RNA Biomarkers to Facilitate the Identification of Aggressive Prostate Cancer

Kathryn L. Pellegrini\textsuperscript{a}, Martin G. Sanda\textsuperscript{a}, and Carlos S. Moreno\textsuperscript{b,\textdagger}\textsuperscript{*}

\textsuperscript{a}Department of Urology at Emory University School of Medicine and Winship Cancer Institute, Atlanta, GA 30322, USA

\textsuperscript{b}Department of Pathology and Laboratory Medicine at Emory University School of Medicine and Winship Cancer Institute, Atlanta, GA 30322, USA

Abstract

A large number of men are diagnosed with prostate cancer each year, but many will not experience morbidity or mortality as a result of their cancers. Therefore, biomarkers for prostate cancer are necessary to carefully select patients for initial diagnostic biopsy or to facilitate care decisions for men who have already been diagnosed with prostate cancer. RNA-based approaches to biomarker discovery allow the investigation of non-coding RNAs, gene fusion transcripts, splice variants, and multi-gene expression panels in tissue, urine, or blood as opportunities to improve care decisions. This review focuses on RNA biomarkers that are available as commercial assays, and therefore already available for potential clinical use, as well as providing an overview of newer RNA biomarkers that are in earlier stages of clinical development.

Keywords

Prostate cancer; Biomarker; RNA; Non-coding RNA; lncRNA; Gene signatures

1. Introduction

Prostate cancer is the most common non-cutaneous cancer diagnosed in men, with over 200,000 diagnosed cases and almost 28,000 deaths per year in the United States (American Cancer Society, 2015). Biomarkers for prostate cancer are generally used in decisions to perform a prostate biopsy, to select the most appropriate treatment, and to monitor patient responses to treatment. One of the biggest obstacles in prostate cancer diagnosis and treatment is identifying patients with clinically significant cancer requiring aggressive treatment versus patients who have indolent cancers that are unlikely to progress.

\textsuperscript{*}Corresponding author Emory University, Whitehead Biomedical Research Building, 615 Michael Street, Room 105J, Atlanta, GA 30322, USA Phone: +1 404-712-2809; Fax: +1 404-727-8538; cmoreno@emory.edu.

Publisher’s Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
Serum levels of prostate specific antigen (PSA) have been widely used in screening for the early detection of prostate cancer, resulting in the identification and treatment of more prostate cancers. However, the recent Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial, and the European Randomised Study of Screening for Prostate Cancer (ERSPC) found little to no reduction in prostate cancer-specific mortality with systematic PSA screening as compared to opportunistic screening (Andriole et al., 2012; Schroder et al., 2012). Despite concerns about higher than anticipated levels of PSA screening occurring in the control group (Miller, 2012), the US Preventive Services Task Force (USPSTF) recommended against PSA-based screening for prostate cancer in all age groups as any ostensible benefits were outweighed by the associated risks of biopsy and over-treatment of indolent disease (Chou et al., 2011; Moyer and U.S. Preventive Services Task Force, 2012; Smith et al., 2015). Even so, serum PSA is still frequently used. Another PSA-based test, the Prostate Health Index (PHI), which combines total PSA, free PSA, and the [-2]proPSA precursor into a single score, was recently approved by the FDA for the detection of prostate cancer in men with total serum PSA values between 4 and 10 ng/mL, and a digital rectal exam (DRE) with no abnormal findings (Catalona et al., 2011). There are over one million prostate biopsies performed annually in Medicare beneficiaries in the US (which consists of men over the age of 65), and the number of biopsy cores retrieved during each procedure has increased from 6 to 12 (Loeb et al., 2011). However, prostate cancer is only found in one third of patients from initial biopsies, although cancer is eventually detected in two thirds of patients after up to four rounds of biopsy (Welch et al., 2007). These data indicate that there is potential to more accurately identify patients at risk of cancer to reduce unnecessary biopsies.

A number of clinical parameters are considered in directing treatment decisions for localized prostate cancer, but patients are generally grouped as having a low, intermediate, or high risk of progression based on serum PSA, tumor staging and Gleason scoring of the biopsy sample (D’Amico et al., 2003). Results from the Scandinavian Prostate Cancer Group Study Number 4 (SPCG-4) and the Prostate Cancer Intervention Versus Observation (PIVOT) trial indicate that observation may be a better approach than radical prostatectomy for patients with low risk prostate cancer, especially for patients over 65 years in age (Bill-Axelson et al., 2014; Wilt et al., 2012). In contrast, men with high risk cancers, or a longer life expectancy are more likely to benefit from the immediate initiation of curative intent therapies such as radical prostatectomy, external beam radiation therapy, or brachytherapy (Bill-Axelson et al., 2014; Wilt et al., 2012). However, complications such as urinary incontinence or sexual dysfunction commonly impact health-related quality of life following these types of curative interventions for prostate cancer, and emphasize the importance of accurately discerning patients at risk of prostate cancer-specific mortality (Potosky et al., 2000; Sanda et al., 2008). While the identification of men with low risk cancers is generally accomplished very well using current clinical parameters, there are many men with intermediate risk cancers that will not experience progression of their tumors. If there were biomarkers available that were better able to distinguish amongst patients diagnosed with intermediate risk prostate cancers to identify men that have low or high probability of experiencing prostate cancer-associated morbidity and mortality, curative interventions could be more specifically directed to avoid over-treatment of indolent cancers.
The measurement of gene expression is an attractive source for the discovery of biomarkers because it allows the evaluation of tumor activity, potentially identifying various molecular subtypes of cancer. Although both protein and RNA provide information about molecular activity, RNA is generally much easier to detect and quantitate, even at low amounts. Additionally, the multiplexing of RNA-based assays is relatively straightforward, which means that many targets can be assessed at once. RNA-based approaches to prostate cancer biomarker discovery include the analysis of non-coding RNAs, gene fusion transcripts, alterations in the presence of splice variants, and multi-gene expression panels based on various tumor cell functions. For many of these targets, expression in urine, blood, or circulating tumor cells have also been investigated as a way to non-invasively detect or track prostate cancer.

2. Non-Coding RNAs

Although the majority of the genome is transcribed, only a fraction of these transcripts are protein-coding and, the ongoing annotation of non-coding RNAs (ncRNAs) means that the non-coding subset of RNA represents a growing resource for the identification of disease-associated genes (Iyer et al., 2015; The ENCODE Project Consortium, 2007). GENCODE v21 contains annotations for 9,534 small non-coding RNAs, which are less than 200 nucleotides in length and include a number of distinct categories such as microRNAs (miRNAs) and small nucleolar RNAs (snoRNAs), while the long non-coding RNA (lncRNA) compartment, defined as being longer than 200 nucleotides, is composed of 15,877 lncRNA genes encoding for 26,414 lncRNA transcripts (Derrien et al., 2012; Harrow et al., 2012). However, recent reports suggest that there may be closer to 60,000 lncRNA genes in the human genome (Iyer et al., 2015). The study of ncRNA biomarkers in prostate cancer has been predominantly focused on the lncRNA PCA3, although the continued discovery and annotation of ncRNAs in the human genome means it is likely that many more ncRNAs will be found to have an association with prostate cancer.

2.1. PCA3

The most well characterized of the next generation of prostate cancer biomarkers is the lncRNA PCA3, which recently received FDA approval for decisions about repeat biopsy (Groskopf et al., 2006). The over-expression of PCA3 in prostate cancer was originally established by differential display analysis of patient-matched malignant and benign tissues collected during radical prostatectomy, where it was found to be 10- to 100-fold higher in the cancerous tissues for 53 of the 56 patients studied (Bussemakers et al., 1999). Another aspect of PCA3 that made it an attractive biomarker candidate was the absence of PCA3 expression in normal or malignant tissues of non-prostate origin, including genitourinary sources such as kidney, bladder, seminal vesicles, and testis (Bussemakers et al., 1999; de Kok et al., 2002). Additionally, PCA3 expression was not increased in patients diagnosed with prostate conditions other than cancer such as benign prostatic hyperplasia (BPH), atypical small acinar proliferation (ASAP), prostatic intraepithelial neoplasia (PIN), inflammation, or chronic prostatitis (de Kok et al., 2002; Deras et al., 2008; Vlaeminck-Guillem et al., 2012). The feasibility of PCA3 as a urinary biomarker was first demonstrated by qRT-PCR analysis of urinary sediments from a cohort of 108 men that had been...
recommended for prostate biopsy based on a serum PSA greater than 3 ng/mL (Hessels et al., 2003).

A clinical grade assay, currently available as the Progensa® PCA3 Assay from Hologic GenProbe™, has now been developed and used to investigate the urinary expression of PCA3 in a number of large cohorts. This assay utilizes transcription-mediated amplification to quantify PSA and PCA3 transcripts in whole urine collected post-DRE, so that a PCA3 score can be calculated (Groskopf et al., 2006). Patients with high urinary PCA3 scores have a greater likelihood of being diagnosed with prostate cancer on biopsy (Aubin et al., 2010; Crawford et al., 2012; Deras et al., 2008; Gittelman et al., 2013; Ramos et al., 2013; van Gils et al., 2007), and urinary PCA3 has shown similar diagnostic performance for patients in either high or low serum PSA ranges (Chevli et al., 2014; Deras et al., 2008; Groskopf et al., 2006; Haese et al., 2008; Stephan et al., 2013; Tinzl et al., 2004). Therefore, the added consideration of urinary PCA3 scores can significantly improve the performance of diagnostic models based on clinical variables such as age, family history, prostate volume, DRE findings, and PSA (Aubin et al., 2010; Gittelman et al., 2013; Hansen et al., 2013; Tombal et al., 2013; Wei et al., 2014). In one recent study, the performance of the PCPT (Prostate Cancer Trials and Prevention) risk calculator for the detection of any prostate cancer and high-grade prostate cancer was significantly improved by including the PCA3 score into the model (Wei et al., 2014). Other studies have found that incorporating PCA3 scores with the PSA-based PHI score shows a higher specificity for the detection of prostate cancer on biopsy than either marker alone or over base models that incorporate standard clinical variables (Ferro et al., 2013; Perdona et al., 2013; Stephan et al., 2013). Although PCA3 has shown improved specificity for the detection of prostate cancer when compared to serum PSA, the lower sensitivity of PCA3 means that it is insufficient as a stand-alone biomarker for decisions regarding initial biopsy. However, as shown in a number of the studies discussed here, the inclusion of PCA3 can improve the performance of other clinical factors, which demonstrates the potential for additional biomarkers to provide complementary information during the diagnosis of prostate cancer.

### 2.2. Additional Non-Coding RNAs Under Investigation

A number of additional ncRNAs, including both miRNAs and lncRNAs, have been studied for their association with prostate cancer and potential use as biomarkers (summarized in Table 1). Although these ncRNAs are still in a much earlier stage of understanding and development than PCA3, they have all been investigated in multiple, independently collected cohorts, some of which include very large numbers of participants. However, their performance in these retrospective studies will need to be carefully considered before attempting to test their performance in prospective multicenter trials. Of these ncRNAs, the current evidence for SChLAP1 is most convincing, given that this lncRNA has shown prognostic performance across multiple prostate cancer outcomes including biochemical recurrence, metastasis, and mortality, and was also able to be detected in urine sediments (Prensner et al., 2014b).
3. Gene Rearrangements and Splice Variants

Chromosomal aberrations such as deletions and translocations are frequently observed in cancer, and can often lead to changes in the expression of oncogenes or tumor suppressors. One of the best characterized gene fusions, BCR-ABL, is the result of a chromosomal translocation (also known as the Philadelphia chromosome) commonly observed in chronic myeloid leukemia that is used as a biomarker during diagnosis, and as a biomarker for monitoring patient response during treatment (Rowley, 2001). In contrast, alternative splicing of genes is a normal molecular event that enables different versions of a gene to be expressed and translated. However, there are a number of splicing abnormalities that occur in cancer, including both transcriptome-wide changes in splicing events during cancer and the expression of particular splice variants in specific cancers (Kalnina et al., 2005; Sowalsky et al., 2015; Venables et al., 2009). Gene rearrangements and splice variants are not always able to be detected on microarray, but the use of RNA sequencing for gene expression profiling means that these types of changes can be more clearly recognized and studied for the identification of biomarkers.

3.1. TMPRSS2:ERG Fusion

The ETS transcription family members and proto-oncogenes, ERG and ETV1, were found to be highly expressed in a subset of prostate cancers by Tomlins et al. (2005) using a unique approach, which they termed cancer outlier profile analysis (COPA), to interrogate the Oncomine database. Genetic analysis of prostate cancer cell lines with over-expression of these genes found the 5′ end of the androgen receptor-regulated gene TMPRSS2 to be aberrantly fused to either ERG or ETV1, which was confirmed by FISH in clinical samples for both TMPRSS2:ETV1 (7 out of 29 cases) and TMPRSS2:ERG (16 out of 29 cases) (Tomlins et al., 2005). Follow up studies have shown TMPRSS2:ERG gene rearrangements to be the most commonly found TMPRSS2:ETS family pairing in prostate cancer, although the frequency of this fusion varies significantly amongst different racial groups (Magi-Galluzzi et al., 2011; Tomlins et al., 2006). In Caucasian cohorts, TMPRSS2:ERG fusion is observed in approximately 50% of prostate cancers, whereas in African American or Asian cohorts, the frequency of TMPRSS2:ERG ranges from 20-30% (Khani et al., 2014; Magi-Galluzzi et al., 2011; Miyagi et al., 2010; Mosquera et al., 2009; Rosen et al., 2012). TMPRSS2:ERG fusions are rarely observed in benign tissue, but are observed in areas of high grade prostatic epithelial neoplasia (HGPIN) in close proximity to a concurrent cancer of the same TMPRSS2:ERG status (Cerveira et al., 2006; Mosquera et al., 2008; Perner et al., 2007; Young et al., 2012). Additionally, ERG expression in HGPIN lesions was found to be indicative of subsequent prostate cancer diagnosis with patients positive for ERG expression in their HGPIN lesions significantly more likely to be diagnosed with prostate cancer within three years (Park et al., 2014). These studies demonstrate the specificity of TMPRSS2:ERG for prostate cancer and indicate a role for TMPRSS2:ERG in the development and progression of prostate cancer, which could be applied as a biomarker to identify patients with HGPIN lesions that have a higher likelihood of progressing to prostate cancer.
The association of the TMPRSS2:ERG fusion with more aggressive tumor characteristics is currently unclear, with some studies claiming significant relationships of TMPRSS2:ERG status with Gleason score, biochemical recurrence, and survival (Attard et al., 2008; Demichelis et al., 2007; Lin et al., 2013; Nam et al., 2007), and others finding no significant differences related to the gene fusion (Hoogland et al., 2012; Minner et al., 2011; Perner et al., 2006; Pettersson et al., 2012). However, the combination of urinary TMPRSS2:ERG with PCA3, serum PSA, or with other clinical risk determinants has been shown to improve the identification of high-grade cancer (Attard et al., 2008; Leyten et al., 2014; Tomlins et al., 2011). In a multicenter cohort of 1,065 men receiving prostate biopsies at community and academic medical centers, patients in the group with the highest combined TMPRSS2:ERG and PCA3 scores had a 3.2-fold higher chance of being diagnosed with cancer, and a 5.9-fold greater chance of having high-grade cancer (Gleason score greater than 6) when compared to the group with the lowest scores (Tomlins et al., 2011). A separate multicenter study of 443 men found that using both PCA3 and TMPRSS2:ERG (with score thresholds of ≥25 and ≥0, respectively), would have avoided 153 biopsies (35%) in men that would have otherwise been recommended for prostate biopsy based on their serum PSA levels alone, while missing only 11 (10%) of the high-grade prostate cancers (Leyten et al., 2014). Additionally, including both the TMPRSS2:ERG and PCA3 scores with the PCPT or ERSPC risk calculators significantly improved the performance of these clinical parameters for the detection of patients at higher risk of having prostate cancer prior on biopsy (Leyten et al., 2014; Tomlins et al., 2011). Now that a urine test for TMPRSS2:ERG is commercially available using the same assay platform as PCA3, the advantages of combining these biomarkers for the detection of prostate cancer can be investigated further.

3.2. AR-V7 Splice Variants

Given the importance of the androgen receptor (AR) in the development, progression, and treatment of prostate cancer, the molecular aspects of the AR have been studied extensively with up to 14 different AR splice variants identified thus far (Dehm et al., 2008; Guo et al., 2009; Hu et al., 2009; Hu et al., 2011; Sun et al., 2010; Watson et al., 2010). The majority of studies have focused on the AR-V7 splice variant (also previously referred to as AR3), which has been confirmed to produce a functional protein that is constitutively localized to the nucleus and therefore able to induce the expression of AR-regulated genes such as PSA and TMPRSS2 in the absence of androgens (Guo et al., 2009; Hu et al., 2009; Hu et al., 2011; Watson et al., 2010). This is important in the context of androgen deprivation therapy (ADT), particularly as it has been shown in human cell lines and xenograft tumor models that AR-V7 expression increases in response to castration or other treatments targeting the AR, and that AR-V7 over-expression promotes androgen-independent growth (Guo et al., 2009; Hu et al., 2012; Li et al., 2012; Watson et al., 2010). In clinical specimens, high levels of AR-V7 expression are more frequently observed in patients that have previously received ADT, but then progressed to metastatic castration resistant prostate cancer (mCRPC) (Antonarakis et al., 2014; Hornberg et al., 2011; Hu et al., 2009; Qu et al., 2015; Zhang et al., 2011). The expression of AR-V7 in primary tumors or metastases has been correlated with significantly worse clinical outcomes, including reduced time to biochemical recurrence (Guo et al., 2009; Hu et al., 2009), and earlier cancer-specific mortality.
More recently, AR-V7 expression in circulating tumor cells (CTCs) was shown to be predictive of patient responses to the AR-targeted treatments enzalutamide or abiraterone for the treatment of mCRPC (Antonarakis et al., 2014). Strikingly, none of the patients positive for AR-V7 expression in their CTCs showed a response to either enzalutamide or abiraterone treatment as measured by percentage change in PSA level, while amongst patients with AR-V7 negative CTCs, the response rate was 53% (10 out of 19 patients) for enzalutamide and 68% (17 out of 25 patients) for abiraterone (Antonarakis et al., 2014). Regardless of treatment group, patients positive for AR-V7 also experienced a significantly reduced progression-free survival time as measured by PSA, clinical or radiographic measures, as well as a significantly reduced overall survival (Antonarakis et al., 2014). Although these findings need to be investigated in a larger cohort, they suggest that the expression of AR-V7 in CTCs could be used as a predictive biomarker to inform treatment decisions in mCRPC, particularly in identifying patients that are unlikely to respond to enzalutamide or abiraterone, and therefore need to consider more appropriate treatment strategies.

4. Multi-Gene mRNA Expression Panels

As high-throughput gene analysis has become more accessible, the development and analysis of multi-gene panels in clinical samples has become more feasible. This is especially important in conditions such as cancer, where there is a high degree of heterogeneity both at the molecular level within the tumor, as well as between individual patients (Fraser et al., 2015; Singh et al., 2002). One way to overcome this heterogeneity is to profile the expression of multiple genes at once to establish the molecular processes occurring in the tumor. Although many multi-gene expression panels have been described, we have mostly focused on those that are commercially available for the analysis of biopsy or prostatectomy tissue as LDTs (Laboratory-Developed Tests) under CLIA (Clinical Laboratory Improvement Amendments) certification. While the three commercially available panels described here (and summarized in Table 2) have all been thoroughly assessed for their prognostic ability in stratifying patients at risk of biochemical recurrence or metastasis after primary treatment, ideally further testing would test their predictive value in decisions about treatment to more clearly link these risk scores with particular therapeutic outcomes, which would allow these tests to be used in a predictive capacity during treatment decisions.

4.1. Cell Cycle Progression Score

The prognostic potential of proliferation markers such as cyclins and Ki67 in prostate cancer has been established in studies utilizing immunoreactivity to determine the expression of these markers in prostatectomy tissues and lymph node metastases (Bubendorf et al., 1996; Cheng et al., 1999; Khor et al., 2009; Mashal et al., 1996). More recently, Cuzick et al. (2011) used a candidate gene approach to select and test a panel of 126 cell cycle and cellular proliferation genes in RNA extracted from FFPE prostate tumor tissues. A refined panel of 31 highly correlated genes was developed to robustly measure proliferation within the tumor, with the information provided from the analysis of these 31 genes termed the cell cycle progression (CCP) score (Cuzick et al., 2011). A retrospective determination of CCP
scores from 366 radical prostatectomy cases and 337 transurethral resection of the prostate cases found that patients with higher CCP scores had a higher probability of biochemical recurrence or death over a 10-year follow-up period (Cuzick et al., 2011). This test, currently commercially available as Prolaris® from Myriad Genetics, also has demonstrated efficacy in prostate tissue obtained from biopsy (Bishoff et al., 2014; Cuzick et al., 2012; Freedland et al., 2013). From these studies, high CCP scores from biopsy were significantly associated with reduced time to biochemical recurrence, development of metastases, or death from prostate cancer (Bishoff et al., 2014; Cuzick et al., 2012). Additionally, the CCP score provided information independent from other clinical variables such as PSA or Gleason score, which meant that patients could be further stratified when the CCP score was combined with these clinical parameters to create an overall risk score for the identification of patients at risk of biochemical recurrence (Cooperberg et al., 2013; Cuzick et al., 2012; Freedland et al., 2013).

4.2. Genomic Prostate Score

In another study utilizing a candidate gene approach, Klein et al. (2014) analyzed the expression of 727 genes in two separate tumor specimens from 441 radical prostatectomy patients, allowing the predictive ability of these genes to be compared in tissues corresponding to either the primary or highest Gleason score. Of the 228 genes found to be associated with biochemical recurrence regardless of Gleason pattern, 81 genes were selected for further evaluation in a separate biopsy cohort to create a final gene signature, which includes genes representative of stromal response, cellular organization, proliferation, and androgen signaling (Klein et al., 2014). Gene expression levels of the 12 genes included in this signature are used to calculate the genomic prostate score (GPS), which has now been analytically validated and is commercially available as the Oncotype DX® Prostate Cancer Assay from Genomic Health (Knezevic et al., 2013). Although GPS distribution was significantly increased in patients scoring higher by other clinical risk determinants such as CAPRA (Cancer of the Prostate Risk Assessment) or NCCN (National Comprehensive Cancer Network) criteria, the GPS values within each clinical risk strata overlapped to a great extent (Cullen et al., 2014; Klein et al., 2014). This meant that the GPS provided additional benefit to traditional risk assessment as demonstrated by the continued significant association of the GPS with the increased likelihood of discovering adverse pathology on prostatectomy, or of biochemical recurrence following prostatectomy, even after adjusting for the inclusion of these clinical risk parameters (Cullen et al., 2014; Klein et al., 2014).

4.3. Genomic Classifier

With the objective of identifying a gene signature that would be predictive of rapid metastasis, Erho et al. (2013) conducted microarray analysis to determine gene expression in the prostate tumor tissues of patients that had developed metastases within five years of biochemical recurrence as compared to a control group of patients that either had no evidence of metastases after biochemical recurrence or had not experienced biochemical recurrence. The final gene signature consisting of 22 genes is combined to create a single Genomic Classifier (GC) score in which higher scores indicate a higher likelihood of metastasis (Erho et al., 2013). This GC score is currently commercially available as the Decipher® Prostate Cancer Classifier from GenomeDx. Higher GC scores were significantly
associated with a greater likelihood of biochemical recurrence and metastasis, even after adjusting for other associated clinical variables such as Gleason score, seminal vesicle invasion, and extraprostatic extension (Den et al., 2014; Erho et al., 2013; Karnes et al., 2013; Klein et al., 2015; Ross et al., 2014). The rapid appearance of metastatic disease is significantly associated with a reduced time until prostate cancer-specific death, which demonstrates the importance of biomarkers such as the GC score that could be used to identify patients at higher risk of prostate cancer-specific mortality that may require more aggressive treatment following prostatectomy (Cooperberg et al., 2015; Klein et al., 2015).

4.4. Additional Gene Signatures of Indolence and Aggression

Aside from the commercially available gene panels, a number of other groups have established gene signatures prognostic for biochemical recurrence or mortality, using methods such as candidate gene approaches to investigate particular pathways (Irshad et al., 2013; Peng et al., 2014), and whole transcriptome analysis either by microarray (Gasi Tandefelt et al., 2013; Yu et al., 2004) or RNA sequencing (Long et al., 2014). A collection of the gene signatures that have been described are summarized in Table 3, and include panels that have demonstrated prognostic ability for either biochemical recurrence or mortality. Although there have been many studies investigating the gene expression profile of prostate cancers, the criteria for inclusion here was that a defined gene signature had to have been identified and validated in cohorts collected at multiple institutions. These expression panels include gene signatures that were developed with a focus on low-risk cancers, subsequently improving the identification of biochemical recurrence or mortality amongst patients with Gleason 7 cancers (Irshad et al., 2013; Penney et al., 2011), as well as a gene signature that outperformed the CCP score in identifying men at risk of biochemical recurrence (Long et al., 2014).

5. Conclusions

Amongst the RNA biomarkers described here are biomarkers with the potential to be used in all stages of prostate cancer management, from detection (PCA3, TMPRSS2:ERG), directing treatment (AR-V7), and prognosis (Prolaris®, Oncotype DX®, Decipher®). Additionally, there are a number of RNAs in earlier stages of investigation with potential application as biomarkers of recurrence, metastasis, and indolence. The continued analysis of these and other biomarkers is important to fully define parameters that could be used to determine the most appropriate course of treatment for each patient, with the understanding that prostate cancer is a very heterogeneous disease. It is likely that a number of factors, including current clinical parameters such as age, family history, and pathology, will need to be combined with additional molecular biomarkers to provide the best characterization of each patient’s cancer and their risk of progression. It is also important to consider that, while patients can be wary of leaving their cancer “untreated,” the best course of action may not necessarily be curative intent treatment. However, if these biomarkers can be clearly linked to particular treatment outcomes, as is the case for the AR-V7 splice variant, patients and physicians will be able to decide upon an appropriate course of action with greater confidence.
Acknowledgements

This work was supported by NCI Cancer Center Support Grant P30CA138292 and NIH U01CA113913.

6. References


Mol Aspects Med. Author manuscript; available in PMC 2016 November 01.


Petrovics G, Zhang W, Makarem M, Street JP, Connelly R, Sun L, Sesterhenn IA, Srikanth V, Moul JW, Srivastava S. Elevated expression of PCGEM1, a prostate-specific gene with cell growth-


Ross RW, Galsky MD, Scher HI, Magidson J, Wassmann K, Lee GS, Katz L, Subbudh SK, Anand A, Fleisher M, Kantoff PW, Oh WK. A whole-blood RNA transcript-based prognostic model in men...


Mol Aspects Med. Author manuscript; available in PMC 2016 November 01.
# Table 1

Non-coding RNAs under investigation as prostate cancer biomarkers.

<table>
<thead>
<tr>
<th>ncRNA</th>
<th>Tissues</th>
<th>Findings</th>
<th>Cohorts</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>miR-141</td>
<td>Prostate</td>
<td>• Increased in PCa and Mets</td>
<td>BCCA, DFCI,</td>
<td>(Brase et al., 2011; Cheng et al., 2013; Mitchell et al., 2008; Nguyen et al., 2013; Selth et al., 2012)</td>
</tr>
<tr>
<td></td>
<td>Serum</td>
<td>• Associated with Gleason score</td>
<td>MSKCC, UMCHE, UM, UW</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Highly expressed in serum from mCRPC patients</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Did not stratify patients for risk of BCR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-375</td>
<td>Prostate</td>
<td>• Not expressed in surrounding stroma</td>
<td>BCCA, DFCI,</td>
<td>(Brase et al., 2011; Cheng et al., 2013; Nguyen et al., 2013; Selth et al., 2012; Wach et al., 2012)</td>
</tr>
<tr>
<td></td>
<td>Serum</td>
<td>• Increased in PCa and Mets</td>
<td>MSKCC, UMCHE, UM, UMu, UR, UW</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Highly expressed in serum from mCRPC patients</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MALAT-1</td>
<td>Prostate</td>
<td>• Increased in PCa and Mets</td>
<td>CH, CZH, WCH</td>
<td>(Ren et al., 2013a; Ren et al., 2013b; Wang et al., 2014)</td>
</tr>
<tr>
<td></td>
<td>Mets</td>
<td>• Associated with Gleason score and T stage</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Serum</td>
<td>• Higher in serum and urine of patients later diagnosed with PCa from biopsy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NEAT1</td>
<td>Prostate</td>
<td>• Increased in PCa and CRPC</td>
<td>MC, WCMC</td>
<td>(Chakravarty et al., 2014)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• High NEAT1 expression associated with greater risk of BCR, and Mets</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCGEM1</td>
<td>Prostate</td>
<td>• Increased in PCa</td>
<td>BCM, CH, MC, UM, USUHS, WCMC, WRAMC</td>
<td>(Petrovics et al., 2004; Prensner et al., 2014a; Srikantan et al., 2000)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• AA and patients with family history more likely to have high PCGEM1 in PCa tissue</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Did not stratify patients for risk of PCSM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SChLAP1</td>
<td>Prostate</td>
<td>• No expression in benign prostate</td>
<td>CC, EMC, MC, UM</td>
<td>(Mehra et al., 2014; Prensner et al., 2011; Prensner et al., 2013; Prensner et al., 2014b)</td>
</tr>
<tr>
<td></td>
<td>Mets</td>
<td>• Increased expression in PCa and Mets</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Urine</td>
<td>• High SChLAP1 associated with greater risk of BCR, Mets, and PCSM</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Detected in urine, higher in patients grouped as intermediate/high risk from biopsy</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

AA, African Americans; BCCA, British Columbia Cancer Agency; BCM, Baylor College of Medicine; BCR, Biochemical recurrence; CC, Cleveland Clinic; CH,anghai Hospital; CRPC, Castration-resistant prostate cancer; CZH, Changzheng Hospital; DFCI, Dana-Farber Cancer Institute; EMC, Erasmus Medical Center; MC, Mayo Clinic; mCRPC, Metastatic castration-resistant prostate cancer; Mets, Metastases; MSKCC, Memorial Sloan-Kettering Cancer Center; PCa, Localized prostate cancer; PCSM, Prostate cancer-specific mortality; UM, University of Michigan; UMu, University of Munster; UMCHE, University Medical Center Hamburg-Eppendorf; UR, University of Regensburg; USUHS, Uniformed Services University of the Health Sciences; UW, University of Washington; WCH, West China Hospital; WCMC, Weill Cornell Medical College; WRAMC, Walter Reed Army Medical Center

*Mol Aspects Med. Author manuscript; available in PMC 2016 November 01.*
### Table 2
Commercially available multi-gene mRNA expression panels for prostate cancer.

<table>
<thead>
<tr>
<th>Panel Name</th>
<th>Number of Genes</th>
<th>Signature Discovery Approach</th>
<th>Reference</th>
</tr>
</thead>
</table>
| **Prolaris®** – Myriad Genetics   | 31 CCP test genes 15 normalizers | • Candidate gene approach  
• 126 cell cycle and cellular proliferation genes tested in prostatectomy tissues | Cuzick et al. (2011) |
| **OncoType DX® Prostate Cancer Assay** – Genomic Health | 12 GPS test genes 5 normalizers | • Candidate gene approach  
• 727 genes tested in prostatectomy tissues  
• Refined panel of 81 genes tested in biopsies | Klein et al. (2014) |
| **Decipher® Prostate Cancer Classifier** – GenomicDx | 22 GC test genes  Whole microarray normalization | • Whole transcriptome analysis  
• Affymetrix GeneChip® Human Exon 1.0 ST Array used to assess prostatectomy tissues | Erho et al. (2013) |

CCP, Cell Cycle Progression; GC, Genomic Classifier; GPS, Genomic Prostate Score
### Table 3
Multi-gene mRNA expression panels with prognostic potential for prostate cancer.

#### PROGNOSTIC FOR BIOCHEMICAL RECURRENCE

<table>
<thead>
<tr>
<th>Reference</th>
<th>Tissue</th>
<th>Cohorts</th>
<th>Signature Discovery Approach</th>
<th>Number of Genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glinsky et al. (2004)</td>
<td>Prostate</td>
<td>Training: DFCI</td>
<td>- Over 12,000 genes measured in training set</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Validation: MSKCC</td>
<td>- Three different gene expression signatures associated with BCR identified</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- The combined panel (14 genes) and the three individual panels all had prognostic value</td>
<td></td>
</tr>
<tr>
<td>Yu et al. (2004)</td>
<td>Prostate</td>
<td>Training: UP</td>
<td>- Comprehensive gene expression analysis of training set performed by microarray</td>
<td>70</td>
</tr>
<tr>
<td>Varambally et al. (2005)</td>
<td>Prostate</td>
<td>Training: UM, UP</td>
<td>- Proteomic analysis followed up with gene expression studies to find common genes</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Validation: MSKCC</td>
<td>- 44 gene signature developed and then refined</td>
<td></td>
</tr>
<tr>
<td>Bismar et al. (2006)</td>
<td>Prostate</td>
<td>Training: UM</td>
<td>- Proteomic analysis followed up with gene expression studies to find common genes</td>
<td>12</td>
</tr>
<tr>
<td>Malhotra et al. (2011)</td>
<td>Prostate</td>
<td>Training: JH, KI, SU</td>
<td>- Analysis of microarray data in training set showed genes involved in the cell cycle</td>
<td>3</td>
</tr>
<tr>
<td>Gasi Tandefelt et al. (2013)</td>
<td>Prostate</td>
<td>Training: EMC</td>
<td>- Microarray analysis of training set</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Validation: MSKCC</td>
<td>- Identified prognostic signature associated with BCR in ERG positive patients</td>
<td></td>
</tr>
<tr>
<td>Irshad et al. (2013)</td>
<td>Prostate</td>
<td>Training: MSKCC, SWW, UP</td>
<td>- 377 genes associated with senescence and aging measured in training set</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Validation: HICCC, MSKCC</td>
<td>- 19 gene panel developed and then refined</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Best performance amongst the low Gleason score group of patients</td>
<td></td>
</tr>
<tr>
<td>Long et al. (2014)</td>
<td>Prostate</td>
<td>Training: AVAMC, MCC</td>
<td>- Training set interrogated by whole transcriptome RNA sequencing</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td></td>
<td>, UT Validation: MSKCC</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### PROGNOSTIC FOR MORTALITY

<table>
<thead>
<tr>
<th>Reference</th>
<th>Tissue</th>
<th>Cohorts</th>
<th>Signature Discovery Approach</th>
<th>Number of Genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Penney et al. (2011)</td>
<td>Prostate</td>
<td>Training: SWW</td>
<td>- 6,100 genes measured to determine expression profiles of Gleason 6 and 8 tumors</td>
<td>157</td>
</tr>
<tr>
<td>Olmos et al. (2012)</td>
<td>Blood</td>
<td>Training: RMH, BWSCC</td>
<td>- Gene expression in blood of CRPC patients measured by microarray</td>
<td>9</td>
</tr>
<tr>
<td>Reference</td>
<td>Tissue</td>
<td>Cohorts</td>
<td>Signature Discovery Approach</td>
<td>Number of Genes</td>
</tr>
<tr>
<td>--------------------</td>
<td>--------</td>
<td>---------------</td>
<td>-----------------------------------------------------------------------------------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Ross et al. (2012)</td>
<td>Blood</td>
<td>Training: DFCI Validation: MSKCC</td>
<td>• Gene expression in blood of CRPC patients measured by microarray</td>
<td>6</td>
</tr>
<tr>
<td>Peng et al. (2014)</td>
<td>Prostate</td>
<td>Training: JH, KI, SU Validation: KI</td>
<td>• 641 embryonic stem cell gene predictors tested in published arrays used as training set  • 34 gene signature developed then refined</td>
<td>3</td>
</tr>
</tbody>
</table>

AVAMC, Atlanta Veteran’s Administration Medical Center; BCR, Biochemical recurrence; BWSCC, Beatson West of Scotland Cancer Center; DFCI, Dana-Farber Cancer Institute; EMC, Erasmus Medical Center; HICCC, Herbert Irving Comprehensive Cancer Center; JH, Johns Hopkins University; KI, Karolinska Institute; MCC, Moffitt Cancer Center; MSKCC, Memorial Sloan-Kettering Cancer Center; PHS, Physician’s Health Study; RMH, Royal Marsden Hospital; SU, Stanford University; SWW, Swedish Watchful Waiting cohort; UM, University of Michigan; UP, University of Pittsburgh; UT, University of Toronto