Integrative Analysis of Methylation and Copy Number Variations of Prostate Adenocarcinoma Based on Weighted Gene Co-expression Network Analysis

Yaxin Hou 1,2†, Junyi Hu 1,2†, Lijie Zhou 1,2, Lilong Liu 1,2, Ke Chen 1,2* and Xiong Yang 1*

1 Department of Urology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, 2 Shenzhen Huazhong University of Science and Technology Research Institute, Shenzhen, China

Prostate adenocarcinoma (PRAD) is the most pervasive carcinoma diagnosed in men with over 170,000 new cases every year in the United States and is the second leading cause of death from cancer in men despite its indolent clinical course. Prostate-specific antigen testing, which is the most commonly used non-invasive diagnostic method for PRAD, has improved early detection rates in the past decade, but its effectiveness for monitoring disease progression and predicting prognosis is controversial. To identify novel biomarkers for these purposes, we carried out weighted gene co-expression network analysis of the top 10,000 variant genes in PRAD from The Cancer Genome Atlas in order to identify gene modules associated with clinical outcomes. Methylation and copy number variation analysis were performed to screen aberrantly expressed genes, and the Kaplan–Meier survival and gene set enrichment analyses were conducted to evaluate the prognostic value and potential mechanisms of the identified genes. Cyclin E2 (CCNE2), rhophilin Rho GTPase-binding protein (RHPN1), enhancer of zeste homolog 2 (EZH2), tonsoku-like DNA repair protein (TONSL), epoxide hydrolase 2 (EPHX2), fibromodulin (FMOD), and solute carrier family 7 member (SLC7A4) were identified as potential prognostic indicators and possible therapeutic targets as well. These findings can improve diagnosis and disease monitoring to achieve better clinical outcomes in PRAD.

Keywords: bioinformatics analysis, prostate adenocarcinoma, biomarker, prognosis, therapeutic target

INTRODUCTION

Prostate adenocarcinoma (PRAD) is one of the most common neoplasms worldwide, ranking 4th among all cancer types in both sexes with an incidence of 7.1% (1). In the United States, PRAD is the most prevalent cancer in men and is estimated to have caused more than 30,000 deaths in 2020 (1, 2).

Cancer was previously considered as a genetic disease, but there is considerable evidence that epigenetic changes contribute to tumorigenesis and tumor progression (3–5). DNA methylation is the most widely studied epigenetic modification in both non-neoplastic and neoplastic diseases including PRAD (6). The methylation of CpG islands, which are often located in the gene promoter,
results in transcriptional silencing (7). Recently, methylation of enhancer regions has also been shown to play an important role in regulating gene expression (8, 9). DNA methyltransferase 1 (DNMT1), DNMT3a, and DNMT3b are upregulated in PRAD tissue compared to normal benign prostatic hyperplastic tissue, and their expression is elevated in cancerous tissue with a higher Gleason score, suggesting a close association between epigenetic alterations and PRAD development and progression (10). Additionally, epigenetic marks are potential biomarkers for PRAD (11) and targets for next-generation drugs.

Copy number variations (CNVs) are the most common genetic alteration in cancers, and CNV burden is associated with the rates of recurrence and death in multiple neoplasms (12). E26 transformation-specific (ETS) genes, tumor protein 53 (TP53), phosphatase and tensin homolog (PTEN), and androgen receptor (AR) are the most frequently altered genes in primary prostate cancer, which leads to dysregulation of phosphoinositide 3-kinase (PI3K)/protein kinase B (AKT), RAS/RAF, and cell cycle signaling pathways; moreover, alterations in AR and TP53 have been linked to castration resistance (13, 14) and worse outcomes (15). Thus, CNVs have prognostic value in PRAD as they can reflect disease progression.

The development and progression of cancers involve gene–gene interactions within a gene co-expression network. In this study, we carried out weighted gene co-expression network analysis (WGCNA) (16) to identify genes associated with clinical outcomes in PRAD and can thus serve as biomarkers. We also investigated CNV and methylation status of genes in key module of the network and assessed their prognostic value for PRAD.

**MATERIALS AND METHODS**

**Data Acquisition**
The expression data matrix of The Cancer Genome Atlas (TCGA) PRAD database comprising 497 tumor and 52 normal tissue samples along with CNVs, DNA methylation, and clinical
Identification of Co-expression Module

Unlike ordinary clustering analysis, clustering criteria of WGCNA have biological significance, so the results obtained by this method have higher credibility. WGCNA clusters genes with similar expression patterns into a module and allows analysis of correlations between module and sample features. In this study, WGCNA was carried out to identify gene module closely related to clinical outcomes in PRAD. To minimize computational burden, the top 10,000 genes with the largest variance were selected. The topological overlap matrix (TOM) was performed to measure the correlation between genes and detection of module, which was able to identify not only the similarity of expression between gene A and gene C, but also the effect of gene A on gene C via gene B. A height of 220 in the sample cluster was used to detect outliers, with two outliers as filters. A power $\beta$ of 8, minimal module size of 30, and branch merge cutoff height of 0.25 were used as the criteria for module construction.

Copy Number Variation Analysis

The TCGA PRAD CNV profiles were originally measured using whole genome microarray at a TCGA genome characterization center, and GISTIC2 method was then conducted to acquire the estimated values to $-2$, $-1$, 0, 1, 2, respectively, representing homozygous deletion, single copy deletion, diploid normal copy, low-level copy number amplification, and high-level copy number amplification (17). The processed data was obtained from https://xenabrowser.net/. In addition, GISTIC2 was conducted to assess the possibility of CNV events in specific chromosomal regions. Genes with changes in frequency $>10\%$ were selected for further analysis. We calculated the Spearman correlation coefficient ($r$) between CNVs and gene expression levels, with $r > 0.4$ as the cutoff value, indicating the significant impact on gene expression of CNV.

Methylation Analysis

The DNA methylation profiles of PRAD from TCGA were available at the University of California, Santa Cruz (UCSC) Xena browser (https://xenabrowser.net/), which were measured experimentally based on the Illumina Infinium HumanMethylation450 platform (Illumina, San Diego, CA, USA). DNA methylation values ($\beta$ values, between 0 and 1) were recorded for every array probe in each sample by virtue of BeadStudio software (Illumina, San Diego, CA, USA), representing the ratio of the intensity of the methylated bead type to the combined locus intensity. The level of methylation evaluated by $\beta$ values were derived from the Johns Hopkins University and University of Southern California TCGA genome characterization center.

Gene Set Enrichment Analysis and Protein–Protein Interaction Network Analysis

Gene Set Enrichment Analysis (GSEA) (18) is a computational method used to determine whether a predefined set of genes can show significant differences between two biological
FIGURE 3 | The level of gene copy number in PRAD samples. The estimated values –2, –1, 0, 1, 2, respectively representing homozygous deletion, single copy deletion, diploid normal copy, low-level copy number amplification, and high-level copy number amplification. The horizontal axis represents PRAD tumor samples in TCGA, whereas the vertical axis represents genes from M7 module with CNV > 10%.

Statuses, which were performed by the GSEA software obtained from http://www.broad.mit.edu/gsea to assess the enrichment of identified genes with distinct CNVs and methylation levels in PRAD, with false discovery rate (FDR) < 25% and nominal p < 0.05 as the cutoff values. Protein–protein interaction (PPI) network analysis of identified genes was completed by an online tool available at https://string-db.org/ to assess possible interactions between their expression products.

Survival and Statistical Analyses
The Kaplan–Meier survival analysis was performed with Prism 7.0 (GraphPad, La Jolla, CA, USA) and the online tool GEPIA (http://gepia.cancer-pku.cn/index.html) (19). Previous study by
| Gene     | pT   | p-value | pN   | p-value |
|----------|------|---------|------|---------|
|          | pT2  | pT3+pT4 |      |         |
| TMEM220  | −1   | 106     | 82   | 96      | 71     | 25  | 0.031 |
|          | 0    | 378     | 216  | <0.001  | 319    | 267 | 52   |
|          | 1    | 150     | 117  | <0.001  | 140    | 103 | 37   | 0.003 |
| SQLE     | 0    | 334     | 181  | 275     | 235    | 40  |
|          | 1    | 150     | 121  | <0.001  | 144    | 107 | 37   | 0.006 |
| RAD54B   | 0    | 330     | 177  | 271     | 231    | 40  |
|          | 1    | 154     | 121  | <0.001  | 144    | 107 | 37   | 0.006 |
| HEXB     | −1   | 89      | 73   | 82      | 61     | 21  | 0.67 |
|          | 0    | 396     | 225  | <0.001  | 333    | 277 | 56   | 0.67 |
| GINS4    | −1   | 160     | 113  | 145     | 113    | 32  |
|          | 0    | 324     | 185  | 0.004   | 270    | 225 | 45   | 0.177 |
| FBXO43   | 0    | 325     | 173  | 266     | 230    | 36  |
|          | 1    | 159     | 125  | <0.001  | 149    | 108 | 41   | <0.001 |
| EZH2     | 0    | 390     | 230  | 332     | 270    | 62  |
|          | 1    | 94      | 68   | 0.017   | 83     | 68  | 15   | 0.900 |
| EPHX2    | −1   | 267     | 194  | 235     | 184    | 51  |
|          | 0    | 217     | 113  | <0.001  | 180    | 154 | 26   | 0.059 |
| DSCC1    | 0    | 335     | 182  | 276     | 235    | 41  |
|          | 1    | 149     | 116  | <0.001  | 139    | 103 | 36   | 0.006 |
| CCNE2    | 0    | 330     | 177  | 271     | 232    | 39  |
|          | 1    | 154     | 121  | <0.001  | 144    | 106 | 38   | 0.003 |
| CBFA2T3  | −1   | 189     | 140  | 169     | 131    | 38  |
|          | 0    | 295     | 158  | <0.001  | 246    | 207 | 39   | 0.088 |
| TBC1D31  | 0    | 334     | 181  | 275     | 234    | 41  |
|          | 1    | 150     | 117  | <0.001  | 140    | 104 | 36   | 0.007 |
| TONSL    | 0    | 350     | 194  | 290     | 246    | 44  |
|          | 1    | 134     | 104  | <0.001  | 125    | 92  | 33   | 0.007 |
| RHPN1    | 0    | 350     | 194  | 290     | 246    | 44  |
|          | 1    | 134     | 104  | <0.001  | 125    | 92  | 33   | 0.007 |

*−1* for deletions and *1* for amplifications.
Li et al. has established a prognostic model and verified with independent datasets after establishing a prognostic model (20); therefore, we downloaded the independent dataset GSE70769 (21) through the National Center for Biotechnology Information Search database (https://www.ncbi.nlm.nih.gov/) and analyzed the impact of the identified genes on the prognosis of prostate cancer patients. Multivariate analyses were carried out with the cox proportional hazards regression model. All data processing was performed using SPSS v22.0 software (SPSS Inc, Chicago, IL, USA) or R software (x64 3.5.1) (22).

The research process is illustrated in Figure 1.

RESULTS
Identification of Co-expression Module in PRAD
The top 10,000 genes with the largest variations in expression level relative to normal tissue were selected for WGCNA. We generated a module–trait association network with 7 clinicopathologic traits and 17 modules and calculated the
Pearson’s correlation coefficients and p-values to evaluate the relationship between clinical traits and feature vectors of genes in the module. The module with highest correlation coefficient and module size >30 (pink module, M7, p < 0.01) was selected for further analysis (Figure 2).

Copy Number Variation Analysis

After analyzing CNV profiles of TCGA PRAD data and combining the results with pink module (M7) from the WGCNA, we selected 111 genes with a variation frequency >10% and constructed a heatmap of the CN of genes in the PRAD samples (Figure 3), which allowed us to identify those with abnormal CN. Because our aim was to identify prognostic biomarkers for PRAD, we examined the pathologic stage associated with the CN variants. The 14 genes with the highest CNV and the corresponding clinicopathologic stage are shown in Table 1. Of these, nine genes with a Spearman correlation coefficient >0.4 were selected to evaluate the association between gene CN and expression level. Positive correlations were observed for the cyclin E2 (CCNE2), DNA replication and sister chromatid cohesion 1 (DCCCI), rhophilin Rho GTPase-binding protein (RHPN1), enhancer of zeste homolog 2 (EZH2), RAD54B, TBC1 domain family member 31 (TBCID31), and tonsoku-like DNA repair protein (TONSL) genes (p < 0.0001), indicating the amplifications of CN events probably correlated with higher gene expression level. However, epoxide hydrolase 2 (EPHX2) and CBFA2/RUNX1 partner transcriptional co-repressor 3 (CBFA2T3) primarily showed deletions of CN events, which leading to the lower level of gene expression (Figure 4). To macro-evaluate the possibility of CNV events in specific chromosomal regions, the deletion and amplification plots based on G scores for CNV were demonstrated in Supplementary Figure 2. The higher G score of a region represents for the greater probability of CNV events in that region.

The Kaplan–Meier Survival Analysis

We performed the Kaplan–Meier analysis to evaluate the relationship between CNV and disease-free survival (DFS). The survival curve indicated that CNV level was significantly associated with the prognosis of patients with PRAD, with lower CNV predicting longer DFS (p = 0.0001; Figure 5). We analyzed the relationship between CNV of the nine above-mentioned genes and patient prognosis and found that lower CNs of CCNE2 [hazard ratio (HR) = 1.6; p < 0.05], RHPN1 (HR = 2; p < 0.05), EZH2 (HR = 2.2; p < 0.001), and TONSL (HR = 1.7; p < 0.05) were associated with better prognosis, whereas the opposite was true for EPHX2 (HR = 0.47; p < 0.001) (Figure 6). A multivariate analysis of CBFA2T3 CN suggested that it may be a protective factor in PRAD, whereas the Kaplan–Meier survival analysis suggested it was not statistically significant in the prognosis of patients with PRAD (Figure 6i and Table 2). The validation of identified biomarkers for prognosis value revealed the similar results as our former analysis, indicating the explicit prognostic significance of CCNE2, SLC7A4, EZH2, etc. (Supplementary Figures 1B–H).

Gene Set Enrichment Analysis and PPI Network Analysis

To identify enriched gene sets in PRAD samples with high CNV and clarify the mechanisms of CNV in tumorigenesis, we performed GSEA to identify relevant biological pathways in the Kyoto Encyclopedia of Genes and Genomes (KEGG) database and Pathway Interaction Database (PID) using FDR < 25% and p < 0.05 as the criteria for significance. For EZH2, TONSL, and CCNE2, the GSEA curves revealed four enriched gene sets including “KEGG–cell cycle,” “KEGG–P53 signaling pathway,” “PID–ataxia–telangiectasia mutated (ATM) pathway,” and “PID–E2F pathway,” which are mainly related to cell cycle regulation, cell apoptosis, and DNA damage repair. Additionally, for EPHX2, two functional gene sets were enriched—namely, “Cell cycle pathway” and “PID–E2F pathway” (Figure 7). The PPI network analysis found that there was a co-expression between EZH2 and CCNE2, both of which play important roles in regulating cell cycle (Supplementary Figure 1A).

Methylation Analysis

After establishing the co-expression module, the DNA methylation level of the genes was examined, and its correlation with gene expression level was evaluated with the Pearson’s correlation coefficient. Differentially methylated genes with the Pearson’s correlation coefficient >0.4 were identified, including fibromodulin (FMO), transmembrane protein 220 (TMEM220), histone H2B type 1-H (HIST1H2BH), zinc finger 334 (ZNF334), RIC3 acetylcholine receptor chaperone (RIC3), and solute carrier family 7 member 2 (SLC7A4); these genes were
all hypermethylated in tumor samples ($n = 336$) compared to normal tissue ($n = 49$) (Figures 8A, 9A). There was a moderate inverse correlation between gene expression and methylation levels ($r > 0.4$, $p < 0.001$; Figure 8B).

The heatmap of DNA methylation revealed significantly higher levels in tumor tissue compared to normal tissue, especially for SLC7A4. The Kaplan–Meier survival curves showed an association between gene expression level and the prognosis of PRAD for FMOD [HR (high) = 0.37; $p < 0.001$] and SLC7A4 [HR (high) = 0.44; $p < 0.001$], with a higher level corresponding to better prognosis (Figures 9B,E), while others were not statistically significant (Figures 9C,D,F,G). The GSEA curves revealed four gene sets that were enriched, including “KEGG–cell cycle,” “PID–E2F pathway,” “Hallmark–E2F target,” and “Hallmark–G2M checkpoint” (Figures 9H–K). These results indicate that FMOD and SLC7A4 are significant genes related to the clinical outcome of PRAD.

**FIGURE 6** | The Kaplan–Meier analysis of nine genes. (A) CCNE2, (B) RHPN1, (C) EZH2, (D) TONSL, (E) EPHX2, (F) DSCC1, (G) TBC1D31, (H) RAD54B, and (I) CBFA2T3. $p < 0.05$ was considered statistically different. DFS, disease-free survival.
DISCUSSION

The number of new cases of PRAD in the United States has shown an increasing trend in the last 3 years, and PRAD is the second leading cause of death in men despite improvements in diagnostic methods and treatments (2, 23). Although magnetic resonance imaging and some biomarkers are used for the diagnosis of PRAD, the standard approach is tissue biopsy (24), which may only be performed at later stages of the disease when therapeutic options are limited.

Copy number variations occur in 4.8–9.5% of the human genome and play a critical role in tumor recurrence (25); and epigenetic modifications such as DNA methylation are potential biomarkers and targets for treatment in cancer (11). Given the increasing rates of PRAD, there is a need for new diagnostic and prognostic biomarkers with high specificity and sensitivity. In this study, we identified five novel genes with high CNV in PRAD by WGCNA (CCNE2, RHPN1, EZH2, TONSL, and EPHX2) along with two hypermethylated genes (FMOD and SLC7A4) that were significantly correlated with the prognosis of patients with PRAD and may thus be clinically useful biomarkers.

Cyclin E2 encodes cyclin E2, a regulatory subunit of cyclin-dependent kinase 2 (CDK2), which controls cell cycle entry from quiescence. Although the gene encoding the other subunit

---

**TABLE 2** | Multivariate analysis of CBFA2T3 CNV and patient survival.

| Variable          | Multivariate analysis |   |
|-------------------|-----------------------|---|
|                   | HR                    | 95% CI    | p-value |
| T stage (≥ T3)    | 2.361                 | 1.175–4.743 | 0.016   |
| N stage           | 1.363                 | 0.494–3.726 | 0.557   |
| M stage           | 1.254                 | 0.713–2.196 | 0.457   |
| Gleason score (≥ 8) | 3.172             | 1.529–6.578 | 0.002   |
| PSA (≥ 10)        | 1.486                 | 0.869–2.432 | 0.119   |
| CBFA2T3           | 0.424                 | 0.218–0.824 | 0.011   |

HR estimated from Cox proportional hazard regression model; p < 0.05 was considered statistically significant.

DFS, disease-free survival; HR, hazard ratio.

---

**FIGURE 7** | Gene Set Enrichment Analysis (GSEA) curve. (A) E2H2, (B) TONEL, (C) CCNE2, and (D) EPHX-2. KEGG, Kyoto Encyclopedia of Genes and Genomes. ATM, ataxia telangiectasia-mutated.
of CDK2, CCNE1, has been linked to poor prognosis in hepatocellular carcinoma (HCC), there is little known about the role of CCNE2 in tumor progression (26). Cyclin E2 was shown to induce the G1-S transition in PC3 prostate cancer cells (27); our results suggest that it may have a similar function in PRAD, given that a lower CCNE2 CN was associated with longer DFS in patients.

Rhophilin Rho GTPase-binding protein is a Rho GTPase-interacting protein that has not been previously reported in PRAD, but is known to modulate the glomerular filtration barrier and podocyte cytoskeletal architecture (28). The long non-coding RNA RPHN1 antisense RNA 1 (RHPN1-AS1) was found to promote the progression of several tumors, including uveal melanoma, cervical cancer, and HCC (29–31). Our results provide the first demonstration that overexpression of RHPN1 is associated with poor prognosis in PRAD.

Enhancer of zeste homolog 2, the catalytic subunit of the DNA methyltransferase polycomb repressive complex 2 (PRC2), is overexpressed in hormone-refractory metastatic PRAD and may be correlated with disease progression and prognosis (32). Consistent with our findings, one study showed that an elevated level of EZH2 was associated with over proliferation of tumor cells and worse prognosis and may have clinical utility for distinguishing indolent PRAD from aggressive disease with a fatal course (31). On the one hand, the utility of EZH2 as a biomarker has been demonstrated in patients with intractable PRAD (33). On the other hand, EZH2 inhibitors have been linked to carcinogenesis and treatment resistance in clinical trials (34), though the detailed mechanisms underlying these effects remain to be determined.

Tonsoku-like DNA repair protein promotes homologous recombination during DNA repair in a complex with MMS22-like (MMS22L) (35). However, the role of TONSL in prostate cancer is unknown. We found that a high level of TONSL was associated with enrichment of genes related to the ATM and E2F pathways—which mediate DNA repair and negatively regulate the cell cycle—and decreased survival time in patients with PRAD. Thus, TONSL is a potential biomarker for the progression of PRAD.
Epoxide hydrolase 2 functions in arachidonic acid and androgen signaling (36–38) and has been linked to the biosynthesis and metabolism of cholesterol in the regulation of testosterone levels (39). EPHX2 silencing induced apoptosis in PRAD cells and enhanced the antiproliferative effect of flutamide (40). In our study, decreased expression of EPHX2 was associated with the enrichment of genes related to the cell cycle and E2F pathway while patients with an elevated level of EPHX2 had better prognosis, suggesting that EPHX2 has a protective role in PRAD.

Fibromodulin (encoded by FMOD) is thought to be involved in the inhibition of tumorigenesis and apoptosis in hematologic malignancies such as B-cell chronic lymphocytic leukemia and mantle cell lymphoma, like other proteoglycans (40). FMOD was shown to be overexpressed in PRAD cell lines and clinical specimens (41), which is supported by our findings. Our analysis revealed that higher expression of FMOD was associated with...
a better clinical outcome, highlighting its potential utility as a biomarker for monitoring disease progression.

Solute carrier family 7 member is a cationic amino acid transporter of unknown function; SLC7A4 expressed in the plasma membrane was insufficient to drive amino acid transport (42). Our results showed that SLC7A4 methylation was higher in tumor specimens than in normal tissue, and that higher SLC7A4 expression was associated with better clinical outcome. We speculate that SLC7A4 inhibits tumor formation in PRAD through regulation of the cell cycle. Thus, SLC7A4 likely has clinical value for monitoring PRAD progression and predicting prognosis.

CONCLUSION

In this study, we used WGCNA to identify seven genes that are potential prognostic biomarkers for PRAD based on CNV (CCNE2, RHPN1, EZH2, TONSL, and EPHX2) and DNA hypermethylation (FMOD and SLC7A4), all of which can serve as indicators of PRAD progression and potential therapeutic targets for the PRAD treatment as well. However, further experiments are needed to elucidate the precise roles and mechanisms of these candidate biomarkers in PRAD and validate their clinical applicability.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

XY and KC: conception, design of study, and revising the manuscript. YH: drafting the manuscript. JH and LZ: analysis and interpretation of data. YH and LL: acquisition of data. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by grant from National Natural Sciences Foundation of China (82072852 and 81672524), Science, Technology and Innovation Commission of Shenzhen Municipality (JCYJ2018030516488833), and Fundamental Research Funds for the Central University (2011fyRCPY0049).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.647253/full#supplementary-material

Supplementary Figure 1 | Protein–protein interaction (PPI) and validation of prognosis for identified biomarkers. (A) A co-expression between enhancer of zeste homolog 2 (EZH2) and Cyclin E2 (CCNE2), both of which play important roles in regulating cell cycle. (B-H) The prognostic value of identified genes in independent dataset GSE70769.

Supplementary Figure 2 | The copy number variation (CNV) events in specific chromosomal regions. The higher G score represents for the greater probability of CNV events in that region. The previously identified genes were marked in the figure (EZH2: 7q36.1; CCNE2: 8q22.1; RHPN1: 8q24.3; TONSL: 8q24.3; EPHX2: 8p21.2-p21.1).

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. (2018) 68:394–424. doi: 10.3322/caac.21492
2. Siegel RL, Miller KD, Jemal A. Cancer statistics, (2020). CA Cancer J Clin. (2020) 70:7–30. doi: 10.3322/caac.21590
3. Jones PA, Baylin SB. The fundamental role of epigenetic events in cancer. Nat Rev Genet. (2002) 3:415–28. doi: 10.1038/nrg116
4. Berger SL, Kouzarides T, Shiekhattar R, Shilatifard A. An operational definition of epigenetics. Genes Dev. (2009) 23:781–3. doi: 10.1101/gad.1787609
5. Jones PA, Baylin SB. The epigenomics of cancer. Cell. (2007) 128:683–92. doi: 10.1016/j.cell.2007.01.029
6. Keil KP, Vezina CM. DNA methylation as a dynamic regulator of development and disease processes: spotlight on the prostate. Epigenomics. (2015) 7:413–25. doi: 10.2217/epi.15.8
7. Bogdanovic O, Veenstra GJ. DNA methylation and methyl-CpG binding proteins: developmental requirements and function. Chromosoma. (2009) 118:549–65. doi: 10.1007/s00412-009-0221-9
8. King AD, Huang K, Rubbi L, Liu S, Wang CY, Wang Y, et al. Reversible regulation of promoter and enhancer histone landscape by DNA methylation in mouse embryonic stem cells. Cell Rep. (2016) 17:289–302. doi: 10.1016/j.celrep.2016.08.083
9. Benetatos L, Vartholomatos G. Enhancer DNA methylation in acute myeloid leukemia and myelodysplastic syndromes. Cell Mol Life Sci. (2018) 75:1999–2009. doi: 10.1007/s00018-018-2783-2
10. Gravina GL, Ranieri G, Muzi P, Marampon F, Mancini A, Di Pasquale B, et al. Increased levels of DNA methyltransferases are associated with the tumorigenic capacity of prostate cancer cells. Oncol Rep. (2013) 29:1189–95. doi: 10.3892/or.2012.2192
11. Valdés-Mora F, Clark SJ. Prostate cancer epigenetic biomarkers: next-generation technologies. Oncogene. (2015) 34:1609–18. doi: 10.1038/onc.2014.111
12. Hieronymus H, Murali R, Tzin A, Yadav K, Abida W, Moller H, et al. Tumor copy number alteration burden is a pan-cancer prognostic factor associated with recurrence and death. Elife. (2018) 7:e37294. doi: 10.7554/eLife.37294.027
13. Robinson D, Van Allen EM, Wu YM, Schultz N, Lonigro RJ, Mosquera JM, et al. Integrative clinical genomics of advanced prostate cancer. Cell. (2015) 161:1215–28. doi: 10.1016/j.cell.2015.05.001
14. Taylor BS, Schultz N, Hieronymus H, Gopalan A, Xiao Y, Carver BS, et al. Integrative genomic profiling of human prostate cancer. Cancer Cell. (2010) 18:11–22. doi: 10.1016/j.ccr.2010.05.006
15. Hamid AA, Gray KP, Shaw G, MacConaill LE, Evan C, Bernard B, et al. Compound genomic alterations of TP53, PTEN, and RB1 tumor suppressors in localized and metastatic prostate cancer. Eur Urol. (2019) 76:89–97. doi: 10.1016/j.eururo.2018.11.045
16. Langelöffel P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinform. (2008) 9:559. doi: 10.1186/1471-2105-9-559
17. Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhim R, Getz G. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. Genome Biol. (2011) 12:R41. doi: 10.1186/gb-2011-12-4-r41
28. Lal MA, Andersson AC, Katayama K, Xiao Z, Nukui M, Hultenby K, Ye P, Shen L, Jiang W, Ye Y, Chen CT, Wu X, et al. Zn-driven discovery of Sonntag R, Giebeler N, Nevzorova YA, Bangen JM, Fahrenkamp D, Lambertz Zarrei M, MacDonald JR, Merico D, Scherer SW. A copy number Litwin MS, Tan HJ. The diagnosis and treatment of prostate cancer: a review. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA Cancer J Clin. Ross-Adams H, Lamb AD, Dunning MJ, Halim S, Lindberg J, Massie CM, et al. Rhophilin-1 is a key regulator of the podocyte cytoskeleton and variation map of the human genome. Nat Rev Genet. 12:3203–17. doi: 10.3390/md12063203 239:116856. doi: 10.1016/j.cell.2019.116856

26. 25:896–905. doi: 10.1073/pnas.1920207118

24. 23:430–40. doi: 10.2174/13816282366669110304549

23. Intramembrane cleavage and normal gene expression profiling and interactive analyses. Nucleic Acids Res. (2017) 45:W98–102. doi: 10.1093/nar/gks247

22. 178:525–36. doi: 10.1016/j.ajpath.2010.10.002

21. 236:472–8. doi: 10.3824/ebio.2019.236.472

20. 25:4630–40. doi: 10.2174/13816282566669110304549

19. 62. doi: 10.1073/pnas.0437724100

18. 63. doi: 10.1073/pnas.0437724100

17. 115:9282–7. doi: 10.1073/pnas.1807155115

16:803–10. doi: 10.1038/nrc.2016.83

15:2532–42. doi: 10.1001/jama.2017.7248

14:924–32. doi: 10.1073/pnas.1701551115

13:287–37. doi: 10.3332/caac.215511

12:3203–17. doi: 10.3390/md12063203

11:1627–35. doi: 10.1002/bja.25055

10:1038/nrg3871

9:7–34. doi: 10.3332/caac.215511

8:226. doi: 10.3390/ijms18010226

7:5–34. doi: 10.3322/caac.215511

6:1197–1197. doi: 10.3824/ebio.2019.6:1197-1197

5:1157–1157. doi: 10.3824/ebio.2019.5:1157-1157

4:309–309. doi: 10.3824/ebio.2019.4:309-309

3:243–243. doi: 10.3824/ebio.2019.3:243-243

2:134–2134. doi: 10.3824/ebio.2019.2:134-2134

1:104–104. doi: 10.3824/ebio.2019.1:104-104

0:1–1. doi: 10.3824/ebio.2019.0:1-1

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Hou, Hu, Zhou, Liu, Chen and Yang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.