# Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to Colman H, Zhang L, Sulman EP et al. A multigene predictor of outcome in glioblastoma.

**Supplementary Appendix for Colman et al., "A multigene predictor of outcome in glioblastoma"**

I. Supplementary Methods
II. Supplementary Results

# I. Supplementary Methods

### Gene Expression Array Data Sets

The meta-analysis was based on Affymetrix gene expression array data derived from frozen samples of newly diagnosed GBM tumors from four independent data sets from individual institutions. Two of these datasets, from the University of California-San Francisco (UCSF) and the University of Texas-MD Anderson Cancer Center (MDA) were generated in the laboratory one of the authors (K.A.) Publicly available Affymetrix GeneChip data (.cel files) were obtained for data sets from the University of California-Los Angeles (UCLA) (1) and Massachusetts General Hospital (MGH).(2) The current analysis only included data from newly diagnosed GBMs with clinical follow-up data sufficient to evaluate for 2-year-survival (either deceased or alive for at least 2 years of follow-up). Samples from patients known to have a prior neurosurgical procedure were excluded.

### Mapping data between two array platforms

Because the data sets studied here involved two different platforms of microarrays (U95Av2 and U133A), extra caution was taken to map the data between the platforms. Although both platforms were developed by Affymetrix using photoliography, the selection of probe sequences followed different algorithms so that there is little overlap between the probe sets used. For our mapping, we first constructed a database of full length mRNA transcripts by merging two publicly available databases: RefSeq(3) and H-InvDB.(4) We performed BLAST searches for each of the probes used in the arrays against the database. Each matched target list was obtained from a BLAST search of a probe sequence against our library of full-length transcripts with the option of filtering the repetitive and low composite sequences turned off. We defined new probe sets by grouping probes that share the same matched target lists. Only exact matches

covering the full-length of a probe were collected in the matched target lists. Detailed information of our probe sets is stored on our web site at http://odin.mdacc.tmc.edu/~zhangli/FLTBP. Our mapping enhances the reproducibility between the two microarray platforms because it ensures that the matching probesets on the two platforms target the same genes.

**Data Normalization and Sample Quality Control**

Probe sets were mapped from the U133A and U95Av2 based on matches to full length mRNA sequences to generate a single output with genes present on both platforms, as described above. The probe signals belonging to the common probe sets were normalized using quantile normalization for each sample from every institution so that the distributions of signals on an array were the same within a platform. Log-expression values were then extracted using the PDNN model (5). The log expression values of probe sets were normalized using quantile normalization so that the distributions of log-expression on each array were the same. Because the PDNN algorithm has a tendency to compress the fold changes(6) the log-expression values were rescaled by multiplying a factor of 2 based on prior comparisons of PDNN-extracted expression values and matched PCR measurements (unpublished data). Finally, the median value within each institution for each probe set was calculated and the measurements were expressed as median ratios within that institution. The last step was found to be critical for eliminating institutional bias in the gene expression data.

Recognizing that inclusion of surrounding non-neoplastic brain tissue would have a confounding effect on the results and interpretation of the expression profiling data, we sought to eliminate samples with an apparent non-neoplastic brain "contamination". We first identified a set of five genes (gamma-aminobutyric acid receptor 5 (GABRA5), neurogranin, somatostatin, synaptotagmin I, and the light polypeptide of neurofilament protein) which we found to be highly overexpressed in non-neoplastic brain relative to malignant glioma samples using a previously published data set (7). A total of 146 cases from the four institutions fit the criteria of newly diagnosed GBM with sufficient follow-up to determine

survival at 2 years. For each of the original 146 samples we calculated a "normal brain expression index" by averaging the expression levels of these five genes. Thirty-six cases exhibited a twofold or greater normal brain expression index of relative to the median, indicating probable “contamination” of the tumor sample by excessive normal brain tissue, and these samples were excluded from subsequent analysis. The number of cases from each of the 4 institutions represented in this set of 36 samples was as follows: UCLA: 18 cases; UCSF: 7 cases; MDA: 8 cases; MGH: 3 cases. Removal of the normal brain contaminated cases left 110 tumors for analysis and a summary of the clinical information of these cases are shown in Supplementary Table 1.

**Statistical Method and Concordance of Survival Association across Institutions**

The overall experimental design used to identify robust survival associated genes from the 4 independent data sets is shown in Supplementary Figure 1. While a variety of methods have been described to identify genes with statistical or biologic significance in individual microarray data sets, it is not well established which test statistic is optimal for the purpose of determining consensus genes across independent datasets from microarray data. This issue is crucial for the identification of consensus prognostic or treatment response biomarkers as use of an inferior approach can result in lists of genes that are highly significant in one data set, but are not prognostic/predictive in independent data sets, and are thus not useful in the clinical arena. We reasoned that the method that resulted in the most consistent ranking of genes across institutions, and which performed best in cross-validation analyses, was most likely to identify a consensus gene expression profile predictive of survival in GBM.

Both fold-change and SAM 2-class analysis were applied to each of the 4 institutional data sets (MGH, MDA, UCLA and UCSF) independently, and genes were ranked from the largest (or most significant) to smallest (or least significant) difference between TS and LTS groups for each statistical method. The standard deviation of the ranks across the 4 institutions for each gene was calculated and plotted against the average rank of each gene for each statistical method (Supplementary Figure 2). This analysis

demonstrated that, in general, the most highly ranked genes showed the lowest standard deviations. We also noted that the consistency of rankings (as measured by the magnitude of the average standard deviation) was continuous as a function of the average rank, but decreased substantially after the top 200 genes (Supplementary Figure 2). It is this relationship that suggested the choice of the top 200 genes within each institution as a threshold for our subsequent analyses. Overall, gene rankings by fold-change resulted in lower standard deviations as a function of rank than when SAM p-value was used (Supplementary Figure 2) or when t-test or Rank Product was used (not shown). These observations are consistent with previous inter-institutional meta-analyses of microarray data from the Microarray Quality Control (MAQC) Project which demonstrated that fold-change was superior to p-value based significance approaches (SAM, t-test) in identifying concordance across studies due to the relatively unstable nature of the variance estimate in the t-statistic (8). Based on these considerations, fold-change was therefore used for subsequent analyses.

**Gene Expression Profiles from Microarray Data Can Predict Survival in Independent Samples of GBM**

Using a cross-validation approach, gene expression profiles from one set of GBM tumor samples identified using fold-change (based on 2-year survival) were evaluated to test whether gene expression data from microarrays could predict survival in GBM. In each round of the analysis, we utilized 3 out of the 4 institutions to form a training set to identify the top genes associated with survival. The genes from these 3 institutions were ranked by fold-change difference of TS versus LTS and the top 200 were selected. The performance of this 200-gene profile was tested using K-means clustering (9) in the remaining test set. The 2 groups defined by the K-means clustering on the test set were then compared for patient outcome. This procedure was repeated for all (n=4) possible combinations of the datasets. The results (Supplementary Figure 3) demonstrated that the survival-associated gene expression profile from the training set showed at least a statistical trend towards survival association in all 4 situations. These

data provided proof-of-principle that an outcome-associated gene expression profile obtained from one set of GBM samples could predict survival in an independent dataset. We therefore attempted to identify a consensus multigene predictor of outcome in GBM from all 4 data sets.

**Calculation of a Metagene Score**

In order to determine the association of the overall gene expression classifier with patient outcome, we calculated a single "metagene" score for each case based on the set of 38 genes by averaging the normalized expression values for all the genes associated with poor prognosis (n=31) and then subtracting the average of the normalized expression values for all the genes associated with good prognosis (n=7) for each case. This resulted in a single numerical score for each tumor, and each tumor was then ranked according to this metagene score.

**False discovery rate of 38-gene concordant set**

To determine whether these observed overlaps of 38 genes across 4 institutions was greater than those expected by chance, the survival times were scrambled and randomly assigned to individual cases, and the same analysis was performed. This analysis was repeated 5 times for graphical representation, and a representative example is shown in Figure 3B of the manuscript. We calculated the expected false discovery rates for the identification of genes common to 4 out of 4 datasets using this approach and found that that there is a 0.3% chance to find 1 common gene among the four lists by chance, and a 99.7% chance that 0 genes would be common to the 4 lists by chance. Thus, the identification of a set of 38 genes associated with survival common to all 4 institutional datasets was highly unlikely to have occurred by chance.

**Quantitative RT-PCR Measurement of Gene Expression from Paraffin Embedded Tissue**

In order to optimize amplification of the fragmented RNA found in FFPE processed tissue, primers were designed with predicted amplicon sizes of 75 base pairs or less (Applied Biosystems, Foster City, CA;

and Roche Applied Sciences, Indianapolis, IN) (Supplementary Table 2). QRT-PCR measurements were performed using a separate set of 68 FFPE GBM samples from the UT MD Anderson Brain Tumor Tissue Bank. The use of the tissue and clinical data for these studies were covered under a protocol approved by the MD Anderson IRB. Samples were examined and dissected if necessary by a neuropathologist (KA) to ensure purity of tumor tissue. RNA was isolated from these samples (Epicentre Biotechnologies, Madison, WI) following deparaffinization and proteinase K treatment. Total tumor RNA was reverse transcribed to single-stranded cDNA using ABI's High Capacity cDNA Archive kit (cat# 4368814) using the maximum allowed concentration of total RNA per manufacturer's instructions (100ng/µl). To determine fold-changes in each gene, qRT-PCR was performed on a Chromo4$^{TM}$ Real-Time PCR Detector from Bio-Rad (Hercules, CA) using the primers and probes shown in Supplementary Table 2. In triplicate, we amplified 1µl cDNA for each sample for each assay in a reaction containing 1X TaqMan$^{®}$ Universal PCR Master Mix without AmpErase UNG and 1X gene expression assay with the following cycling conditions: 10 minutes at 95°C, then 40 cycles of 95°C for 15 seconds and 60°C for 1 minute. The ΔCt values for each gene were calculated by comparison with the average of the Ct values for 2 control genes (GAPDH, GUSB) for each tumor case. To determine the survival association for each gene, the mean ΔCt for the typical survivor (TS) cases was compared with that of the long-term survivor (LTS) cases, and the ΔΔCt representing the difference of these means (TS minus LTS) was determined. Fold-change associated with survival for each gene was determined by raising 2 to the power of the ΔΔCt and taking the reciprocal of this value. Since with qRT-PCR data, a more negative value indicates higher expression, the signs of the ΔCt values were reversed to be consistent with the Affymetrix level (i.e. higher metagene score would predict worse outcome).

Since FFPE-derived RNA can be highly degraded, some quality metrics were employed. First, the RNA was analyzed by nanodrop to assess concentration. The quality metric was to determine the ability to amplify using a control gene (GUSB). This measurement was then used to guide the RT-PCR for the 38

or 9-gene set. Second, Since gene assays with very high Ct levels can mean either low expression of poor quality RNA, we examined the Ct values of the control genes, which were selected for high expression and relatively uniform expression across GBM samples. If the average Ct value of the 2 control genes was over 31, the sample was considered to have RNA quality too low for a reliable measurement. This occurred in less than 5% of the cases. Finally, when the average Ct of the control genes was within our quality metric (31 or less) the finding of particular gene assay with a Ct value of 35 or greater was considered as evidence of low expression of that gene. This occurred in a small proportion (approximately 2%) of gene assays among all the samples.

**Assessment of MGMT methylation**

The FFPE GBM samples were selected and reviewed as described above. The samples were deparaffinized and DNA was extracted using the Epicentre MasterPureTM Complete DNA Purification Kit (Epicentre Biotechnologies, Madison, WI). The extracted DNA then underwent bisulfite treatment to convert unmethylated cytosine to uracil via the Zymo Research EZ DNA Methylation-Gold KitTM. To assess the methylation status of each GBM sample methylation specific qRT-PCR was performed using the eluted bisulfite treated DNA. Methylation specific primers and probes for methylated MGMT, unmethylated MGMT, and collagen (control set) sequences used are listed in the chart below. The PCR reactions were set at 20ul volumes using 100-150ng of bisulfite treated DNA, methylation specific primers and probes, and 1X TaqMan® Universal PCR Master Mix without AmpErase. qRT-PCR was performed on a Chromo4TM Real-Time PCR Detector from Bio-Rad (Hercules, CA) using the following conditions: 10 minutes at 95°C, then 40 cycles of 95°C for 15 seconds, 55°C for 15 seconds, and 60°C for 45 seconds. The resultant curves were evaluated to determine methylation status of MGMT. Primers and probes are shown below.

| Primer/Probe | | Sequence (5'-3') |
|---|---|---|
| MGMT-Methylated | Forward | GCGTTTCGACGTTCGTAGGT |
| | Reverse | CACTCTTCCGAAAACGAAACG' |
| | Probe | DFAM-CGCAAACGATACGCACCGCGA-DTAM |
| MGMT-Unmethylated | Forward | TGTGTTTTGGATATGTTGGGATAGT |
| | Reverse | AACTCCACACTCTTCCAAAAACAA |
| | Probe | 6FAM-TTTTTGTGGTGTGTATTGTT-MGBNFQ |
| Collagen (COL2A1) | Forward | TCTAACAATTATAAACTCCAACCACCAA |
| | Reverse | GGGAAGATGGGATAGAAGGGAATAT |
| | Probe | DFAM-CCTTCATTCTAACCCAATACCTATCCCACCTCTAAA-DTAM |

**Immunohistochemistry**

Immunohistochemistry on archival FFPE samples was performed as previously described.(10) A rabbit polyclonal antibody to CD133 (Abcam) was used at a dilution of 1:250, and a mouse monoclonal antibody to nestin (Santa Cruz Biotechnology) was used at a dilution to 1:20,000. Antigen retrieval was performed by boiling slides in 10mM sodium citrate buffer, pH 6.0 followed by 30 minutes of cooling. Primary antibodies were incubated overnight in the cold. The stain was detected using the Envision kit from Dako.

**Optimization of model for outcome prediction**.

To optimize predictive models, each variable was weighted according to its individual strength of association. In concept, variables with a stronger association with outcome were weighted more heavily than variables with a lesser association. To accomplish this, the estimate of the Cox proportional hazards coefficient (β) was determined in a univariate analysis for each variable. This coefficient was then used as the weighting factor to calculate a metagene score according to the formula

$$\sum_1^i \beta_i V_i$$

where β is the Cox regression coefficient and V is the variable in question.

Variables and weighting factors are shown in the Table below. Note that of the 9 genes, 7 have positive weighting factors, indicating that elevated expression was associated with poor outcome and the remaining 2 have negative weighting factors, indicating that overexpression was associated with improved outcome.

| Variable | β | |
|---|---|---|
| | Overall survival | Progression-free survival |
| AQP1 | 0.15 | 0.10 |
| CHI3L1 | 0.12 | 0.08 |
| EMP3 | 0.13 | 0.15 |
| GPNMB | 0.20 | 0.14 |
| IGFBP2 | 0.22 | 0.18 |
| LGALS3 | 0.13 | 0.09 |
| OLIG2 | -0.15 | -0.19 |
| PDPN | 0.19 | 0.19 |
| RTN1 | -0.20 | -0.20 |

**Supplementary Tables and Figures.**

| Institution | MDA | MGH | UCLA | UCSF |
|---|---|---|---|---|
| Microarray Type | U133A | U95A | U133A | U95A |
| # of Samples | 32 | 24 | 27 | 27 |
| Typical Survivors (<2 yrs) | 20 | 17 | 19 | 21 |
| Long-Term Survivors (≥2 yrs) | 12 | 7 | 8 | 6 |

Supplementary Table 1. Clinical and Microarray Platform Characteristics.

| Gene Symbol | accession # | ABI catalog # | Roche Universal Probe # | Forward primer sequence | Reverse primer sequence |
|---|---|---|---|---|---|
| AQP1 | NM_198098.1 | Hs00166067_m1 | | | |
| CHI3L1 | NM_001276.1 | Hs01072228_m1 | | | |
| COL1A2 | NM_000089.3 | Hs00164099_m1 | | | |
| GABBR1 | NM_001470.1 | Hs00559488_m1 | | | |
| GRIA2 | NM_000826.1 | Hs00181331_m1 | | | |
| GUSB | NM_000181.2 | Hs99999908_m1 | | | |
| IGFBP2 | NM_000597.1 | Hs00167151_m1 | | | |
| IGFBP3 | NM_000598.3 | Hs00426287_m1 | | | |
| LGALS1 | NM_002305.2 | Hs00169327_m1 | | | |
| LGALS3 | NM_002306.1 | Hs00173587_m1 | | | |
| NNMT | NM_006169.1 | Hs00196287_m1 | | | |
| OLIG2 | NM_005806.1 | Hs00377820_m1 | | | |
| RIS1 | NM_015444.1 | Hs00374916_s1 | | | |
| RTN1 | NM_021136.2 | Hs00382515_m1 | | | |
| TIMP1 | NM_003254.1 | Hs00171558_m1 | | | |
| TNC | NM_002160.1 | Hs00233648_m1 | | | |
| ACTN1 | NM_001102.2 | | 42 | TGGCAGAGAAGTACCTGGACA | GGCAGTTCCAACGATGTCTT |
| CLIC1 | NM_001288.4 | | 16 | GACACCAACAAGATTGAGGAATT | GCCAGCTTGGGGTACCTG |
| EMP3 | NM_001425.1 | | 78 | GAGCGAGGGACAAGACTCC | GACATGGCTGCAGTGGAAG |
| FABP5 | NM_001444.1 | | 22 | CAAGAAAATTGAAAGATGGGAAA | CCGAGTACAGGTGACATTGTTC |
| FN1 | NM_002026.2 | | 64 | GCCACTGGAGTCTTTACCACA | CCTCGGTGTTGTAAGGTGGA |
| GAPDH | NM_002046.1 | | 9 | GGGAAGCTTGTCATCAATGG | TTGATTTTGGAGGGATCTCG |
| GPNMB | NM_001005340.1 | | 61 | TGCAAGATTGCCACTTGATG | CCCTCATGTAAGCAGAAGGTCT |
| LDHA | NM_005566.1 | | 47 | GTCCTTGGGGAACATGGAG | GACACCAGCAACATTCATTCC |
| MAOB | NM_000898.3 | | 60 | GAGAGAGCAGCCCGAGAG | GACTGCCAGATTTCATCCTC |
| OMG | NM_002544.3 | | 13 | ACGACACCACGGCTTTGATGG | CCAGGTGTGAGAAACAGAAGG |
| PDPN | NM_001006624.1 | | 20 | GGGTCCTGGCAGAAGGAG | CGCCTTCCAAACCTGTAGTC |
| PLP2 | NM_002668.1 | | 81 | GACCTGCACACCAAGATACC | CGCTATGAGGGTTCGGAAG |
| S100A10 | NM_002966.1 | | 76 | AGTTCCCTGGATTTTTGG | TGGTCCAGGTCCTTCAT |
| SERPINA3 | NM_001085.3 | | 14 | TCACAGGGGCCAGGAACCTA | TGCCCTCCTCAAATACATCAAG |
| SERPINE1 | NM_000602.1 | | 19 | AAGGCACCTCTGAGAACTTCA | CCCAGGACTAGGCAGGTG |
| SERPING1 | NM_000062.1 | | 20 | GACCCTGCTGACCCTCCT | GGAGCTGGTAGCATTTGGAT |
| TAGLN | NM_001001522.1 | | 2 | GGCCAAGGCTCTACTGTCTG | CCATGTCTGGGGAAAGCTC |
| TAGLN2 | NM_003564.1 | | 83 | CCAGCCCGCTTGAAC | CAGGCCATATGCAGGTC |
| TCF12 | NM_003205.3 | | 64 | CCCTGTACAGCAGAGATACTGGAT | AAGCCCCAGATCTTGTCTCA |
| TCTEIL | NM_006520.1 | | 76 | CAGAAGAGCGCATATGGCTT | CTTACGGTACAGGTTCCATC |
| TGFB1 | NM_000358.1 | | 5 | CTTCAAGCATCGTGTTGAGC | GACACCTTTGAGACCCTTCG |
| TMSB10 | NM_021103.2 | | 2 | CTGCCGACCAAAGAGACC | GGGTAGGAAATCCTCCAGG |
| TNR | AB007979.1 | | 6 | GACGATGCACACTTTAATTAGC | GAAGTTGGTTTTTCCTCTCC |
| VEGFA | NM_001025366.1 | | 9 | AGTGTGTGCCCACTGAGGA | GGTGAGGTTTGATCCGCATA |

Supplementary Table 2. Primers/probes used for real-time quantitative RT-PCR for FFPE GBM samples. Reagents were purchased either through the ABI “assay on demand” program (where the sequence is proprietary) or through Roche. When purchased from Roche, the primer sequence is indicated along with the probe #. Genes tested include the 38 genes identified in the microarray analysis plus 2 control genes GAPDH and GUSB).

| Gene symbol | Gene name | Present in list of top 200 survival genes | | | | Expression in poorer survivors |
|---|---|---|---|---|---|---|
| | | MDA | MGH | UCLA | UCSF | |
| C1QL1 | complement component 1, q subcomponent-like 1 | X | X | X | | lower |
| ATP5J2 | ATP synthase, mitochondrial F0 complex, subunit f | X | X | X | | higher |
| PTRF | polymerase I and transcript release factor | X | X | X | | higher |
| NCAM1 | Neural cell adhesion molecule 1 | X | X | X | | lower |
| DKFZP564 | hypothetical protein DKFZp564K0822 | X | X | X | | higher |
| SLC6A1 | solute carrier family 6 | X | X | X | | lower |
| PDGFRA | platelet-derived growth factor receptor, alpha polypeptide | X | X | X | | lower |
| TM4SF2 | transmembrane 4 superfamily member 2 | X | X | X | | lower |
| PDE8B | phosphodiesterase 8B | X | X | X | | lower |
| ALDOC | aldolase C, fructose-bisphosphate | X | X | X | | lower |
| COL3A1 | collagen, type III, alpha 1 | | X | X | X | higher |
| ID1 | inhibitor of DNA binding 1 | | X | X | X | lower |
| COL6A2 | collagen, type VI, alpha 2 | | X | X | X | higher |
| COL6A3 | collagen, type VI, alpha 3 | | X | X | X | higher |
| FABP7 | fatty acid binding protein 7, brain | | X | X | X | higher |
| S100A11 | S100 calcium binding protein A11 | | X | X | X | higher |
| IFITM1 | interferon induced transmembrane protein 1 | | X | X | X | higher |
| COL5A2 | collagen, type V, alpha 2 | | X | X | X | higher |
| PMP22 | peripheral myelin protein 22 | | X | X | X | higher |
| SPP1 | secreted phosphoprotein 1 | | X | X | X | higher |
| ACTA2 | actin, alpha 2, smooth muscle, aorta | | X | X | X | higher |
| WWTR1 | WW domain containing transcription regulator 1 | | X | X | X | higher |
| PLTP | phospholipid transfer protein | | X | X | X | higher |
| PBEF1 | pre-B-cell colony enhancing factor 1 | X | X | | X | higher |
| LTF | lactotransferrin | X | X | | X | higher |
| CHI3L2 | chitinase 3-like 2 | X | X | | X | higher |
| C1S | complement component 1, s subcomponent | X | X | | X | higher |
| CA12 | carbonic anhydrase XII | X | X | | X | higher |
| C1R | complement component 1, r subcomponent | X | X | | X | higher |
| GPR17 | G protein-coupled receptor 17 | X | X | | X | lower |
| ANXA1 | annexin A1 | X | X | | X | higher |
| COL4A1 | collagen, type IV, alpha 1 | X | X | | X | higher |
| BHLHB2 | basic helix-loop-helix domain containing, class B, 2 | X | X | | X | higher |
| VEGF | vascular endothelial growth factor | X | X | | X | higher |
| MSN | moesin | X | X | | X | higher |
| HSPB1 | heat shock 27kDa protein 1 | X | X | | X | higher |
| COL4A2 | collagen, type IV, alpha 2 | X | X | | X | higher |
| ASCL1 | achaete-scute complex-like 1 | X | X | | X | lower |
| SOD2 | superoxide dismutase 2, mitochondrial | X | X | | X | higher |
| VIM | vimentin | X | X | | X | higher |
| GPR51 | G protein-coupled receptor 51 | X | X | | X | lower |
| DPYD | dihydropyrimidine dehydrogenase | X | X | | X | higher |
| MRCL3 | myosin regulatory light chain MRCL3 | X | X | | X | higher |
| PTX3 | pentaxin-related gene, rapidly induced by IL-1 beta | X | X | | X | higher |
| SEC61G | Sec61 gamma subunit | X | | X | X | higher |
| MOXD1 | monooxygenase, DBH-like 1 | X | | X | X | higher |
| EGFR | epidermal growth factor receptor | X | | X | X | higher |
| C10orf56 | chromosome 10 open reading frame 56 | X | | X | X | lower |
| KLRC1 | killer cell lectin-like receptor subfamily C, member 1 | X | | X | X | lower |
| RIPX | rap2 interacting protein x | X | | X | X | lower |
| GJA1 | gap junction protein, alpha 1, 43kDa | X | | X | X | higher |
| SH3GL2 | SH3-domain GRB2-like 2 | X | | X | X | lower |
| MXI1 | MAX interactor 1 | X | | X | X | lower |
| TFRC | transferrin receptor | X | | X | X | higher |
| MT1E | metallothionein 1E | X | | X | X | higher |
| TCF8 | transcription factor 8 | X | | X | X | lower |
| MEOX2 | mesenchyme homeo box 2 | X | | X | X | higher |

Supplementary Table 3. Fifty-seven genes found to be associated with survival in ¾ data sets. Genes present in the list of the top 200 survival genes are shown, listing the datasets in which in each was present. The direction of the survival association (i.e. higher vs. lower expression in poor survivors) is shown.

| Gene name | Fold change (STS/LTS) | Average Ct value |
|---|---|---|
| PDPN | 4.5 | 29 |
| AQP1 | 3.0 | 28 |
| YKL40 | 2.8 | 26 |
| RTN1 | 0.4 | 30 |
| KIAA0510 | 0.4 | 35 |
| S100 | 2.1 | 36 |
| EMP3 | 2.1 | 28 |
| GPNMB | 2.1 | 27 |
| IGFBP2 | 2.0 | 28 |
| OLIG2 | 0.5 | 29 |
| LGALS3 | 2.0 | 29 |
| SERPE3 | 1.9 | 28 |
| NNMT | 1.9 | 30 |
| TNC | 1.8 | 28 |
| VEGFA | 1.8 | 31 |
| GABA | 0.6 | 34 |
| TCTEIL | 1.6 | 30 |
| MAOB | 1.5 | 33 |
| PLP2 | 0.7 | 34 |
| TAGLN2 | 1.5 | 33 |
| OMG | 0.7 | 28 |
| TGFB1 | 1.4 | 29 |
| LGALS1 | 1.4 | 29 |
| SERPG | 1.4 | 30 |
| CLIC1 | 1.4 | 30 |
| TIMP1 | 1.3 | 31 |
| ACTN1 | 1.3 | 32 |
| FABP5 | 1.2 | 31 |
| LDHA | 1.2 | 35 |
| RIS1 | 1.2 | 34 |
| TAGLN | 1.2 | 34 |
| TCF12 | 0.9 | 32 |
| GRIA2 | 0.9 | 32 |
| SERPE | 1.1 | 33 |
| TMSB10 | 0.9 | 29 |
| FN1 | 1.0 | 33 |
| IGFBP3 | 1.0 | 32 |
| COL1A2 | 1.0 | 36 |

Supplementary Table 4. Results of qRT-PCR assay on initial validation set of 68 GBM samples. The fold-change was calculated by averaging the expression level of typical (less than 2 years) versus long term (2 years or greater) survivors). The mean delta Ct value was also calculated. Genes in red (n=9) were selected for further study based on a fold-change level of 2 or above (in either direction) and a mean delta Ct value of 32 or less.

| Gene | HR | P-value |
|---|---|---|
| AQP1 | 2.3 | 0.0013 |
| CHI3L1 | 2.1 | 0.0085 |
| EMP3 | 1.6 | 0.0743 |
| GPNMB | 1.8 | 0.0055 |
| IGFBP2 | 1.9 | 0.0019 |
| LGALS3 | 1.3 | 0.0663 |
| OLIG2 | 0.6 | 0.0212 |
| PDPN | 2.3 | 0.0001 |
| RTN1 | 0.6 | 0.0042 |

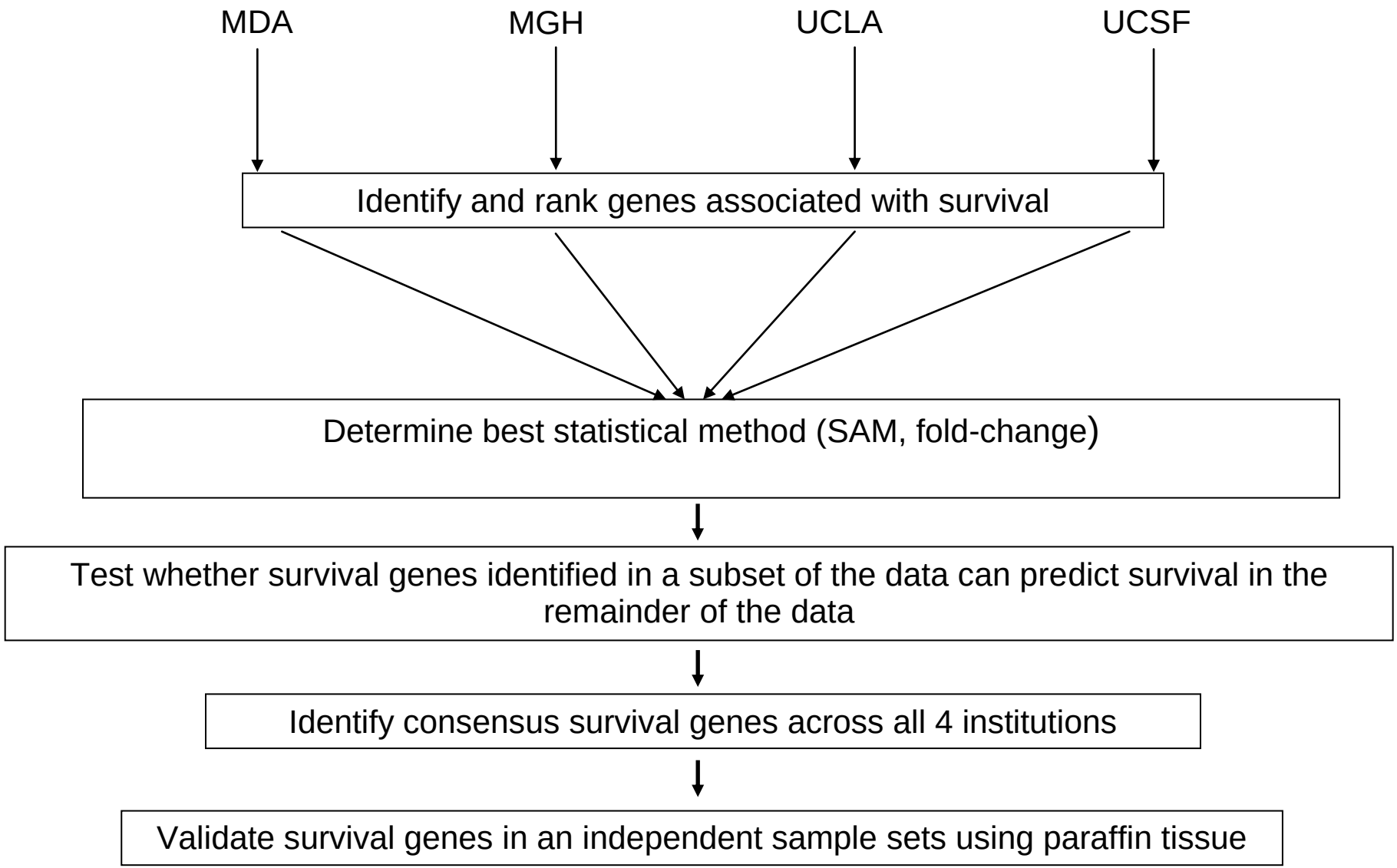
Supplementary Table 5. Results of qRT-PCR analyses for top 9 genes on 68 GBM samples. Univariate Cox analyses were performed for each gene, and the effect size (hazard ratio, or HR) and P-values are shown. To generate comparable hazard ratios, median expression values for each gene were determined and the value was expressed as 0 or 1 based on whether it was lower or higher than the median, respectively.

| Gene | HR | P-value |
|---|---|---|
| AQP1 | 1.8 | 0.0431 |
| CHI3L1 | 1.9 | 0.0477 |
| EMP3 | 1.7 | 0.0435 |
| GPNMB | 1.7 | 0.0394 |
| IGFBP2 | 1.6 | 0.0279 |
| LGALS3 | 1.7 | 0.0521 |
| OLIG2 | 0.6 | 0.0480 |
| PDPN | 1.9 | 0.0452 |
| RTN1 | 0.5 | 0.0091 |

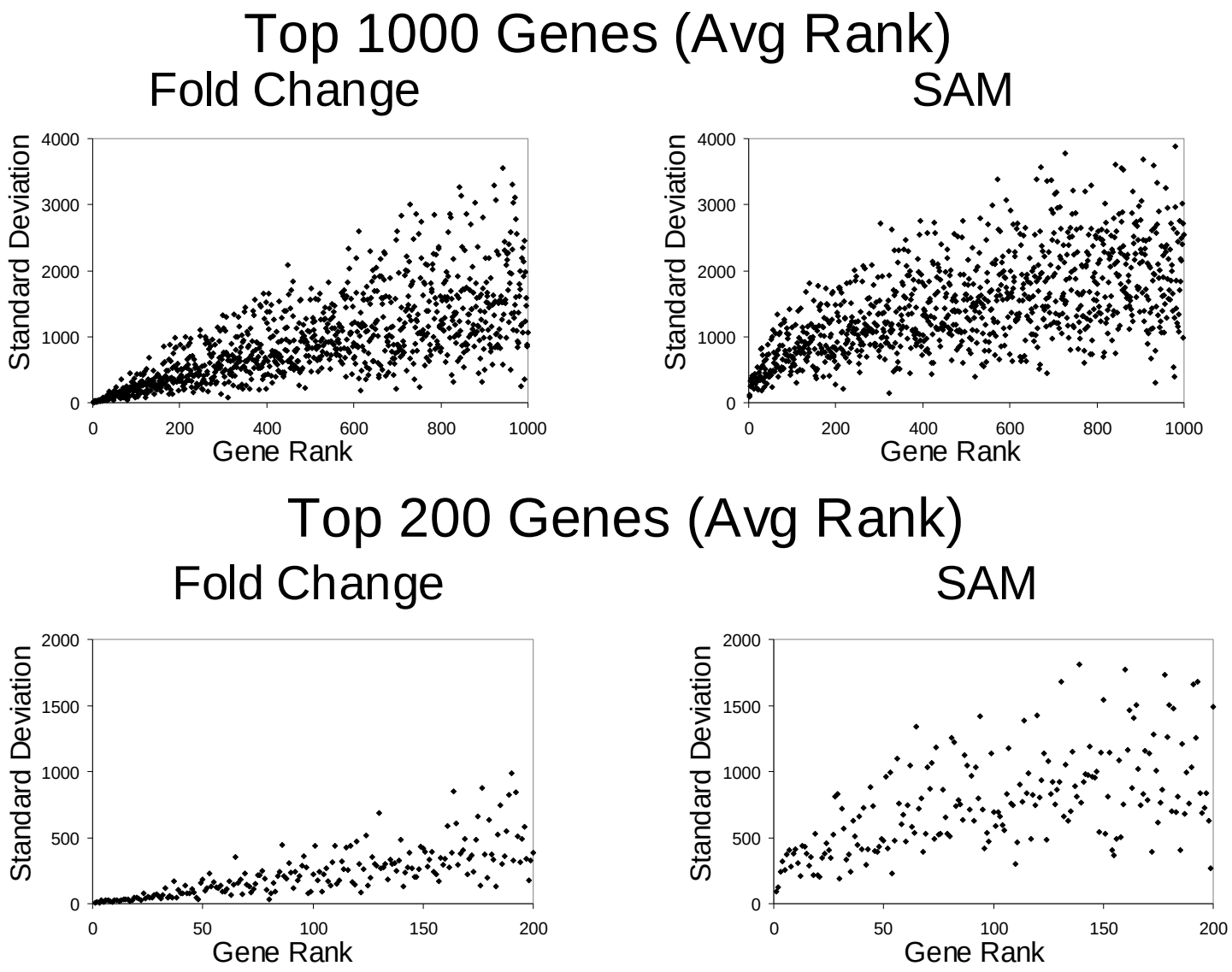
Supplementary Table 6. Results of qRT-PCR analyses for 9-gene profile on 101GBM samples from temozolomide-treated patients. Univariate Cox analyses were performed for each gene, and the effect size (hazard ratio, or HR) and P-values are shown. To generate comparable hazard ratios, median expression values for each gene were determined and the value was expressed as 0 or 1 based on whether it was lower or higher than the median, respectively.

| Case ID | metagene score rank | CD133 | nestin |
|---|---|---|---|
| M234 | 168 | 2 | 2 |
| M229 | 167 | 1 | 1 |
| M201 | 166 | 1 | 1 |
| P84 | 164 | 2 | 2 |
| H72 | 163 | 2 | 2 |
| C207 | 162 | 2 | ND |
| C38 | 160 | 2 | ND |
| C284 | 159 | 2 | 2 |
| H23 | 157 | 2 | 2 |
| H166 | 155 | 2 | ND |
| H6 | 151 | 1 | 2 |
| C22 | 146 | 2 | 2 |
| C114 | 139 | 2 | 2 |
| H103 | 131 | 2 | 2 |
| H29 | 128 | 1 | 2 |
| C27 | 127 | 2 | ND |
| M278 | 123 | ND | 2 |
| M248 | 121 | ND | 2 |
| C167 | 116 | 1 | ND |
| H190 | 113 | 1 | 2 |
| P79 | 97 | 1 | 2 |
| C85 | 95 | 1 | 2 |
| M283 | 94 | ND | 1 |
| C103 | 93 | 2 | ND |
| C176 | 92 | 1 | 2 |
| M205 | 88 | 0 | 1 |
| P87 | 85 | 2 | 1 |
| M244 | 82 | ND | 1 |
| M221 | 80 | 2 | 2 |
| H220 | 79 | 2 | 2 |
| H21 | 78 | 1 | 1 |
| H208 | 76 | 0 | 0 |
| M209 | 74 | 0 | 1 |
| C195 | 71 | 2 | 2 |
| H202 | 69 | 2 | 2 |
| P80 | 68 | 0 | 0 |
| H176 | 57 | 0 | 1 |
| M235 | 56 | 1 | 1 |
| M216 | 55 | 1 | 1 |
| C91 | 52 | 2 | 1 |
| P83 | 51 | 1 | 1 |
| P89 | 46 | ND | 0 |
| P82 | 45 | 0 | 0 |
| C158 | 43 | 1 | 0 |
| P17 | 41 | 0 | 0 |
| M236 | 39 | 1 | 2 |
| C32 | 36 | 2 | 2 |
| C228 | 34 | 0 | ND |
| M232 | 32 | 2 | 1 |
| H215 | 31 | 0 | 2 |
| P78 | 28 | 1 | 2 |
| C110 | 27 | 0 | ND |
| H139 | 26 | 0 | 0 |
| P75 | 23 | 0 | 0 |
| H118 | 18 | 0 | 0 |
| H203 | 17 | 0 | 0 |
| M277 | 16 | ND | 0 |
| H196 | 14 | 0 | 0 |
| M220 | 4 | 0 | 0 |
| H143 | 3 | 1 | 0 |

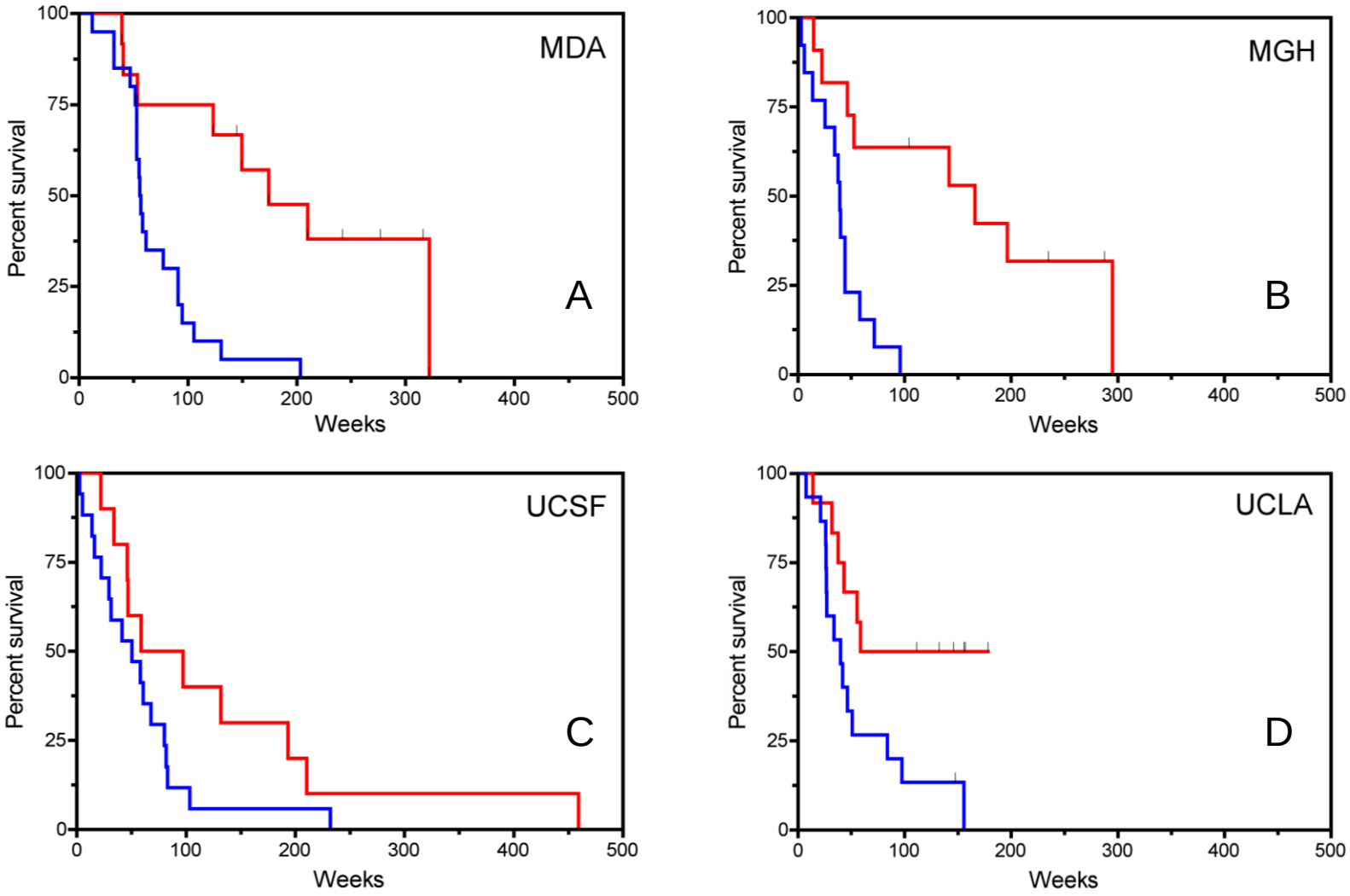
Supplementary Table 7. Expression of CD133 and testing with comparison to metagene scores. Samples for which slides were available were tested by immunohistochemistry for CD133 and nestin and scored semiquantitatively for expression using a 3-tiered system. Scores are shown with comparison to the rank determined by the 9-gene metagene score, where higher rank indicates a prediction of unfavorable outcome. ND=not done.



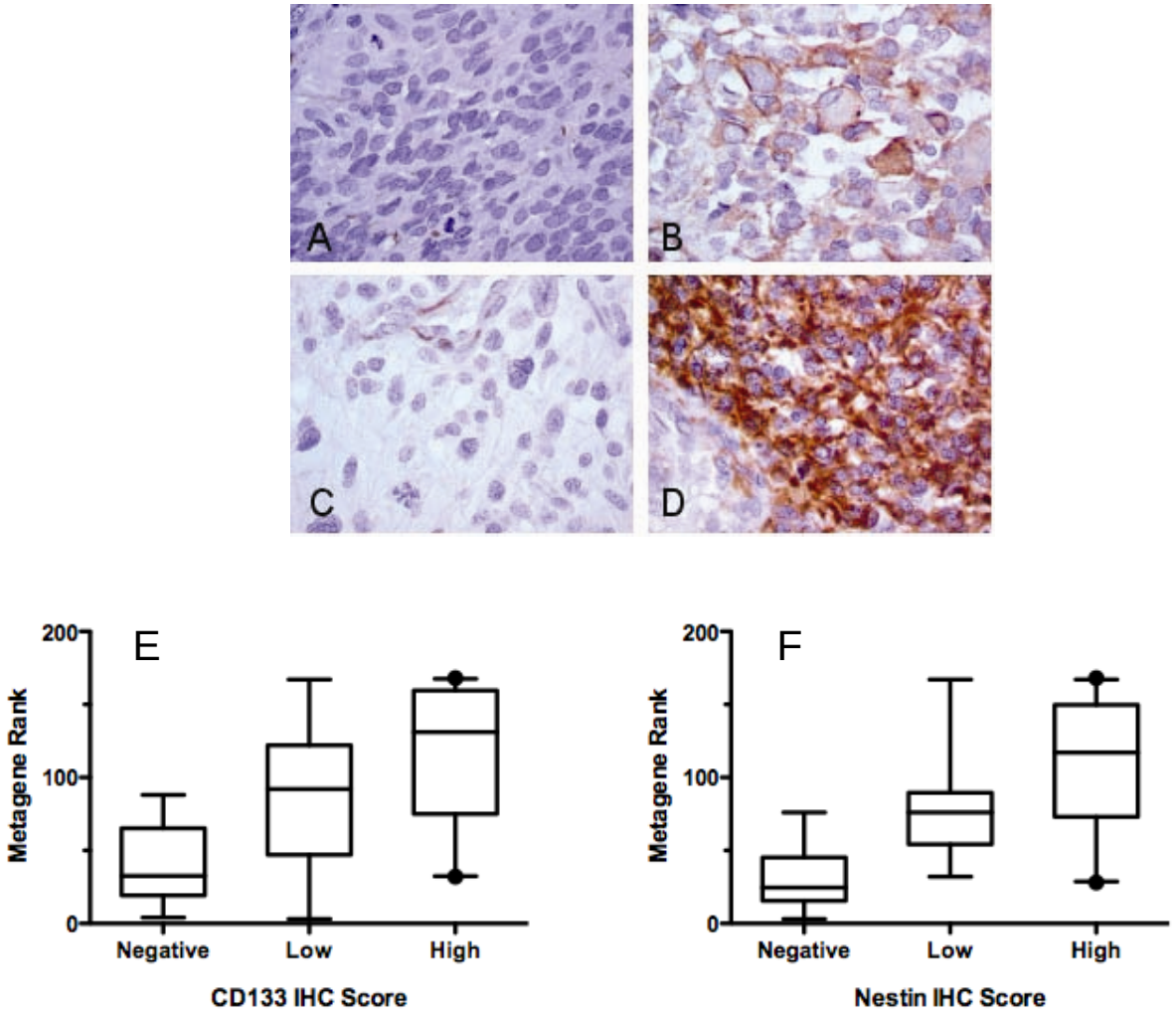

Supplementary Figure 1 Scheme used to identify robust survival genes in independent microarray datasets derived from MD Anderson (MDA), Massachusetts General Hospital (MGH), University of California-Los Angeles (UCLA) and University of California-San Francisco




Supplementary Figure 2. Consistency of Gene Rankings Across Institutions: Individual genes were ranked by fold change or SAM 2-class (TS vs. LTS) within each institution. Average rank and standard deviation of gene ranks across the 4 microarray data sets were calculated. The standard deviation as a function of average gene rank are plotted for the top 1000 genes (top row) or top 200 genes (bottom row) for Fold Change and SAM. The lower standard deviation observed across all rankings using fold change indicated that this method gave more consistent rankings of individual genes across institutions and fold change was thus chosen as the method used to identify the most robust survival genes common to the independent data sets.




Supplementary Figure 3. Test of robustness of gene expression sets among institutions using a cross validation method. Data were combined from 3 institutions into a single dataset, and the list of the top 200 survival genes identified among those 3 institutions (the training set). This list of genes was then used for K-means clustering of the dataset from 4th institution (the test set). The survival times are plotted for the 2 groups that resulted from the clustering analysis. This procedure was repeated for all (n=4) possible combinations of the datasets and the resulting Kaplan-Meier curves for the test set in each case shown in A-D




Supplementary Figure 4. Expression of CD133 and nestin and correlation with metagene score. Immunohistochemistry was performed on GBM tumor samples (see Supplementary Table 5) and scored semiquantitatively. A. and B. CD133 staining, showing a negative case (A) and a positive case with staining at the cell membrane (B). C. and D. nestin staining, showing a negative case (C) and a positive case (D). . E. and F. Box-whisker plots showing the association of CD133 (E) and nestin (F) with metagene score.

References

1. Freije WA, Castro-Vargas FE, Fang Z, Horvath S, Cloughesy T, Liau LM, et al. Gene expression profiling of gliomas strongly predicts survival. Cancer Res 2004;64(18):6503-10.

2. Nutt CL, Mani DR, Betensky RA, Tamayo P, Cairncross JG, Ladd C, et al. Gene expression-based classification of malignant gliomas correlates better with survival than histological classification. Cancer Res 2003;63(7):1602-7.

3. Pruitt KD, Tatusova T, Maglott DR. NCBI Reference Sequence project: update and current status. Nucleic Acids Res 2003;31(1):34-7.

4. Imanishi T, Itoh T, Suzuki Y, O'Donovan C, Fukuchi S, Koyanagi KO, et al. Integrative annotation of 21,037 human genes validated by full-length cDNA clones. PLoS Biol 2004;2(6):e162.

5. Zhang L, Miles MF, Aldape KD. A model of molecular interactions on short oligonucleotide microarrays. Nat Biotechnol 2003;21(7):818-21.

6. Zhang L, Miles MF, Aldape KD. A model of molecular interactions on short oligonucleotide microarrays. Nat Biotechnol 2003.

7. Nigro JM, Misra A, Zhang L, Smirnov I, Colman H, Griffin C, et al. Integrated array-comparative genomic hybridization and expression array profiles identify clinically relevant molecular subtypes of glioblastoma. Cancer Res 2005;65(5):1678-86.

8. Shi L, Reid LH, Jones WD, Shippy R, Warrington JA, Baker SC, et al. The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. Nat Biotechnol 2006;24(9):1151-1161.

9. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 1998;95(25):14863-8.

10. Simmons ML, Lamborn KR, Takahashi M, Chen P, Israel MA, Berger MS, et al. Analysis of complex relationships between age, p53, epidermal growth factor receptor, and survival in glioblastoma patients. Cancer Res 2001;61(3):1122-8.