

The Impact of High-mobility Group Box Mutation of T-cell Factor 4 on Its Genomic Binding Pattern in Non–small Cell Lung Cancer



Hongjian Su, Yahong Qiao, Zhuona Xi, Jifang Wang and Zhen Bao

The Second Department of Respiratory and Critical Care, Henan Provincial Chest Hospital, 450008, China

Abstract

T-cell factor 4 (TCF-4) is determined to play a crucial role in Wnt/ β -catenin signaling pathway activation. The mutations and alternative splice isoforms of TCF-4 can cause cancers and other diseases. The high-mobility group (HMG) box domain of TCF-4 contributes to interacting with DNA motif for transcriptional regulation. However, the impact of the mutations within HMG box of TCF-4 on the genomic binding pattern is poorly investigated. Herein, we generated non–small cell lung cancer (NSCLC) cell line A549 with stably overexpressed TCF-4 with HMG box hot spot mutation (10th exon partial deletion), and conducted TCF-4 and β -catenin chromatin immunoprecipitation sequence to explore the differential genomic binding patterns. Our results revealed that TCF-4 lost 19365 but gained 1724 peaks, and β -catenin lost 4035 but gained 5287 peaks upon mutant TCF-4 compared with the wild type ($\log_2FC > 1$ or < -1 , $FDR < 0.01$). The transcriptional levels of the genes associated with these differential peaks such as H3F3C, KRT1, KRT14, MMP1, and MMP15 were all found to strongly change responding to TCF-4 binding ($P < 0.05$). Furthermore, A549 cells with TCF-4 mutation displayed a more compromising tumor characterization on cell proliferation and invasion. Our data determined the important role of TCF-4 in gene transcription controlling and provided the gain function evidence of TCF-4 caused by the TCF-4 mutation in NSCLC.

Translational Oncology (2020) 13, 79–85

Background

T-cell factor 4 (TCF-4) is a member of the TCF/lymphoid enhancer factor (LEF) family, which can recognize a consensus motif of Ephrussi-box binding site in the promoter region [1]. TCF-4 together with β -catenin coactivator, functions as the major transcriptional mediator of the canonical wingless/integrated (Wnt) signaling pathway, which is one of the pivotal mechanisms involved in the progression of the multiple cancers [2–4]. Usually, TCF-4 is extensively expressed in various cancers and is closely correlated with

cancer progression, and silencing the aberrant expression of TCF-4 can efficiently repress the tumor cell growth and proliferation [5], which indicates that TCF-4 plays an oncogenic role and can be served as a potential therapeutic target for cancer treatment. Previous studies revealed that TCF-4 was causally connected with the development of lung carcinoma [6] by regulating multiple downstream genes, such as MMP15 [7], RUNX3, TMEM88, and APC [8].

TCF-4 consists of three major domains including 1st exon for the β -catenin–binding domain, 10th and 11th exons for the DNA–binding high-mobility group (HMG) boxes, and the 17th exon for C-terminal binding domain [6]. The β -catenin–binding domain contributes to the interaction with β -catenin to form the complex and activate the transcription of target genes. C-terminal binding domain produces a variety of TCF-4 splicing isoforms in different types of cancers [9], which strongly affects the tumor malignancy and patients survival status [10]. HMG box plays a crucial role in DNA motif recognition and binding. As a DNA-binding protein, TCF-4 acts as

Address all correspondence to: Yahong Qiao, No. 1, Weiwu Road, Jinshui District, Zhengzhou City, Henan Province, 450008, China. E-mail: qiaoyahong@yeah.net
Received 24 July 2019; Revised 25 September 2019; Accepted 26 September 2019

© 2019 The Authors. Published by Elsevier Inc. on behalf of Neoplasia Press, Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).
1936-5233/19
<https://doi.org/10.1016/j.tranon.2019.09.012>

an independent transcriptional regulator unless TCF-4 recruits other transcriptional factors to activate gene transcription. One study reported that exon 11 played a crucial role in binding activity on the canonical binding site in lung cancer [6], and the entire loss of exon 11 caused the transcriptional suppression of Wnt signaling targets, and could mimic the antitumor effect of Wnt signaling inhibitor. However, how the mutations of HMG box impact on the DNA motif preference and genomic DNA binding still remain obscure. In this study, we generated the HMG box mutation of TCF-4 and investigated the genomic binding pattern of TCF-4 and β -catenin in non-small cell lung cancer (NSCLC) cells. Our study tried to reveal the function loss and gain of TCF-4 derived from the DNA-binding domain mutation and illustrated the carcinogenic mechanism of Wnt/ β -catenin signaling pathway.

Methods

TCGA Data Mining

TCGA PanCancer Atlas studies with 32 types of cancer (10528 cases samples) [11] were picked out to investigate TCF-4 mutations and compare the relationship with tumor malignancy. The data were output from cBioPortal (<https://www.cbioportal.org/>) [12,13].

Cell Line Preparation

A549 NSCLC cells obtained from the Type Culture Collection of the Chinese Academy of Sciences (Beijing, China) were cultured in Dulbecco's Modified Eagle Medium with 10% fetal bovine serum (Thermo Fisher Scientific, USA) at the incubator with 37 °C, 5% CO₂, and 100% humidity. Full length TCF-4 (1809 bp) was cloned from cDNA of A549 cells and inserted into pEGFP-N1-FLAG (#60360, Addgene, USA). Partial 10th exon (1129–1164 bp) was deleted by mutagenesis polymerase chain reaction (PCR) based on the wild type TCF-4. The plasmids of wild type TCF-4 and TCF-4- $\Delta_{1129-1164}$ were transfected into A549 cells, respectively using the X-tremeGENE HP DNA Transfection Reagent (Roche Diagnostics, Basel, Switzerland) in accordance with the manufacturer's instructions. A549 cells were harvested after 72 h transfection for the next experiments.

Chromatin Immunoprecipitation Sequence

The chromatin immunoprecipitation (ChIP) assay was performed as described [14]. In brief, 10% whole cell lysates were saved as input after genomic DNA was broken into 200–500 bp by sonication. One microgram of immunoprecipitation (IP)-grade antibodies of FLAG or β -catenin were incubated with the rest of the lysate overnight, followed by 2 h protein-A beads incubation at 37 °C for target protein pull down. The FLAG or β -catenin-enriched DNA or input DNA were repaired to 3'-dA overhang and added the ligated adapter. The DNA library eliminated the unligated adapters, and the appropriate size for sequence was selected using an Illumina Hiseq 2000 platform. The raw sequence reads of input and IP were trimmed adapters and filter out low quality reads using Cutadapt (version 1.9.1) and Trimmomatic (version 0.35) [15] and checked the quality of clean reads using Fastqc [16]. Then, clean reads were mapped to the human genome (assembly hg38) using the Bowtie 2 (version 2.2.6) algorithm [17]. The process of peak calling ($P < 0.01$) were performed by MACS 2 (version 2.1.1) [18] and analyzed the different binding domains based on false discovery rate (FDR) value less than 0.05 and annotated by DiffBind [19]. De novo motif were analyzed using the R language and MEME [20–22]. The peaks on certain genomic loci

were visualized by Integrative Genomics Viewer. Gene ontology (GO) analysis was used to interpret the biological function of the genes associated with differential peaks [23]. TCF-4 and β -catenin sequencing data were deposited to sequence read archive (SRA) database assigned with the accession number PRJNA551776.

Real-time PCR

Total RNA of A549 cells were isolated using the TRIzol (Thermo Fisher Scientific). DNA templates taken from RNA reverse transcription was detected by the target gene transcription using the QuantStudio 3 system (Thermo Fisher Scientific). In accordance with the given instructions, 95 °C for 30 s for initial denaturation, followed by 40 cycles at 95 °C for 5 s, appropriate annealing temperatures of 10 s and 72 °C, then 30 s were setup for PCR conditions. Ct values were harvested and calculated by using the delta-delta method. The expression of glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was used as a quality control. All primers used in this study were listed in [Supplementary Table 1](#).

Cell Proliferation CCK-8 Assay

A549 cells after 24 h transfection were cultured in 96-well plate with 80% density, followed by adding 10 μ l CCK-8 solution (Solarbio, China) to incubate for an additional 1 h. Absorption values of 450 nm were examined using Multiskan FC microplate reader (Thermo Fisher Scientific, USA) to construct the regression equation and calculate IC₅₀. The IC₅₀ values evaluated the cell proliferation.

Transwell Cell Migration Assay

1×10^5 A549 cells after 24-h transfection were cultured within 200 μ l suspension in upper and 800 μ l fresh medium in lower transwell chamber (Corning, USA) on 24-well plates for 24 h. The cells at the lower chamber were cross-linked by 1% paraformaldehyde for 10 min and stained by 0.5% crystal violet (Sinopharm Chemical Reagent, China) for 5 min. The stained cells were counted under a light microscope to evaluate the cell invasion.

Statistical Analysis

The experimental data were processed with SPSS 20 software. χ^2 test was used to analyze the correlation between TCF-4 mutation and tumor malignancy. Student's t-test was used for comparison of the difference between wild type and mutant TCF-4 groups. P value less than 0.05 was considered as statistical significance.

Results

The Relationship Between TCF-4 Mutations and Cancer Progression *in vivo*

Initially, we investigated the mutation atlas of the coding region of TCF-4 in cancer patients through TCGA database mining. Except 13 fusion protein, 169 point mutations (1.6%) were found in total 32 types of cancer (10528 cases samples), and 59 were involved in β -catenin-binding domain, 30 were distributed in HMG box whereas the rest of 80 mutations were in other coding region. Furthermore, we analyzed the correlation between the location of TCF-4 mutations and tumor malignancy, and we found that the tumors with TCF-4 mutations at HMG box displayed a lower malignancy of tumor than the ones with the mutations at β -catenin-binding domain ($\chi^2 = 3.979$, $P = 0.038$) and other regions ($\chi^2 = 8.317$, $P = 0.003$) ([Figure 1A](#)). Moreover, the overall survival status was also compared among these three groups of TCF-4

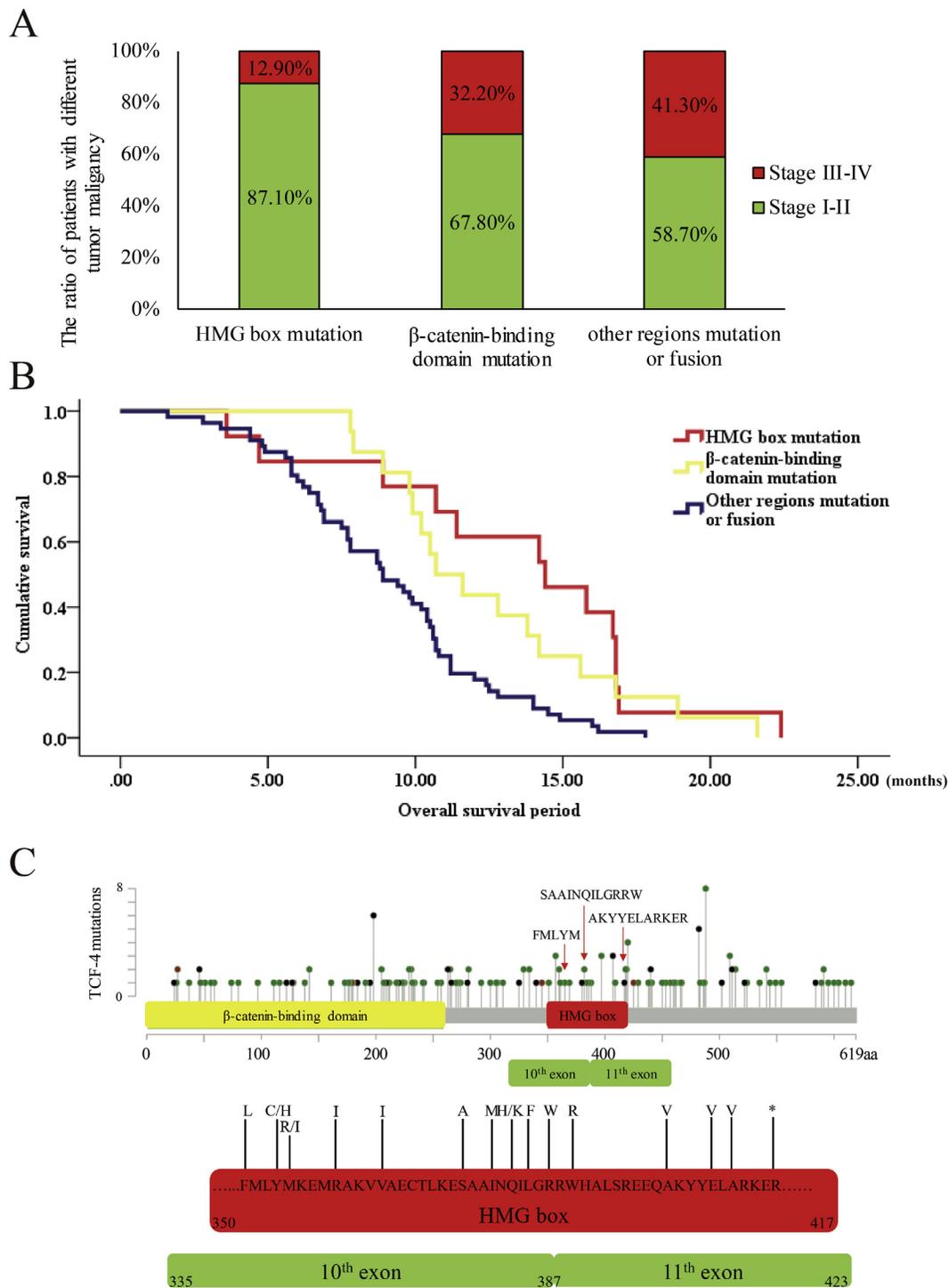


Figure 1. The impact of TCF-4 mutations on human cancer progression. (A) The distribution of human cancer malignancy from TCGA database grouped by TCF-4 mutations in different regions. χ^2 analysis was used to evaluate the difference among these groups. (B) The survival period of cancer patients from TCGA database group by TCF-4 mutations in different regions. (C) The overview mutation landscape of TCF-4 coding region from TCGA database (top), and the zoom-in view at HMG box (below). The black, brown and green spots represent truncation, inframe, and missense mutation, respectively. TCF-4, T-cell factor 4; HMG, high-mobility group.

mutations within β -catenin-binding domain, HMG box domain as well as other regions. We observed that both the patients of TCF-4 mutations within β -catenin-binding domain ($P = 0.011$) and HMG box domain ($P = 0.009$) displayed higher overall survival rate than

the ones of TCF-4 mutations within other regions (Figure 1B). We also noticed that the hot spot mutation at HMG box focused on the key DNA-binding sites (“FMLYM” from 357 to 361 and “SAAINQLGRRW” from 377 to 388 at the 10th exon, although

“AKYYELARKER” from 397 to 407 at the 11th exon) (Figure 1C). Taken together, our results determined that TCF-4 mutations within HMG box was supposed to be benefit for tumor suppression.

Genomic Binding Patterns of HMG Box Mutation of TCF-4

We observed that there are more lung cancer patients with mutations within “SAAINQILGRRW” than other two regions (“SAAINQILGRRW”: 7 cases; “FMLYM”: 3 cases; “AKYYELARKER”: 2 cases). Consequently, we generated the ectopic plasmids of TCF-4 with wild type and “SAAINQILGRRW” deletion (TCF-4 $\Delta_{1129-1164}$) both combined with FLAG-tag. Sanger sequence validated the mutagenesis of TCF-4 (Supplementary Figure 1). A549 NSCLC cells were transfected to overexpress wild type or TCF-4 $\Delta_{1129-1164}$, respectively, and a pull-down assay was subsequently performed using FLAG-tag or β -catenin antibodies to

conduct the deep sequencing for mapping genome-wide TCF-4 profiles in A549 NSCLC cells. Approximately, 82.23 M reads of TCF-4 and 80.37 M reads of β -catenin ChIP sequencing data were collected, respectively (Supplementary Table 2). We observed that the genomic occupancy of TCF-4 was remarkably dependent on these mutant binding sites. TCF-4 $\Delta_{1129-1164}$ substantially weakened the genomic enrichment compared with the wild type group ($P = 0.002$) (Figure 2A). TCF-4 $\Delta_{1129-1164}$ lost the occupancy at 19365 regions but gained additional 1724 peaks compared with wild type TCF-4 ($\log_2FC > 1$ or < -1 , $FDR < 0.01$) (Figure 2B). Nevertheless, 51.77% of total genomic binding regions of TCF-4 $\Delta_{1129-1164}$ were primarily concentrated on promoter regions, which was higher than 21.94% of wild type TCF-4 ($P < 0.001$). However, the overall enrichment of β -catenin on genome seemingly displayed no significant change responding to wild type TCF-4 or TCF-4 $\Delta_{1129-1164}$, although

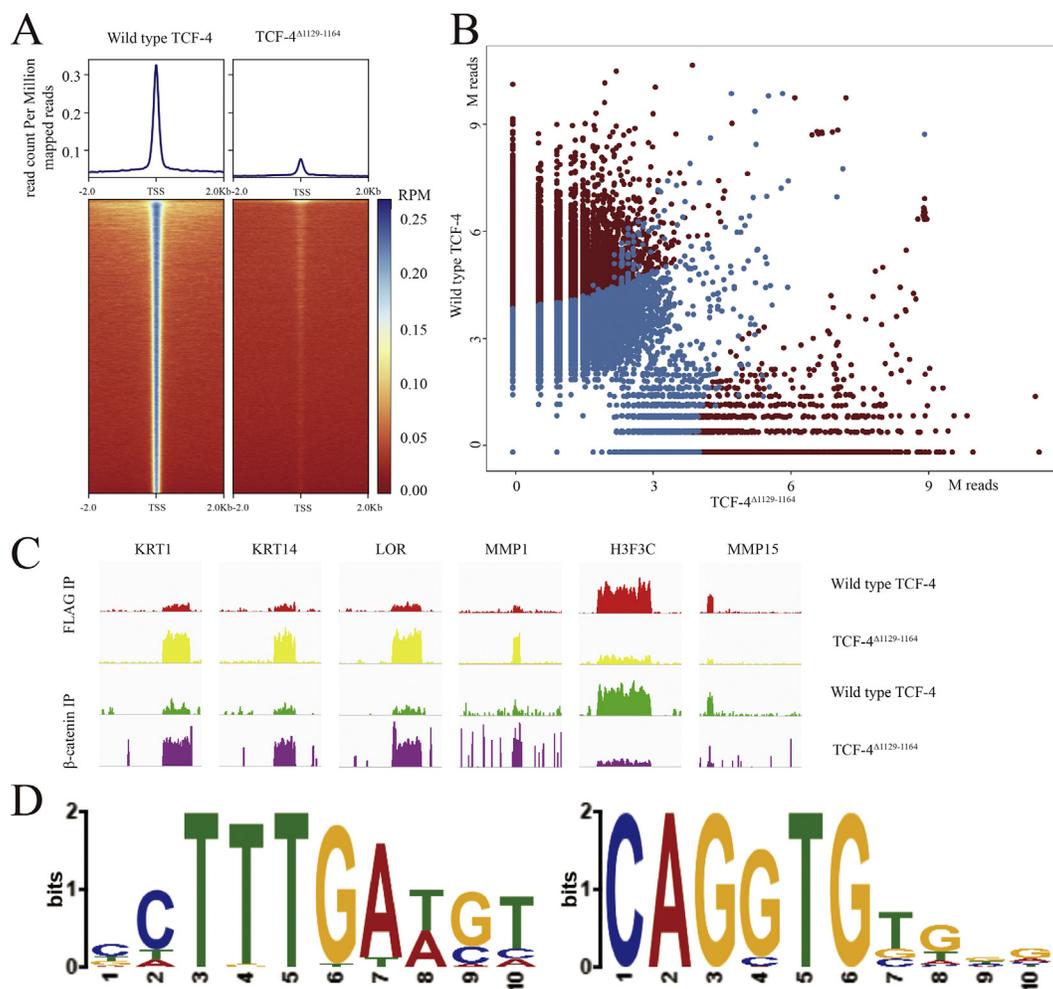


Figure 2. The genomic binding patterns of HMG box mutation of TCF-4. (A) The genome-wide occupancy and heatmap representation of TCF-4 at all annotated gene promoters in A549 cells determined by ChIP-seq. Average TCF-4 enrichment measured by \log_2 (peak P values) in 200-bp bins is shown within genomic regions covering 2 kb up- and downstream of TSSs. (B) The volcano representation of the different peaks compared between wild type TCF-4 and TCF-4 $\Delta_{1129-1164}$. Each brown spot mean a significantly different peak, although blue spot mean the peaks without statistical significance. (C) Enrichment of TCF-4 in A549 cells with wild type TCF-4 (red) or TCF-4 $\Delta_{1129-1164}$ (yellow) overexpression and enrichment of β -catenin in A549 cells with wild type TCF-4 (green) or TCF-4 $\Delta_{1129-1164}$ (purple) overexpression at KRT1, KRT14, LOR, MMP1, H3F3C and MMP15. ChIP-seq data are shown in reads per million with the y-axis floor set to 0.5 reads per million. (D) DNA motifs that is enriched in wild type TCF-4 (Left) and TCF-4 $\Delta_{1129-1164}$ (Right) bound loci in A549 cells. HMG, high-mobility group; TCF-4, T-cell factor 4; ChIP-seq, chromatin immunoprecipitation sequence; TSSs, transcription start sites.

β -catenin lost 4035 peaks but gained another 5287 peaks upon TCF-4 Δ ^{1129–1164} compared with the wild type (Supplementary Figure 2A and B). The examples of two different tendency of TCF-4 binding status such as KRT1, KRT14, LOR, MMP1 as well as H3F3C, MMP15 compared between wild type TCF-4 and TCF-4 Δ ^{1129–1164} were illustrated (Figure 2C). Moreover, we observed that the TCF-4 Δ ^{1129–1164} displayed a varying preferential DNA motif compared with the wild type TCF-4 (Figure 2D). Collectively, our ChIP-seq data determined a distinct pattern of TCF-4 binding at genome upon HMG box partial deletion.

The Tumor Repressive Effect of HMG Box Mutation of TCF-4 on NSCLC Cells

To further investigate the biological impact of differential genomic interaction derived from the TCF-4 Δ ^{1129–1164}, GO analysis was used to analyze the function of genes associated with the differential occupancy of TCF-4 Δ ^{1129–1164}. We observed that the function of angiogenesis, cell growth, cell morphogenesis, and multiple pathways including MAPK, PI3K-Akt, and Wnt were involved in the genes associated with the differential peaks from TCF-4 pull down compared between wild type TCF-4 and TCF-4 Δ ^{1129–1164} (Figure 3A and B). The transcription of associated genes such as KRT1, KRT14, and MMP1 were activated although H3F3C and MMP15 were inactivated upon TCF-4 Δ ^{1129–1164} compared with wild type validated by quantitative polymerase chain reaction (qPCR) (Figure 3C), which was consistent with the data of ChIP-seq (Figure 2C). Finally, the ability of cell proliferation and invasion was explored to evaluate the overall efficacy of TCF-4 Δ ^{1129–1164} in A549 cells by CCK-8 and transwell assay (Figure 3D and E). The presence of the compromising cell proliferation and invasion was observed in A549 cells with TCF-4 Δ ^{1129–1164} compared with wild type, which was coincided with the TCGA data above *in vivo*.

Taken together, our results revealed that TCF-4 Δ ^{1129–1164} affected its genomic binding status and could change the activation

of multiple signaling pathways and induce the tumor repression in NSCLC cells.

Discussion

TCF-4 is broadly expressed in a variety of cell types, including lung, bladder, brain, fat, endometrium, ovary, and placenta [24] and functions in many cell lineage-specific functions, such as development of lymphocytes, neurogenesis, myogenesis, erythropoiesis, and melanogenesis [1,25]. Of all the Wnt signaling pathway components, the TCF/LEF family works with β -catenin as the coactivators to initiate the transcription of Wnt targets. Previous studies have indicated that HMG box encoded by 10th and 11th exons exerted as a DNA-binding domain [9,26]. More studies focused on the protein structure and function of TCF-4/catenin complex [27,28]. However, it remained obscure how the binding activity between TCF-4 protein and target genes was regulated especially when HMG box mutation occurred.

The overactivation of Wnt signaling pathway is well acknowledged to promote cell growth and invasion, thereby strongly linking with carcinogenesis [5,29,30]. However, we noticed that the cancer patients with HMG box mutation of TCF-4 showed a characterization of compromising tumor malignancy and improving survival status compared with the wild type or mutations at other region of TCF-4 (Figure 1A and B). These results indicate that HMG box of TCF-4 may be a crucial part to limit cancer progression more than β -catenin-binding domain and other regions. It makes sense that the Wnt pathway even if activated cannot enhance the transcription of the downstream targets. However, if the HMG box is completely deleted, this pathway is supposed to break down, and the cells fall into the greater chaos. Considering about the pattern of TCF-4 mutation in lung cancer from TCGA database, we generated TCF-4 with HMG partial (SAAINQLGRRW) deletion containing most of the hot spot mutation *in vivo*. Consistently, the ChIP-seq data directly

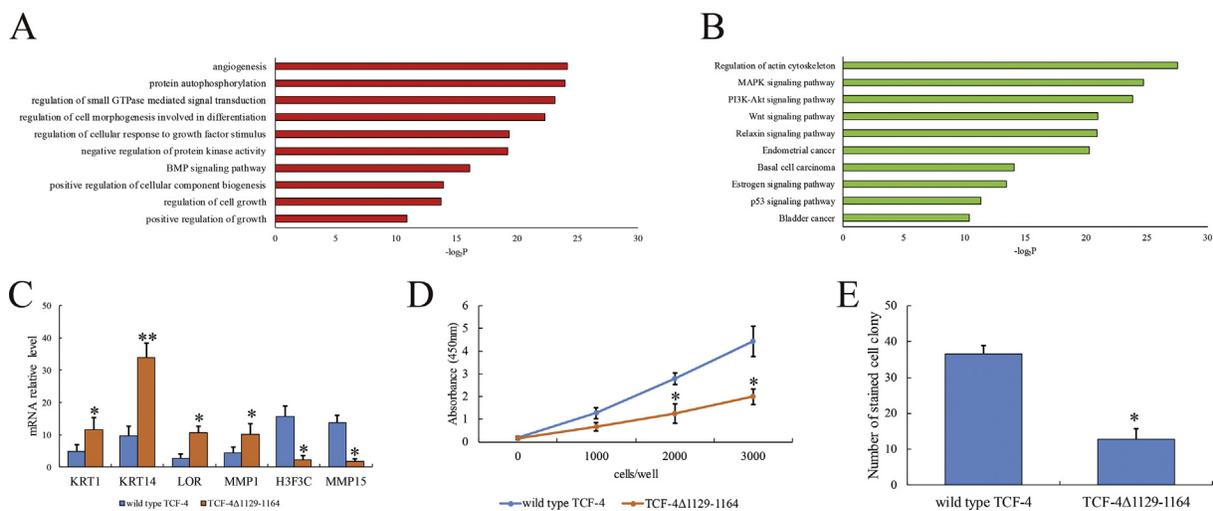


Figure 3. The tumor suppressive gained by TCF-4 Δ ^{1129–1164}. (A) Top ten biological process (GO analysis) of the genes associated with different peaks is presented. (B) Top ten signaling pathway (KEGG analysis) of the genes associated with different peaks is presented. (C) The transcriptional levels of H3F3C, KRT1, KRT14, MMP1, and MMP15 in A549 cells with wild type TCF-4 or TCF-4 Δ ^{1129–1164} overexpression. “**” means the statistical significance with *P* value less than 0.05. Each experiment contains three replicates. (D) The cell growth and (E) the cell invasion of A549 cells with wild type TCF-4 or TCF-4 Δ ^{1129–1164} overexpression. “**” and “***” mean the statistical significance with *P* value less than 0.05 and 0.01, respectively. Each experiment contains three replicates. GO, Gene ontology; TCF-4, T-cell factor 4; KEGG, Kyoto encyclopedia of genes and genomes.

demonstrate that HMG box mutation decreases the occupancy of genomic DNA compared with wild type in our model (Figure 2A).

Unexpectedly, a new preferential DNA motif recognized by TCF-4^{Δ1129–1164} (Figure 2D) was observed compared with wild type 5'-CTTTGWWS-3' suggested by previous study [31]. We speculate that this deletion may cause a changed HMG box of TCF-4. The currently known TCF-4 structures (PDB No.: 1JPW and 2GL7, 1HSM) only displayed a patchy peptide associated with β-catenin-binding domain without HMG box and indicated little information on the protein structure between HMG box and interacted DNA. Given the new genomic regions bound by TCF-4^{Δ1129–1164} compared with wild type from ChIP-seq data, although we cannot figure out the new protein structure based on this mutation, the TCF-4 gain of function is assumed and further studied. Previous study reported that the single nucleotide polymorphism (SNP) on TCF-4 enhanced the binding affinity with high mobility group box 1 (HMGB1) [32], which implied a potential gain of function derived from mutation. Now, our results reveal that TCF-4^{Δ1129–1164} produces a new recognition on the transcription of some genes such as KRT1, KRT14, LOR, and MMP1 (Figure 3A–C), which is consistent with previous studies [7,33]. Therefore, TCF-4^{Δ1129–1164} is supposed to gain a novel function in cancer cells. The overall compromising effect of cell growth and invasion of A549 cells upon TCF-4^{Δ1129–1164} are observed (Figure 3D and E).

In addition, the phenotype of β-catenin on genomic binding and transcription activity is not affected by TCF-4 (Supplementary Figure 2). This way of β-catenin is determined to be partially independent with TCF/LEF on genomic binding and transcriptional activation. The molecular mechanisms underlying β-catenin-independent Wnt signaling cascades and their implications for cell biology, development, and physiology was characterized previously [34]. However, it is hard to figure out the regulatory way of β-catenin in this study because our cell model is based on the ectopic TCF-4 overexpression without the endogenous TCF-4 deletion. We believe the targeted genomic fragments from β-catenin pull down is partially overlapped with endogenous TCF-4 in A549 cells.

Taken together, our results determine an important role of HMG box of TCF-4 in tumor suppression of NSCLC cancer. The HMG box of TCF-4 or the novel downstream genes gained by mutant TCF-4 are worthy of exploitation as the possible tumor treatment targets.

Acknowledgements

This project is not supported by any funds. The authors afforded all the reagents used in this study. All the authors have consented to make the data and materials available and publish this manuscript. The authors declare no competing financial interests.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.tranon.2019.09.012>.

References

- [1] Sepp M, Kannike K, Eesmaa A, Urb M and Timmusk T (2011). Functional diversity of human basic helix-loop-helix transcription factor TCF4 isoforms generated by alternative 5' exon usage and splicing. *PLoS One* **6**(7). e22138.
- [2] Davidsen J, Larsen S, Coskun M, Gogenur I, Dahlgaard K and Bennett EP, et al (2018). The VTI1A-TCF4 colon cancer fusion protein is a dominant

- negative regulator of Wnt signaling and is transcriptionally regulated by intestinal homeodomain factor CDX2. *PLoS One* **13**(7). e0200215.
- [3] Duan J, He Y, Fu X, Deng Y, Zheng M and Lu D (2018 Sep 15). CDK7 activated beta-catenin/TCF signaling in hepatocellular carcinoma. *Exp Cell Res* **370**(2), 461–467.
- [4] Li R, Liu S, Li Y, Tang Q, Xie Y and Zhai R (2018). Long noncoding RNA AFAP1AS1 enhances cell proliferation and invasion in osteosarcoma through regulating miR46955p/TCF4/betacatenin signaling. *Mol Med Rep* **18**(2), 1616–1622.
- [5] Krishnamurthy N and Kurzrock R (2018). Targeting the Wnt/beta-catenin pathway in cancer: update on effectors and inhibitors. *Cancer Treat Rev* **62**, 50–60.
- [6] Yang B, Xu Y, Wang X and Liu Y (2019). A serine in exon 11 determines the full transcriptional activity of TCF-4 in lung carcinoma cells. *Biochem Biophys Res Commun* **508**(3), 675–681.
- [7] Liu Y, Xu Y, Guo S and Chen H (2016). T cell factor-4 functions as a co-activator to promote NF-kappaB-dependent MMP-15 expression in lung carcinoma cells. *Sci Rep* **6**, 24025.
- [8] Stewart DJ (2014). Wnt signaling pathway in non-small cell lung cancer. *J Natl Cancer Inst* **106**(1). djt356.
- [9] Duval A, Rolland S, Tubacher E, Bui H, Thomas G and Hamelin R (2000). The human T-cell transcription factor-4 gene: structure, extensive characterization of alternative splicings, and mutational analysis in colorectal cancer cell lines. *Cancer Res* **60**(14), 3872–3879.
- [10] Kriegl L, Horst D, Reiche JA, Engel J, Kirchner T and Jung A (2010). LEF-1 and TCF4 expression correlate inversely with survival in colorectal cancer. *J Transl Med* **8**, 123.
- [11] Cancer Genome Atlas Research N, Weinstein JN, Collisson EA, Mills GB, Shaw KR and Ozenberger BA, et al (2013). The cancer genome atlas pan-cancer analysis project. *Nat Genet* **45**(10), 1113–1120.
- [12] Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B and Sumer SO, et al (2013). Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal* **6**(269), p11.
- [13] Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO and Aksoy BA, et al (2012). The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* **2**(5), 401–404.
- [14] Fei P, Wang W, Kim SH, Wang S, Burns TF and Sax JK, et al (2004). Bnip3L is induced by p53 under hypoxia, and its knockdown promotes tumor growth. *Cancer Cell* **6**(6), 597–609.
- [15] Bolger AM, Lohse M and Usadel B (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**(15), 2114–2120.
- [16] Andrews S (2013). FastQC a quality control tool for high throughput sequence data. 2013.
- [17] Langmead B and Salzberg SL (2012). Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**(4), 357–359.
- [18] Zhang Y, Liu T, Meyer CA, Eeckhoutte J, Johnson DS and Bernstein BE, et al (2008). Model-based analysis of chip-seq (MACS). *Genome Biol* **9**(9), R137.
- [19] Ross-Innes CS, Stark R, Teschendorff AE, Holmes KA, Ali HR and Dunning MJ, et al (2012). Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* **481**(7381), 389–393.
- [20] Bailey TL, Johnson J, Grant CE and Noble WS (2015). The MEME suite. *Nucleic Acids Res* **43**(W1), W39–W49.
- [21] Li H, Su X, Gallegos J, Lu Y, Ji Y and Mollndrem JJ, et al (2012). dsPIG: a tool to predict imprinted genes from the deep sequencing of whole transcriptomes. *BMC Bioinf* **13**.
- [22] Kong H, Tong P, Zhao X, Sun J and Li H (2018). CAsubtype: an R package to identify gene sets predictive of cancer subtypes and clinical outcomes. *Interdiscip Sci* **10**(1), 169–175.
- [23] Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL and Gillette MA, et al (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**(43), 15545–15550.
- [24] Fagerberg L, Hallstrom BM, Oksvold P, Kampf C, Djureinovic D and Odeberg J, et al (2014). Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol Cell Proteom : MCP* **13**(2), 397–406.
- [25] In 't Hout FEM, van Duren J, Monteferrario D, Brinkhuis E, Mariani N, Westers TM, et al., editors. TCF4 promotes erythroid development. *Experimental hematology*; 2018.

- [26] Shiina H, Igawa M, Breault J, Ribeiro-Filho L, Pookot D and Urakami S, et al (2003). The human T-cell factor-4 gene splicing isoforms, Wnt signal pathway, and apoptosis in renal cell carcinoma. *Clin Cancer Res* **9**(6), 2121–2132.
- [27] Poy F, Lepourcellet M, Shivdasani RA and Eck MJ (2001). Structure of a human Tcf4-beta-catenin complex. *Nat Struct Biol* **8**(12), 1053–1057.
- [28] Sampietro J, Dahlberg CL, Cho US, Hinds TR, Kimelman D and Xu W (2006). Crystal structure of a beta-catenin/BCL9/Tcf4 complex. *Mol Cell* **24**(2), 293–300.
- [29] Taciak B, Pruszyńska I, Kiraga L, Bialasek M and Krol M (2018). Wnt signaling pathway in development and cancer. *J Physiol Pharmacol* **69**(2).
- [30] Tortelote GG, Reis RR, de Almeida Mendes F and Abreu JG (2017). Complexity of the Wnt/betacatenin pathway: searching for an activation model. *Cell Signal* **40**, 30–43.
- [31] Hoverter NP, Zeller MD, McQuade MM, Garibaldi A, Busch A and Selwan EM, et al (2014). The TCF C-clamp DNA binding domain expands the Wnt transcriptome via alternative target recognition. *Nucleic Acids Res* **42**(22), 13615–13632.
- [32] Zhou Y, Oskolkov N, Shcherbina L, Ratti J, Kock KH and Su J, et al (2016). HMGB1 binds to the rs7903146 locus in TCF7L2 in human pancreatic islets. *Mol Cell Endocrinol* **430**, 138–145.
- [33] Chen P, Sun S, Zeng K, Li C, Wen J and Liang J, et al (2018). Exome sequencing identifies a TCF4 mutation in a Chinese pedigree with symmetrical acral keratoderma. *J Eur Acad Dermatol Venereol* **32**(7), 1204–1208.
- [34] Acebron SP and Niehrs C (2016). Beta-catenin-independent roles of Wnt/ LRP6 signaling. *Trends Cell Biol* **26**(12), 956–967.