# Faculty Development Award Project Report:

## Title: Discovery of novel expression of biomarkers for the evaluation of response to immunotherapy in bladder cancer

**Dr. SAURABH GAYEN**, Chemistry & Physics Department, Lincoln University, PA 19352

## Introduction:

Bladder cancer (BlCa) is the sixth most common cancer in men worldwide and the tenth when considering both the genders together [1]. Bladder cancer is estimated to be 83,730 new cases and 4.4% of all new cancer cases in the US in 2021. In terms of costs of cancer care, it is one of the most expensive cancers prevalent in the United States, resulting in 17,200 deaths [2]. The standard treatment is transurethral resection of bladder tumors, followed by intravesical treatment with Bacillus Calmette–Guerin (BCG) or cisplatin based chemotherapy [3]. Despite advances in treatment, over the past decades treatment procedures for BlCa remained relatively unchanged until the emergence of programmed cell death protein (PD-1) and programmed death ligand 1 (PD-L1) immune checkpoint therapies [4–6]. The USA FDA approvals of immunotherapy agents Atezolizumab, Nivolumab, Pembrolizumab, Avelumab, and Durvalumab represent a major paradigm shift in treating BlCa. Immunotherapy has now emerged as an important treatment modality for BlCa but progress towards developing novel biomarkers is a need of the hour. Standardized, reproducible biomarkers are needed to accurately guide treatment decisions to predict responders to immunotherapy. Next generation of predictive biomarkers for immunotherapy should involve targeted gene expression profile with particular attention to immune gene signatures [7].

The present study focuses to identify gene expression profiles that not only play a role in development of BlCa tumors but has the potential to predict responsiveness towards Pembrolizumab based immunotherapy. First, we used publicly available microarray datasets GSE38264 [8], encompassing expression profiles of 10 normal bladder tissue and 28 tumor tissues from BlCa patients, for comprehensive analysis of differential expression using the Transcriptome Analysis Console (TAC) Software (Thermo Fisher Scientific, [9]). Subsequently, Ingenuity Pathway Analysis (IPA, QIAGEN Redwood City, https://digitalinsights.qiagen.com/IPA [10]) helped to delineate the essential cellular pathways involved in the development of this tumor. Then we identified critical genes that are responsible for the responsiveness of the BlCa patients towards Pembrolizumab immunotherapy. For this purpose, differential expression of the publicly available microarray dataset GSE111636 [11] is analysed. Finally, we compared human bladder tumor sample to adjacent normal bladder tissue sample from BlCa patients using Strand NGS (next generation sequencing, [12]) software for the dataset obtained from GSE133624[13].

## Methods:

Analysis of different datasets obtained from GEO database was performed using the following softwares.

### 1. TAC 4.0.1 (Transcriptome Analysis Console):

Using TAC Software we can analyze microarray data set to perform the following:

- Array QC and data normalization
- Statistical tests for differential expression
- Focus on genes or pathways of interest
- Explore interactions between coding and non-coding RNA
- Interpret complex alternative splicing events
- Link out to publicly available annotations
- Obtain sequence information to design validation experiments

**Streamlined Workflow**


### 2. Strand NGS

Strand NGS is an integrated platform that provides analysis, management and visualization tools for next-generation sequencing data. It supports workflows for RNA-Seq, DNA-Seq, ChIP-Seq, Methyl-Seq, MeDIP-Seq and small RNA-Seq experiments. We analyzed RNA-seq data where the workflow includes standard differential expression analysis for different experimental conditions, as well as differential splicing analysis.

### 3. IPA (Qiagen Ingenuity Pathway Analysis)

IPA is an all-in-one, web-based software application that enables analysis, integration, and understanding of data from gene expression, miRNA, and SNP microarrays, as well as metabolomics, proteomics, and RNA-seq experiments. IPA can also be used for analysis of small-scale experiments that generate gene and chemical lists. IPA allows searches for targeted information on genes, proteins, chemicals, and drugs, and building of interactive models of experimental systems. Data analysis and search capabilities help in understanding the significance of data, specific targets, or candidate biomarkers in the context of larger biological or chemical systems.

## 4. Morpheus (Broad Institute)

The core interface in Morpheus is a heat map, in which a matrix of values is mapped to a matrix of colors. Heat maps are one of the most widely used visualizations in the sciences because they allow you to find patterns in your data, compact a large amount of information into a small space, and are a natural representation of a matrix [14].

# Results:

**Part I:** First the data from GSE38264 is analysed using TAC software. The data consists of 10 samples from normal bladder tissues and 28 tumor tissue samples from BlCa patients. 3D PCA plot in **Figure 1** shows that samples from normal bladder tissues and tumor tissues formed clearly separate groups.

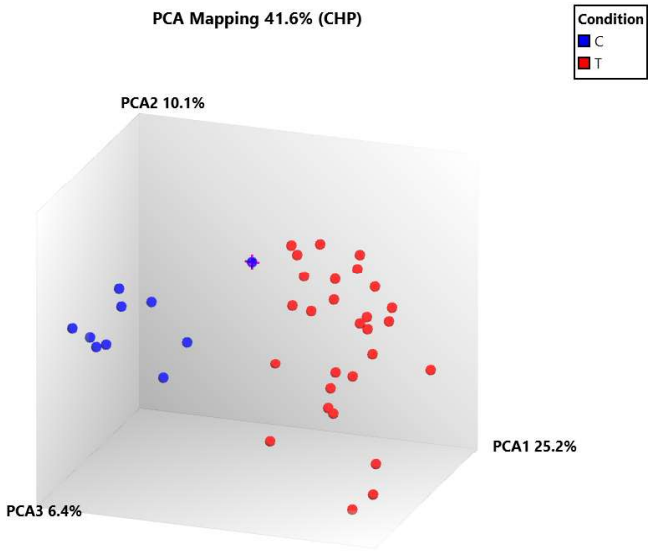
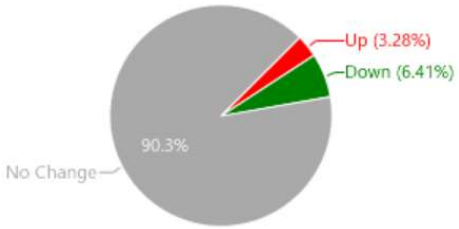



**Figure 1. PCA Plot**

We looked for genes differentially expressed in tumor samples compared to the controls. Following is the summary of our observation with different filters.

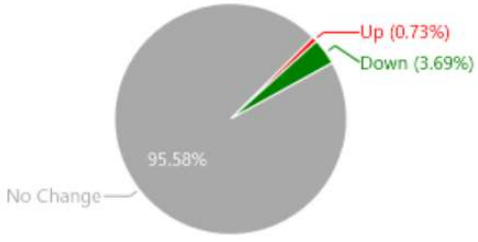
Filter criteria:
- Fold Change: > 1.5 or < -1.5
- P-val: < 0.05

Total number of genes: 29096
- Genes passed filter criteria: 2821 (9.7%)
  - Up-Regulated: 955 (33.85%)
  - Down-Regulated: 1866 (66.15%)




Filter criteria:
- Fold Change: > 2 or < -2
- P-val: < 0.05

Total number of genes: 29096
- Genes passed filter criteria: 1287 (4.42%)
  - Up-Regulated: 212 (16.47%)
  - Down-Regulated: 1075 (83.53%)



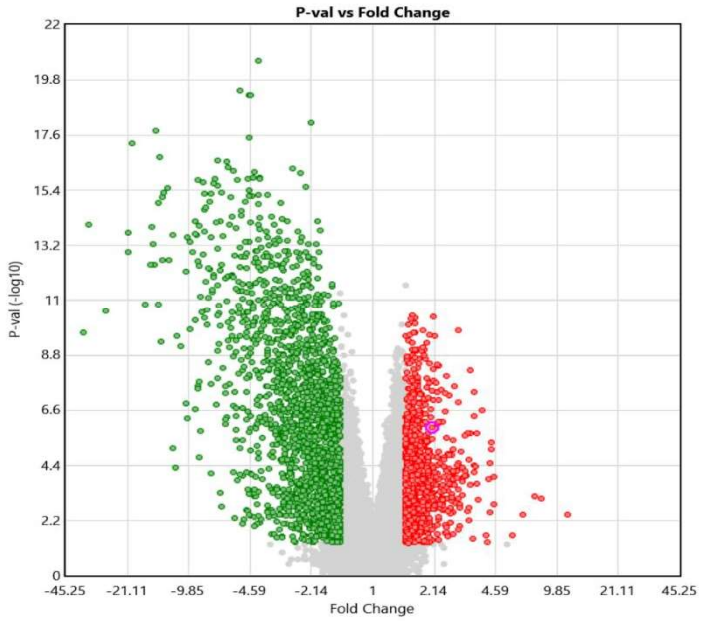



**Figure 2. Volcano plot** of statistically significant genes are shown in this figure. This plot depicts the differentially expressed genes between the tumor versus non tumor groups. The x-axis shows the Fold Change and the y-axis shows the significance of differential expression of each probe (gene), represented as the negative log of the p-value.

**Wiki pathways** in TAC reveals 2 important pathways related to Cancer immunotherapy and Bladder Cancer. We find that all 8 genes involved in Cancer immunotherapy are downregulated. Whereas, the Bladder Cancer pathway involves 7 genes, 6 of them are upregulated and 1 downregulated. The corresponding **heat map** (generated by Morpheus, Broad Institute software) is shown below in **Figure3**.

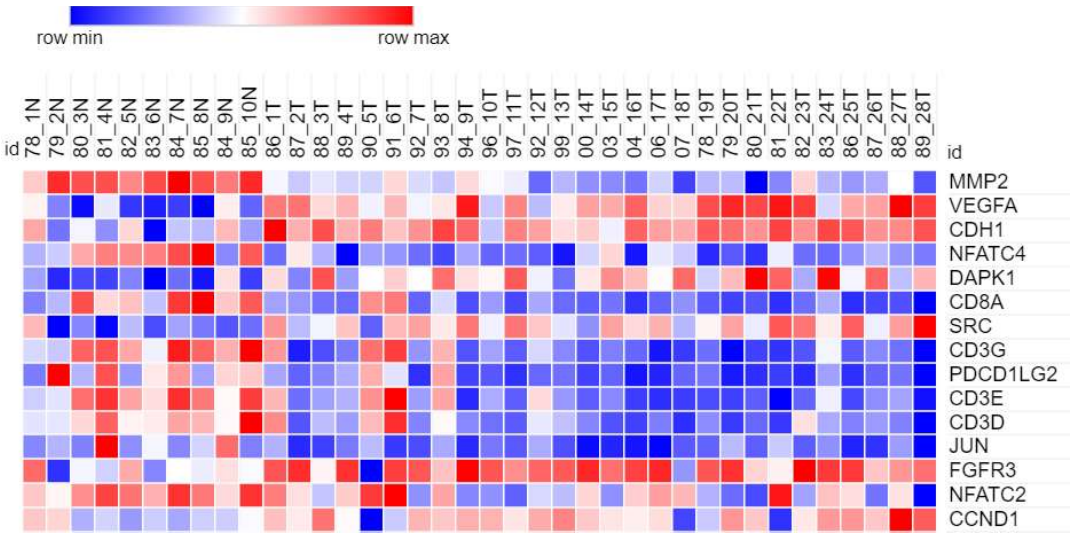


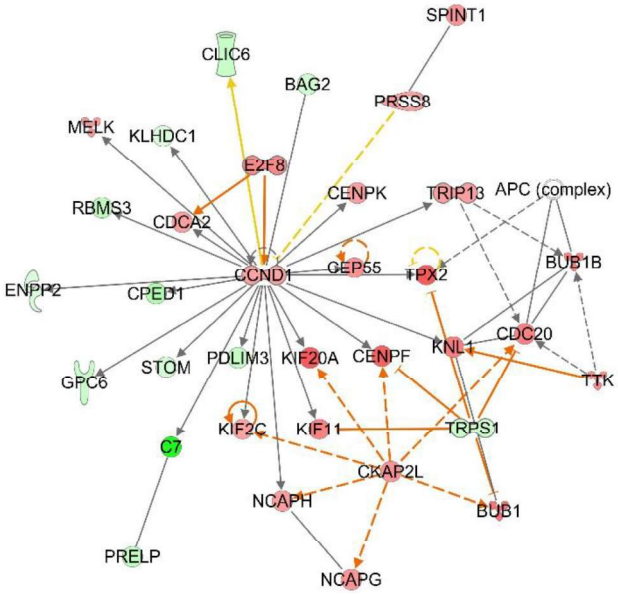

**Figure 3. Heat Map**

Pathway analysis was performed using Ingenuity Pathway Analysis (IPA) software (QIAGEN Inc.,) by loading 1287 probe sets (FC ≥ 2, P-value ≤ 0.05) that were differentially expressed with BlCa tumor group. Ingenuity Canonical Pathway Analysis identified 213 pathways that were significantly [-log(p-value) = 1.3] modified in tumor group, of which top important pathways are depicted in **Table 1**. The top Canonical Pathway revealed in the IPA is Hepatic Fibrosis which is reported to be a common causal factor in many Bladder Cancer patients. Forty-two diseases and functions were modified by differential gene expression with tumor. Important diseases and functions modified in the tumor group are listed in **Table 1**. Cancer is the top modified disease in the analysis. The modification of Cardiotoxicity with tumor included Cardiac Infarction, Arrhythmia, Dilation, Enlargement and Pulmonary Hypertension. Hepatotoxic functions namely Hepatocellular Carcinoma, Liver Hyperplasia/Hyperproliferation, Liver Cirrhosis, Fibrosis, and Inflammation/Hepatitis are altered with the occurrence of tumor. Altered Nephrotoxic functions are Nephrosis, Renal Inflammation, Renal Nephritis, Glomerular Injury, and Renal Hydronephrosis. IPA network analysis has identified networks related to Cancer (**Figure 4**). **Network 1** is related to Cancer, Hematological Disease, Immunological Disease (Score = 39, Focus Molecules 34). **Network 4** is related to Cancer, Cell-To-Cell Signaling and Interaction, Organismal Injury and Abnormalities (Score = 32, Focus Molecules 31). Network 5 involves Cancer, Organismal Injury and Abnormalities, Reproductive System Disease (Score = 30, Focus Molecules 30).

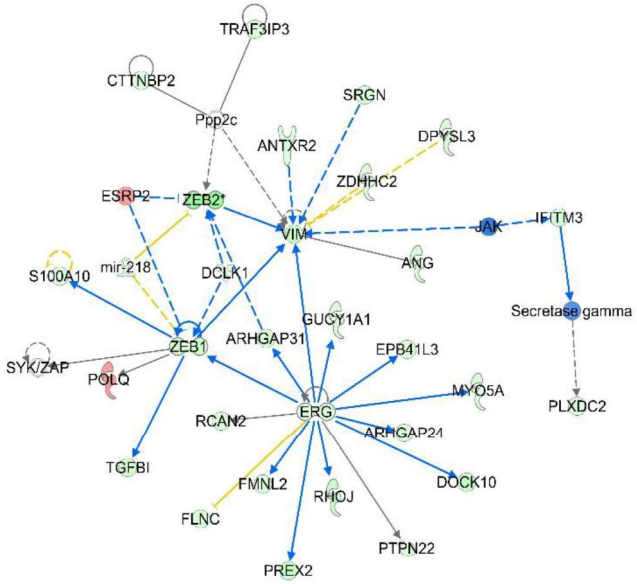
**Table 1.**

| **Ingenuity Canonical Pathways:** | **-log(p-value)** |
|---|---|
| Hepatic Fibrosis / Hepatic Stellate Cell Activation | 17 |
| Cardiac Hypertrophy Signaling (Enhanced) | 11.8 |
| Hepatic Fibrosis Signaling Pathway | 10.9 |
| Atherosclerosis Signaling | 10.1 |
| Tumor Microenvironment Pathway | 9.59 |
| Pulmonary Fibrosis Idiopathic Signaling Pathway | 9.08 |
| Complement System | 7.76 |
| Molecular Mechanisms of Cancer | 6.16 |
| **Diseases and Functions:** | **p-value** |
| Cancer | 1.28E-82 - 4.52E-13 |
| Organismal Injury and Abnormalities | 1.28E-82 - 4.52E-13 |
| Reproductive System Disease | 1.28E-82 - 2.23E-13 |
| Respiratory Disease | 7.58E-41 - 3.29E-14 |
| Hematological Disease | 7.06E-38 - 1.19E-13 |
| Immunological Disease | 8.62E-37 - 2.03E-13 |
| Inflammatory Disease | 8.62E-37 - 1.53E-13 |
| Immune Cell Trafficking | 3.41E-31 - 3.47E-13 |
| Cardiovascular Disease | 3.21E-28 - 2.57E-13 |
| Neurological Disease | 5.92E-26 - 7.7E-15 |

**A. Network 1: Cancer, Hematological Disease, Immunological Disease (Score 39, Focus Molecules 34)**




**B. Network 4: Cancer, Cell-To-Cell Signaling and Interaction, Organismal Injury and Abnormalities (Score 32, Focus Molecules 31)**



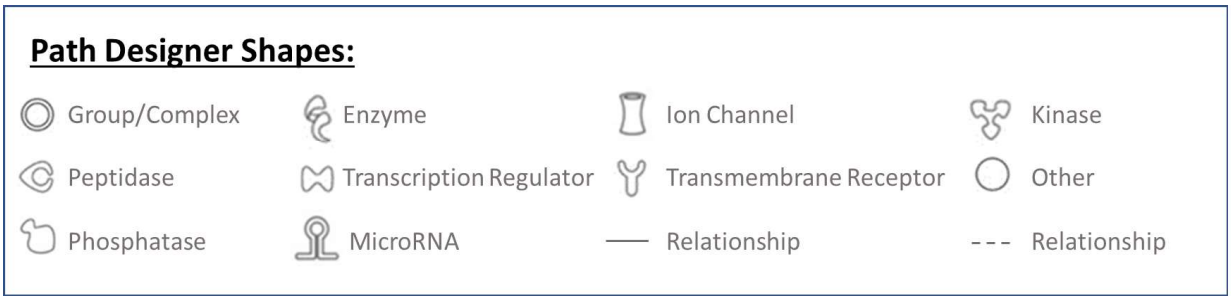




**Figure 4**. Top 2 networks related to Cancer that are picked up by pathway analysis are presented in this figure (panels A, B). Pathway analysis was performed using Ingenuity Pathway Analysis (IPA) software (QIAGEN Inc., https://digitalinsights.qiagen.com/IPA) by loading the 1287 probe sets that were differentially expressed in tumor group. The red filled path designer shapes are upregulated genes and green filled path designer shapes are downregulated genes.

**Part II:** A form of immunotherapy called **bacillus Calmette-Guérin therapy** is typically given after surgery for bladder cancer that hasn't grown into the muscle. However, this is not useful for advanced stage Bladder cancer treatment. In 2014 FDA approved a new class of immunotherapy drugs called **Immune checkpoint inhibitors (ICI)** to treat advanced stage BlCa. The major problem is ICIs may not be very effective for many cancer patients. That is why in this section we try to find out potential biomarkers that could predict the efficacy of ICI used by a patient. Data in GSE111636 comprises of samples from 11 cancer patients collected prior to treatment by ICIs. Depending upon the outcome of immunotherapy, those 11 patients are grouped as responder and progressor. Immunotherapy treatment worked to contain the cancer in the case of responders. Whereas, the progressor didn't benefit from the immunotherapy. There were 5 responders and 6 progressors. We analyzed the gene expressions using TAC software. In **Figure 5**, 3D PCA plot shows clear separation of the progressors from the responders.

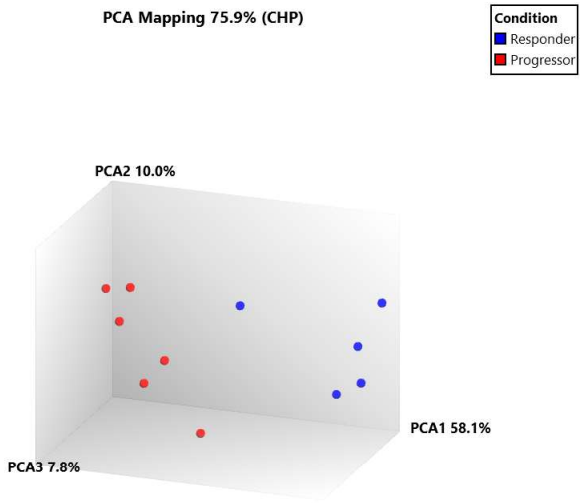



**Figure 5. PCA Plot**

Summary of our analysis with associated **volcano plot in Figure 6** is shown below:

**Progressor vs Responder**

- Progressor: 6 samples, Responder: 5 samples

Filter criteria:

- Fold Change: ≥ 2 or ≤ -2
- P-val: < 0.05

Total number of genes: 67528

- Genes passed filter criteria: 5247 (7.77%)
  - Up-Regulated: 1813 (34.55%)
  - Down-Regulated: 3434 (65.45%)

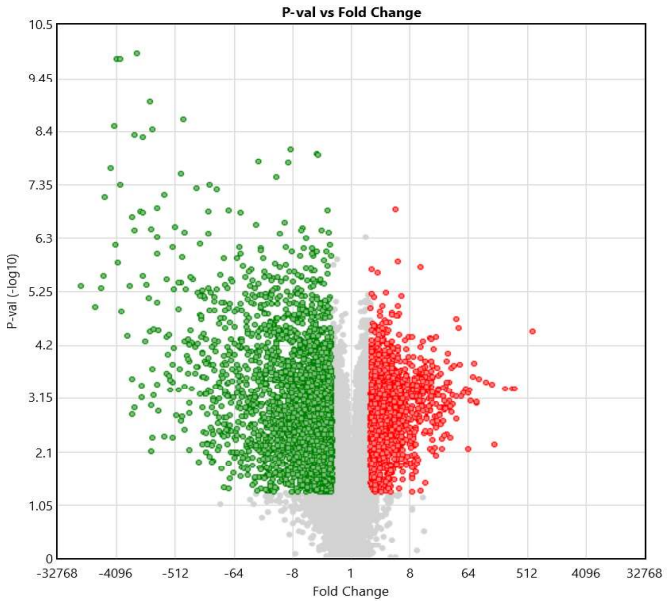



**Figure 6. Volcano Plot**

In addition to that we identified a key oncogene and 5 tumor suppressor genes which could potentially predict the outcome of immunotherapy in BlCa patients. Oncogene TCF3 is upregulated and Tumor suppressor genes TNFAIP3, MDM4, DDX5, CDH1 and ATM are downregulated in progressors compared to responders. TCF3 is involved in several chromosomal translocations that are associated with lymphoid malignancies. MDM4 gene plays a role in apoptosis. It also interacts with another key tumor suppressor gene TP53 which plays a pivotal role in preventing many different forms of cancer. Upregulation of the oncogene and downregulation of tumor suppressor genes in patients could imply potential inefficacy of immunotherapy treatment.

**Part III:** RNA-Seq data from GSE133624 is analysed using Strand NGS software. The data consists of 36 samples of tumors and 29 samples of normal tissues derived from BlCa patients. Average gene expressions of the tumor samples are compared with the non-tumorous samples. 2175 Genes were filtered out with more than 2 fold changes (upregulated or down regulated) and they were identified to be statistically significant with p-values $\leq 0.05$.

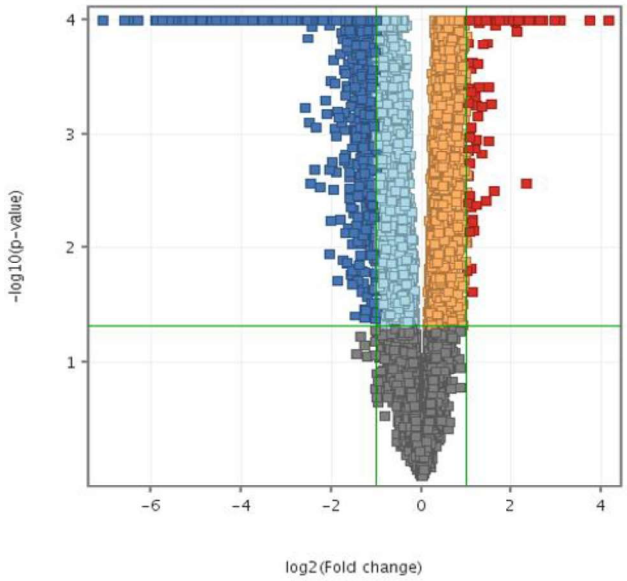



**Figure 7.** The **Volcano plot** of statistically significant differentially expressed genes between 36 samples of tumors and 29 samples of normal tissues derived from BlCa patients are shown in this figure. The x-axis shows the Fold Change and the y-axis shows the significance of differential expression of each gene, represented as the negative log of the p-value.

Pathway analysis performed by IPA software (QIAGEN Inc.), with loading 2175 probe sets (FC $\geq 2$, P-value $\leq 0.05$) that were differentially expressed with tumor, identified 290 Canonical Pathways that were significantly [-log(p-value) = 1.3] modified in BlCa tumor group. Top important pathways are depicted in **Table 2**. In this study group, the top Canonical Pathway identified is also Hepatic Fibrosis which is reported to be a common causal factor in many Bladder Cancer patients. Forty-five diseases and functions were modified by differential gene expression in the BlCa tumor patients. Important diseases and functions modified in the tumor group are listed in **Table 2**. Cancer is identified to be the top modified disease in this analysis. Top Cardiotoxicity altered with tumor included Cardiac Arrhythmia,

Arteriopathy, Infarction, Enlargement and Pulmonary Hypertension. Hepatotoxic functions namely Liver Hyperplasia/Hyperproliferation, Hepatocellular Carcinoma, Liver Cirrhosis, Fibrosis, and Inflammation/Hepatitis are altered with the occurrence of tumor. Altered Nephrotoxic functions are Kidney Failure, Renal Inflammation, Renal Nephritis, Nephrosis and Glomerular Injury. Ingenuity network analysis has identified networks related to Cancer (**Figure 8**). **Network 1** is related to Cell-To-Cell Signaling and Interaction, Cancer, Gastrointestinal Disease (Score 32, Focus Molecules 35). **Network 2** is related to Cancer, Connective Tissue Disorders, Dermatological Diseases and Conditions (Score 30, Focus Molecules 34).

**Table 2.**

| **Ingenuity Canonical Pathways:** | **-log (p-value)** |
|---|---|
| Hepatic Fibrosis / Hepatic Stellate Cell Activation | 21.5 |
| Cardiac Hypertrophy Signaling (Enhanced) | 21 |
| Hepatic Fibrosis Signaling Pathway | 19.4 |
| Axonal Guidance Signaling | 18.8 |
| Pulmonary Fibrosis Idiopathic Signaling Pathway | 15.9 |
| Role of Macrophages, Fibroblasts and Endothelial Cells in Rheumatoid Arthritis | 15.4 |
| Molecular Mechanisms of Cancer | 13.3 |
| Th1 and Th2 Activation Pathway | 13.1 |
| Leukocyte Extravasation Signaling | 12.5 |
| GP6 Signaling Pathway | 11.9 |
| Th1 Pathway | 11.5 |
| Tumor Microenvironment Pathway | 11.3 |
| Neuroinflammation Signaling Pathway | 11.2 |
| Atherosclerosis Signaling | 10.4 |
| Complement System | 9.92 |
| Colorectal Cancer Metastasis Signaling | 9.26 |
| Th2 Pathway | 9 |
| **Diseases and Functions:** | **p-value** |
| Cancer | 9.17E-97 - 5.5E-16 |
| Organismal Injury and Abnormalities | 9.17E-97 - 5.5E-16 |
| Reproductive System Disease | 6.13E-92 - 5.14E-16 |
| Cellular Movement | 1.1E-72 - 3E-16 |
| Gastrointestinal Disease | 2.94E-67-2.14E-16 |
| Neurological Disease | 7.95E-51-1.8E-16 |
| Cell Death and Survival | 4.5E-49-7.44E-17 |
| Hematological Disease | 9.21E-49-1.16E-16 |
| Immunological Disease | 9.21E-49-6.11E-17 |
| Inflammatory Disease | 1.64E-48-5.41E-20 |
| Skeletal and Muscular Disorders | 1.64E-48-3.02E-16 |
| Cardiovascular Disease | 9.17E-45-9.44E-17 |
| Inflammatory Response | 2.26E-43-3E-16 |
| Cell-To-Cell Signaling and Interaction | 2.54E-43-4.75E-17 |
| Cardiovascular System Development and Function | 4.08E-43-4.75E-17 |
| Immune Cell Trafficking | 7.73E-38-3E-16 |
| Respiratory Disease | 1.08E-37-1.79E-16 |

**A. Network 1: Cell-To-Cell Signaling and Interaction, Cancer, Gastrointestinal Disease (Score 32, Focus Molecules 35)**

**B. Network 4: Cancer, Connective Tissue Disorders, Dermatological Diseases and Conditions (Score 30, Focus Molecules 34)**

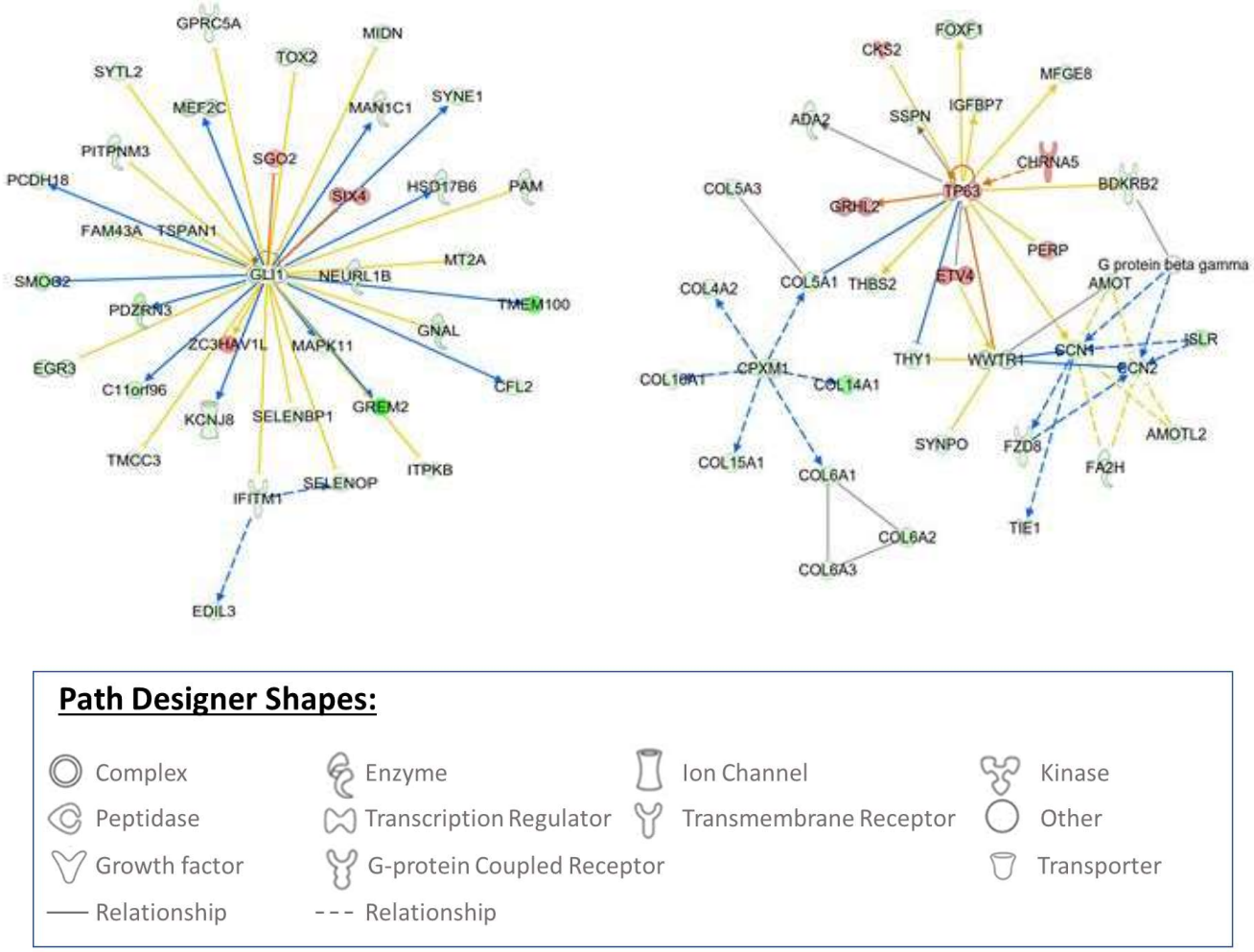



**Figure 8**. Top 2 networks picked up by pathway analysis are related to Cancer. These two networks are presented in this figure (panels A, B). Pathway analysis was performed using Ingenuity Pathway Analysis (IPA) software (QIAGEN Inc., https://digitalinsights.qiagen.com/IPA) by loading the 2175 probe sets that were differentially expressed in BlCa tumor group. The red filled path designer shapes are upregulated genes and green filled path designer shapes are downregulated genes.

**Conclusions:**

In this project we accessed and analysed many publicly available data sets that relates to tumorous and normal bladder cells derived from BlCa patients. In Part I we identified top canonical pathways activated in BlCa Tumor samples. In particular Hepatic Fibrosis/Hepatic Stellate Cell Activation and enhanced Cardiac Hypertrophy Signaling are significantly affected/altered in tumor samples. We also found that Top 2 networks picked up by pathway analysis are related to Cancer. In Part II we looked for a biomarker that could predict the efficacy of immunotherapy treatment using ICIs. We found out upregulation of an

oncogene and downregulation of some tumor suppressor genes can potentially predict whether immunotherapy might be helpful in treating BlCa patients where the disease is at an advanced stage. More work encompassing much larger and diverse data sets is needed to have a definite conclusion in this regard. Next-generation-sequencing (NGS) has revolutionized genomic research. **RNA-Seq** is a sequencing technique which uses NGS to reveal the presence and quantity of RNA in a biological sample at a given moment, analyzing the continuously changing cellular transcriptome. In Part III we used the RNA-Seq data. We found that most of the genes (1872 out of 2175) are downregulated whereas only 303 genes are upregulated. Using IPA software, we found same canonical pathways and same top disease and functions as in Part I. Interestingly, both Part I and Part III studies identified cancer related networks to be scoring either topmost or nearly top.

**Future work:**

Techniques learned and used in this project is very useful in analyzing huge set of biological data related to cancer and other diseases. We are collaborating with experimental groups to analyze their data for breast cancer and other pertinent diseases.

**Acknowledgements:**

This work is supported by faculty development grant from Lincoln University PA. Some of the computational work was done at Cancer Genomics Facility, Sidney Kimmel Cancer Center, Thomas Jefferson University, Philadelphia, PA. The author gratefully acknowledges the discussions with and assistance from Dr. S Addya, cancer genomics specialist working at Thomas Jefferson University, Philadelphia, PA.

**References**

1. Cancer Today (powered by GLOBOCAN 2018) IARC CancerBase No. 15; Ferlay, J.; Ervik, M.; Lam, F.; Colombet, M.; Mery, L.; Pi´neros, M.; Znaor, A.; Soerjomataram, I.; Bray, F. (Eds.) IARC: Lyon, France, 2018; ISBN 978-92-832-0453-4.

2. Surveillance, Epidemiology, and End Results Program. SEER stat fact sheets: kidney and renal pelvis cancer. Bethesda: National Cancer Institute; 2021. at https://seer.cancer.gov/statfacts/html/urinb.html

3. Wołącewicz M., Hrynkiewicz R., Grywalska E., Suchojad T., Leksowski T., Roliński J., Niedźwiedzka-Rystwej P. Immunotherapy in Bladder Cancer: Current Methods and Future Perspectives. Cancers (Basel). 2020 May 7; 12(5):1181.

4. Bellmunt J, Orsola A, Leow JJ, et al. Bladder cancer: ESMO practice guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2014; 25(Suppl 3):iii40–8.

5. Milowsky MI, Rumble RB, Booth CM, et al. Guideline on muscle-invasive and metastatic bladder cancer (European Association of Urology guideline): American Society of Clinical Oncology clinical practice guideline endorsement. J Clin Oncol. 2016; 34:1945–52.

6. Ghasemzadeh A, Bivalacqua TJ, Hahn NM, Drake CG. New strategies in bladder cancer: a second coming for immunotherapy. Clin Cancer Res. 2016; 22:793–801.

7. Audenet F., Attalla K., Sfakianos JP. The evolution of bladder cancer genomics: What have we learned and how can we use it? Urol Oncol. 2018 Jul; 36(7):313-320.

8. Maraver A, Fernandez-Marcos PJ, Cash TP, Mendez-Pertuz M et al. NOTCH pathway inactivation promotes bladder cancer progression. J Clin Invest 2015 Feb; 125(2):824-30.

9.https://www.thermofisher.com/us/en/home/life-science/microarray-analysis/microarray-analysis-instruments-software-services/microarray-analysis-software/affymetrix-transcriptome-analysis-console-software.html

10. Krämer A, Green J, Pollard J, Jr., Tugendreich S. Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinformatics* 2014 February 15, 30(4): 523–530.

11.Carretero-Gonzalez A, Moreno B H, Dueñas M, Martínez-Fernández M, Manneh R, Bernardini A, Rubio C, Villacampa F, Martín R M G, García Escudero R, Peralto J L R, Rueda D, Velasco G D, Castellano D E, Paramio J M. Potential markers of response and resistance to programmed cell death-1 blockade in first-line therapy of cisplatin-inilegible advanced urothelial cancer. Journal of Clinical Oncology 37, no. 7 suppl 2019: 449-449.

12. https://www.strand-ngs.com/

13. Chen X et al. 5-methylcytosine promotes pathogenesis of bladder cancer through stabilizing mRNAs Nature Cell Biology 2019 Aug; 21: 978–990.

14. https://software.broadinstitute.org/morpheus/