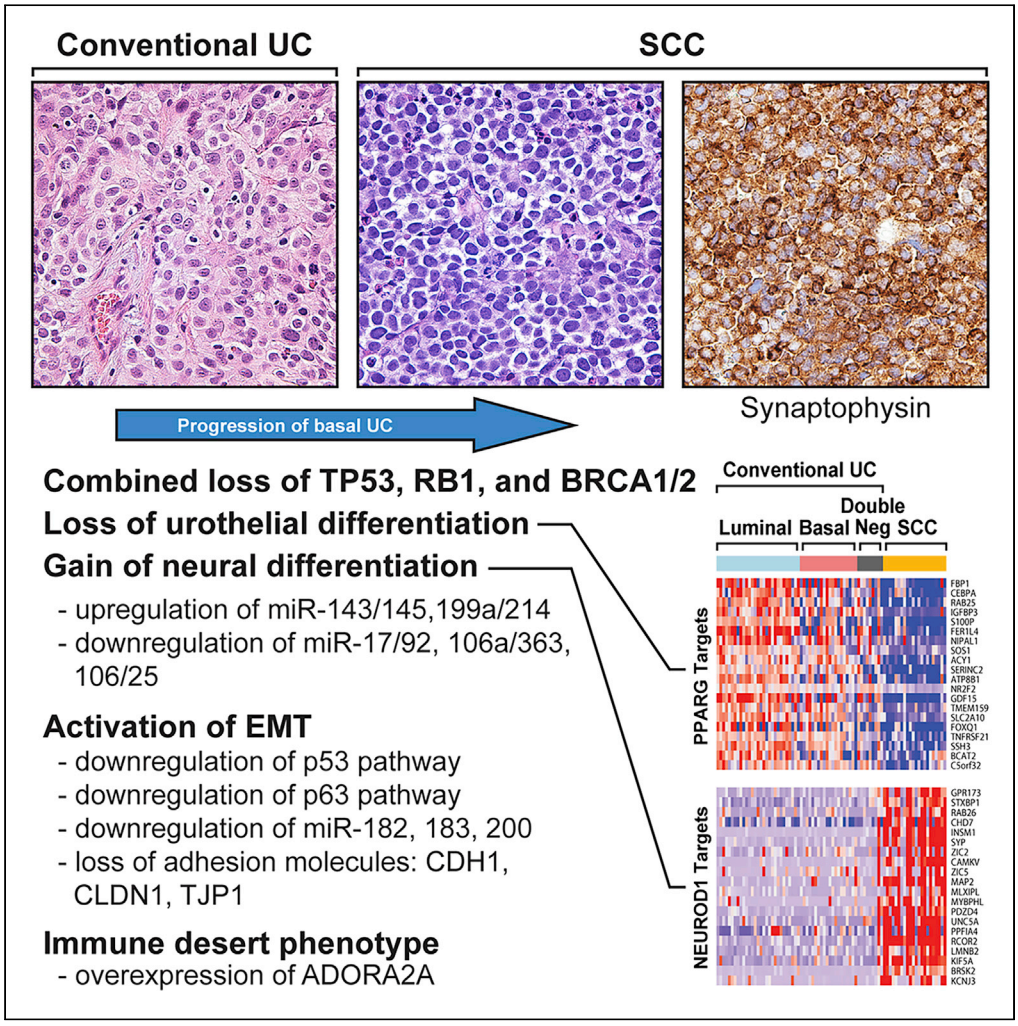
iScience

CellPress
OPEN ACCESS

Article

Urothelial-to-Neural Plasticity Drives Progression to Small Cell Bladder Cancer



Guoliang Yang, Jolanta Bondaruk, David Cogdell, ..., John N. Weinstein, David J. McConkey, Bogdan Czerniak

ccguo@mdanderson.org (C.C.G.)
bczernia@mdanderson.org (B.C.)

HIGHLIGHTS

SCCs show TP53/RB1 loss with mutational signature of BRCA1/2 loss of function

SCCs are driven by neural phenotypic switch with dysregulated EMT network

SCCs show depleted immune phenotype with upregulation of ADORA2A

Yang et al., iScience *23*, 101201
June 26, 2020
https://doi.org/10.1016/j.isci.2020.101201

iScience

CellPress
OPEN ACCESS

Article

Urothelial-to-Neural Plasticity Drives Progression to Small Cell Bladder Cancer

Guoliang Yang,[1,9] Jolanta Bondaruk,[1,9] David Cogdell,[1,9] Ziqiao Wang,[2,9] Sangkyou Lee,[1] June Goo Lee,[1] Shizhen Zhang,[1] Woonyoung Choi,[3] Yan Wang,[1] Yu Liang,[1] Gang Wang,[1] Ying Wang,[2] Hui Yao,[4] Vipulkumar Dadhania,[1] Jianjun Gao,[5] Christopher Logothetis,[5] Arlene Siefker-Radtke,[5] Ashish Kamat,[6] Colin Dinney,[6] Dan Theodorescu,[7] Marek Kimmel,[8] Peng Wei,[2] Charles C. Guo,[1,*] John N. Weinstein,[4] David J. McConkey,[3] and Bogdan Czerniak[1,10,*]

SUMMARY

We report a comprehensive molecular analysis of 34 cases of small cell carcinoma (SCC) and 84 cases of conventional urothelial carcinoma (UC), with The Cancer Genome Atlas cohort of 408 conventional UC bladder cancers used as the reference. SCCs showed mutational landscapes characterized by nearly uniform inactivation of *TP53* and were dominated by Sanger mutation signature 3 associated with loss of *BRCA1/2* function. SCCs were characterized by downregulation of luminal and basal markers and were referred to as double-negative. Transcriptome analyses indicated that SCCs displayed lineage plasticity driven by a urothelial-to-neural phenotypic switch with a dysregulated epithelial-to-mesenchymal transition network. SCCs were depleted of immune cells, and expressed high levels of the immune checkpoint receptor, adenosine receptor A2A (ADORA2A), which is a potent inhibitor of immune infiltration. Our observations have important implications for the prognostication and development of more effective therapies for this lethal bladder cancer variant.

INTRODUCTION

The ability of cells to acquire alternative differentiation-associated phenotypes, referred to as lineage plasticity, is a common feature of embryogenesis and is also intrinsic to malignant transformation. Mild to moderate deviation from the tissue-of-origin terminal differentiation profile as a consequence of either stem cell transformation or dedifferentiation is a general feature of tumor initiation, but major changes in lineage identity can occur in tumor progression (Yuan et al., 2019). Such changes are associated with dramatic, microscopically evident phenotypic switches coupled with the acquisition of clinically aggressive behaviors. Here we show that the progression of urothelial bladder cancer to the clinically aggressive small cell carcinoma (SCC) variant is driven by lineage plasticity signified by the loss of urothelial differentiation and acquisition of neural phenotype and epithelial-to-mesenchymal transition (EMT).

Bladder cancer develops along two distinct tracks, referred to as papillary and non-papillary, that represent clinically and molecularly different forms of the disease (Czerniak et al., 2016). The vast majority of papillary tumors are of luminal molecular subtype characterized by gene expression patterns that are similar to normal intermediate and terminal urothelial differentiation (Choi et al., 2014b). Papillary tumors frequently recur but rarely progress to high-grade invasive carcinomas. On the other hand, non-papillary carcinomas are clinically aggressive, exhibiting a high propensity for invasive growth. A large proportion of them are lethal because of metastatic spread (Kamat et al., 2016). Many of them are of a basal molecular subtype and express genes characteristic of the normal basal urothelial layer and EMT. In addition to conventional urothelial carcinomas (UCs), many microscopically distinct bladder cancer variants have been described and in general are thought to develop via progression of conventional disease (Amin, 2009). The most frequent of these variants are sarcomatoid, small cell, micropapillary, and plasmacytoid. These variants are clinically more aggressive than conventional UCs and require uniquely tailored therapeutic management, which is often unavailable (Amin, 2009; Kamat et al., 2016).

In this report we focus on SCC, which comprises less than 1% of all bladder cancer and often coexists with conventional UC (Amin, 2009). SCC is a highly aggressive disease characterized by early progression to

[1]Department of Pathology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

[2]Department of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

[3]Johns Hopkins Greenberg Bladder Cancer Institute, Johns Hopkins University, Baltimore, MD, USA

[4]Department of Bioinformatics and Computational Biology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

[5]Department of Genitourinary Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

[6]Department of Urology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

[7]Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai, Los Angeles, CA, USA

[8]Department of Statistics, Rice University, Houston, TX, USA

[9]These authors contributed equally

[10]Lead Contact

*Correspondence: ccguo@mdanderson.org (C.C.G.), bczernia@mdanderson.org (B.C.)

https://doi.org/10.1016/j.isci.2020.101201

metastasis and shorter survival compared with conventional UC (Veskimae et al., 2019). We report on the genome-wide characterization of bladder SCC, including microRNA (miRNA), gene expression, and whole-exome mutational profiles of 34 paraffin-embedded SCC and 84 invasive conventional bladder UC samples from an MD Anderson Cancer Center (MDACC) cohort. A cohort of 408 muscle-invasive bladder cancers in The Cancer Genome Atlas (TCGA) was used as a reference (Table S1). These analyses revealed molecular features associated with its aggressive nature that may be relevant for the early detection and treatment of this highly lethal form of bladder cancer.

RESULTS

Mutational Signatures

The mutational profiles of conventional UC were characterized by the presence of statistically significant recurrent somatic mutations in 30 genes (Figure 1A). The 10 most frequently mutated genes were *TP53* (47%), *ARID1A* (25%), *KDM6A* (22%), *PIK3CA* (22%), *RB1* (17%), *EP300* (15%), *FGFR3* (14%), *STAG2* (14%), *ELF3* (12%), and *CREBBP* (11%). The overall mutational landscapes of luminal, basal, and double-negative UC were similar, but several mutated genes were enriched in each molecular subtype. Mutations in *FGFR3*, *ELF3*, *CDKN1A*, and *TSC1* were enriched in luminal tumors, whereas mutations in *TP53*, *RB1*, and *PIK3CA* were enriched in basal tumors (Figure 1A). The mutational landscape of the double-negative subtype was, in general, similar to the basal tumors, which show increased p53 and *RB1* mutations with a low frequency of *FGFR3* mutations. In fact, only one case in the double-negative subtype had *FGFR3* mutation. The most striking difference of double-negative tumors was enrichment for *RB1* mutations (47%) and presence of combined mutational inactivation of p53 and *RB1* (35%). SCC exhibited high overall mutational rates (median mutational frequency = 259, interquartile range = 174), and their significantly mutated genes were similar to those observed in conventional UC (Figure 1B; Table S2). However, the top two most frequently mutated genes in SCCs (*TP53* [93%] and *RB1* [47%]) were mutated at significantly higher frequencies in SCC than they were in conventional UCs of the TCGA cohort ($p < 0.01$). This suggests that SCC evolved from precursor conventional UC carrying these mutations, which may drive the progression process. Consistent with this hypothesis, in the paired cases containing both SCC and conventional UC, nearly all mutations in the conventional UCs were also present in SCCs, indicating that they were clonally related, e.g., the identical mutations of p53 (p.H179Y; p.C176F) and RB1 (p.R798fs) were present in both conventional and small cell components of the same tumor further confirming their clonal evolution (Table S2). Several of the chromatin-remodeling genes that are frequently mutated in conventional UC, including *KDM6A*, *EP300*, *ARID1A*, and *CREBBP*, were not mutated in SCC (Gui et al., 2011). Instead, SCCs carried mutations in *FSCN3* (13%), *BRD4* (13%), *ISLR2* (13%), *MAG* (13%), *MAMDC2* (13%), and *TAF1D* (13%), which are involved in cellular extension, chromatin regulation, cell cycle, and signaling (Abudureyimu et al., 2018; Edwards and Bryan, 1995; Filippakopoulos et al., 2012; Pijnappel et al., 2009; Wierer et al., 2018). The functional significance of mutations in these genes for small cell progression remains unclear, but they are attractive candidates for future mechanistic studies. Interestingly, *FGFR3* mutations, which were present in 14% of conventional UCs, were not present in SCCs.

Mechanisms of Mutagenesis

To further characterize the mutational process associated with progression from conventional UC to SCC, we examined six single-based substitutions (C>A, C>G, C>T, T>A, T>C, and T>G) in all cancer samples (Alexandrov et al., 2013). The results revealed that SCCs were enriched with C>T mutations compared with conventional UCs (Figures 1C–1E). Analyses of Sanger mutational signatures (Faltas et al., 2016) revealed six dominant signatures in the conventional UCs in the TCGA cohorts: signatures 1, 2, 3 (BRCA1/2 mutagenesis), 13 (APOBEC), 19, and 30 (Figures 1F and 1G). Clustering separated the conventional tumors into two subsets (α and β) that were characterized by different levels of signature 13 (APOBEC) prevalence. In contrast, SCCs and the paired precursor conventional UCs were characterized by the uniform dominance of signature 3 (Figure 1H). In addition, clustering segregated SCCs into two subsets that were also characterized by different levels of APOBEC activity. Finally, mutagenesis signatures 16, 21, and 22 were significantly enriched in SCCs compared with conventional UCs (Figure 1I). Overall, these data suggest that SCCs evolve from a distinct subset of conventional UCs.

RNA Expression

Whole-transcriptome expression profiling and unsupervised hierarchical clustering in a combined cohort of SCCs and UCs separated the tumors into two clusters, one dominated by SCCs and the other by

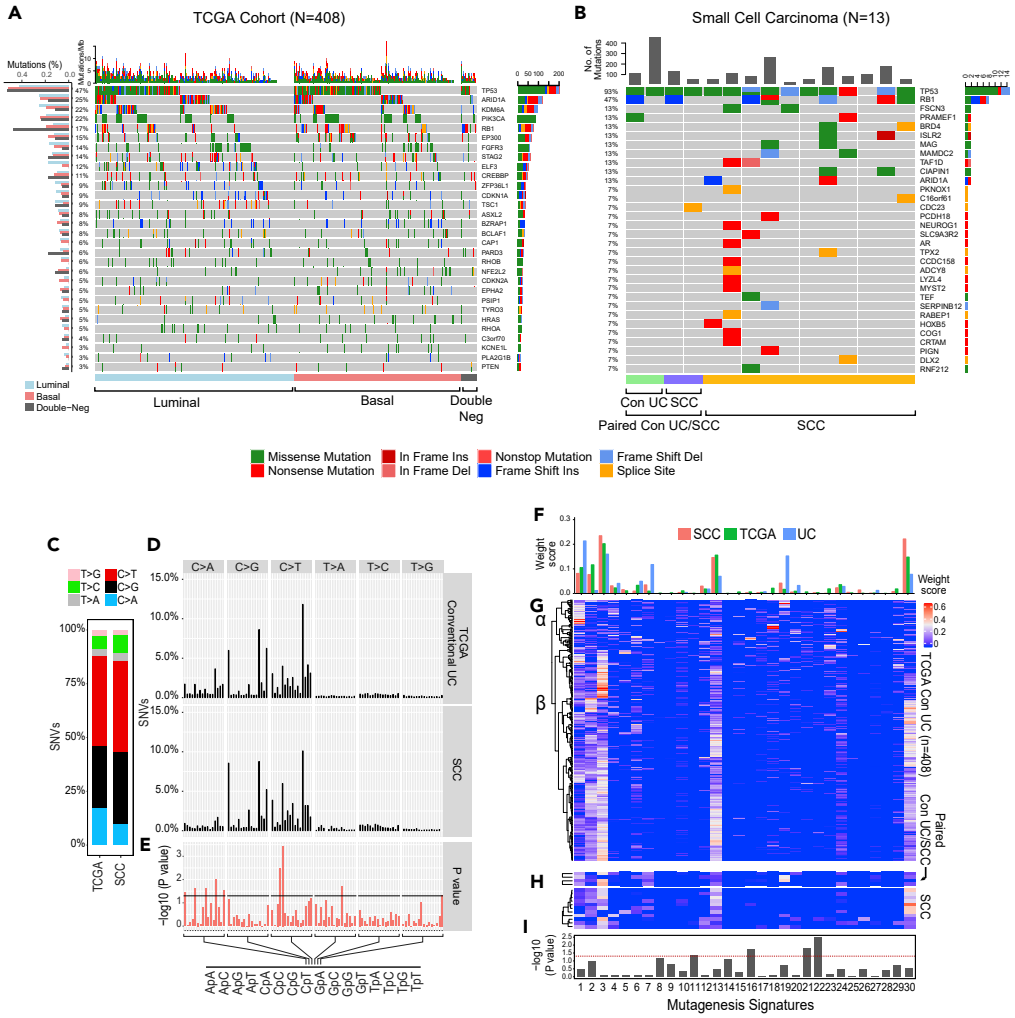


Figure 1. Mutational Landscape of SCC

(A) Mutational landscapes among the molecular subtypes of 408 muscle-invasive bladder cancers from the TCGA cohort showing the frequency of mutations in individual tumors and somatic mutations for significantly mutated genes. The frequencies of mutations of individual genes in the luminal, basal, and double-negative subtypes are shown on the left. Bars on the right show the numbers of specific substitutions for individual genes.

(B) Mutational landscapes of 13 cases of SCC and 2 paired samples of precursor conventional UC showing the frequency of mutations in individual genes and somatic mutations for significantly mutated genes. The frequencies of mutations of individual genes are shown on the left. Bars on the right show the numbers of specific substitutions for individual genes.

(C) Composite bar graphs showing the distributions of all nucleotide substitutions in two sets of samples corresponding to the TCGA cohort and SCC.

(D) Proportion of single-nucleotide variants (SNVs) in specific nucleotide motifs for each category of substitution in two sets of samples as shown in (C).

(E) False discovery rate (FDR) for specific nucleotide motifs in two sets of samples as shown in (C).

(F) Average weight scores of mutagenesis patterns in three sets of samples corresponding to the TCGA cohort, paired precursors conventional UC, and SCC.

(G) Weight scores of mutagenesis patterns in individual tumor samples of the TCGA cohort.

(H) Weight scores of mutagenesis patterns in SCCs and paired precursors conventional UCs.

(I) Statistical significance of mutagenesis patterns (p value) in SCC compared with conventional UC. For (E) and (I), p value was calculated using Wilcoxon rank-sum and Kruskal-Wallis tests, respectively.

conventional UCs (Figure 2A). We recognized that 28% top upregulated and downregulated genes in SCCs were involved in neural differentiation. Similarly, four of the top 10 upregulated genes in the SCC cluster controlled neural development (Figure 2B).

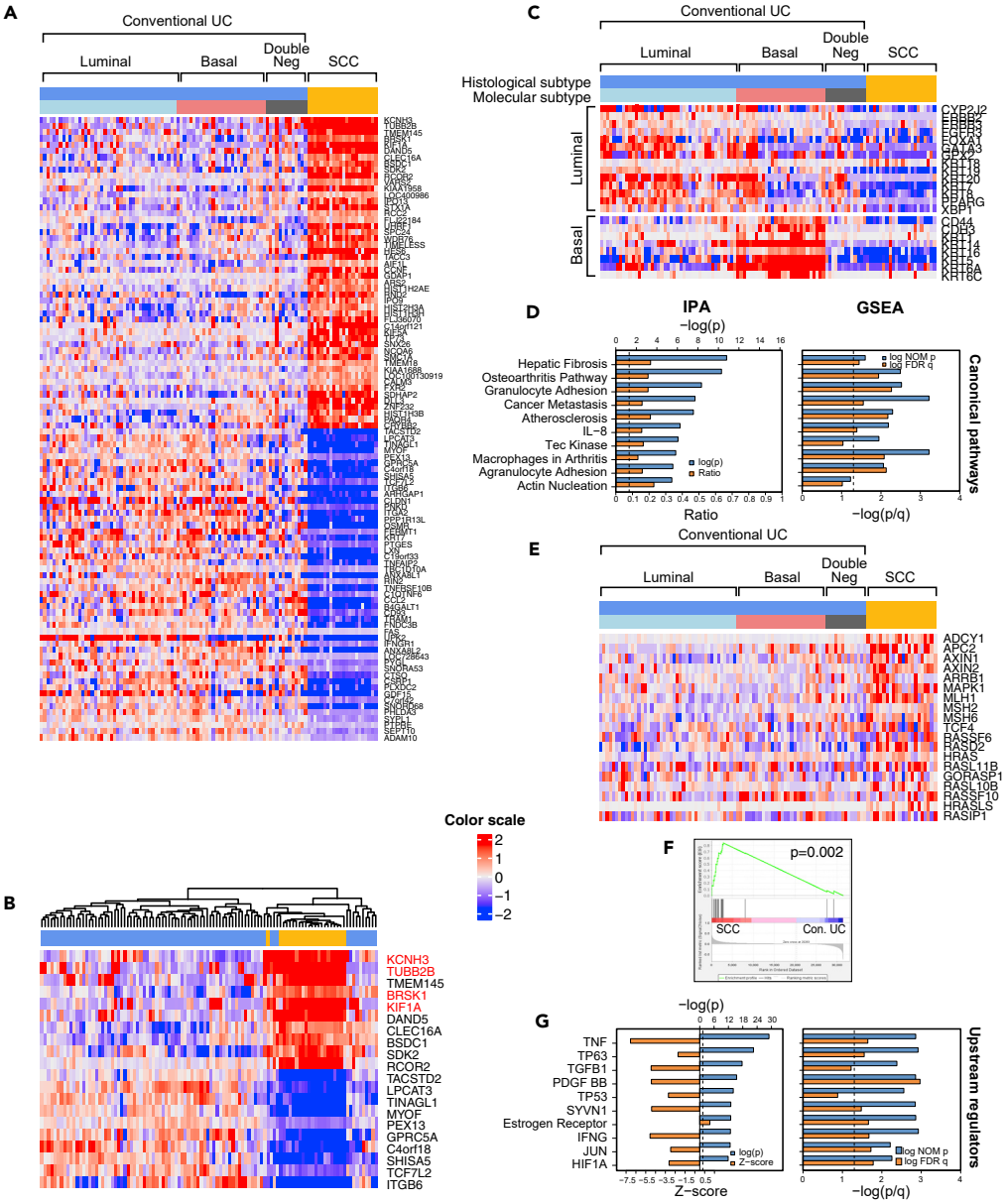


Figure 2. Whole-Genome mRNA Expression Profile of SCC and Conventional UC

(A) The top 50 upregulated and the top 50 downregulated genes in SCC (n = 22) compared with conventional UC (n = 84).

(B) Hierarchical cluster analysis of the cohort shown in (A) using the top 10 upregulated and top 10 downregulated genes identified in SCC.

(C) Intrinsic molecular subtypes identified by the expression of luminal and basal markers in conventional UC (n = 84) and SCC (n = 22).

(D) The top 10 canonical pathways dysregulated in SCC, as revealed by Ingenuity Pathway Analysis (IPA) and Gene Set Enrichment Analysis (GSEA).

(E) Expression patterns of cancer metastasis pathway genes in molecular subtypes of conventional UC and SCC.

(F) GSEA of metastasis-related genes in SCC compared with conventional UC. A p value <0.05 was considered statistically significant. A p value < 0.05 was considered statistically significant.

(G) The top 10 upstream regulators altered in SCC, as revealed by IPA and GSEA. Genes highlighted in red are involved in neural differentiation.

Conventional UCs segregated into 3 subtypes (basal, luminal, and double-negative), whereas the SCCs were uniformly double-negative for basal and luminal genes (Figure 2C). Ingenuity Pathway Analysis revealed that SCCs were highly enriched for genes that promote cancer metastasis, and more than half of the top activated canonical pathways were known to suppress adaptive (i.e., T-cell-mediated) immunity (Figures 2D–2F). Finally, several of the top downregulated pathways promoted inflammation (TNF, IFNG, JUN), stromal infiltration (TGFB1, PDGF BB), and basal bladder cancer biology (TP63, HIF1) (Figures 2D–2G, S1A, and S1B). Consistent with the presence of inactivating p53 mutations in virtually all SCCs, they showed widespread inactivation of the p53 target genes (Figure S1A). Kaplan-Meier analysis has shown that these phenotypic changes were associated with significantly shorter survival of patients with SCC when compared with conventional UC (Figure S2).

More in-depth analyses revealed that these effects were associated with increased expression of transcription factors that were also known to play important roles in neural crest development, pro-neural stem cell differentiation, and cell proliferation, and the Musashi RNA-binding protein family (MSI1 and MSI2) that functions as regulator of neural stem cell differentiation (Siddall et al., 2006) (Figures S3A and S3B). The maintenance of an embryonic/stem cell-like neural phenotype was further reinforced by the upregulation of two adhesion proteins, PROM1 and NCAM1 (Figures S3A and S3B) (Pruszak, 2015).

Comparative analyses of the miRNA expression profiles of conventional UC and SCC reinforced the conclusion that SCC progression involves urothelial-to-neural plasticity. Differential expression analysis revealed that over 200 miRNAs were selectively up- or downregulated in SCCs (Figure 3A), and in unsupervised clustering a subset of the top 20 up- and downregulated miRNAs was sufficient to distinguish SCCs from conventional UCs (Figure 3B). The upregulated miRNAs included the polycistronic miR-143/145 and miR-199a/214 clusters, which control neural crest development and pluripotent neural stem cell maintenance (Celia-Terrassa et al., 2017; Cordes et al., 2009). The downregulated miRNAs included three closely related polycistronic clusters: miR-17/92, miR-106a/363, and miR-106b/25, comprising over 10 individual miRNA species all involved in neural differentiation, including the regulation of neural stem cells (Rajman and Schratt, 2017) (Figure 3C). Interestingly, a subset of luminal miRNAs was downregulated in SCCs, whereas basal miRNAs were upregulated (Figure 3D) (Ochoa et al., 2016).

Analysis of Regulons and EMT

The analysis of both mRNA and miRNA expression profiles suggested that progression to SCC was driven by lineage plasticity characterized by the loss of urothelial differentiation and activation of neural transcriptional programs. To explore this relationship further, we performed additional analyses of regulons implicated in urothelial and neural differentiation (Tables S3 and S4). Most of 23 urothelial regulons displayed reduced enrichment scores in progression to SCC (Figures 4A and 4B). On the other hand, all 15 neural regulons exhibited activation with increases in their enrichment scores (Figures 4C and 4D).

As the development of the central nervous system is a canonical EMT program that is activated during embryogenesis, we hypothesized that dysregulation of EMT might represent an underlying mechanism driving the progression to SCC. In embryogenesis, an EMT-permissive state is activated by the developmental program. In solid tumors, at the core of this circuitry are *TP53* and *RB1*, which negatively regulate EMT and their target genes, and they were coordinately downregulated in virtually all SCCs (Figure 5A). As the activation of EMT is driven by complex regulatory programs involving multiple pathways and their respective genes, we assessed its activation by calculating a quantitative EMT score based on a 76-gene signature developed by Byers et al. (Byers et al., 2013) (Figure 5B). We previously showed that *TP63* controls the expression of basal high-molecular-weight keratins (KRT5, KRT6, and KRT14) and the EMT inhibitor, miR-205, in urothelial cells (Choi et al., 2014b; Tran et al., 2013). The central role of TP63 in the regulation of EMT was confirmed in several solid tumors (Stacy et al., 2017). Consistent with these observations SCCs showed downregulation of *TP53* and *TP63* target genes (Figure 5C). Correspondingly, E-cadherin (CDH1) and other homotypic adhesion molecules, including claudin-1 (CLDN1) and tight junction protein 1 (TJP1), were all downregulated in SCCs (Figure 5D). Surprisingly, none of the major transcriptional EMT regulators of the SNAIL, TWIST, and ZEB families was overexpressed in SCC (Figure 5E). However, several families of miRNAs typically involved in the activation of EMT such as miR-182, miR-183, and members of miR-200 family were downregulated in SCC (Figure 5A). The luminal and basal subtypes of conventional UC were characterized by positive EMT scores reflecting their epithelial phenotype (Figure 5F). The tumors in the double-negative category had intermediate scores indicating their partial EMT, whereas SCCs had the lowest

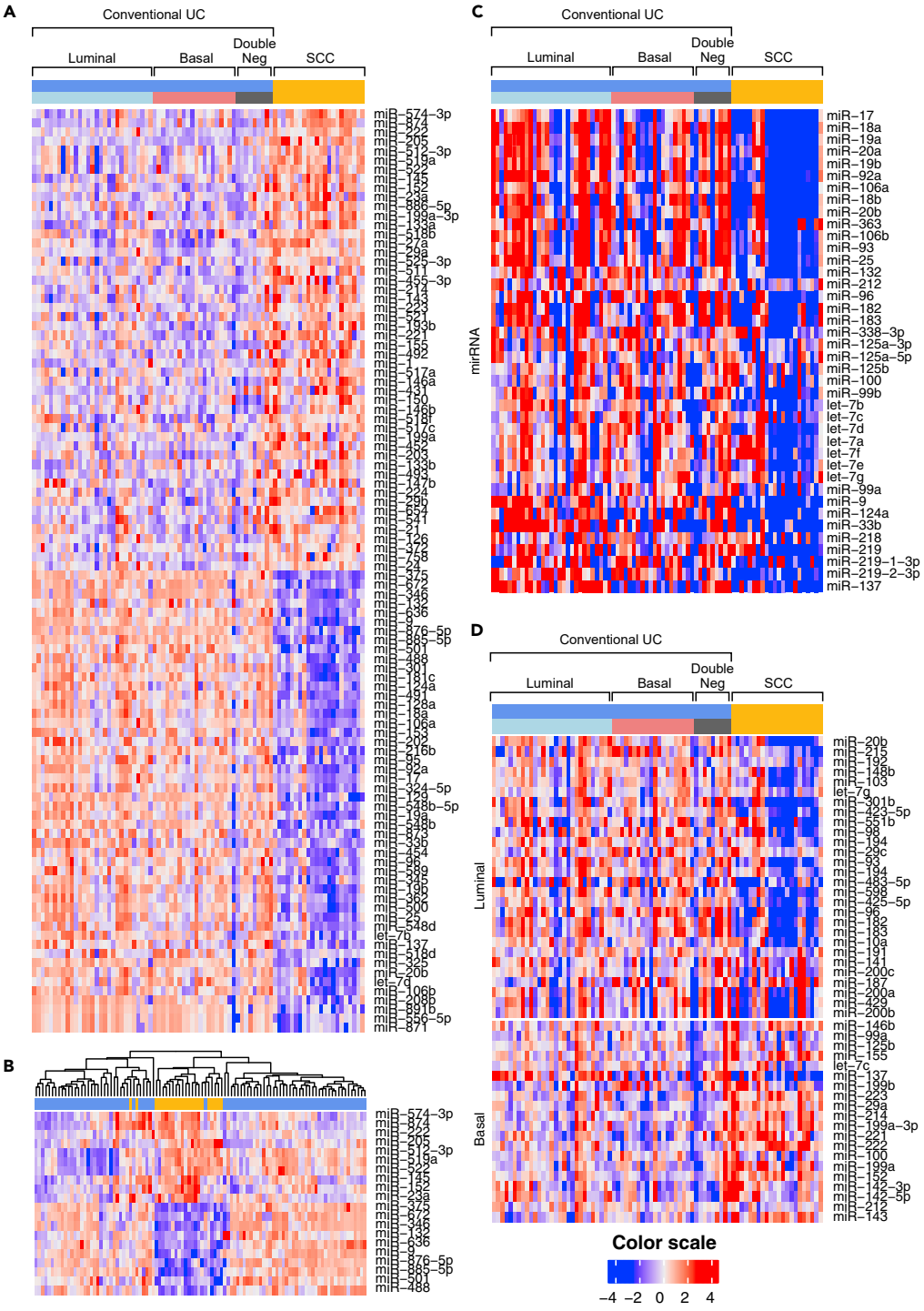


Figure 3. miRNA Expression Profile in SCC and Conventional UC

(A) Top 50 upregulated and downregulated miRNAs in SCC compared with molecular subtypes of conventional UC.

(B) Hierarchical clustering analysis of the cohort shown in (A) using the top 10 up- and downregulated miRNA.

(C) Expression patterns of miRNAs involved in neural development in SCC and molecular subtypes of conventional UC.

(D) Expression patterns of luminal and basal miRNAs in SCC compared with molecular subtypes of conventional UC.

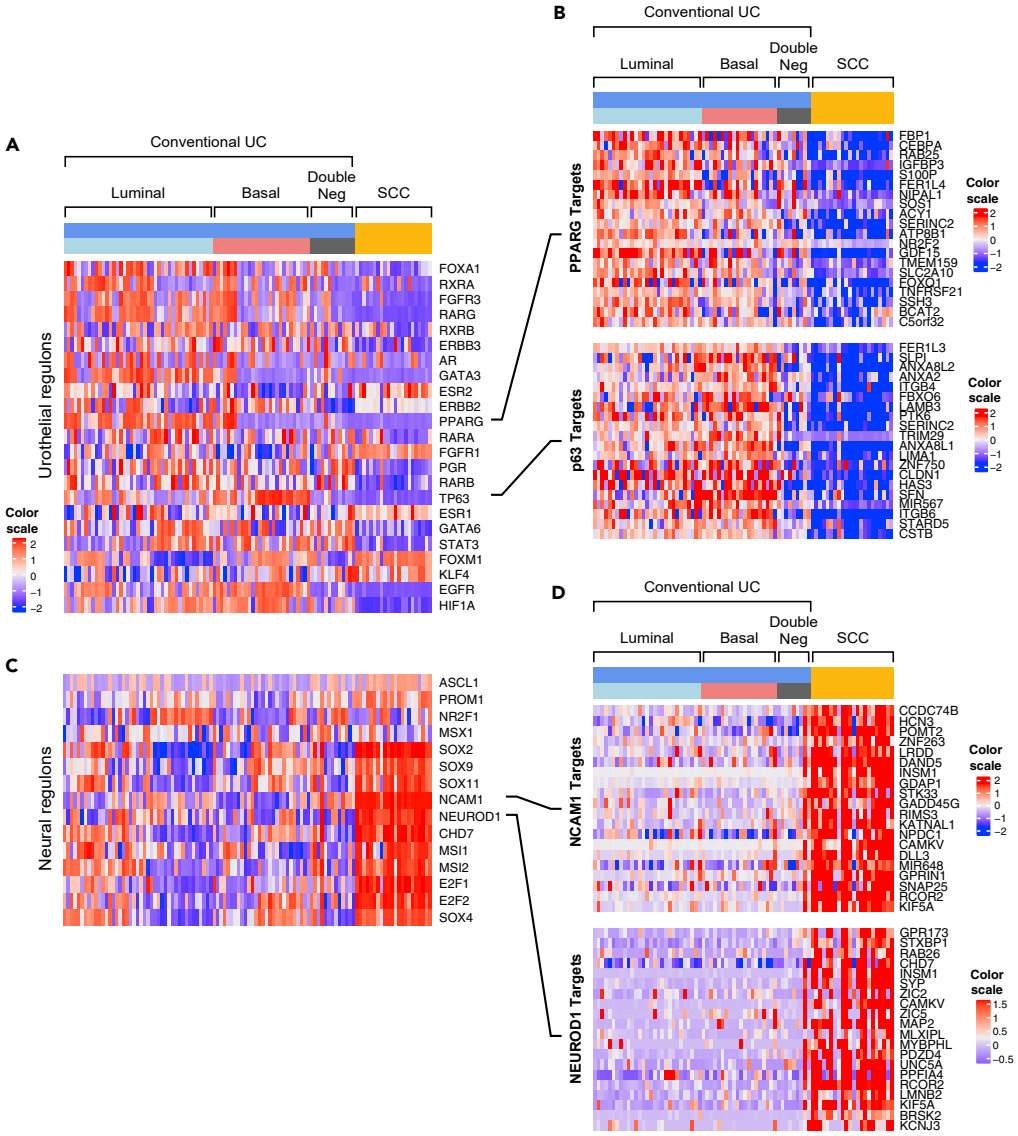


Figure 4. Enrichment Scores of Urothelial and Neural Regulons in Molecular Subtypes of Conventional UC (n = 84) and SCC (n = 22)

(A) Enrichment scores of urothelial regulons.

(B) The expression pattern of selected target genes of PPARG and TP63 regulons.

(C) Enrichment scores of neural regulons.

(D) The expression pattern of selected target genes of NCAM1 and NEUROD1 regulons.

scores confirming their fully activated EMT state. Similar results were obtained by Gene Set Enrichment Analysis using the 175 EMT gene signature developed by Yu et al. (Yu et al., 2013) (Figure 5G). These data confirm that the activation of EMT combined with urothelial-to-neural phenotypic switch is the driving mechanism of progression to SCC.

Immune Infiltrate

Immune checkpoint blockade is clinically active in approximately 20% patients with bladder cancer. The response is positively associated with high mutational burden, infiltration with activated cytotoxic lymphocytes, and specific molecular subtypes (van Dijk et al., 2019). Given the contradictory presence of high mutational rates and immunosuppressive gene expression signatures observed in SCC, we characterized their expression of immune-related genes. In general, conventional carcinomas in the basal subtype were characterized by increased immune

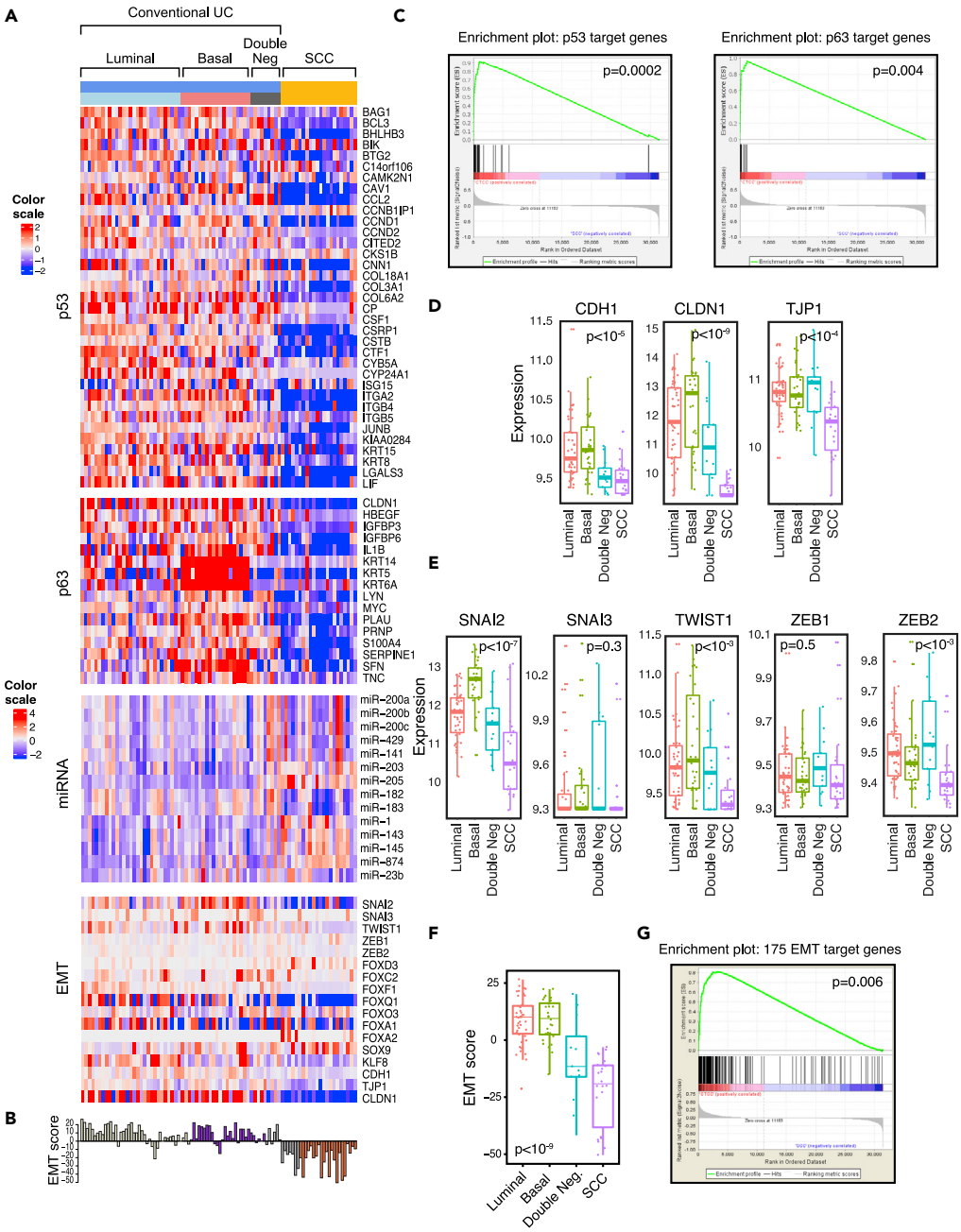


Figure 5. Dysregulation of EMT Network in Molecular Subtypes of Conventional UC and SCC

(A) Expression patterns of representative genes of the EMT regulatory network in molecular subtypes of conventional UC and SCC.

(B) EMT scores in molecular subtypes of conventional UC and SCC.

(C) Gene Set Enrichment Analysis (GSEA) analysis of TP53 (left) and TP63 (right) target genes in SCC compared with conventional UC.

(D) Boxplot analysis of expression levels of CDH1, TJP1, and CLDN1 in molecular subtypes of conventional UC and SCC.

(E) Boxplot analysis of expression levels of SNAI2, SNAI3, TWIST1, ZEB1, and ZEB2 in molecular subtypes of conventional UC and SCC.

(F) Boxplot analysis of EMT scores in molecular subtypes of conventional UC and SCC.

(G) GSEA analysis of 175 EMT target genes in SCC compared with conventional UC. In panels C, D, E, F, and G, a p value < 0.05 was considered statistically significant.

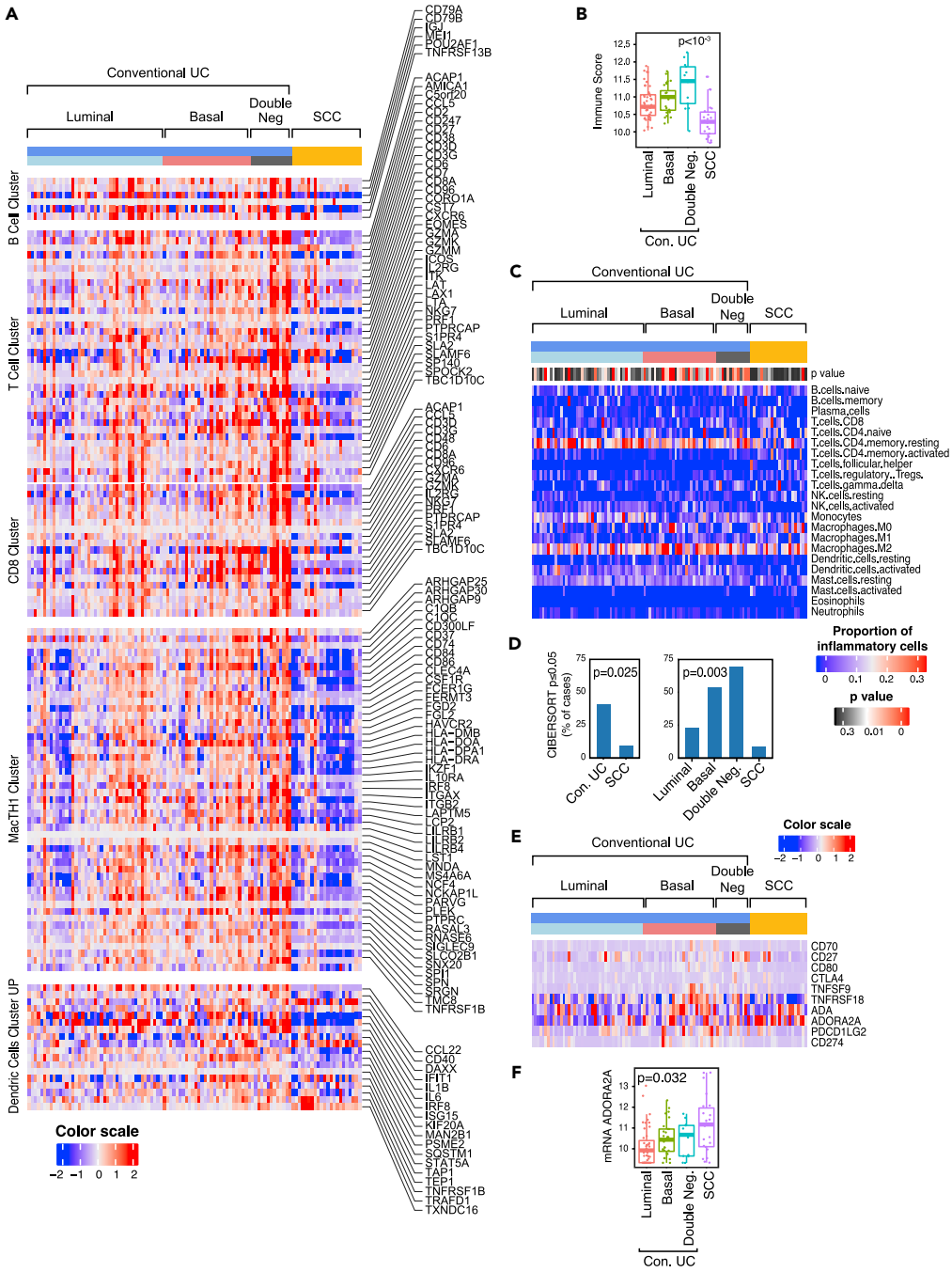


Figure 6. Expression Pattern of Immune Cell Infiltrate and Immune Checkpoint Genes in Molecular Subtypes of Conventional UC and SCC

(A) Expression pattern of immune cell infiltrate. Top to bottom: B cell, T cell, CD8, MacTH1, and dendritic cell expression clusters.

(B) Boxplot of immune scores calculated using the gene expression profile shown in (A) for molecular subtypes of conventional UC and SCC.

(C) Heatmap of CIBERSORT scores for 22 immune cell types in molecular subtypes of conventional UC and SCC. A p value < 0.05 was considered statistically significant.

(D) Proportion of cases with significant CIBERSORT scores in conventional UC and SCC (left). Proportion of cases with significant CIBERSORT scores in molecular subtypes of conventional UC and SCC (right).

Figure 6. *Continued*
(E) Expression of immune checkpoint genes in conventional UC and SCC in relation to their molecular subtypes.
(F) Boxplot of ADORA2A mRNA expression levels in molecular subtypes of conventional UC and SCC. In panels B, D, and F, a p value < 0.05 was considered statistically significant.

gene expression signatures (Figure 6A). In contrast, SCCs were characterized by decrease of immune gene expression. The null immune signature of SCCs was confirmed by significantly reduced immune scores (Figure 6B). More in-depth analysis of immune infiltrate status was performed using the CIBERSORT algorithm, which provided quantitative assessment of expression signatures related to 22 immune cell types (Newman et al., 2015). CIBERSORT confirmed that SCCs were characterized by an immune-null signature when compared with conventional UCs (Figures 6C and 6D). The analysis of the expression signature of immune checkpoint ligands and their receptors (Sharma and Allison, 2015) included CD70 and CD27, CD80 and CTLA4, TNFSF9 and TNFRSF18, ADA and ADORA2A, and PDCD1G2 and CD274 (Figure 6E). SCCs were also characterized by downregulation of nearly all immune checkpoint ligands and their receptors with the exception of ADORA2A, a major suppressor of immune infiltration, which was overexpressed (Figure 6F). This unique feature of SCC potentially explains their null immune phenotype, and ADORA2A may therefore represent an unexplored therapeutic target.

Validation Studies

We verified the expression patterns of signature luminal markers (GATA3), basal markers (KRT5/6, KRT14), and neural markers (Synaptophysin, Chromogranin, and INSM1) by immunohistochemistry (IHC). The SCCs were all negative for GATA3, KRT5/6, and KRT14. In contrast, SCCs were IHC-positive for neural marker such as Synaptophysin, Chromogranin, and INSM1, consistent with the RNA expression data (Figure 7A). IHC stains for CD3 and CD8 T cells confirmed the immune-null phenotype of SCCs, which showed practically no T cell infiltrate (Figure 7B). In contrast, many double-negative conventional UCs showed pronounced infiltrate of CD3 and CD8 T cells. On the other hand, only a sparse T cell infiltrate was present in luminal and basal conventional UC subtypes. Overexpression of synpatophysin, MSI1, E2F1, and ADORA2A with downregulation of E-cadherin in SCCs was confirmed on several fresh frozen tissues by western blotting (Figure 7C).

As the miR17 family is considered to be a regulator of neural differentiation and was consistently downregulated in SCC, we hypothesized that this downregulation played a role in neural switch (Mogilyansky and Rigoutsos, 2013). To verify this hypothesis, we first analyzed the mRNA expression signatures of 30 UC cell lines and found that, similar to patient samples, they could be separated into luminal and basal subtypes. However, a large proportion (33%) of these cell lines was negative for the basal and luminal markers and was considered to be double-negative. Consistent with our observation on patient samples, double-negative UC cell lines showed upregulation of neural markers. As SCCs appeared to be related to basal subtype, we ablated miR17-5p by locked nucleic acid in the basal UC cell line UC6 and confirmed it by RT-qPCR (Figures 8A and S4). The ablation of miR17-5p caused dysregulation of multiple genes enriched for neural markers (Figures 8B and 8C). This included activation of several regulons (TP63, FOXM1, EGFR, PROM1, SOX2, and E2F1) implicated in neural differentiation (Figures 8D–8F). In summary, ablation of miR17 validated the functional role of the miRNA in the induction of neural phenotype.

DISCUSSION

Bladder cancer is the ninth most common cancer worldwide, affecting 430,000 people and resulting in 165,000 deaths per year (Antoni et al., 2017). In the United States it is the fourth most common cancer in men with an estimated incidence of approximately 80,000 cases in 2019 (Siegel et al., 2019). More than 90% bladder cancers are UCs that originate from precursor lesions in the epithelial urothelial layer of the bladder. Progression of conventional UC to the so-called variant UCs such as SCC, sarcomatoid, micropapillary, or plasmacytoid is associated with increased clinical aggressiveness (Amin, 2009). In the following we summarize the key findings concerning molecular characterization of SCCs and suggest how the findings contribute to our understanding of its clinically aggressive behavior as well as how to open more effective therapeutic opportunities.

Similar to conventional UC, SCCs have high mutational rates comparable to those observed in melanomas and non-small cell lung cancers (Hoadley et al., 2018). In SCC, these high mutational rates are associated with mutation signature 3, characteristic of loss of BRCA1/2 function. SCCs can be separated into two subgroups, i.e., with and without mutation signature of the endogenous mutagenic

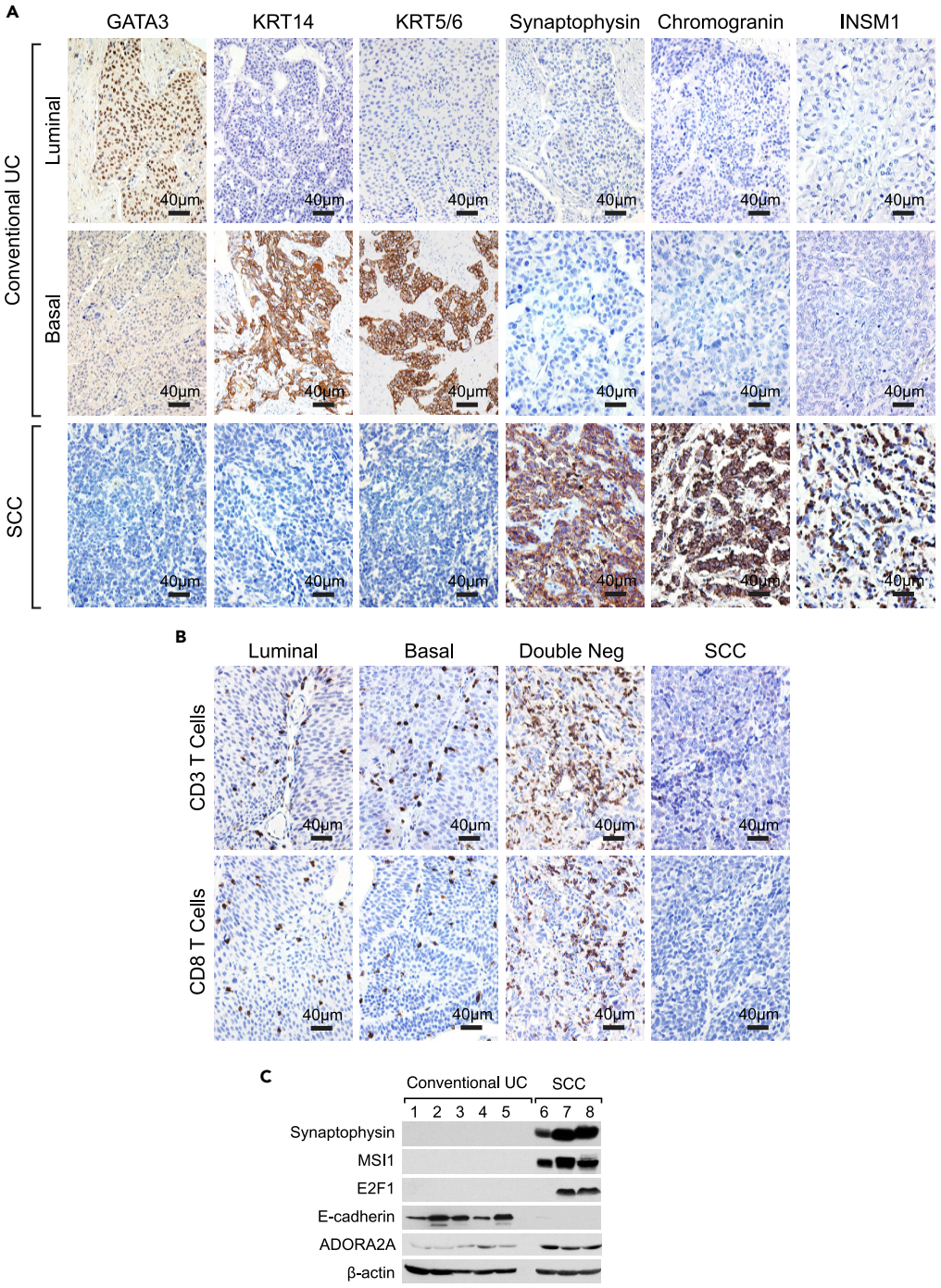


Figure 7. Immunohistochemical Validation Studies of Molecular Features of SCC when Compared with Conventional UC

(A) The immunohistochemical expression signatures of luminal, basal, and neuroendocrine markers in representative luminal and basal cases of conventional UC and SCC on a tissue microarray (TMA) composed of conventional UC (n = 76) and SCC (n = 14). Scale bars, 40 μm.

(B) The immunohistochemical expression of CD3 and CD8 T cells in representative luminal, basal, and double-negative cases of conventional UC and SCC on a TMA composed of conventional UC (n = 76) and SCC (n = 14). Scale bars, 40 μm.

(C) Western blot documenting the overexpression of synaptophysin, MSI1, E2F1, and ADORA2A with loss of E-cadherin in fresh frozen tissue of conventional UC and SCC. All cases of conventional UC are of the basal subtype.

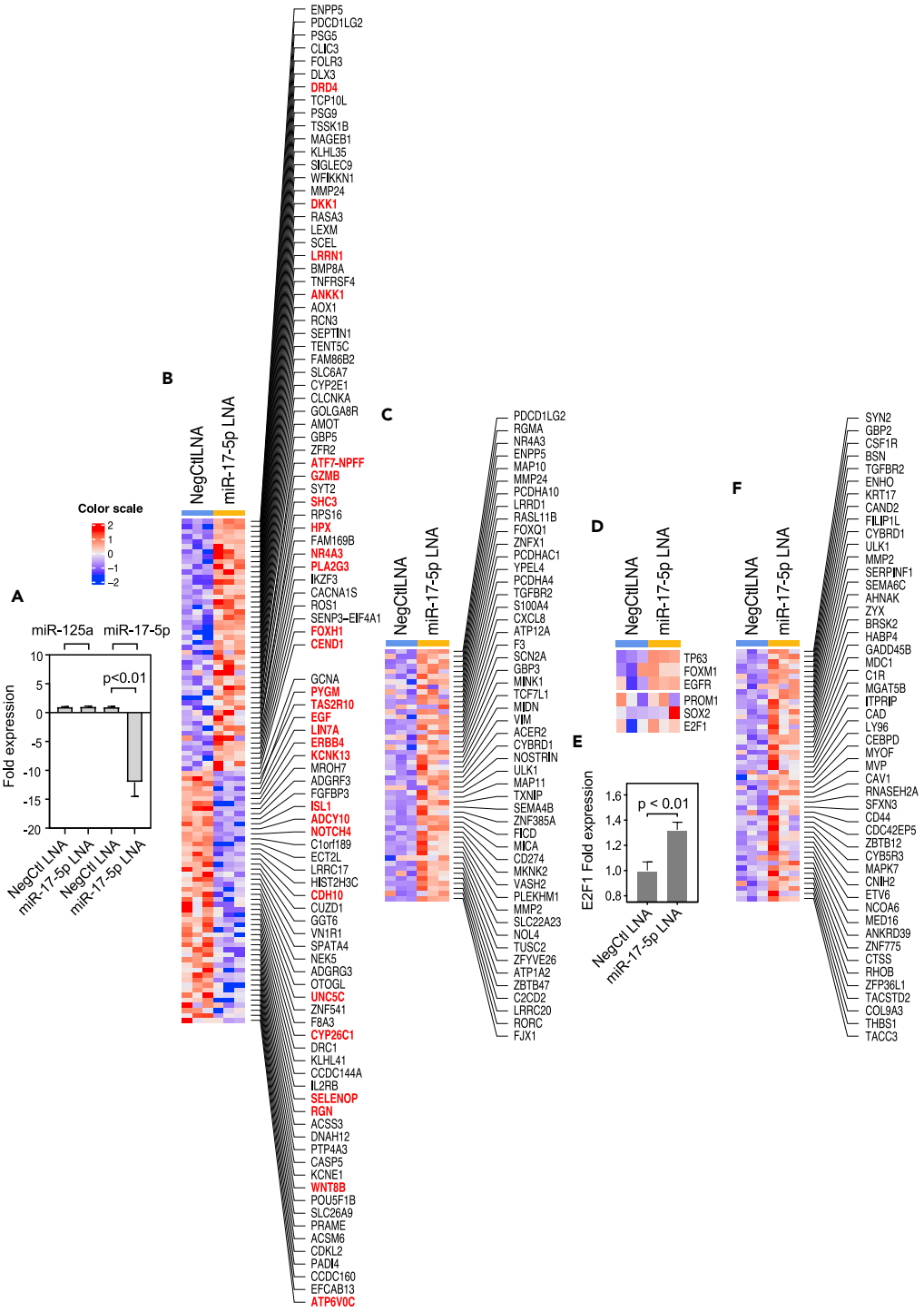


Figure 8. Induction of Neural-like Phenotype by Locked Nucleic Acid (LNA) Ablation of miR-17-5p in UC6 Cell Line

(A) Detection of miR17-5p ablation by RT-qPCR induced by miR17-5p LNA with NegCtl LNA as control and miR-125a as an external reference.

(B) Top 50 up- and downregulated genes induced by miR17-5p LNA by RNA sequencing (RNA-seq). Genes highlighted in red are involved in neural differentiation.

(C) Representative miR17-5p target genes upregulated induced by miR17-5p LNA transfection revealed by RNA-seq.

Figure 8. *Continued*

(D) Representative upregulated urothelial and neural regulons induced by miR17-5p LNA transfection revealed by RNA-seq.

(E) Upregulation of E2F1 expression induced by miR17-5p LNA quantified by RT-qPCR.

(F) Representative upregulated E2F1 targets induced by miR17-5p LNA revealed by RNA-seq. In panels A and E, a p value < 0.05 was considered statistically significant.

enzyme, APOBEC cytidine deaminase. When compared with conventional UC, SCCs are enriched for specific mutations. Practically all of them have inactivating mutations of p53 and many of them show combined inactivation of p53 and RB1, resulting in the combined loss of TP53- and RB1-associated transcriptional pathway activities seen in these tumors. Consistent with their basal origin, none of the SCCs contained the mutations that are typically enriched in the luminal subset of conventional UCs. Another feature of SCCs is the absence of mutations in chromatin-remodeling genes, which are common in conventional UC (Gui et al., 2011), reinforcing the concept that SCCs evolve from a unique subset of conventional UCs.

SCCs show a genome-wide change in their expression profiles affecting nearly 40% of the protein-encoding genome. Much of this expressional change resulted in lineage plasticity, exhibited by global downregulation of genes involved in urothelial differentiation and upregulation of neural differentiation programs. The urothelial-to-neural phenotypic switch was further confirmed by the in-depth analysis of the urothelial and neural regulons, which revealed widespread downregulation of practically all targets controlled by master regulators of the urothelial phenotype and activation of genes in neural regulons. This dramatic phenotypic switch occurred in a background of dysregulated EMT. Although the transcription factors that serve as master regulators of EMT (i.e., the SNAIL, TWIST, and ZEB families) were not activated, SCCs showed downregulation of epithelial micro RNAs (i.e., the miR-200 family and miR-182/183) and homotypic epithelial adhesion molecules such as E-cadherin (CDH1), claudin-1 (CLDN1), and tight junction protein 1 (TJP1). Our quantitative assessment of EMT showed that SCC had the lowest EMT scores, whereas the double-negative subtype of conventional UC had intermediate EMT scores corresponding to their respective complete and partial EMT states. At the core of EMT circuitry was the combined loss of p53 and RB complemented by the downregulation of p63, which positively regulates basal markers and negatively regulates EMT. These features not only activate EMT but also correspond to the essential component of lineage plasticity representing the fundamental biologic feature of progression to SCC.

The neural signature based on limited number of markers was originally identified by the TCGA and Lund groups and was further confirmed on several multi-institutional cohorts (Batista da Costa et al., 2019; Marzouka et al., 2018; Robertson et al., 2017; Sjodahl et al., 2017; Tan et al., 2019). The neural expression profile of these tumors overlap with those of SCC, but the coexpression of urothelial markers and the retention or loss of urothelial phenotype varies in different cohorts. Therefore, it is uncertain whether the tumors in these cohorts represent conventional UCs with upregulation of neural/neuroendocrine genes or they are SCC neuroendocrine variants exhibiting loss of urothelial features with gain of neural differentiation.

Several groups of investigators proposed various classification schemes of bladder cancer based on their molecular profile (Choi et al., 2014b; Damrauer et al., 2014; Robertson et al., 2017; Sjodahl et al., 2012). The original classification proposed by the Lund group divided bladder cancers into five categories (Sjodahl et al., 2012). Similarly the TCGA group identified five molecularly distinct classes of bladder cancer as luminal-papillary, luminal-infiltrated, luminal, basal/squamous, and neuronal (Robertson et al., 2017). Damrauer et al. and our group proposed classifiers dividing bladder cancer into major groups referred to as luminal and basal with a third small subset referred to as double-negative (Choi et al., 2014b; Dadhania et al., 2016; Damrauer et al., 2014). Additional meta- and consensus analyses of bladder cancer were recently published (Kamoun et al., 2020; Tan et al., 2019). Although various groups of investigators propose a different terminology, the set of markers used to classify bladder cancers among the groups is overlapping, and it appears that the top level of dichotomy is between luminal and basal categories. These two major subsets of bladder cancer show features of undifferentiated basal and more differentiated intermediate luminal urothelial cell layers (Choi et al., 2014a, 2014b). A small subset of tumors, referred to as double-negative, appears to be negative for both luminal and basal markers (Dadhania et al., 2016). They can be perceived as a poorly differentiated variant of the basal subtype as they retain the expression of a signature basal marker, CD44. They also show overlapping mutational

signature with the basal subtype (Dadhania et al., 2016). Several animal model studies confirmed the distinct cellular origin of basal and luminal tumors. These studies suggest that basal tumors originate from undifferentiated uroprogenitor cells (Shin et al., 2014; Van Batavia et al., 2014). In contrast, luminal tumors are derived from intermediate uroplakin-expressing cells (Saito et al., 2018). Our combined analyses of the mutational landscape and expression signatures suggest that SCC evolves from basal subtype with double-negative tumors representing an intermediate step in the progression.

We found that SCCs are depleted of immune cell infiltration and are characterized by the so-called immune desert phenotype. In contrast to closely related sarcomatoid carcinomas of the bladder, which show overexpression of the immune checkpoint receptor ligand PD-L1, SCCs express a unique immune checkpoint receptor adenosine receptor A2A (ADORA2A) that has been shown to be a potent inhibitor of immune infiltration and especially of CD8+ cytotoxic lymphocytes (Ma et al., 2017). The overexpression of ADORA2A in the majority of SCCs may explain their overall immune-null phenotype but most importantly represents as of yet unexplored therapeutic target for patients affected by this highly lethal variant of bladder cancer. A phase I clinical trial of A2AR antagonists (PBF-509 and CPI-444) alone or in combination with the immune checkpoint inhibitor atezolizumab is currently recruiting participants to study the efficiency of A2AR blockade for solid tumor including bladder cancer (NCT02655822).

Limitations of the Study

- The number of SCCs analyzed in our study is a limiting factor, and our observations should be validated on larger numbers of cases possibly from the inter-institutional cohorts.
- Immune desert phenotype of SCC and the putative role of ADORA2A may represent therapeutic targets and should be validated in preclinical models.

Resource Availability

Lead Contact

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, B. Czerniak (bczernia@mdanderson.org).

Materials Availability

Requests for materials related to this study should be directed to the Lead Contact of this study.

Data and Code Availability

Data were deposited on SRA (SRA SUB6960718) and GEO (GEO: GSE128192, GSE145260, GSE145259, and GSE128277).

METHODS

All methods can be found in the accompanying Transparent Methods supplemental file.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.isci.2020.101201.

ACKNOWLEDGMENTS

This study was supported by NCI Genitourinary Bladder SPORE Grant P50CA 91846 (Project 1 and Core C to B.C.). G.Y. is a visiting scientist from China, and his permanent institutional affiliation is Department of Urology, Renji Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China.

AUTHOR CONTRIBUTIONS

B.C. conceived and supervised the whole project and wrote the manuscript. G.Y., D.C., J.B., S.L., J.G.L., S.Z., W.C., and D.J.M. performed the experiments and generated data. Z.W., Y.W., and H.Y. performed data analysis. M.K., P.W., and J.N.W. designed and supervised data analysis. J.B., C.C.G., V.D., Y.L., G.W., and Y.W. organized and analyzed pathologic data. A.K., C.D., D.T., A.S.-R., C.L., and J.G. provided and analyzed clinical data.

DECLARATION OF INTERESTS

The authors declare no conflict of interest.

Received: March 25, 2020
Revised: May 8, 2020
Accepted: May 22, 2020
Published: June 26, 2020

REFERENCES

Abudureyimu, S., Asai, N., Enomoto, A., Weng, L., Kobayashi, H., Wang, X., Chen, C., Mii, S., and Takahashi, M. (2018). Essential role of Linx/Islr2 in the development of the forebrain anterior commissure. Sci. Rep. *8*, 7292.

Alexandrov, L.B., Nik-Zainal, S., Wedge, D.C., Aparicio, S.A., Behjati, S., Biankin, A.V., Bignell, G.R., Bolli, N., Borg, A., Borresen-Dale, A.L., et al. (2013). Signatures of mutational processes in human cancer. Nature *500*, 415–421.

Amin, M.B. (2009). Histological variants of urothelial carcinoma: diagnostic, therapeutic and prognostic implications. Mod. Pathol. *22*, S96–S118.

Antoni, S., Ferlay, J., Soerjomataram, I., Znaor, A., Jemal, A., and Bray, F. (2017). Bladder cancer incidence and mortality: a global overview and recent trends. Eur. Urol. *71*, 96–108.

Batista da Costa, J., Gibb, E.A., Bivalacqua, T.J., Liu, Y., Oo, H.Z., Miyamoto, D.T., Alshalalfa, M., Davicioni, E., Wright, J., Dall'Era, M.A., et al. (2019). Molecular characterization of neuroendocrine-like bladder cancer. Clin. Cancer Res. *25*, 3908–3920.

Byers, L.A., Diao, L., Wang, J., Saintigny, P., Girard, L., Peyton, M., Shen, L., Fan, Y., Giri, U., Tumula, P.K., et al. (2013). An epithelial-mesenchymal transition gene signature predicts resistance to EGFR and PI3K inhibitors and identifies Axl as a therapeutic target for overcoming EGFR inhibitor resistance. Clin. Cancer Res. *19*, 279–290.

Celia-Terrassa, T., Liu, D.D., Choudhury, A., Hang, X., Wei, Y., Zamalloa, J., Alfaro-Aco, R., Chakrabarti, R., Jiang, Y.Z., Koh, B.I., et al. (2017). Normal and cancerous mammary stem cells evade interferon-induced constraint through the miR-199a-LCOR axis. Nat. Cell Biol. *19*, 711–723.

Choi, W., Czerniak, B., Ochoa, A., Su, X., Siefker-Radtke, A., Dinney, C., and McConkey, D.J. (2014a). Intrinsic basal and luminal subtypes of muscle-invasive bladder cancer. Nat. Rev. Urol. *11*, 400–410.

Choi, W., Porten, S., Kim, S., Willis, D., Plimack, E.R., Hoffman-Censits, J., Roth, B., Cheng, T., Tran, M., Lee, I.L., et al. (2014b). Identification of distinct basal and luminal subtypes of muscle-invasive bladder cancer with different sensitivities to frontline chemotherapy. Cancer Cell *25*, 152–165.

Cordes, K.R., Sheehy, N.T., White, M.P., Berry, E.C., Morton, S.U., Muth, A.N., Lee, T.H., Miano, J.M., Ivey, K.N., and Srivastava, D. (2009). miR-145 and miR-143 regulate smooth muscle cell fate and plasticity. Nature *460*, 705–710.

Czerniak, B., Dinney, C., and McConkey, D. (2016). Origins of bladder cancer. Annu. Rev. Pathol. *11*, 149–174.

Dadhania, V., Zhang, M., Zhang, L., Bondaruk, J., Majewski, T., Siefker-Radtke, A., Guo, C.C., Dinney, C., Cogdell, D.E., Zhang, S., et al. (2016). Meta-analysis of the luminal and basal subtypes of bladder cancer and the identification of signature immunohistochemical markers for clinical use. EBioMedicine *12*, 105–117.

Damrauer, J.S., Hoadley, K.A., Chism, D.D., Fan, C., Tiganelli, C.J., Wobker, S.E., Yeh, J.J., Milowsky, M.I., Iyer, G., Parker, J.S., et al. (2014). Intrinsic subtypes of high-grade bladder cancer reflect the hallmarks of breast cancer biology. Proc. Natl. Acad. Sci. U S A *111*, 3110–3115.

Edwards, R.A., and Bryan, J. (1995). Fascins, a family of actin bundling proteins. Cell Motil. Cytoskeleton *32*, 1–9.

Faltas, B.M., Prandi, D., Tagawa, S.T., Molina, A.M., Nanus, D.M., Sternberg, C., Rosenberg, J., Mosquera, J.M., Robinson, B., Elemento, O., et al. (2016). Clonal evolution of chemotherapy-resistant urothelial carcinoma. Nat. Genet. *48*, 1490–1499.

Filippakopoulos, P., Picaud, S., Mangos, M., Keates, T., Lambert, J.P., Barsyte-Lovejoy, D., Felletar, I., Volkmer, R., Muller, S., Pawson, T., et al. (2012). Histone recognition and large-scale structural analysis of the human bromodomain family. Cell *149*, 214–231.

Gui, Y., Guo, G., Huang, Y., Hu, X., Tang, A., Gao, S., Wu, R., Chen, C., Li, X., Zhou, L., et al. (2011). Frequent mutations of chromatin remodeling genes in transitional cell carcinoma of the bladder. Nat. Genet. *43*, 875–878.

Hoadley, K.A., Yau, C., Hinoue, T., Wolf, D.M., Lazar, A.J., Drill, E., Shen, R., Taylor, A.M., Cherniack, A.D., Thorsson, V., et al. (2018). Cell-of-Origin patterns dominate the molecular classification of 10,000 tumors from 33 types of cancer. Cell *173*, 291–304.e6.

Kamat, A.M., Hahn, N.M., Efstathiou, J.A., Lerner, S.P., Malmstrom, P.U., Choi, W., Guo, C.C., Lotan, Y., and Kassouf, W. (2016). Bladder cancer. Lancet *388*, 2796–2810.

Kamoun, A., de Reynies, A., Allory, Y., Sjodahl, G., Robertson, A.G., Seiler, R., Hoadley, K.A., Groeneveld, C.S., Al-Ahmadie, H., Choi, W., et al. (2020). A consensus molecular classification of muscle-invasive bladder cancer. Eur. Urol. *77*, 420–433.

Ma, S.R., Deng, W.W., Liu, J.F., Mao, L., Yu, G.T., Bu, L.L., Kulkarni, A.B., Zhang, W.F., and Sun, Z.J. (2017). Blockade of adenosine A2A receptor enhances CD8(+) T cells response and decreases regulatory T cells in head and neck squamous cell carcinoma. Mol. Cancer *16*, 99.

Marzouka, N.A., Eriksson, P., Rovira, C., Liedberg, F., Sjodahl, G., and Hoglund, M. (2018). A validation and extended description of the Lund taxonomy for urothelial carcinoma using the TCGA cohort. Sci. Rep. *8*, 3737.

Mogilyansky, E., and Rigoutsos, I. (2013). The miR-17/92 cluster: a comprehensive update on its genomics, genetics, functions and increasingly important and numerous roles in health and disease. Cell Death Differ. *20*, 1603–1614.

Newman, A.M., Liu, C.L., Green, M.R., Gentles, A.J., Feng, W., Xu, Y., Hoang, C.D., Diehn, M., and Alizadeh, A.A. (2015). Robust enumeration of cell subsets from tissue expression profiles. Nat. Methods *12*, 453–457.

Ochoa, A.E., Choi, W., Su, X., Siefker-Radtke, A., Czerniak, B., Dinney, C., and McConkey, D.J. (2016). Specific micro-RNA expression patterns distinguish the basal and luminal subtypes of muscle-invasive bladder cancer. Oncotarget *7*, 80164–80174.

Pijnappel, W.P., Kolkman, A., Baltissen, M.P., Heck, A.J., and Timmers, H.M. (2009). Quantitative mass spectrometry of TATA binding protein-containing complexes and subunit phosphorylations during the cell cycle. Proteome Sci. *7*, 46.

Pruszak, J. (2015). Neural Surface Antigens from Basic Biology towards Biomedical Applications (Elsevier).

Rajman, M., and Schratt, G. (2017). MicroRNAs in neural development: from master regulators to fine-tuners. Development *144*, 2310–2322.

Robertson, A.G., Kim, J., Al-Ahmadie, H., Bellmunt, J., Guo, G., Cherniack, A.D., Hinoue, T., Laird, P.W., Hoadley, K.A., Akbani, R., et al. (2017). Comprehensive molecular characterization of muscle-invasive bladder cancer. Cell *171*, 540–556.e5.

Saito, R., Smith, C.C., Utsumi, T., Bixby, L.M., Kardos, J., Wobker, S.E., Stewart, K.G., Chai, S., Manocha, U., Byrd, K.M., et al. (2018). Molecular subtype-specific immunocompetent models of high-grade urothelial carcinoma reveal differential neoantigen expression and response to immunotherapy. Cancer Res. *78*, 3954–3968.

Sharma, P., and Allison, J.P. (2015). Immune checkpoint targeting in cancer therapy: toward combination strategies with curative potential. Cell *161*, 205–214.

Shin, K., Lim, A., Odegaard, J.I., Honeycutt, J.D., Kawano, S., Hsieh, M.H., and Beachy, P.A. (2014). Cellular origin of bladder neoplasia and tissue dynamics of its progression to invasive carcinoma. Nat. Cell Biol. *16*, 469–478.

Siddall, N.A., McLaughlin, E.A., Marriner, N.L., and Hime, G.R. (2006). The RNA-binding protein Musashi is required intrinsically to maintain stem cell identity. Proc. Natl. Acad. Sci. U S A *103*, 8402–8407.

Siegel, R.L., Miller, K.D., and Jemal, A. (2019). Cancer statistics, 2019. CA Cancer J. Clin. *69*, 7–34.

Sjodahl, G., Eriksson, P., Liedberg, F., and Hoglund, M. (2017). Molecular classification of urothelial carcinoma: global mRNA classification versus tumour-cell phenotype classification. J. Pathol. *242*, 113–125.

Sjodahl, G., Lauss, M., Lovgren, K., Chebil, G., Gudjonsson, S., Veerla, S., Patschan, O., Aine, M., Ferno, M., Ringner, M., et al. (2012). A molecular taxonomy for urothelial carcinoma. Clin. Cancer Res. *18*, 3377–3386.

Stacy, A.J., Craig, M.P., Sakaram, S., and Kadakia, M. (2017). DeltaNp63alpha and microRNAs: leveraging the epithelial-mesenchymal transition. Oncotarget *8*, 2114–2129.

Tan, T.Z., Rouanne, M., Tan, K.T., Huang, R.Y., and Thiery, J.P. (2019). Molecular subtypes of urothelial bladder cancer: results from a meta-cohort analysis of 2411 tumors. Eur. Urol. *75*, 423–432.

Tran, M.N., Choi, W., Wszolek, M.F., Navai, N., Lee, I.L., Nitti, G., Wen, S., Flores, E.R., Siefker-Radtke, A., Czerniak, B., et al. (2013). The p63 protein isoform DeltaNp63alpha inhibits epithelial-mesenchymal transition in human bladder cancer cells: role of MIR-205. J. Biol. Chem. *288*, 3275–3288.

Van Batavia, J., Yamany, T., Molotkov, A., Dan, H., Mansukhani, M., Batourina, E., Schneider, K., Oyon, D., Dunlop, M., Wu, X.R., et al. (2014). Bladder cancers arise from distinct urothelial sub-populations. Nat. Cell Biol. *16*, 982–991, 981–5.

van Dijk, N., Funt, S.A., Blank, C.U., Powles, T., Rosenberg, J.E., and van der Heijden, M.S. (2019). The cancer immunogram as a framework for personalized immunotherapy in urothelial cancer. Eur. Urol. *75*, 435–444.

Veskimae, E., Espinos, E.L., Bruins, H.M., Yuan, Y., Sylvester, R., Kamat, A.M., Shariat, S.F., Witjes, J.A., and Comperat, E.M. (2019). What is the prognostic and clinical importance of urothelial and nonurothelial histological variants of bladder cancer in predicting oncological outcomes in patients with muscle-invasive and metastatic bladder cancer? A European Association of Urology muscle invasive and metastatic bladder cancer guidelines panel systematic review. Eur. Urol. Oncol. *2*, 625–642.

Wierer, M., Prestel, M., Schiller, H.B., Yan, G., Schaab, C., Azghandi, S., Werner, J., Kessler, T., Malik, R., Murgia, M., et al. (2018). Compartment-resolved proteomic analysis of mouse aorta during atherosclerotic plaque formation reveals osteoclast-specific protein expression. Mol. Cell Proteomics *17*, 321–334.

Yu, M., Bardia, A., Wittner, B.S., Stott, S.L., Smas, M.E., Ting, D.T., Isakoff, S.J., Ciciliano, J.C., Wells, M.N., Shah, A.M., et al. (2013). Circulating breast tumor cells exhibit dynamic changes in epithelial and mesenchymal composition. Science *339*, 580–584.

Yuan, S., Norgard, R.J., and Stanger, B.Z. (2019). Cellular plasticity in cancer. Cancer Discov. *9*, 837–851.

iScience, Volume *23*

Supplemental Information

Urothelial-to-Neural Plasticity Drives Progression to Small Cell Bladder Cancer

Guoliang Yang, Jolanta Bondaruk, David Cogdell, Ziqiao Wang, Sangkyou Lee, June Goo Lee, Shizhen Zhang, Woonyoung Choi, Yan Wang, Yu Liang, Gang Wang, Ying Wang, Hui Yao, Vipulkumar Dadhania, Jianjun Gao, Christopher Logothetis, Arlene Siefker-Radtke, Ashish Kamat, Colin Dinney, Dan Theodorescu, Marek Kimmel, Peng Wei, Charles C. Guo, John N. Weinstein, David J. McConkey, and Bogdan Czerniak

SUPPLEMENTAL FIGURES AND LEGENDS

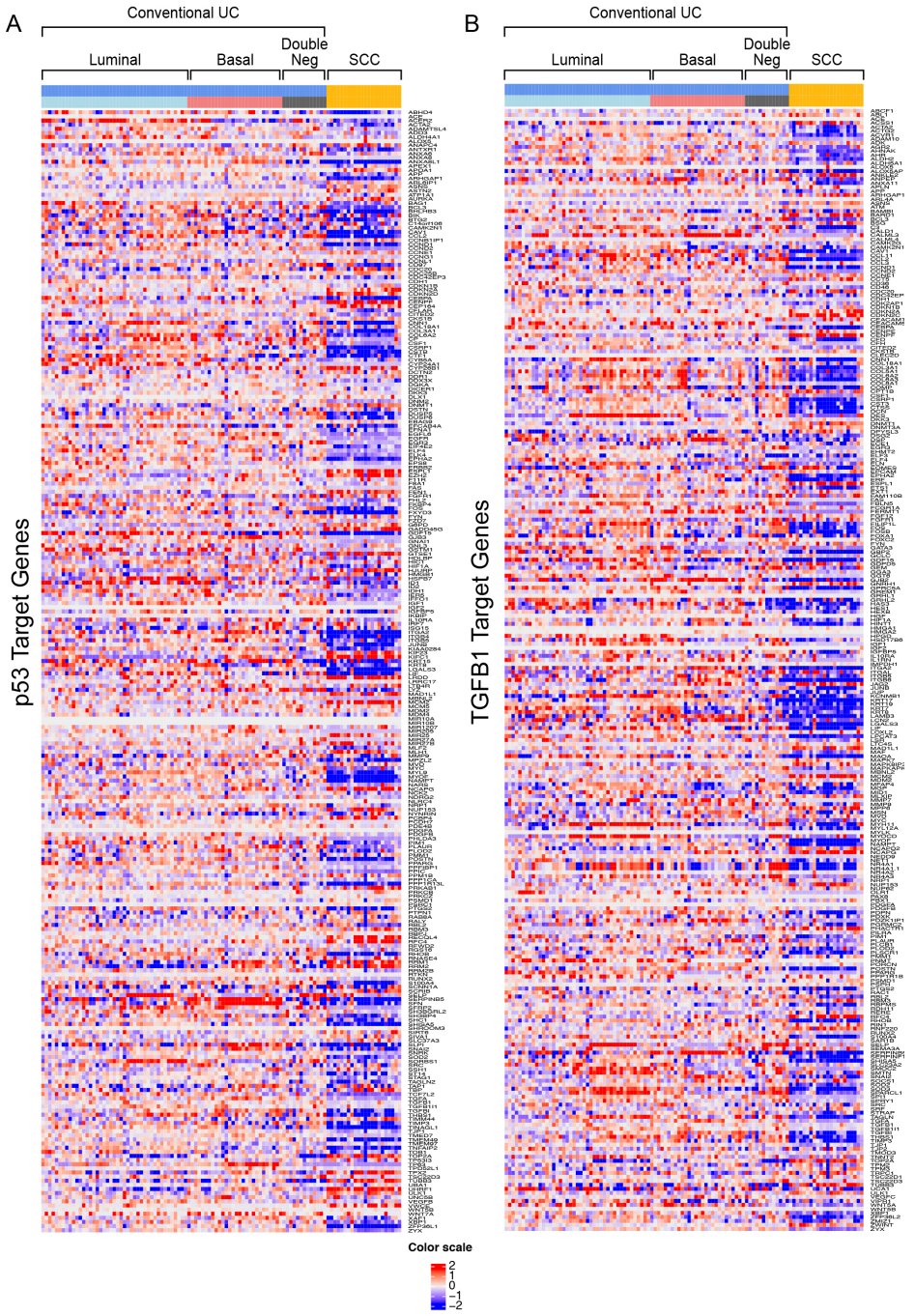


Figure S1: Expression patterns of TP53 and TGFB1 target genes in SCC and conventional UC, Related to Figure 2.

(A) Expression patterns of TP53 target genes in SCC and molecular subtypes of conventional UC. (B) Expression patterns of TGFB1 target genes in SCC and molecular subtypes of conventional UC.

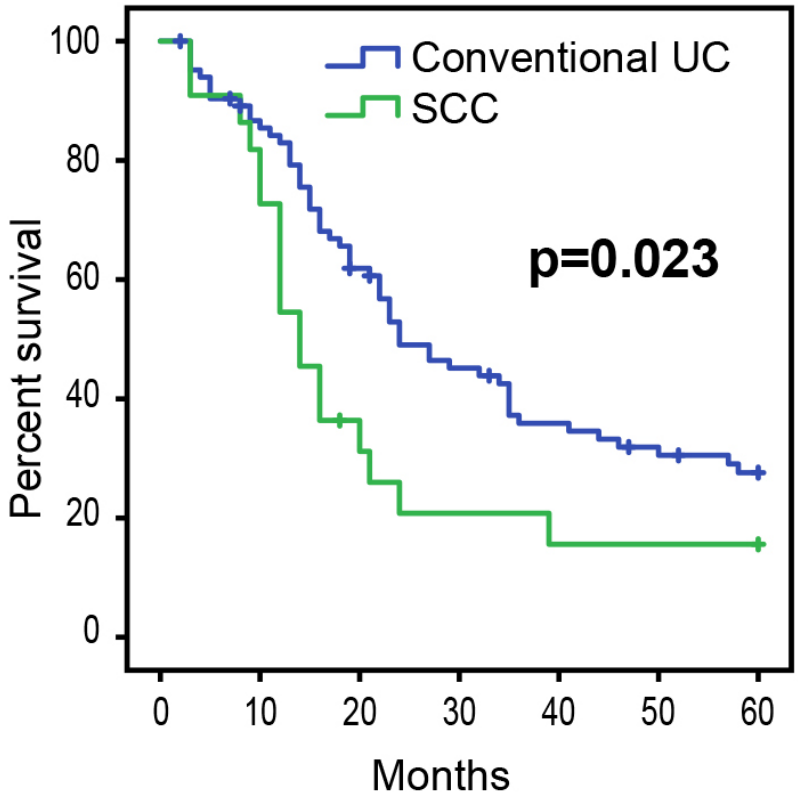


Figure S2: Survival analysis in SCC and conventional UC, Related to Figure 2.
Kaplan Meier analysis of survival in SCC and conventional UC and log-rank testing. A p-value <0.05 was considered statistically significant.

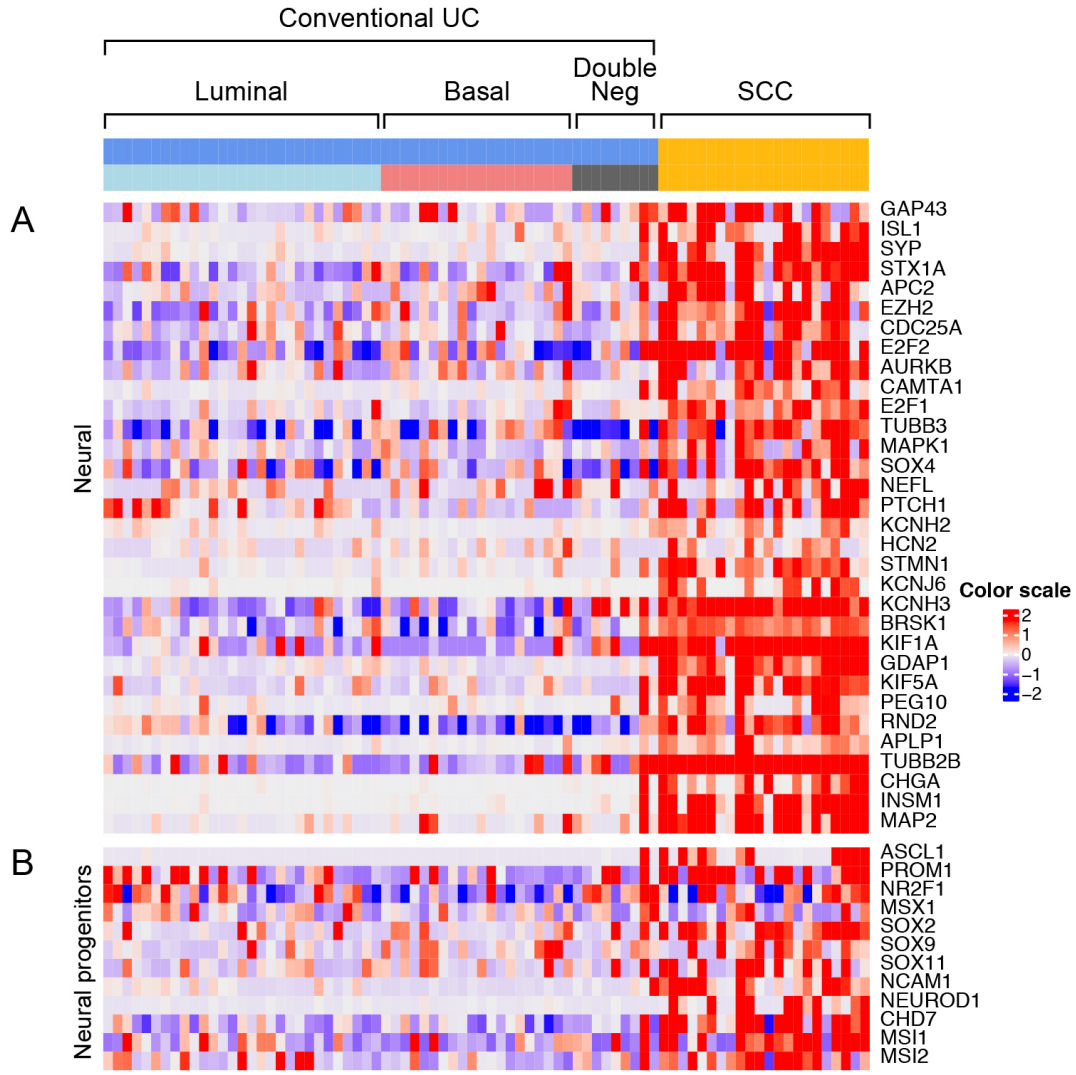


Figure S3: Expression patterns of genes involved in the regulation of neural phenotype, Related to Figure 4.

(A) Expression patterns of neural phenotype-related genes in SCC and molecular subtypes of conventional UC. (B) Expression patterns of transcription factors involved in neural progenitors development in SCC and molecular subtypes of conventional UC.

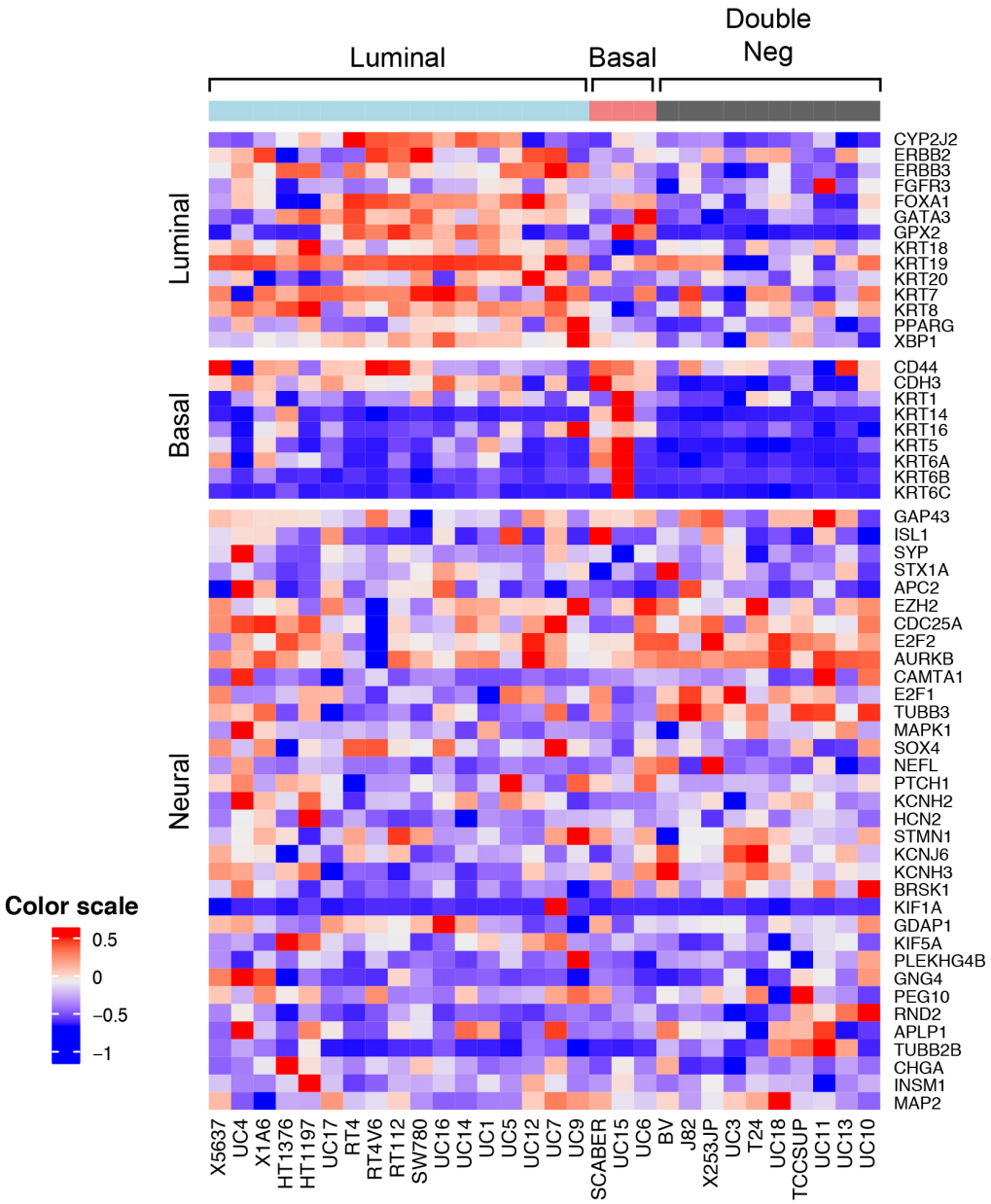


Figure S4. Molecular subtypes in 30 bladder cancer cell lines, Related to Figure 8.
Expression pattern of luminal, basal and neural markers in 30 bladder cancer cell lines revealed by RNASeq analysis.

Table S1: Summary of Clinical and Pathological Data for MD Anderson and TCGA Cohorts

MDACC Small Cell Carcinoma FFPE Cohort, Related to Figure 1.

Stage	Subtype	Gender F/M	Number of samples	Age, yr, mean ± SD
Invasive				
T1	Double negative	1/0	1	N/A
(T2 and higher)	Double negative	2/31	33	67.1 ± 10.8
Total		3/31	34	67.4 ± 10.8

CONVENTIONAL UROTHELIAL CARCINOMA COHORTS

MDACC FFPE Cohort

Invasive	Luminal	7/36	43	70.2 ± 11.7
(T2 and higher)	Basal	10/18	28	68.7 ± 10.9
	Double Negative	2/11	13	67.9 ± 8.8
Total		19/65	84	69.3 ± 11.0

MDACC Tissue Microarray FFPE Cohort

Invasive	Luminal	7/33	40	69.1 ± 11.5
(T2 and higher)	Basal	10/17	27	70 ± 10.3
	Double Negative	1/8	9	63.7 ± 6.1
Total		18/58	76	68.8± 10.6

TCGA Cohort

Invasive	Luminal	48/164	212	68.3 ± 11.0
(T2 and higher)	Basal	56/123	179	68.0 ± 10.1
	Double Negative	3/14	17	65.6 ± 10.2
Total		**107/301**	**408**	**68.1 ± 10.6**

F, female; M, male; yr, year; SD, standard deviation

Table S2: List of Mutations Identified in SCC (n=13) and Paired Samples of Conventional UC, Related to Figure 1.

	Paired Conventional UC/SCC				SCC										
	Conventional UC		SCC												
	1	2	1	2	3	4	5	6	7	8	9	10	11	12	13
	UC1	UC2	SCC1	SCC2	SCC3	SCC4	SCC5	SCC6	SCC7	SCC8	SCC9	SCC10	SCC11	SCC12	SCC13
TP53	p.H179Y	p.C176F	p.H179Y	p.C176F	p.E271K	p.E285K	p.N247I;p.A159fs	p.E286K	p.N210fs	p.E285K	p.R337C	p.R213*		p.P85fs	p.R248W
RB1	p.R798fs		p.R798fs				p.I369fs	p.Y325N; p.R579*;p.M708K			p.E398fs			p.R787*	p.S567L
FSCN3						p.E77Q			p.L259P						
PRAMEF1	p.P456L†											p.S462*			
BRD4											p.Q30E				Splice Site
ISLR2											p.D148N			p.225_226LP>LKRWP	
MAG								p.E178K			p.E572Q				
MAMDC2								p.FDSVLA35fs				p.D131Y			
TAF1D						p.E256*	p.E139del								
CIAPIN1											p.I6N			p.P164Q;p.K163K	
ARID1A					p.YP809fs						p.E1297*				
PKNOX1						p.G367A									
C16orf61															p.V28G
CDC23				p.S311fs											
PCDH18								p.Q744*							
NEUROG1						p.S121*									
SLC9A3R2							p.Q296*								
AR						p.Q488*									
TPX2											Splice Site				
CCDC158						p.S3*									
ADCY8						p.K703K									
LYZL4						p.Q132*									
MYST2						p.S216*									
TEF							p.D13E								
SERPINB12								p.DAPFCLN228fs							
RABEP1						Splice Site									
HOXB5					p.Y19*										
COG1						p.Q139*									
CRTAM						p.S326*									
PIGN								p.Y227*							
DLX2												p.Q195P			
RNF212							p.P147S								

*** Stop codon; fs, Frame shift; †, mutations present in the Conventional UC sample but not in the SCC paired sample**

TRANSPARENT METHODS

Clinical information and tissue samples

The use of human tissue samples for this study was approved by the Institutional Review Board of the UT MD Anderson Cancer Center. Thirty-four archival paraffin-embedded SCC and 84 invasive conventional bladder UC samples from the MD Anderson Cancer Center cohort were analyzed (Table S1). UCs were classified according to the histologic tumor grading system of the World Health Organization. Levels of invasion were defined according to the TNM staging system. All conventional UCs were invasive T2 and above high-grade tumors. The SCC and UC cohorts had similar age distributions with a male predominance. The mean age of the SCC cohort (31 men and 3 women) was 67 years (range, 34–90 years). The mean age of the conventional UC cohort (65 men and 19 women) was 69 years (range, 33–91 years). The median follow-up durations for the SCC and UC cohorts were 14 and 23 months, respectively. Sufficient high-quality DNA was available for 13 SCC cases and 2 paired SCC coexisting with conventional UC cases for whole-exome sequencing. Gene expression profiling was performed on 22 of the cases using Illumina's DASL platform and the data were merged with those obtained from a cohort of 84 conventional UCs. Quantitative RT-PCR was used to analyze the miRNA expression levels in 22 SCC samples and 80 conventional UC samples. To perform immunohistochemical validation studies of selected markers, we created a tissue microarray from the genomic profiling cohort comprising 14 of the cases of SCCs and 76 of the cases of conventional UC. Genomic, clinical, and pathological data from The Cancer Genome Atlas (TCGA) Bladder Cancer (BLCA) cohort of 408 muscle-invasive bladder cancers were used for reference.

DNA and RNA extraction

Genomic DNA and total RNA were extracted from FFPE tissue samples for DNA sequencing and microarray experiments using the MasterPure Complete DNA and RNA Purification Kit (Epicenter Biotechnologies, Madison, WI) according to the manufacturer's instructions as previously described(Guo et al., 2019). In brief, FFPE tissue cylinders were minced, deparaffinized, and digested with 300 μl Proteinase K digestion buffer with 10 μl Proteinase K (50ug/μl, Roche Diagnostics, Mannheim, Germany) at 55 °C overnight. DNA and RNA concentrations and quality were determined using the ND-1000 spectrophotometer (NanoDrop Technologies Inc., Wilmington, DE) and the Quant-iT PicoGreen Kit (Life Technologies, Carlsbad, CA). Sufficient amounts of total RNA for gene expression analysis were extracted from all 47 SCC and 84 conventional UC samples. In addition, sufficient amount of genomic DNA were extracted from 13 cases of SCC, including 2 cases which also contained coexistent conventional precursor UC in the same specimens. DNA extracted from the peripheral blood lymphocytes or normal tissue of the resection specimen from the same patient was used as a control.

Whole-exome sequencing and processing pipeline

Genomic DNA from 13 cases of SCC and two cases of paired conventional UC were used for whole exome sequencing, which was performed on the HiSeq 2000 platform (Illumina, San Diego, CA, USA) at MD Anderson Cancer Center's Genomic Core. The TCGA data on 408 muscle-invasive conventional UC of the bladder were used as a reference set for mutational analyses. BWA-MEM (version 0.7.12) was used to align reads to the hg19 reference genome. Samtools (version 1.4) and Picard (version 2.5.0) were used to sort and convert between formats and remove duplicate reads(Etherington et al., 2015; Li et al., 2009). The Genome Analysis Toolkit (version3.4-46) was used to generate realigned and recalibrated BAM files(McKenna et al., 2010;

Van der Auwera et al., 2013). Somatic variants relative to the normal reference sample were detected by MuTect2(Cai et al., 2016; Callari et al., 2017). Oncotator (version 1.8.0.0) was used to produce gene-based and function-based annotations of the single nucleotide variants (SNVs) and insertions/deletions(Ramos et al., 2015). Similar analyses were performed for the genome-wide expression data from the TCGA cohort (n=408), and tumors were assigned to specific molecular subtypes by applying the sets of luminal and basal markers as described previously(Choi et al., 2014). Mutational data were downloaded from the TCGA portal (https://tcga-data.nci.nih.gov/tcga/). MutSigCV (version1.4; https://www.broadinstitute.org/cancer/cga/mutsig) and used to identify genes that were mutated more often than expected by chance given the background mutation processes(Lawrence et al., 2013). The significant gene list was obtained using a false discovery rate (FDR) cutoff 0.05. The statistical significance of associations between the mutations and the molecular subtypes was assessed by the Fisher's exact test.

Mutagenesis signatures

We used 432 SNVs identified in at least one sample and segregated them into six types of mutations corresponding to the following base pair substitutions: C>A, C>G, C>T, T>A, T>C, T>G. The Fischer's exact test was used to determine the distribution of these mutations in the three groups of samples corresponding to conventional UC in the TCGA cohort, paired UC and SCC cohort. The genomic context of SNVs, referred to as fingerprints which included the two flanking bases on 5' and 3' sides to each position for a total of 96 possible mutational fingerprints, was assembled. Wilcoxon Rank Sum tests were used to test the hypothesis of no difference in the frequency of any fingerprint between any two groups of mucosal samples. The Benjamini and Hochberg method was applied to control the FDR. For each sample, we used its mutational fingerprints (V) and the quadratic programming method to compute a weight score (H) for each of

30 canonical mutational signatures (*W*) available from the Sanger Institute (http://cancer.sanger.ac.uk/cosmic/signatures). We applied the 96-by-30 matrix of canonical signatures (W) and given the 96-by-1 mutational profile of a sample (V), we computed the 30-by-1 vector (H) for each of the canonical signatures' relative contributions to the sample profile by solving the following optimization formula:

$$\min_H (WH - V)^T(WH - V) \text{ such that } h_i \geq 0 \text{ and } \Sigma_i h_i = 1.$$

The optimization problems were solved using "quadprog" (version 1.5-5). Kruskal-Wallis test was used to test against the null hypothesis of no difference in weight scores among three groups of samples: conventional UC, paired UC and SCC.

mRNA expressions and data processing

RNAs from SCCs (n=22), and conventional UCs (n=84) were assessed using Illumina HumanHT-12 DASL Expression BeadChips as per the manufacturer's instructions, and Illumina BeadStudio v3.1.3 (Gene Expression Module V3.3.8) was used for transformation and normalization of the data. Comparisons were carried out using Welch's t-tests and Benjamini-Hochberg -controlled FDR-adjusted p values (<0.05) and fold changes. Unsupervised hierarchical clustering of log ratios was performed in R (Version 3.5.2), and the results were visualized with R package ComplexHeatmap. Pearson's correlation, mean centering, and average linkage were applied in all clustering applications. Genes within 0.5 standard deviations of the log-transformed ratios were discarded. To select specific and robust gene sets associated with SCC, we used the combination analysis with Welch's t-test and fold-change; genes having FDR-adjusted p-values <0.05 and showing fold-change >2.0 were selected. IPA software (Ingenuity Systems, Redwood City, CA) was used to determine dysregulated canonical pathways and predicted upstream

regulators by calculating z-scores and $-\log_{10}$ p values(Jimenez-Marin et al., 2009; Kramer et al., 2014). GSEA was used to evaluate the enrichment probability of the top canonical pathways and upstream regulators identified by IPA(Subramanian et al., 2005). Both SCC and UC samples were classified into luminal, and basal intrinsic molecular subtypes using an algorithm described previously.(Ochoa et al., 2016)

Immune gene expression signatures for SCC and conventional UC were established using unsupervised hierarchical clustering (Pan et al., 2019). Gene dendrogram nodes corresponding to genes characteristically expressed in specific immune cell types were identified and validated through DAVID functional annotation clustering and Ingenuity Systems Analysis (www.ingenuity.com). Immune gene signatures were used as previously reported (De Simone et al., 2016; Iglesia et al., 2014; Sherman et al., 2007; Torri et al., 2010). The immune expression signature was quantitatively assessed by calculating the immune scores for the expression profile of 128 genes shown in Figure 4. Specifically, the immune score for the *i*th sample was defined as m_i-(1/n) $\sum_{j=1}^{n} m_i$, where m_i is the median expression level across the *i*th sample's immune expression profile and (1/n) $\sum_{j=1}^{n} m_i$ is the grand mean of medians across all n samples. Additional analysis of immune infiltrate was performed by the CIBERSORT algorithm (http://cibersort.standford.edu/runcibersort.php). The expression profile of 547 genes using normalized mRNA levels with absolute mode and default parameters was used to assess the presence of 22 immune cell types in the conventional UC and SCCs.(Chen et al., 2018; Gentles et al., 2015) An empirical p-value was calculated using 500 permutations to test against the null hypothesis that no cell type is enriched in each sample. Then a Fisher's exact test was used to test against the null hypothesis of no association between sample types and their statistical significance.

To quantitatively assess the level of EMT (Garg and Singh, 2019; Mak et al., 2016), we calculated the EMT score based on a 76-gene expression signature reported in Byers et al. For each tumor sample, the score was calculated a weighted sum of 76 gene expression levels: $\sum_{i=1}^{76} w_i\, G_{ij}$, where w_i is the correlation coefficient between the *i*th gene expression in the signature and that of E-cadherin and G_{ij} is the *i*th gene's normalized expression in the *j*th tumor sample. We centered the scores by subtracting the mean across all tumor samples so that the grand mean of the score was zero.

miRNA analysis

miRNA analysis was performed on 22 SCC samples and 80 conventional UC samples (Dong et al., 2017). For miRNA cDNA synthesis, 400 ng of total RNA was reverse-transcribed using a miRNA reverse transcription kit (Applied Biosystems; catalogue No. 4366596) in combination with the stem-loop Megaplex primer pool (Applied Biosystems). For each cDNA sample, 381 small RNAs were profiled using TaqMan Human MicroRNA A Cards (Applied Biosystems; catalogue No. 4398965). Fold-change for each microRNA was determined using the ΔCt method and examined using Welch's t-test. An adjusted p value with FDR <0.05 was considered significant.

Regulon analysis

We performed regulon analysis to infer the relative activity of two sets of candidate transcription factors (TFs)(Castro et al., 2016). For a given TF, the set of its putative target genes is defined as a regulon. The first set of TFs were previously reported to be associated with bladder cancer and analyzed for the gene regulatory network: FOXA1, RXRA, FGFR3, RXRB, ERBB3, AR, GATA3, ESR2, ERBB2,PPARG, RARA, FGFR1, PGR, RARB, TP63, ESR1, GATA6,

STAT3, FOXM1, KLF4, EGFR, and HIF1A(Robertson et al., 2017). The second set of TFs were those reported to be related to neuronal phenotyping and manually curated based on the literature, including ASCL1, PROM1, NR2F1, MSX1, SOX2, SOX9, SOX11, NCAM1, NEUROD1, CHD7, MSI1, MSI2, E2F1, E2F2, and SOX4.

To construct the regulons for the TFs of interest, we employed the method developed and implemented in R package RTN1(Castro et al., 2016; Fletcher et al., 2013). In addition, we utilized the IPA to supplement the list of target genes of a TF if the number of target genes identified by the RTN method was too small to allow reliable downstream analyses.

We analyzed each regulon using the two-tailed GSEA(Castro et al., 2016; Subramanian et al., 2005). This method first divided the set of target genes for each TF into positive and negative targets associated with the phenotype of interest (for example, conventional UC vs SCC) using Spearman's correlation coefficient. The distribution of the positive and negative targets was tested respectively, producing enrichment scores (ES) for each sample. The difference between positive ES and negative ES resulted in the differential enrichment score (dES). For further analysis of the difference between the conventional UC and SCC, we extracted the dES for each sample. We assessed how the regulons were associated with bladder cancer subtypes by analyzing the dES with heatmaps and two-tailed GSEA.

Validation studies

Tissue microarrays, immunohistochemistry, and western blotting. The expression levels of selected genes were validated on parallel tissue microarrays (TMAs) comprising FFPE samples from 76 UCs and 14 SCCs. The TMAs were designed and prepared as described previously and profiled by genomic platforms(Sanfrancesco et al., 2016). In brief, TMAs (two 1-mm cores per

case) were constructed with a manual tissue arrayer (Beecher Instruments, Silver Spring, MD). Tissue sections from the TMAs were stained with hematoxylin and eosin to confirm the presence of tumor tissue. Immunohistochemical staining was performed with mouse monoclonal antibodies against human GATA3 (HG3-31 clone, 1:100 dilution; Santa Cruz Biotechnology Inc., Santa Cruz, CA), cytokeratin 5/6 (clone D5/16B4, 1:50 dilution, Dako, Carpinteria, CA), cytokeratin 14 (LL002 clone, 1:50 dilution; BioGenex, Fremont, CA), Synaptophysin, (clone 27G12, 1:600 dilution, Novocastra), Chromogranin (clone LK2H10, 1:600 dilution, Chemicon), and INSM1(clone A-8, 1:300 dilution, Santa Cruz Biotechnology Inc., Santa Cruz, CA). TMAs were also stained with antibodies specific for CD3 (rabbit polyclonal, 1:100 dilution, Dako, Carpinteria, CA), and CD8 Ab-1 (mouse clone C8/144B, 1:50 dilution, Thermo Scientific, Kalamazoo, MI). Immunostaining was performed using the Bond-Max Autostainer (Leica Biosystems, Buffalo Grove, IL). The staining intensities were scored by two pathologists (CCG and BAC) as negative and mildly, moderately, or strongly positive. In addition, the levels of Synaptophysin (D8F6H clone, 1:1000 dilution, Cell Signaling Technology), MSI1 (D46A8 clone, 1:1000 dilution, Cell Signaling Technology), E2F1 (Polyclone, 1:1000 dilution, Cell Signaling Technology), E-Cadherin (4A2 clone, 1:1000 dilution, Cell Signaling Technology), and ADORA2A (7F6-G5-A2 clone, 1:200 dilution, Santa Cruz Biotechnology Inc., Santa Cruz, CA) in SCC were confirmed on selected frozen tumor samples by western blotting.

Locked Nucleic Acid Ablation of miR-17-5p. We first analyzed genomic mRNA profile from 30 bladder cancer cell lines. In brief, total RNA was isolated using the mirVana miRNA isolation kit (Ambion, Inc). RNA purity and integrity were measured by NanoDrop ND-1000 and Agilent Bioanalyzer and only high-quality RNA was used for the cRNA amplification. Direct hybridization assays were performed using the Illumina RNA amplification kit (Ambion, Inc,

Austin, TX) and Illumina HT12 V4 chips (Illumina, Inc., San Diego, CA). Slides were scanned with Bead Station 500X and signal intensities were quantified with GenomeStudio (Illumina, Inc.). Quantile normalization in the Linear Models for Microarray Data (LIMMA) package in the R language environment was used to normalize the data. For the purpose of locked nucleic acid ablation of miR-17-5p experiment we focused our analysis on the expression pattern of luminal, basal, and neural markers. The basal urothelial carcinoma cell line UC6 was selected for the ablation of miR-17-5p. UC6 cells were grown to 70% confluency and were transfected using Lipofectamine 2000 (Thermo Fisher) in triplicates with 20 nM LNA for miR-17-5p (assay ID YI04100215-ADA) and negative control A-LNA (assay ID YI00199006-ADA) (Qiagen). After 24 hours, the cells were washed with PBS and RNA was isolated using Trizol (ThermoFisher). Gene expression profiles of the miR17-5p-LNA and negative control A-LNA in transfected and control UC6 cells were analyzed by RNASeq on an Illumina HiSeq2000. RNA was extracted as before then 76 bp libraries were constructed. mRNA expression profiles were mapped to the hg19 genome using MapSplice (de Matos Simoes et al., 2015). Gene expression was determined by RSEM and normalized with to the upper quartile.

For selected RNAs, expression levels after miR17-5p-LNA transfection were measured by quantitative PCR on an ABI 7900. For miRNA, each sample was reverse transcribed using TaqMan microRNA Reverse Transcription kit with the target specific primer. PCR was then performed to quantify miR-17-5p with miR-125a-3p as control and RNU48 as reference. To determine fold-changes in E2F1, RNA from miR17-5p-LNA and negative control A-LNA transfected UC6 cells were reverse transcribed. PCR was performed for E2F1 and human β actin was used as a reference (Applied Biosystems, Foster City, CA). Software defaults were used to

compute the relative change in expression by the ΔCt method. GraphPad 8.0.0 Prism was used to plot RQ and calculate p value using an unpaired t-test.

General statistical analysis

Survival analyses were performed by Kaplan–Meier analysis and log-rank testing. For genome-wide mRNA and miRNA differential expression analysis, the Benjamini and Hochberg(BH) method was applied to control the false discovery rate(FDR). An adjusted p value with FDR <0.05 was considered statistically significant.

References

Cai, L., Yuan, W., Zhang, Z., He, L., and Chou, K.C. (2016). In-depth comparison of somatic point mutation callers based on different tumor next-generation sequencing depth data. Sci Rep *6*, 36540.

Callari, M., Sammut, S.J., De Mattos-Arruda, L., Bruna, A., Rueda, O.M., Chin, S.F., and Caldas, C. (2017). Intersect-then-combine approach: improving the performance of somatic variant calling in whole exome sequencing data using multiple aligners and callers. Genome Med *9*, 35.

Castro, M.A., de Santiago, I., Campbell, T.M., Vaughn, C., Hickey, T.E., Ross, E., Tilley, W.D., Markowetz, F., Ponder, B.A., and Meyer, K.B. (2016). Regulators of genetic risk of breast cancer identified by integrative network analysis. Nat Genet *48*, 12-21.

Chen, B., Khodadoust, M.S., Liu, C.L., Newman, A.M., and Alizadeh, A.A. (2018). Profiling Tumor Infiltrating Immune Cells with CIBERSORT. Methods Mol Biol *1711*, 243-259.

Choi, W., Czerniak, B., Ochoa, A., Su, X., Siefker-Radtke, A., Dinney, C., and McConkey, D.J. (2014). Intrinsic basal and luminal subtypes of muscle-invasive bladder cancer. Nat Rev Urol *11*, 400-410.

de Matos Simoes, R., Dalleau, S., Williamson, K.E., and Emmert-Streib, F. (2015). Urothelial cancer gene regulatory networks inferred from large-scale RNAseq, Bead and Oligo gene expression data. BMC Syst Biol *9*, 21.

De Simone, M., Arrigoni, A., Rossetti, G., Gruarin, P., Ranzani, V., Politano, C., Bonnal, R.J.P., Provasi, E., Sarnicola, M.L., Panzeri, I.*, et al.* (2016). Transcriptional Landscape of Human Tissue Lymphocytes Unveils Uniqueness of Tumor-Infiltrating T Regulatory Cells. Immunity *45*, 1135-1147.

Dong, F., Xu, T., Shen, Y., Zhong, S., Chen, S., Ding, Q., and Shen, Z. (2017). Dysregulation of miRNAs in bladder cancer: altered expression with aberrant biogenesis procedure. Oncotarget *8*, 27547-27568.

Etherington, G.J., Ramirez-Gonzalez, R.H., and MacLean, D. (2015). bio-samtools 2: a package for analysis and visualization of sequence and alignment data with SAMtools in Ruby. Bioinformatics *31*, 2565-2567.

Fletcher, M.N., Castro, M.A., Wang, X., de Santiago, I., O'Reilly, M., Chin, S.F., Rueda, O.M., Caldas, C., Ponder, B.A., Markowetz, F.*, et al.* (2013). Master regulators of FGFR2 signalling and breast cancer risk. Nat Commun *4*, 2464.

Garg, M., and Singh, R. (2019). Epithelial-to-mesenchymal transition: Event and core associates in bladder cancer. Front Biosci (Elite Ed) *11*, 150-165.

Gentles, A.J., Newman, A.M., Liu, C.L., Bratman, S.V., Feng, W., Kim, D., Nair, V.S., Xu, Y., Khuong, A., Hoang, C.D.*, et al.* (2015). The prognostic landscape of genes and infiltrating immune cells across human cancers. Nat Med *21*, 938-945.

Guo, C.C., Majewski, T., Zhang, L., Yao, H., Bondaruk, J., Wang, Y., Zhang, S., Wang, Z., Lee, J.G., Lee, S.*, et al.* (2019). Dysregulation of EMT Drives the Progression to Clinically Aggressive Sarcomatoid Bladder Cancer. Cell Rep *27*, 1781-1793 e1784.

Iglesia, M.D., Vincent, B.G., Parker, J.S., Hoadley, K.A., Carey, L.A., Perou, C.M., and Serody, J.S. (2014). Prognostic B-cell signatures using mRNA-seq in patients with subtype-specific breast and ovarian cancer. Clin Cancer Res *20*, 3818-3829.

Jimenez-Marin, A., Collado-Romero, M., Ramirez-Boo, M., Arce, C., and Garrido, J.J. (2009). Biological pathway analysis by ArrayUnlock and Ingenuity Pathway Analysis. BMC Proc *3 Suppl 4*, S6.

Kramer, A., Green, J., Pollard, J., Jr., and Tugendreich, S. (2014). Causal analysis approaches in Ingenuity Pathway Analysis. Bioinformatics *30*, 523-530.

Lawrence, M.S., Stojanov, P., Polak, P., Kryukov, G.V., Cibulskis, K., Sivachenko, A., Carter, S.L., Stewart, C., Mermel, C.H., Roberts, S.A.*, et al.* (2013). Mutational heterogeneity in cancer and the search for new cancer-associated genes. Nature *499*, 214-218.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078-2079.

Mak, M.P., Tong, P., Diao, L., Cardnell, R.J., Gibbons, D.L., William, W.N., Skoulidis, F., Parra, E.R., Rodriguez-Canales, J., Wistuba, II*, et al.* (2016). A Patient-Derived, Pan-Cancer EMT Signature Identifies

Global Molecular Alterations and Immune Target Enrichment Following Epithelial-to-Mesenchymal Transition. Clin Cancer Res *22*, 609-620.
McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M.*, et al.* (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res *20*, 1297-1303.
Ochoa, A.E., Choi, W., Su, X., Siefker-Radtke, A., Czerniak, B., Dinney, C., and McConkey, D.J. (2016). Specific micro-RNA expression patterns distinguish the basal and luminal subtypes of muscle-invasive bladder cancer. Oncotarget *7*, 80164-80174.
Pan, S., Zhan, Y., Chen, X., Wu, B., and Liu, B. (2019). Bladder Cancer Exhibiting High Immune Infiltration Shows the Lowest Response Rate to Immune Checkpoint Inhibitors. Front Oncol *9*, 1101.
Ramos, A.H., Lichtenstein, L., Gupta, M., Lawrence, M.S., Pugh, T.J., Saksena, G., Meyerson, M., and Getz, G. (2015). Oncotator: cancer variant annotation tool. Hum Mutat *36*, E2423-2429.
Robertson, A.G., Kim, J., Al-Ahmadie, H., Bellmunt, J., Guo, G., Cherniack, A.D., Hinoue, T., Laird, P.W., Hoadley, K.A., Akbani, R.*, et al.* (2017). Comprehensive Molecular Characterization of Muscle-Invasive Bladder Cancer. Cell *171*, 540-556 e525.
Sanfrancesco, J., McKenney, J.K., Leivo, M.Z., Gupta, S., Elson, P., and Hansel, D.E. (2016). Sarcomatoid Urothelial Carcinoma of the Bladder: Analysis of 28 Cases With Emphasis on Clinicopathologic Features and Markers of Epithelial-to-Mesenchymal Transition. Arch Pathol Lab Med *140*, 543-551.
Sherman, B.T., Huang da, W., Tan, Q., Guo, Y., Bour, S., Liu, D., Stephens, R., Baseler, M.W., Lane, H.C., and Lempicki, R.A. (2007). DAVID Knowledgebase: a gene-centered database integrating heterogeneous gene annotation resources to facilitate high-throughput gene functional analysis. BMC Bioinformatics *8*, 426.
Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S.*, et al.* (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A *102*, 15545-15550.
Torri, A., Beretta, O., Ranghetti, A., Granucci, F., Ricciardi-Castagnoli, P., and Foti, M. (2010). Gene expression profiles identify inflammatory signatures in dendritic cells. PLoS One *5*, e9404.
Van der Auwera, G.A., Carneiro, M.O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J.*, et al.* (2013). From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Curr Protoc Bioinformatics *43*, 11 10 11-33.