Prostate cancer invasion and metastasis: insights from mining genomic data

Bryan D. Hudson, Kristen S. Kulp and Gabriela G. Loots

Advance Access publication date 22 July 2013

Abstract

Prostate cancer (PCa) is the second most commonly diagnosed malignancy in men in the Western world and the second leading cause of cancer-related deaths among men worldwide. Although most cancers have the potential to metastasize under appropriate conditions, PCa favors the skeleton as a primary site of metastasis, suggesting that the bone microenvironment is conducive to its growth. PCa metastasis proceeds through a complex series of molecular events that include angiogenesis at the site of the original tumor, local migration within the primary site, intravasation into the blood stream, survival within the circulation, extravasation of the tumor cells to the target organ and colonization of those cells within the new site. In turn, each one of these steps involves a complicated chain of events that utilize multiple protein–protein interactions, protein signaling cascades and transcriptional changes. Despite the urgent need to improve current biomarkers for diagnosis, prognosis and drug resistance, advances have been slow. Global gene expression methods such as gene microarrays and RNA sequencing enable the study of thousands of genes simultaneously and allow scientists to examine molecular pathways of cancer pathogenesis. In this review, we summarize the current literature that explored high-throughput transcriptome analysis toward the advancement of biomarker discovery for PCa. Novel biomarkers are strongly needed to enable more accurate detection of PCa, improve prediction of tumor aggressiveness and facilitate the discovery of new therapeutic targets for tailored medicine. Promising molecular markers identified from gene expression profiling studies include HPN, CLU1, WTI, WNT5A, AURKA and SPARC.

Keywords: Prostate cancer; metastasis; bone metastasis; biomarkers; microarrays; RNA-seq

INTRODUCTION TO PROSTATE CANCER AND GENOMIC SURVEY OF GENE EXPRESSION

Prostate cancer (PCa) is the second most commonly diagnosed form of cancer and the second leading cause of cancer-related deaths among men worldwide [1]. Most PCa-related deaths result from metastasis, which is defined as the spread of cancer cells from the original site to a new site [2]. The location of the secondary tumor is not random; PCa tends to spread to particular organs or tissues at a rate that is higher than would be expected based on chance. PCa metastasizes almost exclusively to bone and typically forms osteoblastic lesions that are defined by an increase in bone growth [3–5]. Most current treatments for individuals with bone metastases have only palliative effects, with no consequence on long-term survival [3, 6]. Identifying the mechanisms of PCa metastasis is essential for developing new and effective treatments for ameliorating the deadly consequences of the disease and increasing quality of life.

Cancer metastasis is a complex cascade that includes multiple major events: local migration within the primary site and intravasation into the...
blood stream, survival within the circulation, extravasation of the tumor cells to the target organ and colonization of those cells within the new site. Each step of metastatic cancer physiology involves a complicated chain of events that utilize multiple protein–protein interactions, protein signaling cascades and transcriptional changes [4, 7] (Figure 1). Each one of these steps, in turn, involves a myriad of cascade events for their function as well. Comprehensively, metastasis is a very complex set of events that hijack multiple molecular components of the body’s normal physiological functions for its own altruistic purposes. This review presents a general overview of the identification of genes and proteins potentially involved in PCa at each stage of metastasis and their possible uses as biomarkers and/or therapeutic agents/targets.

Considering that PCa metastasis is a multistep event that is likely controlled by many different molecular pathways, the use of high-throughput genomic approaches focused on transcriptional profiling has the potential to unveil key molecular contributors such as biomarkers and therapeutic candidates. Two major experimental methods employed to survey global gene expression in PCa are microarray-based (analog) gene expression profiling technology (microarray) and next-generation sequencing-based gene expression profiling (RNA-Seq) [8]. These methods permit the simultaneous analysis of expression levels for tens of thousands of genes, and, in turn, provide an opportunity to identify differentially expressed genes as a function of cancer progression. In addition, genomics allows for the identification of differential gene expression profiles during multiple experimental designs such as identifying the effectiveness of anti-cancer agents on cancer-related genes in tumor tissue, circulating metastatic cancer in the blood and even what the changes are in PCa cells when exposed to a secondary tumor site, such as the bone. This technology has contributed to the more accurate development of therapeutic strategies and has helped to determine the molecular mechanism(s) of action of metastatic cancer cells. Although microarrays have dominated the past decade of the ‘omics’ era, RNA-Seq is likely to replace this analog technology completely in the near future. Regardless, these techniques are currently being used to elucidate the genomic alterations in both the cancer tumor/cell itself and of the surrounding tissue affected by the cancer during metastasis. To date, many genes/proteins have been identified as key regulators of PCa metastasis using microarray analysis. These findings are summarized below.

### Proliferation and intravasation from the primary tumor

The initial stages of metastasis involve the proliferation of the primary tumor, detachment and migration of malignant cells and their entry into the nearby blood or lymphatic vessels. In the normal prostate gland, epithelial cells have restricted migratory capability due to their relatively strong adherence to each other as well as to the extracellular matrix (ECM). During the process of malignant transformation, however, the adhesiveness of epithelial cells decreases and change from being epithelial in nature to mesenchymal (EMT) [9, 10], thus allowing the cells to have increased migratory properties. In addition to local adhesion and migration, PCa cells degrade the ECM, thus allowing for intravasation into the surrounding tissue and blood circulation.

In an attempt to understand the alterations in gene expression that may signal the initial stages of metastasis, many investigators have studied gene expression correlations that compare the genomic profiles of neoplastic versus non-neoplastic prostate tissue. These tissues can be obtained in a variety of ways including gross dissection, needle biopsies and laser-capture microdissection (LCM) from metastatic PCa patients [11–27].

### Macrodisssection of prostate tissue

The majority of research looking at the genomic profile of PCa has investigated alterations in large sections of surgically removed tissue. For example, Lapointe et al. [15] published a well-designed study in which the authors compared primary prostate tumors with matched normal prostate tissue and unmatched pelvic lymph node metastases. Analysis identified a subset of genes deregulated in high-grade PCa samples as compared with low-grade and normal prostate tissue, including several genes involved in cellular invasion and/or angiogenesis, such as coagulation factor II (thrombin) receptor (F2R) [28, 29], mucin 1 cell surface associated (MUC1) [30–32], neuropilin-1 [33, 34], lipooxygenase (LOX) [35, 36], angiopoetin 2 (ANGPT2) [37, 38] and plexin domain-containing 1 (TEM7) [39, 40]. Of note, MUC1 has been examined in detail over the past decade and has been shown to increase the...
metastatic potential of cancer cells by reducing E-cadherin and integrin-mediated cell adhesion [41, 42]. MUC1 has been shown to be involved in EMT [43] and has been described as a promising tumor-associated antigen for future PCa therapy [44]. In support, ANGPT2, an important gene involved in angiogenic signaling, was also found to be upregulated in a concurrent paper comparing metastatic PCa derived from bone versus PCa that had metastasized to the liver and lymph [45, 46]. Therefore, the genes identified in this article repeatedly have been shown to be involved in PCa metastasis, validating their results and indicating the importance of microarray data in distinguishing PCa-associated factors.

A more recent paper profiled gene expression of normal prostate versus tissue from high-grade (Gleason Score 7) PCa tissue [47]. The authors identified 378 upregulated and 741 downregulated genes between the two groups. Deregulated genes included those that have been found to be involved in cellular invasion, survival and proliferation. Of note, caveolin 1 (CAV1), a scaffold protein involved in T-cell activity and β-catenin recruitment in the WNT signaling pathway [48], was upregulated, whereas B-cell CLL/lymphoma 2 (BCL2), a pro-apoptotic gene [49] and multiple fibulin genes (FBLN1, FBLN4 and FBLN5), ECM proteins involved in cell adhesion and migration [50], were all found to be downregulated. The alteration of

**Figure 1:** Process of PCa bone metastasis and genes altered during each step as identified by genomic profiling. Cancer cells proliferate, invade into surrounding tissues and escape from primary site through the regulation of oncogenes, epithelial to mesenchymal transition, angiogenesis and other factors. Cancer cells enter into the circulation (intravasation), survive and migrate toward bone via upregulation of chemokines and bone-derived factors. Cancer cells then invade through blood vessels (extravasation) and into the bone target. It adheres with bone matrix by integrins and cadherin and breakdown the matrix via MMPs and urokinase-type plasminogen activator (PLAU). Then cancer cells can survive, proliferate and differentiate through molecular interactions with bone microenvironment, which ultimately lead to the establishment of bone metastases.
these genes, and many others, was further validated by real time polymerase chain reaction (rtPCR).

**Needle biopsies**

Microarray-based studies that use tissues acquired from radical prostatectomy specimens and organ donors, such as the studies summarized above, have a number of disadvantages. For instance, surgical manipulation of the malignant tissue often results in altered gene expression of molecules associated with ischemia and hypoxia following devascularization [51] and can therefore introduce tumor-irrelevant changes in transcript levels. Transrectal, ultrasound-guided needle biopsies avoid this particular problem by microdissecting neoplastic prostate epithelium directly from the malignant cancer. The biopsies are frozen immediately after extraction, thus avoiding artifacts due to hypoxia. The first paper published on this technique, Qian et al. [25], reported 954 differentially transcribed genes in PCa samples, none of which were found to be associated with ischemic or surgical manipulation of the prostate gland. Among the cancer-related genes identified in this study were hpsin, v-ets erythroblastosis virus E26 oncogene homolog (avian) (ERG) and α-methylacyl coenzyme A racemase (AMACR); all found to be significantly upregulated, whereas glutathione S-transferase π (GSTP1) expression was found to be downregulated. All these genes were previously found to be altered in independent studies of gross dissections of prostate biopsies [23, 52–54] and have also repeatedly been found to be involved in cancer cell growth and invasion [55–60].

**Laser-capture microdissection**

It is extremely important to note that tumors consist of a highly heterogeneous population of cells and genomic events within the tumor likely vary dramatically depending on the composition of the tissue that is measured during analysis; whole tissue homogenization may not permit the identification of changes critical to metastasis that may occur in a smaller subset of cells potentially more critical in tumor aggressiveness [61]. LCM, first established for gene validation by Shukla et al. [62], allows the separation of benign and malignant epithelial structures from the healthy stromal elements so that they may be examined individually. This process not only reduces contamination of the desired sample by other cellular elements but also maps the identified changes to specific tissue regions. This technique therefore allows the identification of gene changes in an underrepresented number of cells that would otherwise be diluted out within the whole tissue. For instance, since the prostate gland is composed of epithelia and the stroma, each with distinct physiological roles, Gregg et al. [63] investigated transcriptional changes between these two tissues in paired tumor and normal prostate tissues. Although the authors identified 3452 genes differentially expressed in the cancer tissue relative to normal, only 583 of these genes were unique to the PCa epithelial cells and thus comprised unique gene expression patterns found only within the cancer tissue itself and not its surrounding healthy tissue.

Wilms tumor 1 (WT1), a transcription factor involved in E-cadherin expression, was found to be highly upregulated in the neoplastic prostate tissue. WT1 has previously been identified as being upregulated in PCa tissue [64, 65] and in a PCa cell line [65]. In addition, WT1 expression has been shown to be elevated in diverse cancer types, including leukemia [66, 67], breast [64, 67, 68], ovarian [69], mesothelioma and pulmonary adenocarcinomas [70]. A recent paper by Brett et al. [71] showed convincingly that WT1 levels were significantly increased in highly invasive PCa cell lines, that WT1 was sufficient to dampen E-cadherin transcript level and that WT1 can significantly enhance the migration of a non-invasive PCa cell line, LNCaP. It is also important to note that the microarray analysis in the Gregg et al. paper [63] showed that the gene profiles from their microdissected tumor epithelial samples differed from previously published data. For example, SPARC expression had been shown to be upregulated in homogenized PCa biopsy samples [72, 73]. However, in Gregg et al. [63], SPARC was only upregulated in the surrounding stromal tissue. Similarly, CAV1 was also only upregulated in the surrounding stromal tissue, in contrast to the Balacescu et al. paper described previously. These results strongly suggest that some of the tumor genes reported in the literature may be derived from the stromal cell compartment and not the neoplastic epithelial tissue [47].

**RNA-seq**

The advent of next-generation RNA-seq allows for a more detailed window into the transcriptome, including novel transcripts such as noncoding RNAs not measured by conventional analyses [74–76]. With the goal of identifying novel genomic rearrangements, Maher et al. [77] reported using
RNA-seq to interrogate the whole cellular transcriptome of breast cancer cell lines and a small number of PCa cell lines and tumors. The study identified 11 novel PCa-specific gene fusions (6 in PCa cell lines and 5 in primary patient samples). A subsequent study by Pflueger et al. [78] confirmed the upregulation of the TMPRSS2–ERG fusion protein and identified an additional ETS fusion using RNA-seq [androgen-inducible tumor suppressor-(NDRG1)] ERG in human biopsies. A follow-up report by Pflueger et al. [79] investigating non-ETS gene fusions in human PCa reported the importance of IKKB (IKK-b) as being significantly upregulated transcript in metastatic tissue, a finding confirmed by the therapeutic value of IKK inhibitors in various malignancies [80]. Interestingly, recent studies have identified a few mutations in key genes in ETS-negative PCa such as serine peptidase inhibitor, Kazal-type 1 [164] and speckle-type POZ protein [165]. Mutations in these genes increase invasion and suggest a functional role of these genes in metastasis and the formation of a novel subtype of PCa. More recently, Palanisamy et al. [81] published rearrangements for v-raf-1 murine leukemia viral oncogene homolog 1 (RAF) genes in a small percentage of PCa samples, and Beltran et al. [82] found gene amplification of aurora kinase A (AURKA) and the v-myc myelocytomatosis viral-related oncogene, neuroblastoma-derived (avian) (MYCN) in a cohort of prostate adenocarcinoma tissue samples. These oncogenes are critical for cell proliferation and their aberrant expression has been correlated with increased invasiveness of PCa [83–85].

The ability to study noncoding RNA makes RNA-seq a powerful tool in the identification of potential regulators of PCa. Of particular interest, Prensner et al. [76] discovered the prostate-specific noncoding RNA gene, PCAT-1, which was found to be overexpressed in a large subset of metastatic PCa and may contribute to cell proliferation in these tumors. Furthermore, the authors also mention that PCAT-1 resides in the 8q24 ‘gene desert’ locus, a region that has been noted for having a high number of cancer risk single nucleotide polymorphisms (SNPs) and is the location of the c-MYC oncogene.

**Cancer cells in the circulation**

After departing the prostate and upon entering the bloodstream, PCa cells quickly adapt to this new environment by elevating the levels of pro-survival genes and reducing the levels of apoptosis-inducing genes. When a typical epithelial cell leaves its home tissue, it undergoes a particular form of apoptosis called anoikis. Resistance to anoikis and ectopic survival is a prerequisite for invasion and metastasis to occur. Cancer-related factors involved in cell survival, along with the circulating tumor cells (CTCs) themselves, can be identified in the blood stream and can conceivably play an important role in early detection of metastatic PCa in a clinical setting. In a recent paper, Liong et al. [86] attempted to tackle the daunting task of identifying blood-based markers of metastatic cancer from patients with high-grade (Gleason score 7 (4 + 3) – 10) PCa using gene microarrays and discovered seven possible deregulated serum biomarkers, which included cytotoxic and regulatory T-cell molecule (CRTAM), chemokine (C-X-C motif) receptor 3 (CXCR3), Fc receptor-like 3 (FCRL3), KIAA1143, Kruppel-like factor 12 (KLF12), transmembrane protein 204 (TMEM204) and SAM domain, SH3 domain and nuclear localization signals 1 (SAMSN1). These genes are mainly involved in immune responses, chemotaxis and transcriptional regulation in carcinogenesis [87–89]. The Liong study found CRTAM to be significantly underexpressed in blood taken from patients with aggressive PCa, suggesting a possible role for T-cell deficiency. Further support for their conclusions arises from studies that showed altered KLF expression in tumors and during tumor progression [90–93]. CXCR3 was also found to be highly upregulated in human PCa biopsies both at the mRNA and protein level [94]. Furthermore, CXCR3 activation promoted cell motility and invasiveness through a Matrigel matrix barrier in an *in vitro* model using DU-145 and PC3 PCa cell lines via PLCβ3 and μ-calpain activation [94].

PCa cells that invade the blood stream are also known as CTCs. There is an increasing body of evidence showing that CTCs are often present in patients with high-grade PCa [95–98]. Several methods have been developed to collect CTCs for study as markers for disease or metastasis. One approach is to capture the CTCs with a microfiltration system [99]. This system differentiates the PCa-specific CTCs from other cell types based on the size selection of the large epithelial-like cells combined with a lack of the CD45 leukocyte marker. The trapped cells can then be collected individually with a micromanipulator device and further assessed for gene expression profiles. Although this technique has been around for over half a decade, it has only been sparingly applied...
to analyze the transcriptome of circulating breast cancer [100–104] and ovarian cancer cells [105, 106].

Gene expression studies of CTCs collected from PCa patients have been limited to a small subset of genes. For example, a recent study analyzed only 84 EMT-related genes in a microfluidics-based microarray system following CTC isolation and collection [107]. Although somewhat limited in scope, multiple genes involved in PCa survival and invasion were identified including GSK3β, WNT5A, EGFR, MMP9, IGFR1, FOXA2, TCF3, SPP1, FOLH1, PIM2 and ACP5. Increases in GSK3β are known to be associated with pro-survival of PCa through the mammalian target of rapamycin (MTOR) signaling pathway [108]. GSK3β is also a regulator of cellular proliferation by route of being a key member of the canonical WNT signaling pathway and its downstream lymphoid enhancer-binding factor/T-cell specific transcription factors (LEF/TCF) [109]. Furthermore, GSK3β helps regulate the activation of β-catenin. Alterations in Wnt-induced β-catenin have been detected in 85% of all PCa patients with skeletal metastases [110]. Wnt-β-catenin signaling has also been shown to enhance tumorigenicity of phosphatase and tensin homolog (PTEN)-mutated PC3 cells by inducing Akt activity [111]. On a similar note, WNT5A has been shown to be involved in PCa invasion [112, 113], although the mechanism is still under debate and appears to differ greatly based on the experimental model. Another gene identified in the Chen et al. study is SPP1, which encodes for osteopontin, a protein involved in the attachment of osteoclasts to the mineralized bone matrix, via its high-affinity binding to hydroxyapatite. Osteopontin is found at significant levels throughout the bone skeleton [114] and has been reported to promote PCa invasion and to be associated with poor survival [115]. Osteopontin’s role in bone cellular adhesion makes it a particularly attractive candidate as a therapeutic target for blocking metastatic PCa adhesion to the bone matrix.

**Adherence and extravasation**

Exit from the blood stream (extravasation) into the microenvironment of the bone is the most poorly understood step of PCa metastasis. Invasion of PCa cells through the microvasculature is an active process of navigation that requires multiple sequential steps of adhesion, protease degradation of the surrounding matrix and migration/invasion through the surrounding tissue. Factors secreted by PCa cells, such as those found to be altered in the osteolytic PC3 cell line, desmin, I-plastin, laminin-5, insulin-like growth factor-binding protein 4, myosin light chain kinase 2 and clusterin [116], are thought to alter bone homeostasis by disrupting the balance between bone growth and bone resorption [117–120]. Therefore, the interaction between PCa cells and cells within the bone microenvironment is suspected to play a critical role in secondary tumor formation and progression since tumor cells need to physically communicate with the local microenviroment in order to properly adhere, invade and proliferate in the metastatic site [121, 122].

One method for studying the communication between PCa cells and bone is through co-culturing these two cell types. Wang et al. [123] investigated bone marrow stromal cells isolated from the calvaria of neonatal mice co-cultured with the highly invasive, androgen-insensitive PC3 cells. The cells were co-cultured for 48 h either in direct physical contact or separated by a permeable membrane. Microarray analysis showed that several gene families and signal transduction pathways were affected by the interaction between PCa cells and bone cells including: TGFβ superfamily genes (TGFβ receptor and a downstream signal transduction molecule Smad9); colony-stimulating factors (CSF-1 and its receptor, CSF-1R); ECM protein and cell adhesion molecules (collagen types III, IV, VIII and XII); other related cell adhesion molecules (integrin α1 and integrin α2); matrix proteases (cathepsinK, MMP-2, MMP-9, and uPA) and MAPK signal transduction molecules (MKK4, mta-1 and raf-1). As anticipated, the two co-culture conditions (physical versus nonphysical contact) produced different results, although there were several differentially expressed genes shared by both comparisons. These results suggest that some of the genes that regulate the invasive potential of PCa cells may be regulated by factors secreted by the microenvironment of the secondary metastatic site [123].

A potential limitation to this experimental design is that alterations in gene expression levels may be derived from the cancer cells coming into contact with any foreign cell type. Therefore, these changes may not be specific to bone. To address this issue, Zhang et al. [124] compared the gene expression of PC3 cells co-cultured with bone cells to those cultured with normal human prostate stroma. Only two genes were identified as being specific to contact with bone-derived cells: an ECM gene, Collagen
uPA has been shown to facilitate cell migration and invasion not only of malignant cancer cells but during normal, physiological events as well [125–127]. Additionally, uPA overexpression has been shown to be associated with the more aggressive PCA cell lines [128, 129], is correlated with poor patient prognosis [130] and is constitutively secreted by bone-homing PCA cell lines, such as PC3 [125]. Even more convincingly, uPA was found to be expressed at a higher level in PCA metastasis from the bone as compared with PCA metastases from the liver and lymph nodes [45].

An additional co-culture experiment, investigating a less invasive PCA cell line, LNCaP, found alterations in the expression of many of the oncogenic and survival-related genes following the co-culture with bone derived, conditioned media [131]. For instance, the oncogene survivin (BIRC5) was found to be significantly upregulated, a result confirming a previous study in cancer tissue samples [132]. This gene has formerly been found to increase proliferation, motility and invasion of PCA cells through a β-catenin-mediated process [133] and is targeted by miR–203, a microRNA often silenced in malignant PCA [134]. Furthermore, the authors found that pro-survival genes, such as ESPL1, were upregulated, whereas apoptosis mediators, CRADD and BCAP29, were both downregulated, indicating that the bone microenvironment may have the potential to facilitate/promote the survival of PCA cells.

In a different report, Shirevnyamba et al. [135] compared gene expression of tumors that induce bone loss (PC3) to those that promote bone growth (MDA-PCa2b). This report did not identify any significant differences in gene expression between PC3 cells cultured alone and those co-cultured under bilayer conditions. Nor did they find a significant difference in the gene expression of MDA-PCa2b cells under any culture conditions, whatsoever. Only under direct contact were any alterations in gene profiles found, with only nine genes being significantly different. However, four of these genes have previously been found to participate in osteoclastogenesis: Interleukin–1β (IL–1β), Cyclooxygenase-2 (COX-2), IL–6 and C3. Cyclooxygenase-2 is an inducible prostaglandin synthesis enzyme that is upregulated in many cancers and is associated with high mortality in PCA [136]. In addition, COX-2 is indirectly involved in bone resorption and osteoclastogenesis through its interactions with prostaglandin E2 (PGE2) [137] and is known to increase the PCA invasive potential [138]. Interleukin–6 is a mediator of PGE2-induced suppression of the production of bone growth and is thereby important in the breakdown of bone [139, 140]. It has also been found to be associated with more aggressive PCA [141], is upregulated in the majority of PCA bone metastases and PCA soft tissue metastases [142] and is thought to be involved in growth differentiation factor–15 alterations in the invasive potential of PCA [143]. One of the other four upregulated genes not associated with osteoclastogenesis was secreted protein acidic and rich in cysteine (SPARC; also known as osteonectin), a gene that was previously identified as upregulated in human PCA neoplastic tissue [72, 73]. SPARC has been assumed to be important in human PCA bone metastasis as a major bone-derived chemoattractant for PCA cells [144] and has been found to be involved in PCA cell proliferation and matrix invasion [145]. SPARC’s apparent involvement in both osteoclastogenesis at the level of the bone cell (via stimulation of osteoprotegerin [146, 147]) and within PCA tissue implicates this gene as having a greater role in PCA metastasis than previously reported.

THE NEED FOR DATA MINING
The recent explosion in high-throughput methods that survey global gene expression has created a major analytical bottleneck. In each of the aforementioned studies, the overwhelming amount of data severely limit the authors’ ability to discuss every differentially expressed gene and most manuscripts selectively provide details based on the narrow context of the paper and support from the published literature. There are many genes that statistically fit differential expression cut-offs in these experiments, but are omitted in the discussion, therefore it would be extremely important for future studies to revisit some of these data sets and parallel some of the findings from independently generated experiments to highlight new likely participants in PCA metastasis. Knowledge about these genes could greatly contribute to our interpretation of the molecular mechanisms involved in the progression of PCA and help elucidate possible therapeutic venues for treating metastasizing cancer.

The possibility of overlooked data in large-scale genomic experiments has prompted a number of
data-mining investigations aimed at searching a decade of publically stored genomic databases [148–150]. For example, Chen et al. [149] searched public databases storing genomic data of urine collected from PCa patients. They performed three levels of data mining: microarray gene expression data, gene ontology assignment and pathway enrichment, to highlight 19 putative markers among 3964 transcripts found to be upregulated in PCa samples. This list of 19 proteins present in PCa urine samples was subsequently subjected to pathway enrichment analysis producing 10 candidates that were identified based on co-expression, shared protein domains, co-localization and protein physical interaction relationships, including: RBP4 (retinol binding protein 4, plasma), CFH (complement factor H), ITIH4 (inter-alpha (globulin) inhibitor H4) and FTL (ferritin, light polypeptide), (all linked due to co-localization); and APOD (apolipo-protein D), RBP4 and CRABP1 (cellular retinoic acid-binding protein 1) (all found to share protein domains according to Interpro and Pfam databases). In addition, CYP2B6 (cytochrome P450, family 2, subfamily B, polypeptide 6) was found to connect with four putative markers because of co-localization; C6 (complement component 6) was linked to RBP4 and ITIH4 by co-localization and to CFH and RECK (reversion-inducing cysteine-rich protein with kazal motifs) by sharing the same protein domains; transthyretin was found to physically interact with RBP4, CFH and CLU (clusterin) genes; OSBPL1A (oxysterol-binding protein-like 1A) co-localized with APOD and IGSF8 (immunoglobulin superfamily, member 8); CD70 (CD70 molecule) connected with the candidate marker CD27 (CD27 molecule) by co-localization and physical interaction; CD70 shared a protein domain with C1QTNF3 (Clq and tumor necrosis factor-related protein 3). Furthermore, they found that 10 of the 19 urinary markers were closely associated with tumor cell development, growth and proliferation pathways [CLU, CD27, CFB (complement factor B), RBP4, ITIH4, PECAM1 (Platelet endothelial cell adhesion molecule), RECK, MGAT5 (hypothetical LOC151162), APOD and LGALS3 (Lectin, galectoside-binding, soluble, 3)].

Of the genes listed above, only CLU, LGALS3, MGAT5, RECK and SELENBP1 have been previously involved in PCa invasion. Clusterin is a stress-activated and apoptosis-associated molecular chaperone that functions to protect cells from various stressors and is highly expressed in many human cancers including PCa with high Gleason scores [151]. Clusterin confers therapeutic resistance when overexpressed [152], is upregulated in PC3 cells [116] and is currently being investigated in phase trials as a potential cancer therapeutic [153]. Recently, Shiota et al. reported that CLU knockdown dramatically inhibits PCa invasion and metastatic potential of PC3 cells [154, 155] through a TGF-beta/EMT-mediated process. N-acetylgalactosaminyltransferase V (MGAT5) is a glycoprotein that stabilizes matriptase and is highly expressed in PC3 cells (highly invasive), but not in LNCap (non-invasive). Upregulation of MGAT5 in LNCap significantly enhanced the cells’ invasive ability, whereas knockdown of MGAT5 in PC3 cells attenuated the metastatic potential of these cells [156]. Galectin-3 is a carbohydrate-binding protein whose primary function is thought to be as a scaffold [157]. In human PCa, the levels of galectin-3 cleaved by proteases increase with PCa progression. Furthermore, galectin-3 knockdown in PC3 cells resulted in decreased cell migration, cell invasion and suppressed MMP-2 and MMP-9 activity, at least partially due to its role in stabilizing p21-induced apoptosis [158]. A novel drug, Lac-L-Leu, which binds and inhibits galectins, resulted in a dose-dependent inhibition of PCa metastasis [159]. Reversion-inducing cysteine-rich protein with Kazal motifs (RECK) is a membrane-bound protein that is thought to inhibit MMPs and angiogenesis. When RECK levels are elevated in human tumors, there is generally a reduction in local invasion, metastasis and improved prognosis [160, 161]. Interestingly, in a related study, overexpression of RECK was found to cause a significant reduction in the invasion of DU-145 cells and a decrease of pro-MMP-9 and of pro-/active MMP-14 [162]. Why RECK was found to be upregulated in the analyzed microarrays is still under debate, but may be due to differences in their sample populations and in their experimental techniques.

A recent paper comparing RNA-seq data from multiple studies from a reference database demonstrates the value of identifying PCa-specific splicing [163]. The authors compared RNA-seq data sets generated by two independent groups and discovered that the two data sets are extremely divergent, suggesting either heterogeneity in the cancer types or differences in sample preparation. Due to the massive discrepancies, only two splice events were found, from the genes protein phosphatase 3,
catalytic subunit, alpha isozyme (PPP3CA) (upregulated) and solute carrier family 20 (phosphate transporter), member 2 (SLC20A2) (down-regulated) in a tumor-specific fashion; both genes are purported to be involved in calcium transport and calcium-dependent cellular events, but have not been reported to have a function in PCa metastasis until this article.

**CONCLUSIONS**

PCa is a complex disease that involves the activation of many different transcripts and simultaneous inactivation or repression of tumor suppressor genes. Some of these activated genes are more specifically involved in tumor growth at the primary site, whereas other transcripts are activated and may facilitate events orchestrating intravasation, survival, adhesion or extravasation into the bone microenvironment (Figure 1 and Table 1). In this review, we have summarized transcriptional changes identified during various steps of PCa metastasis quantified either by microarrays or RNA-Seq analysis; although not covered in this review due to space limitations, future reviews covering the alterations in microRNA, epigenomics and proteomics during each stage of metastasis would help make an even more cohesive picture of the critical genes involved in the metastatic potential of PCa. Given the similarities in metastatic outcome of the studies summarized above, one would have predicted that the transcriptional changes identified from patient-derived metastatic tissues would partly resemble those from immortalized tumor cell lines explored in animal models. However, this has not yet been the case, and there continues to be a great discrepancy between patient-derived information, in vitro and in vivo animal models or xenografts. Nonetheless, one common thread among all these studies is the emergence of candidate genes that share ontological similarities, suggesting that particular molecular pathways are mechanistically more relevant during the process of PCa metastasis to bone. In general, transcriptome analysis has greatly expanded the repertoire of candidate promoter and suppressor genes implicated in PCa invasion and metastasis. Although genomic tools and applications have greatly expedited the discovery of candidate biomarkers or therapeutic targets, our ability to improve diagnostics, prognosis or to overcome drug resistance has been at a standstill. Future investments need to focus on improving cell-based or animal models that can be used to characterize the metastatic potential of patient-derived primary tumor cells, and the ability of combinatorial chemotherapies to

**Table 1: List of genes altered during PCa growth and metastasis**

<table>
<thead>
<tr>
<th>Step in metastasis</th>
<th>Tissue/technique</th>
<th>Gene name</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adhesion/intravasation</td>
<td>Primary tumor/macrodisssection</td>
<td>F2R</td>
<td>[28]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MUC1</td>
<td>[30–32]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NRPI</td>
<td>[33, 34]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LOX</td>
<td>[35, 36]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ANGPT2</td>
<td>[37, 38]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TEM7</td>
<td>[39, 40]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CAV1, BCL2, FBLN1, FBLN4, FBLN5</td>
<td>[47]</td>
</tr>
<tr>
<td></td>
<td>Primary tumor/needle biopsy</td>
<td>HPN, ERG, AMACR, GSTPI</td>
<td>[23, 25, 52–54]</td>
</tr>
<tr>
<td></td>
<td>Primary tumor/LCM</td>
<td>WTI, SPARC, CAVI</td>
<td>[63]</td>
</tr>
<tr>
<td></td>
<td>Primary tumor/RNA-seq</td>
<td>TMPRSS2-ERG, NDRG1-ERG</td>
<td>[78]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IKBKB</td>
<td>[79]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RAF</td>
<td>[80]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AURKA, MYCN</td>
<td>[82]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PCAT-1</td>
<td>[76]</td>
</tr>
<tr>
<td>Circulation/survival</td>
<td>Blood/ serum collection</td>
<td>CRTAM, CXCR3, FCLR3, KLF2, TMEM204, SAMSNI</td>
<td>[86]</td>
</tr>
<tr>
<td></td>
<td>Cancer cells/CTCs</td>
<td>GSK3β, WNT5A, EGFR, MMP9, IGFR, FOXA2, TCF3, SPI, FOLRI, PIM2, ACP5</td>
<td>[107]</td>
</tr>
<tr>
<td>Adherence/extravasation</td>
<td>Cancer cells/co-culture with bone cells</td>
<td>TGFβ, SMAD9, CSF1, CSF1R, COL3AI, ITGAI, ITG2, CTSK, MMP2, MMP9, PLAU, MAP2K2, MTAI, RAFI</td>
<td>[123]</td>
</tr>
<tr>
<td></td>
<td>Cancer cells/co-culture with conditioned media</td>
<td>COL3AI, PLAU</td>
<td>[124]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IL1B, COX2, IL2, C3, SPARC</td>
<td>[135]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BIRC5, ESPL1, CRADD, BCAP29</td>
<td>[13]</td>
</tr>
</tbody>
</table>

**Bold indicates gene downregulation.**
remove all cancer cells without selecting for drug-resistant clones. The challenge will be to synergize genomic and functional studies in a useful way such that therapeutic information can be derived timely to save a patient’s life. Future studies and therapeutic strategies need to take into account the dynamic molecular complexity of invasion and metastasis implicating cancer cells, their tissue of origin and their metastatic microenvironment.

**Key points**
- PCs is the second leading cause of cancer-related death in men
- Early detection of PCs increases the survival of the patient
- Novel biomarkers are strongly needed to improve detection, prediction and novel therapeutic targets for PCs treatment
- Global gene expression methods such as gene microarrays and RNA-seq are the next step in the identification of genes altered during the metastasis of PCs.

**Acknowledgements**
The authors would like to thank Nick R. Hum for designing Figure 1. This work was done under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract De-AC52-07NA27344 and was supported by LDRD13-ERD-042.

**References**
28. Yoshida E, Tsuchiya K, Sugiki M, et al. Modulation of the receptor for urokinase-type plasminogen activator in...


