Bioinformatics Screening of Cell Surface Receptors to Identify New Stem Cell and Prostate Cancer Markers

Jenica Wang[1], Dr. Xue Gong[2]
Henry M. Gunn High School[1], Stanford University[2]

INTRODUCTION

Prostate cancer is the second most common cancer in men, occurring in 1 of every 7 men, although 1 of every 39 men with prostate cancer dies (Prensner, et al. 2012). Because of its alarming frequency, early detection for prostate cancer has been a focus area for many science researchers. One of the earliest biomarkers for prostate cancer is the prostate-specific antigen (PSA), which is a protein produced by prostate cells. By measuring the amount of PSA in the blood, researchers and doctors could determine a man's risk for prostate cancer. In 1986, the Food and Drug Administration (FDA) approved the PSA Test with the original intent to record prostate cancer progression in prostate cancer patients. Later in 1994, the FDA approved for the PSA Test to be used with a digital rectal exam (DRE) to identify potential prostate cancer patients among men who do not show signs of the disease. Since then, with the advent of the PSA Test, more men have been diagnosed for prostate cancer, about 50% of whom possess the indolent form and are not actually at risk (Prensner, et. al. 2012). Because of this, the PSA Test has led to over diagnosis and overtreatment, doing more harm than good (Mittakanti, et al. 2016).

Since stem cells are characterized by their self-renewal abilities, they play an important role in the initiation and progression of cancer, diseases defined by abnormal cell growth. Due to their longevity and self-renewal ability, many propose that the origin of cancer may come from these stem cells (Smith, et al. 2015). For example, Trop2 is a transmembrane glycoprotein known to be a stem cell marker, a regulator of stem cell self-renewal located on the cell surface. This was proven when Trop2 regenerated prostatic tubules in vivo (Goldstein, et al. 2008). Additionally, expression of Trop2 is found highly elevated in multiple cancers including prostate cancer (Trerotola, et al. 2013). Studies have also shown that tumor cell growth was greatly inhibited when Trop2 was knocked down. Currently, clinical trials are underway using anti-Trop2 antibody-drug conjugate (IMMU-132) for various tumor types, including prostate, breast, and pancreatic cancers (Cardillo, et al. 2015). With this successful example, this implies that by targeting stem cells and identifying stem cell markers, we can find new prostate cancer biomarkers and therapeutic targets.

METHODS

My research is considered applied because I used bioinformatics and other scientific knowledge to investigate and develop a solution to a practical problem: prostate cancer. The data analysis results that I received were quantitative because I found upregulated genes using fold change and false discovery rate (FDR), two concepts applied in statistics. The methods I employed were experimental since I analyzed RNA-Seq and microarray data. The general population I studied was men with prostate cancer; the sample included both published and unpublished sources. I selected my sample based on how much detail was provided about each patient.

Data Sources

1. Fluorescence-activated Cell Sorting (FACS)
   (1) Three sets of ten mouse prostates were used
   (2) Blood cells and prostate cells were distinguished using a marker, which is known to be high in blood cells
   (3) Prostate cells were categorized as luminal, basal, and stromal cells using two other markers, which have varying levels in the different types of prostate cells
   (4) Basal cells were separated into those with high Trop2 expression levels and those with low Trop2 expression levels
2. Microarray preparation
   (1) Platform: Affirmatory GeneChip Mouse Genome 430 2.0 Array
   (2) Conditions:
      1) All probes belonging to Affymetrix controls deleted
      2) Each probe corresponds to a gene
      3) Each probe represents a single gene
      4) All signals from 9 samples detected
      5) All probes with gene names changed in Excel/csv format to dates deleted

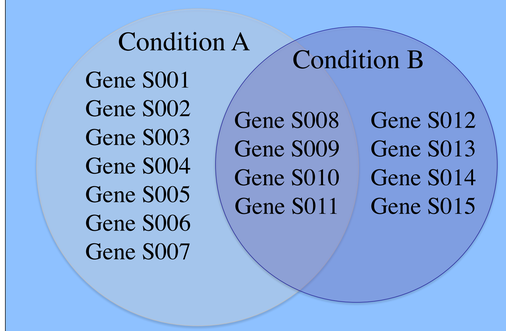
METHODS CONT.

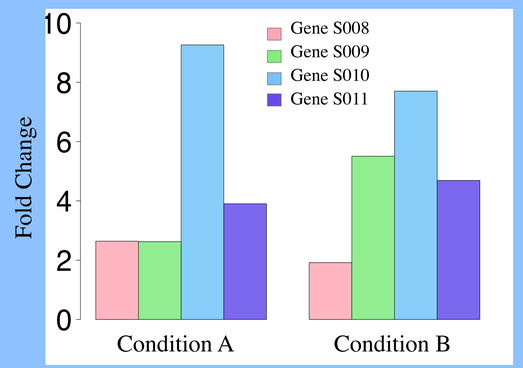
Bioinformatics Workflow

- **Data Download**
  - Microarray (Mouse)
  - RNA-Seq (Human)
- **Data Processing**
  - Conditions A and B
  - 3 Datasets
- **SAM Analysis**
  - Conduct dynamic searching/algorithm
  - Upregulated Genes
- **Gene Ontology Annotations**
  - Cell Surface Receptors

RESULTS

Stem Cell Candidates for Therapeutic Targets





Candidates for therapeutic targets chosen. (A) Venn diagram of upregulated genes that code for cell surface receptors in Condition A and Condition B. (B) Validation for top stem cell marker candidates. Fold change of stem cell markers in microarray data.

Prostate Cancer Data Summary

#ID	Sample Size	cBioPortal Name	Study Name (Title)	PubMed ID	GEO Accession No.
1	176	Prostate Adenocarcinoma (Nelson Lab, Fred Hutchinson CRC 2016)	Substantial interindividual and limited intraindividual genomic diversity among tumors from men with metastatic prostate cancer	26928463	GSE77930
2	61	Prostate Adenocarcinoma, Metastatic (Michigan, Nature 2012)	The mutational landscape of lethal castration-resistant prostate cancer	22722839	GSE35988
3	216	Prostate Adenocarcinoma (MSKCC, Cancer Cell 2010)	Integrative genomic profiling of human prostate cancer	20579941	GSE21032

Public datasets of prostate cancer obtained by RNA-Seq and microarray

REFERENCES

1. Substantial interindividual and limited intraindividual genomic diversity among tumors from men with metastatic prostate cancer. *Nature Medicine* 22, 369-378 (2016)
2. The Mutational Landscape of Lethal Castrate Resistant Prostate Cancer. *Nature*. 2012 July 12;487(7406):239-243. doi 10.1038/nature11125
3. Integrative genomic profiling of human prostate cancer. Cancer Cell. 2010 Jul 13;18(1):11-22
4. Madu CO and Lu Y. Novel diagnostic biomarkers for prostate cancer. Journal of Cancer. 2010; 1:150-177.

RESULTS CONT.

Upregulated Genes that Code for Cell Surface Receptors in RNA-Seq Data

Gene Name	Fold Change	FDR	Dataset
Gene PC001	3.49	0	2
Gene PC002	3.17	0	2
Gene PC003	2.93	0	2
Gene PC004	2.65	0	2
Gene PC005	2.63	2.82 E-05	2
Gene PC006	2.42	0	2
Gene PC007	2.35	0	2
Gene PC008	2.27	3.24 E-07	2
Gene PC009	2.25	0	2
Gene PC010	2.22	2.00 E-04	2
Gene PC011	2.20	0	2
Gene PC012	2.05	0	2
Gene PC013	2.05	4.89 E-05	2
Gene PC014	2.01	5.58 E-04	2
Gene PC015	1.90	7.35 E-02	1
Gene PC016	1.76	9.56 E-02	1
Gene PC017	1.61	3.93 E-02	3

Upregulated genes in metastatic prostate cancer compared to localized prostate cancer

Dataset #2	Dataset #3
Gene PC018	**Gene PC018**
Gene PC019	**Gene PC019**
Gene PC020	**Gene PC020**
Gene PC017	Gene PC023
Gene PC021	
Gene PC022	

Upregulated genes in localized prostate cancer compared to normal samples in Dataset #2 (Fold change > 2, FDR < 0.05) and Dataset #3 (Fold change > 1.5)

CONCLUSIONS

1. Bioinformatics screening identified several genes as top stem cell marker candidates from mouse microarray analysis.
2. Mining public RNA-Seq data reveals several upregulated genes in metastatic prostate cancer.
3. Many genes were discovered as upregulated genes of cell surface receptors in prostate cancer through RNA-Seq analysis.

FUTURE DIRECTIONS

1. Immunohistochemistry (IHC) can be done on the newly discovered stem cell markers.
2. We will test the functional role of stem cell self-renewal in prostate tumorigenesis.
3. We will define the molecular mechanisms through which these genes regulate prostate tumorigenesis.

ACKNOWLEDGEMENTS

I would like to thank the AAR program and Dr. Choe, Ms. Merchant, and Ms. Durquet for providing me this research opportunity. I would also like to thank my mentor Dr. Xue Gong of Stanford University for the year-long guidance and support.