Long non-coding RNAs and their potential impact on diagnosis, prognosis, and therapy in prostate cancer: Racial, ethnic, and geographical considerations


Published in:
Expert review of molecular diagnostics

Document Version:
Publisher's PDF, also known as Version of record

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

Publisher rights
Copyright 2021 the authors.
This is an open access article published under a Creative Commons Attribution-NonCommercial-NoDerivs License (https://creativecommons.org/licenses/by-nc-nd/4.0/), which permits distribution and reproduction for non-commercial purposes, provided the author and source are cited.

General rights
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.
Long non-coding RNAs and their potential impact on diagnosis, prognosis, and therapy in prostate cancer: racial, ethnic, and geographical considerations

Rebecca Morgan, Willian Abraham da Silveira, Ryan Christopher Kelly, Ian Overton, Emma H. Allott & Gary Hardiman

To cite this article: Rebecca Morgan, Willian Abraham da Silveira, Ryan Christopher Kelly, Ian Overton, Emma H. Allott & Gary Hardiman (2021) Long non-coding RNAs and their potential impact on diagnosis, prognosis, and therapy in prostate cancer: racial, ethnic, and geographical considerations, Expert Review of Molecular Diagnostics, 21:12, 1257-1271, DOI: 10.1080/14737159.2021.1996227

To link to this article: https://doi.org/10.1080/14737159.2021.1996227
Long non-coding RNAs and their potential impact on diagnosis, prognosis, and therapy in prostate cancer: racial, ethnic, and geographical considerations

Rebecca Morgan a,b, William Abraham da Silveira a,b,c, Ryan Christopher Kelly d, Ian Overton e, Emma H. Allott b,c,d and Gary Hardiman a,b,c

aFaculty of Medicine, Health and Life Sciences, School of Biological Sciences, Queen’s University Belfast, Belfast, UK; bInstitute for Global Food Security (IGFS), Queen’s University Belfast, Belfast, UK; cFaculty of Medicine, Health and Life Sciences, Patrick G. Johnston Centre for Cancer Research, Queen’s University Belfast, Belfast, UK; dDepartment of Histopathology and Morbid Anatomy, Trinity Translational Medicine Institute, Trinity College Dublin, Dublin, Ireland; eDepartment of Medicine, Medical University of South Carolina (MUSC), Charleston, South Carolina

ABSTRACT

Introduction: Advances in high-throughput sequencing have greatly advanced our understanding of long non-coding RNAs (lncRNAs) in a relatively short period of time. This has expanded our knowledge of cancer, particularly how lncRNAs drive many important cancer phenotypes via their regulation of gene expression.

Areas covered: Men of African descent are disproportionately affected by PC in terms of incidence, morbidity, and mortality. LncRNAs could serve as biomarkers to differentiate low-risk from high-risk diseases. Additionally, they may represent therapeutic targets for advanced and castrate-resistant cancer. We review current research surrounding lncRNAs and their association with PC. We discuss how lncRNAs can provide new insights and diagnostic biomarkers for African American men. Finally, we review advances in computational approaches that predict the regulatory effects of lncRNAs in cancer.

Expert opinion: PC diagnostic biomarkers that offer high specificity and sensitivity are urgently needed. PC specific lncRNAs are compelling as diagnostic biomarkers owing to their high tissue and tumor specificity and presence in bodily fluids. Recent studies indicate that PCA3 clinical utility might be restricted to men of African descent. Further work is required to develop lncRNA biomarkers tailored for men of African descent.

1. Introduction

Prostate cancer (PC) is one of the most prevalent cancers worldwide, particularly affecting men living in a western lifestyle and of African descent, suggesting risk factors that are genetic, environmental, and socioeconomic in nature. Despite major advances in the diagnosis and treatment of PC, the disease remains a global health problem. PC is the second most diagnosed non-skin cancer in men worldwide [1,2]. Currently, there is considerable interest in lncRNAs and their regulatory roles in developmental processes and disease/cancer stages [3,4]. Prior to high-throughput sequencing technologies becoming common place, lncRNAs were simply considered transcriptional ‘noise’ and largely ignored by the scientific community [5–7]. With increased sequencing sensitivity in combination with innovative bioinformatic tools, in the past decade, researchers have started to uncover the functional roles of lncRNAs. Two emerging discoveries include the realization that they act upon a wide range of cell processes and are differentially expressed in disease states, including various cancers [8,9].

Computational predictions have identified over 10,000 lncRNAs, and researchers have been keen to characterize them functionally, identify their regulatory roles, and exploit this knowledge in biomedical research, diagnostics, and therapeutic development [10]. In parallel, bioinformatic methods are also being rapidly improved to better predict the regulatory roles of lncRNAs in cancer and determine their full extent as regulatory modulators. This review provides a detailed overview of lncRNAs, particularly in the context of PC. First, we cover lncRNAs, their history, newly discovered functions, and correlation with cancer. We then review PC, racial differences in disease presentation, and why lncRNAs are relevant in this context. Lastly, we provide a summary of tools dedicated to deciphering the functional roles of lncRNAs.

2. What is a long non-Coding RNA?

lncRNAs are defined as transcripts longer than 200 nucleotides that are putatively not translated into proteins [8]. Both short (<200 base pairs) and lncRNAs may contain coding regions [11]. Short non-coding transcripts, in particular microRNAs (miRNAs), have been studied more extensively than lncRNAs and contribute to the development and progression of many cancers [12]. Prior to the genomic era,
IncRNAs were considered ‘junk RNA’ and were largely ignored by the scientific community. The first IncRNA to be identified was h19 in 1984. Although at first mistakenly categorized as a messenger RNA (mRNA) [13], h19 has subsequently been recognized as an epigenetic regulator, differentially expressed in several cancers where it acts as a tumor suppressor [3,14–16]. In 2002, IncRNAs were finally assigned as functional RNA molecules by the Functional Annotation of the Mammalian Genome (FANTOM) project [17]. Like mRNAs, IncRNAs are transcribed by RNA polymerase II, polyadenylated, capped, and spliced [18,19]. Subcellular localization provides a lens to define IncRNA functional classes [7,20]. The identification and number of IncRNAs is continually increasing; however, many remain poorly annotated and are not experimentally validated [12,21].

As of January 2021, LNCBook, a curated IncRNA database contains 268,848 documented IncRNAs (obtained from publicly accessible databases including GENCODE v27, NONCODE v5.0, LNCipedia v4.1 and MiTranscriptome beta). However, there is a large gap between documented IncRNAs and IncRNAs, which have a known link to a molecular or biological function. Of the 268,848 recorded IncRNAs on LNCBook, only 1,867 have been categorized as experimentally validated (linked with functional mechanisms/biological processes and identified in relevant publications) [21,22]. LncRNAs are classified as either 1) intergenic (lincRNAs), transcribed from intergenic regions located between two genes [5,23,24] 2) bidirectional (divergent IncRNAs) transcribed on the reverse strand of a coding transcript located no more than 1000 base pairs away [5,24] 3) intronic IncRNAs, transcribed only from intronic regions of protein coding genes, 4) antisense IncRNAs transcribed from antisense strands of protein coding genes and 5) sense IncRNAs transcribed from sense strands of protein coding genes [6,25,26].

Understanding how IncRNAs are regulated and how they in turn regulate their target genes is an important research area [27]. Thus far, several functional roles of IncRNAs have been established. The most common IncRNA functional groups include guides (e.g. HOTAIR) [6,28], scaffolds (e.g. TERC) [6,29], signals (e.g. XIST) [6,7] and decoys (e.g. MALAT1) [6,30]. There are also a series of documented multifunctional IncRNAs. An example is NEAT1 [31]. NEAT1 was identified as differentially expressed in numerous tumor contexts, including lung, colorectal, and prostate cancer [31]. NEAT1 is an important mediator in the regulation of clear cell renal cell carcinoma (ccRCC) progression and predicts poor prognosis and decreased survival in patients with ccRCC [7,32–34]. Guide IncRNAs can bind and tether epigenetic protein complexes to a target locus directing gene expression either in cis or in trans. Guide IncRNAs can direct chromatin-modifying enzymes (e.g. histone methylases, acetylases, and deacetylases) to a particular gene target locus, which can lead to the activation or repression of local genes [35,36]. Scaffold IncRNAs recruit epigenetic complexes together, thus acting as a molecular scaffold. They recruit RNA-binding factors to form RNA-protein complexes, which promote or inhibit gene transcription [35,36]. Signal IncRNAs possess signaling properties that promote guidance of epigenetic complexes to specific locations. These IncRNAs possess enhancer functions, which alter the structure of chromatin and guide transcriptional complexes to gene targets enabling transcriptional activity [35,36]. Finally, decoy IncRNAs can lure epigenetic complexes to different genomic locations. These IncRNAs can recruit and decoy a transcriptional complex to a distinct location for suppression of gene expression [35,36]. Many reviews have documented the molecular functions of IncRNAs (for example, reference 23), however, none have considered the computational aspects of predicting IncRNA functions and their targets. Our review addresses this shortcoming.

3. Long Non-coding RNAs and Prostate Cancer

The number of patients presenting with urological cancers (malignancies of the genital and urinary organs of men and the urinary organs of women) is on the increase [37,38]. Currently, urological cancers of the prostate, bladder, and kidney stand among the top 10 most diagnosed cancers in men [39]. Despite their frequency, biomarkers for urological cancers are limited. Those that are available typically lack sensitivity in low-grade disease, e.g. urine cytology for bladder cancer and low specificity, e.g. Prostate Specific Antigen (PSA) [39–41]. IncRNAs are now frequently being linked with urological cancers, and many have been designated as playing roles in the development and progression of these malignancies including prostate cancer [42]. Table 1 summarizes selected IncRNAs that are differentially expressed in urological cancers, which may have potential as biomarkers or therapeutic targets [6,7,31].

Despite major advances in the diagnosis and treatment of PC, the disease remains a global health problem. PC is the second most diagnosed non-skin cancer in men worldwide [60]. Recent estimates indicate that 1 in 7 men will be diagnosed with PC during their lifetime. In 2018 alone, 359,000 recorded deaths from PC occurred globally [61,62]. Racial differences are common in PC. Men of African descent are adversely affected in terms of both incidence and mortality rates compared to European-based counterparts. IncRNAs are...
Table 1. Urological cancer-associated lncRNAs.

<table>
<thead>
<tr>
<th>lncRNA</th>
<th>IncRNA functional group *</th>
<th>Urological cancer type</th>
<th>Sample size</th>
<th>Differential expression</th>
<th>Fold change</th>
<th>Potential cancer impact</th>
<th>Hazard ratio</th>
<th>Study reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>MALAT1</td>
<td>Scaffold [43] Decoy [44]</td>
<td>Bladder</td>
<td>95</td>
<td>Upregulated</td>
<td>Not described</td>
<td>Poor survival</td>
<td>1.26 (0.68–2.13)</td>
<td>[45]</td>
</tr>
<tr>
<td>LINC00346</td>
<td>Not described</td>
<td>Bladder</td>
<td>52</td>
<td>Upregulated</td>
<td>Not described</td>
<td>promotes bladder</td>
<td>Not described</td>
<td>[47]</td>
</tr>
<tr>
<td>KCNQ1OT1</td>
<td>Guide [48] signal [49]</td>
<td>Bladder</td>
<td>30</td>
<td>Upregulated</td>
<td>Not described</td>
<td>KCNQ1OT1 facilitates</td>
<td>Not described</td>
<td>[50]</td>
</tr>
<tr>
<td>LOC400891</td>
<td>Not described</td>
<td>Prostate (All patients underwent radical prostatectomy or laparoscopic radical prostatectomy)</td>
<td>81</td>
<td>Upregulated</td>
<td>Gene chip – 4.1 QPCR – 2.4</td>
<td>Unfavorable prognosis</td>
<td>Multi-variate Tumor stage – 1.823</td>
<td>LOC400891 expression – 2.116</td>
</tr>
<tr>
<td>PCA3</td>
<td>Scaffold [52]</td>
<td>Prostate</td>
<td>56</td>
<td>Upregulated</td>
<td>10–100-fold overexpression</td>
<td>Not described</td>
<td>Not described</td>
<td>[95]</td>
</tr>
<tr>
<td>PCAT1</td>
<td>Decoy [53]</td>
<td>High-risk prostate cancer (Gleason ≥ 7) and metastasis prostate cancer</td>
<td>58 prostate cancer tissues</td>
<td>Upregulated</td>
<td>Not described</td>
<td>Not described</td>
<td>Not described</td>
<td>[104]</td>
</tr>
<tr>
<td>PCAT18</td>
<td>Not described</td>
<td>Prostate</td>
<td>160</td>
<td>Upregulated</td>
<td>8.8–11.1 fold</td>
<td>potential therapeutic target for metastatic prostate cancer</td>
<td>Not described</td>
<td>[110]</td>
</tr>
<tr>
<td>PCAT7</td>
<td>Decoy [115]</td>
<td>Prostate cancer with bone metastasis</td>
<td>57</td>
<td>Upregulated</td>
<td>Not described</td>
<td>promotes PCa bone metastatic</td>
<td>1.85</td>
<td>[115]</td>
</tr>
<tr>
<td>SChLAP1</td>
<td>Decoy [54]</td>
<td>Prostate</td>
<td>208</td>
<td>Upregulated</td>
<td>Not described</td>
<td>SChLAP1 expression is associated with prostate cancer progression</td>
<td>3.97</td>
<td>[118]</td>
</tr>
<tr>
<td>PCGEM1</td>
<td>Scaffold [55]</td>
<td>Prostate (cells of African American patients)</td>
<td>Not described</td>
<td>Upregulated</td>
<td>Not described</td>
<td>increased expression level with high-risk CaP patients and African American race is significantly related to PCGEM1 expression in tumor cells</td>
<td>Not described</td>
<td>[122]</td>
</tr>
<tr>
<td>PCGEM1</td>
<td>Prostate (African American patients)</td>
<td>30 patients (15 African American, 15 (Caucasian-American)</td>
<td>Not described</td>
<td>Upregulated</td>
<td>Not described</td>
<td>Not described</td>
<td>Not described</td>
<td>[124]</td>
</tr>
<tr>
<td>HORASS</td>
<td>Not described</td>
<td>Prostate</td>
<td>Not described</td>
<td>Upregulated</td>
<td>12.51</td>
<td>Poorer clinical outcome in human PCa samples</td>
<td>[169]</td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
now considered to contribute to these racial differences [63,64].

Evidence to date suggests that lncRNAs can alter gene expression at both transcriptional and post-transcriptional levels, thus playing roles in important biological processes including carcinogenesis and metastasis [65]. Some lncRNAs function as tumor suppressors, and possess the ability to interact with oncogenes, inhibit or accelerate proliferation, and influence apoptosis [30,66,67].

Other lncRNAs are commonly aberrantly expressed, deleted, amplified, and/or mutated in multiple cancers [9,68]. lncRNAs have been shown to have roles in numerous forms of cancer [69]. Figure 1 displays lncRNAs correlated with numerous cancer types. Specific cancer types and relevant body tissues are highlighted. Table 2 provides a summary of lncRNAs, which have been identified as differentially expressed in various malignancies.

Presently, three areas that merit particular attention are: 1) Identification of novel diagnostic and prognostic biomarkers for PC, with an emphasis on diagnostic biomarkers for early-stage PC, which differentiates low-risk from high-risk tumors [40,84], 2) An urgent need to better understand biological differences in disease presentation in men of African and Northern European descent [85] and 3) Efficient therapies for advanced and castrate-resistant PC [86]. lncRNAs are emerging as biomarkers or therapeutic targets for these areas, as explained below.

### 4. lncRNAs as prostate cancer biomarkers

It has been widely accepted that the identification of biomarkers should not exclusively focus upon protein coding genes [87]. A key factor for biomarker utility is whether the proposed biomarker is detectable in bodily fluids. Liquid-based biomarkers provide many benefits over traditional tissue-based biomarkers from a ‘no surgery requirement and recovery time’ perspective, painless non-invasive testing, decreased costs, and more rapid diagnostics [88]. lncRNAs are attractive as PC biomarkers owing to their high tissue and tumor specificity [27] and are detectable in bodily fluids, including PCA3 and PCAT18 [88,89].

Biomarkers with high specificity and sensitivity are particularly needed to help address the over-treatment of clinically insignificant PC. The over diagnosis and treatment of low-risk PC remains a common problem, primarily due to the use of Prostate Cancer Specific Antigen (PSA) tests [90] which offer high sensitivity but low specificity [40]. The introduction of the PSA test in the early 1990s was highly criticized for the overdiagnosis and the peak of PC incidence rates [91–93]. Furthermore, PSA tests have poorer diagnostic performance in men of African descent [94]. Therefore, a current need in PC research is the identification of diagnostic biomarkers, which can not only offer high specificity and sensitivity but are also universally reliable for all races.

PC-specific lncRNAs are compelling as diagnostic biomarkers for clinical use [27]. Possibly, the most well-known is PCA3 (Prostate Cancer Specific Antigen 3). PCA3 was first observed to be highly expressed in PC tumors in 1999 and has since been characterized as overexpressed in 95% of all PC cells [95,96]. Urine-based measurements of PCA3 provide greater PC specificity than PSA testing [12]. In 2012, the PROGENSA PCA3 assay developed by Gen-Probe Inc was the first lncRNA PC biomarker to obtain Food and Drug Administration (FDA) approval. This lncRNA molecular test is used to help guide urologists in determining the need for repeat biopsies in men aged 50 years or older who had previously presented with a negative biopsy result [97,98]. However, its practicability as a first use test remains controversial and it cannot predict PC aggressiveness [99,100]. Several studies have also documented PCA3 expression in PC progression; however, further research is required to confirm its feasibility as a PC prognostic biomarker [101]. Additionally, PCA3’s clinical utility might be

---

**Table 1. (Continued).**

<table>
<thead>
<tr>
<th>lncRNA</th>
<th>lncRNA functional group *</th>
<th>Urological cancer type</th>
<th>Sample size</th>
<th>Differential expression</th>
<th>Fold change</th>
<th>Potential cancer impact</th>
<th>Hazard ratio</th>
<th>Study reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CTBP1-AS</strong></td>
<td>Decoy [54]</td>
<td>Prostate</td>
<td>105 prostate cancer samples</td>
<td>Upregulated</td>
<td>Not described</td>
<td>Associated with prostate cancer castration-resistant tumor growth and activates AR signals</td>
<td>Not described</td>
<td>[56,170]</td>
</tr>
<tr>
<td><strong>ARLNC1</strong></td>
<td>Not described</td>
<td>Prostate</td>
<td>14 prostate cancer cell lines</td>
<td>Upregulated</td>
<td>Not described</td>
<td>Strongly associated with AR signaling in prostate cancer progression</td>
<td>Not described</td>
<td>[178]</td>
</tr>
<tr>
<td><strong>PCAT29</strong></td>
<td>Decoy [57]</td>
<td>Prostate</td>
<td>Not described</td>
<td>Downregulated</td>
<td>0.29 ± 0.15 in tumor tissues as compared to normal tissues</td>
<td>Suppresses prostate cancer cell proliferation</td>
<td>Not described</td>
<td>[56]</td>
</tr>
<tr>
<td><strong>GASS</strong></td>
<td>Scaffold [58]</td>
<td>Prostate</td>
<td>Not described</td>
<td>Downregulated</td>
<td>Not described</td>
<td>Proapoptosis</td>
<td>Not described</td>
<td>[59]</td>
</tr>
</tbody>
</table>

* lncRNA functional groups subject to change
restricted to men of European descent. O’Malley, et al. examined the clinical significance of the urinary biomarkers (PCA3 & T2ERG) in 718 African American and non-African American men, and assessed whether they provided increased utility for the detection of PC. Their results indicated that PCA3 and T2ERG enhanced clinical utility; however, this result was only observed in non-African American men [102]. A major limitation of this study was the relatively small number of African American samples [103] in comparison to non-African American men. Further research with larger ethnically diverse datasets would facilitate a more rigorous assessment.

Other studies have yielded varied outcomes in contrast to the initial findings from the O’Malley study. Feibus, et al. conducted a study using 304 patients, 60% of which were African American and noted that measurements of PCA3 enhanced predictability of high-grade PC in African American men [94], while an earlier study by Adam et al. using South African men challenged these results. Of 105 South African men referred for prostate biopsy, Adam et al., found that PCA3 levels were not significantly increased in black men who had tested positive for the disease [103]. Again, larger cohorts would provide greater statistical power.

Several other IncRNAs have been established as possible biomarkers of PC, including PCAT1 (Prostate Cancer Associated Transcript 1). PCAT1 is a highly PC-specific IncRNA [27] and was upregulated in patients presenting with a high-risk (Gleason score ≥7) PC and metastatic disease [27,104]. Prensner et al., 2011 identified 121 unannotated PC ncRNAs through the
analysis of 102 prostate tumor tissues and cell lines, among these was PCAT1 [105]. PCAT1 was significantly overexpressed in most PC tumors but particularly in tumors that had metastasized. The group remarked that PCAT1 was likely contributing to cell proliferation and potentially could serve as a urinary biomarker given the prior success with PCA3 [106]. Interestingly, PCAT1 is located at chromosome 8q24, a chromosomal location frequently correlated with PC risk due to its common amplification and the presence of PC susceptibility single nucleotide polymorphisms (SNPs) [107,108]. Yhan et al. and Prensner et al. concluded that PCAT1 supports cell proliferation in PC with Prensner and colleagues noting that it may represent a suitable candidate for biomarker development [104,109]. PCAT18 is overexpressed in PC. Using tissue samples from patients diagnosed at the Stephanshorn clinic in Switzerland and plasma samples collected in Canada, Crea et al. found that PCAT18 was overexpressed in the plasma of patients with localized and advanced forms of the disease compared to healthy individuals. The results from this study displayed a positive correlation between PCAT18 and PC stage, with the highest levels observed in metastatic castrate resistant PC. Thus, PCAT18 could have prognostic value, including to discriminate between localized and metastatic forms of PC, and is a possible therapeutic target for advanced PC [110]. MALAT1 (Metastasis Associated Lung Adenocarcinoma Transcript 1) is a lncRNA originally correlated with lung cancer; however, it has been found to be overexpressed in several cancers including PC. In 2013, Ren et al. identified MALAT1 as detectable in plasma and serum and suggested that it had the potential to be a diagnostic biomarker [111]. Wang et al. identified MALAT1 as overexpressed in urine samples of Chinese PC patients and noted its potential as a urine-based biomarker [112].

lncRNAs correlated with advanced PC include PCAT7 (Prostate Cancer Associated Transcript 7), SCLAP1 (SWI/SNF Complex Antagonist Associated with Prostate Cancer 1), and PCAT18. PCAT7 has been previously correlated with disease progression in nasopharyngeal and non-small cell lung cancer [113,114]. Lang, et al., 2020 noted that PCAT7 has the potential to be a therapeutic target for men presenting with PC bone metastasis via disruption of the constitutive loop between PCA7 and TGF-β signaling [115]. PCAT7 also has the potential to be a prognostic biomarker. SCLAP1 (SWI/SNF ‘Complex Antagonist Associated With Prostate Cancer’) has been proposed as a biomarker of lethal PC. Cimadamore et al. documented that lncRNA SCLAP1 is critical in the development and progression of PC [116]. Prensner et al. identified that SCLAP1 was highly expressed in ~25% of PCs analyzed (including metastatic PC) and may promote tumor cell invasion (in vitro) as well as metastasis (in vivo) [117–119]. Additionally, Prensner et al. observed SCLAP1 expression in urine and correlated with higher-risk PC patients suggesting its potential as a biomarker to differentiate between low-risk and high-risk tumors [120]. SCLAP1 has also been associated with adverse clinical outcomes in postradical prostatectomy and considered a potential tissue-based prognostic biomarker by Mehra, et al. [118]. Similarly, Chua et al. suggested SCLAP1 expression was linked with increased biochemical relapse and metastasis, cribriform architecture and intraductal carcinoma [121].

PCGEM1 is highly PC specific, and has been linked to PC progression by androgen receptor transcriptional co-regulators and could also decrease apoptosis in PC cell lines [117,122,123]. Petrovics et al. observed that PCGEM1 expression was significantly higher in PC cells of men of African descent compared to men of Northern European descent [122]. An additional study by Lee et al. observed similar results where PCGEM1 was significantly overexpressed in prostate tumors of African American, but not European American men [124].

Despite a better understanding of the roles of lncRNAs and the number identified as potential biomarkers of localized and advanced PC, few lncRNAs have met the success of PCA3. A primary reason for this is that lncRNA research is still in its infancy with thousands of predicted lncRNAs requiring laboratory experimental validation of their utility and confirmation in separate patient cohorts. Moreover, the overwhelming majority of candidate biomarkers fail to achieve clinical utility [89]. The candidate PC biomarkers summarized above require analytical and clinical validation to better assess their utility. To date, numerous articles have been published detailing lncRNAs directly related to PC (refer to references 38, 51, 53, 54, 55, and 68); however, our review differs in two specific ways: 1) more in-depth detail relating to PC and racial differences, and 2) computational methods and software now available for identifying potential lncRNA biomarkers in the disease.

5. Racial differences in Prostate Cancer

Racial differences are common among cancers, and examples include multiple myeloma, colorectal, and breast cancer. Multiple myeloma is the most common hematologic cancer in patients of African descent [125]. Although it is a rare disease, diagnoses are twice as high in patients of African descent compared to those of European descent. Reasons for these differences are currently unknown and the disease remains incurable [126–128]. In the USA, colorectal cancer patients of African descent experience decreased stage-specific survival in comparison to patients of European descent irrespective of socioeconomic background, age, treatment, and stage at diagnosis [129]. Invasive breast cancer mortality rates also differ between patients of African and Northern European descent. In the USA, African American women are more likely to die of the disease than European American women [130]. African American women also experience worse prognosis and decreased survival times compared to European American women regardless of their socioeconomic background, age, and stage at diagnosis [131].

Prostate cancer has one of the largest racial disparities of all cancer types. In the U.S.A., African Americans have among the highest mortality rates and lowest survival rates of any race or ethnic group [132]. PC is also the top cause of cancer mortality among men of African descent living in Caribbean and sub-Saharan Africa [133]. Similar disparities exist in the UK, where men of African descent are 2–3 times more likely to develop PC than Caucasian men and have a 30% higher mortality rate [134]. Nonbiological factors contributing to PC racial disparities include diminished societal trust in the medical
community, level of education attained, and financial concerns. Access to health care plays an important role in PC mortality. However, African American men have higher PSA values and higher tumor grades even at the same cancer stage, compared with non-African American men within the same health-care system, such as the Veterans Administration in the U.S.A. [135]. The reasons underlying these racial differences remain an area of active research. Many biological explanations have been presented, including dietary factors such as alcohol intake, vitamin D deficiency, and lycopene and isoflavone consumption [135–139]. Biological factors include genetic alterations – African Americans have higher rates of variations in tumor suppressor genes, such as EPHB2, and are susceptible to higher prevalence of chromosome 8q24 variants [64,85,107,140–142].

An emerging area of research is the role of IncRNAs in cancer development and progression in the context of racial differences. Frequently, IncRNAs are reported to be differentially expressed in numerous cancers and are known to be involved in tumorigenesis and metastasis [65]. However, there is a lack of research investigating IncRNA differences between racial groups.

Much remains to be uncovered about the mechanisms by which IncRNAs regulate transcriptional and translational processes that drive oncogenesis. The IncRNA previously found to be significantly correlated with both PC and men of African ancestry is PCGEM1 (Prostate Cancer Gene Expression Marker 1). PCGEM1 is highly PC specific and has been linked to PC progression by androgen receptor transcriptional co-regulators and attenuates apoptotic processes in PC cell lines [117,122,123]. Petrovics et al. identified PCGEM1 expression as being significantly higher in PC cells of men of African descent compared to men of Northern European descent [122]. An additional study by Lee et al., 2016, observed similar results where PCGEM1 was significantly overexpressed in prostate tumors of African American but not in European American men [124].

Recently, Yuan, et al., 2020, examined genomic and transcriptomic differences among African American and European American PC patients. The 470 patients that were analyzed consisted of RNA-seq transcriptomes from 57 African American and 413 European American patients from The Cancer Genome Atlas project (TCGA) prostate adenocarcinoma (PRAD) portal. The significantly lower number of African American samples represents a common challenge in health disparities research, with genomic studies typically being constrained to populations of European descent [143].

Yuan, et al., identified 1868 (31%) of the differentially expressed transcripts between African American and European American patients as IncRNAs. Patient and disease characteristics (age and Gleason score) were accounted for [64]. Of these differentially expressed IncRNAs, all were significantly enriched and regulated at least one target gene identified by the Long Noncoding RNA Heterogeneous Regulatory Network integrator (LongHorn) algorithm. LongHorn predictions are based on reverse-engineered canonical interactions, determined experimentally as part of the Encyclopedia of DNA Elements (ENCODE) project including the crosslinking and immunoprecipitation assay (eCLIP) and chromatin immunoprecipitation sequencing (ChIP-seq) data [30,64]. The researchers also showed that nearly 90% of IncRNAs differentially expressed between racial groups remain uncharacterized, underlining the need for continued IncRNA research. Of the 28 characterized IncRNAs differentially expressed between African and European American men, 13 were PC specific including PCA1 and PCAT10. The researchers noted that the differentially expressed IncRNAs potentially contribute to disease progression and influence PC racial differences. Further work is required to understand the regulatory functions of uncharacterized IncRNAs differentially expressed between racial groups [64].

Given the evidence highlighting the potential roles of IncRNAs in PC tumorigenesis, there is an urgent need to understand the functional consequences of IncRNA differences between African and Northern European PC patients. One of the pressing challenges for integrative computational biology and statistical genetics in racial disparities research is predicting genotype-to-phenotypes. The ability to identify the key drivers among a multitude of interacting molecules is challenging. Additional obstacles include the rapid growth of data, the and unavailability of data through issues with incompleteness, inaccuracies, heterogeneity, and data silos [144,145]. LncRNAs may be prioritized according to their impact upon gene regulatory networks through graph theoretic and systems biology approaches [146,147]. Understanding the biological underpinnings of racial differences in PC will ultimately provide an opportunity to help eliminate them by identifying possible PC diagnostic and prognostic biomarkers as well as potential therapeutic targets for men of African descent.

6. LncRNAs as potential therapeutic targets for advanced and castrate resistant Prostate Cancer

Advanced PC is traditionally defined as a disease, which has spread (metastasized) outside the prostate gland [148]. Cancers generally metastasize preferentially to specific distant organ sites [149]. In most cases of advanced PC, metastasis targets the bones [150]. Androgen Deprivation Therapy (ADT) is clinically used as a palliative treatment for patients diagnosed with advanced PC [151–153]. In healthy men, the androgen testosterone and its derivative dihydrotestosterone are required for normal functioning of the prostate, however, researchers have identified it is also vital for PC tumor growth [154]. Despite initial success with ADT, nearly all patients relapse and develop castrate-resistant PC (CRPC) after a mean duration of 2–3 years [155,156]. CRPC is defined as the progression of PC despite ADT and may present as one or any combination of a steady growth in PSA levels, progression of preexisting disease, or the emergence of new metastases [157]. Currently CRPC remains incurable and most patients with non-metastatic tumors will develop metastasis within 36 months of diagnosis, 90% of which will experience metastasis to the bone [158,159]. The survival probability of metastatic PC patients depends on a variety of circumstances from which organ sites the disease has metastasized to, the patient response to treatment, and the extent of the metastasis [160]. However, once the disease has spread, survival rates
are bleak with 29% five-year survival [161]. Drug resistance in cancer therapies continuously present as a major obstacle [162]. lncRNAs are recognized as possible drivers of drug resistance and may be modulated to enhance drug sensitivity in cancer therapy [162]. Drug resistance refers to tolerance against pharmaceutical drugs [163]; drug sensitivity may be restored after tolerance has been developed [163][162].

In PC, there remains an urgent need to identify therapeutic targets for CRPC and advanced forms of the disease. lncRNAs are being proposed as alternative therapeutic targets for slowing the progression of PC, several of which include HOTAIR (Hox Transcript Antisense Intergenic RNA), MALAT1 (Metastasis Associated Lung Adenocarcinoma Transcript 1), and GASS (Growth Arrest Specific Transcript 5) [87]. The majority of lncRNAs identified as potential CRPC and metastatic PC therapeutic targets are specifically associated with the androgen receptor (AR) signaling pathway [164,165]. Notably, the AR is vital in all stages of PC because it stimulates cell growth and survival [166]. Thus, most newly suggested therapeutic targets for PC center around this [167,168].

Several lncRNAs proposed as candidate therapeutic targets for CRPC and metastatic PC are primarily associated with the AR, including HORASS (LINC00161) [169], CTBP1-AS (C-Terminal Binding Protein 1 Antisense RNA) [170], and PCAT29 (Prostate Cancer Associated Transcript 29) [166]. lncRNAs can be directly influenced by androgen regulation [171], and many have been documented to participate in the advancement of PC through direct contact with the AR [172].

Parolli et al., 2019 found that HORASS facilitated the progression of CRPC in an AR-dependent manner by regulating AR mRNA stability [169] and noted that reduction in HORASS expression had a knock-on effect on the expression of the AR and its associated cancerous AR targets such as KIAA0101 [169]. lncRNA CTBP1-AS has been observed to have similar results for both CRPC and hormone dependent PC principally by promoting tumor growth [170]. Yakayama, et al. found that CTBP1-AS can suppress expression levels of CTBP1, a transcriptional corepressor known to regulate expression of tumor suppressors, and several genes involved with cell apoptosis [170,173]. CTBP1-AS had androgen-dependent functions, which could hinder the expression of tumor suppressors in both hormone dependent PC and CRPC suggesting potential as a therapeutic target [170]. PCAT29 is a PC-specific lncRNA, which has also been linked to the AR [166] Malik, et al. identified PCAT29, as a tumor suppressor of PC, which is suppressed

Figure 2. Androgen receptor (AR) signaling pathway associated PC lncRNAs (Adapted from Aird et al. [149]). This schematic provides a summary on key lncRNAs and their role in up- and down-regulation of the AR. The role of these key lncRNAs in promoting and inhibiting carcinogenesis and the interplay with miRNAs and proteins is highlighted.
by the AR. PCAT29 has PC suppressive effects including hindering cellular proliferation, tumor growth, cellular migration, and metastasis [166]. Lower expression of PCAT29 is correlated with decreased survival in PC patients [166].

Several other studies have identified IncRNAs linked to PC and the AR. An associated IncRNA is HOTAIR [174]. HOTAIR is categorized as ubiquitously expressed and associated with multiple illnesses, such as cardiovascular diseases, and cancers including breast, gastric, and colorectal [175]. Zhang, et al., 2015 identified that HOTAIR is repressed by androgens and overexpressed in CRPC [176]. Further research supports the hypothesis that HOTAIR contributes to the progression from hormone dependent to CRPC [177]. Taken together, these results suggest HOTAIR is a key player in numerous cancers including CRPC [176]. Another IncRNA recently associated with the AR is ARLN1 (AR-regulated long non-coding RNA 1). Zhang, et al., 2018 identified ARLN1 was strongly correlated with AR signaling in the advancement of PC and deemed it a possible therapeutic target with further detailed study [178].

Several additional IncRNAs have been correlated with the androgen receptor (AR) signaling pathway in either to inhibit or promote tumor growth. Therefore, IncRNAs can act as key drivers of PC and have potential as therapeutic targets. Figure 2 presents a schematic overview of AR associated IncRNAs that either inhibit or promote PC pathophysiology [164]. This schematic summarizes key IncRNAs, their role in up- and down-regulation of the AR, the role of these key IncRNAs in promoting and inhibiting carcinogenesis, and interplay with miRNAs and proteins.

7. LncRNA Interaction catalogs and prediction tools

Exploratory studies suggest that IncRNAs have an extensive interactome; however, current validation data are insufficient; therefore, many IncRNAs lack known interaction partners and their function remains elusive. IncRNA interactions may be acquired from public databases, some of which are considered below. However, existing annotations do not capture all the possible mechanisms and contexts in which IncRNAs may control gene expression. Experimental determination of IncRNA interactions is labor-intensive and costly, not least due to the large space of potential interactions, and so IncRNA function prediction software is required.

7.1. Databases related to IncRNAs

Presented by Zhou and colleagues, the experimentally validated functional IncRNA database (EVlncRNA) 2.0 is an updated version of the EVlncRNA database, which comprises only functional IncRNAs which have been experimentally validated through low-throughput techniques [179]. The update, which involved the manual curation of over 19,000 IncRNA-associated publications, includes information on 4,010 curated IncRNAs across 124 species, 1,082 diseases, and 6,244 interactions. Importantly, the interaction network integrated into the tool’s online platform (https://www.sdklab-biophysics-dzu.net/EVLncRNAs2/) provides users with the ability to easily browse information retained within the database.

Despite the quality of IncRNA annotated within EvlncRNA 2.0, it could be considered much less comprehensive than alternative databases, such as RNAInter (RNA Interactome Database) or RISE (RNA Interactome from sequencing experiments) which rely heavily on data from high-throughput studies and thus possess substantially more IncRNA interactions [180–182].

IncTarD database (http://bio-bigdata.hrmbu.edu.cn/ LncTarD/) is a manually curated resource retaining information corresponding to 2822 IncRNA-target regulatory interactions that have proven relevance in 177 human diseases. For prostate cancer specifically, the database possesses data on 73 IncRNA-target interactions along with the mechanism of interaction and its functional consequences [105]. LIVE (IncRNA Interaction Validated Encyclopedia) is a database of manually curated and experimentally validated IncRNA interactions [183]. Consequently, LIVE is small with just 572 IncRNA interactions at the time of writing. Importantly, the database contains information relating to IncRNA interactions with multiple cellular components, including proteins and transcription factors, in addition to disease associations.

IncCaNet (LncRNA–Cancer gene co-expression Network) was first introduced in 2016 and is currently available as a web-based tool (http://lncanet.bioinfo-minzhao.org/) [184]. It operates as a database containing IncRNA-gene co-expression predictions for 11 cancers (colon, lung, breast, etc.), with co-expression data sourced from Mitranscriptome, an online data repository of long polyadenylated RNA transcripts identified by RNA-seq [185]. IncCaNet includes IncRNA-gene expression correlations for 2,544 cancer genes and 9,641 IncRNAs.

7.2. Prediction tools

In the absence of experimental evidence, predictive algorithms are used; however, benchmarking has revealed room for improved predictive performance [186,187]. Most RNA–RNA prediction algorithms must take into consideration the sequence of the RNA molecule when determining RNA prediction interactions. For example, both IntaRNA and LncRIResearch methods utilize minimum free energy, a thermodynamic principle affected by the number, type, and arrangement of nucleotides of an RNA molecule, to predict interactions [188–190]. In addition, the RiScoPre tool is capable of extracting RNA–RNA interactions through literature text mining, which is highly valuable given that findings are likely to be experimentally validated [191].

IncRNA–DNA interactions mostly occur through RNA–DNA triplex formation, and estimations of such structures are useful substitutes in lieu of experimental evidence. For example, the recently developed Triplex Forming Potential Prediction (TriplexFP) is a machine learning-based model, which considers triplex formation capacity of an IncRNA based on computational methods in addition to triplex formation sites in DNA based on experimental data in a machine learning model to predict triplex formation [191]. In theory, such a method would be able to outperform more traditional methods such as Triplexator as it is able to consider beyond the limits of the
computational model through the inclusion of experimental data \[191,192]\.

Capsule-LPI tool considers sequence, motif, physicochemical and secondary structural features from IncRNAs and proteins within a IncRNA-protein pair to determine IncRNA–protein interactions \[193\]. Despite its complex mathematical architecture, the authors have compiled Capsule-LPI into a web-based tool (http://csbg-jlu.site/lpc/predict) allowing users to capture IncRNA–protein interactions by providing up to 100 IncRNA and protein sequences in Fasta format. Importantly, under 5-fold cross validation, Capture-LPI was found to possess the highest precision and the second highest sensitivity, highlighting the tool's utility and its room for improvement.

Interestingly, methods are available to translate interaction information into functional predictions. For example, MechRNA is a pipeline for IncRNA functional prediction based on RNA–RNA and RNA–protein interaction data. This approach utilizes the IntaRNA2 algorithm to predict RNA–RNA interactions along with publicly available CLIP-seq (cross-linking immunoprecipitation sequencing) data and the GraphProt computational framework to determine RNA–protein interactions before combining data and predicting function based on several predetermined possibilities \[194,195\].

Possibly, the most well-known method for IncRNA regulatory predictions is the LongHorn algorithm, which derives IncRNA interactions based on four predefined models of IncRNA regulation. These models include activities such as decoys, miRNA sponges, cofactors/guides, and switches \[30\]. In a seminal publication, Chiu et al. applied the algorithm to TCGA data and predicted IncRNA–gene interaction predictions for individual cancers of interest. The IncRNA–gene predictions are available presently for 13 tumor contexts including prostate adenocarcinoma, lung adenocarcinoma, and ovarian serous cystadenocarcinoma. For PC, Longhorn was applied to TCGA data including RNA, miRNA expression and copy number profiles of 371 prostate adenocarcinoma tumors. PLAIDOH (Predicting IncRNA Activity through Integrative Data-driven Omics and Heuristics) was applied to identify IncRNAs associated with non-Hodgkin lymphoma \[196\]. PLAIDOH offers users the opportunity to analyze both publicly and privately owned datasets, analyze statically small datasets, and ultimately rank IncRNAs in order of importance based on customizable metrics \[196\]. One limitation is that it only predicts IncRNA targets that are located within a genomic window limit of 800 kb. Using PLAIDOH, researchers have validated predictions by identifying well-recognized IncRNA gene regulatory pairs, such as HOTAIR matching with HOX genes and familiar IncRNA-protein binding pairs such as NEAT1 and NONO \[196\], new IncRNA interactions have also been uncovered and subsequently validated, demonstrating the utility of this tool.

IncRNA-screen is an interactive platform that operates as a fully functional analysis pipeline which performs all major steps involved with typical RNA-sequencing processing, and its main objective is to facilitate the computational discovery of IncRNA candidates that can be further examined by functional experiments \[181\]. This pipeline can automatically download and prepare raw FASTQ (sequencing) files from multiple public repositories and performs data quality control, pre-processing, alignment, annotation, differential expression, and novel IncRNA identification and classification. After analysis, IncRNA-screen provides users with an interactive HTML summary report that can be searched and filtered to suit users' individual needs. Information includes locus, number of exons, a genome snapshot and IncRNA–mRNA interactions. IncRNA-screen is limited to IncRNA-gene predictions by physical location (neighboring genes).

These databases and computational tools may be applied to understand the regulatory impact of IncRNAs upon oncogenes and tumor suppressors \[197,198\]. The resultant predictions are useful for the identification of candidate cancer biomarkers and therapeutic targets, in concert with traditional biochemical methods \[199\].

8. Concluding remarks

IncRNAs are now established regulators of a vast number of biological processes as well key mediators for the development and progression of various cancers. This is remarkable, considering that IncRNAs were once branded as ‘junk’ with no regulatory capabilities. Despite significant progress, there is still much to learn about IncRNAs and their potential value as cancer biomarkers or in therapy.

9. Expert opinion

At present, there are three areas of prostate cancer research that require focus: 1) understanding fundamental biological reasons as to why men of African descent are significantly more prone to disease morbidity and mortality, 2) improved therapeutic treatments for advanced and castrate-resistant PC, and 3) the necessity for biomarkers, which can differentiate low-risk and high-risk diseases and assess PC racial disparities. Elucidating the functional roles of IncRNAs will help drive advances in each of these areas and will lead to better diagnosis and treatment of PC.

Despite these advances, opinions in the scientific community as to whether IncRNAs provide realistic diagnostic, prognostic, or diagnostic targets are varied at this time. This will add to the complexity of implementing clinical testing based on IncRNA diagnostics. The Clinical Laboratory Improvement Amendments (CLIA) federal regulatory standards require that clinical laboratories determine and detail their own performance specifications for laboratory-developed tests to guarantee accurate and precise results prior to implementation of the test. Characteristics that must be established typically include accuracy and precision, range of reporting, analytical sensitivity and specificity, and reference intervals. Although both a positive and negative attribute, the sheer number of IncRNAs typically predicted by computational tools presents challenges for both clinical and research practice. The hundreds to thousands of IncRNAs that may be identified computationally cannot be feasibly validated in a clinical setting. This elevates the need for more sophisticated computational
algorithms based on deep learning and other approaches, which not only accurately predict IncRNAs but also facilitate stringent thresholds and intelligent approaches to triage-specific IncRNAs subsets for laboratory-based testing and ultimately clinical implementation.

As precision medicine approaches evolve and become more inclusive, an important impact will be the advancement of knowledge of racial differences among prostate and other cancers. Several studies have determined that IncRNAs may play a role in differences in racial outcomes in cancer. There remain considerable and persistent racial disparities in prostate cancer outcomes. Reduced access to health-care services contributes to racial disparities in PC outcomes, but even in equal access health-care systems, such as the Veterans Administration (VA) in the USA, African American Veterans have higher PSA levels and higher-grade tumors compared to European American Veterans even when presenting with the same stage of disease. Unfortunately, medical mistrust issues and their associated indicators among African American patients have an adverse effect in recruiting sufficient sample numbers of ethnic minorities and their participation in clinical trials. Distress in the health-care system and fear associated with detection and treatment are responsible for late diagnosis in African American men. There is a paucity of samples from racial minorities in public genomic databases to detect moderately common genomic alterations in African Americans, which is contributing to the gap in PC disparities. We anticipate that as the IncRNA research field develops and matures, greater emphasis will be placed on identifying IncRNAs that play key roles in the development and progression of cancer across all ethnic groups.

This field has already significantly progressed in a relatively short period of time, from the first IncRNA identified (h19) to the realization that IncRNAs do in fact process regulatory capabilities. With novel IncRNAs being identified daily and associated with cancer, the pressure to experimentally validate these emerging candidates is increasing. In the future, we expect IncRNA computational tools to become an essential tool for both research and clinical laboratories and for a variety of diseases beyond cancer. We also expect significant growth in curated IncRNA databases with improved functional predictions of IncRNAs. We anticipate increased commercial activities based around IncRNA therapeutics, prognostics, and diagnostics.

**Funding**

This paper was funded by support from NIH/NIDA 1U01DA045300-01A1, U54MD010706-CHH and start-up funding from Queens University Belfast to G Hardiman.

**Declaration of Interest**

IO has undertaken a consultancy for Mevox Ltd. The other authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

**ORCID**

Rebecca Morgan [http://orcid.org/0000-0003-3589-9143](http://orcid.org/0000-0003-3589-9143)

Ian Overton [http://orcid.org/0000-0003-1158-8527](http://orcid.org/0000-0003-1158-8527)

Gary Hardiman [http://orcid.org/0000-0003-4558-0400](http://orcid.org/0000-0003-4558-0400)

**References**

Papers of special note have been highlighted as either of interest (•) or of considerable interest (••) to readers.


Detailed review of IncRNAs in different tissue context.


Extensive review on IncRNAs and their association with prostate cancer.


