Advances in Prognostic Methylation Biomarkers for Prostate Cancer

Dilys Lam¹, Susan Clark¹,², Clare Stirzaker¹,²,* and Ruth Pidsley ¹,²,#†

¹ Epigenetics Research Laboratory, Genomics and Epigenetics Division, Garvan Institute of Medical Research, Sydney, 2010, New South Wales, Australia.
² St. Vincent's Clinical School, UNSW Sydney, 2010, New South Wales, Australia.
* equal
† Correspondence: r.pidsley@garvan.org.au; Tel.: +61 2 92958315

Simple Summary: Prostate cancer is a major cause of cancer-related death in men world-wide. There is an urgent clinical need for improved prognostic biomarkers to better predict the likely outcome and course of the disease and thus inform the clinical management of these patients. Currently, clinically recognized prognostic markers lack sensitivity and specificity in distinguishing aggressive from indolent disease, particularly in patients with localised, intermediate grade prostate cancer. Thus, there is major interest in identifying new molecular biomarkers to complement existing standard clinicopathological markers. DNA methylation is a frequent alteration in the cancer genome and offers potential as a reliable and robust biomarker. In this review, we provide a comprehensive overview of the current state of DNA methylation biomarker studies in prostate cancer prognosis. We highlight advances in this field that have enabled the discovery of novel prognostic genes and discuss the potential of methylation biomarkers for non-invasive liquid-biopsy testing.

Abstract: There is a major clinical need for accurate biomarkers for prostate cancer prognosis, to better inform treatment strategies and disease monitoring. Current clinically recognised prognostic factors, including prostate-specific antigen (PSA) levels, lack sensitivity and specificity in distinguishing aggressive from indolent disease, particularly in patients with localised intermediate grade prostate cancer. There has therefore been a major focus on identifying molecular biomarkers that can add prognostic value to existing markers, including investigation of DNA methylation, which has a known role in tumorigenesis. In this review, we will provide a comprehensive overview of the current state of DNA methylation biomarker studies in prostate cancer prognosis, and highlight the advances that have been made in this field. We cover the numerous studies into well-established candidate genes, and explore the technological transition that has enabled hypothesis-free genome-wide studies and the subsequent discovery of novel prognostic genes.

Keywords: 1; DNA methylation, 2; epigenetics, 3; biomarkers, 4; circulating DNA, 5; cfDNA, 6; prostate cancer, 7; early detection, 8; prognosis

1. Introduction

Prostate cancer (PCa) is the most commonly diagnosed non-cutaneous cancer in men and one of the leading causes of cancer death in males. Globally, 1,276,106 new cases were diagnosed in 2018 alone [1,2], and this number is projected to rise by ~80%, to more than two million new cases a year, by 2040 [3-5]. Currently, PCa diagnosis is achieved through assessment of blood prostate-specific antigen (PSA) levels, digital rectal examination (clinical T-stage) and histological examination of needle biopsies (Gleason Score (GS)/ISUP Grade) [6]. PSA-based screening was introduced in the late 1980s, and has significantly increased the early detection of localized disease [7-9]. Diagnosis at this early organ-confined stage of disease is crucial as it is potentially curable by radical prostatectomy (RP), a procedure to surgically remove the whole prostate gland. While this is curative for most
prostate cancers, approximately 30% of patients treated by RP experience biochemical recurrence (BCR) [10], and 17-22% of these relapsed patients progress to metastatic-lethal PCa [11-13]. There is therefore a need to identify the men at high-risk of metastatic progression, so that additional interventions can be offered earlier (e.g. adjuvant therapies such as chemotherapy and radiotherapy) [14].

On the other hand, many men diagnosed with PCa have an indolent form of the disease, which is characterised by slow progression with no eventual clinical manifestation [15,16]. For these men RP represents an overtreatment given the risk of unnecessary side effects and compromised quality of life [17,18]. Thus strategies such as ‘watchful waiting’ and ‘active surveillance’ have emerged for men diagnosed with low grade disease, in which regular monitoring is used to detect tumour progression, with the aim of delaying RP until it is clinically necessary [19]. However, 13-45% of low-risk men on active surveillance develop a BCR and progress to surgery [20-22], indicating that they may have been inappropriately assigned to monitoring, and should have been treated earlier.

PCa has a heterogeneous clinical course which makes it challenging to decide the most appropriate treatment or monitoring strategy for individual patients [23]. There is a major unmet clinical need for specific prognostic biomarkers that can accurately differentiate indolent from aggressive tumours that are likely to metastasise and lead to lethal disease [24]. The ability to identify the risk of progression at initial diagnosis would inform decisions about personalized treatment and/or monitoring strategies, as well as the use of adjuvant therapy, for improved clinical management and enhanced outcomes for PCa patients (Figure 1).

![Figure 1](https://via.placeholder.com/150)

**Figure 1.** Schematic showing the pathway for the application of biomarkers for improved prognostic stratification in prostate cancer patients. Prostate cancer is a heterogeneous disease and identifying those patients at diagnosis that have aggressive versus indolent disease is critical in informing the clinical management of these patients. Biological sampling includes tissue biopsies, and blood and urine samples. These are assayed using molecular biomarkers, including DNA mutations, DNA methylation and transcription, and clinicopathological markers, such as GS/ISUP grade and tumour staging. The ultimate goal is to improve the prognostic stratification of patients to inform the optimal treatment strategies for prostate cancer patients. TNM: tumour/nodes/metastasis staging; PSA: prostate-specific antigen.

1.1. Current clinicopathological prognostic markers

The existing scoring systems and nomograms that are used to identify patients at high-risk of progression are based solely on routine clinical and pathological variables at the time of diagnosis or surgery, including pre-operative (pre-op) PSA levels, GS/ISUP grade scoring, tumour staging (clinical or pathological T-stage) and margin status (Figure 1) [11,25-30]. These tools, whilst useful in the diagnostic assessment of the tumour, lack sensitivity and specificity in classifying the risk of an individual patient’s disease [6,31-33]. One problem with the current prognostic approach is that a number of scoring systems exist, with no single accepted standard system being used in clinical practice [29,34]. Furthermore, these prognostic variables are based on histological assessment, which may not capture the underlying biology of a tumour and its potential to progress to an aggressive, lethal cancer. Advances in molecular profiling techniques mean that we can now access another layer of biological information in tumours, including early molecular changes that may precede
histologically visible alterations [23,35]. Much research is now being dedicated to investigating whether molecular information can better inform clinical decisions about individualised treatment and/or monitoring strategies, as well as providing new biological insights to guide the development of new therapies [35].

1.2. Molecular biomarkers for prognosis

There is increasing evidence for the value of prognostic molecular biomarkers to complement existing standard clinicopathological markers. Molecular prognostication of PCa has been investigated in various contexts, including at the level of genetic, epigenetic and gene expression alterations [36,37], from both tumour tissue [37,38] and liquid biopsies [39,40] (Figure 1). To date, the major success stories of prognostic molecular biomarkers are the commercialised tissue-based tests centred on panels of gene expression signatures. These tests include Prolaris (Myriad Genetics), Oncotype DX Prostate (Genomic Health) and Decipher (GenomeDX Biosciences) [37]. The Prolaris test determines risk of progression (BCR, cancer-specific death) by measuring a proliferation signature of 31 cell cycle progression transcripts [41], whilst Oncotype DX predicts adverse pathology (high grade/stage disease) or poor outcome (BCR) based on 12 genes [42]. Decipher is based on 22 markers which have been trained to predict early metastasis and aggressive PCa [43]. Whilst these tests demonstrate the potential for molecular biomarkers, they have yet to be integrated into the standard clinical routine following the initial diagnosis [37].

1.3. DNA methylation biomarkers

DNA methylation is one of the earliest, most stable, and frequent alterations in the cancer genome and has been extensively investigated as a source of molecular biomarkers [36,44-46]. DNA methylation is an epigenetic modification, in which a methyl group is added to the cytosine base of a cytosine-guanine (CpG) dinucleotide. It is associated with gene regulation and function, with promoter-associated clusters of marks, termed CpG islands, often linked to gene silencing [47,48]. In the context of PCa, aberrant DNA methylation is a key feature observed during early tumorigenesis, as well as in its progression and metastatic development [49-51] and occurs at a much higher frequency and more consistently than genetic mutations [52]. Additionally, DNA methylation has been shown to outperform gene expression in detecting cancer from prostate biopsies [53]. The feasibility of using DNA methylation as a biomarker is further supported by the fact that DNA is more stable than RNA [54] and its methylation patterns are retained following long-term storage of clinical material, including as formalin-fixed paraffin-embedded tissue (FFPET). DNA methylation assays can also be easily integrated into routine clinical practice as many diagnostic labs already have the infrastructure in place, due to their similarity with DNA-sequence-based biomarker approaches [55]. There is currently one validated epigenetic test commercially available, ConfirmMDx (MDxHealth), designed for diagnostic rather than prognostic purposes, that uses the methylation profile of 3 genes (**APC, RASSF1, GSTP1**) to detect cancer in histologically negative biopsies [56]. PCa-derived aberrant DNA methylation patterns have also been detected in liquid biopsies such as blood and urine samples, paving the way to the development of non-invasive molecular tests [57-60].

2. Current state of prognostic methylated biomarkers

A plethora of studies have been conducted over the last two decades investigating DNA methylation-based biomarkers to aid PCa prognosis. The majority of these studies have interrogated primary prostate tumours extracted from RP tissue, whilst others have used prostate tissue from core or needle biopsies, transurethral resection of the prostate (TURP) specimens, as well as tumour-adjacent and benign non-neoplastic prostate specimens. This section will primarily review studies performed on RP tissue, and studies that have used other types of prostate specimens will be noted accordingly. Initial studies in this field, limited by the laboratory techniques available, took the traditional *a priori* approach of examining genes that had been implicated in biological pathways of PCa. These studies used targeted methylation profiling techniques including methylation-specific
PCR (MSP) [61], quantitative methylation-specific PCR (qMSP) [62], pyrosequencing [63,64] and mass spectrometry (MassARRAY EpiTYPER) [65] to assess the methylation profile at a specific gene of interest (Figure 2). Figure 2 lists other targeted approaches that have also been used to assess methylation in cancer, including ddPCR [66], COBRA [67], high resolution melt curve [68] and headloop MSP [69-71]. With recent advances in the technological capabilities to interrogate the methylome more broadly, the field has transitioned into conducting hypothesis-free genome-wide screens for novel prognostic methylation biomarkers. The candidate and genome-wide studies reviewed below use a range of outcome measures to assess the prognostic value of methylated genes, including low vs high grade cancers, localised vs advanced disease, and clinical outcomes such as BCR, metastatic relapse and PCa-specific death (PCa death). Importantly, the clinical outcome most often studied is BCR, defined by an increase in serum PSA levels post-RP. However, there is increasing evidence that BCR is not a sufficient indicator of progression to aggressive lethal disease, with metastatic relapse instead being the clinically relevant endpoint for predicting survival [72,73]. This requires long-term follow-up (≥ 15 years) for metastatic relapse and PCa death to manifest [16], which many study populations lack, thus reducing their ability to fully evaluate and assess molecular biomarkers of PCa prognosis.

Figure 2. Methylation profiling approaches for biomarker discovery and validation. (a) Targeted candidate gene approaches used in a priori studies and for the validation of novel candidate markers. (b) Microarray and sequencing-based genome-wide approaches used for the discovery of novel biomarkers – comparison of the methodology, number of CpGs and/or genes targeted, and coverage of the methylome across the different platforms. PCR: Polymerase chain reaction; MSP: Methylation-specific PCR; ddPCR: Droplet digital PCR; COBRA: Combined bisulphite restriction analysis; qMSP: quantitative methylation-specific PCR; HM: Human Methylation; RRBS: Reduced representation bisulphite sequencing; MBDCap-Seq: Methyl-CpG binding domain capture sequencing; WGBS: Whole genome bisulphite sequencing.
2.1. Candidate (a priori) markers

To date, DNA methylation at over 60 candidate genes has been investigated. To summarise the top candidate markers with the greatest prognostic evidence, this review focuses on genes that have been investigated in at least three studies and used well-defined prognostic outcomes. Applying these criteria, we identified the 26 studies detailed in Table 1. Figure 3 visualises the markers examined in these studies and specifies which of these were found to have potential prognostic value in univariate (U) and/or multivariate (M) statistical models (adjusted for clinicopathological variables) of prognosis. Below we discuss the genes that have been most extensively studied and validated across independent studies: GSTP1, APC, RARB, PITX2, CCND2 and PTGS2.

![Figure 3. Genes studied in ≥ 3 prognostic candidate marker studies.](image)

Each row represents a candidate gene study, and each column represents a gene for which three or more studies have investigated its methylation profile as a prognostic marker of PCa. The number in brackets indicates the number of studies the gene has been examined in. For each of the 26 studies, genes that were investigated have been highlighted: grey = examined in study but no significant associations were found; faded green = significant associations found in univariate analysis only; light green = significant associations observed in multivariate analysis only; green = significant associations reported in both univariate and multivariate analysis.
<table>
<thead>
<tr>
<th>Study [ref]</th>
<th>Primary outcomes examined</th>
<th>Total cohort size</th>
<th>Additional cohort details</th>
<th>Follow-up (years)</th>
<th>Method</th>
<th>Genes examined</th>
<th>Genes validated&lt;sup&gt;b,d&lt;/sup&gt;</th>
<th>Results&lt;sup&gt;f&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yegnasubraman et al. 2004 Cancer Res [74]</td>
<td>BCR</td>
<td>n = 36</td>
<td>n = ns BCR</td>
<td>Range: 8 - 13</td>
<td>qMSP</td>
<td>GSTP1, APC, RASSF1, PTGS2, MDR1, HIC1, EDNRB, ESR1, CDKN2a, CDKN2b, p14/QRF, MGMT, hMLH1, TIMP3, DAPK1, CDH1</td>
<td>U: PTGS2 M: PTGS2 [+ GS + pathological T-stage]</td>
<td>U: HR 2.82 (1.07-7.44), P = 0.04 M: HR 4.26 (1.36-13.36), P = 0.01</td>
</tr>
<tr>
<td>Rosenbaum et al. 2005 Clin Cancer Res [75]</td>
<td>Primary: Progression</td>
<td>n = 74</td>
<td>n = 37 no progression, n = 37 progression (n = 14 BCR, n = 16 metastatic relapse, n = 7 PCa death)</td>
<td>Min: 7, Median: 9</td>
<td>qMSP</td>
<td>APC, CCND2, GSTP1, TIG1, RASSF1, RARB</td>
<td>Progression: U: GSTP1 M: [+ age &gt; 60] (1) GSTP1 + APC (2) GSTP1 + APC and/or CCND2</td>
<td>Metastatic relapse and/or PCa death: U: None M: None</td>
</tr>
<tr>
<td>Woodson et al. 2006 Epigenetics [76]</td>
<td>BCR</td>
<td>n = 60</td>
<td>n = 49 no BCR, n = 11 BCR</td>
<td>Mean (SD): No BCR- 6.5 (3.2) BCR- 4.7 (2.8)</td>
<td>qMSP</td>
<td>GSTP1, RARB, CD44, PTGS2</td>
<td>U: (1) CD44; (2) PTGS2 M: CD44 + PTGS2 [+ GS]</td>
<td>U: (1) OR 6.83 (1.67-27.99), P = 0.008; (2) OR 4.38 (1.13-17.40), P = 0.04 M: CD44 + PTGS2 [OR 8.87 (1.85-42.56), P = 0.006]</td>
</tr>
<tr>
<td>Henrique et al. 2007 Clin Cancer Res [77]</td>
<td>Primary: PCa death</td>
<td>n = 83</td>
<td>n = 15 PCa death, n = 37 BCR</td>
<td>Median (range): 3.7 (0.5 - 5)</td>
<td>qMSP</td>
<td>APC, CCND2, GSTP1, RARB, RASSF1</td>
<td>PCa death: U: APC M: APC [+ Clinical T-stage] BCR: U: (1) APC; (2) GSTP1; (3) RASSF1 M: APC [+ Clinical T-stage]</td>
<td>U: Log-rank P = 0.010 M: 3.51 (1.23-9.96), P = 0.018 BCR: U: Log-rank (1) P = 0.002; (2) P = 0.047; (3) P = 0.019 M: OR 2.58 (1.29-5.16), P = 0.008</td>
</tr>
<tr>
<td>Ellinger et al. 2008 Urology [78]</td>
<td>BCR</td>
<td>n = 41</td>
<td>n = 28 no BCR, n = 13 BCR</td>
<td>Mean; median (range): 2.3; 1.7 (0.5 - 6)</td>
<td>qMSP</td>
<td>Annexin2, APC, EDNRB, GSTP1, PTGS2, MDR1, RARB, Reprimo, TIG1</td>
<td>U: (1) APC + Reprimo, (2) &gt; 5 genes hypermethylated together M: None</td>
<td>U: Log-rank (1) P = 0.0078; (2) P = 0.0074 M: non-sig</td>
</tr>
<tr>
<td>Alumkal et al. 2008 Urology [79]</td>
<td>BCR</td>
<td>n = 151</td>
<td>n = 104 no BCR, n = 47 BCR</td>
<td>At least 5 years</td>
<td>Nested MSP</td>
<td>GSTP1, MGMT, ASC, CDKN2A, EDNRB, CDH13, CD44, TIMP3, RUNX3, APC, WIF-1</td>
<td>U: CDKN2A M: CDH13 and/or ASC [+ GS + extra capsular penetration + seminal vesicle involvement + margin status]</td>
<td>U: OR 0.43 (0.19-0.98), P = 0.05 M: CDH13 [OR 5.51 (1.34-22.67), P = 0.02], CDH13 and/or ASC [OR 5.64 (1.47-21.7), P = 0.01, sensitivity = 72.3%, specificity = 48%]</td>
</tr>
<tr>
<td>Study</td>
<td>Year</td>
<td>Journal</td>
<td>Marker/Significance</td>
<td>BCR</td>
<td>No BCR</td>
<td>BCR Rate</td>
<td>Method</td>
<td>qMSP</td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
<td>---------</td>
<td>---------------------</td>
<td>-----</td>
<td>--------</td>
<td>----------</td>
<td>--------</td>
<td>------</td>
</tr>
<tr>
<td>Weiss et al. 2009</td>
<td>The Journal of Urology [80]</td>
<td>BCR</td>
<td>n = 605</td>
<td>n = 540 no BCR, n = 65 BCR</td>
<td>Median: 5.5</td>
<td>qMSP</td>
<td>ABHD9, Chr3-EST, GPB7, HIST2H2BF, CCND2, PITX2</td>
<td>U: (1) ABHD9; (2) Chr3-EST; (3) GPR7; (4) HIST2H2BF; (5) PITX2 (also in GS7 only and GS8 only patients)</td>
</tr>
<tr>
<td>Vanaja et al. 2009</td>
<td>Cancer Invest [81]</td>
<td>Recurrence within 5 years</td>
<td>n = 64</td>
<td>n = 32 no recurrence, n = 32 recurrence (n = 10 BCR, n = 10 local, n = 12 metastatic relapse)</td>
<td>Range: 0 - 5</td>
<td>MassARRA Y EpiTYPER</td>
<td>FLNC, EFS, ECGR4, RARB, PITX2, GSTP1, PDLIM4, KCNMA1</td>
<td>U: (1) FLNC (6 CpG units), (2) GSTP1 (3 CpG units), (3) PITX2 (1 CpG unit)</td>
</tr>
<tr>
<td>Richiardi et al. 2009</td>
<td>J Clin Oncol [82]</td>
<td>PCa death</td>
<td>1980s cohort: n = 216</td>
<td>1990s cohort: n = 243</td>
<td>1980s cohort: n = 95 no PCa death, n = 121 PCa death</td>
<td>1990s cohort: n = 167 no PCa death, n = 76 PCa death</td>
<td>Median (range): 1980s cohort - 3.1 (0 - 14) 1990s cohort - 6.3 (0 - 14)</td>
<td>MSP</td>
</tr>
<tr>
<td>Banez et al. 2010</td>
<td>J Urol [83]</td>
<td>BCR</td>
<td>n = 476</td>
<td>n = 370 no BCR, n = 106 BCR</td>
<td>ns</td>
<td>qMSP</td>
<td>PITX2</td>
<td>U: PITX2 (also in GS7 patients only)</td>
</tr>
<tr>
<td>Vasiljevis et al. 2011</td>
<td>Disease Markers [84]</td>
<td>Low vs High GS</td>
<td>n = 48</td>
<td>n = 9 GS6, n = 23 GS7, n = 7 GS8, n = 9 GS9-10</td>
<td>NA</td>
<td>Pyro sequencing</td>
<td>RARB, GSTP1, HIN1, APC, BCL2, CCND2, CDH13, EGFR5, NKX2-5, RASSF1, DPYS, MDR1, PTGS2, EDNRB, MAL, PDLIM4, HLAa, TIG1, ESR1, SLIT2, CDKN2A, MCAM, SFN, THR8, CDH1, TWIST1</td>
<td>U: SFN, SLIT2, SERPINB5 (pairwise measures)</td>
</tr>
</tbody>
</table>

P-values not specified in this study
<table>
<thead>
<tr>
<th>Study</th>
<th>Follow-up</th>
<th>Survival Analysis</th>
<th>qMSP</th>
<th>Training Cohort</th>
<th>Testing Cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rosenbaum et al. 2012 The Prostate [85]</td>
<td>n = 95 n = 47 no progression, n = 48 progression (n = 22 BCR only), n = 17 metastatic relapse, n = 9 PCa death [all GS 7 (3+4)]</td>
<td>All: At least 8 years no progression (Median (range)) - 10 (8 - 14) Progression (Median) - 8</td>
<td>qMSP</td>
<td>RBM6, MT1G, CDH1, AIM1, KIF1A, PAK3</td>
<td>U: A1M1 [+ age at diagnosis + lymph node status] M: Metastatic relapse and/or PCa death: U: None M: None</td>
</tr>
<tr>
<td>Vasiljevic et al. 2012 Prostate Cancer Prostatic Dis [86]</td>
<td>n = 349 n = 258 no PCa death, n = 91 PCa death</td>
<td>Median (max): 9.5 (20) Pyro sequencing</td>
<td>HSPB1</td>
<td>U: HSPB1 (in all samples and in subgroup of GS &lt; 7) M: HSPB1 [+ GS + extent of disease (T score), PSA + HSPB1 x GS interaction term]</td>
<td>Training cohort: U: PITX2 M: Not conducted</td>
</tr>
<tr>
<td>Dietrich et al. 2013 J Mol Diagn [87]</td>
<td>BCR n = 157 Training cohort: n = ns BCR Testing cohort: n = 523 [same cohort as Banez et al. 2010 [83]]</td>
<td></td>
<td>ns qMSP</td>
<td>PITX2</td>
<td>Training cohort: U: 3.479 (1.259-9.614), P-value not given M: NA</td>
</tr>
<tr>
<td>Richiardi et al. 2013 PloS ONE [88]</td>
<td>PCa death n = 157 [Non-neoplastic tissue adjacent to prostate tumour]</td>
<td>Median (range): 6.8 (0.03 - 24.1) qMSP</td>
<td>APC, GSTP1</td>
<td>U: (1) APC, (2) GSTP1 M: APC + GSTP1 [+ age at diagnosis + year of diagnosis + source of tumour tissue + methylation in prostate tumour tissue + GS] (also in restricted analyses of first 5 years of follow-up)</td>
<td>Training cohort: U: 2.38 (1.23-4.61), P-value not given; (2) HR 2.92 (1.49-5.74), P-value not given M: APC + GSTP1 [HR 2.40 (1.15-5.01), P = 0.032]; first 5 yrs follow-up: HR 3.29 (1.27-8.52), P = 0.019</td>
</tr>
<tr>
<td>Study</td>
<td>Year</td>
<td>Journal</td>
<td>n</td>
<td>Clinical Stage</td>
<td>nPCR</td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
<td>---------</td>
<td>---</td>
<td>----------------</td>
<td>-----</td>
</tr>
<tr>
<td>Moritz et al. 2013 Anticancer Res [89]</td>
<td>BCR</td>
<td>n = 84</td>
<td>n = 31 no BCR, n = 53 BCR [GS 5 - 7]</td>
<td>Mean; median (range): 4; 1.8 (0 - 10.9)</td>
<td>qMSP</td>
</tr>
<tr>
<td>Vasiljevic et al. 2014 BMC Cancer [90]</td>
<td>PCa death</td>
<td>n = 367</td>
<td>n = 90 no PCa death, n = 45 PCa death [all GS ≤ 7] [subset of cohort from Vasiljevic et al. 2014 BMC Cancer [90]]</td>
<td>Median (range): 9.5 (0.7 - 19.6)</td>
<td>Pyro sequencing</td>
</tr>
<tr>
<td>Vasiljevic et al. 2014 Biomarkers in Medicine [91]</td>
<td>PCa death</td>
<td>n = 135</td>
<td>n = 193 no progression, n = 259 progression</td>
<td>No PCR death (Median) - 7.8 PCa death (Median (max)) - 15.3 (20)</td>
<td>Pyro sequencing</td>
</tr>
<tr>
<td>Maldonado et al. 2014 J Urol [92]</td>
<td>Progression</td>
<td>n = 452</td>
<td>n = ns BCR</td>
<td>No BCR (Median (range)) - 3.4 (0.2 - 5.5)</td>
<td>qMSP</td>
</tr>
<tr>
<td>Daniunaite et al. 2014 J Urol [93]</td>
<td>BCR</td>
<td>n = 149</td>
<td>Training cohort: n = 117 no CF, n = 30 CF Validation cohort: n = 58 no CF, n = 13 CF [High-risk PCA patients: Clinical stage ≥ T3a, GS</td>
<td>Median (range): Training cohort - 6.8 (0.1 - 12.8) Validation cohort - 11.5 (1.4 - 18.8)</td>
<td>Multiplex qMSP</td>
</tr>
<tr>
<td>Litovkin et al. 2015 PloS ONE [94]</td>
<td>Clinical Failure</td>
<td>Training cohort: n = 147 Validation cohort: n = 71</td>
<td>Training cohort: n = 117 no CF, n = 30 CF Validation cohort: n = 58 no CF, n = 13 CF [High-risk PCA patients: Clinical stage ≥ T3a, GS</td>
<td>Median (range): Training cohort - 6.8 (0.1 - 12.8) Validation cohort - 11.5 (1.4 - 18.8)</td>
<td>Multiplex qMSP</td>
</tr>
<tr>
<td>Study</td>
<td>Time of Death</td>
<td>n</td>
<td>PCa Specimens details</td>
<td>Study Details</td>
<td>Detection Methods</td>
</tr>
<tr>
<td>-------</td>
<td>--------------</td>
<td>---</td>
<td>----------------------</td>
<td>---------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Carozzi et al. 2016 J Cancer Res Clin Oncol [95]</td>
<td>PCa death (within 10 yrs of diagnosis)</td>
<td>n = 129</td>
<td>n = 91 alive &gt; 10 yrs, n = 38 died ≤ 10 years [Needle biopsy specimens]</td>
<td></td>
<td>Pyro sequencing</td>
</tr>
<tr>
<td>Rybicki et al. 2016 Int J Cancer [96]</td>
<td>BCR</td>
<td>n = 353</td>
<td>n = 262 no BCR, n = 91 BCR (White: n = 152 no BCR, n = 54 BCR African American: n = 110 no BCR, n = 37 BCR) [Benign prostate specimens – patients who eventually developed PCa]</td>
<td></td>
<td>Nested MSP</td>
</tr>
<tr>
<td>Holmes et al. 2016 Clin Epigenetics [97]</td>
<td>BCR</td>
<td>Cohort 1: (TCGA): n = 498 Cohort 2: n = 300</td>
<td></td>
<td>Cohort 1: n = ns BCR Cohort 2: n = ns BCR Mean; median: Cohort 1 - 1.83; 1.3 Cohort 2 - 5.5; 5.2 Cohort 1 - HM450K Microarray</td>
<td>(1) PITX3; (2) PITX3 + PITX2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Cohort 2 - qMSP</td>
</tr>
<tr>
<td>Ahmad et al. 2016 Oncotarget [98]</td>
<td>PCa death</td>
<td>n = 385</td>
<td>n = 328 no PCa death, n = 57 PCa death [low (0-2) to Median (IQR): 11.36 (6.20 - 14.72)]</td>
<td>Pyro sequencing</td>
<td>HSPB1, CCND2, TIG1, DPYS, PITX2, MAL</td>
</tr>
</tbody>
</table>

- 8-10 and/or PSA >20ng/ml
- M: (1) HR 4.27 (1.03-17.72), P = 0.046, C-index = 0.80 (vs 0.75 for stage + GS + PSA only); (2) HR 0.19 (0.05-0.79), P = 0.022; (3) HR 3.81 (1.09-13.34), P = 0.036
- U: 2nd quartile [OR 1.54 (0.56-4.23)], 3rd quartile [HR 2.42 (0.91-6.49), P = 0.0474 M: non-sig

- Carozzi et al. 2016 J Cancer Res Clin Oncol [95]
- Rybicki et al. 2016 Int J Cancer [96]
- Holmes et al. 2016 Clin Epigenetics [97]
- Ahmad et al. 2016 Oncotarget [98]
<table>
<thead>
<tr>
<th>Study</th>
<th>Cohort 1:</th>
<th>Cohort 2:</th>
<th>qMSP</th>
<th>U: Methylation score</th>
<th>M: Methylation score [+ CAPRA]</th>
<th>44% (vs CAPRA only: 68%/44%)</th>
<th>At 10 yr follow-up: AUC = 0.74 (vs CAPRA only: 0.62)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uhl et al. 2017 J Mol Diagn [99]</td>
<td>BCR</td>
<td>ISUP grade group (as surrogate for survival)</td>
<td>n = 208</td>
<td>no BCR, n = 52 BCR</td>
<td>[same cohort as Cohort 2 in Holmes et al. 2016 [97]]</td>
<td>Cohort 1: U: PITX2 M: None</td>
<td>Cohort 1: U: HR 1.77 (1.01-3.10), P = 0.046 M: non-sig</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Abbreviations:**
- AUC = area under the curve; CF = clinical failure; df = degrees of freedom; GS = Gleason Score; HR = hazard ratio; IQR: Interquartile range; LR = likelihood ratio; M = multivariate analysis; MSP = methylation-specific PCR; NA = not applicable; non-sig = non-significant; ns = not specified; OR = odds ratio; PCa = prostate cancer; PSA = prostate-specific antigen; qMSP = quantitative methylation-specific PCR; r = correlation coefficient; U = univariate analysis. **Definitions:**
- BCR: Biochemical recurrence - PSA elevations ≥ 0.2ng/ml post-RP, except [77] > 0.4ng/ml and [89]> 0.1ng/ml; Clinical Failure: either of local recurrence or metastatic relapse; Clinical T-stage: Tumour staging based on results of digital rectal examination, PSA levels and GS. Local recurrence: cancer observed on prostatic bed, confirmed by histological examination of biopsies; Metastatic relapse: metastatic deposits (visceral, bony metastasis) confirmed by positive biopsies or cT/bone scans; Pathological T-stage: Tumour staging based on pathological examination of surgically removed prostate tissue; PCa death: Prostate cancer-specific death; Progression: either of BCR, metastatic relapse or PCa death; Recurrence: either of BCR, local recurrence or metastatic relapse. * All studies are on prostate cancer tissues from radical prostatectomy, unless specified. b Univariate (U) or Multivariate (M) analyses. * plus (+) sign indicates variables in the same multivariate model or methylation score together. * square bracket ([]) indicate the clinicopathological factors adjusted for in each multivariate model. * genes validated may have been impacted by high number of missing cases. * number in brackets following HR or OR indicate the 95% confidence interval.
The glutathione S-transferase pi gene (GSTP1) is the most well-studied DNA methylation biomarker of PCa, particularly in diagnosis [100]. It encodes glutathione S-transferase, a detoxifying enzyme and tumour suppressor involved in drug metabolism and protecting DNA from oxidative damage [101]. Hypermethylation (increased methylation) of GSTP1 is observed frequently in PCa tissue but rarely in histologically negative prostate tissues [102]. We identified 15 candidate prognostic biomarker studies that have studied GSTP1 methylation, with its prognostic value validated in eight of these studies (Table 1). The earliest of these studies assessed GSTP1 methylation using qMSP in a cohort of GS 7 (3 + 4) patients \((n = 74)\), and reported that GSTP1 hypermethylation was significantly associated with time to progression (any of BCR, metastatic relapse and/or PCa death) in univariate analysis, and as independent predictor in multivariate analysis with other candidate genes [75]. Subsequent studies of RP tissue, using qMSP, have again found GSTP1 to be an independent prognostic factor. Briefly, Maldonado and colleagues used a large cohort \((n = 452)\) to show that GSTP1 methylation was a significant independent prognostic factor of progression in a multivariate model adjusting for age at surgery, pre-op PSA, surgery year, surgical margins, pathological T-stage and GS; but only in samples from early, organ-confined disease \((n = 183)\) [92]. Litovkin and colleagues examined clinical failure (defined as local recurrence (cancer observed on prostatic bed) and/or metastatic relapse) as the prognostic outcome in two cohorts (Training: \(n = 147\), Validation: \(n = 71\)) of high-risk PCa patients (Clinical stage ≥ T3a, GS 8-10 and/or PSA > 20ng/ml) and found only trichotomised GSTP1 methylation (cut-offs at 15% and 50%) to be an independent prognostic predictor across both cohorts in a multivariate model adjusting for GS, pathological T-stage and pre-op PSA levels [94]. In a study of non-neoplastic tissue adjacent to the prostate tumour from PCa patients with follow-up up to 24 years \(\text{\(n = 157\)\)}\), the presence of GSTP1 methylation increased risk of PCa death by almost 3-fold, and remained an independent prognostic factor in multivariate models in combination with APC methylation, GS, age at diagnosis, year of diagnosis, source of tumour tissue, and methylation in matched tumour tissue [88]. Another study assessing PCa death, using pyrosequencing of TURP tissue \((n = 367)\), also found univariate associations with GSTP1 methylation [90].

Two small studies with BCR as the clinical endpoint observed associations with GSTP1 methylation: 1) in a univariate analysis, using sextant biopsy cores \((n = 83)\) [77] and 2) as part of a multigene signature, using RP tissue \((n = 41)\) [78]. Another small study \((n = 64)\) investigating the broader outcome of recurrence (any of BCR, local recurrence or metastatic relapse) observed associations with GSTP1 methylation at 3 CpG units in univariate analysis only [81]. Importantly, of the 15 studies assessing GSTP1 methylation as a prognostic biomarker, seven studies did not find any prognostic value in GSTP1 methylation in the prediction of BCR [74,76,79,89,93], low vs high GS cancers [84] and PCa death [77,82].

2.1.2. APC

APC is a tumour suppressor gene which encodes the adenomatous polyposis coli (APC) protein, with a known role in the cellular processes of tumourigenesis [103]. Hypermethylation of APC is observed in PCa tumours [104] and a number of studies have demonstrated its prognostic potential (Table 1). In two of these studies, APC hypermethylation was included in multivariate models alongside GSTP1 methylation to predict progression in GS 7 patients [75], and PCa death (using DNA from non-neoplastic adjacent tissue) [88]. APC methylation has also been observed to be an individual methylation marker of BCR [77,96] and PCa death [77,82]. In the study by Henrique and colleagues, methylation levels at five genes \((\text{APC, CCND2, GSTP1, RARB, RASSF1})\) were quantified by qMSP of DNA from sextant biopsies \((n = 83)\), of which APC was the only gene significantly associated with both BCR and PCa death in univariate and multivariate analyses with other clinicopathological factors. [77]. A separate study that used MSP to quantify APC, RUNX3, GSTP1 reported that APC was the only independent prognostic gene in the prediction of PCa death across two cohorts of RP patients, one before PSA-testing was widespread (1980s cohort: \(n = 216\)) and one after the introduction
of PSA-testing (1990s cohort: n = 243); adjusting for source of tumour tissue, GS and follow-up duration [82]. One study looking at BCR examined APC methylation using nested MSP in benign prostate specimens (needle biopsy or TURP) from patients who eventually developed PCa (n = 353) and found associations between APC and risk of BCR in White patients only (n = 206), adjusting for age, tumour stage, GS, PSA level and treatment type [96]. Two other studies of APC methylation levels reported only univariate associations with risk of BCR [78] and PCa death in TURP tissues [90]. It should be noted that, while these studies provide evidence for the potential prognostic value of APC methylation, a number of other studies did not observe such associations with predicting risk of BCR [74,79,89], low vs high GS [84], clinical failure [94], progression [92] and PCa death within 10 years of diagnosis [95].

2.1.3. RARB

The RARB gene encodes retinoic acid receptor beta protein, a nuclear transcriptional regulator important in cellular signalling in cell growth and differentiation processes, and often silenced and hypermethylated in PCa [105]. Five of the twelve studies examining RARB methylation have reported its potential prognostic utility, of which two studies have observed RARB methylation levels to be an independent prognostic variable in multivariate models including other clinicopathological factors [94,96]. Briefly, the qMSP study by Litovkin and colleagues examining prediction of clinical failure in high-risk patients found dichotomised RARB methylation, significant in univariate and multivariate analyses (adjusted for GS, pathological T-stage and PSA) in a small validation cohort (n = 41), but not in the training cohort (n = 71) [94]. In their more recent study, methylation of the RARB promoter region significantly increased risk of BCR in African American patients (n = 147), but only when another gene was methylated (APC, CTNND2, RASSF1 or MGMT) and no inflammation was present in the prostate specimen [96]. Other studies reported only univariate associations between methylation and prognosis with the clinical endpoints of PCa death [90] and BCR if hypermethylated alongside four or more candidate genes [78] or in RP patients with GS ≤ 7 [89]. Other studies assessing BCR [76,93], recurrence or progression [75,81,92] and PCa death [77] found no associations with RARB methylation.

2.1.4. PITX2

PITX2 encodes the paired-like homeodomain transcription factor 2, induced by the WNT pathway to activate growth-regulating genes required for cell-type specific proliferation [106]. Aberrant PITX2 methylation has been observed in multiple tumour types including breast [107] and PCa [80]. Of nine studies on PITX2 methylation in Table 1, all but one study, which assessed methylation in benign prostate specimens [96], reported significant associations with risk of progression. The first study by Weiss and colleagues, observed PITX2 methylation, quantified by qMSP, as the strongest and only independent predictor of BCR, providing additional prognostic information to existing clinicopathological factors of GS and pathological T-stage [80]. A subsequent study of n = 476 patients confirmed the association between PITX2 hypermethylation, also quantified by qMSP, and increased BCR risk in a multivariate model with GS, pathological T-stage, surgical margin, age and PSA levels [83]. Another study, using samples from the same cohort as [83], plus a smaller training cohort (Training cohort: n = 157, Testing cohort: n = 523) observed that PITX2 methylation added prognostic information to GS, pathological T-stage and surgical margins in the prediction of BCR in multivariate cox analysis. Vanaja and colleagues constructed a methylation score consisting of 11 CpG units across 5 genes (from the EpiTYPER MassARRAY platform) including sites in the PITX2 promoter region to predict recurrence within 5 years, in a model combined with GS, pre-op PSA, seminal vesicle involvement and margin status [81]. In a more recent study investigating PCa death as the clinical endpoint in a large cohort (n = 385), a prognostic model was built on six methylation biomarkers including PITX2 (alongside HSPB1, CCND2, TIG1, NPYS and MAL), profiled by pyrosequencing in TURP tissues from patients categorised with low-to-intermediate risk by Cancer of the Prostate Risk Assessment (CAPRA) scores [98]. The resulting methylation score combined with the CAPRA score improved sensitivity of prognostic test (compared to CAPRA...
alone), predicting aggressive PCa at 10 years follow-up with an AUC of 0.74 [98]. A smaller study of risk of PCa death in patients with GS ≤ 7 (n = 135, median follow-up of 15 years, TURP tissue specimens), also using pyrosequencing, reported a significant association with increased PITX2 methylation levels [91]. Two additional studies from the same lab likewise reported univariate associations between PITX2 methylation and BCR, with no multivariate analysis with clinicopathological conducted [97,99]. Additional analyses showed that combinations of both PITX2 and PITX3 methylation were associated with BCR [97] and that there was a strong correlation between PITX2 methylation and ISUP grade group in core needle biopsy specimens [99].

2.1.5. CCND2 and PTGS2

CCND2 and PTGS2 have been studied in eight and six studies respectively, with four studies reporting modest evidence of potential prognostic value for each gene. Hypermethylation of CCND2 methylation has been reported as an independent prognostic marker of clinical failure [94], and for prediction of progression [75] and PCa death [90,98] in combination with other markers. Three small studies (each n ≤ 60) observed associations between higher PTGS2 methylation and increased risk of BCR [74,76,78] with one study reporting a nine-fold increased risk when combined with CD44 methylation [76], whilst a larger two-cohort study reported associations with clinical failure in the training cohort only [94].

2.1.6. Other candidate genes

Other candidate genes that have been investigated in at least three studies include RASSF1, TIG1, EDNRB, MGMT, MDR1, CDKN2A, TIMP3, CDH1, PDLIM4, DPYS, MAL, SLIT2, SFN and HSPB1 (Table 1). Methylation at four of these genes (MGMT, CDKN2A, TIMP3, CDH1) was not found to have any prognostic utility, whilst EDNRB, MDR1 and PDLIM4 methylation was only reported to have significant univariate associations with disease risk in one study each (Figure 3). Surprisingly, RASSF1, frequently hypermethylated in PCa and part of the ConfirmMDX panel [56], has only been reported to have associations with risk of BCR in two small studies [77,93]. TIG1 was not observed to be an independent prognostic factor alone, only in combination with other genes for prediction of PCa death in low-to-intermediate-risk patients [98]. Pairwise combinations of high SLIT2, SFN and SERPINB5 methylation were able to classify high from low GS patients in random forest modelling of a small cohort (n = 48) [84]. DPYS, MAL and HSPB1 emerge as potential prognostic biomarkers, with HSPB1 in particular validating as an independent prognostic factor of PCa death in the three studies it has been investigated in [86,90,98], SERPINB5 and AIM1, the only markers found to have significant associations with progression [95] and overall survival [85] have only been investigated in two small independent studies each thus far, and require further validation in larger cohorts. And finally, new candidate genes from different biological pathways have been explored recently (e.g. PD-1, PD-L1, CD81, TFF3, ZNF660) [108-112] and also require further validation.

In summary, the evidence for the prognostic value of the most extensively studied candidate genes (e.g. GSTP1, APC, RARB, CCND2, PTGS2) is conflicting, potentially due to differences in study designs including diversity in sample type, cohort size, clinical endpoints examined, methylation profiling methodologies, analytical approach and clinicopathological factors adjusted for in multivariate analyses. Thus far, PITX2 methylation has shown the most robust evidence of providing additive prognostic information to traditional clinicopathological markers, in particular in the prediction of BCR progression.

2.2. Genome-wide prognostic biomarker discovery studies

Technological advancements in microarray and next-generation sequencing technologies over the last decade have enabled hypothesis-free, genome-wide screening for new prognostic methylation biomarkers. Platforms that have been used for epigenome-wide screens include restriction enzyme-based, capture-based, and microarray-based platforms, together with next-generation sequencing, summarised in Figure 2. In this section, we summarise genome-wide
prognostic methylation biomarker discovery studies in primary PCa tumours (Table 2). We focus on those studies that used measures of disease risk to identify potential biomarkers (for example, comparing methylation of patients of different GS, or different survival outcomes), rather than studies that compared methylation differences between benign and tumour tissue to identify disease-specific biomarkers, with assessment of their prognostic value as only a secondary step (for example, [113-119]).
<table>
<thead>
<tr>
<th>Study [ref]</th>
<th>Approach (Discovery)</th>
<th>Cohort size (Discovery)</th>
<th>Follow-up (Discovery) [years]</th>
<th>Approach (Validation)</th>
<th>Cohort size (Validation)</th>
<th>Method (Validation)</th>
<th>Follow-up (Validation) [years]</th>
<th>Final markers identified (Validation)</th>
<th>Key results (Validation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lin et al. 2013 Neoplasia [121]</td>
<td>Enhanced Reduced Representation Bisulphite Sequencing</td>
<td>n = 7 indolent, n = 6 advanced</td>
<td>Indolent (range) - 5 - 6 years</td>
<td>Indolent vs aggressive PCa</td>
<td>n = 16 indolent, n = 8 advanced</td>
<td>MassARRAY EpiTYPER</td>
<td>Indolent - 3 - 7 years</td>
<td>Panel of 13 CpG islands: KCNC2, ZDHHC1, TBX1, CAPG, RARRES2, GRASP, SAC3D1, TPM4, GSTP1, NKX2-1, FAM107A, SLC13A3, FILIP1</td>
<td>U: AUC = 0.975 (Sensitivity = 95%; Specificity = 95%); M: NA</td>
</tr>
<tr>
<td>Cottrell et al. 2007 J Urol [120]</td>
<td>Restriction-based methylation sequencing platforms</td>
<td>• Low GS (2-6 with no grade 4 or 5 patterns) vs High GS (8-10)</td>
<td>Range: 2 - 4</td>
<td>• Low GS vs High GS</td>
<td>Cohort 1: n = 304 (130 low GS, 96 high GS; 88 no BCR, 63 BCR)</td>
<td>1. Custom methylation oligonucleotide microarray</td>
<td>Range: 2 - 4</td>
<td>Low GS vs High GS: Cohort 1 - U: (1) ABDH9, (2) Chr3-EST, (3) GPR7, (4) NOTCH, (5) KBTBD6; M: not conducted</td>
<td>Low GS vs High GS: Cohort 1 - U: AUC (all Wilcoxon’s P &lt; 0.001) (1) 0.71, (2) 0.70, (3) 0.72, (4) 0.71, (5) 0.71; M: NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Early BCR (&lt;2 yrs post-RP) vs no early BCR, n = 5 early BCR,</td>
<td></td>
<td>• Early BCR vs no early BCR</td>
<td>Cohort 2: n = 233 (28 low GS, 27 high GS; 134 no BCR, 59 BCR)</td>
<td>2. MethyLight (qMSP)</td>
<td></td>
<td>Cohort 2 - U: (1) ABDH9, (2) ABDH9 (intermediate GS 6, 7 patients only) (3) Chr3-EST, (4) Chr3-EST (intermediate GS), (5) GPR7, (6) GPR7 (intermediate GS); M: not conducted</td>
<td>Early BCR vs no early BCR: Cohort 1 - U: (1) ABDH9, (2) ABDH9, (3) Chr3-EST, (4) Chr3-EST (intermediate GS), (5) GPR7, (6) GPR7 (intermediate GS); M: not conducted</td>
</tr>
</tbody>
</table>

### Table 2. Candidate (a priori) prognostic methylated tissue-based biomarker studies.

**1. Methylation sensitive arbitrarily primed PCR and methylated CpG island amplification**

<table>
<thead>
<tr>
<th>Study [ref]</th>
<th>Approach (Discovery)</th>
<th>Cohort size (Discovery)</th>
<th>Follow-up (Discovery) [years]</th>
<th>Approach (Validation)</th>
<th>Cohort size (Validation)</th>
<th>Method (Validation)</th>
<th>Follow-up (Validation) [years]</th>
<th>Final markers identified (Validation)</th>
<th>Key results (Validation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cottrell et al. 2007 J Urol [120]</td>
<td>Restriction-based methylation sequencing platforms</td>
<td>• Low GS (2-6 with no grade 4 or 5 patterns) vs High GS (8-10)</td>
<td>Range: 2 - 4</td>
<td>• Low GS vs High GS</td>
<td>Cohort 1: n = 304 (130 low GS, 96 high GS; 88 no BCR, 63 BCR)</td>
<td>1. Custom methylation oligonucleotide microarray</td>
<td>Range: 2 - 4</td>
<td>Low GS vs High GS: Cohort 1 - U: (1) ABDH9, (2) Chr3-EST, (3) GPR7, (4) NOTCH, (5) KBTBD6; M: not conducted</td>
<td>Low GS vs High GS: Cohort 1 - U: AUC (all Wilcoxon’s P &lt; 0.001) (1) 0.71, (2) 0.70, (3) 0.72, (4) 0.71, (5) 0.71; M: NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Early BCR (&lt;2 yrs post-RP) vs no early BCR, n = 5 early BCR,</td>
<td></td>
<td>• Early BCR vs no early BCR</td>
<td>Cohort 2: n = 233 (28 low GS, 27 high GS; 134 no BCR, 59 BCR)</td>
<td>2. MethyLight (qMSP)</td>
<td></td>
<td>Cohort 2 - U: (1) ABDH9, (2) ABDH9 (intermediate GS 6, 7 patients only) (3) Chr3-EST, (4) Chr3-EST (intermediate GS), (5) GPR7, (6) GPR7 (intermediate GS); M: not conducted</td>
<td>Early BCR vs no early BCR: Cohort 1 - U: (1) ABDH9, (2) ABDH9, (3) Chr3-EST, (4) Chr3-EST (intermediate GS), (5) GPR7, (6) GPR7 (intermediate GS); M: not conducted</td>
</tr>
</tbody>
</table>

2. Enhanced Reduced Representation Bisulphite Sequencing

<table>
<thead>
<tr>
<th>Study [ref]</th>
<th>Approach (Discovery)</th>
<th>Cohort size (Discovery)</th>
<th>Follow-up (Discovery) [years]</th>
<th>Approach (Validation)</th>
<th>Cohort size (Validation)</th>
<th>Method (Validation)</th>
<th>Follow-up (Validation) [years]</th>
<th>Final markers identified (Validation)</th>
<th>Key results (Validation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lin et al. 2013 Neoplasia [121]</td>
<td>Enhanced Reduced Representation Bisulphite Sequencing</td>
<td>n = 7 indolent, n = 6 advanced</td>
<td>Indolent (range) - 5 - 6 years</td>
<td>Indolent vs aggressive PCa</td>
<td>n = 16 indolent, n = 8 advanced</td>
<td>MassARRAY EpiTYPER</td>
<td>Indolent - 3 - 7 years</td>
<td>Panel of 13 CpG islands: KCNC2, ZDHHC1, TBX1, CAPG, RARRES2, GRASP, SAC3D1, TPM4, GSTP1, NKX2-1, FAM107A, SLC13A3, FILIP1</td>
<td>U: AUC = 0.975 (Sensitivity = 95%; Specificity = 95%); M: NA</td>
</tr>
</tbody>
</table>

---

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 October 2020
doi:10.20944/preprints202010.0154.v1
## Table 2. continued.

<table>
<thead>
<tr>
<th>Study [ref]</th>
<th>Approach (Discovery)</th>
<th>Cohort size (Discovery)</th>
<th>Follow-up (Discovery) [years]</th>
<th>Approach (Validation)</th>
<th>Cohort size (Validation)</th>
<th>Method (Validation)</th>
<th>Follow-up (Validation) [years]</th>
<th>Final markers identified (Validation)</th>
<th>Key results (Validation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bhasin et al. 2015 Cell Rep [122]</td>
<td>Low GS (6) vs High GS (8-10)</td>
<td>n = 6 Low GS, n = 9 High GS</td>
<td>NA</td>
<td>Low GS vs High GS</td>
<td>n = 46 Low GS, n = 203 High GS (TCGA)</td>
<td>HM450K Microarray</td>
<td>NA</td>
<td>U: 101 DMRs including at CD14, PCDHGA11, EYA1, CCDC8, HOXC4; M: not conducted</td>
<td>U: LIMMA P = 2.81 x 10^-8 - 0.05 (range); M: NA</td>
</tr>
<tr>
<td>Kron et al. 2009 PLoS ONE [123]</td>
<td>Low GS (6 (3+3)) vs High GS (8 (4+4))</td>
<td>n = 10 Low GS, n = 10 High GS</td>
<td>NA</td>
<td>1. Low GS vs High GS</td>
<td>n = 20 low GS vs n = 19 high GS (MethyLight)</td>
<td>MethyLight</td>
<td>NA</td>
<td>U: HOXD3 (detected in n = 2 GS6 vs n = 6 GS8); M: not conducted</td>
<td>Sample size too low for statistical validity; U: 17.3% difference in methylation; M: NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2. Kron et al. 2010 Lab Invest [124]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>GS≤6 vs GS7 - U: HOXD3; M: not conducted</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• GS6 vs GS7 vs High GS</td>
<td>n = 232 (n = 101 GS ≤6, n = 107 GS ≥7, n = 147 no BCR, n = 85 BCR)</td>
<td>MethyLight</td>
<td>Mean (range): 4.4 (0.2-9.5)</td>
<td>GS≤6 vs GS7 - U: HOXD3; M: not conducted</td>
<td>GS≤6 vs GS7 - U: HOXD3; M: not conducted</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• GS≤6 vs GS7 • BCR</td>
<td></td>
<td></td>
<td></td>
<td>GS≤6 vs GS7 vs High GS: U: HOXD3; M: not conducted</td>
<td>GS≤6 vs GS7 vs High GS: U: HOXD3; M: not conducted</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3. Liu et al. 2011 Int J Cancer [125]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BCR - U: HOXD3; M: HOXD3 x pathological T-stage interaction term [+ GS + pathological T-stage + margin status]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• GS≤6 vs GS7 • BCR</td>
<td>(reduced cohort from Kron et al. 2010 [124])</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4. Jeyapala et al. 2018 Eur Urol Oncol [126]</td>
<td>Cohort 1: n = 435, n = 43 BCR;</td>
<td>Cohort 1:</td>
<td>Mean (range): 1.9 (0 - 12.6)</td>
<td>Cohort 1 – U: GBX2; M: not conducted</td>
<td>Cohort 1 – U: Mann-Whitney Test cg09094860 [P = 0.003], cg00302494 [P = 0.01]; M: not conducted</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• BCR</td>
<td>HM450K Microarray (TCGA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Capture-based methylation sequencing platforms

**MBD (methyl-CpG binding domain)-isolated genome sequencing (MiGS)**

- **1. Agilent Human CpG Island Microarray**
    - Low GS (6 (3+3)) vs High GS (8 (4+4))
    - n = 10 Low GS, n = 10 High GS
    - NA
    - 1. Low GS vs High GS
    - n = 20 low GS vs n = 19 high GS (MethyLight)
    - MethyLight
    - Mean (range): 4.4 (0.2-9.5)
    - U: HOXD3 (detected in n = 2 GS6 vs n = 6 GS8); M: not conducted
    - Sample size too low for statistical validity; U: 17.3% difference in methylation; M: NA
  - Kron et al. 2010 Lab Invest [124]
    - GS≤6 vs GS7
    - BCR
    - n = 232 (n = 101 GS ≤6, n = 107 GS ≥7, n = 147 no BCR, n = 85 BCR)
    - MethyLight
    - Mean (range): 4.4 (0.2-9.5)
    - GS≤6 vs GS7 - U: HOXD3; M: not conducted
    - BCR - U: HOXD3; M: HOXD3 x pathological T-stage interaction term [+ GS + pathological T-stage + margin status]
  - Liu et al. 2011 Int J Cancer [125]
    - GS≤6 vs GS7
    - BCR
    - (reduced cohort from Kron et al. 2010 [124])
    - MethyLight
    - Mean (range): 4.4 (0.2-9.5)
    - GS≤6 vs GS7 - U: HOXD3; M: not conducted
    - BCR - U: HOXD3; M: HOXD3 x pathological T-stage interaction term [+ GS + pathological T-stage + margin status]
  - Jeyapala et al. 2018 Eur Urol Oncol [126]
    - Cohort 1: n = 435, n = 43 BCR;
    - Cohort 1: HM450K Microarray (TCGA)
    - Mean (range): 1.9 (0 - 12.6)
    - Cohort 1 – U: GBX2; M: not conducted
    - Cohort 1 – U: Mann-Whitney Test cg09094860 [P = 0.003], cg00302494 [P = 0.01]; M: not conducted

### Microarray-based platforms

- **1. Agilent Human CpG Island Microarray**
    - Low GS (6 (3+3)) vs High GS (8 (4+4))
    - n = 10 Low GS, n = 10 High GS
    - NA
    - 1. Low GS vs High GS
    - n = 20 low GS vs n = 19 high GS (MethyLight)
    - MethyLight
    - Mean (range): 4.4 (0.2-9.5)
    - GS≤6 vs GS7 - U: HOXD3; M: not conducted
    - BCR - U: HOXD3; M: HOXD3 x pathological T-stage interaction term [+ GS + pathological T-stage + margin status]
  - Kron et al. 2010 Lab Invest [124]
    - GS≤6 vs GS7
    - BCR
    - n = 232 (n = 101 GS ≤6, n = 107 GS ≥7, n = 147 no BCR, n = 85 BCR)
    - MethyLight
    - Mean (range): 4.4 (0.2-9.5)
    - GS≤6 vs GS7 - U: HOXD3; M: not conducted
    - BCR - U: HOXD3; M: HOXD3 x pathological T-stage interaction term [+ GS + pathological T-stage + margin status]
  - Liu et al. 2011 Int J Cancer [125]
    - GS≤6 vs GS7
    - BCR
    - (reduced cohort from Kron et al. 2010 [124])
    - MethyLight
    - Mean (range): 4.4 (0.2-9.5)
    - GS≤6 vs GS7 - U: HOXD3; M: not conducted
    - BCR - U: HOXD3; M: HOXD3 x pathological T-stage interaction term [+ GS + pathological T-stage + margin status]
  - Jeyapala et al. 2018 Eur Urol Oncol [126]
    - Cohort 1: n = 435, n = 43 BCR;
    - Cohort 1: HM450K Microarray (TCGA)
    - Mean (range): 1.9 (0 - 12.6)
    - Cohort 1 – U: GBX2; M: not conducted
    - Cohort 1 – U: Mann-Whitney Test cg09094860 [P = 0.003], cg00302494 [P = 0.01]; M: not conducted
<table>
<thead>
<tr>
<th>Study [ref]</th>
<th>Approach (Discovery)</th>
<th>Cohort size (Discovery)</th>
<th>Follow-up (Discovery) [years]</th>
<th>Approach (Validation)</th>
<th>Cohort size (Validation)</th>
<th>Method (Validation)</th>
<th>Follow-up (Validation) [years]</th>
<th>Final markers identified (Validation)</th>
<th>Key results (Validation)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cohort 2:</td>
<td>n = 254 (n = 202, n</td>
<td>Cohort 2: MethyLight</td>
<td>Cohort 2: 5.7 (0.1-12.3)</td>
<td></td>
<td>Cohort 2 –</td>
<td>GBX2 (and in IDC/C-negative patients only);</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>= 52 BCR, n = 58 IDC/C-positive, n = 196 IDC/C-negative)</td>
<td></td>
<td></td>
<td></td>
<td>M: (1) GBX2 [+ GS + Pathological T-stage + pre-op PSA], 2) GBX2 [+ pre-op PSA]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cohort 1:</td>
<td>n = 254 (n = 202 no</td>
<td>Median (range):</td>
<td>Cohort 1: 6.7 (0.1-12.8);</td>
<td>Cohort 2: 6.7 (0.2-18.6);</td>
<td></td>
<td>4-G model: HOXD3, TGFβ2, CRIP3, APC (candidate)</td>
<td></td>
<td>Salvage RT/hormone therapy: U: 4-G model; M: Integrative model</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BCR, n = 52 BCR, n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>= 205 no salvage RT, n = 42 salvage RT, n = 226 no hormone therapy, n = 21 hormone therapy)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cohort 2:</td>
<td>n = 199 (n = 159 no</td>
<td></td>
<td>Cohort 1 - BCR- U: 4-G model; M: Integrative model</td>
<td></td>
<td></td>
<td></td>
<td>Salvage RT/hormone therapy: U: 4-G model; M: Integrative model</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>BCR, n = 40 BCR, n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>= 180 no salvage RT, n = 19 salvage RT, n = 177 no hormone therapy, n = 22 hormone therapy)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5. Jeyapala et al. 2019 Urol Oncol [127]

- BCR
- salvage RT/hormone therapy

Cohort 1 –
BCR- U: HR 2.72 (1.77-4.17), P < 0.001, Sensitivity = 90.9%, Specificity = 35.2%, AUC = 0.740; M: HR 1.49 (1.12-1.99), P = 0.006, Sensitivity = 92.9%, Specificity = 43.4%, AUC = 0.846
Salvage RT/hormone therapy: U: HR 2.20 (1.48-3.29), P < 0.001; M: HR 1.34 (1.03-1.75), P = 0.027

Cohort 2 –
BCR- U: Mann-Whitney Test P = 0.0001, IDC/C-negative patients: Log-rank P = 0.002;
M: (1) HR 1.02 (1.006-1.034), P = 0.004, (2) C-index 0.78 (vs 0.71 for PSA alone)
Salvage RT/hormone therapy - U: HR 1.97 (1.21-3.21), P < 0.001, Sensitivity = 91.2%, Specificity = 27.4%, AUC = 0.636; M: HR 1.17 (0.79-1.72), P = 0.441, AUC = 0.731 (vs 0.723 for CAPRA-S alone)
### Table 2 continued.

| Study [ref] | Approach (Discovery) | Cohort size (Discovery) | Follow-up (Discovery) [years] | Approach (Validation) | Cohort size (Validation) | Method (Validation) | Follow-up (Validation) [years] | Final markers identified (Validation) | Key results (Validation)
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>6. Savio et al. 2019</strong></td>
<td>JMDI [128]</td>
<td>BCR; Late BCR (5 and 7 yrs post-RP); salvage RT/hormone therapy [Biopsy specimens pre-RP]</td>
<td>n = 86 (n = 61 no BCR, n = 25 BCR, n = 75 no salvage RT, n = 11 salvage RT, n = 70 no hormone therapy, n = 15 hormone therapy)</td>
<td>MethyLight</td>
<td>Median (range): 5.1 (0.1-16)</td>
<td>BCR - U: none; M: 4-G model [+ pre-op PSA] Late BCR - U: none; M: (1) 4-G model [+ pre-op PSA] (5 yrs), (2) 4-G model [+ pre-op PSA] (7 yrs) Salvage RT/hormone therapy: U: 4-G model; M: (1) 4-G model [+ pre-op PSA], (2) 4-G model [+ CAPRA]</td>
<td><strong>BCR: U: non-sig; M: Sensitivity = 78.6%, Specificity = 64.7%, AUC = 0.714</strong> Late BCR: U: non-sig; M: (1) Sensitivity = 80%, Specificity= 60.5%, AUC = 0.705 (vs 0.667 for PSA alone), (2) Sensitivity = 76.9%, Specificity = 62.9%, AUC = 0.688 (vs 0.6 for PSA alone) Salvage RT/hormone therapy - U: Sensitivity = 66.7%, Specificity = 75%, AUC = 0.699; M: (1) Sensitivity = 75%, Specificity = 61.1%, AUC = 0.699, (2) Sensitivity = 76.9%, Specificity = 58.3%, AUC = 0.797</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>2. Illumina GoldenGate Methylation Microarrays</strong></td>
<td>Goh et al. 2014 PloS ONE [129]</td>
<td>Low GS (6) vs High GS (8-10); Overall survival</td>
<td>n = 87 (n = 19 GS6, n = 48 GS8-10, n = ns death)</td>
<td>Median (range): 4 (0 - 11.8)</td>
<td>GS (6-9)</td>
<td>n = 59 (n = 23 GS 6, n = 22 GS 7, n = 13 GS 8-10, n = 18 for BCR analysis)</td>
<td>GoldenGate</td>
<td>No BCR – 5 (1 - 13)</td>
<td><strong>“PHYMA” signature: 55 probes targeting CpG loci of 46 genes, including at ALOX12, PDGFRB</strong> GS: U: PHYMA (GS 6-8) M: not conducted BCR: U: none M: not conducted</td>
</tr>
<tr>
<td>Study [ref]</td>
<td>Approach (Discovery)</td>
<td>Cohort size (Discovery)</td>
<td>Follow-up (Discovery) [years]</td>
<td>Approach (Validation)</td>
<td>Cohort size (Validation)</td>
<td>Method (Validation)</td>
<td>Follow-up (Validation) [years]</td>
<td>Final markers identified (Validation)(^{c,d,e})</td>
<td>Key results (Validation)(^f)</td>
</tr>
<tr>
<td>------------</td>
<td>----------------------</td>
<td>------------------------</td>
<td>-----------------------------</td>
<td>-----------------------</td>
<td>------------------------</td>
<td>------------------</td>
<td>-----------------------------</td>
<td>--------------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Angulo et al. 2016 Urol Int [130]</td>
<td>BCR</td>
<td>n = 26 no BCR, n = 32 BCR</td>
<td>Mean ± SD (range): 6.3 ± 3 (0.8-13.8)</td>
<td>No validation</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>Discovery only: BCR: U: (1) Gene hypermethylation profile of cluster 3 patients (GSTM2, GSTP1, RARB, ALOX12, APC, PDGFRB, SCLG3A1, CFTR, MT1A, PENK, NEU1, CCNA1, MET, KLK10, RARA, MFAP4, TERT, TXB1, TAL1, ERG, MMP14, EYA4, COL18A1, GADD45A, POMC, GSI1, SEPT9, MYCL2), (2) MT1A, (3) ALOX12, (4) GSTM2, (5) APC, (6) MYCL2, (7) RARB, (8) GSTM2 + MYCL2 M: (1) Gene hypermethylation profile [+ D’Amico classification]; (2) GSTM2 [+ ns]; (3) MYCL2 [+ ns]</td>
<td>PCa Death: U: GSTM2 + MYCL2 M: not conducted</td>
</tr>
<tr>
<td>Kobayashi et al. 2011 Genome Res [131]</td>
<td>BCR</td>
<td>n = 86</td>
<td>Range: 0 - 5.5</td>
<td>No validation</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>Discovery only: U: KCNK4, WDR86, OAS2, TMEM179 (FDR ≤ 1%; hypermethylated) M: not conducted</td>
<td>Discovery only: U: ns M:</td>
</tr>
</tbody>
</table>

### 3. **Illumina Infinium HumanMethylation 27K Microarray**
<table>
<thead>
<tr>
<th>Study [ref]</th>
<th>Approach (Discovery)</th>
<th>Cohort size (Discovery)</th>
<th>Follow-up (Discovery) [years]</th>
<th>Approach (Validation)</th>
<th>Cohort size (Validation)</th>
<th>Method (Validation)</th>
<th>Follow-up (Validation) [years]</th>
<th>Final markers identified (Validation)</th>
<th>Key results (Validation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mahapatra et al. 2012</td>
<td>Clin Cancer Res [132]</td>
<td>Indolent vs aggressive disease:</td>
<td>• No recurrence vs recurrence</td>
<td>No recurrence - 6.2 ± 1.5</td>
<td>BCR: 5.9 ± 1.4</td>
<td>Local recurrence: 4.2 ± 1.7</td>
<td>Metastatic relapse: 4.4 ± 4.0</td>
<td>Pyro sequencing</td>
<td>(Mean ± SD):</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• n = 75 no recurrence, n = 123 recurrence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• n = 43 BCR, n = 80 clinical recurrence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Indolent vs aggressive disease:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Local recurrence vs metastatic relapse</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• No recurrence vs recurrence</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• BCR vs clinical recurrence</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Local recurrence vs metastatic relapse</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>n = 20 no recurrence, n = 20 BCR, n = 20 local recurrence, n = 20 metastatic relapse</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(Mean ± SD):</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• No recurrence - 6.5 ± 1.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• BCR - 5.45 ± 1.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Local recurrence - 4.5 ± 1.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Metastatic relapse - 3.5 ± 1.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No recurrence vs recurrence:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>U: Sensitivity/Specificity, AUC, t-test (1) 65.0%/65.6%, 0.727, P = 0.0139; (2) 70.4%/75.3%, 0.788, P = 0.0018; (3) 65.0%/60.0%, 0.773, P = 0.0115; (4) 70.3%/60.4%, 0.660, P = 0.0835; (5) 75.7%/55.2%, 0.682, P = 0.0515</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BCR vs clinical recurrence:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>U: Sensitivity/Specificity, AUC, t-test (1) 65.6%/65.0%, 0.73, P = 0.0129; (2) 60.8%/75.6%, 0.692, P = 0.0373; (3) 75.4%/60.3%, 0.761, P = 0.0047</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Local vs metastatic relapse:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>U: Sensitivity/Specificity, AUC, t-test (1) 75.2%/60.0%, 0.741, P = 0.0091; (2) 65.5%/70.7%, 0.701, P = 0.0295; (3) 70.6%/75.4%, 0.748, P = 0.0071; M: NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>No recurrence vs recurrence:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>U: (1) CRIP1; (2) RUNX3; (3) HS3ST2; (4) FLNC; (5) RASGRF2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>M: Not conducted</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BCR vs clinical recurrence:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>U: Sensitivity/Specificity, AUC, t-test (1) 65.6%/65.0%, 0.73, P = 0.0129; (2) 60.8%/75.6%, 0.692, P = 0.0373; (3) 75.4%/60.3%, 0.761, P = 0.0047</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Local recurrence vs metastatic relapse:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>U: Sensitivity/Specificity, AUC, t-test (1) 75.2%/60.0%, 0.741, P = 0.0091; (2) 65.5%/70.7%, 0.701, P = 0.0295; (3) 70.6%/75.4%, 0.748, P = 0.0071; M: NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study [ref]</td>
<td>Approach (Discovery)</td>
<td>Cohort size (Discovery)</td>
<td>Follow-up (Discovery) [years]</td>
<td>Approach (Validation)</td>
<td>Cohort size (Validation)</td>
<td>Method (Validation)</td>
<td>Follow-up (Validation) [years]</td>
<td>Final markers identified (Validation)c,d,e</td>
<td>Key results (Validation)f</td>
</tr>
<tr>
<td>------------------------</td>
<td>----------------------</td>
<td>-------------------------</td>
<td>------------------------------</td>
<td>-----------------------</td>
<td>--------------------------</td>
<td>----------------------</td>
<td>-------------------------------</td>
<td>--------------------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>Geybels et al. 2016 Clin Epigenetics [133]</td>
<td>Low GS (≤ 6) vs High GS (8-10)</td>
<td>n = 65 Low GS, n = 88 High GS (TCGA)</td>
<td>NA</td>
<td>Progression</td>
<td>n = 323 no progression, n = 108 progression</td>
<td>HM450K Microarray</td>
<td>Mean (SD): 8.0 (4.2) years</td>
<td>Signature consisting of 52 CpG sites (32 unique genes) including at RRM2, VWA3B, MFSD9, ANO7, GALNT12, SEMA3F, ATXN7, SLC15A2, MME, USP17, KIAA0922, FOX11, URGCP, PTRN2, RPI, MRPS28, MKI67, CPT1A, KCNMB4, TMEM132D, ARHGEF7, CTAGES, EIF2AK4, C15orf26, PLK1, YPEL3.</td>
<td>U: (1) Signature; (2) Signature (GS7 only); (3) Signature (GS7 (3+4) only)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>M: [+ GS+ pathological T-stage + pre-op PSA] (1) Signature; (2) Signature (GS7 only); (3) Signature (GS7 (3+4) only)</td>
<td></td>
</tr>
<tr>
<td>Zhao et al. 2017 Clin Cancer Res [134]</td>
<td>Metastatic-lethal progression</td>
<td>n = 304 no progression vs n = 24 metastatic-lethal</td>
<td>Mean: •Metastatic relapse - 8.1 • Survival - 12.2</td>
<td>1. Metastatic-lethal progression</td>
<td>n = 41 no progression, n = 24 metastatic-lethal</td>
<td>HM450K Microarray</td>
<td>Mean: 9</td>
<td>U: (1) ALKBH5; (2) ATP11A; (3) FHAD1; (4) KLHL8; (5) PI15; (6) Intergenic region (chr1); (7) Intergenic (chr16); (8) Intergenic (chr17)</td>
<td>M: [+GS] (1) ALKBH5; (2) FHAD1; (3) KLHL8; (4) PI15</td>
</tr>
<tr>
<td>Study [ref]</td>
<td>Approach (Discovery)</td>
<td>Cohort size (Discovery)</td>
<td>Follow-up (Discovery) [years]</td>
<td>Approach (Validation)</td>
<td>Cohort size (Validation)</td>
<td>Method (Validation)</td>
<td>Follow-up (Validation) [years]</td>
<td>Final markers identified (Validation)</td>
<td>Key results (Validation)</td>
</tr>
<tr>
<td>------------</td>
<td>---------------------</td>
<td>------------------------</td>
<td>-----------------------------</td>
<td>---------------------</td>
<td>------------------------</td>
<td>---------------------</td>
<td>-----------------------------</td>
<td>-----------------------------------</td>
<td>-------------------------</td>
</tr>
<tr>
<td>Zhao et al. 2018</td>
<td>The Prostate [135]</td>
<td>Metastatic-lethal progression</td>
<td>Training dataset (from Zhao et al. 2016 [134]): n = 344 no recurrence, n = 48 metastatic-lethal; Testing dataset: n = 11 no recurrence, 23 metastatic-lethal</td>
<td>Pyrosequencing</td>
<td>Training (Mean (minimum)) - 8 (5)</td>
<td>Methylation score: ALKBH5 + ATP11A + FHAD1 + KLHL8 + PI15 [+ GS]</td>
<td>Training - U: not conducted; M: Methylation score Testing - U: not conducted; M: Methylation score</td>
<td>Training – U: NA; M: Logistic regression β-coefficient - ALKBH5 [-0.75], ATP11A [-0.7], FHAD1 [-0.72], KLHL8 [-0.33], PI15 [0.70], GS [1.13] Testing - U: NA; M: Mean difference = 2.49 (P = 6.8 x 10^-10), OR 4.0 (1.8-14.3), P = 0.006, AUC = 0.91 (vs 0.87 for GS alone), pAUC = 0.037 (vs 0.025 for GS alone), sensitivity at 95% specificity = 74% (vs 53% for GS alone)</td>
<td></td>
</tr>
<tr>
<td>Mundbjerg et al. 2017</td>
<td>Genome Biology [136]</td>
<td>Aggressiveness (individual PCa foci vs matched lymph node metastasis)</td>
<td>n = 14 (n = 92 samples: multiple tumour foci, adjacent normal tissue, lymph node metastases and normal lymph nodes)</td>
<td>HM450K Microarray</td>
<td>n = 351 (TCGA)</td>
<td>Mean: 3.2 years</td>
<td>Aggressiveness classifier: 25 probes, including in NXPH2, NCAPH1, TRIB1, PCDHA1-PCDHA8, C3orf37, C3orf3, CPN1, TCF7L2, ROBO1, GFPT2, FBXQ7, SKI, HDAC9, CARS, SLC6A17, BCA1, GAS1, RAI1</td>
<td>U: Specificity = 97.4%, Sensitivity = 96.2%, Negative predictive value = 76%, Positive predictive value = 99.7%, Lymph node metastases [Fishers exact test P = 9.2 x 10^-10], pathological stage T3 [Fishers exact test P = 2.2 x 10^-7] M: NA</td>
<td></td>
</tr>
</tbody>
</table>
Table 2 continued.

<table>
<thead>
<tr>
<th>Study [ref]a</th>
<th>Approach (Discovery)</th>
<th>Cohort size (Discovery)</th>
<th>Follow-up (Discovery) [years]</th>
<th>Approach (Validation)</th>
<th>Cohort size (Validation)</th>
<th>Method (Validation)</th>
<th>Follow-up (Validation) [years]</th>
<th>Final markers identified (Validation)b,c,d,e</th>
<th>Key results (Validation)f</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toth et al. 2019 Clinical Epigenetics [137]</td>
<td>Good prognosis vs Poor prognosis</td>
<td>n = 35 good prognosis, n = 35 poor prognosis (80% training set, 20% testing set)</td>
<td>Range: 3 - 5</td>
<td>• Good prognosis vs poor prognosis (Cohort 1, 2) • BCR (Cohort 1, 2, 3)</td>
<td>Cohort 1: n = 222 (n = 63 for prognosis analysis, n = ns BCR) (ICGC cohort [138]) Cohort 2: n = 477 (n = 27 good prognosis, 57 poor prognosis, n = ns BCR) (TCGA) Cohort 3: n = 12,581 (n = 3612 BCR) (for ZIC2 immunostaining analysis only)</td>
<td>HM450K Microarray</td>
<td>Cohort 1 &amp; 2 - At least 5 years Cohort 3 (Median (range)) = 4 (0.08 - 20.08)</td>
<td>Signature consisting of 598 CpG sites. Top 20: CCT8L2, NOP56, FCRL1, OR5W2, ZFP36L2, PRMT8, SLC1A6, DOK5, CCT8L2, ZFP36L2, MMP16, ESRI, ZIC2, GPR137B, NANOS1, LCE3A, C11orf87, PEG3, ZIM2, CTSC, CD84, MOS, RND2</td>
<td>Good vs poor prognosis: Cohort 1 - U: AUC = 0.997; M: NA Cohort 2 - U: AUC = 0.775; M: NA BCR: Cohort 1 - U: Log-rank P &lt; 0.0001, M: NA Cohort 2 - U: Log-rank P &lt; 0.0001; M: Cox regression P = 0.011 Cohort 3 - U: Log-rank P &lt; 0.0001; M: ns Good vs good prognosis: Cohort 1 - U: (1) Signature, M: not conducted Cohort 2 - U: Signature; M: not conducted BCR: Cohort 1 - U: Signature, M: not conducted Cohort 2 - U: Signature; M: Signature [+ GS] Cohort 3 - U: ZIC2 protein; M: ZIC2 protein [+ GS + pathological T-stage + nodal stage + PSA]</td>
</tr>
</tbody>
</table>

**Abbreviations:** AUC = area under the curve; av. = average, CAPRA-S = Cancer of the Prostate Risk Assessment Score; CRPC = Castration-Resistant Prostate Cancer; GS = Gleason score; HR = hazard ratio; IDC/C = Intraductal Carcinoma and Cribriform Architecture; M = multivariate analysis; NA = not applicable; non-sig = non-significant; ns = not specified; OR = odds ratio; pAUC = partial area under the cure; PCa = prostate cancer; PMR = percent methylated ratio; PSA = prostate-specific antigen; RT = Radiotherapy; U = univariate analysis.

**Definitions:** BCR: Biochemical recurrence: PSA elevations ≥ 0.2ng/ml post-RP, except [130] > 0.4ng/ml and [131] > 0.07ng/ml; Clinical recurrence = local recurrence or metastatic relapse; Good prognosis: organ-confined disease (pT2) and lack of BCR for at least 5 years; Local recurrence: cancer observed on prostatic bed, confirmed by histological analysis of biopsies; Metastatic relapse: metastatic deposits (visceral, bony metastasis) confirmed by positive biopsies or cT/bone scans; Metastatic-lethal progression = metastatic relapse or PCa death; Pathological T-stage: Tumour staging based on pathological examination of surgically removed prostate tissue; PCa death: Prostate cancer-specific death; Poor prognosis: systemic presence of metastatic disease, indicated by recurrence within 3 years and no response to local radiation therapy; Progression: either of BCR, metastatic relapse or PCa death; Recurrence: either of BCR, local recurrence or metastatic relapse. a All studies are on prostate cancer tissues from radical prostatectomy, unless specified. b MethylLight is a quantitative methylation-specific PCR (qMSP) platform. c Univariate (U) or Multivariate (M) analyses. d plus (+) sign indicates variables in the same multivariate model or methylation score together. e square bracket ([]) indicate the clinicopathological factors adjusted for in each multivariate model. f number in brackets following HR or OR indicate the 95% confidence interval.
2.2.1. Restriction-based methylation sequencing studies

The first genome-wide prognostic methylation biomarker discovery study in 2007 used methylation sensitive arbitrarily primed PCR [139] and methylation CpG island amplification [140], to find markers that could distinguish between patients with high GS (8 - 10, n = 5) vs those with low GS (2 - 6 with no grade 4 or 5 patterns; n = 5), as well as markers that could predict early BCR following RP (n = 5 no BCR (> 4 yrs), n = 5 early BCR (< 2 yrs)) [120]. In this approach, methylation-(in)sensitive restriction enzymes were used to digest DNA, and the resulting fragments were screened for methylation differences between patient groups, and sequenced if differences were found. The top 51 markers, along with 11 candidate markers, were validated in two large, independent cohorts, using a custom methylation oligonucleotide microarray (Cohort 1, n = 304) and MethyLight qMSP assays (Cohort 2, n = 233) [120]. GRP7, ABHD9 and Chr3-EST were significantly hypermethylated in high GS patients and could distinguish between no BCR and early BCR, independent of patient GS in the Cohort 1 [120]. These associations were validated in Cohort 2 where increased methylation of ABHD9 and Chr3-EST correlated with high grade disease and early BCR, even after adjusting for GS, pathological T-stage and margin status [120]. A subsequent study reported only univariate associations between ABHD9 methylation and BCR in a larger cohort (n = 605) [80], whilst another saw no ABHD9 methylation difference with BCR status (n = 407) [141].

A second restriction enzyme-based method used in genome-wide screening is Enhanced Reduced Representation Bisulphite Sequencing (ERRBS) platform. Similarly, ERRBS involves enzymatic digestion of DNA at CpG sites using MspI, followed by size selection and bisulphite sequencing, with the main advantage of enabling single base pair resolution profiling of CpG sites in GC-rich genomic regions such as promoter CpG islands [142,143]. Only one study has used this platform, profiling the methylome of a small discovery set of PCa patients, comparing indolent (localised disease with no recurrence; n = 7) vs advanced cancers (aggressive castration-resistant PCa; n = 6) [121]. A series of differentially methylated CpG islands were identified using linear model analysis, and a prognostic panel of 13 hypermethylated CpG islands (see Table 2) was built using random forest classification. This panel successfully discriminated between indolent and advanced cancers in the validation cohort (n =16 indolent, n = 8 advanced, MassARRAY EpiTYPER) with an AUC of 0.975 in 10-fold cross validation [121]. Of the genes included in the panel, GSTP1 has been widely studied as a diagnostic and prognostic biomarker of PCa (see above) [44,45,144], whilst GRASP and TPM4 have been previously shown to be differentially methylated in PCa compared to normal prostate tissue [132].

2.2.2. Capture-based methylation sequencing studies

Capture-based DNA methylation sequencing for PCa prognostic studies were first used in 2015. The approach involves the capture of methylated sequences by Methyl-CpG binding domain (MBD) protein after shearing of genomic DNA, followed by massive parallel sequencing of enriched sequences [145,146]. A limitation of this approach is that it does not provide single nucleotide resolution, instead identifying regions with multiple methylated CpGs. Only one study has used this method to profile and identify methylation differences between low and high GS tumours (n = 6 vs 9) [122]. They reported hypermethylation of 4,932 regions, in high grade disease [122]. Extensive genomic and functional characterisation of these Differentially Methylated Regions (DMRs) were conducted, including comparison with publicly available data from The Cancer Genome Atlas (TCGA) project [147]. This allowed them to validate the association between methylation and high grade disease at 101 DMRs, and show that a subset of these DMRs correlated with gene expression changes that associated with poorer survival in PCa patients, including the CCDC8 and HOXD4 gene [122].
2.2.3. DNA methylation microarray studies

Microarrays have become the most popular technology for genome-wide DNA methylation profiling for biomarker discovery. For this method DNA is first treated to enable later distinction between methylated and unmethylated sites (using methylated DNA immunoprecipitation, methylation-specific restriction enzymes or bisulphite-conversion). The DNA is then hybridised to unique oligonucleotide probes of CpGs on arrays, labelled with a fluorescent dye and imaged, and the signal is used to determine single nucleotide resolution CpG methylation. Several DNA methylation microarrays have been produced, including the Agilent Human CpG Island Microarray [148] (237,220 CpGs) (Figure 2) and Illumina DNA methylation microarrays, with successive array updates interrogating a broader range and number of CpG sites across the genome: GoldenGate Cancer Panel I (1,505 CpGs) [149], Infinium HumanMethylation 27K Microarray (HM27K: 27,578 CpGs) [150] and Infinium HumanMethylation 450K Microarray (HM450K: 485,577 CpGs) [151] (Figure 2).

2.2.3.1. Agilent Human CpG Island Microarray

An initial genome-wide study in 2009 using the Agilent Human CpG Island Microarray [123] laid the foundation for a number of subsequent validation studies [124-128], leading to the identification of several robust methylation markers with prognostic value. The original study screened for methylation differences between patients with GS 6 (n = 10) and GS 8 (n = 10), finding 493 CpG sites (223 individual genes) that could distinguish between the patient groups. A candidate marker, HOXD3, was selected and assessed in an independent set of samples (n = 20 GS6 vs n = 19 GS8), validating HOXD3 hypermethylation in GS8 compared to GS6 patients. Further studies investigated HOXD3 [124] and GBX2 [126] as individual markers of progression, using BCR as their outcome of interest. HOXD3 was associated with BCR in univariate analysis (n = 147 no BCR, n = 85 BCR), and a methylation score combining HOXD3 with pathological T-stage was found to be an independent predictor of BCR [124]. GBX2 methylation was assessed in methylation data from TCGA (n = 435 no BCR, n = 43 BCR), and in a second cohort (n = 202 no BCR, n = 52 BCR). In both cohorts, associations were observed between GBX2 methylation and BCR, with GBX2 methylation shown to have potential as an additive predictor when combined with PSA levels at diagnosis (Cohort 2) [126].

The same research team also investigated HOXD3 in combination with other markers [125,127]. One study combined HOXD3, TGFβ2 (another differentially methylated gene from the original discovery study) and APC, an a priori candidate marker [125]. Using a cohort from an earlier study [124] (n = 219), they found that this multigene panel improved prediction of BCR over any individual markers, and was independent of existing clinicopathological variables [125]. In another study they used a penalized cox regression method to develop a 4-gene (4-G) prognostic model for BCR, consisting of APC, HOXD3, TGFβ2 and CRIP3 [127]. The 4-G model associated with BCR as well as progression to post-surgical therapies (hormone and salvage radiotherapy) in two large cohorts (Cohort 1: n = 202 no BCR, n = 52 BCR; Cohort 2: n = 159, n = 40 BCR) [127]. Most recently, the prognostic ability of the 4-G model was investigated in pre-surgery prostate biopsy specimens (n = 61 no BCR, n = 25 BCR) [128]. The 4-G methylation panel was able to prognosticate BCR, late recurrence (BCR 5 & 7 yrs post-RP) and eventual progression to post-surgery treatments [127]. Additionally, a study from another team found strong evidence of HOXD3 hypermethylation in BCR progression (n = 303 no BCR, n = 104 BCR) [141]. These studies provide strong support for HOXD3 hypermethylation as a robust marker for BCR progression. Further studies examining the prognostic utility of HOXD3 and other genes in the panel are warranted, particularly in cohorts with more clinically relevant endpoints for aggressive disease, such as metastatic relapse and PCa death.

2.2.3.2. GoldenGate Cancer Panel I Microarray platform

Between 2014 - 2016 two research groups published studies using Illumina’s GoldenGate Cancer Panel I Microarray platform to identify novel biomarkers for disease risk. In the first study, Support Vector Machine was used to build a classification model, generating a signature consisting of 55
probes across 46 genes (including ALOX12, PDGFRB) [129]. The signature, termed “PHYMA”, was able to distinguish between low and high GS tumours, and a high PHYMA score associated with poorer survival outcome (adjusted for clinical T-stage and GS) (n = 87). Trending associations between PHYMA score and GS, but not BCR, was observed in a separate cohort (n = 59) [129]. The second study using the GoldenGate Cancer Panel reported that a gene hypermethylation profile based on hierarchical clustering of patients (see Table 2 for details), as well as hypermethylation at individual markers GSTM2 and MCLY2, independently predicted BCR risk. The concurrent methylation of the two markers was also associated with PCa death, however no further validation of the study findings in a separate independent cohort was conducted. [130].

2.2.3.3. HM27K platform

Two studies published in 2011 - 2012 used the HM27K Microarray platform for prognostic discovery biomarker studies of PCa. One study (n = 86) reported increased methylation at 4 CpGs (KCNK4, WDR86, OAS2, TMEM179) in patients with a shorter time to BCR, although no further validation of these markers was conducted [131]. Another study performed a number of binary comparisons of methylation levels: no recurrence (n = 75) vs recurrence (BCR or clinical recurrence (local recurrence or metastatic relapse), n = 123); BCR (n = 43) vs clinical recurrence (n = 80); and local (n = 44) vs metastatic relapse (n = 36). The discovery analysis found 75, 16 and 68 genes significantly methylated in each analysis respectively. Several markers from each group of analyses were assessed by pyrosequencing in 80 patients (n = 20 per clinical endpoint), validating many of the nominated candidate genes (see Table 2 for details), including RUNX3, a candidate gene previously studied in a priori prognostic biomarker studies [79,82].

2.2.2.4. HM450K platform

The HM450K extended the HM27K probe design to provide coverage of a more diverse set of genomic categories and regions [150,151]. The platform has been widely used for prognostic biomarker discovery (since 2016) and in the generation of publicly available data, including the TCGA dataset of nearly 500 PCa methylomes, which is commonly used by researchers for biomarker discovery or validation [147]. A study by Geybels and colleagues used TCGA HM450K methylation data to identify methylation differences between PCa patients with low GS (≤ 6, n = 65) vs high GS (8 - 10, n = 88) [133]. The elastic net method was used to build a signature consisting of 52 CpG sites across 32 genes, many of which were novel prognostic candidates. The signature was then tested in HM450K data from a larger validation cohort (n = 523) for its ability to predict disease progression (any of BCR, metastases and/or PCa death) and was found to be an independent predictor of progression in multivariate analysis including GS, pathological T-stage, diagnostic PSA level, in all patients and in a subset of GS 7 patients [134]. Another study from the same lab used the HM450K data from 430 primary PCa tissues, to identify 42 DNA methylation biomarkers that could predict the more serious endpoint of metastatic-lethal progression [134]. Eight of these CpG sites validated in a small validation cohort (n = 65, HM450K), with methylation at four of these sites observed to complement GS in discriminating between non-recurrent and metastatic-lethal patients [134]. A follow-up study by the same research group used pyrosequencing to technically validate five of the eight differentially methylated CpG sites (in ALKBH5, ATP11A, FHAD1, KLHL8, PI15) [135]. They then used a training cohort (n = 392) to build a model, based on the five sites, from which they calculated a prognostic methylation score for prediction of metastatic-lethal progression. In a multivariate model with GS, a four-fold increase in risk of metastatic-lethal progression was reported in the testing cohort (n = 34), for each unit increase in the methylation score, and the methylation score outperformed prediction by GS alone [135].

A novel approach by Mundbjerg and colleagues used the HM450K to profile multiple samples per patient (different tumour foci, adjacent normal tissue, lymph node metastases and normal lymph nodes) from a cohort of patients who had undergone RP for multifocal disease (n=14 patients, n= 92 samples). They then used a GLMnet algorithm to categorise the aggressiveness of individual PCa foci based on how well they matched the methylation profile of the lymph node metastasis. The resulting
aggressiveness classifier consisted of 25 CpG sites (including in \textit{NXPH2}, \textit{TRIB1} and \textit{PCDHA1-PCDHA8}), and was successfully validated in the TCGA cohort (n = 351) through accurate prediction of lymph node metastases and invasive pathological stage T3 tumours [136]. Finally, the most recent HM450K study which aligns with our criteria of prognostic discovery, used random-forest based modelling to identify markers that could differentiate between good prognosis, defined as organ-confined disease (pT2) and no BCR for at least 5 years (n = 35), and poor prognosis, defined as systemic metastatic disease with recurrence within 3 years (n = 35) [137]. A DNA methylation-based classifier consisting of 598 sites was developed, validating in two independent cohorts of patients with publicly available methylation data, based on the same selection criteria (ICGC cohort n = 63, TCGA cohort: n = 84) [137]. Further analyses highlighted the independent prognostic value of a gene overlapping one of the 598 sites, with immunostaining analysis reporting significant association between loss of ZIC2 protein expression and poorer prognosis (adjusted for GS, pathological T- stage, nodal stage and PSA) [137].

In summary, technological advances now mean that many 100,000s of CpG sites can be profiled simultaneously, which has provided a more complete view of the complexity and heterogeneity of the PCa methylome. This has enabled the discovery of more accurate and novel biomarkers for PCa prognosis, that aid or outperform existing clinicopathological factors. These range from individual markers (e.g. \textit{ABHD9}, \textit{HOXD3}, \textit{GBX2}, \textit{RASGRF2}) to methylation signatures (e.g. 4-G model, PHYMA). However, with an average follow-up of just ~5 years across the studies, most focus on short-term clinical endpoints such as BCR. To discover and validate novel DNA methylation biomarkers for the most important clinical endpoints of metastasis and PCa specific mortality, further research needs to be conducted on large independent cohorts with extensive long-term follow-up data (≥ 15 years) [16]. Furthermore, the genome-wide methods described above are still limited in the number of CpGs assessed (Figure 2), and have a strong bias towards targeting methylation in CpG-rich regions of gene promoter and CpG islands [145]. More recent techniques, such as Illumina’s EPIC microarray, cover more distal regulatory genomic regions [151], and the ‘gold standard’ Whole Genome Bisulphite Sequencing (WGBS) technique can profile all ~28 million CpGs in the methylome [152] (Figure 2). An expanded search of the methylome will enable comprehensive discovery of novel biomarkers for PCa prognosis.

3. Non-invasive detection of prognostic DNA methylation markers in liquid biopsies

There is widespread interest in using ‘liquid biopsies’ as a minimally invasive means to improve the accuracy and safety of cancer diagnosis, risk-stratification and disease monitoring. Liquid biopsies include bodily fluids, such as blood, urine, saliva and cerebrospinal fluid, which can be sampled for the presence of circulating tumour cells, cell-free (cf)DNA (released from tumour cells by apoptosis, necrosis and active secretion) and tumour-secreted exosomes containing RNAs, DNAs and proteins. A liquid biopsy offers the opportunity to gain a more comprehensive profile of the heterogeneous molecular landscape of the tumour at diagnosis and during tumour evolution over the course of the disease and treatment. This is particularly relevant in PCa, as the majority of patients have multifocal disease, meaning that the information from a single tissue biopsy may not reflect the dynamics of all tumour foci in the prostate, which can have variable aggressiveness and progression [136].

DNA methylation biomarkers are particularly pertinent in the liquid biopsy setting. In contrast to the limited number of recurrent genetic mutations in cancer, aberrant DNA methylation events tend to be tissue and cancer-type specific and occur across larger genomic regions, allowing DNA methylation to be easily targeted for measurement [153]. The recent development of new technologies has greatly contributed to the ability to sensitively measure DNA methylation [55]. This is highly relevant in liquid biopsy samples where tumour DNA may be present at very low concentrations i.e. < 0.01% of the total DNA content [154,155]. For example, \textit{GSTP1} hypermethylation, one of the most common epigenetic events in PCa tumour specimens, has been readily detected in liquid biopsy samples from PCa patients, such as urine, semen, blood serum and plasma samples [156].
3.1. Urine-based methylated biomarkers

Early studies showed the detection of GSTP1 methylation in urine from patients with PCa; however diagnostic sensitivity was poor (less than 30%) [157]. Expanding the panel to a three-gene signature (GSTP1, APC, and RARB) improved sensitivity to 60% [158]. More recently a six-gene methylation panel has been developed, termed ‘Epigenetic Cancer of the Prostate Test in Urine’ (epiCaPture), which targets GSTP1, SFRP2, IGFBP3, IGFBP7, APC and PTGS2 [159]. epiCaPture was applied to urine samples of men with PCa and showed significant associations between DNA methylation and disease aggressiveness, with AUC of 0.64, 0.86 and 0.83 for detecting PCa, high-grade PCa, and high-risk PCa, respectively. Overall, the study concluded that epiCaPture can accurately determine risk compared to two widely used risk stratification systems, D’Amico [26] and CAPRA [160]. In another study, a two-gene methylation panel (HOXD3 and GSTP1) was developed called Prostate Cancer Urinary Epigenetic (ProCuR) [161]. When applied to urine samples, the positive predictive value of this panel was 59.4-78%, higher than PSA (38.2-72.1%), for all risk category comparisons. In addition, Moreira-Barbosa and colleagues assessed methylation of two different gene panels comprising (miR-193b/miR-34b/c) and (APC, GSTP1, RARB) in tissue and urine; they showed that a combination of methylation measurements from the two panels in urine independently predicted shorter disease-specific survival [162]. Hypermethylation of the RASSF1 promoter has also been reported for its prognostic value as a urine-based methylated biomarker [93]. In this study a multivariate model of RASSF1 methylation together with pathological T-stage was the most significant predictor of BCR in patients (GS 6), in both tissue and urine samples [93]. Overall, these studies highlight the potential of DNA methylation as a urine-based prognostic biomarker in PCa.

3.2. Blood-based methylated prognostic biomarkers for cfDNA testing

In cancer patients a proportion of circulating cfDNA is derived from tumour cells, i.e. circulating tumour DNA (ctDNA). cfDNA can be isolated from blood plasma or serum and is present at very low concentrations, ranging from ~0-50 ng/ml in healthy individuals; in cancer patients, the proportion of ctDNA can vary between 0.01% to more that 90% of the cfDNA.

A number of studies have shown that methylated GSTP1 in circulating cfDNA has prognostic value [163-165]; for example, Mahon and colleagues showed that GSTP1 methylation in cfDNA was associated with overall survival and response to chemotherapy in men with advanced PCa [164]. Importantly, this study demonstrated that GSTP1 methylation levels prior to and after one chemotherapy cycle were stronger predictors of overall survival than changes in PSA levels at three months post-chemotherapy. More recently, Hendriks and colleagues reported hypermethylation of GSTP1 and APC in plasma cfDNA, together with the concentration of cfDNA, to be statistically significant as a prognostic biomarker for overall survival in castration-resistant PCa [166]. Further cfDNA studies report the prognostic utility of GSTP1 methylation in combination with other frequently methylated genes; for example, GSTP1 and RASSF2A methylation [59] and GSTP1, RASSF1 and RARB methylation [167].

Other genes have shown promise as prognostic methylation biomarkers in cfDNA in PCa. Horning and colleagues showed that promoter hypermethylation of SRD5A2 and CYP11A1 was associated with BCR and poorer prognosis [168]. In another study, cfDNA methylation of the APC, GSTP1, RASSF1, MDRI and PTGS2 genes was associated with overall survival time in men with advanced PCa [60]. Additionally, MSP on a cohort of n = 117 patient serum samples showed that PCDH8 methylation was an independent predictive risk factor for BCR-free survival (p< 0.007) in low GS (< 7) PCa patients after surgery [169]. Overall, these studies highlight the potential value of DNA methylation biomarkers in cfDNA as prognostic indicators of relapse.

4. Conclusions and Future Directions

Research over the last two decades has shown the potential of DNA methylation as a biomarker for PCa prognosis. DNA methylation biomarker discovery has accelerated rapidly with the
emergence of affordable, scalable, whole-genome profiling techniques. However, the ongoing technological advancements are bringing new analytical challenges, such as establishing the best way to aggregate methylation data across genomic regions, control for multiple tests; combine methylation with other ‘omic’ data types, and select and prioritise the most important prognostic features from which to build predictive models [170]. A number of new analytical approaches have been made to address these issues (for example, [171-175]) Given the ever-increasing sophistication of technologies, and thus growing number of high-dimensional datasets, bioinformatic method development will continue to be a high-priority research area.

Even with the most sophisticated laboratory and bioinformatic tools for biomarker discovery, the ultimate test of whether a prospective methylation biomarker is prognostic is through validation in multiple, appropriately sized, independent cohorts. One of the obstructions to validation, and therefore translation of new DNA methylation biomarkers to the clinic, is the dearth of suitable publicly available methylation datasets with adequate clinical follow-up data. Indeed, the flagship TCGA PCa methylation dataset has only short-term follow-up clinical data, and so cannot be used to fully assess the prognostic value of putative methylation biomarkers. For existing methylation datasets such as this, their utility for prognostic research would be increased through the continued collection of long-term follow-up data.

A notable problem is that studies frequently use the same few public cohorts for discovery and/or validation, which may be leading to biased results across the field. To advance the field, new PCa methylation public datasets need to be generated. An emphasis should be placed on using the very latest laboratory techniques which allow for full genome screening, such as the comprehensive EPIC microarray or WGBS, which will increase the likelihood of identifying novel biomarker regions. Many of the highest impact journals now have an open data policy. Going forward, this open data ethos should be adopted by more researchers and publications, as it not only provides new resources for other researchers to use in their validation efforts, but also allows transparency in the research method, which ultimately improves the quality of research in the field overall.

Finally, we have discussed the advances in non-invasive DNA methylation prognostic biomarker research. Looking ahead, clinical translation of this research will be a priority as liquid biopsies offer a number of advantages over tissue-based methods, such as reducing side-effects like infection and surgical complications [17,18], and allowing serial collection of samples during the course of monitoring or therapy to provide opportunities for timely therapeutic interventions [176]. This should be paired with the utilisation of DNA methylation assays suitable for application in clinical settings (for example, Multiplex Bisulfite PCR Sequencing [177,178]) which are cost-effective, scalable, reproducible, and capable of sensitively detecting methylated tumour DNA in limited clinical material such as liquid biopsies. In conclusion, DNA methylation shows great potential as a prognostic biomarker and could thus transform the clinical management of PCa patients.

Author Contributions: DL: SJC, RP, CS all contributed to the writing of the manuscript.

Funding: This work is supported by the National Health and Medical Research Council (NHMRC) project grants (1128912 – SJC, CS; 1106870 - SJC, RP) and NHMRC fellowships (1063559 & 1156408 – SJC); Cancer Council NSW (RG-18-09 - RP, SJC), Cancer Institute of New South Wales Fellowship (CINSW) to RP (14/ECF/1-23), CINSW Translational Program Grant # TPG172146 (SJC). The contents of the published material are solely the responsibility of the administering institution and individual authors and do not reflect the views of the NHMRC.

Conflicts of Interest: The authors declare no conflict of interest.

References


cancer in urine sediments and miR-34b/c or miR-129-2 promoter methylation define subsets of clinically

Jarmalaite, S. Clinical significance of miRNA host gene promoter methylation in prostate cancer. *Hum Mol

S.; Borre, M.; Pedersen, J.S., et al. RHCG and TCAF1 promoter hypermethylation predicts biochemical
recurrence in prostate cancer patients treated by radical prostatectomy. *Oncotarget 2017*, 8, 5774-5788,

Mouritzen, P.; Ulhøi, B.; Orntoft, T., et al. Aberrant DOCK2, GRASP, HIF3A and PKF2 Hypermethylation

118. Li, X.; Zhang, W.; Song, J.; Zhang, X.; Ran, L.; He, Y. SLCO4C1 promoter methylation is a potential
biomarker for prognosis associated with biochemical recurrence-free survival after radical prostatectomy.


Haefliger, C.; Weiss, G. Discovery and validation of 3 novel DNA methylation markers of prostate cancer

H.; Rubin, M.A.; Elemento, O. Epigenomic alterations in localized and advanced prostate cancer. *Neoplasia

Methylyome-wide Sequencing Detects DNA Hypermethylation Distinguishing Indolent from Aggressive

Fleshner, N.; van der Kwast, T., et al. Discovery of novel hypermethylated genes in prostate cancer using

124. Kron, K.J.; Liu, L.; Pethe, V.V.; Demetrashvili, N.; Nesbitt, M.E.; Trachtenberg, J.; Ozcelik, H.; Fleshner,
N.E.; Briollais, L.; van der Kwast, T.H., et al. DNA methylation of HOXD3 as a marker of prostate cancer

125. Liu, L.; Kron, K.J.; Pethe, V.V.; Demetrashvili, N.; Nesbitt, M.E.; Trachtenberg, J.; Ozcelik, H.; Fleshner,
N.E.; Briollais, L.; van der Kwast, T.H., et al. Association of tissue promoter methylation levels of APC,
TGFbeta2, HOXD3 and RASSF1A with prostate cancer progression. *Int J Cancer 2011*, 129, 2454-2462,

126. Jeyapala, R.; Savio, A.J.; Olkhov-Mitsel, E.; Kamdar, S.; Zhao, F.; Cuizon, C.; Liu, R.S.; Zlotta, A.; Fleshner,
N.; van der Kwast, T., et al. GBX2 Methylation Is a Novel Prognostic Biomarker and Improves Prediction
of Biochemical Recurrence Among Patients with Prostate Cancer Negative for Intraductal Carcinoma and

N.; van der Kwast, T., et al. An integrative DNA methylation model for improved prognostication of
postsurgery recurrence and therapy in prostate cancer patients. *Urol Oncol 2020*, 38, 39 e31-39 e39,

N.E.; van der Kwast, T., et al. Methylation Markers in Prostate Biopsies Are Prognosticators for Late
Biochemical Recurrence and Therapy after Surgery in Prostate Cancer Patients. *J Mol Diagn 2020*, 22, 30-39,

Huang, H.H., et al. Diagnostic and prognostic utility of a DNA hypermethylated gene signature in prostate
130. Angulo, J.C.; Lopez, J.I.; Dorado, J.F.; Sanchez-Chapado, M.; Colas, B.; Ropero, S. A DNA
Hypermethylation Profile Independently Predicts Biochemical Recurrence Following Radical

R.M.; Sherlock, G. DNA methylation profiling reveals novel biomarkers and important roles for DNA

Global methylation profiling for risk prediction of prostate cancer. *Clin Cancer Res* 2012, 18, 2882-2895,

Nelson, P.S., et al. Epigenetic signature of Gleason score and prostate cancer recurrence after radical

134. Zhao, S.; Geybels, M.S.; Leonardson, A.; Rubicz, R.; Kolb, S.; Yan, Q.; Klotzle, B.; Bibikova, M.; Hurtado-
Biomarkers of Metastatic-Lethal Progression in Men Diagnosed with Clinically Localized Prostate Cancer.

with radical prostatectomy for prostate cancer. *Prostate* 2018, 10.1002/pros.23667,

M.; Siegmund, K.D.; Ukimura, O.; Aron, M., et al. Identifying aggressive prostate cancer foci using a DNA

C.; Minner, S.; Schlomm, T., et al. Random forest-based modelling to detect biomarkers for prostate cancer

Molecular Risk Markers and Clinical Trajectories. *Cancer Cell* 2018, 34, 996-1011, e1018,

139. Liang, G.; Gonzalez, M.L.; Salem, C.; Jones, P.A. Identification of DNA methylation differences during
tumorigenesis by methylation-sensitive arbitrarily primed polymerase chain reaction. *Methods* 2002, 27,

140. Toyot, M.; Issa, J.P. Methylated CpG island amplification for methylation analysis and cloning

141. Stott-Miller, M.; Zhao, S.; Wright, J.L.; Kolb, S.; Bibikova, M.; Klotzle, B.; Ostrander, E.A.; Fan, J.B.; Feng,
Z.; Stanford, J.L. Validation study of genes with hypermethylated promoter regions associated with
prostate cancer recurrence. *Cancer epidemiology, biomarkers & prevention : a publication of the American
Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* 2014,

profoundly divergent epigenetic landscapes in acute myeloid leukemia. *PLoS Genet* 2012, 8, e1002781,
doi:10.1371/journal.pgen.1002781.

Figueroa, M.E.; Melnick, A.M. Enhanced reduced representation bisulfite sequencing for assessment of

144. Hopkins, T.G.; Burns, P.A.; Routledge, M.N. DNA methylation of GSTP1 as biomarker in diagnosis of


146. Serre, D.; Lee, B.H.; Ting, A.H. MBD-isolated Genome Sequencing provides a high-throughput and
comprehensive survey of DNA methylation in the human genome. *Nucleic Acids Res* 2010, 38, 391-399,


