Published in final edited form as:

Prostate. 2015 February ; 75(2): 218–224. doi:10.1002/pros.22908.

Epithelial-mesenchymal transition in prostate cancer is associated with quantifiable changes in nuclear structure

James E. Verdone¹, Princy Parsana^{1,5}, Robert W. Veltri¹, and Kenneth J. Pienta^{1,2,3,4,5}

¹The James Buchanan Brady Urological Institute, Department of Urology, The Johns Hopkins School of Medicine

²Department of Oncology, The Johns Hopkins School of Medicine

³Department of Pharmacology and Molecular Sciences, The Johns Hopkins School of Medicine

⁴Departments of Chemical and Biomolecular Engineering, Johns Hopkins University, Baltimore, MD, USA

⁵Computer Science, Johns Hopkins University, Baltimore, MD, USA

Abstract

Background—Prostate cancer progression is concomitant with quantifiable nuclear structure and texture changes as compared to non-cancer tissue. Malignant progression is associated with an epithelial-mesenchymal transition (EMT) program whereby epithelial cancer cells take on a mesenchymal phenotype and dissociate from a tumor mass, invade, and disseminate to distant metastatic sites. The objective of this study was to determine if epithelial and mesenchymal prostate cancer cells have different nuclear morphology.

Methods—Murine tibia injections of epithelial PC3 (PC3-Epi) and mesenchymal PC3 (PC3-EMT) prostate cancer cells were processed and stained with H&E. Cancer cell nuclear image data was obtained using commercially available image-processing software. Univariate and multivariate statistical analysis were used to compare the two phenotypes. Several nonparametric classifiers were constructed and permutation-tested at various training set fractions to ensure robustness of classification between PC3-Epi and PC3-EMT cells *in vivo*.

Results—PC3-Epi and PC3-EMT prostate cancer cells were separable at the single cell level in murine tibia injections on the basis of nuclear structure and texture remodeling associated with an epithelial-mesenchymal transition. Support vector machine and multinomial logistic regression models based on nuclear architecture features yielded AUC-ROC curves of 0.95 and 0.96, respectively, in separating PC3-Epi and PC3-EMT prostate cancer cells *in vivo*.

Conclusions—Prostate cancer cells that have undergone an epithelial-mesenchymal transition demonstrated an altered nuclear structure. The association of nuclear changes and a mesenchymal phenotype demonstrates quantitative morphometric image analysis may be used to detect cancer

Correspondence: Kenneth J. Pienta, The James Buchanan Brady Urological Institute, Departments of Urology, Oncology, and Pharmacology and Molecular Sciences, Johns Hopkins School of Medicine; Department of Chemical and Biomolecular Engineering, Johns Hopkins University, 600 N Wolfe St, 105 Marburg, Baltimore, MD 21287, Phone: 410-502-3137, Fax: 410-955-0833, kpienta1@jhmi.edu.

Conflicts of Interest: Nothing to declare.

cells that have undergone epithelial-mesenchymal transition. This morphometric measurement could provide valuable prognostic information in patients regarding the likelihood of [future] metastatic disease.

Keywords

EMT; imaging; nuclear morphology; prostate cancer; morphometry

Introduction

Prostate cancer (PCa) is the most commonly diagnosed non-skin cancer in men in the United States [1]. Despite the frequency of prostate cancer diagnosis, there are few prognostic biomarkers available for clinical use [2-4].

High-order organization and compartmentalization within the cell nucleus is critical for cellular processes including proliferation, cell division, and transcription [5-8]. The cell nucleus in cancer is subject to genetic instability from DNA double-strand breaks leading to significant alterations in nuclear size, shape, and chromatin texture that may promote proliferative signaling, cancer progression, and malignancy [9]. Quantitative analysis of cancer cell nuclear morphology has been used to distinguish cancer aggressiveness, Gleason grade, and heterogeneity on the basis of this cancer-associated nuclear remodeling [10-16].

Throughout the course of progression to metastasis, malignant cancer cells undergo a series of reversible transitions between intermediate phenotypic states bounded by pure epithelium and pure mesenchyme. The transformation from epithelium to mesenchyme is known as epithelial-mesenchymal transition (EMT) and the reverse program is known as mesenchymal-epithelial transition (MET) [17]. The EMT program is fundamentally characterized by the loss of epithelial markers such as E-cadherin and the gain of mesenchymal markers such as vimentin and N-cadherin. Recent studies investigating the EMT program in malignant cancers show overexpression of EMT transcription factors (EMT-TFs) such as ZEB1/2, Twist, Snail, and Slug is indicative of invasive tumors with poor patient outcome [18-21]. EMT molecular programmatic shift allows, albeit with some degree of transiency, epithelial invasion and dissemination to distant metastatic sites and appears to confer some attributes of stem-like cells [22-24]. As distant metastases are the leading cause of cancer-related death in men with PCa, observing EMT alterations in circulating tumor cells or cancer cells from prostate biopsies could be a promising prognostic marker.

It remains difficult to identify mesenchymal from epithelial cancer cells in tissue on the basis of morphologic features alone without the application of special markers. This limits our ability to understand how and when EMT and MET are occurring and how these processes might affect patient prognosis. In this study, we demonstrate that a quantitative measurement of nuclear morphology distinguishes prostate cancer cells that have undergone EMT from their epithelial counterparts. These mesenchymal cancer cells were well discriminated from epithelial cancer cells on the basis of nuclear structure and texture remodeling at the single cell level.

Materials and Methods

Cell Culture and Reagents

PC3-Epithelial (PC3-Epi) and PC3-Mesenchymal (PC3-EMT) were derived from parental PC3 prostate cancer cells as previously described [25]. Cells were maintained in RPMI-1640 media (Gibco, Grand Island, NY) supplemented with 10% fetal bovine serum (FBS) (Sigma Life Sciences, St. Louis, MO) and penicillin/streptomycin (1% [v/v]; Mediatech, Inc., Manassas, VA) at 37°C and 5% CO₂ in a humidified incubator.

NOD-SCID Mouse Intra-tibia Injection Model and Slide Preparation

6-8 week old male NOD-SCID mice were purchased from the Johns Hopkins University Sidney Kimmel Comprehensive Cancer Center Immune Compromised Colony II (JHU SKCCC Office of Research Services, Animal Resources, Baltimore, MD). All animal experiments were performed in accordance with an institutionally approved protocol under the provisions of the Johns Hopkins Institutional Animal Care and Usage Committee and in accordance with local and federal law. For intra-tibia injection, 2×10^5 PC3-Epi or 2×10^5 PC3-EMT cells suspended in 100 μ L PBS were inoculated intra-tibially and grown for two weeks (5 mice per group). Tibias were extracted and fixed in 10% neutral buffered formalin solution (Sigma, St. Louis, MO) for 48 hours. Tibias were then rinsed in PBS at 4°C for 2 hours and decalcified in Immunocal decalcification solution (Decal Chemical Corporation, Tallman, NY) overnight. Decalcified tibias were then rinsed in PBS for 15 minutes and paraffin-embedded for serial sectioning at 5 μ m thicknesses. Serially sectioned slides were stained at every third slide with hematoxylin (Dako North America, Inc., Carpinteria, CA) and eosin Y (Fisher Scientific Company LLC, Kalamazoo, MI) (H&E).

Image Acquisition of H&E Cancer Nuclei in Tibia Model

Images of H&E-stained tibia sections were captured at 400X using a Nikon Eclipse E400 microscope with Nikon DigitalSight image capture attachment (Nikon Instruments Inc., Melville, NY). Image-Pro Premier Software (MediaCybernetics, Rockville, MD) extracted various nuclear measurements from images and data was pooled across replicates. Nuclear biometric data was then exported for statistical analysis; a complete list of measured features can be found in Table S1.

Statistical Analysis and Data Processing

All data processing and statistical analysis was performed using the MATLAB R2014a numerical computing software suite (MathWorks, Natick, MA). Mann-Whitney *U*-tests and 2-sample Kolmogorov-Smirnov tests were used for non-parametric univariate feature analysis. P-values were adjusted using Bonferroni correction to minimize family wise error rate. Features with Bonferroni corrected p-value < 0.05 were considered significant. A number of supervised and unsupervised machine learning methods were applied to cluster, classify, and evaluate robustness of classification of PC3-Epi and PC3-EMT phenotypes on the basis of nuclear biometric features including *k*-nearest neighbors, support vector machines, and multinomial logistic regression models.

Results

Epithelial to Mesenchymal Transition Directs Structural Changes in Prostatic Cancer Cell Nuclei Detectable at the Population Level *in vivo*

H&E-stained PC3-Epi and PC3-EMT cells display distinct phenotypes when inoculated in murine tibias (Fig. 1 A-B). We used the univariate Mann-Whitney *U*-test on nuclear biometric data to identify differential nuclear morphological features between the two phenotypes. 49 out of the 72 features obtained from Image-Pro Premier system were significant (Bonferroni corrected P-value <0.05) (Table S1). Principal component analysis (PCA) was performed using 49 features to select a subset characterized by a high correlation with more prominent principal components and a low inter-feature correlation to eliminate feature redundancy and reduce data complexity while maximizing total information retained; this subset of 5 features is listed in Table 1. Nuclear texture, nuclear size, and nuclear content features figured most prominently in discriminating PC3-Epi and PC3-EMT at the population level as tested by Mann-Whitney *U*-tests. These features also confirmed separation of PC3-Epi and PC3-EMT when evaluated with two-sample Kolmogorov-Smirnov Tests (See Table 1 and Fig. 2 A-E, respectively). PCA was performed on the subset and confirmed non-redundancy of features (Fig. 3).

Expanding from univariate analysis to multivariate analysis, feature data was linearly transformed into canonical variate space (CV space) to maximize variance between *a priori* labeled classes. A Gaussian mixture model in CV space predicts two distinct cluster-means in unlabeled data, implying separability of transformed group-labeled data (Fig. 4).

Epithelial and mesenchymal cells can be discriminated at the Single Cell Level *in vivo*

Supervised data classification was performed with *k*-nearest neighbors (k-NN), a support-vector machine in canonical variate space (SVM-CV), and multinomial logistic regression (MLR). Non-parametric classifier methods were selected for simplicity and to avoid model-assumed knowledge of underlying feature distributions. Supervised learning methods employed the reduced set of 5 features shown in Table 1. *k* for *k*-nearest neighbors was chosen to minimize *n*-fold cross validation loss (mean-squared error) (Fig. S1). Training and testing sets were randomly generated by permutation and divided into 80% training and 20% testing sets. 100-fold cross validation was performed to ensure stability of classification output.

Area under the representative receiver operating characteristic curves (AUC-ROC) for SVM-CV and MLR classifiers demonstrated a high degree of separation between PC3-Epi and PC3-EMT at the single cell level (Fig. 5, Table 2). Additionally, the stringent k-NN classifier performed moderately well (AUC = 0.82), consistent with these results. To demonstrate classification robustness, AUC-ROC was calculated for 100-fold cross validation of random permutation-selected data sets at 50% training and 50% testing, 60%:40%, 70%:30%, and 90%:10% training and testing set, respectively (Fig. S2). These results indicate that these classifiers are robust and successfully discriminate individual instance of epithelial-mesenchymal transition derived mesenchymal cancer cells from epithelial cancer cells in prostate cancer.

Discussion

Quantitative nuclear morphometry is a powerful technique for analyzing grade and stage of prostate cancer. Changes in cancer cell nuclear compartmentalization and chromatin structure parallel cancer grade and stage [10,12,14]. Unfortunately for the majority of patients with mid-grade prostate cancer, prognosis is uncertain. Distinguishing lethal prostate cancers from nonlethal cancers is a major goal in the field. Lethal cancer corresponds with the large-scale genomic remodeling tied to the epithelial-mesenchymal transition program: overexpression of EMT-TFs Snail, Slug, ZEB1/2, and Twist that negatively regulate E-cadherin correlate strongly with patients that have poor outcomes [18-21].

EMT-TFs drive the EMT program by various reversible epigenetic alterations at varying degrees of stability [26,27]. EMT-TF machinery may employ a host of epigenetic mechanisms to silence epithelial genes including de-acetylation of histone H3 lysine acetylation sites (H3Kac), trimethylation of histone H3 K27, H3 K9, and H3 K4 (H3K27me3, H3K9me3, H3K4me3 respectively) on nucleosomes proximal to epithelial gene promoter regions, and DNA CpG methylation of epithelial gene promoters [17,28-34]. The latter corresponds with the conversion of euchromatin or facultative heterochromatin (loosely packed chromatin) to constitutive heterochromatin (densely packed chromatin) and consequent stable blockage of epithelial genes. As changes in chromatin organization are expected to alter nuclear appearance, we propose the nuclear texture alterations observed in PC-EMT cells originate from EMT-TF driven epigenetic modifications of chromatin structure [30]. Our data is consistent with this framework, as mesenchymal PC3 cells were most strongly distinguishable at the population level using univariate nuclear texture features.

Recent technological advances in scanning microscopy and computer science have permitted advances in quantitative nuclear morphometric approaches to overcome obstacles in the diagnostic space. A current difficulty in the area of diagnostic pathology of prostate and breast biopsy specimens is the differential diagnosis of the very small volume of tumor samples. Recent studies with nuclear morphometry techniques successfully detect the cancer field effect in adjacent tissue in both prostate and breast cancers allowing for detection of cancer without direct observation of a lesion [35-38]. Applying these types of technical advancements to the prognostic space, future studies using nuclear morphometry to detect the degree of mesenchymal transformation or EMT fraction of circulating tumor cells (CTCs) within a prostate biopsy may have considerable value in predicting patient outcome and appropriate course of treatment.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We would like to thank Dr. Donald S. Coffey for his support and continued mentorship on this project.

Financial support: This work was supported by NIH grant U54CA163214, 1P01CA093900, U01CA143055, and U54CA143803. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute or the National Institutes of Health.

References

1. Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics, 2014. *CA: a cancer journal for clinicians*. 2014; 64(1):9–29. [PubMed: 24399786]
2. Van der Kwast TH. Prognostic prostate tissue biomarkers of potential clinical use. *Virchows Archiv: an international journal of pathology*. 2014; 464(3):293–300. [PubMed: 24487790]
3. Tombal B, Alcaraz A, James N, Valdagni R, Irani J. Can we improve the definition of high-risk, hormone naïve, non-metastatic prostate cancer? *BJU international*. 2014; 113(2):189–199. [PubMed: 24112742]
4. Bhavsar T, McCue P, Birbe R. Molecular diagnosis of prostate cancer: are we up to age? *Seminars in oncology*. 2013; 40(3):259–275. [PubMed: 23806492]
5. Croft JA, Bridger JM, Boyle S, Perry P, Teague P, Bickmore WA. Differences in the localization and morphology of chromosomes in the human nucleus. *The Journal of cell biology*. 1999; 145(6): 1119–1131. [PubMed: 10366586]
6. Ferreira J, Paolella G, Ramos C, Lamond AI. Spatial organization of large-scale chromatin domains in the nucleus: a magnified view of single chromosome territories. *The Journal of cell biology*. 1997; 139(7):1597–1610. [PubMed: 9412456]
7. Osborne CS, Chakalova L, Brown KE, Carter D, Horton A, Debrand E, Goyenechea B, Mitchell JA, Lopes S, Reik W, Fraser P. Active genes dynamically colocalize to shared sites of ongoing transcription. *Nature genetics*. 2004; 36(10):1065–1071. [PubMed: 15361872]
8. Sadoni N, Langer S, Fauth C, Bernardi G, Cremer T, Turner BM, Zink D. Nuclear organization of mammalian genomes. Polar chromosome territories build up functionally distinct higher order compartments. *The Journal of cell biology*. 1999; 146(6):1211–1226. [PubMed: 10491386]
9. Zink D, Fischer AH, Nickerson JA. Nuclear structure in cancer cells. *Nature reviews Cancer*. 2004; 4(9):677–687.
10. Ali S, Veltri R, Epstein JI, Christudass C, Madabhushi A. Adaptive energy selective active contour with shape priors for nuclear segmentation and gleason grading of prostate cancer. *Medical image computing and computer-assisted intervention: MICCAI International Conference on Medical Image Computing and Computer-Assisted Intervention*. 2011; 14(Pt 1):661–669.
11. Veltri RW, Christudass CS, Isharwal S. Nuclear morphometry, nucleomics and prostate cancer progression. *Asian journal of andrology*. 2012; 14(3):375–384. [PubMed: 22504875]
12. Veltri RW, Marlow C, Khan MA, Miller MC, Epstein JI, Partin AW. Significant variations in nuclear structure occur between and within Gleason grading patterns 3, 4, and 5 determined by digital image analysis. *The Prostate*. 2007; 67(11):1202–1210. [PubMed: 17525934]
13. Epstein JI, Berry SJ, Eggleston JC. Nuclear roundness factor. A predictor of progression in untreated Stage A2 prostate cancer. *Cancer*. 1984; 54(8):1666–1671. [PubMed: 6478405]
14. Partin AW, Steinberg GD, Pitcock RV, Wu L, Piantadosi S, Coffey DS, Epstein JI. Use of nuclear morphometry, gleason histologic scoring, clinical stage, and age to predict disease-free survival among patients with prostate cancer. *Cancer*. 1992; 70(1):161–168. [PubMed: 1606538]
15. Mohler JL, Partin AW, Epstein JI, Lohr WD, Coffey DS. Nuclear roundness factor measurement for assessment of prognosis of patients with prostatic carcinoma. II. Standardization of methodology for histologic sections. *The Journal of urology*. 1988; 139(5):1085–1090. [PubMed: 3283380]
16. Mohler JL, Partin AW, Lohr WD, Coffey DS. Nuclear roundness factor measurement for assessment of prognosis of patients with prostatic carcinoma. I. Testing of a digitization system. *The Journal of urology*. 1988; 139(5):1080–1084. [PubMed: 3361653]
17. Tam WL, Weinberg RA. The epigenetics of epithelial-mesenchymal plasticity in cancer. *Nature medicine*. 2013; 19(11):1438–1449.

18. Blanco MJ, Moreno-Bueno G, Sarrio D, Locascio A, Cano A, Palacios J, Nieto MA. Correlation of Snail expression with histological grade and lymph node status in breast carcinomas. *Oncogene*. 2002; 21(20):3241–3246. [PubMed: 12082640]
19. Savagner P, Yamada KM, Thiery JP. The zinc-finger protein slug causes desmosome dissociation, an initial and necessary step for growth factor-induced epithelial-mesenchymal transition. *The Journal of cell biology*. 1997; 137(6):1403–1419. [PubMed: 9182671]
20. Wellner U, Schubert J, Burk UC, Schmalhofer O, Zhu F, Sonntag A, Waldvogel B, Vannier C, Darling D, zur Hausen A, Brunton VG, Morton J, Sansom O, Schuler J, Stemmler MP, Herzberger C, Hopt U, Keck T, Brabletz S, Brabletz T. The EMT-activator ZEB1 promotes tumorigenicity by repressing stemness-inhibiting microRNAs. *Nature cell biology*. 2009; 11(12):1487–1495.
21. Yang J, Mani SA, Donaher JL, Ramaswamy S, Itzykson RA, Come C, Savagner P, Gitelman I, Richardson A, Weinberg RA. Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. *Cell*. 2004; 117(7):927–939. [PubMed: 15210113]
22. Morel AP, Lievre M, Thomas C, Hinkal G, Ansieau S, Puisieux A. Generation of breast cancer stem cells through epithelial-mesenchymal transition. *PloS one*. 2008; 3(8):e2888. [PubMed: 18682804]
23. Mani SA, Guo W, Liao MJ, Eaton EN, Ayyanan A, Zhou AY, Brooks M, Reinhard F, Zhang CC, Shipitsin M, Campbell LL, Polyak K, Briskin C, Yang J, Weinberg RA. The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell*. 2008; 133(4):704–715. [PubMed: 18485877]
24. Polyak K, Weinberg RA. Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits. *Nature reviews Cancer*. 2009; 9(4):265–273.
25. Roca H, Hernandez J, Weidner S, McEachin RC, Fuller D, Sud S, Schumann T, Wilkinson JE, Zaslavsky A, Li H, Maher CA, Daignault-Newton S, Healy PN, Pienta KJ. Transcription factors OVOL1 and OVOL2 induce the mesenchymal to epithelial transition in human cancer. *PloS one*. 2013; 8(10):e76773. [PubMed: 24124593]
26. Bednarz-Knoll N, Alix-Panabieres C, Pantel K. Plasticity of disseminating cancer cells in patients with epithelial malignancies. *Cancer metastasis reviews*. 2012; 31(3-4):673–687. [PubMed: 22733306]
27. Chao Y, Wu Q, Acquafondata M, Dhir R, Wells A. Partial mesenchymal to epithelial reverting transition in breast and prostate cancer metastases. *Cancer microenvironment: official journal of the International Cancer Microenvironment Society*. 2012; 5(1):19–28. [PubMed: 21892699]
28. Peinado H, Ballestar E, Esteller M, Cano A. Snail mediates E-cadherin repression by the recruitment of the Sin3A/histone deacetylase 1 (HDAC1)/HDAC2 complex. *Molecular and cellular biology*. 2004; 24(1):306–319. [PubMed: 14673164]
29. von Burstin J, Eser S, Paul MC, Seidler B, Brandl M, Messer M, von Werder A, Schmidt A, Mages J, Pagel P, Schnieke A, Schmid RM, Schneider G, Saur D. E-cadherin regulates metastasis of pancreatic cancer in vivo and is suppressed by a SNAIL/HDAC1/HDAC2 repressor complex. *Gastroenterology*. 2009; 137(1):361–371. 371 e361-365. [PubMed: 19362090]
30. Veltri RW, Christudass CS. Nuclear morphometry, epigenetic changes, and clinical relevance in prostate cancer. *Advances in experimental medicine and biology*. 2014; 773:77–99. [PubMed: 24563344]
31. Bracken AP, Dietrich N, Pasini D, Hansen KH, Helin K. Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. *Genes & development*. 2006; 20(9):1123–1136. [PubMed: 16618801]
32. Cao Q, Yu J, Dhanasekaran SM, Kim JH, Mani RS, Tomlins SA, Mehra R, Laxman B, Cao X, Yu J, Kleer CG, Varambally S, Chinnaiyan AM. Repression of E-cadherin by the polycomb group protein EZH2 in cancer. *Oncogene*. 2008; 27(58):7274–7284. [PubMed: 18806826]
33. Zouridis H, Deng N, Ivanova T, Zhu Y, Wong B, Huang D, Wu YH, Wu Y, Tan IB, Liem N, Gopalakrishnan V, Luo Q, Wu J, Lee M, Yong WP, Goh LK, Teh BT, Rozen S, Tan P. Methylation subtypes and large-scale epigenetic alterations in gastric cancer. *Science translational medicine*. 2012; 4(156):156ra140.
34. Coolen MW, Stirzaker C, Song JZ, Statham AL, Kassir Z, Moreno CS, Young AN, Varma V, Speed TP, Cowley M, Lacaze P, Kaplan W, Robinson MD, Clark SJ. Consolidation of the cancer

- genome into domains of repressive chromatin by long-range epigenetic silencing (LRES) reduces transcriptional plasticity. *Nature cell biology*. 2010; 12(3):235–246.
35. Gann PH, Deaton R, Amatya A, Mohnani M, Rueter EE, Yang Y, Ananthanarayanan V. Development of a nuclear morphometric signature for prostate cancer risk in negative biopsies. *PloS one*. 2013; 8(7):e69457. [PubMed: 23922715]
36. Bista RK, Wang P, Bhargava R, Uttam S, Hartman DJ, Brand RE, Liu Y. Nuclear nano-morphology markers of histologically normal cells detect the “field effect” of breast cancer. *Breast cancer research and treatment*. 2012; 135(1):115–124. [PubMed: 22706633]
37. Veltri RW, Khan MA, Miller MC, Epstein JI, Mangold LA, Walsh PC, Partin AW. Ability to predict metastasis based on pathology findings and alterations in nuclear structure of normal-appearing and cancer peripheral zone epithelium in the prostate. *Clinical cancer research: an official journal of the American Association for Cancer Research*. 2004; 10(10):3465–3473. [PubMed: 15161703]
38. Makarov DV, Isharwal S, Sokoll LJ, Landis P, Marlow C, Epstein JI, Partin AW, Carter HB, Veltri RW. Pro-prostate-specific antigen measurements in serum and tissue are associated with treatment necessity among men enrolled in expectant management for prostate cancer. *Clinical cancer research: an official journal of the American Association for Cancer Research*. 2009; 15(23):7316–7321. [PubMed: 19934305]

Figure 1.
PC3-Epi and PC3-EMT display distinct phenotypes in murine tibia injections (A) PC3-Epi
(B) PC3-EMT

Figure 2.

PC3-Epi and PC3-EMT are statistically separable by univariate analysis at the population level by two sample Kolmogorov-Smirnov tests. (A) M2 Colocalization (B) Heterogeneity (C) Minimum Intensity (D) Mean Caliper (E) Box-XY. “D+/-” are Kolmogorov-Smirnov statistics quantifying if PC3-Epi is stochastically greater (D+) or less (D-) than PC3-EMT for a selected feature; one-tailed p-values are given.

Figure 3. Principal components analysis of reduced feature set for nuclear size, texture, and nuclear content features or correlates thereof

Gaussian Mixture Model in Canonical Variate Space

Figure 4.

Gaussian mixture model predicts two clusters in pooled phenotype data. Probability density contour overlaid on phenotype data plotted in maximal separation canonical variate space. Presence of two peaks indicates separability of data.

Representative Receiver Operating Characteristic Curve – 80% Training Set

Figure 5.

Receiver operating characteristic (ROC) curves for three classifier methods show high separation between PC3-Epi and PC3-EMT. *k*-nearest neighbors shows strong performance despite stringent nature of classifier method. *k* chosen by minimizing *n*-fold validation loss (mean square error) (see Fig. S1). Multinomial logistic regression and support vector machine (SVM) in canonical variate space showed best performance. Area under ROC curve (AUC-ROC) listed in Table 2.

Table 1Reduced Feature Set for *in vivo* Analysis

#	Feature Name	Feature Type	P-Value (Mann-Whitney <i>U</i>)
1	M2 Colocalization	Nuclear Texture/ Nuclear Content	7.98E-69
2	Intensity, Min (lum)	Nuclear Texture	2.41E-53
3	Heterogeneity	Nuclear Texture	1.53E-34
4	Caliper, Mean (pix)	Nuclear Size	5.06E-22
5	Box XY	Nuclear Size	6.31E-15

Table 2

Performance of classifier methods demonstrates separation of PC3-Epi and PC3-EMT; AUC-ROC - area under receiver-operating characteristic curve

<i>in vivo</i> (80% training set, 100 permutations)	
Classification Method	AUC-ROC
k-Nearest Neighbors (k = 20)	0.820 $\pm$ 0.029
Support Vector Machine in Canonical Variate Space	0.953 $\pm$ 0.014
Multinomial Logistic Regression	0.955 $\pm$ 0.013