Genome-wide copy number analysis reveals candidate gene loci that confer susceptibility to high-grade prostate cancer

Prevathe Poniah, B.Sc. a,b,*, Shamsul Mohd Zain, Ph.D. b, Azad Hassan Abdul Razack, M.B.B.S., Ph.D. a, Shanggar Kuppusamy, M.B.B.S., Ph.D. a, Shankar Karuppayah, Ph.D. c, Hooi Sian Eng, Ph.D. b, Zahurin Mohamed, Ph.D. b,*

a Department of Surgery, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia
b Department of Pharmacology, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia
c National Advanced IPv6 Centre, Universiti Sains Malaysia, Pulau Pinang, Malaysia

Received 3 August 2016; received in revised form 8 April 2017; accepted 23 April 2017

Abstract

Background: Two key issues in prostate cancer (PCa) that demand attention currently are the need for a more precise and minimally invasive screening test owing to the inaccuracy of prostate-specific antigen and differential diagnosis to distinguish advanced vs. indolent cancers. This continues to pose a tremendous challenge in diagnosis and prognosis of PCa and could potentially lead to overdiagnosis and overtreatment complications. Copy number variations (CNVs) in the human genome have been linked to various carcinomas including PCa. Detection of these variants may improve clinical treatment as well as an understanding of the pathobiology underlying this complex disease.

Methods: To this end, we undertook a pilot genome-wide CNV analysis approach in 36 subjects (18 patients with high-grade PCa and 18 controls that were matched by age and ethnicity) in search of more accurate biomarkers that could potentially explain susceptibility toward high-grade PCa. We conducted this study using the array comparative genomic hybridization technique. Array results were validated in 92 independent samples (46 high-grade PCa, 23 benign prostatic hyperplasia, and 23 healthy controls) using polymerase chain reaction–based copy number counting method.

Results: A total of 314 CNV regions were found to be unique to PCa subjects in this cohort ($P < 0.05$). A log$_2$ ratio-based copy number analysis revealed 5 putative rare or novel CNV loci or both associated with susceptibility to PCa. The CNV gain regions were 1q21.3, 15q15, 7p12.1, and a novel CNV in PCa 12q23.1, harboring ARNT, THBS1, SLC5A8, and DDC genes that are crucial in the p53 and cancer pathways. A CNV loss and deletion event was observed at 8p11.21, which contains the SFRP1 gene from the Wnt signaling pathway. Cross-comparison analysis with genes associated to PCa revealed significant CNVs involved in biological processes that elicit cancer pathogenesis via cytokine production and endothelial cell proliferation.

Conclusion: In conclusion, we postulated that the CNVs identified in this study could provide an insight into the development of advanced PCa.

Keywords: Copy number variation; Prostate cancer; Genome-wide analysis

1. Introduction

Prostate cancer (PCa) is a slow growing, noncutaneous cancer that continues to be the most common malignancy and the second leading cause of cancer-related deaths among men in developed countries [1]. Cancer registry in Malaysia ranks PCa as the fourth most common cancer among men [2]. However, the PCa prevalence in Asians has been reported to be 20 times lower than in Western countries [3]. Although
substantial research on PCa has been conducted, the exact etiology and pathogenesis of PCa have yet to be fully elucidated. To date, only prostate-specific antigen (PSA) test in conjunction with digital rectal examination has been widely used in clinical practice to identify patients who require transrectal ultrasound biopsy for detection of PCa [4]. Nonetheless, PSA test comes with well-known limitations. Although PSA test contributes toward reducing PCa-related death by 20%, overdiagnosis has been a huge concern, as it is often not possible to differentiate a benign condition from that of a cancerous one [5]. Therefore, histopathology remains the gold standard for PCa diagnosis.

Cancer genetics can offer ways to identify predictive markers for cancers, and such efforts have been deemed successful. This is proven by mounting evidence from studies conducted in hereditary breast and ovarian cancers regarding the BRCA1 and BRCA2 genes [6], the KRAS, p53, and epidermal growth factor receptor (EGFR) genes being notable biomarkers in colorectal, esophageal, liver, and pancreatic cancers [7]; hypermethylation of the DAP-kinase in bladder cancer [8]; and mutation or polymorphisms in SRD5A2 that contribute toward androgen metabolism in PCa [9,10]. A deletion copy number variation (CNV) at 20p13 is associated with aggressiveness of PCa in a genome-wide study [11]. CNVs have also been characterized in several other PCa-related studies [12–14]. Nonetheless, findings of these studies have been inconclusive due to ethnicity, lifestyle, and environment differences. Unlike familial breast cancer, in which high penetrance of the BRCA1 mutation results in 80% lifetime risk of developing breast cancer among individuals with the mutation, PCa on the contrary, is a complex disease with multiple risk factors involved and reduced penetrance, hence providing challenges in predicting clinical outcomes.

In this study, we compared patients with PCa with high-grade (Gleason 8–10) cancer, the stage in which cells are moderately to poorly differentiated and which would be a more suitable group to assess the prognosis of PCa [15] to benign prostatic hyperplasia (BPH). PCa of lower stages were excluded from the study owing to possibilities of coexistence with BPH because lower-grade PCa often resembles foci of BPH [16]. Here, we described a pilot study designed to detect somatic CNV loci in Chinese patients within the Malaysian population, with high-grade PCa. Although the literature is mostly focused on the Western population owing to its high disease prevalence, relatively little is known about this condition in Asians. In addition, the Malaysian Chinese differ in the population structure from that of the Chinese Han Beijing, commonly reported in the HapMap [17].

2. Materials and Methods

The materials and methods and any associated references are available in the online version of the paper as Supplementary S3 at the following URL http://dx.doi.org/10.1016/j.urolonc.2017.04.017.

3. Results

Within the study cohort, patients’ ages ranged from 60 to 84 years (mean age: 69.9 ± 8.45) with PSA value ranging from 7.67 to 159 ng/ml. Subjects’ data are presented in Table 1. All samples passed strict quality control during sample processing, data extraction, and analysis. The array comparative genomic hybridization (aCGH) method follows the protocol of pooling DNA samples from patients and controls. This allows for subjects to be cross-compared to normalize common copy changes in the control samples. As such, CNV calls detected via aCGH principle are designed to be patient-specific.

3.1. CNV identification

Using the aCGH platform, somatic CNVs within the genome of every individual were identified in this study cohort, comprising 18 PCa cases and 18 matched controls. We ran these samples in parallel whereby the test DNA (patient) and reference (control) were assessed simultaneously. We detected a total of 339 CNVs (P < 0.05) spanning from 1.47 kb to 10.88 Mb in size with 92.6% CNVs mapped to gene-specific loci. Copy number amplifications were more commonly observed compared to deletions with 237 (69.9%) CNV gains and 102 (30.1%) CNV losses (gain:loss ratio = 2.3:1). Mean size of the CNVs was 289.66 kb, with a median of 5.65 kb, encompassing 164.4 Mb that is 5.08% of the genome. Twenty-five (7.37%) CNVs were found on the sex chromosomes. However, we excluded all markers from the sex
chromosome, as the frequency of amplification or deletion events was less than 5%, hence not fulfilling the CNV calling criteria resulting in 314 autosomal CNVs. This includes 225 CNV gains and 89 CNV losses (Fig. 1). CNVs that were included for analysis in this study were those that are present with a frequency of at least 10%, amounting to 140 (44.5%) autosomal CNVs (101 gains and 39 losses) with an average of 7.8 CNVs per individual.

The most frequent CNV gain (54.5%) was observed in the chromosome region 1q21.3 that codes for the ARNT gene (aryl hydrocarbon receptor nuclear translocator) (Table 2). A few other regions that were amplified (14q32.32-q32.33 and 15q14-q15.1) contained several genes that have been previously associated with PCa, namely the BAG5 and the THBS1. The most frequently deleted region (45.5%) was 8p11.21 that harbors the secreted frizzled-related protein 1 (SFRP1) and Ankyrin 1 (ANK1) genes that are widely reported in many cancers including PCa [7,18]. There were several other CNV regions with the reasonably high frequency of deletion detected, which includes 12q23.1 (27.3%). Region 12q23.1 contains the solute carrier family 5 (SLC5A8) gene responsible for tumor-suppressor and antitumor activity in PCa [19–21]. Out of the 314 CNVs, amplification events were more common (71.7%) compared to deletions (Fig. 2). Approximately 101 CNVs from the amplified region presented with a frequency of more than 10%. On the contrary, among the 89 deleted regions, 43.8% had a CNV frequency of more than 10%.

### 3.2. Identification of functional implication of CNVs associated with PCa

CNVs unique to patients with PCa were identified by a cross-comparison method with known CNVs from the database of genomic variation (DGV). The assessment revealed 20 rare or novel CNVs (DGV coverage <50%) or both that were present in at least 10% of the samples. Out of these, 13 were classified as rare (DGV coverage <50%), whereas 7 others were novel (DGV coverage 0%) (Supplementary Material S1 Table). A chi-square test confirmed the significance of the association of these CNVs with PCa ($P < 0.05$) as compared to BPH. Evaluation of overlap between CNVs reported in DGV database with CNVs found in this study revealed 5 CNVs that were present in at least 10% of the samples found to be associated with PCa (Table 3). Of these, 3 were rare CNVs (DGV

**Table 2**

<table>
<thead>
<tr>
<th>Cytoband</th>
<th>CNV frequency (%)</th>
<th>DGV coverage (%)</th>
<th>Start</th>
<th>End</th>
<th>Size (kb)</th>
<th>Number of genes within region</th>
<th>Candidate gene(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Amplification</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1q21.3</td>
<td>54.5</td>
<td>&lt;50</td>
<td>150838873</td>
<td>150852662</td>
<td>13.79</td>
<td>1</td>
<td>ARNT</td>
</tr>
<tr>
<td>7q31.33</td>
<td>45.5</td>
<td>100</td>
<td>125833251</td>
<td>125837583</td>
<td>4.3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7p12.1</td>
<td>45.5</td>
<td>&lt;50</td>
<td>51562878</td>
<td>51565730</td>
<td>2.85</td>
<td>5</td>
<td>DDCC</td>
</tr>
<tr>
<td>14q32.32-q32.33</td>
<td>36.4</td>
<td>&lt;50</td>
<td>103792091</td>
<td>104925483</td>
<td>1133.39</td>
<td>25</td>
<td>BAG5, MIR20, IGHG3</td>
</tr>
<tr>
<td>15q14-q15.1</td>
<td>36.4</td>
<td>&lt;50</td>
<td>39877118</td>
<td>40174089</td>
<td>296.9</td>
<td>3</td>
<td>THBS1</td>
</tr>
<tr>
<td>5q35.3</td>
<td>36.4</td>
<td>100</td>
<td>180343369</td>
<td>180345129</td>
<td>1.7</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td><strong>Deletion</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10q21.3</td>
<td>45.4</td>
<td>100</td>
<td>66976925</td>
<td>66985010</td>
<td>8.1</td>
<td>0</td>
<td>CTNNA3</td>
</tr>
<tr>
<td>8p11.2</td>
<td>45.4</td>
<td>&lt;50</td>
<td>39351357</td>
<td>39505315</td>
<td>153.9</td>
<td>4</td>
<td>SFRP1, ANKI</td>
</tr>
<tr>
<td>11q22.3</td>
<td>36.4</td>
<td>100</td>
<td>106744337</td>
<td>106747979</td>
<td>3.6</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td><strong>Amplification and deletion</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2p16.3</td>
<td>36.4</td>
<td>100</td>
<td>52604016</td>
<td>52637451</td>
<td>33.4</td>
<td>3</td>
<td>NRRX1, CCDC88A</td>
</tr>
</tbody>
</table>

**CNV =** copy number variation; **CNV frequency =** CNVs that are frequently observed as either amplifications or deletions in each sample.
coverage <50%; 1q21.3, 15q15, 7p12.1), 1 novel CNV in PCa (12q23.1), and 1 common CNV (8p11.21). To further assess the likelihood of the involvement of these CNVs in PCa, the genes located within these regions were identified, and their involvement in shared biological processes with known PCa-related genes was assessed. First, we profiled the genes within the chromosomal regions that are bounded by the 20 rare or novel CNVs or both, where genes such as chemokine (C-X-C motif) receptor 2 (CXCR2) and cytochrome P450, family 27, subfamily A, polypeptide 1 (CYP27A1) are located. A list of genes known to be associated with PCa was then obtained. We then performed gene ontology (GO) enrichment and Kyoto encyclopedia of genes and genomes (KEGG) pathway analysis using the Database for Annotation, Visualization, and Integrated Discovery (DAVID) gene annotation tool for the 2 sets of genes (genes within 1q21, 15q15, 12q23.1, 7p12.1, and 8p11.21 regions and known genes associated with PCa) (see Supplementary Material S2 Table for full list of genes associated with PCa). This analysis revealed several metabolic and biological processes that could potentially be linked to PCa such as the p53 signaling pathway, Wnt signaling pathway, as well as other cancer pathways involved in the pathogenesis of the disease (Tables 4 and 5).
3.3. Validation of candidate CNVs

We validated another subset of samples comprising 46 cases and 46 controls for 3 randomly selected CNV regions. Each is a patient-control matched pair. This selection was done based on comparative hybridization method whereby samples were matched and randomly assigned as test references. We validated 3 randomly selected CNV regions representing different copy number status (the CNVs at 1q21.3 and 7p12.1 were rare copy number gains, and 8p11.21 was a copy number loss) (Fig. 3). Genomic regions with events of amplifications and deletions were defined based on differences between patient’s copy number and the wild-type copy number. Genomic aberrations were identified by applying log2 intensity ratios of sample to reference, log2-ratios more than 0.3 for copy number gain, and less than −0.6 for copy number loss [22]. Wells with VIC Ct greater than 32.0 were excluded from the analysis. The default filtering threshold to categorize samples with relatively weak, background amplification as having zero copies of the target was 4.0.

Validation of the 3 randomly selected CNV regions was successfully executed. Amplification of the region 1q21.3 harboring ARNT gene in the array was found to be

<table>
<thead>
<tr>
<th>Cytoband</th>
<th>Sample frequency (%)</th>
<th>DGV coverage (%)</th>
<th>Start</th>
<th>Size (kb)</th>
<th>Number of genes within demarcated region</th>
<th>Candidate gene(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amplification</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rare</td>
<td>1q21.3</td>
<td>33.3</td>
<td>&lt;50</td>
<td>150838873</td>
<td>150852662</td>
<td>13.79</td>
</tr>
<tr>
<td></td>
<td>15q15</td>
<td>22.2</td>
<td>&lt;50</td>
<td>39877118</td>
<td>40174089</td>
<td>296.9</td>
</tr>
<tr>
<td></td>
<td>7p12.1</td>
<td>27.8</td>
<td>&lt;50</td>
<td>50526134</td>
<td>50633154</td>
<td>2.853</td>
</tr>
<tr>
<td>Novel</td>
<td>12q23.1</td>
<td>16.7</td>
<td>0</td>
<td>100626837</td>
<td>100629891</td>
<td>3.05</td>
</tr>
<tr>
<td>Deletion</td>
<td>8p11.21</td>
<td>27.8</td>
<td>&lt;50</td>
<td>40894110</td>
<td>40898625</td>
<td>4.516</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gene symbol</th>
<th>Gene name</th>
<th>KEGG_pathway</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARNT</td>
<td>Aryl hydrocarbon receptor nuclear translocator</td>
<td>hsa05200: Pathways in cancer, hsa05211: Renal cell carcinoma</td>
</tr>
<tr>
<td>DDC</td>
<td>Dopa decarboxylase</td>
<td>hsa00340: Histidine metabolism, hsa00350: Tyrosine metabolism, hsa00360: Phenylalanine metabolism, hsa00380: Tryptophan metabolism</td>
</tr>
<tr>
<td>SFRP1</td>
<td>secreted frizzled-related protein 1</td>
<td>hsa04115: g53 signaling pathway, hsa04350: TGF-beta signaling pathway, hsa04510: Focal adhesion, hsa04512: ECM-receptor interaction, hsa05219: Bladder cancer</td>
</tr>
<tr>
<td>THBS1</td>
<td>thrombospondin 1</td>
<td>hsa04115: g53 signaling pathway, hsa04350: TGF-beta signaling pathway, hsa04510: Focal adhesion, hsa04512: ECM-receptor interaction, hsa05219: Bladder cancer</td>
</tr>
</tbody>
</table>

Table 3 Rare, novel, and common CNVs in PCa

Table 4 Enriched GO terms associated with PCa

Table 5 KEGG pathway associated with candidate genes from CNV regions

GO_BP, gene ontology biological process; GO_CC, gene ontology cellular component; GO_MF, gene ontology molecular function.

"Modified Fisher’s exact test, P ≤ 0.05."
Fig. 3. qPCR validation performed on selected regions. (A) Results for region 1q21.3. (B) Results for region 7p12.1. (C) Results for region 8p11.21. Copy number of approximately 2 was deemed to be indicating wild-type (no CNV). Copy number 1 or below is indicative of copy loss, whereas those that are 3 and above are copy number gains. Black bars indicate cases and gray bars are controls. Missing bar indicates deletion event. Error bars denote mean ± SD (n = 4). qPCR = quantitative polymerase chain reaction. (Color version of figure is available online.)
significantly higher in the PCa cases compared to the controls from the validation cohort, \( P = 0.018 \). Similar finding was observed for the deletion at 8p11.21, odds ratio \( = 6.071 \) (95% CI: 1.137–32.411); \( P = 0.055 \). However, amplification at 7p12.1, odds ratio \( = 2.727 \) (95% CI: 0.520–14.925); \( P = 0.392 \) was not significant between cases and controls. The \( P \) values for all 3 regions lost its significance after conservative Bonferroni correction for multiple testing (0.05/140). These warrant further investigation.

4. Discussion

Detection and quantification of genome-wide structural alterations often become a challenging process, wherein 12 prostate biopsy cores are obtained, and variations in tumor grading could occur if areas without cancer are captured instead of tumor tissues. The availability of blood samples from patients provides an opportunity to analyze disease prognosis as tumor-derived DNA are found in the circulatory system of cancer patients, and it has been a viable, noninvasive approach in studies related to risk and progression of the disease [23]. This study represents a whole-genome CNV analysis in the Malaysian cohort of age-matched Chinese patients with PCa against BPH. We identified 5 CNVs that are rare or novel or both that could potentially link with PCa clinical outcome.

In this study, the most frequently amplified region was 1q21.3 enriched with the \( ARNT \) gene. The \( ARNT \) gene is also known as the hypoxia-inducible factor (\( HIF-1\alpha \)). Approximately 1% of the genome comprises hypoxia-regulated genes (\( HIF-1\alpha \) and \( HIF-1\beta \)) [24]. \( HIF-1\alpha \) forms a heterodimer with the \( ARNT \) gene corresponding to the various levels of oxygen in the tumor microenvironment, thus promoting cell survival as well as angiogenesis [25]. These genes are known to significantly influence cancer cells in adapting to hypoxic environment and glycolysis. Increasing levels of hypoxia are known to be associated with higher clinical stages in PCa. Up-regulation of \( HIF-1\alpha \) was also significantly higher in high-grade prostate intraepithelial neoplasia [26]. High-grade prostate intraepithelial neoplasia is a precursor to PCa development. Therefore, amplification of this CNV could be indicative of an early event in PCa [27]. A hypoxic microenvironment promotes genetic alterations subsequently leading to more aggressive cell phenotype and disease progression in PCa [28]. \( ARNT \) was found to be expressed constitutively under nonhypoxic conditions in PCa cells suggestive to be sharing common signaling pathways with the \( HIF-1\alpha \) proteins for nuclear protein accumulation in PCa [29]. In this study, the somatic CNV at 1q21.3 region was highly amplified potentially contributing to increased expression of \( ARNT \) at the site of the tumor, suggesting its candidacy in being a potential marker for PCa pathogenesis. These data taken together suggest that locus 1q21.3 may explain susceptibility to PCa and serve as a promising target as well. However, there were limited publications that have reported the importance of this region in prostate diseases. Hence, further studies on this locus will be required to confirm findings reported here.

Regions 8p11.21 and 10q21.3 were most frequently deleted among PCa subjects in this study. The 8p11.21 locus harbors the \( SFRP1 \), a tumor suppressor gene that could potentially have implications on PCa. The KEGG pathway analysis revealed the biological process of \( SFRP1 \) gene in the Wnt signaling pathway, a well-known pathway in carcinogenesis [30,31]. Data from a genome-wide association study (GWAS) conducted in a white population identified tumors with loss of heterozygosity event in the 8p11 subregion containing \( SFRP1 \) gene associated with PCa aggressiveness [32]. Loss of region at chromosome 8p is a common event in PCa [33,34]. Apart from \( SFRP1 \), the \( ANK1 \), a putative tumor suppressor gene, was also found within the same region that was most frequently deleted. Allelic losses have been reported previously in this region in PCa tumors [35]. CNV region on chromosome 10q21.3 found in this study harbors \( CTNN3 \), a \( \alpha \)-catenin (cadherin-associated) protein. This \( \alpha \)-catenin protein was found to be significantly down-regulated in PC-3 cell line and has been reported to be a promising putative marker in distinguishing aggressive PCa from indolent ones [36]. To the best of our knowledge, there are no reports on GWAS or CNVs identified for this region in PCa so far.

In this pilot study, we identified 5 putative somatic CNVs that are crucial in PCa. Of these, 4 were amplified wherein 3 were rare (1q21.3, 15q15, and 7p12.1), 1 CNV novel to PCa, 12q23.1 and 1 deletion event at 8p11.21. Findings of regions 7p12.1 gain and 8p11.21 loss in our study coincide with data from a study that reported genomic alterations in Chinese men compared to the Western population [3]. The potential significance of these loci was successfully established through chi-square tests against another loci and validation by quantitative polymerase chain reaction. This way we could confirm that, despite the small sample size, our study was able to produce biologically meaningful and reproducible results. Although a significant number of CNVs were identified in other PCa studies, an interesting aberration at the 12q23.1 locus is a novel CNV found in our study. The \( SLC5A8 \) gene within this region has been implicated in several cancers including PCa [37]. Pathway analysis indicates that this gene is actively involved in PCa progression. Although this region has been reported for CNV associated with other diseases [38–40], this is the first time an amplification event was observed for PCa at 12q23.1. Another aberration at region 7p12.1 harbors the i-dopa decarboxylase (\( DDC \)) gene that is significantly up-regulated in patients with PCa [41]. \( DDC \) is considered an important protein coding gene and appears in more than 30% of PCa cases involving neuroendocrine differentiation [42]. The messenger RNA expression levels of \( DDC \) were found to be overexpressed in prostate tissues with higher Gleason scores that correspond to high-grade
PCa, compared to BPH [43]. Overexpression of DDC has been shown to increase androgen receptor activity and plays a crucial biological role in prostate gland physiology and progression of PCa [44]. A study conducted using radical prostatectomy specimens shows copy number gains of the EGFR gene also spanning the 7p12 chromosome region. Amplification of EGFR gene copies suggests an association with EGFR protein expression with grade, stage, patient prognosis, and PSA recurrence in PCa [45]. Findings reporting copy number gains involving DDC from our study and EGFR from a previous study indicate the importance of structural genomic aberrations at region 7p12 that is often associated with poor prognosis in PCa.

A CNV, which has been identified in cases when compared to controls, becomes more meaningful if it contributes to the pathogenesis of the disease. The search for putative biomarkers was further refined by stretching our analysis to KEGG pathway analysis and gene ontology functional enrichment for genes within the CNV regions and those that are known to be involved in PCa. Here, we observed pathways that are commonly shared between CNVs identified in this study and the list of genes reported previously for PCa. Interestingly, through KEGG pathway analysis, we found region 15q15 to be involved in the p53 signaling pathway and transforming growth factor (TGF)-beta signaling pathway (Table 5). The CNV gain identified at 15q15 harbors the THBS1 gene, a major activator of TGFβ-1. TGFβ plays important roles in PCa by promoting tumor-suppressor and tumor-promoting activities [46]. Hence, any misregulation of TGFβ can induce cell proliferation and tumorigenesis [47]. The p53 gene is noted to be a tumor suppressor gene that activates THBS1 gene promoter, thereby leading THBS1 to act as an angiogenesis inhibitor in cancers. The p53 gene is found to increase THBS1 expression [48]. High levels of THBS1 are secreted by tumors to inhibit tumor growth.

Although this study has a relatively smaller sample size, it is noteworthy that several CNVs discovered in this study have been previously reported in PCa by other GWAS studies, thus further strengthening our findings [3]. This preliminary screening was successful in presenting structural aberrations specific to PCa. Apart from that, we also found several CNVs that are involved in pathways that could potentially explain disease progression, which to the best of our knowledge has not been reported before in Malaysian men, thereby adding value to the current knowledge of PCa studies in the Asian population. Albeit reporting our findings based on somatic DNA events, however, the study could have been further supported with evidence of circulating tumor cells if tumor tissues are available. Despite the small numbers, which are influenced by patients’ availability, we identified CNVs that are statistically significant. Namely, region 1q21.3, which were one of the most frequently amplified CNVs in our study, was also found to be a copy number gain in another PCa study reported by Robbins et al. [49]. Similarly, another CNV at region 14q32.33, one of the top regions of copy number gain for PCa in this study, was also reported by Laitinen et al. [50] to be associated with PCa in high-risk African American families. Regions harboring CNVs that were identified are biologically relevant to PCa development, hence worthwhile for future investigations. We hypothesize that these CNVs could provide a better understanding of PCa pathophysiology, especially in the Asian population. We believe our study that has been successful in identifying rare and novel CNVs would pave the way for a more comprehensive investigation in a larger cohort taking into consideration other factors that may potentially contribute toward variations in this population which may explain functional interpretations of CNVs reported here. This could also provide useful information in understanding the disease better in a different ethnic group as compared to the Western population.

In our study, most of the CNVs that were detected are common CNVs (DGV coverage 100%). Common CNVs may potentially have minimal effect on the pathogenesis of diseases. However, it has to be taken into consideration that the definition of “common” in CNV selection is also debatable. This is because not all CNVs reported in DGV are specific to PCa (non-PCa studies), and in any case, it is unlikely that these were obtained from the Malaysian Chinese population. The Pubmed database search revealed that only 1.3% of total published CNV studies had reported on PCa indicating the need for more comprehensive studies focusing on CNVs in PCa in order to explain common CNVs. Having said that, the present study has limitations that need to be acknowledged, the first of which is its small sample size. Recruitment of subjects in this study is from the University of Malaya Medical Centre (UMMC). Ethnic distribution also needs to be taken into account carefully as the catchment area consists mainly of the Chinese ethnic group. Therefore, to minimize the effect of ethnicity as a confounding factor, and to strengthen the possibilities of outcomes in our study despite small sample size, we have restricted our sample selection to only Chinese patients who are diagnosed with high-grade PCa (Gleason score: 8–10).

Our study includes only cases (PCa) and the benign group (BPH) as controls. We did not include any age-matched normal healthy controls. Limiting our controls only to the BPH group posed strength in our study, whereby controls recruited were all confirmed negative for PCa, histologically confirmed for BPH, and had adequate follow-up. Selection of this group allowed us to rule out the possibility of the healthy control coexisting within the BPH arm, given that BPH is a common occurrence among men from this age group as our subjects were all more than the age of 60. Furthermore, PCa subjects included in this study are of advanced cancers with Gleason scores 8 to 10. High-grade PCa occurs in the peripheral zone of the prostate [16], thereby minimizing any overlap with unidentified BPH, which commonly arises in the transition zone. Hence, the selection of BPH as our control group would point to the fact that the CNVs discovered are indicative of a predisposition to PCa and reduce the
possibilities of the same CNVs being present in BPH conditions. CNVs involved in PCa predisposition may be the key toward targeted cancer therapies with increased effectiveness. The validation results also managed to confirm that there were no significant differences in CNVs between BPH and healthy controls. However, the frequency of event was much higher in the discovery cohort as compared to the validation cohort. This could be due to the following reasons: (1) the number of samples in the discovery cohort was relatively low, and (2) cases from the discovery cohort were stringently matched and less so in the validation cohort.

This study has successfully identified several CNVs that are being reported for the first time in PCa, and therefore may be useful in diagnosis and prognosis of PCa. In contrast to presently available commercial molecular tests that rely upon tissue samples involving biopsy procedures, this work demonstrates a minimally invasive approach in diagnosing PCa. CNVs detected through this study are able to provide a clear insight into understanding PCa carcinogenesis and progression, thereby aiding in patient management. Hence, our study suggests a method of improving PCa diagnosis and prognosis by clearly distinguishing patients who are susceptible to developing the disease.

5. Conclusion

The use of aCGH technology in CNV identification has made high-resolution CNV detection possible and more reliable than previous technologies. We have done manual verification on genes that have been associated with PCa genotyping, expression, or protein studies. Beyond the possible prognostic significance of CNVs, our work is a prelude to further studies on CNV in PCa. Therefore, a more comprehensive analysis of a larger study cohort and additional functional studies regarding genes and their roles in PCa will have to be conducted before a valid conclusion can be done to support these findings.

Acknowledgments

The authors thank all the patients and volunteers for their assistance and contribution in performing the above-described studies.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.urolonc.2017.04.017.

References


