HHS Public Access

Author manuscript

Cell Rep. Author manuscript; available in PMC 2024 July 23.

Published in final edited form as:

Cell Rep. 2024 May 28; 43(5): 114146. doi:10.1016/j.celrep.2024.114146.

Loss of *LPAR6* and *CAB39L* dysregulates the basal-to-luminal urothelial differentiation program, contributing to bladder carcinogenesis

Sangkyou Lee[1], **Jolanta Bondaruk**[1], **Yishan Wang**[2], **Huiqin Chen**[2], **June Goo Lee**[1], **Tadeusz Majewski**[1], **Rachel D. Mullen**[3], **David Cogdell**[4], **Jiansong Chen**[2], **Ziqiao Wang**[5], **Hui Yao**[6], **Pawel Kus**[7], **Joon Jeong**[1,12], **Ilkyun Lee**[1,13], **Woonyoung Choi**[8], **Neema Navai**[9], **Charles Guo**[1], **Colin Dinney**[9], **Keith Baggerly**[2], **Cathy Mendelsohn**[10], **David McConkey**[8], **Richard R. Behringer**[3], **Marek Kimmel**[11], **Peng Wei**[2], **Bogdan Czerniak**[1,14,*]

[1]Department of Pathology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

[2]Department of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

[3]Department of Genetics, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

[4]Department of Neurosurgery, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

[5]Department of Biostatistics, Johns Hopkins University, Baltimore, MD 21218, USA

[6]Department of Bioinformatics & Computational Biology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

[7]Department of Systems Biology and Engineering, Silesian University of Technology, Gliwice, Poland

[8]Johns Hopkins Greenberg Bladder Cancer Institute, Johns Hopkins University, Baltimore, MD 21218, USA

[9]Department of Urology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

*Correspondence: bczernia@mdanderson.org.

AUTHOR CONTRIBUTIONS

B.C. conceived and supervised the whole project, interpreted the data, and wrote the manuscript; S.L. and J.G.L. developed animal models and performed the experiments; J.B. performed methylation experiments; T.M., D.C., J.J., and I.L. performed SNP mapping experiments; R.D.M. and R.R.B. performed and supervised the immunofluorescence imaging; R.R.B. supervised the development of germline knockout animal models; N.N. and C.D. provided and organized clinical data for the MD Anderson patient cohorts and whole- organ mapping data; W.C. and D.M. provided and analyzed molecular taxonomy data on bladder cancer cell lines; C.G. performed pathologic analysis; C.M. interpreted the molecular data related to luminal urothelial differentiation; Y.W., H.C., Z.W., H.Y., J.C., and P.K. analyzed the data; K.B., M.K., and P.W. supervised the data analysis.

DECLARATION OF INTERESTS

The authors declare no competing interests.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2024.114146.

[10]Department of Urology, Genetics & Development and Pathology, Columbia University, New York, NY 10032, USA

[11]Department of Statistics, Rice University, Houston, TX 77005, USA

[12]Present address: Department of Surgery, Gangnam Severance Hospital, Seoul, Korea

[13]Present address: Department of Surgery, International St. Mary's Hospital, Incheon, Korea

[14]Lead contact

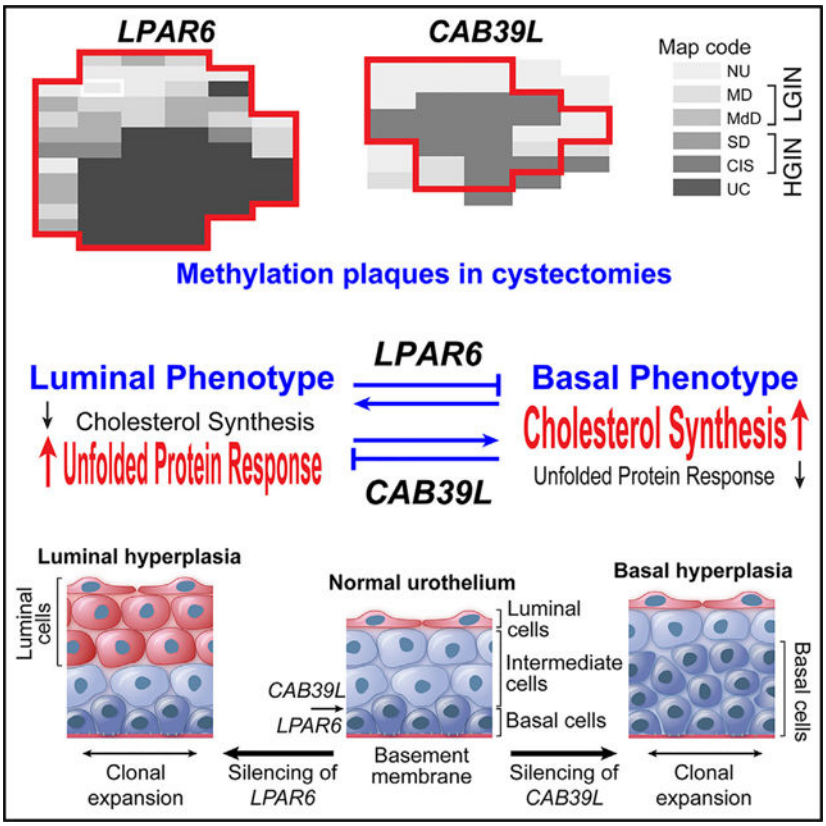
SUMMARY

We describe a strategy that combines histologic and molecular mapping that permits interrogation of the chronology of changes associated with cancer development on a whole-organ scale. Using this approach, we present the sequence of alterations around *RB1* in the development of bladder cancer. We show that *RB1* is not involved in initial expansion of the preneoplastic clone. Instead, we found a set of contiguous genes that we term "forerunner" genes whose silencing is associated with the development of plaque-like field effects initiating carcinogenesis. Specifically, we identified five candidate forerunner genes (*ITM2B*, *LPAR6*, *MLNR*, *CAB39L*, and *ARL11*) mapping near *RB1*. Two of these genes, *LPAR6* and *CAB39L*, are preferentially downregulated in the luminal and basal subtypes of bladder cancer, respectively. Their loss of function dysregulates urothelial differentiation, sensitizing the urothelium to N-butyl-N-(4-hydroxybutyl) nitrosamine-induced cancers, which recapitulate the luminal and basal subtypes of human bladder cancer.

In brief

Lee et al. report that FR genes mapping near *RB1* are associated with the development of field effects initiating bladder carcinogenesis. Two of these genes are distinctively downregulated in luminal and basal cancers. Their loss contributes to carcinogenesis by dysregulating urothelial differentiation mediated by cholesterol and the unfolded protein reaction.

Graphical abstract



INTRODUCTION

The prevailing view is that common epithelial cancers develop from microscopically recognizable precursor lesions such as dysplasia and carcinoma *in situ*.[1] However, new evidence suggests that dysplasia and carcinoma *in situ* are not the initiating events of carcinogenesis but represent the evolution of preexisting occult disease caused by the field effects of carcinogens and/or chronic inflammation, which exhibits minimal phenotypic deviation from normal tissue but harbors complex molecular alterations that trigger carcinogenesis. The areas influenced by field effects may form large plaques involving extensive mucosal areas of the affected organ.[2–4] The molecular mechanisms driving this initiating phase of carcinogenesis are largely unknown, but preliminary evidence has demonstrated that a set of genes distinct from major known oncogenes and tumor suppressors may be involved.[5,6]

Bladder cancer is a unique model of human carcinogenesis, facilitating studies of mucosal field effects because of the simple anatomy of the organ, which permits mapping of *in situ* preneoplastic conditions geographically across the entire mucosa.[7] We developed an approach referred to as whole-organ histologic and genomic mapping (WOHGM) that facilitates the identification of genes and their downstream pathways whose dysregulation is associated with a growth advantage, allowing for tracking of the evolution of bladder cancer, from occult mucosal field effects to invasive disease.[8–14]

Using WOHGM, we previously identified several potential target genes referred to as "forerunner" (FR) genes, which generally map to genomic regions that also contain well-established tumor-suppressor genes.[5] The FR genes are related to tumor suppressors in the sense that they contribute to tumorigenesis via loss of function, but their inactivation precedes functional loss of the neighbor tumor suppressor. We hypothesized that loss of function of the FR genes contributed to the initial growth advantage of a preneoplastic clone, whereas subsequent inactivation of tumor suppressors and activation of oncogenes represented a transforming event associated with clonal evolution to severe dysplasia/carcinoma *in situ* progressing to invasive disease along the luminal and basal pathways.[5,6]

Herein, we present the results of our studies focused on the 13q14 region around the model tumor suppressor *RB1*, in which we identified five candidate FR genes (*ITM2B*, *LPAR6*, *MLNR*, *CAB39L*, and *ARL11*) and provided evidence that their silencing, via hypermethylation or, less frequently, mutations, is associated with the initial expansion of intraurothelial neoplasia that sets the stage for subsequent events of carcinogenesis. We validated the involvement of FR genes in several independent bladder cancer cohorts that were profiled according to a luminal/basal molecular taxonomy.[15–17] These studies demonstrated that the FR genes *LPAR6* and *CAB39L* are distinctively involved in the luminal and basal molecular subtypes of bladder cancer, respectively. We present *in vitro* and *in vivo* evidence that downregulation of these two genes contributes to bladder carcinogenesis by dysregulating the urothelial differentiation program. In animal models, loss of function of *Lpar6* and *Cab39l* induced distinctive luminal and basal types of urothelial hyperplasia, respectively. This in turn sensitized the urothelium to N-butyl-N-(4-hydroxybutyl)nitrosamine (BBN)-induced cancers, which recapitulated the luminal and basal subtypes of human bladder cancers.

RESULTS

The overall strategic plan for our studies concerning the role of FR genes in the development of bladder cancer is summarized in Figure S1. Initially, we performed high-resolution single-nucleotide polymorphism (SNP)-based WOHGM studies focused on the 13q14 region around the model tumor suppressor *RB1*, which identified the candidate FR genes *ITM2B*, *LPAR6*, *MLNR*, *CAB39L*, and *ARL11*. We confirmed their involvement in bladder cancer using human samples and bladder cancer cell lines with custom-made NimbleGen tiling expression arrays spanning an approximately 5-Mb region around *RB1*. We also confirmed the silencing of candidate FR genes by methylation and, occasionally, by inactivating mutations in the early phases of bladder carcinogenesis using WOHGM and validated it in several publicly available human bladder cancer cohorts (The University of Texas MD Anderson Cancer Center, The Cancer Genome Atlas [TCGA], and Lund University) comprising 905 tumor samples.[18,19] We selected the *LPAR6* and *CAB39L* genes for more in-depth mechanistic studies, as they were distinctively involved in molecular subtypes of bladder cancer (i.e., *LPAR6* was preferentially involved in the luminal subtype, whereas *CAB39L* was predominantly involved in the basal subtype). We studied their biologic effects in CRISPR ablation systems with bladder cancer cell lines and in germline ablation mouse models using embryonic fibroblasts, bladder urothelium, and BBN-induced bladder tumors. Finally, we validated the phenotypic effects of *LPAR6* and *CAB39L* loss of

function on the same publicly available bladder cancer cohorts and showed that the luminal and basal molecular subtypes of the disease exhibited *LPAR6*- and *CAB39L*-depleted phenotypes, respectively.

Mapping and expression studies of the candidate FR genes around *RB1*

We performed our whole-organ mapping studies focused on the 13q14 region containing the model tumor suppressor *RB1* with 902 geographically annotated mucosal samples of 20 whole-organ cystectomy samples (Figure S2A). The samples were microscopically classified as normal urothelium (NU), *in situ* preneoplastic conditions referred to as low- and high-grade intraurothelial neoplasia (LGIN and HGIN), and urothelial carcinoma (UC) (Figure S2B). The clinical and pathologic data for the 20 cystectomy samples used in this study are summarized in Table S1. We performed the SNP-based WOHGM studies focused on a 27-Mb segment around *RB1* using 234 samples from the first five cystectomy maps (maps 1–5) (Figure 1A). We superimposed loss of polymorphism (LOP) identified by informative SNPs on the whole-organ histologic maps and compared it to the *RB1* mutational status. This permitted the identification of chromosomal segments with LOPs associated with intraurothelial expansion forming large plaques in the bladder mucosa. Integration of such data from all five whole-organ maps revealed that such losses clustered in an approximately 5-Mb region surrounding *RB1*. The alignment of LOPs in individual whole-organ maps identified a minimal deleted segment involving *RB1* and its flanking regions spanning 1.7 Mb that contained 20 putative positional candidate genes, which may have contributed to the initial clonal expansion of intraurothelial neoplasia. This clonal expansion represented a mucosal field effect that was not accompanied by inactivation of the second *RB1* allele and, most importantly, extended to mucosal areas with minimal or no deviation from normal urothelial morphology (Figure 1B). In cases in which loss of one *RB1* allele was followed by mutational inactivation of the second allele, we found homozygous inactivation of *RB1* to be a later event corresponding to the onset of severe dysplasia/carcinoma *in situ* that eventually progressed to invasive carcinoma (Figure 1B, map 2). Such a pattern of loss and *RB1* gene inactivation provided the foundation for the concept of alternative tumor-suppressor genes termed FR genes that may drive the initial expansion of intraurothelial neoplasia and contribute to the development of mucosal field effects initiating bladder carcinogenesis.

To narrow the list of positional candidate FR genes, we performed expression studies of all genes mapping to a 5-Mb region around *RB1* using custom-made NimbleGen tiling arrays containing 385,000 DNA expression probes.[20] With this approach, we obtained data on all expressed sequences in the region, including the expression patterns for all annotated genes at exon-scale resolution in 9 bladder cancer cell lines and 22 primary human bladder cancers (Figures S3A–S3D). Analysis of these data identified the candidate FR genes *ITM2B*, *LPAR6*, *MLNR*, *CAB39L*, and *ARL11*, whose expression was downregulated at least 2-fold in more than 50% of the tested samples.

To identify the putative mechanisms underlying the observed downregulation of the FR genes, we used TCGA bladder cancer data to analyze the correlation between the expression of the genes and their methylation levels. In these analyses, we considered only the regions

with negative correlation, with coefficient correlation values ranging from −0.21 to −0.67.[18] For the *ITM2B*, *MLNR*, and *ARL11* genes, we explored the level of hypermethylation of the CpG islands in the promoter regions. We also identified the CpG sites in the *LPAR6*, *CAB39L*, and *ARL11* sequences, which are located in either the first exon or the 5′ untranslated region of this exon, in which the level of methylation was negatively correlated with the expression (Figure S4A).

To verify this correlative observation and determine the contributions of epigenetic silencing, we used methylation-specific polymerase chain reaction and sequencing along with expression profiling with 12 bladder cancer cell lines and 5 normal urothelial cell samples (Figure S4B). To quantify the proportion of methylated DNA, we generated standard titration methylation curves for all five candidate FR genes (Figures S4C and S4D). We noticed a negative correlation between the status of methylation and the level of expression for the FR candidate genes in almost all analyzed cell lines (Figures S4E and S4F). In addition, we observed that expression of the FR genes in the selected cell lines could be restored by demethylation after treatment with decitabine (Figure S4G).

We also explored the development of mucosal field effects by FR gene downregulation and the silencing of FR genes by methylation using WOHGM. We initially examined the methylation or mutational status of our candidate FR genes in representative tumor samples from all 20 whole-organ maps. For the six maps in which hypermethylation or mutation of candidate FR genes was identified, we extended testing to all mucosal samples. This approach identified hypermethylation of *ITM2B*, *LPAR6*, *MLNR*, *CAB39L*, and *ARL11* in maps 3, 10, 11, 17, and 21. In addition, in map 16, we identified mutational inactivation of *LPAR6* combined with LOP. In all instances, the areas of hypermethylation or LOP combined with mutational inactivation of *LPAR6* formed contiguous plaques involving large areas of bladder mucosa. Examples of plaque-like methylation patterns for these genes identified via WOHGM are shown in Figure 1C. In each instance, the hypermethylation or mutational inactivation formed a plaque that extended to mucosal areas with minimal or no deviation from normal urothelial morphology. This distribution pattern was consistent with the involvement of the candidate FR genes in early incipient phases of intraurothelial neoplasia and development of diffuse mucosal field effects.

To determine the frequency of FR gene silencing in bladder cancer, we performed mutation, methylation, expression, and LOP studies with the MD Anderson (n = 189), TCGA (n = 408), and Lund University (n = 308) bladder cancer cohorts (Figures S5A–S5E and S6A–S6C). The clinical and pathologic data for these three cohorts are summarized in Table S2. Analysis of the data available for the TCGA cohort showed that mutations of the FR genes were extremely rare and that the predominant mechanism of downregulation was hypermethylation (Figure S5A). Among the candidate FR genes, *LPAR6* exhibited nucleotide substitutions causing an amino acid change in 2% of the bladder cancers. Therefore, in our subsequent studies, we performed mutational analysis of *LPAR6* for the MD Anderson cohort and focused on methylation analysis of all five candidate FR genes. Specifically, we performed quantitative real-time polymerase-chain-reaction-based methylation studies of the FR genes for this cohort. For 111 of the MD Anderson samples, we performed parallel copy-number-loss studies of the *RB1* gene and its flanking regions

spanning approximately 5 Mb (Table S3). The results of the methylation studies for the 189 MD Anderson bladder tumor samples are summarized in Figure S6A. These studies demonstrated that *LPAR6* and *CAB39L* were the most significantly hypermethylated FR genes. The results of parallel LOP studies of *RB1* and its flanking regions in a subset of 111 tumor samples in the same cohort are depicted in Figure S6B. The results of combined methylation mutational analysis of the five FR genes in relation to *RB1* status documented using immunohistochemistry and mutational analysis for the same 111 samples are shown in Figure S6C. Also, the results of mutational, methylation, and LOP studies of *RB1* and the FR genes are summarized in Table S3. We examined the loss of genetic material by allelotyping SNPs mapping to the 5-Mb region around *RB1*, detecting LOP in 67% of bladder cancers and observing that nearly half (49%) of the tumors involved the 1.7-Mb region containing the candidate FR genes. The promoter regions of *ITM2B*, *MLNR*, and *ARL11* were hypermethylated in 42%, 22%, and 24% of the primary human bladder tumor samples, respectively (Figure 1D). In addition, the methylation of the CpG sites located in the sequences of *LPAR6* and *CAB39L* was inversely correlated with their expression levels; these genes were methylated in 72% and 92% of the bladder tumor samples, respectively. When we combined the methylation or mutational status of candidate FR genes with LOP, implicating their homozygous inactivation, the range for individual candidate FR genes varied from 3% to 27% (Figure 1E). Furthermore, we identified methylation involving greater than 50% of DNA, implicating homozygous silencing of at least one FR gene without associated LOP in 51% of the cases. We identified LOP combined with methylation or mutation of at least one FR gene in 27% of the cases, as well. Taking these analyses together, we documented homozygous inactivation of at least one FR gene in 36% of the bladder tumor samples. At least one FR gene was lost in more than 40% of bladder cancers. We observed combined loss of FR genes and *RB1* in 15% of the samples (Figure 1F).

To identify the frequency of involvement of candidate FR genes in molecular subtypes of bladder cancer, we analyzed their expression and methylation in the TCGA (n = 408) and Lund University (n = 308) cohorts profiled according to luminal/basal taxonomy (Table S2; Figures S6A–S6E).[18,19] Both cohorts had somewhat consistent patterns of methylation and downregulation of FR genes in relation to molecular subsets of bladder cancer. First, these studies confirmed that FR genes were hypermethylated and downregulated in about 50% of luminal and basal bladder cancers. Moreover, downregulation and hypermethylation of *CAB39L* were enriched in the basal subtype. The remaining candidate FR genes were downregulated or hypermethylated at almost equal levels in about 50% of both the luminal and the basal subtypes. *LPAR6* was more frequently downregulated in the luminal subtype in the Lund University cohort, which included a large proportion of low-grade superficial papillary tumors, which are almost exclusively luminal. In our subsequent studies, we focused on the biologic effects of *LPAR6* and *CAB39L* loss of function, which appeared to be distinctively involved in the development of luminal and basal tumors.

Effects of *LPAR6* and *CAB39L* loss of function on bladder cancer cell lines

To study the biologic effects and underlying mechanisms of *LPAR6* and *CAB39L* loss of function, we ablated these genes in selected bladder cancer cell lines. Because our initial observations provided circumstantial evidence that these genes may regulate the urothelial

differentiation program, we asked whether their loss of function may change the luminal and basal phenotypes of bladder cancer cell lines. To answer this question, we first classified bladder cancer cell lines according to the previously developed luminal and basal classifier including the quantitative basal-to-luminal transition (BLT) score (Figures S7A and S7B).[17] We observed that luminal cell lines uniformly formed well-developed urospheres exhibiting the expression and microscopic features of luminal differentiation when they grew in an adhesion-free medium. On the other hand, basal cell lines did not form urospheres in an adhesion-free medium and grew in a predominantly dispersed pattern, occasionally forming small, loosely arranged cell clusters that did not exhibit features of luminal differentiation. Given these data, we selected prototypic basal (UC6) and luminal (UC7) cell lines to ablate *LPAR6* and *CAB39L*, respectively, using CRISPR technology (Figures S8A and S8B). We predicted that ablation of *LPAR6* in the UC6 cell line could activate luminal differentiation. In contrast, we expected that ablation of *CAB39L* in the UC7 cell line could suppress its differentiation and activate basal features.

Effects of LPAR6 *loss on the UC6 basal bladder cancer cell line*

As expected, loss of *LPAR6* in UC6 cells induced the formation of well-developed urospheres exhibiting luminal differentiation as documented using light and electron microscopy as well as quantitative assessment of the size and number of the urospheres (Figures 2A–2I). $LPAR6^{-/-}$ cells had upregulation of *GATA3* according to immunohistochemistry, whereas dispersedly growing parental $LPAR6^{+/+}$ cells were negative for this signature luminal marker (Figures 2G and 2H). RNA sequencing demonstrated that loss of *LPAR6* in UC6 cells dysregulated the expression of 8,929 genes (4,446 upregulated and 4,483 downregulated). Among the top upregulated genes associated with loss of *LPAR6* were *LAMB3* and *EREG*, which are involved in epithelial differentiation; *IL36G*, *IL1A*, and *PLEKHG5*, which activate the NFKB1 signaling pathway; and *STAT4* (Figure 2J). The downregulated genes included *FZD4*, which is involved in regulation of the β-catenin signaling pathway; *DUSP4*, which negatively regulates the MAPK/ERK pathway; and *CAMK2N1*, which negatively regulates the ERK1/ERK2 cascade (Figure 2J). The formation of urospheres was accompanied by activation of the signature luminal transcriptional factors *GATA3* and *PPARγ* as well as their downstream target genes (Figures 2K–2M). In addition, activation of the luminal differentiation program was confirmed by a significant increase in the BLT score (Figure 2N). Activation of luminal differentiation was associated with downregulation of cholesterol synthesis and upregulation of the unfolded protein response (Figures 2O and 2P). However, gain of luminal differentiation was not associated with suppression of the basal phenotype, as these cells had upregulation of *p63*, a key transcriptional factor involved in maintenance of the basal phenotype of urothelial cells, and its downstream target genes.[21,22] In fact, *p63* was among the top upregulated genes in $LPAR6^{-/-}$ UC6 cells.

Effects of CAB39L *loss on the UC7 luminal bladder cancer cell line*

Loss of *CAB39L* in UC7 cells suppressed their luminal phenotype and activated their undifferentiated basal features. $CAB39L^{-/-}$ cells lost their ability to form well-organized urospheres and to express *GATA3* (Figures 3A–3I). RNA sequencing demonstrated that $CAB39L^{-/-}$ UC7 cells had dysregulation of 1,106 genes (514 upregulated and 592

downregulated). Among the top upregulated genes were *PEG10*, *THBS1*, *RHOB*, and *KRT5* (Figure 3J). The downregulated genes included *DMXL2*, *FGFRL1*, and *TNFRSF1B* (Figure 3J). Loss of urosphere formation was accompanied by downregulation of the GATA3 and PPARγ luminal transcription factors and their target genes (Figure 3K). These features were complemented by activation of the p63 basal transcription factor and its respective downstream target genes as well as a significant reduction of the BLT score (Figures 3L–3N). Loss of luminal differentiation in these cells was associated with activation of cholesterol synthesis and suppression of unfolded protein response (Figures 3O and 3P).

Effects of *Lpar6* and *Cab39l* loss of function in mouse models

To determine the biologic effects of *Lpar6* and *Cab39l* loss of function *in vivo*, we developed mouse models with germline knockout of these genes. The generation of germline $Lpar6^{-/-}$ and $Cab39l^{-/-}$ knockout mice is depicted in Figures S8C–S8F.

Effects of Lpar6 *and* Cab39l *loss on mouse embryonic fibroblasts*

To verify the biologic effects of *Lpar6* and *Cab39l* loss observed in bladder cancer cell lines on a genetic background of normal proliferating cells, we used cultured mouse embryonic fibroblasts (MEFs) from germline ablation models of these genes. $Lpar6^{-/-}$ MEFs dysregulated the expression of 5,639 genes (2,741 upregulated and 2,898 downregulated), whereas loss of *Cab39l* dysregulated the expression of 1,106 genes (514 upregulated and 592 downregulated).

Among the top upregulated genes in $Lpar6^{-/-}$ cells were *Afp* and *Dkk3*. The top downregulated genes included *Pcna-ps2*, *Pmaip1*, *Perp*, *Hoxb8*, and *Apc2* (Figure S9A). $Lpar6^{-/-}$ MEFs exhibited activation of luminal differentiation, resulting in a shift from negative to positive BLT scores and activation of unfolded protein response (Figures S9B–S9E).

$Cab39l^{-/-}$ MEFs had upregulation of the *Arhgap22*, *Hbefg*, *Il1rl1*, and *Msx2* genes, among others (Figure S9F). The top downregulated genes in the $Cab39l^{-/-}$ MEFs included *Afp*, *IiGP1*, *Wt1*, *Hoxc10*, and *Ifit3* (Figure S9F). The $Cab39l^{-/-}$ MEFs exhibited suppression of the luminal differentiation program, downregulation of unfolded protein response, upregulation of basal markers, and a shift from positive to negative BLT scores (Figures S9F–S9J). $Lpar6^{-/-}$ and $Cab39l^{-/-}$ MEFs did not have any significant differences in the expression patterns for genes involved in cholesterol synthesis. Although the top dysregulated genes in $Lpar6^{-/-}$ and $Cab39l^{-/-}$ MEFs differed from those observed in bladder cancer cell lines with CRISPR ablation of these genes, the overall biologic effects were similar.

Effects of Lpar6 *and* Cab39l *loss on mouse bladder urothelium*

The effects of *Lpar6* and *Cab39l* loss of function on mouse bladder urothelium were consistent with their role in the urothelial differentiation program. In both instances, they induced urothelial hyperplasia, which exhibited distinctive features. These features could be documented using light and electron microscopy as well as immunohistochemically by showing the expression pattern for two signature basal (KRT14) and luminal (KRT18)

keratins (Figures 4A–4D). *Lpar6*$^{-/-}$ bladder urothelium exhibited hyperplasia, with cells undergoing a gradual transition from basal through intermediate to luminal urothelial layers reminiscent of that in *Lpar6*$^{+/+}$ urothelium. In contrast, the urothelial hyperplasia in *Cab39l*$^{-/-}$ urothelium resulted in distinct expansion of the basal undifferentiated layer.

To provide more in-depth insight into the biologic effects of *Lpar6* and *Cab39l* loss of function on the urothelium, we performed single-cell RNA sequencing of urothelial cell suspensions harvested from control NU wild-type (NU-wt), *Lpar6*$^{-/-}$, and *Cab39l*$^{-/-}$ mice. Because we did not have enough cells from an individual mouse bladder to perform single-cell sequencing, we pooled cells from the three groups of samples listed above from 15 animals. After quality control and filtering of a total of 6,117, 6,677, and 7,503 cells from NU-wt, *Lpar6*$^{-/-}$, and *Cab39l*$^{-/-}$ mice, respectively, we segregated via unsupervised clustering and visualized using uniform manifold approximation and projection (UMAP). We annotated the cell clusters according to the expression of marker genes known to be differentially expressed in basal intermediate and urothelial cells and of marker genes for various stromal and immune cells (Figures S10A and S10B). Because our focus was on the biologic effects of *Lpar6* and *Cab39l* ablation on urothelial cells, we filtered out nonurothelial cells. This provided 4,693, 6,254, and 7,012 high-quality urothelial cells in the three sets of samples for downstream analysis. We then reclustered the pure urothelial cells and again visualized them using UMAP. To preserve the identity of cell clusters, we initially performed unsupervised clustering with the combined samples from the three groups and then visualized NU-wt, *Lpar6*$^{-/-}$, and *Cab39l*$^{-/-}$ separately (Figure 4E). The expression levels for the top differential markers across the *Lpar6*$^{-/-}$ and *Cab39l*$^{-/-}$ urothelial cell clusters are shown in Figures S11A–S11C. The urothelial cells were segregated into 18 clusters, with each cluster representing a different cell population. In analyzing the expression patterns for the basal and luminal markers, we could separate the clusters into three major compartments corresponding to basal, intermediate, and luminal urothelial cells (Figures 4F–4H). The dominant cell population in NU-wt mice corresponded to intermediate cells, accounting for about 60% of the cell total. Ablation of Lpar6 reinforced luminal differentiation, and in *Lpar6*$^{-/-}$ animals, luminal cells were dominant, accounting for more than 60% of the urothelial cells. In contrast, *Cab39l* loss of function was associated with expansion of the basal compartment, which accounted for about 70% of the cells (Figure 4H). These results were in unity with quantitative assessment of urothelial differentiation according to the BLT score. The BLT scores demonstrated dominance of cells with positive BLT scores in *Lpar6*$^{-/-}$ animals and that of cells with negative BLT scores in *Cab39l*$^{-/-}$ animals (Figure 4I).

Consistent with the changes in the overall luminal and basal differentiation patterns *in vitro* induced by the ablation of *Lpar6* and *Cab39l*, single-cell RNA sequencing revealed major changes in the activation of urothelial regulons (Figures S12A and S12B). *Lpar6*$^{-/-}$ cells exhibited expansion of the activated regulons in luminal clusters, whereas *Cab39l*$^{-/-}$ cells exhibited expansion of the activated regulons in basal clusters. Because our *in vitro* studies showed that the biologic effects of *Lpar6* and *Cab39l* ablation were mediated by the unfolded protein response and cholesterol synthesis, we analyzed their activation patterns using single-cell sequencing, which demonstrated that in *Lpar6*$^{-/-}$ cells, genes of the unfolded protein response were activated in luminal clusters (Figures S13A and

S13B). In contrast, *Cab39l*$^{-/-}$ cells exhibited activation of the unfolded protein response in basal clusters. We observed a somewhat similar pattern of dysregulation in genes involved in cholesterol synthesis, as *Lpar6*$^{-/-}$ cells exhibited activation of cholesterol synthesis in luminal clusters, whereas *Cab39l*$^{-/-}$ cells exhibited activation of it in basal clusters.

To better analyze the effect of *Lpar6* and *Cab39l* on urothelial differentiation, we compared the differentiation trajectories of NU-wt, *Lpar6*$^{-/-}$, and *Cab39l*$^{-/-}$ cells (Figure S14A). Similar to previously published data, the trajectories of urothelial cell differentiation were not on a single path.[23] Instead, we identified four distinct trajectories in NU-wt cells and three distinct trajectories in *Lpar6*$^{-/-}$ and *Cab39l*$^{-/-}$ cells. Surprisingly, in NU-wt cells, the major trajectories (paths 1–3) represented basal-to-intermediate transitions, with only a small proportion of basal cells progressing through intermediate to terminally differentiated luminal cells (path 4). Consistent with the enhancement of luminal differentiation, *Lpar6*$^{-/-}$ cells exhibited all three paths (paths 1–3) progressing from basal-to-luminal differentiation. In contrast, *Cab39l*$^{-/-}$ cells had a major transition (path 1) progressing from basal-to-intermediate differentiation, whereas paths 2 and 3 exhibited progression to terminal luminal differentiation. In NU-wt cells, paths 2 and 3 had the highest activity of urothelial regulons. In both *Lpar6*$^{-/-}$ and *Cab39l*$^{-/-}$ cells, path 2 had the highest activity of urothelial regulons (Figure S14B). Furthermore, *Lpar6*$^{-/-}$ and *Cab39l*$^{-/-}$ cells had distinctive involvement of molecular pathways and gene ontologies across their differentiation trajectories (Figure S14C). Ablation of both genes changed the intercellular communication networks in the urothelium (Figures S14D, S15A, and S15B). *Lpar6*$^{-/-}$ cells exhibited activation of Egfr and Notch intercellular communication networks among intermediate cells. *Cab39l*$^{-/-}$ intermediate cells retained the active Notch pathway but had reduction of Egfr activity. We observed similar patterns of activation of the Notch pathway in the *Lpar6*$^{-/-}$ and *Cab39l*$^{-/-}$ basal cells. Overall, the number of activated intercellular communication networks in both *Lpar6*$^{-/-}$ and *Cab39l*$^{-/-}$ cells was increased (Figure S15B).

Single-cell-sequencing-based analysis of the cell cycle demonstrated greater proliferation activity in the *Lpar6*$^{-/-}$ and *Cab39l*$^{-/-}$ cells than in the NU-wt cells[24] (Figure S16A). The distribution of cells with high G2/M-phase and S-phase scores in *Lpar6*$^{-/-}$ and *Cab39l*$^{-/-}$ cells was different from that in NU-wt cells (Figure S16B). In both instances, we observed expansion of proliferating cells predominantly in the basal and intermediate cell clusters. Analysis of the G2/M-phase scores across the cell trajectories demonstrated that path 3 in both *Lpar6*$^{-/-}$ and *Cab39l*$^{-/-}$ cells had the highest proliferative activity (Figure S16C). The distribution of signature luminal (*Upk3a*) and basal (*Krt15*) markers demonstrated that their expansion and high expression levels were consistent with the increased luminal differentiation in *Lpar6*$^{-/-}$ cells and expansion of undifferentiated basal *Cab39l*$^{-/-}$ cells (Figure S16D).

Effects of Lpar6 *and* Cab39l *loss on BBN-induced mouse bladder cancers*

BBN exposure in *Lpar6*$^{-/-}$ and *Cab39l*$^{-/-}$ mice induced distinct forms of intraurothelial precursor lesions. The *in situ* precursor lesions in *Lpar6*$^{-/-}$ animals induced carcinoma *in situ* that exhibited retention of KRT5/6 basal cells at the base of the urothelium and were consistent with the so-called luminal subtype of carcinoma *in situ* (Figure 5A). In contrast,

after exposure to BBN, *Cab39l*$^{-/-}$ animals initially experienced development of carcinoma *in situ* that was characterized by the proliferation of undifferentiated KRT5/6-positive cells, consistent with the basal variant of carcinoma *in situ* (Figure 5B). The loss of function of both genes sensitized urothelium to BBN-induced bladder cancers, and in both instances, the loss of function was associated with the development of significantly higher numbers of tumors of distinct pathogenetic and molecular subtypes than in control wild-type animals (Figures 5C and 5D). *Lpar6*$^{-/-}$ animals had 79 tumors (45 low-grade superficial noninvasive papillary and 34 high-grade invasive urothelial carcinomas), whereas *Lpar6*$^{+/+}$ animals had only 17 tumors (2 low-grade superficial noninvasive papillary and 15 high-grade invasive urothelial carcinomas). All superficial papillary tumors were of the luminal subtype and were strongly positive for luminal KRT20. Invasive carcinomas in *Lpar6*$^{-/-}$ mice were also positive for KRT20, and RNA-sequencing analysis of their three representative tumors revealed their positive BLT scores consistent with the luminal phenotype (Figure 5E). *Cab39l*$^{-/-}$ animals had 24 tumors (1 low-grade superficial noninvasive papillary and 23 high-grade invasive urothelial carcinomas), whereas *Cab39l*$^{+/+}$ had only 9 tumors, all of which were high-grade invasive urothelial carcinomas (Figures 5C and 5D). Notably, the invasive carcinomas in *Cab39l*$^{-/-}$ mice were all poorly differentiated, with focal squamous features recapitulating the histologic hallmarks of human basal urothelial carcinomas of the bladder. Invasive carcinomas in the *Cab39l*$^{-/-}$ animals were strongly positive for basal KRT14, and RNA-sequencing analysis of three representative tumors demonstrated negative, low BLT scores consistent with their basal phenotype (Figures 5F and 5G). Overall, the numbers of tumors in both *Lpar6*$^{-/-}$ and *Cab39l*$^{-/-}$ were significantly much higher than in their respective wild-type controls ($p < 0.001$). These results demonstrated that loss of both *Lpar6* and *Cab39l* sensitized the urothelium to BBN-induced carcinogenesis. These experiments also supported the concept that loss of *Lpar6* sensitizes the bladder urothelium to primarily luminal carcinogenesis, whereas loss of *Cab39l* sets carcinogenesis on the basal track after exposure to the tobacco-related carcinogen BBN.

Expression of LPAR6 *and* CAB39L *target genes in human bladder cancers*

To validate the involvement of *LPAR6* and *CAB39L* in human bladder cancer development, we analyzed the expression of their target genes in two large publicly available bladder cancer cohorts. Specifically, we analyzed the expression patterns for upregulated and downregulated genes identified in *Lpar6*$^{-/-}$ and *Cab39l*$^{-/-}$ MEFs in the TCGA ($n = 408$) and Lund University ($n = 308$) bladder cancer cohorts classified according to luminal and basal taxonomy (Figures S17A, S17B, S18A, and S18B). Analysis of upregulated and downregulated *LPAR6* and *CAB39L* target genes demonstrated that in both cohorts, luminal tumors were characterized by an *LPAR6*-depleted phenotype. On the other hand, basal tumors had a *CAB39L*-depleted phenotype. These observations confirmed the key role of *LPAR6* and *CAB39L* in the development of luminal and basal human bladder cancers, respectively.

DISCUSSION

In this report, we present evidence that FR gene inactivation results in clonal expansion of cellular plaques in the bladder mucosa, contributing to the initiation of bladder

tumorigenesis. The FR genes that we identified are located near the established tumor suppressor *RB1*, but they were inactivated much earlier in tumor progression than was *RB1*. Multiple lines of evidence, including that obtained via genetic mapping and mutational and methylation analyses, support the concept that FR gene inactivation plays a contributory role in early incipient phases of bladder tumorigenesis. Functional studies of selected FR genes such as *LPAR6* and *CAB39L* established that their inactivation produces effects on cell proliferation and differentiation that may account for the clonal expansion that occurs as a result of field effects in the initiated urothelium.

The FR gene concept postulates at least three sequential steps involving four "hits" that lead to inactivation of an FR gene and its neighboring tumor-suppressor gene during cancer initiation and progression (Figure 6A). Tumorigenesis begins with silencing of the FR gene, most often by somatic homozygous hypermethylation (hits 1 and 2) or a combination of loss of one copy and hypermethylation of the remaining FR gene allele. The loss of one FR gene copy is usually synchronous with the loss of a neighboring tumor suppressor, such as *RB1* (hits 2 and 3). The second step may have several substeps in which additional FR genes are homozygously inactivated. Available evidence suggests that, in rare instances, steps 1 and 2 can be reversed, as inactivation of the first FR gene allele can be accomplished by germline nucleotide substitutions, including population-based polymorphisms.[5,6] Homozygous inactivation of FR genes is associated with clonal expansion of the *in situ* preneoplastic clone. In the third and final step, the remaining allele of a contiguous tumor suppressor is inactivated, most commonly by a mutation (hit 4). This step is associated with clonal evolution into the transformed phenotype with features of carcinoma *in situ* progressing to invasive bladder cancer. This scenario represents an oversimplification of FR genes' involvement in tumorigenesis. In reality, multiple FR genes not necessarily located around the *RB1* gene interact with tumor suppressors and oncogenes throughout the genome. Our recent whole-organ genomic studies demonstrated that about 100–200 genes are involved in the development of mucosal field effects that initiate bladder carcinogenesis and that these effects overwhelmingly result from epigenetic modifications. These alterations dysregulate more than 100 pathways involving immunity, differentiation, and proliferation.[25,26] Whole-organ-scale analysis and time modeling identified the three major waves of mutations distinctively involved in the evolution of bladder cancer from field effects.[25,26] The most frequent were low-frequency mutations involving small mucosal areas that targeted individual uroprogenitor cells. They gradually developed over several decades. Another group consisted of low-frequency mutations that expanded across the bladder mucosa and signified the advent of the progressive phase of urothelial carcinogenesis, which lasted about 5 years. Finally, the progression to invasive disease was driven by a group of high-frequency, clonally expanding progressive mutations during the last 2–3 years of bladder carcinogenesis. These data clearly indicate that the biologic effects of FR genes should be perceived in the context of complex alterations of the mucosal environment.

Of the five candidate FR genes (*ITM2B*, *LPAR6*, *MLNR*, *CAB39L*, and *ARL11*) identified by expression, mutation, and methylation analyses combined with whole-organ mapping, we focused our functional studies on *LPAR6* and *CAB39L*, which appeared to be distinctively involved in the luminal and basal subsets of bladder cancer. The *LPAR6* gene is located within intron 17 of the *RB1* gene in the reverse orientation and encodes a 39-kDa G-protein-

coupled receptor that binds the lipid signaling molecule lysophosphatidic acid.[27,28] Several factors prompted us to select *LPAR6* for ablation in cell lines and develop its germline knockout in a mouse model. First, among the candidate FR genes, *LPAR6* was the only one silenced by missense mutations, and an *LPAR6* polymorphism was a strong risk factor for bladder cancer in smokers.[5,6] In addition, germline mutations of *LPAR6* were associated with a novel familial cancer syndrome.[6] Other groups demonstrated that germline mutations of *LPAR6* were associated with the rare familial woolly hair autosomal recessive syndrome.[28,29] *LPAR6* also has been shown to be involved in other cancer types, including hepatocellular, pancreatic, and colon cancers.[30–35] Of note, *LPAR6* has protumorigenic oncogene-type effects on hepatocellular and pancreatic cancers,[30,31,33–35] whereas it has an antitumorigenic tumor-suppressor effect on colon and breast cancers.[32,35] In comparison, the *CAB39L* gene is located about 3 Mb downstream of *RB1* and encodes 39 kDa, which is a core component of the LKB1 tumor-suppressor complex that activates a subset of serine/threonine protein kinases.[36–38] Researchers have implicated *CAB39L* as having a role in the development of several solid and hematopoietic cancers, including gastric, colorectal, and renal carcinomas, as well as acute monocytic leukemia, among others.[39–41]

We showed that *LPAR6* and *CAB39L* play a role in urothelial differentiation and are the upstream regulators of luminal (GATA3 and PPARγ) and basal (p63) transcription factors. Their silencing dysregulates the basal-to-luminal differentiation program and causes distinct urothelial hyperplasia of luminal and basal types, respectively (Figure 6B). Silencing of *LPAR6* sensitizes the urothelium to the development of BBN-induced luminal bladder cancer, whereas silencing of *CAB39L* sensitizes the urothelium to the development of BBN-induced basal bladder cancer. Validation of these observations using *in vitro* and *in vivo* models in major human bladder cancer cohorts demonstrated that luminal tumors are characterized by an *LPAR6*-depleted phenotype, whereas basal tumors have a *CAB39L*-depleted phenotype. This confirmed the major role of these genes in human bladder cancer development.

Although recent whole-genome characterizations provided remarkable insight into the genetic and epigenetic characteristics of established cancers, much less is known about the earliest events that initiate human carcinogenesis. Past attempts to characterize field effects that produce these early changes tended to employ candidate gene approaches that identified alterations of known oncogenes and tumor suppressors.[42–45] More recent genomic studies of field effects initiating bladder carcinogenesis disclosed a complex mutational landscape with remarkable heterogeneity across clones and individuals. The majority of these mutational changes were of low variant allele frequency and were unlikely to contribute to the clonal expansion of preneoplastic clones.[46–49] Our whole-organ mapping studies demonstrated that, although the mutational landscape in microscopically normal bladder urothelium is complex, only a small proportion of the mutations participate in clonal expansion, and the background field change initiating bladder carcinogenesis is predominantly epigenetic.[25,26] Furthermore, work by others has indicated that early events of carcinogenesis also involve metabolic dysregulation, such as glycolysis and oxidative phosphorylation, resulting in the selection of early clonal expansion with increased lactic acid fermentation and the Warburg phenotype.[50–52]

Our results support the conclusion that field effects involving the entire mucosal membrane underlie the growth advantage of the initiated urothelial clone that triggers carcinogenesis in the bladder. However, the molecular targets of these field effects are not the major tumor-suppressor genes themselves but rather the FR genes located in close proximity to them. This structural organization suggests that the FR genes have conserved biological functions within the pathways controlled by the tumor-suppressor genes they are adjacent to.

Limitations of the study

The limitations of this study are related to our rudimentary understanding of the complexity of the field effects initiating bladder carcinogenesis and the exact roles of *LPAR6* and *CAB39L* in the maintenance of a mucosal microenvironment. The biologic effects of *LPAR6* and *CAB39L* on the urothelial differentiation program are mediated by unfolded protein reactions and cholesterol synthesis, but the exact links between the FR genes and these molecular pathways are unknown. Although we postulate that silencing of FR genes occurs before dysregulation of tumor suppressors and oncogenes in early phases of carcinogenesis, the synergistic biologic effects of FR genes, tumor suppressors, and oncogenes are uncertain. In this study we focused on the two prototypic genes, *LPAR6* and *CAB39L*, and the roles of other candidate FR genes mapping around *RB1* are unknown. In addition, it is very likely that FR-like genes exist in other parts of the genome, and they may contribute to the development of bladder cancer by dysregulating other aspects of the mucosal microenvironment. The development of double-knockout or transgenic models involving FR genes and tumor suppressors or oncogenes is needed to determine their synergistic biologic effects. Such studies, combined with more detailed characterization of the molecular microenvironments of the mucosal field effects, will define the complexity of molecular mechanisms of incipient phases of bladder carcinogenesis, providing targets to intercept the process in its early phases and preventing the development of invasive, clinically aggressive cancer.

STAR★METHODS

RESOURCE AVAILABILITY

Lead contact—Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Bogdan A Czerniak (bczernia@mdanderson.org).

Materials availability—Knock-out strains and reagents generated in this study will be shared by the lead contact upon request.

Data and code availability

- This paper analyzes existing, publicly available data. These accession numbers for the datasets are listed in the key resources table. Summarized statistic results data reported in this paper will be shared by the lead contact upon request.
- This paper does not report original code.

- Any additional information required to reanalyze the data reported in this work paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Cell lines and culture conditions—All UM-UC cells, RT-4 (RRID:CVCL_0036), RT-112 (RRID:CVCL_1670), Scaber (RRID:CVCL_3599), and all knock-out cell lines cultured in MEM medium, were maintained as previously described.[5,6] TERT-NHU (RRID:CVCL_JX41) was obtained from Dr. Margaret A Knowles (University of Leeds) and were maintained in KSFM supplemented with the supplied bovine pituitary extract and epidermal growth factor plus 30 ng/ml cholera toxin in a humidified atmosphere of 5% CO_2 at 37°C.[57,58]

All cell lines were confirmed to be mycoplasma-free.

Mice—All animal experiments were conducted in accordance with the guidelines of the Institutional Animal Care and Use Committee under approved protocols (898-RN04). Prior to the start of cancer-related experiments, the mice were in good health, exhibiting bright, alert, and responsive behavior. The mice were housed in an animal facility equipped with an Individually Ventilated Caging system, which maintained a regulated temperature of approximately 22°C and a humidity level of 45 ± 10%. The mice had free access to water and food and were maintained on a 12-h light/dark cycle (lights on from 06:30 to 18:30). Male and Female C57BL/6J mice were purchased from the National Laboratory Animal Center. Lpar6 knock-out mouse was generated by electroporating the targeting vector, while the knock-out first allele was purchased for generating Cab39L knock-out mouse. Bladder tumors on mice at the age of 8 weeks were induced by the treatments with BBN. At the end of the experiment, mice were euthanized by carbon dioxide asphyxiation.

Human samples—Human samples and clinical data were collected and archived according to a laboratory protocol approved by the Institutional Review Board of The University of Texas MD Anderson Cancer Center. Whole-organ histologic and genetic mapping was performed using radical cystectomy samples from 20 patients with a mean age of 69 years (range, 47.0–86.0 years) who had high-grade muscle-invasive (T3) UC (Table S1).

METHOD DETAILS

Human tissue samples—All human tissues used in this study were collected, processed, and labeled under protocols reviewed and approved by the Institutional Review Board of The University of Texas MD Anderson Cancer Center. Urothelial carcinomas (UC) were classified according to the histologic tumor grading system of the World Health Organization[59] and were dichotomized as low-grade or high-grade tumors. The growth pattern of papillary versus nonpapillary or solid tumors and the depth of invasion were also recorded. Levels of invasion were defined according to the TNM staging system.[60] T_1 tumors were substaged as T_{1a} or T_{1b} to dichotomize them as superficial (T_a -T_{1a}) or invasive (T_{1b} and higher).[61,62] The mucosal samples were histologically classified as containing: NU, normal urothelium; MD, mild dysplasia; MdD, moderate dysplasia;

SD, severe dysplasia; CIS, carcinoma *in situ*; and UC. For analytical purposes, the precursor intraurothelial lesions were dichotomized into two groups designated as low-grade intraurothelial neoplasia (LGIN) by combining samples with features of MD and MdD and as high-grade intraurothelial neoplasia (HGIN) by combining samples with features of SD and CIS as previously described.[9]

The following tissue specimens and their respective patient cohorts as well as bladder cancer cell lines were used in this study: (1) whole organ cystectomy specimens from patients with bladder cancer (n=20); (2) fresh frozen primary tumor samples and paired peripheral blood DNA from patients with bladder cancer (n=189), (3) TCGA bladder cancer cohort (n=408), (4) Lund bladder cancer cohort (n=308); and (5) bladder cancer cell lines (n=30).[15] The clinical and pathological data of cystectomy specimens used for WOHGM are summarized in Table S1. The clinical and pathological data for the MDACC, TCGA, and Lund cohorts are summarized in Table S2.

Whole-organ mapping—Freshly collected radical cystectomy specimens from 20 patients with untreated sporadic high-grade invasive UC were used for whole-organ histologic genetic mapping (WOHGM). The patients comprised one woman and 19 men, ranging in age from 47 to 86 years (mean 69.0 years ± 10.9 standard deviation [SD]). These specimens provided 900 DNA samples corresponding to areas of bladder mucosa with the following microscopic characteristics: microscopically normal-appearing urothelium (NU; n=227), LGIN (n=282), HGIN (n=171), and invasive UC (n=220) (Table S1). WOHGM was carried out by allelotyping 661 SNPs that mapped within a 27-Mb segment around *RB1* of 234 DNA samples from five cystectomy specimens (maps 1–5) as previously described in detail.[5,6,9] In brief, each fresh cystectomy specimen was opened longitudinally along the anterior wall of the bladder and pinned down to a paraffin block. The entire bladder mucosa was divided into 1-cm × 2-cm samples, which were frozen in OCT compound. Several 5-μm sections were cut from each frozen block. One frozen section from each block was stained with hematoxylin and eosin to evaluate the distribution of *in situ* lesions and bladder cancer microscopically. The remaining unstained sections were stored at −70°C until used for immunohistochemical studies.

For DNA extraction, the mucosal samples were defrosted and washed three times in phosphate-buffered saline solution (PBS). The mucosal surface of each sample was scraped with a razor blade, and the scraped cells were transferred into a conical tube containing PBS. In samples with invasive bladder cancer, the tumor was dissected from the frozen block to minimize contamination with non-tumor cells. Only samples that yielded more than 90% microscopically recognizable intact urothelial tumor cells were used for extraction. DNA extracted from peripheral blood lymphocytes and/or from normal tissue in the resected specimen of each patient was used as a control. Allelotyping of SNPs was performed and analyzed using an automated pyrosequencing instrument, PSQ96MA (Biotage AB). The minimal deleted region associated with clonal expansion of *in situ* neoplasia was identified by the alignment of clonal plaque-like deleted segments from individual cystectomy specimens.

Analysis of the transcriptional activity around RB1 by tiling arrays—The transcriptional activity in the *RB1* region was analyzed on nine bladder cancer cell lines and 22 fresh frozen samples of bladder cancer using a custom designed NimbleGen tiling expression array. The patients comprised five women and 17 men (mean age 68.6 years ± 11.5 SD), and the specimens constituted 10 low-grade and 12 high-grade tumors comprising 13 superficial (T_a –T_{1a}) papillary urothelial carcinomas and nine invasive (T_{1b} and higher) carcinomas. The Roche NimbleGen Tiling Array containing 385,000 tiling expression probes covering the 5-Mb segment around *RB1*, from 45.10 Mb to 50.17 Mb, was used to analyze the expression patterns of the *RB1* gene and its flanking segments. Total RNA samples were transformed into cDNA using the SuperScript Double-Stranded cDNA Synthesis Kit (Invitrogen, Inc.). The arrays were scanned using a GenePix 4000B microarray scanner (Molecular Devices, Inc.), and the raw data were initially filtered by a signal-to-noise subtraction algorithm. Expression levels of individual exons of all annotated genes in the region were compared to their respective controls. The expression levels of bladder tumor samples were compared to the expression levels of normal urothelial cells obtained from ureters from nephrectomy specimens, while the expression levels of bladder cancer cell lines were compared to those of TERT-NHUC cells. Quantile normalization was initially applied to align values across all arrays, and the signals from the probes corresponding to exons of each gene were summarized into a curve as follows: the threshholded log2 values after subtracting the background were used to define weights for a kernel fit. For each tumor sample, we computed the difference between the tumor curve and the average of the control curves. Within each exon, we estimated the kernel density at the center of the exon and then recorded the largest magnitude exon value as the summary for the gene. The final data were presented as the mean relative expression levels of exon-based probes of all annotated genes in the region and were summarized as a heatmap.

Methylation analyses of *ITM2B, LPAR6, MLNR, CAB39L and ARL11*—Methylation analysis was performed on the MDACC cohort of 189 fresh frozen bladder tumor samples. We used quantitative methylation-specific PCR to identify whether methylation of CpG islands in the putative promoter regions of *ARL11*, *MLNR*, and *ITM2B* played a role in their silencing. Methylation studies were performed on 189 bladder tumor samples included a cohort of 111 samples analyzed for LOP around *RB1* by SNPlex technology, and 20 whole-organ cystectomy specimens. Normal human urothelial cells from nephrectomy specimens served as normal controls. For whole-organ cystectomy specimens, the methylation status was first tested on representative tumor samples; when an increased level of methylation was identified, the remaining mucosal samples of the entire cystectomy map were tested.

The bisulfite treatment of DNA and design of the assays were performed as previously described.[63,64] In brief, we analyzed regions around the reading frames of *ARL11*, *MLNR*, and *ITM2B* and identified their potential CpG islands (Figure S4A). For *ARL11*, we concentrated our analysis on CpG3 located near the transcription initiation site of exon1, as previously published data indicated that it controlled the expression of this gene.[65] First, we amplified fragments of bisulfite-treated DNA containing the sequences of the CpG islands in the promoter regions of *ARL11*, *MLNR*, and

ITM2B using the following sets of primers: TTGAGTTGTTA AAAAATGGAGTTG (sense) and CATCCCTACTCCCAAAAATAAA (antisense) for *ARL11*, TAGTTAGTGGAGAGGGAAGYGTT (sense) and CAAACACAAACACACAACRATC (antisense) for *MLNR*, and GGTGTTGTTAAAGGTTAGTTTTTTGT (sense) and CRCAACCTC CCAACTCCC (antisense) for *ITM2B*. We used bladder cancer cell lines and control urothelium as PCR templates for bisulfite-treated DNA sequencing. We also validated the methylation status of CpG islands in selected cancer cell lines by DNA sequencing of bisulfite-treated DNA after subcloning into the TA vector pCR 4-TOPO (Invitrogen) (Figure S4B).

The amplified bisulfite-modified DNA fragments used for sequencing were subjected to methylation-specific quantitative PCR using two methylation-specific minor groove-binder Taq probes. Probes specific for methylated and unmethylated sequences of the *ARL11, MLNR*, and ITM2B are listed in Table S4. The amplifications of methylated and unmethylated alleles were monitored for each sample during the same RT-PCR reaction. For quantification of the ratio of methylated and unmethylated alleles, we calculated ΔC_T = C_T of methylated allele – C_T of unmethylated allele. The ΔC_T values were compared to the standard curves produced by titration of fully methylated control DNA (placental DNA after M.Sssl methylase (New England Biolabs) treatment) and unmethylated control placental DNA. For each gene, we generated nine samples with defined ratios of methylated and unmethylated target DNA sequences (Extended Data Figure 2C). To verify that the results were dependent on DNA concentration, we tested several concentrations of DNA and verified that ΔC_T was almost unchanged. The ΔC_T value of each sample was calculated and methylation ratio was deduced from the standard curve.

We used sequencing analysis of bisulfite treated DNA to check the methylation status of CpG site (cg16071219) located in the body of 1st exon of *LPAR6* and of CpG site (cg17777592) located in 5'UTR of 1st exon of *CAB39L* (Figure S4A). The chosen cg sites were showing negative correlation between methylation and expression data in the TCGA cohort. We designed the following set of primers to amplify the regions comprising CpG sites: cg16071219: TGTTTGTGTTTGGGTTAATATTT (sense) and ATTTTCAAAACAAACTTCT AAAACA(antisense); cg17777592: AAGGTTTGGAATTTGAGAGTAA (sense) and TCACCAAATATTACAAAAATACCTAA (antisense). We generated nine samples with defined ratios of methylated and unmethylated DNA as described above and after sequencing we created the standard curves using the measurement of the surface under the peak of the cytosine sequence trace (Figures S4C and S4D).

To correlate the methylation status of CpG islands to the expression of *ARL11*, *MLNR*, *ITM2B*, *LPAR6 and CAB39L* we performed quantitative RT-PCR on bladder cancer cell lines and control urothelial tissues using TaqMan gene expression assays (Figure S4E).

To verify whether the expression of *ARL11*, *MLNR*, *ITM2B, LPAR6 and CAB39L* could be restored by demethylation of their promoters, selected cell lines were treated with various concentrations of 5-aza-2′-deoxycytidine or PBS as a control for 3 days, and the expression levels of all three genes were tested by quantitative RT-PCR (Figures S4F and S4G).

Mutational analysis of candidate FR genes—To identify whether nucleotide substitutions may represent the potential cause of FR gene silencing we analyzed their mutational data in the TCGA cohort (Figure S6A). It showed that *ITM2B, MLNR, CAB39L*, and *ARL11* are mutated in 1% or even less of human bladder cancers. Among the five candidate FR genes *LPAR6* showed nucleotide substitutions causing the amino acid substitution and potentially altering the function of the gene could be detected in 2% of bladder cancer tumor samples. This data provided the preliminary information that in a small fraction of bladder cancers *LPAR6* may be silenced by nucleotide substitutions and provided incentive for its sequencing in the MDACC cohort comprising of 189 bladder tumor samples on which additional methylation studies of all five FR genes were performed. In addition, on 111 bladder tumor samples of the same cohort we performed the LOP analysis around *RB1* by SNPlex technology. For 20 whole-organ cystectomy specimens, a representative tumor sample was initially tested for the presence of mutations and then, if mutations were present, the testing was performed on all remaining samples of the map. Sequencing was performed on PCR-amplified DNA samples using exon-specific primers, and sequencing reactions were analyzed with an ABI 3730 sequencer (Applied Biosystems, Inc.). Mutations were identified by using the Sequencer software (Gene Codes Corporation) and were confirmed by visual inspection.

Testing of LOP in the *RB1* region by SNPlex technology—The loss of polymorphism (LOP) around *RB1* was tested by SNPlex on fresh frozen bladder tumor samples from 111 patients comprising 86 men and 25 women (mean age 66.7 years ± 10.5 SD). There were 35 low-grade and 76 high-grade tumors comprising 39 superficial papillary urothelial carcinomas (T_a –T_{1a}) and 72 invasive (T_{1b} and higher) carcinomas. The allelotyping of 124 SNPs on chromosome 13 that mapped between the *SPERT* and *KCNRG* genes and spanning approximately 5 Mb around *RB1* was carried out by SNPlex according to the manufacturer's protocol (available at http://docs.appliedbiosystems.com/pebiodocs/04360857.pdf). Allelotypes were analyzed using the Prism 3730 Sequencer and Genemapper 3.5 software (Applied Biosystems, Inc.).

RNA sequencing and data analysis—The RNA sequencings were performed on the samples of *Lpar6* KO UC6 and *Cab39L* KO UC7 cell lines, *Lpar6* KO and *Cab39L* KO mouse embryonic fibroblast cells, and BBN-induced mouse bladder tumors in *Lpar6* KO and *Cab39L* KO mice. The RNA integrity was assessed using a 2100 Bioanalyzer (Agilent Technologies). The RNA concentration was determined using RiboGreen quantification (Quant-iT RiboGreen RNA Assay Kit; Invitrogen). RNA samples meeting a quantity threshold of 1 μg and with an RNA integrity number of at least 7 were analyzed in the Advanced Technology Genomics Core at MD Anderson. Prior to RNA library construction, ribosomal RNA was removed from total RNA preparations, and cDNA synthesis using oligo d(T) and random hexamers was performed. The library was made up of random fragments representing the entire sample. It was created by shearing DNA into 150–400 base fragments that were ligated to specific adapters. Following a sample cleanup step, the resulting library was quantified using quantitative PCR and checked for quality using a TapeStation (Agilent Technologies). The analyses were performed with 30 RNA samples (12 samples from cell lines, 12 samples from mouse embryonic fibroblasts, and 6 samples from

BBN-induced bladder tumors). Quality control for RNA was conducted using RSeQC[53] and FastQC software. Sequencing reads were aligned to the GRCh38 reference genome using STAR (version 2.7.3a)[66] with GENCODE (release 32) transcript annotations.[67] Read counts for individual genes were obtained using featureCounts software in the Subread package.[54] On average, 27 million reads per sample associated with more than 58,000 unique genes, both coding and noncoding, were obtained. From this set of genes, 44,000 with more than 10 reads in at least one sample were selected and used in the subsequent analyses. Genes differentially expressed in specific sample groups were identified using DESeq2 software (version 1.26.0)[55] with the Wald test by using a design formula that included batch effect correction. Benjamini-Hochberg correction was used for multiple testing in all instances[68]. The normalized read counts were transformed into $\log_2$ ratios between the signal of KO samples and the signal of wild type control samples.

To evaluate the molecular subtypes of bladder cancer tumor samples and bladder cancer cell lines we analyzed the mRNA expression in TCGA and Lund cohorts as well as 30 bladder cancer cell lines as previously described. [15–17] In brief, for assessment of the luminal and basal phenotypes, the expression levels for 14 luminal and 9 basal marker genes identified previously were used. For quantitative assessment of these two molecular subtypes of bladder cancer, the previously developed BLT score was used.[17] In brief, for quantitative assessment of the luminal and basal phenotype we used previously developed 28 luminal and 20 basal marker genes. [15–17] Linear discriminant analysis was performed to determine the power of individual markers in identifying molecular subtypes of bladder cancer.[69] The unidimensional BLT score was defined as $\sum W_i * E_i$, where W_i is the negative coefficient of the linear discriminant and E_i is the expression of marker genes. The least absolute shrinkage and selection operator analysis was used to identify the best 16 luminal and 12 basal markers.[70] The TCGA cohort was used as a training set to build a linear discriminant analysis model with the 28 selected genes.

Generation of *LPAR6* and *CAB39L* knockout cell lines using CRISPR/CAS9— The CRISPR/CAS9 knockout cell lines were generated by the Genome Engineering & iPSC center, Washington University in St. Louis. gRNA sequences selected for knocking 5'gRNA and 3'gRNA of *LPAR6* in UM-UC-6 cell lines were TTTCCGCTGGGTTC TTCAACNGG and GTCAATGACCGCATAAACGANGG and these two gRNAs were designed to delete the entire DNA of a single *LPAR6* exon. gRNA sequence selected for knocking out *CAB39L* in UM-UC-7 cell line was ATCCTGTTTATGCTCCTCAANGG and this gRNA was designed to make the indel mutations in the exon 3. The gene alterations of all selected clones were confirmed by sequencing and the absence of the expression of the encoded protein was confirmed by Western blotting.

Generation of *Lpar6* KO and *Cab39l* KO mice— *Lpar6* has only one exon and the entire exon including 5' UTR and 3' UTR was cloned in the targeting vector. The targeting vector was electroporated into 129/Sv mouse ES cells, and recombinant ES cell clones were selected with G418 as described. [71,72] Mouse chimeras were generated by injecting correctly targeted ES clones into C57BL/6J mouse blastocysts. *Lpar6* KO mouse was generated by crossing with Cre-mice (Strain #:006054, The Jackson Laboratory) and the deletion of *Lpar6*

exon was confirmed by Southern blot and the ablation of the encoded protein was confirmed by Western blot.

Cab39L knockout first allele was purchased from the EMMA mouse repository and the exon 6 of *Cab39L* was floxed in this mouse. LacZ reporter gene and neo selection cassette were removed by mating with FLP mice and the *Cab39L* KO mouse was generated by crossing the resulting floxed mice and Cre-mice (Strain #:008454, The Jackson Laboratory). The exon 6 deletion of *Cab39L* was confirmed by Southern blot and the ablation of the encoded protein was confirmed by Western blot.

Tumor Induction in mouse models by N-butyl-N-(4-hydroxybutyl)nitrosamine—*Cab39l* KO and *Lpar6* KO mouse groups were compared to their wild type mice, respectively, for BBN treatments. N-butyl-N-(4-hydroxybutyl)nitrosamine (BBN) was solved in water and administered as a 0.05% solution in drinking water for 12 weeks. After finishing the BBN treatments, water was changed to the normal drinking water. Five mice from each group were sacrificed for the analysis at six time points from the discontinuation of BBN exposure as follows: 1, 2, 4, 8, 12, and 16 weeks. After euthanizing the mice, the abdomen was opened, and the base of bladder was tied with a stitch to prevent liquid leaking. About 200 μl of 10% Formalin was injected into the bladder with 30 gauge needle syringe, and the inflated bladder was harvested, and immersed in the 10% formalin to fix the whole organ for 24 hours. Then the bladder was stored in 70% ethanol solution until use. The frozen tumor of the bladder was prepared by snap freezing the tumor cut with OCT in liquid nitrogen at the harvest and the tissue was stored at −80°C freezer.

Electron microscopy—For scanning electron microscopy cell suspensions were fixed in 3% glutaraldehyde with 2% paraformaldehyde and post fixed in 1% osmium tetroxide in 0.1M cacodylate buffer, pH 7.3. The cells were dehydrated with a graded series of increasing concentrations of ethanol, in increasing concentration of hexamethyldisilazane and air dried overnight. The samples were mounted on double-stick carbon tabs (Ted Pella, Inc., Redding, CA) and coated using Balzer MED 010 evaporator (Technotrade International, Manchester, NH) with platinum alloy and carbon. The samples were examined and imaged in a JSM-5910 scanning electron microscope (JEOL, USA, Inc., Peabody, MA) at an accelerating voltage of 5kV.

For transmission electron microscopy cell suspensions and tissue samples were fixed in 3% glutaraldehyde with 2% paraformaldehyde in 0.1M cacodylate buffer, pH 7.3, post fixed in 1% buffered osmium, and stained *en bloc* with 1% uranyl acetate. The samples were dehydrated in increasing concentrations of ethanol and embedded in LX-112 medium. Ultrathin sections were cut in a Leica Ultracut microtome (Leica, Deerfield, IL) stained with uranyl acetate and lead citrate and examined in a JEM 1010 transmission electron microscope (JEOL, USA, Inc., Peabody, MA) at an accelerating voltage of 80 kV. Images were obtained using AMT Imaging System (Advanced Microscopy Techniques Corp, Danvers, MA).

Spheroid cell culture—UM-UC-6 and UM-UC-7 cell lines were maintained in Minimum Essential Media supplemented with 10% fetal bovine in a humidified incubator under an

atmosphere of 5% CO2 in air. For formation of spheres cells were washed with PBS and plated on the low attachment dishes (Corning CLS3262) in serum free medium Corning 16–405-CV supplemented with 20 ng/ml of EGF (ThermoFisher PHG0313), 20 ng/ml bFGF (ThermoFisher PHG0021) and B27 (ThermoFisher 17504–044). Images, spheroid counts, and measurements of their size were performed with Olympus cellSens or GIMP 2.6.12.

Western blots—The following antibodies were used to detect proteins by Western blots: LPAR6 (Santa Cruz sc-20126), Rb (QED Bioscience 3101–3107), GATA3 (Santa Cruz sc-268), Ppar gamma (cell signaling 2435), p63 (cell signaling 13109), β-actin (Sigma A5441), CAB39L (abcam ab197922), and KRT14 (abcam ab181595).

Immunohistochemistry—The following antibodies were used to visualize luminal and basal proteins of the urothelial cells. The luminal markers included mouse monoclonal antibody against human GATA3 (HG3–31 clone, 1:100 dilution; Santa Cruz Biotechnology Inc., Santa Cruz, CA), KRT20 (Ks20.8 clone, 1:400 dilution, Dako). The basal markers included KRT5/6 (D5/16B4 clone, 1:50 dilution, Dako), and KRT14 (LL002 clone, 1:50 dilution; BioGenex, Fremont, CA. Immunohistochemical stains were performed using the Bond-Max Autostainer (Leica Biosystems, Buffalo Grove, IL). The bound primary antibodies were detected with the visualization reagent linked to a dextran polymer backbone with DAB (3, 3-diaminobenzidine) as a chromogen. Then, the slides were counterstained with Mayer's hematoxylin.

Immunofluorescent staining—Mouse bladders were fixed in 10% Formalin for 24 hours for the paraffin sections according to standard protocols. The fixed tissues were routinely stained with Hematoxylin and Eosin and the double immunofluorescent stainings were performed using antibody pair: cytokeratin 5 (EPR1600Y clone, 1:100 dilution, abcam) and cytokeratin 18 (C-04 clone, 1:100 dilution, abcam), and cytokeratin 14 (1:100 dilution, Covance) and cytokeratin 18 (C-04 clone, 1:100 dilution, abcam). Alexa fluor 488 and 546 conjugated secondary antibodies (1:500 dilution, Thermo Fisher Scientific) were used as green and red fluorochromes, respectively and images were taken using Nikon 80i upright microscope.

Single cell sequencing—Single cell sequencing of the bladder urothelium was performed on cells harvested from wild type, *Lpar6*$^{-/-}$ and *Cab39l*$^{-/-}$ mice. For each group a pooled sample of single urothelial cell suspension harvested from 15 mice was prepared. The bladders were harvested after ligating the urethra and injecting 50 μl of 0.25% Trypsin in PBS. The cells were incubated in Trypsin at 37°C for 30 minutes. The urothelial cell clusters and tissue fragments were removed by passing the cell suspension through the 70μm cell strainer. The red blood cells were removed by Ficoll gradient centrifugation. The collected single cell suspensions were washed twice with PBS containing 0.04% BSA and resuspended in 200μl of PBS. Single cell sequencing was performed at the Advanced Technology Genomics Core (ATGC) of UT MD Anderson Cancer Center. Cell suspensions were assessed for cell concentration and viability using Life Technologies Countess II FL cell counter using 0.4% trypan blue exclusion staining. Samples accepted for single cell sequencing had to have target capture of at least 8000 cells per sample and cell viability

of at least 70% or higher. QC steps after cDNA amplification and library preparation steps were carried out by running ThermoFisher Qubit HS dsDNA Assay along with Agilent HS DNA Bioanalyzer for concentration and quality assessments. Equal amounts of each uniquely-indexed sample library were pooled together. The resultant pool was verified for concentration via qPCR using a KAPA Biosystems KAPA Library Quantification Kit. The samples were sequenced using NovaSeq6000 SP flow cells. The run parameters used were 28 cycles for read 1, 91 cycles for read2, 8 cycles for index1, and 0 cycles for index2 as stipulated in the protocol mentioned above. Raw sequencing data (fastq file) was demultiplexed and analyzed using 10X Genomics Cell Ranger software utilizing standard default settings and the cellranger count command to generate html QC metrics and cloupe files for each sample. The initial analyses were performed by using the cloupe files in 10X Genomics Loupe Browser software.

Raw data comprising 8,085, 8,293, and 9,406 cells, respectively, from pooled samples of 15 mouse bladders from three groups of animals (wildtype, *Lpar6*−/−, and *Cab39l*−/−) were filtered for low quality cells using Seurat v4.[73] Briefly, cells with mitochondrial reads >25% of total reads were filtered out, leaving 6,117, 6,677, and 7,503 high quality cells in the three groups of samples, respectively, for the downstream analysis. For each sample group, the gene expression count data was normalized and scaled. Since there was no sequencing technical variation between the three groups, we merged all samples without batch correction. The merged data was further normalized and scaled. Using the 3,000 top-most variable genes of the merged data we computed principal components (PCs). The top 18 PCs were used to cluster the cells at resolution 0.6, which were visualized by Uniform Manifold Approximation and Projection (UMAP). Clustering optimization was performed by a shared nearest neighbor (SNN) modularity optimization based clustering algorithm. The clusters were annotated by marker gene expression known to be differentially expressed in basal, intermediate and luminal urothelial cells. Additional markers were used to identify various types of stromal and immune cells. Since we were interested in the effects of *Lpar6* and *Cab39l* loss of function on urothelial cells, all nonurothelial clusters were removed. In addition, the poor-quality urothelial clusters with the cells that had low unique gene reads or high mitochondrial reads were removed. This led to 4,693, 6,254, and 7,012 high quality urothelial cells which were re-clustered by UMAP using the top 40 new PCs. The final UMAP clusters comprised 18 cell clusters, including 5 basal urothelial cell clusters, 5 luminal urothelial cell clusters, and 8 intermediate urothelial cell clusters identified by 10 basal urothelial cell marker genes and 14 luminal urothelial cell marker genes. For every UMAP cluster of the three sample groups, the expression of the most differentially expressed gene, the unfolded protein response marker genes, and the cholesterol synthesis marker genes were visualized by dot plots, respectively. The cell proportion of each sample group across 18 clusters and the cell proportion of each urothelial cell type across the three sample groups were visualized by bar plots, respectively. The cell-level basal-to-luminal (BLT) score is a weighted sum of the urothelial marker gene expression and was calculated based on 6 basal urothelial cell marker genes with weight −1 and 14 luminal urothelial cell marker genes with weight 1.[17] The BLT scores were sorted using ascending order and visualized by bar plots across the three sample groups.

The trajectory analysis was performed by using the workflow of Monocle3.[74] Briefly, the UMAP coordinates were transferred from Seurat to Monocle3 to preserve the cluster identities, and the trajectories across the cells were generated by using gene expressions that signified cell state transitions within the dataset. The trajectories were learned by learn_graph function. It created several distinct trajectory paths in the three sample groups indicating multiple outcomes of cell transitions. For each trajectory path, we manually chose a starting point from basal urothelial cells and treated it as the root cell. The cells across a trajectory path were ordered in pseudotime by order_cells function. The lowest pseudotime score was assigned to the root basal urothelial cell and the ordering of cells across a trajectory was assigned based on a pseudotime score in the Monocle3 tool's perception of the cell state transitions. The 200 top genes that changed as the cell progress across a trajectory path were identified by graph_test function, and they were subjected to the enrichment analysis performed by g:Profiler.[75] For each sample group, the significantly enriched KEGG pathways were compared across different trajectory paths by their $-\log_{10}$ p-value plotted on a heatmap.

The cell-cell interaction analysis was performed by CellPhoneDB [76] on the normalized gene counts of the three sample groups to determine significant ligand-receptor interactions with a p-value cutoff of 0.0001. Ligand-receptor pairs involving secreted factors were assessed for signaling the interactions between the intermediate or basal urothelial cells and the rest two urothelial cell types respectively across the three sample groups. The outputs of CellPhoneDB analysis were processed by InterCellar.[77] A circos plot described in iTALK [78] was used to visualize the interactions between the intermediate urothelial cells and the rest two urothelial cell types respectively across the three sample groups.

The cell cycle and proliferation status of single cells was analyzed by calculating cell S scores and G2M scores using a set of genes which were reported to be involved in cell cycle and include 43 maker genes for S phase and 53 marker genes for G2M phase.[79] We supplied these marker genes to Seurat CellCycleScoring function to calculate G2M and S scores and define cell phases. The estimated probability distribution densities of S scores and G2M scores of the three urothelial cell types across the three sample groups were visualized by ridge plots. We used the Wilcoxon rank-sum test to compare the distributions of G2M scores of the basal urothelial cells with the rest two urothelial cell types, respectively, across the three sample groups. In addition, G2M scores were visualized across every trajectory path in the three sample groups using the scatter plots with the smoothing curve generated by loess method in ggplot in R.

We used regulon analysis[80] to analyze the activity of regulons across different cluster groups in basal, intermediate, and luminal urothelial cells. We started with the potential transcription factors (TFs) with possible target genes to form the regulons. We analyzed the first group of 23 urothelial regulons originally analyzed in the TCGA cohort (n=408) of bladder cancer,[81] which included the following TFs: FOXA1, RXRA, FGFR3, RARG, RXRB, ERBB3, AR, GATA3, ESR2, ERBB2, PPARG, RARA, FGFR1, PGR, RARB, TP63, ESR1, GATA6, STAT3, FOXM1, KLF4, EGFR, and HIF1A. The second group of 74 regulons were previously identified by single cell sequencing of the human urothelium.[23] Then, we implemented the method in R package RTN to create the regulons for the target

TFs.[80,82] Using Spearman's correlation coefficient, this method first classified the list of target genes for each TF into positive and negative targets related to the phenotype of interest. The positive and negative target distributions were tested, yielding enrichment scores (ES) for each sample. The differential enrichment score (dES) was determined by the difference between positive and negative ES. We extracted the dES for each sample to further investigate the difference in the activities of the cells among the basal, intermediate, and luminal urothelial clustering groups by the two-sample t-test. We examined how the regulons were related to basal and luminal subtypes using heatmaps and two-tailed gene set enrichment analysis (GSEA).[56] Out of 97 regulons, we filtered out 24 regulons with no significant difference in the activities of the cells among the different subtype clustering groups.

QUANTIFICATION AND STATISTICAL ANALYSIS

All plotted data are presented as the mean ± SD, unless otherwise specified, and described in the corresponding section of method details and figure legend. Statistical differences between the two groups were compared and determined by Student's t-test for parametric analysis, Mann-Whitney U-test for non-parametric analysis, and log-rank test for survival analysis. p values were calculated using software R or GraphPad Prism 10.0.1. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$, ns = non-significant. Bioinformatics analyses are described in the above sections.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

This study was supported by NCI Genitourinary Bladder SPORE grant P50CA 91846 (Project 1 and Core C) to B.C.; Cancer Prevention & Research Institute of Texas grants RP220021 to B.C. and RP230166 to B.C. and P.W.; the MD Anderson Nathan W. Lassiter Endowment; and MD Anderson Donor Funds to B.C.

REFERENCES

1. Henson DE, and Albores-Saavedra J. (2001). Pathology of Incipient Neoplasia, Third Edition (Oxford University Press).
2. Czerniak B. (2008). Incipient events in human caecinogenesis: A concept of forerunner genes. In Regulation of Gene Expression in the Tumor Environment, Bar-Eli M, ed. (Springer Netherlands), pp. 125–146. 10.1007/978-1-4020-8341-9_7.
3. Czerniak B. (2011). Molecular pathology and biomarkers of bladder cancer. Cancer Biomark 9, 159–176. 10.3233/CBM-2011-0175.
4. Dinney CPN, McConkey DJ, Millikan RE, Wu X, Bar-Eli M, Adam L, Kamat AM, Siefker-Radtke AO, Tuziak T, Sabichi AL, et al. (2004). Focus on bladder cancer. Cancer Cell 6, 111–116. 10.1016/j.ccr.2004.08.002. [PubMed: 15324694]
5. Lee S, Jeong J, Majewski T, Scherer SE, Kim MS, Tuziak T, Tang KS, Baggerly K, Grossman HB, Zhou JH, et al. (2007). Forerunner genes contiguous to RB1 contribute to the development of in situ neoplasia. Proc. Natl. Acad. Sci. USA 104, 13732–13737. 10.1073/pnas.0701771104. [PubMed: 17702869]
6. Majewski T, Lee S, Jeong J, Yoon DS, Kram A, Kim MS, Tuziak T, Bondaruk J, Lee S, Park WS, et al. (2008). Understanding the development of human bladder cancer by using a whole-

organ genomic mapping strategy. Lab. Invest 88, 694–721. 10.1038/la-binvest.2008.27. [PubMed: 18458673]

7. Koss LG, Tiamson EM, and Robbins MA (1974). Mapping cancerous and precancerous bladder changes. A study of the urothelium in ten surgically removed bladders. JAMA 227, 281–286. [PubMed: 4859656]
8. Czerniak B, Chaturvedi V, Li L, Hodges S, Johnston D, Roy JY, Luthra R, Logothetis C, Von Eschenbach AC, Grossman HB, et al. (1999). Superimposed histologic and genetic mapping of chromosome 9 in progression of human urinary bladder neoplasia: implications for a genetic model of multistep urothelial carcinogenesis and early detection of urinary bladder cancer. Oncogene 18, 1185–1196. 10.1038/sj.onc.1202385. [PubMed: 10022124]
9. Czerniak B, Li L, Chaturvedi V, Ro JY, Johnston DA, Hodges S, and Benedict WF (2000). Genetic modeling of human urinary bladder carcinogenesis. Genes Chromosomes Cancer 27, 392–402. [PubMed: 10719370]
10. Kim MS, Jeong J, Majewski T, Kram A, Yoon DS, Zhang RD, Li JZ, Ptaszynski K, Kuang TC, Zhou JH, et al. (2006). Evidence for alternative candidate genes near RB1 involved in clonal expansion of in situ urothelial neoplasia. Lab. Invest 86, 175–190. 10.1038/labinvest.3700378. [PubMed: 16402033]
11. Kram A, Li L, Zhang RD, Yoon DS, Ro JY, Johnston D, Grossman HB, Scherer S, and Czerniak B. (2001). Mapping and genome sequence analysis of chromosome 5 regions involved in bladder cancer progression. Lab. Invest 81, 1039–1048. 10.1038/labinvest.3780315. [PubMed: 11454992]
12. Spiess PE, and Czerniak B. (2006). Dual-track pathway of bladder carcinogenesis: practical implications. Arch. Pathol. Lab Med. 130, 844–852. 10.5858/2006-130-844-DPOBCP. [PubMed: 16740038]
13. Tuziak T, Jeong J, Majewski T, Kim MS, Steinberg J, Wang Z, Yoon DS, Kuang TC, Baggerly K, Johnston D, and Czerniak B. (2005). High-resolution whole-organ mapping with SNPs and its significance to early events of carcinogenesis. Lab. Invest 85, 689–701. 10.1038/labinvest.3700270. [PubMed: 15908911]
14. Yoon DS, Li L, Zhang RD, Kram A, Ro JY, Johnston D, Grossman HB, Scherer S, and Czerniak B. (2001). Genetic mapping and DNA sequence-based analysis of deleted regions on chromosome 16 involved in progression of bladder cancer from occult preneoplastic conditions to invasive disease. Oncogene 20, 5005–5014. 10.1038/sj.onc.1204612. [PubMed: 11526485]
15. Choi W, Porten S, Kim S, Willis D, Plimack ER, Hoffman-Censits J, Roth B, Cheng T, Tran M, Lee IL, et al. (2014). Identification of distinct basal and luminal subtypes of muscle-invasive bladder cancer with different sensitivities to frontline chemotherapy. Cancer Cell 25, 152–165. 10.1016/j.ccr.2014.01.009. [PubMed: 24525232]
16. Dadhania V, Zhang M, Zhang L, Bondaruk J, Majewski T, Siefker-Radtke A, Guo CC, Dinney C, Cogdell DE, 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. 10.1016/j.ebiom.2016.08.036. [PubMed: 27612592]
17. Guo CC, Bondaruk J, Yao H, Wang Z, Zhang L, Lee S, Lee JG, Cogdell D, Zhang M, Yang G, et al. (2020). Assessment of Luminal and Basal Phenotypes in Bladder Cancer. Sci. Rep 10, 9743. 10.1038/s41598-020-66747-7. [PubMed: 32546765]
18. Robertson AG, Kim J, Al-Ahmadie H, Bellmunt J, Guo G, Cher-niack AD, Hinoue T, Laird PW, Hoadley KA, Akbani R, et al. (2017). Comprehensive Molecular Characterization of Muscle-Invasive Bladder Cancer. Cell 171, 540–556.e25. 10.1016/j.cell.2017.09.007. [PubMed: 28988769]
19. Sjödahl G, Lauss M, Lövgren K, Chebil G, Gudjonsson S, Veerla S, Patschan O, Aine M, Fernö M, Ringnér M, et al. (2012). A molecular taxonomy for urothelial carcinoma. Clin. Cancer Res. 18, 3377–3386. 10.1158/1078-0432.CCR-12-0077-T. [PubMed: 22553347]
20. Bertone P, Stolc V, Royce TE, Rozowsky JS, Urban AE, Zhu X, Rinn JL, Tongprasit W, Samanta M, Weissman S, et al. (2004). Global identification of human transcribed sequences with genome tiling arrays. Science 306, 2242–2246. 10.1126/science.1103388. [PubMed: 15539566]
21. Fishwick C, Higgins J, Percival-Alwyn L, Hustler A, Pearson J, Bastkowski S, Moxon S, Swarbreck D, Greenman CD, and Southgate J. (2017). Heterarchy of transcription factors driving basal and luminal cell phenotypes in human urothelium. Cell Death Differ. 24, 809–818. 10.1038/cdd.2017.10. [PubMed: 28282036]

22. Karni-Schmidt O, Castillo-Martin M, Shen TH, Gladoun N, Domingo-Domenech J, Sanchez-Carbayo M, Li Y, Lowe S, Prives C, and Cordon-Cardo C. (2011). Distinct expression profiles of p63 variants during urothelial development and bladder cancer progression. Am. J. Pathol 178, 1350–1360. 10.1016/j.ajpath.2010.11.061. [PubMed: 21356385]
23. Fink EE, Sona S, Tran U, Desprez PE, Bradley M, Qiu H, Eltemamy M, Wee A, Wolkov M, Nicolas M, et al. (2022). Single-cell and spatial mapping Identify cell types and signaling Networks in the human ureter. Dev. Cell 57, 1899–1916.e6. 10.1016/j.devcel.2022.07.004. [PubMed: 35914526]
24. Cheng X, Lai H, Luo W, Zhang M, Miao J, Song W, Xing S, Wang J, and Gao WQ (2021). Single-cell analysis reveals urothelial cell heterogeneity and regenerative cues following cyclophosphamide-induced bladder injury. Cell Death Dis. 12, 446. 10.1038/s41419-021-03740-6. [PubMed: 33953164]
25. Majewski T, Yao H, Bondaruk J, Chung W, Lee S, Lee JG, Zhang S, Cogdell D, Yang G, Choi W, et al. (2019). Whole-Organ Genomic Characterization of Mucosal Field Effects Initiating Bladder Carcinogenesis. Cell Rep. 26, 2241–2256.e4. 10.1016/j.celrep.2019.01.095. [PubMed: 30784602]
26. Bondaruk J, Jaksik R, Wang Z, Cogdell D, Lee S, Chen Y, Dinh KN, Majewski T, Zhang L, Cao S, et al. (2022). The origin of bladder cancer from mucosal field effects. iScience 25, 104551. 10.1016/j.isci.2022.104551.
27. Yanagida K, Masago K, Nakanishi H, Kihara Y, Hamano F, Tajima Y, Taguchi R, Shimizu T, and Ishii S. (2009). Identification and characterization of a novel lysophosphatidic acid receptor, p2y5/LPA6. J. Biol. Chem 284, 17731–17741. 10.1074/jbc.M808506200. [PubMed: 19386608]
28. Pasternack SM, von Kügelgen I, Al Aboud K, Lee YA, Rüschendorf F, Voss K, Hillmer AM, Molderings GJ, Franz T, Ramirez A, et al. (2008). G protein-coupled receptor P2Y5 and its ligand LPA are involved in maintenance of human hair growth. Nat. Genet 40, 329–334. 10.1038/ng.84. [PubMed: 18297070]
29. Shimomura Y, Wajid M, Ishii Y, Shapiro L, Petukhova L, Gordon D, and Christiano AM (2008). Disruption of P2RY5, an orphan G protein-coupled receptor, underlies autosomal recessive woolly hair. Nat. Genet 40, 335–339. 10.1038/ng.100. [PubMed: 18297072]
30. Gnocchi D, Afonso MB, Cavalluzzi MM, Lentini G, Ingravallo G, Sabbà C, Rodrigues CMP, and Mazzocca A. (2023). Inhibition of lysophosphatidic acid receptor 6 upregulated by the choline-deficient l-amino acid-defined diet prevents hepatocarcinogenesis in mice. Mol. Carcinog 62, 577–582. 10.1002/mc.23516. [PubMed: 36752344]
31. Ishii S, Hirane M, Fukushima K, Tomimatsu A, Fukushima N, and Tsujiuchi T. (2015). Diverse effects of LPA4, LPA5 and LPA6 on the activation of tumor progression in pancreatic cancer cells. Biochem. Biophys. Res. Commun 461, 59–64. 10.1016/j.bbrc.2015.03.169. [PubMed: 25849892]
32. Lei J, Guo S, Li K, Tian J, Zong B, Ai T, Peng Y, Zhang Y, and Liu S. (2022). Lysophosphatidic acid receptor 6 regulated by miR-27a-3p attenuates tumor proliferation in breast cancer. Clin. Transl. Oncol 24, 503–516. 10.1007/s12094-021-02704-8. [PubMed: 34510318]
33. Mazzocca A, Dituri F, De Santis F, Filannino A, Lopane C, Betz RC, Li YY, Mukaida N, Winter P, Tortorella C, et al. (2015). Lysophosphatidic acid receptor LPAR6 supports the tumorigenicity of hepatocellular carcinoma. Cancer Res. 75, 532–543. 10.1158/0008-5472.CAN-14-1607. [PubMed: 25589345]
34. Mazzocca A, Dituri F, Lupo L, Quaranta M, Antonaci S, and Giannelli G. (2011). Tumor-secreted lysophostatidic acid accelerates hepatocellular carcinoma progression by promoting differentiation of peritumoral fibroblasts in myofibroblasts. Hepatology 54, 920–930. 10.1002/hep.24485. [PubMed: 21674557]
35. Takahashi K, Fukushima K, Onishi Y, Inui K, Node Y, Fukushima N, Honoki K, and Tsujiuchi T. (2017). Lysophosphatidic acid (LPA) signaling via LPA(4) and LPA(6) negatively regulates cell motile activities of colon cancer cells. Biochem. Biophys. Res. Commun 483, 652–657. 10.1016/j.bbrc.2016.12.088. [PubMed: 27993681]
36. Li W, Wong CC, Zhang X, Kang W, Nakatsu G, Zhao Q, Chen H, Go MYY, Chiu PWY, Wang X, et al. (2018). CAB39L elicited an anti-Warburg effect via a LKB1-AMPK-PGC1alpha axis to inhibit gastric tumorigenesis. Oncogene 37, 6383–6398. 10.1038/s41388-018-0402-1. [PubMed: 30054562]

37. Filippi BM, de los Heros P, Mehellou Y, Navratilova I, Gourlay R, Deak M, Plater L, Toth R, Zeqiraj E, and Alessi DR (2011). MO25 is a master regulator of SPAK/OSR1 and MST3/MST4/YSK1 protein kinases. EMBO J. 30, 1730–1741. 10.1038/emboj.2011.78. [PubMed: 21423148]
38. Zeqiraj E, Filippi BM, Deak M, Alessi DR, and van Aalten DMF (2009). Structure of the LKB1-STRAD-MO25 complex reveals an allosteric mechanism of kinase activation. Science 326, 1707–1711. 10.1126/science.1178377. [PubMed: 19892943]
39. Zhang PY, Zhang WG, He AL, Wang JL, and Li WB (2009). Identification and functional characterization of the novel acute monocytic leukemia associated antigen MLAA-34. Cancer Immunol. Immunother 58, 281–290. 10.1007/s00262-008-0552-z. [PubMed: 18592235]
40. Choi MR, An CH, Yoo NJ, and Lee SH (2016). Frameshift Mutations of CAB39L, an Activator of LKB1 Tumor Suppressor, in Gastric and Colorectal Cancers. Pathol. Oncol. Res 22, 225–226. 10.1007/s12253-015-9973-0. [PubMed: 26306467]
41. Wu Y, Xu Z, Chen X, Fu G, Tian J, Shi Y, Sun J, and Jin B. (2023). Bioinformatics Prediction and Experimental Verification Identify CAB39L as a Diagnostic and Prognostic Biomarker of Kidney Renal Clear Cell Carcinoma. Medicina (Kaunas) 59, 716. 10.3390/medicina59040716. [PubMed: 37109674]
42. Pierceall WE, Kripke ML, and Ananthaswamy HN (1992). N-ras mutation in ultraviolet radiation-induced murine skin cancers. Cancer Res. 52, 3946–3951. [PubMed: 1617670]
43. Takahashi T, Nau MM, Chiba I, Birrer MJ, Rosenberg RK, Vinocour M, Levitt M, Pass H, Gazdar AF, and Minna JD (1989). p53: a frequent target for genetic abnormalities in lung cancer. Science 246, 491–494. 10.1126/science.2554494. [PubMed: 2554494]
44. Liggett WH Jr., and Sidransky D. (1998). Role of the p16 tumor suppressor gene in cancer. J. Clin. Oncol 16, 1197–1206. 10.1200/JCO.1998.16.3.1197. [PubMed: 9508208]
45. Wang X, Christiani DC, Mark EJ, Nelson H, Wiencke JK, Gunn L, Wain JC, and Kelsey KT (1999). Carcinogen exposure, p53 alteration, and K-ras mutation in synchronous multiple primary lung carcinoma. Cancer 85, 1734–1739. [PubMed: 10223567]
46. Lawson ARJ, Abascal F, Coorens THH, Hooks Y, O'Neill L, Latimer C, Raine K, Sanders MA, Warren AY, Mahbubani KTA, et al. (2020). Extensive heterogeneity in somatic mutation and selection in the human bladder. Science 370, 75–82. 10.1126/science.aba8347. [PubMed: 33004514]
47. Li R, Du Y, Chen Z, Xu D, Lin T, Jin S, Wang G, Liu Z, Lu M, Chen X, et al. (2020). Macroscopic somatic clonal expansion in morphologically normal human urothelium. Science 370, 82–89. 10.1126/science.aba7300. [PubMed: 33004515]
48. Strandgaard T, Nordentoft I, Lamy P, Christensen E, Thomsen MBH, Jensen JB, Dyrskjøt L, Jensen JB, Bjerggaard Jensen J, Dyrskjot L, and Dyrskjot L. (2020). Mutational Analysis of field Cancerization i Bladder Cancer. Bladder Cancer 6, 253–264.
49. Thomsen MBH, Nordentoft I, Lamy P, Vang S, Reinert L, Mapendano CK, Høyer S, Ørntoft TF, Jensen JB, and Dyrskjøt L. (2017). Comprehensive multiregional analysis of molecular heterogeneity in bladder cancer. Sci. Rep 7, 11702. 10.1038/s41598-017-11291-0. [PubMed: 28916750]
50. Damaghi M, West J, Robertson-Tessi M, Xu L, Ferrall-Fairbanks MC, Stewart PA, Persi E, Fridley BL, Altrock PM, Gatenby RA, et al. (2021). The harsh microenvironment in early breast cancer selects for a Warburg phenotype. Proc. Natl. Acad. Sci. USA 118, e2011342118. 10.1073/pnas.2011342118.
51. Gnocchi D, Nikolic D, Paparella RR, Sabbà C, and Mazzocca A. (2023). Cellular Adaptation Takes Advantage of Atavistic Regression Programs during Carcinogenesis. Cancers 15, 3942. 10.3390/cancers15153942. [PubMed: 37568758]
52. Jia D, Lu M, Jung KH, Park JH, Yu L, Onuchic JN, Kaipparettu BA, and Levine H. (2019). Elucidating cancer metabolic plasticity by coupling gene regulation with metabolic pathways. Proc. Natl. Acad. Sci. USA 116, 3909–3918. 10.1073/pnas.1816391116. [PubMed: 30733294]
53. Wang L, Wang S, and Li W. (2012). RSeQC: quality control of RNA-seq experiments. Bioinformatics 28, 2184–2185. 10.1093/bioinformatics/bts356. [PubMed: 22743226]

54. Liao Y, Smyth GK, and Shi W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30, 923–930. 10.1093/bioinformatics/btt656. [PubMed: 24227677]
55. Love MI, Huber W, and Anders S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550. 10.1186/s13059-014-0550-8. [PubMed: 25516281]
56. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, and Mesirov JP (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. USA 102, 15545–15550. 10.1073/pnas.0506580102. [PubMed: 16199517]
57. Chapman EJ, Williams SV, Platt FM, Hurst CD, Chambers P, Roberts P, and Knowles MA (2009). Integrated genomic and transcriptional analysis of the in vitro evolution of telomerase-immortalized urothelial cells (TERT-NHUC). Genes Chromosomes Cancer 48, 694–710. 10.1002/gcc.20672. [PubMed: 19405089]
58. Chapman EJ, Hurst CD, Pitt E, Chambers P, Aveyard JS, and Knowles MA (2006). Expression of hTERT immortalises normal human urothelial cells without inactivation of the p16/Rb pathway. Oncogene 25, 5037–5045. 10.1038/sj.onc.1209513. [PubMed: 16619045]
59. Cancer, I.A.f.R.o., Moch H, and Reuter VE (2016). WHO Classification of Tumours of the Urinary System and Male Genital Organs (International Agency for Research on Cancer).
60. Brierley JD, Gospodarowicz MK, and Wittekind C. (2016). TNM Classification of Malignant Tumours, 8th Edition (Wiley-Blackwell).
61. Richter J, Jiang F, Görög JP, Sartorius G, Egenter C, Gasser TC, Moch H, Mihatsch MJ, and Sauter G. (1997). Marked genetic differences between stage pTa and stage pT1 papillary bladder cancer detected by comparative genomic hybridization. Cancer Res. 57, 2860–2864. [PubMed: 9230190]
62. Sen S, Zhou H, Zhang RD, Yoon DS, Vakar-Lopez F, Ito S, Jiang F, Johnston D, Grossman HB, Ruifrok AC, et al. (2002). Amplification/overexpression of a mitotic kinase gene in human bladder cancer. J. Natl. Cancer Inst. 94, 1320–1329. 10.1093/jnci/94.17.1320. [PubMed: 12208897]
63. Fackler MJ, McVeigh M, Mehrotra J, Blum MA, Lange J, Lapides A, Garrett E, Argani P, and Sukumar S. (2004). Quantitative multiplex methylation-specific PCR assay for the detection of promoter hypermethylation in multiple genes in breast cancer. Cancer Res. 64, 4442–4452. 10.1158/0008-5472.CAN-03-3341. [PubMed: 15231653]
64. Zeschnigk M, Böhringer S, Price EA, Onadim Z, Masshöfer L, and Lohmann DR (2004). A novel real-time PCR assay for quantitative analysis of methylated alleles (QAMA): analysis of the retinoblastoma locus. Nucleic Acids Res. 32, e125. 10.1093/nar/gnh122. [PubMed: 15353561]
65. Calin GA, Trapasso F, Shimizu M, Dumitru CD, Yendamuri S, Godwin AK, Ferracin M, Bernardi G, Chatterjee D, Baldassarre G, et al. (2005). Familial cancer associated with a polymorphism in ARLTS1. N. Engl. J. Med 352, 1667–1676. 10.1056/NEJMoa042280. [PubMed: 15843669]
66. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, and Gingeras TR (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21. 10.1093/bioinformatics/bts635. [PubMed: 23104886]
67. Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, Aken BL, Barrell D, Zadissa A, Searle S, et al. (2012). GEN-CODE: the reference human genome annotation for The ENCODE Project. Genome Res. 22, 1760–1774. 10.1101/gr.135350.111. [PubMed: 22955987]
68. Benjamini Y, and Hochberg Y. (1995). Controlling the False Discovery Rate - A Practical and Powerful Approach to Multiple Testing. J. Roy. Stat. Soc. B 57, 289–300.
69. Solberg HE (1978). Discriminant analysis. CRC Crit. Rev. Clin. Lab. Sci 9, 209–242. 10.3109/10408367809150920. [PubMed: 401370]
70. Li Z, Liu H, and Tu W. (2018). Model selection in multivariate semiparametric regression. Stat. Methods Med. Res 27, 3026–3038. 10.1177/0962280217690769. [PubMed: 28164740]
71. Dudas M, Sridurongrit S, Nagy A, Okazaki K, and Kaartinen V. (2004). Craniofacial defects in mice lacking BMP type I receptor Alk2 in neural crest cells. Mech. Dev 121, 173–182. 10.1016/j.mod.2003.12.003. [PubMed: 15037318]
72. Kaartinen V, Voncken JW, Shuler C, Warburton D, Bu D, Heisterkamp N, and Groffen J. (1995). Abnormal lung development and cleft palate in mice lacking TGF-beta 3 indicates defects of

epithelial-mesenchymal interaction. Nat. Genet 11, 415–421. 10.1038/ng1295-415. [PubMed: 7493022]

73. Hao Y, Hao S, Andersen-Nissen E, Mauck WM 3rd, Zheng S, Butler A, Lee MJ, Wilk AJ, Darby C, Zager M, et al. (2021). Integrated analysis of multimodal single-cell data. Cell 184, 3573–3587.e29. 10.1016/j.cell.2021.04.048. [PubMed: 34062119]
74. Cao J, Spielmann M, Qiu X, Huang X, Ibrahim DM, Hill AJ, Zhang F, Mundlos S, Christiansen L, Steemers FJ, et al. (2019). The single-cell transcriptional landscape of mammalian organogenesis. Nature 566, 496–502. 10.1038/s41586-019-0969-x. [PubMed: 30787437]
75. Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, and Vilo J. (2019). g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). Nucleic Acids Res. 47, W191–W198. 10.1093/nar/gkz369. [PubMed: 31066453]
76. Efremova M, Vento-Tormo M, Teichmann SA, and Vento-Tormo R. (2020). CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. Nat. Protoc 15, 1484–1506. 10.1038/s41596-020-0292-x. [PubMed: 32103204]
77. Interlandi M, Kerl K, and Dugas M. (2022). InterCellar enables interactive analysis and exploration of cell-cell communication in single-cell transcriptomic data. Commun. Biol 5, 21. 10.1038/s42003-021-02986-2. [PubMed: 35017628]
78. Wang Y, Wang R, Zhang S, Song S, Jiang C, Han G, Wang M, Ajani J, Futreal A, and Wang L. (2019). iTALK: an R Package to Characterize and Illustrate Intercellular Communication. Preprint at bioRxiv. 10.1101/507871.
79. Tirosh I, Izar B, Prakadan SM, Wadsworth MH 2nd, Treacy D, Trombetta JJ, Rotem A, Rodman C, Lian C, Murphy G, et al. (2016). Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science 352, 189–196. 10.1126/science.aad0501. [PubMed: 27124452]
80. Castro MAA, de Santiago I, Campbell TM, Vaughn C, Hickey TE, Ross E, Tilley WD, Markowetz F, Ponder BAJ, and Meyer KB (2016). Regulators of genetic risk of breast cancer identified by integrative network analysis. Nat. Genet 48, 12–21. 10.1038/ng.3458. [PubMed: 26618344]
81. Yang G, Bondaruk J, Cogdell D, Wang Z, Lee S, Lee JG, Zhang S, Choi W, Wang Y, Liang Y, et al. (2020). Urothelial-to-Neural Plasticity Drives Progression to Small Cell Bladder Cancer. iScience 23, 101201. 10.1016/j.isci.2020.101201.
82. Fletcher MNC, Castro MAA, Wang X, de Santiago I, O'Reilly M, Chin SF, Rueda OM, Caldas C, Ponder BAJ, Markowetz F, and Meyer KB (2013). Master regulators of FGFR2 signalling and breast cancer risk. Nat. Commun 4, 2464. 10.1038/ncomms3464. [PubMed: 24043118]

Highlights

- Silencing of FR genes is associated with mucosal field effects in bladder carcinogenesis
- *LPAR6* and *CAB39L* are distinctively downregulated in luminal and basal bladder cancers
- Effects of *LPAR6* and *CAB39L* are mediated by cholesterol and the unfolded protein reaction
- *LPAR6* and *CAB39L* loss dysregulates urothelial differentiation, contributing to carcinogenesis

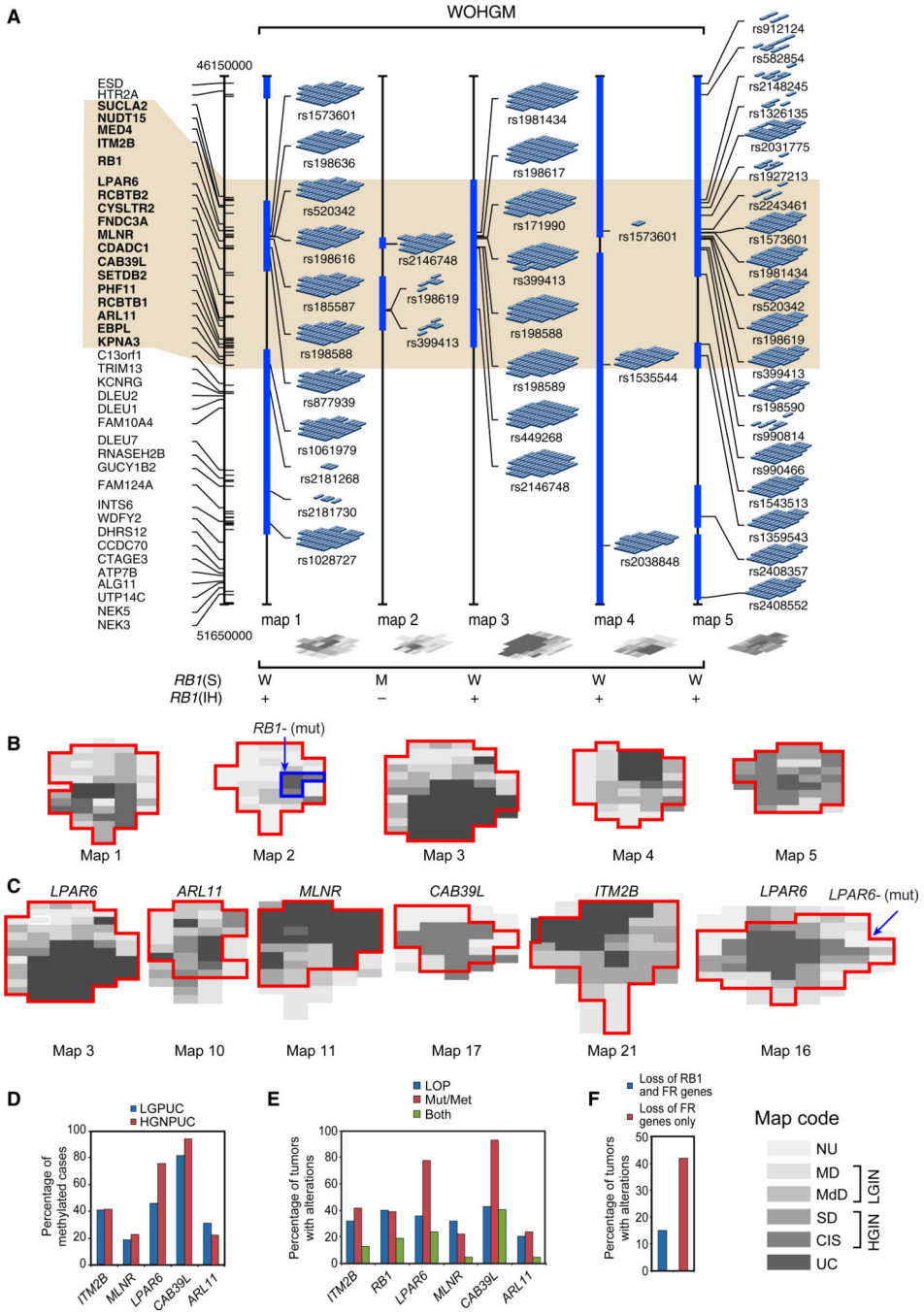


Figure 1. Mapping and expression studies of the 13q14 region around *RB1*

(A) Regions of loss of genetic material associated with clonal expansion identified using WOHGM in five cystectomy samples. The results of *RB1* sequencing (*RB1*(S)) and immunohistochemical studies for RB protein expression (*RB1*(IH)) are tabulated below the individual maps (W, RB1 wild type; M, RB1 mutant). The presence or absence of immunohistochemically detectable RB protein is designated by (+/−). Patterns of genetic loss in individual cystectomy samples (maps 1–5) reconstructed from the data on LOP of individual SNPs (dark-gray-shaded blocks) are depicted by the blue solid bars. The

light-brown-shaded area designates the minimal deleted region associated with plaque-like clonal expansion of intraurothelial neoplasia.

(B) The distribution of loss of genetic material identified by allelotyping of SNPs superimposed over histologic maps of cystectomy samples. The solid red lines outline the areas of loss involving *RB1* and its flanking regions identified using WOHGM with allelotyping of SNPs. The blue line on map 2 shows the area with mutated *RB1* (C1666T, R566STOP).

(C) Distribution of methylation in the promoter region combined with loss of *LPAR6* (map 3), *ARL11* (map 10), *MLNR* (map 11), *CAB39L* (map 17), and *ITM2B* (map 21) genetic material. In addition, for *LPAR6* (map 16), the distribution of gene silencing by mutation (G886del with a frameshift to a stop codon) and loss of genetic material is shown. The solid red line outlines an area of hypermethylation of the promoter region combined with loss of genetic material and mutations for *LPAR6* as described above.

(D) Percentage of low-grade papillary urothelial carcinoma (LGPUC) and high-grade nonpapillary urothelial carcinoma (HGNPUC) with hypermethylated candidate genes in the MD Anderson cohort of 111 bladder cancers.

(E) Percentage of tumors with combined LOP and mutations/methylation of FR genes in the cohort of 111 bladder cancers.

(F) Percentage of cases with combined loss of the *RB1* and FR genes and percentage of tumor cases with loss of FR genes only in the cohort of 111 bladder cancers. Loss of gene function is defined as combined loss of genetic material (LOP) and methylation or mutation. Map code: NU, normal urothelium; MD, mild dysplasia; MdD, moderate dysplasia; SD, severe dysplasia; CIS, carcinoma *in situ*; UC, urothelial carcinoma; LGIN, low-grade intraurothelial neoplasia; HGIN, high-grade intraurothelial neoplasia.

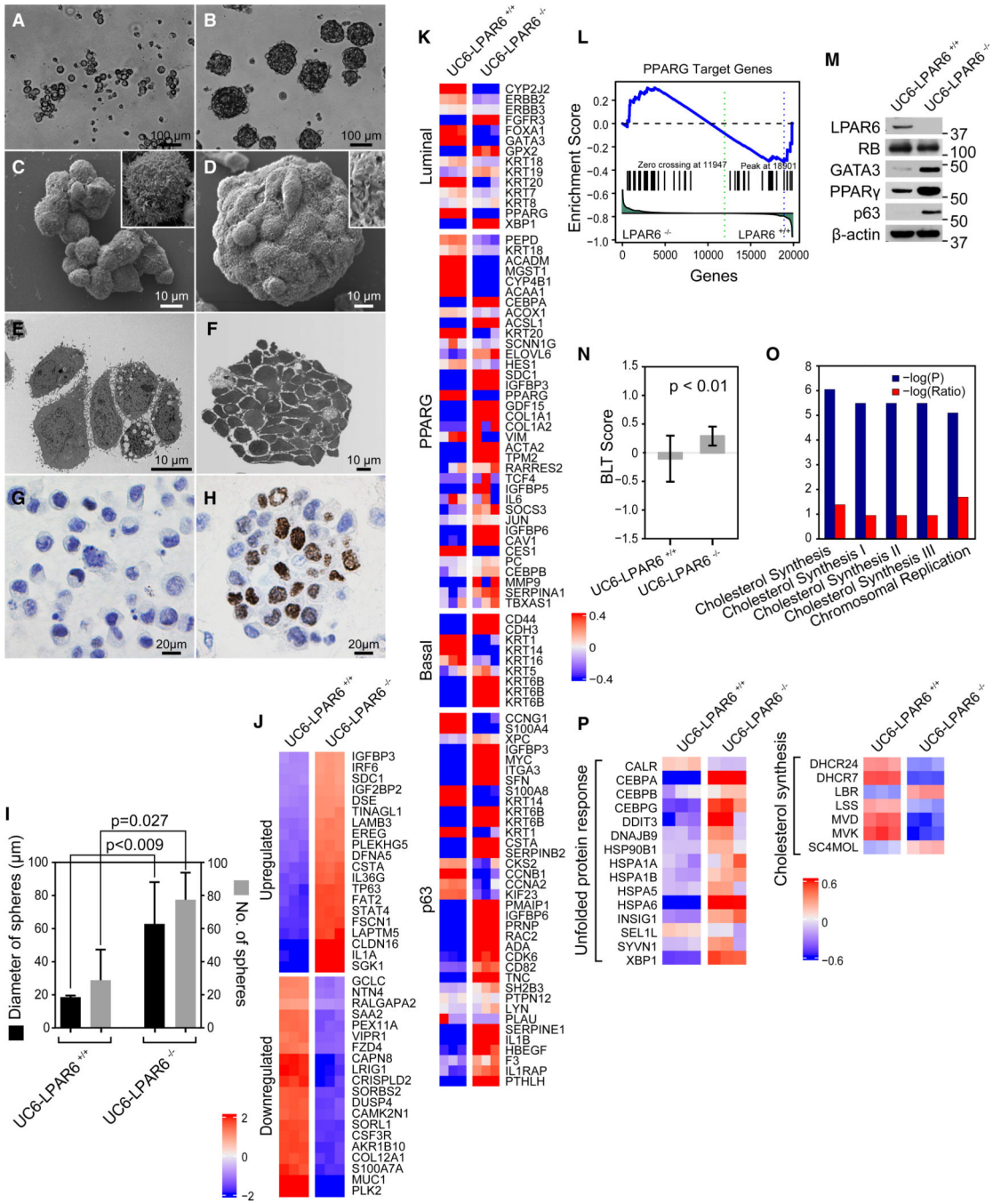


Figure 2. Effects of *LPAR6* ablation on the UC6 basal bladder cancer cell line

(A) Dispersed growth pattern of *LPAR6*$^{+/+}$ UC6 cells. Scale bar, 100 μm.

(B) Formation of urospheres in *LPAR6*$^{-/-}$ UC6 cells. Scale bar, 100 μm.

(C) Scanning electron micrograph showing disorganized, loosely arranged clusters of *LPAR6*$^{+/+}$ UC6 cells. Inset: the surface of a poorly differentiated urothelial cell with microvilli. Scale bar, 10 μm.

(D) Scanning electron micrograph showing a well-developed urosphere in *LPAR6*$^{-/-}$ UC6 cells. Inset: the surface of the urothelial cells outlining a urosphere with cell membrane ridges. Scale bar, 10 μm.
(E) Transmission electron micrograph showing disorganized, loosely arranged clusters of *LPAR6*$^{+/+}$ UC6 cells. Scale bar, 10 μm.
(F) Transmission electron micrograph showing a well-developed urosphere in *LPAR6*$^{-/-}$ UC6 cells. Scale bar, 10 μm.
(G) Negative staining for GATA3 in *LPAR6*$^{+/+}$ UC6 cells. Scale bar, 20 μm.
(H) Strong positive nuclear staining for GATA3 in a urosphere of *LPAR6*$^{-/-}$ UC6 cells. Scale bar, 20 μm.
(I) Quantitative analysis of urosphere size and number in *LPAR6*$^{+/+}$ and *LPAR6*$^{-/-}$ UC6 cells. Bars represent mean + standard error. Number of replicates (n)= 3. The p values are based on two-tailed Student's t test.
(J) Heatmap showing the expression pattern for the top 20 upregulated and downregulated genes in *LPAR6*$^{+/+}$ and *LPAR6*$^{-/-}$ UC6 cells.
(K) Heatmap showing the pattern of luminal and basal marker expression in *LPAR6*$^{+/+}$ and *LPAR6*$^{-/-}$ UC6 cells.
(L) Gene set enrichment analysis of *PPARγ* target genes in *LPAR6*$^{-/-}$ UC6 cells compared with *LPAR6*$^{+/+}$ cells.
(M) Western blots showing ablation of LPAR6 protein in *LPAR6*$^{-/-}$ UC6 cells, intact expression of RB protein, and expression of representative luminal (GATA3 and PPARγ) and basal (p63) proteins.
(N) BLT score for *LPAR6*$^{+/+}$ and *LPAR6*$^{-/-}$ UC6 cells. Bars represent mean ± standard error. Number of replicates (n) = 3. The p value is based on two-tailed Student's t test.
(O) Top five dysregulated pathways in *LPAR6*$^{-/-}$ UC6 cells.
(P) Expression pattern for representative unfolded protein response (left) and cholesterol synthesis genes (right) in *LPAR6*$^{+/+}$ and *LPAR6*$^{-/-}$ UC6 cells.

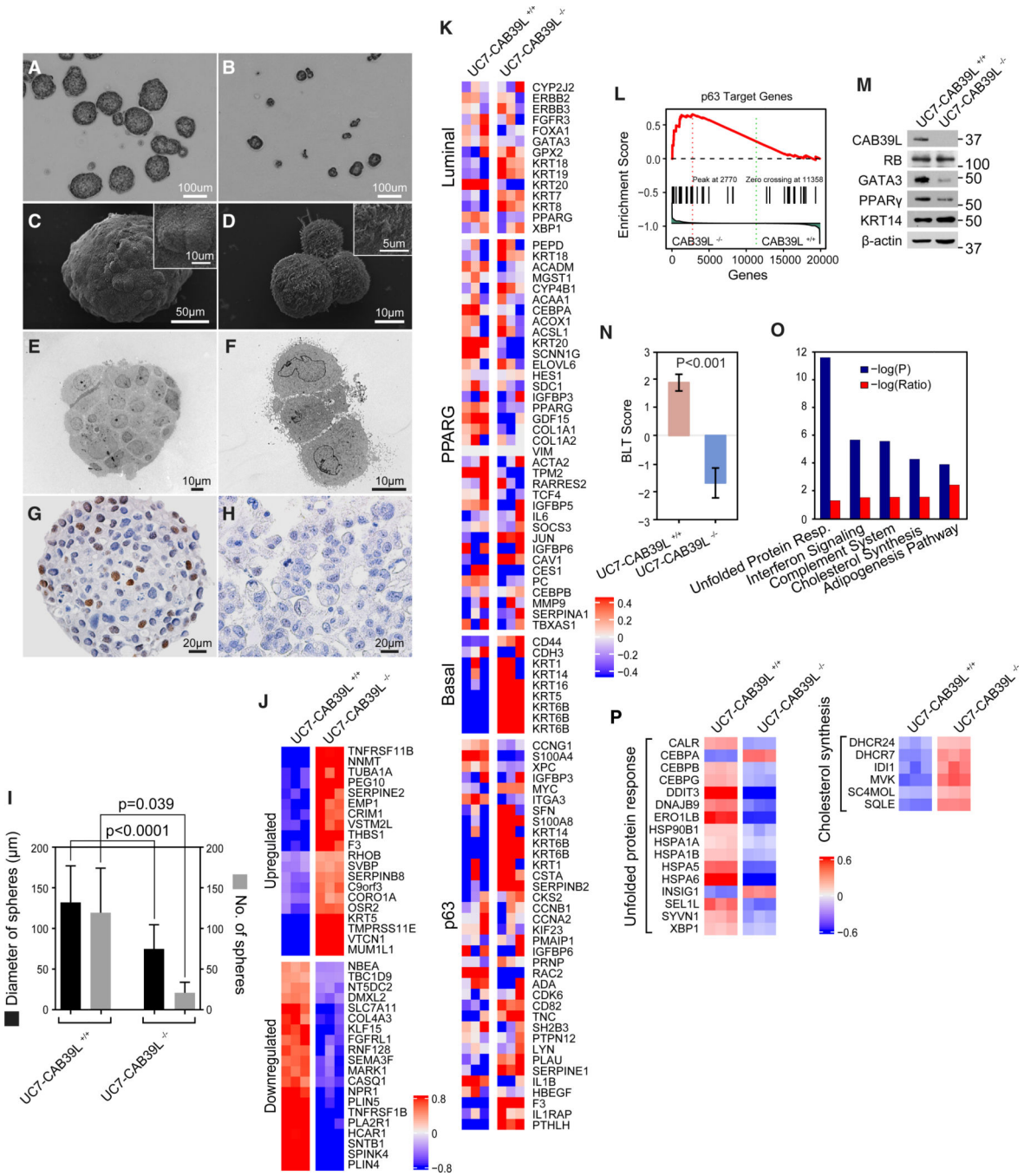


Figure 3. Effects of *CAB39L* ablation on the UC7 luminal bladder cancer cell line

(A) Formation of well-developed urospheres in *CAB39L*$^{+/+}$ UC7 cells. Scale bar, 100 μm.

(B) Dispersed growth pattern with formation of poorly developed clusters of *CAB39L*$^{-/-}$ UC7 cells. Scale bar, 100 μm.

(C) Scanning electron micrograph showing a well-developed urosphere in *CAB39L*$^{+/+}$ UC7 cells. Inset: the surface of the urothelial cells outlining a urosphere with cell membrane ridges. Scale bar, 50 μm.

(D) Scanning electron micrograph showing poorly developed, loosely arranged clusters of *CAB39L*$^{-/-}$ UC7 cells. Inset: the surface of a poorly differentiated urothelial cell with microvilli. Scale bar, 10 μm.
(E) Transmission electron micrograph showing a well-developed exosphere in *CAB39L*$^{+/+}$ UC7 cells. Scale bar, 10 μm.
(F) Transmission electron micrograph showing a poorly developed, loosely arranged cluster of *CAB39L*$^{-/-}$ UC7 cells. Scale bar, 10 μm.
(G) Positive staining for GATA3 in the nuclei of *CAB39L*$^{+/+}$ UC7 cells forming a urosphere. Scale bar, 20 μm.
(H) Negative staining for GATA3 in *CAB39L*$^{-/-}$ UC7 cells that lost the ability to form urospheres and grow in a dispersed pattern. Scale bar, 20 μm.
(I) Quantitative analysis of urosphere size and number in *CAB39L*$^{+/+}$ and *CAB39L*$^{-/-}$ UC7 cells. Bars represent mean + standard error. Number of replicates (n) = 3. The p values are based on two-tailed Student's t test.
(J) Heatmap showing the expression pattern for the top 20 upregulated and downregulated genes in *CAB39L*$^{+/+}$ and *CAB39L*$^{-/-}$ UC7 cells.
(K) Heatmap showing the expression pattern for luminal and basal markers in *CAB39L*$^{+/+}$ and *CAB39L*$^{-/-}$ UC7 cells.
(L) Gene set enrichment analysis of *p63* target genes in *CAB39L*$^{-/-}$ UC7 cells compared with *CAB39L*$^{+/+}$ cells.
(M) Western blots showing ablation of CAB39L protein in *CAB39L*$^{-/-}$ UC7 cells, intact expression of RB protein, and expression of representative luminal (GATA3 and PPARγ) and basal (KRT14) proteins.
(N) BLT scores for *CAB39L*$^{+/+}$ and *CAB39L*$^{-/-}$ UC7 cells. Bars represent mean ± standard error. Number of replicates (n) = 3. The p value is based on two-tailed Student's t test.
(O) Top five dysregulated pathways in *CAB39L*$^{-/-}$ UC7 cells.
(P) Expression pattern for representative unfolded protein response (left) and cholesterol synthesis genes (right) in *CAB39L*$^{+/+}$ and *CAB39L*$^{-/-}$ UC7 cells.

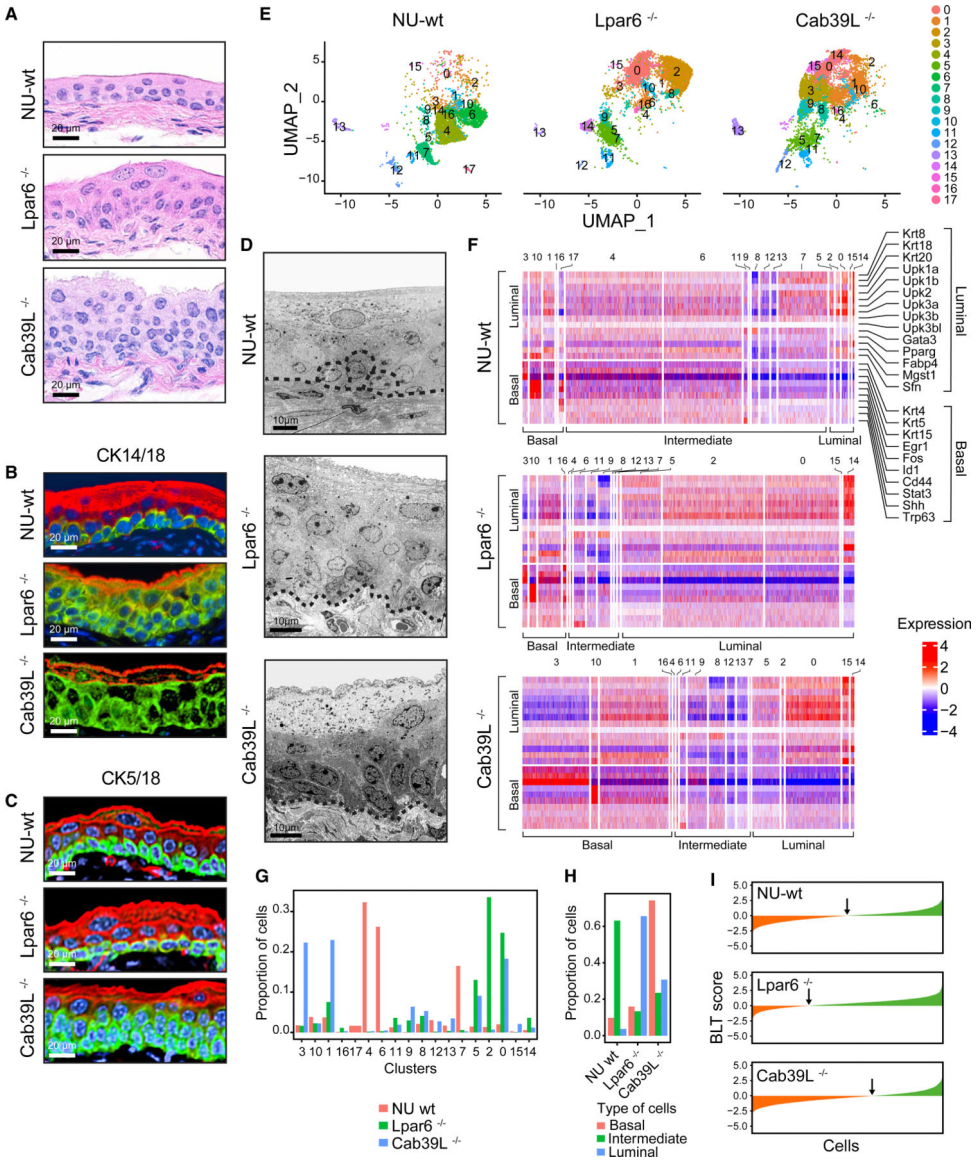


Figure 4. Single-cell analysis of the effects of *Lpar6* and *Cab39l* ablation on mouse urothelium

(A) Histology of the bladder urothelium in a wild-type mouse (NU-wt; top), *Lpar6*$^{-/-}$ mouse (middle), and *Cab39l*$^{-/-}$ mouse (bottom). Scale bars, 20 μm.

(B) Dual staining for cytokeratins 14 and 18 in an NU-wt mouse (top), *Lpar6*$^{-/-}$ mouse (middle), and *Cab39l*$^{-/-}$ mouse (bottom). Scale bars, 20 μm.

(C) Dual staining for cytokeratins 5 and 18 in an NU-wt mouse (top), *Lpar6*$^{-/-}$ mouse (middle), and *Cab39l*$^{-/-}$ mouse (bottom). Scale bars, 20 μm.

(D) Top: transmission electron micrograph of NU-wt mouse bladder urothelium with outlined basement membrane and basal as well as peribasal cells. Middle: transmission electron micrograph of the urothelium in an *Lpar6*$^{-/-}$ mouse with outlined basement membrane showing urothelial hyperplasia and gradual differentiation from basal-to-luminal cells. Bottom: transmission electron micrograph of the urothelium in a *Cabl39l*$^{-/-}$ mouse

with outlined basement membrane showing pronounced hyperplasia of undifferentiated basal cells. Scale bars, 10 μm.

(E) Single-cell sequencing and UMAP clustering of mouse bladder urothelial cells in NU-wt, *Lpar6*$^{-/-}$, and *Cab39l*$^{-/-}$ mice after filtration of nonurothelial cells.

(F) Expression pattern for luminal and basal genes in UMAP urothelial clusters of NU-wt, *Lpar6*$^{-/-}$, and *Cab39l*$^{-/-}$ mice.

(G) Proportion of cells in UMAP clusters of NU-wt, *Lpar6*$^{-/-}$, and *Cab39l*$^{-/-}$ mice.

(H) Proportion of basal intermediate and luminal cells in NU-wt, *Lpar6*$^{-/-}$, and *Cab39l*$^{-/-}$ mice.

(I) BLT score for NU-wt, *Lpar6*$^{-/-}$, and *Cab39l*$^{-/-}$ mice.

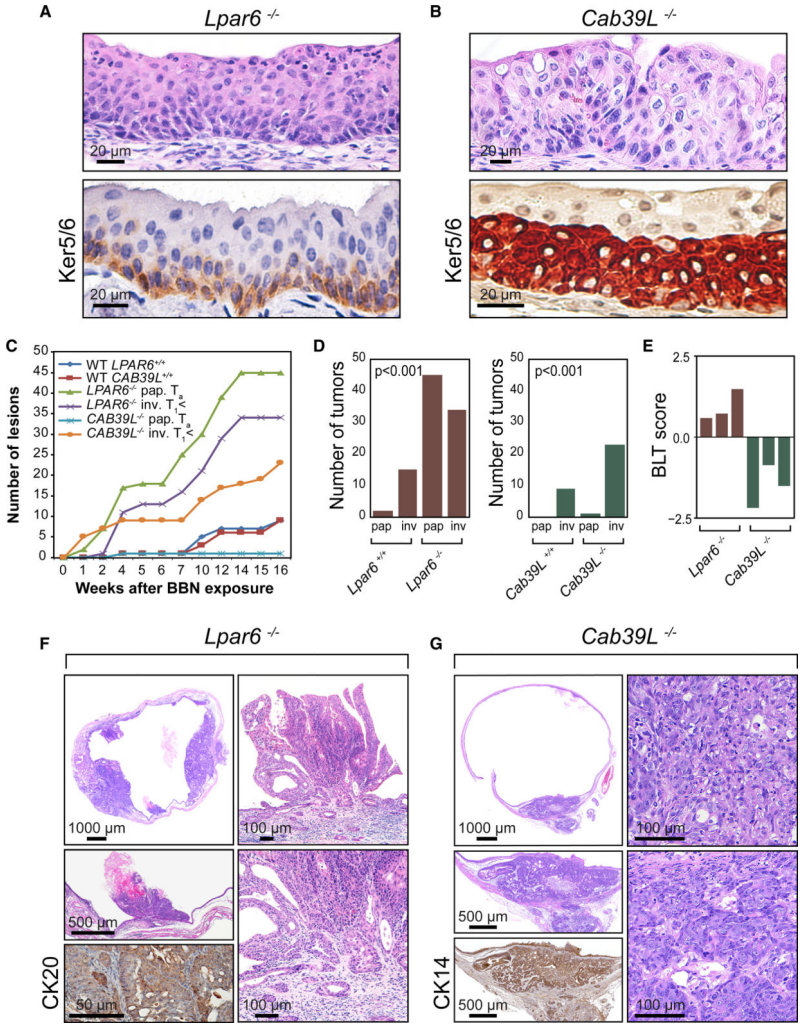


Figure 5. BBN-induced bladder cancers in *Lpar6*$^{-/-}$ and *Cab39l*$^{-/-}$ mice

(A) Top: high-grade intraurothelial neoplasia in an *Lpar6*$^{-/-}$ mouse 3 weeks after discontinuation of BBN exposure. Bottom: staining for keratins 5/6 in the same animal. Scale bars, 20 μm.

(B) Top: high-grade intraurothelial neoplasia in a *Cab39l*$^{-/-}$ mouse 3 weeks after discontinuation of BBN exposure. Bottom: staining for keratins 5/6 in the same animal. Scale bars, 20 μm.

(C) Frequency curves of BBN-induced tumors in *Lpar6*$^{-/-}$ and *Cab39l*$^{-/-}$ mice compared with *Lpar6*$^{+/+}$ and *Cab39l*$^{+/+}$ mice.

(D) Number of tumors developing after BBN exposure in *Lpar6*$^{-/-}$ compared with *Lpar6*$^{+/+}$ mice (left) and *Cab39l*$^{-/-}$ compared with *Cab39l*$^{+/+}$ mice (right). The *p* values are based on two-tailed Student's t test.

(E) BLT scores for the BBN-induced tumors in *Lpar6*$^{-/-}$ and *Cab39l*$^{-/-}$ mice.

(F) Upper left: development of multifocal, high-grade, invasive papillary tumors in an *Lpar6*$^{-/-}$ mouse 2 months after discontinuation of BBN exposure. Upper right: higher magnification image of a papillary tumor shown in the previous photomicrograph. Lower left: (top) another example of a papillary tumor; (bottom) positive staining for cytokeratin

20. Lower right: higher magnification of the base of the tumor in the upper right showing an invasive growth pattern. Scale bars, 1,000, 500, 100, and 50 μm.
(G) Upper left: development of multifocal, high-grade, invasive nonpapillary tumors 2 months after discontinuation of BBN exposure. Upper right: higher magnification image showing poorly differentiated invasive carcinoma. Left middle: lower magnification image showing an invasive growth pattern for carcinoma. Left bottom: expression of cytokeratin 14 in the same tumor. Right bottom: invasive carcinoma with focal squamous differentiation. Scale bars, 1,000, 500, and 100 μm.

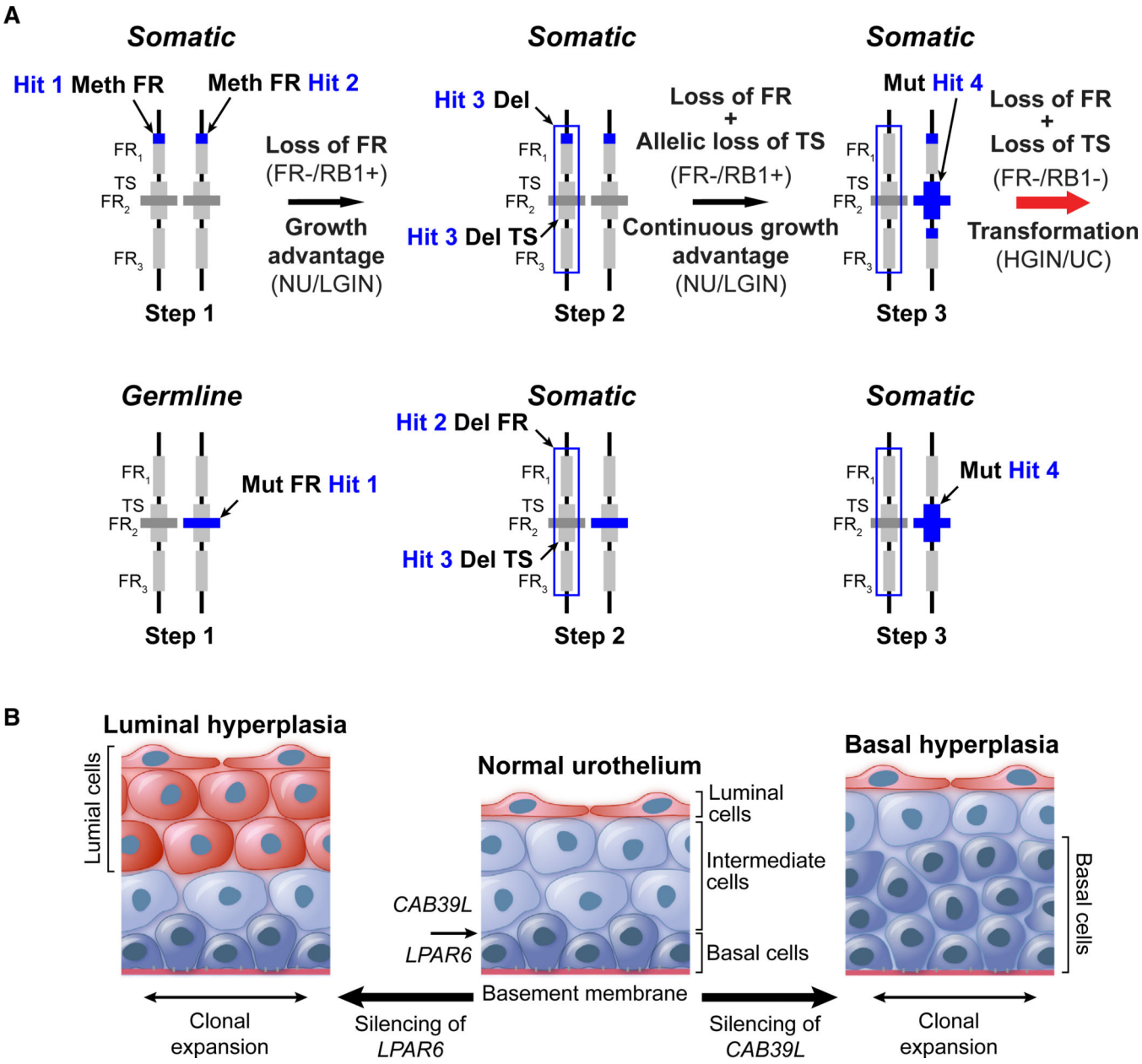


Figure 6. Bladder cancer initiation and progression induced by sequential inactivation of the FR genes and a tumor suppressor

(A) The four-hit, three-step mechanism of inactivation for neighboring FR genes (FR_1, FR_2, and FR_3) and tumor suppressor (TS) contributing to the initial expansion of low-grade intraurothelial neoplasia (LGIN) with progression to high-grade intraurothelial neoplasia (HGIN) and invasive cancer (urothelial carcinoma). Tumorigenesis begins with silencing of the FR gene, most often by somatic homozygous hypermethylation (hits 1 and 2) or a combination of loss of one copy and hypermethylation of the remaining gene allele. The loss of one alternative FR gene copy is usually synchronous with the loss of a neighboring tumor suppressor, such as *RB1* (hits 2 and 3). The second step may have several substeps in which additional FR genes are homozygously inactivated. Available

evidence suggests that, in rare instances, steps 1 and 2 can be reversed, as inactivation of the first FR gene allele can be accomplished by germline nucleotide substitutions, including population-based polymorphisms. Homozygous inactivation of FR genes is associated with clonal expansion of the *in situ* preneoplastic clone. In the third and final step, the remaining allele of a contiguous tumor suppressor is inactivated, most commonly by a mutation (hit 4). This step is associated with clonal evolution into the transformed phenotype (high-grade intraurothelial neoplasia) exhibiting features of carcinoma *in situ* and progressing to invasive bladder cancer (urothelial carcinoma).

(B) Biologic effects of *Lpar6* and *Cab39l* silencing in mouse urothelium causing luminal and basal hyperplasia, respectively.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-Lpar6	Santa Cruz	Cat# sc-20126; RRID: AB_653698
Anti-Rb1	QED Bioscience	Cat# 3101–3107
Anti-GATA3	Santa Cruz	Cat# sc-268; RRID: AB_2108591
Anti-PPARgamma	Cell signaling	Cat# 2435; RRID: AB_2166051
Anti-p63	Cell signaling	Cat# 13109; RRID: AB_2637091
Anti-beta actin	Sigma	Cat# A5441; RRID: AB_676744
Anti-Cab39L	abcam	Cat 197922
Anti-Krt14	abcam	Cat# 181595; RRID: AB_2811031
Anti-Krt18	abcam	Cat# 668; RRID: AB_305647
Anti-Krt5	abcam	Cat# 75869; RRID: AB_1310119
Anti-Krt5/6	Dako	Cat# GA78061–2
Anti-Krt20	Dako	Cat# GA77761–2
Secondary antibody- Alexa Fluor 488	Thermo Fisher Scientific	Cat# A-11034; RRID:AB_2576217
Secondary antibody- Alexa Fluor 546	Thermo Fisher Scientific	Cat# A-11030; RRID:AB_2534089
Anti-mouse IgG1	SouthernBiotech	Cat# 1070–01; RRID:AB_2794408
Biological samples		
Whole organ cystectomy specimens from patients with bladder cancer	UT MD Anderson Cancer Center	N/A
Fresh frozen primary tumor samples and paired peripheral blood DNA from patients with bladder cancer	UT MD Anderson Cancer Center	N/A
Chemicals, peptides, and recombinant proteins		
N-butyl-N-(4-hydroxybutyl) nitrosamine	Selleck Chemicals	E0199
5-Aza-2′-deoxycytidine	Sigma-Aldrich	A3656
Critical commercial assays		
Whole Exome Sequencing on Illumina NovaSeq 6000 Platform	Illumina	N/A

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Infinium MethylationEPIC BeadChip Kit	Illumina	GPL21145
Single cell sequencing	10X Genomics	N/A
Deposited data		
RNAseq raw and analyzed data	This paper	SRA:PRJNA891706
Nimblegen expression tiling array	This paper	GEO:GSE49024
Single cell sequencing	This paper	GEO:GSE235455
Lund cohort data	(Sjodahl et al.)[19]	GEO:GSE32894
RNA sequencing of UC cell lines	(Choi et al.) [15]	GEO:GSE47992
Human reference NCBI build 38, GRCh38	Genome Reference Consortium	https://www.ncbi.nlm.nih.gov/assembly/grc/human
Human Genome Browser, GRCh38	University of California Santa Cruz Genomic Institute	https://genome.ucsc.edu/cgi-bin/hgGateway
Experimental models: cell lines		
UM-UC-1	Dr. Colin Dinney at MDACC	RRID:CVCL_2743
UM-UC-2	Dr. Colin Dinney at MDACC	RRID:CVCL_8155
UM-UC-3	Dr. Colin Dinney at MDACC	RRID:CVCL_1783
UM-UC-6	Dr. Colin Dinney at MDACC	RRID:CVCL_2751
UM-UC-7	Dr. Colin Dinney at MDACC	RRID:CVCL_2752
UM-UC-9	Dr. Colin Dinney at MDACC	RRID:CVCL_2753
UM-UC-10	Dr. Colin Dinney at MDACC	RRID:CVCL_2744
UM-UC-12	Dr. Colin Dinney at MDACC	RRID:CVCL_S992
UM-UC-13	Dr. Colin Dinney at MDACC	RRID:CVCL_2746
UM-UC-15	Dr. Colin Dinney at MDACC	RRID:CVCL_R742
UM-UC-16	Dr. Colin Dinney at MDACC	RRID:CVCL_2748
UM-UC-17	Dr. Colin Dinney at MDACC	RRID:CVCL_R743
SV-HUC-1	Dr. Catherine A. Reznikoff	RRID:CVCL_3798
TERT-NHUC	Dr. Margaret Knowles	RRID:CVCL_JX41
RT-4	Dr. Colin Dinney at MDACC	RRID:CVCL_0036
RT-112	Dr. Colin Dinney at MDACC	RRID:CVCL_1670
SCaBER	Dr. Colin Dinney at MDACC	RRID:CVCL_3599

REAGENT or RESOURCE	SOURCE	IDENTIFIER
UM-UC-6-Lpar6 KO	This paper	N/A
UM-UC-7-Cab39L KO	This paper	N/A
RT-4- Cab39L KO	This paper	N/A
RT-112- Cab39L KO	This paper	N/A
SCaBER- Lpar6 KO	This paper	N/A
UM-UC-15- Lpar6 KO	This paper	N/A
Experimental models: Organisms/strains		
Mouse C57BL/6NCr/LPAR6 KO	This paper	*Lpar6* −/−
Mouse C57BL/6NCr /Cab39L KO	This paper	*Cab39L* −/−
Oligonucleotides		
gRNA to knock out of *LPAR6* TTTCCGCTGGGTTCTTCAACNGG	This paper	N/A
gRNA to knock out *LPAR6* GTCAATGACCGCATAAACGANGG	This paper	N/A
gRNA to knock out *CAB39L* ATCCTGTTTATGCTCCTCAANGG	This paper	N/A
Oligonucleotides for methylation analysis are provided in Table S4	This paper	N/A
Recombinant DNA		
the TA vector pCR 4-TOPO	Invitrogen	K4575–02
Software and algorithms		
RSeQC software	(Wang et al.) [53]	http://rseqc.sourceforge.net/
FastQC software	Babraham Bioinformatics	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
FeatureCounts software in the Subread package	(Liao et al.)[54]	http://subread.sourceforge.net/
DESeq2 software (version 1.26.0)	(Love et al.)[55]	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
Illumina GenomeStudio software (version 2.0)	Illumina	https://support.illumina.com/downloads/genomestudio-2-0.html
Gene set enrichment analysis (GSEA)	(Subramanian et al.)[56]	https://software.broadinstitute.org/cancer/software/gsea/wiki/index.php/Main_Page

Cell Rep. Author manuscript; available in PMC 2024 July 23.