



Histone methyltransferase KMT2D sustains prostate carcinogenesis and metastasis via epigenetically activating LIFR and KLF4

Shidong Lv^{1,2} · Liyan Ji² · Bin Chen³ · Shuqiang Liu² · Chengyong Lei¹ · Xi Liu² · Xiaoxiao Qi² · Ying Wang² · Elaine Lai-Han Leung⁴ · Hongyi Wang¹ · Lin Zhang¹ · Xiaoming Yu¹ · Zhongqiu Liu^{2,4} · Qiang Wei¹ · Linlin Lu^{2,4}

Received: 26 May 2017 / Revised: 14 October 2017 / Accepted: 19 October 2017 / Published online: 22 December 2017
© The Author(s) 2018. This article is published with open access

Abstract

Abnormalities in epigenetic modifiers are emerging as driving events in prostate cancer (PCa). The histone methyltransferase KMT2D, a frequently aberrant epigenetic modifier in various tumors, has an undefined role in PCa. Moreover, little is known regarding KMT2D's mutation in Chinese patients or its downstream signaling pathways and targets. Here, we profiled the mutational spectrum of 32 significantly PCa-associated genes by using disease-targeted sequencing, and found that KMT2D was highly mutated (63.04%, 29/46) in Chinese patients. Moreover, high KMT2D transcription was also associated with poor prognosis in an independent cohort ($n = 51$). In KMT2D-knockdown PC-3 and DU145 cells, cell proliferation ($P < 0.01$), invasion ($P < 0.001$), and migration ($P < 0.01$) were consequently suppressed. KMT2D depletion effectively suppressed tumor growth by 92.21% *in vivo*. Notably, integrative analyses of RNAseq and ChIPseq characterized two crucial genes downregulated by KMT2D, leukemia inhibitory factor receptor (LIFR) and Kruppel-like factor-4 (KLF4), which are regulators in PI3K/Akt and EMT, respectively. Our present study revealed that KMT2D epigenetically activates PI3K/Akt pathway and EMT by targeting LIFR and KLF4 and thus serves as a putative epigenetic-based target for treating PCa.

Shidong Lv and Liyan Ji contributed equally to this work.

Electronic supplementary material The online version of this article (<https://doi.org/10.1038/s41388-017-0026-x>) contains supplementary material, which is available to authorized users.

- ✉ Zhongqiu Liu
liuzq@gzucm.edu.cn
- ✉ Qiang Wei
qwei@smu.edu.cn
- ✉ Linlin Lu
lllu@gzucm.edu.cn

- ¹ Department of Urology, Nanfang Hospital, Southern Medical University, Guangzhou, Guangdong, China
- ² International Institute for Translational Chinese Medicine, Guangzhou University of Chinese Medicine, Guangzhou, Guangdong, China
- ³ Department of Science and Training, General Hospital of Guangzhou Military Command of People's Liberation Army, Guangzhou, Guangdong, China
- ⁴ State Key Laboratory of Quality Research in Chinese Medicine/Macau Institute For Applied Research in Medicine and Health, Macau University of Science and Technology, Macau (SAR), China

Introduction

Prostate cancer (PCa) is a commonly diagnosed malignancy and is the predominant cause of cancer-related mortality in men worldwide [1]. Epidemiological evidence has revealed that ethnicity-related heterogeneity at the genetic, epigenetic, and phenotypic level is directly related to incidence, prognosis, and the chemotherapy response of PCa [2–5]. Recent sequencing studies have revealed that the somatically acquired genetic, epigenetic, and transcriptional alterations in PCa, including oncogenic gene PIK3CA, ETS fusion gene, and tumor suppressor gene (TSG) PTEN, SPOP, CHD1 [6–9]. A further genotyping study has revealed that the frequency of PTEN loss is 29.7 vs. 5.4% in Caucasian and Chinese patients, respectively [3], thus suggesting that Caucasian patients with PCa have a poor prognosis and less sensitivity to abiraterone [4, 10]. Therefore, discovering and validating PCa-associated genomic aberrations in Chinese patients, especially prevalent significant genes, would provide potential therapeutic targets for PCa.

Histone lysine methyltransferase 2D (KMT2D/MLL2), an important epigenetic modifier that catalyzes the

monomethylation of H3K4 (H3K4me1) [11], has been found to be mutated from 0–12% in prostate cancers [12]. KMT2D is enriched at enhancers of multiple genes and epigenetically stimulates gene expression in various pathways, including the p53 pathway, cholestasis signaling, and cAMP-mediated signaling [13]. In lymphoma, KMT2D, a strong TSG, effectively inhibits tumorigenesis and metastasis [14]. In solid tumors such as pancreatic adenocarcinoma, breast, and colorectal cancer, KMT2D is essential for tumor cell proliferation, and its depletion increases chemotherapeutic sensitivity [15, 16]. A recent study has revealed that the mutation of KMT2D is PCa-specific and might be involved in the malignant progression from high grade prostatic intraepithelial neoplasia to PCa [17]. However, whether KMT2D is an oncogene or TSG in PCa remains unclear, and its downstream pathways still must be comprehensively determined.

Herein, we present our discovery of a highly mutated and overexpressed epigenetic gene, KMT2D, where its expression was positively correlated with metastasis and predicts prognosis. We also identified putative KMT2D/PI3K and KMT2D/EMT oncogenic pathways that promote prostate cancer outgrowth and metastasis by epigenetically transcriptional activation of novel targets, namely LIFR and KLF4. These findings suggested epigenetic regulator KMT2D might have therapeutic efficacy for treating PCa.

Results

Identification of recurrent KMT2D alterations in PCa through disease-targeted sequencing

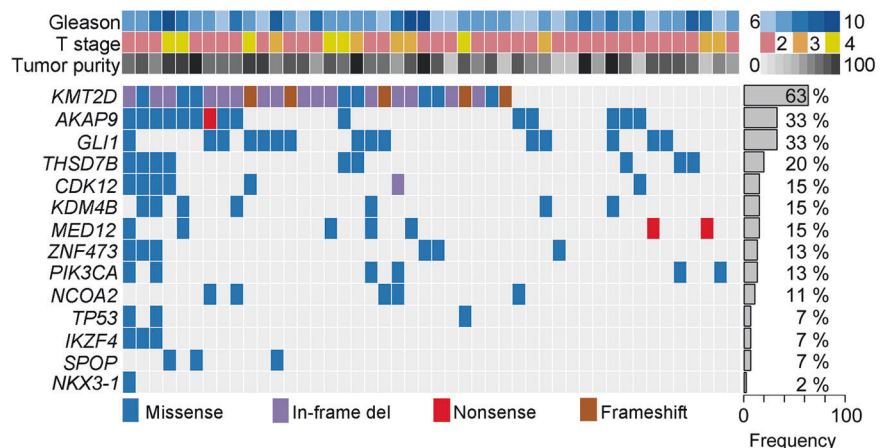
To identify the mutational spectrum of PCa pathogenesis-associated genes in Chinese patients, 32 significantly mutated genes were analyzed by using disease-targeted sequencing in 46 FFPE samples (patient information see

Supplementary Table 1). Sequencing depths were $3527 \times$ on average (range, $1156 \times -5419 \times$). Each specimen harbored an average of five mutations with an allele frequency $\geq 3\%$ (range, 1–34). We identified a total of 245 somatic non-synonymous mutations, including 199 missense, 3 nonsense, 24 in-frame dels, 1 nonstop, and 18 frameshift in 27 genes (Supplementary Table 2). Among them, 99 mutations (40.57%) were expected to be harmful for the structure or function of the encoded proteins, according to a SIFT/PolyPhen analysis (Ion Reporter). In an independent platform (Ion Torrent PGM), 91.5% of these high-confidence mutations were validated (Supplementary Table 2).

The prevalence of mutations in high-ranking genes was different among the mutational spectrum in the Caucasian population and our cohort of Chinese patients ($n=46$). Compared with the Caucasian population, the significantly mutated genes in our cohort were KMT2D (29/46, 63.04%), AKAP9 (15/46, 32.61%), GLI1 (15/46, 32.61%), THSD7B (9/46, 19.57%), CDK12 (7/46, 15.22%), KDM4B (7/46, 15.22%), MED12 (7/46, 15.22%), ZNF473 (6/46, 13.04%), PIK3CA (6/46, 13.04%), and NCOA2 (5/46, 10.87%) (Fisher's test, $P < 0.05$; Fig. 1). In contrast, frequently mutated genes in the Caucasian population, such as PTEN, CDKN1B, CDKN2A, KLF6, MYC, and SYNE3, were rarely identified in the 46 Chinese patients ($P < 0.05$, Fisher's test; Supplementary Fig. 1a). Furthermore, only two significantly mutated genes (TP53 and SPOP) overlapping between the Chinese cohort and three independent Caucasian populations, whereas 10 highly mutated genes were uniquely identified in Chinese patients (Supplementary Fig. 1b).

Given the high mutation frequency of KMT2D, a correlation analysis was performed to determine whether functional interactions existed on the basis of the KMT2D variant frequency. We found that the allele frequency of KMT2D showed a positive trend toward gene abnormalities

Fig. 1 Integrative molecular profiling and major clinical characterization in prostate cancer (PCa) samples. A heatmap showing genes (rows) with somatic mutations in 46 PCa patients (columns). Clinicopathological variables are shown in the header, including tumor stage, Gleason score, and tumor purity estimated by HE staining. Variant classification is indicated by color coding. Right, mutation percentage in 46 PCa specimens



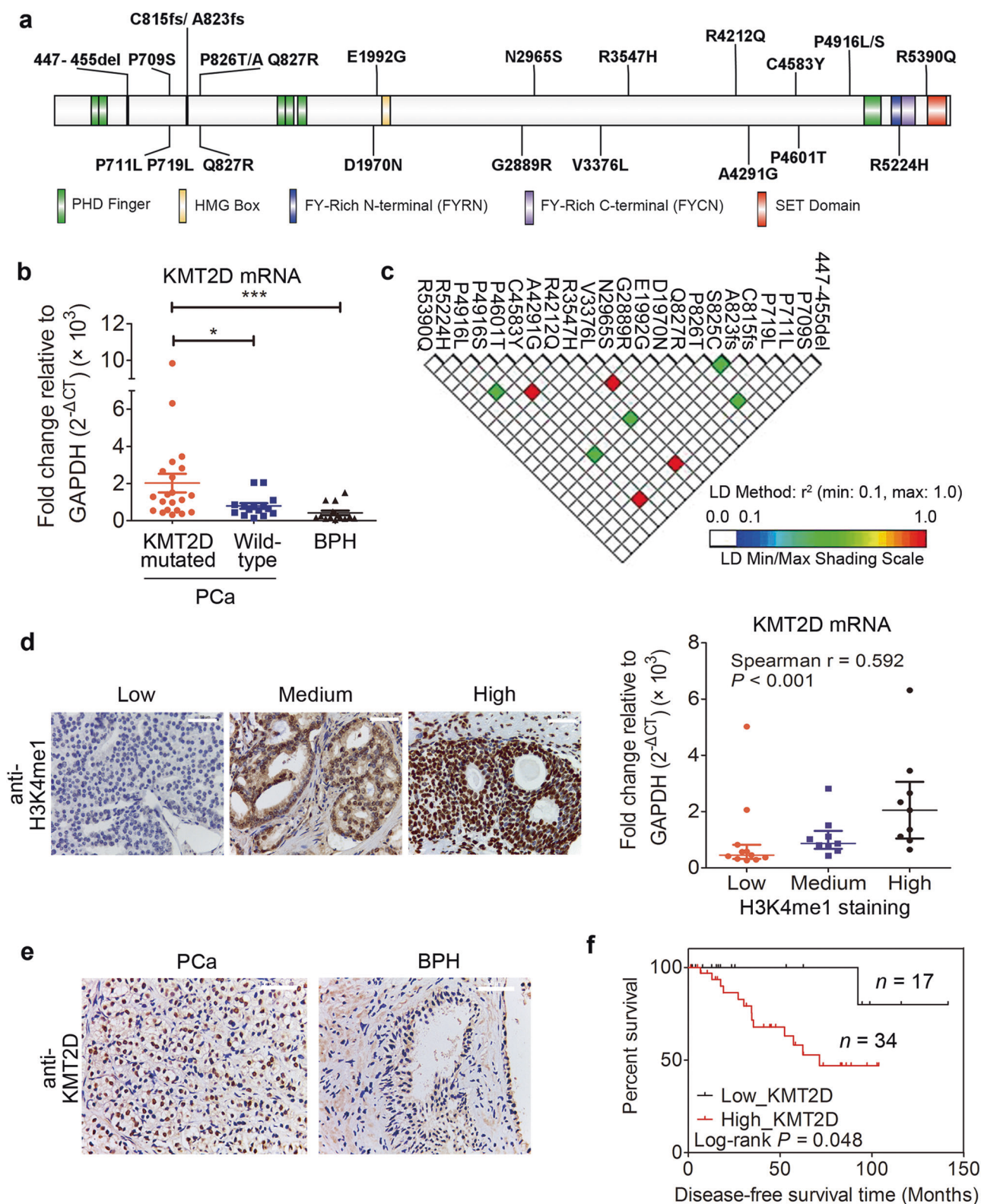


Fig. 2 High expression of KMT2D in PCa is associated with recurrent mutations and leads to poor prognosis. **a** Schematic representation of KMT2D mutations detected in 63.04% (29/46) of prostate tumors. **b** mRNA expression of KMT2D in KMT2D mutated, wild-type cases, and benign prostate hyperplasia. **c** Linkage disequilibrium plot of KMT2D mutations in the protein-coding region. The LD plot is based on pairwise r^2 value. **d** KMT2D gene expression classified by

H3K4me1 protein levels in PCa patients. Scale bars: 50 μ m. **e** Representative images of immunohistochemistry staining against KMT2D in benign prostate hyperplasia ($n = 3$) and PCa samples ($n = 51$). Scale bars: 50 μ m. **f** KM survival analysis of PCa patients based on protein expression of KMT2D ($n = 51$). Values are shown as median \pm quartile

Table 1 H3K4me1 expression in different KMT2D mutated tissues ($n = 35$)

	Strong expression	Medium expression	Low expression	Fisher's test P value	
Wild-type ($n = 13$)	3	6	4	–	
Missense ($n = 8$)	3	3	2	0.861 ^a	
In-frame del ($n = 10$)	3	5	2	1 ^a	
Frameshift ($n = 4$)	0	0	4	0.085 ^a	0.0212 ^b

^aFisher's test between wild-type and KMT2D-mutated (missense, in-frame del and frameshift) cases

^bFisher's test between Frameshift and other KMT2D mutated (missense and in-frame del) cases

in PI3K pathway ($P = 0.064$; Supplementary Fig. 1c), thus suggesting that functional interplay is involved.

High expression of KMT2D in PCa is associated with recurrent mutations and leads to poor prognosis

In 29 out of 46 Chinese patients, 46 KMT2D mutations, including 23 missense, 18 in-frame del, and 5 frameshift, were found to be distributed over 23 loci (Fig. 2a). Owing to multiple subclones in the sequencing specimens and the long length of the KMT2D gene (~19 kb), more than one KMT2D mutation (maximum 5) was identified in 28.26% of our PCa samples (Supplementary Table 2).

Compared to BPH, KMT2D was significantly over-expressed in PCa tissues in our cohort 1 and TCGA data (Supplementary Fig. 2a and 2b). Meanwhile, KMT2D transcription in mutation cases were 2.73-fold and 5.15-fold higher, respectively than that in wild-type and BPH (Fig. 2b). Linkage disequilibrium among variations is predicted to affect gene expression [18]. To determine the potential link between KMT2D mutations and its transcription, an allelic imbalance analysis was performed. Four strong linkage disequilibrium events among the KMT2D mutations were observed, including NM_003482:c.C13801A (p.P4601T) and NM_003482:c.G10640A (p.R3547H), NM_003482:c.C13801A (p.P4601T) and NM_003482:c.C2132T (p.P711L), NM_003482:c.G10640A (p.R3547H) and NM_003482:c.C2132T (p.P711L), as well as NM_003482:c.G10126C (p.V3376L) and NM_003482:c.A5975G (p.E1992G) ($r^2 = 1$; Fig. 2c). Whereas no significant correlation between KMT2D gene expression and copy number variation (CNV) was observed (Supplementary Fig. 2c). Besides, a weak negative correlation between DNA methylation and KMT2D transcription was found using TCGA data ($r = -0.27$, $P < 0.001$; Supplementary Fig. 2d), implying that rather than CNV, KMT2D variants, and epigenetic methylation may associate with its transcriptional regulation.

Furthermore, H3K4me1 abundance, marker for mirroring KMT2D activity [14, 19], was examined in PCa tissues with

KMT2D mutations and wild-type. As shown in Fig. 2d, KMT2D transcription was positively correlated with H3K4me1 levels ($r = 0.592$, $P < 0.001$), suggesting KMT2D gene expression was consistent with its catalytic activity. Moreover, in KMT2D frameshift cases, the percentage of H3K4me1-positive cells was consistently decreased ($n = 4$, $P = 0.085$; Table 1), and compared to other KMT2D mutations, KMT2D catalytic activity in KMT2D frameshift cases were significantly reduced ($P = 0.021$; Table 1). These suggested that in our cohort 1, frameshift might probably be associated with the reduced KMT2D catalytic activity.

In cohort 2, KMT2D was abundantly expressed in 66.67% (34/51) of PCa specimens but was not detected in any (0/3) of BPH tissues ($P = 0.037$; Fig. 2e and Supplementary Table 1). KMT2D was predominantly expressed in nuclei and exhibited a heterogeneous pattern within the same PCa specimens (Fig. 2e). Notably, higher KMT2D protein levels were correlated with poorer disease-free survival (DFS) time in 51 Chinese patients (log-rank $P = 0.048$; Fig. 2f). Similarly, a negative correlation was also observed between KMT2D transcription and prognosis in TCGA dataset ($P = 0.031$; Supplementary Fig. 2e). These findings indicated that high KMT2D expression may serve as a putative prognostic biomarker for Chinese PCa patients.

KMT2D depletion inhibits cell proliferation and induces cell apoptosis in PCa

To investigate KMT2D function in PCa tumorigenesis, cell proliferation and apoptosis were evaluated in KMT2D-depleted PCa cells. Since KMT2D abundance in LNCaP cells was relatively low, PC-3 and DU145 cells were selected for further knockdown ($P < 0.001$; Supplementary Fig. 3a). Both the transcriptional and translational levels of KMT2D were effectively silenced by two KMT2D-specific siRNAs (#1 and #2; Supplementary Fig. 3b and 3c).

As expected, siKMT2D transfections significantly suppressed cell viabilities (all $P < 0.001$; Fig. 3a). In KMT2D-depleted PC-3 cells, the percentage of EdU-positive cells

was decreased by 59.67%, significant reduction was also observed in KMT2D-silenced DU145 cells ($P = 0.005$; Fig. 3b). Moreover, with KMT2D deficiency, the early and total apoptotic ratios were significantly increased ($P < 0.05$; Fig. 3c), and subsequently, the expression of anti-apoptotic proteins, Bcl-2 and Bcl-xl, were dramatically decreased in KMT2D-silenced PC-3 and DU145 cells ($P < 0.001$; Fig. 3d).

To further verify the role of KMT2D in PCa development in vivo, KMT2D-depleted PC-3 xenograft model was developed (Supplementary Fig. 3d). Tumor growth was effectively inhibited, and tumor volume was remarkably suppressed by 92.21% ($P < 0.001$, $n = 5$; Figs. 3e, f) with KMT2D knockdown. Hence, KMT2D depletion restricted PCa progression, partially by inhibiting tumor cell proliferation and inducing cell apoptosis.

KMT2D knockdown attenuates tumor metastatic capacity in vitro

Considering approximately 90% of PCa patients develop metastasis [20], to further determine the role of KMT2D in tumor metastasis process, wound healing and transwell assays were performed with KMT2D-depleted cells. Our results showed that the migratory ratio was significantly decreased after KMT2D silencing (range, 23.07–54.21%, $P < 0.01$; Fig. 4a). In addition, KMT2D knockdown markedly inhibited the invasiveness of PC-3 and DU145 cells (range, 65.61–87.40%, $P < 0.001$; Fig. 4b). KMT2D depletion remarkably increased the expression of E-cadherin (epithelial marker [21]; Fig. 4c), while decreased that of N-cadherin (mesenchymal marker [21]; Fig. 4d) in PC-3 and DU145 cells.

By analyzing transcriptional data in the MSKCC cohort [22], we found that the KMT2D mRNA level in metastatic tissues was significantly higher than that in primary tissues ($P = 0.029$; Supplementary Fig. 4a). Similarly, in our cohort 1, KMT2D gene expression also exhibited a positive trend relating to high aggression (Supplementary Fig. 4b and 4c).

KMT2D sustains the expression of a set of genes in PCa cells

To further determine the mechanism underlying the biological consequences of KMT2D knockdown, RNAseq were performed. After silencing KMT2D with siRNA, 1442 and 1310 differentially expressed genes (DEGs) were detected, and 928 overlapping DEGs were also found (Fig. 5a and Supplementary Table 3). Among the top 10% of the overlapping DEGs, 90.2% of the genes were

decreased after KMT2D depletion. Through KEGG and GSEA analysis, the 928 overlapping DEGs were clustered into the PI3K/Akt pathway (Fisher's test, $P = 4.56E-06$; Figs. 5b, c). In validating whether the PI3K/Akt pathway is involved in KMT2D function, we found a significant decrease in the phosphorylation of Akt at S473 and T308 in KMT2D-depleted DU145 and PC-3 cells (Fig. 5d). Moreover, two downstream targets of p-Akt, p-BRCA1 and p-CREB, were also markedly decreased, while Akt-inactivated p-GSK3 β was increased in KMT2D-depleted cells (Fig. 5d). GSEA enrichment revealed that disrupted genes by KMT2D-knockdown were negatively correlated with prostate carcinogenesis and Akt upregulated genes, while positively correlated with "breast cancer metastasis downregulated" (FDR < 0.05 ; Fig. 5e). Hence, the subsequently altered genes enriched in the PI3K/Akt and metastasis pathways may partially contribute to the oncogenic role of KMT2D in tumor proliferation and metastasis.

KMT2D modulates activation of PI3K/Akt and EMT-associated pathways by epigenetically stimulating LIFR and KLF4

As an effective histone methyltransferase catalyzing monomethylation of H3K4, KMT2D plays an important role in cancer initiation, progression, and metastasis [14, 23]. To confirm the catalytic function of KMT2D, the abundance of H3K4me1 was evaluated with immunofluorescence staining. As expected, in KMT2D-silenced PC-3 and DU145 cells, the protein expression of H3K4me1 was markedly suppressed in the nucleus (Fig. 6a).

To further investigate KMT2D-mediated downstream genes, ChIPseq was performed using the H3K4me1 antibody in PC-3 cells. We identified 33,527 and 28,466 H3K4me1-binding sites in siControl and siKMT2D transfections, respectively. In the analysis of 230 differential binding events, 45.7% were distributed in intergenic regions, whereas the others were located in introns (53.3%), exons (0.7%), or promoters (0.4%) of 115 protein-coding genes (Fig. 6b and Supplementary Table 4).

According to the clustered PI3K/Akt pathway and EMT in KMT2D-depleted PC-3 cells, LIFR and KLF4 were selected, owing to their significant discrepancy in H3K4me1 binding ability (Supplementary Fig. 5 and Supplementary Table 5). LIFR, an upstream gene in the PI3K/Akt pathway, is essential for tumor progression through regulating cell proliferation [24]. KLF4 is a crucial protein involved in PCa metastasis [25]. After KMT2D knockdown, we observed a substantial loss of H3K4me1-binding sites in the gene body of LIFR and downstream (~643 kb) of KLF4 (Fig. 6c). In addition, we

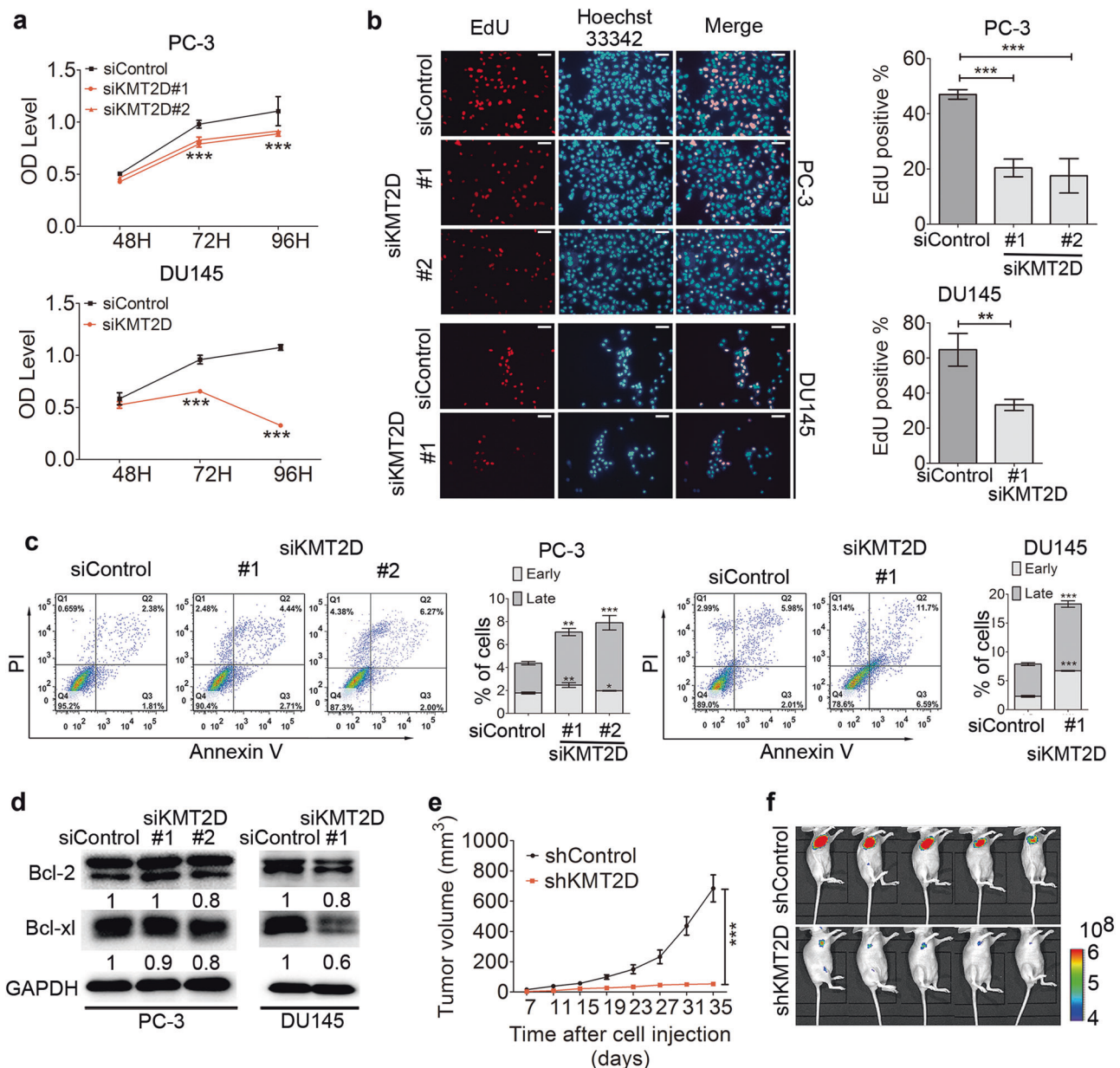


Fig. 3 KMT2D is a prerequisite for tumor proliferation and induces cell apoptosis in PCA. **a** Cell viability analysis of PC-3 and DU145 cells containing siControl or siKMT2D at 48, 72, and 96 h post-transfection. Values represent the mean OD of triplicate measurements. *** $P < 0.001$, compared with siControl cells. **b** Representative images of an EdU incorporation assay (20 \times objective) for PC-3 and DU145 cells with siControl or siKMT2D. The EdU positive percentage is shown as bar graphs. ** $P < 0.01$, *** $P < 0.001$, compared with siControl cells. **c** Cell apoptosis was evaluated with flow cytometry using Annexin V-FITC/PI staining in PC-3 and DU145 cells with siControl or siKMT2D. Quantification is shown as a bar graph in the

right panel (analyzed with FlowJo software). * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, compared with siControl cells. **d** Western blot analysis of the apoptosis-related proteins Bcl-2 and Bcl-xl after KMT2D knockdown in PC-3 and DU145 cells. **e** shControl and shKMT2D PC-3 cells were injected into the left flanks of mice ($n = 5$ each group), and tumor volumes were recorded and plotted for 35 days. **f** Images showing the size of the tumors that developed in mice of both groups; tumors were detected with IntegriSense 750 by IVIS at day 35. Color scale depicts fluorescence in units of radiance (photons/sec/cm²/sr/ μ m²). All values are mean \pm SD. Data are representative of at least three independent experiments. sr steradian radium

observed that both LIFR and KLF4 genes contained the consensus motif that was identified in the H3K4me1 binding peaks in the setting of KMT2D depletion (Fig. 6d).

To validate whether LIFR and KLF4 were downstream target genes for KMT2D modification, qRT-PCR was conducted with KMT2D-depleted cells. We observed that the gene expression of LIFR and KLF4 was consistently

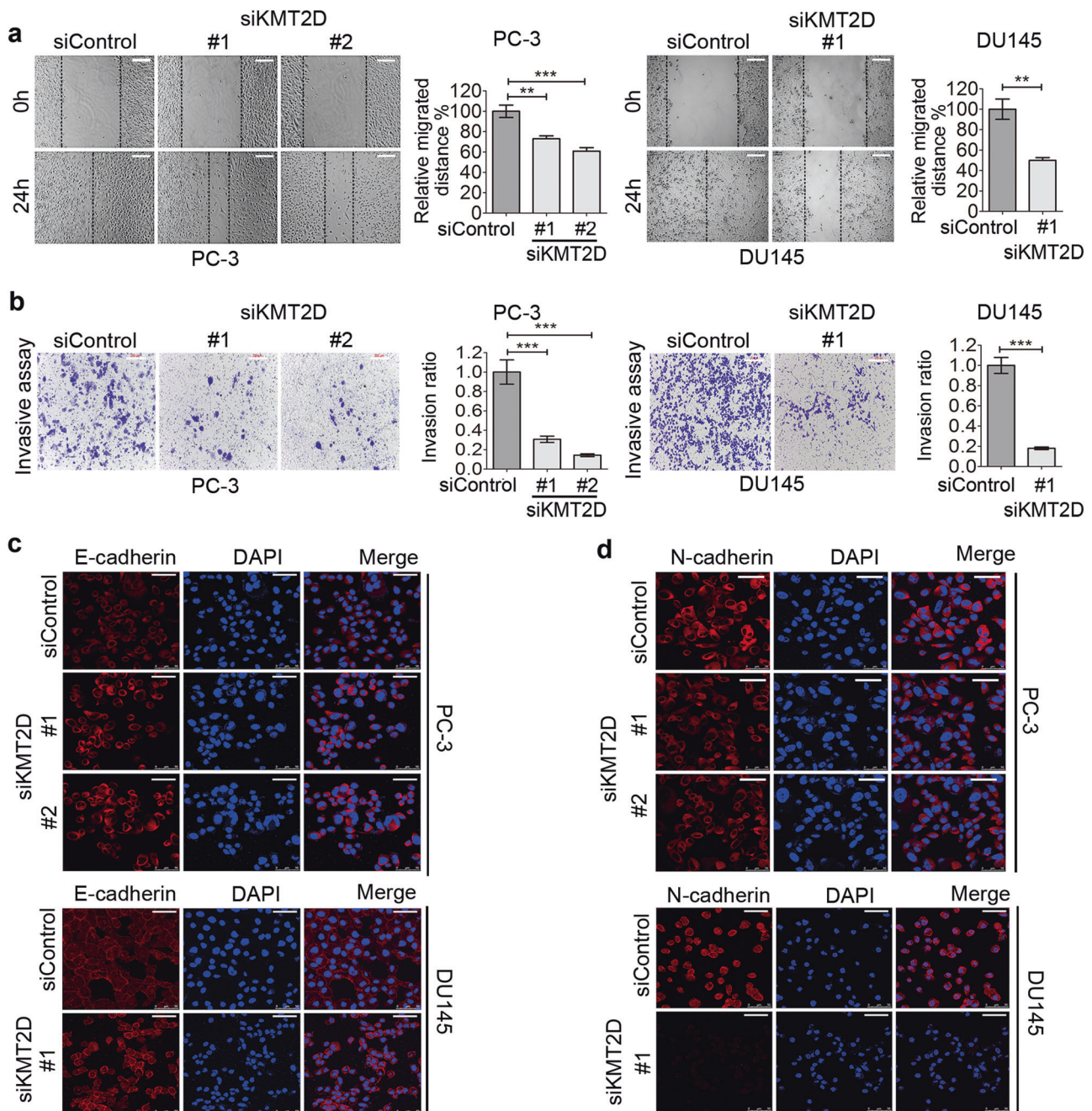


Fig. 4 Loss of KMT2D attenuates cell migration and invasion. **a** Cell migration capacity measured with a wound healing test in PC-3 and DU145 cells transfected with siControl or siKMT2D for 48 h ($10\times$ objective). The wound healing percentage is shown as a bar graph and was quantified with ImageJ software. $**P < 0.01$, $***P < 0.001$, compared with siControl cells. **b** The invasive ability of cells, as measured with a transwell matrigel assay ($10\times$ objective). Invading

cells were stained with violet and dissolved in DMSO. The OD levels were quantified with microplate reader and are shown as a bar graph. $***P < 0.001$, compared with siControl cells. Immunofluorescence staining analysis of E-cadherin and N-cadherin in PC-3 **c** and DU145 **d** cells treated with siControl or siKMT2D (scale bars: $50\mu\text{m}$). All values are mean \pm SD. Data are representative of three independent experiments

decreased (range, 0.31–0.70-fold) after KMT2D silencing (Fig. 6e). Furthermore, using cohort 1 and TCGA dataset, we also found that the KMT2D mRNA level was positively correlated with that of LIFR ($P < 0.001$; Fig. 6f) and KLF4

($P < 0.001$; Fig. 6g). These results suggested that the KMT2D-mediated alterations in the PI3K/Akt pathway and EMT may be ascribed to epigenetic activation of LIFR and KLF4.

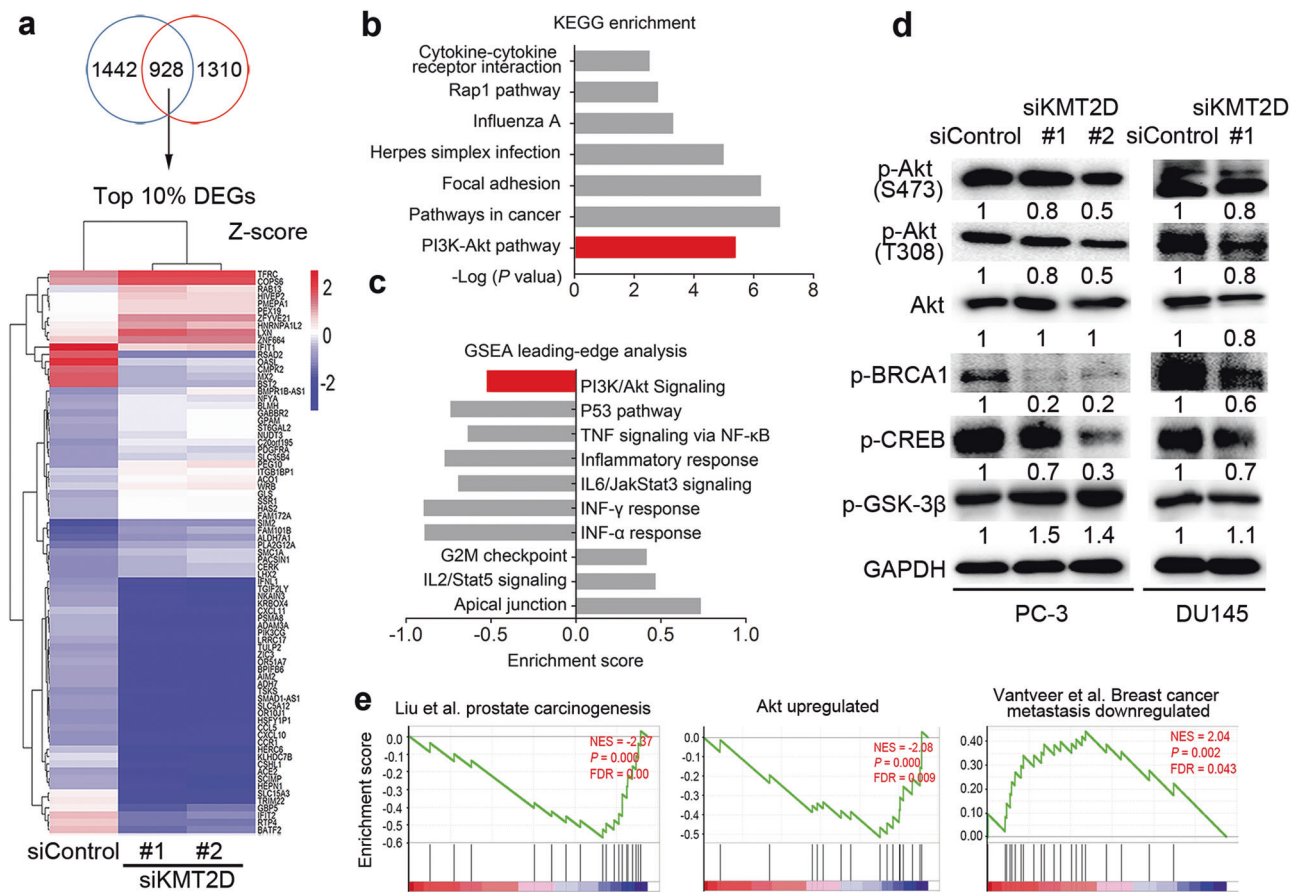


Fig. 5 RNAseq identifies transcriptional changes enriched in PI3K signaling following loss of KMT2D. **a** A Venn diagram depicting the overlapping DEGs in PC-3 cells transduced with two siRNAs against KMT2D vs. siControl-treated PC-3 cells (top) and heatmaps showing the top 10% of overlapping DEGs (bottom). **b, c** Pathway enrichment identified by KEGG **b** and GSEA **c** analysis of the DEGs between

KMT2D-depleted PC-3 and siControl cells. **d** Western blot analysis of Akt and phospho-Akt (S473 and T308), p-BRCA1, p-CREB, and p-GSK3 β after KMT2D knockdown in PC-3 and DU145 cells. **e** GSEA analysis of the DEGs enriched in prostate carcinogenesis, Akt upregulated, and breast cancer metastasis downregulated pathway

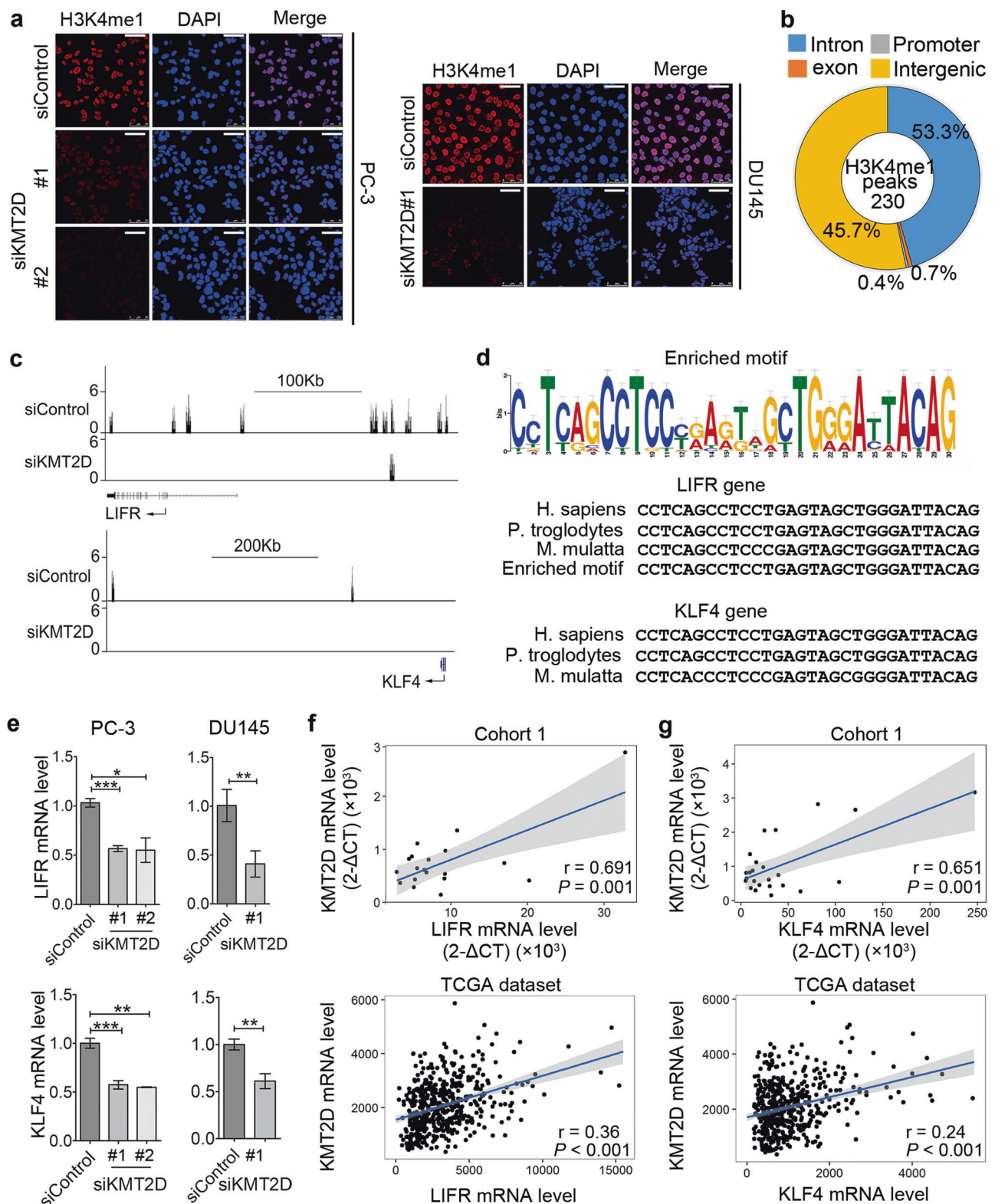
Discussion

In this study, we profiled the mutational spectrum of 32 PCa-associated significantly mutated genes in Chinese PCa patients (cohort 1, $n = 46$) and found KMT2D was the most recurrently mutated gene (variant frequency 63.04%) and high transcription. In an independent cohort ($n = 51$), KMT2D expression was negatively correlated with PCa prognosis. By generating KMT2D-depleted PCa cells, KMT2D-mediated biological consequences and corresponding signaling pathways were determined. After KMT2D knockdown, tumor cell proliferation, invasiveness, and migratory capacities were suppressed, whereas apoptosis was enhanced. Combinatorial analysis of RNAseq and ChIPseq revealed that KMT2D markedly activated the PI3K/Akt pathway and EMT by epigenetically stimulating LIFR and KLF4, respectively (Fig. 7).

Our study revealed the mutation profiling of KMT2D in Chinese PCa patients. The distinct mutational spectrum of

32 significant genes also laid a foundation for further and extensive discoveries regarding the ethnicity-specific mutational landscape. In contrast to frequently mutated genes in the Caucasian population, such as PTEN, CDKN1B, and CDKN2A [6], in Chinese patients, the high-ranking mutated genes (>20%) were KMT2D (63.04%), AKAP9 (32.61%), and GLLI (32.61%). The variant frequency of KMT2D, top mutated gene in our study, was much higher than the frequencies in Michigan (10.1%) [8], Robinson (12%) [26] and TCGA (6%) [12]. Moreover, KMT2D transcription was higher in PCa specimens than that in BPH ($P < 0.001$), and KMT2D high expression was correlated with poor prognosis ($P = 0.048$). These suggested KMT2D could serve as an alternative prognostic biomarker for PCa patients and further indicated Chinese-specific molecular subclassification and targeted therapy are urgently needed for Chinese PCa patients.

We provided a comprehensive portrait for KMT2D in genomic, transcriptional, translational, and especially



functional levels in PCa patients. Mutation within gene body (or exons) is thought to affect protein integrity and/or activity. While unexpectedly, in our study, KMT2D transcription was increased in PCa tissues harboring KMT2D mutations. Recent study revealed that for certain genes

(RYR2, NRIP1, and IRF8), their transcription could be triggered by mutations in coding exon regions [27]. Meanwhile, KMT2D-silencing was reported to cause genomic instability [28]. And this kind of genomic instability, induced by duplication of chromosome harboring the mutated sites,

◀ **Fig. 6** KMT2D up-regulates LIFR and KLF4 by methylating H3K4me1. **a** Immunofluorescence staining analysis using anti-H3K4me1 (red) in PC-3 and DU145 cells treated with siControl or siKMT2D (nuclear staining with DAPI; scale bars: 50 μ m). **b** Genomic distribution of different peaks localized in the functional region (exon, intron, promoter, intergenic region) of nearest genes by ChIPseq using anti-H3K4me1 in PC-3 cells transfected with siControl or siKMT2D. **c** Normalized read density tracks of H3K4me1 ChIPseq peaks by UCSC in PC-3 cells transfected with siControl or siKMT2D. **d** The consensus motif of the binding sites loss upon KMT2D silencing: H3K4me1 ($P = 1.56E-15$). The alignment of LIFR and KLF4 gene sequences from Human (*Homo sapiens*), Chimpanzee (*Pan troglodytes*), and Rhesus (*Macaca mulatta*). **e** qRT-PCR confirmed the decreased gene expression of LIFR and KLF4 after KMT2D knockdown in PC-3 and DU145 cells. $**P < 0.01$, compared with scramble control cells. **f, g** KMT2D mRNA levels positively correlates with LIFR **f** and KLF4 **g** gene expression in PCa tissues in our cohort 1 and TCGA dataset. All values are mean \pm SD. Data are representative of three independent experiments

might promote certain gene amplification through structural variation [29]. Our TCGA analysis showed that DNA methylation could also potentially contribute to KMT2D gene regulation ($r = -0.27$, $P < 0.001$), suggesting an alternative mechanism underlying KMT2D transcription might be involved. Frameshift and nonsense usually led to truncated protein and decreased KMT2D activity in small cell lung cancer [30]. In our study, KMT2D frameshift mutation may probably associate with the reduced KMT2D catalytic activity compared to other mutations such as missense and in-frame del ($P = 0.0212$, Table 1). But given the small size of frameshift cases ($n = 4$), unfortunately, no significance was observed compared that to wild-type cases. Additionally, positive correlation between KMT2D gene expression and H3K4me1 level was also observed in PCa tissues ($r = 0.749$, $P < 0.001$), suggesting the transcription of KMT2D was consistent with its catalytic activity.

In addition, our findings provide substantial evidence that KMT2D is a putative oncogene in PCa progression. In KMT2D-knockdown PCa cells, cell proliferation was reduced while mitochondria-dependent apoptosis was induced, and KMT2D deficiency dramatically restricted tumor growth in xenograft model. Furthermore, KMT2D silencing significantly decreased invasion and migration accompanied by the increase of E-cadherin and decrease of N-cadherin. In contrast to the role of KMT2D as a TSG in lymphoma [14], the potential oncogenic function of KMT2D in PCa illustrated that the biological consequences of KMT2D are dependent on cancer type. By addressing the distinct function of KMT2D in PCa progression and metastasis, our present study strengthens the current understanding of KMT2D-linked epigenetic alterations in tumorigenesis and further expands the list of promising chemotherapeutic targets [31].

The most interesting discovery in our study is the undescribed identification of certain signaling pathways linked to

KMT2D, such as PI3K/Akt pathway and EMT. Furthermore, integrative analysis of RNAseq and ChIPseq identified novel targets (LIFR and KLF4) in each pathway. Our current study provides clues that complement the signaling pathways linked to H3K4 methyltransferase and insights improving understanding of the mechanism underlying KMT2D-induced tumorigenesis in PCa: (i) Since approximately 25% of PCa tumors harbor PI3K abnormalities, and nearly 90% of PCa patients develop metastasis [20, 32], KMT2D-mediated activation of PI3K/Akt pathway strongly suggests that development of KMT2D inhibitors is critical either for PCa progression, or metastasis. (ii) Emerging evidence suggests that KMT2D is closely linked with nuclear receptors (ER α , RXR) [33, 34], transcription factors (Hox and Sox2) [34, 35] and inflammation pathways (CD40, JAK-STAT, and Toll-like receptor) in distinct tissues [14]. By identifying two novel KMT2D downstream targets in PCa, LIFR and KLF4, our study indicates that multiple cancer-specific targets are involved in KMT2D-mediated tumorigenesis. (iii) Previous studies have revealed that LIFR, through binding with IL-6, mediates several critical signaling pathways that regulate cell proliferation and survival, including JAK/STAT3, PI3K/Akt, and ERK1/2 [24]. KLF4 overexpression is reported to determine migration and proliferation of PCa cells, whereas its depletion attenuates tumor metastasis [25]. In line with these studies, the KMT2D-mediated epigenetic modification of LIFR and KLF4 observed in our study further supports these observations.

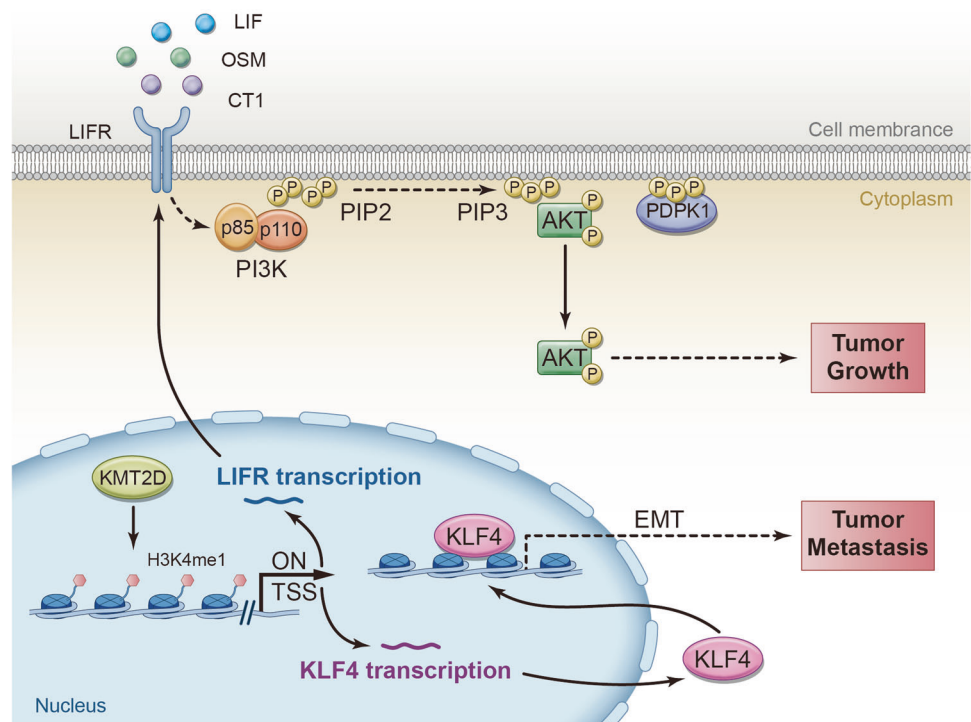
In summary, our current study deepens understanding of KMT2D in following aspects: (i) The distinct mutational spectrum of significant genes in Chinese PCa patients. (ii) By identifying numerous KMT2D-regulated pathways and novel downstream targets, combined with functional studies, we robustly defined the oncogenic role of KMT2D in PCa. (iii) The study revealed a direct correlation between PI3K/Akt signaling, the predominant aberrant pathway in PCa, and KMT2D, thus establishing an extensive basis and providing a rationale for studying the role of KMT2D in PCa tumorigenesis. (iv) LIFR and KLF4 identification complemented KMT2D-associated downstream genes in tumorigenesis. Our study lays a foundation for further understanding of the role of KMT2D in PCa and provides evidence for epigenetic therapy in treating PCa.

Materials and methods

Patients

This study was approved by Medical Ethics Committee of Nanfang hospital, and the species were used with the informed written consent in accordance with the

Fig. 7 The proposed mechanistic model of KMT2D in prostate cancer. The KMT2D sustains cell H3K4me1 level in prostate cancer and facilitates the expression of a set genes, including LIFR and KLF4. LIFR upon activation of PI3K pathway, and activated AKT phosphorylates downstream target such as CREB and BRCA1, promoting tumor growth. KLF4 facilitates transcription of EMT pathways, leading to tumor metastasis



Declaration of Helsinki. The study comprised a total of 17 benign prostate hyperplasia samples and 97 subjects with PCa including 46 in target next generation sequencing group (cohort 1; Supplementary Table 1) and 51 subjects in immunohistochemical staining of KMT2D (cohort 2; Supplementary Table 1). All samples have been histologically confirmed as PCa by two independent pathologists. Cases received pre-operative hormone therapy, chemotherapy, and radiotherapy were excluded.

DNA sample preparation and disease-targeted sequencing

The samples of disease-targeted sequencing cohort were dissected from 4 μ m formalin-fixed paraffin-embedded (FFPE) sections and genomic DNA was extracted by a Qiagen GeneRead DNA FFPE Tissue kit (Qiagen, Hilden, Germany) with Uracil-N-glycosylase according to the manufacturer's protocol. A prostate specific panel was brought by Qiagen containing 32 common frequently mutated genes, including AKAP9, APC, AR, CDK12, CDKN1B, CDKN2A, GLI1, IKZF4, KDM4B, KLF6, KMT2D, MED12, MYC, NCOA2, NIPA2, NKX3-1, NRCAM, OR5L1, PDZRN3, PIK3CA, PTEN, RB1, SCN11A, SPOP, SYNE3, TBX20, TFG, THSD7B, TP53, ZFH3, ZNF473, and ZNF595. The extracted DNA was quantified using a Qubit fluorimeter (Invitrogen, CA, USA)

and QIAseq DNA QuantiMIZE Assay Kit (Qiagen, Hilden, Germany). Genomic DNA with a yield >40 ng without low fragmentation and damage was used for library preparation. Amplicon enrichment was performed using the prostate cancer panel within the GeneRead DNaseq Targeted Panels V2 (Qiagen, Hilden, Germany).

Library preparation was prepared with the GeneRead DNA Library Core Kit for Ion Torrent instrument (Qiagen, Hilden, Germany) and barcode with the Ion Xpress™ Barcode Adapters (Thermo Fisher, MA, USA). Final bar-coded pools were sequenced using the Ion Torrent Proton. An average of 6,243,762 mapped reads were finally obtained for each sample. The sequencing reads were mapped to hg19 using TMAP algorithm. Single-nucleotide variants were detected using a combinational online software consisted of the Cloud-Based DNaseq Sequence Variant Analysis (Qiagen, Hilden, Germany), custom scripts of Qiagen, and Ion Reporter (Thermo Fisher, MA, USA). We validated the variants by sequencing genomic DNA in three PCa samples with sufficient amounts of DNA using Ion Torrent PGM platform. Significantly mutated genes were identified by MutSigCV software [36].

For CNVs prediction, peripheral blood mononuclear cells from seven independently healthy males were collected and sequenced for CNVs baseline detection. CNVs calling was performed by Cloud-Based DNaseq Sequence Variant Analysis (Qiagen, Hilden, Germany). For the

targeted genes, CNVs were defined as clustered-amplicons with significant change of normalized log₂-read-depth, as previously reported [37].

Immunohistochemical staining

Consecutive 4- μ m sections were obtained from FFPE blocks of the immunohistochemically staining group (86 subjects with PCa and three BPH). Sections were immunostained by anti-KMT2D antibody (Santa Cruz, CA, USA, 1:200) or H3K4me1 (Abcam, #ab8895, 1:1000) using the immunoperoxidase technique as previous study [38]. All sections stained with target protein were assessed by two independent pathologists.

RNA interference

Two specific siRNA targeting KMT2D and scrambled siRNA were bought from GE Dharmacon (# LQ-004828-00-0010, sequences see Supplementary Table 6). For siRNA transfection, a final concentration of 100 pmol siRNA for each with 1 μ l DharmaFECT 2 or DharmaFECT 1 Transfection Reagent (GE Dharmacon, Lafayette, USA) was used to transfect PC-3 and DU145 cells (1×10^5), respectively. Lentivirus expressing scrambled vector and KMT2D shRNA were packaged by Genepharma (Shanghai, China, sequences see Supplementary Table 6). PC-3 cells were infected with the lentivirus at 100 MOI by polybrene (Genepharma, Shanghai, China) at 5 μ g/ml. After 2–5 days incubation, transfected cells were used for further analysis.

MTT and EdU cell proliferation assay

Cells (1×10^4 /well) were seeded in 96-well plates. After transfection with siRNAs for 2–4 days, the alive cells were detected by MTT assay.

At 48 h post-transfection, cell proliferation was also assessed by EdU incorporation assay (Ribobio, Guangzhou, China).

Quantitative PCR

RNA from 46 PCa samples and 14 BPH tissues was isolated by a RNeasy FFPE Kit (Qiagen, Hilden, Germany) as the manufacturer's protocol. RNA from cells was isolated by Trizol Reagent (Invitrogen, CA, USA). The extracted RNA was quantified by Qubit fluorimeter (Thermo Fisher, MA, USA). Reverse transcription (RT) was performed with 1 μ g RNA using QuantiTect Reverse Transcription Kit (Qiagen, Hilden, Germany). The cDNA was amplified with gene-specific primers (Supplementary Table 6) and SYBR Premix Ex Taq II kit (TaKaRa, Shiga, Japan). Data were analyzed using a $2^{-\Delta\Delta C_t}$ method [39].

Cell culture

Human PCa cell lines PC-3, DU145 and LNCaP were purchased from cellcook biological technology Co., Ltd. (Guangzhou, China). Cells were cultured in RPMI1640 medium supplemented with 10% FBS (Thermo Fisher Scientific, Waltham, MA, USA). All cell lines were authenticated by STR profiling and tested negative for mycoplasma contamination.

Immunoblotting

Cells (3×10^5 /well) were plated in 6-well plates and transfected with siControl or siKMT2D for 72 h. Whole protein lysates were prepared by RIPA buffer containing 1% PMSF and phosphatase inhibitor cocktail (Roche). As describe previously [14], the blots were probed with primary antibodies to: KMT2D (Santa Cruz, # sc-68671, 1:200), Bcl2 (CST, #2872, 1:1000), Bcl-xL (CST, #2764, 1:1000), Akt (CST, #4691, 1:1000), p-Akt (Thr308, CST, #13038, 1:1000), p-Akt (Ser473, CST, #4060, 1:1000), p-GSK-3 β (CST, #5558, 1:1000), p-BRCA1 (CST, #9009, 1:1000), p-CREB1 (Santa Cruz, # sc-7978, 1:200), GAPDH (CST, #2118, 1:1000).

Flow cytometry analysis

Cells (3×10^5 /well) were plated in 6-well plates and transfected with siRNA vector or siKMT2D. Cell cycle distribution was analyzed with PI staining (BD Biosciences, Auckland, New Zealand); For cell apoptotic assay, cells were transfected for 48 h and assessed by Annexin V-Propidium Iodide kit (BD Biosciences, Auckland, New Zealand). The stained cells were acquired by flow cytometry (BD Biosciences, San Diego, CA, USA) and analyzed by FlowJo v7.6 software.

Xenograft tumor model

All animal studies were approved by Animal Care and Use Committee (IACUC) in Guangzhou University of Chinese Medicine. Female Balb/c-nude mice (4–6 weeks, 18–20 g) were purchased from Laboratory Animal Center of Sun Yat-Sen University (Guangzhou, China) and maintained in specific pathogen free environment in Guangzhou University of Chinese Medicine. Animals were fed with food and water freely and housed with a 12-dark:12-light cycle. Mice were randomized by weight and sample sizes were estimated according to "Resource Equation".

PC-3 cells transfected with shRNA expressing lentivirus ($n = 5$) or shKMT2D ($n = 5$) were injected subcutaneously (6×10^6 cells per mouse) into left flanks of mice. Tumors were measured with a caliper every 4 days and tumor

volumes were calculated as formula: $TV \text{ (mm}^3\text{)} = (L \times W^2)/2$. When the tumor volume reached 800 mm^3 , mice were intravenously administered with IntegriSense750 (PerkinElmer, MA, USA), which specially labeled the tumors. Then the tumor load was evaluated by fluorescence imaging using IVIS (PerkinElmer, MA, USA) after 24 h.

Wound healing assay

Cells (3×10^5 /well) were cultured in 6-well plate and transfected with siControl or siKMT2D. After growing up to 100% confluence, the cell monolayer was scratched with a 10 μ l pipette and incubated with 1% serum-containing medium for another 24 h. Photomicrographs were taken by contrast phase microscope (Leica, Jena, Germany) and migratory percentage was analyzed by ImageJ software.

Transwell assay

Cells (1×10^4 /well) transfected with siRNAs control or siKMT2D were seeded in upper chamber of transwell plate coated with Matrigel. The upper chamber contained 1% serum medium, while 10% serum medium was added to the low chamber. After 48 hrs of incubation, invasive cells were stained with 0.1% crystal violet, photographed and quantified by colorimetry.

Immunofluorescence staining

Cells (1×10^5 /well) were plated on sterile glass cover slips in 24-well plates. After transfected with siRNAs control or siKMT2D for 72 h, cells were fixed, permeabilized and blocked. Cells were stained overnight with primary antibodies to H3K4me1 (Abcam, #ab8895, 1:500), E-cadherin (Santa Cruz, # sc-7870, 1:50), N-cadherin (Santa Cruz, # sc-7939, 1:50). Subsequently, cells were incubated with secondary fluorescent antibody conjugated with PE (Santa Cruz, 1:200) and DAPI (5 μ g/ml).

RNA-seq and sequencing analysis

PC-3 cells (3×10^5 /well) were seeded in 6-well plates and transfected with siControl or siKMT2D#1 or siKMT2D#2. Cells were collected and total RNA for RNA sequencing (RNA-seq) was isolated using Trizol (Invitrogen, CA, USA). RNA quantification and integrity was measured by BioAnalyzer 2100 (Agilent, CA, USA). Oligo (dT) magnetic beads (Invitrogen, CA, USA) were used to enrich mRNA. Then mRNA was fragmented, reversed into cDNA and amplified by Phusion High Fidelity DNA polymerase as previously reported [40]. Each PCR products were single-

end (50 bps) sequenced on a BGISEQ-500 sequencer (BGI, Shenzhen, China).

Average 658 M raw reads were aligned to hg19 human genome by Bowtie2 [41] and HISAT [42] algorithm. Gene expression level was quantified by RSEM software [43].

We used false-discovery rate (FDR, Benjamini Hochberg method) ≤ 0.001 and cut-off of Log2 fold change ≥ 1 as the threshold to judge the significance of differential gene expression between siKMT2Ds vs. siControl. The DAVID online software were used for KEGG pathway enrichment of the selected DEGs [44]. Gene set enrichment analysis (GSEA) was also used to evaluate the related pathway affected by leading-edge genes [45].

ChIP-seq and sequencing analysis

H3K4me1 ChIP were carried out using a Pierce ChIP assay protocol (Thermo Fisher Scientific, Waltham, MA). Briefly, PC-3 cells (2×10^5 /well) were plated in 6-well plates and transfected with siRNA and siKMT2D. Cells were fixed with 1% formaldehyde, lysed and digested by micrococcal nuclease. Precleared DNA were immunoprecipitated with 4 μ l of H3K4me1 antibody (Abcam, Cambridge, UK) overnight and eluted. ChIP DNA was then prepared for sequencing library with fragment ranging 100–300 bps. Finally, prepared libraries were sequenced on BGISEQ-500 sequencer as 50 bps single-end read runs.

The ChIP-seq data was mapped to the hg19 human genome by SOAP2 [46]. Peaks calling was performed by SICER [47] and differential peaks were identified between siControl and siKMT2D samples by MAnorm algorithm [48]. The significant regions were picked up if $|MI| \geq 1$ and p -value $\leq 10^{-5}$. The peak-related genes were annotated by ANNOVAR software using UCSC known gene database [49]. Sequences centered on 500 bps of peak region were submitted to MEME online software to identify the consensus motif enriched in H3K4me1 [50].

Statistical analysis

Kaplan–Meier (KM) analysis was performed as previously report [51]. Disease-free survival (DFS) was defined as time from diagnosis to recurrence or to last follow-up data. Log-rank test was used to calculate the P values of KM.

Other data were statistically calculated using Student's *t*-test (two groups) or one-way ANOVA followed by the LSD post-hoc test (for more than two groups) or Mann–Whitney *U* test (nonparametric analysis). $*P < 0.05$, $**P < 0.01$, $***P < 0.001$ are determined as significance. All the experiments were performed at least in triplicates. Value presented as the means \pm standard deviation (SD) by GraphPad Prism software (GraphPad Software, CA, USA).

Data availability

RNA-seq (GSE94807) and ChIP-seq (GSE94817) are available in GEO dataset. Targeted sequencing data has been deposited into SRA dataset (SRA527454).

Funding This work was mainly supported by National Natural Science Foundation of China (81720108033), Natural Science Foundation of Guangdong Province (2015A030312012, 2016A050502052, and 2015B020233015), the Science and Technology Planning Project of Guangdong Province (2016A020215122), Bureau for Science and Information Technology of Guangzhou Municipality (201509010004).

Author contributions L.L., Q.W., Z.L., had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Study concept and design: S.L., L.J., L.L., Q.W., and Z.L. Acquisition of data: S.L., S.L., X.L., H.W., L.Z., and X.Y. Analysis and interpretation of data: S.L. and L.J. Drafting of the manuscript: S.L., L.J., and L.L. Critical revision of the manuscript for important intellectual content: L.L., Q.W., Z.L., G.D., and E.L.-H. L. Statistical analysis: S.L. and L.J. Obtaining funding: L.L., Q.W. and Z.L. Administrative, technical, or material support: S.L., B.C., C. L., X.Q., and Y.W. Supervision: L.L., Q.W., and Z.L.

Compliance with ethical standards

Conflict of interest The authors declare that they have no competing interests.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, and provide a link to the Creative Commons license. You do not have permission under this license to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *CA Cancer J Clin*. 2015;65:87–108.
2. Ito K. Prostate cancer in Asian men. *Nat Rev Urol*. 2014;11:197–212.
3. Mao X, Yu Y, Boyd LK, Ren G, Lin D, Chaplin T, et al. Distinct genomic alterations in prostate cancers in chinese and western populations suggest alternative pathways of prostate carcinogenesis. *Cancer Res*. 2010;70:5207–12.
4. Fukagai T, Namiki TS, Carlile RG, Yoshida H, Namiki M. Comparison of the clinical outcome after hormonal therapy for prostate cancer between japanese and caucasian men. *BJU Int*. 2006;97:1190–3.
5. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, et al. Cancer statistics in China, 2015. *CA Cancer J Clin*. 2016;66:115–32.
6. Cancer Genome Atlas Research N. The molecular taxonomy of primary prostate Cancer. *Cell*. 2015;163:1011–25.
7. Barbieri CE, Baca SC, Lawrence MS, Demichelis F, Blattner M, Theurillat JP, et al. Exome sequencing identifies recurrent SPOP, FOXA1 and MED12 mutations in prostate cancer. *Nat Genet*. 2012;44:685–9.
8. Grasso CS, Wu YM, Robinson DR, Cao X, Dhanasekaran SM, Khan AP, et al. The mutational landscape of lethal castration-resistant prostate cancer. *Nature*. 2012;487:239–43.
9. Fraser M, Sabelnykova VY, Yamaguchi TN, Heisler LE, Livingstone J, Huang V, et al. Genomic hallmarks of localized, non-indolent prostate cancer. *Nature*. 2017;541:359–64.
10. Bader DA, Stossi F, Li H, Samanta S, Cao X, Tsouko E, et al. PTEN protein loss and clinical outcome from castration-resistant prostate cancer treated with abiraterone acetate. *Nat Commun*. 2015;6:795–802.
11. Calo E, Wysocka J. Modification of enhancer chromatin: what, how, and why? *Mol Cell*. 2013;49:825–37.
12. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal*. 2013;6:11.
13. Guo C, Chang CC, Wortham M, Chen LH, Kernagis DN, Qin X, et al. Global identification of MLL2-targeted loci reveals MLL2's role in diverse signaling pathways. *Proc Natl Acad Sci USA*. 2012;109:17603–8.
14. Ortega-Molina A, Boss IW, Canela A, Pan H, Jiang Y, Zhao C, et al. The histone lysine methyltransferase KMT2D sustains a gene expression program that represses B cell lymphoma development. *Nat Med*. 2015;21:1199–208.
15. Dawkins JB, Wang J, Maniati E, Heward JA, Koniali L, Kocher HM, et al. Reduced expression of histone methyltransferases KMT2C and KMT2D correlates with improved outcome in pancreatic ductal adenocarcinoma. *Cancer Res*. 2016;76:4861–71.
16. Ford DJ, Dingwall AK. The cancer COMPASS: navigating the functions of MLL complexes in cancer. *Cancer Genet*. 2015;208:178–91.
17. Jung SH, Shin S, Kim MS, Baek IP, Lee JY, Lee SH, et al. Genetic progression of high grade prostatic intraepithelial neoplasia to prostate cancer. *Eur Urol*. 2016;69:823–30.
18. Sun C, Southard C, Witonsky DB, Olopade OI, Di Rienzo A. Allelic imbalance (AI) identifies novel tissue-specific cis-regulatory variation for human UGT2B15. *Hum Mutat*. 2010;31:99–107.
19. Guo C, Chen LH, Huang Y, Chang CC, Wang P, Pirozzi CJ, et al. KMT2D maintains neoplastic cell proliferation and global histone H3 lysine 4 monomethylation. *Oncotarget*. 2013;4:2144–53.
20. Carlin BI, Andriole GL. The natural history, skeletal complications, and management of bone metastases in patients with prostate carcinoma. *Cancer*. 2000;88:2989–94.
21. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011;144:646–74.
22. 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.
23. Zhang J, Dominguez-Sola D, Hussein S, Lee JE, Holmes AB, Bansal M, et al. Disruption of KMT2D perturbs germinal center B cell development and promotes lymphomagenesis. *Nat Med*. 2015;21:1190–8.
24. Liu SC, Tsang NM, Chiang WC, Chang KP, Hsueh C, Liang Y, et al. Leukemia inhibitory factor promotes nasopharyngeal carcinoma progression and radioresistance. *J Clin Invest*. 2013;123:5269–83.
25. Shin SH, Kwon YW, Heo SC, Jeong GO, Kim BR, Seo EJ, et al. Kruppel-like factor 4 mediates lysophosphatidic acid-stimulated migration and proliferation of PC3M prostate cancer cells. *Exp Mol Med*. 2014;46:e104.
26. 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.

27. Mathelier A, Lefebvre C, Zhang AW, Arenillas DJ, Ding J, Wasserman WW, et al. Cis-regulatory somatic mutations and gene-expression alteration in B-cell lymphomas. *Genome Biol.* 2015;16:84.
28. Kantidakis T, Saponaro M, Mitter R, Horswell S, Kranz A, Boeing S, et al. Mutation of cancer driver MLL2 results in transcription stress and genome instability. *Genes Dev.* 2016;30:408.
29. Preudhomme C, Warot-Loze D, Roumier C, Grardel-Duflos N, Garand R, Lai JL, et al. High incidence of biallelic point mutations in the Runt domain of the AML1/PEBP2 alpha B gene in Mo acute myeloid leukemia and in myeloid malignancies with acquired trisomy 21. *Blood.* 2000;96:2862–9.
30. Augert A, Zhang Q, Bates B, Cui M, Wang X, Wildey G, et al. Small cell lung cancer exhibits frequent inactivating mutations in the histone methyltransferase KMT2D/MLL2: CALGB 151111 (Alliance). *J Thorac Oncol.* 2017;12:704–13.
31. Roychowdhury S, Chinnaiyan AM. Advancing precision medicine for prostate cancer through genomics. *J Clin Oncol.* 2013;31:1866–73.
32. Cancer Genome Atlas Research N. The molecular taxonomy of primary prostate cancer. *Cell.* 2015;163:1011–25.
33. Mo R, Rao SM, Zhu YJ. Identification of the MLL2 complex as a coactivator for estrogen receptor alpha. *J Biol Chem.* 2006;281:15714–20.
34. Ansari KI, Hussain I, Shrestha B, Kasiri S, Mandal SS. HOXC6 is transcriptionally regulated via coordination of MLL histone methylase and estrogen receptor in an estrogen environment. *J Mol Biol.* 2011;411:334–49.
35. Zhao MH, Liang S, Kim NH, Cui XS. MLL2 is essential for porcine embryo development in vitro. *In Vitro Cell Dev Biol Anim.* 2016;52:699–704.
36. Lawrence MS, Stojanov P, Polak P, Kryukov GV, Cibulskis K, Sivachenko A, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature.* 2013;499:214–8.
37. Grasso C, Butler T, Rhodes K, Quist M, Neff TL, Moore S, et al. Assessing copy number alterations in targeted, amplicon-based next-generation sequencing data. *The Journal of molecular diagnostics: JMD* 2015;17:53–63.
38. Jiang Y, Zhang Q, Hu Y, Li T, Yu J, Zhao L et al. ImmunoScore signature: a prognostic and predictive tool in gastric cancer. *Ann Surg.* 2016. doi: 10.1097/SLA.0000000000002116.
39. Schmittgen TD, Livak KJ. Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc.* 2008;3:1101–8.
40. Zhang B, Zhang W, Nie RE, Li WZ, Segraves KA, Yang XK, et al. Comparative transcriptome analysis of chemosensory genes in two sister leaf beetles provides insights into chemosensory speciation. *Insect Biochem Mol Biol.* 2016;79:108–18.
41. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 2009;10:R25.
42. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods.* 2015;12:357–60.
43. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinform.* 2011;12:323.
44. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc.* 2009;4:44–57.
45. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA.* 2005;102:15545–50.
46. Li R, Yu C, Li Y, Lam TW, Yiu SM, Kristiansen K, et al. SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics.* 2009;25:1966–7.
47. Zang C, Schones DE, Zeng C, Cui K, Zhao K, Peng W. A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. *Bioinformatics.* 2009;25:1952–8.
48. Shao Z, Zhang Y, Yuan GC, Orkin SH, Waxman DJ. MAnorm: a robust model for quantitative comparison of ChIP-Seq data sets. *Genome Biol.* 2012;13:R16.
49. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 2010;38:e164.
50. Komai Y, Sakai Y, Gotohda N, Kobayashi T, Kawakami S, Saito N. A novel 3-dimensional image analysis system for case-specific kidney anatomy and surgical simulation to facilitate clampless partial nephrectomy. *Urology.* 2014;83:500–6.
51. Budczies J, Klauschen F, Sinn BV, Gyorffy B, Schmitt WD, Darb-Esfahani S, et al. Cutoff Finder: a comprehensive and straightforward Web application enabling rapid biomarker cutoff optimization. *PLoS ONE.* 2012;7:e51862.