Overexpression of NCAPH is upregulated and predicts a poor prognosis in prostate cancer

FEILUN CU, JIANPENG HU, ZHIPENG XU, JIAN TAN and HUAMING TANG

Department of Urology, Affiliated People's Hospital of Jiangsu University, Zhenjiang, Jiangsu 212002, P.R. China

Received March 6, 2018; Accepted March 18, 2019

DOI: 10.3892/ol.2019.10260

Correspondence to: Dr Jianpeng Hu, Department of Urology, Affiliated People's Hospital of Jiangsu University, 8 Dianli Road, Zhenjiang, Jiangsu 212002, P.R. China
E-mail: 249056614@qq.com

Key words: prostate cancer, non-SMC condensin I complex subunit H, prognosis, biomarker, cell cycle
lung squamous cell carcinoma (LuSC), lung adenocarcinoma (KIRC), liver hepatocellular carcinoma (LIHC), lung squamous cell carcinoma (LUSC), lung adenocarcinoma (LUAD), kidney renal papillary cell carcinoma (KIRP), stomach adenocarcinoma (STAD) and PCa were downloaded from TCGA (https://tcga-data.nci.nih.gov/tcga/) data portal. The RNA expression data (level 3) were generated from a HiSeq 2000 sequencing platform (Illumina, Inc., San Diego, CA, USA) by RNASeqV2 post-processing pipelines and demonstrated as RNA-Seq by Expectation-Maximization normalized count data. Patient clinical data, including age at diagnosis, days to last follow-up, pathological T stage and pathological N stage, were retrospectively abstracted from patient records, which were downloaded from cBioPortal (http://www.cbioportal.org/). A total of 112 normal breast samples and 1100 BRCA samples, 44 normal head and neck samples and 522 HNSC samples, 41 normal colon samples and COAD samples, 72 normal kidney renal clear cell samples and 534 KIRC samples, 50 normal liver samples and 373 LIHC samples, 32 kidney renal papillary cell samples and 291 KIRP samples, 35 normal stomach samples and 415 STAD samples, 51 lung squamous cell samples and 501 LUSC samples, and 59 lung samples and 517 LUAD samples were used in the present study. All the patients were staged according to the 2009 TNM classification of the American Joint Committee on Cancer/International Union Against Cancer (15).

The expression levels of NCAPH were downloaded as TCGA data from cBioPortal (http://www.cbioportal.org/) (16). The TCGA PCa dataset was analyzed to evaluate the expression pattern of NCAPH in PCa samples. A total of 38 normal and 497 PCa samples were included in this dataset. Normal prostate tissue samples analyzed from the TCGA cohort were obtained from the marginal area of tumours derived from the same patient and were therefore paired samples. The median age of the patients in the TCGA PCa dataset was 62 years (range, 41-78 years). NCAPH protein expression levels in the PCa tissues and normal prostate tissues were reviewed using the Human Protein Atlas (http://www.proteinatlas.org/) (17).

Microarray data. Three public datasets GSE21034 (18), GSE17951 (19) and GSE55945 (20) were download from the Gene Expression Omnibus (GEO) database (www.ncbi.nlm.nih.gov/geo/). GSE21034 (Taylor dataset) was reported by Taylor et al (18) and includes 370 PCa tissue samples and cell line samples. The GSE17951 dataset was reported by Jia et al (19) and contains 70 normal prostate and 67 PCa samples. The GSE55945 dataset was reported by Arredouani et al (20) and contains eight normal prostate and 13 PCa samples. Unfortunately, the clinicopathological features of the patients were not available in the three GEO datasets. Genes with a fold-change ≥2 and P<0.05 were considered to be significantly differentially expressed.

Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis. To identify functions of NCAPH in PCa, GO functional enrichment analysis was performed for three functional ontologies: Biological process, cellular component and molecular function. KEGG pathway enrichment analysis was also performed to identify pathways enriched in PCa using the Database for Annotation, Visualization and Integrated Discovery (DAVID; https://david.ncifcrf.gov/home.jsp) (21). The P-value was calculated by hypergeometric distribution, and a pathway with P<0.05 was considered as significant.

Protein-protein interaction (PPI) network construction. Using a NCAPH interacting proteins dataset which was download from NCBI (https://www.ncbi.nlm.nih.gov/gene/23397), a total of 51 candidate proteins were identified. The Search Tool for the Retrieval of Interacting Genes/Proteins (STRING; string-db.org/) database is a pre-computed global resource for the investigation and analysis of associations between proteins (22). STRING was used to analyze the interactions between NCAPH-associated proteins, which were visualized with PPI networks. In the present study, only experimentally validated interactions with a combined score ≥0.4 were selected as significant. PPI networks were constructed using Cytoscape 3.6.1 software (http://www.cytoscape.org/) (23).

Statistical analysis. SPSS 17.0 software (SPSS, Inc., Chicago, IL, USA) was used for all statistical analysis. Pearson's χ² test and Fisher's exact test were used to analyze the associations of NCAPH expression with clinicopathological characteristics. Paired and independent Student's t-tests were used to compare continuous variables in two groups. Receiver operating curve (ROC) analysis was performed for evaluation of the specificity and sensitivity of NCAPH expression levels for discriminating patients with PCa from normal patients. Associations between continuous variables and categorical variables were evaluated using Mann Whitney U tests for two groups. Kaplan-Meier and Cox regression analyses were utilized to assess the associations between NCAPH and overall survival, as well as the prognosis of patients with PCa. A log-rank test was used to compare differences in survival times. All tests were two-sided and P<0.05 was considered to indicate a statistically significant difference. All numerical data are presented as the mean ± standard deviation of at least three independent experiments.

Results

NCAPH is upregulated in PCa. The expression patterns of NCAPH in ten types of human cancer types, including BRCA, HNSC, COAD, KIRC, LIHC, LUSC, LUAD, KIRP, STAD and PCa, were analyzed. It was identified that NCAPH was significantly overexpressed in cancer tissues compared with normal samples (Figs. 1A-I and 2A), which suggests that NCAPH serves as an oncogene in human cancer. The present study then focused on the clinical roles of NCAPH in PCa.

As presented in Fig. 2A and B and Table I, TCGA analysis demonstrated that NCAPH was upregulated in PCa tissues compared with non-tumor tissues. In addition, it was identified that NCAPH could serve as a specific and sensitive biomarker for PCa (area under the curve, 0.769; 95% confidence interval, 0.701-0.836; P<0.0001; Fig. 2C). To further validate this result, two independent public datasets GSE17951 (24) and GSE55945 (20) were analyzed. NCAPH was revealed to be overexpressed in PCa compared with normal samples in the GSE17951 dataset (P<0.001; Fig. 2D) and the GSE55945 dataset (P<0.05; Fig. 2E).
NCAPH is associated with pathological stage in PCa. The possible prognostic values of NCAPH were then evaluated using TCGA RNA-seq data. Associations between NCAPH mRNA expression and the clinicopathological characteristics of PCa are summarized in Table II. Positive expression of NCAPH was associated with age (Fig. 3A; Table II) and pathological stage (Fig. 3B and C; Table II). As presented in Fig. 3, NCAPH expression was significantly upregulated in T3/T4 PCa samples compared with T2 samples (Fig. 3B) and in N1 stage PCa samples compared with N0 samples (Fig. 3C). However, univariate analysis indicated that age, T stage and N stage were not associated with overall survival in PCa (Table III). These results indicate that NCAPH is associated with the pathological stage of PCa.

Upregulation of NCAPH predicts a poor prognosis in PCa. Furthermore, Kaplan-Meier analysis was performed to determine whether NCAPH expression is significantly associated with overall survival of patients with PCa. Two independent public datasets were analyzed, including the Taylor dataset (18) and TCGA dataset. The median NCAPH mRNA expression level in PCa tissues was used as the cut-off point to divide all cases into NCAPH-high (Taylor dataset, n=71; TCGA dataset, n=248) and NCAPH-low (Taylor dataset, n=71; TCGA dataset, n=247) groups. The samples with NCAPH expression higher than median value were identified as NCAPH-high, and the samples with NCAPH expression lower than median value were identified as NCAPH-low. In TCGA analysis, the overall survival rate was significantly higher in the NCAPH-low group compared with the NCAPH-high group (P=0.0012; Fig. 4A). In addition, similar results were observed with the Taylor dataset (P=0.0099; Fig. 4B). These results indicate that NCAPH expression could serve as a biomarker to predict the outcome of PCa.

Construction of NCAPH mediated PPI networks in PCa. The molecular functions of NCAPH in PCa remain largely unclear. Therefore, the current study performed PPI network analysis to reveal the potential roles of NCAPH. Using a NCAPH interacting proteins dataset which was download from NCBI (https://www.ncbi.nlm.nih.gov/gene/23397), a total of 51 candidate proteins were identified. By applying STRING analysis, 10 of the 51 proteins [NCAPG, NCAPD2, structural maintenance of chromosomes 4 (SMC4), structural maintenance of chromosomes 2 (SMC2), Aurora kinase A (AURKA), Aurora kinase B (AURKB), cyclin-dependent kinase 1 (CDK1), H2A, H2B, histone H3, histone H4] were identified as potential interactors of NCAPH. These results suggest that NCAPH may play a role in the regulation of cell cycle and DNA repair in PCa.
Table I. Expression levels of NCAPH in paired prostate cancer samples by using The Cancer Genome Atlas database.

<table>
<thead>
<tr>
<th>ID</th>
<th>Gene symbol</th>
<th>FC</th>
<th>P-value</th>
<th>Total samples</th>
<th>Samples unchanged</th>
<th>Samples upregulated</th>
<th>Samples downregulated</th>
</tr>
</thead>
<tbody>
<tr>
<td>23397</td>
<td>NCAPH</td>
<td>3.522</td>
<td>2.59x10^{-17}</td>
<td>38</td>
<td>10</td>
<td>27</td>
<td>1</td>
</tr>
</tbody>
</table>

NCAPH, non-SMC condensin I complex subunit H; FC, fold-change.

Figure 2. NCAPH expression level is upregulated in prostate cancer. (A) The Cancer Genome Atlas data analysis revealed that NCAPH expression levels were significantly upregulated in prostate cancer samples compared with normal prostate tissues. (B) The Cancer Genome Atlas data analysis revealed that NCAPH expression levels were significantly upregulated in 38 prostate cancer samples compared with match normal tissues. (C) ROC analysis revealed NCAPH could serve as a specific and sensitive biomarker in prostate cancer. (D) GSE17951 dataset analysis demonstrated that NCAPH expression levels were significantly higher in prostate tumor samples compared with normal prostate tissues (E) GSE55945 data analysis revealed NCAPH expression levels were significantly upregulated in prostate cancer samples compared with normal prostate tissues. *P<0.05, **P<0.001. NCAPH, non-SMC condensin I complex subunit H; ROC, receiver operating curve; AUC, area under the curve; CI, confidence interval.
Table II. Associations of NCAPH expression with clinicopathological characteristics of patients with prostate cancer in The Cancer Genome Atlas dataset.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>NCAPH expression level</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤62</td>
<td>147</td>
<td>121</td>
</tr>
<tr>
<td>&gt;62</td>
<td>102</td>
<td>127</td>
</tr>
<tr>
<td>Total</td>
<td>249</td>
<td>248</td>
</tr>
<tr>
<td>T stage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2</td>
<td>122</td>
<td>65</td>
</tr>
<tr>
<td>T3/4</td>
<td>123</td>
<td>180</td>
</tr>
<tr>
<td>Total</td>
<td>244</td>
<td>246</td>
</tr>
<tr>
<td>N stage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N0</td>
<td>174</td>
<td>171</td>
</tr>
<tr>
<td>N1</td>
<td>27</td>
<td>52</td>
</tr>
<tr>
<td>Total</td>
<td>201</td>
<td>223</td>
</tr>
</tbody>
</table>

NCAPH, non-SMC condensin I complex subunit H.

Table III. Univariate analysis of overall survival for patients with prostate cancer.

<table>
<thead>
<tr>
<th>Variable</th>
<th>$\chi^2$</th>
<th>df</th>
<th>Log-rank P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.741</td>
<td>1</td>
<td>0.389</td>
</tr>
<tr>
<td>N stage</td>
<td>3.101</td>
<td>1</td>
<td>0.078</td>
</tr>
<tr>
<td>T stage</td>
<td>2.068</td>
<td>1</td>
<td>0.150</td>
</tr>
</tbody>
</table>

df, degree freedom.

Figure 3. NCAPH expression is upregulated in high pathological T/N stage prostate cancer. (A) NCAPH expression was upregulated in patients with prostate cancer with an age >62 years compared with patients ≤62 years old. (B) NCAPH expression was significantly upregulated in T3/4 stage prostate cancer samples compared with T2 stage prostate cancer samples. (C) NCAPH expression was significantly upregulated in N1 stage prostate cancer samples compared with N0 stage prostate cancer samples. *P<0.05, **P<0.01, ***P<0.001. NCAPH, non-SMC condensin I complex subunit H.

Figure 4. Kaplan-Meier survival analysis of the association of overall survival rate with NCAPH expression in prostate cancer. Compared with patients with high NCAPH expression, the overall survival rates were higher for patients with low NCAPH expression in the (A) TCGA and (B) Taylor datasets. NCAPH, non-SMC condensin I complex subunit H; TCGA, The Cancer Genome Atlas.

histone family member Z (H2AFZ), POC1 centriolar protein A (POC1A) and histone cluster 2 H2A family member C (HIST2H2AC) were identified to interact with NCAPH with a combined score ≥0.4 and selected as seed genes.

PPI networks were then generated with the NCAPH interacting seed genes by applying STRING analysis. A total of 417 single nodes and 10,071 edges were included in the NCAPH-mediated PPI network. According to a previous study
by Zhang et al (25), the ‘K value’ of the PPI network was calculated to identify the hub networks in PCa. Finally, a total of six hub networks were revealed. As presented in Fig. 5, hub-network 1 contained 62 nodes and 1,325 edges (Fig. 5A), hub-network 2 consisted of 57 nodes and 627 edges (Fig. 5B), hub-network 3 contained 55 nodes and 418 edges (Fig. 5C), hub-network 4 consisted of 17 nodes and 57 edges (Fig. 5D), hub-network 5 included six nodes and 15 edges (Fig. 5E) and hub-network 6 was composed of 25 nodes and 65 edges (Fig. 5F).

Enrichment analysis of NCAPH in PCa. To investigate the biological characteristics of NCAPH in PCa, enrichment analysis based on the GO annotation system was performed. A total of 417 genes in the NCAPH-mediated PPI network were classified according to GO terms using DAVID. GO analysis demonstrated that NCAPH is predominantly involved in regulating PCa cell cycle progression and cell proliferation (Fig. 6A). A series of associated biological process terms, including ‘Cell division’, ‘Mitotic nuclear division’, ‘G2/M transition of mitotic cell cycle’, ‘G1/S transition of mitotic cell cycle’, ‘Nucleosome assembly’, ‘DNA repair’, ‘Negative regulation of gene expression, epigenetic’, ‘DNA replication’ and ‘Sister chromatid cohesion’ were enriched in GO analysis. Molecular function analysis demonstrated NCAPH is associated with ‘Protein binding’, ‘Histone binding’ and ‘Chromatin binding’ (Fig. 6B).

Discussion

PCa is one of the most frequently diagnosed cancer types worldwide (8,26). However, to the best of our knowledge, there remains a lack of accurate biomarkers to predict the outcome of PCa. The present study identified that NCAPH is significantly upregulated in tumor samples and could serve as a biomarker for PCa (14). To reveal the potential roles of NCAPH in PCa, NCAPH-mediated PPI networks were constructed. Bioinformatics analysis was also performed to predict the potential functional roles of NCAPH in PCa.

NCAPH, NCAPG and NCAPD2 form the non-SMC condensin I complex, which serves crucial roles in chromosome-wide gene regulation in mitotic and meiotic cell cycles by controlling chromosome organization (1). Previous studies have indicated the important roles of the condensin I complex in human cancer (4-6). NCAPH has been reported to be upregulated in melanoma and colon cancer (5,6). The present study demonstrated that NCAPH is upregulated in PCa samples by analyzing three public datasets, including TCGA data, GSE17951 and GSE55945. In addition, it was identified that NCAPH is upregulated in T3/T4 PCa samples compared with T2 samples and upregulated in N1 stage PCa samples compared with N0 samples. Furthermore, Kaplan-Meier analysis revealed that patients with PCa with low NCAPH expression exhibited longer overall survival times compared with patients with high NCAPH expression. The present analysis was similar to a previous study regarding human colon cancer (5). Yin et al (5) also observed that NCAPH expression is higher in colon cancer tissues compared with adjacent normal tissues. However, the higher NCAPH expression in colon cancer was associated with an improved prognosis and longer survival time. One limitation of the current study should be noted; the detailed clinical information of the patients in the four public datasets was not completed. To the best of our knowledge, the present study reported for the first time that NCAPH is associated with the prognosis of PCa.

A number of studies have demonstrated that oncogenes or tumor-suppressive genes cooperate in interconnected pathways and molecular networks at different levels to regulate cancer progression (27,28). The present study constructed a NCAPH-mediated PPI network to reveal the
associations of NCAPH in PCa, which included 417 single nodes and 10,071 edges. Abnormal cell cycle progression is a hallmark of human cancer (29,30). Therefore, targeting cell-cycle checkpoints could provide novel strategies for cancer treatment. Notably, GO and KEGG analysis revealed that NCAPH predominantly participates in regulating the PCa cell cycle and proliferation by affecting a series of associated biological processes, including cell division, mitotic nuclear division, G2/M transition of mitotic cell cycle and G1/S transition of mitotic cell cycle. Of note, Yin et al (5) also demonstrated that knockdown of NCAPH could induce cell cycle arrest at the G2/M phase. The present study together with other previous studies indicate that NCAPH-mediated cell cycle regulation is a mechanism involved in PCa proliferation. These results suggest NCAPH may be a therapeutic target for PCa.

In conclusion, the present study demonstrated for the first time that NCAPH was significantly upregulated in PCa. Furthermore, Kaplan-Meier analysis revealed that overexpression of NCAPH is associated with poor survival of patients with PCa. Bioinformatics analysis revealed NCAPH is involved in regulating PCa cell cycle by interacting with ten proteins, including NCAPD2, NCAPG, SMC4, SMC2, AURKA, AURKB, CDK1, H2A-FZ, POC1A, HIST2H2AC. In summary, the current results suggest NCAPH may be a potential beneficial and novel diagnostic and therapeutic marker in PCa.

Acknowledgements
Not applicable.

Funding
The present study was supported by the Social Development Plan of Jiangsu Province-Standardization of Key Disease Diagnosis and Treatment Projects (grant no. BE2016715) and the Jiangsu Province Youth Medical Key Talent Program (grant no. QNR2016457).

Availability of data and materials
The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authors’ contributions
FC and JH were responsible for the study conception and design. FC, JH and HT developed the methodology. FC, JT and ZX analyzed and interpreted the data. FC and JH wrote, reviewed and revised the manuscript.

Ethics approval and consent to participate
Not applicable.

Patient consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

References


