



March 4, 2021

Differential Expression Analysis for the Identification of Survival Associated Genes in Primary Bladder Cancer using Microarray Data

AMOS OCHIENG OKUTSE

Joma Kenyatta University of Agriculture and Technology, Kenya, amos.ochieng@students.jkuat.ac.ke

KELVIN W. NYONGESA

Kenyatta National Hospital, Kenya, kelvinyongesa@gmail.com

Abstract

Bladder cancer (BC) is a highly malignant and malevolent type of tumor whose incidence and mortality worldwide remains high. This paper aimed to identify the differentially expressed genes (DEGs) between normal and primary bladder cancer samples, which might be of prognostic value, and evaluate their association to clinical outcomes (survival) in patients with primary bladder cancer. A sample of 256 gene expression profiles, of which 10 were normal bladder mucosae, whereas 165 were primary bladder cancer tissues, were downloaded from the gene expression omnibus (GEO). The limma package in R was used to perform screening for DEGs after variance reduction through quantile normalization and log₂ transformation. Gene Ontology (GO), KEGG enrichment analysis, and Network analysis were performed to determine the functions of these DEGs. Cox proportional hazard regression model was then used to determine the association between the top selected DEGs and overall primary bladder cancer survival. Selected genes were then encoded as either low, middle, or high expression, and Kaplan-Meier plots were then used to visualize the effect of these expression levels for the selected five genes on bladder cancer survival probability. This study identified a total of 5 979 DEGs where 2 878 were upregulated, 3 101 were downregulated, and 37 169 were not significantly expressed differently. Cox proportional hazard stepwise regression models revealed several DEGs that have a statistically significant association with bladder cancer. The selected five DEGs associated with survival were *EDNRB* (HR=1.9858, 95%CI 1.4965–2.6350), *COL14A1* (HR=1.5335, 95%CI 1.1670–2.0151), *MOXD1* (HR=1.4878, 95%CI 1.1070–1.9996), *FAM107A* (HR=1.4970, 95%CI 1.1162–2.0077), and *REM1* (HR=1.5999, 95%CI 1.2422–2.0605). After encoding these genes as either low, middle, or high expression, high expression of both *EDNRB*, *COL14A1*, *FAM107A*, and *REM1* were found to result in statistically significant reductions in survival probabilities of patients ($p < 0.05$). Contrastingly, high expression of *MOXD1* did not contribute to a statistically significant reduction in patients' survival probability ($p > 0.05$). The genes found in this study could provide useful insight into understanding the association between the identified DEGs and primary bladder cancer survival, which can further inform future studies and clinical therapies

Keywords

Gene expression, Primary bladder cancer survival, Prognosis, Biological markers, Transcriptomics, Bioinformatics, Survival analysis, Microarray data analysis

Peer Review

This work has undergone a double-blind review by a minimum of two faculty members from institutions of higher learning from around the world. The faculty reviewers have expertise in disciplines closely related to those represented by this work. If possible, the work was also reviewed by undergraduates in collaboration with the faculty reviewers.

Introduction

Cancer is a “genetic disorder driven by the progressive accumulation of multiple genetic and epigenetic changes” (W. J. Kim et al., 2010). These hereditary and epigenetic vicissitudes often at the molecular level result in the proliferation of cells, the spread of these cells to other parts of the organism, and depreciated cell death. The changes that occur in an individual’s gene expression profile play a vital role in determining tumor biology, including such aspects as the growth of the tumor, progression, recurrence, and metastasis, all of which might have statistically significant effects on the overall patient survival probability. Additionally, changes in an individual's gene expression profile are suggestive of probable underlying gene mutations (Daneshmand, 2019; W. J. Kim et al., 2010).

Globally, based on recently published estimates, bladder cancer is among the top ten cancers that are most common in the world, with new cases estimated at 550 000 per annum (Richters et al., 2019). While an extravagant portion of the burden of this cancer falls, for the most part, in the developed countries, shifts are expected in the course of the next few decades due to the changing world demographics, including growing and aging populations in the African continent (Richters et al., 2019). Previously conducted studies have found high bladder cancer incidences in North America, Western Europe, and Southern Europe, whereas high mortalities have been reported in Northern Africa (Wong et al., 2018). Malats and Real (2015) reported high incidences in older men results which conform with the published bladder cancer estimates from the American Cancer Society (2020), where the average age at diagnosis has been reported as seventy-three years, with men reporting significantly higher incidences (62, 100 new male cases against 19, 300 cases in women) and mortality

than women (13, 050 male deaths against 4, 930 deaths).

Throughout the years, technological advancements have resulted in the development of novel and high throughput technologies such as gene microarray sequencing technologies that allow the analysis of several patient gene expression profiles, particularly in carcinogenesis and progression of cancers (Huang et al., 2018). Additionally, the identification of novel biomarkers of disease (cancer) progression (Botling et al., 2013; W. T. Kim et al., 2010; Sun et al., 2010) can be carried out through the use of gene expression profiles as well as clinical cases (Shukla et al., 2017; Tang et al., 2017). Over the last decade, the systematic profiling of samples from humans with cancers has been established and authenticated as an excellent method of obtaining the molecular characteristics of various tumors and their subtypes (Martinez-Romero et al., 2018).

In previously published studies on bladder cancer, evidence of the effect of genetic mutations has been reported ranging from significant associations between gene mutations and bladder cancer (Mikhailenko & Nemtsova, 2016; Zhang & Zhang, 2015) to non-substantial evidence (Allory et al., 2014). Additionally, some papers have reported both significant and null associations between some genetic mutations and bladder cancer (K. Wang et al., 2015). While studies have tried to explore the genetic variations resulting in atypical gene expression profiles in individuals with bladder cancer, most studies have focused on differentially expressed genes (DEGs) and the identification of genetic biomarkers with little or no focus on the associations between these genes and overall cancer-specific survival in samples obtained from patients with primary bladder cancer (Aljabery et al., 2018; Gao et al., 2018; Shi & Tian, 2019). In

studies in which survival was considered (Buttiglierio et al., 2017; Chen et al., 2019; W. J. Kim et al., 2010; Seiler et al., 2017), specific focus on the influence of these specific top differentially expressed genes on primary bladder cancer survival is either obscured or non-existent or focused on the analysis of a single gene (Theodorescu et al., 2004).

In the current paper, an analysis of differentially expressed genes in primary bladder tumor samples from microarray experiments was carried out to identify the top genes expressed differently between normal bladder samples and primary bladder samples. A statistical analysis of these genes was then carried out to identify whether these genes statistically significantly influenced overall survival in primary bladder cancer patients from whom the samples were obtained. Identifying the specific impact of the DEGs on bladder cancer survival may play a crucial role in prognosis and the discovery of novel drug targets for the disease. Additionally, this analysis may enable the discovery of the effect of low, middle, and high gene expression of the selected top prognostic genes on primary bladder cancer patient survival.

Methodology

Data source. This study sought to identify the associations between the top differentially expressed genes between healthy and primary bladder cancer patients and the association between these genes, bladder cancer prognosis, and overall survival using microarray gene expression profiles. A total of 256 samples prepared using the Illumina Human-6 expression BeadChip Version 2.0 from which 165 were primary bladder cancer samples, ten normal bladder mucosae, 23 recurrent non-muscle invasive tumor samples, and 58 normal-looking bladder mucosae were downloaded from the Gene Expression Omnibus (GEO),

with accession number [GSE13507](#) (W. J. Kim et al., 2010).

Data Preprocessing. Samples from the downloaded dataset with no clinical information were filtered out of the downloaded dataset to remain with 165 samples derived from patients with primary bladder cancer and ten healthy bladder samples with survival information. Variance reduction between the microarrays was carried out using quantile normalization (Bolstad et al., 2003; Calza et al., 2008). Gene expression values were then log₂-transformed to enhance normality and reduce heterogeneity (Quackenbush, 2002).

Differential gene expression analysis. Analysis of differentially expressed genes from the selected samples was carried out using the limma package in R (Ritchie et al., 2015). Linear models for each of the genes in the given series of arrays was fit using the 'lmFit' function. The fitted microarray linear model was then used to calculate and derive modulated statistics, including t-statistics, log-odds of differential expression, and F-statistics using the empirical Bayes statistics for differential expression (eBayes) with 0.01 as an assumed proportion of DEGs.

Upregulated and downregulated DEGs were selected based on the rule: $|\log(\text{foldchange})| > 0$ and $|\log(\text{foldchange})| < 0$ respectively. Similarly, the threshold, $\log(\text{foldchange}) = 0$, was used to determine DEGs that were not statistically significantly differentially expressed between normal and primary bladder cancer samples. The limma function 'topTable' was used to retrieve the top-ranked differentially expressed genes from the empirical Bayesian model. Benjamin and Hochberg (False Discovery Rate, FDR) adjustments were applied to the p-values.

Functional annotation and enrichment analysis. ClusterProfiler, an R package for

the analysis and visualization of the functional annotations for genes and clusters, was used to perform enrichment analyses (Yu et al., 2012). The 'enrichKEGG' function from the package was used to perform enrichment analysis based on the Kyoto Encyclopedia for Genes and Genome (KEGG). In contrast, Gene Ontology (GO) categories were returned using the 'enrichGO' function with False Discovery Rate (FDR) control. Probability values were Benjamini-Hochberg (BH) adjusted with the cut-off for the most significant pathway analysis results for KEGG and GO categories obtained using the thresholds $p < 0.05$ and $p < 0.01$, respectively. Visualizations of the enrichment results for both GO and KEGG were done using dot plots. Additionally, networks showing gene linkages and biological functions were constructed using the 'cnetplot' function from the package 'enrichplot' in R (Yu, 2019). These visualizations are significant in visualizing the enriched pathways and genes belonging to several annotation classes.

Cox proportional hazard model for the identification of survival associated genes in primary bladder cancer. The log-transformed probe intensity data were Z-score transformed to allow comparability of the results from multiple experiments regardless of the initial hybridization concentrations after standardization (Cheadle et al., 2003). The survival package (Borgan, 2001) and the RegParallel package (Blighe, 2019) from Bioconductor were used to construct a multivariable Cox proportional hazard model and allow analogous processing over the considerably large microarray gene expression dataset, respectively.

The DEGs which satisfied the threshold $\text{LogRank} < 0.01$ (top hits) were retrieved and annotated using the biomaRt package (Durinck et al., 2009). Attributes

including platform information, Entrez gene IDs, gene biotype, and the external gene names for the survival associated genes were retrieved for further analysis. The DEGs that were associated with bladder cancer survival were prepared for downstream analysis and encoded with regards to their expression levels as either high expression (Z-score cut-off ≥ 1.0), low expression (Z-score cut-off ≤ -1.0), or median expression (Z-score = 0). The survminer package (Alboukadel et al., 2018) was used to construct Kaplan-Meier (KM) plots for visualizing the effect of the selected top differentially expressed genes on the overall primary bladder cancer survival.

Statistical analysis and data availability.

All statistical analysis procedures were carried out using the R statistical programming environment version 3.6.3 (<https://www.r-project.org/>). The packages used in this analysis are available and freely downloadable from Bioconductor (<https://www.bioconductor.org/>) and The Comprehensive R Archive Network (CRAN) (<https://cran.r-project.org/>). The dataset for this analysis and the related clinical information for the 165 primary bladder cancer samples were downloaded from the Gene Expression Omnibus (GEO) with accession number GSE13507 available at <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi>.

Results

Differential Gene Expression Analysis. 2,878 genes were significantly upregulated ($\log(\text{foldchange}) > 0$), whereas 3,101 genes were significantly downregulated ($\log(\text{foldchange}) < 0$) between normal and primary bladder cancer samples (**Table 1**). These results were visualized using a volcano plot (**Figure 1**).

Differential Expression Analysis for the Identification of Survival Associated Genes in Primary Bladder Cancer using Microarray Data

Table 1: Top 10 most significantly up or downregulated differentially expressed genes (DEGs) between normal and primary bladder cancer samples.

Gene ID	Gene Symbol	Official Gene Name*	LogFC	Adj. P-value
Upregulated genes				
991	CDC20	Cell division cycle 20	0.4716	6.28E-09
5373	PMM2	Phosphomannomutase	0.1711	9.80E-09
128239	IQGAP3	IQ motif containing GTPase activating protein 3	0.4125	1.26E-08
2932	GSK3B	glycogen synthase kinase 3 beta	0.1896	8.40E-08
51537	MTFP1	Mitochondrial fission process 1	0.2168	1.64E-07
51203	NUSAP1	nucleolar and spindle associated protein 1	0.3719	3.11E-07
7153	TOP2A	DNA topoisomerase II alpha	0.4546	3.63E-07
2810	SFN	Stratifin	0.3100	8.07E-07
51599	LSR	Lipolysis stimulated lipoprotein receptor	0.2140	1.15E-06
9212	AURKB	Aurora kinase B	0.3141	1.31E-06
Downregulated genes				
9890	PLPPR4	phospholipid phosphatase related 4	-0.3804	2.14E-22
54360	CYTL1	cytokine like 1	-0.2691	6.44E-19
11126	CD160	CD160 molecule	-0.1475	8.68E-18
121601	ANO4	anoctamin 4	-0.1705	9.34E-17
5126	PCSK2	proprotein convertase subtilisin/kexin type 2	-0.1268	7.05E-16
9745	ZNF536	zinc finger protein 536	-0.2280	7.31E-16
55022	PID1	phosphotyrosine interaction domain containing 1	-0.3029	1.79E-15
10699	CORIN	corin, serine peptidase	-0.1868	3.19E-15
79625	NDNF	neuron derived neurotrophic factor	-0.3949	1.31E-14
4916	NTRK3	neurotrophic receptor tyrosine kinase 3	-0.3849	6.79E-14

*the official gene names are based on the Hugo Gene Nomenclature Committee (HGNC) retrieved from <https://www.genenames.org/>

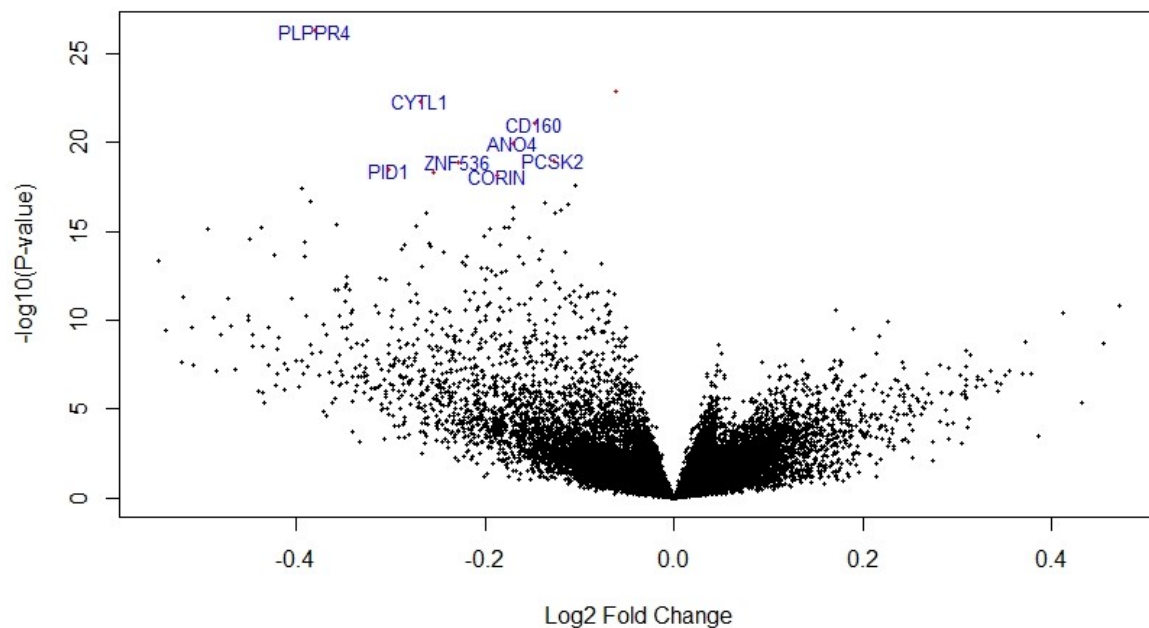


Figure 1: A volcano plot showing the differentially expressed genes (DEGs). Gene symbols for the top genes that are expressed differently between normal and bladder cancer tissue samples. Genes in the upper quadrants are statistically significantly differentially expressed.

Pathway enrichment analysis and network construction results. The top 300 DEGs' biological functions were identified through enrichment analysis, and functional annotation was based on the GO terms. Gene Ontology enabled the exploration of the role distribution of these genes in primary bladder cancer. KEGG pathway enrichment analysis for the top 300 DEGs allowed identifying the most significant pathways involved in primary bladder cancer. GO categories are separated into three groups, namely, molecular function (MF), biological processes (BP), and cellular component (CC). The top 5 most significant GO terms under each category are summarized in **Table 2**. For molecular function, the most significantly enriched GO terms were receptor activity (GO:0004872, $p=6.8123E-$

05) and carbohydrate-binding (GO:0030246, $p=1.0526E-04$), extracellular matrix organization for biological process (GO:0030198, $p=4.0049E-05$) and extracellular region for cellular component (GO:0005576, $p=9.1054E-11$).

After KEGG enrichment analysis, the most significant pathways were retrieved using a p -value cut-off of 0.05. The topmost significantly enriched pathways (Figure 3) were cell adhesion molecules ($p=3.3007E-04$), PI3K-Akt signaling pathway ($p=0.0021$), Basal cell carcinoma ($p=0.0034$), Pathways in cancer ($p=0.0069$), and Ether lipid metabolism ($p=0.0099$). The KEGG enrichment results are summarized in **Table 3**.

Differential Expression Analysis for the Identification of Survival Associated Genes in Primary Bladder Cancer using Microarray Data

Table 2: The top 5 most significantly enriched Gene Ontology (GO) terms for each GO category

Term	Gene Count	P-value	Fold Enrichment	Benjamini	FDR*
Biological Process (BP)					
GO:0030198~extracellular matrix organization	14	4.0049E-5	4.0936	0.0684	0.0678
GO:0014068~positive regulation of phosphatidylinositol 3-kinase signalling	8	1.3149E-4	7.0536	0.1098	0.2226
GO:0007155~cell adhesion	21	1.6393E-4	2.6221	0.0921	0.2774
GO:0007399~nervous system development	15	5.2444E-4	2.9953	0.2070	0.8850
GO:0060070~canonical Wnt signalling pathway	8	6.0123E-4	5.5239	0.1917	1.0140
Cellular Component (CC)					
GO:0005576~extracellular region	66	9.1054E-11	2.3493	2.4402E-8	1.1958E-7
GO:0005615~extracellular space	51	2.7182E-7	2.1698	3.6423E-5	3.5698E-4
GO:0005886~plasma membrane	111	5.1777E-7	1.5436	4.6253E-5	6.7999E-4
GO:0005578~proteinaceous extracellular matrix	18	4.9476E-6	3.8491	3.3144E-4	0.0065
GO:0009897~external side of plasma membrane	16	5.1036E-6	4.3048	2.7352E-4	0.0067
Molecular Function (MF)					
GO:0004872~receptor activity	14	6.8123E-5	3.8896	0.0309	0.0969
GO:0030246~carbohydrate binding	13	1.0526E-4	3.9988	0.0239	0.1497
GO:0008201~heparin binding	11	3.3416E-4	4.1449	0.0499	0.4745
GO:0042803~protein homodimerization activity	24	0.0024	1.9821	0.2409	3.3541
GO:0005518~collagen binding	6	0.0031	6.0289	0.2482	4.3166
*FDR=False Discovery Rate					

Table 3: Enriched KEGG pathways for the DEGs in primary bladder cancer

Term	Gene Count	P-value	Benjamini	FDR*
hsa04514: Cell adhesion molecules (CAMs)	11	3.3007E-4	0.0571	0.4050
hsa04151: PI3K-Akt signalling pathway	16	0.0021	0.1678	2.5069
hsa05217: Basal cell carcinoma	6	0.0034	0.1812	4.0566
hsa05200: Pathways in cancer	16	0.0069	0.2682	8.2648
hsa00565: Ether lipid metabolism	5	0.0099	0.3001	11.5932
hsa05144: Malaria	5	0.0134	0.3295	15.2689
hsa05215: Prostate cancer	6	0.0251	0.4761	26.8442
hsa04014: Ras signaling pathway	10	0.0269	0.4551	28.5048
hsa04360: Axon guidance	7	0.0328	0.4826	33.6102
hsa04510: Focal adhesion	9	0.0408	0.5242	40.1328
hsa04550: Signaling pathways regulating pluripotency of stem cells	7	0.0489	0.5559	46.0266
hsa04978: Mineral absorption	4	0.0498	0.5313	46.6434

*FDR=False Discovery Rate

Network Construction. Network construction was based on the top 300 DEGs and was constructed using the ClusterProfiler cnetplot function. The network was constructed based on the GO enrichment analysis results and was used to

visualize the links between the DEGs and their biological functions. Through the network results, genes, including *LAMA2*, *BCL2*, and *FLRT2*, were observed to belong to various annotation groups (**Figure 2**).

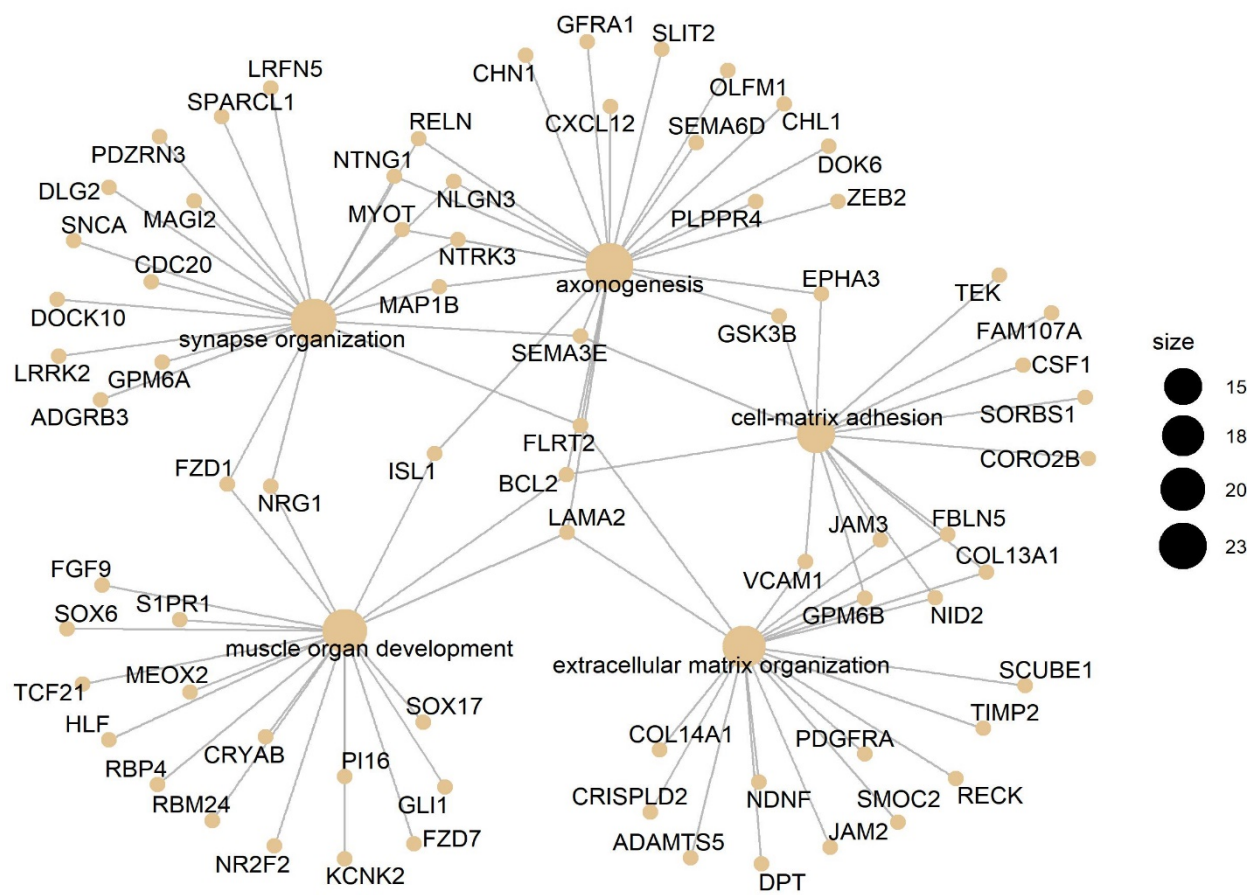


Figure 2: A visual representation of the linkages of the differentially expressed genes (DEGs) and their biological functions as a network. Each node represents an enriched pathway with its size corresponding to the number of genes in that enriched pathway. The edges are linked to each of the genes and overlap where the genes belong to multiple annotation categories. Genes that belong to the same annotation category tend to be clustered around that node.

Proportional Cox hazard survival model.

Cox proportional hazard survival models were constructed to identify the roles of the Differentially Expressed Genes (DEGs) in overall primary cancer survival. Multiple survival analysis using RegParallel revealed a total of 8,746 statistically significantly DEGs associated with survival in primary bladder cancer “Supplementary file 1: Cox proportional hazard model results.” 2,623 DEGs satisfied the set threshold $\text{LogRank} < 0.01$ “Supplementary file 2: Reduced proportional hazard model with $\text{LogRank} < 0.01$ ” and were annotated using the biomaRt package. The list of the retrieved DEGs that were associated with survival in primary bladder cancer after annotation were

summarized in “Supplementary file 3: biomaRt annotated survival associated genes.” Of these genes, 50 genes corresponded to those found in the top 300 differentially expressed genes identified through differential expression analysis using the limma package. These genes were extracted and summarized as presented in **Table 4** and “Supplementary file 4: Top 50 survival associated genes in primary bladder cancer.”

The top 5 DEGs out of the 50 that this study identified to be significantly differentially expressed between normal and cancerous cell samples and as markers of overall survival in primary bladder cancer

were then encoded to visualize their effects using Kaplan Meier plots. The top five of these genes were *EDNRB*, *COL14A*, *MOXD1*, *FAM107A*, and *REM1*.

High expression of *EDNRB* was statistically significantly associated with reduced survival in primary bladder cancer ($p=0.0003$) (**Figure 3A**). Additionally, the high expression of *COL14A1* (**Figure 3B**) and *FAM107A* were statistically significantly associated with a reduced survival probability in patients with primary bladder cancer ($p=0.0052$ and $p=0.0066$, respectively). **Figure 3C** presents the effects of low, middle, and high expression of *FAM107A*, respectively.

High expression of *REM1* was also associated with a statistically significant reduction in the survival probability of patients with primary bladder cancer ($p=0.0091$; **Figure 3D**). Contrastingly, the high expression of the DEG *MOXD1* was not associated with significant reductions in the survival probability of patients with primary bladder cancer ($p=0.2200$; **Figure 4**).

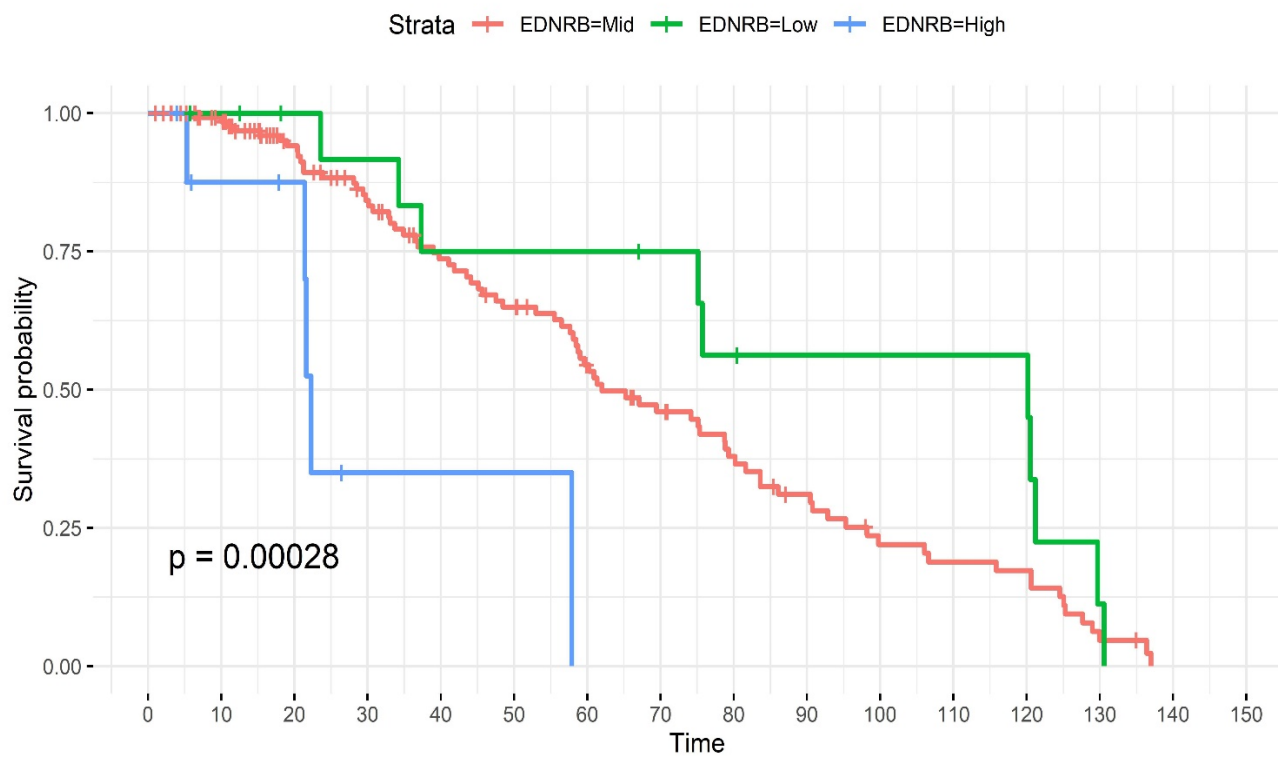
Discussion

Technological advancements over the years have made it possible for efforts geared towards the particular understanding of the genetic links between several cancers. While studies have identified biomarkers previously, there is still a need for new survival markers in primary bladder cancer. In this paper, survival associated genetic biomarkers that are differentially expressed

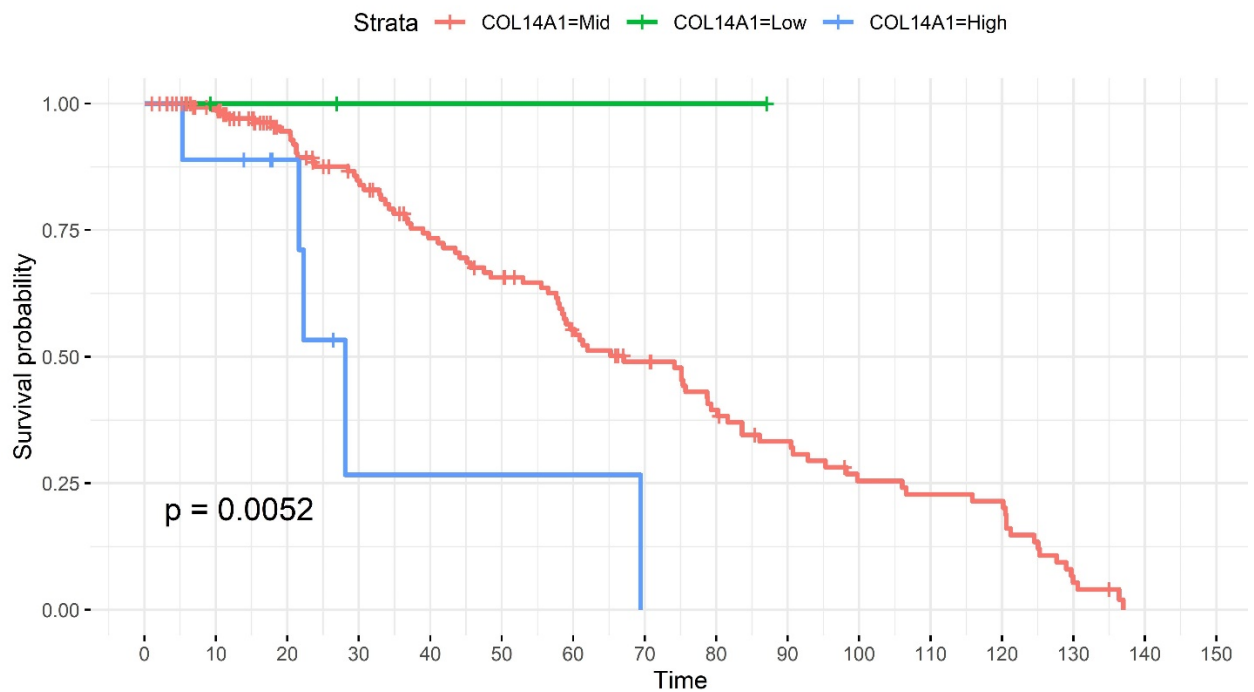
between normal and primary bladder cancer samples have been proposed using a dataset downloaded from the GEO database. The Cox proportional hazard model was used to determine the particular genes that influence overall survival in patients with tumors of the bladder.

Analysis of gene expression profiles revealed a total of 5 979 differentially expressed genes between normal and primary bladder cancer samples. The top five upregulated genes were *CDC20*, *PMM2*, *IQGAP3*, and *GSK3B*, whereas *PLPPR4*, *CYTL1*, *CD160*, *ANO4*, and *PCSK2* were the top five most downregulated genes. *CDC20*, which was identified as the most significantly upregulated gene, has previously been proposed as a potential, quiescent and novel ameliorative target notably in cancer doctoring (L. Wang et al., 2015; Z. Wang et al., 2013), having been previously obtained as a gene of interest in human bladder carcinoma (Choi et al., 2013). *CDC20* was also identified in advanced distant metastasis and poor overall primary bladder cancer survival in humans (Choi et al., 2013). On the other hand, the top downregulated DEG was *PLPPR4* has no prognostic role in primary bladder cancer and results from genomic mutations. *CYTL1*, the second-top most DEG, has been previously highlighted as being differently methylated in bladder urothelial carcinoma (X. Wang et al., 2019).

Differential Expression Analysis for the Identification of Survival Associated Genes in Primary Bladder Cancer using Microarray Data

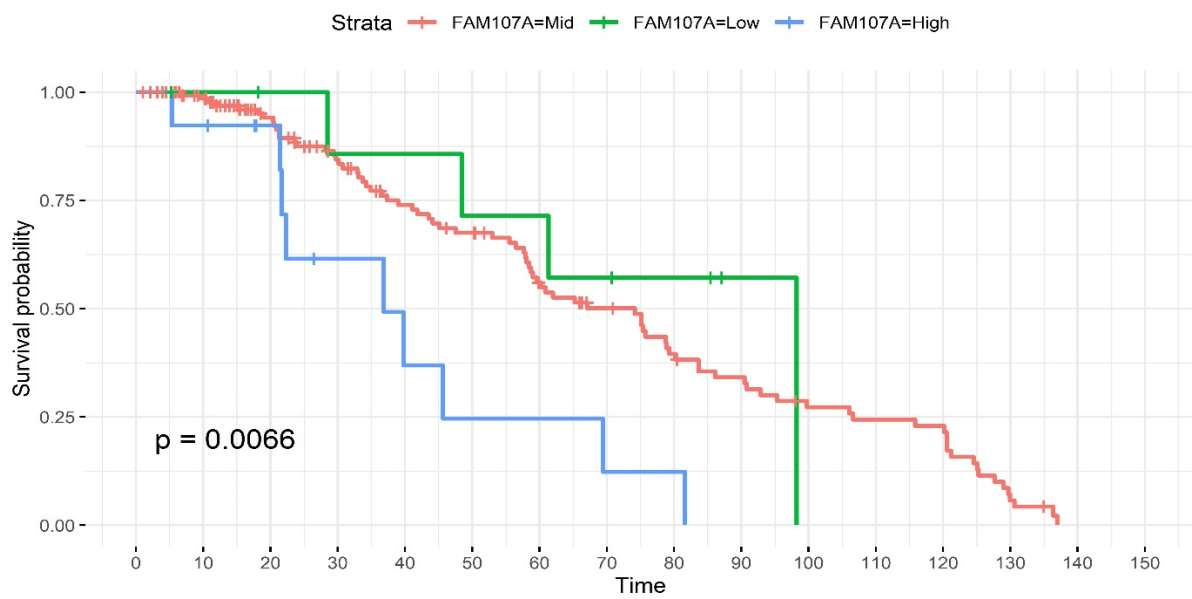


A

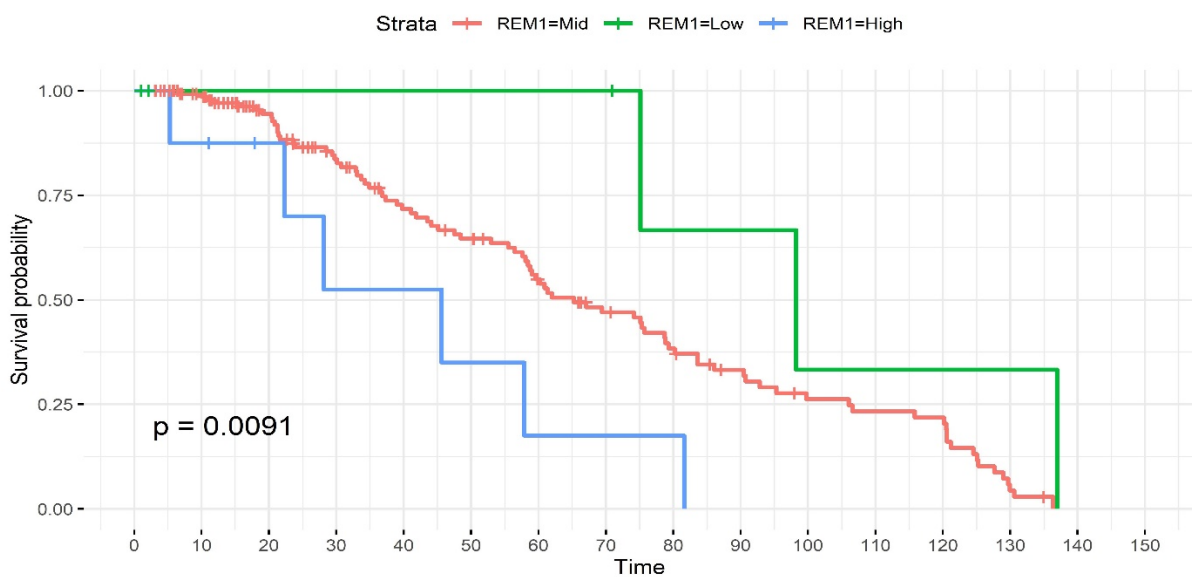


B

Differential Expression Analysis for the Identification of Survival Associated Genes in Primary Bladder Cancer using Microarray Data



C



D

Figure 3: Kaplan-Meier plots of the effect of the selected top differentially expressed genes (DEGs) on primary bladder cancer survival. (A) *EDNRRB*, (B) *COL14A1*. Kaplan-Meier plots of the effect of the top selected top differentially expressed genes (DEGs) on primary bladder cancer survival. (C) *FAM107A*, (D) *REM1*. High expression of *EDNRRB*, *COL14A1*, *FAM107A*, and *REM1* was associated with a statistically significant reduction in overall primary bladder cancer survival.

Differential Expression Analysis for the Identification of Survival Associated Genes in Primary Bladder Cancer using Microarray Data

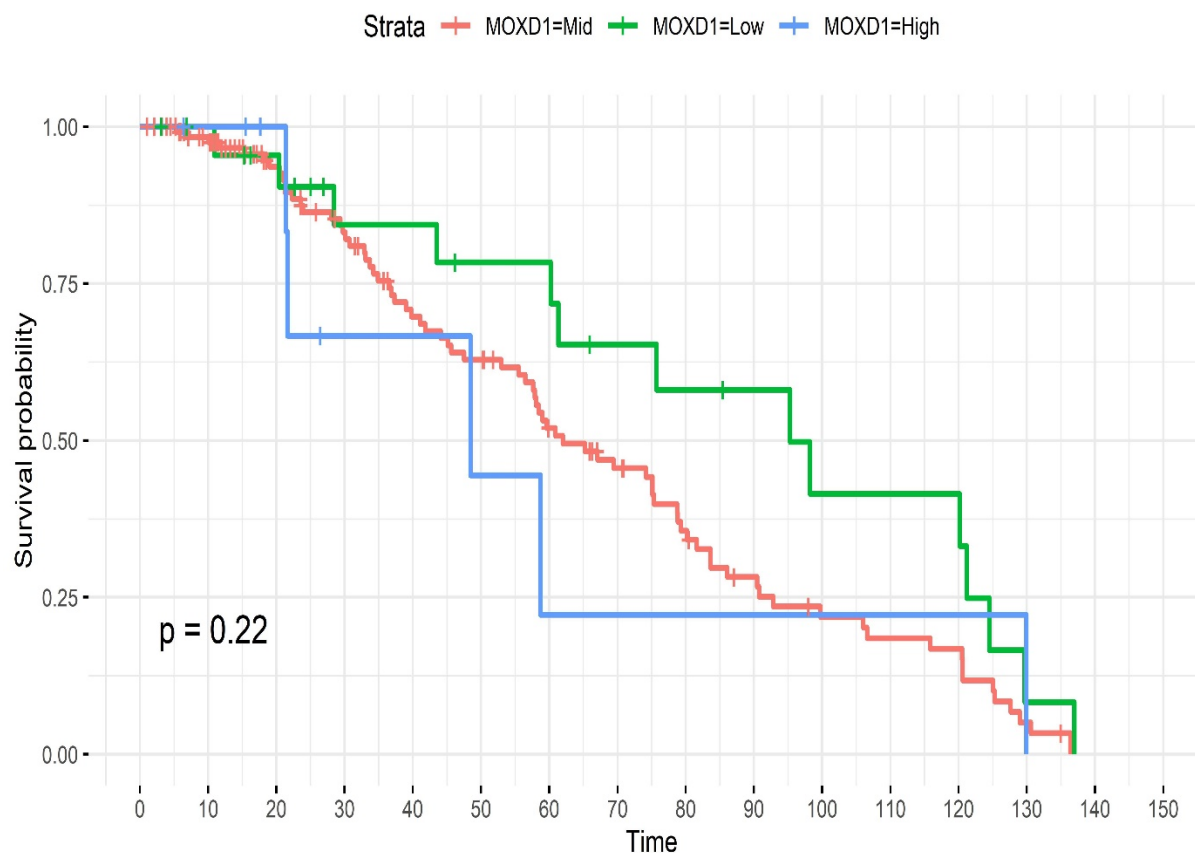


Figure 4: The impact of high, middle, and low expression of *MOXD1* on primary bladder cancer survival. High expression of *MOXD1* was not statistically significantly associated with reducing the survival of primary bladder cancer patients.

Table 4: Top 10 survival-associated Differentially Expressed Genes (DEGs) for primary bladder cancer

GENE SYMBOL	Beta	Standard Error	Z*	P#	LRT!	Wald	LogRank	HR	HRLower	HRUpper
EDNRB	0.6860	0.1443	4.7535	0.0000	0.0000	0.0000	0.0000	1.9858	1.4965	2.6350
COL14A1	0.4276	0.1393	3.0684	0.0022	0.0080	0.0022	0.0019	1.5335	1.1670	2.0151
MOXD1	0.3973	0.1508	2.6336	0.0084	0.0128	0.0084	0.0078	1.4878	1.1070	1.9996
FAM107A	0.4035	0.1497	2.6943	0.0071	0.0127	0.0071	0.0064	1.4970	1.1162	2.0077
REM1	0.4699	0.1291	3.6400	0.0003	0.0012	0.0003	0.0002	1.5999	1.2422	2.0605
FBLN5	0.3350	0.1247	2.6871	0.0072	0.0134	0.0072	0.0067	1.3980	1.0949	1.7850
CADM3	0.5496	0.1816	3.0270	0.0025	0.0066	0.0025	0.0027	1.7325	1.2138	2.4729
VCAM1	0.5882	0.1245	4.7246	0.0000	0.0001	0.0000	0.0000	1.8008	1.4109	2.2984
PAPPA	0.4509	0.1653	2.7277	0.0064	0.0095	0.0064	0.0062	1.5698	1.1353	2.1704
RBP4	0.4383	0.1356	3.2335	0.0012	0.0055	0.0012	0.0008	1.5501	1.1884	2.0218
PI16	0.3508	0.1331	2.6352	0.0084	0.0218	0.0084	0.0061	1.4202	1.0940	1.8435
IL6ST	0.4100	0.1300	3.1543	0.0016	0.0020	0.0016	0.0016	1.5068	1.1679	1.9440
TLR1	0.4130	0.1303	3.1706	0.0015	0.0028	0.0015	0.0013	1.5114	1.1708	1.9510
F10	0.5270	0.1453	3.6271	0.0003	0.0020	0.0003	0.0002	1.6938	1.2741	2.2518
OLFML1	0.4209	0.1504	2.7989	0.0051	0.0070	0.0051	0.0051	1.5234	1.1345	2.0457
LAMA2	0.4954	0.1387	3.5705	0.0004	0.0009	0.0004	0.0003	1.6411	1.2504	2.1540
CHN1	0.5258	0.1276	4.1201	0.0000	0.0001	0.0000	0.0000	1.6918	1.3174	2.1725
GNG2	0.4218	0.1347	3.1318	0.0017	0.0072	0.0017	0.0016	1.5247	1.1710	1.9853
RHOJ	0.3405	0.1048	3.2479	0.0012	0.0029	0.0012	0.0009	1.4056	1.1445	1.7262
SCN2B	0.3551	0.1127	3.1501	0.0016	0.0063	0.0016	0.0017	1.4263	1.1436	1.7790

* Z score, # Probability value, ! Likelihood Ratio Test

KEGG and GO enrichment analysis and functional annotation was performed to demonstrate the functions of the DEGs. The most significantly enriched KEGG pathways based on the top 300 DEGs were cell adhesion molecules (CAMs) and the P13K-Akt signaling pathway. As such, anomalous pathways can be regarded as one of the causes of bladder tumors, as CAMs have been previously linked to invasion and metastasis in human malignancies, and in particular, tumors of the genitourinary tract. CAMs are significant in the interaction between cells as well as the communication amid cells and other apparatuses of the extracellular matrix (Bryan, 2015; Cohen et al., 1997). Furthermore, the P13K-Akt signaling pathway is responsible for regulating cell growth, metastasis, and survival. Alterations in these pathway functions might result in the development of cancers (Sathe & Nawroth, 2018).

DEGs were significantly enriched in GO terms of 'extracellular matrix organization,' 'extracellular region' and 'receptor activity.' Studies have previously reported the role of these terms in bladder cancer (Brunner & Tzankov, 2007; Dozmorov et al., 2006).

The 300 of the selected top differentially expressed gene expression profiles obtained from microarray data were then subjected to proportional Cox hazard models to identify the differently expressed genes and their effect on overall primary bladder cancer survival. DEGs were retrieved based on the defined threshold defined in the previous section. The twenty of the genes selected as estimators of survival in primary bladder cancer were *EDNRB*, *COL14A1*, *MOXD1*, *FAM107A*, *REM1*, *FBLN5*, *CADM3*, *VCAM1*, *PAPPA*, *RBP4*, *PI16*, *IL6ST*, *TLR1*, *F10*, *OLFML1*, *LAMA2*, *CHN1*, *GNG2*, *RHOJ*, and

SCN2B. Five genes, that is, *EDNRB*, *COL14A1*, *MOXD1*, *FAM107A*, *REM1*, were selected for further exploration.

The endothelin B receptor, also called *EDNRB*, is a "G protein-coupled receptor" responsible for activating the second phosphatidylinositol-calcium messenger system. The role of *EDNRB* in cancer has been previously studied, with findings by Davenport et al. (2016) suggesting the downregulation of the gene by promoter hypermethylation during tumor development (tumorigenesis). The downregulation of *EDNRB* results in the alteration of the ET1 signaling pathway (Davenport et al., 2016). The subsequent modifications in this pathway have also been reported to result in the proliferation of tumors, metastasis, and angiogenesis (Lahav et al., 1999; Mallat et al., 1995; Zhao et al., 2009).

This paper found that high expression of *EDNRB* results in a significant reduction in overall bladder cancer survival based on the proportional Cox hazard regression model results. This is particularly due to the impact that *EDNRB* has on the ET1 signaling pathway, whose alliteration results in the propagation of the growth of tumors.

Similarly, collagen, type XIV, alpha 1 (*COL14A1*) was found to significantly reduce the overall survival of patients with primary bladder cancer. The role of *COL14A1* in cancer development has been reported elsewhere (Morris et al., 2010). In this paper, we explain its role in bladder cancer tumorigenesis and progression based on its impact on the process of DNA methylation. Methylation frequency was observed to increase with the overexpression of *COL14A1* in bladder cancer. Methylation results in the addition of a methyl group to a

DNA molecule. This situation can result in a change in a DNA section's activity without the sequence being altered.

DNA methylation on a gene promoter contributes to the repression of gene transcription. In tumorigenesis, altered methylation patterns have been implicated. In particular, the destruction of genes involved in suppressing the growth of tumors upsets the balance that facilitates cell proliferation regulation. This situation results in the propagation of tumors into complete malevolent status resulting in reduced overall patient survival (Wajed et al., 2001). Additionally, *COL14A1* has been previously reported to interact with decorin, which results in the downregulation of receptors (including IGF-I and LRP) that have been implicated in reduced survival and increased growth of cancers (Brandan et al., 2006; Schaefer et al., 2007).

Additionally, 'family with sequence similarity 107 member A' (*FAM107A*) was found to contribute significantly to increases in overall survival of patients with primary bladder cancer in low expression. This can be argued to result from the effect that the gene has in terms of reducing cell growth even though the role of the gene remains controversial (Nakajima et al., 2012). In studies in which the role of *FAM107A* was investigated, varied evidence ranging from the gene suppressing tumor development through inducing apoptosis to the gene driving tumor invasion in gliomas have been reported (Le et al., 2010; Liu et al., 2009). According to a previously conducted study in which the diagnostic value of *FAM107A* in discriminating bladder cancer from hematuria was investigated, high ratios of this gene were associated with bladder cancer (Xu et al., 2019).

On the other hand, the high expression of 'Monooxygenase DBH like 1' (*MOXD1*) did not significantly affect the

overall survival probability of patients with primary bladder cancer. This gene was also found to be significantly downregulated between normal and primary bladder cancer samples. In previous studies, the knockdown of the gene *MOXD1* has been reported to annihilate and suppress the spread of cancerous cells in osteosarcoma through the induction of cellular death (apoptosis) (Han et al., 2019). In this way, the role of *MOXD1* in enhancing the survival of patients with bladder cancer in high expression can be argued to be a result of the induction of the death of cancerous cells that the gene promotes.

High expression of 'GTP-binding protein REM 1' was also associated with reduced survival probability in bladder cancer. The role of *REM1* in cancer has been studied previously, where the gene has been linked to angiogenesis and Ca^{2+} signaling. Similarly, the role that angiogenesis plays with respect to tumor growth and metamorphosis and that of calcium signaling in the growth of cancerous cells has been tackled elsewhere (Capiod et al., 2007; Folkman, 2002).

While studies have previously used the GSE13507 dataset, their findings have been quite limited and varied in terms of the genes that they reported. For instance, W. J. Kim et al. (2010) focused on identifying genes that could be used in the prediction of bladder cancer progression. The authors reported genes, including *S100A8*, *HMOX1*, *MTAP*, *KIF1A*, *COCH*, *CELSR3*, *MGCIT624*, and *PFKFB4*, as possible predictors of muscle-invasive bladder cancer progression. On the other hand, ee et al. (20) focused on E2F1 as well as its associated genes as predictors of progression of bladder cancer from superficial to invasive. In this paper, the same dataset has been used with attention paid to the effect of genes (including *EDNRB*, *COL14A1*, *MOXD1*,

FAM107A, and *REMT1*), which have had a minimal investigation in terms of their influence on the overall survival of patients with primary bladder cancer. Patterns in the expression of these genes in primary bladder cancer that had previously been obscured are unmasked. Additionally, this study adds to the fund of knowledge on prognostic and survival genetic markers of primary bladder cancer.

In summary, this paper involved a comprehensive analysis of DEGs between normal and primary bladder cancer samples. Proportional Cox hazard regression models were then used to study the effect of these DEGs on the overall survival of patients with primary bladder cancer. The role of the selected DEGs in cancer has also been enumerated. Additionally, the function of the selected pathways in bladder cancer tumorigenesis and progression have been discussed. The genes identified in this paper are likely to inform future research aimed at understanding their genetic link to bladder cancer and overall effects on patient survival probability. Also, biological and therapeutic information can be derived from this research as gene expression analysis has been known from previous studies to inform clinical and therapeutic technologies. For instance, gene therapy, such as Gendicine™, was made possible following findings from genetic and genomic studies (Raty et al., 2010). While this study adds substantial knowledge on survival markers for primary bladder cancer, further research should be carried out employing considerably large sample sizes to confirm the significance of these survival-associated DEGs.

Conclusion

This study identified 5 979 genes expressed differently between normal and primary bladder cancer samples derived from microarray gene expression profiling and downloaded from the Gene Expression

Omnibus. 50 DEGs were identified as having a statistically significant effect on the overall survival probability of patients with primary bladder cancer. These results will provide invaluable insight for future research studies in the field, in addition to aiding clinical therapies as markers of bladder cancer survival.

Ethics and consent

Not applicable.

Competing interests

The authors have no competing interests to declare.

Author's contributions

AOO performed the study design, data collection, literature searches, statistical analysis, data interpretation, and manuscript preparation. KWN oversaw the statistical analysis, data interpretation, and manuscript review. All author's read and approved the final manuscript.

References

- Alboukadel, K., Marcin, K., Przemyslaw, B., & Scheipl, F. (2018). Drawing Survival Curves using “ggplot2” [R package survminer version 0.4.3]. *R Package Version 0.4.3*.
- Aljabery, F., Shabo, I., Gimm, O., Jahnsen, S., & Olsson, H. (2018). The expression profile of p14, p53, and p21 in tumor cells is associated with disease-specific survival and the outcome of postoperative chemotherapy treatment in muscle-invasive bladder cancer. *Urologic Oncology: Seminars and Original Investigations*. <https://doi.org/10.1016/j.urolonc.2018.05.025>
- Allory, Y., Beukers, W., Sagraera, A., Flández, M., Marqués, M., Márquez, M., Van Der Keur, K. A., Dyrskjot, L., Lurkin, I., Vermeij, M., Carrato, A., Lloreta, J.,

- Lorente, J. A., Carrillo-De Santa Pau, E., Masius, R. G., Kogevinas, M., Steyerberg, E. W., Van Tilborg, A. A. G., Abas, C., ... Real, F. X. (2014). Telomerase reverse transcriptase promoter mutations in bladder cancer: High frequency across stages, detection in urine, and lack of association with outcome. *European Urology*. <https://doi.org/10.1016/j.eururo.2013.08.052>
- Blighe, K. (2019). *RegParallel: Standard regression functions in R enabled for parallel processing over large data-frames*. R package version 1.4.0. <https://github.com/kevinblighe/RegParallel>
- Bolstad, B. M., Irizarry, R. A., Åstrand, M., & Speed, T. P. (2003). A comparison of normalization methods for high-density oligonucleotide array data based on variance and bias. *Bioinformatics*. <https://doi.org/10.1093/bioinformatics/19.2.185>
- Borgan, Ørnulf. (, 2001). Modeling Survival Data: Extending the Cox Model. Terry M. Therneau and Patricia M. Grambsch, Springer-Verlag, New York, 2000. No. of pages: xiii + 350. Price: \$69.95. ISBN 0-387-98784-3. *Statistics in Medicine*. <https://doi.org/10.1002/sim.956>
- Botling, J., Edlund, K., Lohr, M., Hellwig, B., Holmberg, L., Lambe, M., Berglund, A., Ekman, S., Bergqvist, M., Pontén, F., König, A., Fernandes, O., Karlsson, M., Helenius, G., Karlsson, C., Rahnenführer, J., Hengstler, J. G., & Micke, P. (2013). Biomarker discovery in non-small cell lung cancer: Integrating gene expression profiling, meta-analysis, and tissue microarray validation. *Clinical Cancer Research*. <https://doi.org/10.1158/1078-0432.CCR-12-1139>
- Brandan, E., Retamal, C., Cabello-Verrugio, C., & Marzolo, M. P. (2006). The low-density lipoprotein receptor-related protein functions as an endocytic receptor for decorin. *Journal of Biological Chemistry*. <https://doi.org/10.1074/jbc.M602919200>
- Brunner, A., & Tzankov, A. (2007). The Role of Structural Extracellular Matrix Proteins in Urothelial Bladder Cancer (Review). *Biomarker Insights*. <https://doi.org/10.4137/bmi.s294>
- Bryan, R. T. (2015). Cell adhesion and urothelial bladder cancer: The role of cadherin switching and related phenomena. In *Philosophical Transactions of the Royal Society B: Biological Sciences*. <https://doi.org/10.1098/rstb.2014.0042>
- Buttigliero, C., Tucci, M., Vignani, F., Scagliotti, G. V., & Di Maio, M. (2017). Molecular biomarkers to predict response to neoadjuvant chemotherapy for bladder cancer. In *Cancer Treatment Reviews*. <https://doi.org/10.1016/j.ctrv.2017.01.002>
- Calza, S., Valentini, D., & Pawitan, Y. (2008). Normalization of oligonucleotide arrays based on the least-variant set of genes. *BMC Bioinformatics*. <https://doi.org/10.1186/1471-2105-9-140>
- Capiod, T., Shuba, Y., Skryma, R., & Prevarskaya, N. (2007). Calcium signaling and cancer cell growth. *Subcellular Biochemistry*. https://doi.org/10.1007/978-1-4020-6191-2_15
- Cheadle, C., Vawter, M. P., Freed, W. J., & Becker, K. G. (2003). Analysis of microarray data using Z score transformation. *Journal of Molecular Diagnostics*. [https://doi.org/10.1016/S1525-1578\(10\)60455-2](https://doi.org/10.1016/S1525-1578(10)60455-2)
- Chen, Z., Liu, G., Hossain, A., Danilova, I.

- G., Bolkov, M. A., Liu, G., Tuzankina, I. A., & Tan, W. (2019). A co-expression network for differentially expressed genes in bladder cancer and a risk score model for predicting survival. *Hereditas*. <https://doi.org/10.1186/s41065-019-0100-1>
- Choi, J. W., Kim, Y., Lee, J. H., & Kim, Y. S. (2013). High expression of spindle assembly checkpoint proteins CDC20 and MAD2 is associated with poor prognosis in urothelial bladder cancer. *Virchows Archiv*. <https://doi.org/10.1007/s00428-013-1473-6>
- Cohen, M. B., Griebeling, T. L., Ahaghotu, C. A., Rokhlin, O. W., & Ross, J. S. (1997). Cellular adhesion molecules in urologic malignancies. In *American Journal of Clinical Pathology*. <https://doi.org/10.1093/ajcp/107.1.56>
- Daneshmand, S. (2019). Epidemiology and risk factors of urothelial (transitional cell) carcinoma of the bladder. *International Journal of Urology*.
- Davenport, A. P., Hyndman, K. A., Dhaun, N., Southan, C., Kohan, D. E., Pollock, J. S., Pollock, D. M., Webb, D. J., & Maguire, J. J. (2016). Endothelin. In *Pharmacological Reviews*. <https://doi.org/10.1124/pr.115.011833>
- Dozmorov, M. G., Kyker, K. D., Saban, R., Knowlton, N., Dozmorov, I., Centola, M. B., & Hurst, R. E. (2006). Analysis of the interaction of the extracellular matrix and phenotype of bladder cancer cells. *BMC Cancer*. <https://doi.org/10.1186/1471-2407-6-12>
- Durinck, S., Spellman, P. T., Birney, E., & Huber, W. (2009). Mapping identifiers for the integration of genomic datasets with the R/ Bioconductor package biomaRt. *Nature Protocols*. <https://doi.org/10.1038/nprot.2009.97>
- Folkman, J. (2002). Role of angiogenesis in tumor growth and metastasis. In *Seminars in oncology*. <https://doi.org/10.1053/sonc.2002.37263>
- Gao, X., Chen, Y., Chen, M., Wang, S., Wen, X., & Zhang, S. (2018). Identification of key candidate genes and biological pathways in bladder cancer. *PeerJ*. <https://doi.org/10.7717/peerj.6036>
- Han, X., He, J., Wang, Z., Bai, Z., Wang, W., & Li, X. (2019). Monooxygenase DBH, like 1 (MOXD1) knockdown, suppresses the proliferation and xenograft tumor growth of osteosarcoma cells via inducing apoptosis. *Panminerva Medica*. <https://doi.org/10.23736/S0031-0808.19.03717-0>
- Huang, Z., Yang, Q., & Huang, Z. (2018). Identification of critical genes and five prognostic biomarkers associated with colorectal cancer. *Medical Science Monitor*. <https://doi.org/10.12659/MSM.907224>
- Kim, W. J., Kim, E. J., Kim, S. K., Kim, Y. J., Ha, Y. S., Jeong, P. P., Kim, M. J., Yun, S. J., Lee, K. M., Moon, S. K., Lee, S. C., Cha, E. J., & Bae, S. C. (2010). Predictive value of progression-related gene classifier in primary non-muscle-invasive bladder cancer. *Molecular Cancer*. <https://doi.org/10.1186/1476-4598-9-3>
- Kim, W. T., Yun, S. J., Park, C., Kim, I. Y., Moon, S. K., Kwon, T. G., Choi, Y. H., & Kim, W. J. (2010). Identification of C16orf74 as a marker of progression in primary non-muscle-invasive bladder cancer. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0015260>
- Lahav, R., Heffner, G., & Patterson, P. H. (1999). An endothelin receptor B antagonist inhibits growth and induces cell death in human melanoma cells in vitro and in vivo. *Proceedings of the*

Differential Expression Analysis for the Identification of Survival Associated Genes in Primary Bladder Cancer using Microarray Data

- National Academy of Sciences of the United States of America.*
<https://doi.org/10.1073/pnas.96.20.11496>
- Le, P. U., Angers-Loustau, A., De Oliveira, R. M. W., Ajlan, A., Brassard, C. L., Dudley, A., Brent, H., Siu, V., Trinh, G., Mölenkamp, G., Wang, J., Seyed Sadr, M., Bedell, B., Del Maestro, R. F., & Petrecca, K. (2010). DRR drives brain cancer invasion by regulating cytoskeletal-focal adhesion dynamics. *Oncogene*.
<https://doi.org/10.1038/onc.2010.216>
- Lee, J. S., Leem, S. H., Lee, S. Y., Kim, S. C., Park, E. S., Kim, S. B., Kim, S. K., Kim, Y. J., Kim, W. J., & Chu, I. S. (2010). Expression signature of E2F1 and its associated genes predict superficial to the invasive progression of bladder tumors. *Journal of Clinical Oncology*.
<https://doi.org/10.1200/JCO.2009.25.0977>
- Liu, Q., Zhao, X. Y., Bai, R. Z., Liang, S. F., Nie, C. L., Yuan, Z., Wang, C. T., Wu, Y., Chen, L. J., & Wei, Y. Q. (2009). Induction of tumor inhibition and apoptosis by a candidate tumor suppressor gene DRR1 on 3p21.1. *Oncology Reports*.
https://doi.org/10.3892/or_00000538
- Malats, N., & Real, F. X. (2015). Epidemiology of Bladder Cancer. In *Hematology/Oncology Clinics of North America*.
<https://doi.org/10.1016/j.hoc.2014.10.001>
- Mallat, A., Fouassier, L., Preaux, A. M., Serradeil-Le Gal, C., Raufaste, D., Rosenbaum, J., Dhumeaux, D., Jouneaux, C., Mavier, P., & Lotersztajn, S. (1995). Growth inhibitory properties of endothelin-1 in human hepatic myofibroblastic Ito cells. An endothelin B receptor-mediated pathway. *Journal of Clinical Investigation*.
<https://doi.org/10.1172/JCI118052>
- Martinez-Romero, J., Bueno-Fortes, S., Martín-Merino, M., Ramirez De Molina, A., & De Las Rivas, J. (2018). Survival marker genes of colorectal cancer derived from consistent transcriptomic profiling. *BMC Genomics*.
<https://doi.org/10.1186/s12864-018-5193-9>
- Mikhailenko, D. S., & Nemtsova, M. V. (2016). Point somatic mutations in bladder cancer: key carcinogenesis events, diagnostic markers, and therapeutic targets. In *Urologiia (Moscow, Russia : 1999)*.
- Morris, M. R., Ricketts, C., Gentle, D., Abdulrahman, M., Clarke, N., Brown, M., Kishida, T., Yao, M., Latif, F., & Maher, E. R. (2010). Identification of candidate tumor suppressor genes frequently methylated in renal cell carcinoma. *Oncogene*.
<https://doi.org/10.1038/onc.2009.493>
- Nakajima, H., Koizumi, K., Tanaka, T., Ishigaki, Y., Yoshitake, Y., Yonekura, H., Sakuma, T., Fukushima, T., Umehara, H., Ueno, S., Minamoto, T., & Motoo, Y. (2012). Loss of HITS (FAM107B) expression in cancers of multiple organs: A tissue microarray analysis. *International Journal of Oncology*.
<https://doi.org/10.3892/ijco.2012.1550>
- Quackenbush, J. (2002). Microarray data normalization and transformation. In *Nature Genetics*.
<https://doi.org/10.1038/ng1032>
- Raty, J., Pikkarainen, J., Wirth, T., & Yla-Herttuala, S. (2010). Gene Therapy: The First Approved Gene-Based Medicines, Molecular Mechanisms, and Clinical Indications. *Current Molecular Pharmacology*.
<https://doi.org/10.2174/1874467210801010013>
- Richters, A., Aben, K. K. H., & Kiemeny, L. A. L. M. (2019). The global burden of urinary bladder cancer: an update. In the *World Journal of Urology*.

- <https://doi.org/10.1007/s00345-019-02984-4>
- Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*. <https://doi.org/10.1093/nar/gkv007>
- Sathe, A., & Nawroth, R. (2018). Targeting the PI3K/AKT/mTOR pathway in bladder cancer. In *Methods in Molecular Biology*. https://doi.org/10.1007/978-1-4939-7234-0_23
- Schaefer, L., Tsalastra, W., Babelova, A., Baliova, M., Minnerup, J., Sorokin, L., Gröne, H. J., Reinhardt, D. P., Pfeilschifter, J., Iozzo, R. V., & Schaefer, R. M. (2007). Decorin-mediated regulation of fibrillin-1 in the kidney involves the insulin-like growth factor-I receptor and mammalian target of rapamycin. *American Journal of Pathology*. <https://doi.org/10.2353/ajpath.2007.060497>
- Seiler, R., Ashab, H. A. D., Erho, N., van Rhijn, B. W. G., Winters, B., Douglas, J., Van Kessel, K. E., Fransen van de Putte, E. E., Sommerlad, M., Wang, N. Q., Choerung, V., Gibb, E. A., Palmer-Aronsten, B., Lam, L. L., Buerki, C., Davicioni, E., Sjö Dahl, G., Kardos, J., Hoadley, K. A., ... Black, P. C. (2017). Impact of Molecular Subtypes in Muscle-invasive Bladder Cancer on Predicting Response and Survival after Neoadjuvant Chemotherapy [Figure presented]. *European Urology*. <https://doi.org/10.1016/j.eururo.2017.03.030>
- Shi, S., & Tian, B. (2019). Identification of biomarkers associated with progression and prognosis in bladder cancer via co-expression analysis. *Cancer Biomarkers*. <https://doi.org/10.3233/CBM-181940>
- Shukla, S., Evans, J. R., Malik, R., Feng, F. Y., Dhanasekaran, S. M., Cao, X., Chen, G., Beer, D. G., Jiang, H., & Chinnaiyan, A. M. (2017). Development of an RNA-seq based prognostic signature in lung adenocarcinoma. *Journal of the National Cancer Institute*. <https://doi.org/10.1093/jnci/djw200>
- Sun, X., Cheng, G., Hao, M., Zheng, J., Zhou, X., Zhang, J., Taichman, R. S., Pienta, K. J., & Wang, J. (2010). CXCL12 / CXCR4 / CXCR7 chemokine axis and cancer progression. In *Cancer and Metastasis Reviews*. <https://doi.org/10.1007/s10555-010-9256-x>
- Tang, H., Wang, S., Xiao, G., Schiller, J., Papadimitrakopoulou, V., Minna, J., Wistuba, I. I., & Xie, Y. (2017). Comprehensive evaluation of published gene expression prognostic signatures for biomarker-based lung cancer clinical studies. In *Annals of Oncology*. <https://doi.org/10.1093/annonc/mdw683>
- Theodorescu, D., Sapinoso, L. M., Conaway, M. R., Oxford, G., Hampton, G. M., & Frierson, H. F. (2004). Reduced expression of metastasis suppressor RhoGDI2 is associated with decreased survival for patients with bladder cancer. *Clinical Cancer Research*. <https://doi.org/10.1158/1078-0432.CCR-03-0653>
- Wajed, S. A., Laird, P. W., & DeMeester, T. R. (2001). DNA methylation: An alternative pathway to cancer. In *Annals of Surgery*. <https://doi.org/10.1097/00000658-200107000-00003>
- Wang, K., Liu, T., Liu, C., Meng, Y., Yuan, X., Liu, L., Ge, N., Liu, J., Wang, C., Ren, H., Yan, K., Hu, S., Xu, Z., Fan, Y., & Xu, D. (2015). TERT Promoter Mutations and TERT mRNA but Not FGFR3 Mutations Are Urinary Biomarkers in Han Chinese Patients With Urothelial Bladder Cancer. *The*

Differential Expression Analysis for the Identification of Survival Associated Genes in Primary Bladder Cancer using Microarray Data

- Oncologist*.
<https://doi.org/10.1634/theoncologist.2014-0391>
- Wang, L., Zhang, J., Wan, L., Zhou, X., Wang, Z., & Wei, W. (2015). Targeting Cdc20 as a novel cancer therapeutic strategy. In *Pharmacology and Therapeutics*.
<https://doi.org/10.1016/j.pharmthera.2015.04.002>
- Wang, X., Li, T., Cheng, Y., Wang, P., Yuan, W., Liu, Q., Yang, F., Liu, Q., Ma, D., Ding, S., Wang, J., & Han, W. (2019). CYTL1 inhibits tumor metastasis with decreasing STAT3 phosphorylation. *OncoImmunology*.
<https://doi.org/10.1080/2162402X.2019.1577126>
- Wang, Z., Wan, L., Zhong, J., Inuzuka, H., Liu, P., Sarkar, F. H., & Wei, W. (2013). Cdc20: A Potential Novel Therapeutic Target for Cancer Treatment. *Current Pharmaceutical Design*.
<https://doi.org/10.2174/1381612811319180005>
- Wong, M. C. S., Fung, F. D. H., Leung, C., Cheung, W. W. L., Goggins, W. B., & Ng, C. F. (2018). The global epidemiology of bladder cancer: A joinpoint regression analysis of its incidence and mortality trends and projection. *Scientific Reports*.
<https://doi.org/10.1038/s41598-018-19199-z>
- Xu, Y., Kim, Y. H., Jeong, P., Piao, X. M., Byun, Y. J., Kang, H. W., Kim, W. T., Lee, J. Y., Kim, I. Y., Moon, S. K., Choi, Y. H., Yun, S. J., & Kim, W. J. (2019). Diagnostic value of combined IQGAP3/BMP4 and IQGAP3/FAM107A expression ratios in urinary cell-free DNA for discriminating bladder cancer from hematuria. *Urologic Oncology: Seminars and Original Investigations*.
<https://doi.org/10.1016/j.urolonc.2018.10.023>
- Yu, G. (2019). *enrichplot: Visualization of Functional Enrichment Result*. R package version 1.6.1.
<https://github.com/YuLab-SMU/enrichplot>
- Yu, G., Wang, L. G., Han, Y., & He, Q. Y. (2012). ClusterProfiler: An R package for comparing biological themes among gene clusters. *OMICS A Journal of Integrative Biology*.
<https://doi.org/10.1089/omi.2011.0118>
- Zhang, X., & Zhang, Y. (2015). Bladder Cancer and Genetic Mutations. *Cell Biochemistry and Biophysics*.
<https://doi.org/10.1007/s12013-015-0574-z>
- Zhao, B. J., Sun, D. G., Zhang, M., Tan, S. N., & Ma, X. (2009). Identification of aberrant promoter methylation of EDNRB gene in esophageal squamous cell carcinoma. *Diseases of the Esophagus*.
<https://doi.org/10.1111/j.1442-2050.2008.00848.x>