Supplemental Materials and Methods:

Animal procedures. Nkx3.1-deficient mice used in this study were previously described (Abdulkadir et al., 2002). Conditional *Pten* knockout mice and prostate-specific c-myc overexpression transgenic mice were originally generated and kindly provided by Hong Wu (UCLA) and Charles Sawyers (UCLA), respectively. *Pten*$^{loxp/loxp}$ mice were crossed to *PSA-Cre* mice that express Cre exclusively in prostate epithelial cells beginning at ~6 weeks of age. PCR primers and conditions for genotyping all the genetically modified mice with tail genomic DNA were used as provided (Ellwood-Yen et al., 2003; Wang et al., 2003). For all castration-testosterone replacement experiments, mice were castrated at 8 weeks of age and allowed to recover. At 14 days post-castration, testosterone was replaced by implanting 2 cm silastic capsules (1.5 mm inside diameter, Fisher Scientific, Pittsburgh, PA) filled with 50 mg of testosterone propionate (Sigma, St. Louis, MO). At the indicated time points after testosterone replacement, prostates were harvested to obtain total RNA for qRT-PCR.

Cell Culture, Plasmid Construction, Infection, and Transfection. LAPC4 cells were cultured in ISCOVE supplemented with 10% FBS and 100 µg/ml penicillin-streptomycin-glutamine (Sigma, St. Louis, MO). PC-3 (CRL-1435) cells were purchased from ATCC and cultured according to their instructions. pRNS-1-1 (non-tumorigenic human prostate epithelial cell line) was kindly provided by J. Rhim (University of Health Sciences, Bethesda, MD). To generate VP16-NKX3.1, the VP16 acidic domain (residues 446-490) was fused to the n-terminus of NKX3.1 by PCR. For generation of adenoviruses expressing VP16-NKX3.1 or NKX3.1, the AdEasy system was used

according to manufacturer's instructions (He et al., 1998). Adenovirus titers of 10^{11-12} infectious particles per ml were obtained. Cells were infected with Adenovirus (MOI = 50 to 100) for 2 hr and washed twice with medium. Protein or mRNA was prepared at designated time points after infection by standard methods. To generate lentivirus expressing VP16-NKX3.1-ER, we fused the ER^{TM} module, a mutated ligand binding domain of the human estrogen receptor, to the Nkx3.1 carboxy terminus and cloned this fragment into the FUIV lentivirus vector (Araki et al., 2004). Lentiviruses were produced by standard methods. The cells were infected with lentivirus for 24 hr. After 3 days, the cells were sorted by FACS based on EGFP fluorescence intensity to obtain cells with a similar level of expression of the gene of interest (e.g. VP16-NKX3.1). Transient expression in LAPC4 cells was achieved using Lipofectamine-mediated transfection. Luciferase assays were performed 2 days after transfection by standard methods. The luciferase activity was normalized to β-galactosidase activity expressed from the co-transfected CMV-lacZ plasmid.

Quantitative RT-PCR. For validation of microarray gene expression data, qRT-PCR analyses were performed using a SDS 7700 real-time thermal cycler (Applied Biosystems, Foster City, California). Independent LCM RNA (n=3 or 4) and total RNA from each lobe of prostates (n=5 or 6), or total prostates in castration-TR experiments (n=5 in each time point and each genotype) were used to corroborate selected candidate genes identified from microarray data. Amplicons were detected using Syber Green I fluorescence (Molecular Probes) as described elsewhere (Svaren et al., 2000). Target genes were analyzed using standard curves to determine relative levels of gene

expression. Individual RNA samples were normalized according to the levels of GAPDH mRNA or 18S RNA (Zhang et al., 2007).

Co-analysis using gene expression profiles from the castration-testosterone replacement (TR) paradigm. Fifty non-EST probes (45 genes) from the loss-of-Nkx3.1 signature were aligned with their counterparts on the MU74v2 microarray data sets previously generated from castration-TR experiments in Nkx3.1-deficient mice (Magee et al., 2003). The post-TR day 7 and 14 time points of the castration-TR paradigm were selected to perform cross-examination with the LCM-based early prostate cancer progression datasets because allelic loss of Nkx3.1 in luminal epithelial cells leads to dysregulated proliferation/differentiation at these time points. Probes corresponding to loss-of-Nkx3.1 signature genes from Figure 1 without presence call were not presented in the heat-map in Figure 4. Genes that were not present on the earlier MU74v2 microarrays were directly investigated using qRT-PCR with appropriate TR samples to examine their expression patterns in the castration-TR paradigm.

Meta-analysis of expression profile data from mouse models of prostate adenocarcinoma. Probe-level datasets for gene expression profiles were obtained from the primary investigators (Ellwood-Yen et al., 2003; Majumder et al., 2003; Wang et al., 2003) and used as provided. Because different microarray formats were used for each study, each dataset was independently analyzed and then transformed to produce similar output across data sets for further analysis. The meta-analysis applied here was modified from the bioinformatic approach previously described (Pilarsky et al., 2004). Briefly, SAM (Tusher et al., 2001) was used to identify probe sets that were significantly

overexpressed or underexpressed in prostate tissue specimens from each different state of malignant progression relative to normal prostate tissues. A false discovery rate (FDR) of 5% was applied to Affymetrix Genechip data sets, Nkx3.1 (MOE430 2.0), and c-myc and Akt (MU74v2). The FDR for the Pten data set, which was obtained using custom cDNA microarrays, was 0.1% so that similar numbers of genes were obtained for this dataset. We assigned the value of 1 if they were overexpressed or -1 if they were underexpressed in the tumor and/or PIN. If a probe was not differentially expressed, we assigned a value of 0. Genes that were either overexpressed or underexpressed in more than one data set were used for the meta-analysis.

Supplemental Reference:

Abdulkadir SA, Magee JA, Peters TJ, Kaleem Z, Naughton CK, Humphrey PA and Milbrandt J. (2002). *Mol Cell Biol*, **22,** 1495-503.

Araki T, Sasaki Y and Milbrandt J. (2004). *Science*, **305,** 1010-3.

Ellwood-Yen K, Graeber TG, Wongvipat J, Iruela-Arispe ML, Zhang J, Matusik R, Thomas GV and Sawyers CL. (2003). *Cancer Cell*, **4,** 223-38.

He TC, Zhou S, da Costa LT, Yu J, Kinzler KW and Vogelstein B. (1998). *Proc Natl Acad Sci U S A*, **95,** 2509-14.

Magee JA, Abdulkadir SA and Milbrandt J. (2003). *Cancer Cell*, **3,** 273-83.

Majumder PK, Yeh JJ, George DJ, Febbo PG, Kum J, Xue Q, Bikoff R, Ma H, Kantoff PW, Golub TR, Loda M and Sellers WR. (2003). *Proc Natl Acad Sci U S A*, **100,** 7841-6.

Pilarsky C, Wenzig M, Specht T, Saeger HD and Grutzmann R. (2004). *Neoplasia*, **6,** 744-50.

Svaren J, Ehrig T, Abdulkadir SA, Ehrengruber MU, Watson MA and Milbrandt J. (2000). *J Biol Chem*, **275,** 38524-31.

Tusher VG, Tibshirani R and Chu G. (2001). *Proc Natl Acad Sci U S A*, **98,** 5116-21.

Wang S, Gao J, Lei Q, Rozengurt N, Pritchard C, Jiao J, Thomas GV, Li G, Roy-Burman P, Nelson PS, Liu X and Wu H. (2003). *Cancer Cell*, **4,** 209-21.

Zhang B, Jain S, Song H, Fu M, Heuckeroth RO, Erlich JM, Jay PY and Milbrandt J. (2007). *Development*, **134,** 3191-201.

Supplemental Figure Legends:

Supplemental Figure S1. Sprr2a is highly expressed in hyperplasic and PIN lesions of *Nkx3.1*$^{-/-}$ mice.

(A) qRT-PCR analysis of Sprr2a expression in NOR (normal glands from 2 month old *Nkx3.1*$^{+/+}$ mice), HYP (hyperplastic lesions from 2 month old *Nkx3.1*$^{-/-}$ mice), and PIN (PIN lesions from 5 month old *Nkx3.1*$^{-/-}$ mice) RNA samples obtained by LCM. (B) Sprr2a expression by qRT-PCR using total prostate RNA from 2 month old wildtype and *Nkx3.1*$^{-/-}$ mice. (C) *In situ* hybridization with ^{35}S-labeled antisense probe for Sprr2a. Arrows in the middle and bottom panels indicate representative hyperplastic cells and PIN lesions where Sprr2a is highly induced, respectively. Sprr2a is not expressed in normal glands in wildtype prostate. The right panel with H&E staining shows histology of an adjacent section of Sprr2a *in situ* hybridization. 2M(+/+), 2M(-/-), and 5M(-/-)

represent 2 month old wildtype, 2 month old Nkx3.1-deficient, and 5 month old Nkx3.1-deficient prostates, respectively.

Supplemental Figure S2. Cathepsin E is highly expressed in hyperplasic and PIN lesions in *Nkx3.1*$^{-/-}$ mice.

(A) qRT-PCR analysis of cathepsin E expression in NOR (normal glands from 2 month old *Nkx3.1*$^{+/+}$ mice), HYP (hyperplastic lesions from 2 month old *Nkx3.1*$^{-/-}$ mice), and PIN (PIN lesions from 5 month old *Nkx3.1*$^{-/-}$ mice) RNA samples obtained by LCM. (B) Cathepsin E expression by qRT-PCR using total prostate RNA from 2 month old wildtype and *Nkx3.1*$^{-/-}$ mice.

(C) *In situ* hybridization with ^{35}S-labeled antisense probe for cathepsin E. Arrows in the middle and bottom panels indicate representative hyperplastic cells and PIN lesions where cathepsin E is highly induced, respectively. Cathepsin E is not expressed in normal glands in wildtype prostate. The right panel with H&E staining shows histology of an adjacent section of Sprr2a *in situ* hybridization. 2M(+/+), 2M(-/-), and 5M(-/-) represent 2 month old wildtype, 2 month old Nkx3.1-deficient, and 5 month old Nkx3.1-deficient prostates, respectively.

Supplemental Figure S3. Quiescin Q6 is highly expressed in high grade PIN lesions of *Pten*-deficient mice.

Quiescin Q6 *in situ* hybridization was performed on ventral (VP) and dorsolateral (DLP) prostates in prostate-specific Pten-deficient mice. Note that quiescin Q6 is overexpressed only in glands with high-grade PIN both in VP and DLP. Arrows and arrowheads

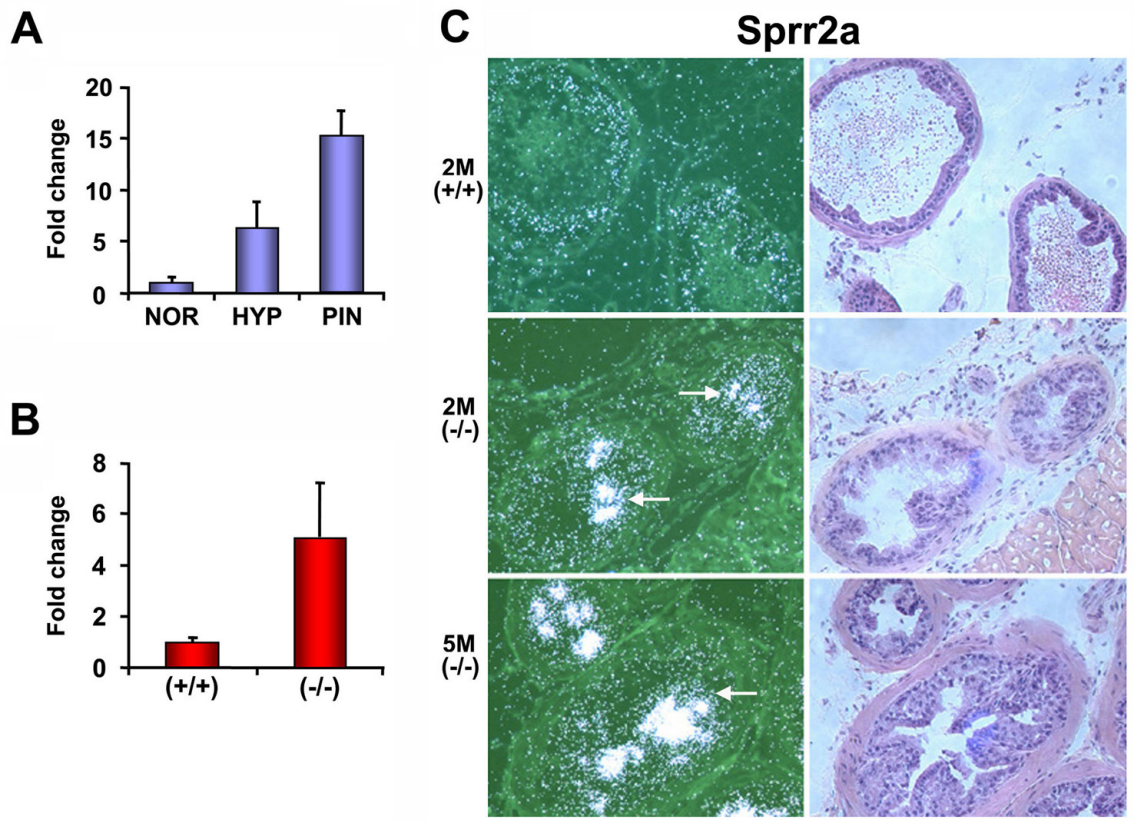
indicate representative high grade PIN lesions that show quiescin Q6 overexpression. Top panel represents bright field images of sections adjacent to those used for *in situ* hybridization.

Supplemental Figure S4. Clusterin mRNA levels are not elevated in human primary prostate tumors.

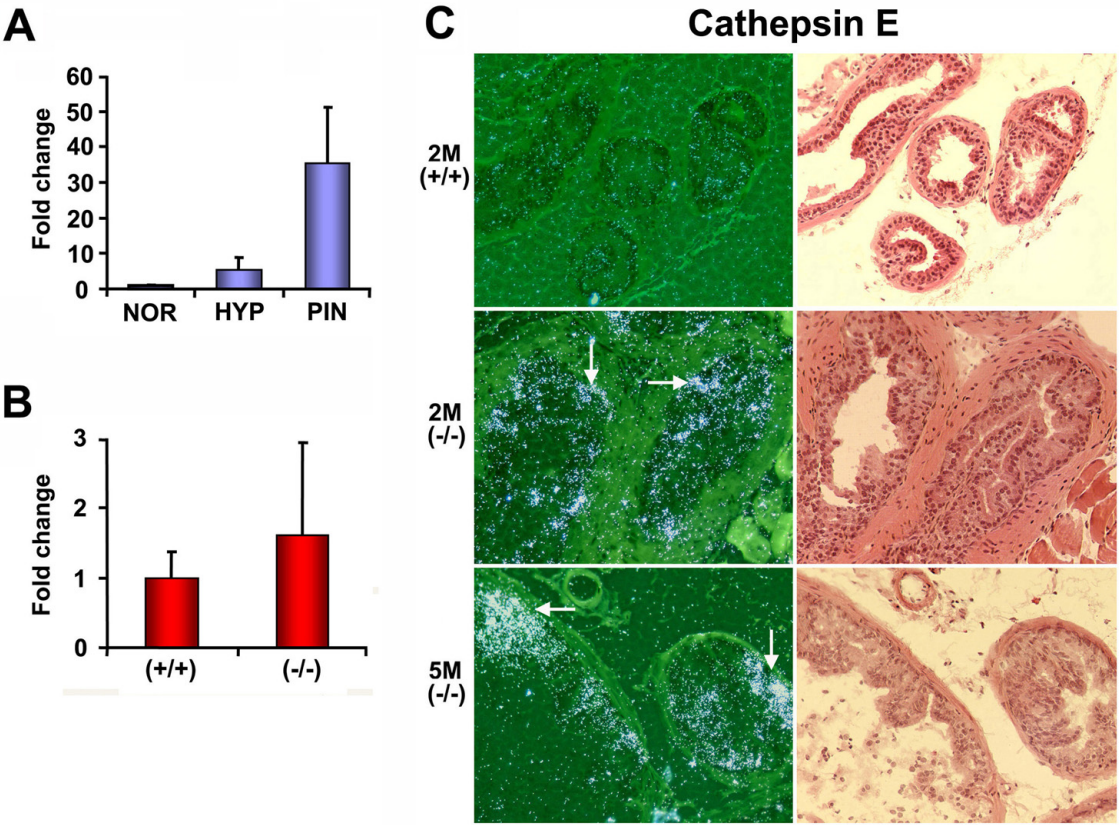
(A) qRT-PCR analysis of clusterin expression in total RNA samples of normal prostate and primary prostate tumor. Note that there is no statistical significance between them. (B) Values of human clusterin mRNA levels from microarrays performed with stage-specific human RNA samples obtained by LCM. "n" indicates the number of total RNA samples (A) and microarray chips with RNA obtained by LCM (B). Norm, PIN, and Inv represent normal prostate, prostatic intraepithelial neoplasia, invasive adenocarcinoma, respectively.

Supplemental Table S1. Functional categories of differentially expressed genes listed in Figure 1 during the progression of prostate tumorigenesis in Nkx3.1 deficient mice.

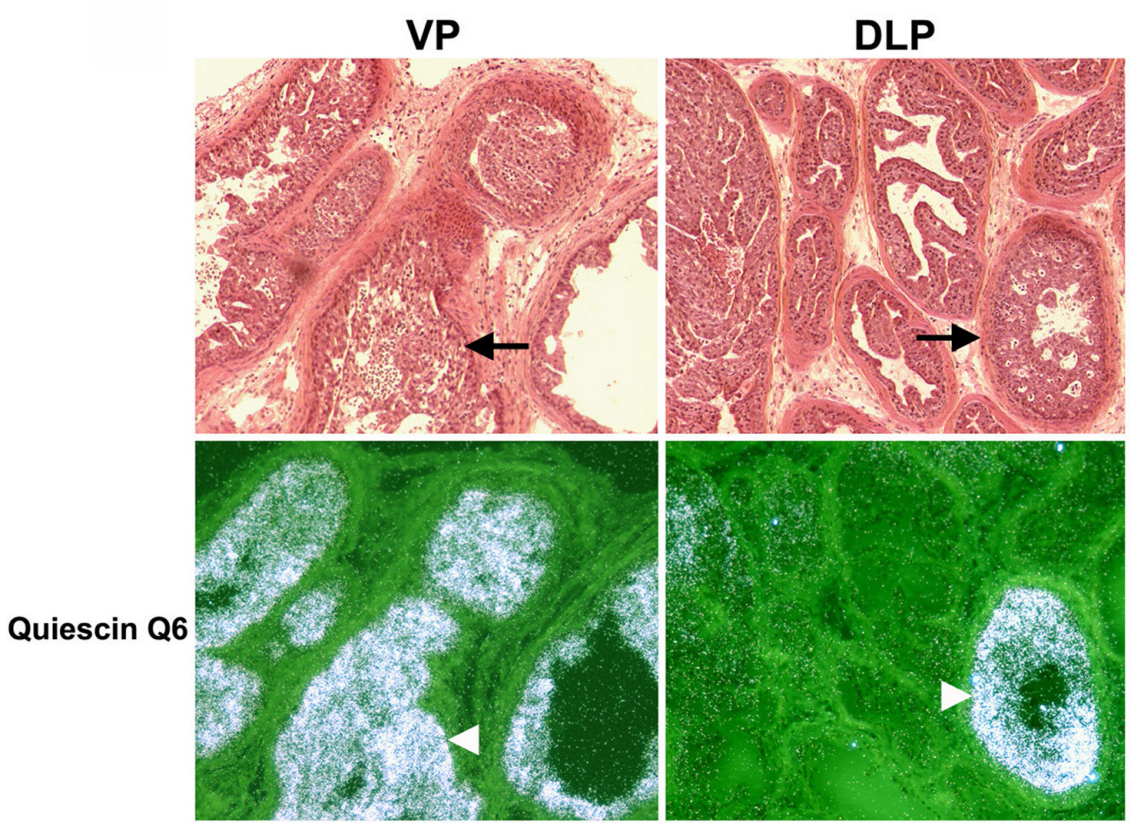
Supplemental Table S2. A list of genes which are either up- and/or down-regulated in at least two out of four mouse models used for meta analysis in Figure 8. Values (-1, 0, 1) represent down-regulation, no change, and up-regulation, respectively.



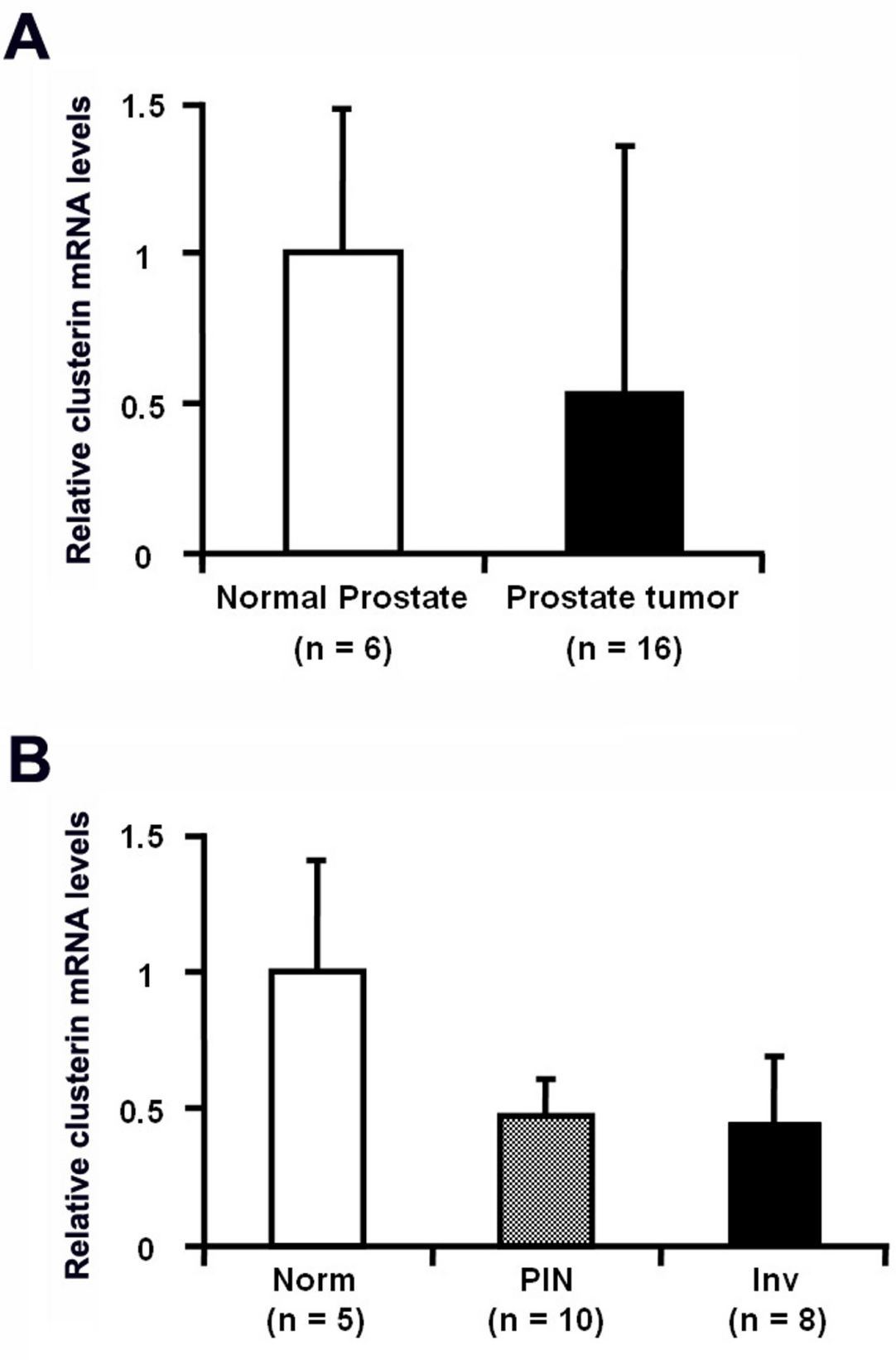
Supplemental Figure S1
Song et al.



Supplemental Figure S2
Song et al.



Supplemental Figure S3
Song et al.



Song et al.
Supplemental Figure S4