

Article

Not peer-reviewed version

---

# Alterations in the Epigenetic Machinery Associated With Prostate Cancer Health Disparities

---

[Jenna Craddock](#) , [Jue Jiang](#) , [SEAN PATRICK](#) , Shingai Mutambirwa , [Phillip D Stricker](#) , [Riana Bornman](#) , [Weerachai Jaratlerdsiri](#) , [Vanessa M Hayes](#) \*

Posted Date: 26 May 2023

doi: 10.20944/preprints202305.1875.v1

Keywords: prostate cancer; somatic alteration; epigenomics; epigenetic machinery; African ancestry; southern Africa; health disparity



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

# Alterations in the Epigenetic Machinery Associated with Prostate Cancer Health Disparities

Jenna Craddock <sup>1</sup>, Jue Jiang <sup>2</sup>, Sean M. Patrick <sup>1</sup>, Shingai B.A. Mutambirwa <sup>3</sup>, Phillip D. Stricker <sup>4</sup>, M.S. Riana Bornman <sup>1</sup>, Weerachai Jaratlerdsiri <sup>2</sup> and Vanessa M. Hayes <sup>1,2,5,\*</sup>

<sup>1</sup> School of Health Systems and Public Health, Faculty of Health Sciences, University of Pretoria, Pretoria, South Africa

<sup>2</sup> Ancestry and Health Genomics Laboratory, Charles Perkins Centre, School of Medical Sciences, Faculty of Medicine and Health, University of Sydney, Camperdown, NSW, Australia

<sup>3</sup> Department of Urology, Sefako Makgatho Health Science University, Dr George Mukhari Academic Hospital, Medunsa, South Africa

<sup>4</sup> Department of Urology, St. Vincent's Hospital, Darlinghurst, NSW, Australia

<sup>5</sup> Manchester Cancer Research Centre, University of Manchester, Manchester M20 4GJ, United Kingdom

\* Correspondence: vanessa.hayes@sydney.edu.au

**Simple Summary:** African ancestry is a significant risk factor for aggressive prostate cancer (PCa), with southern African ethnicity conferring a near 3-fold increased global risk for associated mortality. It is well understood that epigenetic alterations drive PCa initiation and progression, coupled with somatic alterations in genes encoding epigenetic enzymes. However, differences in the somatic alterations in these genes in African *versus* European derived prostate tumors and how they may contribute to PCa health disparities has yet to be investigated, which forms the objective of this study. With current PCa care almost exclusively based on and tailored for men of European ancestry, the identification of African-specific novel PCa epigenetic cancer drivers ( $n = 18$ ), including therapeutic potential (6/18), offers clinical significance with the possibility of improving healthcare approaches and health outcomes for men of African ancestry.

**Abstract:** Prostate cancer is driven by acquired genetic alterations, including those impacting the epigenetic machinery. With African ancestry a significant risk factor for aggressive disease, we hypothesize that dysregulation among the roughly 656 epigenetic genes may contribute to prostate cancer health disparities. Interrogating prostate tumor genomic data from 109 men of southern African and 56 men of European Australian ancestry, we found African-derived tumors to present with a longer tail of epigenetic driver gene candidates (72 *versus* 10). Biased towards African-specific drivers (63 *versus* 9 shared), many are novel to prostate cancer (18/63) including several putative therapeutic targets (*CHD7*, *DPF3*, *POLR1B*, *SETD1B*, *UBTF* and *VPS72*). Through clustering of all variant types and copy number alterations, we describe two epigenetic PCa taxonomies capable of differentiating patients by ancestry and predicted clinical outcomes. We identified top genes in African and European-derived tumors that represent a multifunctional “generic machinery”, alteration to which may be instrumental in epigenetic dysregulation and prostate tumorigenesis. In conclusion, numerous somatic alterations in the epigenetic machinery drive prostate carcinogenesis but African-derived tumors appear to achieve this with greater diversity amongst such alterations. The greater novelty observed in African-derived tumors illustrates the significant clinical benefit to be derived from a much needed African tailored approach to prostate cancer healthcare aimed at reducing prostate cancer health disparities.

**Keywords:** prostate cancer; somatic alteration; epigenomics; epigenetic machinery; African ancestry; southern Africa; health disparity

## 1. Introduction

The epigenetic machinery comprises genes that encode proteins involved in regulating chromatin organization, histone modifications, DNA methylation, non-coding RNA and RNA methylation [1,2]. Many of these genes are directly or indirectly linked to epigenetic regulation of gene expression. The components of the epigenetic machinery can establish and maintain the epigenetic programming of a cell or they can dynamically alter it, thereby affecting both the identity and function of cells. Either way, proteins of the epigenetic machinery collectively interact with complex interdependence and interactivity. Cancer genome sequencing has increased our understanding of epigenetic dysregulation as a feature of tumor development. Independent of genomic aberrations, prostate tumors display DNA methylation patterns that differ from normal tissue. For example, the tumor suppressor *GSTP1* promoter region is typically hypermethylated in prostate cancer (PCa), resulting in the loss of its expression [3,4]. Androgen stimulation in PCa has the capacity to recruit histone modifiers, triggering changes in the chromatin states of PCa cells [5,6]. Ultimately, these epigenetic changes confer a more active or inactive chromatin state, dysregulating gene expression, thereby causing downregulation or silencing of tumor suppressor genes or a loss of regulation of genes that promote carcinogenesis. Although aberrant DNA methylation and disordered chromatin organization have long been recognized as features of cancer, the exact mechanisms driving epigenetic dysregulation are only beginning to be understood.

A number of studies have shown that driver gene mutations in several cancer types are enriched for epigenetic machinery genes, including PCa. While most individual genes are mutated infrequently, epigenetic machinery genes, as a class, are some of the most frequently mutated in PCa [7,8]. However, there are a number of epigenetic modulators revealed to carry frequent and recurrent mutations. In PCa, variants of this nature have been identified in mediators of DNA methylation (e.g., *TET2*, *MBD1*), histone acetylation (e.g., *KAT6B*, *ARID4A*), histone methylation (e.g., *KMT2C*, *SETD2*), as well as in chromatin remodelers (e.g., *ARID1A*, *SMARCA1*) [7–11]. The putative driver mutations are often truncating or missense [8], giving rise to non-functional proteins or proteins with altered functionality. Besides small somatic variants, prostate tumors are prone to acquire more complex variation, including structural variations (SVs) and copy number (CN) aberrations [12]. Of relevance to epigenetic machinery is double-stranded breaks, and the alteration of CpG methylation, local histone methylation and chromatin structure at DNA repair sites, with consequent altered gene expression [13–15]. Zhang *et al.*, 2019 [16] showed overall SV burden to be associated with global hypomethylation and increased expression of methyltransferase genes across cancer types. Epigenetic dysregulation by imbalanced genomic rearrangements has also been demonstrated in tumors, with the consequences of such rearrangements often being local, in that CN changes of a gene will alter the DNA methylation or gene expression of genomic regions close by [17].

What remains to be considered is the potential contribution of somatic alterations within the epigenetic machinery and relevance to PCa health disparities. Notably, genetic ancestry is a significant risk factor for aggressive PCa, specifically African ancestry. Within the United States, African American men are 1.7 times more likely to be diagnosed, and over twice as likely to die from PCa than European ancestral American men, reaching 3.1-fold for men younger than 65 years at diagnosis [18]. Globally, mortality rates are 2.7-fold greater for men from Sub-Saharan Africa [19]. While both genetic (common and rare variants) [20,21] and non-genetic (socioeconomic and cultural) contributing factors have been proposed [22], studies focused within populations from Sub-Saharan Africa have been scarce [23,24]. Conversely, ancestral differences in epigenetic aberrations have been observed for PCa, including genome-wide aberrant methylation patterns [25–28]. Gene-specific examples include hypermethylation of *CD44* in prostate tumors derived from African-American compared to European ancestral Americans which was positively correlated with tumor grade [29], and hypermethylation of *RARB* which was significantly associated with a higher risk of PCa in African-American over European ancestral American men [30]. While genomic aberrations in epigenetic machinery components have been studied previously for PCa [7,8,11], the contribution to ancestral differences associated with health outcomes is yet to be investigated, specifically within the context of Sub-Saharan Africa. Using a unique resource of prostate tumor genomic data derived from

113 men of Southern African and 53 men of European Australian ancestry [31], we set out to decipher whether genomic aberrations in epigenetic machinery components could, at least in part, explain the ancestral disparity observed for PCa.

## 2. Materials and Methods

### 2.1. Patient clinical characteristics and genomic data

Patients were recruited as part of the Southern African Prostate Cancer Study (SAPCS) with approval granted by the University of Pretoria Faculty of Health Sciences Research Ethics Committee (with US Federal wide assurance FWA00002567 and IRB00002235 IORG0001762) in South Africa (43/2010). In Australia, participant recruitment was approved by the St Vincent's Human Research Ethics Committee (HREC) (SVH/12/231), with genomic data generation approved by the St. Vincent's HREC (SVH/15/227). Additional study-specific approval was granted by the University of Pretoria Faculty of Health Sciences Research Ethics Committee (504/2022). Fresh blood-tumor paired deep sequenced whole genome data was generated as previously published [31]. In brief, data was generated using  $2 \times 150$  cycle paired-end Illumina HiSeq/NovaSeq sequencing, with reads aligned to the GRCh38 reference genome with alternative contigs, achieving a mean depth of coverage of 90X (range 28-139X) for tumor and 46X (range 30-97X) for blood. Both germline and somatic variants were called as previously described [31], including single nucleotide variants (SNVs) and small (< 50 bases) insertions and deletions (indels), structural variants (SVs, > 50 bases) and somatic copy number alterations (CNAs). While somatic variant frequencies and CNAs were used to determine tumor purities, which ranged from 13% to 88%, germline 7,472,833 biallelic SNVs were used to determine patient genetic ancestral fractions using fastSTRUCTURE v.1.0 population sub-structure analyses [32]. From the 183 patients included in the Jaratlerdsiri *et. al.*, 2022 study [31], patients were excluded if they lacked a positive PCa diagnosis ( $n = 6$ ), were Brazilian ( $n = 7$ ) or of admixed ancestry (defined as < 85% genetic contribution from a single ancestral identifier,  $n = 3$ ) or if their tumors were hypermutated (defined as > 30 mutations per Mb,  $n = 2$ ). Of a total of 165 treatment-naïve PCa patients included in our study, 4/109 (3.7%) of the African and 3/56 (5.4%) of the European ancestral patients lacked any genomic aberrations within the epigenetic machinery, leaving 105 and 53 for further interrogation, respectively. Both cohorts were biased towards aggressive disease, representing the International Society of Urological Pathology (ISUP) Group Grading 4 and 5 in 73% of African and 85% of European derived tumors (see Table S1 for a summary of clinical characteristics).

### 2.2. Epigenetic process group classification

Using the PathCards database to identify genes that map to epigenetic process pathways [33], we identified 656 epigenetic process-related genes. A SuperPath represents a cluster of one or several pathways that are grouped together based on the similarity of their associated genes (Table S2). Based on a review of the literature conducted in July 2022, additional epigenetic process-related genes were included for frequent reference and/or previous mention of relation to PCa [1,7,8,34–36]. As several epigenetic processes regulate the chromatin state, we further subdivided the 656 genes into their Epigenetic Process Group (EPG). EPG 1 genes are involved in chromatin organization and regulation ( $n = 530$  genes, Table S3), EPG 2 genes in histone modifications ( $n = 240$ , Table S4), EPG 3 genes in DNA methylation ( $n = 101$ , Table S5), EPG 4 genes in RNA regulation ( $n = 136$ , Table S6) and EPG 5 genes in epigenetic regulation of gene expression ( $n = 253$ , Table S7). Due to the multifunctional nature of these genes, a number have been assigned to multiple EPGs, while others are exclusive to a single EPG.

### 2.3. Tumor mutational burden (TMB), damaging variant detection and mutational frequency analysis

Whole genome tumor mutational burden (TMB) was calculated for each patient by taking the total number of small somatic variants (SNVs and indels), divided by the total genome size (3,088 Mbp). For each EPG classification, we defined mutational burden as the total number of small somatic coding variants present in a respective collection of epigenetic machinery genes, divided by the total

coding size (Mbp) of that gene collection. The damaging variant mutational burden was defined as the total number of potentially damaging variants present in a respective collection of epigenetic machinery genes, as per functional impact prediction, divided by the total coding size (Mbp) of that gene collection. Coding genome size for each EPG was mined using the Ensembl v.108 BioMart online data retrieval tool [37,38].

For each EPG, two approaches were used to identify potentially damaging variants and genes including functional impact prediction and mutational recurrence. Specifically, SIFT [39] and PolyPhen [40] scores for epigenetic process coding gene variants were determined using the SNPnexus v.4 annotation tool [41]. A variant was considered to be potentially damaging if identified by SIFT as “Deleterious” or “Deleterious - Low Confidence”, or if identified by PolyPhen as “Possibly Damaging” or “Probably Damaging”.

For recurrently mutated genes, we applied the computational tool DrGaP (driver genes and pathways) [42] to synonymous and non-synonymous somatic variants to determine the probability of each variant occurring by chance. DrGaP defines driver genes as those for which the non-synonymous mutation rate is significantly higher than the background mutation rate (BMR), while integrating biological variables such as the length of protein-coding regions. Using DrGaP, we defined significantly altered genes as those with a false discovery rate (FDR) < 5% using the Benjamini-Hochberg (BH) method.

Of the potentially damaging variants and genes found, we further identified genes that overlap with the Pan Cancer Analysis of Whole Genomes (PCAWG) compendium of mutational drivers [43]. To visualize patients’ overall somatic variant landscape, we used the maftools R package [44] to generate summary oncoplots for each of the EPGs.

#### 2.4. Integrative analysis of epigenetic machinery-driven prostate cancer subtypes

We performed integrative clustering of three genomic data types (small somatic variants, SVs and somatic CNAs) overlapping epigenetic process-related genes for 158 patients using the MOVICS (Multi-Omics integration and Visualization in Cancer Subtyping) R package [45]. We ran the optimal cluster number identification function with clusters ranging from 2 to 8, and MOVICS arbitrarily assigning an optimal cluster number of 8 for the variant data. The Cluster Prediction Index (CPI) and Gap statistic encouraged consideration of a cluster number of 3 as optimal (Figure S1). We executed ten classical clustering algorithms to subtype patients with different molecular features (Table S8), from which the resulting consensus matrix and silhouette plot ultimately demonstrated an optimal cluster number of 3 (Figure S2 and Figure S3), with which we proceeded for our analyses. Our prior research also informed this decision, where through whole tumor genome interrogation, we identified four PCa taxonomies which differentiated patients by their ancestries, termed Global Mutational Subtypes (GMSs) [31]. We then visualized the consensus clustering result with a heatmap and annotated top features based on posterior probability for each genomic feature as a driver, estimated by iClusterBayes [46–48], while additional hierarchical clustering was performed for CNAs [49]. The relationship between cancer subtypes and various clinical features was assessed.

#### 2.5. Statistical analyses

For continuous variables, group means were compared using the Mann-Whitney U test while for categorical variables, groups were compared using Fisher’s exact test. A  $p$ -value < 0.05 was regarded as statistically significant. Biochemical relapse (BCR) survival probability and cancer survival probability for European patients were analyzed with the Kaplan-Meier method, followed by group comparison with a log-rank test for significance.

### 3. Results

#### 3.1. Tumor mutational landscape

We have previously shown that African derived prostate tumors present with an elevated TMB and potentially damaging somatic variants [31,50]. When considering the epigenetic machinery

genes, 73.3% of African derived tumors and 62.3% of European derived tumors harbored altered genes, representing an average of 1.8 (range 1 to 20) and 1.4 (range 1 to 3) damaging variants, respectively. Compared to patient-matched whole genome data [31], we found the same trend (although not significant) when considering epigenetic machinery restricted mutational burden (Figure S4 and Table S9) and damaging variants (Table 2, Figure S5 and Table S10) for all genes and EPGs 1, 2, 4 and 5. Notably, each of the five EPGs displayed a higher mean mutational burden than can be expected when compared to the mean TMB observed for the whole genome, irrespective of patient ancestry. In contrast, EPG 3 showed the highest and equal mean mutational burdens (Figure S4) and number of damaging variants (Figure S5) for both Africans and Europeans (Figure S4).

To substantiate the potential carcinogenic nature of identified damaging genes (Table 3), we correlated for recurrence using the PCAWG compendium of mutational drivers [43]. Interestingly, *KMT2C*, identified here as a potentially damaging gene and by the PCAWG as a candidate cancer driver, although not predicted to contain functionally impactful variants, showed the highest recurrence rate, irrespective of ancestry (7.1% European, 5.5% African). The following most recurrent genes in African derived tumors were *CHD3* and *ARID1B* (4.6% and 3.7%, respectively), and in Europeans, were *BRD7*, *KDM6A*, *KDM6B*, *KMT2A*, *KMT2B* and *RANBP2* (each 3.6%). Seven genes presented in tumors from three or more patients of African ancestry, while being notably absent from European derived tumors and included *CHD3* (5 patients), *ARID1B* (4), *HDAC4* (3), *ARID1A* (3), *CHD1* (3), *PRDM16* (3) and *STAG2* (3). Conversely, genes exclusively altered in two or more European derived tumors included *KMT2B* (2 patients) and *RANBP2* (2). In each EPG, African derived tumors displayed a greater mutation frequency for most genes (likely reflective of a greater sample number), along with richer diversity in variant types than European derived tumors (Figure 1). In terms of mutation frequency, the top gene(s) in EPG 1 and EPG 2 is *KMT2C*; *CBX2*, *DNMT3B* and *TDG* in EPG 3; and *CHD3* in EPG 4 and EPG 5.

**Table 1.** For each epigenetic process group, genes in Africans and Europeans containing damaging variants based on functional impact prediction.

	Africans (unique)	Africans & Europeans (shared)	Europeans (unique)
EPG 1	<i>AIRE</i> , <i>ARID1A</i> <sup>2</sup> , <i>ARID1B</i> <sup>2</sup> , <i>BAZ1B</i> , <i>BRD1</i> , <i>BRD7</i> , <i>BRWD1</i> <sup>2</sup> , <i>CARM1</i> , <i>CHD1</i> <sup>1,2</sup> , <i>CHD3</i> <sup>2</sup> , <i>CXXC1</i> , <i>DNMT1</i> , <i>EHMT1</i> , <i>EHMT2</i> , <i>ELL</i> , <i>GLI1</i> , <i>HCFC1</i> , <i>HDAC4</i> <sup>1,2</sup> , <i>HJURP</i> , <i>KAT8</i> , <i>KDM2B</i> , <i>KDM6B</i> <sup>1</sup> , <i>KMT2D</i> <sup>2</sup> , <i>MBD5</i> , <i>MED12</i> , <i>MEN1</i> , <i>NCOR1</i> , <i>NUP188</i> , <i>NUP93</i> , <i>PAX3</i> , <i>PELP1</i> , <i>POLR1A</i> , <i>POLR3A</i> , <i>PRDM16</i> , <i>PRMT6</i> , <i>PSIP1</i> <sup>2</sup> , <i>RBBP5</i> , <i>REST</i> , <i>RNF20</i> , <i>RTF1</i> , <i>SATB2</i> , <i>SETD2</i> <sup>2</sup> , <i>SMARCA1</i> <sup>2</sup> , <i>SMARCA4</i> <sup>1</sup> , <i>SRCAP</i> , <i>SSRP1</i> , <i>STAG2</i> <sup>1,2</sup> , <i>TADA2B</i> <sup>1,2</sup> , <i>TAF1</i> , <i>TAF6L</i> , <i>TPR</i> , <i>VPS72</i> , <i>WAPL</i> , <i>WHSC1L1</i> <sup>1</sup>	<i>ATRX</i> <sup>2</sup> , <i>BRD4</i> , <i>CHD4</i> , <i>DNMT3B</i> , <i>NCOR2</i> <sup>2</sup> , <i>PADI1</i> , <i>PBRM1</i> <sup>1</sup> , <i>SETD1B</i> <sup>2</sup>	<i>AR</i> , <i>CBX2</i> , <i>CLOCK</i> , <i>ELOA</i> , <i>ERCC3</i> , <i>HNRNPA2B1</i> , <i>JAK2</i> , <i>KAT2B</i> , <i>KAT6A</i> , <i>KDM1B</i> , <i>KDM6A</i> , <i>KMT2B</i> <sup>3</sup> , <i>KMT5C</i> , <i>PRDM5</i> , <i>RANBP2</i> , <i>RB1</i> , <i>RNF40</i> , <i>SIRT3</i>
EPG 2	<i>AIRE</i> , <i>ARID1A</i> <sup>2</sup> , <i>ARID1B</i> <sup>2</sup> , <i>CARM1</i> <sup>2</sup> , <i>CHD3</i> <sup>2</sup> , <i>DNMT1</i> , <i>EHMT1</i> , <i>EHMT2</i> <sup>2</sup> , <i>HDAC4</i> <sup>1,2</sup> , <i>KDM2B</i> , <i>KDM6B</i> <sup>1,2</sup> , <i>KMT2D</i> <sup>2</sup> , <i>MBD5</i> , <i>NCOR1</i> , <i>PRDM16</i> <sup>2</sup> , <i>PRMT6</i> , <i>RBBP5</i> <sup>2</sup> , <i>REST</i> <sup>2</sup> , <i>SETBP1</i> , <i>SETD2</i> <sup>2</sup> , <i>SMARCA4</i> <sup>1,2</sup> , <i>SRCAP</i> , <i>WHSC1L1</i> <sup>1</sup>	<i>BRD4</i> , <i>CHD4</i> , <i>DNMT3B</i> , <i>NCOR2</i> <sup>2</sup> , <i>PBRM1</i> <sup>1</sup> , <i>SETD1B</i> <sup>2</sup>	<i>JAK2</i> , <i>KDM1B</i> , <i>KDM6A</i> <sup>4</sup> , <i>KMT2B</i> <sup>3</sup> , <i>KMT5C</i> , <i>PRDM5</i> , <i>SETD5</i>
EPG 3	<i>DNMT1</i> , <i>MBD5</i>	<i>DNMT3B</i> , <i>TDG</i>	<i>CBX2</i>
EPG 4	<i>ALG13</i> , <i>BAZ1B</i> , <i>CHD3</i> <sup>2</sup> , <i>DNMT1</i> , <i>EHMT2</i> , <i>GSK3B</i> , <i>POLR1A</i> , <i>POLR1B</i> , <i>TDRD7</i> , <i>UBTF</i>	<i>CHD4</i> , <i>DNMT3B</i>	<i>ERCC3</i> , <i>KAT2B</i> , <i>SF3B1</i>

	<i>AIRE, ALG13, BAZ1B, CHD3</i> <sup>2</sup> , <i>DNMT1, EHMT2, FOXO3, GSK3B, CHD4, DNMT3B, NCOR2, TDG, TP53</i>	<i>ERCC3, HDAC9, KAT2B, RANBP2,</i> <i>RXRβ, SF3B1, SIRT3, STAT3</i>
EPG 5	<i>HDAC4</i> <sup>1,2</sup> , <i>MXD1, NCOR1, POLR1A,</i> <i>POLR1B, RANGAP1, REST, SRCAP,</i> <i>UBTF, XRCC6</i>	

EPG, epigenetic process group. <sup>1</sup> Genes that contain more than one damaging variant in an EPG (Africans). <sup>2</sup> Genes identified as potentially damaging in an EPG, based on recurrent somatic variant identification (Africans). <sup>3</sup> Genes that contain more than one damaging variant in an EPG (Europeans). <sup>4</sup> Genes identified as potentially damaging in an EPG, based on recurrent somatic variant identification (Europeans).

**Table 2.** Damaging variant summary for each epigenetic process group, for Africans and Europeans, based on functional impact prediction.

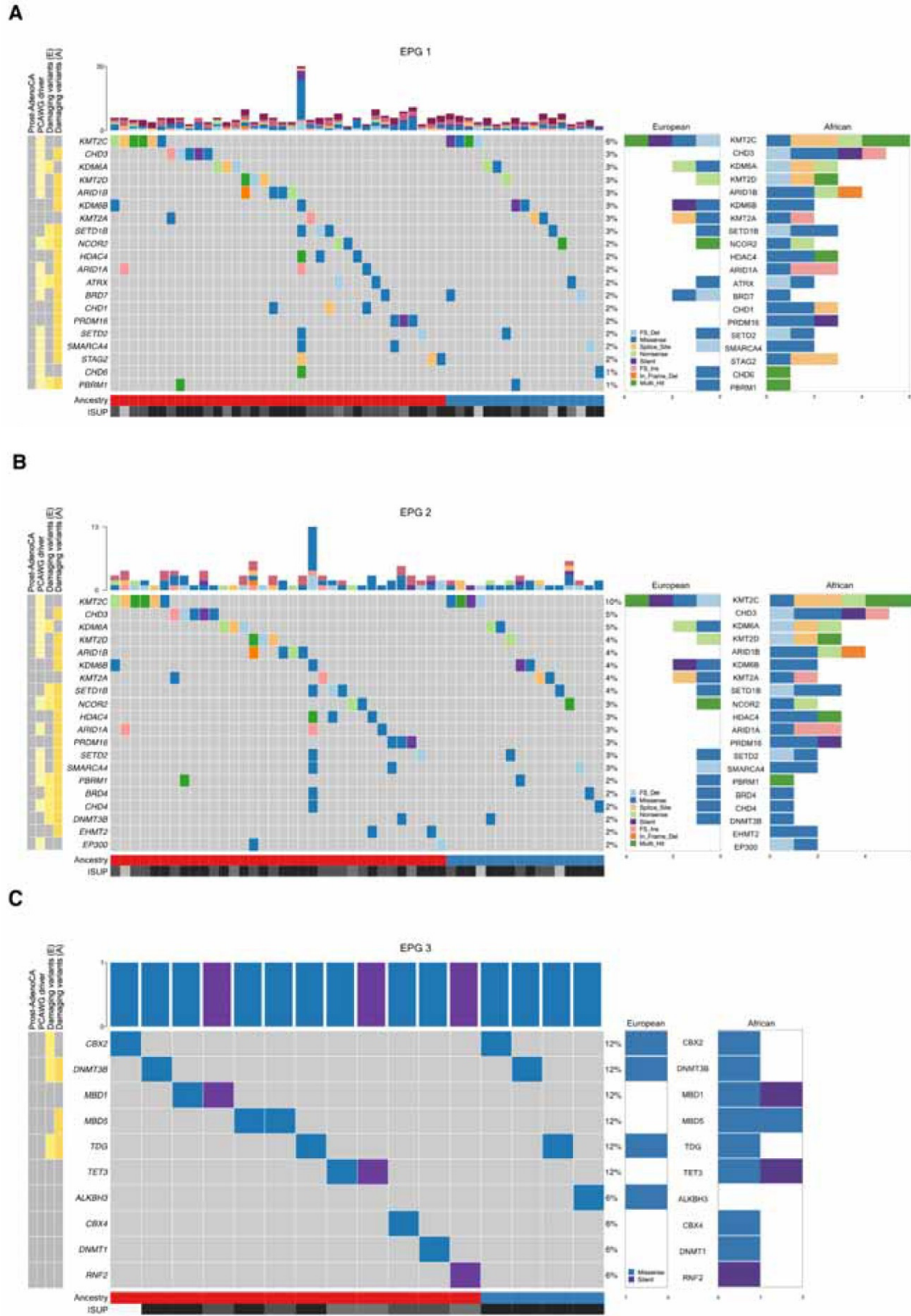
	Africans (n = 109)			Europeans (n = 56)		p-value *
	Total genes	Total damaging variants	Samples that contain damaging variants	Total damaging variants	Samples that contain damaging variants	
EPG 1	530	71	37	26	22	0.4988
EPG 2	240	35	21	13	12	0.8375
EPG 3	101	5	5	3	3	1
EPG 4	136	12	9	5	4	1
EPG 5	253	31	23	15	13	0.8426

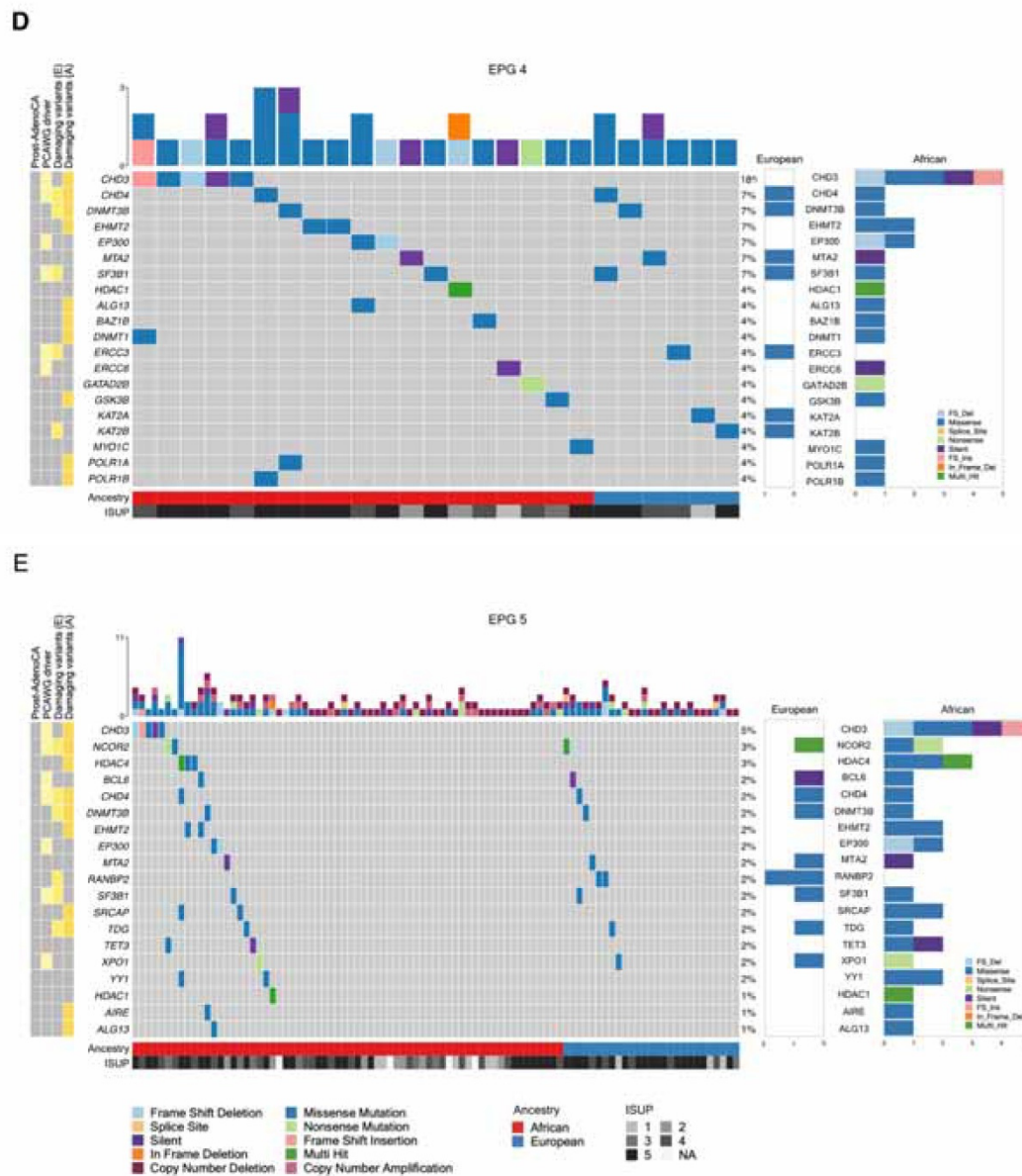
\* Fisher's exact test to compare variance in number of samples containing damaging variants between Africans and Europeans.

**Table 3.** For each epigenetic process group, genes in Africans and Europeans identified as potentially damaging based on recurrent somatic variants.

	Africans (unique)	Africans & Europeans (shared)	Europeans (unique)
EPG 1	<i>ARID1A</i> <sup>1,2</sup> , <i>ARID1B</i> <sup>1,2</sup> , <i>ATRX</i> <sup>1,2</sup> , <i>BRWD1</i> <sup>1,2</sup> , <i>CHD1</i> <sup>2</sup> , <i>CHD3</i> <sup>1,2</sup> , <i>CHD7, DPF3, ELP2, EP300</i> <sup>1</sup> , <i>GATAD2B, GLI3, HDAC1, HDAC3</i> <sup>1</sup> , <i>HDAC4</i> <sup>2</sup> , <i>KMT2A, KMT2D</i> <sup>1,2</sup> , <i>NCOR2</i> <sup>1,2</sup> , <i>NUP35, PSIP1</i> <sup>1,2</sup> , <i>SETD1B</i> <sup>2</sup> , <i>SETD2</i> <sup>1,2</sup> , <i>SMARCA1</i> <sup>1,2</sup> , <i>STAG2</i> <sup>1,2</sup> , <i>TADA2B</i> <sup>2</sup> , <i>XPO1</i> <sup>1</sup>	<i>KMT2C</i> <sup>1</sup>	—
EPG 2	<i>ARID1A</i> <sup>1,2</sup> , <i>ARID1B</i> <sup>1,2</sup> , <i>ARID5B</i> <sup>1</sup> , <i>BRMS1, CARM1</i> <sup>2</sup> , <i>CHD3</i> <sup>1,2</sup> , <i>EHMT2</i> <sup>2</sup> , <i>EP300</i> <sup>1</sup> , <i>GATAD2B, HDAC1, HDAC3</i> <sup>1</sup> , <i>HDAC4</i> <sup>2</sup> , <i>KDM6B</i> <sup>2</sup> , <i>KMT2A, KMT2D</i> <sup>1,2</sup> , <i>NCOR2</i> <sup>1,2</sup> , <i>PRDM16</i> <sup>2</sup> , <i>RBBP5</i> <sup>2</sup> , <i>REST</i> <sup>2</sup> , <i>SETD1B</i> <sup>2</sup> , <i>SETD2</i> <sup>1,2</sup> , <i>SMARCA4</i> <sup>1,2</sup>	<i>KDM6A</i> <sup>1,3</sup> , <i>KMT2C</i> <sup>1</sup>	—
EPG 3	—	—	—
EPG 4	<i>CHD3</i> <sup>1,2</sup> , <i>HDAC1</i>	—	—
EPG 5	<i>CHD3</i> <sup>1,2</sup> , <i>HDAC1, HDAC4</i> <sup>2</sup>	<i>TP53</i> <sup>1,2,3</sup>	—

EPG, epigenetic process group. <sup>1</sup> Pan Cancer Analysis of Whole Genomes candidate cancer mutational driver. <sup>2</sup> Contains damaging variants in Africans based on functional impact prediction. <sup>3</sup> Contains damaging variants in Europeans based on functional impact prediction.





**Figure 1.** Somatic alteration landscape for each epigenetic process group. The top bar graph shows the number of non-synonymous variants and copy number alterations per tumor. The middle gene panel reports synonymous and non-synonymous variants in a maximum of 20 top altered genes. The bottom panels annotate sample ancestries and ISUP grades. The right-hand bar plots display European and African mutation frequencies, respectively. The left-hand panel indicates whether top genes identified in the oncoplot overlap with candidate cancer mutational drivers identified by the Pan Cancer Analysis of Whole Genomes and if so, whether the gene was identified as a candidate driver in prostate adenocarcinoma. Finally, the left-hand panel additionally indicates whether genes displayed in the oncoplot contain damaging variants based on functional impact prediction, as described in Table 1. Yellow tiles indicate 'yes', grey tiles indicate 'no'. Due to the hypermutated nature of these genes and their indirect epigenetic involvement in chromatin state regulation, *TP53*, *SPOP* and *FOXA1* genes were excluded from the oncoplots. (A) Epigenetic process group 1; (B) epigenetic process group 2; (C) epigenetic process group 3; (D) epigenetic process group 4; (E) epigenetic process group 5. A, African; E, European; FS\_Del, frameshift deletion; FS\_Ins, frameshift insertion; In\_Frame\_Del, in-frame deletion; ISUP, International Society of Urologic Pathologists; PCAWG, Pan Cancer Analysis of Whole Genomes; Prost-AdenoCA, prostate adenocarcinoma.

### 3.2. Integrative clustering analysis

Hierarchical consensus clustering through joint analysis of somatic variant data (small variants, SVs and CNAs) for epigenetic machinery genes, identified three PCa subtypes (Epigenetic Cancer Subtype 1-3 i.e., ECS1-ECS3). These three subtypes are presented in Figure 2A for 158 patients (columns). Among genes with significantly different alteration frequencies between the cancer subtypes, ECS1 demonstrated a higher frequency of small somatic alterations and SVs compared to ECS2 and ECS3 (Table S11 and Table S12), though all three ECSs were rather mutationally quiet. Overall, SV demonstrated similarity between the subtypes, while CNAs appeared to be the strongest indicator for subtype clustering. While ECS1 showed both CN gains (median 1 gene gained per tumor, range 0 to 38 genes gained) and losses (median 0 genes deleted per tumor, range 0 to 8 genes deleted), ECS2 and ECS3 were characterized by substantial CN gains (median 9 genes gained per tumor, range 0 to 187 genes gained) and CN losses (median 5.5 genes deleted per tumor, range 0 to 23 genes deleted), respectively, and were both African-predominant (91% and 67% African, respectively;  $p < 0.001$ ; Table S13). The four European patients allocated to ECS2 resided within Australia ( $n = 3$ ) or South Africa ( $n = 1$ ).

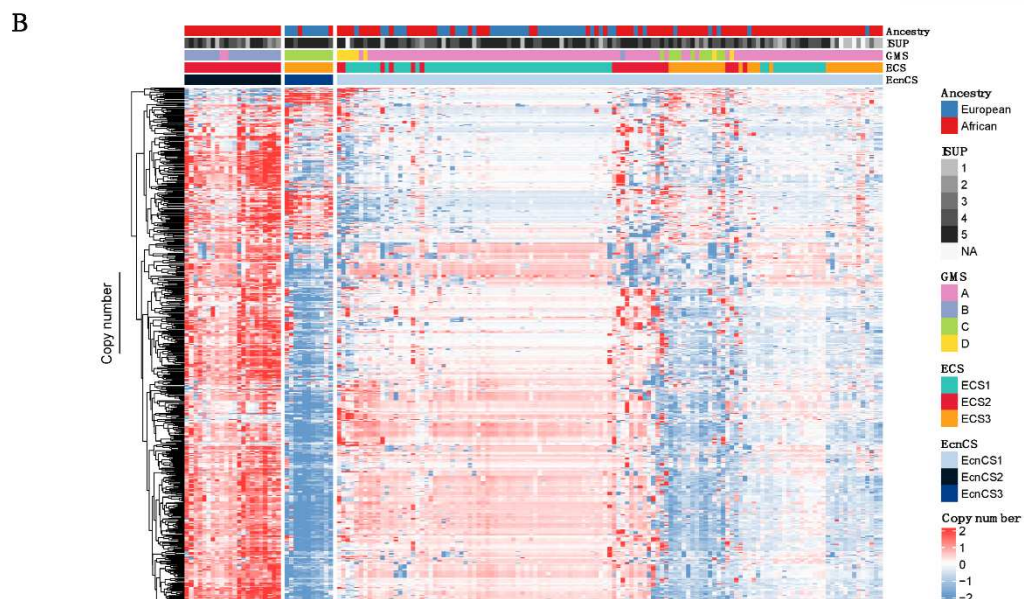
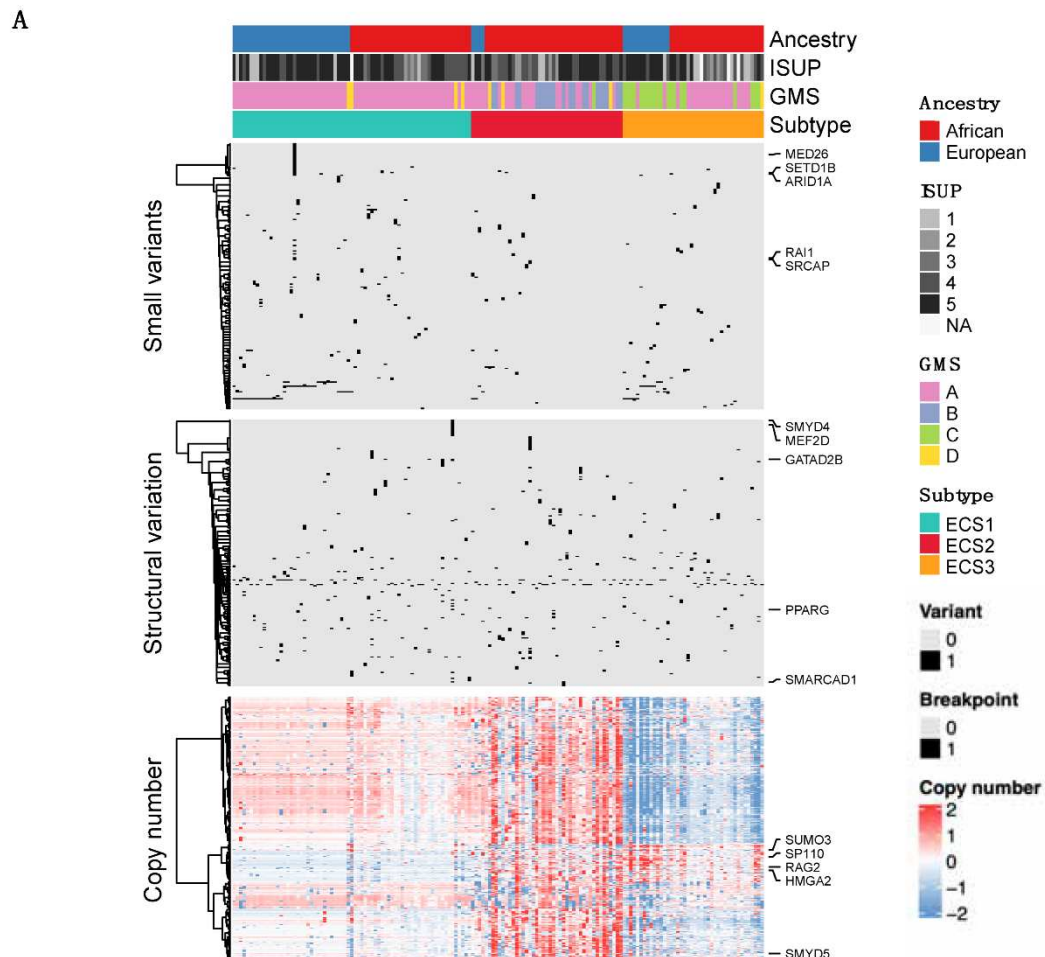
Feature selection identifies genomic features that make important contributions the oncogenic processes and that drive the integrative clustering. Using iClusterBayes, we calculated the posterior probability for each genomic feature as a driver. We found the top 5 identified features were dominated by genes belonging to EPG 1, followed by EPG 2 and EPG 5. Posterior probability was only high for small somatic mutation-identified features, with 30 potential drivers identified in total (posterior probability  $> 0.5$ , Table S14), a number of which were in agreement with identified damaging genes (Table 1 and Table 3). Irrespective of ancestry, potential drivers and damaging variant-containing genes included *SETD1B*, *CHD4* and *BRD4*, while we observed a longer tail of unique drivers within Africans including *SRCAP*, *ARID1A*, *HCFC1*, *BRD1*, *POLR1B*, *STAG2*, *VPS72*, *UBTF*, *MXD1*, *KDM6B*, *HDAC4*, *ELL*, *RANGAP1*, *SMARCA4*, *KDM2B* and *NCOR1*. This suggests different epigenetic mechanisms, especially chromatin organization and regulation, play a role in PCa among Africans.

Through whole-genome PCa molecular taxonomy of the same sample source, all patients were classified into one of four recently described GMSs [31]. In brief, both GMS-A and GMS-C are ethnically diverse, marked by a mutationally quiet landscape and substantial CN losses, respectively. Conversely, GMS-B and GMS-D are African-predominant, with GMS-B demonstrating substantial CN gains, and GMS-D a mutationally noisy landscape including CN gains and losses. We found significant correlation between our ECSs and GMSs ( $p < 0.001$ , Table S13). Of the ECS1 tumors, 94% had previously clustered with GMS-A, while the ECS2 tumors were dominated by GMS-B (47%), mapping exclusively to African-specific ECS2 tumors, and GMS-A (44%). In contrast, all European-derived ECS2 tumors belong to GMS-A. ECS3, characterized by near-equal contributions from GMS-A (50%) and GMS-C (48%), demonstrated the African-European GMS-C almost exclusively. Finally, the African-specific GMS-D tumors were the least represented of the subtypes, showing no favorable clustering with distribution across ECS1 to ECS3.

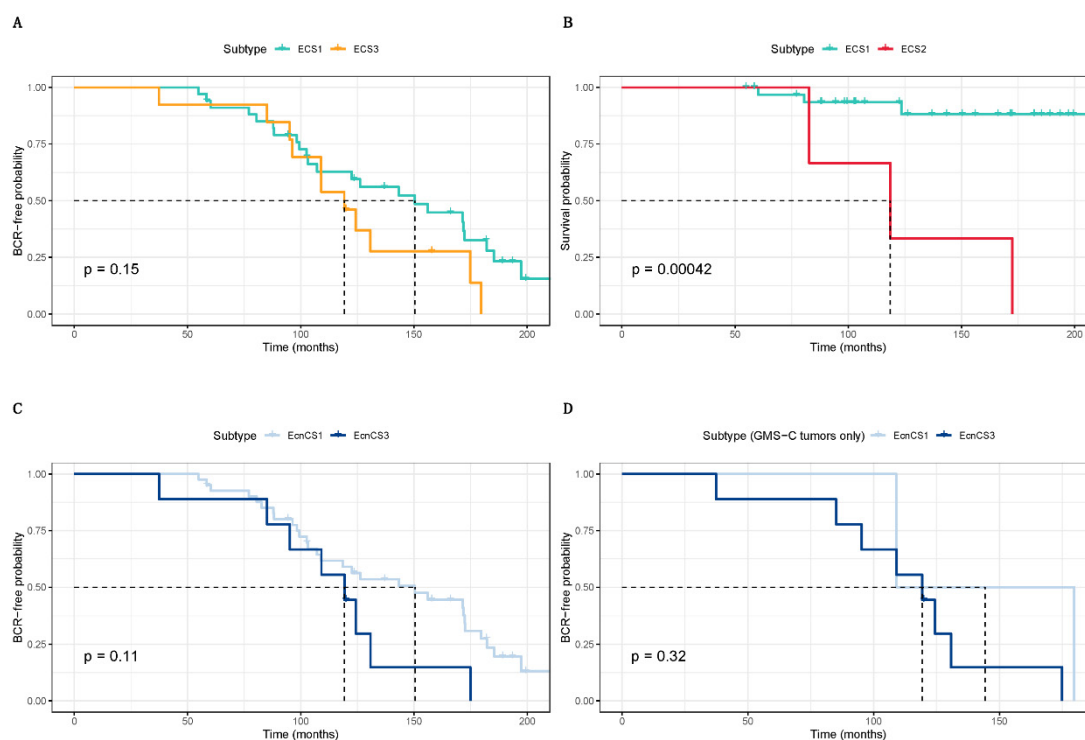
As CNAs appear to be the strongest determinant for ECS clustering, we performed hierarchical clustering on the CN data alone (Figure 2B). Recognizing three Epigenetic CN Cancer Subtypes (EcnCS, Table S15), we found EcnCS2 to be African-exclusive and dominated by GMS-B (91%), while EcnCS3 was European-dominant (82%), representing only GMS-C tumors. The ancestry diverse subtype EcnCS1 was predominated by GMS-A tumors (85%). Overall, the EcnCSs correlated significantly with our previously-identified GMSs ( $p < 0.001$ ) and ECSs ( $p < 0.001$ ).

Through the availability of extensive follow-up data for our European Australian patients (mean 127.4 months, range 37.4 to 214.3 months), we previously predicted for significantly better clinical outcomes for patients presenting with GMS-A over GMS-C tumors, defined as no biochemical relapse (BCR) and/or survival [31]. Here we sought to correlate our identified ECSs and EcnCSs with clinical outcomes in our patient-matched European cohort. BCR-free probability revealed better clinical outcomes for ECS1 over ECS3 tumors, although not significant (Figure 3A, log-rank test,  $p = 0.15$ ), and ECS2 tumors predicted poorer survival probability than ECS1 tumors (Figure 3B, log-rank test,

$p < 0.001$ ). Notably, of the seven deaths recorded, all three Australian European men presenting with ECS2 died within 172.5 months of their surgery. Although not significant, when considering the CN subtypes, BCR-free probability showed better clinical outcome for patients presenting with EcnCS1 over EcnCS3 (Figure 3C, log-rank test,  $p = 0.11$ ), even when considering only the GMS-C tumors (Figure 3D, log-rank test,  $p = 0.32$ ), which are characterized by poor clinical outcome.



**Figure 2.** (A) Consensus clustering heatmap based on 10 multi-omics integrative clustering algorithms, for somatic data (small variants, structural variation and copy number alterations) spanning epigenetic machinery genes in 105 African and 53 European derived prostate tumors. For each variant data, the top 5 features are listed. Feature selection identifies complex cross-talk between different variant data, which may allude to biological significance driving cancer heterogeneity. (B) Hierarchical clustering heatmap based only on somatic copy number alteration data spanning epigenetic machinery genes, for 105 African and 53 European derived prostate tumors. ECS, epigenetic cancer subtype; EcnCS, epigenetic copy number cancer subtype; GMS, global mutational subtype; ISUP, International Society of Urologic Pathologists; NA, not available.



**Figure 3.** Kaplan-Meier curves of consensus clustering results for European patients. The probability estimates, 95% confidence intervals and  $p$ -values (log-rank test) are indicated. (A) Kaplan-Meier curve of biochemical relapse (BCR)-free probability for ECS1 ( $n = 34$ ) and ECS3 ( $n = 13$ ) tumors. (B) Kaplan-Meier curve of the cancer survival probability for ECS1 ( $n = 34$ ) and ECS2 ( $n = 3$ ) tumors. (C) Kaplan-Meier curve of BCR-free probability for EcnCS1 ( $n = 41$ ) and EcnCS3 ( $n = 9$ ) tumors. (D) Kaplan-Meier curve of BCR-free probability for EcnCS1 ( $n = 2$ ) and EcnCS3 ( $n = 9$ ) tumors allocated to GMS-C. BCR, biochemical relapse; ECS, epigenetic cancer subtype; EcnCS, epigenetic copy number cancer subtype; GMS, global mutational subtype.

#### 4. Discussion

Overall, compared with European derived tumors, African derived prostate tumors presented with a higher burden of variants and potentially damaging variants across epigenetic machinery genes. Although our findings were in line with our previous work, demonstrating a whole-genome African-elevated TMB [50], the epigenetic burden within Africans was not significantly higher than that of Europeans. When considering all epigenetic machinery genes, the African derived tumors demonstrated a higher overall mutational frequency than the European derived tumors, although not significant. In contrast to a recent genome-wide study which found ~20% of prostate tumors harbored driver mutations across 12 epigenetic machinery genes [8], here we report frequencies of 52.3% for African and 50.0% for European derived tumors, which may be explained by a larger inclusivity of epigenetic regulators in our study (656 versus only 12 genes). Irrespective of patient ancestry, we found *KMT2C* to be the most frequently mutated PCa epigenetic regulator gene,

concurring with previous studies reporting frequencies of ~5%-8% [7,8]. The type 2 histone lysine methyltransferase *KMT2C* is one component of chromatin remodeling machinery responsible for DNA promoter and enhancer regulation, ultimately promoting active chromatin conformations. With strong links to numerous cancer types, mutations in these components confirm their roles as tumor suppressors [51]. Irrespective of potentially damaging (Table 1) and recurrent (Table 3) driver gene classification, as was observed for the whole genome, African derived tumors showed a longer tail of African-specific epigenetic gene candidates. Besides *KMT2C*, the only recurrent driver genes to be shared between the ancestries are the well-known tumor suppressor genes *KDM6A* and *TP53*. In contrast to a lack of European-specific recurrent drivers, 35 African-specific recurrent driver genes were observed. The latter included putative loss-of-function PCa mutations previously reported for *ARID1A*, *ATRX*, *CHD1*, *CHD3*, *HDAC4*, *KMT2A*, *KMT2D*, *SETD2* and *SMARCA1* [7], with *BRMS1*, *CARM1*, *EHMT2*, *GLI3*, *HDAC1*, *KDM6B*, *PRDM16*, *RBBP5* and *REST* having known roles in PCa [52–60]. *ARID1B*, *ARID5B*, *BRWD1*, *EP300*, *HDAC3*, *NCOR2*, *PSIP1*, *SMARCA4*, *STAG2* and *XPO1*, although reported by PCAWG [43] are new to PCa, leaving *CHD7*, *DPF3*, *ELP2*, *GATAD2B*, *NUP35*, *SETD1B* and *TADA2B* as novel candidate drivers.

Taking a closer look at the epigenetic processes, in contrast to our whole genome data and EPGs 1, 2, 4 and 5 with an African ancestry-elevated burden, we consistently showed EPG3 alterations to be similar between the ancestries. Overall this group of DNA methylation gene regulators appears to be highly conserved, as previously reported [61], with no recurrent drivers (Table 3) and potentially damaging variants in only two genes, *DNMT3B* and *TDG* (Table 1). *DNMT3B* is a DNA methyltransferase (DNMT) enzyme responsible for establishing and maintaining methylation at satellite sequences and gene bodies [62,63]. DNMT polymorphisms are associated with PCa progression by means of downregulatory tumor suppressor gene promoter methylation [64] and elevated *DNMT3B* expression in aggressive *versus* non-aggressive PCa cell lines [65]. Similarly, a damaging variant in *DNMT1* was observed in an African sample. *TDG*, or Thymine DNA Glycosylase, plays a key role in active DNA demethylation, and as a tumor suppressor, several polymorphisms in *TDG* are associated with increased risk for cancer, although this gene has also been found to act as an oncogene, promoting tumorigenesis [66–68]. It remains to be determined whether the *TDG* damaging variants identified in our study have gain-of-function or loss-of-function properties. Ultimately, aberrant DNA methylation is a hallmark of cancer progression, and dysregulation of the DNA methylation machinery may lead to a reprogramming of the epigenomic landscape in cancer.

Using hierarchical consensus clustering for all somatic mutational types (small variants, SVs and CNAs), we describe two epigenetic PCa taxonomies (ECS and EcnCS), which independently showed significant agreement with our previously-reported GMSs [31]. Showing extensive overlap amongst Europeans, both ECS3 and GMS-C tumors predicted poorer clinical outcome over ECS1 and GMS-A tumors, respectively, demonstrating the bias of each GMS to an ECS. As such, our identified ECSs validate the whole genome-derived GMSs and are able to relatively distinguish those global subtypes based on just a subset of the genome, indicating a significant role for epigenetic mechanisms in PCa development. While numbers for recorded PCa-associated death are arguably small (7/50 Australians), it is notable that all three Australian European derived tumors presenting with ECS2 succumbed to PCa, i.e., 42.9% of PCa deaths were associated with ECS2 tumors. Furthermore, as ECS2 is otherwise characterized by African predominance, specifically ISUP group grading > 3 PCa (78% of African derived ECS2 tumors *versus* 73% of all African derived tumors), our data suggests that ECS2 is a predictor of poor outcome.

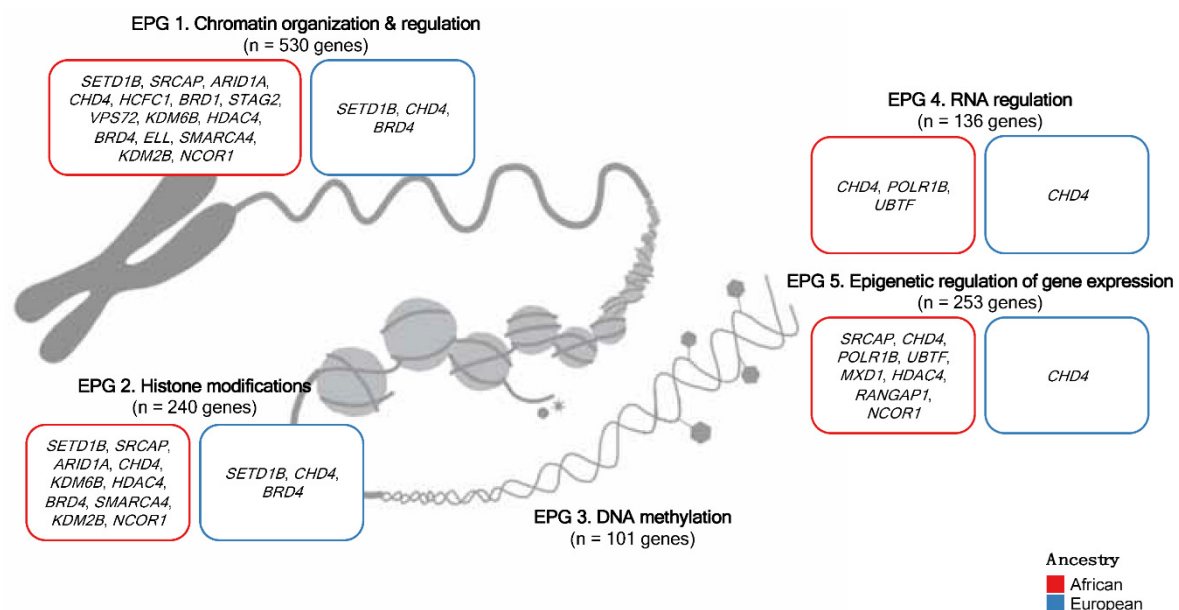
More aligned with the whole genome-derived GMSs [31], the EcnCSs showed ancestral distinction including an African-specific subtype (EcnCS2), a European-predominant subtype (EcnCS3) and a shared subtype (EcnCS1). Notably, EcnCS2 defined by significant CN gain further defines ECS2, while almost exclusively incorporates all the African-specific GMS-B tumors (95.2%, 20/21). EcnCS3 further distinguished ECS3 and the poor outcome-associated GMS-C, as a singular cluster defined by epigenetic gene CN loss. Of the GMS-C tumors, EcnCS3 presented with a higher predominance of ISUP group grading 5 PCa (81.8%) over EcnCS1 (60.0%), with EcnCS3 predicting

poorer outcome for BCR than EcnCS1, indicating a more aggressive presentation for EcnCS3-GMS-C tumors. Suggesting that epigenetic CNAs alone have the potential to predict patient outcomes in our study, the relationship between CNAs and DNA methylation in cancer has been examined previously [17], although not at length. However, it is generally understood that a gene's CNAs affect DNA methylation of genomic regions nearby. These two processes may be negatively associated (i.e., DNA methylation decreases with copy number gain and *vice versa*), in which case the effect is localized to CpG islands; or they may be positively associated, in which case the open sea (genomic region beyond 4 kb from a CpG island border) is affected. Either way, it has been suggested that genome-wide DNA methylation changes in response to CNA events are likely initiated and maintained by some "generic" machinery. This is supported by the Sun *et al.*, (2018) [17] finding that CNA events and their association with altered DNA methylation are similar across cancer types. Another observation common to several cancer types is ancestral differences in DNA methylation patterns. This has been observed in PCa, in which African-American tumors display a higher prevalence of DNA hypermethylation at disease-related loci compared to European-American tumors [26]. Therefore, each of the epigenetic (copy number) cancer subtypes, with their distinct CNA events, likely give rise to distinct aberrant DNA methylation patterns. Whether those DNA methylation patterns would cluster in agreement with the CN patterns is yet to be determined. Of course, aberrant DNA methylation does not only arise in response to CN gain/loss events. However, inclusion of patient-matched DNA methylation data could decipher this.

As a function of hierarchical clustering, feature selection identified the top five genes for each variant type for potential driver gene classification. In rank order based on posterior probability, the top five features for small somatic variant data were *RAI1*, *SETD1B*, *SRCAP*, *ARID1A* and *MED26*, for SV data were *SMYD4*, *GATAD2B*, *PPARG*, *MEF2D* and *SMARCA1*, and for CNAs, *HMG2A*, *SMYD5*, *SUMO3*, *SP110* and *RAG2*. Of the top five selected features for SV and CNA data, and for the 30 small somatic mutation-identified drivers (Table S14), the genes that appear new to PCa lacking in PCAWG and African-specific include *SP110*, *GATAD2B*, *RAI1*, *MED26*, *BRD1*, *POLR1B*, *VPS72*, *ELP5*, *UBTF*, *MXD1*, *DR1* and *ELL*. Formerly considered to be a transcriptional regulator of circadian clock components in neuronal tissue, a recent study found *RAI1* to act as a tumor suppressor in esophageal cancer; prior to this finding, the functional role of *RAI1* in tumors was unknown [69]. *SETD1B*, an essential component of a histone methyltransferase complex, believed to have essential, even housekeeping, functions within cells [70], although having no clear role in malignancy, has been reportedly mutated in gastric and colorectal cancers [71]. *MED26*, belonging to the Mediator complex (MED) gene family, while implicated in several cancer types does not include PCa [72,73]. Additionally, a number of the epigenetic regulators specific to African tumors have been identified as potential therapeutic targets. Chromatin remodeler *CHD7*, a somatic driver candidate in colorectal cancer (CRC), promotes CRC cell growth by binding target gene promoters encouraging an open chromatin conformation and subsequent transcription, whereas *CHD7* knockdown inhibits CRC cell growth [74]. Similarly, *POLR1B* knockdown induces lung cancer cell apoptosis [75], *VPS72* knockdown inhibits the proliferation, invasion and migration of hepatocellular carcinoma (HCC) cells [76], and *UBTF* silencing suppresses melanoma cell proliferation [77]. *DPF3*, a chromatin remodeling cofactor significantly downregulated in breast cancer tissue, promoting the proliferation of breast cancer cells, has been suggested as a novel therapeutic target for breast cancer therapy [78]. Increased *SETD1B* expression in HCC positively correlated with tumor size, clinical stage and liver cirrhosis. Decreased *SETD1B* expression was associated with increased patient survival times, identifying this histone methyltransferase as a potential therapeutic target in HCC [79]. While several of the African-specific drivers show clinical relevance, the remaining genes are not well-studied as therapeutic targets in cancer [80–83].

Ultimately, alterations to genes encoding epigenetic machinery components are increasingly recognized in many cancer types, including PCa. From this study, based on genes containing potentially damaging variants as per functional impact prediction and/or recurrence as well as putative driver gene status as defined by feature selection during hierarchical clustering, we have summarized top genes in African and European derived tumors per EPG, that may be instrumental

in epigenetic dysregulation and subsequent development and/or progression of PCa (Figure 4). Identifying a number of putative drivers, *ARID1A*, *CHD4*, *HCFC1*, *STAG2*, *SMARCA4* and *NCOR1* are known cancer driver genes [43]. Notably, there is extensive ancestral overlap amongst the top genes in all the EPGs. The assignment of numerous epigenetic machinery genes to more than one EPG is owed to the multifunctional nature of these genes. For example, *CHD4*, or Chromodomain Helicase DNA Binding Protein 4, is the main component of the nucleosome remodeling and deacetylase (NuRD) complex that plays an important role in epigenetic transcriptional repression. *CHD4*/NuRD also regulates RNA synthesis [84]. As such, the multifunctionality of *CHD4* warrants its inclusion in EPGs 1, 2, 4 and 5. Rather than the top genes being epigenetic machinery components exclusive to a single EPG, this broad overlap is reminiscent of the previously-discussed CN-DNA methylation aberration events common to many cancer types arising from some “generic machinery”. Indeed, epigenetic regulators are well-conserved and mutated infrequently. However, should epigenetic regulation be disrupted, as a class, perhaps the genomic alteration of a common core group of multifunctional epigenetic regulators will achieve this, promoting tumorigenesis. Many of our top-identified genes have well-established roles across cancer types, further supporting the representation of these altered genes as a “generic machinery” promoting cancer. Yegnasubramanian describes alterations in epigenetic reprogramming to be almost universal in human cancers [7]. However, it is clear that African derived tumors present with many more (ancestry-specific) possible cancer drivers than do European derived tumors, highlighting the diversity by which epigenetic dysregulation and consequent tumorigenesis may arise in Africans.



**Figure 4.** Top epigenetic machinery genes somatically altered per epigenetic process group, in African and European derived tumors, that may be instrumental in epigenetic dysregulation and consequent prostate cancer oncogenesis. The total number of genes in each epigenetic process group is also displayed. Top genes were selected based on iClusterBayes feature selection with posterior probability > 0.5 as well as presence of potentially damaging variants based on functional impact prediction and/or recurrence. No top genes were identified for epigenetic process group 3, epigenetic process group.

## 5. Conclusions

Alterations to epigenetic machinery components dysregulate epigenetic programming, chromatin structure and consequent transcription, a feature of PCa development and progression that is becoming increasingly understood. Here we describe somatic alterations within the epigenetic machinery and their relevance to PCa health disparities, with African derived tumors demonstrating

a longer tail of African-specific epigenetic driver gene candidates, a number of which are novel to PCa (*BRD1*, *DR1*, *ELL*, *ELP2*, *ELP5*, *GATAD2B*, *MED26*, *MXD1*, *NUP35*, *RAI1*, *SP110*, *TADA2B*) and are putative therapeutic targets (*CHD7*, *DPF3*, *POLR1B*, *SETD1B*, *UBTF*, *VPS72*). Here we also described two epigenetic PCa taxonomies (ECS and EcnCS) that differentiate patients by ancestry, predict clinical outcomes, resemble whole genome derived global subtypes and identify more African-specific putative drivers, ultimately indicating a significant role for epigenetic mechanisms in PCa development. Identifying many more African-specific (*versus* European-specific) potentially novel PCa drivers highlights the urgency for African inclusion in precision medicine-informed healthcare approaches to ultimately reduce PCa health disparities and improve health outcomes for African men.

**Supplementary Materials:** The following supporting information can be downloaded at the website of this paper posted on Preprints.org. Figure S1: Optimal cluster number identification; Figure S2: Consensus heatmap for variant data overlapping epigenetic machinery genes based on results from ten multi-omics integrative clustering algorithms with the assigned cluster number of (A)  $k = 3$  and (B)  $k = 8$ ; Figure S3: Silhouette plot quantifying sample similarity based on results from ten multi-omics integrative clustering algorithms with the assigned cluster number of (A)  $k = 3$  and (B)  $k = 8$ ; Figure S4: Mutational burden in African and European derived tumors; Figure S5: Damaging variant mutational burden in African and European derived tumors; Table S1: Patient summary for African and European study participants; Table S2: SuperPaths and their associated pathways included in this study for relation to epigenetic processes; Table S3: List of genes assigned to Epigenetic Process Group 1 (chromatin organization and regulation); Table S4: List of genes assigned to Epigenetic Process Group 2 (histone modifications); Table S5: List of genes assigned to Epigenetic Process Group 3 (DNA methylation); Table S6: List of genes assigned to Epigenetic Process Group 4 (RNA regulation); Table S7: List of genes assigned to Epigenetic Process Group 5 (epigenetic regulation of gene expression); Table S8: MOVICS clustering results; Table S9: Statistical summary for tumour mutational burden (per Mb) based on all coding variants in epigenetic machinery genes in African and European derived tumors; Table S10: Statistical summary for tumour mutational burden (per Mb) based only on damaging variants (as per functional impact prediction) in epigenetic machinery genes in African and European derived tumors; Table S11: Independent test between epigenetic cancer subtype (ECS) and small somatic mutation to compare mutation frequency; Table S12: Independent test between epigenetic cancer subtype (ECS) and structural variation to compare structural variation frequency; Table S13: Clinical summary based on hierarchical clustering results, with epigenetic cancer subtype (ECS) as the grouping variable; Table S14: Top features, posterior probability and rank order for joint analysis of small somatic mutation, somatic structural variant and somatic copy number alteration data identified by iClusterBayes; Table S15: Clinical summary based on hierarchical clustering results for somatic copy number alteration data only, with epigenetic copy number cancer subtype (EcnCS) as the grouping variable.

**Author Contributions:** Conceptualization, V.M.H.; methodology and software, J.C., J.J., W.J., and V.M.H.; formal analysis, J.C.; investigation, J.C.; resources, S.B.A.M., P.D.S., and M.S.R.B.; data curation, J.J.; writing—original draft preparation, J.C.; writing—review and editing, all authors; visualization, J.C.; study supervision, W.J., and V.M.H.; project administration, S.M.P., M.S.R.B. and V.M.H.; funding acquisition, M.S.R.B. and V.M.H. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by grants from the U.S.A. Congressionally Directed Medical Research Programs (CDMRP) Prostate Cancer Research Program (PCRP) Idea Development Award (PC200390, TARGET Africa Project) and Health Equity Research Outcomes Integrity Consortium (HEROIC) Award (PC210168, HEROIC PCaPH Africa1K), the National Health and Medical Research Council (NHMRC) of Australia Ideas Grant (APP2001098) and by a Cancer Association of South Africa (CANSa) Development Grant. Craddock is supported by the National Research Foundation of South Africa (PMDS22070633683), and Hayes by the Petre Foundation, Australia.

**Institutional Review Board Statement:** The study was conducted in accordance with the Declaration of Helsinki, and approved by the University of Pretoria Faculty of Health Sciences Research Ethics Committee (with US Federal wide assurance FWA00002567 and IRB00002235 IORG0001762) in South Africa for SAPCS (approval 43/2010) and SAPCS PhD study (approval 504/2022), while further study approval for genomic analyses was granted by the St. Vincent's Hospital Human Research Ethics Committee in Sydney, Australia (HREC approval SVH/15/227) and the Human Research Protection Office (HRPO) of the US Army Medical Research and Development Command as part of TARGET Africa (approval E02371) and HEROIC PCaPH Africa1K (approval E03280).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** Data used in this study was published by Jaratlerdsiri et al., 2022 and made accessible via the European Genome-Phenome Archive (EGA; <https://ega-archive.org>) under study accession EGAS00001006425 and dataset accession EGAD00001009067 (Southern African Prostate Cancer Study, SAPCS) and EGAD00001009066 (Garvan/St Vincent's Prostate Cancer Study).

**Acknowledgments:** The authors are grateful to the study participants and their families for their contribution to this study, as well as the many clinical contributors over the years to the Southern African Prostate Cancer Study (SAPCS) and the St Vincent's Garvan Institute of Medical Research Prostate Cancer Biobank. We also acknowledge the Sydney Informatics Hub for computational infrastructure support and data storage.

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

## References

- Fahrner, J.A.; Bjornsson, H.T. Mendelian disorders of the epigenetic machinery: postnatal malleability and therapeutic prospects. *Hum. Mol. Genet.* **2019**, *28*, 2–254.
- Nilsson, E.E.; Sadler-Riggelman, I.; Skinner, M.K. Environmentally induced epigenetic transgenerational inheritance of disease. *Environ. Epigenet.* **2018**, *4*, 1–13.
- Lee, W.H.; Morton, R.A.; Epstein, J.I.; Brooks, J.D.; Campbell, P.A.; Bova, G.S.; Hsieh, W.S.; Isaacs, W.B.; Nelson, W.G. Cytidine methylation of regulatory sequences near the pi-class glutathione S-transferase gene accompanies human prostatic carcinogenesis. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 11733–11737.
- Martignano, F.; Gurioli, G.; Salvi, S.; Calistri, D.; Costantini, M.; Gunelli, R.; de Giorgi, U.; Foca, F.; Casadio, V. GSTP1 methylation and protein expression in prostate cancer: diagnostic implications. *Dis. Markers* **2016**, *2016*.
- Wang, Q.; Carroll, J.S.; Brown, M. Spatial and temporal recruitment of androgen receptor and its coactivators involves chromosomal looping and polymerase tracking. *Mol. Cell* **2005**, *19*, 631–642.
- Wissmann, M.; Yin, N.; Müller, J.M.; Greschik, H.; Günther, T.; Metzger, E.; Schüle, R.; Fodor, B.D.; Jenuwein, T.; Vogler, C.; et al. Cooperative demethylation by JMJD2C and LSD1 promotes androgen receptor-dependent gene expression. *Nat. Cell Biol.* **2007**, *9*, 347–353.
- Yegnasubramanian, S.; de Marzo, A.M.; Nelson, W.G. Prostate cancer epigenetics: from basic mechanisms to clinical implications. *Cold Spring Harb. Perspect. Med.* **2019**, *9*, a030445.
- Armenia, J.; Wankowicz, S.A.M.; Liu, D.; Gao, J.; Kundra, R.; Reznik, E.; Chatila, W.K.; Chakravarty, D.; Han, G.C.; Coleman, I.; et al. The long tail of oncogenic drivers in prostate cancer. *Nat. Genet.* **2018**, *50*, 645–651.
- Jones, S.; Li, M.; Williams Parsons, D.; Zhang, X.; Wesseling, J.; Kristel, P.; Schmidt, M.K.; Markowitz, S.; Yan, H.; Bigner, D.; et al. Somatic mutations in the chromatin remodeling gene ARID1A occur in several tumor types. *Hum. Mutat.* **2012**, *33*, 100–103.
- Zhang, Y.; Yan, L.; Yao, W.; Chen, K.; Xu, H.; Ye, Z. Integrated analysis of genetic abnormalities of the histone lysine methyltransferases in prostate cancer. *Med. Sci. Monit.* **2019**, *25*, 193.
- Grasso, C.S.; Wu, Y.M.; Robinson, D.R.; Cao, X.; Dhanasekaran, S.M.; Khan, A.P.; Quist, M.J.; Jing, X.; Lonigro, R.J.; Brenner, J.C.; et al. The mutational landscape of lethal castration-resistant prostate cancer. *Nature* **2012**, *487*, 239–243.
- Abeshouse, A.; Ahn, J.; Akbani, R.; Ally, A.; Amin, S.; Andry, C.D.; Annala, M.; Aprikian, A.; Armenia, J.; Arora, A.; et al. The molecular taxonomy of primary prostate cancer. *Cell* **2015**, *163*, 1011–1025.
- Morano, A.; Angrisano, T.; Russo, G.; Landi, R.; Pezone, A.; Bartollino, S.; Zuchegna, C.; Babbio, F.; Bonapace, I.M.; Allen, B.; et al. Targeted DNA methylation by homology-directed repair in mammalian cells. Transcription reshapes methylation on the repaired gene. *Nucleic Acids Res.* **2014**, *42*, 804–821.
- Russo, G.; Landi, R.; Pezone, A.; Morano, A.; Zuchegna, C.; Romano, A.; Muller, M.T.; Gottesman, M.E.; Porcellini, A.; Avvedimento, E.V. DNA damage and repair modify DNA methylation and chromatin domain of the targeted locus: mechanism of allele methylation polymorphism. *Sci. Rep.* **2016**, *6*, 1–14.
- Allen, B.; Pezone, A.; Porcellini, A.; Muller, M.T.; Masternak, M.M. Non-homologous end joining induced alterations in DNA methylation: a source of permanent epigenetic change. *Oncotarget* **2017**, *8*, 40359.
- Zhang, Y.; Yang, L.; Kucherlapati, M.; Hadjipanayis, A.; Pantazi, A.; Bristow, C.A.; Lee, E.A.; Mahadeshwar, H.S.; Tang, J.; Zhang, J.; et al. Global impact of somatic structural variation on the DNA methylome of human cancers. *Genome Biol.* **2019**, *20*, 1–24.
- Sun, W.; Bunn, P.; Jin, C.; Little, P.; Zhabotynsky, V.; Perou, C.M.; Hayes, D.N.; Chen, M.; Lin, D.Y. The association between copy number aberration, DNA methylation and gene expression in tumor samples. *Nucleic Acids Res.* **2018**, *46*, 3009–3018.
- Siegel, R. L.; Miller, K. D.; Wagle, N. S.; Jemal, A. Cancer statistics, 2023. *CA Cancer J. Clin.* **2023**, *73*, 17–48.

19. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **2021**, *71*, 209–249.
20. Gheybi, K.; Jiang, J.; Mutambirwa, S.B.A.; Soh, P.X.Y.; Kote-Jarai, Z.; Jaratlerdsiri, W.; Eeles, R.A.; Bornman, M.S.R.; Hayes, V.M. Evaluating germline testing panels in southern African males with advanced prostate cancer. *J. Natl. Compr. Canc. Netw.* **2023**, *21*, 289–296.
21. Conti, D.V.; Darst, B.F.; Moss, L.C.; Saunders, E.J.; Sheng, X.; Chou, A.; Schumacher, F.R.; Olama, A.A.A.; Benlloch, S.; Dadaev, T.; et al. Trans-ancestry genome-wide association meta-analysis of prostate cancer identifies new susceptibility loci and informs genetic risk prediction. *Nat. Genet.* **2021**, *53*, 65–75.
22. Lowder, D.; Rizwan, K.; McColl, C.; Paparella, A.; Ittmann, M.; Mitsiades, N.; Kaochar, S. Racial disparities in prostate cancer: a complex interplay between socioeconomic inequities and genomics. *Cancer Lett.* **2022**, *531*, 71–82.
23. Popejoy, A.B.; Fullerton, S.M. Genomics is failing on diversity. *Nature* **2016**, *538*, 161–164.
24. Hayes, V.M.; Gong, T.; Mutambirwa, S.B.A.; Jaratlerdsiri, W.; Bornman, M.S.R. African inclusion in prostate cancer genomic studies provides the first glimpses into addressing health disparities through tailored clinical care. *Clin. Transl. Med.* **2023**, *13*.
25. Enokida, H.; Shiina, H.; Urakami, S.; Igawa, M.; Ogishima, T.; Pookot, D.; Li, L.C.; Tabatabai, Z.L.; Kawahara, M.; Nakagawa, M.; et al. Ethnic group-related differences in CpG hypermethylation of the GSTP1 gene promoter among African-American, Caucasian and Asian patients with prostate cancer. *Int. J. Cancer* **2005**, *116*, 174–181.
26. Kwabi-Addo, B.; Wang, S.; Chung, W.; Jelinek, J.; Patierno, S.R.; Wang, B.D.; Andrawis, R.; Lee, N.H.; Apprey, V.; Issa, J.P.; et al. Identification of differentially methylated genes in normal prostate tissues from African American and Caucasian men. *Clin. Cancer Res.* **2010**, *16*, 3539–3547.
27. Devaney, J.M.; Wang, S.; Furbert-Harris, P.; Apprey, V.; Ittmann, M.; Wang, B.D.; Olender, J.; Lee, N.H.; Kwabi-Addo, B. Genome-wide differentially methylated genes in prostate cancer tissues from African-American and Caucasian men. *Epigenetics* **2015**, *10*, 319–328.
28. Rubicz, R.; Zhao, S.; Geybels, M.; Wright, J.L.; Kolb, S.; Klotzle, B.; Bibikova, M.; Lance, R.; Ostrander, E.A.; Feng, Z.; et al. DNA methylation profiles in African American prostate cancer patients in relation to disease progression. *Genomics* **2016**, *111*, 10–16.
29. Woodson, K.; Hayes, R.; Wideroff, L.; Villaruz, L.; Tangrea, J. Hypermethylation of GSTP1, CD44, and E-cadherin genes in prostate cancer among US Blacks and Whites. *The Prostate* **2003**, *55*, 199–205.
30. Tang, D.; Kryvenko, O.N.; Mitrache, N.; Do, K.C.; Jankowski, M.; Chitale, D.A.; Trudeau, S.; Rundle, A.; Belinsky, S.A.; Rybicki, B.A. Methylation of the RARB gene increases prostate cancer risk in black Americans. *J. Urol.* **2013**, *190*, 317–324.
31. Jaratlerdsiri, W.; Jiang, J.; Gong, T.; Patrick, S.M.; Willet, C.; Chew, T.; Lyons, R.J.; Haynes, A.M.; Pasqualim, G.; Louw, M.; et al. African-specific molecular taxonomy of prostate cancer. *Nature* **2022**, *609*, 552–559.
32. Raj, A.; Stephens, M.; Pritchard, J.K. fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics* **2014**, *197*, 573–589.
33. Belinky, F.; Nativ, N.; Stelzer, G.; Zimmerman, S.; Stein, T.I.; Safran, M.; Lancet, D. PathCards: multi-source consolidation of human biological pathways. *Database* **2015**, *2015*.
34. Roy, D.M.; Walsh, L.A.; Chan, T.A. Driver mutations of cancer epigenomes. *Protein Cell* **2014**, *5*, 265–296.
35. Guo, M.; Peng, Y.; Gao, A.; Du, C.; Herman, J.G. Epigenetic heterogeneity in cancer. *Biomark. Res.* **2019**, *7*, 1–19.
36. Fahrner, J.A.; Bjornsson, H.T. Mendelian disorders of the epigenetic machinery: tipping the balance of chromatin states. *Annu. Rev. Genom. Hum. Genet.* **2014**, *15*, 269–293.
37. Cunningham, F.; Allen, J.E.; Allen, J.; Alvarez-Jarreta, J.; Amode, M.R.; Armean, I.M.; Austine-Orimoloye, O.; Azov, A.G.; Barnes, I.; Bennett, R.; et al. Ensembl 2022. *Nucleic Acids Res.* **2022**, *50*, D988–D995.
38. Kinsella, R.J.; Kähäri, A.; Haider, S.; Zamora, J.; Proctor, G.; Spudich, G.; Almeida-King, J.; Staines, D.; Derwent, P.; Kerhornou, A.; et al. Ensembl BioMart: a hub for data retrieval across taxonomic space. *Database* **2011**, *2011*.
39. Vaser, R.; Adusumalli, S.; Leng, S.N.; Sikic, M.; Ng, P.C. SIFT missense predictions for genomes. *Nat. Protoc.* **2016**, *11*, 1–9.
40. Adzhubei, I.; Jordan, D.M.; Sunyaev, S.R. Predicting functional effect of human missense mutations using PolyPhen-2. *Curr. Protoc. Hum. Genet.* **2013**, *76*, 7–20.
41. Oscanoa, J.; Sivapalan, L.; Gadaleta, E.; Dayem Ullah, A.Z.; Lemoine, N.R.; Chelala, C. SNPnexus: a web server for functional annotation of human genome sequence variation (2020 update). *Nucleic Acids Res.* **2020**, *48*, W185–W192.
42. Hua, X.; Xu, H.; Yang, Y.; Zhu, J.; Liu, P.; Lu, Y. DrGaP: a powerful tool for identifying driver genes and pathways in cancer sequencing studies. *Am. J. Hum. Genet.* **2013**, *93*, 439–451.
43. Pleasance, E.D.; Pan-cancer analysis of whole genomes. *Nature* **2020**, *578*, 82–93.

44. Mayakonda, A.; Lin, D.C.; Assenov, Y.; Plass, C.; Koeffler, H.P. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res.* **2018**, *28*, 1747–1756.
45. Lu, X.; Meng, J.; Zhou, Y.; Jiang, L.; Yan, F. MOVICS: an R package for multi-omics integration and visualization in cancer subtyping. *Bioinformatics* **2020**, *36*, 5539–5541.
46. Monti, S.; Tamayo, P.; Mesirov, J.; Golub, T. Consensus clustering: a resampling-based method for class discovery and visualization of gene expression microarray data. *Mach. Learn.* **2003**, *52*, 91–118.
47. Mo, Q.; Shen, R.; Guo, C.; Vannucci, M.; Chan, K.S.; Hilsenbeck, S.G. A fully Bayesian latent variable model for integrative clustering analysis of multi-type omics data. *Biostatistics* **2018**, *19*, 71–86.
48. Gu, Z.; Eils, R.; Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **2016**, *32*, 2847–2849.
49. Kolde, R. pheatmap: pretty heatmaps. R package version 1.0.12. **2019**.
50. Jaratlerdsiri, W.; Chan, E.K.F.; Gong, T.; Petersen, D.C.; Kalsbeek, A.M.F.; Venter, P.A.; Stricker, P.D.; Bornman, M.S.R.; Hayes, V.M. Whole-genome sequencing reveals elevated tumor mutational burden and initiating driver mutations in African men with treatment-naïve, high-risk prostate cancer. *Cancer Res.* **2018**, *78*, 6736–6746.
51. Fagan, R.J.; Dingwall, A.K. COMPASS ascending: emerging clues regarding the roles of MLL3/KMT2C and MLL2/KMT2D proteins in cancer. *Cancer Lett.* **2019**, *458*, 56–65.
52. Grypari, I.M.; Logotheti, S.; Zolota, V.; Troncoso, P.; Efstathiou, E.; Bravou, V.; Melachrinou, M.; Logothetis, C.; Tzelepi, V.; Zhang, H. The protein arginine methyltransferases (PRMTs) PRMT1 and CARM1 as candidate epigenetic drivers in prostate cancer progression. *Medicine* **2021**, *100*.
53. Fan, H.T.; Shi, Y.Y.; Lin, Y.; Yang, X.P. EHMT2 promotes the development of prostate cancer by inhibiting PI3K/AKT/mTOR pathway. *Eur. Rev. Med. Pharmacol. Sci.* **2019**, *23*, 7808–7815.
54. Burleson, M.; Deng, J.J.; Qin, T.; Duong, T.M.; Yan, Y.; Gu, X.; Das, D.; Easley, A.; Liss, M.A.; Renee Yew, P.; et al. GLI3 is stabilized by SPOP mutations and promotes castration resistance via functional cooperation with androgen receptor in prostate cancer. *Mol. Cancer Res.* **2022**, *20*, 62–76.
55. Burdelski, C.; Ruge, O.M.; Melling, N.; Koop, C.; Simon, R.; Steurer, S.; Sauter, G.; Kluth, M.; Hube-Magg, C.; Minner, S.; et al. HDAC1 overexpression independently predicts biochemical recurrence and is associated with rapid tumor cell proliferation and genomic instability in prostate cancer. *Exp. Mol. Pathol.* **2015**, *98*, 419–426.
56. Cao, Z.; Shi, X.; Tian, F.; Fang, Y.; Wu, J.B.; Mrdenovic, S.; Nian, X.; Ji, J.; Xu, H.; Kong, C.; et al. KDM6B is an androgen regulated gene and plays oncogenic roles by demethylating H3K27me3 at cyclin D1 promoter in prostate cancer. *Cell Death & Disease* **2021**, *12*, 2.
57. Xie, H.; Zhao, J.; Wan, J.; Zhao, J.; Wang, Q.; Yang, X.; Yang, W.; Lin, P.; Yu, X. Long non-coding RNA AC245100.4 promotes prostate cancer tumorigenesis via the microRNA-145-5p/RBBP5 axis. *Oncol. Rep.* **2021**, *45*, 619–629.
58. Chang, Y.T.; Lin, T.P.; Campbell, M.; Pan, C.C.; Lee, S.H.; Lee, H.C.; Yang, M.H.; Kung, H.J.; Chang, P.C. REST is a crucial regulator for acquiring EMT-like and stemness phenotypes in hormone-refractory prostate cancer. *Sci. Rep.* **2017**, *7*, 42795.
59. Zhang, H.M.; Qiao, Q.D.; Xie, H.F.; Wei, J.X. Breast cancer metastasis suppressor 1 (BRMS1) suppresses prostate cancer progression by inducing apoptosis and regulating invasion. *Eur. Rev. Med. Pharmacol. Sci.* **2017**, *21*, 68–75.
60. Zhu, S.; Xu, Y.; Song, M.; Chen, G.; Wang, H.; Zhao, Y.; Wang, Z.; Li, F. PRDM16 is associated with evasion of apoptosis by prostatic cancer cells according to RNA interference screening. *Mol. Med. Rep.* **2016**, *14*, 3357–3361.
61. Goll, M.G.; Bestor, T.H. Eukaryotic cytosine methyltransferases. *Annu. Rev. Biochem.* **2005**, *74*, 481–514.
62. Tao, Q.; Huang, H.; Geiman, T.M.; Lim, C.Y.; Fu, L.; Qiu, G.H.; Robertson, K.D. Defective de novo methylation of viral and cellular DNA sequences in ICF syndrome cells. *Hum. Mol. Genet.* **2002**, *11*, 2091–2102.
63. Yang, X.; Han, H.; De Carvalho, D.D.; Lay, F.D.; Jones, P.A.; Liang, G. Gene body methylation can alter gene expression and is a therapeutic target in cancer. *Cancer Cell* **2014**, *26*, 577–590.
64. Singal, R.; Das, P.; Manoharan, M.; Reis, I.; Schlesselman, J. Polymorphisms in the DNA methyltransferase 3b gene and prostate cancer risk. *Oncol. Rep.* **2005**, *14*, 569–573.
65. Gravina, G.L.; Ranieri, G.; Muzi, P.; Marampon, F.; Mancini, A.; Di Pasquale, B.; Di Clemente, L.; Dolo, V.; D'Alessandro, A.M.; Festuccia, C. Increased levels of DNA methyltransferases are associated with the tumorigenic capacity of prostate cancer cells. *Oncol. Rep.* **2013**, *29*, 1189–1195.
66. Reddy Parine, N.; Alanazi, I.O.; Shaik, J.P.; Aldhaian, S.; Aljebreen, A.M.; Alharbi, O.; Almadi, M.A.; Azzam, N.A.; Alanazi, M.; Yang, Q. TDG gene polymorphisms and their possible association with colorectal cancer: a case control study. *J. Oncol.* **2019**, 2019.

67. Ruczinski, I.; Jorgensen, T.J.; Shugart, Y.Y.; Schaad, Y.B.; Kessing, B.; Hoffman-Bolton, J.; Helzlsouer, K.J.; Kao, W.H.L.; Wheless, L.; Francis, L.; et al. A population-based study of DNA repair gene variants in relation to non-melanoma skin cancer as a marker of a cancer-prone phenotype. *Carcinogenesis* **2012**, *33*, 1692–1698.
68. Xu, X.; Yu, T.; Shi, J.; Chen, X.; Zhang, W.; Lin, T.; Liu, Z.; Wang, Y.; Zeng, Z.; Wang, C.; et al. Thymine DNA glycosylase is a positive regulator of Wnt signaling in colorectal cancer. *J. Biol. Chem.* **2014**, *289*, 8881–8890.
69. Chen, P.; Li, S.; Zhang, K.; Zhao, R.; Cui, J.; Zhou, W.; Liu, Y.; Zhang, L.; Cheng, Y. N 6-methyladenosine demethylase ALKBH5 suppresses malignancy of esophageal cancer by regulating microRNA biogenesis and RAI1 expression. *Oncogene* **2021**, *40*, 5600–5612.
70. Meeks, J.J.; Shilatifard, A. Multiple roles for the MLL/COMPASS family in the epigenetic regulation of gene expression and in cancer. *Annu. Rev. Cancer Biol.* **2017**, *1*, 425–446.
71. Choi, Y.J.; Oh, H.R.; Choi, M.R.; Gwak, M.; An, C.H.; Chung, Y.J.; Yoo, N.J.; Lee, S.H. Frameshift mutation of a histone methylation-related gene SETD1B and its regional heterogeneity in gastric and colorectal cancers with high microsatellite instability. *Hum. Pathol.* **2014**, *45*, 1674–1681.
72. Fowler, T.; Ghatak, P.; Price, D.H.; Conaway, R.; Conaway, J.; Chiang, C.M.; Bradner, J.E.; Shilatifard, A.; Roy, A.L. Regulation of MYC expression and differential JQ1 sensitivity in cancer cells. *PLoS One* **2014**, *9*, e87003.
73. Tan, W.; Peng, S.; Li, Z.; Zhang, R.; Xiao, Y.; Chen, X.; Zhu, J.; Li, B.; Lv, X. Identification of therapeutic targets and prognostic biomarkers among genes from the mediator complex family in the hepatocellular carcinoma tumour-immune microenvironment. *Comput. Math. Methods Med.* **2022**, *2022*.
74. Zhang, X.; Zhou, Y.; Shi, Z.; Liu, Z.; Chen, H.; Wang, X.; Cheng, Y.; Xi, L.; Li, X.; Zhang, C.; et al. Integrated analysis of genes encoding ATP-dependent chromatin remodellers identifies CHD7 as a potential target for colorectal cancer therapy. *Clin. Transl. Med.* **2022**, *12*, e953.
75. Yang, F.; Liu, H.; Zhao, J.; Ma, X.; Qi, W. POLR1B is upregulated and promotes cell proliferation in non-small cell lung cancer. *Oncol. Lett.* **2020**, *19*, 671–680.
76. Chen, T.; Tu, Y.; Lv, D.; Lin, K.; Tang, H.; Huang, W. Vacuolar protein sorting-associated protein 72 homolog (VPS72) binding to lysine acetyltransferase 5 (KAT5) promotes the proliferation, invasion and migration of hepatocellular carcinoma through regulating phosphatidylinositol 3-kinase (PI3K)/protein kinase B (AKT) signaling pathway. *Bioengineered* **2022**, *13*, 9197–9210.
77. Zhang, J.; Zhang, J.; Liu, W.; Ge, R.; Gao, T.; Tian, Q.; Mu, X.; Zhao, L.; Li, X. UBTF facilitates melanoma progression via modulating MEK1/2-ERK1/2 signalling pathways by promoting GIT1 transcription. *Cancer Cell Int.* **2021**, *21*, 1–16.
78. Lin, W.; Dai, W.; Xu, X.; Yu, Q.; Zhang, B.; Li, J.; Li, H. Downregulation of DPF3 promotes the proliferation and motility of breast cancer cells through activating JAK2/STAT3 signaling. *Biochem. Biophys. Res. Commun.* **2019**, *514*, 639–644.
79. Chen, D.; Li, T.; Wang, C.; Lei, G.; Wang, R.; Wang, Z.; Yu, L.; Yan, J.; Zhang, P.; Wang, X.; et al. High-level SETD1B gene expression is associated with unfavorable prognosis in hepatocellular carcinoma. *Mol. Med. Rep.* **2019**, *19*, 1587–1594.
80. Hu, Z.; Zhou, J.; Jiang, J.; Yuan, J.; Zhang, Y.; Wei, X.; Loo, N.; Wang, Y.; Pan, Y.; Zhang, T.; et al. Genomic characterization of genes encoding histone acetylation modulator proteins identifies therapeutic targets for cancer treatment. *Nat. Commun.* **2019**, *10*, 733.
81. Moradi Marjaneh, R.; Hassanian, S.M.; Ghobadi, N.; Ferns, G.A.; Karimi, A.; Jazayeri, M.H.; Nasiri, M.; Avan, A.; Khazaei, M. Targeting the death receptor signaling pathway as a potential therapeutic target in the treatment of colorectal cancer. *J. Cell. Physiol.* **2018**, *233*, 6538–6549.
82. Grzeskowiak, C.L.; Kundu, S.T.; Mo, X.; Ivanov, A.A.; Zagrodna, O.; Lu, H.; Chapple, R.H.; Tsang, Y.H.; Moreno, D.; Mosqueda, M.; et al. In vivo screening identifies GATAD2B as a metastasis driver in KRAS-driven lung cancer. *Nat. Commun.* **2018**, *9*, 2732.
83. Korakiti, A.M.; Moutafi, M.; Zografos, E.; Dimopoulos, M.A.; Zagouri, F. The genomic profile of pregnancy-associated breast cancer: a systematic review. *Front. Oncol.* **2020**, *10*, 1773.
84. Zhao, Z.; Sentürk, N.; Song, C.; Grummt, I. lncRNA PAPAS tethered to the rDNA enhancer recruits hypophosphorylated CHD4/NuRD to repress rRNA synthesis at elevated temperatures. *Genes Dev.* **2018**, *32*, 836–848.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.