus cDNA, 5 end /clone=IMAGE-7
99564_at	Cluster Incl D87908:Mus musculus mRNA for nuclear protein np95, complete cd
97960_at	Cluster Incl AW125800:UI-M-BH2.2-aql-h-05-0-UI.s1 Mus musculus cDNA, 3 er
102632_at	Cluster Incl AF062378:Mus musculus calmodulin-binding protein SHA1 (Sha1) m
93099_f_at	Cluster Incl U01063:Polo-like kinase homolog, (Drosophila) /cds=(27,1838) /gb=
99578_at	Cluster Incl U01915:Topoisomerase (DNA) II alpha /cds=(0,829) /gb=U01915 /gi
93164_at	Cluster Incl Y12783:M.musculus mRNA for ring1B, partial /cds=(0,608) /gb=Y12
103735_at	Cluster Incl M89800:Wingless-related MMTV integration site 6 /cds=(0,1094) /gb
94055_at	Cluster Incl U03184:Cortactin /cds=(130,1770) /gb=U03184 /gi=414990 /ug=Mm
103040_at	Cluster Incl AI837100:UI-M-AK0-adc-d-02-0-UI.s1 Mus musculus cDNA, 3 end /
93294_at	Cluster Incl M70642:Fibroblast inducible secreted protein /cds=(137,1183) /gb=M
100901_at	Cluster Incl U80821:Mus musculus C1 transcription factor gene, complete cds /c
160501_at	Cluster Incl Y09632:Mus musculus mRNA for rabkinesin-6 /cds=(302,2965) /gb=
92434_at	Cluster Incl U77844:Mus musculus mTRIP (mTRIP) mRNA, complete cds /cds=(
97173_f_at	Cluster Incl M27134:Histocompatibility 2, K region locus 2 /cds=(0,731) /gb=M27
104423_at	Cluster Incl AW212532:uo08d02.x1 Mus musculus cDNA, 3 end /clone=IMAGE-
98489_at	Cluster Incl AI510131:vx97a06.y1 Mus musculus cDNA, 5 end /clone=IMAGE-1
103201_at	Cluster Incl M86377:Mouse esk kinase mRNA, complete cds /cds=(37,2607) /gb
98319_at	Cluster Incl AJ000328:Mus musculus mRNA for type 2 desmoglein /cds=(0,245)
103553_at	Cluster Incl AA867646:vx12d04.r1 Mus musculus cDNA, 5 end /clone=IMAGE-1
97238_at	Cluster Incl AW209238:uo07g04.x1 Mus musculus cDNA, 3 end /clone=IMAGE-
100010_at	Cluster Incl U36340:Mus musculus CACCC-box binding protein BKLF mRNA, cc
96784_at	Cluster Incl AW123269:UI-M-BH2.1-apg-f-01-0-UI.s1 Mus musculus cDNA, 3 er
101521_at	Cluster Incl AB013819:Mus musculus mRNA for TIAP, complete cds /cds=(108,5
97411_at	Cluster Incl L11316:Ect2 oncogene /cds=(530,2746) /gb=L11316 /gi=293331 /ug
161000_i_at	Cluster Incl AA275196:vc05b12.r1 Mus musculus cDNA, 5 end /clone=IMAGE-7
95032_at	Cluster Incl AA856349:vw99g06.r1 Mus musculus cDNA, 5 end /clone=IMAGE-
99541_at	Cluster Incl AJ223293:Mus musculus mRNA for kinesin-related mitotic motor pro
102934_s_at	Cluster Incl L16926:Cell division cycle control protein 25C /cds=(481,1878) /gb=L
99149_at	Cluster Incl AI851230:UI-M-BH0-ajx-h-10-0-UI.s1 Mus musculus cDNA, 3 end /c
103539_at	Cluster Incl X55663:Cytoplasmic tyrosine kinase, Dscr28C related (Drosophila) /
161856_f_at	Cluster Incl AV059766:AV059766 Mus musculus cDNA /clone=1810060C09 /clo
96168_at	Cluster Incl AI591702:vq88c03.y1 Mus musculus cDNA, 5 end /clone=IMAGE-1
160699_at	Cluster Incl AI877184:uc56g02.r1 Mus musculus cDNA, 5 end /clone=IMAGE-1
103746_at	Cluster Incl AA797843:vy04c12.r1 Mus musculus cDNA, 5 end /clone=IMAGE-1
95549_at	Cluster Incl D13545:DNA primase, p58 subunit /cds=(144,1661) /gb=D13545 /gi=
93250_r_at	Cluster Incl X67668:High mobility group protein 2 /cds=(135,752) /gb=X67668 /g
104593_at	Cluster Incl AI849396:UI-M-AJ1-ahc-b-12-0-UI.s1 Mus musculus cDNA, 3 end /c
104322_at	Cluster Incl AI121796:uc42g10.r1 Mus musculus cDNA, 5 end /clone=IMAGE-1
100116_at	Cluster Incl AI122538:uc62f07.r1 Mus musculus cDNA, 5 end /clone=IMAGE-14
97527_at	Cluster Incl AA681998:vr45a06.s1 Mus musculus cDNA, 5 end /clone=IMAGE-1
100128_at	Cluster Incl M38724:Cell division cycle control protein 2a /cds=(24,917) /gb=M38
160973_at	Cluster Incl AA590345:vn58c04.r1 Mus musculus cDNA, 5 end /clone=IMAGE-1
102313_at	Cluster Incl L09737:GTP cyclohydrolase 1 /cds=(122,847) /gb=L09737 /gi=2933
101920_at	Cluster Incl AF036898:Mus musculus DNA polymerase epsilon small subunit mR

104716_at Cluster Incl X60367:Retinol binding protein 1, cellular /cds=UNKNOWN /gb=X60
104644_at Cluster Incl D12646:Kinesin heavy chain member 4 /cds=(731,4426) /gb=D1264
101946_at Cluster Incl AA840463:vx93d09.r1 Mus musculus cDNA, 5 end /clone=IMAGE-1
102935_at Cluster Incl U15562:Cell division cycle control protein 25C /cds=(319,1662) /gb=
103005_s_ Cluster Incl X66084:CD44 antigen /cds=(9,1976) /gb=X66084 /gi=53679 /ug=Mn
103203_f_ Cluster Incl AW259499:up28d02.x1 Mus musculus cDNA, 3 end /clone=IMAGE-
104476_at Cluster Incl U27177:Retinoblastoma-like 1 (p107) /cds=(63,3254) /gb=U27177 /g
103308_at Cluster Incl AI450597:mq87e05.x1 Mus musculus cDNA, 3 end /clone=IMAGE-5
94294_at Cluster Incl X66032:M.musculus mRNA for cyclin B2 /cds=(101,1297) /gb=X660
94088_at Cluster Incl AW228429:up22e01.y1 Mus musculus cDNA, 5 end /clone=IMAGE-
92759_at Cluster Incl U43298:Laminin, beta 3 /cds=(156,3662) /gb=U43298 /gi=1151214 /
93550_at Cluster Incl D88792:Mouse mRNA for double LIM protein-1, complete cds /cds=(
103444_at Cluster Incl AI272489:uk06e09.y1 Mus musculus cDNA, 5 end /clone=IMAGE-1
94079_at Cluster Incl X61452:M.musculus mRNA for H5 clone /cds=(35,1471) /gb=X6145
103506_f_ Cluster Incl AW228162:up14b05.x1 Mus musculus cDNA, 3 end /clone=IMAGE-
160464_s_ Cluster Incl U60593:Mus musculus cytoplasmic protein Ndr1 (Ndr1) mRNA, com
102976_at Cluster Incl U32446:Breast cancer 1 /cds=(116,5554) /gb=U32446 /gi=969171 /u
92861_i_at Cluster Incl M13805:Mouse type I epidermal keratin mRNA, clone pkSCC-50, 3
99186_at Cluster Incl X75483:Cyclin A2 /cds=(121,1389) /gb=X75483 /gi=414061 /ug=Mm
96596_at Cluster Incl U52073:Mus musculus TDD5 mRNA, complete cds /cds=(135,1280)
93918_at Cluster Incl AA673500:vp49g08.r1 Mus musculus cDNA, 3 end /clone=IMAGE-1
103069_at Cluster Incl AI853682:UI-M-BH0-ajq-h-03-0-UI.s1 Mus musculus cDNA, 3 end /c
92782_at Cluster Incl U39074:Thymopoietin /cds=(237,1595) /gb=U39074 /gi=1335840 /ug
103418_at Cluster Incl AW122092:UI-M-BH2.2-aoo-g-10-0-UI.s1 Mus musculus cDNA, 3 e
101961_at Cluster Incl U67327:Mus musculus WD40-repeat type I transmembrane protein A
AFFX-Gap Cluster Incl M32599 Mouse glyceraldehyde-3-phosphate dehydrogenase mRNA, complete c
97740_at Cluster Incl AI642662:vw01d07.x1 Mus musculus cDNA, 3 end /clone=IMAGE-1
93909_f_a Cluster Incl X04120:Intracisternal A particles, Thbd linked /cds=(0,2444) /gb=X04
98446_s_a Cluster Incl U06834:Mus musculus Balb/c eph-related receptor protein tyrosine k
92247_at Cluster Incl U67160:Mus musculus p190-B gene, complete cds /cds=(0,4505) /g
92639_at Cluster Incl U80932:Serine/threonine kinase 6 /cds=(48,1235) /gb=U80932 /gi=1
95543_at Cluster Incl AI843046:UI-M-AK1-aeq-g-04-0-UI.s1 Mus musculus cDNA, 3 end /
160894_at Cluster Incl X61800:CCAAT/enhancer binding protein (C/EBP), delta /cds=(315,
103797_at Cluster Incl AB019388:Cell division cycle 7-like 1 /cds=(86,1780) /gb=AB019388
102853_at Cluster Incl Y15128:Mus musculus mRNA for novel leucine zipper protein, Mmip
93836_at Cluster Incl AF041054:Mus musculus E1B 19K/Bcl-2-binding protein homolog (N
92735_at Cluster Incl X74266:Phospholipase A2, group IIA (platelets, synovial fluid) /cds=(
104218_s_ Cluster Incl AI507524:vl53f03.x1 Mus musculus cDNA, 3 end /clone=IMAGE-97
97421_at Cluster Incl U42385:Fibroblast growth factor inducible 16 /cds=(45,1187) /gb=U4
93228_at Cluster Incl U25691:Helicase, lymphoid specific /cds=(39,1850) /gb=U25691 /gi=
161332_f_ Cluster Incl AV114755:AV114755 Mus musculus cDNA /clone=2610036L13 /clo
101906_at Cluster Incl AA032310:MTA.G11.085.A Mus musculus cDNA, 5 end /clone=MTA
97124_at Cluster Incl U42384:Fibroblast growth factor inducible 15 /cds=(569,832) /gb=U4
96772_at Cluster Incl J04620:Mouse primase p49 subunit (priA) mRNA, complete cds /cds
100890_at Cluster Incl AI173038:uh16b06.r1 Mus musculus cDNA, 5 end /clone=IMAGE-1
95785_s_a Cluster Incl Y13361:RAB7, member RAS oncogene family, pseudogene 1 /cds=(
93023_f_a Cluster Incl M32459:Mouse histone H3 (H3.2-221) gene, complete cds /cds=(21,
160192_at Cluster Incl AF031568:Mus musculus heterogeneous nuclear ribonucleoprotein C
98761_i_at Cluster Incl L20450:Mus musculus DNA-binding protein mRNA, complete cds /c
160286_at Cluster Incl AA733594:vu73d10.r1 Mus musculus cDNA, 5 end /clone=IMAGE-1
94296_s_a Cluster Incl AF043220:Mus musculus TFII-I protein short form mRNA, alternative

100612_at Cluster Incl K02927:Ribonucleotide reductase M1 /cds=(81,2459) /gb=K02927 /g
100955_at Cluster Incl AA989957:ua55a07.r1 Mus musculus cDNA, 5 end /clone=IMAGE-1
160901_at Cluster Incl V00727:FBJ osteosarcoma oncogene /cds=(152,1294) /gb=V00727
99338_at Cluster Incl AA674798:vm74h03.s1 Mus musculus cDNA, 5 end /clone=IMAGE-
98946_at Cluster Incl AF033186:Mus musculus WSB-1 mRNA, complete cds /cds=(0,1265
93356_at Cluster Incl D26091:Mus musculus mRNA for mCDC47, complete cds /cds=(106
98982_at Cluster Incl AW046443:UI-M-BH1-akp-e-04-0-UI.s1 Mus musculus cDNA, 3 enc
98996_at Cluster Incl L29480:Mus musculus serine/threonine kinase (sak-b) mRNA, comp
160234_at Cluster Incl AI848382:UI-M-AH1-ago-f-10-0-UI.s1 Mus musculus cDNA, 3 end /c
104743_at Cluster Incl AB022100:Mus musculus mRNA for T-cadherin, complete cds /cds=
103529_at Cluster Incl AW125505:UI-M-BH2.2-aqm-e-07-0-UI.s1 Mus musculus cDNA, 3 e
160156_at Cluster Incl AA981015:vx55c11.r1 Mus musculus cDNA, 5 end /clone=IMAGE-1
92777_at Cluster Incl M32490:Insulin-like growth factor binding protein 10 /cds=(189,1328)
103821_at Cluster Incl AJ223087:Mus musculus mRNA for Cdc6-related protein /cds=(103,
160890_at Cluster Incl AW214587:uo49c11.x1 Mus musculus cDNA, 3 end /clone=IMAGE-
93908_f_at Cluster Incl X16670:Mouse RNA for type IIB intracisternal A-particle (IAP) eleme
96810_at Cluster Incl AI154017:ud58a11.r1 Mus musculus cDNA, 5 end /clone=IMAGE-1
98038_at Cluster Incl AF022465:Mus musculus high mobility group protein homolog HMG4
101781_f_a Cluster Incl V00754:Mouse gene coding for embryonic H3 histone /cds=(0,407) /
160546_at Cluster Incl AW121134:UI-M-BH2.3-aoh-e-02-0-UI.s1 Mus musculus cDNA, 3 e
104249_g_ Cluster Incl AW227650:up12b09.y1 Mus musculus cDNA, 5 end /clone=IMAGE-
96319_at Cluster Incl AW061324:UI-M-BH1-anw-e-09-0-UI.s1 Mus musculus cDNA, 3 enc
97896_r_at Cluster Incl AW125218:UI-M-BH2.1-aps-c-06-0-UI.s1 Mus musculus cDNA, 3 en
97295_at Cluster Incl AW122331:UI-M-BH2.2-aos-f-09-0-UI.s1 Mus musculus cDNA, 3 en
103385_at Cluster Incl U64033:Mus musculus Tera (Tera) mRNA, complete cds /cds=(81,9
94217_f_at Cluster Incl AC002397:Mouse chromosome 6 BAC-284H12 (Research Genetics
100023_at Cluster Incl X70472:Myeloblastosis oncogene-like 2 /cds=(142,2256) /gb=X7047
97909_at Cluster Incl AI838080:UI-M-AL0-abv-e-12-0-UI.s1 Mus musculus cDNA, 3 end /c
96072_at Cluster Incl M17516:Mouse lactate dehydrogenase A-4 (LDH-A) mRNA, complet
103525_at Cluster Incl AA981725:ua28h11.r1 Mus musculus cDNA, 5 end /clone=IMAGE-1
101069_g_ Cluster Incl AA656621:vr49a10.s1 Mus musculus cDNA, 5 end /clone=IMAGE-1
97334_at Cluster Incl AW048812:UI-M-BH1-amg-e-09-0-UI.s1 Mus musculus cDNA, 3 en
96255_at Cluster Incl AF067395:Mus musculus NIX (Nix) mRNA, nuclear gene encoding m
97948_at Cluster Incl M26391:Retinoblastoma 1 /cds=(98,2863) /gb=M26391 /gi=200452 /
103071_at Cluster Incl AI843655:UI-M-AO1-aen-h-11-0-UI.s1 Mus musculus cDNA, 3 end /
104536_at Cluster Incl U60530:MAD homolog 2, (Drosophila) /cds=(59,1462) /gb=U60530 /
93975_at Cluster Incl AI853531:UI-M-BH0-ajd-f-01-0-UI.s1 Mus musculus cDNA, 3 end /c
96606_at Cluster Incl AB025217:Mus musculus mRNA for Sid470p, complete cds /cds=(55
97870_s_at Cluster Incl AA798624:vy03d11.r1 Mus musculus cDNA, 5 end /clone=IMAGE-1
160876_at Cluster Incl X78684:B-cell receptor-associated protein 29 /cds=(55,777) /gb=X78
99632_at Cluster Incl U83902:Mus musculus mitotic checkpoint component Mad2 mRNA,
160896_at Cluster Incl D13003:Reticulocalbin /cds=(34,1011) /gb=D13003 /gi=220581 /ug=
103429_i_a Cluster Incl AW125330:UI-M-BH2.1-apy-e-12-0-UI.s1 Mus musculus cDNA, 3 en
103234_at Cluster Incl M35131:Neurofilament, heavy polypeptide /cds=(153,3371) /gb=M35
93543_f_at Cluster Incl J03952:Mouse, glutathione transferase GT8.7 mRNA, complete cds
102327_at Cluster Incl AF078705:Amine oxidase, copper containing 3 /cds=(0,2297) /gb=AF
100952_at Cluster Incl U47323:Stromal interaction molecule 1 /cds=(193,2250) /gb=U47323
101441_i_a Cluster Incl AF031127:Mus musculus inositol trisphosphate receptor type 2 (Itpr2
97487_at Cluster Incl X70296:Serine protease inhibitor 4 /cds=(148,1341) /gb=X70296 /gi=
161822_at Cluster Incl AV378993:AV378993 Mus musculus cDNA, 3 end /clone=9230001C
103743_at Cluster Incl AI851348:UI-M-BH0-akh-d-01-0-UI.s1 Mus musculus cDNA, 3 end /

101075_f_a Cluster Incl AB016592:Mus musculus mRNA for GOB-4, complete cds /cds=(51,
100443_at Cluster Incl AF031467:Branched chain aminotransferase 2, mitochondrial /cds=(
160535_at Cluster Incl X78709:Nuclear factor, erythroid derived 2,-like 1 /cds=(44,2269) /gb
96240_at Cluster Incl AI851309:UI-M-BH0-akf-h-03-0-UI.s1 Mus musculus cDNA, 3 end /c
94232_at Cluster Incl AI849928:UI-M-BG0-ahz-e-10-0-UI.s1 Mus musculus cDNA, 3 end /
101442_f_a Cluster Incl AF031127:Mus musculus inositol trisphosphate receptor type 2 (Itpr2
161522_i_a Cluster Incl AV218205:AV218205 Mus musculus cDNA, 3 end /clone=2900087E
161026_s_ Cluster Incl AB025259:Mus musculus mRNA for granuphilin-b, complete cds /cds
160516_at Cluster Incl U49351:Glucosidase, alpha, acid /cds=(112,2973) /gb=U49351 /gi=1
96057_at Cluster Incl AI647493:uk42h02.x1 Mus musculus cDNA, 3 end /clone=IMAGE-1
102338_at Cluster Incl AF031147:Mus musculus cGMP phosphodiesterase (PDE9A*1) mRI
93482_at Cluster Incl AI117835:uc41b01.r1 Mus musculus cDNA, 5 end /clone=IMAGE-1
100139_at Cluster Incl AI841733:UI-M-AL0-abo-f-12-0-UI.s1 Mus musculus cDNA, 3 end /c
103367_at Cluster Incl U18975:Ganglioside expression 2 /cds=(3,1604) /gb=U18975 /gi=94
99633_at Cluster Incl AB017608:Mus musculus gene for neurochondrin-1, complete cds /c
93928_f_at Cluster Incl AF043689:Mus musculus Mouse mammary tumor virus clone 66B en
100345_f_a Cluster Incl W65964:me10a08.r1 Mus musculus cDNA, 5 end /clone=IMAGE-38
94839_at Cluster Incl M96823:Mouse nucleobindin mRNA, complete cds /cds=(18,1385) /g
96346_at Cluster Incl AI854020:UI-M-BH0-ajm-b-01-0-UI.s1 Mus musculus cDNA, 3 end /
97511_at Cluster Incl AI846600:UI-M-AQ1-aef-e-03-0-UI.s1 Mus musculus cDNA, 3 end /
160264_s_ Cluster Incl AI836143:UI-M-AQ0-aai-c-05-0-UI.s2 Mus musculus cDNA, 3 end /c
104280_at Cluster Incl AF017255:Mus musculus persyn mRNA, complete cds /cds=(68,439
162332_f_a Cluster Incl AV091649:AV091649 Mus musculus cDNA /clone=2310065C07 /clo
103955_at Cluster Incl AW050325:UI-M-BH1-ang-b-04-0-UI.s1 Mus musculus cDNA, 3 enc
161085_r_ Cluster Incl AA682038:vu13a07.s1 Mus musculus cDNA, 3 end /clone=IMAGE-
101587_at Cluster Incl U89491:Mus musculus microsomal expoxide hydrolase (Eph1) mRN
102094_f_a Cluster Incl AI841270:UI-M-AM0-adu-d-10-0-UI.s1 Mus musculus cDNA, 3 end
93090_at Cluster Incl M23362:Fibroblast growth factor receptor 2 /cds=(0,1037) /gb=M233
96713_at Cluster Incl AF052453:3 -phosphoadenosine 5 -phosphosulfate synthase 2 /cds=
94515_at Cluster Incl AW208628:uo62a08.x1 Mus musculus cDNA, 3 end /clone=IMAGE-
102020_at Cluster Incl AB013345:Mus musculus mRNA for cTBAK, complete cds /cds=(135
102916_s_ Cluster Incl AB010266:Mus musculus Creb-rp, Tnx and Cyp21 genes for cAMP r
98477_s_a Cluster Incl L40632:Ankyrin 3, epithelial /cds=(56,5941) /gb=L40632 /gi=710548
97112_at Cluster Incl AI157017:ud10h08.r1 Mus musculus cDNA, 5 end /clone=IMAGE-1
98476_at Cluster Incl L40631:Ankyrin 3, epithelial /cds=(647,3931) /gb=L40631 /gi=71054
103235_at Cluster Incl AI848386:UI-M-AH1-ago-g-03-0-UI.s1 Mus musculus cDNA, 3 end /
97990_at Cluster Incl D85923:Myosin heavy chain 11, smooth muscle /cds=(104,6022) /gb
93351_at Cluster Incl U44389:Mus musculus NAD(+)-dependent 15-hydroxyprostaglandin
104477_at Cluster Incl AW047643:UI-M-BH1-alo-a-02-0-UI.s1 Mus musculus cDNA, 3 end
161907_s_ Cluster Incl AV375788:AV375788 Mus musculus cDNA, 3 end /clone=9130202H
98480_s_a Cluster Incl M32352:Mouse renin (Ren-1-d) gene, complete cds /cds=(0,1208) /g
101539_f_a Cluster Incl AW226939:um64a07.y1 Mus musculus cDNA, 5 end /clone=IMAGE
100882_at Cluster Incl AF003525:Defensin beta 1 /cds=(42,251) /gb=AF003525 /gi=219707
160899_at Cluster Incl X17320:Purkinje cell protein 4 /cds=(15,203) /gb=X17320 /gi=53607
98308_at Cluster Incl AJ002522:Mus musculus mRNA for myosin heavy chain 2X, partial /
103355_at Cluster Incl AI120671:ub71d01.r1 Mus musculus cDNA, 5 end /clone=IMAGE-1
94339_at Cluster Incl AI841330:UI-M-AM0-adv-b-11-0-UI.s1 Mus musculus cDNA, 3 end /
101538_i_a Cluster Incl AW226939:um64a07.y1 Mus musculus cDNA, 5 end /clone=IMAGE

20_GENES BMI_1 pathway
94200_at Cluster Incl Z48800:Gastrulation brain homeobox 2 /cds=(421,1467)
99457_at Cluster Incl X82786:Antigen identified by monoclonal antibody Ki 67
160159_at Cluster Incl X64713:M.musculus mRNA for cyclin B1 /cds=(75,1367)
99564_at Cluster Incl D87908:Mus musculus mRNA for nuclear protein np95, c
104097_at Cluster Incl AF002823:Budding inhibited by benzimidazoles 1 (S. cer
93441_at Cluster Incl AI595322:va04a10.y1 Mus musculus cDNA, 5 end /clon
97960_at Cluster Incl AW125800:UI-M-BH2.2-aql-h-05-0-UI.s1 Mus musculus
102632_at Cluster Incl AF062378:Mus musculus calmodulin-binding protein SH
100901_at Cluster Incl U80821:Mus musculus C1 transcription factor gene, com
93164_at Cluster Incl Y12783:M.musculus mRNA for ring1B, partial /cds=(0,60
103553_at Cluster Incl AA867646:vx12d04.r1 Mus musculus cDNA, 5 end /clon
160264_s_ Cluster Incl AI836143:UI-M-AQ0-aai-c-05-0-UI.s2 Mus musculus cDN
98477_s_a Cluster Incl L40632:Ankyrin 3, epithelial /cds=(56,5941) /gb=L40632
98476_at Cluster Incl L40631:Ankyrin 3, epithelial /cds=(647,3931) /gb=L4063
93090_at Cluster Incl M23362:Fibroblast growth factor receptor 2 /cds=(0,1037
96713_at Cluster Incl AF052453:3 -phosphoadenosine 5 -phosphosulfate syntl
104477_at Cluster Incl AW047643:UI-M-BH1-alo-a-02-0-UI.s1 Mus musculus cD
103355_at Cluster Incl AI120671:ub71d01.r1 Mus musculus cDNA, 5 end /clon
94339_at Cluster Incl AI841330:UI-M-AM0-adv-b-11-0-UI.s1 Mus musculus cD
101538_i_a Cluster Incl AW226939:um64a07.y1 Mus musculus cDNA, 5 end /cl

"Stemness" BMI-1pathway 11 gene-signature

GENE	UniGene (Homo sapiens)	Unigene (Mus Musculus)	GenBank	Description
Gbx2	Hs.184945	Mm.1306	Z48800	Gastrulation brain hom
KI67	Hs.80976	Mm.4078	X82786	Antigen identified by m
Cyclin B1	Hs.23960	Mm.22569	X64713	cyclin B1
BUB1	Hs.287472	Mm.2185	AF002823	Budding inhibited by b
HEC	Hs.414407	Mm.6642	AI595322	cDNA
KIAA1063	Hs.12064	Mm.30602	AW125800	cDNA
HCFC1	Hs.83634	Mm.57203	U80821	C1 transcription factor
RNF2	Hs.124186	Mm.5312	Y12783	ring1B
ANK3	Hs.440478	Mm.3526	L40632	Ankyrin 3, epithelial
FGFR2	Hs.404081	Mm.16340	M23362	Fibroblast growth factc
CES	Hs.499222	Mm.88078	AW226939	cDNA

"Stemness" BMI-1pathway 11 gene-signature

GENE	UniGene (Homo sapiens)	MG-U74A probe set	GenBank	MG-U74A probe set
Gbx2	Hs.184945	94200_at	Z48800	94200_at
KI67	Hs.80976	99457_at	X82786	99457_at
Cyclin B1	Hs.23960	160159_at	X64713	160159_at
BUB1	Hs.287472	104097_at	AF002823	104097_at
HEC	Hs.414407	93441_at	AI595322	93441_at
KIAA1063	Hs.12064	97960_at	AW125800	97960_at
HCFC1	Hs.83634	100901_at	U80821	100901_at
RNF2	Hs.124186	93164_at	Y12783	93164_at
ANK3	Hs.440478	98477_s_at	L40632	98477_s_at
FGFR2	Hs.404081	93090_at	M23362	93090_at
CES	Hs.499222	101538_i_at	AW226939	101538_i_at

"Stemness" BMI-1pathway 11 gene-signature

GENE	Log10_METS_TRAMP	Log10_PNS_BMI_1+/+vsBMI_1-/-	Pearson	P
Gbx2	1.505997382	2.31386722	0.989714348	<0.0001
KI67	1.429651363	2.935003151		
Cyclin B1	1.228849308	2.596597096		
BUB1	1.065859694	2.089905111		
HEC	1.018076064	2.041392685		
KIAA1063	0.960530233	1.886490725		
HCFC1	0.748188027	1.322219295		
RNF2	0.722222464	1.643452676		
ANK3	-0.402261382	-0.681562541		
FGFR2	-0.653815282	-0.557256137		
CES	-1.660034973	-2.698970004		

neobox 2
nonoclonal antibody Ki 67

enzimidazoles 1 (S. cerevisi

gene

or receptor 2

HG-U95Av2	HG-U133A
33688_at	210560_at
418_at	212022_s_at
34736_at	214710_s_at
41081_at	216277_at
40041_at	204162_at
39866_at	216964_at
37910_at	202473_x_at
33484_at	205215_at
36967_g_at	209442_x_at
1143_s_at	208228_s_at
37203_at	209616_s_at

CNS neurosphere-derived 11-gene BMI-1-pathway signature

Gene	UniGene (Homo sapiens)	Unigene (Mus Musculus)	GenBank
CTGF	Hs.410037	Mm.1810	M70642
VEGF	Hs.73793	Mm.282184	M95200
RBP1	Hs.101850	Mm.302504	X60367
KRT7	Hs.23881	Mm.30142	AA755126
BNIP3	Hs.79428	Mm.2159	AF041054
DDX39 (DDXL; BAT1)	Hs.311609	Mm.28222	AA754887
BGN	Hs.821	Mm.2608	X53928
ANK3	Hs.440478	Mm.235960	L40632
CIR (CBF1)	Hs.89421	Mm.18422	AI120671
SNCG	Hs.349470	Mm.282800	AF017255
FOXA1 (HNF3A)	Hs.163484	Mm.4578	U44752

11 genes	Log10_METS_TRAMP	Log10_CNS_BMI_1+/+vsBMI_1-/-	Pearson
93294_at	1.581352221	1.380211242	0.993901805
103520_at	0.833704768	0.779317142	P<0.0001
104716_at	0.740362689	0.769792659	
97920_at	0.856577858	0.753327667	
93836_at	0.630936119	0.464346967	
97832_at	0.55479191	0.413256767	
96049_at	0.597283727	0.411101455	
98477_s_at	-0.402261382	-0.579833719	
103355_at	-0.949999543	-1.556302497	
104280_at	-0.95754743	-1.633468455	
92697_at	-1.574393029	-2.222716468	

U95Av2	MG-U74A
36638_at	93294_at
36101_s_at	103520_at
38634_at	104716_at
41293_at	97920_at
38010_at	93836_at
35292_at	97832_at
38126_at	96049_at
36967_g_at	98477_s_at
37980_at	103355_at
36555_at	104280_at
37141_at	92697_at

14-gene MTSS/PNS/CNS signature

U95Av2_Affymetrix		TRAMP_METS	Log10_METS	Log10_WT_KO	Pearson
36638_at	93294_at	38.1375	1.581352221	1.342269625	0.969545
2090_i_at	103735_at	21.7375	1.337209595	1.505149978	
40957_at	103746_at	8.69375	0.939207147	0.698970004	
37056_at	103539_at	8.6375	0.93638806	0.783480084	
38634_at	104716_at	5.5	0.740362689	0.56391756	
41085_at	101920_at	5.275	0.722222464	0.567297885	
39927_at	92247_at	4.975	0.696793085	0.411728293	
38010_at	93836_at	4.275	0.630936119	0.394125526	
32747_at	96057_at	0.127693536	-0.893831088	-0.319946984	
2017_s_at	94232_at	0.127591707	-0.894177554	-0.295176984	
36555_at	104280_at	0.11026878	-0.95754743	-0.425968732	
38790_at	101587_at	0.031207334	-1.505743335	-0.501959329	
39660_at	100882_at	0.006520499	-2.185719179	-1.01302051	
39655_at	98480_s_at	0.004258264	-2.370767448	-0.903089987	

GeneBank	UniGene	UniGene	Description
X78947	Hs.75511	Hs.410037	connective tissue growth factor
H12458	EST		similar to WNT-6 PROTEIN
D63881	Hs.197803	Hs.462732	KIAA0160
D29767	Hs.89656	Hs.479670	Tec protein-tyrosine kinase
M11433	Hs.101850	Hs.529571	cellular retinol-binding protein 1
AF025840	Hs.99185	Hs.162777	DNA polymerase epsilon subunit B (DPE2)
U17032	Hs.37604	Hs.525287	p190-B (p190-B)
AF002697	Hs.79428	Hs.79428	E1B 19K/Bcl-2-binding protein Nip3
X05409	Hs.195432	Hs.436437	mitochondrial aldehyde dehydrogenase I ALDH I
M64349	Hs.523852	Hs.523852	cyclin D1
AF044311	Hs.63236	Hs.63236	gamma-synuclein
L25879	Hs.89649	Hs.89649	p53/HEH epoxide hydrolase (EPHX)
AI309115	Hs.32949	Hs.32949	defensin, beta 1
M26901	Hs.3210	Hs.3210	renin

Table S3

	Hs.184945	Observed (	Recurrence (months)
T16__tum	**0.266558**	**0**	**49**
T24__tum	**0.251275**	**0**	**54**
T46__tum	**0.232846**	**0**	**57**
T26__tum	**0.218995**	**1**	**14**
T04__tum	**0.204687**	**1**	**46**
T45__tum	**0.203994**	**1**	**6**
T17__tum	**0.155608**	**1**	**3**
T29__tumc	0.103776	0	51
T25__tumc	0.047901	0	52
T22__tumc	-0.036278	0	54
T54__tumc	-0.071442	0	51
T13__tumc	-0.074067	0	54
T23__tumc	-0.145422	1	37
T33__tumc	-0.161217	0	15
T55__tumc	-0.227092	0	66
T59__tumc	-0.298114	1	26
T60__tumc	-0.316119	0	55
T01__tumc	-0.443361	0	55
T62__tumc	-0.459042	1	30
T10__tumc	-0.488777	0	50
T57__tumc	-0.575547	1	4

	Hs.80976	Observed (	Recurrence
T46__tum	**0.38364**	**0**	**57**
T17__tum	**0.245273**	**1**	**3**
T13__tum	**0.236535**	**0**	**54**
T33__tum	**0.210746**	**0**	**15**
T22__tum	**0.169258**	**0**	**54**
T16__tum	**0.132176**	**0**	**49**
T23__tum	**0.129879**	**1**	**37**
T59__tum	**0.097624**	**1**	**26**
T45__tum	**0.041576**	**1**	**6**
T26__tum	**0.02694**	**1**	**14**
T24__tum	**-0.020181**	**0**	**54**
T04__tum	**-0.085342**	**1**	**46**
T29__tum	**-0.108236**	**0**	**51**
T57__tum	**-0.138342**	**1**	**4**
T01__tum	**-0.240406**	**0**	**55**
T55__tum	**-0.244204**	**0**	**66**
T62__tum	**-0.349782**	**1**	**30**
T54__tumc	-0.479144	0	51
T10__tumc	-0.620618	0	50
T60__tumc	-0.783935	0	55
T25__tumc	-1.58936	0	52

Gbx2

Comparison of Survival Curves

Logrank Test

Chi squar	1.677
df	1
P value	0.1953
P value si	ns
Are the si	No

Median survival

Data 1:Dat	46
Data 1:Dat	Undefined

Hazard Ratio

Ratio	2.423
95% CI of	0.5958 to 12.61

KI67

Comparison of Survival Curves

Logrank Test

Chi squar	2.572
df	1
P value	0.1088
P value si	ns
Are the si	No

Median survival

Data 1:Dat	Undefined
Data 1:Dat	Undefined

Hazard Ratio

Ratio	Undefined

e (months)		Hs.23960	Observed (	Recurrence (months)
	T17__tumc	0.413666	1	3
	T62__tumc	0.225597	1	30
	T45__tumc	0.184696	1	6
	T59__tumc	0.174083	1	26
	T10__tumc	0.156816	0	50
	T01__tumc	0.085157	0	55
	T16__tumc	0.081143	0	49
	T54__tumc	0.075869	0	51
	T60__tumc	0.010349	0	55
	T33__tumc	0.009398	0	15
	T25__tumc	-0.012553	0	52
	T04__tumc	-0.050705	1	46
	T57__tumc	-0.101653	1	4
	T46__tumc	-0.291156	0	57
	T29__tumc	-0.300765	0	51
	T23__tumc	-0.418318	1	37
	T22__tumc	-0.459711	0	54
	T24__tumc	-0.518103	0	54
	T55__tumc	-0.576324	0	66
	T13__tumc	-0.736229	0	54
	T26__tumc	-0.882887	1	14

Cyclin B1

Comparison of Survival Curves

Logrank Test

Chi squar	14.94
df	1
P value	0.0001
P value su	***
Are the su	Yes

Median survival

Data 1:Dat	16
Data 1:Dat	Undefined

Hazard Ratio

Ratio	8.986
95% CI of	9.990 to 1139

		Hs.287472	Observed (
	T17__tumc	0.46502	1
	T26__tumc	0.388444	1
	T46__tumc	0.374758	0
	T23__tumc	0.360627	1
	T54__tumc	0.219038	0
	T01__tumc	0.104	0
	T62__tumc	0.081028	1
	T04__tumc	0.056204	1
	T33__tumc	0.00569	0
	T10__tumc	-0.319602	0
	T22__tumc	-0.337499	0
	T13__tumc	-0.380299	0
	T25__tumc	-0.43476	0
	T24__tumc	-0.471432	0
	T60__tumc	-0.535201	0
	T57__tumc	-0.548704	1
	T59__tumc	-0.560286	1
	T55__tumc	-0.669095	0
	T16__tumc	-0.748276	0
	T29__tumc	-0.85433	0
	T45__tumc	-1.014544	1

BUB1

Comparison of Survival Curves

Logrank Test

Chi squar	3.547
df	1
P value	0.0596
P value su	ns
Are the su	No

Median survival

Data 1:Dat	25.5
Data 1:Dat	Undefined

Hazard Ratio

Ratio	3.599
95% CI of	0.9253 to 49.33

Recurrence (months)		Hs.414407 Observed (	Recurrence (months)		Hs.12064
3	**T46__tumc**	**0.385506**	**0**	**57**	**T01__tumc 0.167094**
14	**T60__tumc**	**0.283876**	**0**	**55**	**T22__tumc 0.141043**
57	**T45__tumc**	**0.262055**	**1**	**6**	**T46__tumc 0.131985**
37	**T10__tumc**	**0.247762**	**0**	**50**	**T33__tumc 0.107572**
51	**T29__tumc**	**0.233405**	**0**	**51**	**T13__tumc 0.091716**
55	**T17__tumc**	**0.166529**	**1**	**3**	**T16__tumc 0.086636**
30	**T04__tumc**	**0.139552**	**1**	**46**	**T54__tumc 0.070297**
46	**T54__tumc**	**0.136654**	**0**	**51**	**T55__tumc 0.058829**
15	**T59__tumc**	**0.097686**	**1**	**26**	**T25__tumc 0.055542**
50	**T57__tumc**	**-0.035622**	**1**	**4**	**T29__tumc 0.052349**
54	T25__tumc	-0.124259	0	52	**T62__tumc 0.002326**
54	T33__tumc	-0.213288	0	15	**T57__tumc -0.003177**
52	T16__tumc	-0.230795	0	49	**T04__tumc -0.004567**
54	T62__tumc	-0.244563	1	30	**T60__tumc -0.017116**
55	T22__tumc	-0.314061	0	54	**T23__tumc -0.026906**
4	T01__tumc	-0.375692	0	55	**T45__tumc -0.034396**
26	T26__tumc	-0.525101	1	14	**T24__tumc -0.048795**
66	T55__tumc	-0.580673	0	66	**T59__tumc -0.0657**
49	T24__tumc	-0.831011	0	54	**T10__tumc -0.123276**
51	T13__tumc	-0.94865	0	54	**T26__tumc -0.281693**
6	T23__tumc	-0.955132	1	37	**T17__tumc -0.856214**

	HEC	KIAA1063
Comparison of Survival Curves		Comparison of Surviva
Logrank Test		
Chi squar	1.205	10.79
df	1	1
P value	0.2722	**0.001**
P value su	ns	**
Are the su	No	Yes
Median survival		
Data 1:Dat	56	Undefined
Data 1:Dat	Undefined	30
Hazard Ratio		
Ratio	2.183	0
95% CI of	0.5400 to 8.894	0.02282 to

Observed (	Recurrence (months)	Hs.83634		Observed (	Recurrence (months)	
0	55	T45__tumc	0.299611	1	6	**T10__tume**
0	54	T46__tumc	0.269453	0	57	**T17__tume**
0	57	T01__tumc	0.256677	0	55	**T54__tume**
0	15	T60__tumc	0.212374	0	55	**T46__tume**
0	54	T16__tumc	0.190828	0	49	**T59__tume**
0	49	T54__tumc	0.190661	0	51	**T01__tume**
0	51	**T62__tume**	**0.086986**	**1**	**30**	**T26__tume**
0	66	**T55__tume**	**0.079941**	**0**	**66**	**T57__tume**
0	52	**T59__tume**	**0.075289**	**1**	**26**	**T62__tume**
0	51	**T57__tume**	**0.002521**	**1**	**4**	**T23__tume**
1	30	**T25__tume**	**-0.005265**	**0**	**52**	T24__tumc
1	4	**T22__tume**	**-0.00592**	**0**	**54**	T60__tumc
1	46	**T23__tume**	**-0.043708**	**1**	**37**	T45__tumc
0	55	**T33__tume**	**-0.065112**	**0**	**15**	T55__tumc
1	37	**T29__tume**	**-0.207738**	**0**	**51**	T22__tumc
1	6	**T04__tume**	**-0.235976**	**1**	**46**	T25__tumc
0	54	**T24__tume**	**-0.244749**	**0**	**54**	T16__tumc
1	26	**T10__tume**	**-0.260721**	**0**	**50**	T33__tumc
0	50	**T13__tume**	**-0.281238**	**0**	**54**	T04__tumc
1	14	**T17__tume**	**-0.918555**	**1**	**3**	T29__tumc
1	3	**T26__tume**	**-1.373722**	**1**	**14**	T13__tumc

HCFC1

ıl Curves	Comparison of Survival Curves		Comparisoı
	Logrank Test		Logrank Te
	Chi squar	1.416	Chi squar
	df	1	df
	P value	0.2341	P value
	P value su	ns	P value su
	Are the su	No	Are the su
	Median survival		Median sur
	Data 1:Dat	Undefined	Data 1:Dat
	Data 1:Dat	Undefined	Data 1:Dat
	Hazard Ratio		Hazard Rat
	Ratio	0.3024	Ratio
0.3850	95% CI of	0.09145 to 1.795	95% CI of

Hs.124186 Observed (	Recurrence (months)	
0.207081	**0**	**50**
0.17234	**1**	**3**
0.171502	**0**	**51**
0.160162	**0**	**57**
0.105766	**1**	**26**
0.099701	**0**	**55**
0.091704	**1**	**14**
0.090694	**1**	**4**
0.07752	**1**	**30**
0.049908	**1**	**37**
0.044131	0	54
0.035223	0	55
0.033688	1	6
0.023579	0	66
-0.076391	0	54
-0.127017	0	52
-0.131773	0	49
-0.135729	0	15
-0.150158	1	46
-0.530839	0	51
-0.873262	0	54

Hs.440478 Observed (	Recurrence (months)	
T13__tume	**0.335338**	**0**	**54**
T26__tume	**0.292844**	**1**	**14**
T62__tume	**0.237081**	**1**	**30**
T16__tume	**0.215708**	**0**	**49**
T46__tume	**0.16284**	**0**	**57**
T23__tume	**0.148412**	**1**	**37**
T25__tumc	0.128396	0	52
T04__tumc	0.083404	1	46
T29__tumc	0.067383	0	51
T01__tumc	-0.034027	0	55
T59__tumc	-0.064256	1	26
T55__tumc	-0.199653	0	66
T60__tumc	-0.255494	0	55
T17__tumc	-0.267579	1	3
T54__tumc	-0.317385	0	51
T33__tumc	-0.377424	0	15
T45__tumc	-0.424727	1	6
T24__tumc	-0.482619	0	54
T22__tumc	-0.576523	0	54
T57__tumc	-0.66847	1	4
T10__tumc	-1.161934	0	50

RNF2

n of Survival Curves

est

	3.821
	1
	0.0506
ns	
No	

vival

	33.5
Undefined	

tio

	4.286
0.9962 to 16.91	

ANK3

Comparison of Survival Curves

Logrank Test

Chi squar	0.2175
df	1
P value	0.641
P value su	ns
Are the su	No

Median survival

Data 1:Dat	51.5
Data 1:Dat	Undefined

Hazard Ratio

Ratio	1.402
95% CI of	0.3149 to 6.534

	Hs.404081	Observed (	Recurrence (months)
T55__tumc	0.312585	0	66
T23__tumc	0.306006	1	37
T54__tumc	0.268343	0	51
T13__tumc	0.265983	0	54
T25__tumc	0.212053	0	52
T26__tumc	0.188859	1	14
T22__tumc	0.11772	0	54
T59__tumc	0.034044	1	26
T17__tumc	-0.052818	1	3
T01__tumc	-0.080072	0	55
T16__tumc	-0.128002	0	49
T62__tumc	-0.149024	1	30
T33__tumc	-0.191473	0	15
T29__tumc	-0.222095	0	51
T46__tumc	-0.268141	0	57
T24__tumc	-0.502307	0	54
T45__tumc	**-0.627907**	**1**	**6**
T57__tumc	**-0.710812**	**1**	**4**
T60__tumc	**-0.776425**	**0**	**55**
T04__tumc	**-0.893361**	**1**	**46**
T10__tumc	**-0.991493**	**0**	**50**

FGFR2

Comparison of Survival Curves

Logrank Test

Chi squar	1.273
df	1
P value	0.2592
P value si	ns
Are the si	No

Median survival

Data 1:Dat	Undefined
Data 1:Dat	46

Hazard Ratio

Ratio	0.449
95% CI of	0.06887 to 2.056

	Hs.499222	Observed (	Recurrence
T46__tumc	**0.419495**	**0**	**57**
T54__tumc	**0.361282**	**0**	**51**
T45__tumc	**0.347256**	**1**	**6**
T55__tumc	**0.202781**	**0**	**66**
T25__tumc	**0.195466**	**0**	**52**
T13__tumc	**0.160339**	**0**	**54**
T33__tumc	**0.13878**	**0**	**15**
T04__tumc	**0.051301**	**1**	**46**
T60__tumc	**0.014152**	**0**	**55**
T17__tumc	**-0.033823**	**1**	**3**
T16__tumc	**-0.067775**	**0**	**49**
T29__tumc	**-0.079452**	**0**	**51**
T01__tumc	**-0.091451**	**0**	**55**
T10__tumc	**-0.21798**	**0**	**50**
T59__tumc	**-0.456791**	**1**	**26**
T26__tumc	**-0.537447**	**1**	**14**
T22__tumc	**-0.556316**	**0**	**54**
T23__tumc	**-0.635652**	**1**	**37**
T24__tumc	**-0.652648**	**0**	**54**
T57__tumc	**-0.821103**	**1**	**4**
T62__tumc	**-1.058851**	**1**	**30**

CES

Comparison of Survival Curves

Logrank Test

Chi squar	4.634
df	1
P value	**0.0314**
P value si	*
Are the si	Yes

Median survival

Data 1:Dat	Undefined
Data 1:Dat	30

Hazard Ratio

Ratio	0.2391
95% CI of	0.03860 to 0.8586

e (months)

Table S4: Support Vector Machine (SVN) Classification Using the 11-gene MTTS/PNS Signature

Support Vector Machine (SVM) classification

32 samples MSKCC prostate cancer metastases (9) and primary prostate adenocarcinomas (23) data set

SVM Mode: Training and Classification

Total Number of Experiments: 32

Positive Experiments
# of Experiments initially selected as positive examples: 9
# of Experiments classified as positive (Total Positives): 9
# of Experiments retained in positive class (True Positives): 8
# of Experiments recruited into positive class from Negatives (False Negatives): 1

Negative Experiments
# of Experiments initially selected as negative examples: 23
# of Experiments classified as negative (Total Negatives): 23
# of Experiments retained in negative class (True Negatives): 22
# of Experiments recruited into negative class from Positives (False Positives): 1

Support Vector Machine (SVM) classification

21 samples MIT prostate cancer recurrence data set

SVM Mode: Training and Classification

Total Number of Experiments: 21

Positive Experiments
# of Experiments initially selected as positive examples: 8
# of Experiments classified as positive (Total Positives): 10
# of Experiments retained in positive class (True Positives): 7
# of Experiments recruited into positive class from Negatives (False Negatives): 3

Negative Experiments
# of Experiments initially selected as negative examples: 13
# of Experiments classified as negative (Total Negatives): 11
# of Experiments retained in negative class (True Negatives): 10
# of Experiments recruited into negative class from Positives (False Positives): 1

Support Vector Machine (SVM) classification

79 samples MSKCC prostate cancer recurrence data set

SVM Mode: Training and Classification

Total Number of Experiments: 79

Positive Experiments
# of Experiments initially selected as positive examples: 37
# of Experiments classified as positive (Total Positives): 32
# of Experiments retained in positive class (True Positives): 26
# of Experiments recruited into positive class from Negatives (False Negatives): 6

Negative Experiments
# of Experiments initially selected as negative examples: 42
# of Experiments classified as negative (Total Negatives): 47
# of Experiments retained in negative class (True Negatives): 36
# of Experiments recruited into negative class from Positives (False Positives): 11

Bladder Cancer Set

31 samples (10 non-recurrent and 21 recurrent cancer)

SVM Mode: Training and Classification

Total Number of Experiments: 31

Positive Experiments
# of Experiments initially selected as positive examples: 21
# of Experiments classified as positive (Total Positives): 26
# of Experiments retained in positive class (True Positives): 20
# of Experiments recruited into positive class from Negatives (False Negatives): 6

Negative Experiments
# of Experiments initially selected as negative examples: 10
# of Experiments classified as negative (Total Negatives): 5
# of Experiments retained in negative class (True Negatives): 4
# of Experiments recruited into negative class from Positives (False Positives): 1

Breast Cancer Data Set

19 samples (12 metastatic recurrence and 7 no recurrence)

SVM Mode: Training and Classification

Total Number of Experiments: 19

Positive Experiments
# of Experiments initially selected as positive examples: 12
# of Experiments classified as positive (Total Positives): 11
# of Experiments retained in positive class (True Positives): 10
# of Experiments recruited into positive class from Negatives (False Negatives): 1

Negative Experiments
# of Experiments initially selected as negative examples: 7
# of Experiments classified as negative (Total Negatives): 8
# of Experiments retained in negative class (True Negatives): 6
# of Experiments recruited into negative class from Positives (False Positives): 2

Breast Cancer Data Set

78 samples (34 metastatic recurrence and 44 no recurrence)

SVM Mode: Training and Classification

Total Number of Experiments: 78

Positive Experiments
# of Experiments initially selected as positive examples: 34
# of Experiments classified as positive (Total Positives): 30
# of Experiments retained in positive class (True Positives): 23
# of Experiments recruited into positive class from Negatives (False Negatives): 7

Negative Experiments
# of Experiments initially selected as negative examples: 44
# of Experiments classified as negative (Total Negatives): 48
# of Experiments retained in negative class (True Negatives): 37
# of Experiments recruited into negative class from Positives (False Positives): 11

Breast Cancer Data Set

97 samples (46 metastatic recurrence and 51 no recurrence)

SVM Mode: Training and Classification

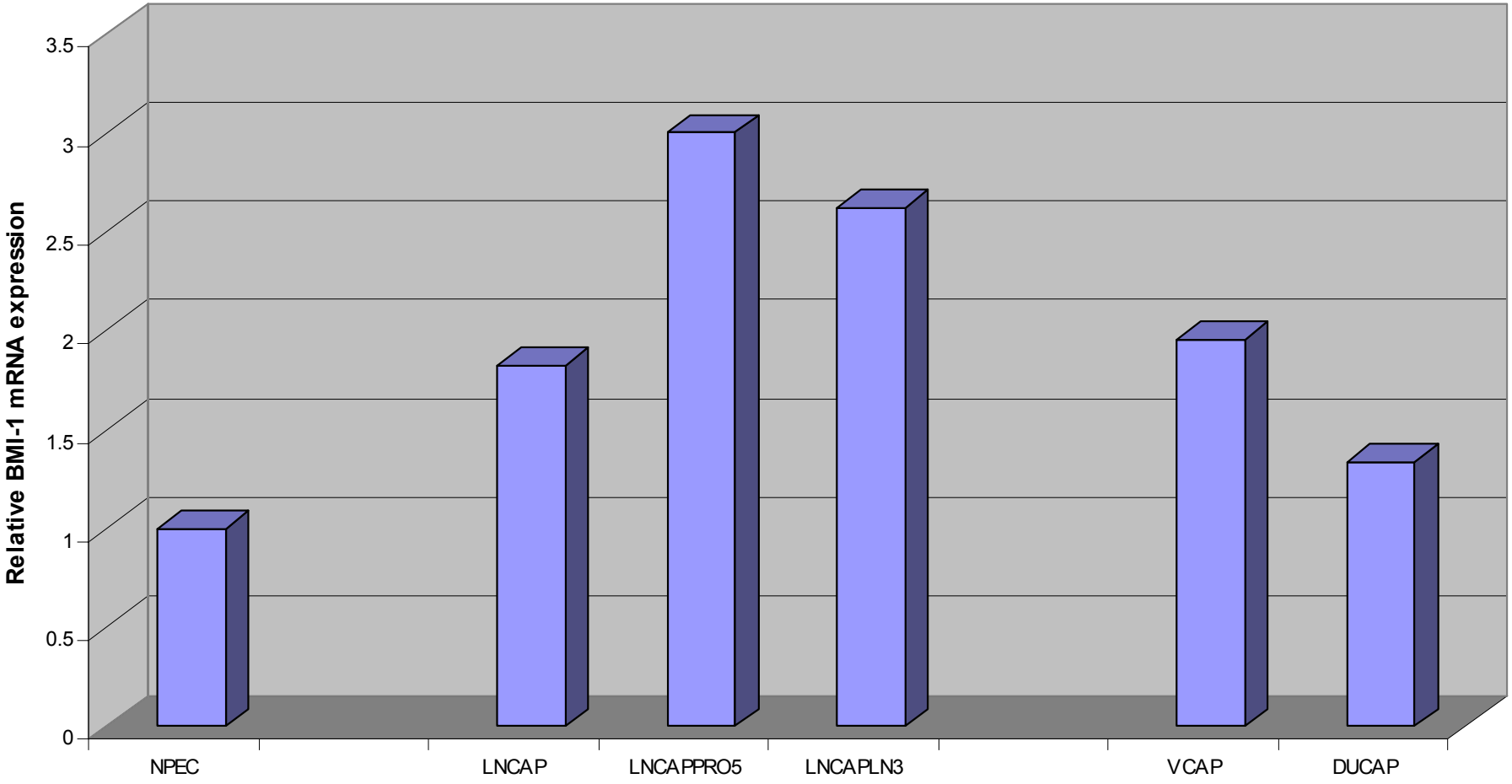
Total Number of Experiments: 97

Positive Experiments
# of Experiments initially selected as positive examples: 46
# of Experiments classified as positive (Total Positives): 46
# of Experiments retained in positive class (True Positives): 32
# of Experiments recruited into positive class from Negatives (False Negatives): 14

Negative Experiments
# of Experiments initially selected as negative examples: 51
# of Experiments classified as negative (Total Negatives): 51
# of Experiments retained in negative class (True Negatives): 37
# of Experiments recruited into negative class from Positives (False Positives): 14

Figure S1. Q-RT-PCR analysis reveals increased expression of *BMI-1* mRNA in multiple human prostate cancer cell lines established from metastatic tumors compared to normal human prostate epithelial cells. NPEC, normal human prostate epithelial cells.

A. Q-RT-PCR analysis of the BMI-1 mRNA expression in established human prostate carcinoma cell lines



B. Q-RT-PCR analysis of the BMI-1 mRNA expression in established human prostate carcinoma cell lines

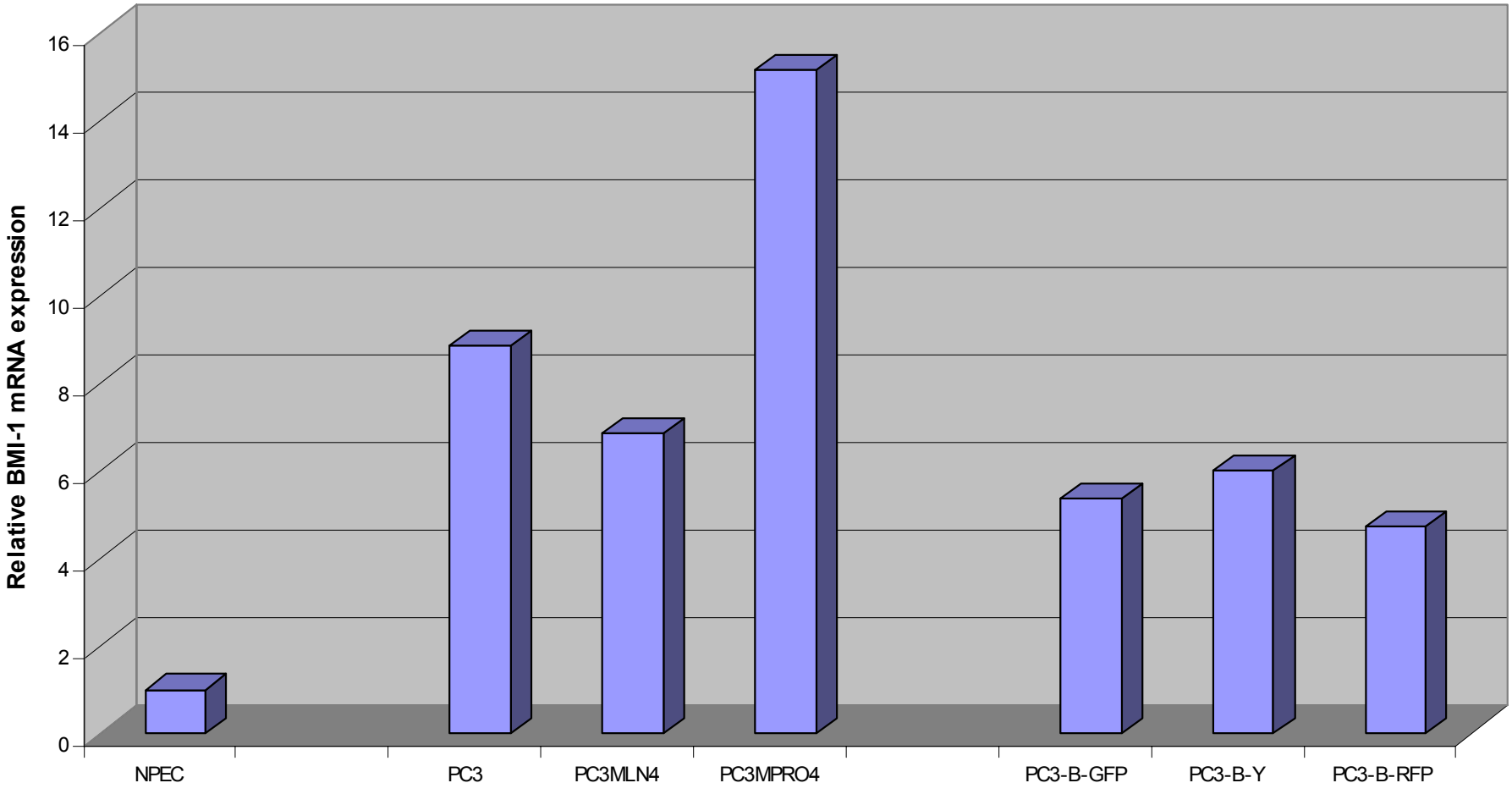
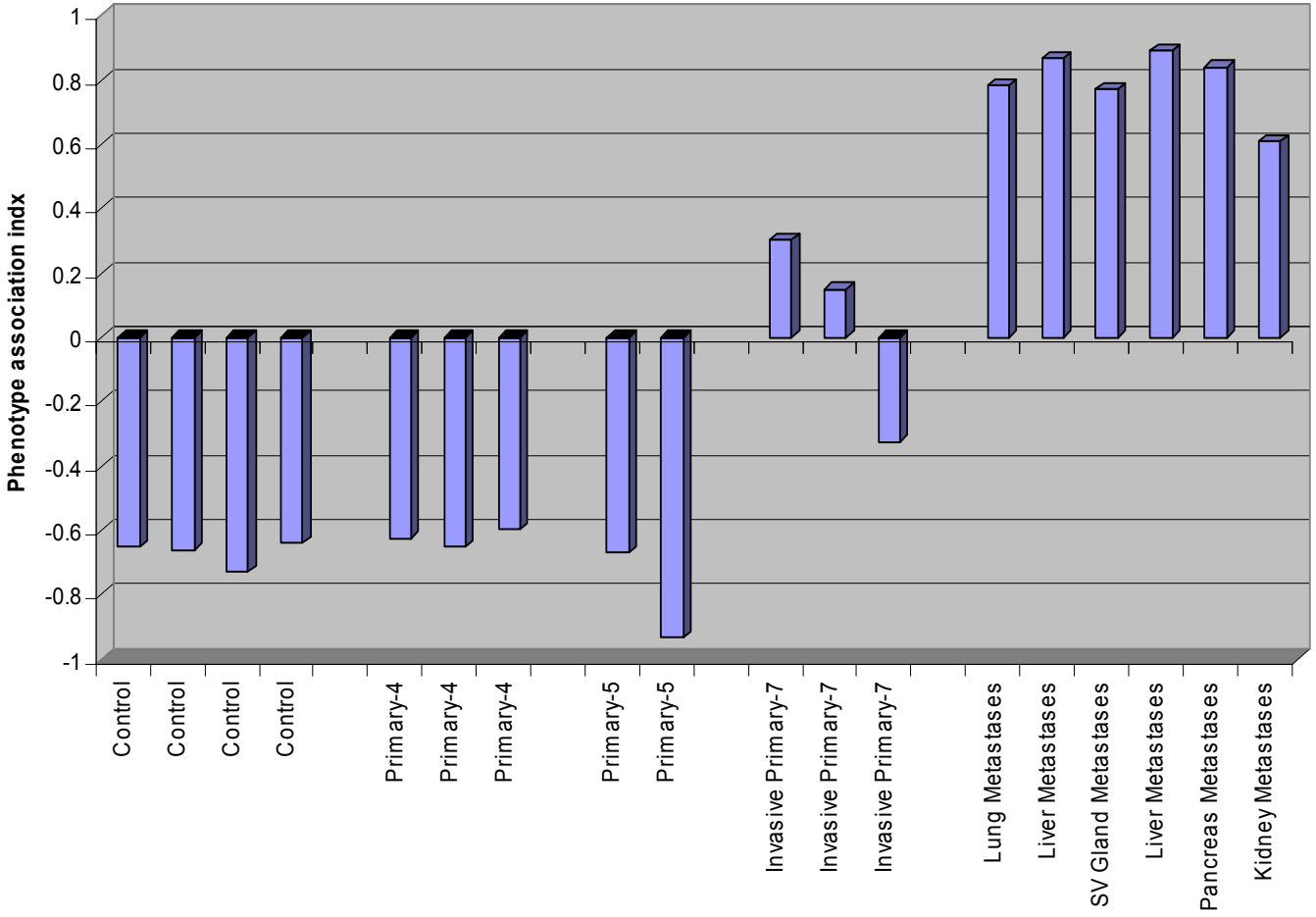
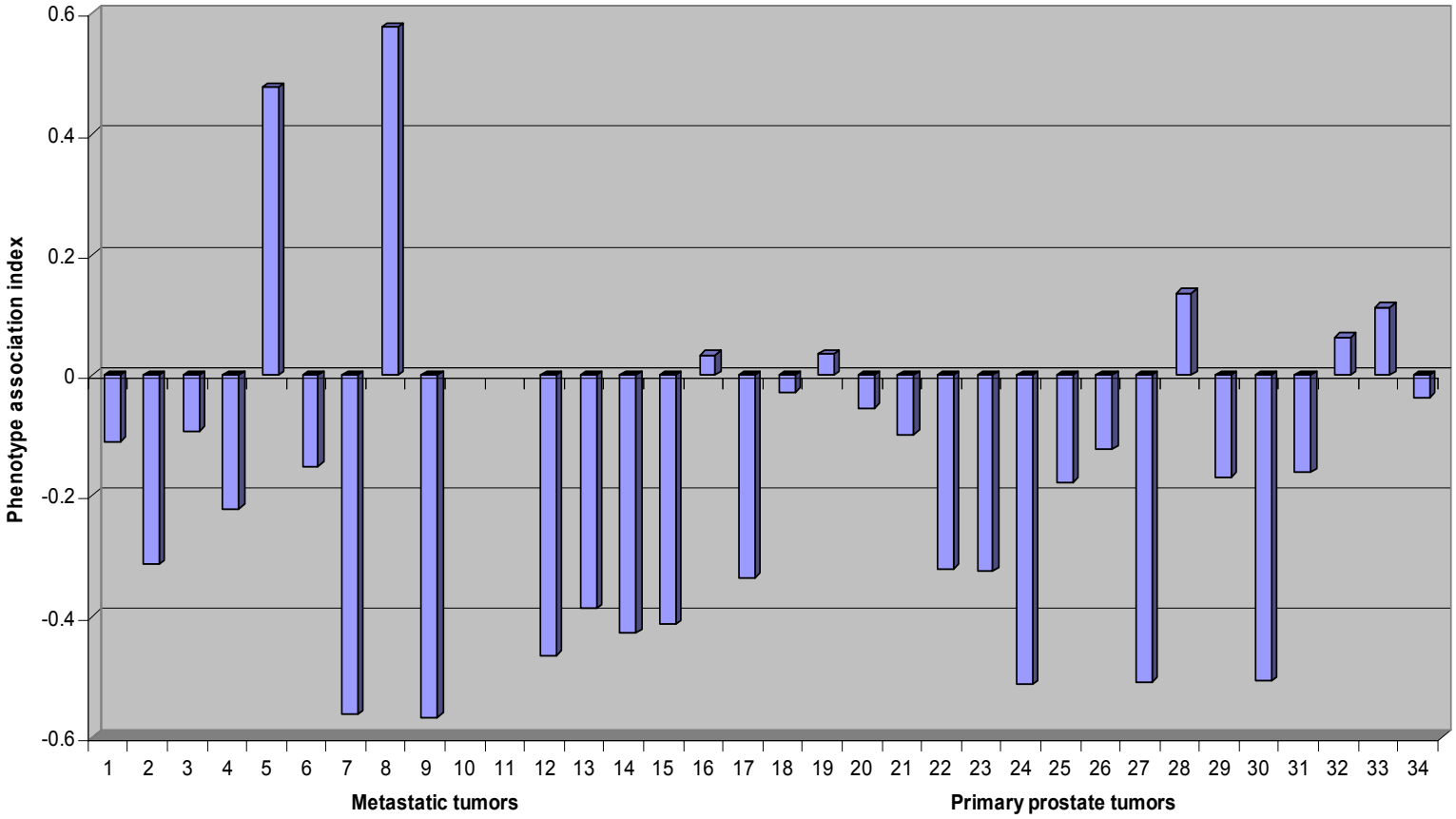


Figure S2. Validation analysis of the CNS neurosphere-derived 11-gene *BMI-1*-pathway signature

A. Expression profiles of the CNS neurosphere-derived 11-gene BMI-1-pathway signature in distant metastatic lesions and primary tumors of the TRAMP transgenic mouse model of prostate cancer



B. Expression profiles of the CNS neurosphere-derived 11-gene *BMI-1*-pathway signature in 9 metastatic lesiona and 23 primary prostate tumors



C. Relapse-free survival of prostate cancer patients with distinct expression profiles of the CNS neurosphere-derived 11-gene BMI-1-pathway signature

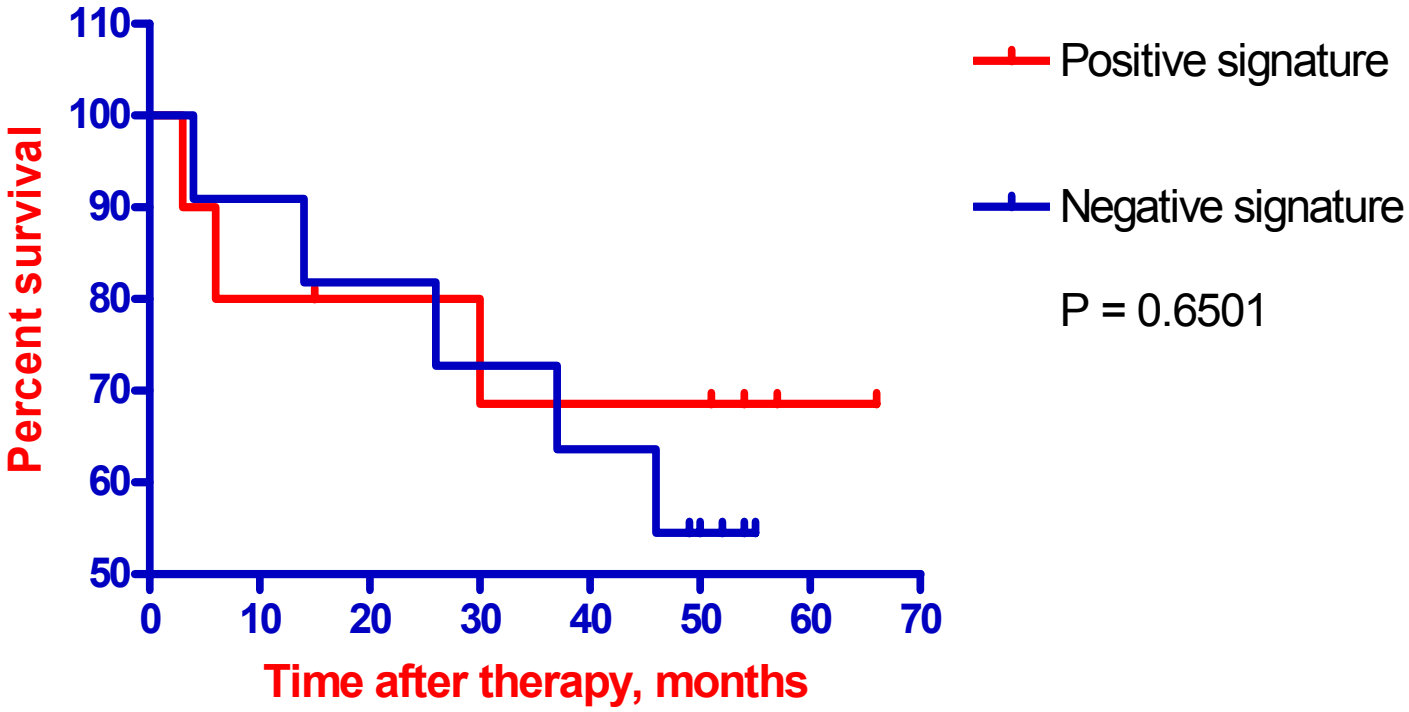
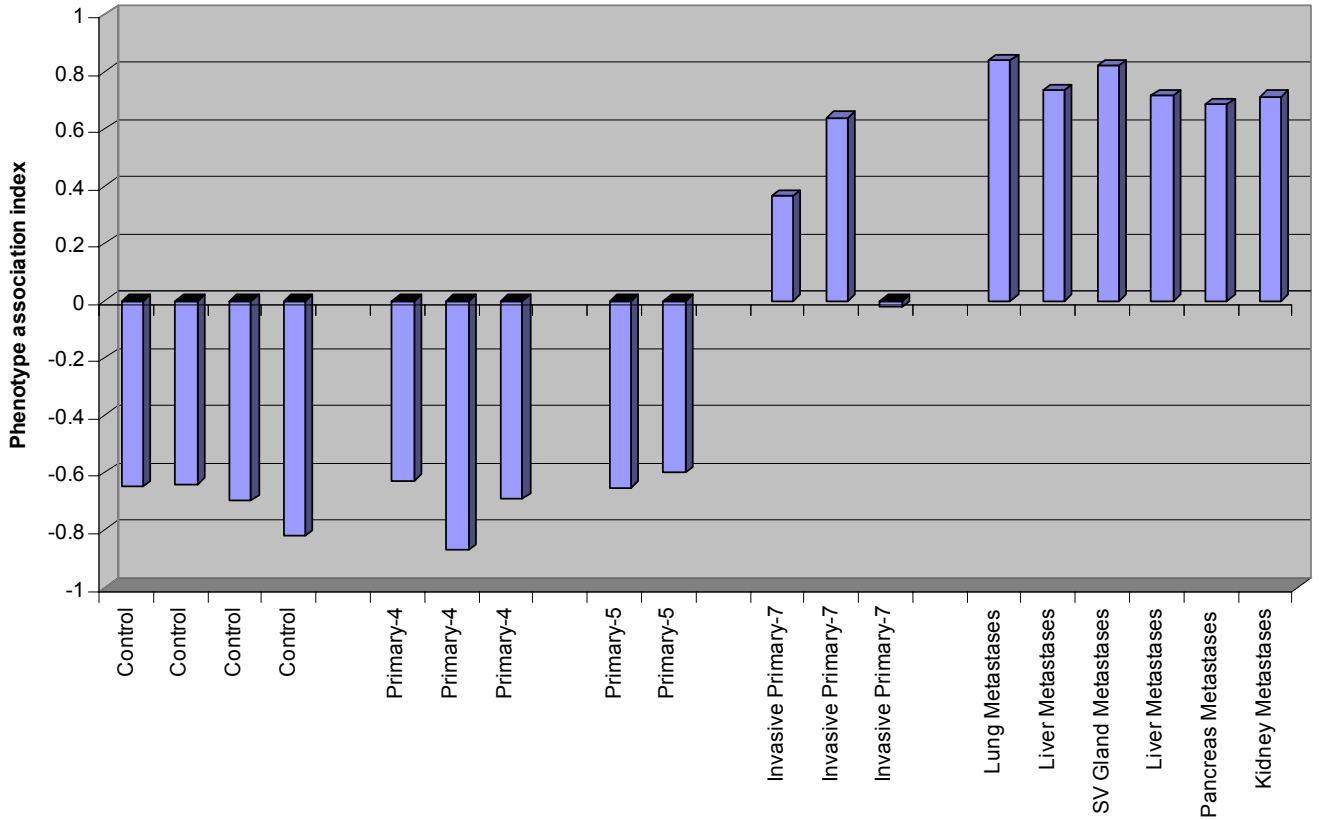
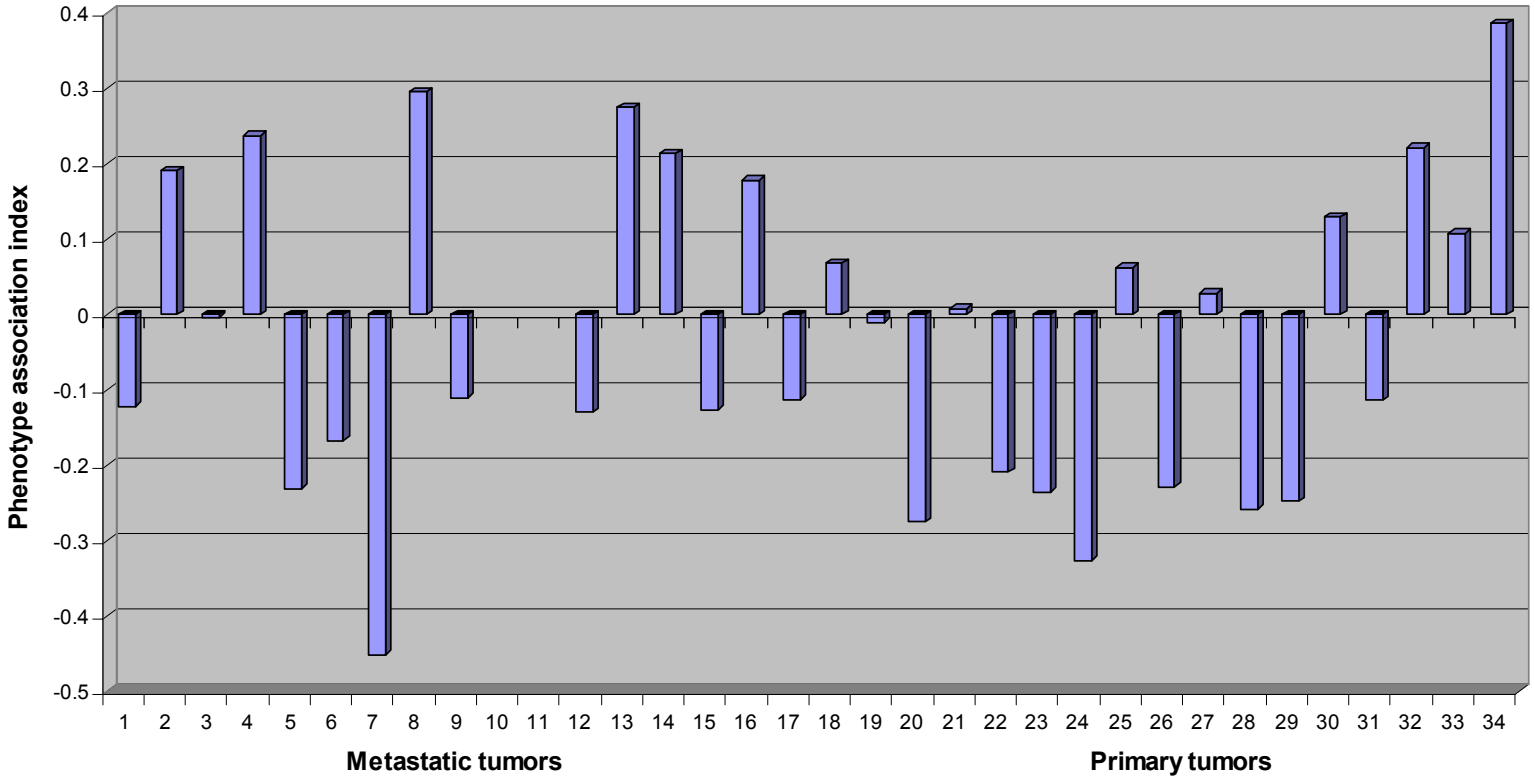


Figure S3. Validation analysis of the common PNS/CNS neurosphere-derived 14-gene *BMI-1*-pathway signature

A. Expression profiles of the common PNS and CNS neurosphere-derived 17-gene BMI-1-pathway signature in distant metastatic lesions and primary tumors of the TRAMP transgenic mouse model of prostate cancer



B. Expression profiles of the common PNS/CNS neurosphere-derived 14-gene BMI-1-pathway signature in 9 metastatic lesiona and 23 primary prostate tumors



C. Relapse-free survival of prostate cancer patients with distinct expression profiles of the common PNS/CNS neurosphere-derived 14-gene BMI-1-pathway signature