SOX4 and SMARCA4 as Cell Cycle Regulators of Triple Negative Breast Cancer

Rushabh Mehta[1], Gaurav Mehta[2], Michael L. Gatza[2]

[1]Department of Genetics, Rutgers University-New Brunswick;

[2]Department of Radiation Oncology, Rutgers Cancer Institute of New Jersey

Introduction

Triple negative breast cancer (TNBC) remains a leading cause of cancer deaths amongst women in the United States due to a distinct lack of targeted therapies as well as the aggressive behavior of the cancer[1,2]. BRG1, the ATPase subunit of the SWI/SNF human chromatin remodeling complex, is overexpressed in approximately 90% of TNBC tumors and has been shown to contribute to tumor development and progression; however, the mechanisms by which BRG1 mediates these processes and the required cofactors remain unknown[3]. Recent work by the Gatza lab has determined that BRG1 interacts with the oncogenic transcription factor SOX4 and that this complex is responsible for regulating PI3K signaling in this subset of tumors[4,5]. To better understand the downstream effects of BRG1-SOX4 signaling in TNBC, the large-scale tumor datasets produced by The Cancer Genome Atlas (TCGA) Program as well as the Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) study were interrogated. Basal-like breast tumors, a molecularly identified subtype of breast cancer that is largely synonymous with TNBC[1], will be investigated using these datasets. Overall, the *in silico* analysis reveals distinct signaling pathways active in tumors with gene expression differentially correlated with SOX4 and SMARCA4. Genomic as well as proteomic analysis provide the foundation for the assertion that SOX4 and SMARCA4 cooperate to promote cell proliferation through the downstream transcriptional control of cell cycle genes.

Research Objective

The goal of this project is to determine the mechanistic roles of SMARCA4 and SOX4 in promoting tumor growth and proliferation. Previous studies by the Gatza Lab and others have demonstrated the role of these genes as oncogenic in triple negative breast cancer, yet the mechanisms by which they control tumorigenesis are unknown. As both genes modulate the PI3K pathway, which has been shown to be highly druggable, understanding mechanisms of proliferation control create the potential to inform novel regulatory targets for therapy.

Genes positively and negatively correlated with SOX4/SMARCA4 expression show distinct signaling pathways

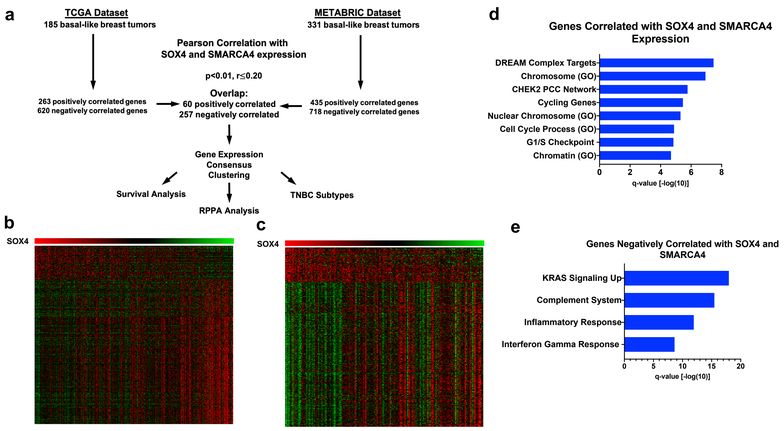

Figure 1. (a) Workflow used to identify genes positively and negatively correlated with SOX4 and SMARCA4 expression in basal-like tumors. 60 positively correlated and 257 negatively correlated genes shown in a heatmap from the **(b)** TCGA and **(c)** METABRIC datasets. Tumors are ordered according to SOX4 expression. Red indicated high expression of genes and green indicates low expression. GSEA results from **(d)** positively correlated genes show several cell cycle-related gene sets and **(e)** negatively correlated genes show immune-related gene sets.

Consensus clustering reveals 3 classes of TNBC tumors

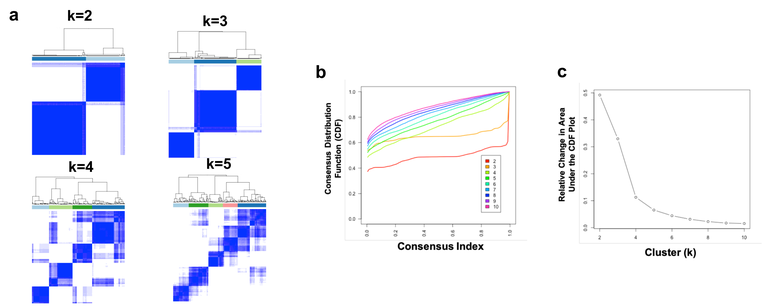

Figure 2. Identification of correlated gene expression-based classes within TNBC tumors. **(a)** Consensus matrices from k=2 to k=5 show the cluster confidence for each sample. K-means clustering was performed 1000 times with 80% item resampling and 100% gene resampling. **(b)** The CDF plot shows the cumulative distribution from consensus matrices at different k values. **(c)**The relative change in area under the CDF plot informs the ideal number of k clusters. K=4 is an optimal number according to this plot as the area does not change drastically following this cluster number. However, the consensus matrix appears to have less cluster confidence at this number, therefore k=3 was chosen to be the number of classes based on expression.

Consensus clusters show differential enrichment for TNBC subtypes and overall survival

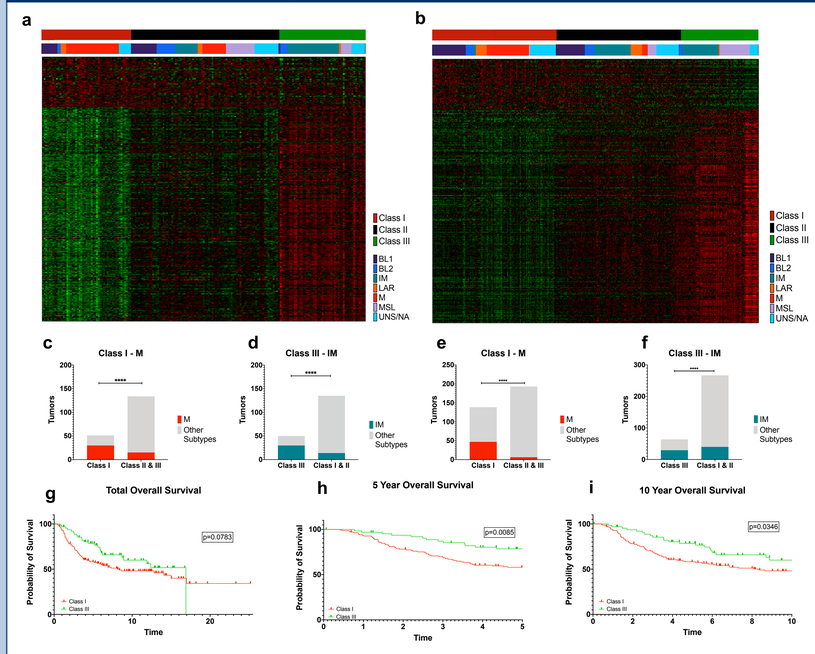

Figure 3. (a) TCGA and **(b)** METABRIC k-means clustering. Tumor classes, defined by consensus clustering, are shown. Class I tumors have high SOX4/SMARCA4 correlated gene expression, class II has mixed expression, and class III has high SOX4/SMARCA negatively correlated gene expression. TNBCtype[6] was used to identify the TNBC subtypes of the tumors. Fisher's exact test shows **(c)** significant enrichment of mesenchymal subtype in class I tumors and **(d)** enrichment of immunomodulatory subtype in class III tumors. In the METABRIC clusters, similar enrichment is observed **(e, f)**. Overall survival rates differ between class I and class III tumors. Median overall survival for class I and class III tumors is 8.30 years and 16.90 years, respectively. Difference in total overall survival **(g)** is not significant using Logrank test, but significant for **(h)** 5-year and **(i)** 10-year overall survival.

Cell cycle-related genes are overexpressed in tumors with SOX4/SMARCA4 correlative gene expression

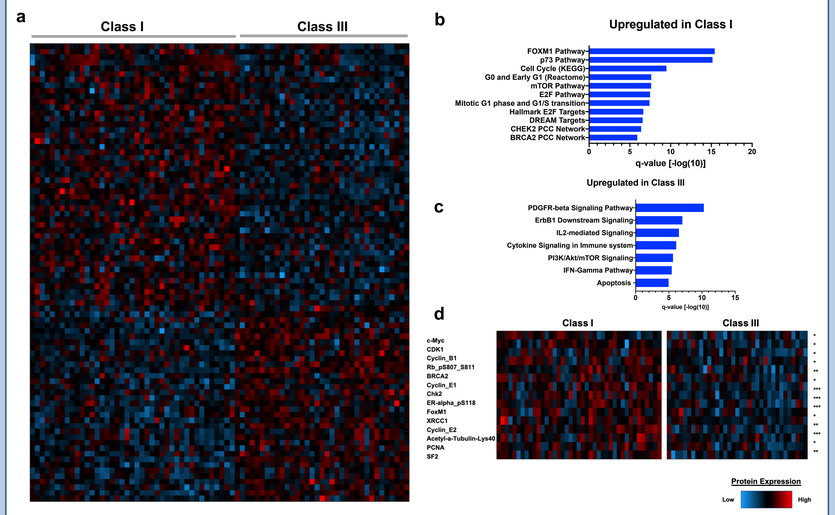

Figure 4 Significant protein level changes reflect genetic expression in TNBC tumors. **(a)** 83 proteins from TCGA RPPA dataset differentially expressed identified by ANOVA (p<0.05). 47 proteins overexpressed in class I tumors and 35 overexpressed in class 3 tumors. **(b)** Overexpressed proteins in class I tumors show enrichment of cell cycle-related gene sets and **(c)** overexpressed proteins in class III tumors show enrichment in immune signaling-related gene sets. **(d)** Across gene sets, 14 important cell cycle regulators are overexpressed in class I tumors. Significance levels are as follows: <0.05 (*), <0.01 (**), and <0.001 (***).

SOX4 knockdown leads to downregulation of cell cycle genes

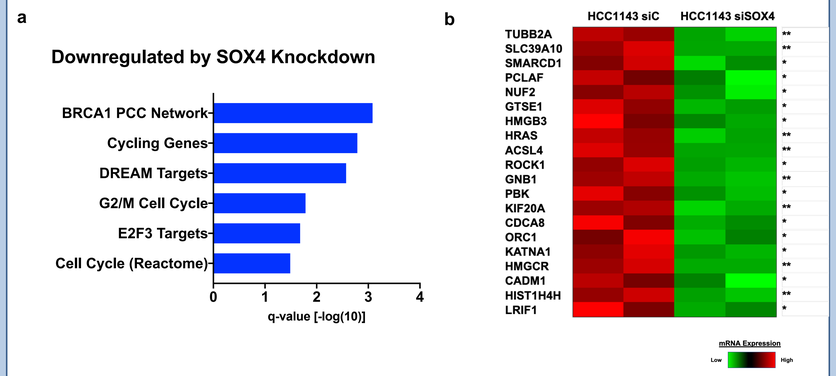

Figure 5. RNA-Seq of siRNA-mediated SOX4 knockdown in HCC1143 cell line. **(a)** 165 downregulated genes show over-enrichment for cell cycle-related gene sets. **(b)** Across gene sets, 20 genes related to cell cycle are differentially expressed, including *ORC1* and *CDCA8*.

Conclusion

SOX4 and SMARCA4 are canonical regulators of cell proliferation and correlated with poor prognosis in many cancers. Specifically in triple negative breast cancer, SOX4 and SMARCA4 have been shown to promote tumor growth and proliferation. This study provides preliminary evidence as to the role of these genes in tumor proliferation through the regulation of cell cycle genes. Genes that are positively correlated with SOX4 and SMARCA4 expression show an enrichment for cell cycle related genes, which is reflected in proteomic datasets as well. Clustering positively and negatively expressed genes segregates tumors into three classes, with class I and class III tumors having the most differential expression of these genes between them. The enrichment of the mesenchymal triple negative subtype as well as the poor prognosis of patients within this subtype provide context for the enriched cell signaling gene sets. Overall, this project demonstrates the role of SOX4 and SMARCA4 in driving tumor cell proliferation through the downstream transcriptional regulation of various cell cycle genes.

Future Direction

Interrogation of larger proteomic and phospho-proteomic datasets would provide more insight about the active pathways in TNBC with SOX4 and SMARCA4 overexpression past the genomic level. Cell cycle assays would help inform the specific stages during which the two oncogenic regulators are active. Finally, coimmunoprecipitation of the SOX4-SMARCA4 complex suggested by the Gatza Lab would allow for the discovery of other cofactors involved in cell proliferation control.

Acknowledgements

This summer research project was supported by the generous contribution of the Human Genetics Institute of New Jersey and the MacMillan Foundation. I would further like to thank Dr. Michael Gatza and Dr. Gaurav Mehta for their continued support and mentoring through this project.

References

1. Cancer Genome Atlas Network. (2012). Comprehensive molecular portraits of human breast tumours. *Nature*, *490*(7418), 61.
2. Miller, K. D., Nogueira, L., Mariotto, A. B., Rowland, J. H., Yabroff, K. R., Alfano, C. M., Jemal, A., Kramer, J. L., & Siegel, R. L. (2019). Cancer treatment and survivorship statistics, 2019. *CA: a cancer journal for clinicians*, *69*(5), 363–385.
3. Wu, Q., Madany, P., Akech, J., Dobson, J. R., Douthwright, S., Browne, G., Colby, J. L., Winter, G. E., Bradner, J. E., Pratap, J., Sluder, G., Bhargava, R., Chiosea, S. I., van Wijnen, A. J., Stein, J. L., Stein, G. S., Lian, J. B., Nickerson, J. A., & Imbalzano, A. N. (2015). The SWI/SNF ATPases Are Required for Triple Negative Breast Cancer Cell Proliferation. *Journal of cellular physiology*, *230*(11), 2683–2694.
4. Gatza, M. L., Angus, S. P., Khella, C. A., Tong, K., Verzi, M., & Mehta, G. A. (2019). Abstract P2-03-03: BRG1-SOX4 mediates a novel and essential signaling network that activates PI3K/Akt signaling in TNBC.
5. Mehta, G. A., Parker, J. S., Silva, G. O., Hoadley, K. A., Perou, C. M., & Gatza, M. L. (2017). Amplification of SOX4 promotes PI3K/Akt signaling in human breast cancer. *Breast cancer research and treatment*, *162*(3), 439-450.
6. Chen, X., Li, J., Gray, W. H., Lehmann, B. D., Bauer, J. A., Shyr, Y., & Pietenpol, J. A. (2012). TNBCtype: A Subtyping Tool for Triple-Negative Breast Cancer. *Cancer Informatics.*