


RESEARCH

Open Access

Subtype-associated epigenomic landscape and 3D genome structure in bladder cancer



Tejaswi Iyyanki^{1,2†}, Baozhen Zhang^{2,3†}, Qixuan Wang^{2†}, Ye Hou^{2†}, Qiushi Jin², Jie Xu^{1,2}, Hongbo Yang², Tingting Liu², Xiaotao Wang², Fan Song^{1,2}, Yu Luan², Hironobu Yamashita^{4,5}, Ruby Chien⁶, Huijue Lyu², Lijun Zhang⁷, Lu Wang², Joshua Warrick^{4,5}, Jay D. Raman⁵, Joshua J. Meeks⁸, David J. DeGraff^{4,5*} and Feng Yue^{2,9*} 

* Correspondence: ddegraff@pennstatehealth.psu.edu; yue@northwestern.edu

[†]Tejaswi Iyyanki, Baozhen Zhang, Qixuan Wang and Ye Hou contributed equally to this work.

⁴Department of Pathology and Laboratory Medicine, The Pennsylvania State University, College of Medicine, Hershey, PA, USA

²Department of Biochemistry and Molecular Genetics, Feinberg School of Medicine Northwestern University, Chicago, IL, USA
Full list of author information is available at the end of the article

Abstract: Muscle-invasive bladder cancers are characterized by their distinct expression of luminal and basal genes, which could be used to predict key clinical features such as disease progression and overall survival. Transcriptionally, FOXA1, GATA3, and PPARG are shown to be essential for luminal subtype-specific gene regulation and subtype switching, while TP63, STAT3, and TFAP2 family members are critical for regulation of basal subtype-specific genes. Despite these advances, the underlying epigenetic mechanisms and 3D chromatin architecture responsible for subtype-specific regulation in bladder cancer remain unknown.

Result: We determine the genome-wide transcriptome, enhancer landscape, and transcription factor binding profiles of FOXA1 and GATA3 in luminal and basal subtypes of bladder cancer. Furthermore, we report the first-ever mapping of genome-wide chromatin interactions by Hi-C in both bladder cancer cell lines and primary patient tumors. We show that subtype-specific transcription is accompanied by specific open chromatin and epigenomic marks, at least partially driven by distinct transcription factor binding at distal enhancers of luminal and basal bladder cancers. Finally, we identify a novel clinically relevant transcription factor, Neuronal PAS Domain Protein 2 (NPAS2), in luminal bladder cancers that regulates other subtype-specific genes and influences cancer cell proliferation and migration.

Conclusion: In summary, our work identifies unique epigenomic signatures and 3D genome structures in luminal and basal urinary bladder cancers and suggests a novel link between the circadian transcription factor NPAS2 and a clinical bladder cancer subtype.

Introduction

Urinary bladder cancers (BLCA) are the second most commonly diagnosed urologic malignancy in the USA, with over 81,400 total new cases diagnosed in 2019 [1, 2]. As BLCA is a morbid disease that is costly to treat, increased molecular understanding is required [3]. Expression of luminal (*FOXA1*, *GATA3*, *PPARG*, etc.) and basal (*KRT1*, *KRT5*, *KRT6A*, etc.) [4, 5] genes have been used to molecularly characterize muscle



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

invasive BLCA. In particular, the presence of basal BLCA, which is often enriched for squamous differentiation, is associated with significant morbidity, disease progression, and lower survival [6, 7].

Recent studies have identified both luminal (FOXA1, GATA3, and PPARG [8]) and basal (TP63 [9–12], STAT3 [4, 13–15], TFAP2A, and TFAP2C [16]) -specific transcription factors (TFs) with functional roles in BLCA. For instance, we previously reported that FOXA1 and GATA3 cooperate with PPARG activation to drive trans-differentiation of basal BLCA cells to luminal subtype [8]. This observation is in agreement with the role of these factors in maintaining urothelial cell differentiation [17, 18] and supports a role for inactivation of FOXA1, GATA3, and/or PPARG during BLCA progression to a basal subtype. Subtype-specific TFs also appear to drive programs in oncogenesis and tumor progression. For example, basal factors STAT3, TP63, and TFAP2A/C have been shown to promote tumor cell invasion and/or metastasis [12, 13, 16]. Similarly loss of FOXA1 [19] and GATA3 [20] have been shown to promote cell migration and invasion. Therefore, specific TFs play a key role in subtype specification.

In addition to directly regulating transcription, studies show that TFs regulate gene expression through epigenetic histone modifications and open chromatin accessibility in breast cancers [21–24]. However, the degree to which the specific repertoire of TFs, epigenetic open chromatin TF accessibility, histone modifications, and 3D genome architecture cooperate for subtype expression is unknown. Therefore, we performed the most comprehensive set of genome-wide experiments to systematically map the epigenome, transcriptome, TF binding, and 3D chromatin loops. To our knowledge, this is the first report identifying 3D genome architecture in bladder cancer. Our work highlights the relevance of epigenetic modifications, open chromatin accessibility, and TF repertoire and identifies a novel identified basic helix loop helix (bHLH) TF NPAS2, all of which cooperate in the coordination of subtype-specific gene expression in bladder cancer.

Results

Comprehensive epigenomic profiling in both BLCA lines and primary tumors

In this project, we performed RNA-Seq, ChIP-Seq for Histone 3 lysine 27 acetylation (H3K27ac), Assay for Transposase-Accessible Chromatin using sequencing (ATAC-Seq), and genome-wide chromatin conformation capture experiments (Hi-C) on 4 bladder cancer cell lines (Fig. 1a), two of which (RT4 and SW780) were previously annotated as luminal and the two others (SCABER and HT1376) that were characterized as basal [8, 25]. Based on the RNA-Seq data generated in this study, we used a previously reported molecular subtyping approach [26] to confirm assignment to luminal and basal states. Our results confirmed RT4 and SW780 as belonging to the Luminal-papillary subtype, while SCABER and HT1376 belong to the Basal/squamous subtype (Additional file 1: Table S1). Each experiment in bladder cancer cell lines has at least two biological replicates (Additional file 2: Table S2) and we observed a high correlation between the two replicates (Additional file 3: Table S3). More importantly, we performed the same set of experiments on four patient muscle-invasive bladder tumors as well. By using the same molecular subtyping method, we determined their subtypes as the following: T1 is Luminal-papillary, T3 is Stroma-rich, and T4 and T5 are basal/squamous.

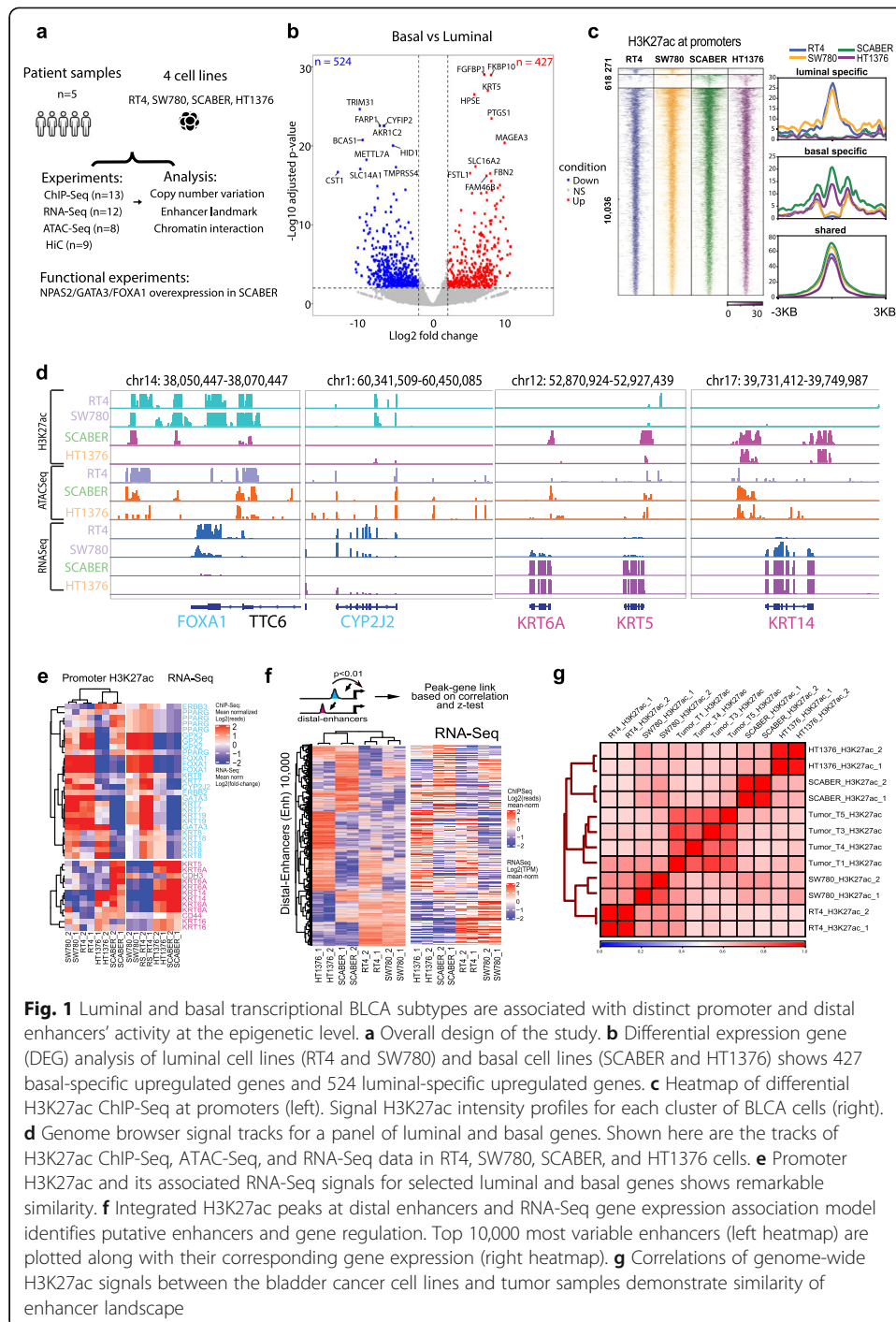


Fig. 1 Luminal and basal transcriptional BLCA subtypes are associated with distinct promoter and distal enhancers' activity at the epigenetic level. **a** Overall design of the study. **b** Differential expression gene (DEG) analysis of luminal cell lines (RT4 and SW780) and basal cell lines (SCABER and HT1376) shows 427 basal-specific upregulated genes and 524 luminal-specific upregulated genes. **c** Heatmap of differential H3K27ac ChIP-Seq at promoters (left). Signal H3K27ac intensity profiles for each cluster of BLCA cells (right). **d** Genome browser signal tracks for a panel of luminal and basal genes. Shown here are the tracks of H3K27ac ChIP-Seq, ATAC-Seq, and RNA-Seq data in RT4, SW780, SCABER, and HT1376 cells. **e** Promoter H3K27ac and its associated RNA-Seq signals for selected luminal and basal genes shows remarkable similarity. **f** Integrated H3K27ac peaks at distal enhancers and RNA-Seq gene expression association model identifies putative enhancers and gene regulation. Top 10,000 most variable enhancers (left heatmap) are plotted along with their corresponding gene expression (right heatmap). **g** Correlations of genome-wide H3K27ac signals between the bladder cancer cell lines and tumor samples demonstrate similarity of enhancer landscape

Luminal and basal transcriptional BLCA subtypes are associated with distinct promoter and distal enhancers activity at the epigenetic level

Enrichment of H3K27ac signals has been used to predict both active promoters and distal enhancers [27, 28]. Therefore, we first performed ChIP-Seq for H3K27ac in all four cell types and four patient samples. We observed that biologic replicates following H3K27ac ChIP-seq always clustered together, indicating our results are highly reproducible (Additional file 4: Figure S1A). Further, we found that two luminal subtypes (RT4

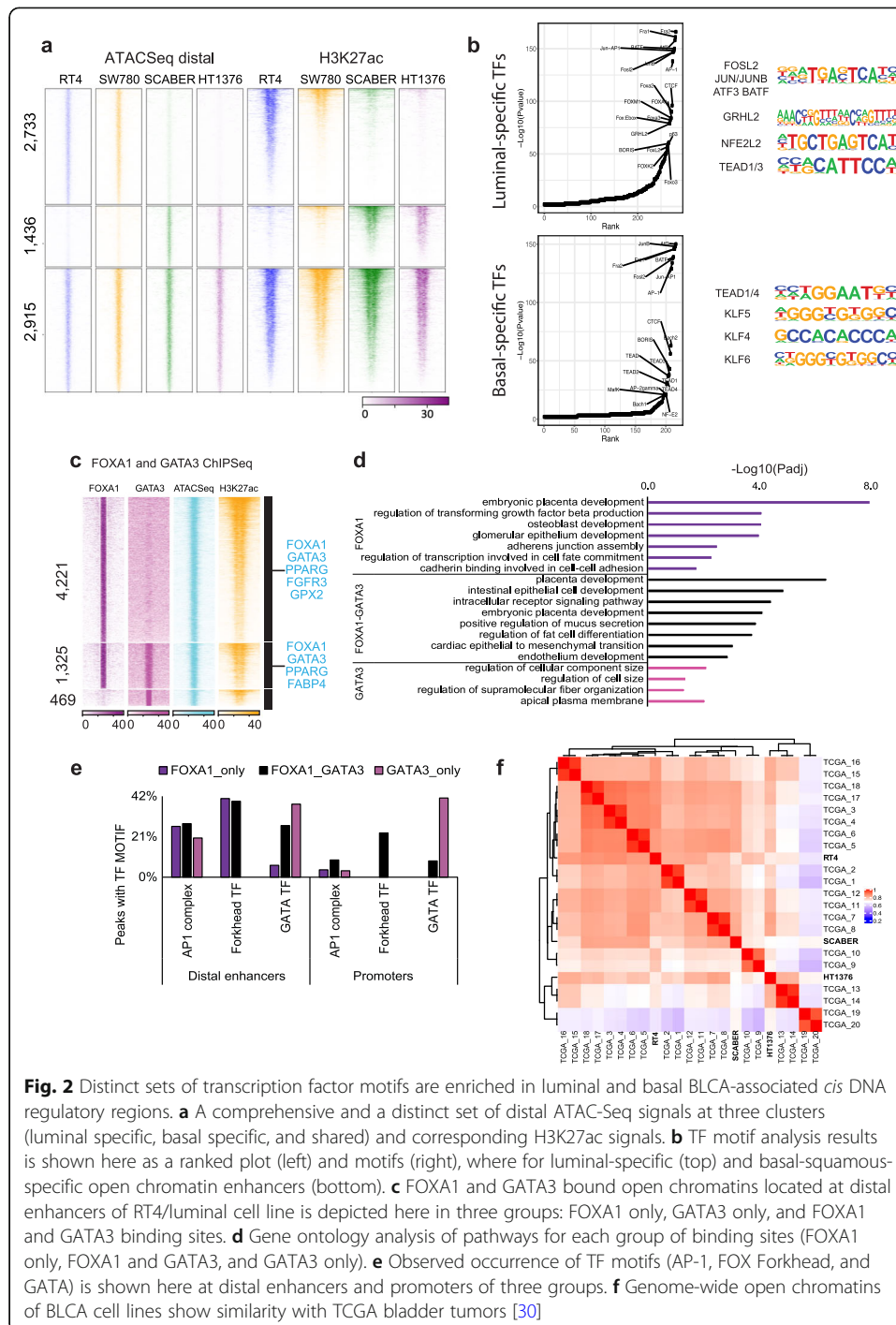
and SW780) clustered together, while two basal (SCABER and HT1376) cell lines are grouped together as well (Additional file 4: Figure S1A). These clustering results suggest global epigenomic profiles accurately reflect cell identity. The hierarchical clustering in the cell lines based on H3K27ac signals was also mirrored by global mRNA expression by RNA-Seq data (Additional file 4: Figure S1B). We performed differential gene expression analysis on the two groups of cell types (RT4 and SW780 vs. SCABER and HT1376) and identified 427 basal-specific (Additional file 5: Table S4) and 524 luminal-specific genes (Fig. 1b, Additional file 6: Table S5).

Next, we examined promoter usage based on H3K27ac signals at known genes. We confirmed that promoter H3K27ac intensities are remarkably similar to gene expression (Fig. 1c), and clustering analysis based on promoter H3K27ac intensity was able to distinguish luminal and basal models of BLCA (Additional file 4: Figure S1C). For example, we observed that two luminal subtype BLCA cell lines RT4 and SW780 have similar H3K27ac patterns at luminal genes *FOXA1*, *GATA3*, and *PPARG* (Fig. 1d, e), while the two basal cell lines share similar promoter marks at genes encoding the basal/squamous markers *KRT5/14*. Interestingly, although based on global gene expression, HT1376 is classified as a basal/squamous subtype, it shows a similar promoter H3K27ac pattern at luminal genes (*GATA3*, *KRT7/8/18*, Fig. 1e).

Distal H3K27ac peaks from gene promoter regions have been used as markers for active enhancers [27, 29]. We took the same approach here, and on average, we predicted 59,466 (40,731–78,506) enhancers in each cell line (Additional file 7: Table S6). To link the distal enhancers to their target genes, we performed a correlation-based distal-enhancer peak-gene association as described in [30] and identified the top 10,000 variable distal enhancers that show significant correlation to its linked gene (correlation ≥ 0.5 , $p < 0.01$; a total of 58,509 satisfied our criteria; Fig. 1f and Additional file 8: Table S7). We observed that the enhancers show clear clustering according to different cell types, and their target genes show similar cell-type-specific patterns (Fig. 1f and Additional file 4: Figure S1D). Moreover, to understand the clinical relevance of our findings, we performed H3K27ac ChIP-Seq in four muscle invasive bladder patient samples. Our results show a remarkable correlation of tumor cell lines (Fig. 1g). In summary, we show in these cell lines and in a limited tumor cohort that epigenetic regulation is correlated with molecular subtype assignment.

Distinct sets of transcription factor motifs are enriched in luminal and basal BLCA-associated *cis* DNA regulatory regions

We performed ATAC-Seq in RT4, SW780, SCABER, and HT1376 cell lines to evaluate their open chromatin status in the genome. On average, in each cell line, we identified 32,000 open chromatin regions (Fig. 2a and Additional file 9: Table S8). Among them, 40.8% of open chromatin regions were located at promoter regions, while 59.2% were located at distal regions. Overall, > 90% of the open chromatin promoter regions overlap with H3K27ac (Additional file 4: Figure S2A, S2C–D). The overlap of distal ATAC-Seq peaks and H3K27ac is lower (Additional file 4: Figure S2A and Additional file 10: Table S9), at least partially due to the different numbers of peaks in different datasets. Genome-wide correlation of ATAC-Seq showed that HT1376 and SCABER clustered together with 80% similarity (Additional file 4: Figure S2E) compared to luminal RT4



(~ 65%). We noted that this observation agrees with the RNA-Seq-based clustering and H3K27ac-based clustering (Additional file 4: Figure S1A and B).

Next, we performed motif analysis of these open chromatin regions (Additional file 11: Table S10). We observed that binding sites for CTCF and AP-1 complex are enriched in all cell lines (Fig. 2b and Additional file 4: Figure S2G). Further ranking of TF motifs by enrichment *p*-value revealed luminal open chromatin regions (shared between RT4 and SW780) were enriched with binding motifs for GRHL2, TP53, and

TP63 while basal open chromatin (shared between SCABER and HT1376) were enriched for TEAD1/4 and KLF factor (Fig. 2b) binding motifs. GRHL2 [31] was previously reported to be a luminal gene, thereby validating our findings. Interestingly, binding motifs for AP-1 complex proteins FOSL1/2, JUN/JUNB, ATF3, and BATF TFs [32] were the topmost enriched motifs for both luminal and basal-squamous open chromatin. We then comprehensively mapped all the enriched TF motifs in luminal, basal-squamous and shared open chromatin of distal enhancers to examine the relationship between TFs and BLCA subtypes (Additional file 11: Table S10). We discovered that at distal enhancers, the luminal BLCA subtypes are associated with previously reported steroid hormone receptor TFs. On the other hand, basal-squamous open chromatin areas at enhancers show enrichment of previously unreported factors MADS box TF MEF2C and the homeobox TF OTX2. Not surprisingly, luminal pioneering TFs such as forkhead transcription factors (FOXA1/2/3, FOXF1, FOXK1, FOXM1), and GATA TFs (GATA3/4/6) were enriched in luminal-associated enhancers with an open chromatin conformation. More surprisingly, forkhead and GATA motifs were also identified as being associated with open chromatin at enhancer elements across cell lines (Additional file 11: Table S10). While FOXA1 and GATA3 are known to have low expression in basal bladder cancer cell lines and tumors, the enrichment of forkhead and GATA motifs in open chromatin across BLCA cell lines suggests compensation by Forkhead and GATA factors other than FOXA1 and GATA3. In addition, Forkhead and GATA motif enrichment across cell lines in areas of open chromatin may indicate luminal-specific TFs are poised to bind to these areas of open chromatin. Furthermore, FOXA1 and GATA3 are known to play a role in the development of urothelium [31] suggesting that their binding sites may be primed early during development. We also discovered that the stem-cell-associated pioneering TFs such as KLF factors (KLF10/14), ATF factors (ATF1/2/4/7), and NANOG were enriched in basal-associated enhancers. This is interesting because there exists a progenitor cell population within basal urothelium that can contribute to urothelial development and differentiation [33, 34].

FOXA1 and GATA3 bind at luminal open chromatin at regulatory distal enhancers to drive expression of luminal-specific genes

We hypothesized that TFs such as FOXA1 and GATA3 bind at the open chromatin region to pioneer luminal enhancers and activate associated gene expression. To test this hypothesis, we performed GATA3 ChIP-Seq in the RT4 luminal BLCA cell line and obtained FOXA1 ChIP-Seq in RT4 cells from our previously published work (Additional file 12: Table S11) [8]. As predicted, luminal TFs FOXA1 and GATA3 showed enriched binding at the open chromatin loci of luminal-associated (*FOXA1*, *GATA3*, *PPARG*, *FGFR3*, and *FABP4*) distal enhancers (Fig. 2c). More specifically, we discovered 1325 distal enhancers that show co-binding of both FOXA1 and GATA3 in RT4 (Fig. 2c). Similarly, FOXA1 and GATA3 showed enriched binding at open chromatin loci of luminal marker genes (*FOXA1*, *ERBB3*, *KRT19*, *GPX2*, and *FABP4*) promoters (Additional file 4: Figure S2F).

GO term analysis of genes proximal to these distal enhancer sites showed regulation of TGF beta production, epithelium development, regulation of transcription involved in cell fate commitment, and cell-cell adhesion biological processes (cadherin binding

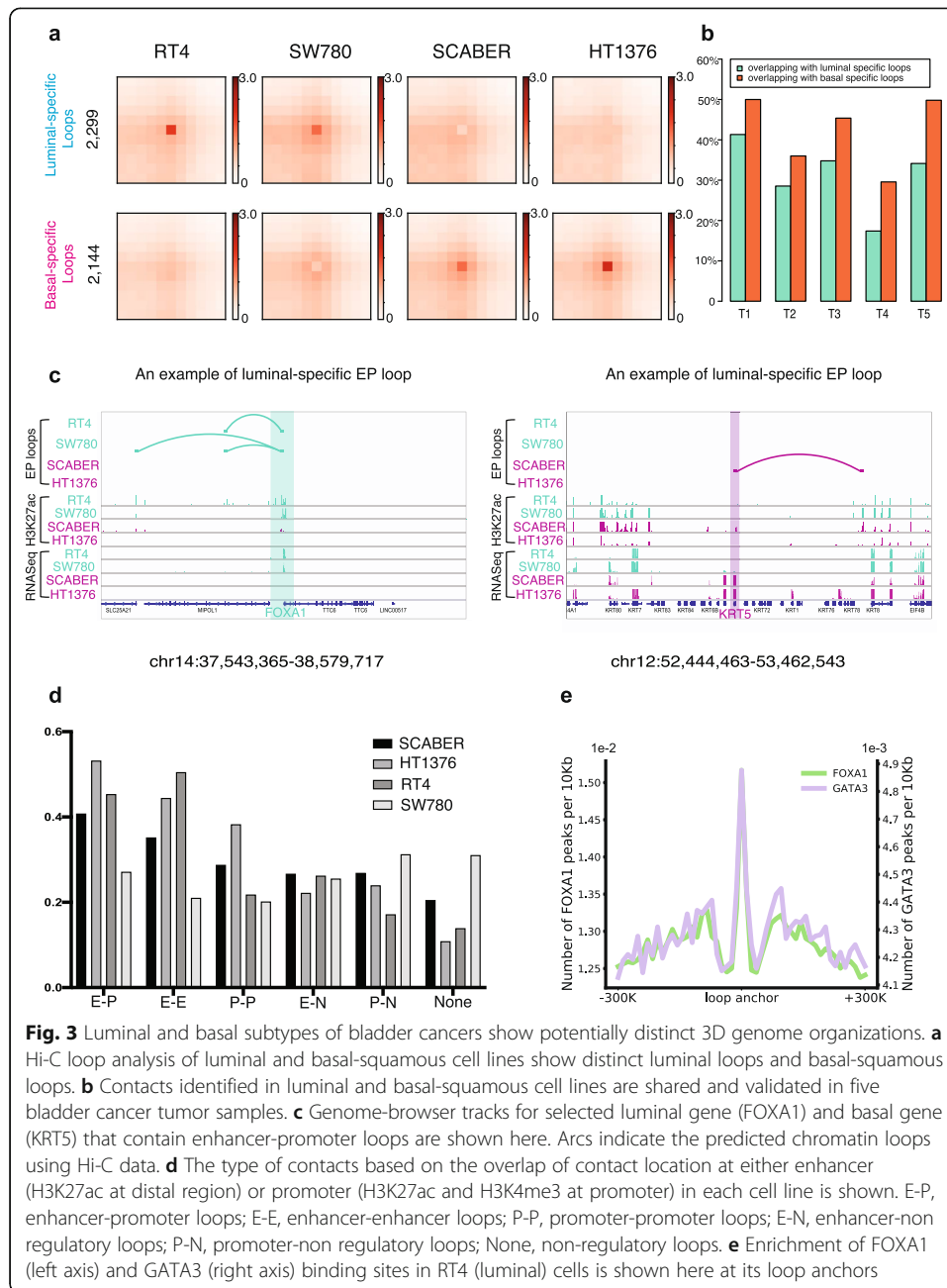
and adherens junction assembly) as terms associated with FOXA1. In addition, regulation of cellular component, cell size, and apical plasma membrane biological processes were terms identified with GATA3-bound genes proximal to these distal enhancers, suggesting a strong involvement of both TFs in commitment to cell fate and luminal differentiation (Fig. 2d). In regard to proximal genes associated with distal enhancers bound by both FOXA1 and GATA3, terms identified were associated with various developmental processes and the regulation of mucus secretion and fat cell differentiation, both important metabolic attributes of differentiated urothelium (Fig. 2d).

We then proceeded with the motif analysis of FOXA1 only, GATA3 only, and co-bound sites. Surprisingly, AP1-complexes were enriched specifically in all distal enhancers in addition to FOXA or GATA motifs (Fig. 2e). The order of binding of these three factors remains to be investigated. Finally, to understand the clinical relevance of our findings, we compared our four BLCA cell lines to the TCGA muscle-invasive bladder tumor ATAC-Seq data [30] and discovered that the genome-wide open chromatin profile in our cell lines is clustered with distinct clusters of tumors (Fig. 2f), suggesting that the open chromatin regions in these cell lines share similar patterns with patient tumors.

Luminal and basal subtypes of BLCA show potentially distinct 3D genome organizations

Previous studies have shown that 3D chromatin organization is associated with epigenetic activation or silencing of genes in cells [35]. For example, the majority of heterochromatin is known to be compressed in nuclei and located near the lamina-associated periphery of the nuclear envelope [35]. To obtain initial insights into the genome-wide 3D landscape of luminal and basal BLCA, we performed high-resolution Hi-C experiments on all four cell lines (at least 800 M reads, each) and five bladder tumor patients (> 800 M reads, each) (Additional file 4: Figure S3). We used our recently developed software, Peakachu [36], which is a machine learning-based chromatin loop detection approach, to predict loops at 10Kb bin resolution. First, we identified an average of 56,315 loops (range between 38,271 and 69,032) in the four cell lines (prob> 0.8; Additional file 13: Table S12). Then, by using the probability score output from Peakachu, we assigned subtype-specific chromatin loops as shown in the Aggregate Peak Analysis (APA, Fig. 3a and Additional file 14: Table S13) [37]. Based on our approach, we observed more potentially luminal-specific loops in RT4 and SW780 (2299) relative to the basal BLCA models SCABER and HT1376 (2144). We then compared each of these categories with loops detected in five patient samples (Fig. 3b): ~30–40% of luminal-assigned and basal-assigned 3D chromatin loops identified in the cell lines were observed in these five tumor samples.

Finally, we examined enhancer and promoter loops in each category for their association with subtype-specific gene expression. Examples are shown in Fig. 3c, in which we found that the luminal gene *FOXA1* and the basal gene *KRT5* showed increased number of enhancer-promoter loops in luminal and basal cell lines, respectively. Overall, we observed that ~40% of the chromatin loops exist between enhancers and promoters (Fig. 3d). Furthermore, we found a significant enrichment of FOXA1 and GATA3 binding sites at these loop anchors, indicating the involvement of these pioneer factors in the regulation of the 3D genome (Fig. 3e). This finding is in agreement with



previous studies reporting the enrichment of FOXA1 binding sites in enhancer-promoter loops [38].

Copy number variation (CNV) and chromatin loops in bladder cancer

A hallmark of cancer is large structural variations (SVs), which includes inversions, deletions, duplications, and translocations. Recently, it has been shown that alteration in CNVs and SVs can lead to the alterations in 3D genome structure, including the formation of new topologically associated domains (“neo-TADs”) [39] and resultant “enhancer hijacking [40].” Neo-TADs refer to scenarios where an SV event leads to the

formation of new chromatin domains, which can in turn affect the expression profiles of the genes located in those regions. In the “enhancer-hijacking” model, altered 3D genome organization results in abnormal enhancer interaction, with enhancers brought in close proximity to the wrong target gene (usually an oncogene) resulting in inappropriate target activation.

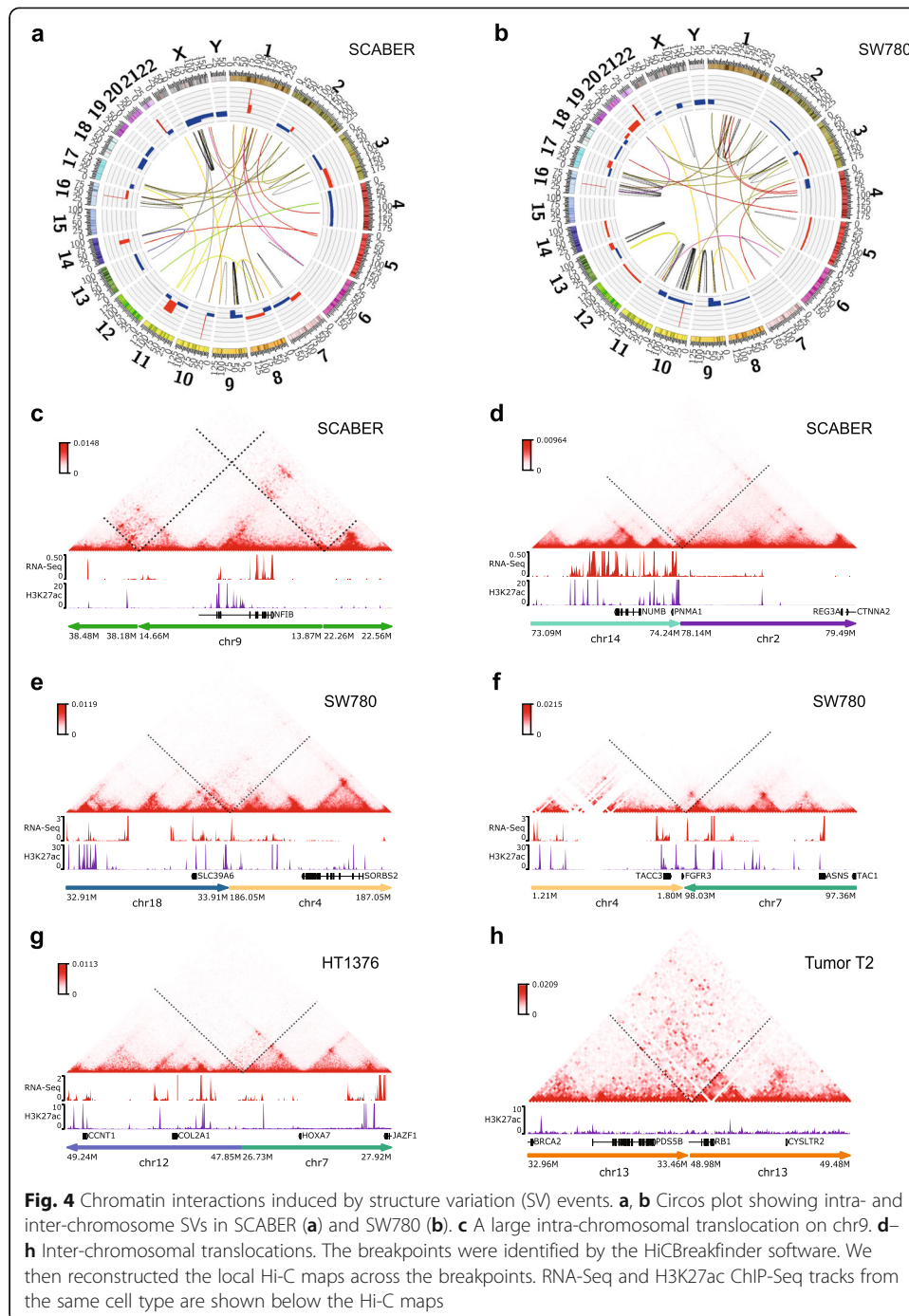
We first systematically identified copy number variations (CNVs) and SV events using the Hi-C data with HiNT [41] and the Hi-Cbreakfinder [42] software. We identified tens of large SVs, including inversions, deletions, and translocations (Fig. 4a, b, Additional file 4: Figures S4–S5, Additional file 15: Table 14). As might be expected, we observed fewer CNVs in the patient samples than in cell lines. More importantly, we were able to re-construct the local Hi-C map surrounding the breakpoints of the SVs. We can observe interesting enhancer-hijacking events and the formation of neo-TADs in these local Hi-C maps (Fig. 4c–h). These observations provide an important resource to further study the function of the re-arranged enhancers in the context of bladder cancer.

Neuronal PAS Domain Protein 2 (NPAS2) is a novel luminal BLCA TF which regulates luminal gene expression and cell migration

Genome-wide open chromatin analysis of BLCA cell lines provides an ideal platform for the identification of novel transcriptional regulators of BLCA cell fate and phenotype. Here we performed motif analysis of luminal-associated, basal-associated, and shared open chromatin regions, resulting in the identification of distinct TFs in each cluster. Among them, many represent known families of subtype-specific regulators, such as the GATA, FOX, and ETS families at luminal-associated ATAC-Seq peaks. Among them, we noticed a potential novel bHLH containing regulator, NPAS2, which is enriched in the luminal-associated and shared clusters, but not enriched in basal-associated ATAC-Seq peaks (Fig. 5a). We examined its binding profile using the latest ENCODE data (HEPG2 cells) [43] and found that NPAS2 binds at the FOXA1 promoter region (Fig. 5b), but not at regulatory regions for basal marker genes. This suggests the possibility that NPAS2 may be an upstream regulator of FOXA1. We then checked the TCGA data and found that high expression level of NPAS2 is significantly correlated to overall patient survival (Fig. 5c).

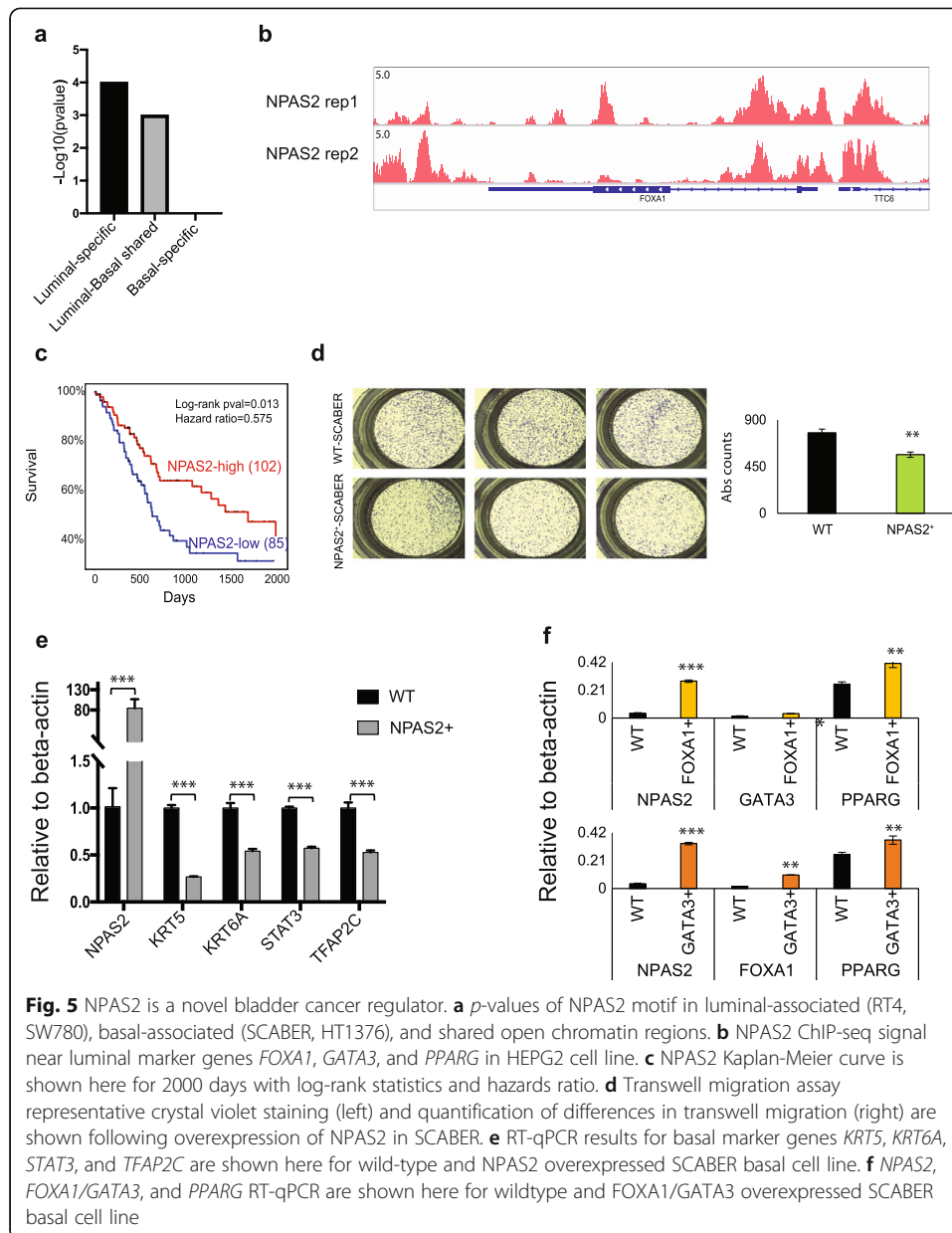
To further determine whether NPAS2 expression influences the downstream target expression and phenotype, we overexpressed NPAS2 in the basal-squamous BLCA cell line SCABER. First, we performed trans-well migration assays and found that overexpression of NPAS2 in SCABER cells decreased cell trans-well migration (Fig. 5d). We then performed RT-qPCR experiments and found that the basal marker genes (such as *KRT5*, *KRT6A*, and *TFAP2C*) are significantly downregulated (Fig. 5e) following NPAS2 overexpression, suggesting NPAS2 represses the expression of a subset of basal marker genes.

Because our functional genomics analysis suggests that FOXA1 and GATA3 cooperate to regulate luminal target genes [8], we individually overexpressed FOXA1 and GATA3 in SCABER cells to test their ability to regulate NPAS2 expression. We observed increased expression of NPAS2 by both FOXA1 and GATA3 overexpression (Fig. 5f).



Discussion

Muscle invasive BLCA is a morbid and expensive disease to treat [3, 44–46]. However, with recent development of immunotherapies such as anti-PD-1 [47] and PD-L1 [48], as well as targeted approaches including FGFR3 inhibitors, clinical management has been revolutionized [49]. However, response rates to these and other standard approaches are suboptimal, suggesting the need for increased molecular understanding. In keeping with this, recent National Comprehensive Cancer Network (NCCN)



guidelines have encouraged biomarker and molecular-based subtype studies to further stratify patients for recent targeted therapies [50].

It has been suggested that RNA-Seq-based molecular subtyping of BLCA is prognostic of clinical outcomes in patients [6–8]. While TCGA and other studies have identified mRNA-based molecular subtypes, the epigenetic differences underlying these expression subtypes are unknown. The Encyclopedia of DNA Elements (ENCODE) Consortium has contributed greatly to the current understanding of how epigenetic modifications in multiple tissues vary to regulate tissue-specific gene expression [29]. Histone modification states such as H3K27ac, among various other epigenetic states, mark enhancers and promoters that form a complex interacting network hub to regulate gene expression [51, 52]. TCGA has incorporated DNA methylation to

understanding epigenetic states [53]. DNA methylation states have been shown to be coupled with histone modifications particularly at the CpG sites at promoters [54]. However large changes in epigenetic histone modification states that influence gene expression lie in distal regions in enhancers and other sites that orchestrate the 3D genome (CTCF) [29, 38, 55–57]. Hence, our study utilized large-scale genomic experiments such as ATAC-Seq and H3K27ac ChIP-Seq, as well as FOXA1 and GATA3 ChIP-Seq and Hi-C combined with RNA-Seq to construct a comprehensive molecular map of luminal and basal BLCA in both cell line models as well as patient tumors. We have further utilized TCGA datasets to orthogonally validate and derive inferences for clinical importance of our findings.

We found evidence for regulation of luminal and basal bladder cancer genes by proximal-promoters and distal enhancers that form long-range chromatin loops and potentially drive oncogenic programs. Our findings are largely in agreement with previously known work on the role of FOXA1 and GATA3 in the regulation and maintenance of oncogenic programs in luminal bladder cancers [8, 19–21]. Interestingly, we found a novel co-regulation of FOXA1 and GATA3 with a common binding partner of AP-1 complex like breast cancers [58, 59] that appears to drive activity of distal enhancers, but not promoters. Our comprehensive 3D genome map shows distinct chromatin loop interaction networks available to luminal and basal BLCA. To our knowledge, this is the first report of a comprehensive 3D genomic map of bladder tumor patients. Our analysis further demonstrates the regulation of distal enhancers with promoters of BLCA subtype-specific genes through physical loops as reported in other studies [37, 38, 52]. Through our analysis, we have identified a novel bHLH TF regulator of luminal BLCA—NPAS2—whose expression correlates with overall survival of BLCA patients included in the TCGA cohort. Through several biological experiments, we showed that NPAS2 regulates the expression of several genes which serve as markers of basal-squamous BLCA, and further diminishes the migration ability of basal BLCA cells. Most importantly, our work highlights how these TFs can cooperatively regulate molecular subtypes and drive clinical associations.

The clinical implication of identification of potential regulators of primary luminal differentiation such as NPAS2 is that, once identified, these factors can be leveraged or targeted. In cases in which the cancers are less luminal and more basal, shifting the biology of the tumor to a more luminal gene expression subtype by activating NPAS2 (and other required factors) could slow the growth of a basal tumors. Alternatively, 30% of muscle invasive tumors are luminal, with upregulated luminal pathways, and blocking the function of luminal tumors could potentially improve survival in patients with luminal BLCA.

Although our studies provide an excellent starting point by identifying associations between epigenetic landscape and 3D genome architecture with tumor subtype, increased numbers of tumor specimens will be required. However, the cost of sequencing limited our current capability to include a large set of patient tumors. An additional limitation of our current study is the lack of a consortium-level analysis of ChIP-Seq data for histone modifications and all major TFs. Such an approach would increase precision in our regulatory analysis. However, previous studies were limited to the understanding of single or combination of few TFs in the context of gene regulation. Therefore, we believe that our study will emerge to be a solid comprehensive resource to

launch a further series of hypothesis-driven biological experiments based on gene and epigenetic perturbations to unveil both novel molecular targets as well as biomarkers.

Materials and methods

Cell lines and patient tumor samples

Bladder cancer cell lines—RT4, SCABER, SW780, and HT1376—were obtained from ATCC and cultured as previously described [8]. Bladder tumor samples were obtained from Penn State Hershey, College of Medicine's biobank storage at the Institute of Personalized Medicine (IPM) with appropriate protocol approval from the institutional review board (IRB Number: STUDY00001117). The samples from Northwestern University were also obtained with proper approval from the institutional board (IRB number: STU00088853). Samples were selected based on its availability (50 mg) for several rounds of sequencing experiments.

Cell culture

Bladder cancer cell lines were cultured in growth medium containing media + 10% fetal bovine serum and 1% penicillin and streptomycin (Corning). RT4, SW780, SCABER, and HT1376 cells were cultured in McCoy's 5A (Gibco), RPMI-1640 (Corning), Eagle minimum essential medium (MEM; GE life sciences), and Eagle MEM with 1% non-essential amino acids (Corning), respectively. Cells were plated in tissue culture plates (TCPs, Corning)—T-25, T-75 or 15-cm dishes to further grow and expand in 5% CO₂ humidified incubator for several different sequencing experiments. For future storage, cells were preserved in 5% DMSO containing growth medium in the vapor of liquid nitrogen. For passaging, cells were washed with phosphate-buffered saline (PBS, Corning) and trypsinized for 5 min to detach cells from the TCPs. They were further spun down at 200×g to pellet and washed with PBS for further experiments.

RNA-Seq

For RNA-Seq, RNA was extracted from frozen cell pellets using RNeasy Mini Kit (Qiagen). Extracted RNA was quantitated using NanoDrop (Thermo Scientific). SureSelect strand-specific RNA library preparation kit (Agilent) was used to generate cDNA libraries where polyA RNA was pulled down using 2 µg of oligo (dT) beads. Extracted cDNA was then fragmentation, reverse transcribed, end repaired, 3'-end adenylated, adaptor ligation, and subsequently amplified and beads purified (Beckman Coulter). Barcode sequences were thus used to multiplex high-throughput sequencing. The cDNA library was QC'ed for size distribution and concentration using BioAnalyzer (Agilent) High Sensitivity DNA Kit (Agilent) and Kapa Library Quantification Kit (Kapa Biosystems). Final libraries were then pooled and diluted to 2 nM and subsequently sequenced using Illumina HiSeq, X Ten, or NovaSeq platform (Illumina).

Chromatin crosslinking and ChIP-Seq library preparation

Each cell line was grown in 15-cm dishes (× 4) with 25-mL growth medium where they were detached from the TCPs using trypsin as described above and further pelleted. Approximately 10–20 M cell pellets (2x biological replicates) were crosslinked immediately with 1% formaldehyde in PBS at RT for 10 min and subsequently quenched with

0.125 M glycine for 5 min. Crosslinked cells were then washed in PBS and 100 μ L freshly prepared lysis buffer (1% SDS, 50 mM Tris-HCl pH 8, 20 mM EDTA, and 1x complete protease inhibitor) was added. Lysed cells were then diluted in 900 μ L TE buffer and sonicated using focused beam ultrasound sonicator (COVARIS). Sonicated samples were repeated for extended periods of time (up to 1.5 h) until the chromatin size distribution of \sim 200–300 bp was achieved. Sonicated DNA-chromatin complexes were then pulled down with anti-H3K27ac antibody and washed several times with RIPA buffer to remove non-specific bindings. Pulled-down samples as well as input controls were all de-crosslinked at 65 °C overnight. Samples were treated with RNase and Proteinase K digestion at 37 °C and 55 °C, respectively, followed by further DNA library extraction using phenol-chloroform method. The library was then prepared using Kapa Hyper Prep Kit (KAPA) and further amplified using Hi-fidelity KAPA PCR kit (KAPA) for 6–11 cycles and purified with Kapa pure beads (KAPA). The final library was quantified using Qubit high sensitivity DNA assay (ThermoFisher) and then sequenced using Illumina's HiSeq platform 2500, X Ten, or NovaSeq platform (Illumina).

Nuclei extraction and ATAC-Seq library preparation

Each cell line was pelleted following detachment from the TCPs, as described above. Cells were washed with PBS, counted and kept on ice as a pellet. Fifty thousand cells were used for ATAC-Seq library preparation as described by Greenleaf [60]. First, pelleted cells were reconstituted in a freshly made lysis buffer to remove unwashed mitochondria DNA, spun down, and the buffer was discarded. Then, the pelleted nuclei were tagmented with Tn5 transposase (Illumina ref: 15027866 & 15027865) in 50 μ L volume for 30 min at 37 °C and the DNA was subsequently purified using Qiagen MinElute kit (Qiagen). Then, the library was amplified using Hi-fidelity KAPA PCR kit (KAPA) with Nextera's PCR nonbarcoded Ad1 and barcoded Ad2.* primers for 6 cycles and purified using Kapa pure beads (KAPA). The library was then quantified using Qubit high sensitivity DNA assay (ThermoFisher) and further assessed for quality using bioanalyzer (Agilent). Libraries were either sequenced using Illumina's HiSeq platform 2500, X Ten, or NovaSeq platform (Illumina).

Chromatin crosslinking and HiC library preparation

Each cell line was grown in T-25 flasks (\times 4) with 5-mL growth medium, trypsinized as described above, and pelleted. Approximately 4–5 M cells as pellets were crosslinked immediately with 2% formaldehyde in PBS at RT for 10 min. Crosslinked cells were then washed in PBS and frozen as 1–1.5 M aliquots in -80 °C for several months before library preparation. We used ARIMA's Hi-C kit for making libraries (ARIMA Genomics). As per their protocol, we tested and did QC on samples and sequenced 300 M–600 M reads for each sample using Illumina's NovaSeq platform (Illumina).

Overexpression of NPAS2

Overexpression of FLAG (DYKKK) tagged NPAS2 protein was performed using genscript plasmids, which were expanded following transformation in competent *Escherichia coli* and picking clones. We then transiently transfected SCABER cell lines with

2 µg plasmid/well in 6-well plates (2x wells). For transfection, we used Lipofectamine 3000 reagent (Invitrogen) as per manufacturer protocol and allowed cells to be cultured up to 2 days before collecting it for various analysis.

Quantitative reverse transcription PCR (RT-qPCR)

We used the following primers for RT-qPCR to detect mRNA levels.

NPAS2: 5'-ACACCCTTCAAGACCTTGCC-3' (F); 5'-AGGTTCGTCAACTATG CACATTT-3' (R)

FOXA1: 5'-CGCTTCGCACAGGGCTGGAT-3' (F); 5'-TGCTGACCGGGACGGA GGAG-3' (R)

GATA3: 5'-ATACACCACCTACCCGCCTAC-3' (F); 5'-ACTCCCTGCCTTCTGT GCT-3' (R)

PPARG: 5'-GATCTCCAGTGATATCGACCAGC-3' (F); 5'-GATGGCCACCTCTT TGCTCTG-3' (R)

TFAP2C: 5'- ATCGAAAAATGGAGGCCGGT-3' (F); 5'-CGGCTTCACAGACA TAGGCA-3' (R)

STAT3: 5'- GAGGACTGAGCATCGAGCA-3' (F); 5'-CATGTGATCTGACACC CTGAA-3' (R)

KRT5: 5'- GATCGCCACTTACCGCAAGC-3' (F); 5'-ACTGCCATATCCAGAGGA AACA-3' (R)

KRT6A: 5'-AAGTGTTGTGAACCCCCACCC-3' (F); 5'-AGCAATTGCAAACA GCGAAGAG-3' (R)

beta-actin: 5'-CATGTACGTTGCTATCCAGGC-3' (F); 5'-CTCCTTAATGTAC GCACGAT-3' (R)

RNA was extracted from frozen cell pellets using RNeasy Mini Kit (Qiagen). Samples were treated with DNase to digest any additional DNA extracted during the process. DNase-free RNA was then further converted to cDNA using reverse transcriptase kit (Invitrogen) according to the manufacturer's protocol. Reverse transcribed cDNA was then assayed for RT-qPCR using KAPA SYBR FAST qPCR Master Mixes (Roche) at 60 °C melting temperature and quantitated using BioRad quantitative PCR system. CT values obtained through the quantitation were then normalized to beta-actin and further transformed to relative expression shown in plots.

Transwell migration assay

Transwell migration assay was performed using 8 µm PVDF inserts (Corning) in a transwell chamber fitting into 24-well plates (Corning). Each cells (control and overexpressing NPAS2 for 2 days) were seeded with 50,000 cells in a transwell (× 3 replicates) in a FBS-free medium containing 1%PS. Cells were allowed to migrate through the transwell inserts for 24 h into medium containing regular 10% FBS and 1%PS. Transwell chambers were removed, washed with PBS and stained with crystal violet. Cells that were not migrated through the insert were removed using a Q-tip. Migrated cells were then visualized in microscope and scored for number of stained spots and compared with different experiment groups. *T*-test was used to calculate significance between groups.

Computational analysis methods are provided in Supplementary information part as Additional file 16.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13059-021-02325-y>.

Additional file 1: Table S1. Classification results for each sample using consensusMIBC.

Additional file 2: Table S2. Cell Lines and Experiment table.

Additional file 3: Table S3. RNA-Seq correlation between cell lines.

Additional file 4: Figure S1. Epigenetic landscape analysis of histone modifications in luminal and basal bladder cancers. **a** Genome-wide H3K27ac signals show that biological replicates and molecular subtypes (basal and luminal) cluster together. **b** Hierarchical clustering of genome-wide RNA-Seq results for 4 cell lines recapitulate the luminal and basal gene expression based molecular subtypes. **c** Integrated H3K27ac peaks at promoters and RNA-Seq gene expression association model identifies putative promoter and gene regulation. Top 10,000 most variable promoters (left heatmap) are plotted along with their corresponding gene expression (right heatmap). Luminal (cyan) and basal (magenta) genes are highlighted for their specific linked enhancers. **d** Corresponding enhancer H3K27ac and its linked RNA-Seq signals based on our predicted model for selected luminal and basal genes shows remarkable similarity. **Figure S2.** Epigenetic landscape analysis of open-chromatins in luminal and basal subtypes of bladder cancers. **a** Genome-wide overlap of ATAC-Seq peaks with H3K27ac ChIP-Seq is shown here for each cell line at either promoter, enhancer or all locations. **b** Genome-wide overlap of H3K27ac ChIP-Seq peaks with ATAC-Seq is shown here for each cell line at either promoter, enhancer or all locations. **c** Overlap between distal H3K27ac and ATAC-Seq peaks. **d** ATAC-Seq signal at distal enhancers compared with distal H3K27ac signal. **e** Genome-wide correlation of ATAC-Seq signals between cell lines recapitulate enhancer/promoter and RNA-Seq based clustering. **f** FOXA1 and GATA3 ChIP-Seq binding sites overlapped at promoters are shown here as genome-wide tag plot in three groups. **g** A comparison of top 3 motifs enriched *p*-values in each open-chromatins that does not overlap with any H3K27ac signals within its cell lines are shown. **Figure S3.** Hi-C maps of luminal and basal subtypes of bladder cancers and bladder tumors. Genome-wide chromosome view of Hi-C map is shown for RT4 (**a**), SW780 (**b**), SCABER (**c**), HT1376 (**d**), tumor T1 (**e**), tumor T2 (**f**), tumor T3 (**g**), tumor T4 (**h**) and tumor T5 (**i**) at 10 MB resolution. **Figure S4.** Copy number profiles for four bladder cancer cell lines (HT1376, RT4, SW780, and SCABER) and five tumor samples (Tumor T1, Tumor T2, Tumor T3, Tumor T4, and Tumor T5). CNVs were computed using Hi-C data. **Figure S5.** Intra- and inter-chromosome structure variation (SV) events. Circos plot showing intra- and inter-chromosome SVs in HT1376 (**a**), RT4 (**b**), Tumor T1 (**c**), Tumor T2 (**d**), Tumor T3 (**e**), Tumor T4 (**f**) and Tumor T5 (**g**).

Additional file 5: Table S4. Basal up-regulated genes from differentially expressed genes (DESeq2) analysis.

Additional file 6: Table S5. Luminal up-regulated genes from differentially expressed genes (DESeq2) analysis.

Additional file 7: Table S6. H3K27ac ChIP-Seq peaks called in each cell line and its proximal/distal location w.r.t transcription start site (TSS).

Additional file 8: Table S7. Epigenetic landscape of H3K27ac signals correlated and linked to nearby genes.

Additional file 9: Table S8. ATAC-Seq peaks called in each cell line and its proximal/distal location w.r.t transcription start site (TSS).

Additional file 10: Table S9. ATAC-Seq clusters representing Luminal-specific, Basal-specific and shared open-chromatins at distal-enhancers.

Additional file 11: Table S10. TF motif analysis of luminal specific, basal specific and shared open-chromatin. Clusters at distal-enhancers.

Additional file 12: Table S11. FOXA1 and GATA3 ChIP-Seq peaks called in RT4 cells.

Additional file 13: Table S12. Peakachu Loops called in each bladder cancer cell line.

Additional file 14: Table S13. Luminal-specific and Basal-specific loops.

Additional file 15: Table S14. Number of large structural variations (SVs) in bladder cancer cell lines and tumor samples.

Additional file 16. Computational data analysis methods.

Additional file 17. Review history.

Review history

The review history is available as Additional file 17.

Peer review information

Anahita Bishop was the primary editor of this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

Authors' contributions

The study was conceived and supervised by F.Y. and D.D. T.J., B.Z., Y.H., Q.J., H. Yang, T. L, J.X., and H. Yamashita performed the experiments. T.J., Q.W., X.W., F.S., L.Z., R.C., Y.L., and H.L. performed the data analysis. B.Z. did all the work while she was in Northwestern University. T.J., B.Z., and Y.H. performed the ATAC-Seq, ChIP-Seq, and Hi-C experiments in both cell lines and primary tissues. J.W., J.D.R., and J.J.M. provided clinical insights. T.J., B.Z., Q.W., Y.H., D.D., and F.Y. wrote the manuscript. All authors approved the final version for publication.

Funding

F.Y. is supported by 1R35GM124820, R01HG009906, U01CA200060, and R24DK106766. DJD is supported by RSG1723301TBE (American Cancer Society) and the Bladder Cancer Support Group at Penn State Health.

Availability of data and materials

Our data is available in gene expression omnibus (GEO) at GSE148079 [61].

Link to data: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE148079>.

Raw fastQ files for patient samples are available in European Genome-phenome Archive (EGA) at EGAS00001005071 [62]. Link to data: <https://ega-archive.org/studies/EGAS00001005071>.

The TCGA ATAC-seq datasets in Fig. 2f are from [30]. The TCGA datasets we used in Fig. 5c are from [53].

The analysis code for figures could be found in GitHub repository (<https://github.com/Qixuan771/Source-code-for-bladder-cancer-project>, under GNU General Public License v3.0) [63] and zenodo <https://zenodo.org/record/4396080#X-oRtOKgY0>, DOI: <https://doi.org/10.5281/zenodo.4396080>) [64].

Declarations

Ethics approval and consent to participate

Bladder tumor samples were obtained from Penn State Hershey, College of Medicine's biobank storage at the Institute of Personalized Medicine (IPM) with appropriate protocol approval from the institutional review board (IRB Number: STUDY00001117). The samples were also obtained from Northwestern University with proper approval from the institutional board (IRB number: STU00088853). All patients had previously provided written informed consent for tumor collection and subsequent analysis. This study was performed in compliance with the Helsinki Declaration.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Biochemistry and Molecular Biology, Penn State School of Medicine, Hershey, PA, USA. ²Department of Biochemistry and Molecular Genetics, Feinberg School of Medicine Northwestern University, Chicago, IL, USA. ³Present address: Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education/Beijing), Division of Etiology, Peking University Cancer Hospital and Institute, Beijing, China. ⁴Department of Pathology and Laboratory Medicine, The Pennsylvania State University, College of Medicine, Hershey, PA, USA. ⁵Department of Surgery, Division of Urology, The Pennsylvania State University, College of Medicine, Hershey, PA, USA. ⁶University of Illinois College of Medicine, Chicago, IL, USA. ⁷Department of Biochemistry and Molecular Biology, The Pennsylvania State University, College of Medicine, Hershey, PA, USA. ⁸Department of Urology, Feinberg School of Medicine and The Robert H. Lurie Comprehensive Cancer Center, Northwestern University, Chicago, IL, USA. ⁹Robert H. Lurie Comprehensive Cancer Center of Northwestern University, Chicago, IL, USA.

Received: 18 April 2020 Accepted: 25 March 2021

Published online: 15 April 2021

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin.* 2020;70(1):7–30. <https://doi.org/10.3322/caac.21590>.
2. Ferlay J, Colombet M, Soerjomataram I, Mathers C, Parkin DM, Pineros M, Znaor A, Bray F. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *Int J Cancer.* 2019;144:1941–53.
3. Sievert KD, Amend B, Nagele U, Schilling D, Bedke J, Horstmann M, Hennenlotter J, Kruck S, Stenzl A. Economic aspects of bladder cancer: what are the benefits and costs? *World J Urol.* 2009;27(3):295–300. <https://doi.org/10.1007/s00345-009-0395-z>.
4. Choi W, Porten S, Kim S, Willis D, Plimack ER, Hoffman-Censits J, Roth B, Cheng T, Tran M, Lee IL, Melquist J, Bondaruk J, Majewski T, Zhang S, Pretzsch S, Baggerly K, Siefker-Radtke A, Czerniak B, Dinney CPN, McConkey DJ. Identification of distinct basal and luminal subtypes of muscle-invasive bladder cancer with different sensitivities to frontline chemotherapy. *Cancer Cell.* 2014;25(2):152–65. <https://doi.org/10.1016/j.ccr.2014.01.009>.
5. Seiler R, Ashab HAD, Erho N, van Rhijn BWG, Winters B, Douglas J, Van Kessel KE, van de Putte EE F, Sommerlad M, Wang NQ, et al. Impact of molecular subtypes in muscle-invasive bladder cancer on predicting response and survival after Neoadjuvant chemotherapy. *Eur Urol.* 2017;72(4):544–54. <https://doi.org/10.1016/j.eururo.2017.03.030>.
6. Choi W, Czerniak B, Ochoa A, Su X, Siefker-Radtke A, Dinney C, McConkey DJ. Intrinsic basal and luminal subtypes of muscle-invasive bladder cancer. *Nat Rev Urol.* 2014;11(7):400–10. <https://doi.org/10.1038/nrurol.2014.129>.
7. WeinStein JNAR, Broom BM, Wang W, Verhaak RG, McConkey D, Lerner S. Comprehensive molecular characterization of urothelial bladder carcinoma. *Nature.* 2014;507:315–22.
8. Warrick JI, Walter V, Yamashita H, Chung E, Shuman L, Amponsa VO, Zheng Z, Chan W, Whitcomb TL, Yue F, Iyyanki T, Kawasawa YI, Kaag M, Guo W, Raman JD, Park JS, DeGraff DJ. FOXA1, GATA3 and PPAR cooperate to drive luminal subtype in bladder cancer: a molecular analysis of established human cell lines. *Sci Rep.* 2016;6(1):38531. <https://doi.org/10.1038/srep38531>.
9. Karni-Schmidt O, Castillo-Martin M, Shen TH, Gladoun N, Domingo-Domenech J, Sanchez-Carbayo M, Li Y, Lowe S, Prives C, Cordon-Cardo C. Distinct expression profiles of p63 variants during urothelial development and bladder cancer progression. *Am J Pathol.* 2011;178(3):1350–60. <https://doi.org/10.1016/j.ajpath.2010.11.061>.
10. Choi W, Shah JB, Tran M, Svatek R, Marquis L, Lee IL, Yu D, Adam L, Wen S, Shen Y, Dinney C, McConkey DJ, Siefker-Radtke A. p63 expression defines a lethal subset of muscle-invasive bladder cancers. *PLoS One.* 2012;7(1):e30206. <https://doi.org/10.1371/journal.pone.0030206>.

11. Tran MN, Choi W, Wszolek MF, Navai N, Lee IL, Nitti G, Wen S, Flores ER, Siefker-Radtke A, Czerniak B, et al. The p63 protein isoform DeltaNp63alpha inhibits epithelial-mesenchymal transition in human bladder cancer cells: role of MIR-205. *J Biol Chem.* 2013;288(5):3275–88. <https://doi.org/10.1074/jbc.M112.408104>.
12. Palmboos PL, Wang Y, Bankhead Iii A, Kelleher AJ, Wang L, Yang H, Ahmet ML, Gumkowski ER, Welling SD, Magnuson B, et al. ATDC mediates a TP63-regulated basal cancer invasive program. *Oncogene.* 2019;38(18):3340–54. <https://doi.org/10.1038/s41388-018-0646-9>.
13. Gatta LB, Melocchi L, Bugatti M, Missale F, Lonardi S, Zanetti B, Cristinelli L, Belotti S, Simeone C, Ronca R, Grillo E, Licini S, Bresciani D, Tardanico R, Chan SR, Giuriso E, Calza S, Vermi W. Hyper-activation of STAT3 sustains progression of non-papillary basal-type bladder cancer via FOXL1 regulome. *Cancers (Basel).* 2019;11(9). <https://doi.org/10.3390/cancers11091219>.
14. Ho PL, Kurtova A, Chan KS. Normal and neoplastic urothelial stem cells: getting to the root of the problem. *Nat Rev Urol.* 2012;9(10):583–94. <https://doi.org/10.1038/nrurol.2012.142>.
15. Chan KS, Espinosa I, Chao M, Wong D, Ailles L, Diehn M, Gill H, Presti J Jr, Chang HY, van de Rijn M, et al. Identification, molecular characterization, clinical prognosis, and therapeutic targeting of human bladder tumor-initiating cells. *Proc Natl Acad Sci U S A.* 2009;106(33):14016–21. <https://doi.org/10.1073/pnas.0906549106>.
16. Yamashita H, Kawasawa YI, Shuman L, Zheng Z, Tran T, Walter V, Warrick JI, Chen G, Al-Ahmadie H, Kaag M, et al. Repression of transcription factor AP-2 alpha by PPARgamma reveals a novel transcriptional circuit in basal-squamous bladder cancer. *Oncogenesis.* 2019;8(12):69. <https://doi.org/10.1038/s41389-019-0178-3>.
17. Osei-Amponsa V, Buckwalter JM, Shuman L, Zheng Z, Yamashita H, Walter V, Wildermuth T, Ellis-Mohl J, Liu C, Warrick JI, Shantz LM, Feehan RP, al-Ahmadie H, Mendelsohn C, Raman JD, Kaestner KH, Wu XR, DeGraff DJ. Hypermethylation of FOXA1 and allelic loss of PTEN drive squamous differentiation and promote heterogeneity in bladder cancer. *Oncogene.* 2020;39(6):1302–17. <https://doi.org/10.1038/s41388-019-1063-4>.
18. Liu C, Tate T, Baturina E, Truschel ST, Potter S, Adam M, Xiang T, Picard M, Reiley M, Schneider K, Tamargo M, Lu C, Chen X, He J, Kim H, Mendelsohn CL. Pparg promotes differentiation and regulates mitochondrial gene expression in bladder epithelial cells. *Nat Commun.* 2019;10(1):4589. <https://doi.org/10.1038/s41467-019-12332-0>.
19. DeGraff DJ, Clark PE, Cates JM, Yamashita H, Robinson VL, Yu X, Smolkin ME, Chang SS, Cookson MS, Herrick MK, et al. Loss of the urothelial differentiation marker FOXA1 is associated with high grade, late stage bladder cancer and increased tumor proliferation. *PLoS One.* 2012;7(5):e36669. <https://doi.org/10.1371/journal.pone.0036669>.
20. Li Y, Ishiguro H, Kawahara T, Kashiwagi E, Izumi K, Miyamoto H. Loss of GATA3 in bladder cancer promotes cell migration and invasion. *Cancer Biol Ther.* 2014;15(4):428–35. <https://doi.org/10.4161/cbt.27631>.
21. Theodorou V, Stark R, Menon S, Carroll JS. GATA3 acts upstream of FOXA1 in mediating ESR1 binding by shaping enhancer accessibility. *Genome Res.* 2013;23(1):12–22. <https://doi.org/10.1101/gr.139469.112>.
22. Kong SL, Li G, Loh SL, Sung WK, Liu ET. Cellular reprogramming by the conjoint action of ERalpha, FOXA1, and GATA3 to a ligand-inducible growth state. *Mol Syst Biol.* 2011;7(1):526. <https://doi.org/10.1038/msb.2011.59>.
23. Glont SE, Chernukhin I, Carroll JS. Comprehensive genomic analysis reveals that the pioneering function of FOXA1 is independent of hormonal signaling. *Cell Rep.* 2019;26:2558–2565.e2553.
24. Zhang G, Zhao Y, Liu Y, Kao LP, Wang X, Skery B, Li Z. FOXA1 defines cancer cell specificity. *Sci Adv.* 2016;2(3):e1501473. <https://doi.org/10.1126/sciadv.1501473>.
25. Earl J, Rico D, Carrillo-de-Santa-Pau E, Rodriguez-Santiago B, Mendez-Pertuz M, Auer H, Gomez G, Grossman HB, Pisano DG, Schulz WA, et al. The UBC-40 Urothelial bladder cancer cell line index: a genomic resource for functional studies. *BMC Genomics.* 2015;16(1):403. <https://doi.org/10.1186/s12864-015-1450-3>.
26. Kamoun A, de Reynies A, Allory Y, Sjadahl G, Robertson AG, Seiler R, Hoadley KA, Groeneveld CS, Al-Ahmadie H, Choi W, et al. A consensus molecular classification of muscle-invasive bladder cancer. *Eur Urol.* 2020;77(4):420–33. <https://doi.org/10.1016/j.eururo.2019.09.006>.
27. Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA, Wysocka J. A unique chromatin signature uncovers early developmental enhancers in humans. *Nature.* 2011;470(7333):279–83. <https://doi.org/10.1038/nature09692>.
28. Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim TK, Koche RP, Lee W, Mendenhall E, O'Donovan A, Presser A, Russ C, Xie X, Meissner A, Wernig M, Jaenisch R, Nusbaum C, Lander ES, Bernstein BE. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature.* 2007;448(7153):553–60. <https://doi.org/10.1038/nature06008>.
29. Dunham IKA, Aldred SF, Collins PJ, Davis CA, Doyle, Lochovsky L, Min R, Mu XJ, Rozowsky J, Yan KK, Yip KY, Birney EF, Epstein CB, Fritze S, Harrow J, Kaul R, Khatun J, Lajoie BR, Landt SG, Lee BK, Pauli F, Rosenbloom KR, Sabo P, Safi A, Sanyal A, Shores N, Simon JM, Song L, Trinklein ND, Altshuler RC, Birney E, Brown JB, Cheng C, Djebali S, Dong X, Dunham I, Ernst J, Furey TS, Gerstein M, Giardine B, Greven M, Hardison RC, Harris RS, Herrero J, Hoffman MM, Iyer S, Kellis M, Khatun J, Kheradpour P, Kundaje A, Lassmann T, Li Q, Lin X, Marinov GK, Merkel A, Mortazavi A, Parker SC, Reddy TE, Rozowsky J, Schlesinger F, Thurman RE, Wang J, Ward LD. An integrated encyclopedia of DNA elements in the human genome. *Nature.* 2012;489:57–74.
30. Corces MR, Granja JM, Shams S, Louie BH, Seoane JA, Zhou W, Silva TC, Groeneveld C, Wong CK, Cho SW, Satpathy AT, Mumbach MR, Hoadley KA, Robertson AG, Sheffield NC, Felau I, Castro MAA, Berman BP, Staudt LM, Zenklusen JC, Laird PW, Curtis C, The Cancer Genome Atlas Analysis Network, Greenleaf WJ, Chang HY. The chromatin accessibility landscape of primary human cancers. *Science.* 2018;362(6413):eaav1898. <https://doi.org/10.1126/science.aav1898>.
31. Fishwick C, Higgins J, Percival-Alwyn L, Hustler A, Pearson J, Bastkowski S, Moxon S, Swarbrick D, Greenman CD, Southgate J. Hierarchy of transcription factors driving basal and luminal cell phenotypes in human urothelium. *Cell Death Differ.* 2017;24(5):809–18. <https://doi.org/10.1038/cdd.2017.10>.
32. Garces de Los Fayos Alonso I, Liang HC, Turner SD, Lagger S, Merkel O, Kenner L. The role of Activator Protein-1 (AP-1) family members in CD30-positive lymphomas. *Cancers (Basel).* 2018;10(4):93. <https://doi.org/10.3390/cancers10040093>. PMID: 29597249; PMCID: PMC5923348.
33. Sano T, Kobayashi T, Ogawa O, Matsuda M. Gliding basal cell migration of the urothelium during wound healing. *Am J Pathol.* 2018;188(11):2564–73. <https://doi.org/10.1016/j.ajpath.2018.07.010>.
34. Gandhi D, Molotkov A, Baturina E, Schneider K, Dan H, Reiley M, Laufer E, Metzger D, Liang F, Liao Y, Sun TT, Aronow B, Rosen R, Mauney J, Adam R, Rosselot C, van Batavia J, McMahon A, McMahon J, Guo JJ, Mendelsohn C. Retinoid signaling in progenitors controls specification and regeneration of the urothelium. *Dev Cell.* 2013;26(5):469–82. <https://doi.org/10.1016/j.devcel.2013.07.017>.

35. Reddy KL, Zullo JM, Bertolino E, Singh H. Transcriptional repression mediated by repositioning of genes to the nuclear lamina. *Nature*. 2008;452(7184):243–7. <https://doi.org/10.1038/nature06727>.
36. Salameh TJ, Wang X, Song F, Zhang B, Wright SM, Khunsriraksakul C, Yue F. A supervised learning framework for chromatin loop detection in genome-wide contact maps. *Nat Commun*. 2020;11(1):3428. <https://doi.org/10.1038/s41467-020-17239-9>. PMID: 32647330; PMCID: PMC7347923.
37. Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, Aiden EL. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*. 2014;159(7):1665–80. <https://doi.org/10.1016/j.cell.2014.11.021>.
38. Rhie SK, Perez AA, Lay FD, Schreiner S, Shi J, Polin J, Farnham PJ. A high-resolution 3D epigenomic map reveals insights into the creation of the prostate cancer transcriptome. *Nat Commun*. 2019;10(1):4154. <https://doi.org/10.1038/s41467-019-12079-8>.
39. Franke M, Ibrahim DM, Andrey G, Schwarzer W, Heinrich V, Schopflin R, Kraft K, Kempfer R, Jerkovic I, Chan WL, et al. Formation of new chromatin domains determines pathogenicity of genomic duplications. *Nature*. 2016;538(7624):265–9. <https://doi.org/10.1038/nature19800>.
40. Ooi WF, Nargund AM, Lim KJ, Zhang S, Xing M, Mandoli A, Lim JQ, Ho SWT, Guo Y, Yao X, Lin SJ, Nandi T, Xu C, Ong X, Lee M, Tan AL, Lam YN, Teo JX, Kaneda A, White KP, Lim WK, Rozen SG, Teh BT, Li S, Skanderup AJ, Tan P. Integrated paired-end enhancer profiling and whole-genome sequencing reveals recurrent CCNE1 and IGF2 enhancer hijacking in primary gastric adenocarcinoma. *Gut*. 2020;69(6):1039–52. <https://doi.org/10.1136/gutjnl-2018-317612>. Epub 2019 Sep 21. PMID: 31542774.
41. Wang S, Lee S, Chu C, Jain D, Kerpedjiev P, Nelson GM, Walsh JM, Alver BH, Park PJ. HiNT: a computational method for detecting copy number variations and translocations from Hi-C data. *Genome Biol*. 2020;21(1):73. <https://doi.org/10.1186/s13059-020-01986-5>.
42. Dixon JR, Xu J, Dileep V, Zhan Y, Song F, Le VT, Yardimci GG, Chakraborty A, Bann DV, Wang Y, et al. Integrative detection and analysis of structural variation in cancer genomes. *Nat Genet*. 2018;50(10):1388–98. <https://doi.org/10.1038/s41588-018-0195-8>.
43. Partridge EC, Chhetri SB, Prokop JW, Ramaker RC, Jansen CS, Goh ST, Mackiewicz M, Newberry KM, Brandsmeier LA, Meadows SK, Messer CL, Hardigan AA, Coppola CJ, Dean EC, Jiang S, Savic D, Mortazavi A, Wold BJ, Myers RM, Mendenhall EM. Occupancy maps of 208 chromatin-associated proteins in one human cell type. *Nature*. 2020;583(7818):720–8. <https://doi.org/10.1038/s41586-020-2023-4>.
44. Yabroff KR, Lund J, Kepka D, Mariotto A. Economic burden of cancer in the United States: estimates, projections, and future research. *Cancer Epidemiol Biomark Prev*. 2011;20(10):2006–14. <https://doi.org/10.1158/1055-9965.EPI-11-0650>.
45. Botteman MF, Pashos CL, Redaelli A, Laskin B, Hauser R. The health economics of bladder cancer: a comprehensive review of the published literature. *Pharmacoeconomics*. 2003;21(18):1315–30. <https://doi.org/10.1007/BF03262330>.
46. Yeung C, Dinh T, Lee J. The health economics of bladder cancer: an updated review of the published literature. *Pharmacoeconomics*. 2014;32(11):1093–104. <https://doi.org/10.1007/s40273-014-0194-2>.
47. Balar AV, Castellano D, O'Donnell PH, Grivas P, Vuky J, Powles T, Plimack ER, Hahn NM, de Wit R, Pang L, Savage MJ, Perini RF, Keefe SM, Bajorin D, Bellmunt J. First-line pembrolizumab in cisplatin-ineligible patients with locally advanced and unresectable or metastatic urothelial cancer (KEYNOTE-052): a multicentre, single-arm, phase 2 study. *Lancet Oncol*. 2017;18(11):1483–92. [https://doi.org/10.1016/S1470-2045\(17\)30616-2](https://doi.org/10.1016/S1470-2045(17)30616-2).
48. Balar AV, Galsky MD, Rosenberg JE, Powles T, Petrylak DP, Bellmunt J, Loriot Y, Necchi A, Hoffman-Censits J, Perez-Gracia JL, Dawson NA, van der Heijden M, Dreicer R, Srinivas S, Retz MM, Joseph RW, Drakaki A, Vaishampayan UN, Sridhar SS, Quinn DI, Durán I, Shaffer DR, Eigel BJ, Grivas PD, Yu EY, Li S, Kadel EE 3rd, Boyd Z, Bourgon R, Hegde PS, Mariathasan S, Thåström A, Abidoye OO, Fine GD, Bajorin DF, IMvigor210 Study Group. Atezolizumab as first-line treatment in cisplatin-ineligible patients with locally advanced and metastatic urothelial carcinoma: a single-arm, multicentre, phase 2 trial. *Lancet*. 2017;389(10064):67–76. [https://doi.org/10.1016/S0140-6736\(16\)32455-2](https://doi.org/10.1016/S0140-6736(16)32455-2).
49. Loriot Y, Necchi A, Park SH, Garcia-Donas J, Huddart R, Burgess E, Fleming M, Rezazadeh A, Mellado B, Varlamov S, Joshi M, Duran I, Tagawa ST, Zakharia Y, Zhong B, Stuyckens K, Santiago-Walker A, de Porre P, O'Hagan A, Avadhani A, Siefker-Radtke AO. Erdafitinib in locally advanced or metastatic urothelial carcinoma. *N Engl J Med*. 2019;381(4):338–48. <https://doi.org/10.1056/NEJMoa1817323>.
50. Flaig TW. NCCN guidelines updates: management of muscle-invasive bladder cancer. *J Natl Compr Cancer Netw*. 2019;17(5.5):591–3. <https://doi.org/10.6004/jnccn.2019.5017>.
51. Creighton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, Hanna J, Lodato MA, Frampton GM, Sharp PA, Boyer LA, Young RA, Jaenisch R. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci U S A*. 2010;107(50):21931–6. <https://doi.org/10.1073/pnas.1016071107>.
52. Huang J, Li K, Cai W, Liu X, Zhang Y, Orkin SH, Xu J, Yuan GC. Dissecting super-enhancer hierarchy based on chromatin interactions. *Nat Commun*. 2018;9(1):943. <https://doi.org/10.1038/s41467-018-03279-9>.
53. Robertson AG, Kim J, Al-Ahmadie H, Bellmunt J, Guo G, Cherniack AD, Hinoue T, Laird PW, Hoadley KA, Akbani R, et al. Comprehensive molecular characterization of muscle-invasive bladder cancer. *Cell*. 2017;171:540–556.e525.
54. Szulwach KE, Jin P. Integrating DNA methylation dynamics into a framework for understanding epigenetic codes. *Bioessays*. 2014;36(1):107–17. <https://doi.org/10.1002/bies.201300090>.
55. Heitzman ND, Hon GC, Hawkins RD, Kheradpour P, Stark A, Harp LF, Ye Z, Lee LK, Stuart RK, Ching CW, Ching KA, Antosiewicz-Bourget JE, Liu H, Zhang X, Green RD, Lobanenkov VV, Stewart R, Thomsson JA, Crawford GE, Kellis M, Ren B. Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature*. 2009;459(7243):108–12. <https://doi.org/10.1038/nature07829>.
56. Ong CT, Corces VG. CTCF: an architectural protein bridging genome topology and function. *Nat Rev Genet*. 2014;15(4):234–46. <https://doi.org/10.1038/nrg3663>.
57. Song L, Zhang Z, Grasfeder LL, Boyle AP, Giresi PG, Lee BK, Sheffield NC, Graf S, Huss M, Keefe D, Liu Z, London D, McDaniell RM, Shibata Y, Showers KA, Simon JM, Vales T, Wang T, Winter D, Zhang Z, Clarke ND, Birney E, Iyer VR, Crawford GE, Lieb JD, Furey TS. Open chromatin defined by DNaseI and FAIRE identifies regulatory elements that shape cell-type identity. *Genome Res*. 2011;21(10):1757–67. <https://doi.org/10.1101/gr.121541.111>.

58. Bi M, Zhang Z, Jiang YZ, Xue P, Wang H, Lai Z, Fu X, De Angelis C, Gong Y, Gao Z, Ruan J, Jin VX, Marangoni E, Montaudon E, Glass CK, Li W, Huang TH, Shao ZM, Schiff R, Chen L, Liu Z. Enhancer reprogramming driven by highorder assemblies of transcription factors promotes phenotypic plasticity and breast cancer endocrine resistance. *Nat Cell Biol.* 2020;22(6):701–15. <https://doi.org/10.1038/s41556-020-0514-z>. Epub 2020 May 18. PMID: 32424275; PMCID: PMC7737911.
59. Lupien M, Eeckhoute J, Meyer CA, Wang Q, Zhang Y, Li W, Carroll JS, Liu XS, Brown M. FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription. *Cell.* 2008;132(6):958–70. <https://doi.org/10.1016/j.cell.2008.01.018>.
60. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods.* 2013;10(12):1213–8. <https://doi.org/10.1038/nmeth.2688>.
61. Tejaswi I, Baozhen Z, Wang Q, Ye H, Qiusi J, Xu J, Yang H, Tingting L, Wang X, Fan S, Yu L, Hironobu Y, Ruby C, Huijue L, Lijun Z, Lu W, Joshua W, Jay RD, Joshua MJ, David DGJ, Feng Y. Subtype-associated epigenomic landscape and 3D genome structure in bladder cancer. *Gene Expr Omnibus.* 2020; <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE148079>.
62. Tejaswi I, Baozhen Z, Wang Q, Ye H, Qiusi J, Xu J, Yang H, Tingting L, Wang X, Fan S, Yu L, Hironobu Y, Ruby C, Huijue L, Lijun Z, Lu W, Joshua W, Jay RD, Joshua MJ, David DGJ, Feng Y. Subtype-associated epigenomic landscape and 3D genome structure in bladder cancer. *Eur Genome Phenome Arch.* 2021; <https://ega-archive.org/studies/EGAS00001005071>.
63. Tejaswi I, Baozhen Z, Wang Q, Ye H, Qiusi J, Xu J, Yang H, Tingting L, Wang X, Fan S, Yu L, Hironobu Y, Ruby C, Huijue L, Lijun Z, Lu W, Joshua W, Jay RD, Joshua MJ, David DGJ, Feng Y. Subtype-associated epigenomic landscape and 3D genome structure in bladder cancer. *Github.* 2020; <https://github.com/Qixuan771/Source-code-for-bladder-cancer-project>.
64. Tejaswi I, Baozhen Z, Wang Q, Ye H, Qiusi J, Xu J, Yang H, Tingting L, Wang X, Fan S, Yu L, Hironobu Y, Ruby C, Huijue L, Lijun Z, Lu W, Joshua W, Jay RD, Joshua MJ, David DGJ, Feng Y. Subtype-associated epigenomic landscape and 3D genome structure in bladder cancer. *Zenodo.* 2020. <https://zenodo.org/record/4396080#X-oRtOLkY0/>. <https://doi.org/10.5281/zenodo.4396080>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

