Rafael Stroggilos[1,3], Maria Frantzi[2], Jerome Zoidakis[1], Emmanouil Mavrogeorgis[1÷], Marika Mokou[2], Maria G Roubelakis[3,4], Harald Mischak[2], Antonia Vlahou[1*]

1. Systems Biology Center, Biomedical Research Foundation of the Academy of Athens, Athens, Greece.
2. Mosaiques Diagnostics GmbH, Hannover, Germany.
3. Laboratory of Biology, National and Kapodistrian University of Athens, School of Medicine, Athens, Greece
4. Cell and Gene Therapy Laboratory, Biomedical Research Foundation of the Academy of Athens, Athens, Greece.

÷current address: Mosaiques Diagnostics GmbH, Hannover, Germany

* Corresponding author:

Antonia Vlahou, PhD

Biomedical Research Foundation, Academy of Athens

Soranou Efessiou 4,

11527, Athens, Greece

Tel: +30 210 6597506

Fax: +30 210 6597545

E-mail: vlahoua@bioacademy.gr

TITLE

Gene expression and coexpression alterations marking evolution of bladder cancer

ABSTRACT

Despite advancements in therapeutics, Bladder Cancer (BLCA) constitutes a major clinical burden, with locally advanced and metastatic cases facing poor survival rates. Aiming at expanding our knowledge of BLCA molecular pathophysiology, we integrated 1,508 publicly available, primary, well-characterized BLCA transcriptomes and investigated alterations in gene expression with stage (T0-Ta-T1-T2-T3-T4). We identified 157 genes and several pathways related prominently with cell cycle, showing a monotonically up- or down-regulated trend with higher disease stage. Genome wide coexpression across stages further revealed intrinsic and microenvironmental gene rewiring programs that shape BLCA evolution. Novel associations between epigenetic factors (CBX7, ZFP2) and BLCA survival were validated in external data. T0 together with advanced stages were heavily infiltrated with immune cells, but of distinct populations. We found AIF1 to be a novel driver of macrophage-based immunosuppression in T4 tumors. Our results suggest a continuum of alterations with increasing malignancy.

INTRODUCTION

Bladder Cancer (BLCA) accounts for approximately 200,000 annual deaths worldwide and is considered the most expensive cancer to manage[1]. Advances in imaging technologies and drug discovery have improved patient survival and quality of life [1]. However, early-stage incidents [classified as non-muscle invasive (NMI)], still suffer from high rates of disease recurrence, whereas advanced stage [classified as muscle invasive (MI)], and metastatic cancers face poor outcomes [2].

Advances in state-of-art molecular profiling technologies have enabled deeper investigations of BLCA, expanding our current understanding of its molecular pathology. According to the dual track model of bladder carcinogenesis [3], papillary-NMI and MI disease develop from different sets of molecular alterations. Studies performing mutational profiling suggest that low grade Ta tumors arise from activating mutations in either FGFR3 or HRAS, which typically result in over-activation of the downstream Akt/PIK3CA/mTOR and RTK/MAPK growth pathways [3]. In contrast, MIBC tumors are thought to develop from dysplastic Tis lesions with non-functional (mutated) TP53 or RB1 tumor suppressive pathways [3]. However, it remains unclear how these mutational signatures translate to different gene expression routes. Moreover, the dual track model cannot explain adequately molecular events driving transformation of a papillary-NMI tumor to MI, nor in the case of MIBC, alterations happening before the dissemination to detrusor muscle.

In an effort to better understand molecular pathogenesis, various BLCA molecular subtypes have been described [4-15]. Data supporting a continuum of alterations that likely drive bladder carcinogenesis came from MI patients having tumors with a mosaic of both intraepithelial (Tis) and papillary growth patterns, and from MI cancers having traits of papillary-NMI related mutations. Approximately 22% of MI tumors present with activating mutations in PI3KCA and homozygous deletion of the CDKN2A locus, respectively [9], while CDKN2A deletion in MIBC has been observed to occur more frequently in FGFR3 mutated than wild type tumors [16]. Interestingly, comparative mutational analysis between low grade NMI, high grade NMI and MIBC revealed smooth increments or declines in the frequency of mutations in driver genes (FGFR3, KDM6A, TP53, CDKN2A) with increasing malignancy [17]. Additionally, multi-omics analysis of NMIBC identified dysfunctional TP53 and RB1 pathways in about ~25% of both Ta and T1 tumors [15].

The clinical distinction between NMI and MI diseases and the current understanding of their molecular determinants cannot describe adequately the events driving tumor evolution. On the other hand, molecular subtypes may have important utilities in the diagnosis, prognosis, and decision making, but unfortunately, they represent static entities and their dynamics can only be studied between baseline diagnosis and future recurrences/progressions. In contrast, stage as an ordinal variable reflecting tumor size and depth of invasion, offers a better opportunity for studying the progressive alterations marking tumor initiation, growth, dissemination to detrusor muscle and metastasis.

Gene expression studies comparing the stage profiles of BLCA typically involve small sample sizes and are often limited to comparing NMI and MI. To obtain insight into the trajectories of cancer evolution in BLCA, we collected publicly available transcriptomes from bulk tissue samples, and performed a comprehensive stage analysis of 1,508 subjects with primary BLCA, including a novel integrated pathway-to-network analysis. As the disease progresses gradually to higher stages, our results indicate tumor dependencies on concerted alterations of gene expression, with most prominent those involved in cell cycle regulation.

METHODS

Dataset mining

A comprehensive data mining strategy was employed to retrieve studies applying -omics technologies in BLCA. The overall workflow is summarized in **Figure 1**. All genomic urothelial cancer data from cBioportal (including The Cancer Genome Atlas) were downloaded (5/1/2020). Gene Expression Omnibus (GEO) was queried for transcriptomics, additional genomics or protein array datasets using the search terms “bladder cancer” and “urothelial carcinoma”. We also queried ArrayExpress using the special filter “Array express data only” to obtain any additional datasets missing from GEO. All cohort data published or updated between 2010 and 2019, annotated as Homo sapiens, coming from tissue samples with sample size $\geq$10, were initially retrieved (25/1/2020). All used datasets were published and downloaded anonymized.

Integration of the transcriptomes

Microarray data were summarized to gene level with the package *oligo* [18]. Affymetrix were normalized with the RMA method, while Illumina were filtered based on detection p value < 0.05, followed by quantile normalization, addition of 1, and transformation to the natural logarithmic scale. Data were annotated using *biomaRt* (v2.42.1) [19] and mircroarray probes matching to multiple genes were excluded from downstream analysis. The probe with the highest mean across arrays was selected as representative in cases where multiple probes matched to the same gene. Merging of expression matrices was based on the Hugo Gene Symbol using the intersection of genes between studies.

Adjustment for batch effects

To correct for batch effects across different studies in the discovery data, *ComBat* [20], *removeBatchEffect* [21], and *naiveRandRuv* [22] were evaluated. ComBat performed best and was chosen for further use (manuscript in preparation). The quality of corrected data was assessed with *BatchQC* [23] (**Supplementary Figures 1-2)**, with boxplots of expression distribution per sample and principal component analysis plot of sample relationships ,with gene expression comparisons of housekeeping to other genes, and with a set of 12 BLCA markers with known regulation across normal-NMI-MI or across normal-low grade- high grade disease (**Supplementary Figure 3**).

Differential expression, monotonicity and functional annotation

To identify genes that form a continuum of changes across BLCA stages, each disease stage was compared against normal samples. Genes significantly affected (Mann-Whitney $p < 0.05$ and had also same orientation of change in all comparisons), were extracted (n = 3,108 genes). We refer to this set as Concordantly Differentially Expressed Genes (CDEGs). Monotonicity for a gene was defined as being a CDEG and additionally having a continuously larger/smaller fold change with increasing stage (fold change, as defined in each disease stage versus normal control). Functional annotation and enrichment were performed using PubMed and the online tool GeneCards (https://ga.genecards.org/), respectively.

Monotonicity in pathway activation and analysis for gene coexpression

We utilized CDEGs to infer pathway activation scores and to create stage coexpression networks. Pathway activation scores per sample were calculated with the *ssGSEA-GSVA* method [24], using the Molecular Signature Database libraries of Hallmark, Canonical Pathways (Reactome subset), C3 (GO biological processes subset) and C5 (GTRD subset of transcription factor targets). *Dorothea* (https://github.com/saezlab/dorothea) [25] was utilized to assess regulon activity. Activation scores across stages were compared with Mann-Whitney tests, and direction of change was defined based on fold change (= Mean of stage – Mean of T0). Pathways showing a monotonal change in the activation scores with increasing stage were extracted. Monotonicity for a pathway was defined similarly to the previous section, i.e. being significantly different in all stage comparisons to T0, and also having a continuously larger/smaller fold change with increasing stage. Stromal infiltration scores were imputed with the ESTIMATE algorithm [26].

Gene-pair coexpression weights (non-negative) were approximated with ensemble learning, using GENIE3 [27] and direction of coexpression (positive/negative) was determined by Spearman's coefficient. Stage specific networks were constructed with igraph using the top 5,600 positively correlated gene pair interactions with the highest weights for each of the stages. Networks were analysed with Louvain clustering [28] to identify local modules of significantly higher coexpression relatioships (communities) within each stage. The top 5 in size (= number of genes) communities of coxpressed genes per stage were analysed for Gene Ontology Biological Processes [29]. Significance was thresholded at $p<0.05$ after FDR adjustment according to Benjamini-Hochberg. Hub genes in the monotonal networks were defined based on the betweenness centrality metric.

Statistical analysis

For continuous variables, significance was defined at Mann-Whitney $p < 0.05$, unless stated otherwise. Categorical variables were investigated for significance with the Chi-squared test and were adjusted for multiple hypothesis (CRAN package *RVAideMemoire*). All statistical tests were two sided. Gene Ontology and Reactome pathway enrichment analysis was conducted with functions from the package *ClusterProfiler* that performs a two-sided hypergeometric test. All reported correlation scores correspond to the Spearman's Rank coefficient. Kaplan-Meier analysis and cox proportional hazards regression were performed with the CRAN packages *survminer* and *survival*, and statistical significance was determined with the log-rank method. CIBERSORT analysis was conducted in the web platform https://cibersort.stanford.edu/, and only samples with successful deconvolution ($p < 0.05$, n = 350 samples) were further used for the comparisons of relative immune populations among stages. All processing, analyses and visualizations were created in the R programming language (version 4.0.2), using predominantly Bioconductor libraries.

Technology	Platform	Dataset ID	T0	Ta	T1	T2	T3	T4	Tis	Study size
Affymetrix	GPL17586	GSE121711	10	3	2	3				18
Affymetrix	GPL17586	GSE93527		79						79
Affymetrix	GPL570	E-MTAB-1940	4	41	41					86
Affymetrix	GPL570	GSE31684		5	10	16	41	18		90
Affymetrix	GPL6244	GSE104922		19	7	10	5	0		41
Affymetrix	GPL6244	GSE128959		13	39				1	53

Affymetrix	GPL6244	GSE83586		13	44	241	1	1	1	301
Illumina	GPL14951	GSE48276			2	6	23	6		37
Illumina	GPL14951	GSE52219				15	7	1		23
Illumina	GPL14951	GSE69795			3	7	27			37
Illumina	GPL6102	GSE13507	67	23	80	31	19	12		232
Illumina	GPL6947	GSE48075		33	34	42	23	8	2	142
		Stage size	81	229	262	371	146	46	4	1139

Table 1: Stage distribution across the microarray datasets used for the discovery phase.

RESULTS

Cohort description and data quality controls

To perform a comprehensive investigation of publicly available omics studies, we collected 105 datasets comprising more than 8,000 individuals (**Figure 1**). Minimum inclusion criteria for datasets were defined as having at least stage information per subject. To maintain high integrity, we selected microarray data quantified by commercially available single-color channel vendors (Affymetrix and Illumina). Samples or datasets collected after administration of neoadjuvant chemotherapy (NAC) or not clearly annotated as primary were excluded, generating a final dataset corresponding to 1,508 patients. This included in total 12 microarray studies which were employed as a discovery set (**Table 1**) and in addition, the TCGA 2017 BLCA project based on RNA-seq analysis was used as a validation set (**Figure 1**).

The compiled discovery cohort comprised 1,058 primary bladder cancer tumor transcriptomes of treatment-naïve patients without prior cancer history, along with 81 adjacent to the cancer bladder samples, a total of 1,139 gene expression profiles (**Table 1**). The ratio of men:woman was 3.5:1, with equal distribution among NMI and MI disease ($p = 0.99$) and similar mean ages at baseline diagnosis (68 years, $p = 0.81$). Percentages of normal samples, NMI, and MI in the dataset were *7.1%*, *43.5%,* and *49.4%,* respectively (**Table 1**), with the grade distribution being as follows: *16.5%* low grade, *48.8%* high grade disease with the remaining samples lacking available grade information. Detailed histological records were missing for 71.5% of the cohort, with the most frequently reported histology among the available records being urothelial/papillary (= 23.3%) with squamous differentiation being the most frequent variant (= 1.3%). Annotations included mutation status for FGFR3, PI3KCA, RAS, RB1 and TP53.

Removing batch effects with ComBat was assessed with Relative Log Expression and Principal Component Analysis plots (**Figure 2A**), as well as with comparative analysis of expression levels between houekeeping and non-housekeeping genes (**Figure 2B**), with the algorithm BatchQC (**Supplementary Figures 1 and 2**) [23], or using a set of 12 positive BLCA markers with known regulation across normal-NMI-MI or normal-low-high grade disease (**Supplementary Figure 3**). ComBat successfully eliminated batch effects while maintaining the biological variability (**Figures 2A-C**). BatchQC reports indicated that the variability in the corrected data was generally explained by stage rather than the batch variable **(Supplementary Figure 2)**, allowing for an in-depth downsteam analysis.

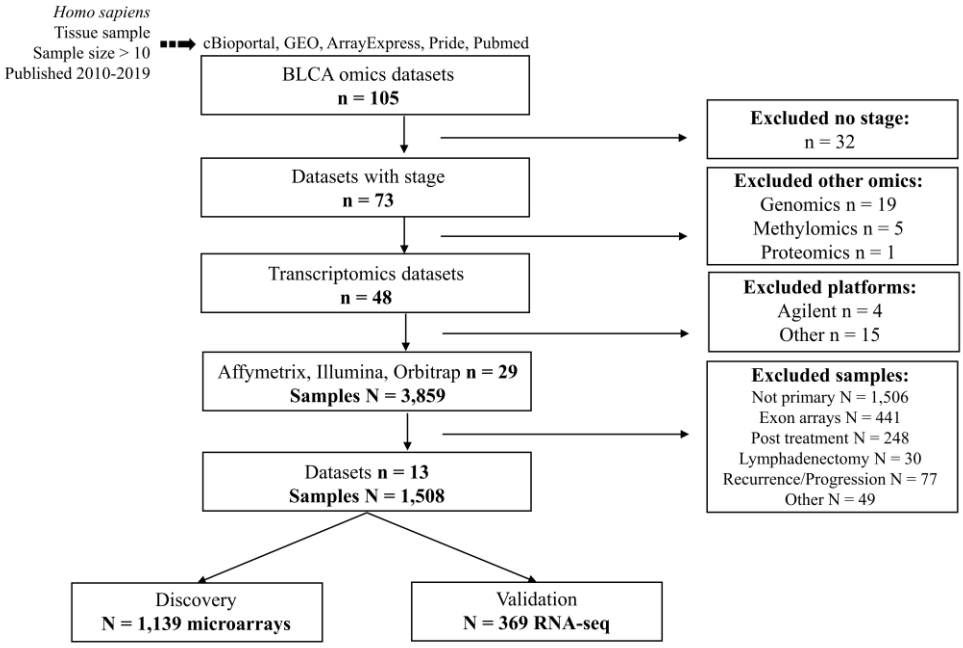

Figure 1: Study design and workflow for the integrative analysis of primary BLCA transcriptomes.

Monotonically changing genes in BLCA regulate growth and cell-cycle progression

For initial assessment of the gene expression relations between normal adjacent urothelium (referred to as T0) and cancerous samples, we performed differential expression analysis of T0 versus NMI and T0 versus MI cancers. When compared to T0, a strong positive correlation ($r=0.83$, $p < 2.1\times10^{-16}$) was observed between the two sets of log fold change values, suggesting mutual concordance in abundance change and directionality between NMI and MI tumors. To investigate transcriptional changes associated with increasing malignancy, CDEGs were extracted from the comparisons Ta vs T0, T1 vs T0, T2 vs T0, T3 vs T0 and T4 vs T0. A total of 157 genes were identified having a monotonal (i.e. continuously increasing or decreasing in comparison to T0) change in expression across stages, of which 118 were up- and 39 were downregulated with increasing stage (**Table 2**). Functional annotation revealed that for 46 of these genes, experimental evidence on mediating cell cycle progression exists (Table 2; **Supplementary Table 1**). Upregulated cell-cycle associated genes (n = 44) were not phase specific and included cyclins, DNA polymerases, regulators of the cohesin complex and kinetochore components. The two downregulated cell-cycle associated genes, BTG2 and CDC14B, are both tumor suppressors. The list included 23 genes involved in signal transduction (**Table 2**), 6 of which (ARHGAP11A, AURKA, CDKN3, PBK, PLK1, RRM2), promote cell-cycle progession and were all upregulated with increased stage. The data also indicated an overactivation of the Wnt pathway with increasing disease stage, with its upstream inhibitor APCDD1 being downregulated and its activating ligand WNT2 upregulated (**Supplementarty Table 1**). Fourteen of the 157 genes were transcriptional or translational regulators (**Table 2**), including genes with known upregulation in bladder cancer (transcription factors E2F1, DEPDC1 [30, 31]). Based on the monotonal changes with

increasing stage, increased androgen receptor activity may be predicted, as both its translational enhancer BUD31 [32] and its downstream transcritption factor ELK1 [33] were upregulated. Four (HTR2C, LRP8, NENF, NMU) of the 157 genes are involved in neurotransmission or neuronal development, all upregulated (**Table 2**). Among the 157 genes, 21 genes were of not well described or unknown function (**Supplementary Table 1**), including the oncogenic factor TRIM65 [34] found upregulated with bladder cancer stage. Additionally, upregulation was detected for the cisplatin resistance gene CLPTM1L, as well as for SRD5A1, which catalyzes the conversion of testosterone and progesterone to their 5-alpha forms [35]. Further functional enrichment using GeneCards for the 157 genes verified their involvement in cell-cycle pathways, with top hits being related to the regulation of the Anaphase-promoting (APC) complex (score = 31.53), to PLK1 (score = 24.47) and Aurora B (score = 20. 95) signaling, as well as to TP53 (score = 19.06) and RB1 (score = 17.85) cell cycle checkpoint control (**Supplementary Table 2**). Using the same tool, analysis for compounds with potentially therapeutic benefit provided as top hit the DNA alkylating agent Bendamustine (score = 19.23) (**Supplementary Table 3**) [36]. Univariate cox regression analysis indicated 29 genes with potentially prognostic impact at $p < 0.01$ (**Supplementary Table 4**).

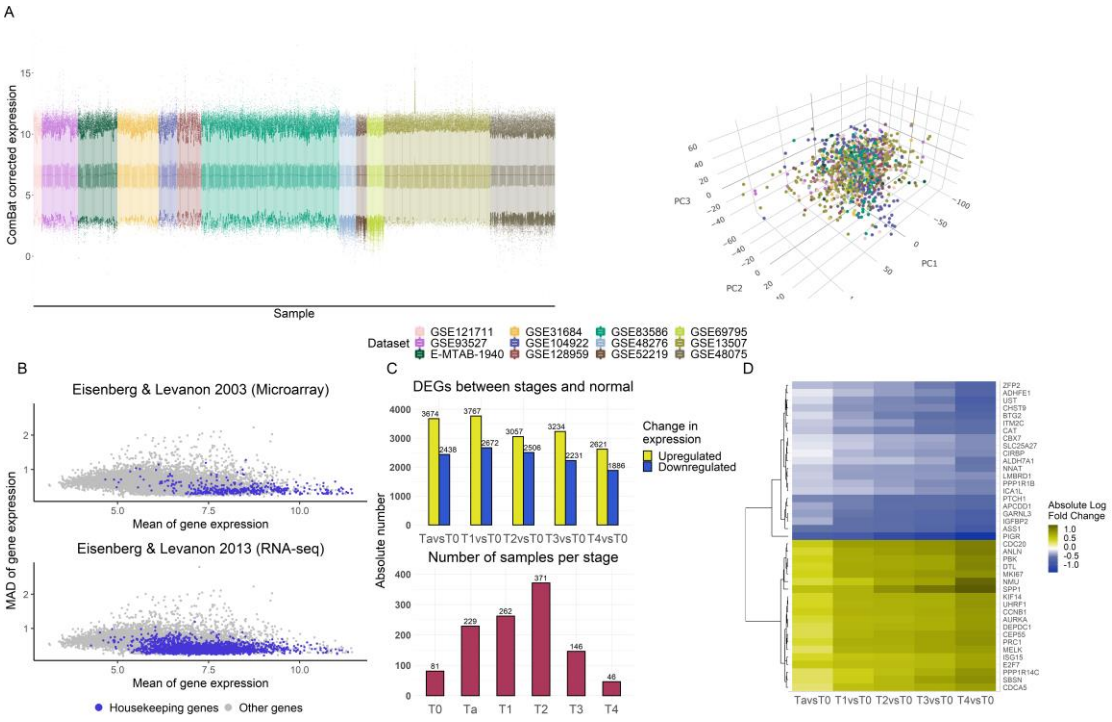

Figure 2: ComBat corrected expression data of the discovery cohort and respective differential gene expression analysis A) Relative Log Expression (right) and Principal Component Analysis plots (left) showing gene expression levels per sample and sample relationships, respectively; the different colors denote the different datasets B) dispersion-to-expression plots illustrating preservation of housekeeping gene properties (higher mean expression and lower median absolute deviation (MAD) compared to non-housekeeping genes) in the corrected data, regardless of microarray (top) or RNAseq (bottom) defined housekeeping genes; C) numbers of differentially expressed genes (DEGs) between disease stages and normal adjacent urothelium (top), along with respective group sizes (bottom), E) heatmap of the top (based on fold change) 20 downregulated and top 20 upregulated monotonically changing genes with increasing disease stage.

Function	Genes
Cell Cyle	ANLN, ARHGAP11A, AURKA, CDC20, CDCA3, CDCA5, CDCA8, CDK2AP2, CDKN3, CENPA, CENPN, CEP72, KIF23, KIFC1, NEIL3, OIP5, PRC1, PTTG1, RAE1, RUVBL2, SAC3D1, SMC4, FBXO5, CCNA2, CCNB1, CCDC124, CEP55, MKI67, MND1, GINS2, MCM10, MCM4, KIF11, KIF14, NDC80, POLD1, POLE2, PBK, PLK1, RRM2, E2F1, DAXX, E2F7, DTL, BTG2, CDC14B
Signal transduction	ARHGAP11A, AURKA, CDKN3, HTR2C, ITPR3, LRP8, MELK, NENF, NMU, NOX4, PBK, PLK1, PPP1R14C, RRM2, WNT2, AKAP7, APCDD1, IGFBP2, KIT, MTMR9, PPP1R1B, PTCH1, TAPT1
Transcriptional regulation	DAXX, DEPDC1, E2F1, E2F7, ELK1, FOXA3, HOXC6, HOXC9, MED19, TEAD4, AEBP2, CBX7
DNA damage response	CDCA5, EXO1, NDC80, NTHL1, POLD1, POLE2, RAD51AP1, UBE2T
Ubiquitin mediated proteolysis	DTL, FBXW2, PSMB3, PSMC1, TRAF2, DET1
Immunity	ISG15, SPP1, GBP6, IFNAR1, PIGR
Metabolism	CAD, DAGLA, ENO1, IMPDH1, NUDT1, SRD5A1, ADHFE1, ALDH7A1, ASS1, CAT, CHST9, CYP4V2, MGST1
Other roles	ARF5, ARPC5L, BUD31, CLPTM1L, EIF4EBP1, FAM91A1, GP6, PACSIN3, XPO5, CIRBP, GALNT12, RSL1D1, RWDD3, SLC25A27, UST
Homeostasis	ARHGDIA, BOP1, CTSA, GTPBP4, LMNB2, LYAR, MRPL37, NIP7, NXT1, POLR2D, POP7, RBM28, SRPRB, CTSO
Uncertain/Unknown	FAM50A, IGFL2, KIAA2013, MLF2, PAQR4, RECQL4, RP9, SBSN, SLC22A15, SNX8, TRIM65, YIF1A, FYCO1, GARNL3, ICA1L, ITM2C, LONRF1, NHS, NNAT, ZFP2, ZNF181

Table 2: Classification of the 157 monotonically changing genes with increasing disease BLCA stage to functional categories. Red and blue colors correspond to upregulation and downregulation, respectively. More details per gene are provided in **Supplementary Table 1**.

The expanded dataset of 3,108 CDEGs links cell-cycle, ECM remodelling and metabolic alterations with increasing disease stage

To increase coverage of the BLCA alterations, we further extracted Concordantly Differentially Expressed Genes (defined in Methods; CDEGs, n = 3,108) from the comparisons of cancerous stages to normal adjacent urothelium and investigated the differentially activated pathways they represent, using ssGSEA. Similar to the analysis of monotonicity, we sought to identify pathways with monotonically up- or down- regulated activation scores across T0 and BLCA stages. Results indicated gradually stronger activations of several mitotic processes, positive regulation of the canonical Wnt pathway, mTORC1 signaling, expression of MYC targets, degradation of anaphase inhibitors, metabolism of nucleotides, mobility of formins, and the TNFR2/non-canonical NF-kB pathway (**Figure 3, Supplementary Table 5**). Conversely, consistently reduced expression versus controls was observed in genes involved in lipid and fatty acid catabolic processes, in the metabolism of heme, in pathways promoting muscle cell differentiation and, interestingly, in genes regulating the circadian clock (**Figure 3**). Regulon activity per sample was estimated and respective scores between

cancerous stages and normal tissue were compared. GATA3 and GLI2 were the only regulons that significantly associated with increasing malignancy with their activity being downregulated (**Figure 3**).

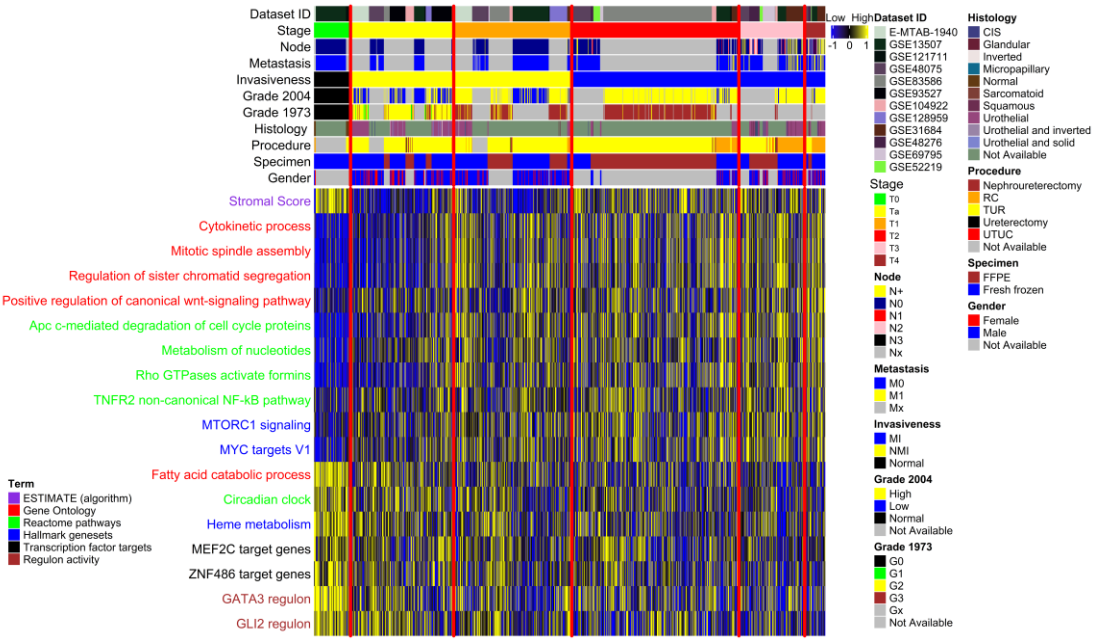

Figure 3: Excerpt of the pathways showing a monotonal increase or decline in their activation scores with higher stage.

To further investigate genome wide alterations and coexpression patterns associated with cancer progression, an integrated network-pathway analysis of the different disease stages was performed, using the 3,108 CDEGs. Out of 9,659,664 weighted gene pair interactions, the top 5,600 were used as the most representative to build the corresponding networks, applying Louvain clustering to identify topologically connected communities (modules). Retrieved gene clusters have high degree of coexpression, and thus changes in their composition across stages reflect gains/losses in pathway rewiring. We used the Gene Ontology – Biological Process library to identify molecular processes affected by changes in coexpression inside the top 5 largest in size (based on number of genes) communities (**Figure 4, Supplementary Tables 6-11**). The analysis revealed large differences in gene coexpression between T0 and cancerous samples with 4 out of the 5 largest in size communities associating clearly with specific biological processes.

Three out of the top 5 communities were consistently detected in all BLCA stage networks. Based on examination of their enriched processes, these were labeled as 1) the cell-cycle community, 2) the ECM and developmental community 3) the metabolic and translational community (**Figures 4A, 4B**)

When investigating the cell cycle community and specifically the processes of G1/S transition of mitotic cell cycle, cell-cycle G2/M phase transition, regulation of cell cycle phase transition, regulation of chromosome organization and nuclear division, most of the participating genes were found to be coexpressed in all BLCA stages (**Figure 4B**). However, the gene size of the cell cycle community increased with higher stage (Ta n = 118, T1 n = 148

and MIBC n = 168-170), and interestingly, the proportion of the genes with unknown or unclear function (lacking GO annotation) also increased (11.9% for Ta, 29.8% for T4). Among the genes with cell cycle annotation (n=184), 80 genes were coexpressed in all cancerous stages and were also upregulated in malignancy compared to T0, possibly forming the backbone of cell cycle progression in BLCA (Supplementary Table 13). However, only 36 of them had a monotonal increase in the expression levels with higher stage (Supplementary Table 13). To detect potential drivers of variation in the cell cycle co-expression profiles of stages, we looked into the betweenness centrality scores of the genes inside each of the stage cell cycle communities (Figure 4C), and found CDCA5 and KIF2C to be hub genes in Ta tumors, FOXM1 and AURKB in T1, CDT1 and SMC4 in T2, CCNB1 and RRM2 in T3 and KIF14 in T4 tumors (Figure 4C).

The community of ECM and developmental processes was enriched in cell-cell communication and cell-matrix interactions, in responses to microenvironmental stress, as well as in differentiation programs of epithelial, mesenchymal and stem cells (**Figure 5B**). The biological process of extracellular matrix organization included coexpressions of 15-36 genes of which COL13A1, FGFR4, FOXF2 and SCUBE1 were coexpressed only in T0 compared to disease stages, whereas 26 genes (including COL6A1/A2, COL16A1, MFAP5, MMP11) were coexpressed in malignancy but not in T0. In line with recent observations [37], we noticed that T0 presented with an active ECM remodeling profile. Sixteen of the ECM associated genes were coexpressed both in T0 and in the NMIBC stages, including genes promoting ECM degradation and tumor cell migration (MMP2, CTSK, PDPN), fibrotic collagens (COL1A2, COL6A3, COL14A1, COL15A1), and pro-angiogenic factors (PDGFRA, RECK). Notably, coexpression in the T0 samples was predicted to be driven by the cancer stem cell marker ALDH1A2 and MFAP4 (**Figure 4C**), while in Ta tumors, by COL16A1 and CLIP3. Since CLIP3 interacts with both AKT1 and AKT2 [38], CLIP3 may have an important role in the early AKT/PI3K/mTOR axis of carcinogenesis. For the same process of extracellular matrix organization, compared to Ta, T1 tumors had gains in ADAMTSL2, while compared to T1, T2 tumors had gains in MMP11 and LRP1. Similarly compared to T2, T3 tumors had coexpression gains in ITGA10 and in genes of the endocytic pathway (ABL1, CYP1B1) whereas compared to T3, T4 tumors had gains in DDR2 and JAM2.

The community of metabolism and translation encompassed mitochondrial, translational and multiple metabolic processes being activated during carcinogenesis, and was more profound in the T1 and more advanced tumors. The processes of cellular respiration, translational initiation, mRNA catabolic process, nonsense mediated decay and protein targeting to ER were consistently enriched in most of BLCA stages. The results highlighted a set of 12 genes commonly co-expressed across stages for these processes, including COX7B, DLD, NDUFS4, UQCRFS1, PAIP2, RPL15, RPL30, RPL7, RPS23, RPS27, RPS27A, RPS4X, with RPL36A and YTHDF3 being T2-specific, COQ10B, EIF4E3, MTIF2, NCPB1 as T3-specific while EIF4E1, ISCU and GSPT2 were T4-specific.

Besides the abovementioned consistently detected communities in BLCA, a community enriched in processes of immune cell activation and cytokine secretion, was identified in the T0 and the MIBC stages and involved both innate and adaptive responses, as well as, processes of immune cell adhesion and migration. It was the least variable community with respect to different processes and respective genes coexpressed across the different stages. Strikingly, T0 and the MIBC stages had similar profiles. Out of the 17 genes of the process of

T-cell activation that were commonly coexpressed at the MIBC stages, 15 were also coexpressed in the T0 samples. To further investigate these observations, the transcriptome data per sample were deconvoluted into relative abundances of immune cell populations using CIBERSORT, and cell fractions between disease stages were compared. Significant results were obtained for the following populations: CD8+, activated CD4+, activated NK, Monocytes, Macrophages M2 and activated Dendritic cells (**Supplementary Figure 4**). Results indicated differential commitment of immune cells to T0 and BLCA stages. T0 (n = 37) were significantly more infiltrated with CD8+ (p = 0.046) and monocytes (p = 4.6E-4) than tumor samples (n = 313), verifying the presence of the immune community in the T0 samples. However, compared to tumor, T0 samples had significantly less abundance of activated CD4+ cells (p = 5.28E-3), of macrophages (p = 16.8E-5), of activated dendritic cells (p = 0.0002) and of activated natural killer cells (p = 0.015). Generally, NMIBC had lower infiltration burden than MIBC, while compared to other BLCA stages, Ta tumors (n = 34) were significantly more infiltrated with activated dendritic cells (p = 0.024). Abundance of CD8+, of activated NK cells, and of M2 macrophages increased linearly with higher malignancy. Interestingly, AIF1 a gene that promotes macrophage survival [39], was found to be driver of immune coexpression in the T4 tumors.

A

Communities of coexpressed genes	T0	Ta	T1	T2	T3	T4
Cell Cycle	-	+ (105/118)	+ (124/148)	+ (134/170)	+ (130/170)	+ (118/168)
Extracellular matrix	+ (105/121)	+ (119/170)	+ (145/192)	+ (122/179)	+ (84/141)	+ (144/201)
Metabolism and translation	-	+ (67/134)	+ (119/220)	+ (65/123)	+ (44/165)	+ (95/221)
Immune activation	+ (87/108)	-	-	+ (95/124)	+ (108/138)	+ (97/123)

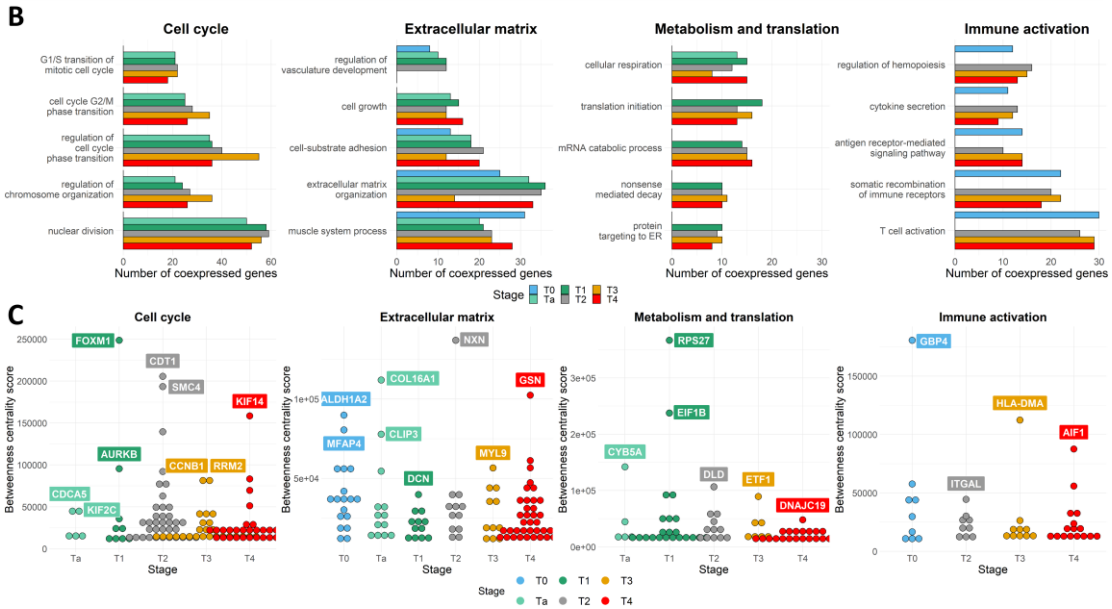

Figure 4: Biological process analysis of the largest in size coexpressed communities identified in each BLCA stage, using the 3,108 CDEGs between T0 and cancerous stages. **A)** Coherent communities identified and characterized across T0 and disease stages. Numbers in parentheses show the fraction of genes with existing Biological Process annotation with respect to the total number of genes found to be coexpressed in each of the community **B)** Barplots of the most significantly enriched biological processes per community depicting

number of coexpressed genes for each. **C**) Hub genes identified across the studied conditions based on the betweenness centrality scores (y axis).

RNA-seq independent analysis validates presence of communities and prognostic genes

In lack of another dataset comprising all the disease stage spectrum of primary BLCA incidents, the observed alterations in the discovery set were investigated for their reproducibility in the TCGA-BLCA publication [9]. Although the validation set is not perfectly suited since it includes only stages T2-T3-T4, we used it due to the RNA-seq technology which offers a good reliability at the validated findings. Differential expression analysis among the available stage comparisons (T3 vs T2 and T4 vs T3) in the TCGA data validated 74 of the 157 monotonal genes in the comparison T3vsT2, and none in the comparison T4vsT3 (**Supplementary Table 13**). Cox regression analysis of the 157 genes in the TCGA data validated 12 genes having a prognostic value (**Supplementary Table 4**), including IMPDH1, MRPL37, MED19, ENO1, ANLN, GTPBP4, MLF2, higher levels of which associating with worse survival and CBX7, ZFP2, AKAP7, ICA1L, CDC14B higher levels of which associating with better survival probability. To validate as possible the coexpression analysis findings, stage specific coexpression networks were also created using the TCGA data, and were clustered with the Louvain algorithm. GO-Biological process analysis of the communities validated the differential segregation of the cell-cycle, extracellular matrix and immune activation processes to distinct communities (**Supplementary Figure 5**).

DISCUSSION

Early integration of raw BLCA data has been previously performed, in the context of characterizing molecular subtypes [40], or validating results of either scRNA-seq [41] or RNA-seq re-analysis [42]. In this study we performed an early integration meta-analysis of normal looking urothelium and BLCA stages, aiming to identify continuous gene expression alterations with increasing malignancy. To our knowledge, this is the first attempt to associate molecular alterations with clinical classification, by analyzing more than one thousand well-characterized, primary samples. Instead of focusing onto molecular subtypes, we increased power and addressed the disease as a continuum, under the assumption that individual samples reflect different snapshots of the whole process. Our novel design based on the hypothesis of continuous evolution through stages, has been successfully applied here and resulted in novel findings on gene regulation associated with cancer progression.

First, we identified a set of 157 genes with a monotonal change in expression across T0 and BLCA stages. Almost half of these genes were components of the cell cycle machinery, or kinases signaling positively for it, or transcription factors responsible for the expression of cell cycle genes. Twelve of the 157 genes had also prognostic value in external RNA-seq data. Of these, apart from ENO1 higher levels of which had been previously linked with worse BLCA outcome [43], all the remaining 11 associations to survival are novel. IMPDH1 plays a role in cancer progression by promoting cell growth, and higher expression levels of this gene have been observed in MIBC compared to normal tissue [44]. Similarly, MED19 a component of the mediator complex that regulates the transcription of RNA-polymerases, was found overexpressed by IHC in human BLCA compared to normal tissues, and its knockdown in the T24 and 5637 bladder cell lines resulted in cell-cylce arrest at the G0/G1

checkpoint and also in attenuation of cell growth [45]. The involvement of MRPL37, GTPBP4 and MLF2 in BLCA development has not been characterized, but oncogenic properties have been attributed to these genes in other contexts. ANLN and AKAP7 are thought to regulate bladder cell growth and apoptosis in a TP53 independent manner [46]. ICA1L is naturally expressed in sperm cells and its implication in BLCA has not been investigated. The CDC14B gene is located on the 9q chromosome, a region that is often deleted in BLCA. This might also explain its downregulation in malignancy, as observed in the discovery set. CDC14B is believed to dephosphorylate TP53 [47], but the functional consequence on the mitotic or DNA damage repair pathways is not yet well clarified [48]. CBX7 is a component of the chromatin modifier PRC1-complex and is required for the propagation of the transcriptionally repressive state of multiple genes through cell-division during embryonic development [49], including Hox genes [50]. Expectedly, we noticed that HOXC6 and HOXC9 were both monotonically upregulated with increasing malignancy (**Supplementary Table 1**). ZFP2 is a probable transcription factor and evidence suggest an epigenetic role as well [51], while recently, high load of mutations in another member of the ZFP family (ZFP36) were associated with upper tract urothelial carcinoma [52].

Our analysis highlighted expected landmark pathways, such as mTORC1 pathway [53] and MYC targets [54] which were upregulated, but also novel downregulated pathways such us the circadian clock and the metabolism of heme. These results associate for the first time BLCA progression to the disruption of the circadian homeostasis and to iron metabolism deficiencies, events that are thought to be tumorigenic [55, 56], but their exact mechanism of action is not well understood. We detected an “incomplete” Warburg effect in which the monotonical downregulation (or deactivation) of lipid and fatty acids metabolism, was not accompanied by any upregulation of glycolysis. Additionally, to the GATA3 regulon which is a known driver of luminal biology, we detected a novel progressive downregulation of the GLI2 regulon. GLI proteins are transcription factors of the Sonic hedgehog (Shh) pathway and although GLI2 expression levels positively correlate with more invasive BLCA cell lines, Shh genes do not behave accordingly [57]. Our results validate these observations, as the entire regulon of the GLI2 TF was inactivated with increasing BLCA malignancy, suggesting no potential therapeutic effect in its inhibition.

A change in the expression levels of a single gene may not always have a functional consequence, mainly due to the complementary action of other genes of that pathway. Instead, a concerted change in the expression levels of multiple genes is often required for the cell to successfully transit between phenotypes. We investigated this type of changes using a network coexpression analysis followed by intra-stage graph clustering, and were able to detect both intrinsic and microenvironmental gene rewiring programs that shape tumor evolution. These gene programs were found to be organized into distinct biological communities (which were validated in the TCGA 2017 study), and consisted of genes stably coexpressed among all BLCA stages, but also of genes specifically coexpressed (gained or lost) in a particular stage only. We assume that the ECM remodelling relies mostly on this last property of gains and losses in coexpression, since the bladder consists of heterogeneous anatomical tissues with stiff biological barriers (basement membrane, muscle layer), which in turn require specialized cell functions to be dissolved. The immune activation community was present in both T0 and MIBC. Results of the CIBERSORT analysis showed that inside the bladder tumor and with increasing stage, most monocytes preferentially differentiate into macrophages with M2 polarization. This is in line with the recent findings from Chen et al. [41] in which authors analyzed scRNA-seq data of BLCA

patients and observed a similar pattern of differentiation for monocytes. As the CIBERSORT algorithm has not been designed to recognize signatures of T-cell exhaustion, CD8+ cells appeared at higher number in T0 samples compared to tumor, however we cannot exclude the existence of biological barriers in the normal adjacent urothelium that forbid CD8+ cells from reaching tumor. T4 samples had the highest percent of M2-macrophages, which could partially explain their immunosuppressive state, and as a novel finding, T4 coexpression in the immune cells was driven by AIF1, a gene that ensures the survival and proliferation of macrophages. We hypothesize that immune evasion in BLCA is likely promoted by M2-macrophages that actively express AIF1, but further work is required to validate these observations.

Our study has its limitations, which include the retrospective nature of the analysis, and the usage of batch correction methods to integrate different microarrays. The batch correction method unavoidably eliminates some of the biological variability leading to increased false negative rate; our validation of the applied method via checking expression of known genes, in combination to focusing on positive signals, suggests that sufficient biological variability is maintained. In addition, restrictions in the validation set imposed by lack of samples from all disease stages did not allow validating the observations particularly at the NMIBC. Finally, clinical stage is known to have varying rates of misdiagnosis. However, the high number of samples used in each of the stage categories is expected to balance out to some extent misdiagnosed cases while increasing power of the received results.

Data availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Code availability

All custom R scripts created for analysis and visualizations are available by the first and the corresponding author upon reasonable request.

Acknowledgements

MM is supported by the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 898260 (H2020-MSCA-IF-2019, ReDrugBC, Grant agreement ID: 898260).

Contributions

HM and AV conceived and designed the investigation, RS, MF, MZ and EM processed and analyzed data, RS and EM wrote software for analysis and visualizations, RS, MF, MZ HM and AV interpreted results, RS wrote the manuscript, MM, MR, HM and AV provided critical revisions.

Ethics declarations/ Competing interests

HM is the co-founder and co-owner of Mosaiques Diagnostics, MF, EM, and MM are employed by Mosaiques Diagnostics.

References

[1] Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians. 2018;68:394-424.

[2] Knowles MA, Hurst CD. Molecular biology of bladder cancer: new insights into pathogenesis and clinical diversity. Nature reviews Cancer. 2015;15:25-41.

[3] Czerniak B, Dinney C, McConkey D. Origins of Bladder Cancer. Annual review of pathology. 2016;11:149-74.

[4] Dyrskjot L, Kruhoffer M, Thykjaer T, Marcussen N, Jensen JL, Moller K, et al. Gene expression in the urinary bladder: a common carcinoma in situ gene expression signature exists disregarding histopathological classification. Cancer research. 2004;64:4040-8.

[5] Sjodahl G, Lauss M, Lovgren K, Chebil G, Gudjonsson S, Veerla S, et al. A molecular taxonomy for urothelial carcinoma. Clinical cancer research : an official journal of the American Association for Cancer Research. 2012;18:3377-86.

[6] Damrauer JS, Hoadley KA, Chism DD, Fan C, Tiganelli CJ, Wobker SE, et al. Intrinsic subtypes of high-grade bladder cancer reflect the hallmarks of breast cancer biology. Proceedings of the National Academy of Sciences of the United States of America. 2014;111:3110-5.

[7] Choi W, Porten S, Kim S, Willis D, Plimack ER, Hoffman-Censits J, et al. Identification of distinct basal and luminal subtypes of muscle-invasive bladder cancer with different sensitivities to frontline chemotherapy. Cancer cell. 2014;25:152-65.

[8] Rebouissou S, Bernard-Pierrot I, de Reynies A, Lepage ML, Krucker C, Chapeaublanc E, et al. EGFR as a potential therapeutic target for a subset of muscle-invasive bladder cancers presenting a basal-like phenotype. Science translational medicine. 2014;6:244ra91.

[9] Robertson AG, Kim J, Al-Ahmadie H, Bellmunt J, Guo GW, Cherniack AD, et al. Comprehensive Molecular Characterization of Muscle-Invasive Bladder Cancer. Cell. 2017;171:540-+.

[10] Hedegaard J, Lamy P, Nordentoft I, Algaba F, Hoyer S, Ulhoi BP, et al. Comprehensive Transcriptional Analysis of Early-Stage Urothelial Carcinoma. Cancer cell. 2016;30:27-42.

[11] Hurst CD, Alder O, Platt FM, Droop A, Stead LF, Burns JE, et al. Genomic Subtypes of Non-invasive Bladder Cancer with Distinct Metabolic Profile and Female Gender Bias in KDM6A Mutation Frequency. Cancer cell. 2017;32:701-15 e7.

[12] Sjodahl G, Eriksson P, Liedberg F, Hoglund M. Molecular classification of urothelial carcinoma: global mRNA classification versus tumour-cell phenotype classification. The Journal of pathology. 2017;242:113-25.

[13] Mo Q, Nikolos F, Chen F, Tramel Z, Lee YC, Hayashi K, et al. Prognostic Power of a Tumor Differentiation Gene Signature for Bladder Urothelial Carcinomas. Journal of the National Cancer Institute. 2018;110:448-59.

[14] Kamoun A, de Reynies A, Allory Y, Sjodahl G, Robertson AG, Seiler R, et al. A Consensus Molecular Classification of Muscle-invasive Bladder Cancer. European urology. 2020;77:420-33.

[15] Lindskrog SV, Prip F, Lamy P, Taber A, Groeneveld CS, Birkenkamp-Demtroder K, et al. An integrated multi-omics analysis identifies prognostic molecular subtypes of non-muscle-invasive bladder cancer. Nature communications. 2021;12:2301.

[16] Rebouissou S, Herault A, Letouze E, Neuzillet Y, Laplanche A, Ofualuka K, et al. CDKN2A homozygous deletion is associated with muscle invasion in FGFR3-mutated urothelial bladder carcinoma. Journal of Pathology. 2012;227:315-24.
[17] Nassar AH, Umeton R, Kim J, Lundgren K, Harshman L, Van Allen EM, et al. Mutational Analysis of 472 Urothelial Carcinoma Across Grades and Anatomic Sites. Clinical Cancer Research. 2019;25:2458-70.
[18] Carvalho BS, Irizarry RA. A framework for oligonucleotide microarray preprocessing. Bioinformatics. 2010;26:2363-7.
[19] Durinck S, Spellman PT, Birney E, Huber W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. Nature protocols. 2009;4:1184-91.
[20] Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics. 2007;8:118-27.
[21] Smyth GK, Speed T. Normalization of cDNA microarray data. Methods. 2003;31:265-73.
[22] Jacob L, Gagnon-Bartsch JA, Speed TP. Correcting gene expression data when neither the unwanted variation nor the factor of interest are observed. Biostatistics. 2016;17:16-28.
[23] Manimaran S, Selby HM, Okrah K, Ruberman C, Leek JT, Quackenbush J, et al. BatchQC: interactive software for evaluating sample and batch effects in genomic data. Bioinformatics. 2016;32:3836-8.
[24] Hanzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-Seq data. Bmc Bioinformatics. 2013;14.
[25] Garcia-Alonso L, Holland CH, Ibrahim MM, Turei D, Saez-Rodriguez J. Benchmark and integration of resources for the estimation of human transcription factor activities. Genome Res. 2019;29:1363-75.
[26] Yoshihara K, Shahmoradgoli M, Martinez E, Vegesna R, Kim H, Torres-Garcia W, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. Nature communications. 2013;4.
[27] Huynh-Thu VA, Irrthum A, Wehenkel L, Geurts P. Inferring Regulatory Networks from Expression Data Using Tree-Based Methods. PloS one. 2010;5.
[28] Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. J Stat Mech-Theory E. 2008.
[29] Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. Omics : a journal of integrative biology. 2012;16:284-7.
[30] Lee SR, Roh YG, Kim SK, Lee JS, Seol SY, Lee HH, et al. Activation of EZH2 and SUZ12 Regulated by E2F1 Predicts the Disease Progression and Aggressive Characteristics of Bladder Cancer. Clinical cancer research : an official journal of the American Association for Cancer Research. 2015;21:5391-403.
[31] Kanehira M, Harada Y, Takata R, Shuin T, Miki T, Fujioka T, et al. Involvement of upregulation of DEPDC1 (DEP domain containing 1) in bladder carcinogenesis. Oncogene. 2007;26:6448-55.
[32] Hsu CL, Liu JS, Wu PL, Guan HH, Chen YL, Lin AC, et al. Identification of a new androgen receptor (AR) co-regulator BUD31 and related peptides to suppress wild-type and mutated AR-mediated prostate cancer growth via peptide screening and X-ray structure analysis. Molecular oncology. 2014;8:1575-87.
[33] Kawahara T, Shareef HK, Aljarah AK, Ide H, Li Y, Kashiwagi E, et al. ELK1 is up-regulated by androgen in bladder cancer cells and promotes tumor progression. Oncotarget. 2015;6:29860-76.
[34] Wei WS, Chen X, Guo LY, Li XD, Deng MH, Yuan GJ, et al. TRIM65 supports bladder urothelial carcinoma cell aggressiveness by promoting ANXA2 ubiquitination and degradation. Cancer letters. 2018;435:10-22.

603 [35] Pham H, Ziboh VA. 5 alpha-reductase-catalyzed conversion of testosterone to
604 dihydrotestosterone is increased in prostatic adenocarcinoma cells: suppression by 15-
605 lipoxygenase metabolites of gamma-linolenic and eicosapentaenoic acids. The Journal of
606 steroid biochemistry and molecular biology. 2002;82:393-400.
607 [36] Leoni LM, Hartley JA. Mechanism of action: the unique pattern of bendamustine-
608 induced cytotoxicity. Seminars in hematology. 2011;48 Suppl 1:S12-23.
609 [37] Wullweber A, Strick R, Lange F, Sikic D, Taubert H, Wach S, et al. Bladder Tumor Subtype
610 Commitment Occurs in Carcinoma In Situ Driven by Key Signaling Pathways Including ECM
611 Remodeling. Cancer research. 2021;81:1552-66.
612 [38] Ding J, Du K. ClipR-59 interacts with Akt and regulates Akt cellular
613 compartmentalization. Molecular and cellular biology. 2009;29:1459-71.
614 [39] Egana-Gorrono L, Chinnasamy P, Casimiro I, Almonte VM, Parikh D, Oliveira-Paula GH, et
615 al. Allograft inflammatory factor-1 supports macrophage survival and efferocytosis and limits
616 necrosis in atherosclerotic plaques. Atherosclerosis. 2019;289:184-94.
617 [40] Tan TZ, Rouanne M, Tan KT, Huang RY, Thiery JP. Molecular Subtypes of Urothelial
618 Bladder Cancer: Results from a Meta-cohort Analysis of 2411 Tumors. European urology.
619 2019;75:423-32.
620 [41] Chen Z, Zhou L, Liu L, Hou Y, Xiong M, Yang Y, et al. Single-cell RNA sequencing highlights
621 the role of inflammatory cancer-associated fibroblasts in bladder urothelial carcinoma.
622 Nature communications. 2020;11:5077.
623 [42] Chen X, Chen H, He D, Cheng Y, Zhu Y, Xiao M, et al. Analysis of Tumor
624 Microenvironment Characteristics in Bladder Cancer: Implications for Immune Checkpoint
625 Inhibitor Therapy. Frontiers in immunology. 2021;12:672158.
626 [43] Ji M, Wang Z, Chen J, Gu L, Chen M, Ding Y, et al. Up-regulated ENO1 promotes the
627 bladder cancer cell growth and proliferation via regulating beta-catenin. Bioscience reports.
628 2019;39.
629 [44] Lee SJ, Lee EJ, Kim SK, Jeong P, Cho YH, Yun SJ, et al. Identification of pro-inflammatory
630 cytokines associated with muscle invasive bladder cancer; the roles of IL-5, IL-20, and IL-28A.
631 PloS one. 2012;7:e40267.
632 [45] Zhang H, Jiang H, Wang W, Gong J, Zhang L, Chen Z, et al. Expression of Med19 in
633 bladder cancer tissues and its role on bladder cancer cell growth. Urologic oncology.
634 2012;30:920-7.
635 [46] da Silva GN, Evangelista AF, Magalhaes DA, Macedo C, Bufalo MC, Sakamoto-Hojo ET, et
636 al. Expression of genes related to apoptosis, cell cycle and signaling pathways are
637 independent of TP53 status in urinary bladder cancer cells. Molecular biology reports.
638 2011;38:4159-70.
639 [47] Li L, Ljungman M, Dixon JE. The human Cdc14 phosphatases interact with and
640 dephosphorylate the tumor suppressor protein p53. The Journal of biological chemistry.
641 2000;275:2410-4.
642 [48] Dietachmayr M, Rathakrishnan A, Karpiuk O, von Zweydorf F, Engleitner T, Fernandez-
643 Saiz V, et al. Antagonistic activities of CDC14B and CDK1 on USP9X regulate WT1-dependent
644 mitotic transcription and survival. Nature communications. 2020;11:1268.
645 [49] Moussa HF, Bsteh D, Yelagandula R, Pribitzer C, Stecher K, Bartalska K, et al. Canonical
646 PRC1 controls sequence-independent propagation of Polycomb-mediated gene silencing.
647 Nature communications. 2019;10:1931.
648 [50] Grossniklaus U, Paro R. Transcriptional silencing by polycomb-group proteins. Cold
649 Spring Harbor perspectives in biology. 2014;6:a019331.
650 [51] Sobocinska J, Molenda S, Machnik M, Oleksiewicz U. KRAB-ZFP Transcriptional
651 Regulators Acting as Oncogenes and Tumor Suppressors: An Overview. International journal
652 of molecular sciences. 2021;22.

653 [52] Su X, Lu X, Bazai SK, Comperat E, Mouawad R, Yao H, et al. Comprehensive integrative
654 profiling of upper tract urothelial carcinomas. Genome biology. 2021;22:7.
655 [53] Ching CB, Hansel DE. Expanding therapeutic targets in bladder cancer: the
656 PI3K/Akt/mTOR pathway. Lab Invest. 2010;90:1406-14.
657 [54] Mahe M, Dufour F, Neyret-Kahn H, Moreno-Vega A, Beraud C, Shi M, et al. An
658 FGFR3/MYC positive feedback loop provides new opportunities for targeted therapies in
659 bladder cancers. EMBO molecular medicine. 2018;10.
660 [55] Fu L, Kettner NM. The circadian clock in cancer development and therapy. Progress in
661 molecular biology and translational science. 2013;119:221-82.
662 [56] Chen Y, Fan Z, Yang Y, Gu C. Iron metabolism and its contribution to cancer (Review).
663 International journal of oncology. 2019;54:1143-54.
664 [57] Mechlin CW, Tanner MJ, Chen M, Buttyan R, Levin RM, Mian BM. Gli2 expression and
665 human bladder transitional carcinoma cell invasiveness. The Journal of urology.
666 2010;184:344-51.

667

668