Washington University in St. Louis
Washington University Open Scholarship

Arts & Sciences Electronic Theses and Dissertations

Arts & Sciences

Spring 5-15-2023

Identification and Characterization of Targets of Metastasis in High-Grade Serous Ovarian Cancer

Emilee Nicole Kotnik
Washington University in St. Louis

Follow this and additional works at: https://openscholarship.wustl.edu/art_sci_etds

Recommended Citation

Kotnik, Emilee Nicole, "Identification and Characterization of Targets of Metastasis in High-Grade Serous Ovarian Cancer" (2023). *Arts & Sciences Electronic Theses and Dissertations*. 2933.
https://openscholarship.wustl.edu/art_sci_etds/2933

This Dissertation is brought to you for free and open access by the Arts & Sciences at Washington University Open Scholarship. It has been accepted for inclusion in Arts & Sciences Electronic Theses and Dissertations by an authorized administrator of Washington University Open Scholarship. For more information, please contact digital@wumail.wustl.edu.

WASHINGTON UNIVERSITY IN ST. LOUIS

Division of Biology and Biomedical Sciences
Molecular Genetics and Genomics

Dissertation Examination Committee:
Katherine C. Fuh, Chair
Jason Weber, Co-Chair
Graham Colditz
John Edwards
Christopher Maher

Identification and Characterization of Targets of Metastasis in High-Grade Serous Ovarian Cancer

by

Emilee Nicole Kotnik

A dissertation presented to
Washington University in St. Louis
in partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

May 2023
St. Louis, Missouri

© 2023, Emilee Nicole Kotnik

Table of Contents

List of Figures

List of Tables

Acknowledgments

First, I would like to thank my mentor, Katherine Fuh. It has been an honor to be your first graduate student. I'd like to thank my lab mentors, Hollie Noia, Daniel Wilke, and Elena Lomonosova for their teachings, support, and for answering my daily questions. This degree would not have been possible without the support and encouragement from my fellow graduate students labmates, Alyssa Oplt, Angela Schab, Favour Akinjiyan, Isabelle Seppa, and Judith Sokei. A special thanks goes to my undergraduate mentee, Shirley Lopez De Leon. It was a pleasure to work with a brilliant, budding scientist and I thank you for your assistance on these projects. I must acknowledge the Center for Reproductive Health Sciences leadership and administrative assistants, Sarah England, Dineo Khabele, Magdalena Nelson, Kirsten Biggs, and Audrey Johnston. Their guidance, help, and encouragement were instrumental to my thesis work and they all foster an excellent environment for female scientists. Additionally, I want to recognize the Ovarian Cancer Research Innovation Fund that the department provided that helped with my funding in the last couple years of my program.

I'd like to thank the student groups that I worked within: ProSPER, GALNACS/SACNAS, NSPN, IMSD, and YSP. Being a part of these groups and their leadership teams taught me so many valuable skills that I will benefit from for the rest of my career. Acknowledgement goes to the Cell-to-Society Pathway, led by Graham Colditz and Susan Dutcher and granted by the Burroughs Wellcome Fund, that provided the funding of my student stipend for several years. I was very fortunate to join the Molecular Genetics and Genomics program under program directors, Jim Skeath and John Edwards. In particular, I would also like to thank John Edwards,

my thesis committee chair for his support and advice throughout these projects. A very special thanks and acknowledgement goes to Jason Weber and his lab members for welcoming me into their lab for the final months of my thesis work.

Lastly, I need to thank Mom, Dad, Donny, Caitlin, Patrick, Becky, my family, and friends for loving and supporting me throughout this PhD program. My friends Megan Richters, Kiona Elliott, and Ryan Friedman provided feedback, help, advice, and much love during the preparation of this dissertation. I am very lucky to have such amazing people in my life.

Emilee Nicole Kotnik

Washington University in St. Louis

May 2023

Dedicated to my nieces,

Josephine, Caroline, and Anna

ABSTRACT OF THE DISSERTATION

Identification and Characterization of Targets of Metastasis in High-Grade Serous Ovarian Cancer

by

Emilee Nicole Kotnik

Doctor of Philosophy in Biology and Biomedical Sciences

Molecular Genetics and Genomics

Washington University in St. Louis, 2023

Dr. Katherine C. Fuh. Chair

Dr. Jason Weber, Co-Chair

High-Grade Serous Ovarian Cancer (HGSC) is a highly metastatic cancer with the majority of patients presenting in advanced stages. Currently there are limited targeted treatment options for patients, especially for women with high metastatic tumor burdens. In order to improve patient outcomes, we have aimed to identify and characterize novel targets of metastasis utilizing sequencing from patient tumors and a functional genomic screen.

First, we identified genetic alterations of metastasis and short survival by characterizing the genomic and transcriptomic alterations of primary and metastatic tumors in HGSC patients from 23 short-term survivors (overall survival (OS) <3.5 years) and 16 long-term survivors (OS >5 years). We compared somatic mutations, copy number alternations, mutational burdens, differential expression, immune cell fractions, and gene fusion predictions between the primary and metastatic tumors of the ST and LT survival groups. From this project we identified a TP53 R273H mutation, which may have a gain-of-function in ovarian cancer, suggested from evidence in triple negative breast cancer. We aimed to determine if this mutation can be targetable with

PARP inhibitor combination treatments and if this mutation is gain-of-function in ovarian cancer cell lines.

Third, we performed a siRNA functional genomic screen on 719 kinases in an attachment assay utilizing a collagen fibronectin matrix plated with primary ovarian fibroblasts and GFP-labeled ovarian cancer cell lines. From this initial screen, we have identified 4 candidate kinases to validate. Candidate kinases were validated by siRNA knock-down in ovarian cancer cell lines and assessed on their ability to migrate in a wound-healing assay.

Chapter 1: Introduction

1.1 Ovarian Cancer Overview

As of 2019, there were approximately 233,565 women in the United States living with ovarian cancer; less than half of those women will survive more than 5 years after their diagnosis.[1] Although ovarian cancer is less prevalent than other cancer types, like breast cancer, patients diagnosed with ovarian cancer are more likely to die of their disease. [2] An estimated 12,810 patients died from ovarian cancer in 2022. [1] High-Grade Serous Ovarian Cancer (HGSC) is the most lethal ovarian cancer histotype, the second most common gynecologic malignancy, and the most deadly. [3, 4] Epithelial ovarian tumors are categorized into five classes: serous, endometrioid, mucinous, clear cell, and undifferentiated. [5, 6] Serous tumors are the most common and can be broken down further into low-grade and high-grade tumors, depending on the appearance of their nuclei, mitosis, and molecular abnormalities. [7]

1.2 Current HGSC Treatment and Patient Survival

The first line of treatment for patients with HGSC is chemotherapy and an initial debulking surgery, where surgeons remove as much visible tumor as possible. Debulking surgery is usually then followed by chemotherapy. [8] In advanced cases, neoadjuvant chemotherapy is sometimes used to reduce tumor size prior to debulking surgery. [9-11] The typical treatment regimen is platinum chemotherapy with a taxane, usually these are carboplatin and paclitaxel, which are either given through an IV or by injecting directly into the peritoneal cavity. [8] Another type of treatment currently being used are PARP inhibitors (PARPi), which are inhibitors that target PARP proteins involved in detection and repair of single-strand DNA breaks. PARP inhibitors work effectively in patients that have mutations in the homologous recombination (HR) repair pathway genes, such as *BRCA1* or *BRCA2*, because the

combination of a defective HR pathway and inhibiting single-strand break repair will lead to an accumulation of DNA damage that cannot be repaired and will result in cell death. [12-15] Clinical trials on the PARP inhibitors Olaparib and Niraparib have shown that combination therapy with these drugs can improve progression free survival in ovarian cancer patients with germline and somatic *BRCA1/2* mutations or tumors with homologous recombination deficiency (HRD). [12, 13] Therefore, PARP inhibitors are beneficial for patients with HRD defects, but that only accounts for up to half of patients with HGSC [12, 16]

Over the past few decades, chemotherapy and combination treatments have been the standard of care, and the overall patient survival has not significantly improved. [17] This problem is partially due to recurrent chemoresistant tumors, which contributes the most to ovarian cancer deaths. [14] Patients are more likely to survive longer than 5 years after diagnosis if their cancer is caught early, however about 80% of cases present at late stages of disease, when the tumor has already metastasized. [1, 18] Women who are diagnosed at more advanced stages have only a 29.2% chance of surviving longer than 5 years, hence the need for a better understanding of ovarian cancer metastasis and for genetic targets as biomarkers or for drug development in more aggressively metastatic tumors. [1]

1.3 Ovarian Cancer Metastasis and Tumor Microenvironment

HGSC tumors most commonly metastasize to other areas in the intrabdominal cavity, also known as the peritoneal cavity, such as the local reproductive organs, the sigmoid colon, and the omentum. [2] The omentum is a fat pad that drapes over the bowel and abdominal cavity and is particularly prone to tumor metastasis, where the tumor burden can lead to obstruction of the stomach and bowels. [2, 6] The ovarian cancer tumor microenvironment in the peritoneum is unique because its designed to surround and protect the abdominal organs, but is also composed of surface epithelium that is exposed within the cavity. [2, 6, 8] In the peritoneal cavity, the surface epithelium and serosal membrane, made up of

mesothelium, sits on top of a basement membrane, which is composed of collagen types I and IV, laminin, and fibronectin. [2, 6] Metastasis in HGSC happens more readily compared to other cancers like colon or breast cancer and ovarian tumors usually only invade the surface mesothelial cell layers of the peritoneum. [2]

In the process of metastasis, tumor cells must undergo epithelial-to-mesenchymal transition (EMT) to leave the primary tumor site. When cells begin EMT, they lose E-cadherin-mediated cell-cell interactions and then interact with the surrounding stromal cells to induce mesenchymal signaling to enact a cascade of events that lead to the cleavage of the E-cadherin ectodomain. [2] This cleavage event frees the cells to shed as single cells and spheroids into ascites, the surrounding fluid, where the cancer cells can easily follow the movement of the peritoneal fluid and come into contact with the peritoneum or omentum. [2, 3] It is still unknown if single cells detach from the primary site and clump with other single cells in the ascites before attaching to a secondary site, or if spheroids shed as a clump from the primary site to attach elsewhere to seed other tumors. Once split from the primary tumor, a tumor cell will attach to the mesothelial cell layer of peritoneal cavity. [2] Although these earlier stages of ovarian carcinoma metastasis have been well studied, less is known about the tumor cell behavior after they have implanted onto the mesothelium and invade into the stroma.

1.4 Genetic and Genomic Landscape of HGSC Tumors

In 2011, The Cancer Genome Atlas (TCGA) Research Network published one of the first large-scale genomic characterizations of ovarian cancer, analyzing microarray and exome sequencing data from over 300 HGSC patient primary tumors. [19] They originally observed that 96% of their tumor cohort harbored *TP53* mutations, but a rereview acknowledged that the wildtype *TP53* cases were not typical of HGSC and concluded that 100% of HGSC cases have *TP53* mutations. [19, 20]

TP53 is one of the most mutated tumor suppressor genes in cancer and in general lead to tumors with an aggressive phenotype that are more invasive, metastatic, proliferative, and poorly differentiated. [21-23] When mice are deficient in their version of this gene, *Trp53*, they develop spontaneous tumors. [24] Additionally, human germline mutations in *TP53* leads to Li-Fraumeni syndrome where people have a high-risk of early-onset cancers. [25] A study that examined *TP53* gain-of-function mutations in ovarian cancer saw correlations with more distant metastases and platinum treatment resistance. [26] However, it is still unclear if specific *TP53* mutations are therapeutically targetable in HGSC. [23, 27, 28] *TP53* mutations are thought to arise early in HGSC development and it is debated if these mutations are a consequence or cause of high genomic instability in HGSC tumors. [19, 23]

TCGA observed that HGSC tumors have high copy number alterations and a low prevalence of other recurrently mutated genes. [16] This observation is consistent with another study that compared oncogenic signatures among different cancer types in TCGA, which found that ovarian cancer had one of the most densely packed copy number alterations signatures, but a more moderate mutational burden. [29, 30] Other significantly mutated genes in HGSC include *BRCA1, BRCA2, CSMD3, NF1, CDK12, FAT3, GABRA6,* and *RB1*. [19] Overall, about 20% of ovarian cancer tumors have a germline or somatic mutations in *BRCA1* or *BRCA2*, and about 11% of patients show DNA methylation of *BRCA1*. [16, 17, 31] Additionally, TCGA reported that about half of their cohort tumors had HRD. [16] Similar to the use of PARP inhibitors for patients with *BRCA1/2* mutations, there are targeted therapies for both *NF1* and *RB1* mutations that have been used for other cancer types that frequently have mutations in these genes. [32]

TCGA's microarray clustering analysis found four expression subtypes, which they grouped as immunoreactive, differentiated, proliferative, and mesenchymal based on the gene contents of each cluster. [16] Prior to TCGA, other studies used GeneChip microarrays from normal ovary, primary and metastatic tumors of HGSC patients, with the goal of identifying potential biomarkers with gene

expression data. Both studies found a small number of differentially expressed genes between the primary and metastatic tumors, which they determined was evidence for high clonality between HGSC primary and metastatic tumors. [33, 34]

Recent studies have attempted to identify structural genomic signatures and patterns in HGSC tumors. A *Nature Genetics* article, Macintrye et al., examined copy number signatures in ovarian cancer by analyzing shallow WGS of primary and relapsed samples from 132 patients for absolute copy number profiles. [17] They were able to identify seven copy number signatures and proposed mechanisms that caused each signature. Good survival correlated with signatures that had higher enrichment for HRD, both with and without *BRCA1/2*, while poor survival correlated with signatures which had higher mutations in *NF1*, the RAS signaling pathway, and CDK12 mutations with tandem duplications. [17] Similarly, another study examined WGS data from 80 primary HGSC tumors and matched normal samples; the authors found one genomic rearrangement signature that correlated with poor prognosis. This signature was characterized by medium to large deletions and tandem duplications. Additionally, tumors with a higher contribution of this signature demonstrated poor overall survival, which was confirmed in a TCGA validation set. [35] Interestingly, other studies have described ovarian, breast, and endometrial cancers as having a tandem duplicator phenotype (TDP), meaning that they detected tandem duplication configurations throughout the genomes of these gynecological tumors. [36-39] Further classifications of the TDP showed that the ovarian cancer cohort had the most enrichment for the TDP and more than 20% of the ovarian cancer tumors analyzed had tandem duplications that spanned about 11 kb in size. They also showed this span size correlated with *TP53* and *BRCA1* mutations, and that abrogating *TP53* and *BRCA1* in mice could drive the development of tumors with a similar TDP observed in human tumors. [39] These studies demonstrate the distinct structural changes that characterize ovarian cancer genomes, but they fail to examine the genomes of HGSC metastatic tumors.

In October of 2018, Yang et al. characterized clinical and genomic features that correlated with good and poor survival using whole-exome sequencing (WES) and RNA-seq data. [32] Their cohort of HGSC patients consisted of primary tumors from 20 long-term survivors that lived greater than 10 years after diagnosis and 21 short-term survivors that lived between 6 months to 2 years after diagnosis. They found clinical characteristics such as younger age at diagnosis, low CA125 levels in the blood after chemotherapy, and no residual disease post debulking surgery correlated with long-term survival. Long-term survivor tumors had a higher somatic mutational burden, loss of heterozygosity, and biallelic inactivation of *BRCA1* and *BRCA2*. They were also enriched for activated CD4+, CD8+ T cells, and effector memory CD4+ T cells. Short-term survivor tumors had more focal copy number gains of *CCNE1* and had lower HRD scores. [32] While this study does provide evidence that there are clinical and genetic characteristics of patients with different survival outcomes, it does not characterize metastatic tumors from HGSC patients and only focuses on the genetic characteristics from WES. An integrative study that examined the genomes of over 500 metastatic tumors from several major cancer types, but not including ovarian cancer, determined that profiling metastatic tumors can provide clinically relevant and dimensional views of the molecular landscape of metastatic cancer, thus stressing the need to profile more genomes from metastatic cancers. [34]

The most recent study in HGSC patient survivorship was completed by the Bowtell Lab, where they examined the whole-genome sequencing, RNA-sequencing, and methylome profiling from primary tumors of 60 patients that lived more than 10 years compared to 66 patients with short or moderate survival. They found that multiple alterations in HDR and DNA repair related genes correlated with long survivorship. They also saw that markers of proliferation correlated with long survivorship, possibly because proliferation sensitizes the cells to chemotherapy treatment. The tumors they analyzed could be stratified by survival into genomic and immune cell signatures, thus providing evidence that survivorship does have distinct characteristics. [40]

1.5 Gene Fusions in Cancer

The Yang et al. study analyzed gene fusion predictions from the RNA-sequencing transcripts in their patient cohort. They identified 4 recurrent gene fusions, one of which involved the genes *ESR1* and *CCDC170* that was recurrent in two short-term survivors. [32] The ESR1>CCDC170 gene fusion has also been characterized in aggressive estrogen-receptor positive breast cancers, along with other *ESR1* gene fusions that have been shown to drive endocrine therapy resistance, EMT, and metastasis. [33, 41] Lei et al. demonstrated that CDK4/6 inhibitors were able to suppress growth that was driven by *ESR1* gene fusions, indicating that gene fusion driven cancers are treatable. [41] Perhaps the most well-known example of treatable tumor gene fusions arose from the discovery of the Philadelphia Chromosome, a gene fusion between the genes *BCR* and *ABL* in chronic myeloid leukemias, whose discovery lead to the development of ABL inhibitors. [33] Additionally, the identification of EML4>ALK gene fusions in non-small cell lung cancer paved the way for the development of ALK inhibitors. Finally, drugs targeting tumors of any cancer type with gene fusions involving *NTRK* genes have recently been approved by the FDA. [33, 42] Studies have characterized gene fusions across cancer types, such as Gao et al. which analyzed over 9600 tumor samples from TCGA with multiple gene fusion tools. They identified over 1000 fusions involving kinases with intact kinase domains and they suggested that gene fusions can drive the development of 16.5% cancer cases. They also found that many gene fusions involve genes that are already druggable and suggest that fusions could expand immunogenic peptide predictions and provide more targets for drugs and immunotherapies. [43] Gene fusions such as BCAM-AKT2 and CDKN2D>WDFY2 have been implicated in ovarian cancer previously, suggesting that gene fusion predictions might be targetable biomarkers for HGSC tumors. [44, 45]

1.6 Role of Kinases in Ovarian Cancer

Kinases are a class of proteins that activate signal transduction, usually through phosphorylation, and induce many biological processes like proliferation, migration, mitosis, and apoptosis. [46] They have the potential to be therapeutically targetable because we can develop inhibitors that bind to their receptors and mimic their ligands to influence their functionality. Several kinases, such as tyrosine kinases and Aurora kinases, have previously been implicated in ovarian cancer and seem to influence tumor cell metastasis. [47, 48] For example, Aurora kinases are serin-threonine kinases that regulate the formation of the mitotic spindle and are highly expressed in ovarian cancers and other cancer types. Clinical trials on Aurora kinase inhibitors in combination with other treatments have shown to improve progression free survival. [49] The Fuh Lab has studied DDR2 and AXL and their role in metastasis and the use of drug inhibitors to target these kinases. [50-52] Other groups have shown that reducing the expression of DDR2 in mouse cells can increase the production of CD8+ T cells and has been shown to increase sensitivity to PDL-1 inhibitors. [53] There are also several small molecule inhibitors that target DDR1 and DDR2 that are in development, again showing that these kinases are targetable. [54] AXL is another tyrosine kinase has been shown to be highly expressed in ovarian cancer and particularly in metastatic tumors. Genetic inhibition of AXL *in vivo* prevented the development of further metastatic tumors and decreased invasion. [51] Therefore, there is evidence that AXL inhibition and therapy can prevent metastatic tumor progression.

1.7 Ovarian Cancer Immune Cell Landscape

The immune cells in the tumor microenvironment play an important role in how HGSC develops, metastases, and becomes chemoresistant. Many groups have characterized the immune cell landscapes of HGSC with a variety of methods including immunohistochemistry, RNA sequencing, and single-cell sequencing of patient tumors, blood, and ascites. This work has established that lymphocytes,

macrophages, and neutrophils play a large role in the cell-to-cell interactions between cancer cells and the tumor microenvironment. In general, these immune cells form a network with cytokines and chemokines that surrounds tumor cells to create local immunosuppression. As a consequence, those cancer cells survive and become metastatic. [55]

Recent ovarian cancer research has drawn attention to tumor infiltrating lymphocytes (TILs). [56, 57] TILs consist of white blood cells, like T-cells, B-cells, macrophages, and natural killer cells, that have localized into the tumor stroma. [57] Meta analyses and primary studies have found that TILs, especially CD8+ TILs, are a prognostic factor for survival in ovarian cancer. One meta-analysis on TCGA and the Gene Expression Omnibus (GEO) database of ovarian tumor RNA-seq found that patients with the best outcomes had a tumor microenvironment high in macrophage M1 and T cell follicular helper cells. They determined that the follicular helper T cells, mast cell activation, and M1 and M2 macrophages influence the patient's immunotherapy. [58] Meanwhile, regulatory T-Cells (Tregs) have been associated with poor outcomes and more advanced stages and grades of disease [59-61], possibly because their presence in the tumor microenvironment decreases the immune system's ability to attack cancer cells. [57, 62]

Tumor-associated macrophages (TAMs) are the major immune cell population for primary ovarian tumors and ascites. [55] These TAMs mostly take on an M2-like phenotype that support the HGSC tumor growth, metastasis, and resistance to chemotherapy. They encourage tumor growth at several stages of cancer development including immune escape of tumor cells, cell invasion, migration, and angiogenesis. The M2-like TAMs also correlate with poor prognoses in patient survival. [55, 63] Several groups have been working to develop and test therapies to target TAMs by blocking the migration and recruitment of the macrophages, re-polarizing the macrophages from M2 to M1-like phenotype, or by blocking the immune checkpoint PD-L and PD-L1. Some of these anti-TAMs therapies are in clinical trials. [55] TAMs are adaptable to changes in the tumor microenvironment, and there are complex immune cell interactions occurring that we still do not fully understand. [55, 64-67]

Most recently, Izar et al. performed single-cell sequencing on 22 ascites samples from 11 HGSC patients for a total of 35,957 cell profiles. A highlight from this study is their comparison between the cluster expression of their cohort to the four TCGA expression subtypes: differentiated proliferative, mesenchymal, and immunoreactive. They observed that the mesenchymal and immunoreactive expression signatures show low expression in cancer cell clusters but were highly expressed in the macrophage and cancer-associated fibroblast (CAF) cell clusters. This correlation suggests that signatures derived from bulk RNA sequencing may reflect impure sampling of malignant cells and may actually reflect the composition of immune cells and CAFs in the tumor microenvironment. They also provided extensive evidence that the JAK/STAT pathway would be targetable in HGSC. They then performed a drug screen and identified that JSI-124 inhibits the JAK/STAT pathway to effectively inhibit OVCAR4 cell viability. Therefore, this group demonstrates that single-cell sequencing of cancer cells and immune cell populations from HGSC tumor samples and ascites can provide insights into the tumor microenvironment and discover treatments that can be utilized in HGSC. [68]

1.8 Thesis Rationale and Significance

HGSC is a rapidly evolving cancer that easily metastasizes, which leads to a high proportion of diagnoses being made at advanced stages that are more difficult to treat. Characterization studies of metastases, like Robinson et al. [34], have stressed the necessity to examine metastatic tumors to find genetic drivers of cancer. Yang et al. [32] has demonstrated that genomic differences between long-term and short-term survivors exist among HGSC primary tumors. Most studies, including Yang et al. [32], Adib et al. [69], and Patch et al. [31], thus far have yet to identify mechanisms that drive ovarian cancer metastasis and are predictive of poor prognosis outcomes. Therefore, understanding the genetic variation in HGSC metastases that exist between patients with different prognoses can better inform treatments and may identify new targets for drug development. In order to improve patient outcomes, there is also a need

to identify actionable targets that influence process of metastasis, so we can develop drugs that not only target metastatic tumors but can also serve as maintenance therapies to help prevent further metastasis and mediate chemoresistance.

I predict that there are distinct genetic characteristics that distinguish between primary and metastatic tumors of HGSC patients with different prognoses. I expect to identify recurrently mutated genes, mutations, gene fusions, copy altered genes, and/or differentially expressed gene in metastatic tumors and the tumors of HGSC patients with short survival. I also hypothesize that there are genes important to promoting metastasis that will be discovered using genetic screens and will significantly promote or suppress tumor cell attachment and invasion within the context of the tumor microenvironment. These revelations will lead to a deeper understanding of the mechanisms behind HGSC metastasis and have the potential to identify targets that can be further studied for treatment development.

1.9 References

1. Howlader N, N.A., Krapcho M, Miller D, Brest A, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ, Cronin KA (eds), *SEER Cancer Statistics Review.* National Cancer Institute, 1975-2017.
2. Lengyel, E., *Ovarian cancer development and metastasis.* Am J Pathol, 2010. **177**(3): p. 1053-64.
3. McPherson, A., et al., Divergent modes of clonal spread and intraperitoneal mixing in high-grade serous ovarian cancer. Nat Genet, 2016. **48**(7): p. 758-67.
4. Bowtell, D.D., The genesis and evolution of high-grade serous ovarian cancer. Nat Rev Cancer, 2010. **10**(11): p. 803-8.
5. Ducie, J., et al., Molecular analysis of high-grade serous ovarian carcinoma with and without associated serous tubal intra-epithelial carcinoma. Nat Commun, 2017. **8**(1): p. 990.
6. Rosen, D.G., et al., *Ovarian cancer: pathology, biology, and disease models.* Front Biosci (Landmark Ed), 2009. **14**(6): p. 2089-102.
7. Levanon, K., C. Crum, and R. Drapkin, New insights into the pathogenesis of serous ovarian cancer and its clinical impact. J Clin Oncol, 2008. **26**(32): p. 5284-93.
8. Cojocaru, E., C.A. Parkinson, and J.D. Brenton, *Personalising Treatment for High-Grade Serous Ovarian Carcinoma.* Clin Oncol (R Coll Radiol), 2018. **30**(8): p. 515-524.
9. Fagö-Olsen, C.L., et al., Does neoadjuvant chemotherapy impair long-term survival for ovarian cancer patients? A nationwide Danish study. Gynecol Oncol, 2014. **132**(2): p. 292-8.
10. Fagotti, A., et al., Randomized trial of primary debulking surgery versus neoadjuvant chemotherapy for advanced epithelial ovarian cancer (SCORPION-NCT01461850). Int J Gynecol Cancer, 2020. **30**(11): p. 1657-1664.
11. Stewart, C., C. Ralyea, and S. Lockwood, *Ovarian Cancer: An Integrated Review.* Semin Oncol Nurs, 2019. **35**(2): p. 151-156.

12. Mirza, M.R., S. Pignata, and J.A. Ledermann, *Latest clinical evidence and further development of PARP inhibitors in ovarian cancer.* Ann Oncol, 2018. **29**(6): p. 1366-1376.

13. Del Campo, J.M., et al., Niraparib Maintenance Therapy in Patients With Recurrent Ovarian Cancer After a Partial Response to the Last Platinum-Based Chemotherapy in the ENGOT-OV16/NOVA Trial. J Clin Oncol, 2019. **37**(32): p. 2968-2973.

14. McLachlan, J., A. George, and S. Banerjee, *The current status of PARP inhibitors in ovarian cancer.* Tumori, 2016. **102**(5): p. 433-440.

15. Ledermann, J.A., *PARP inhibitors in ovarian cancer.* Ann Oncol, 2016. **27 Suppl 1**: p. i40-i44.

16. Martín-Cameán, M., et al., *The role of surgery in advanced epithelial ovarian cancer.* Ecancermedicalscience, 2016. **10**: p. 666.

17. Macintyre, G., et al., Copy number signatures and mutational processes in ovarian carcinoma. Nat Genet, 2018. **50**(9): p. 1262-1270.

18. Testa, U., et al., Ovarian Cancers: Genetic Abnormalities, Tumor Heterogeneity and Progression, Clonal Evolution and Cancer Stem Cells. Medicines (Basel), 2018. **5**(1).

19. Bell, D., et al., *Integrated genomic analyses of ovarian carcinoma.* Nature, 2011. **474**(7353): p. 609-615.

20. Ciriello, G., et al., *Emerging landscape of oncogenic signatures across human cancers.* Nature Genetics, 2013. **45**(10): p. 1127-1133.

21. Muller, P.A.J. and K.H. Vousden, *p53 mutations in cancer.* Nature Cell Biology, 2013. **15**(1): p. 2-8.

22. Olivier, M., M. Hollstein, and P. Hainaut, *TP53 mutations in human cancers: origins, consequences, and clinical use.* Cold Spring Harb Perspect Biol, 2010. **2**(1): p. a001008.

23. Silwal-Pandit, L., A. Langerød, and A.L. Børresen-Dale, *TP53 Mutations in Breast and Ovarian Cancer.* Cold Spring Harb Perspect Med, 2017. **7**(1).

24. Donehower, L.A., et al., Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. Nature, 1992. **356**(6366): p. 215-21.

25. Malkin, D., et al., Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. Science, 1990. **250**(4985): p. 1233-8.

26. Kang, H.J., et al., Clinical relevance of gain-of-function mutations of p53 in high-grade serous ovarian carcinoma. PLoS One, 2013. **8**(8): p. e72609.

27. Nakamura, M., et al., The Association and Significance of p53 in Gynecologic Cancers: The Potential of Targeted Therapy. Int J Mol Sci, 2019. **20**(21).

28. Zhu, J., et al., Gain-of-function p53 mutants co-opt chromatin pathways to drive cancer growth. Nature, 2015. **525**(7568): p. 206-11.

29. Fuller, E. and V. Howell, Culture Models to Define Key Mediators of Cancer Matrix Remodeling. Frontiers in Oncology, 2014. **4**.

30. Schumacher, T.N. and R.D. Schreiber, *Neoantigens in cancer immunotherapy.* Science, 2015. **348**(6230): p. 69-74.

31. Patch, A.M., et al., Whole-genome characterization of chemoresistant ovarian cancer. Nature, 2015. **521**(7553): p. 489-94.

32. Yang, S.Y.C., et al., Landscape of genomic alterations in high-grade serous ovarian cancer from exceptional long- and short-term survivors. Genome medicine, 2018. **10**(1): p. 81-81.

33. Veeraraghavan, J., et al., Recurrent ESR1-CCDC170 rearrangements in an aggressive subset of oestrogen receptor-positive breast cancers. Nat Commun, 2014. **5**: p. 4577.

34. Robinson, D.R., et al., *Integrative clinical genomics of metastatic cancer.* Nature, 2017. **548**(7667): p. 297-303.

35. Hillman, R.T., et al., Genomic Rearrangement Signatures and Clinical Outcomes in High-Grade Serous Ovarian Cancer. J Natl Cancer Inst, 2018. **110**(3): p. 265-72.

36. Nik-Zainal, S., et al., Landscape of somatic mutations in 560 breast cancer whole-genome sequences. Nature, 2016. **534**(7605): p. 47-54.

37. Popova, T., et al., Ovarian Cancers Harboring Inactivating Mutations in CDK12 Display a Distinct Genomic Instability Pattern Characterized by Large Tandem Duplications. Cancer Res, 2016. **76**(7): p. 1882-91.

38. Menghi, F., et al., *The tandem duplicator phenotype as a distinct genomic configuration in cancer.* Proc Natl Acad Sci U S A, 2016. **113**(17): p. E2373-82.

39. Menghi, F., et al., The Tandem Duplicator Phenotype Is a Prevalent Genome-Wide Cancer Configuration Driven by Distinct Gene Mutations. Cancer Cell, 2018. **34**(2): p. 197-210.e5.

40. Garsed, D.W., et al., The genomic and immune landscape of long-term survivors of high-grade serous ovarian cancer. Nature Genetics, 2022. **54**(12): p. 1853-1864.

41. Lei, J.T., et al., Functional Annotation of ESR1 Gene Fusions in Estrogen Receptor-Positive Breast Cancer. Cell Rep, 2018. **24**(6): p. 1434-1444.e7.

42. Cocco, E., M. Scaltriti, and A. Drilon, *NTRK fusion-positive cancers and TRK inhibitor therapy.* Nat Rev Clin Oncol, 2018. **15**(12): p. 731-747.

43. Gao, Q., et al., Driver Fusions and Their Implications in the Development and Treatment of Human Cancers. Cell Rep, 2018. **23**(1): p. 227-238.e3.

44. Kannan, K., et al., CDKN2D-WDFY2 is a cancer-specific fusion gene recurrent in high-grade serous ovarian carcinoma. PLoS Genet, 2014. **10**(3): p. e1004216.

45. Kannan, K., et al., Recurrent BCAM-AKT2 fusion gene leads to a constitutively activated AKT2 fusion kinase in high-grade serous ovarian carcinoma. Proc Natl Acad Sci U S A, 2015. **112**(11): p. E1272-7.

46. Katopodis, P., et al., *Kinase Inhibitors and Ovarian Cancer.* Cancers (Basel), 2019. **11**(9).

47. Do, T.V., et al., Aurora kinase A mediates epithelial ovarian cancer cell migration and adhesion. Oncogene, 2014. **33**(5): p. 539-49.

48. Morotti, M., et al., *Targeting tyrosine-kinases in ovarian cancer.* Expert Opin Investig Drugs, 2013. **22**(10): p. 1265-79.

49. Pérez-Fidalgo, J.A., et al., *Aurora kinases in ovarian cancer.* ESMO Open, 2020. **5**(5): p. e000718.

50. Quinn, J.M., et al., Therapeutic Inhibition of the Receptor Tyrosine Kinase AXL Improves Sensitivity to Platinum and Taxane in Ovarian Cancer. Mol Cancer Ther, 2019. **18**(2): p. 389-398.

51. Rankin, E.B., et al., AXL is an essential factor and therapeutic target for metastatic ovarian cancer. Cancer Res, 2010. **70**(19): p. 7570-9.

52. Grither, W.R., et al., TWIST1 induces expression of discoidin domain receptor 2 to promote ovarian cancer metastasis. Oncogene, 2018. **37**(13): p. 1714-1729.

53. Tu, M.M., et al., Targeting DDR2 enhances tumor response to anti-PD-1 immunotherapy. Sci Adv, 2019. **5**(2): p. eaav2437.

54. Elkamhawy, A., et al., The Journey of DDR1 and DDR2 Kinase Inhibitors as Rising Stars in the Fight Against Cancer. Int J Mol Sci, 2021. **22**(12).

55. Nowak, M. and M. Klink, The Role of Tumor-Associated Macrophages in the Progression and Chemoresistance of Ovarian Cancer. Cells, 2020. **9**(5).

56. Zhang, L., et al., Intratumoral T cells, recurrence, and survival in epithelial ovarian cancer. N Engl J Med, 2003. **348**(3): p. 203-13.

57. Santoiemma, P.P. and D.J. Powell, Jr., *Tumor infiltrating lymphocytes in ovarian cancer.* Cancer Biol Ther, 2015. **16**(6): p. 807-20.

58. Li, X., et al., Immune Cell Infiltration Landscape of Ovarian Cancer to Identify Prognosis and Immunotherapy-Related Genes to Aid Immunotherapy. Front Cell Dev Biol, 2021. **9**: p. 749157.

59. Kryczek, I., et al., Relationship between B7-H4, regulatory T cells, and patient outcome in human ovarian carcinoma. Cancer Res, 2007. **67**(18): p. 8900-5.

60. Wolf, D., et al., The expression of the regulatory T cell-specific forkhead box transcription factor FoxP3 is associated with poor prognosis in ovarian cancer. Clin Cancer Res, 2005. **11**(23): p. 8326-31.

61. Barnett, J.C., et al., Ovarian cancer tumor infiltrating T-regulatory (T(reg)) cells are associated with a metastatic phenotype. Gynecol Oncol, 2010. **116**(3): p. 556-62.

62. Sehouli, J., et al., Epigenetic quantification of tumor-infiltrating T-lymphocytes. Epigenetics, 2011. **6**(2): p. 236-46.

63. Lan, C., et al., Expression of M2-polarized macrophages is associated with poor prognosis for advanced epithelial ovarian cancer. Technol Cancer Res Treat, 2013. **12**(3): p. 259-67.

64. Krishnan, V., et al., *Tumor associated macrophages in gynecologic cancers.* Gynecol Oncol, 2018. **149**(1): p. 205-213.

65. Mantovani, A. and P. Allavena, The interaction of anticancer therapies with tumor-associated macrophages. J Exp Med, 2015. **212**(4): p. 435-45.

66. Cheng, H., et al., Macrophage Polarization in the Development and Progression of Ovarian Cancers: An Overview. Front Oncol, 2019. **9**: p. 421.

67. Zheng, X., et al., Redirecting tumor-associated macrophages to become tumoricidal effectors as a novel strategy for cancer therapy. Oncotarget, 2017. **8**(29): p. 48436-48452.

68. Izar, B., et al., A single-cell landscape of high-grade serous ovarian cancer. Nat Med, 2020. **26**(8): p. 1271-1279.

69. Adib, T.R., et al., *Predicting biomarkers for ovarian cancer using gene-expression microarrays.* British journal of cancer, 2004. **90**(3): p. 686-692.

Chapter 2: Genetic Characterization of Primary and Metastatic Tumors with Survival in High-Grade Serous Ovarian Cancer

Declarations

Emilee N. Kotnik contributed to consolidating, analyzing, interpreting all genomics data, creating figures and writing and preparing the manuscript. Mary M Mullen contributed to interpretation of the data and collecting patient clinical data. Nicholas C. Spies contributed to analyzing and interpreting the SNV and copy number data. Tiandao Li contributed to initial tumor sequencing processing and genomic data analysis. Matthew Inkman and Jin Zhang analyzed RNA-seq data with INTEGRATE for gene fusion analysis. Fernanda Martins-Rodrigues contributed code and helped with DE analysis. Ian S. Hagemann, Carolyn K. McCourt, Premal H. Thaker, Andrea R. Hagemann, Matthew A Powell, and David G. Mutch, all contributed to surgically collecting tumors and patient data collection for the study. Christopher A. Maher helped interpret gene fusion and genomic data. Christopher A. Miller, Elaine R. Mardis, and Gregory Longmore contributed to advising on genomic analyses, interpreting genomic data, and revising this manuscript. Katherine C. Fuh is the senior author and contributed to the conception, design of the project, and preparation of the manuscript. The authors have read and approved of the final manuscript.

2.1 Abstract

High-grade serous ovarian cancer (HGSC) is the most lethal histotype of ovarian cancer and the majority of cases present with metastasis and late stage disease. Over the last few decades, the overall survival for patients has not significantly improved and there are limited targeted treatment options. We aimed to better characterize the distinctions between primary and metastatic tumors based on short- or long-term survival. We characterized 39 matched primary and metastatic tumors by whole exome and

RNA sequencing. Of these, 23 were short-term (ST) survivors (overall survival (OS) <3.5 years) and 16 were long-term (LT) survivors (OS >5 years). We compared somatic mutations, copy number alterations, mutational burden, differential gene expression, immune cell infiltration, and gene fusion predictions between the primary and metastatic tumors and between ST and LT survivor cohorts.

We identified more distinct genomic and transcriptomic alterations between ST and LT survivors. Overall, there were few differences in RNA expression between paired primary and metastatic tumors, but significant differences between the transcriptomes of LT and ST survivors in both their primary and metastatic tumors. We identified gene fusions in ST Survivors involving *ESR1* and collagen genes. Additionally, there were DE lncRNAs between ST and LT survivors.

2.2 Introduction

High-Grade Serous Cancer (HGSC) of ovary, fallopian tube, or peritoneum is the most lethal ovarian cancer histotype and the second most common gynecologic malignancy [1, 2]. Over the past few decades, chemotherapy has been the standard of care, yet overall survival (OS) has not significantly improved [3, 4]. PARP inhibitors, which target base-excision DNA repair mechanisms and cause genetic lethality in tumors of patients harboring *BRCA1* or *BRCA2* mutations or homologous recombination deficiencies (HRD), can be used as maintenance therapies, but are only applicable for ~ 50% of HGSC patients [5, 6]. The majority of cases, about 80%, present with late stage disease, when the tumor has already metastasized [3, 4, 7]. These patients have only a 29.2% chance of surviving longer than 5 years [4]. Therefore, to improve patient survival, we sought to better characterize the genomic and transcriptomic landscapes of matched primary and metastatic ovarian cancers and to identify novel targets for drug development, especially in a more aggressive metastatic disease setting.

Large-scale tumor characterizations, by consortia such as The Cancer Genome Atlas (TCGA), have established that primary HGSC tumors harbor ubiquitous *TP53* mutations and copy number

alterations, and a low prevalence of other recurrently mutated genes [8]. Prior genomic studies of ovarian cancers have only included large numbers of primary tumors rather than comparing matched primary and metastatic disease in the context of outcomes. A recent study, Yang et al., characterized clinical and genomic features from HGSC primary tumors that correlated with short-term (ST, OS < 2 years) and long-term (LT, OS > 10 years) survival [9]. While this study does provide evidence that there are clinical and genomic characteristics unique to patient survival in primary tumors, metastatic tumors were not included. Study design is further nuanced by the context of survival duration, since exceptional survivors (>10 years of survival), have a high prevalence of *BRCA* mutations, and are known to respond well to standard therapy [10]. One study has examined genomic and transcriptomic sequencing from matched primary and metastatic tumors in the context of response to chemotherapy or surgical resection [11]. Another study identified a transcriptome signature that distinguished between primary and metastatic tumors but did not relate this to survival [12].

Here, we sought to determine whether there are unique features in the genomes and transcriptomes of metastatic tumors from short-term survivors when compared to their matched primary tumors and/or to primary/metastasis paired tumors from long-term survivors. Our cohort design examines these differences between tumors from patients within the median survival time for ovarian cancer [13]. In this context, we compared somatic variants, copy number alterations, mutational burden, differential expression, immune cell infiltrates, and gene fusion predictions between chemo naïve primary and metastatic tumors from 23 HGSC short-term (ST, OS < 3.5 years) survivors and 16 HGSC long-term (LT, OS >5 years) survivors using whole exome sequencing (WES) and RNA sequencing (RNA-seq).

2.3 Materials and Methods

2.3.1 Patient cohort sample criteria

We collected normal tissue, primary tumor, and metastatic tumor samples from a total of 39 patients diagnosed with FIGO stage III or higher HGSC. Normal tissue samples consisted of adjacent non-malignant omentum or peritoneum. All tumors were collected during primary cytoreductive surgery, prior to any chemotherapy treatment, and were stored as either fresh frozen (FF) or formalin-fixed paraffin-embedded (FFPE). These patients were separated into two groups, based on their overall survival. Patients who lived less than 3.5 years after their diagnosis were considered short-term (ST) survivors and patients who lived more than 5 years after diagnosis were considered long-term (LT) survivors (Table 1). Other clinical characteristics of patients are shown in Table 1. All patients received standard regimens of carboplatin and paclitaxel following cytoreductive surgery. More LT survivors received intraperitoneal (IP) chemotherapy than ST survivors (1 ST survivors, 5 LT survivors, p-value=0.042), Otherwise there were no differences in the use of bevacizumab or PARP inhibitor treatments between the two cohorts. All 23 ST patients and 12 LT patients had matched DNA and RNA extracted and sequenced. An additional 4 LT patients had tumor sequencing performed: Patients 031 and 035 had matched primary and metastatic tumors DNA sequenced and Patients 032 and 040 only had RNA-sequencing from their matched primary and metastatic tumor tissue.

2.3.2 Exome and RNA sequencing

All tumors were examined by a pathologist to determine tumor cellularity and necrosis and only samples of 60% tumor cellularity or higher with <20% necrosis were sequenced. DNA and RNA were extracted from FF or FFPE tissues using Qiagen's DNeasy Blood & Tissue Kit and RNeasy kit. Whole exome sequencing of DNA from matched primary tumor, metastatic tumors, and normal tissue samples was completed for 39 patients with the NimbleGen VCRome exome capture kit (NimbleGen Roche)

according to the manufacturer's protocol. Paired-end Illumina 151 bp reads were generated for normal samples to a minimum of depth of 65x, while tumor samples were sequenced to a minimum of 139x, with the average coverage of ~300x. A coverage table provides per-sample coverage details (Supplementary Table 1). RNA sequencing of primary and metastatic tumor samples was performed using the Illumina TruSeq stranded Total RNA library kit following the Manufacturer-recommended protocol. Paired-end Illumina sequencing of 151 bp read length yielded an average of approximately 125 million paired reads per sample and an average of approximately 134 million reads mapped per sample. Quality Control metrics for the RNA-seq samples were generated using MultiQC and are reported in Supplementary File 2 [14].

2.3.3 Variant calling and genomic analysis

Exome sequencing data were aligned to human reference build GRCh37 using BWA-mem and deduplicated with Picard version 1.113 [15, 16]. Somatic variants were called from combined data using the Genome Modeling System pipeline [17, 18]. In brief, variants were called from the union of 4 callers which included Samtools version r932, Somatic Sniper version 1.0.4, VarScan version 2.3.6, Strelka version 1.0.11, and Mutect v1.1.4 [17, 19-23]. Indels were detected from the union of 4 callers; GATK somatic-indel version 5336, Pindel version 0.5, VarScan version 2.3.6, and Strelka version 1.0.11 [17, 20, 21, 24, 25]. Further variant filtering was applied as described in Ghobadi et al. [17]. Briefly, SNVs and indels were discarded if they had below 20x coverage, appeared as artifacts in a panel of 905 normal exomes, or exceeded 0.1% frequency in the 1000 genomes or NHLBI exome sequencing projects [26, 27]. A Bayesian classifier (https://github.com/genome/genome/blob/master/lib/perl/Genome/

Model/Tools/Validation/IdentifyOutliers.pm) was also applied and variants that classified as somatic with a binominal log-likelihood of at least 5 were retained [17]. All called variants compared in this study are provided in Supplementary Table 2. Mutational burden was calculated as the number of variants called per megabase for all variants that passed the QC filtering. The waterfall plot (Figure 1A-B) depicting

frequently mutated genes from TCGA-OV was generated using GenVisR [8, 28]. Mutational clinical significance for somatic and germline BRCA mutations was determined from ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/) (Tables 3-5) [29]. CLOVAR signatures were calculated according to parameters defined in Verhaak et al. [30]. Copy-altered segments were identified from VarScan (Supplementary Figure 5) [21]. Significant copy-altered segments were identified for all tumors, all tumors from ST survivors, all tumors from LT survivors, metastatic tumors, primary tumors, and only the metastatic tumors of ST survivors using the GISTIC 2.0 version 6.15.28 Module on the AWS GenePattern cloud (https://cloud.genepattern.org/gp). Default parameters and reference genome Human_Hg19.mat were used to run GISTIC 2.0 analyses [31]. We used the wide peak region analyses from GISTIC 2.0 to calculate the total number of genes amplified or deleted within those regions. The correlation between CNA and RNA-seq expression was completed using the thresholded CNA values GISTIC 2.0 calculated based on each sample's segment files [31, 32]. The violin plot was created by binning all CNA threshold values from every gene for every sample and plotting that with their corresponding log2(FPKM) values (Supplementary Figure 5E).

2.3.4 Differential expression analysis

Normalization and quality control: Transcript read counts were obtained using Kallisto version: v0.43.1 and gene-level read counts were calculated using GRCh37 in Ensembl [33]. Quality control and normalization of the raw count data were performed using the R/Bioconductor package edgeR version 3.28 [34]. For our comparison of LT survivor samples to ST survivor samples, we removed genes with less than 1 Count Per Million (CPM) mapped reads in at least half of the samples to ensure that a gene was retained if expressed in only one of the two groups. For our comparison of primary to metastatic tumors, genes with less than 1 CPM mapped reads in at least half the samples were removed. Normalization factors were calculated using the Trimmed Mean of M-values (TMM) normalization method in edgeR to account for compositional biases in libraries between each pair of samples.

Removal of batch effects: Due to technical artifacts introduced by the use of FFPE that can affect gene expression analyses, we performed batch effect correction prior to differential expression analysis for the comparison of LT to ST survival samples [35, 36]. We used the SVA function of the R/Bioconductor package SVA version 3.34.0 to estimate and remove surrogate variables for unwanted and unknown batch effects and other sources of variation present in the data [37]. The SVA function estimated surrogate variables for each subset analysis, which was adjusted for within the statistical model applied in the edgeR package in downstream analyses of differential gene expression. After batch effect correction, samples were analyzed by a Principal Component Analysis (PCA) using the R function "dist" on regularized log-transformed (rlog) data to calculate the Euclidian distance between samples. Plotting of the first (PC1) and second (PC2) principal components revealed that expression values from the same patient are more related to one another than between groups *(*Supplementary Figure 3*)*. We also observed 4 potential outlier samples, which were removed from downstream analyses because of their distance from the other samples in the PCA plot after normalizing and batch correcting transcript counts. These 4 removed samples are highlighted in Supplementary Figure 3A and were all collected within the same year, but their exclusion could mean we are missing out on some biological features of these tumor samples.

Differential gene expression (DGE) analysis: DGE analysis was performed using edgeR version 3.28.0, which implements a negative-binomial general linear model [34]. We performed 4 comparisons: ST survival samples versus LT survival samples for all tumors in the study; ST survival versus LT survival among metastatic tumors; ST survival versus LT survival among primary tumors; and primary tumors versus their matched metastatic tumor. The surrogate variables estimated with SVA were included in the model used for the LT versus ST survival comparison. To normalize gene-level variance, the biological coefficient of variation (BCV) was calculated using Cox-Reid dispersion for negative binomial general linear models. The p-values of differential expression tests were corrected for multiple-hypothesis testing

using Benjamini-Hochberg false discovery rate (FDR) correction. The threshold for significance was set to FDR Q-value < 0.01. We further curated our differentially expressed genes (DEGs) by limiting to protein coding genes that were listed in Ensembl genes 100 Human genes (GRCh38.p13) protein_coding transcript type on BioMart. All DEGs discussed in this paper are listed in Supplementary File 1. Pathway analysis was applied to the DEG and gene fusion gene lists using the PANTHER classification system 16.0 (http://pantherdb.org/), with the organisms set as 'Homo sapiens' and performing a statistical overrepresentation test using Fisher's Exact test and calculating a False Discovery Rate [38]. We used all Gene Ontology (GO) terms (Biological Processes, Molecular Function, and Cellular Components), PANTHER pathways, and Reactome pathways annotation sets [39-41]. We used DAVID to identify enrichment for KEGG and Biocarta pathways. [42, 43]

2.3.5 Immune cell abundance estimates

We used Cibersort (https://cibersort.stanford.edu/) to estimate the abundance of infiltrating immune cell types using our tumor RNA-seq data [44]. We generated a mixture file for our cohort of tumor samples based on the gene abundance counts generated from the RNA-seq reads using Kallisto [33]. We used the LM22 gene signature, which calculated immune cell fractions for 22 immune cell types, and ran our Cibersort analysis with 500 permutations under the relative mode.

2.3.6 Gene fusion predictions

Gene fusion predictions for each tumor sample were produced using INTEGRATE v0.2.6 to analyze the tumor RNA-sequencing data [45]. Full-length raw reads and a set of reads trimmed to remove potentially low-quality bases were each aligned to human reference genome GRCh38 (r90) using STAR v2.5.3a with a minimum chimeric segment length of 18 and chimeric alignments output to a separate SAM file [46]. The chimeric alignments were then used as inputs for INTEGRATE fusion with default parameters for fusion discovery with tumor RNA-seq only. Fusion predictions from the full and trimmed reads were

then merged and manually reviewed to ensure all fusion calls were valid. Since normalization of FFPE and FF tumor samples is more challenging for gene fusions, we characterized the predicted gene fusions as independent events regardless of sample preparation.

2.4 Results

2.4.1 Characterization of Genomic Landscape

We compared somatic variants, copy number alterations, and mutational burden between the primary and metastatic tumors of the ST and LT survival groups. Our cohort of patient tumors exhibited characteristics typical of those seen in previously sequenced HGSC tumors, such as nearly ubiquitous *TP53* mutations, high numbers of copy alterations, and a low number of recurrently mutated genes (Figure 1A-B). As was found in the Yang et al study, our cohort of LT survivors also exhibited a significantly greater mutational burden than the ST survivors (Mann-Whitney-Wilcoxon Test stat (MW) statistic = 611, p-value = 5.1e-6) [9]. There was no statistical difference between the mutational burden of paired primary and metastatic tumors (MW stat = 611, p-value = 0.4299) (Figure 2C).

Interestingly, in comparison to recurrently mutated genes identified in the TCGA cohort, we also observed that *RB1* mutations were found only in primary and metastatic tumors of LT survivors. This is a finding consistent with studies analyzing exceptional HGSC survivors [8, 47]. *CDK12* was only mutated in the LT survivors. We confirmed previously published findings that tumors of LT survivors were more likely to have *BRCA1* alterations and copy-altered segments when compared to ST survivors [9]. In total, we identified an average of 723 somatic mutations per LT survivor (10,124 SNVs total / 14 patients) and an average of 591 somatic mutations per ST survivor (13,599 SNVs total/ 23 patients). Four patients exhibited somatic *BRCA1* mutations with VAF >30 (Table 1). Six patients exhibited germline *BRCA1* mutations and 1 LT survivor had a benign germline *BRCA2* mutation (VAF >30), all of which had mixed ClinVar-based clinical interpretations of pathogenic significance (Tables 4-5) [29].

We applied "Classification of Ovarian Cancer" (CLOVAR) signatures to our data set and observed no statistically significant differences based on survivorship or tumor type among the signatures. The most prevalent signatures among our cohort were Mesenchymal and Differentiated (Figure 1B) [30]. To compare the types of variants found in each of the tumors, we calculated the percentage of total variants identified in each patient that can be contextualized by the human cancer mutational signatures [48] [49]. In our cohort, the most common signatures were for 5-methylcytosine deamination, mismatch repair, and double-strand break repair, along with a large number of mutations contributing to the unknown signature. There were no statistically significant differences between the signature percentages when comparing matched primary and metastatic tumors or ST and LT survivors. However, there is a higher percentage contribution to the mismatch repair signature in LT survivors compared to ST survivors, although not statistically significant (Supplementary Figure 1). We also calculated percentages of the mutational nucleotide transitions and transversions. Transitions from C>T account for the largest percentage of total mutations in most of the tumor samples. The percent of total mutation ratios remain relatively the same between the primary and metastatic tumors in the majority of patients, but large shifts in the mutational transitions and transversions can be seen in Patients 15, 20, 30 and 34. There were no statistically significant differences when comparing the mutational percentages of transitions and transversions between ST and LT survivors and primary and metastatic tumors (Supplementary Figure 2).

2.4.2 ST survivors exhibited a higher percentage of shared variants

For each patient, we calculated the percentage of called variants that were unique to the primary or to the metastatic tumor, or were shared between the two tumors. We observed that there were higher percentages of shared variants between the primary and metastatic tumors for ST survivors compared to LT survivors, although this was not statistically significant (MW statistic = 117.0, p-value = 0.0866) (Figure 2A-B). Of note, all of our LT survivor samples were FFPE whereas the ST survivors included FF specimens. There was no FFPE/FF specific variant filtering applied in our variant calling pipeline, but

each sample did undergo quality control and log likelihood ratio (LLR) filtering (https://github.com/genome/docker-somatic-llr-filter/blob/master/somatic_llr_filter.py). Thus, the differences seen in the number of shared variants of the ST and LT cohorts could be affected by FFPE artifacts from sample preparation.

Within our cohort, all but 2 patients had tumors that harbored *TP53* variants (Table 2). Of the 35 patients that carried *TP53* mutations, all but one patient shared the same *TP53* mutation between their primary and metastatic tumor and 4 patients carried multiple *TP53* mutations. The majority of *TP53* mutations were missense or frame-shift deletions within the DNA-binding domain. One known hotspot mutation, R273H, was present in 3 patients, 2 of which were ST survivors. The 2 ST survivors that harbored this hotspot mutation had an overall survival ranging between 17-19.6 months, whereas the LT survivor lived more than 147 months after their diagnosis. In the TCGA-OV patient cohort, 2% (11/489) of tumor samples also had the TP53 R273H mutation, compared to the 8% (3/37) in our cohort [8, 50, 51].

2.4.3 LT survivors exhibited more copy number alterations

Concordant with findings from the TCGA-OV project, copy number alterations (CNAs) were abundant in these data, with CNAs observed in every sample, and the number of copy-altered segments ranged from 33 to 739, with segment lengths ranging from 3,636 to 229,754,969 nucleotides. The average segment length was 3,997,903 nucleotides (Supplementary Figure 5A-B) [8]. The LT survivors had a greater proportion of copy-altered segments ($p = .03$, 95% CI 0.004 to 0.11), driven by a greater proportion of amplifications ($p = .01$, 95%CI 0.01-0.1). There was no significant difference between primary tumor samples and metastases, nor were there significant differences in mean estimated ploidy between groups.

We identified recurrent CNAs in our cohort overall and subsets of the ST survivors, LT survivors, primary tumors, and metastatic tumors (Supplementary Figure 5C-D) [31]. Overall, our cohort exhibited a total of 254 recurrent copy-altered segments, including 85 amplified segments and 169 deleted segments with a 90% confidence interval. We identified 2333 genes within the amplified peaks and 4904 genes within the deleted segments. Region 20q13.12, previously identified in other ovarian tumors, was amplified in our cohort, along with other genes that have previously found amplified in ovarian cancer such as *CCNE1, ERBB2, RSF1* and deleted genes like *BRCA1* [8, 52]. Among the cohort and subset analyses, there were more recurrently deleted segments than amplified segments and regions within 8q, 3q, and 19q were among the most recurrently amplified segments while peaks on 9q, 15q, 16q, 17q were among the most recurrently deleted segments. We correlated the CNA and our RNA-seq data for all samples in our cohort, utilizing the threshold values GISTIC2.0 calculates with their corresponding RNA-seq FPKM values for the genes involved in altered regions (Supplementary Figure 5E). The relationship between copy number and expression is not simple, but the medians of the data suggest that more amplified genes trend toward having higher RNA expression.

Our GISTIC2.0 analysis results comparing ST to LT survivors revealed that there were more recurrent copy-altered segments among the ST survivors (ST=79 amplified, 198 deleted; LT= 60 amplified, 101 deleted). Comparing the primary and metastatic tumor analyses showed that the metastatic tumors had more recurrent segments (primary= 39 amplified, 135 deleted; met= 63 amplified, 157 deleted). Consistent with published data [9], *CCNE1* was amplified in ST survivors, primary tumors, and metastatic tumors sample subsets, but not among the LT survivor samples. Both the primary tumor and metastatic tumor subsets were significantly amplified at 19q12 and at 20q13.12, while the ST subset had 19q12 amplified and the LT subset had 20q13.12 significantly amplified.

2.4.4 Differentially expressed genes correlate with survival

We calculated differential expression (DE) of genes between the ST and LT survivors in both primary and metastatic tumors. Overall, there were distinct transcriptomes that correlated with ST or LT survival, both within the tumor cohorts separately and when combining all patients regardless of tumor type (Figure 3A-B).

Within the metastatic tumor cohort, there were 4792 DE genes (DEGs) between ST and LT survivors, with an FDR <0.01, after selecting for only protein-coding genes. Genes such as *SZRD1* and *ERV3-1* were upregulated in the metastatic tumors of ST survivors, as previously reported in other solid cancer types such as cervical and colorectal cancer [53, 54].

In order to identify DEGs that were specifically associated with survival in the metastatic tumors, we filtered out any genes that were also differentially expressed between the ST and LT survivors' primary tumors. This revealed 325 genes only DE in metastatic tumors, with 295 of these (90.7%) downregulated in ST survivors (Figure 3A). The DEGs unique to metastatic tumors of ST survivors are enriched for several Biological Processes GO terms with an FDR <0.05, such as "regulation of cellular biosynthetic process" (Supplementary Figure 4A) [38-41]. A GO enrichment analysis on the 30 upregulated DEGs unique to metastatic tumors showed enrichment with an FDR <0.05 for several Molecular Function GO terms associated with DNA binding and transcriptional activity (Figure 3C) [38, 40, 41]. This enrichment is most likely due to the 13/30 of those upregulated DEGs that are in the zinc finger family, and many of which have previously been implicated in cancer. We also used DAVID to find enrichment of the KEGG and Biocarta pathways within our DEGs, and although there were no significantly enriched Biocarta terms, there were 9 KEGG pathways enriched. Some of these included "Adherens junction" and "protein processing in endoplasmic reticulum". (Supplementary Figure 4C) Of note, *FOXL2NB* and *PTCH2* have correlated with poor survival in other cancer types [55, 56]. There is evidence that *OGN* plays a role in *EMT,* and *PRDX1* has been studied as prognostic marker in lung cancer

[57, 58]. We calculated DE between genes of ST and LT survivors within the primary tumors. We found that there was a total of 4248 DE genes with FDR <0.01. After filtering for protein-coding genes, we narrowed our list of DE genes to 3694, with 502 DEGs that were specifically differentially expressed only in primary tumors (Figure 3B).

When all tumors are included in the DE analysis, there were a total of 7304 protein-coding DEGs between ST and LT survivors (Supplementary Table 1). The top 50 upregulated and top 50 downregulated DEGs are included in Supplementary Figure 4B. Additionally, we calculated DEGs between primary and metastatic tumors and identified only 4 DEGs with an FDR <0.01. When we lower the FDR filter to <0.1, the number of DEGs increased to 15. Of those, 5 genes *(WIPF3, STAR, SCUBE1, PEG3, CNTNAP2*) were also found in the top 100 DEGs identified by Sallinen et al, which compared DEGs between 10 matched primary and metastatic ovarian tumors having an FDR <0.1 [12].

2.4.5 Differentially expressed lncRNAs correlate with survival

From the RNA-sequencing data, we identified several long-noncoding RNA transcripts (lncRNAs) that were among the top differentially expressed transcripts between the ST and LT survivors in both metastatic and primary tumors. Within the metastatic tumors, we identified 11 lncRNAs that were differentially expressed and all but one was upregulated in ST survivors (Figure 3D). This set of lncRNAs included *ARRDC1-AS1*, which was shown to be a part of a potential lncRNA prognostic signature in breast cancer [59]. Among the primary tumors, we identified 36 lncRNAs of which 35 were upregulated in ST survivors. Of these 36 lncRNAs, 9 lncRNAs (25%), (*FAR2P1, ARRDC1-AS1, MIRLET7BHG, OVCH1-AS1, C11orf72, FLJ22447, LACTB2-AS1, ALOX12-AS1*, and *C5orf56*) overlapped with the lncRNAs identified in our metastatic tumor cohort.

2.4.6 Tumors from ST survivors harbored recurrent gene fusion predictions

A total of 1,164 gene fusions were predicted among our tumor sample cohort, 35 of which were recurrent (seen in at least 2 samples) and unique to ST survivors (Table 6). The higher number of gene fusions identified in ST survivors was due to a higher level of quality in the RNA-sequencing since this subgroup included FF tumor samples, whereas all LT survivors were FFPE samples.

INTEGRATE detected several *ESR1* gene fusions in our tumor cohort, which have previously been implicated in breast and ovarian cancer [9, 60]. In particular, the *ESR1>CCDC170* recurrent gene fusion identified by Yang et al. was present in 2 of our tumor samples, 1 ST primary tumor (5 reads) and 1 LT metastatic tumor (7 reads). Interestingly, we also noticed that a total of 33 gene fusions involved collagen genes, 32 of which were identified in ST survivors, 21 were in metastatic tumors, and 20 are in-frame fusions (Table 7). Pathway analysis on the genes involved in recurrent gene fusions in our cohort were significantly enriched for terms related to "collagen chain trimerization" and "Collagen degradation" in the PANTHER reactome set, which is interesting given the known role of collagen in the ovarian cancer tumor microenvironment [38-41].

2.4.7 Immune cell populations abundances

We used the program Cibersort to estimate the fraction of immune cell types in our tumor samples (Figure 4A). The immune cell groups CD4 T-Cells, macrophages, and monocytes had the highest fractions in many of the tumor samples. There was much variability in immune cell type fractions across patients, but there were few significant differences in immune cell fractions between primary or metastatic tumors or between ST and LT survivors among the 22 immune cell types (Supplementary Figure 6). Of note, CD4 naïve T-cells (higher fractions in ST), follicular helper T-cells (higher fractions in LT), regulatory T-cells (higher fractions in LT), and activated dendritic cells (higher fractions in LT) were

significantly different between ST and LT survivors with Mann-Whitney statistical p-values <0.05. Between the primary and metastatic tumors, the CD8 T-cells, activated CD4 memory T-cells, and neutrophils were significantly higher in metastatic tumors based on Wilcoxon statistical p-values <0.05. A chart of the statistical differences between all of the subsets for the 22 immune cell fractions is in Supplementary Figure 6W.

Lee et al. found significant abundance differences of M2 macrophages and monocytes between their R0 and NACT patient groups, and a significant difference between the abundance of resting CD4 memory T cells between primary and metastatic tumors, but these patterns did not appear in our dataset [12]. Thorsson et al. performed immunogenomic analysis across cancer types in TCGA and identified six immune subtypes [61]. In their analysis, the ovarian cancer cohort correlated the most with their C2 IFN-y dominant signature, which is defined by having high M1 and M2 macrophage polarization and strong CD8 signal. This is consistent with the higher fraction of macrophages we found in our Cibersort analysis. The ovarian cohort also had representation of their C1 wound healing and C4 lymphocyte depleted signatures, but did not have representation for their C3 inflammatory, C5 immunologically quiet, or C6 TGF-B Dominant signatures. The lack of these signatures is consistent with our cohort's low immune cell fractions for several lymphocytes and the variability between patient samples. Since our cohort included metastatic tumors that are not represented in TCGA, perhaps a more specific immunogenomic analysis with more metastatic tumors for ovarian cancer is necessary to better understand the immune landscape in these tumors [61].

2.5 Discussion

HGSC can rapidly metastasize before patients experience symptoms, therefore many patients are diagnosed at late stages and have limited treatment options. Despite many studies of the genetics of HGSC tumors, we have yet to fully characterize and identify genetic biomarkers of HGSC metastatic

tumors, especially those with poor survival outcomes. In this study we built on previous studies to better characterize the genomic features of matched primary and metastatic tumors in the context of patient survival, so we might identify unique features of ST metastatic tumors.

We found supporting data for *RB1* mutations as a marker for long survivorship as previously discovered, since *RB1* mutations were identified exclusively in our LT survivor cohort [8, 47]. In our study we found that there was a higher percentage of shared variants between the primary and metastatic tumors of ST survivors compared to LT survivors. Although this difference was not significant, it can suggest that tumors from ST survivors may be more clonal and genetically similar than tumors from LT survivors. This could mean that tumors from ST survivors are inherently more resistant to treatments, since both their primary and metastatic tumors are genetically similar. However, other studies of the clonality of HGSC tumors have yet to find a correlative pattern between clonality and survival [1, 7, 62-64], hence, many more tumor samples will need to be analyzed to answer this question. Shared variants that are likely to be present in all clones of the tumor may be the best suited for targeted therapies. With the advent of single-cell sequencing, we may now be able to answer more questions about the heterogeneity and clonal development of HGSC tumors [65].

TP53 mutations are a hallmark of high-grade serous ovarian cancer and *TP53* gene has known hotspot mutations across cancer types. One of these hotspot mutations, R273H, was identified in 3 patients within our cohort, two of whom had an overall survival ranging 17-34 months. In TCGA Genomic Data Commons Portal, there are a total of 99 cases across cancer types that harbor a mutation at this position in *TP53*, 9 of which are in ovarian cancer samples. Recent functional studies have shown that this particular mutation results in a p53 gain-of-function that may promote metastasis in colorectal, esophageal, and breast cancers. Additionally, breast cancer cell lines with a R273H gain-of-function have been found to have improved response to combination PARPi and a DNA-damaging agent [66-68]. If this is also seen in ovarian cancer cells, this may lead to additional patients receiving PARP inhibitor and

combination treatments in the future. However, further work in characterizing the therapeutic potential of this specific *TP53* mutation in ovarian cancer is needed.

Yang et al. demonstrated genomic differences between HGSC primary tumors of ST and LT extreme survivors [9]. Our study focused on paired primary and metastatic tumors within the median survival range of ovarian cancer. Yang et al. demonstrated that more than 50% of tumors with *BRCA* mutations are LT survivors with an OS >10 years. This is consistent since HGSC patients with *BRCA* mutations respond better to chemotherapy [10]. Therefore, our study was better able to characterize the genomic features of tumors from patients with more moderate survival to poor survival.

Recently gene fusions have proven to be useful drug targets for cancer. For example, the identification of EML4>ALK gene fusions in non-small cell lung cancer paved the way for the development of ALK inhibitors and recently drugs targeting tumors of any cancer type with gene fusions involving *NTRK* genes have been approved by the FDA [69, 70]. In our analyses, we identified an *ESR1-CCDC170* gene fusion in our cohort as previously described in another cohort [9]. There is evidence that *ESR1* gene fusions in estrogen receptor-positive breast cancer promote endocrine therapy resistance and metastasis, thus *ESR1* gene fusions may have a role in ovarian cancer progression [60]. Lei et al. demonstrated that CDK4/6 inhibitors were able to suppress growth that was driven by *ESR1* gene fusions, indicating that gene fusion driven cancers are treatable [60]. We found a higher number of gene fusion predictions in our tumors from ST survivors and these could be a potential source for new drug development, but additional work will be needed to identify recurrent gene fusions that are targetable in ovarian cancer

Previous studies have demonstrated that HGSC primary and metastatic tumors have similar transcriptomes. Two such studies using microarrays identified few differentially expressed genes between the HGSC primary and metastatic tumors [71, 72]. In this study, we also identified few DEGs between

primary and metastatic tumors. However, when we analyzed primary or metastatic tumors separately to find DEGs between ST and LT survivors, we found DEGs unique to metastatic tumors from patients with ST survival. This demonstrates that clinical outcome can be used to identify DEGs specific to metastatic tumors. We found several DEGs in the zing finger family that were upregulated in the ST survivor metastatic, suggesting that these tumors have more transcriptionally active genes that could be promoting metastatsis or could be used as markers for poor prognosis, like FOXL2NB [55] and PTCH2 [56] which have correlated with poor survival in other cancer types. The large number of DEGs that identified in our DE analyses are a resource for future studies for biomarkers given their correlation with poor prognosis in ovarian and other cancer types and because we filtered for genes unique to the metastatic tumors in our cohort.

Additionally, we identified lncRNAs that were differentially expressed between survival groups. LncRNAs have only recently been studied for their role in cancer development and prognosis and have not been extensively studied in ovarian cancer yet [73, 74]. There are some lncRNAs, such as RP11-190D6.2, that have shown to be tumor suppressors or oncogenes in ovarian cancer cell lines [74, 75]. Given that we found several lncRNAs having increased expression in ST survivors, they could serve as potential targets or biomarkers for future treatment development.

It should be noted that all of our LT survivor samples are from FFPE, while ST tumors were not. Though we have applied rigorous quality control and filtering to our variant calling, we cannot exclude the possibility that sample preparation has some effect on the results. It is possible that the batch correction from FFPE and FF samples reduced the number of DEGs that were able to be identified in our cohort between the primary and metastatic tumors. The SVA batch correction may have over accounted for unknown variation or it may be introduced variation, but was still necessary so we could include all tumor samples in our DE analysis, regardless of RNA sample preparation.This dataset, like many using

patient samples, has limitations but provides insights into the differences between HGSC primary and metastatic tumors in the context of moderate survival outcomes.

2.6 Conclusions

In this study we characterized the exomes and transcriptomes of a unique data set of matched primary and metastatic tumors in the context of patient survival. We were able to confirm many of the genomic features seen in previous studies [8, 9, 11, 12, 30]. We observed that the transcriptomes of primary and metastatic tumors were similar to each other, compared to the transcriptomes of tumors from ST and LT survivors that had more DEGs and DE lncRNAs. Our gene fusion analysis revealed previously identified and novel fusions that have the potential to be new targets in HGSC and could warrant further functional studies.

Our findings suggest that there are more defining genomic features of tumors from ST survivors rather than unique features of metastatic tumors or primary tumors. Further analysis and characterization of metastatic tumors of ST survivors could lead to additional treatment developments and maintenance therapies. Our results identified several unique expression patterns for HGSC metastatic tumors from ST survivors. This increased understanding of genetic variation in HGSC metastases that exist between patients with different prognoses can better inform treatments and may identify new targets for drug development.

2.7 Figures and Tables

	Short-Term (ST)	Long-Term (LT)
Patients (no.)	n=23	n=16
Tumors	46	32
Primary	23	16
Metastatic	23	16
Age (years)	61.5 ± 19.5	57 ± 8.2
FIGO Stage		
IIIA	2	0
IIIC	14	16
IV	7	0
FIGO Grade		
Moderately-differentiated	3	0
Poorly-differentiated	20	16
Histology		
Serous	22	16
Endometrioid	1	0
Median Overall Survival (OS) (months)	21 (range: 0-41)	111 (range: 82-195)
Fresh Frozen (FF) Samples		
Primary	17	0
Metastatic	11	0
Paraffin-Fixed (FFPE) Samples		
Primary	6	16
Metastatic	12	16
Whole-Exome Sequenced	46	14
RNA-sequenced	46	14

Table 2.1 Clinical Characteristics of Patient Cohort

Clinical Characteristics of Patients diagnosed with HGSC at FIGO stage III-IV. Tumor samples were collected from patients during primary cytoreductive surgery in Washington University in St. Louis.

Patient ID	Clonal Variant	Variant Type	Nucleotide Position	Amino Acid Change	p53 Protein Domain	Primary VAF	Metastatic VAF
ST_001	yes	missense	c.536T>C	p.H179R	DNA-binding domain	55.70	23.85
ST_002	yes	missense	c.423G>C	p.C141W	DNA-binding domain	76.04	55.11
ST_007	yes	mssense	c.707T>C	p.Y236C	DNA-binding domain	49.59	36.27
ST_009	yes	missense	c.659T>C	p.Y220C	DNA-binding domain	72.30	50.00
ST_011	yes	missense	c.818C>T	p.R273H	DNA-binding domain	92.86	64.00
ST_013	yes	frame shift deletion	C.636A>-	p.R213fs	DNA-binding domain	92.54	78.82
ST_014	yes	missense	c.818C>T	p.R273H	DNA-binding domain	47.57	25.00
ST_015	yes	frame shift deletion	c.250C>-	p.A84fs	NA	38.18	29.93
ST_015	no	5' UTR	c.-7579816G>T	NULL	NA	6.59	NA
ST_016	yes	frame shift deletion	c.297_294GGAA>-	p.S99fs	DNA-binding domain	65.69	39.93
ST_017	yes	missense	c.487A>T	p.Y163N	DNA-binding domain	53.65	77.71
ST_018	yes	frame shift insertion	c.371_370->A	p.C124fs	DNA-binding domain	53.23	50.00
ST_019	yes	in-frame deletion	c.478_473TGGCGC>-	p.RAM158in_frame_delL	DNA-binding domain	NA	NA
ST_020	yes	missense	c.830C>A	p.C277F	DNA-binding domain	21.28	72.41
ST_021	yes	splice site	c.560-1C>T	e5-1	NA	30.63	25.90
ST_023	yes	nonsense	c.499G>A	p.Q167*	DNA-binding domain	18.60	50.00
ST_025	yes	missense	c.490T>C	p.K164E	DNA-binding domain	49.77	62.31
ST_026	yes	missense	c.580G>A	p.L194F	DNA-binding domain	53.37	55.75
ST_027	yes	frame shift deletion	c.1024G>-	p.R342fs	tetramer	51.16	54.20
ST_028	no	splice site	c.376-2T>C	e4-2	NA	95.05	NA
ST_029	yes	missense	c.764A>C	p.I255S	DNA-binding domain	76.10	89.09
ST_030	yes	missense	c.817G>A	p.R273C	DNA-binding domain	76.00	22.22
LT_031	yes	missense	c.722G>A	p.S241F	DNA-binding domain	63.24	46.01
LT_033	yes	missense	c.524C>T	p.R175H	DNA-binding domain	73.97	67.80
LT_033	yes	intronic	c.96+41_96+34 CTCCAGGT>-	e2+41	NA	NA	NA
LT_034	yes	missense	c.818C>T	p.R273H	DNA-binding domain	23.85	15.97
LT_035	yes	splice site	c.672+1C>1	e5+1	NA	57.79	45.45
LT_036	yes	missense	c.584A>G	p.I195T	DNA-binding domain	90.16	81.64
LT_037	yes	nonsense	c.493G>A	p.Q165*	DNA-binding domain	67.61	80.99
LT_038	yes	missense	c.743C>T	p.R248Q	DNA-binding domain	83.19	61.97
LT_039	yes	nonsense	c.592C>A	p.E198*	DNA-binding domain	70.54	41.90
LT_041	yes	missense	c.748G>A	p.P250S	DNA-binding domain	62.39	53.60
LT_041	yes	missense	c.747C>A	p.R249S	DNA-binding domain	62.39	53.28
LT_041	no	intronic	c.96+41_96+34CTCCAGGT>-	e2+41	NA	NA	NA
LT_042	yes	frame shift deletion	c.366_350 CACAGACTTGGCTGTCC>-	p.G117fs	DNA-binding domain	NA	NA
LT_043	yes	frame shift deletion	c.801C>-	p.N268fs	DNA-binding domain	82.26	83.78
LT_043	yes	intronic	c.96+41_96+34 CTCCAGGT>-	e2+41	NA	NA	NA
LT_044	yes	missense	c.488T>G	p.Y163S	DNA-binding domain	41.98	96.08
LT_045	yes	missense	c.638C>A	p.R213L	DNA-binding domain	81.82	77.00
LT_046	yes	frame shift insertion	c.853_852->T	p.E284fs	DNA-binding domain	66.67	17.59

Table 2.2 Somatic TP53 mutations

Somatic TP53 mutations identified in tumor cohort from WES.

Patient ID	Clonal Variant	Variant Type	Nucleotide Position	Amino Acid Change	Primary VAF	Metastatic VAF	ClinVar Clinical Significance
ST_009	yes	missense	c.5344A>C	p.F1782V	72.48	46.46	not provided
ST_015	yes	frame-shift deletion	c.2419	p.A807fs	36.96	42.76	pathogenic
LT_039	no	missense	c.3419C>G	p.S1140T	56.67	NA	uncertain significance
LT_042	yes	frame-shift deletion	c.2284	p.R762fs	92.04	75.77	pathogenic

Table 2.3 BRCA1 Somatic Mutations

All BRCA1 somatic mutations identified in tumor cohort for WES with a VAF>30. Nucleotide position and amino acid changes are based on GRCh37.

Patient ID	Variant Type	Nucleotide Position	Amino Acid Change	VAF	ClinVar Clinical Significance
ST_021	missense	c.4798G>C	p.P1600A	53.77	conflicting interpretations of pathogenicity
ST_027	missense	c.1487C>T	p.R496H	43.96	benign
ST_027	missense	c.455A>G	p.L152P	59.51	conflicting interpretations of pathogenicity
ST_029	missense	c.4931G>C	p.A1644G	51.67	pathogenic/likely pathogenic
LT_031	nonsense	c.2035T>A	p.K679*	57.81	pathogenic
LT_031	missense	c.736A>C	p.L246V	55.29	benign
LT_037	missense	c.5396T>C	p.D1799G	42.86	conflicting interpretations of pathogenicity
LT_044	nonsense	c.2138G>C	p.S713*	37.01	pathogenic

Table 2.4 BRCA1 Germline Mutations

All BRCA1 germline mutation identified in patient cohort from WES with a VAF>30. Nucleotide position and amino acid changes are based on GRCh37.

Patient ID	Variant Type	Nucleotide Position	Amino Acid Change	VAF	ClinVar Clinical Signficiance
LT_039	missense	c.3568C>T	p.R1190W	47.89	benign

Table 2.5 BRCA2 Germline Mutations

All BRCA2 germline mutations identified in patient cohort from WES with a VAF>30. Nucleotide position and amino acid changes are based on GRCh37.

Recurrent gene fusions unique to ST survivors	Number of samples	Number of patients	Number of primary samples	Number of met samples
LINC01176>>ZNRF2	11	10	4	7
ZRANB1>>AL731577.2	10	9	4	6
CCDC32>>CBX3	9	9	6	3
AC060766.7>>SLFN12	9	9	5	4
AL035409.1>>ST6GALNAC5	9	8	5	4
TIAM2>>SCAF8	10	8	6	4
SLC35A1>>AL109930.1/AL354993.1	7	6	3	4
SLC35A1>>AL354993.1	6	5	2	4
AC025178.1>>GOLPH3	5	5	3	2
BMPR1B>>PDLIM5	4	4	3	1
AL049697.1/SLC35A1>>AL109930.1/AL354993.1	3	3	2	1
SLC35A1>>AL109930.1	4	3	1	3
BACH1-IT1>>BACH1	3	3	2	1
FP236383.1>>PDE3A	3	3	2	1
HYDIN>>AC239809.3	3	3	2	1
AL596220.1>>C1orf27	3	3	1	2
PPP1CB>>PLB1	3	2	2	1
COL14A1>>DEPTOR	3	2	1	2
CCDC6>>ANK3	2	2	1	1
SLC35A1>>AL354993.1/AL109930.1	2	2	1	1
TRIO>>ARHGEF3	3	2	2	1
FP236383.1>>COL6A1	2	2	0	2
RNU6-4P>>SNORA20/TCP1	2	2	1	1
ERBB4>>IKZF2	2	2	1	1
TPTE2>>MRPS31P2	2	2	2	0
CATSPERB>>TC2N	2	2	2	0
PSPC1>>ZMYM5	3	2	1	2
RNU6-2>>SNORD10	2	2	2	0
COL4A2>>COL4A1	2	2	1	1
NCALD>>GRHL2	2	2	2	0
LINC00032>>EQTN	2	2	1	1
MECOM>>PHC3	2	2	2	0
RAD9A>>KDM2A	2	2	2	0
SKIL>>PRKCI	2	2	1	1
CYP1B1-AS1>>RMDN2	2	2	2	0

Table 2.6 INTEGRATE Recurrent Gene Fusions

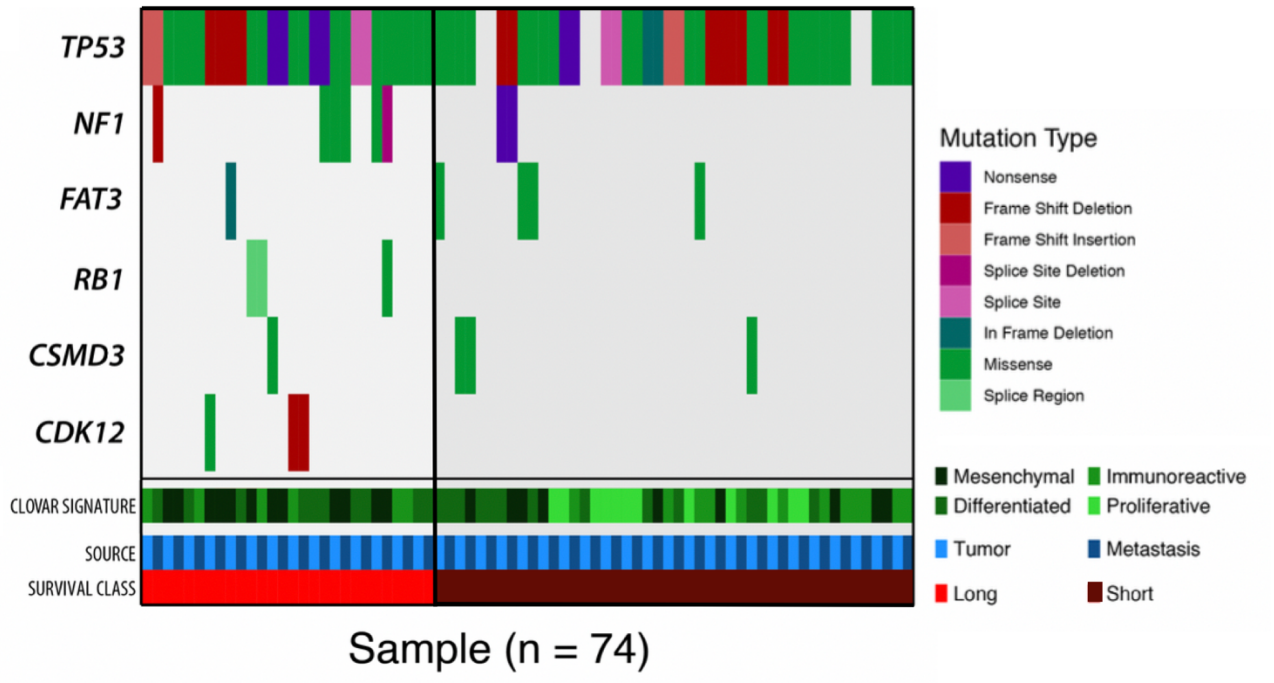
Chart showing the recurrent gene fusions predicted by INTEGRATE for the ST survivor tumor samples. The chart displays a breakdown of number and types of samples each gene fusion was identified among the ST survivors.

Gene fusion (5'>>3')	Patient sample	Encompassing and spanning reads detected in sample	In-Frame prediction
COL1A1>>7SK/RN7SK	ST.015.M.FF	8	out-of-frame
COL1A1>>GSN	ST.001.M.FF	4	in-frame
BROX>>COL1A1	ST.015.M.FF	2	in-frame
MALAT1>>COL1A2	ST.018.M.FFPE	76	in-frame
COL1A2>>NEAT1	ST.029.M.FFPE	40	out-of-frame
C3>>COL1A2	ST.027.P.FF	20	in-frame
RPS11>>COL1A2	ST.018.P.FF	9	in-frame
COL1A2>>PABPC1	ST.001.M.FF	7	in-frame
COL1A2>>PCBP2/AC023509.1	ST.002.M.FF	4	in-frame
AC093752.1>>COL1A2	ST.001.M.FF	3	in-frame
COL1A2>>TBC1D4	ST.018.P.FF	3	out-of-frame
COL1A2>>PCBP2	ST.002.M.FF	2	in-frame
COL1A2>>TBCCD1	ST.001.M.FF	2	in-frame
COL1A2>>YWHAB	ST.029.P.FF	2	in-frame
HM13>>COL1A2	ST.021.P.FF	2	in-frame
MALAT1>>COL3A1	ST.028.M.FFPE	169	in-frame
FN1>>COL3A1	ST.018.P.FF	54	in-frame
COL3A1>>RN7SL2	ST.001.M.FF	52	out-of-frame
COL3A1>>RMRP	ST.021.P.FF	10	out-of-frame
CLU>>COL3A1	ST.029.P.FF	4	in-frame
STX7>>COL3A1	ST.029.P.FF	3	out-of-frame
HAX1>>COL3A1	ST.018.P.FF	3	in-frame
MAP2>>COL3A1	ST.018.P.FF	3	in-frame
COL4A2>>COL4A1	ST.001.M.FF ST.011.P.FF	5 2	in-frame in-frame
RPL15>>COL4A2	ST.030.M.FF	2	out-of-frame
FP236383.1>>COL6A1	ST.023.M.FFPE ST.028.M.FFPE	15 28	out-of-frame
COL6A3>>ARID1B	ST.021.P.FF	3	out-of-frame
COL6A2>>FP236383.3	LT.041.M.FFPE	7	out-of-frame
ORC4>>COL6A3	ST.011.M.FF	2	out-of-frame
COL6A6>>TTLL9	ST.027.P.FF	6	out-of-frame
RPS24>>COL12A1	ST.018.P.FF	4	out-of-frame
COL14A1>>DEPTOR	ST.007.P.FF ST.007.M.FF ST.030.M.FF	49 3 39	in-frame
TFDP2>>COL18A1	ST.013.M.FF	3	in-frame

Table 2.7 Predicted Collagen Gene Fusions

Chart displaying the 33 gene fusions involving collagen genes, the sample the fusion was identified in, the number of encompassing and spanning reads that identified the fusion, and INTEGRATE's prediction if the reads were in- or out-of-frame.

A



B

Sample (n = 74)

Figure 2.1: Mutational Landscape of primary and metastatic tumors from ST and LT survivors

A. Samples are organized as matched tumor pairs, with primary first followed by metastic tumor. Somatic mutations in OV TCGA frequently mutated genes, colors indicate mutational type
B. Status of samples source (primary tumor (light blue), metastatic tumor (dark blue)), survival class (ST (dark red), LT (red)), and CLOVAR signature (mesenchymal, differentiated, immunoreactive, and proliferative) (Green shades)

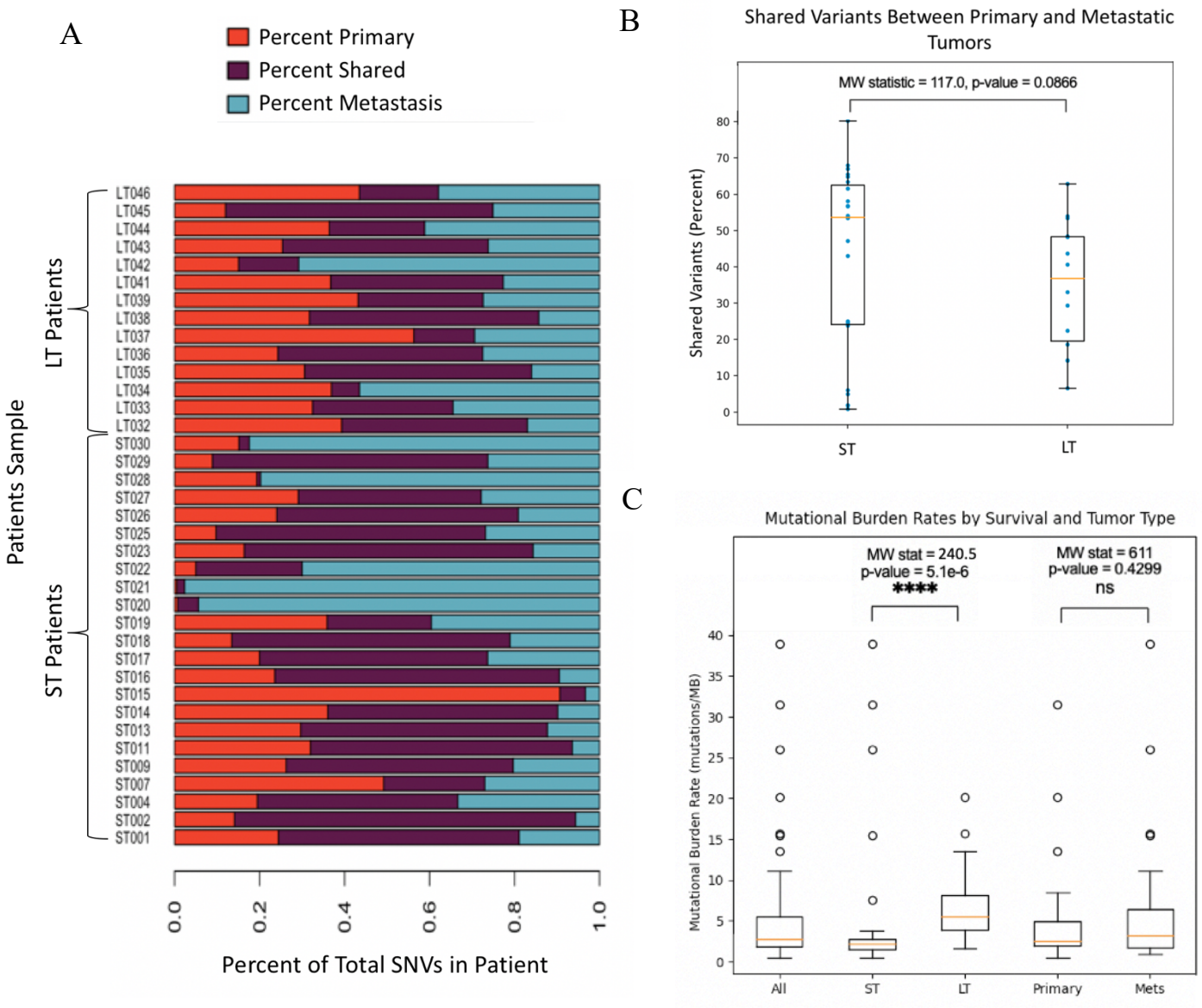
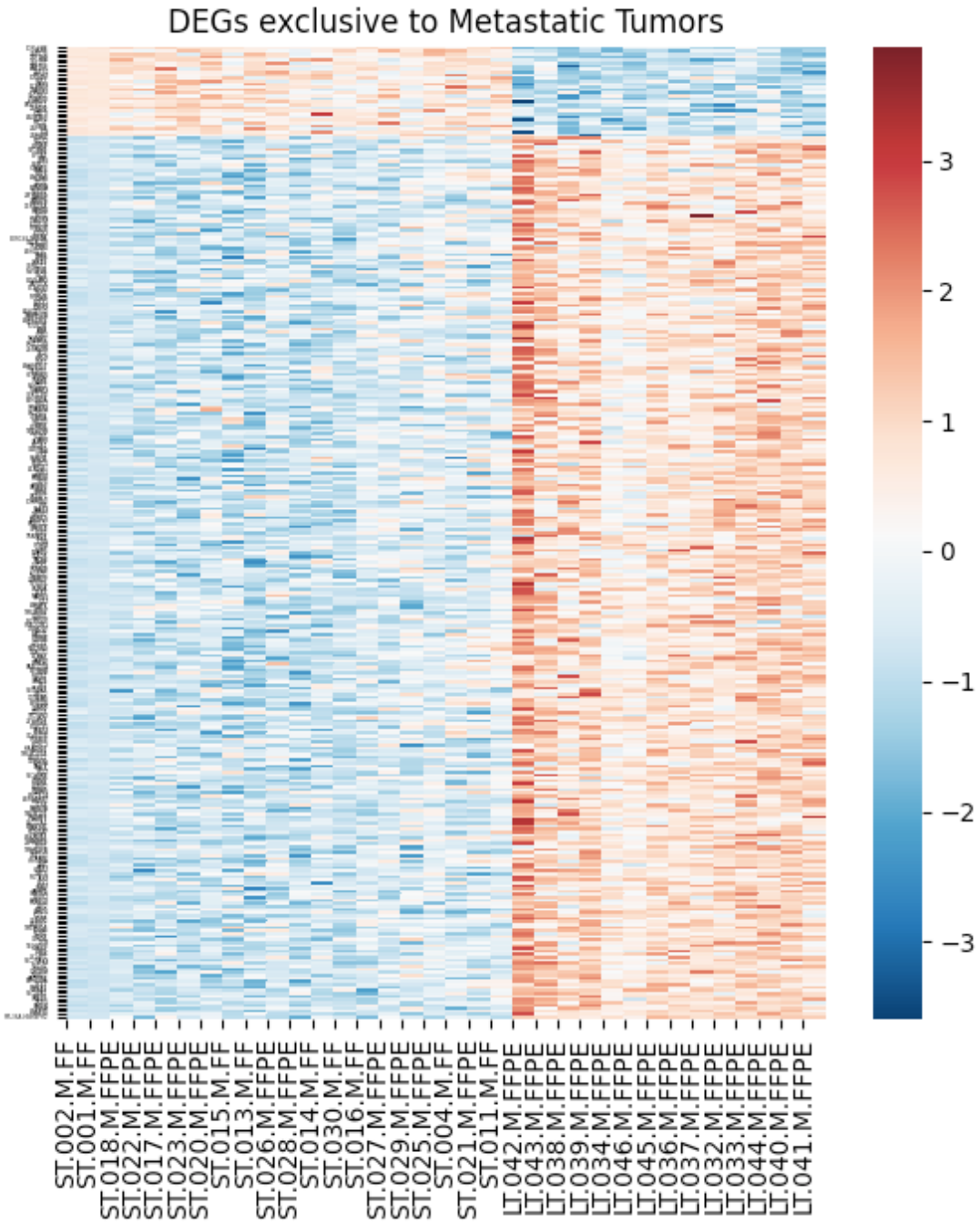

A
Percent Primary
Percent Shared
Percent Metastasis
Patients Sample
LT Patients
ST Patients
LT046
LT045
LT044
LT043
LT042
LT041
LT039
LT038
LT037
LT036
LT035
LT034
LT033
LT032
ST030
ST029
ST028
ST027
ST026
ST025
ST023
ST022
ST021
ST020
ST019
ST018
ST017
ST016
ST015
ST014
ST013
ST011
ST009
ST007
ST004
ST002
ST001
0.0
0.2
0.4
0.6
0.8
1.0
Percent of Total SNVs in Patient
B
Shared Variants Between Primary and Metastatic Tumors
MW statistic = 117.0, p-value = 0.0866
Shared Variants (Percent)
0
10
20
30
40
50
60
70
80
ST
LT
C
Mutational Burden Rates by Survival and Tumor Type
MW stat = 240.5
p-value = 5.1e-6
****
MW stat = 611
p-value = 0.4299
ns
Mutational Burden Rate (mutations/MB)
0
5
10
15
20
25
30
35
40
All
ST
LT
Primary
Mets

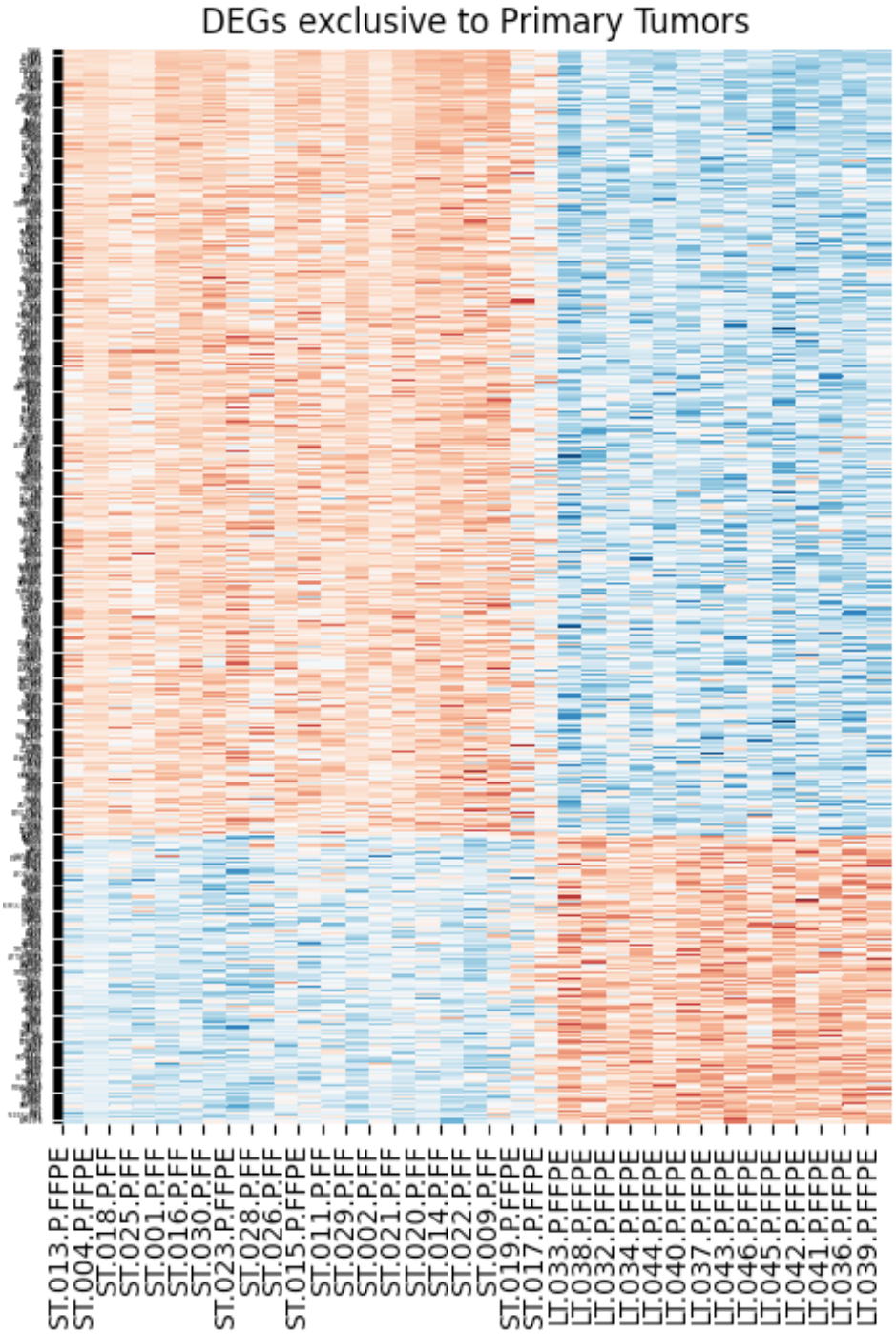
Figure 2.2: Shared variants between primary and metastatic tumors.

A. Percentage of somatic mutations unique to primary (Red) and metastatic tumor (Blue) and shared (Purple) between samples.
B. Comparison of percentage of shared variants among primary and metastatic tumor between ST and LT survivors. (ST survivors n= 23 LT survivors n= 14)
C. Boxplot displaying the mutational burden rates (mutations/MB) for all tumors and subsets. (All n= 74, ST survivors n= 23, LT survivors n= 17, primary tumors n= 37, metastatic tumors n= 37)

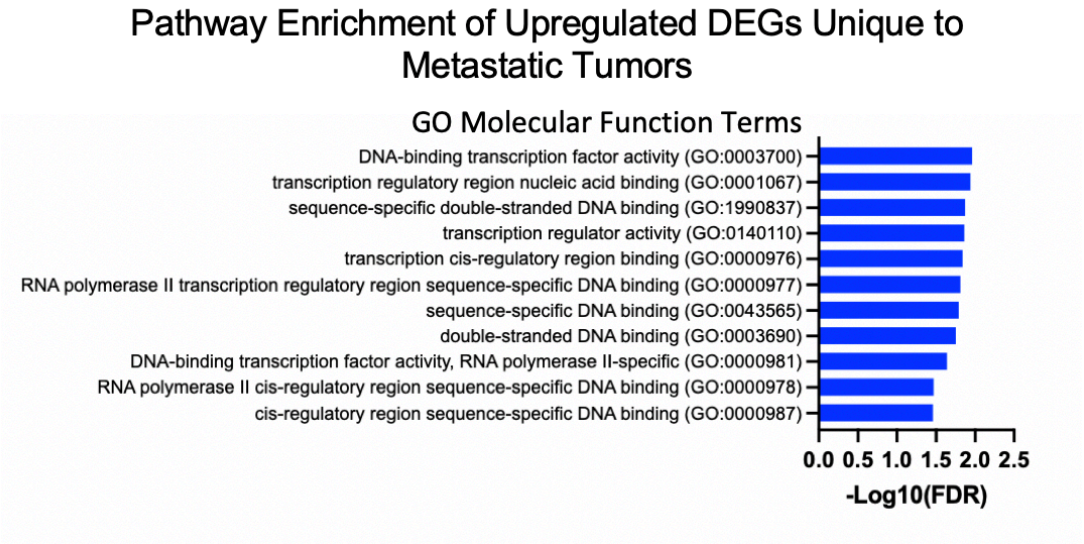
A


DEGs exclusive to Metastatic Tumors
3
2
1
0
−1
−2
−3
ST.002.M.FF
ST.001.M.FF
ST.018.M.FFPE
ST.022.M.FFPE
ST.017.M.FFPE
ST.023.M.FFPE
ST.020.M.FFPE
ST.015.M.FF
ST.013.M.FF
ST.026.M.FFPE
ST.028.M.FFPE
ST.014.M.FF
ST.030.M.FF
ST.016.M.FF
ST.027.M.FFPE
ST.029.M.FFPE
ST.025.M.FFPE
ST.004.M.FF
ST.021.M.FFPE
ST.011.M.FF
LT.042.M.FFPE
LT.043.M.FFPE
LT.038.M.FFPE
LT.039.M.FFPE
LT.034.M.FFPE
LT.046.M.FFPE
LT.045.M.FFPE
LT.036.M.FFPE
LT.037.M.FFPE
LT.032.M.FFPE
LT.033.M.FFPE
LT.044.M.FFPE
LT.040.M.FFPE
LT.041.M.FFPE

B


DEGs exclusive to Primary Tumors
ST.013.P.FFPE
ST.004.P.FFPE
ST.018.P.FF
ST.025.P.FF
ST.001.P.FF
ST.016.P.FF
ST.030.P.FF
ST.023.P.FFPE
ST.028.P.FF
ST.026.P.FF
ST.015.P.FFPE
ST.011.P.FF
ST.029.P.FF
ST.002.P.FF
ST.021.P.FF
ST.020.P.FF
ST.014.P.FF
ST.022.P.FF
ST.009.P.FF
ST.019.P.FFPE
ST.017.P.FFPE
LT.033.P.FFPE
LT.038.P.FFPE
LT.032.P.FFPE
LT.034.P.FFPE
LT.044.P.FFPE
LT.040.P.FFPE
LT.037.P.FFPE
LT.043.P.FFPE
LT.046.P.FFPE
LT.045.P.FFPE
LT.042.P.FFPE
LT.041.P.FFPE
LT.036.P.FFPE
LT.039.P.FFPE

C


Pathway Enrichment of Upregulated DEGs Unique to Metastatic Tumors
GO Molecular Function Terms
DNA-binding transcription factor activity (GO:0003700)
transcription regulatory region nucleic acid binding (GO:0001067)
sequence-specific double-stranded DNA binding (GO:1990837)
transcription regulator activity (GO:0140110)
transcription cis-regulatory region binding (GO:0000976)
RNA polymerase II transcription regulatory region sequence-specific DNA binding (GO:0000977)
sequence-specific DNA binding (GO:0043565)
double-stranded DNA binding (GO:0003690)
DNA-binding transcription factor activity, RNA polymerase II-specific (GO:0000981)
RNA polymerase II cis-regulatory region sequence-specific DNA binding (GO:0000978)
cis-regulatory region sequence-specific DNA binding (GO:0000987)
0.0 0.5 1.0 1.5 2.0 2.5
-Log10(FDR)

D

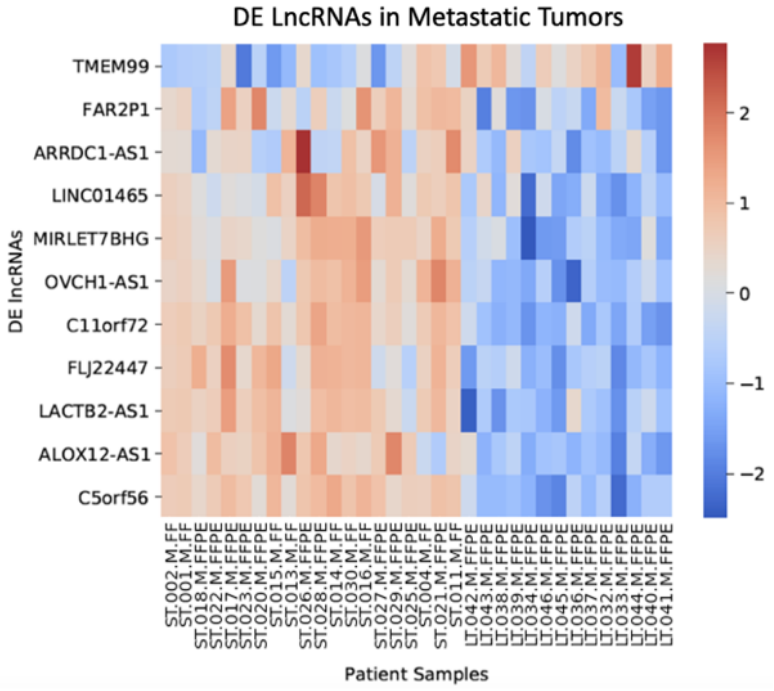

DE LncRNAs in Metastatic Tumors
TMEM99
FAR2P1
ARRDC1-AS1
LINC01465
MIRLET7BHG
OVCH1-AS1
C11orf72
FLJ22447
LACTB2-AS1
ALOX12-AS1
C5orf56
DE lncRNAs
2
1
0
−1
−2
Patient Samples

Figure 2.3 DEGs and lncRNAs in tumors from ST and LT survivors

A. Heatmap displaying the 325 significant protein-coding DEGs between ST (left) and LT survivors (right) unique to metastatic tumors in our patient cohort. (ST survivors n= 20, LT survivors n= 14) Red indicates upregulated DEGs while blue indicates downregulated DEGs with FDR <0.01.
B. Heatmap displaying the 502 significant protein-coding DEGs between ST (left) and LT survivors (right) unique to primary tumors in our patient cohort. (ST survivors n= 21, LT survivors n= 14) Red indicates upregulated DEGs while blue indicates downregulated DEGs with FDR <0.01.
C. Barplot displaying the GO Molecular Function Terms that are statistically overrepresented in the 30 upregulated DEGs unique to Metastatic tumors. Plot displays their -Log10(FDR) value for each GO term.
D. Heatmap displaying 11 significantly (FDR <0.01) differentially expressed lncRNAs between the ST (left) and LT (right) survivors in metastatic tumors. (ST survivors n= 20, LT survivors n= 14)

A

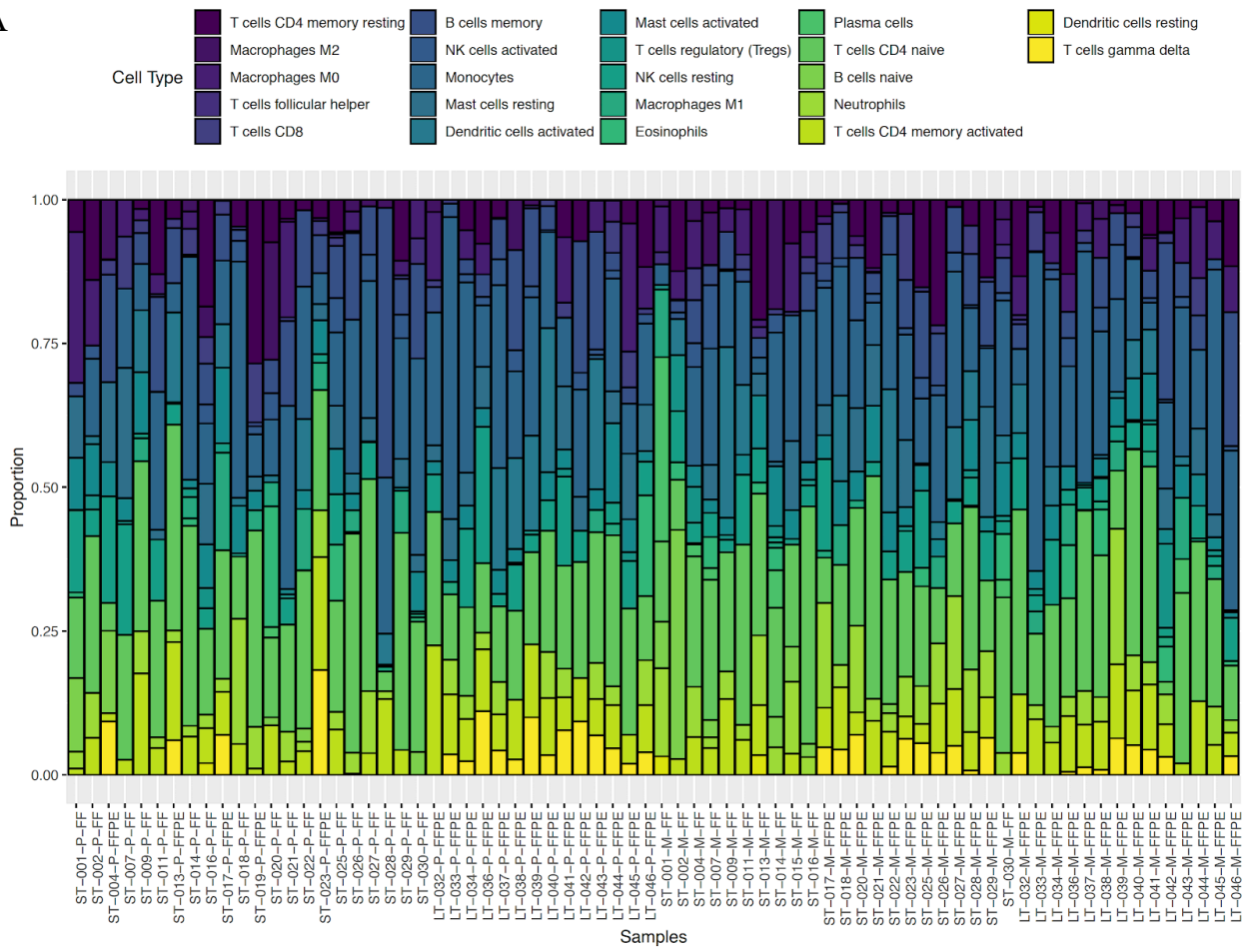

Cell Type
T cells CD4 memory resting
Macrophages M2
Macrophages M0
T cells follicular helper
T cells CD8
B cells memory
NK cells activated
Monocytes
Mast cells resting
Dendritic cells activated
Mast cells activated
T cells regulatory (Tregs)
NK cells resting
Macrophages M1
Eosinophils
Plasma cells
T cells CD4 naive
B cells naive
Neutrophils
T cells CD4 memory activated
Dendritic cells resting
T cells gamma delta
Proportion
1.00
0.75
0.50
0.25
0.00
Samples

Figure 2.4: Cibersort Immune Cell Fractions

A. Stacked bar plot of the proportion of the 22 immune cell types expressed in each tumor sample. Samples are organized by tumor type, then by survival. (from left to right: ST Primary n= 22, LT Primary n= 14, ST Metastatic n= 22, LT Metastatic n= 14) Annotation colors are shown in legend above barplot.

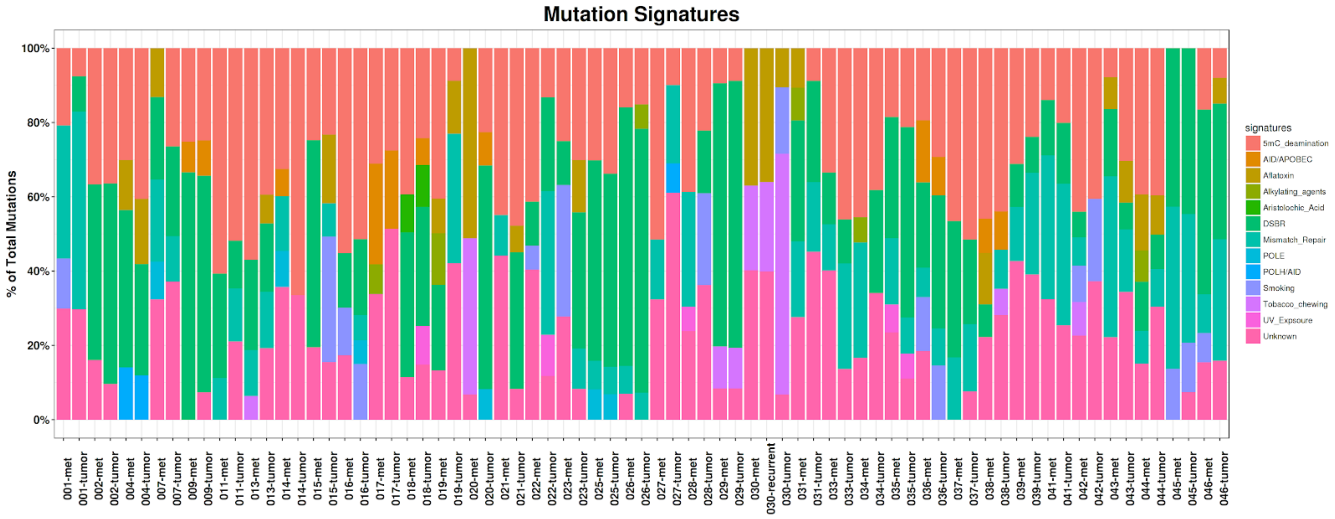
2.8 Supplementary Figures and Tables

Exome Sequencing Coverage			
Patient ID	**Normal**	**Primary**	**Metastatic**
ST_001	148	244	251
ST_002	136	258	309
ST_004	79	388	225
ST_007	139	231	274
ST_009	97	325	267
ST_011	113	179	180
ST_013	131	365	234
ST_014	101	248	246
ST_015	259	309	559
ST_016	114	266	271
ST_017	107	309	272
ST_018	146	235	316
ST_019	157	279	295
ST_020	131	279	390
ST_021	145	231	306
ST_022	127	264	350
ST_023	142	384	309
ST_025	150	272	356
ST_026	112	250	266
ST_027	237	258	266
ST_028	117	279	339
ST_029	131	230	279
ST_030	65	331	139
LT_031	133	244	442
LT_033	226	300	316
LT_034	169	428	248
LT_035	142	278	355
LT_036	210	299	332
LT_037	213	263	336
LT_038	120	344	288
LT_039	175	389	218
LT_041	167	275	319
LT_042	186	338	181
LT_043	90	439	364
LT_044	241	331	278
LT_045	138	396	344
LT_046	187	283	331

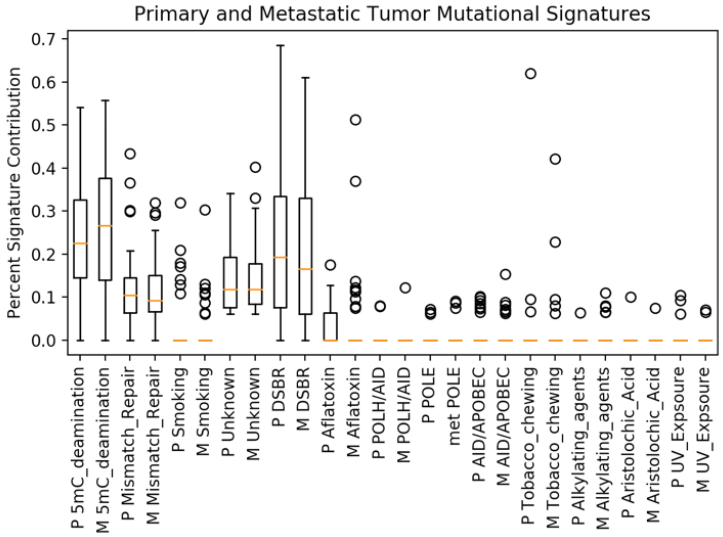
Supplementary Table 2.8: Tumor Sample WES coverage

Exome sequencing coverage for each patient's normal, primary, and metastatic tumor. Coverage is expressed as the number of times a base would be covered by sequencing reads.

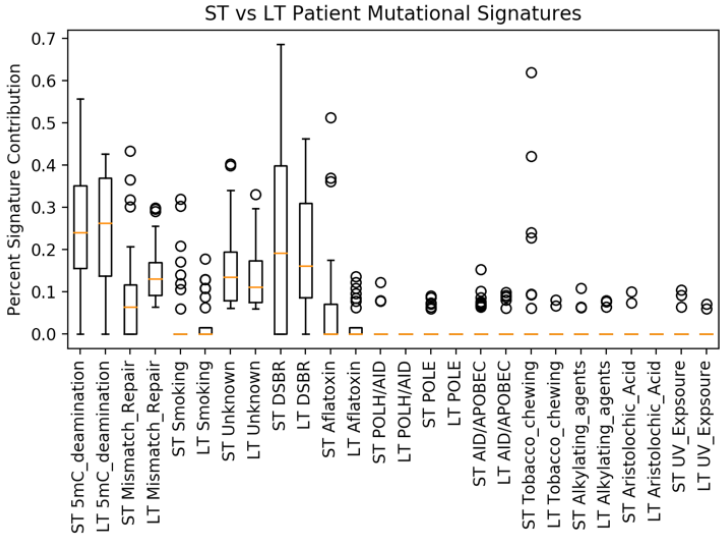
A


Mutation Signatures
% of Total Mutations
100%
80%
60%
40%
20%
0%
signatures
5mC_deamination
AID/APOBEC
Aflatoxin
Alkylating_agents
Aristolochic_Acid
DSBR
Mismatch_Repair
POLE
POLH/AID
Smoking
Tobacco_chewing
UV_Expsoure
Unknown

B


Primary and Metastatic Tumor Mutational Signatures
Percent Signature Contribution
P 5mC_deamination
M 5mC_deamination
P Mismatch_Repair
M Mismatch_Repair
P Smoking
M Smoking
P Unknown
M Unknown
P DSBR
M DSBR
P Aflatoxin
M Aflatoxin
P POLH/AID
M POLH/AID
P POLE
met POLE
P AID/APOBEC
M AID/APOBEC
P Tobacco_chewing
M Tobacco_chewing
P Alkylating_agents
M Alkylating_agents
P Aristolochic_Acid
M Aristolochic_Acid
P UV_Expsoure
M UV_Expsoure

C

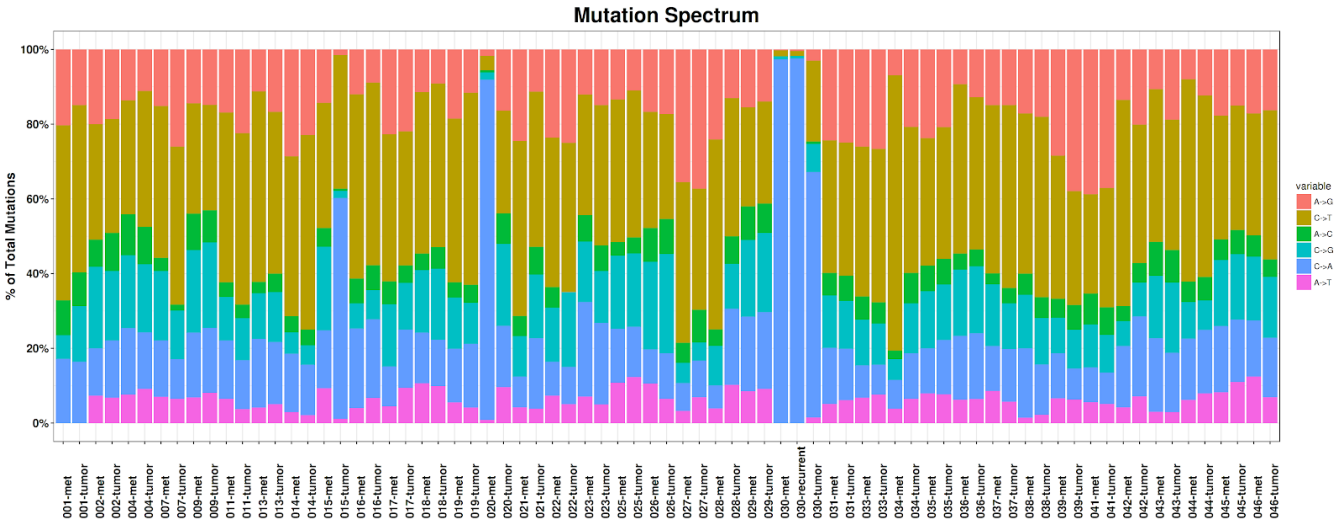

ST vs LT Patient Mutational Signatures
Percent Signature Contribution
ST 5mC_deamination
LT 5mC_deamination
ST Mismatch_Repair
LT Mismatch_Repair
ST Smoking
LT Smoking
ST Unknown
LT Unknown
ST DSBR
LT DSBR
ST Aflatoxin
LT Aflatoxin
ST POLH/AID
LT POLH/AID
ST POLE
LT POLE
ST AID/APOBEC
LT AID/APOBEC
ST Tobacco_chewing
LT Tobacco_chewing
ST Alkylating_agents
LT Alkylating_agents
ST Aristolochic_Acid
LT Aristolochic_Acid
ST UV_Expsoure
LT UV_Expsoure

Supplementary Figure 2.5: Mutational Signatures

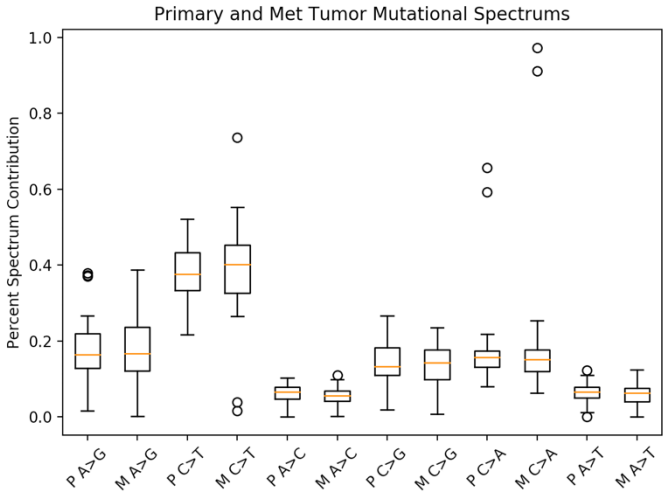
A. Stacked bar plot displaying the percent of total mutations that fit into the mutational signatures shown on the right. Each patient is ordered by patient number and their metastatic and primary tumor with ST survivors on the left (001-met – 030 tumor) and LT survivors on the right (031-met – 046-tumor).

B-C. Box and whisker plots showing the percentage of total mutations for each signature, grouped by the patients tumor as a ST or LT survivor (B) or as a primary (P) and metastatic tumor (M) (C)

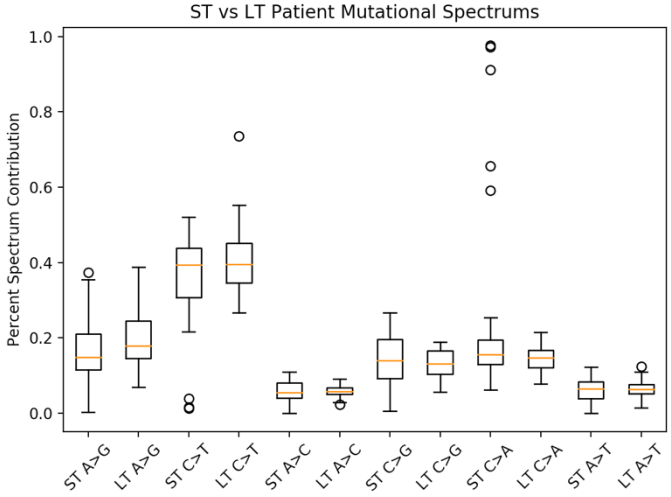
A


Mutation Spectrum
% of Total Mutations
100%
80%
60%
40%
20%
0%
variable
A->G
C->T
A->C
C->G
C->A
A->T

B


Primary and Met Tumor Mutational Spectrums
Percent Spectrum Contribution
1.0
0.8
0.6
0.4
0.2
0.0
P A>G
M A>G
P C>T
M C>T
P A>C
M A>C
P C>G
M C>G
P C>A
M C>A
P A>T
M A>T

C

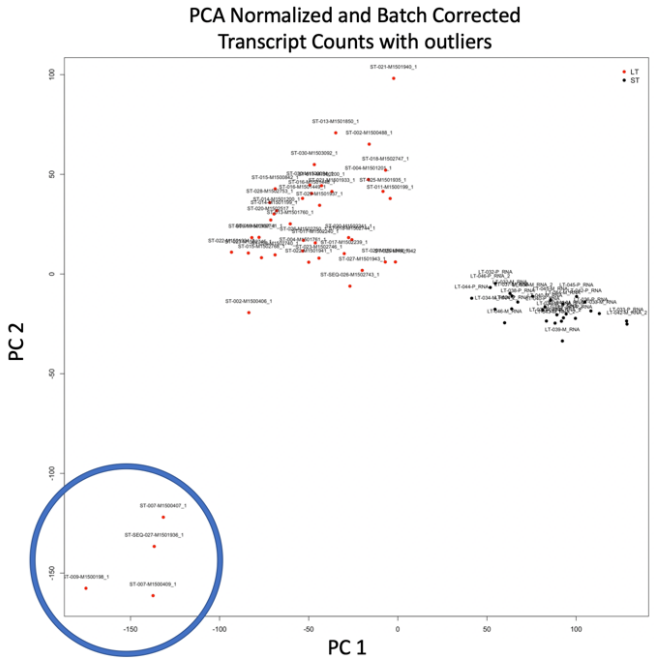

ST vs LT Patient Mutational Spectrums
Percent Spectrum Contribution
1.0
0.8
0.6
0.4
0.2
0.0
ST A>G
LT A>G
ST C>T
LT C>T
ST A>C
LT A>C
ST C>G
LT C>G
ST C>A
LT C>A
ST A>T
LT A>T

Supplementary Figure 2.6: Mutational Spectrum

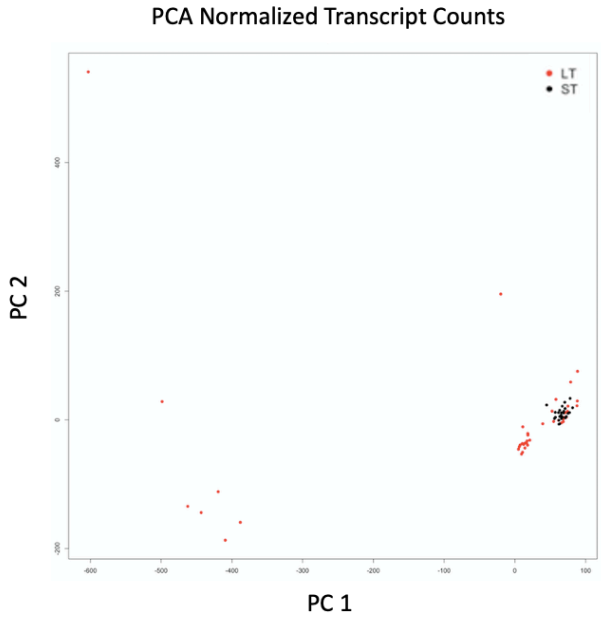
A. Stacked bar plot displaying the percent of total mutations that were nucleotide transitions or transversions, according to the key on the right. Each patient is ordered by patient number and their metastatic and primary tumor with ST survivors on the left (001-met – 030 tumor) and LT survivors on the right (031-met – 046-tumor).

B-C. Box and whisker plots showing the percentage of total mutations for each signature, grouped by the patients tumor as a ST or LT survivor (B) or as a primary (P) and metastatic tumor (M) (C)

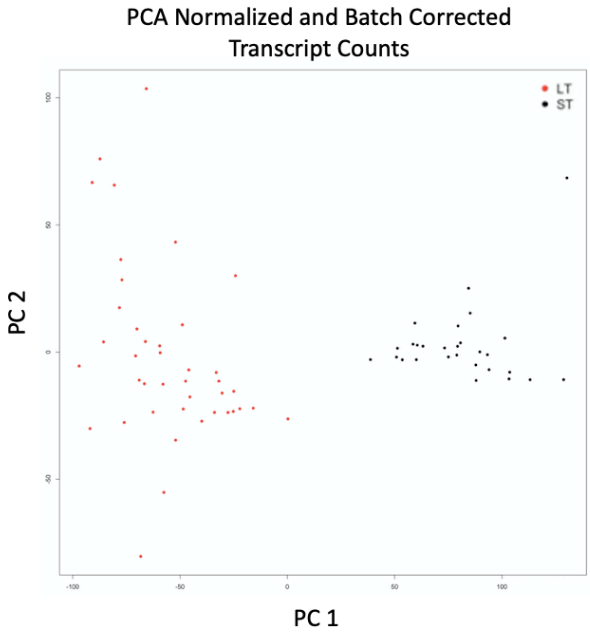
A


PCA Normalized and Batch Corrected
Transcript Counts with outliers
LT
ST
PC 2
PC 1

B


PCA Normalized Transcript Counts
LT
ST
PC 2
PC 1

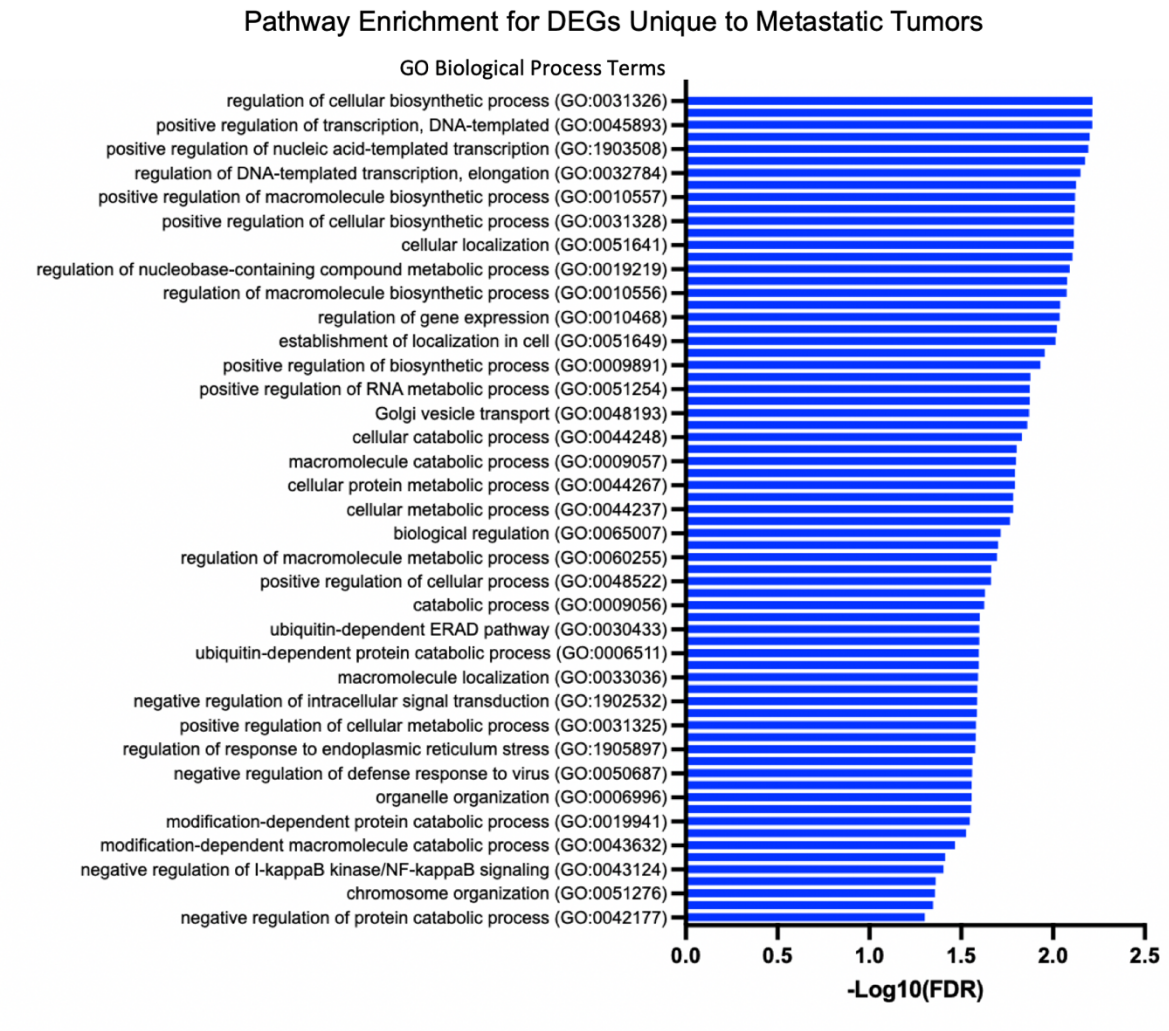
C


PCA Normalized and Batch Corrected
Transcript Counts
LT
ST
PC 2
PC 1

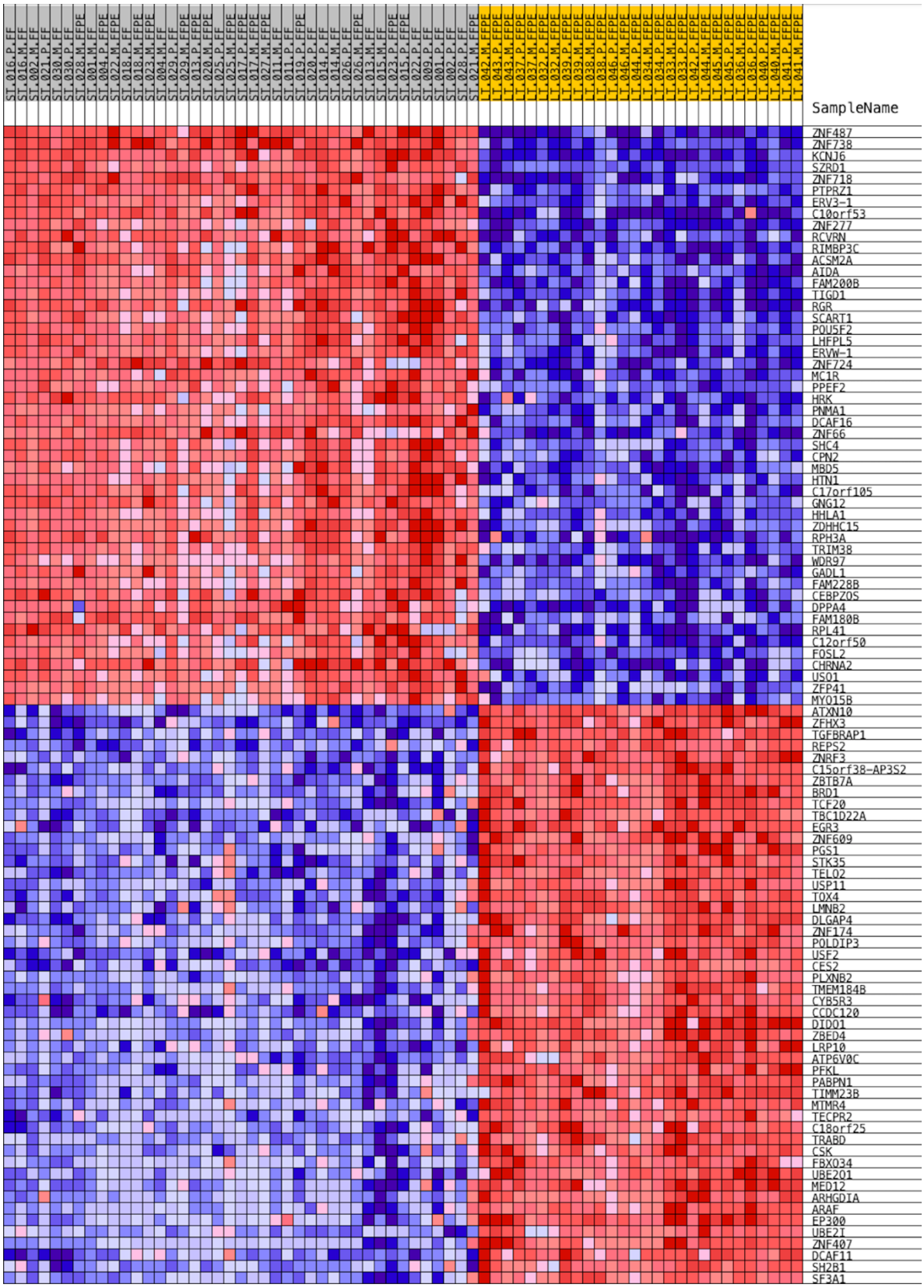
Supplementary Figure 2.7: PCA Plots of Log CPM counts in Differential Expression Analysis

A. PCA of cohort samples RNA-sequencing transcript counts with all tumor samples after normalization and batch effect correction with SVA. The 4 samples circled in the plot were removed as outliers.
B. PCA of cohort samples RNA-sequencing transcript counts after removal of the 4 outlier samples and normalization.
C. PCA of cohort samples RNA-sequencing transcript counts after removal of the 4 outlier samples, normalization, and batch effect correction with SVA.

A


Pathway Enrichment for DEGs Unique to Metastatic Tumors
GO Biological Process Terms
regulation of cellular biosynthetic process (GO:0031326)
positive regulation of transcription, DNA-templated (GO:0045893)
positive regulation of nucleic acid-templated transcription (GO:1903508)
regulation of DNA-templated transcription, elongation (GO:0032784)
positive regulation of macromolecule biosynthetic process (GO:0010557)
positive regulation of cellular biosynthetic process (GO:0031328)
cellular localization (GO:0051641)
regulation of nucleobase-containing compound metabolic process (GO:0019219)
regulation of macromolecule biosynthetic process (GO:0010556)
regulation of gene expression (GO:0010468)
establishment of localization in cell (GO:0051649)
positive regulation of biosynthetic process (GO:0009891)
positive regulation of RNA metabolic process (GO:0051254)
Golgi vesicle transport (GO:0048193)
cellular catabolic process (GO:0044248)
macromolecule catabolic process (GO:0009057)
cellular protein metabolic process (GO:0044267)
cellular metabolic process (GO:0044237)
biological regulation (GO:0065007)
regulation of macromolecule metabolic process (GO:0060255)
positive regulation of cellular process (GO:0048522)
catabolic process (GO:0009056)
ubiquitin-dependent ERAD pathway (GO:0030433)
ubiquitin-dependent protein catabolic process (GO:0006511)
macromolecule localization (GO:0033036)
negative regulation of intracellular signal transduction (GO:1902532)
positive regulation of cellular metabolic process (GO:0031325)
regulation of response to endoplasmic reticulum stress (GO:1905897)
negative regulation of defense response to virus (GO:0050687)
organelle organization (GO:0006996)
modification-dependent protein catabolic process (GO:0019941)
modification-dependent macromolecule catabolic process (GO:0043632)
negative regulation of I-kappaB kinase/NF-kappaB signaling (GO:0043124)
chromosome organization (GO:0051276)
negative regulation of protein catabolic process (GO:0042177)
0.0
0.5
1.0
1.5
2.0
2.5
-Log10(FDR)

B


ST.016.P.FF
ST.016.M.FF
ST.002.M.FF
ST.021.P.FF
ST.030.M.FF
ST.030.P.FF
ST.028.M.FFPE
ST.001.M.FF
ST.004.P.FFPE
ST.022.M.FFPE
ST.018.P.FF
ST.018.M.FFPE
ST.023.M.FFPE
ST.004.M.FF
ST.029.P.FF
ST.029.M.FFPE
ST.013.P.FFPE
ST.020.M.FFPE
ST.025.P.FF
ST.025.M.FFPE
ST.017.P.FFPE
ST.017.M.FFPE
ST.027.M.FFPE
ST.011.M.FF
ST.011.P.FF
ST.019.P.FFPE
ST.020.P.FF
ST.014.P.FF
ST.014.M.FF
ST.026.P.FF
ST.026.M.FFPE
ST.013.M.FF
ST.015.M.FF
ST.023.P.FFPE
ST.015.P.FFPE
ST.022.P.FF
ST.009.P.FF
ST.001.P.FF
ST.002.P.FF
ST.028.P.FF
ST.021.M.FFPE
LT.042.M.FFPE
LT.043.P.FFPE
LT.043.M.FFPE
LT.037.P.FFPE
LT.037.M.FFPE
LT.032.P.FFPE
LT.032.M.FFPE
LT.039.P.FFPE
LT.039.M.FFPE
LT.038.M.FFPE
LT.038.P.FFPE
LT.046.P.FFPE
LT.046.M.FFPE
LT.044.P.FFPE
LT.034.M.FFPE
LT.034.P.FFPE
LT.033.M.FFPE
LT.033.P.FFPE
LT.042.P.FFPE
LT.044.M.FFPE
LT.045.M.FFPE
LT.045.P.FFPE
LT.036.M.FFPE
LT.036.P.FFPE
LT.040.P.FFPE
LT.040.M.FFPE
LT.041.P.FFPE
LT.041.M.FFPE
SampleName
ZNF487
ZNF738
KCNJ6
SZRD1
ZNF718
PTPRZ1
ERV3-1
C10orf53
ZNF277
RCVRN
RIMBP3C
ACSM2A
AIDA
FAM200B
TIGD1
RGR
SCART1
POU5F2
LHFPL5
ERVW-1
ZNF724
MC1R
PPEF2
HRK
PNMA1
DCAF16
ZNF66
SHC4
CPN2
MBD5
HTN1
C17orf105
GNG12
HHLA1
ZDHHC15
RPH3A
TRIM38
WDR97
GADL1
FAM228B
CEBPZOS
DPPA4
FAM180B
RPL41
C12orf50
FOSL2
CHRNA2
USO1
ZFP41
MYO15B
ATXN10
ZFHX3
TGFBRAP1
REPS2
ZNRF3
C15orf38-AP3S2
ZBTB7A
BRD1
TCF20
TBC1D22A
EGR3
ZNF609
PGS1
STK35
TELO2
USP11
TOX4
LMNB2
DLGAP4
ZNF174
POLDIP3
USF2
CES2
PLXNB2
TMEM184B
CYB5R3
CCDC120
DIDO1
ZBED4
LRP10
ATP6V0C
PFKL
PABPN1
TIMM23B
MTMR4
TECPR2
C18orf25
TRABD
CSK
FBXO34
UBE2O1
MED12
ARHGDIA
ARAF
EP300
UBE2I
ZNF407
DCAF11
SH2B1
SF3A1

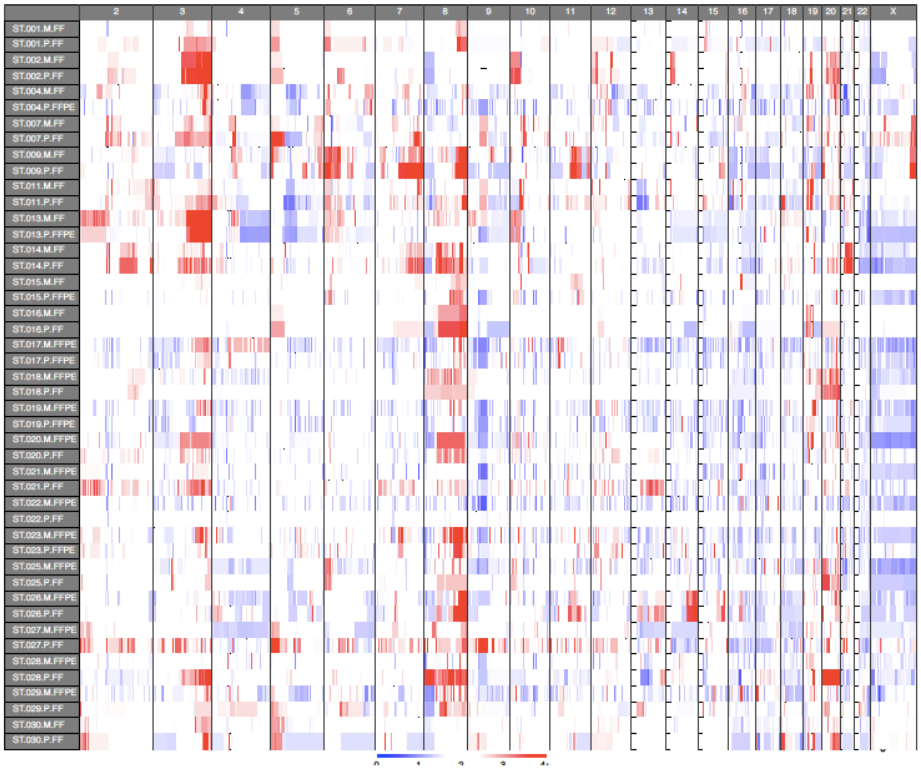
C

Category	Term	RT	Genes	Count	%	P-Value	Fold Enrichment	Benjamini	FDR	Fisher Exact
KEGG_PATHWAY	Adherens junction	RT		6	1.9	5.2E-3	5.3	6.2E-1	6.2E-1	8.9E-4
KEGG_PATHWAY	Protein processing in endoplasmic reticulum	RT		9	2.8	5.6E-3	3.3	6.2E-1	6.2E-1	1.6E-3
KEGG_PATHWAY	SNARE interactions in vesicular transport	RT		4	1.3	1.5E-2	7.6	1.0E0	1.0E0	1.8E-3
KEGG_PATHWAY	Alcoholic liver disease	RT		7	2.2	2.5E-2	3.1	1.0E0	1.0E0	7.4E-3
KEGG_PATHWAY	mTOR signaling pathway	RT		7	2.2	3.7E-2	2.8	1.0E0	1.0E0	1.2E-2
KEGG_PATHWAY	mRNA surveillance pathway	RT		5	1.6	6.7E-2	3.2	1.0E0	1.0E0	1.9E-2
KEGG_PATHWAY	Phospholipase D signaling pathway	RT		6	1.9	8.5E-2	2.5	1.0E0	1.0E0	3.1E-2
KEGG_PATHWAY	Insulin resistance	RT		5	1.6	9.1E-2	2.9	1.0E0	1.0E0	2.9E-2
KEGG_PATHWAY	MAPK signaling pathway	RT		9	2.8	9.4E-2	1.9	1.0E0	1.0E0	4.5E-2

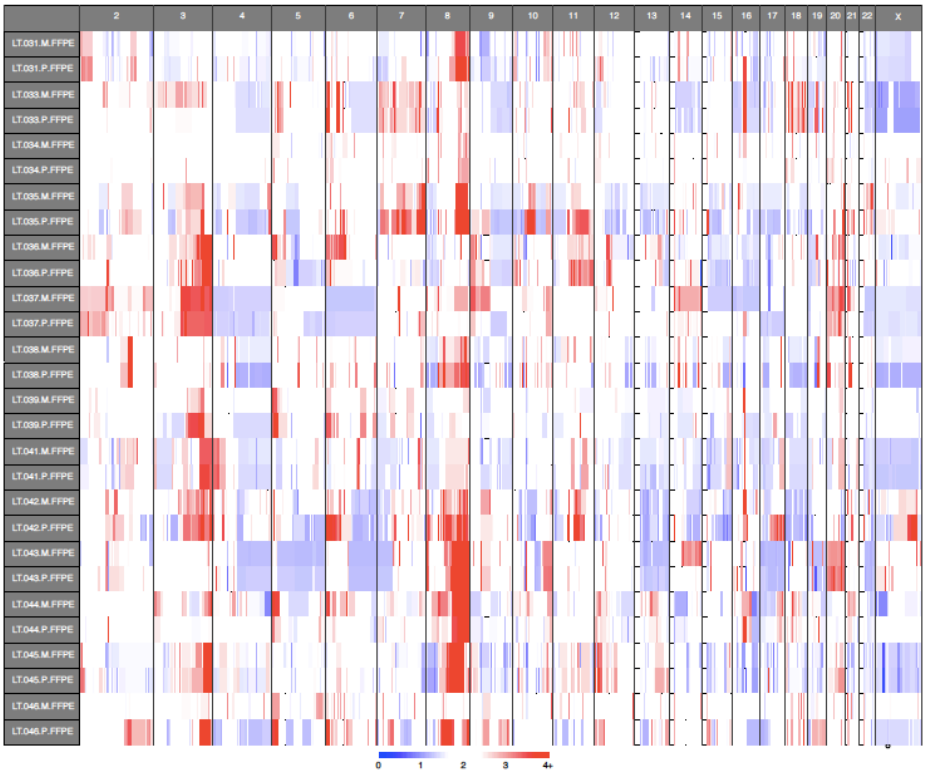
Supplementary Figure 2.8: Pathway Enrichment and DEGs

A. Barplot displaying the -Log10(FDR) values for each Biological Processes GO term that was statistically overrepresented among the DEGs unique to metastatic tumors
B. Heatmap displaying the top 50 upregulated and top 50 downregulated DEGs between the all ST and LT survivors.
C. KEGG Pathways enriched in DEGs as identified by DAVID

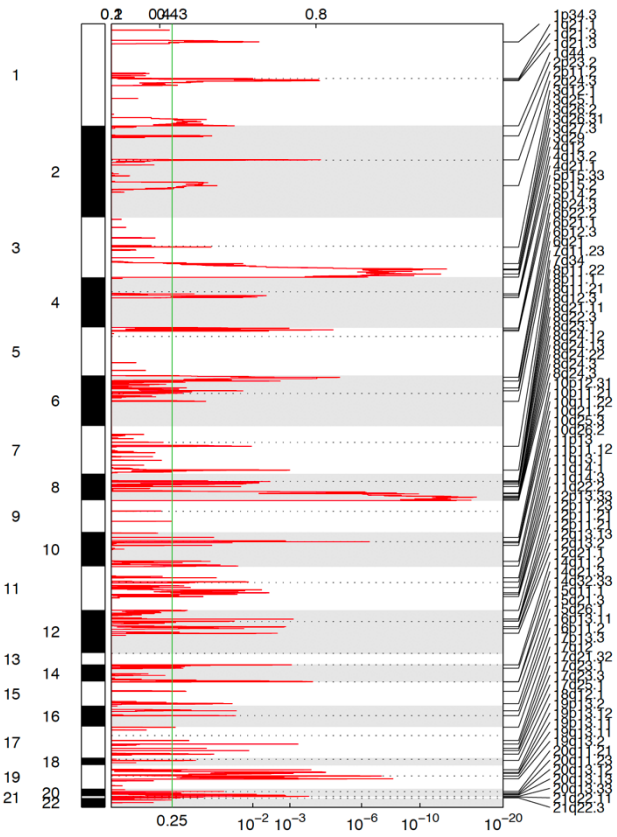
A


2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X
ST.001.M.FF
ST.001.P.FF
ST.002.M.FF
ST.002.P.FF
ST.004.M.FF
ST.004.P.FFPE
ST.007.M.FF
ST.007.P.FF
ST.009.M.FF
ST.009.P.FF
ST.011.M.FF
ST.011.P.FF
ST.013.M.FF
ST.013.P.FFPE
ST.014.M.FF
ST.014.P.FF
ST.015.M.FF
ST.015.P.FFPE
ST.016.M.FF
ST.016.P.FF
ST.017.M.FFPE
ST.017.P.FFPE
ST.018.M.FFPE
ST.018.P.FF
ST.019.M.FFPE
ST.019.P.FFPE
ST.020.M.FFPE
ST.020.P.FF
ST.021.M.FFPE
ST.021.P.FF
ST.022.M.FFPE
ST.022.P.FF
ST.023.M.FFPE
ST.023.P.FFPE
ST.025.M.FFPE
ST.025.P.FF
ST.026.M.FFPE
ST.026.P.FF
ST.027.M.FFPE
ST.027.P.FF
ST.028.M.FFPE
ST.028.P.FF
ST.029.M.FFPE
ST.029.P.FF
ST.030.M.FF
ST.030.P.FF

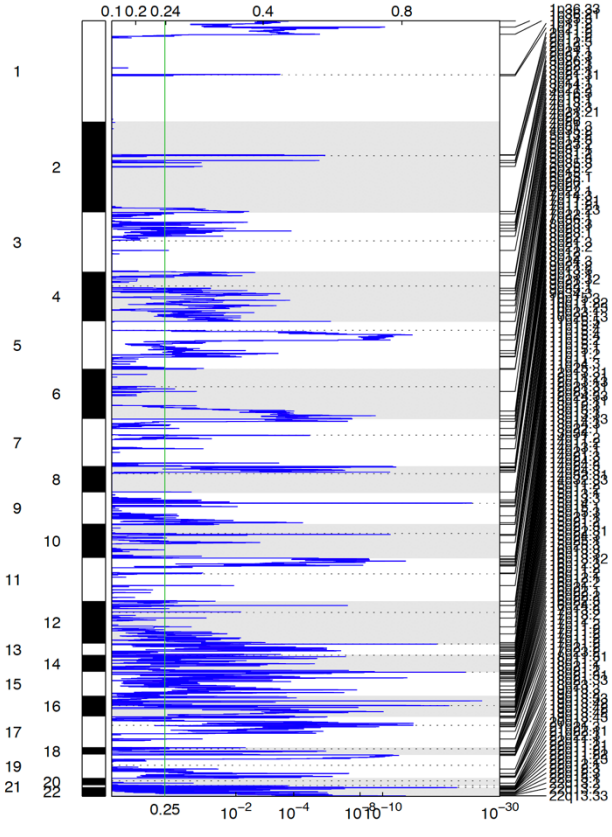
B

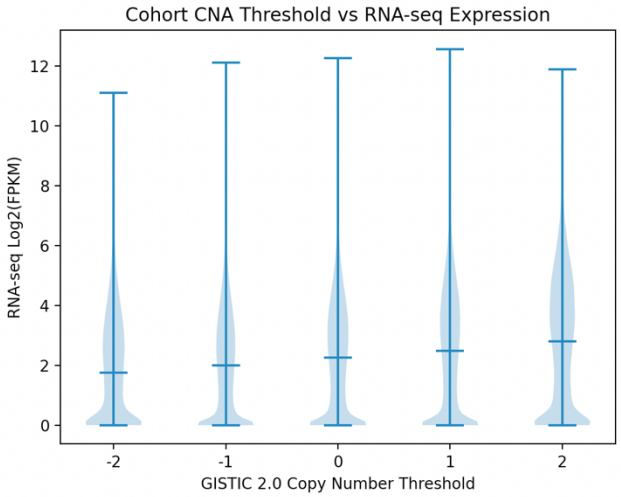

2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X
LT.031.M.FFPE
LT.031.P.FFPE
LT.033.M.FFPE
LT.033.P.FFPE
LT.034.M.FFPE
LT.034.P.FFPE
LT.035.M.FFPE
LT.035.P.FFPE
LT.036.M.FFPE
LT.036.P.FFPE
LT.037.M.FFPE
LT.037.P.FFPE
LT.038.M.FFPE
LT.038.P.FFPE
LT.039.M.FFPE
LT.039.P.FFPE
LT.041.M.FFPE
LT.041.P.FFPE
LT.042.M.FFPE
LT.042.P.FFPE
LT.043.M.FFPE
LT.043.P.FFPE
LT.044.M.FFPE
LT.044.P.FFPE
LT.045.M.FFPE
LT.045.P.FFPE
LT.046.M.FFPE
LT.046.P.FFPE
0 1 2 3 4+

C



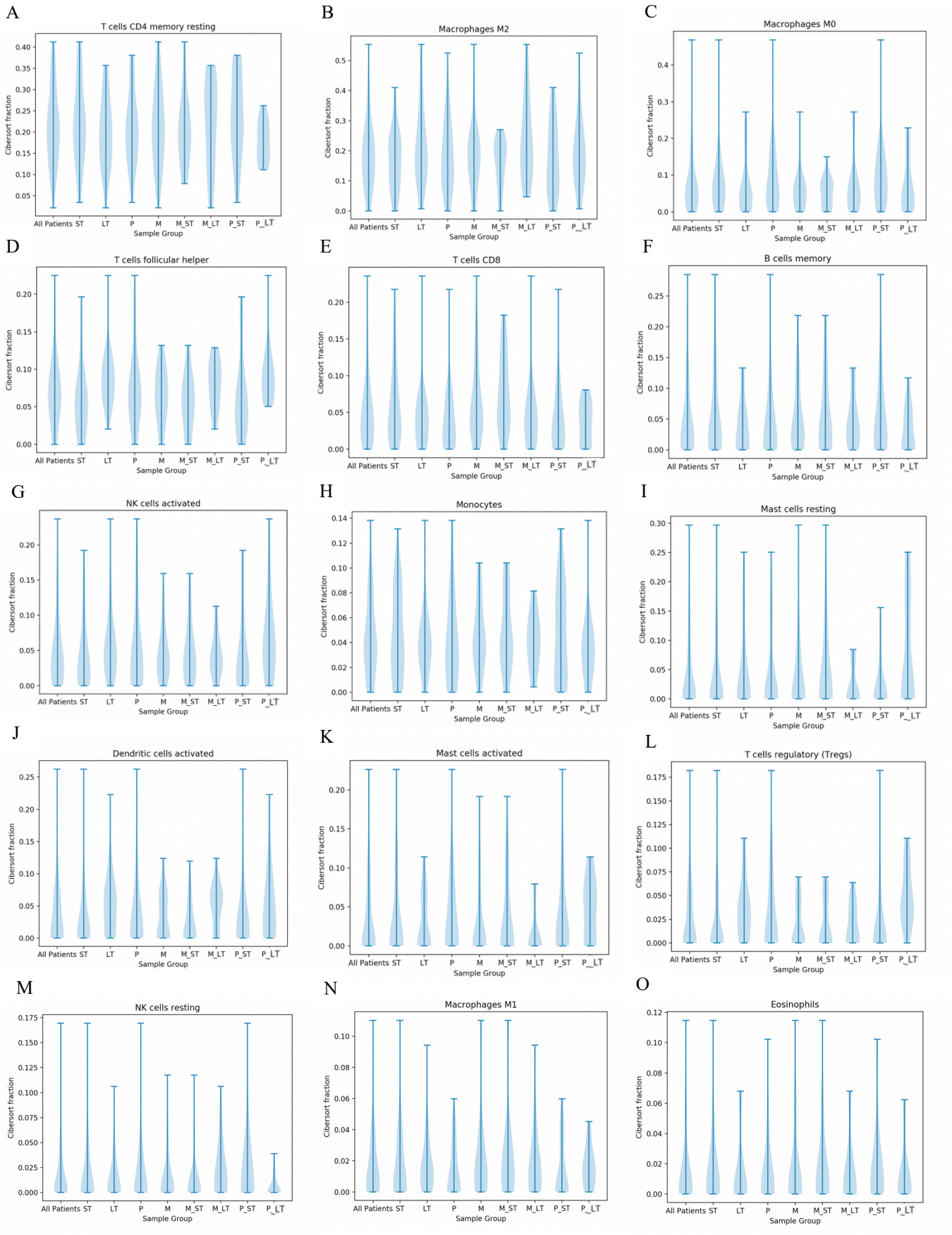
D



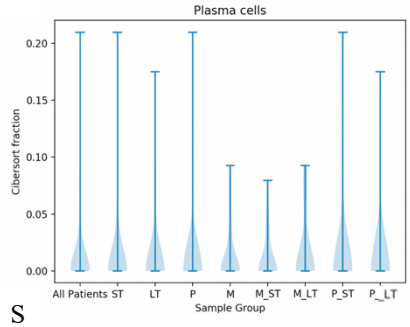

Cohort CNA Threshold vs RNA-seq Expression
RNA-seq Log2(FPKM)
GISTIC 2.0 Copy Number Threshold

Supplementary Figure 2.9: Copy number alterations in tumors of ST and LT survivors

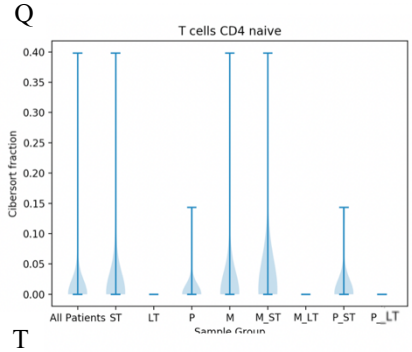
A. Copy-altered segments in primary and metastatic tumors of ST survivors across the genome. Amplifications are depicted in red while deletions are seen in blue.
B. Copy-altered segments in primary and metastatic tumors of LT survivors across the genome. Amplifications are depicted in red while deletions are seen in blue.
C. GISTIC Qplot of significantly amplified lesions across genome of all patient tumors (n=74) in cohort with 90% Confidence.
D. GISTIC Qplot of significantly deleted lesions across genome of all patient tumors (n=74) in cohort with 90% Confidence.
E. Violin plot of the GISTIC2.0 CNA value for every gene in every patient primary and metastatic sample plotted against their corresponding Log2(FPKM) value from RNA-seq. The medians of the FPKM values and the CNA threshold values has a spearmanr correlation value of 0.999 and p-value of 1.4x10^-24). (Number of genes in Copy Number Threahold groups from left to right: -2 n= 24316, -1 n= 416616, 0 n=280688, 1 n=346144 , 2 n=51856)


A
T cells CD4 memory resting
B
Macrophages M2
C
Macrophages M0
D
T cells follicular helper
E
T cells CD8
F
B cells memory
G
NK cells activated
H
Monocytes
I
Mast cells resting
J
Dendritic cells activated
K
Mast cells activated
L
T cells regulatory (Tregs)
M
NK cells resting
N
Macrophages M1
O
Eosinophils
Cibersort fraction
Sample Group
All Patients ST
LT
P
M
M_ST
M_LT
P_ST
P_LT

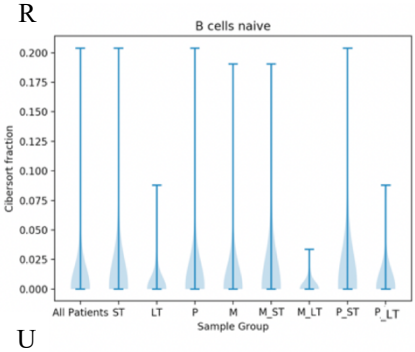
P



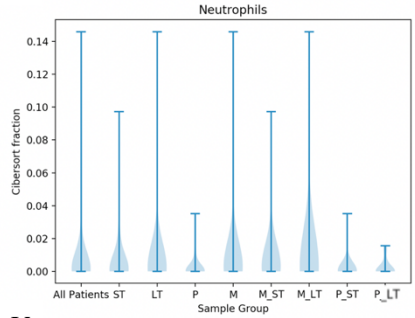
Q



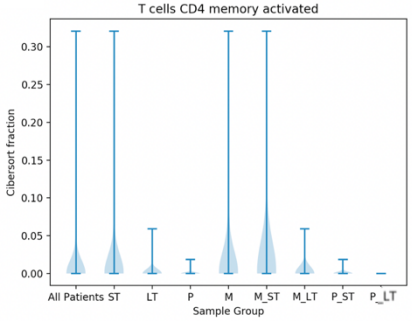
R



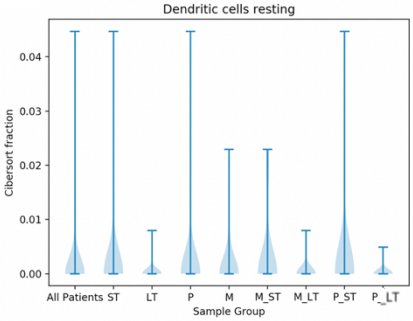
S



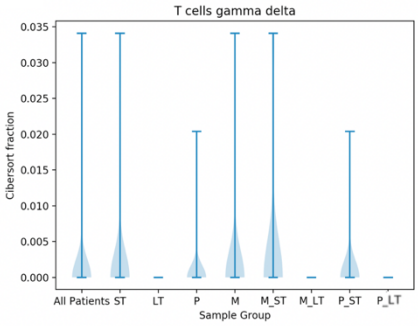
T



U



V



W

	ST v LT survivors		Primary (P) v Metastatic (M)		ST survivors - PvM		LT survivors - PvM		Primary tumors - ST v LT		Metastatic tumors - ST v LT	
Immune cell type	Mann-Whitney Statistic	p-value	Wilcoxon Statistic	p-value	Wilcoxon Statistic	p-value	Wilcoxon Statistic	p-value	Mann-Whitney Statistic	p-value	Mann-Whitney Statistic	p-value
B cells naive	690	0.28222397	51	0.379374669	32	0.58291955	3	0.46520882	177	0.37074399	167	0.593786144
B cells memory	647	0.72285701	196	0.203540906	82	0.60087101	24	0.13295729	164	0.75343757	158	0.909458667
Plasma cells	696.5	0.24182612	52	0.408044206	32	0.58291955	4	0.71500065	189.5	0.11864898	157	0.922974403
T cells CD8	733	0.17680624	187	0.036035175	49	0.01032448	45	0.97212533	154.5	1	217.5	0.040789302
T cells CD4 naive	784	0.00287582	23	0.646462203	23	0.6464622	NA	NA	189	0.06265679	203	0.022749719
T cells CD4 memory resting	690	0.3958775	312	0.741458844	104	0.48259115	24	0.07849121	208	0.08255079	133	0.505906885
T cells CD4 memory activated	658	0.37963095	2	0.025061844	1	0.04639946	0	0.17971249	161	0.45884883	169	0.495912321
T cells follicular helper	356	0.00271976	259	0.359021078	100	0.5900638	34	0.26757813	84	0.02411617	93	0.049561507
T cells regulatory (Tregs)	375	0.00298062	144	0.279641534	43	0.86130426	12	0.00854492	53	0.0004148	145	0.770623805
T cells gamma delta	658	0.16619833	1	0.285049407	1	0.28504941	NA	NA	161	0.45884883	168	0.269900195
NK cells resting	581	0.65898278	141	0.797068494	26	0.17295492	4	0.00992549	205.5	0.05963849	80	0.009823086
NK cells activated	490.5	0.1478352	259	0.925510561	75	0.42090992	27	0.19605096	105	0.11421541	149	0.883753087
Monocytes	676	0.49160656	301	0.615148209	119	0.82372856	40	0.46313477	168	0.66061466	164	0.757858154
Macrophages M0	691	0.38808951	202	0.06419212	67	0.05424547	39	0.64964403	181	0.38921972	163.5	0.769031965
Macrophages M1	505	0.19618414	223	0.304232069	68	0.27724064	48	0.80773926	109	0.13385961	150.5	0.9222738
Macrophages M2	499	0.17838381	321	0.850463631	115	0.72618389	41	0.50158691	134	0.52685427	115	0.211546541
Dendritic cells resting	664	0.41330065	34	0.694886602	15	0.67442407	5	1	162	0.70775726	170	0.466676921
Dendritic cells activated	452	0.05303165	227	0.909930768	72	0.55657941	29	0.43276758	134	0.51125239	93	0.0464126
Mast cells resting	606	0.90155873	111	0.875710309	25	0.15195619	6	0.09289194	139.5	0.60513707	166	0.676859087
Mast cells activated	625	0.91640813	137	0.211555387	75	0.94339206	6	0.02841686	124.5	0.32455326	183.5	0.301292617
Eosinophils	715	0.24089382	180	0.828812374	85	0.98262729	19	0.67840238	174	0.51689552	182	0.354247644
Neutrophils	587.5	0.71791202	78	0.039672486	21	0.04798975	22	0.57506213	155	0.98413058	143	0.718726623

Supplementary Figure 2.10: CIBERSORT immune cell fraction violin plots

A-V. Breakdown of the immune cell fractions calculated by CIBERSORT for each of the 22 immune cell signatures broken down by tumor type subgroup (all Patient tumors n= 72, ST survivors n= 44, LT survivors n= 28, Primary tumors n = 36, Metastatic tumors n= 36, ST survivor metastatic tumors n= 22, LT survivor metastatic tumors n=14, ST survivor primary tumors n= 22, and LT survivor metastatic tumors n= 14)

2.9 References

1. McPherson, A., et al., *Divergent modes of clonal spread and intraperitoneal mixing in high-grade serous ovarian cancer.* Nat Genet, 2016. **48**(7): p. 758-67.
2. Bowtell, D.D., *The genesis and evolution of high-grade serous ovarian cancer.* Nat Rev Cancer, 2010. **10**(11): p. 803-8.
3. Stewart, C., C. Ralyea, and S. Lockwood, *Ovarian Cancer: An Integrated Review.* Semin Oncol Nurs, 2019. **35**(2): p. 151-156.
4. Howlader N, N.A., Krapcho M, Miller D, Brest A, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ, Cronin KA (eds), *SEER Cancer Statistics Review.* National Cancer Institute, 1975-2017.
5. McLachlan, J., A. George, and S. Banerjee, *The current status of PARP inhibitors in ovarian cancer.* Tumori, 2016. **102**(5): p. 433-440.
6. Ledermann, J.A., *PARP inhibitors in ovarian cancer.* Ann Oncol, 2016. **27 Suppl 1**: p. i40-i44.
7. Testa, U., et al., *Ovarian Cancers: Genetic Abnormalities, Tumor Heterogeneity and Progression, Clonal Evolution and Cancer Stem Cells.* Medicines (Basel), 2018. **5**(1).
8. Bell, D., et al., *Integrated genomic analyses of ovarian carcinoma.* Nature, 2011. **474**(7353): p. 609-615.
9. Yang, S.Y.C., et al., *Landscape of genomic alterations in high-grade serous ovarian cancer from exceptional long- and short-term survivors.* Genome medicine, 2018. **10**(1): p. 81-81.
10. Alsop, K., et al., *BRCA mutation frequency and patterns of treatment response in BRCA mutation-positive women with ovarian cancer: a report from the Australian Ovarian Cancer Study Group.* J Clin Oncol, 2012. **30**(21): p. 2654-63.
11. Lee, S., et al., *Molecular Analysis of Clinically Defined Subsets of High-Grade Serous Ovarian Cancer.* Cell Rep, 2020. **31**(2): p. 107502.

12. Sallinen, H., et al., *Comparative transcriptome analysis of matched primary and distant metastatic ovarian carcinoma.* BMC Cancer, 2019. **19**(1): p. 1121.

13. Knisely, A.T., et al., *Trends in Primary Treatment and Median Survival Among Women With Advanced-Stage Epithelial Ovarian Cancer in the US From 2004 to 2016.* JAMA Netw Open, 2020. **3**(9): p. e2017517.

14. Ewels, P., et al., *MultiQC: summarize analysis results for multiple tools and samples in a single report.* Bioinformatics, 2016. **32**(19): p. 3047-8.

15. *"Picard Toolkit"*. 2019, GitHub Repository: Broad Institute.

16. Li, H. and R. Durbin, *Fast and accurate short read alignment with Burrows-Wheeler transform.* Bioinformatics, 2009. **25**(14): p. 1754-60.

17. Ghobadi, A., et al., *Shared cell of origin in a patient with Erdheim-Chester disease and acute myeloid leukemia.* Haematologica, 2019. **104**(8): p. e373-e375.

18. Griffith, M., et al., *Genome Modeling System: A Knowledge Management Platform for Genomics.* PLoS Comput Biol, 2015. **11**(7): p. e1004274.

19. Cibulskis, K., et al., *Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples.* Nat Biotechnol, 2013. **31**(3): p. 213-9.

20. Saunders, C.T., et al., *Strelka: accurate somatic small-variant calling from sequenced tumor-normal sample pairs.* Bioinformatics, 2012. **28**(14): p. 1811-7.

21. Koboldt, D.C., et al., *VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing.* Genome Res, 2012. **22**(3): p. 568-76.

22. Larson, D.E., et al., *SomaticSniper: identification of somatic point mutations in whole genome sequencing data.* Bioinformatics, 2011. **28**(3): p. 311-317.

23. Li, H., *A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data.* Bioinformatics, 2011. **27**(21): p. 2987-93.

24. Ye, K., et al., *Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads.* Bioinformatics, 2009. **25**(21): p. 2865-71.

25. DePristo, M.A., et al., *A framework for variation discovery and genotyping using next-generation DNA sequencing data.* Nature Genetics, 2011. **43**(5): p. 491-498.

26. Fairley, S., et al., *The International Genome Sample Resource (IGSR) collection of open human genomic variation resources.* Nucleic Acids Res, 2020. **48**(D1): p. D941-d947.

27. *NHLBI GO Exome Sequencing Project (ESP)*, E.V. Server, Editor.: Seattle, WA.

28. Skidmore, Z.L., et al., *GenVisR: Genomic Visualizations in R.* Bioinformatics, 2016. **32**(19): p. 3012-4.

29. Landrum, M.J., et al., *ClinVar: improving access to variant interpretations and supporting evidence.* Nucleic acids research, 2018. **46**(D1): p. D1062-D1067.

30. Verhaak, R.G., et al., *Prognostically relevant gene signatures of high-grade serous ovarian carcinoma.* J Clin Invest, 2013. **123**(1): p. 517-25.

31. Mermel, C.H., et al., *GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers.* Genome Biology, 2011. **12**(4): p. R41.

32. Shao, X., et al., *Copy number variation is highly correlated with differential gene expression: a pan-cancer study.* BMC Med Genet, 2019. **20**(1): p. 175.

33. Bray, N.L., et al., *Near-optimal probabilistic RNA-seq quantification.* Nature Biotechnology, 2016. **34**(5): p. 525-527.

34. Robinson, M.D. and A. Oshlack, *A scaling normalization method for differential expression analysis of RNA-seq data.* Genome Biol, 2010. **11**(3): p. R25.

35. Parker, H.S., et al., *Preserving biological heterogeneity with a permuted surrogate variable analysis for genomics batch correction.* Bioinformatics, 2014. **30**(19): p. 2757-63.

36. Viljoen, K.S. and J.M. Blackburn, *Quality assessment and data handling methods for Affymetrix Gene 1.0 ST arrays with variable RNA integrity.* BMC Genomics, 2013. **14**: p. 14.

37. Leek, J.T., et al., *The sva package for removing batch effects and other unwanted variation in high-throughput experiments.* Bioinformatics, 2012. **28**(6): p. 882-3.

38. Mi, H., et al., *PANTHER version 16: a revised family classification, tree-based classification tool, enhancer regions and extensive API.* Nucleic Acids Research, 2020. **49**(D1): p. D394-D403.

39. Mi, H. and P. Thomas, *PANTHER pathway: an ontology-based pathway database coupled with data analysis tools.* Methods Mol Biol, 2009. **563**: p. 123-40.

40. Ashburner, M., et al., *Gene ontology: tool for the unification of biology. The Gene Ontology Consortium.* Nat Genet, 2000. **25**(1): p. 25-9.

41. *The Gene Ontology resource: enriching a GOld mine.* Nucleic Acids Res, 2021. **49**(D1): p. D325-d334.

42. Huang da, W., B.T. Sherman, and R.A. Lempicki, *Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources.* Nat Protoc, 2009. **4**(1): p. 44-57.

43. Sherman, B.T., et al., *DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update).* Nucleic Acids Res, 2022. **50**(W1): p. W216-21.

44. Newman, A.M., et al., *Robust enumeration of cell subsets from tissue expression profiles.* Nature Methods, 2015. **12**(5): p. 453-457.

45. Zhang, J., et al., *INTEGRATE: gene fusion discovery using whole genome and transcriptome data.* Genome Res, 2016. **26**(1): p. 108-18.

46. Dobin, A., et al., *STAR: ultrafast universal RNA-seq aligner.* Bioinformatics, 2013. **29**(1): p. 15-21.

47. Peng, G. and G.B. Mills, *Surviving Ovarian Cancer: An Affair between Defective DNA Repair and RB1.* Clin Cancer Res, 2018. **24**(3): p. 508-510.

48. Alexandrov, L.B., et al., *Deciphering signatures of mutational processes operative in human cancer.* Cell Rep, 2013. **3**(1): p. 246-59.

49. Alexandrov, L.B., et al., *Signatures of mutational processes in human cancer.* Nature, 2013. **500**(7463): p. 415-21.

50. Cerami, E., et al., *The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data.* Cancer Discov, 2012. **2**(5): p. 401-4.

51. Gao, J., et al., *Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal.* Sci Signal, 2013. **6**(269): p. pl1.

52. Etemadmoghadam, D., et al., *Integrated genome-wide DNA copy number and expression analysis identifies distinct mechanisms of primary chemoresistance in ovarian carcinomas.* Clin Cancer Res, 2009. **15**(4): p. 1417-27.

53. Wang-Johanning, F., et al., *Expression of human endogenous retrovirus k envelope transcripts in human breast cancer.* Clin Cancer Res, 2001. **7**(6): p. 1553-60.

54. Lee, S.H., et al., *Elevation of human ERV3-1 env protein expression in colorectal cancer.* J Clin Pathol, 2014. **67**(9): p. 840-4.

55. Tian, S., G. Meng, and W. Zhang, *A six-mRNA prognostic model to predict survival in head and neck squamous cell carcinoma.* Cancer management and research, 2018. **11**: p. 131-142.

56. Geng, H., et al., *Survival prediction for patients with lung adenocarcinoma: A prognostic risk model based on gene mutations.* Cancer Biomark, 2020. **27**(4): p. 525-532.

57. Song, C., et al., *PRDX1 stimulates non-small-cell lung carcinoma to proliferate via the Wnt/β-Catenin signaling.* Panminerva Med, 2020.

58. Chen, H., L. Yang, and W. Sun, *Elevated OGN expression correlates with the EMT signature and poor prognosis in ovarian carcinoma.* Int J Clin Exp Pathol, 2019. **12**(2): p. 584-589.

59. Liu, H., et al., *Long non-coding RNAs as prognostic markers in human breast cancer.* Oncotarget, 2016. **7**(15): p. 20584-96.

60. Lei, J.T., et al., *Functional Annotation of ESR1 Gene Fusions in Estrogen Receptor-Positive Breast Cancer.* Cell Rep, 2018. **24**(6): p. 1434-1444.e7.

61. Thorsson, V., et al., *The Immune Landscape of Cancer.* Immunity, 2018. **48**(4): p. 812-830.e14.

62. Lee, J.Y., et al., *Tumor evolution and intratumor heterogeneity of an epithelial ovarian cancer investigated using next-generation sequencing.* BMC Cancer, 2015. **15**: p. 85.

63. Zhang, A.W., et al., *Interfaces of Malignant and Immunologic Clonal Dynamics in Ovarian Cancer.* Cell, 2018. **173**(7): p. 1755-1769.e22.

64. Rojas, V., et al., *Molecular Characterization of Epithelial Ovarian Cancer: Implications for Diagnosis and Treatment.* International journal of molecular sciences, 2016. **17**(12): p. 2113.

65. Hornburg, M., et al., *Single-cell dissection of cellular components and interactions shaping the tumor immune phenotypes in ovarian cancer.* Cancer Cell, 2021.

66. Xiao, G., et al., *Gain-of-Function Mutant p53 R273H Interacts with Replicating DNA and PARP1 in Breast Cancer.* Cancer Res, 2020. **80**(3): p. 394-405.

67. Qiu, W.G., et al., *Identification, validation, and targeting of the mutant p53-PARP-MCM chromatin axis in triple negative breast cancer.* NPJ Breast Cancer, 2017. **3**.

68. Polotskaia, A., et al., *Proteome-wide analysis of mutant p53 targets in breast cancer identifies new levels of gain-of-function that influence PARP, PCNA, and MCM4.* Proc Natl Acad Sci U S A, 2015. **112**(11): p. E1220-9.

69. Koivunen, J.P., et al., *EML4-ALK fusion gene and efficacy of an ALK kinase inhibitor in lung cancer.* Clin Cancer Res, 2008. **14**(13): p. 4275-83.

70. Cocco, E., M. Scaltriti, and A. Drilon, *NTRK fusion-positive cancers and TRK inhibitor therapy.* Nat Rev Clin Oncol, 2018. **15**(12): p. 731-747.

71. Adib, T.R., et al., *Predicting biomarkers for ovarian cancer using gene-expression microarrays.* British journal of cancer, 2004. **90**(3): p. 686-692.

72. Hibbs, K., et al., *Differential gene expression in ovarian carcinoma: identification of potential biomarkers.* The American journal of pathology, 2004. **165**(2): p. 397-414.

73. Ma, Y., Y. Lu, and B. Lu, *MicroRNA and Long Non-Coding RNA in Ovarian Carcinoma: Translational Insights and Potential Clinical Applications.* Cancer Invest, 2016. **34**(9): p. 465-476.

74. Wang, J.Y., A.Q. Lu, and L.J. Chen, *LncRNAs in ovarian cancer.* Clin Chim Acta, 2019. **490**: p. 17-27.

75. Tong, W., et al., *A new tumor suppressor lncRNA RP11-190D6.2 inhibits the proliferation, migration, and invasion of epithelial ovarian cancer cells.* Onco Targets Ther, 2017. **10**: p. 1227-1235.

Chapter 3: Assessment of TP53 R273H as a Gain-of-Function mutation in High-Grade Serous Ovarian cancer

Declarations

All experiments were designed, performed, and analyzed by Emilee Kotnik. Elena Lomonosova, Hollie Noia, and Pat Kung advised on experimental design for the plasmid and lentiviral transfections and drug treatments for cell viability assays. Daniel Wilke helped design the clonogenics protocol. Shirley De Leon-Lopez assisted with cell maintenance and sanger sequencing.

3.1 Introduction

TP53 mutations are a hallmark of high-grade serous ovarian cancer, as virtually all High-Grade Serous Ovarian Cancer (HGSC) tumors harbor them [1]. *TP53* encodes for a transcription factor, known as p53, that plays a major role in maintaining the integrity of the genome and controls the expression of hundreds of genes in response to stress signals in cells. It is also one of the most commonly mutated tumor suppressor genes across cancer types. [2, 3]

TP53 has several known hotspot mutations across cancer types and in HGSC there are 9 reported hotspot mutations at loci R273, R248, R175, Y220, I195, C176, G245, S241, and Y163. [4] The majority of alterations to *TP53* are missense mutations in the DNA-binding domain and cause loss-of-function (LOF). The LOF effect means that the cell losses the anticancer protection of wild type p53, but some time the mutant allele actually masks the function of the wild-type copy of the gene. [3, 5] There are also gain-of-function (GOF) p53 mutations identified in cancers that contribute to metastasis and

chemoresistance. [4] For example, Zhou et al showed that mutant TP53 mutations R175H, G245C, and R282W promoted invasive cell growth in head and neck cancer cells and that mutant p53 could influence cancer cell metabolism through inhibiting AMPK Activation. [6] Mutant p53 has also been shown to modify histones and affect chromatin modeling and gene expression. [7]

Recent functional studies have shown that the TP53 R273H mutation results in a p53 GOF that may promote metastasis in colorectal [8], esophageal [9], and breast cancer [10-12]. It has been shown that mice expressing human TP53 R273H have more metastatic disease compared to mice without p53 or with WT p53 [13]. Studies on breast cancer cell lines have linked the mutant p53 to PARP1 and DNA repair pathway mechanisms [11, 12]. Their work suggests that cancer cells with this mutation could be treated more effectively with combination therapy with Temozolomide and PARP inhibitors (PARPi) [11]. Between 1.5-2.25% of ovarian cancer patients harbor the TP53 R273H mutation, making this a targetable mutation that has the potential to be clinically relevant. [1] If this biology were to hold true in ovarian cancer, it would potentially qualify more patients for PARPi and combination treatments in the future. Therefore, we aimed to gather preliminary evidence that the TP53 R273H mutation causes a GOF in ovarian cancer cells and to determine if the TP53 mutation sensitizes cells to PARPi treatment in combination with Temozolomide.

3.2 Methods and Materials

3.2.1 Cell culture and maintenance

CAOV-3 cells were obtained from Wendy Fantl (Stanford University) and maintained in DMEM (Sigma-Aldrich) media supplemented with 10% FBS and 1% penicillin (pen) and streptomycin (strep) antibiotics. DPBS (Gibco) and 0.25% trypsin-EDTA (Gibco) were used to wash and lift cells for expanding. Cell counting was performed on a Countess II (Life Technologies) according to manufacturer's protocols. OVCA-420 cells were obtained from Joyce Lui (Dana-Faber Canccer Institute)

and cultured with RPMI 1640 media (Sigma-Aldrich) supplemented with 10% FBS, 10 mM HEPES, and 0.5% pen/strep antibiotics. MDA-MB-468 cells were maintained in DMEM media (Sigma-Aldrich) with 10% FBS and 1% pen/strep antibiotics and were a gift from the Jason Weber Lab (Washington University in St. Louis). LentiX 293T cells were maintained in DMEM with 10% FBS. All cells were incubated at 37°C with 5% CO2.

3.2.2 Plasmid transfections

Plasmids were obtained from Addgene for the empty vector control (pCMV-Neo-Bam plasmid #16440), mutant TP53 R273H (pCMV-Neo-Bam p53 R273H plasmid #16239 [14]), and wildtype TP53 (pCMV-Neo-Bam p53 wt plasmid #16434 [14]). They were provided in bacteria as agar stabs. Plasmids were maxi-prepped according manufacturer's protocols using the HiSpeed Plasmid Maxi Kit (REF#12662).

Plasmid transfections were attempted with 3 transfection reagents to see which performed the most effectively with either lipofectamine 2000 (ThermoFisher), TransIT-LT1 (Mirus Bio), and TurboFect (ThermoFisher) in alignment with their individual manufacturer's protocols. All transfections were performed with CAOV-3 cells in DMEM media with 10% FBS without antibiotic supplements. Plasmids and transfection reagents were diluted in Opti-MEM (Gibco) for the Lipofectamine and TransIT-LT1 protocols. These plasmids contained the neomycin resistance gene and were selected for with the addition of G418 solution (Millipore Sigma #G8168) at a concentration of 0.5 mg/ml. We determined that TurboFect transfected cells exhibited stable mutant p53 expression after antibiotic selection.

3.2.3 Lentiviral transductions

We obtained 3rd generation lentiviral plasmids from Addgene: scramble shRNA (Plasmid #1864 [15]), mutant TP53 R273H (pLenti6/V5-p53_R273H Plasmid #22934 [16]), and wildtype TP53

(pLenti6/V5-p53_wt p53 Plasmid #222945 [16]). We used viral packaging plasmids pMDLg/pRRE (Addgene #12251), pRSV-Rev (Addgene #12253), and envelope plasmid pCMV-VSV-G (Cell Biolabs RV-110). These plasmids were received in a bacterial stab and were maxi prepped using the Qiagen HiSpeed Plasmid Maxi Kit (REF#12662).

LentiX 293T cells were seeded at a concentration of 2x10^6 in a 10 cm dish for viral packaging. The following day when they were 70-80% confluent, media was replaced on the cells 2 hours prior to transfection and then transfected with the transfection plasmids and reagent. We used 60µl of TransIT-LT1 diluted in 85 µl of opti-MEM that was incubated with 5µg of pMDLg/PRRE, 2.5 µg of pRSV-Rev, pCMV-VSV-G, and 10 µg of either the empty vector, mutant p53, or wildtype p53 plasmid for 30 minutes. This solution was added to the 293T cells dropwise and replaced to the incubation overnight for 18 hours. Transfection media was replaced with 15 ml of high serum growth media consisting of DMEM, 30% FBS, and 1% pen/strep for 24 hours. To harvest, media was spun down at 1250 rpm for 5 minutes and the supernatant was filtered through a 45 um PES filter. The viral media was added directly to the CAOV-3 cells that were seeded the day before at 4x10^6 cells/well in a 10 cm plate. This harvesting process was repeated and the virus was again used to the transduce the cells directly after collection. Cells were selected with 6 µg/ml of blasticidin (Invitrogen cat #R210-01) for 10 days post-transfection.

3.2.4 MTS survival assay

CAOV-3, OVCA-420, or MDA-MB-468 were plated at 5000 cells/well in a 96-well plate (TPP) and were treated the next day with Temozolomide (Selleck Chemicals), Olaparib (Selleck Chemicals) Talazoparib (Selleck Chemicals, BMN 673) alone or in combination with each other. All MTS assays were plated in triplicate and the highest dose formulated with media at concentrations of 2 mM Temozolomide, 20 µM of Talazoparib, and 100 µM of Olaparib and then serially diluted by half to create cell viability curves. All drugs stocks were diluted in DMSO and heated to get the drug into solution.

After 72 hours of treatment, 20 μl of MTS (3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium) solution (Promega) was added to each well, incubated for 4 hours, and then the absorbance was measured at 490 nm with a 96-well plate reader (Tecan infinite M200 Pro). If PARPi treatment was used, plates were treated for 72 hours and then washed and changed to full media without treatment and read with MTS 72 hours after the media change. Alongside all treatments, a DMSO control was used to validate that the DMSO volume was not affecting cell viability. Cell viability percentages were calculated by subtracting the average absorption reading of the media and MTS and then dividing by the no treatment well for each triplicate and converting to a percentage. Statistics were calculated with GraphPad Prism 9 software and the cell viability percentages were log transformed. A log inhibitor vs response 4 parameter nonlinear regression fit test was used.

3.2.5 Clonogenics

CAOV-3 and OVCA-420 cells were plated at a concentration of 5000 cells/well in a 6-well plate (TPP) and treated with Temozolomide, Talazoparib, Olaparib, or a combination of those drugs. Dose concentrations were determined based on the MTS cell viability curves IC50 values. Cells were treated for 3 consecutive days with drug combinations and then replaced with full media without drug and incubated for 7 days before staining. Staining consisted of washing the with cold DPBS, and fixing with 100% MeOH for 20 minutes. We then stained with 2 ml of crystal violet for 5 minutes and washed with water twice. Plates were then covered with foil and dried overnight. Images were taken on a ChemiDoc (Bio-Rad Laboratories). The plates were then de-stained and the amount of crystal violet was measured by absorbance.

3.2.6 Western blot

Cells were lysed using 8 molar urea lysis buffer and DTT and then sonicated for a few seconds before being spun down in a centrifuge at 4C for 15 min at 13000 rpm. Protein amounts were quantified by

Bradford assay and 50-70 ug of lysates were mixed with BME. Lysates were subjected to SDS-PAGE. The gel was run for 30 min at 61V and another 60-90 minutes at 120V. The protein was then transferred to a nitrocellulose membrane at 85 V for 80 minutes. Membranes were blocked with 10% milk solution dissolved in TBST, washed with TBST, and probed with primary antibodies for TP53 (Cell Signaling, #2527, diluted 1:1000 in BSA) to assess knockdown and B-actin (Millipore Sigma A1978, diluted 1:3000 in milk) or GAPDH (Cell Signaling, diluted 1:3000 in milk) for 1 night at 4C. The blot was then washed with and incubated in secondary mouse antibody (Thermo Fisher Scientific A-11001) for approximately 2 hours and washed with TBST. Primary and secondary antibodies were diluted in 5% milk and TBST. Protein signals detected with the Pierce ECL Westering Blotting Substrate and imaged using chemiluminescence on a ChemiDoc (Bio-Rad Laboratories).

3.2.7 Sanger Sequencing

Primers flanking the TP53 R273H mutation in CAOV-3 cells were ordered from IDT, with the forward sequence 5'TAACTGCACCCTTGGTCTCC3' and Reverse sequence 5' GGCTTTGGGACCTCTTAACC'. Plasmid primers were also from IDT with forward sequence 5' cctcaccatcatcacactgg 3' and reverse sequence as 5' tggacagtgctcgcttagtg 3'. The TP53 gene was amplified in PCR with master mix (x) and annealing temp 60C. The PCR reaction was run on a 1.5% agarose gel and bands were cut and purified with the Qiagen gel purification kit according to manufacturer's protocol. The isolated DNA was diluted with nuclease free water and sent for sanger sequencing through GeneWiz according to their plasmid or purified PCR product guidelines. Chromatograms were viewed using FinchTV.

3.3 Results

3.3.1 Identification of the TP53 R273H mutation in HGSC patient cohort

In Chapter 2, our lab whole-exome sequenced primary and metastatic tumors from 39 HGSC patients and identified the mutant TP53 R273H in the primary and metastatic tumors of 3 patients. Two of these patients were short-term survivors (OS <3 years after diagnosis). (Figure 1A) The OS of the ST survivors ranged between 17-19.6 months and the LT survivor lived more than 147 months after their diagnosis. In The Cancer Genome Atlas (TCGA) database, there are a total of 99 cases across cancer types that harbor a TP53 R273H mutation, 9 of which are in ovarian cancer tumors.

3.3.2 Creation of an ovarian cancer cell line that expresses the TP53 R273H mutation

In order to determine if the TP53 R273H mutation creates a GOF p53 protein and is therapeutically targetable, first we created an ovarian cancer cell line that expresses the mutant p53. We transfected CAOV-3 cells with a plasmid with the TP53 R273H mutated gene, an empty vector or wildtype TP53 gene. CAOV-3 is an ovarian cancer cell line that does not express *TP53* because of a nonsense mutation early in the gene. First, we validated that CAOV-3 cells do no express p53 at the protein level via western blot. (Figure 1B) We also demonstrated that the ovarian cancer cell line OVCA-420 and breast cancer cell line MDA-MB-468 both endogenously the express TP53 R273H mutation. (Figure 1B)

We attempted to transfect and transduce mutant p53 plasmids into CAOV-3 cells several times. Figure 1C highlights the p53 protein expression of those cell lines. The CAOV-3 cells replicate slowly and took several weeks to recover from being transfected with the mutant p53 plasmid after selection and reselection with Neomycin (G418). We also verified via Sanger Sequencing that the plasmid transfected

into the CAOV-3 cells harbors the R273H resulting mutation from a G>A. (Figure 1D) The CAOV-3 cells transfected with Turbofect had the most stable mutant p53 protein expression and we moved forward with treatment experiments with those cells. We attempted lentiviral plasmid transfections 2 times, but did not see much mutant p53 expression and did not use them for further experiments. (Figure 1C) All cells transfected or transduced with wildtype p53 plasmids died, unsurprisingly given the function of p53 as tumor suppressor. Even if the cancer cells successfully took up the plasmid, p53 expression most likely led to apoptosis for those cells.

3.3.3 TP53 R273H does not influence sensitivity to PARP inhibitors and Temozolomide treatment

After establishing mutant p53 expression in ovarian and breast cancer cells, we next determined whether expression of the mutation TP53 R273H sensitizes cells to PARP inhibitors and is clinically treatable. First, we wanted to recapitulate combination treatments that other groups have published. [11, 12]

To measure cell viability in response to treatment, we performed MTS assays and treated cells with PARPi (Talazoparib or Olaparib) in combination with Temozolomide, a DNA alkylating agent. First, we treated MDA-MB-468 using treatment amounts based on the published Qiu et al paper. [11] We treated with 20 μM of Talazoparib and 1 mM of Temozolomide in combination and alone as our highest treatments, then serially diluted by half for 11 dilutions and compared to no treatment. (Figure 2A) The MDA-MB-468 cells were most sensitive to combination treatment at levels comparable to the breast cancer studies that demonstrated sensitivity to the combination of Talazoparib and Temozolomide in an XTT cell viability assay (Figure 2B).

After validating the MTS methods with previously published results, we used this same approach to test whether TP53 R273H mutations altered the sensitivity of PARPi using OVCA-420 and the

transfected CAOV-3 cells. We expected that the OVCA-420 and CAOV-3 cells with mutant p53 would be more sensitive to PARPi treatment than the parental or empty vector controls.

We again used the same treatment amounts of Talazoparib and Temozolomide from the MDA-MB-468 experiments as a starting point for these ovarian cell lines. The CAOV-3 mutant p53, empty vector control, parental cell lines were the most sensitive to combination treatments of Talazoparib and Temozolomide than either drug alone. (Figure 3A-C) If we compare each of the CAOV-3 cells combination treatment response (Figure 3D), they all have comparable IC50 values and show that the CAOV-3 cells are sensitive to the combination of drugs regardless of the presence of mutant p53.

Next, we wanted to see if using a PARPi that is more commonly used in ovarian cancer patients would show any differences in sensitization with Temozolomide in the context of the mutant p53. Figure 3E-H shows the MTS cell viability curve comparison between the CAOV-3 parental, empty vector control, and mutant p53 cells treated with the combination of Olaparib and Temozolomide. We used 100 μM as the highest dose of Olaparib and 2 μM as the highest dose of Temozolomide, which were then diluted by half and normalized to no treatment. Again, there were no differences between these cell lines in response to the combination treatment, so we can determine that the mutant p53 R273H in the CAOV-3 does not improve the sensitivity of the cells more to this combination treatment.

Additionally, we sought to further investigate cell survival to complement our MTS cell viability assay by treating the cells in a clongenics assay. The transfected CAOV-3 cells and parental cells were treated with 8 μM of Olaparib, 10 μM of Olaparib, 200 μM of Temozolomide, or the combination of those treatment doses. In this assay, the CAOV-3 cells were the most sensitive to the Temozolomide treatment, therefore we could not conclude that the combination treatments with Temozolomide led to any further sensitization at these treatment doses. All the CAOV-3 cell lines were sensitive to Olaparib as well and showed no further sensitivities to the PARPi in the presence of mutant p53. (Figure 3I)

We treated the OVCA-420 cells with combinations of either Talazoparib or Olaparib, and Temozolomide. The cell viability of the OVCA-420 were also the most sensitive to combination treatment of Talazoparib and Temozolomide, but this combination was not significantly different from the sensitivity to Temozolomide alone and the combination of Olaparib and Temozolomide. However, at the same concentration of drugs used in the previous MTS assays, the cells viability never showed less than 50% viability even at the highest combinations of the treatments. At these high doses, the cells began to die from the high volume of DMSO added from the drug solutions so we could not further increase our treatment doses. This shows that the OVCA-420 cells are not as sensitive to these drugs as the other cell lines tested. (Figure 4A). We also performed a clonogenic assay to measure cell survival of the OVCA-420 cells in response to Olaparib and Temozolomide treatments. The results did show that the cells were most sensitive to the combination treatment compared to the controls (Figure 4B).

3.4 Discussion

We set out to determine if mutant TP53 R273H can sensitize cells to combination treatments between PARPi and Temozolomide because of a GOF in p53 in ovarian cancer. We were able to replicate the published results by Qui et al showing that the MDA-MB-468 cell line that harbored the R273H mutation are most sensitive to the combination treatment of Temozolomide and Talazoparib. We therefore expected this result to hold true in the OVCA-420 cell line and transfected CAOV-3 cells with mutant p53. However, our results showed that the combination of PARP inhibitors Talazoparib and Olaparib with Temozolomide was effective in eliminating ovarian cancer cell lines, independent of the TP53 R273H mutation.

To further determine if the p53 mutation causes the increased sensitivity to combination treatment, we could have knocked-down or knocked-out p53 expression in the OVCAR-420 or MDA-MB-468 cells. This would have shown if sensitivity to the combination treatments occurred without the presence of any

p53 expression. We also could have used different DNA alkylating reagents other than Temozolomide. Further, we showed that the OVCA-420 cells were overall less sensitive to PARPi and Temozolomide treatment, individually and in combination, compared to the other breast and ovarian cell lines. Perhaps these cells were not a proper comparison to treatment sensitivity and the mutant p53 could still have more oncogenic functions that have yet to be explored. Our results could also be explained by the amount of mutant p53 in the transfected ovarian cells. According to our western blots, the amount of p53 expression in the transfected cells was much less than the OVCA-420 cells. This suggests that perhaps a higher expression of the mutant p53 would be needed to show an effect on treatment sensitivity.

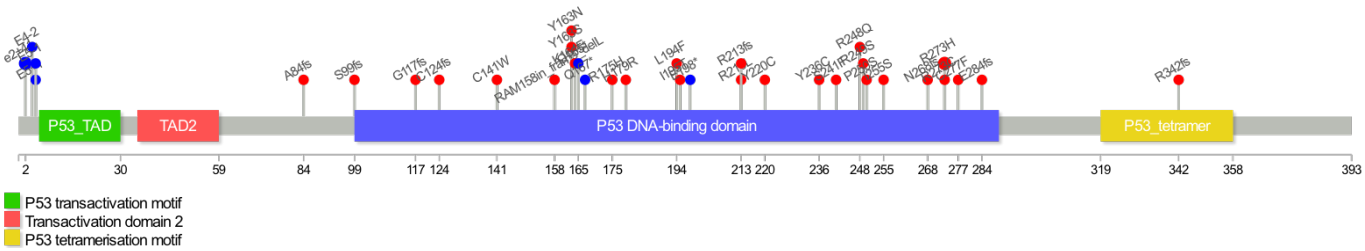
If we were to investigate this mutant p53 regardless of treatment response, we would have gathered evidence if this mutation causes the cells to have a more metastatic phenotype, for example if the cells were more proliferative or invasive. Proliferation assays measure how many cells divide, and therefore proliferate, over time while an invasion assay measures how a cell line is able to invade through a layer of Matrigel that mimics the tumor microenvironment. We hypothesize that the cells expressing mutant R273H p53 would proliferate and invade significantly more than their wildtype counterparts if mutant p53 were to have a distinct function from wildtype.

3.5 Conclusions

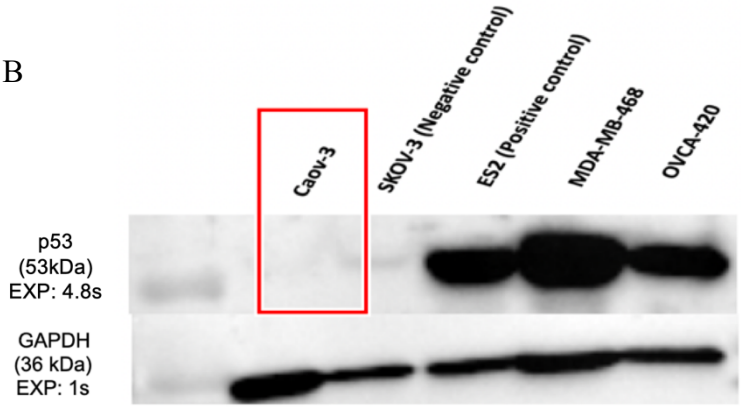
Ultimately, we did not gather enough evidence to support that this mutation is a GOF in ovarian cancer cells and this question warrants further study. We decided that since cells that contain the mutation are not susceptible to targetable treatments than it might not be translational for patients and will be unlikely to improve patient survival in the near future.

3.6 Figures

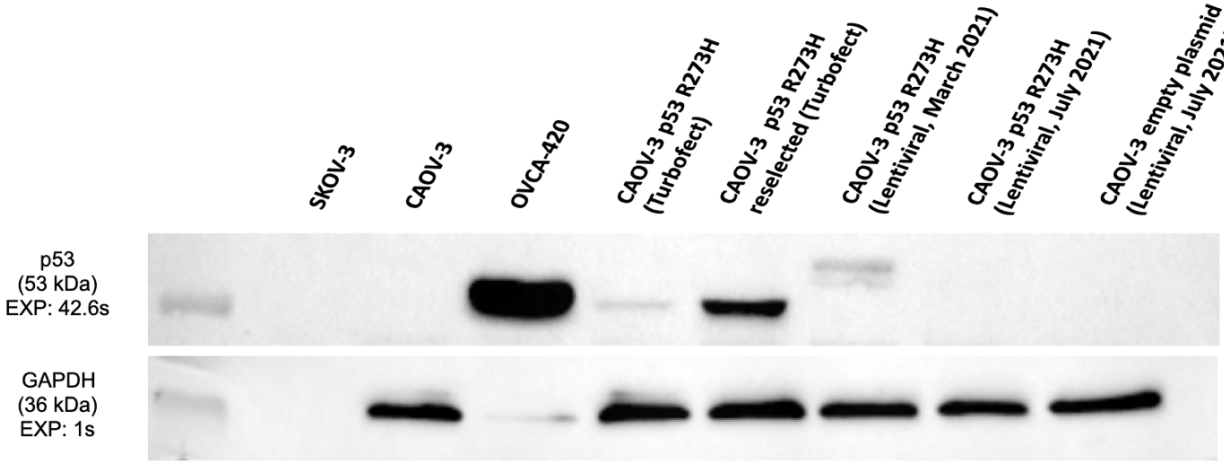
A



B



C



D

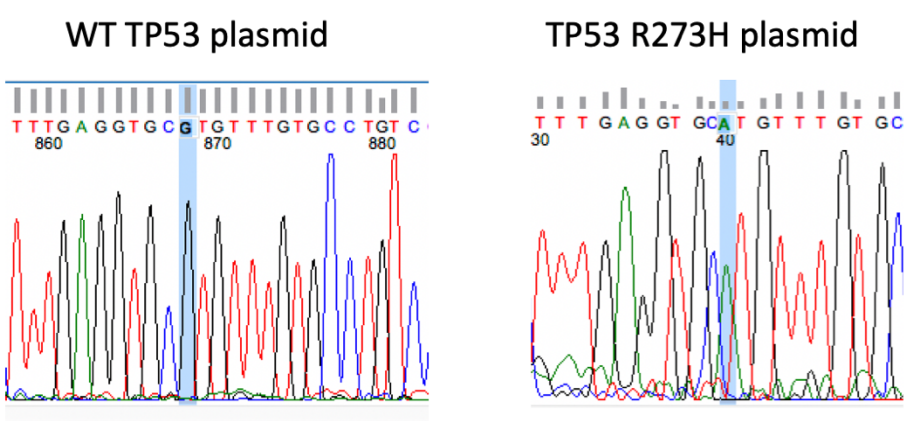
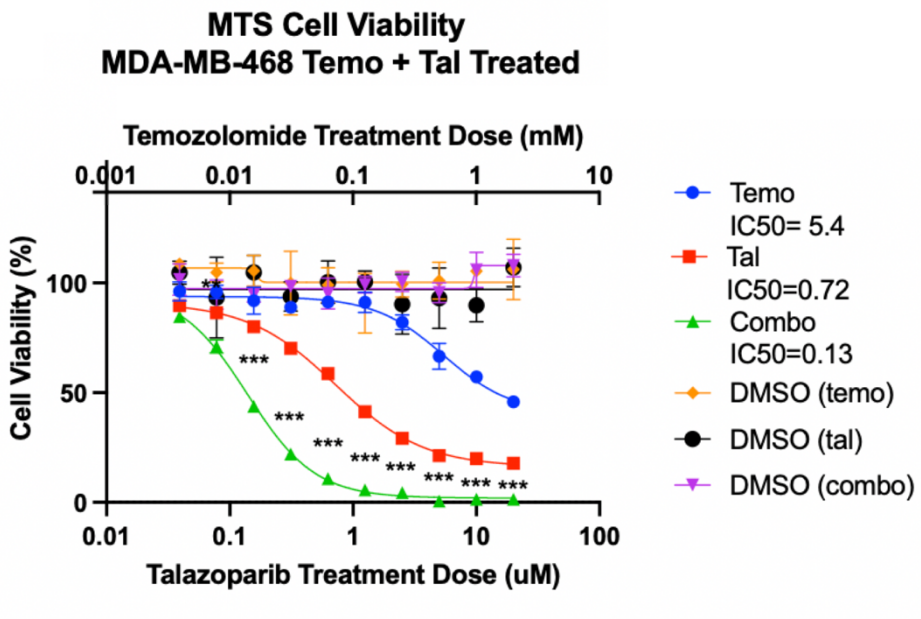


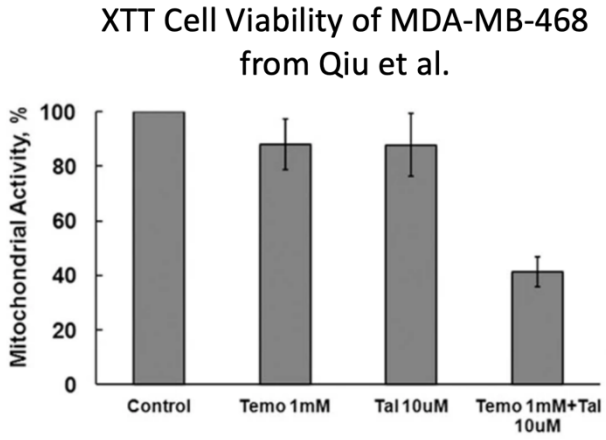
Figure 3.1: TP53 mutations in HGSC patients and p53 protein in ovarian and breast cancer cell lines

A. Lollipop plot displaying all TP53 mutations present in 39 HGSC patients that were WES, highlighting that 3 patients had the same R273H hotspot mutation in the DNA binding domain
B. Western blot displaying presence or absence of p53 protein in cell lines CAOV-3, SKOV-3, ES2, MDA-MB-458, and OVCA-420. CAOV-3 cells do not express a p53 protein compared to other ovarian cancer cell lines as negative (SKOV-3) and positive (ES2) controls for p53 protein expression.
C. Western blot for p53 with plasmid and lentiviral transfected CAOV-3 cell lines created to expression TP53 R273H
D. Sanger sequencing chromatograms validating the presence of the R273H mutation in the plasmid from the standard transfection

A


MTS Cell Viability
MDA-MB-468 Temo + Tal Treated
Temozolomide Treatment Dose (mM)
0.001
0.01
0.1
1
10
Cell Viability (%)
100
50
0
0.01
0.1
1
10
100
Talazoparib Treatment Dose (uM)
***
Temo
IC50= 5.4
Tal
IC50=0.72
Combo
IC50=0.13
DMSO (temo)
DMSO (tal)
DMSO (combo)

B


XTT Cell Viability of MDA-MB-468
from Qiu et al.
Mitochondrial Activity, %
100
80
60
40
20
0
Control
Temo 1mM
Tal 10uM
Temo 1mM+Tal
10uM

C

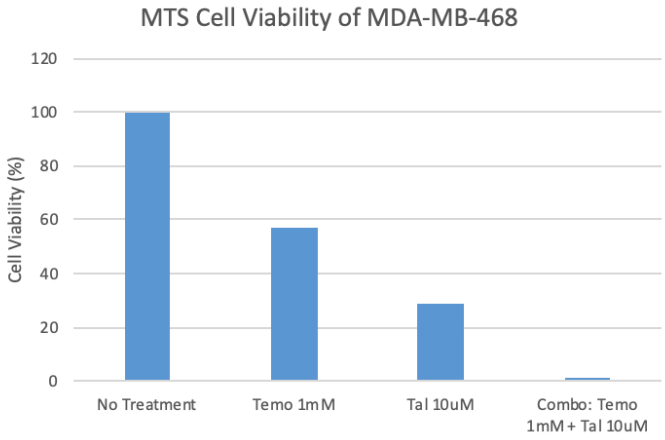
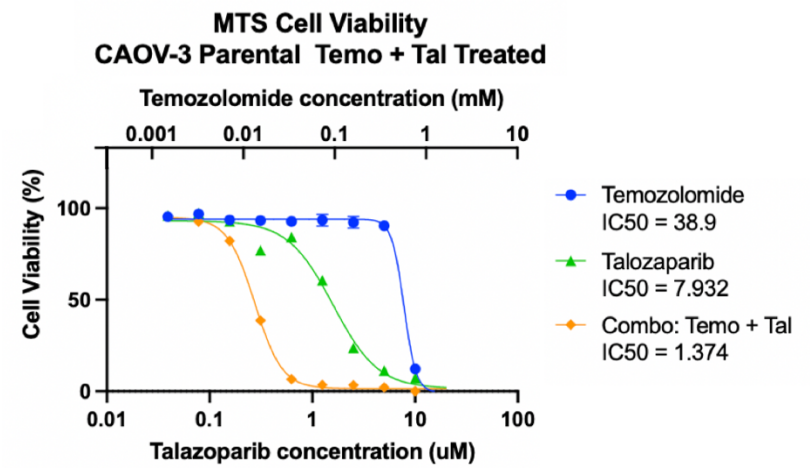

MTS Cell Viability of MDA-MB-468
120
100
80
60
40
20
0
Cell Viability (%)
No Treatment
Temo 1mM
Tal 10uM
Combo: Temo
1mM + Tal 10uM

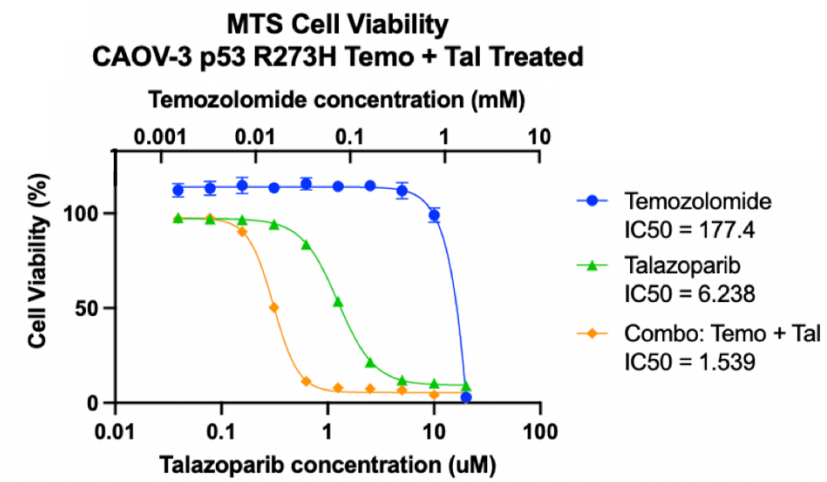
Figure 3.2: MAD-MB-468 are most sensitive to the combination of Talazoparib and Temozolomide treatment

A. MTS cell viability curve for MDA-MB-468 cells treated with Talazoparib, Temozolomide, and their combination
B. Bar graphs comparing the cell viability from our MTS assay to published XTT cell viability at published treatment doses

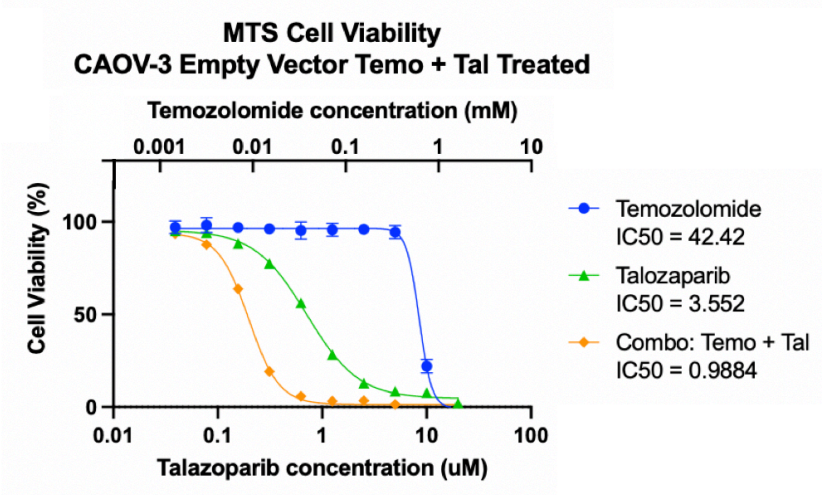
A


MTS Cell Viability
CAOV-3 Parental Temo + Tal Treated
Temozolomide concentration (mM)
0.001
0.01
0.1
1
10
Cell Viability (%)
100
50
0
0.01
0.1
1
10
100
Talazoparib concentration (uM)
Temozolomide
IC50 = 38.9
Talozaparib
IC50 = 7.932
Combo: Temo + Tal
IC50 = 1.374

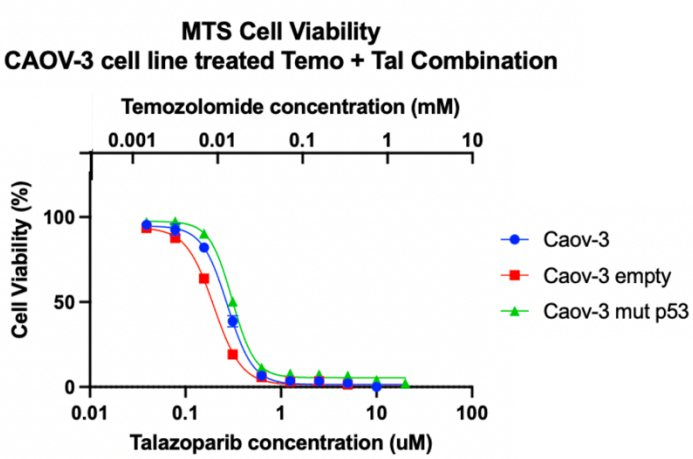
B


MTS Cell Viability
CAOV-3 p53 R273H Temo + Tal Treated
Temozolomide concentration (mM)
0.001
0.01
0.1
1
10
Cell Viability (%)
100
50
0
0.01
0.1
1
10
100
Talazoparib concentration (uM)
Temozolomide
IC50 = 177.4
Talazoparib
IC50 = 6.238
Combo: Temo + Tal
IC50 = 1.539

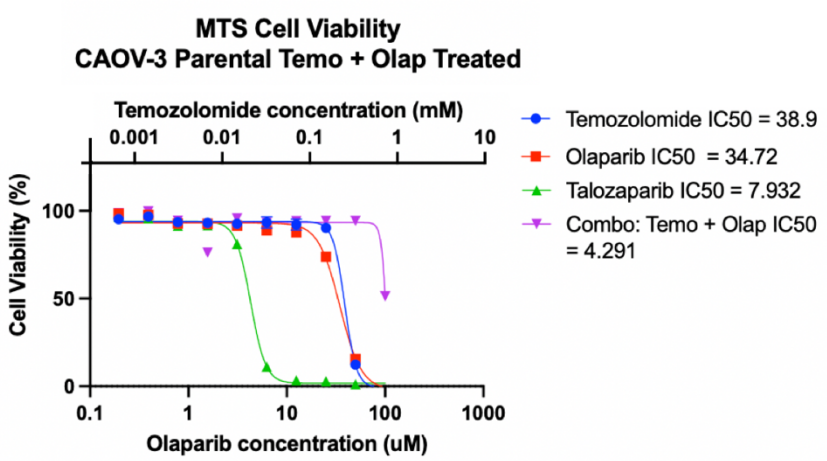
C


MTS Cell Viability
CAOV-3 Empty Vector Temo + Tal Treated
Temozolomide concentration (mM)
0.001
0.01
0.1
1
10
Cell Viability (%)
100
50
0
0.01
0.1
1
10
100
Talazoparib concentration (uM)
Temozolomide
IC50 = 42.42
Talozaparib
IC50 = 3.552
Combo: Temo + Tal
IC50 = 0.9884

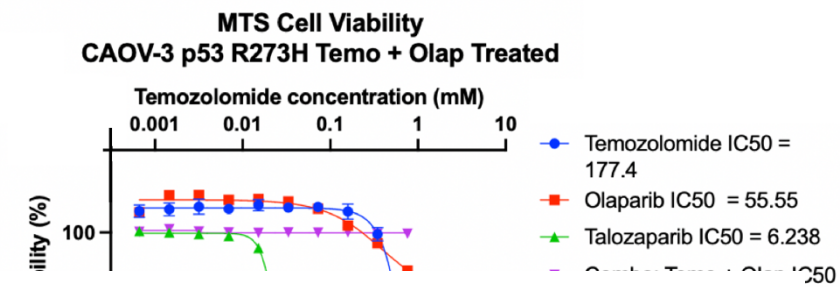
D


MTS Cell Viability
CAOV-3 cell line treated Temo + Tal Combination
Temozolomide concentration (mM)
0.001
0.01
0.1
1
10
Cell Viability (%)
100
50
0
0.01
0.1
1
10
100
Talazoparib concentration (uM)
Caov-3
Caov-3 empty
Caov-3 mut p53

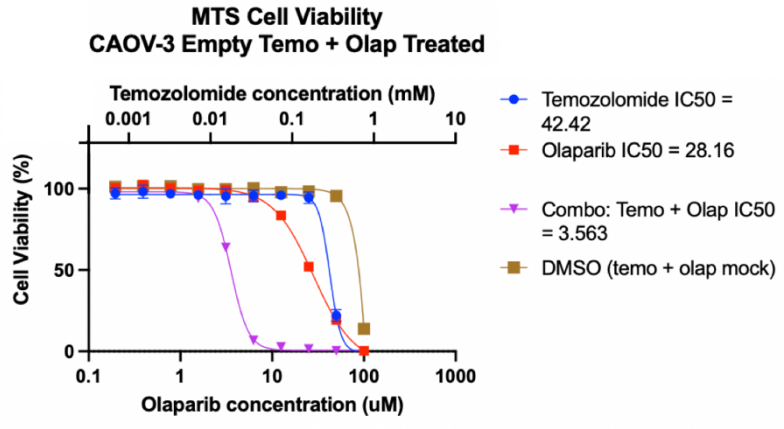
E


MTS Cell Viability
CAOV-3 Parental Temo + Olap Treated
Temozolomide concentration (mM)
0.001
0.01
0.1
1
10
Cell Viability (%)
100
50
0
0.1
1
10
100
1000
Olaparib concentration (uM)
Temozolomide IC50 = 38.9
Olaparib IC50 = 34.72
Talozaparib IC50 = 7.932
Combo: Temo + Olap IC50 = 4.291

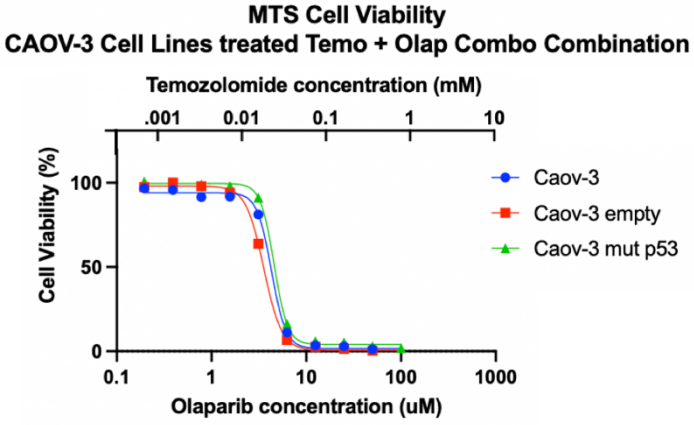
F


MTS Cell Viability
CAOV-3 p53 R273H Temo + Olap Treated
Temozolomide concentration (mM)
0.001
0.01
0.1
1
10
100
Temozolomide IC50 = 177.4
Olaparib IC50 = 55.55
Talozaparib IC50 = 6.238

G


MTS Cell Viability
CAOV-3 Empty Temo + Olap Treated
Temozolomide concentration (mM)
0.001
0.01
0.1
1
10
Cell Viability (%)
100
50
0
0.1
1
10
100
1000
Olaparib concentration (uM)
Temozolomide IC50 = 42.42
Olaparib IC50 = 28.16
Combo: Temo + Olap IC50 = 3.563
DMSO (temo + olap mock)

H


MTS Cell Viability
CAOV-3 Cell Lines treated Temo + Olap Combo Combination
Temozolomide concentration (mM)
.001
0.01
0.1
1
10
Cell Viability (%)
100
50
0
0.1
1
10
100
1000
Olaparib concentration (uM)
Caov-3
Caov-3 empty
Caov-3 mut p53

I

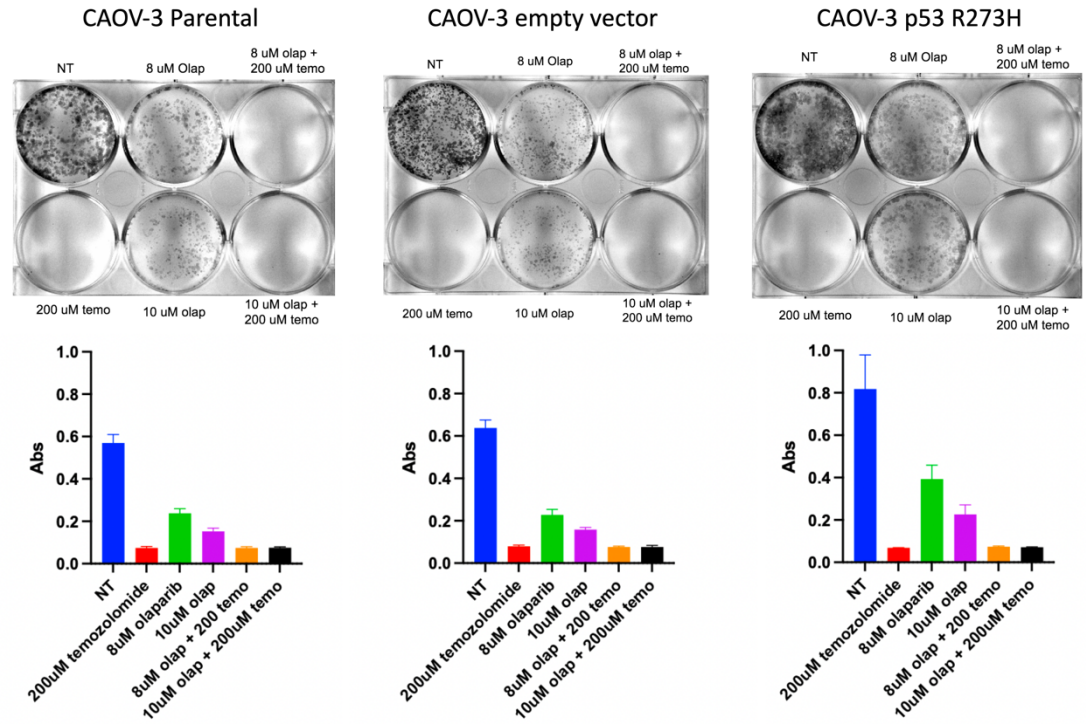
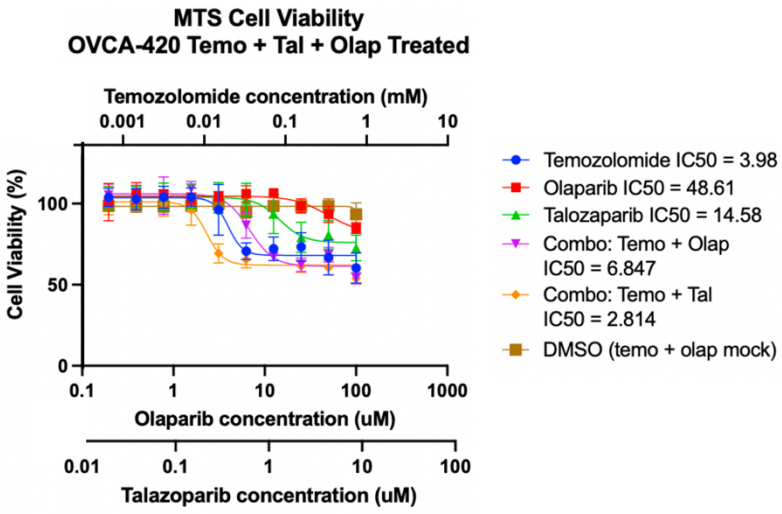

CAOV-3 Parental
CAOV-3 empty vector
CAOV-3 p53 R273H
NT
8 uM Olap
8 uM olap + 200 uM temo
200 uM temo
10 uM olap
10 uM olap + 200 uM temo
Abs
1.0
0.8
0.6
0.4
0.2
0.0
NT
200uM temozolomide
8uM olaparib
10uM olap
8uM olap + 200 temo
10uM olap + 200uM temo

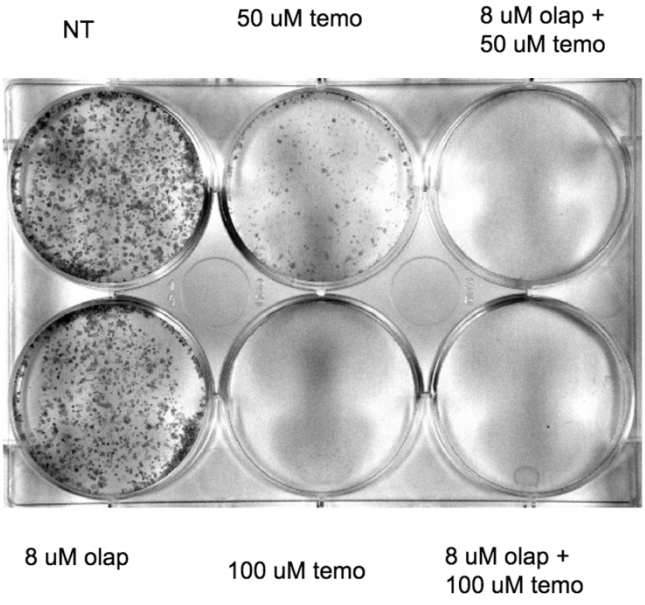
Figure 3.3: Cell Viability of transfected CAOV-3 cells treated with Talazoparib, Olaparib, and Temozolomide

A. MTS cell viability curve of parental CAOV-3 cells treated with Talazoparib and Temozolomide
B. MTS cell viability curve of CAOV-3 cells transfected with empty vector treated with Talazoparib and Temozolomide
C. MTS cell viability curve of CAOV-3 cells transfected with mutant TP53 R273H treated with Talazoparib and Temozolomide
D. MTS cell viability curve of CAOV-3 parental, empty vector, and mutant TP53 R273H comparison treated with combination of Talazoparib and Temozolomide
E. MTS cell viability curve of parental CAOV-3 cells treated with Olaparib and Temozolomide
F. MTS cell viability curve of CAOV-3 cells transfected with empty vector treated with Olaparib and Temozolomide
G. MTS cell viability curve of CAOV-3 cells transfected with mutant TP53 R273H treated with Olaparib and Temozolomide
H. MTS cell viability curve of CAOV3 parent, empty vector, and mutant TP53 R273H treated with combination Olaparib and Temozolomide
I. Clonogenic plates and corresponding bar graph showing cell viability of CAOV-3 parental, empty vector, or mutant p53 treatment response to Olaparib, Temozolomide or their combination

A


MTS Cell Viability
OVCA-420 Temo + Tal + Olap Treated
Temozolomide concentration (mM)
0.001
0.01
0.1
1
10
Cell Viability (%)
100
50
0
0.1
1
10
100
1000
Olaparib concentration (uM)
0.01
0.1
1
10
100
Talazoparib concentration (uM)
Temozolomide IC50 = 3.98
Olaparib IC50 = 48.61
Talozaparib IC50 = 14.58
Combo: Temo + Olap
IC50 = 6.847
Combo: Temo + Tal
IC50 = 2.814
DMSO (temo + olap mock)

B


NT
50 uM temo
8 uM olap +
50 uM temo
8 uM olap
100 uM temo
8 uM olap +
100 uM temo

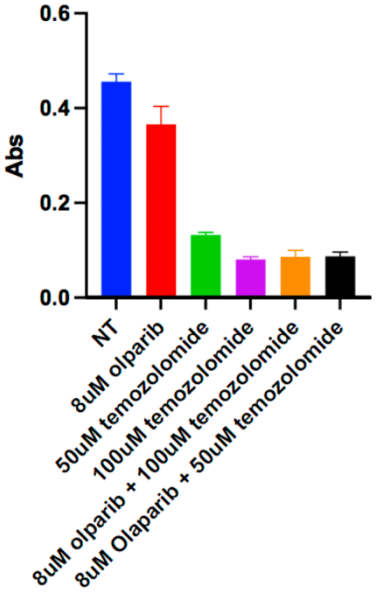

0.6
0.4
0.2
0.0
Abs
NT
8uM olparib
50uM temozolomide
100uM temozolomide
8uM olparib + 100uM temozolomide
8uM Olaparib + 50uM temozolomide

Figure 3.4: Cell Viability of OVCA-420 cells treated with combinations of Olaparib, Talazoparib, and Temozolomide

A. MTS cell viability curve of OVCA-420 cells treated with Olaparib, Talazoparib, Temozolomide, and their combination

Clonogenic plate and corresponding bar graph comparing OVCA-420 cells treated with Olaparib, Temozolomide and their combinations

3.7 References

1. Bell, D., et al., *Integrated genomic analyses of ovarian carcinoma.* Nature, 2011. **474**(7353): p. 609-615.
2. Lane, D. and A. Levine, *p53 Research: the past thirty years and the next thirty years.* Cold Spring Harb Perspect Biol, 2010. **2**(12): p. a000893.
3. Silwal-Pandit, L., A. Langerød, and A.L. Børresen-Dale, *TP53 Mutations in Breast and Ovarian Cancer.* Cold Spring Harb Perspect Med, 2017. **7**(1).
4. Tuna, M., et al., *Clinical relevance of TP53 hotspot mutations in high-grade serous ovarian cancers.* Br J Cancer, 2020. **122**(3): p. 405-412.
5. Turner, N., A. Tutt, and A. Ashworth, *Hallmarks of 'BRCAness' in sporadic cancers.* Nat Rev Cancer, 2004. **4**(10): p. 814-9.
6. Zhou, G., et al., *Gain-of-function mutant p53 promotes cell growth and cancer cell metabolism via inhibition of AMPK activation.* Mol Cell, 2014. **54**(6): p. 960-974.
7. Zhu, J., et al., *Gain-of-function p53 mutants co-opt chromatin pathways to drive cancer growth.* Nature, 2015. **525**(7568): p. 206-11.
8. Nakayama, M. and M. Oshima, *Mutant p53 in colon cancer.* J Mol Cell Biol, 2019. **11**(4): p. 267-276.
9. Kang, N., et al., *Mutant TP53 G245C and R273H promote cellular malignancy in esophageal squamous cell carcinoma.* BMC Cell Biol, 2018. **19**(1): p. 16.
10. Polotskaia, A., et al., *Proteome-wide analysis of mutant p53 targets in breast cancer identifies new levels of gain-of-function that influence PARP, PCNA, and MCM4.* Proc Natl Acad Sci U S A, 2015. **112**(11): p. E1220-9.
11. Qiu, W.G., et al., *Identification, validation, and targeting of the mutant p53-PARP-MCM chromatin axis in triple negative breast cancer.* NPJ Breast Cancer, 2017. **3**.

12. Xiao, G., et al., *Gain-of-Function Mutant p53 R273H Interacts with Replicating DNA and PARP1 in Breast Cancer.* Cancer Res, 2020. **80**(3): p. 394-405.

13. Olive, K.P., et al., *Mutant p53 gain of function in two mouse models of Li-Fraumeni syndrome.* Cell, 2004. **119**(6): p. 847-60.

14. Baker, S.J., et al., *Suppression of human colorectal carcinoma cell growth by wild-type p53.* Science, 1990. **249**(4971): p. 912-5.

15. Sarbassov, D.D., et al., *Phosphorylation and regulation of Akt/PKB by the rictor-mTOR complex.* Science, 2005. **307**(5712): p. 1098-101.

16. Junk, D.J., et al., *Different mutant/wild-type p53 combinations cause a spectrum of increased invasive potential in nonmalignant immortalized human mammary epithelial cells.* Neoplasia, 2008. **10**(5): p. 450-61.

Chapter 4: Functional genomic screen of kinases in high-grade-serous ovarian cancer metastasis

Declarations

Emilee Kotnik performed analyzed the patient data and screen data selection of candidates and designed and performed all secondary screen experiments including transfections, scratch assays, western blots, qRT-PCRs, and cell viability assays. Daniel Wilke performed the original genomic kinase screen with High Throughput Screening Center (Washington University in St. Louis).

4.1 Introduction

High-Grade Serous Ovarian Cancer (HGSC) is a highly metastatic cancer with the majority of patients presenting in advanced stages. Currently there are limited targeted treatment options for patients, especially for women with high metastatic tumor burden. HGSC tumors must overcome certain barriers to metastasize. In brief, tumor cells must break away from the original tumor, attach to the mesothelial layer of cells in the perineal cavity, clear away those mesothelial cells, and then invade into the stroma [1] In order to improve patient outcomes, there is a need to identify actionable, druggable targets that influence these steps of metastasis. This way we more precisely target metastatic tumors and develop maintenance therapies to help prevent further metastasis and mediate chemoresistance.

Thus far, the field has characterized primary tumors in HGSC and in Chapter 2, we discussed our attempt to further characterize metastatic tumors, which have not been extensively studied in ovarian cancer. Patient tumor databases are a powerful and essential tool for discovery-based projects, but they also require a lot of coordination, consistency, deep sequencing, and power in order to find novel targets and unique features that can be used for precision medicine. [2] In the lab, performing high-throughput

screen is another approach we can use to complement patient data and expand the possibilities of finding targets and drug compounds with high-powered, systematic, controlled methodology. [3]

Kinases serve as actionable targets because inhibitors can be developed that mimic their ligands or block their receptors to reduce their functioning and signaling. Several kinases, such as tyrosine kinases and aurora kinases, have previously been implicated in ovarian cancer and seem to influence tumor cell metastasis. [4-6] For example, Aurora kinases are serine-threonine kinases that regulate the mitotic spindle formation. They are highly expressed in ovarian cancers, and other cancer types, and clinical trials on Aurora kinase inhibitors in combination with other treatments have shown to improve progression free survival. [4] In particular, the Fuh Lab has studied tyrosine kinases DDR2 and AXL. We have studied their role in metastasis and the use of drug inhibitors to target these kinases. [7-9] Reducing the expression of DDR2 in OVCAR4 cells can increase the production of CD8+ T cells and has been shown to increase sensitivity to PDL-1 inhibitors. [10] There are several small molecule inhibitors that target DDR1 and DDR2 that are in development. [11] AXL is another tyrosine kinase has been shown to be highly expressed in ovarian cancer and particularly in metastatic tumors. Inhibiting AXL genetically *in vivo* prevented the formation of new metastatic tumors and decreased metastatic invasion. Therefore, there is evidence that AXL inhibition and therapy can prevent metastatic tumor progression. [8]

Our study aims to identify actionable targets that influence the steps of metastasis, so we can develop drugs that target metastatic tumors and serve as maintenance therapies to help prevent further metastasis, mediate chemoresistance, and improve patient outcomes. We have utilized a functional genomic screen for tumor cell attachment and invasion to identify additional kinases that may promote or suppress tumor cell attachment and invasion. We then designed a secondary screen to validate the results of this original screen to identify candidate kinases that promote metastasis, which is measured by their ability to affect tumor cell migration.

4.2 Methods and Materials

4.2.1 Cell Culture and Maintenance

OVCAR8 cells were obtained from the National Cancer Institute. OVCAR8 cells were cultured in RPMI 1640 media (Sigma-Aldrich) supplemented with 10% FBS and 1% penicillin (pen) and streptomycin (strep) antibiotic when culturing and expanding cells. RPMI media with 10% FBS and no antibiotics was used during transfection and seeding cells prior to transfection. ES2 cells were cultured in McCoy's 5A Medium (1X, Gibco) media supplemented with 10% FBS and 1% pen/strep antibiotics. Transfections with ES2 cells were done with McCoy's with only 10% FBS and no antibiotics and cells were kept in antibiotic free media for the duration of the cell seeding, the transfection, and during the scratch assay. All Cells are incubated at 37°C with 5% CO_2.

4.2.2 Functional genomic screen design, statistical analysis, and controls

Primary fibroblasts cultured from the human omentum were plated as one component to recapitulate the tumor microenvironment. An arrayed small-interfering RNA (siRNA) panel targeting 719 genes encoding kinases was used. The screen was performed on tumor cells derived from human, high-grade serous, epithelial ovarian cancer cells stably labeled with GFP. Following exposure to siRNA, ovarian cancer cells are overlaid upon a mixed submesothelial (NOFs, collagen I, and fibronectin) matrix. We did this process in 3 ovarian cancer cell lines, including OVCAR8, ES2, and A2780 across 9 panels for two libraries. Both libraries had distinct siRNAs for each kinase, and each well had a pool of 2 oligos per kinase, adding up to a total of 4 siRNAs per kinase. The functional readout is a decrease in attachment/early invasion. Cell death was controlled by staining with resazurin. Data was analyzed by expressing the effect of increased or decreased attachment using the median average deviation (MAD).

In order to account for background fluorescence, the fluorescent background was subtracted from each plate which was represented by the wells which did not contain GFP labeled cancer cells. Values were then normalized to the siNegative controls on a plate by plate basis. Finally, it was determined how many Median Absolute Deviations a given value was from the overall experiment wide median. Thresholds to identify outliers were determined for each cell line based on the dynamic range of the measured fluorescence values. Therefore, MAD values were calculated by dividing each fluorescence value by the mean fluorescence of transfection reagent mock to normalize. Then those values are averaged for all 3 replicates. The following equations were used:

$$\text{median absolute deviation: } \text{median}(\text{abs}(\text{value}-\text{median}_1)) = \text{median}_{\text{abs}}.$$

$$\text{average deviation} = (\text{mean}-\text{median}_1)/\text{median}_{\text{abs.}}$$

For controls, one panel included the fibroblasts alone without any GFP labeled cell lines, giving a baseline for no florescence. We used RNAiMAX Mock with all the transfection reagent to demonstrate that the transfection wouldn't affect basal level tumor cell attachment and we used siRNA negative to measure for baseline tumor cell attachment. We used two positive controls, with integrin and cantharidin, which are both known to inhibit tumor cell attachment so these controls served as a comparison of regents that should block tumor cell attachment. Lastly, siDeath was used as a control to measure transfection efficiency. The RNAiMAX Mock control wells were used to normalize the Alamar blue staining, so we could better determine if lack of tumor cell attachment or invasion was due to the siRNA knockdown of a kinase or due to cell death from transfection or technical errors. The MAD values for these controls are included in Figure 2.

4.2.3 Secondary screen siRNA transfections

For siRNA transfections, dsiRNAs from IDT were used to knockdown candidate genes and controls via transfection reagent lipofectamine 2000 (Invitrogen 1mg/ml REF#11668-027). The name and

reference numbers for the dsiRNAs used for our candidate genes are listed in Table 1. DsiRNAs were diluted in nuclease-free dulex buffer (IDT cat#11-01-03-01) to a 100 μM stock concentration and were heated at 94°C for 2 minutes, then allowed to cool to room temperature. A 20 μM working stock was made from the 100 μM stock and both dilutions were stored at –20°C.

For the transfection protocol, 24-hours prior to transfection OVCAR8 cells were plated in 10 cm plates (TPP) at a concentration of 1.8x10^6 cells per plate using RPMI media 10% FBS without antibiotics. Cells were counted with a Countess II (Life Technologies). Transfections were performed according to the Lipofectamine 2000 manufacturer's protocol. We added 0.1 nmol of dsiRNA was added to opti-MEM media (Gibco, REF#31985-070). We then added the volume of dsiRNA and the lipofectamine solution together, mixed gently by pipetting, and allowed to incubate for 20 minutes at room temperature. OVCAR8 cells were washed with DPBS (Gibco, REF# 14190-136) and received 5ml of fresh RPMI media with 10% FBS. After the incubation period, 3 ml of the transfection reagent and dsiRNA solution were added dropwise to OVCAR8 cells in 10 cm plates containing 5 ml of 10% FBS containing RPMI media. After 24 hours of transfection, RPMI media was supplemented again with 10% FBS and 1% pen/strep antibiotics until the scratch assay was performed. After scratches were made, cells were kept in serum free RPMI media with 1% pen/strep antibiotics.

For transfections with ES2 cells, the same protocol as described was followed, but prior to transfection cells were plated at a concentration of 4x10^5 cells/well in 6-well plates in triplicate and McCoy's media was used for culturing. When scratches were made, ES2 cells were kept in serum free and antibiotic free McCoy's.

4.2.4 Secondary screen scratch assay and analysis

Twenty-four hours post transfection, the cells were washed with DPBS lifted with 0.5% trypsin-EDTA (Gibco, REF# 25300-054) and plated in triplicate in 12-well plates (TPP) at a concentration of

4.5x10^5 cells/well for OVCAR8 cells and at 5x10^5 cells/well for ES2 cells. Twenty-four hours after plating (which is also 48 hours post-transfection) we used an unfiltered 200 μl pipet tip (Molecular BioProducts) was used to create a scratch down the center of each well. Images of the center of the scratch were taken for each well at timepoints 0, 8, or 24 hours after the scratch was created under 4x magnification using either a Visual Dynamix EXI-300 microscope or a Nikon Eclipse Ti microscope using QCapture Pro 6.0 software.

An ImageJ plugin was used to measure scratch areas and widths from the images taken at various timepoints, which is called "wound healing size tool" (https://github.com/AlejandraArnedo/Wound-healing-size-tool/wiki). [12] This software is able to measure the scratch space in pixels or measurements, depending on the type of file images being analyzed. Images take with Visual Dynamix EXI-300 microscope were measured in pixels, and images with the Nikon Eclipse Ti microscope were measured in inches from tif images. The percent change in area and/or width of the scratch over an 8- or 24- hour time period was calculated by taking the area measurement at 0 hours minus the 12-hour area measurement, dividing by the 0-hour measurement and multiplying by 100 to get a percentage. If 3 images of the same scratch were taken, then the average percentage change of those 3 images was used. To calculate the statistical differences between the controls and candidate genes scratches, we used GraphPad Prism 7 software to calculate a Welch's correlation.

4.2.5 RT-qPCR

RNA was collected from cells 48-hours post-transfection using the Qiagen RNeasy plus mini kit. RNA was quantified by nanodrop and between 0.5-1 μg of RNA was converted to cDNA using the SuperScript IV First-Strand Synthesis System (Invitrogen, REF#18091050) according to the manufacturer's protocols. SYBR Green PCR Master Mix (Applied Biosystems) was used in the ABI detection system (Applied Biosystems) for real-time PCR reactions. Each target RNA was measured in

triplicate and quantified using the 2-delta-deltaCT methods. All primer pairs for gene targets are listed in Table 2.

4.2.6 Western Blotting

Cells that transfected but not utilized for the scratch assay triplicates were plated separately and harvested at 48 hours post-transfection to be analyzed. Cells were lysed using 8M urea lysis buffer and DTT and sonicated for a few seconds before being spun down in a centrifuge at 4C for 15 min at 13000 rpm. Protein amounts were quantified by Bradford assay and 50-70 μg of lysates were mixed with BME and subjected to SDS-PAGE. The gel was run for 30 min at 61V and another 60-90 minutes at 120V. The protein was then transferred to a nitrocellulose membrane at 85 V for 80 minutes. Membranes were blocked with a 10% milk solution diluted in TBST, washed with TBST, and probed with primary antibody (AXL, Cell Signaling #4977 and B-actin, Sigma #A1978) for 1 night at 4C. The blot was then washed and incubated with secondary mouse antibody for approximately 2 hours and washed with TBST. Primary and secondary antibodies were diluted in 5% milk and TBST. Protein signals were detected with the Pierce ECL Westering Blotting Substrate and imaged using chemiluminescence on a ChemiDoc (Bio-Rad Laboratories).

4.2.7 CellTiter Glo Assay

CellTiter Glo 2.0 luminescent cell viability assay (Promega, REF#G924C) was used to measure cell proliferation in the OVCAR8 cells after siRNA transfection of candidate kinases. OVCAR8 cells were seeded in RPMI media with 10% FBS and no antibiotic at a concentration of 1x10^4 cells/well in a 96-well plate. The next day the cells were transfected according to the Lipofectamine 2000 protocol. 48-hours post-transfection, 200 μl of the CellTiter Glo 2.0 reagent was added to each well. The plate was placed on an orbital shaker for 2 minutes, covered in foil, and them allowed to incubate at room

temperature for 10 minutes before reading luminescence on the GloMax Navigator instrument (Promega) with software version 3.1.0.

4.3 Results

4.3.1 Functional genomic screen identifies multiple targets for tumor cell attachment and invasion

Our lab performed a functional genomic screen on 719 kinases in a tumor cell attachment and invasion assay to identify kinases that effect tumor cell invasion and attachment. In this screen, we cultured primary fibroblasts from patient omentum and plated them on top of a collagen and fibronectin matrix, to mimic the tumor microenvironment. We added siRNA transfection reagents atop the fibroblasts and plated GFP-labeled ovarian cancer cells on top. The florescence of the ovarian cancer cells was measured and used to calculate a median average deviations (MAD) value. (Figure 1A) Kinases that were knocked-down and demonstrated a negative MAD value, attached to or invaded through the tumor microenvironment cell layers less than controls. Conversely, kinases that were knocked-down and demonstrated a positive MAD value, attached to or invaded through the tumor microenvironment cell layers more than controls. From this screen, we had several significant kinases with a total of 17 kinases had a MAD value of 3 and 19 hits that were below at -3 MAD value (Figure 1B).

In the functional genomic screen, both the GFP labeled ovarian cancer cells and patient fibroblasts were exposed to the siRNA transfection reagents, so knockdown of the kinase could have occurred in both cell types. This design was done purposefully so we could elevate kinases on both the tumor cell and stromal cell level at the same time. Additionally, since the cells were incubated for a couple days and allowed to invade, the screen MAD value results could also measure tumor cell proliferation and

attachment to the Matrigel as well as invasion. So in order to identify kinases that influence tumor cell metastasis, we decided to design a secondary screen to further investigate candidates kinases.

4.3.2 Screen results and prioritization of candidates for secondary screen

Considering the above limitations, we used a known marker of invasion that was also screened to find candidate kinases to investigate in a secondary screen. We focused on kinases that had negative MAD values, because this would mean that when that kinase is knocked-down there is less tumor cell attachment and invasion. Therefore, this knockdown mimics the function of an inhibitor drug and has more potential to be therapeutically translational. The Fuh Lab has previously shown that the tyrosine kinase, DDR2, can promote tumor cell invasion and mesothelial cell clearance in ovarian cancer [9-11] In the screen, DDR2 had a MAD of -1.56 and -1.47 in ES2 cells (values form each library performed during the functional genomic screen). We then estimated that kinases in a range around -1.5 would be a priority and give us a more relaxed cut-off to find more potentially interesting candidates.

We identified 19 kinases with MAD values less than -1.5 in ES2 cells and 25 kinases in OVCAR8 cells. We narrowed this list further looking at their expression levels in ES2 and OVCAR8 cells in cbioportal and with published literature to find kinases with functions that could be involved in cancer mechanisms. [13, 14] Table 4 details the top 10 candidates we identified to pursue further and lists their function from GeneCards. [15, 16] In the end, we were able to test 4 candidate kinases that were selected based on their cell viability comparison to DDR2. These kinases were BUB1, PFKP, ITPK1, and PAK6.

4.3.3 Secondary screen design and troubleshooting in ES2 cells

We utilized a wound-healing (henceforth called a scratch assay) to determine if the candidate kinases influence ovarian cell migration when silenced by siRNAs. We transfected ES2 and OVCAR8 cells with siRNA and assessed the cell's ability to migrate into the scratch area over time. We

hypothesized that kinases that promote metastasis and had negative MAD values in the functional genomic screen would show less tumor cell migration when knocked-down by siRNA.

To develop our secondary screen protocol, we decided to use AXL as a positive control, since it is expressed in both ES2 and OVCAR8 and has previously been shown by our lab to promote metastasis and migration in Boyden chamber assays [7, 8, 17] Figure 2A shows the scratch assay performed 48-hours post siRNA transfection at time points 0, 8, and 24 hours of the scratch through confluence cells. Replicates of the percentage of wound-closure area over the 8-hour time period are shown in the box and whisker in Figure 2B. There were no significant differences between the percent change in wound areas between the siRNA knockdown of AXL and the non-transfected ES2 cells or the ES2 transfected with a non-targeting siRNA, despite sufficient knockdown of AXL at the RNA and protein level shown in Figures 2C and 2D.

Next, we attempted the screen with ES2 cells again and testing one candidate, ULK2. During transfection, many cells died but we were able to perform the scratch assay in duplicate and showed significant knockdown with siRNA against AXL and ULK2 (Figure 2C) 24-hours post-transfection and at the protein level for AXL 48-hours post-transfection (Figure 3D). After 8 hours, there visually was less wound-closure in samples with knockdown of AXL compared to the non-targeting control, although this was not statistically significant. Visually, there did not seem to be less migration between ULK2 knockdown and the negative control. After 24 hours, the non-targeting control scratch had completely closed, while the ULK2 knockdown was nearly closed and AXL knockdown sample still had a visual scratch. These experiments showed that we could transfect the ES2 with siRNAs and use AXL as a positive control (Figures 3A and 3B).

We observed that the ES2 cells' morphology changes when they are stressed by creating a wound and they do not tolerate serum free media conditions well. We wanted to keep cells in serum free media after scratching to be a true measure of migration, rather cell migration. So we ultimately decided from

these observations and the experiments described above that ES2 cells were not conducive to a wound-healing assay.

4.3.4 Secondary screen results in OVCAR8 cells

Instead, we adapted our scratch protocol to OVCAR8 cells. We tested some preliminary candidate kinases first, which consisted of AURKB, CHEK1, FGFR1, and ULK2. Ultimately, we decided not to pursue investigating these candidate kinases further since multiple testing with them changed the cell morphology, likely due to cell cycle dysregulation, too much to be able to properly assess migration. AURKB and CHEK1 have already been implicated in ovarian cancer previously in the literature [18, 19] The siRNA transfections of these candidates were successful and showed significant knockdown according to the RT-qPCR on cDNA collected from the cells at the time of the scratch was made, which is 48-hours post-transfection (Figure 5B). Lysates collected at the same time, 48-hours post-transfection, showed that our positive control AXL was almost completely knocked down (Figure 5C). Images from these scratches at timepoints 0, 24, and 48 hours post-scratch showed variable results across the scratches and cells were too connected to be able to measure wound-closure with our imageJ plugin. (Figure 5A) Therefore we showed that we can get sufficient knockdown with siRNA transfection in these cells, but we decided to pursue other candidate kinases that didn't affect the cell cycle and cell morphology as much. We performed a cell viability assay, CellTiter Glo, to see if siRNA knockdown of our other top 10 candidates (Table 3) affected cell proliferation. We picked 4 candidates to test in the secondary screen based their cell proliferation results comparable to AXL, our positive control.

In the final iteration of the secondary screen, we knocked-down BUB1, PFKP, ITPK1, and PAK6 in OVCAR8 cells. Although the siRNA transfections were inconsistent between gene targets, according to the RT-qPCR results, there was significant knockdown of gene expression in BUB1, ITPK1, and PAK6. (Figure 6D) While the transfections for AXL and PFKP showed reduced gene expression, they were not significantly less expressed. With transfection efficiency taken into account, we cannot make conclusions

about the percent change in area for the candidate gene PFKP. The candidate gene ITPK1 had a moderate reduction in gene expression and showed significant wound closure after 8 hours, but not after 24 hours. The loss of significance between 8- and 24- hours could be explained by the siRNA transfection waning over time. Candidates BUB1 and PAK6 showed highly significant knockdown by siRNA and significantly less change in their scratch wound closure over both the 8-hour and 24-hour time periods when compared to the negative control with the non-targeting siRNA. (Figure 6A-C) These candidates we can determine that when knocked-down, the ovarian cancer cells showed less migration over a 24-hour time period.

4.3.5 Overlap of patient data and functional genomic screen results to prioritize candidates

Additionally, we also attempted to find patterns between the patient sequencing data our lab collected and the screen results. From Chapter 2 of this dissertation, we had analyses from the WES and RNA-seq data from 37 patients with matched primary and metastatic tumors collected during their debulking surgery. We attempted to see how many kinases screened were among the recurrent single nucleotide variants (SNVs), differentially expressed genes, and recurrent gene fusions identified in the patient sequencing analyses. In total, there were 90 kinases screened with genes that were recurrently mutated among the patient data (61 kinases were from ST survivors, 20 were from metastatic tumor samples, and 9 kinases in tumors that were both metastatic and ST survivor samples). There were 15 kinases screened that were differentially expressed between ST and LT survivors in metastatic tumors and 23 kinases screened that were differentially expressed between ST and LT survivors among the primary tumors. Among the recurrent gene fusions predicted among the patient sequencing data, there were 15 kinases that were screen. To further narrow these potential candidates, we focused on kinases that had extremely high or low MAD values in at least 2 of the cell lines used in the functional genomic screen.

This filtering resulted in a total of 6 kinases (3 from recurrent SNVs, 2 from differentially expressed genes, and 1 from the recurrent gene fusion data). (Figure 1C)

These 6 kinase genes were *PIK3C2B, WNK2, CDKL2, MAST4, PACSIN2,* and *PI4KA* and a description of their function, results the patient data and screen, and a summary of the literature of them are displayed in Table 3. Most of these potential candidates were not good targets to pursue for further study, since they had conflicting evidence between the patient data and their results from the screen or there was a lack of literature on their function or relevance to cancer mechanisms. However, PIK3C2B has been studied in ovarian cancer and may be therapeutically targetable. [20] We determined that these candidates did not warrant further validation and that having hard cut off for MAD values and a smaller cohort of patient data meant that prioritizing candidates in this manner wouldn't be logical at the time.

4.4 Discussion

We performed a functional genomic screen on over 700 kinases to identify kinases that have the potential to influence metastasis in the context of the tumor microenvironment. This initial screen had limitations in its design, namely that results could have been affected by the fibroblast kinase knockdown and could have measured proliferation as well as migration and invasion. We attempted to prioritize candidates using patient data described in Chapter 2 of this dissertation, however those candidates had a lack of supporting evidence to be translational targets. We chose 4 candidates, based on having significantly negative MAD values in at least two cells lines in the screen, peer-reviewed literature, and cell proliferation results. Kinases BUB1, PFKP, ITPK1, and PAK6 were validated in a secondary screen that measures migration in a wound-healing scratch assay.

We acknowledge that a scratch migration assay does not consider the context of the tumor microenvironment like the original screen. This assay is performed in a dish without a cell matrix, but as an initial validation does show that some of our candidate kinases can influence cell migration and have

the potential to be targeted for treatments. The migration assay is also a cost-effective assay that could gather evidence to support the investigation into these kinases and it was conducive to the time sensitivity of siRNA knockdowns. In our case, the siRNA transfections are better than CRISPR knock-outs because they better recapitulate the activity of kinase inhibitors.

When designing this secondary screen, we observed variable results with positive controls and candidate kinases in ES2 cells, and instead performed the assay with OVCAR8 cells. We completed the screen and determined that BUB1 and PAK6 kinases showed less cell migration when knocked down in OVCAR8 cells and could promote metastasis in ovarian cancer cell lines. For reproducibility, the positive controls and candidates PFKP and ITPK1 should be repeated.

BUB1 is necessary for the spindle assembly checkpoint during mitosis and contributes to chromosome alignment and segregation. [21] There has been published literature on BUB1 in ovarian cancer. BUB1 mRNA is broadly expressed in ovarian tumors and is co-expressed with AURKA and AURKB kinases, which as previously stated are also related in function to spindle formation. BUB1 expression is lower in post-chemotherapy ovarian tumors. [22] Downregulating BUB1 in hepatocellular carcinoma cells can inhibit cell proliferation, migration, and invasion [23] and in breast cancer BUB1 expression has correlated with poor clinical outcomes and metastases [24]. Another kinase, KIF4A has been associated with BUB1, because when knocked-down, BUB1 expression was also reduced and OVCAR3 cells showed less migration and cell viability. [25]

Our secondary screen provides further evidence that BUB1 can promote metastasis in ovarian cancer. Further studies would be necessary to investigate BUB1 in more ovarian cancer cell lines and in other steps of metastasis, like invasion, has it has been for other cancer types. After defining BUB1's ability to influence metastasis, further studies would focus on how its regulation of spindle formation and mitosis lead to the development of metastases and if therapeutically targeting BUB1 can be beneficial to HGSC patients.

PAK6 is a p21-activating kinase in the serine/threonine kinase family. So far, PAK6 has not been linked to ovarian cancer, but in cervical cancer tumor microarrays showed elevated expression of PAK6. Additionally, when overexpressed in HeLA cells, PAK6 promoted metastasis and showed less cell migration when knocked-down. This group suggests that PAK6 plays a role in regulating Wnt/B-catenin signaling. [26] An earlier study on prostate cancer showed evidence that PAK6 is required for cell-to-cell dissociation in prostate cancer cells and that it interaction with junctional proteins, suggesting that it can drive junction disassembly. [27] Recent studies in gastric cancer have associated PAK6 with chemo resistance through its role in DNA damage response, specifically in activating homologous recombination through ATR. They saw correlations between high expression of PAK6 with poorer stages of gastric cancer and with metastatic tumors in the lymph nodes. [28]

Our findings provide support evidence that PAK6 can influence migration in ovarian cancer and is the first time PAK6 has been associated with ovarian cancer. Like BUB1, further evidence is needed to define its role in all the steps of metastasis in more ovarian cancer cell lines and to see if it can be therapeutically targeted. Homologous recombination and DNA damage response are important mechanisms for ovarian cancer tumorigenesis , so linking PAK6 with cell junctions or cell migration and metastasis has potential to be useful in combination with other DNA damaging drugs like PARP inhibitors.

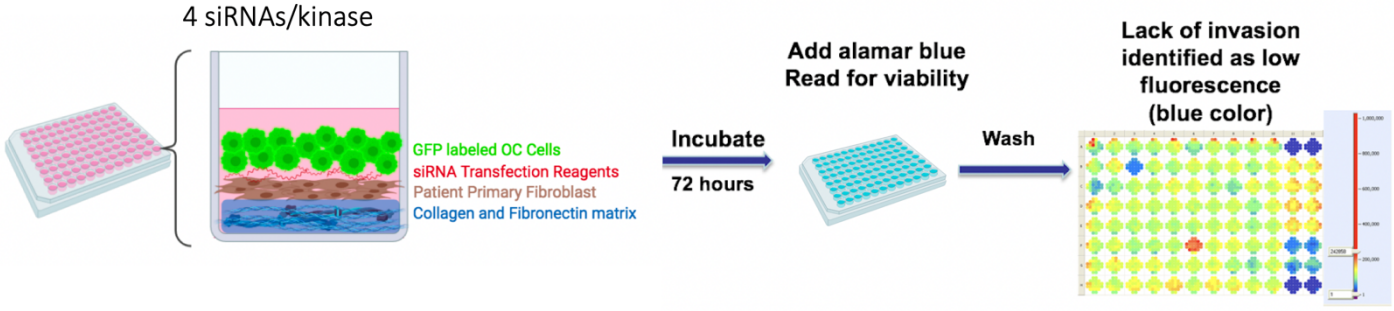
The results shown here for ITPK1 and PFKP were inconclusive, given that their siRNA knock-down was only moderately successful. ITPK1 has not yet been associated with cancer, but it does have a role in regulating Ca^{2+}-activated chloride channels in the plasma membrane [29] and seems to be required for necroptosis [30]. PFKP has been implicated as a mediator of cancer cell metabolism through its role in glycolysis [31, 32] and in particular, it may influence cell invasion and metastasis in triple negative breast cancer [33]. Despite our unsatisfying results, these kinases still have potential to be targetable kinases and warrant further consideration and study.

4.5 Conclusions

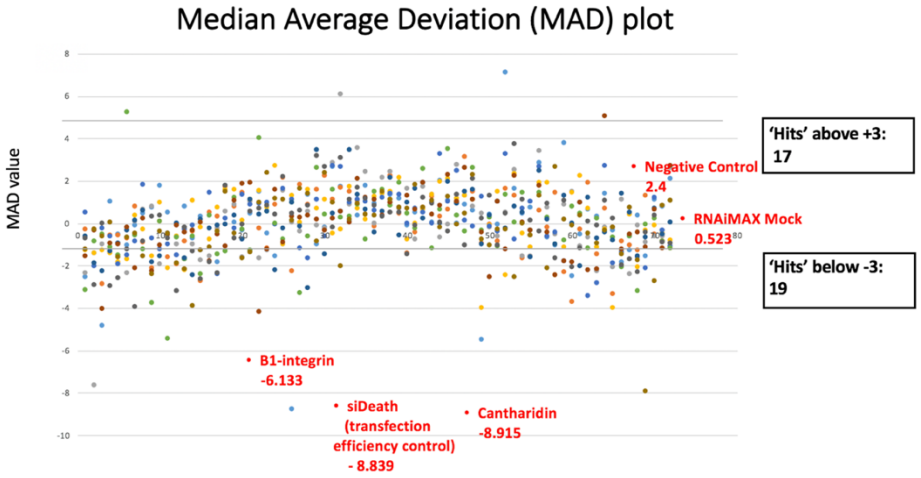
In this chapter, we used screens to identify kinases that regulate HGSC metastasis. We identified BUB1 and PAK6 as potential targets that affect tumor cell migration when knocked-down with siRNAs. These kinases will need to be further studied to elucidate how they influence migration and if they can regulate the other steps of tumor cell metastasis, like cell clearance and invasion. If promising and validated *in vitro* and *in vivo*, these kinases could be used as a potential target for drug inhibitors for HGSC patients.

4.6 Figures and Tables

A



B



C

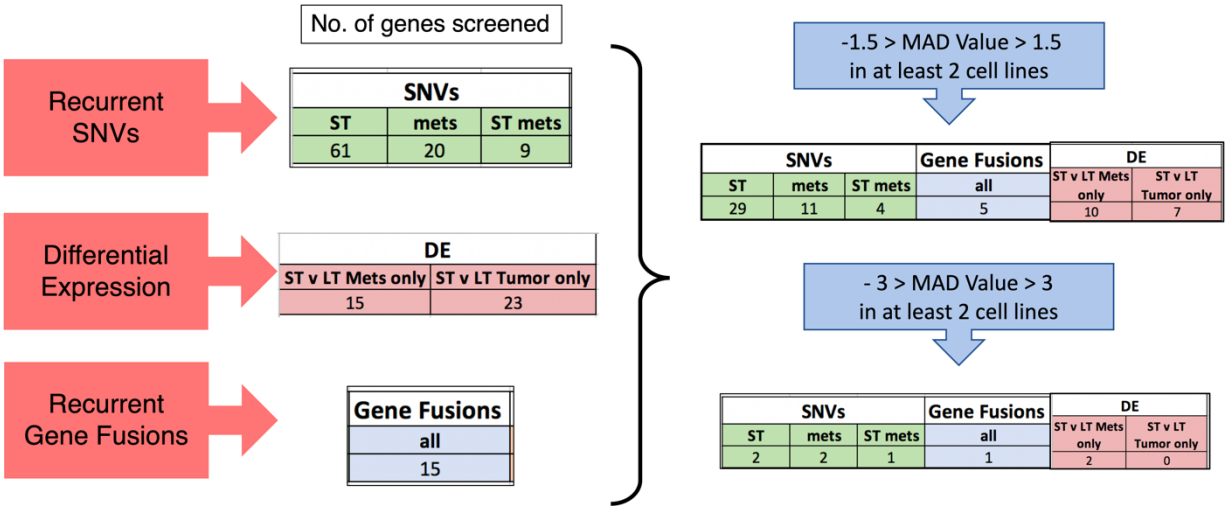
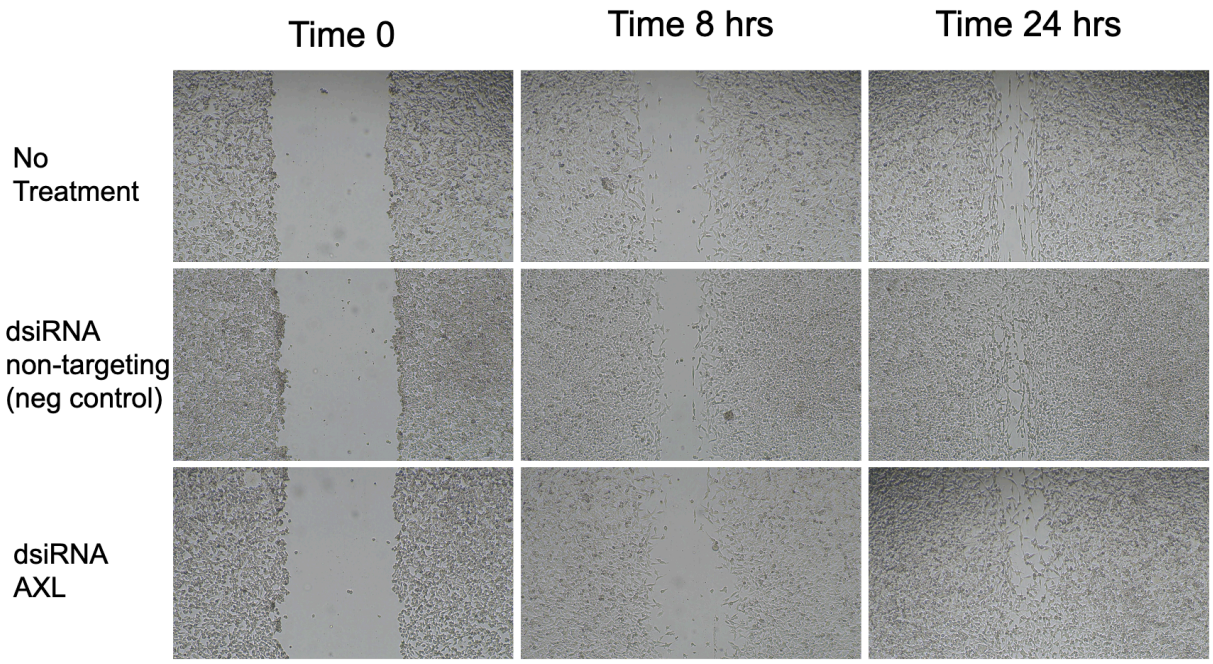


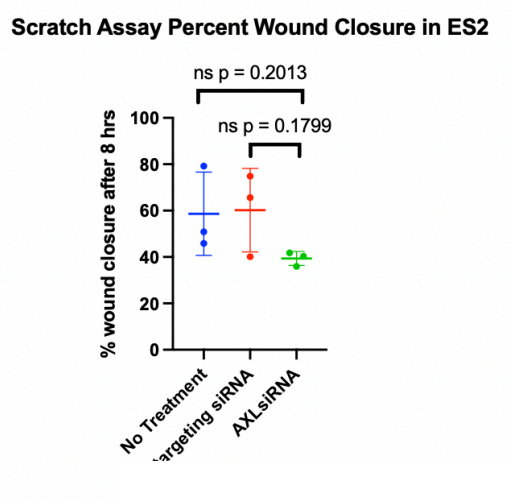
Figure 4.1: Functional genomic screen MAD values and patient overlap data

A. Graphical schematic of functional kinase screen design
B. Each dot represents a knocked-down kinase's MAD value. Screen control conditions are highlighted in red.
C. Schematic demonstrating the filtering and overlap between patient data and screen MAD value results

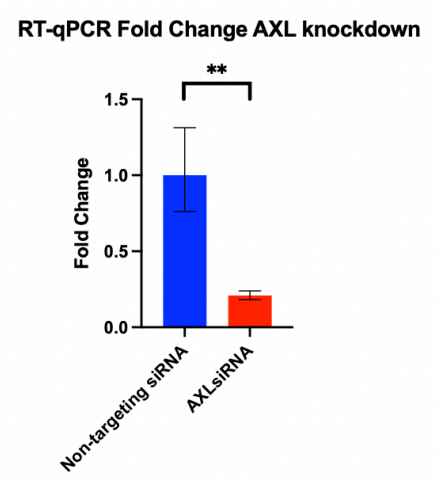
A


Time 0
Time 8 hrs
Time 24 hrs
No Treatment
dsiRNA non-targeting (neg control)
dsiRNA AXL

B


Scratch Assay Percent Wound Closure in ES2
ns p = 0.2013
ns p = 0.1799
% wound closure after 8 hrs
100
80
60
40
20
0
No Treatment
AXLsiRNA

C


RT-qPCR Fold Change AXL knockdown
**
Fold Change
1.5
1.0
0.5
0.0
Non-targeting siRNA
AXLsiRNA

D

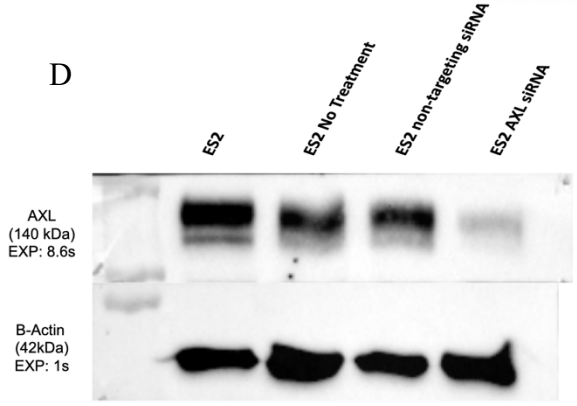
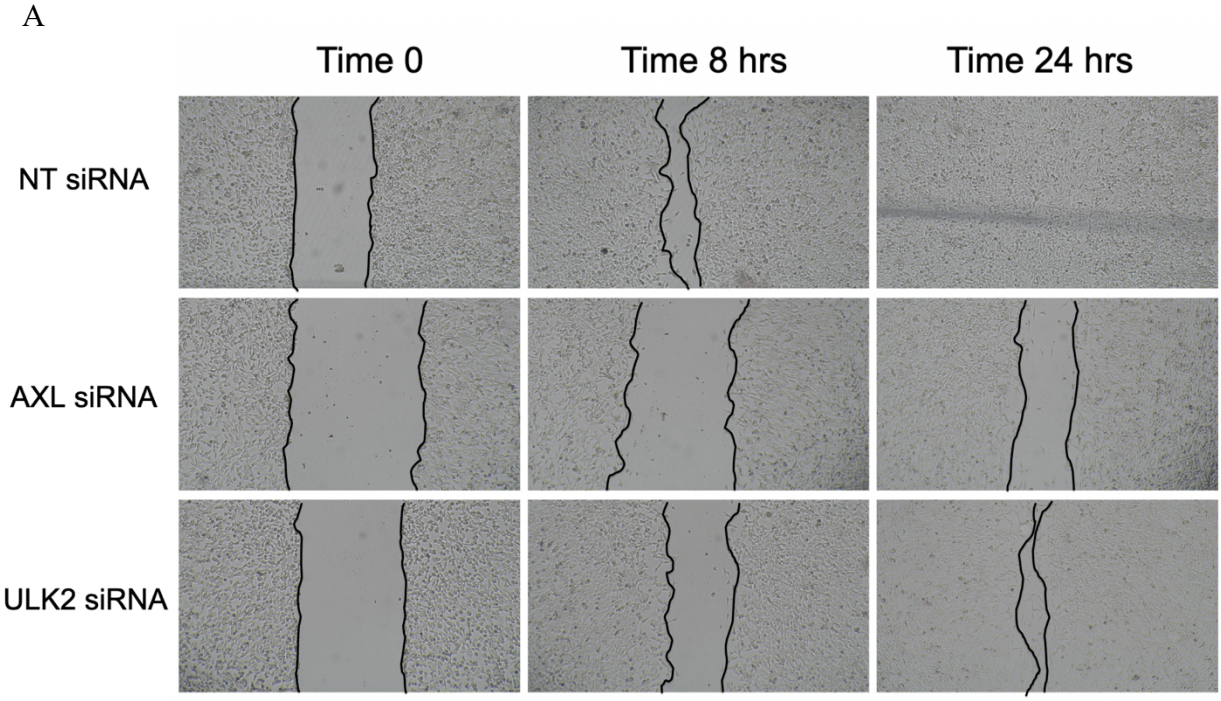

ES2
ES2 No Treatment
ES2 non-targeting siRNA
ES2 AXL siRNA
AXL (140 kDa) EXP: 8.6s
B-Actin (42kDa) EXP: 1s

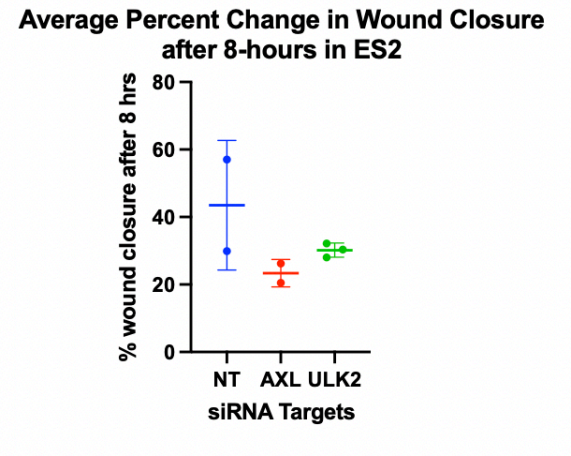
Figure 4.2: Secondary Screen with ES2 cells with AXL positive control

A. ES2 cell images at 4x showing scratch at time points 0, 8, and 24 hours. Time 0 is 48 hours post-transfection for the non-targeting siRNA and AXL siRNA
B. Box and whisker of percent change in the area of the scratch over the 8 hour time course
C. Fold change of the siRNA knockdown 24 hours post-transfection, prior to scratch
D. Western blot for AXL in cells 48 hours post-transfection, at time 0 for the scratch assay

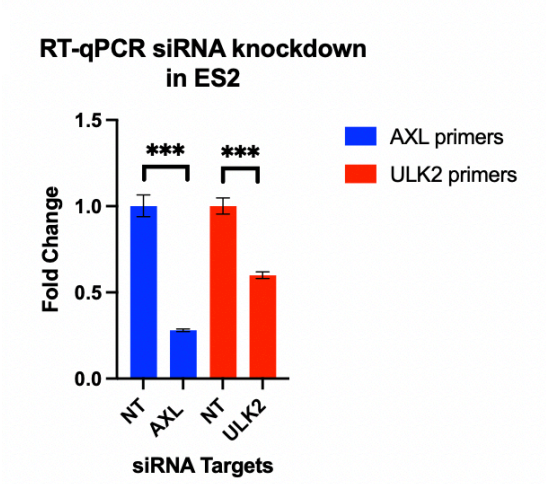
A


Time 0
Time 8 hrs
Time 24 hrs
NT siRNA
AXL siRNA
ULK2 siRNA

B


Average Percent Change in Wound Closure after 8-hours in ES2
% wound closure after 8 hrs
80
60
40
20
0
NT
AXL
ULK2
siRNA Targets

C


RT-qPCR siRNA knockdown in ES2
Fold Change
1.5
1.0
0.5
0.0
***
***
AXL primers
ULK2 primers
NT
AXL
NT
ULK2
siRNA Targets

D

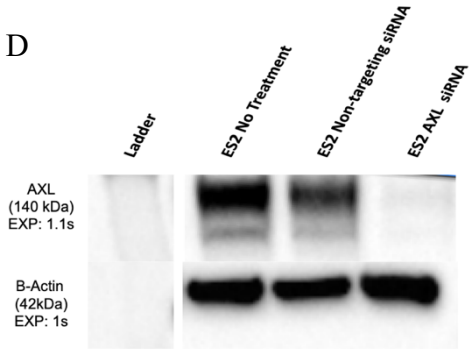
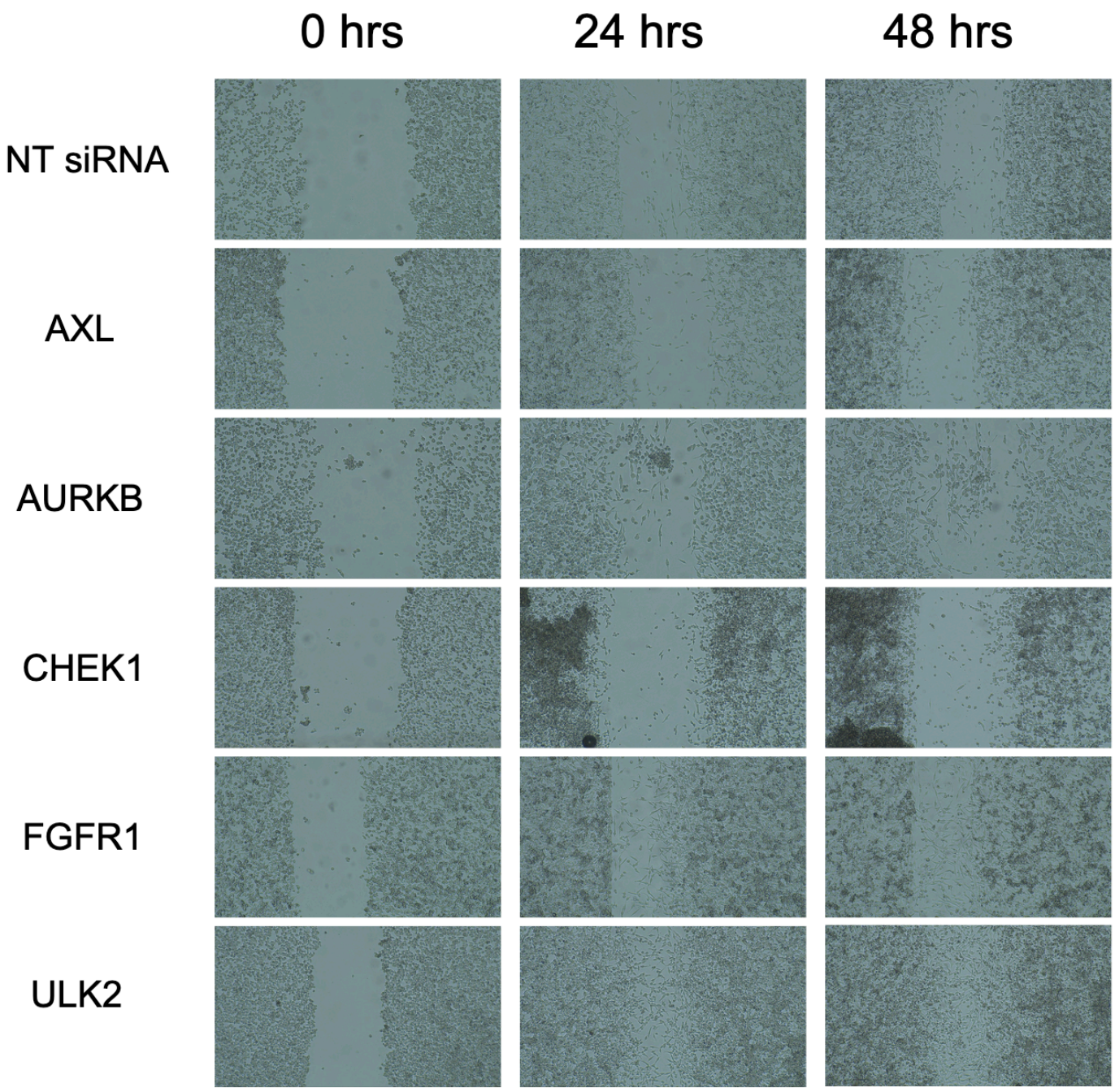

Ladder
ES2 No Treatment
ES2 Non-targeting siRNA
ES2 AXL siRNA
AXL
(140 kDa)
EXP: 1.1s
B-Actin
(42kDa)
EXP: 1s

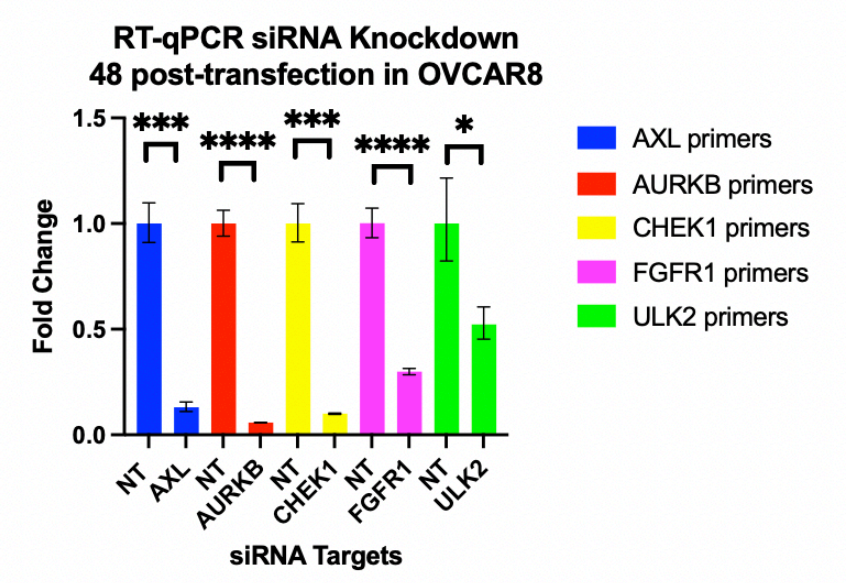
Figure 4.3: Secondary Screen with ES2 with knockdown of AXL and ULK2

A. ES2 cell images at 4x showing scratch at time points 0, 8, and 24 hours. Time 0 is 48 hours post-transfection for the non-targeting siRNA, AXL siRNA, and candidate ULK2 siRNA.
B. Box and whisker of percent change in the area of the scratch over the 8-hour time course. NT and AXL conditions were plated in duplicate, while ULK2 was plated in triplicate. 3 images per each replicate was taken and the percent change in scratch area was averaged for those 3 images to get one percentage per replicate.
C. Fold change of the siRNA knockdown 24 hours post-transfection, prior to scratch
D. Western blot for AXL in cells 48 hours post-transfection, at time 0 for the scratch assay

A



B


RT-qPCR siRNA Knockdown
48 post-transfection in OVCAR8
Fold Change
1.5
1.0
0.5
0.0
***
****
***
****
*
NT
AXL
NT
AURKB
NT
CHEK1
NT
FGFR1
NT
ULK2
siRNA Targets
AXL primers
AURKB primers
CHEK1 primers
FGFR1 primers
ULK2 primers

C

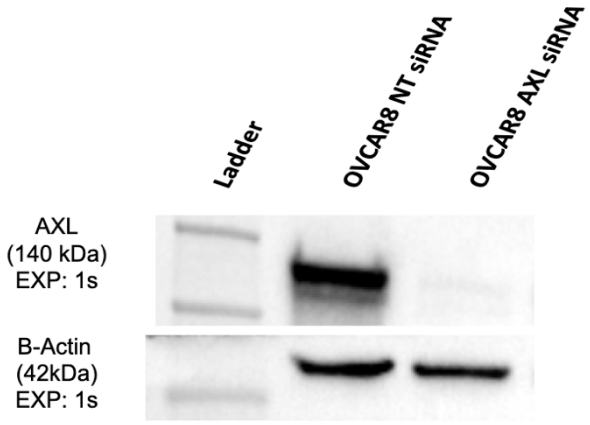
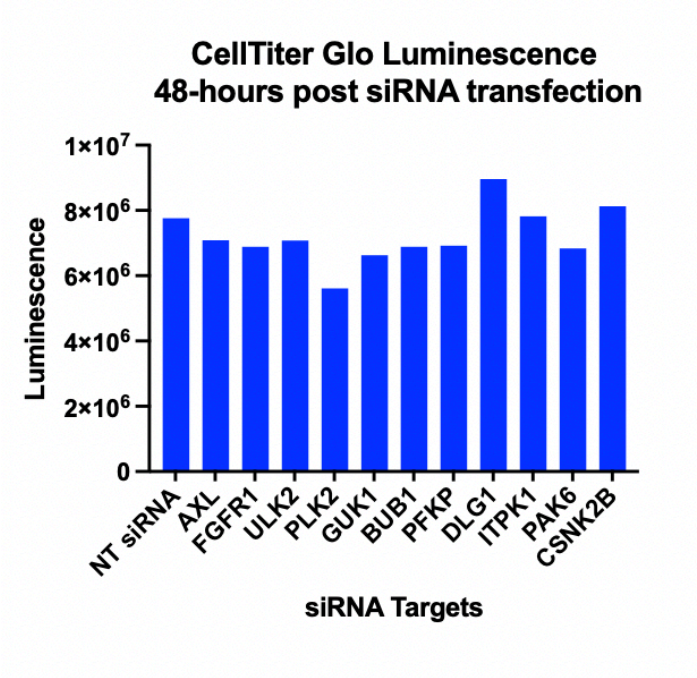

Ladder
OVCAR8 NT siRNA
OVCAR8 AXL siRNA
AXL
(140 kDa)
EXP: 1s
B-Actin
(42kDa)
EXP: 1s

Figure 4.4: Secondary screen in OVCAR8 with knockdown of AURKB, CHEK1, FGFR1, and ULK2

A. OVCAR8 cell images at 4x showing scratch at time points 0, 24, and 48 hours. Time 0 is 48 hours post siRNA transfection of NT siRNA, AXL, AURKB, CHEK1, FGFR1, and ULK2
B. Fold change of the siRNA knockdown 48 hours post-transfection, at time 0 for the scratch assay
C. Western blot for AXL in cells 48 hours post-transfection, at time 0 for the scratch assay

A


CellTiter Glo Luminescence
48-hours post siRNA transfection
Luminescence
1×10^7
8×10^6
6×10^6
4×10^6
2×10^6
0
NT siRNA
AXL
FGFR1
ULK2
PLK2
GUK1
BUB1
PFKP
DLG1
ITPK1
PAK6
CSNK2B
siRNA Targets

B

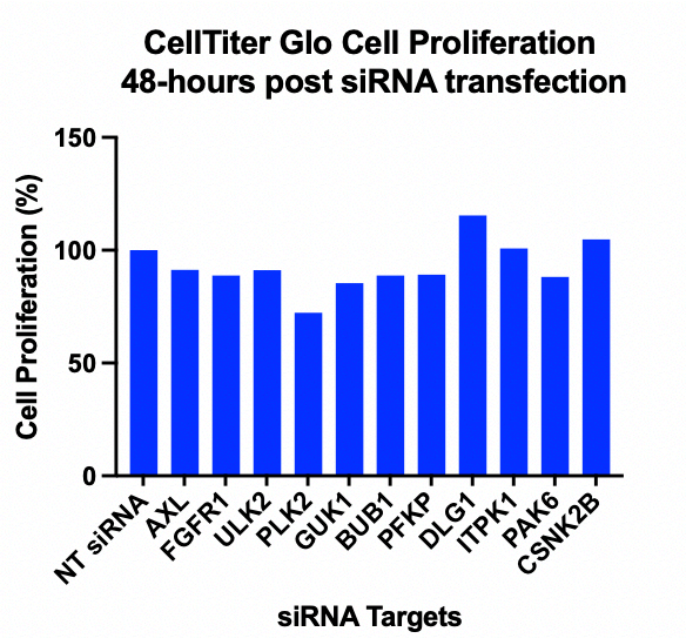
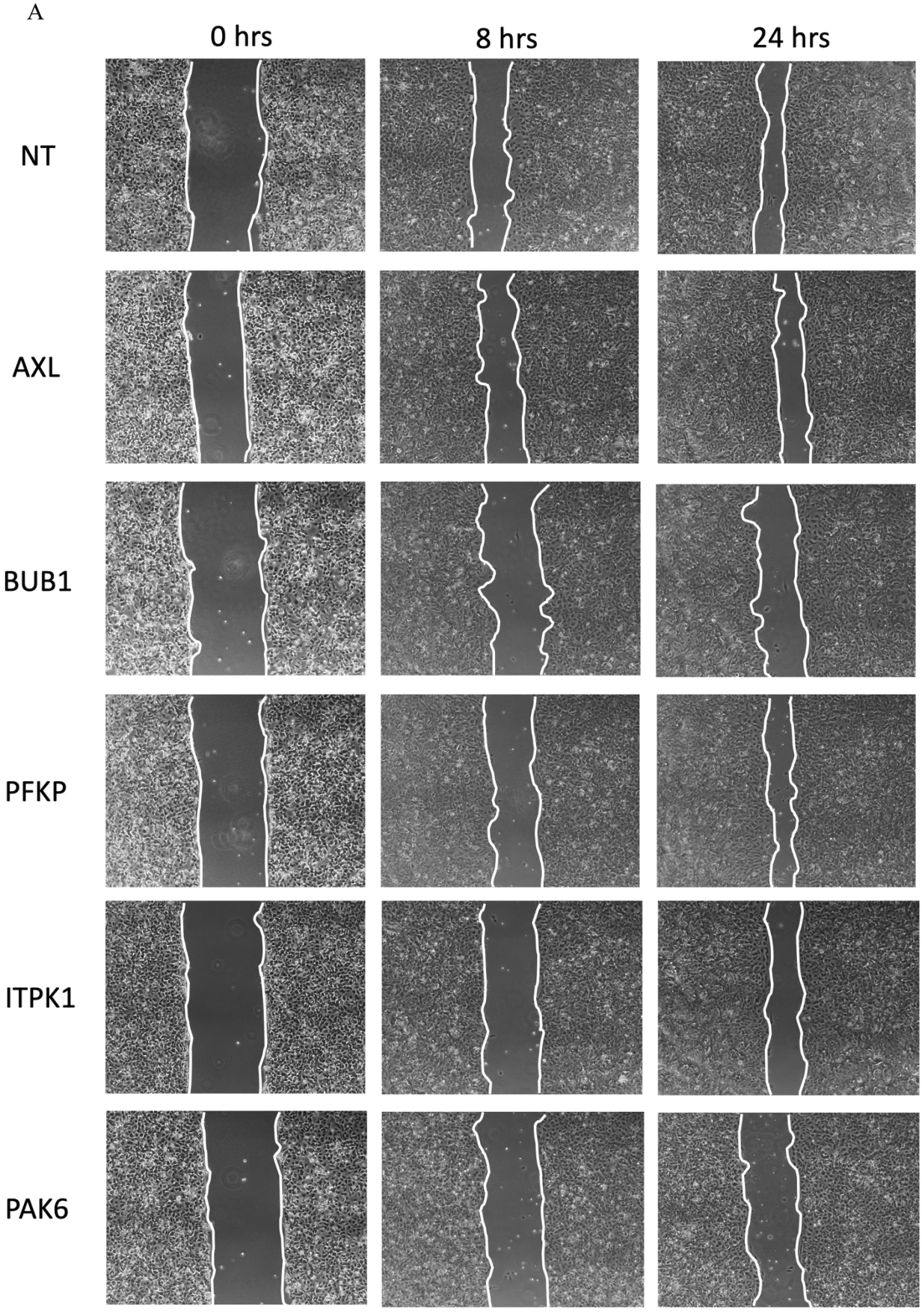

CellTiter Glo Cell Proliferation
48-hours post siRNA transfection
Cell Proliferation (%)
150
100
50
0
NT siRNA
AXL
FGFR1
ULK2
PLK2
GUK1
BUB1
PFKP
DLG1
ITPK1
PAK6
CSNK2B
siRNA Targets

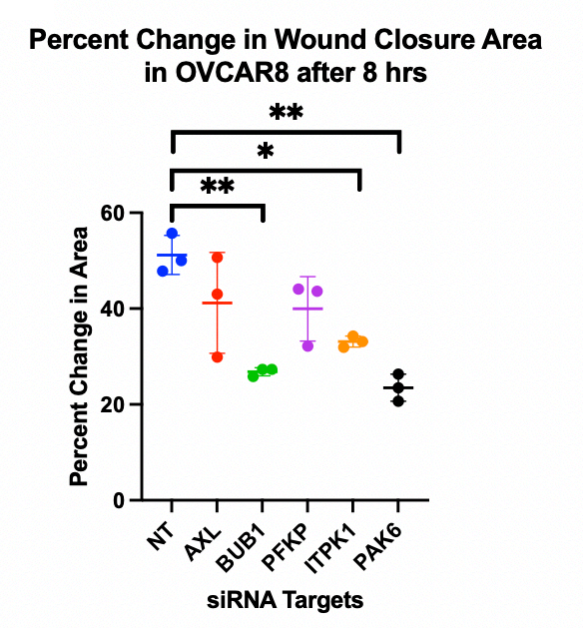
Figure 4.5: CellTiter Glo assay on OVCAR8 cells transfected with siRNA

A. Bar graph displaying the luminescence measurements taken from OVCAR8 cells 48-hours after they were transfected with siRNAs against the top 10 candidate kinases
B. Bar graph displaying the cell proliferation as a percent, which is the luminescence measurements normalized to the NT siRNA control

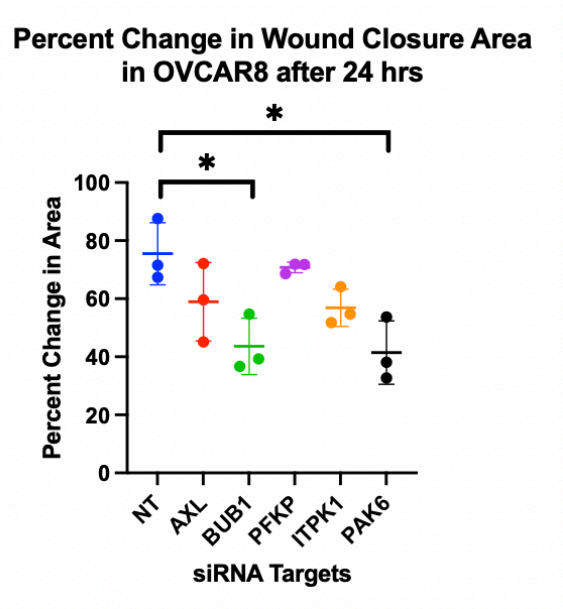
A



B


Percent Change in Wound Closure Area
in OVCAR8 after 8 hrs
**
*
**
Percent Change in Area
60
40
20
0
NT
AXL
BUB1
PFKP
ITPK1
PAK6
siRNA Targets

C


Percent Change in Wound Closure Area
in OVCAR8 after 24 hrs
*
*
Percent Change in Area
100
80
60
40
20
0
NT
AXL
BUB1
PFKP
ITPK1
PAK6
siRNA Targets

D

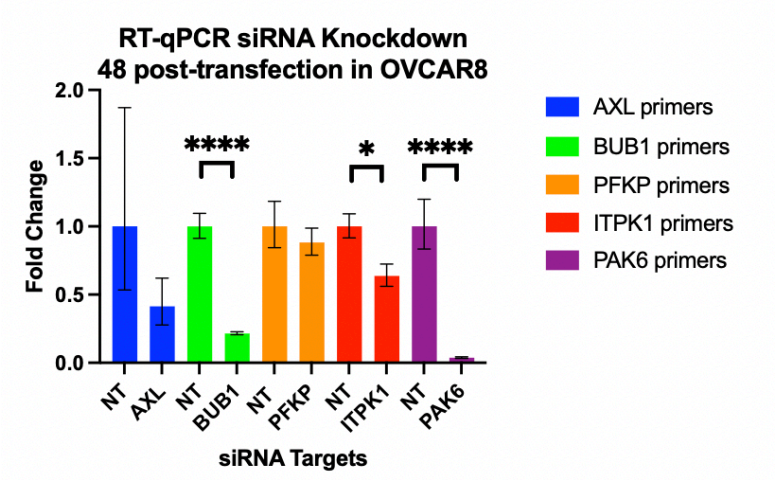

RT-qPCR siRNA Knockdown
48 post-transfection in OVCAR8
Fold Change
2.0
1.5
1.0
0.5
0.0
****
*
****
NT
AXL
NT
BUB1
NT
PFKP
NT
ITPK1
NT
PAK6
siRNA Targets
AXL primers
BUB1 primers
PFKP primers
ITPK1 primers
PAK6 primers

Figure 4.6: Secondary screen in OVCAR8 with knockdown of candidate kinases BUB1, PFKP, ITPK1, and PAK6

A. OVCAR8 cell images at 4x showing scratch at time points 0, 8, and 24 hours. Time 0 is 48 hours post siRNA transfection of NT siRNA, AXL, BUB1, PFKP, ITPK1, and PAK6
B. Box and whisker of percent change in the area of the scratch over the 8-hour time course. All conditions were plated in triplicate and a single image in the center of the scratch was take for each replicate.
C. Box and whisker of percent change in the area of the scratch over the 24-hour time course.
D. Fold change of the siRNA knockdown 48 hours post-transfection, at time 0 for the scratch assay

TABLE 4.1: IDT dsiRNA details

Candidate Gene	Reference #	Batch #	Description	Product Amount
AURKB	318486628	460702464	hs.Ri.AURKB.13.1	2 nmol
CHEK1	318486631	460702465	hs.Ri.CHEK1.13.2	2 nmol
FGFR1	408511126	494468348	hs.Ri.FGFR1.13.2	10 nmol
ULK2	408511129	492783403	hs.Ri.ULK2.13.1	10 nmol
PLK2	408511132	494468349	hs.Ri.PLK2.13.1	10 nmol
GUK1	408511135	492627404	hs.Ri.GUK1.13.1	10 nmol
BUB1	408511138	492627405	hs.Ri.BUB1.13.2	10 nmol
PFKP	408511141	492627406	hs.Ri.PFKP.13.2	10 nmol
DLG1	408511144	492627407	hs.Ri.DLG1.13.2	10 nmol
CSNK2B	408511147	492627408	hs.Ri.CSNK2B.13.1	10 nmol
ITPK1	408511150	492627409	hs.Ri.ITPK1.13.1	10 nmol
PAK6	408511153	492627410	hs.Ri.PAK6.13.2	10 nmol
AXL	408511156	492627411	hs.Ri.AXL.13.2	10 nmol

TABLE 4.2: RT-qPCR primer sequences

Target gene	qPCR F1 sequence	qPCR R1 sequence
ULK2	5′-CTCCTCAGGTTCTCCAGTGC-3′	5′-TTGGTGGGAGAAGTTCCAAG-3′
AURKB	5’-CGCAGAGAGATCGAAATCCAG-3	5’-AGATCCTCCTCCGGTCATAAAA-3’
CHEK1	5'-TGTTGGATGAAAGGGATAAC-3'	5'-AAACATCAACTGGTTCTGC-3'
FGFR1	5'-GAC CAC AGA ATT GGA GGC TA - 3'	5' - ATG AAC TCC ACG TTG CTA CC - 3'
GAPDH	AGCCACATCGCTCAGACA	GCCCAATACGACCAAATCC
PLK2	5′-TCAgCAACCCAgCAAACACAgg-3′	5′-TTTCCAgACATCCCCgAAgAACC-3′
GUK1	5′-CGG CCC ATC TAC ATC TCT GT-3	5′-CGG TCC TTT GAG CTT TCT TG-3′
BUB1	-TGGGAAAGATACATACAGTGGGT-3	5'-AGGGGATGACAGGGTTCCAAT-3
PFKP	TGACGTGGACATCCGCAAAG	CTGGAAGGTGGACAGCGAGG
DLG1	5′- AATTCAGGGCTTGGTTTCAG-3′	5′-ATACAGTCATTGACCCGCAA-3′
ITPK1	CATCCTTGAAGCCGACCAGAATG	CTCATAGGACTTGGAGCGGTCA
PAK6	GACTCCATCCTGCTGACCCTC	CACCTCAGTGGCATACAAAGACC
ULK2	AATCTTGCCCAGTCCCAGTG	GCTTCTCCACCATCCCTTCC
CSNK2B	TGAGCAGGTCCCTCACTACC	GTAGCGGGCGTGGATCAAT

Table 4.3: Patient Data and Functional Genomic Screen Result Top Candidates

Candidate Kinase	Cellular Function	Relevance to Cancer	Patient Data	Screen MAD values	Limitations
PACSIN2	This gene is a member of the protein kinase C and casein kinase substrate in neurons family. The encoded protein is involved in linking the actin cytoskeleton with vesicle formation by regulating tubulin polymerization	• Not associated with cancer in published literature	DEG unique to met – downregulated in ST survivors	-2.5 OVCAR8 Lib1 -0.7 ES2 Lib1 -1.8 A2780 Lib1 3.2 OVCAR8 Lib2 5.6 ES2 Lib2 1.8 A2780 Lib2	patient data and screen data don't correlate well together. Also lack of supporting literature and relevance to cancer.
PI4KA	This gene encodes a phosphatidylinositol (PI) 4-kinase which catalyzes the first committed step in the biosynthesis of phosphatidylinositol 4,5-bisphosphate.	• Overexpression of phosphatidylinositol 4-kinase type IIIα is associated with undifferentiated status and poor prognosis of human hepatocellular carcinoma	DEG unique to met – downregulated in ST survivors	3.2 OVCAR8 Lib1 1.6 ES2 Lib1 -0.2 A2780 Lib2 2.6 OVCAR8 Lib2 2.2 ES2 Lib2 -0.7 A2780 Lib2	patient data and screen data don't correlate well together . Not much literature relating to cancer. No reagents (antibodies) available.
MAST4	A serine/threonine protein kinase that part of a family of microtubule-associated proteins.[41]	• Not associated with cancer in published literature.	MAST4 was predicted to be part of a gene fusion with CA10 identified in a LT survivor.	-0.4 OVCAR 8 Lib1 0.3007 ES2 Lib1 0.3453 A2780 Lib1 -1.5 OVCAR8 Lib2 -3.8 ES2 Lib2 -3.2 A2780 Lib2	lack of relevance to cancer in literature and only identified in non-recurrent gene in LT survivors
CDKL2	CDKL2 is a cyclin dependent kinase-like protein, and is a serine/threonine protein kinase in the mitogen-activated protein kinase (MAPK) family.[41]	• CDKL2 was shown to promote EMT in breast cancer.[51] • Loss of CDKL2 is predictive of poor prognosis in gastric cancer and when overexpressed, it suppresses cell growth and invasion.[52]	Mutated only in ST survivors. CDKL2 was mutated in the primary and met of 1 patient, and the recurrent tumor of another.	-0.5 OVCAR 8 Lib1 -0.6 ES2 Lib1 1.1 A2780 Lib1 1.6 OVCAR8 Lib2 3.1 ES2 Lib2 3.1 A2780 Lib2	patient data and screen data don't correlate well together
PIK3C2B	PIK3C2B is a member of the PI3K protein family, so it plays a role in the signaling pathways that regulate cell proliferation, oncogenic transformation, cell survival and migration, and intracellular protein trafficking.[41]	• One study found that decreased expression of this protein inhibited metastasis in an ovarian cancer cell line.[48] • Another study has shown that there are copy number gains of this gene in ovarian cancer, similar to other PI3K family kinases.[53]	Mutated in the metastatic tumors of 2 ST survivors.	1.2 OVCAR8 Lib 1 3.9 ES2 Lib1 4.5 A2780 Lib1 -3.0 OVCAR8 Lib2 -2.9 ES2 Lib2 -3.7 A2780 Lib2	Mutations in ST survivors suggest loss of function, while literature supports higher expression that promotes metastasis
WNK2	WNK2 is a serine/threonine kinases that is important for the regulation of electrolyte homeostasis, cell signaling, survival, and proliferation.[41]	• Previous studies have identified WNK2 as a driver and risk factor for early recurrence of hepatocellular carcinoma.[54] • It has been shown that a microRNA was able to promote cell growth of WNK2 in breast cancer.[55] • WNK2 is a known tumor suppressor in glioma, which acts by upregulating MMP2 and JNK.[56]	Mutated in two metastatic tumors only, 1 ST survivor and 1 LT survivor.	5.2 OVCAR8 Lib1 3.6 ES2 Lib1 1.7 A2780 Lib1 0.9 OVCAR 8 Lib2 0.05 ES2 Lib2 -0.3 A2780 Lib2	mutations in both ST and LT survivors, and tumor suppressors are more difficult to develop targeted therapies for

Table 4.4: Top 10 candidate kinases selected for secondary screen validation

Candidate Kinase	Function
ITPK1 (inositol-tetrakisphosphate 1-kinase)	Regulates synthesis of inositol tetraphosphate, metabolism
ULK2 (UNC-51 like autophagy activating kinase 2)	Autophagy; mTORC1 regulation
FGFR1 (fibroblast growth factor receptor 1)	Regulates development, cell prolif, differentiation, and migration. Downstream of PI3K and MAPK pathways
PAK6 (P21 (RAC1) activated kinase 6)	Regulation of gene transcription
PLK2 (polo like kinase 2)	Serine/threonine protein kinase, may have role in cell division
GUK1 (guanylate kinase 1)	Phosphorylation of GMP to GDP
BUB1 (budding uninhibitied by benzimidazoles 1 homolog/mitotic checkpoint ser/thr kinase)	Phosphorylate mitotic checkpoint complex and spindle checkpoint, may have function in DNA damage response
PFKP (phosphofructokinase)	Glycolysis regulation
CSNK2B (casein kinase 2 beta)	Regulates metabolic pathways, signal transduction, transcription, translation, replication
DLG1 (discs, large homolog 1)	Scaffolding protein, may play role in junction formation, signal transduction, cell prolif

4.7 References

1. Lengyel, E., *Ovarian cancer development and metastasis.* Am J Pathol, 2010. **177**(3): p. 1053-64.
2. Ning, M. and E.H. Lo, *Opportunities and challenges in omics.* Transl Stroke Res, 2010. **1**(4): p. 233-7.
3. Aherne, G.W., E. McDonald, and P. Workman, *Finding the needle in the haystack: why high-throughput screening is good for your health.* Breast Cancer Research, 2002. **4**(4): p. 148.
4. Pérez-Fidalgo, J.A., et al., *Aurora kinases in ovarian cancer.* ESMO Open, 2020. **5**(5): p. e000718.
5. Do, T.V., et al., *Aurora kinase A mediates epithelial ovarian cancer cell migration and adhesion.* Oncogene, 2014. **33**(5): p. 539-49.
6. Morotti, M., et al., *Targeting tyrosine-kinases in ovarian cancer.* Expert Opin Investig Drugs, 2013. **22**(10): p. 1265-79.
7. Quinn, J.M., et al., *Therapeutic Inhibition of the Receptor Tyrosine Kinase AXL Improves Sensitivity to Platinum and Taxane in Ovarian Cancer.* Mol Cancer Ther, 2019. **18**(2): p. 389-398.
8. Rankin, E.B., et al., *AXL is an essential factor and therapeutic target for metastatic ovarian cancer.* Cancer Res, 2010. **70**(19): p. 7570-9.
9. Grither, W.R., et al., *TWIST1 induces expression of discoidin domain receptor 2 to promote ovarian cancer metastasis.* Oncogene, 2018. **37**(13): p. 1714-1729.
10. Tu, M.M., et al., *Targeting DDR2 enhances tumor response to anti-PD-1 immunotherapy.* Sci Adv, 2019. **5**(2): p. eaav2437.
11. Elkamhawy, A., et al., *The Journey of DDR1 and DDR2 Kinase Inhibitors as Rising Stars in the Fight Against Cancer.* Int J Mol Sci, 2021. **22**(12).

12. Suarez-Arnedo, A., et al., *An image J plugin for the high throughput image analysis of in vitro scratch wound healing assays.* PLoS One, 2020. **15**(7): p. e0232565.

13. Gao, J., et al., *Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal.* Sci Signal, 2013. **6**(269): p. pl1.

14. Cerami, E., et al., *The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data.* Cancer Discov, 2012. **2**(5): p. 401-4.

15. Safran, M., et al., *The GeneCards Suite*, in *Practical Guide to Life Science Databases*, I. Abugessaisa and T. Kasukawa, Editors. 2021, Springer Nature Singapore: Singapore. p. 27-56.

16. Stelzer, G., et al., *The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses.* Curr Protoc Bioinformatics, 2016. **54**: p. 1.30.1-1.30.33.

17. Guo, Z., et al., *Axl inhibition induces the antitumor immune response which can be further potentiated by PD-1 blockade in the mouse cancer models.* Oncotarget, 2017. **8**(52): p. 89761-89774.

18. Parmar, K., et al., *The CHK1 Inhibitor Prexasertib Exhibits Monotherapy Activity in High-Grade Serous Ovarian Cancer Models and Sensitizes to PARP Inhibition.* Clin Cancer Res, 2019. **25**(20): p. 6127-6140.

19. Chen, Y.-J., et al., *Overexpression of Aurora B is associated with poor prognosis in epithelial ovarian cancer patients.* Virchows Archiv, 2009. **455**(5): p. 431.

20. Gu, L., et al., *Polyphyllin I inhibits the growth of ovarian cancer cells in nude mice.* Oncol Lett, 2016. **12**(6): p. 4969-4974.

21. Kim, T. and A. Gartner, *Bub1 kinase in the regulation of mitosis.* Anim Cells Syst (Seoul), 2021. **25**(1): p. 1-10.

22. Davidson, B., et al., *BUB1 mRNA is significantly co-expressed with AURKA and AURKB mRNA in advanced-stage ovarian serous carcinoma.* Virchows Arch, 2014. **464**(6): p. 701-7.

23. Xu, B., et al., *MiR-490-5p Suppresses Cell Proliferation and Invasion by Targeting BUB1 in Hepatocellular Carcinoma Cells.* Pharmacology, 2017. **100**(5-6): p. 269-282.

24. Tang, D., et al., *Identification of hub genes to regulate breast cancer metastasis to brain by bioinformatics analyses.* J Cell Biochem, 2019. **120**(6): p. 9522-9531.

25. Jin, W. and L. Ye, *KIF4A knockdown suppresses ovarian cancer cell proliferation and induces apoptosis by downregulating BUB1 expression.* Mol Med Rep, 2021. **24**(1).

26. Yang, Q., et al., *PAK6 promotes cervical cancer progression through activation of the Wnt/β-catenin signaling pathway.* Oncol Lett, 2020. **20**(3): p. 2387-2395.

27. Fram, S., et al., *A PAK6-IQGAP1 complex promotes disassembly of cell-cell adhesions.* Cell Mol Life Sci, 2014. **71**(14): p. 2759-73.

28. Huang, W., et al., *PAK6 promotes homologous-recombination to enhance chemoresistance to oxaliplatin through ATR/CHK1 signaling in gastric cancer.* Cell Death Dis, 2022. **13**(7): p. 658.

29. Saiardi, A. and S. Cockcroft, *Human ITPK1: a reversible inositol phosphate kinase/phosphatase that links receptor-dependent phospholipase C to Ca2+-activated chloride channels.* Sci Signal, 2008. **1**(4): p. pe5.

30. Dovey, C.M., et al., *MLKL Requires the Inositol Phosphate Code to Execute Necroptosis.* Mol Cell, 2018. **70**(5): p. 936-948.e7.

31. Shen, J., et al., *PFKP is highly expressed in lung cancer and regulates glucose metabolism.* Cell Oncol (Dordr), 2020. **43**(4): p. 617-629.

32. Lang, L., et al., *PFKP Signaling at a Glance: An Emerging Mediator of Cancer Cell Metabolism.* Adv Exp Med Biol, 2019. **1134**: p. 243-258.

33. Peng, M., et al., *Intracellular citrate accumulation by oxidized ATM-mediated metabolism reprogramming via PFKP and CS enhances hypoxic breast cancer cell invasion and metastasis.* Cell Death Dis, 2019. **10**(3): p. 228.

Chapter 5: Conclusions and Future Directions

At the onset of this dissertation work, I set out to complete discovery-based projects that would identify markers of metastasis to characterize and utilize for translational purposes. I hypothesized that I would discover genetic targets from both patient sequencing data and innovatively designed screens that would promote metastasis in HGSC and become sources for new projects in the Fuh Lab.

In Chapter 2, I used sequencing from HGSC patient data to characterize the genetic differences between primary and metastatic tumors and survivorship. I aimed to find unique genetic features of metastatic tumors, however these analyzes revealed that there were more genetic differences between ST and LT survivors than unique features distinguishing metastatic tumors from primary tumors in our cohort. This finding is interesting considering that many studies have focused on longer survivorship differences, usually comparing between patients that lived less than 5 years and over 10 years, while our survival groups were between less than 2 years and over 5 years after their diagnoses. Overall, a larger cohort of patients would be needed to identify more targets or defining features of metastatic tumors. Our study found a few differences, but that could be due to a lack of sampling or biases in race and demographics of this small cohort of patients. Further analysis and characterization of metastatic tumors of ST survivors could lead to additional treatment developments and maintenance therapies.

One of the original goals of this study was to define clonal evolution features between the primary and metastatic tumors, in order to identify gene targets that aid in the development of metastases. However, with the progression of the tumor sequencing over the last decade, I would have needed more spatial samples for each of the primary and metastatic tumors pairs to define clonal origins. Perhaps with the emergence and prevalence of single-cell sequencing, this kind of sampling will be more feasible now and in the near future. We also now know that whole genome sequencing is vital to understanding the

genetic features of ovarian cancer, since it is a cancer characterized by large structural and copy number alterations that cannot be confidently detected with only whole exome sequencing.

There are several findings from this study that could warrant further investigation. I identified several gene fusions involving *ESR1*. Some of these gene fusions have also been identified in the literature with other HGSC patient tumor cohorts, but they have not been investigated for their role in cancer metastasis or as a target for treatments. I identified gene fusions predictions involving several collagen genes that could be explored more as either an artifact to the RNA-sequencing or verified gene fusions that could be involved in the structure of the tumor microenvironment. For example, DDR2's ligand is collagen 1A, so perhaps these collagen gene fusions could lead to receptor activation and contribute to tumorigenesis. Additionally, in my differential expression analyses I identified several lncRNAs and their role in ovarian cancer has also largely been understudied.

One of the findings from Chapter 2 that I did further investigate was the *TP53* R273H mutation found in 3 patient's tumors and discussed in Chapter 3. I determined that in the HGSC ovarian cancer cell lines that endogenously expressed the *TP53* R273H mutation and the cells transfected with the mutation showed no more sensitivity to Temozolomide and PARP inhibitor treatments compared to controls without the mutation. I did not however determine or gather any further evidence that the mutation is a gain-of-function. If I were to further investigate the mutation as a gain-of-function, we would first need to establish if the mutant causes a different, perhaps more aggressive, phenotype in ovarian cancer cell lines. For instance, I have shown that the OVCA-420 were less susceptible to PARPi and temozolomide treatment than the other ovarian cancer cell lines discussed in Chapter 3. To further this, I would design experiments to test if cells with the mutation are more proliferative, migratory, and/or invasive. If we observed that cells with the mutation were significantly more proliferative or invasive compared to wildtype counterparts, or by knocking-down p53 expression in the OVCA-420 cells, that would be evidence to support a distinct function from wildtype p53. Had any of these combination treatments

sensitized the cells specifically because of the TP53 mutation, we had planned to move these experiments *in vivo* and test the same hypothesizes and treatment combinations in mice. We would then use the same combination treatments as stated with Temozolomide, PARPi, and carboplatin to see which combinations reduce tumor burden the most *in vivo*. With the expectation that PARPi and Temozolomide and PARPi with carboplatin will show the least amount of tumor burden after treatment in the mice with cells that express the mutant p53. Perhaps if the mutation did cause a more aggressive phenotype *in vitro* and *in vivo*, we could learn more about the mechanisms behind the hotspot mutation and discover a way to target it therapeutically beyond what was tested here.

Another approach I used to identify targets that promote metastasis is with a high-throughput screen. My lab designed a functional screen in the context of the tumor microenvironment that sought to identify kinases that regulate HGSC metastasis. We silenced over 700 kinases in 3 GFP-labeled ovarian cancer cell lines to identify kinases that either increased or decreased tumor cell attachment and/or invasion to primary fibroblasts and a matrix of collagen and fibronectin. I first attempted to prioritize candidates using patient data described in Chapter 2, however those candidates had a lack of supporting evidence to be translational targets. However, those targets could still be investigated for their role in metastasis in the future.

I was able to validate 2 kinases from this screen, BUB1 and PAK6, in a secondary screen measuring cell migration. There were several other kinases I selected as candidates, ITPK1, PFKP, DLG1, GUK1, ULK2, PLK2, CSNK2B, and FGFR1 that were not tested in the final iteration of the secondary screen. It will be necessary to improve the secondary screen design for these additional candidates by having at least 2 dsiRNA's pooled for each candidate to ensure proper knockdown during siRNA transfection and to verify the protein expression of the knockdown. BUB1 and PAK6 will need to be further studied to elucidate how they are able to influence migration and if they can regulate the other

steps of tumor cell metastasis, like cell clearance and invasion. If promising and validated *in vitro* and *in vivo*, these kinases could be used as a potential target for drug inhibitors for HGSC patients.

I have shown the culmination of using both patient tumor data and *in vitro* screens as tools to discover new targets that promote metastasis in high grade serous ovarian cancer. I have learned that both methods can present different challenges and limitations, but both are useful discovery tools. Ultimately, the goal has always been to find genes and proteins that can be therapeutically targeted so we can provide more precision-based treatments for patients and improve their survival. Every patient and every tumor is unique and the more our field can characterize and understand the mechanisms of metastasis, the more hope we have at developing treatments for individuals to prevent their cancer from metastasizing and reduce their tumor burden. Although there is much more work to be done to investigate the discoveries made here, this work presents targets that have potential to improve patient survival in the future.

Appendix

Implementation of Families Accelerating Cascade Testing Toolkit for Hereditary Cancers

The following is a final class report for the Dissemination and Implementation course I took as part of the requirements for the Cell-to-Society pathway.

Quality gap

In 2018, breast, ovarian, endometrial, and colon cancers together accounted for over 200,000 deaths in the United States. [1] It is estimated that 5-30% of breast, ovarian, endometrial, and colon cancers are caused by genetic mutations that cause hereditary breast and ovarian cancer (HBOC) or Lynch Syndrome (LS). [2] Fortunately, genetic testing can reduce this cancer risk through preventative medicine strategies such as screening, early detection, and risk-reducing surgeries. There has also been a rise of tailored treatment plans based on genetics, such as the use of PARP inhibitors for HBOC, thus providing more incentive for genetic testing in cancer patients. Patients with inherited mutations (probands) are encouraged to share their results to at-risk first-degree relatives so they can undergo genetic testing (cascade genetic testing) and subsequent preventative care.[3,4] However, there are numerous barriers faced by clinicians, probands, and their family members that prevent cascade genetic testing. For example, clinicians often do not have enough time to educate their patients on cascade testing, patients often lack the proper knowledge to inform their family members of the benefits of genetic testing, and family members bear the burden to reach out to health care providers for testing. In order to improve cascade genetic testing rates, we aim to mitigate the challenges of cascade genetic testing by implementing the Families Accelerating Cascade Testing Toolkit (FACTT) in the Siteman Cancer Center.

Thus far, there has been a lack of data available for cascade genetic testing rates. One study surveyed 115 probands with BRCA1/2 mutations and found that 77% of the probands informed their family members of their results, but only half of those relatives underwent genetic testing. Other studies

have reported as few as 9% of high-risk family members underwent testing.[5] The Washington University Gynecologic Oncology Division surveyed 103 probands with HBOC and LS between 2011 and 2016. Although the surveys reported that 87% of the first-degree relatives were informed of their family members' positive result, only 40% of those relatives underwent testing. This data emphasizes the need for more consistent data collection of cascade genetic testing and for better communication between clinicians, probands, and family members on the benefits of lifesaving genetic testing, which the FACTT can help facilitate.

Evidence-Based Intervention

The FACTT is built on evidence-based tools and is designed to facilitate communication between cancer providers and their patients to educate probands' family members about the importance of genetic testing. If family members test positive for cancer-risk mutations, the toolkit offers them resources for genetic counseling and subsequent preventative treatment, if necessary. FACTT does this by offering a variety of materials for probands and their relatives, such as an online pedigree tool, informational videos, letters to send to family members, brochures on genetic testing, and a letter for family members to give to healthcare providers. The chart to the right lists all of the tools currently provided in FACTT.

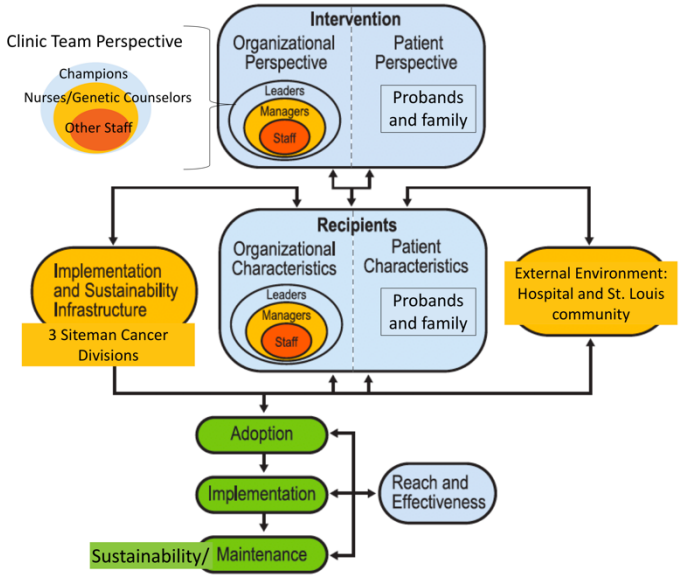
Family Communication Packet	
For probands	• Online pedigree tool • Video on importance of cascade testing for family • Siteman Cancer Center website
For family members	• Letter to family member • Proband's test results in a sealed envelope • Myriad gene information sheet • Siteman Cancer Center website with support resources and testing options (e.g., InformedDNA, Myriad, link to National Society of Genetic Counselors "find a genetic counselor" page, *Facing Our Risk of Cancer Empowered* or *Lynch Syndrome International* websites). • Video on importance of testing • Letter for healthcare provider about importance of testing and insurance coverage information

Other studies are implementing other interventions to improve cascade genetic testing rates, such as directly sending genetic testing kits to relatives. [3] However, thus far the efficacy of a toolkit such as FACTT has not been evaluated nor has there been studies evaluating strategies to implement a genetic testing kit for family members in cancer clinics. FACTT is currently being tested in the Siteman Cancer center to test its efficacy and effectiveness in increasing the number of family members of probands that

undergo genetic testing, compared to the current standard of care. This on-going study also aims to improve the toolkit based on feedback from patient advocacy groups and surveys from patients, while also gathering more information of the barriers for patients and family members for genetic testing and preventative treatments. The study will identify barriers for the proband disseminating the information about their diagnosis to their relatives and for the family members in getting tested and undergoing preventative measures. These barriers will be helpful in the future for shaping the implementation of the toolkit and adjusting its materials to better suit the needs of probands and their family. For the purposes of this project, we will assume that the toolkit is effective in improving the number of proband family members that undergo genetic testing. This study is designed to test and report on implementation strategies to effectively adopt the toolkit in the Siteman cancer clinics and it will use the effectives of the toolkit as one measure of the implementation outcomes.

Framework/Theoretical model

We will use the model PRISM (Practical, robust implementation and sustainability model), which incorporates some elements from the RE-AIM Model. [4] This model will be useful because it incorporates both organizational and patient perspectives into implementation and examines the organizational perspective at multiple levels. This will be used in the study to consider the opinions of staff and the implementation team at their different hierarchies, so we can better measure outcomes for how well the toolkit can be adapted and maintained within the existing infrastructure of the clinics. PRISM as a model will evaluate how FACTT will influence adoption, implementation, maintenance, reach and effectiveness and will be useful to guide measurements for implementation outcomes. [4] The following graphic and table depict a the model applied to the context of this study and the PRISM elements that are activated.



Program (Interventions)	Organizational Perspective • Addresses barriers of frontline staff • Coordination across departments and specialties • Usability and adaptability • Ability to observe results Patient Perspective • Patient centeredness • Addresses patient barriers • Service and access • Feedback of results
External Environment	• Community resources
Implementation and Sustainability Infrastructure	• Dedicated team • Adopter training and support • Relationship and communication with adopters • Adaptable protocols and procedures • Plan for sustainability
Recipients	Organizational Perspective • Management support and communication • Shared goals and cooperation • Clinical leadership • Expectation of sustainability Patient Perspective • Demographics • Disease burden (in this case, I interpreted this as uptake of toolkit)

PRISM Elements Activated:

Study Context

This study will occur within the same 3 clinics in the Siteman Cancer Center that are testing the efficacy of the FACTT toolkit currently. These oncology clinics include the Gynecology Oncology, Breast Surgery, and Colorectal Surgery Divisions. In this context, probands that will be enrolled in the study will be from a variety of socioeconomic backgrounds and will be dealing with their own diagnosis and treatments while bearing the burden of disseminating the FACTT materials and information to their family members. The toolkit will hopefully help reduce the burden on probands and surveys for patient feedback and reporting will be designed to be read and understood by patients from all backgrounds, along with the patient's health and treatment timeline in mind. Surveys will be completed online in their own time, in the clinic, or by mail depending on the patient's access to internet and other resources. Patient advocacy groups will be consulted to advise when and how to best conduct surveys.

On the clinical side, oncologists, nurses, genetic counselors and staff all have very busy schedules and will have to coordinate their time together for proper implementation of FACTT. A dedicated team to

conduct the implementation will help coordinate and facilitate the study around everyone's schedules. Any surveys given to these employees will be online so they can complete them at their discretion.

Stakeholders

- **Target population:** Probands diagnosed with ovarian, breast, or colorectal cancer and their first-degree relatives.
- **End users of the EBI:** Probands, which are defined as patients that have been diagnosed with cancer and carry a mutation in a gene that would increase the risk of developing cancer for family members. Probands will receive the toolkit and disseminate the information, letters, and resources to their first-degree relatives so they may choose to seek genetic testing and preventative care.
- **Implementers/interventionists:** Oncologists, champions, nurses, genetic counselors, and research staff within the Divisions clinics.
- **Other key decision makers:** Additional faculty, like nurses and genetic counselors, will have to be hired to form implementation teams in each clinic and conduct the study, therefore the Division chiefs and directors may be involved in supervising or hiring team members.
- **Community members:** Patient advocacy group members will be consulted to advise the study and make sure we are sensitive to the needs of patients and family members.

Adaptations

The implementation team will assess the progress and implementation strategy procedures 6 months into the study, to make adjustments and improvements to the delivery or consultation of the toolkit or to re-train personnel (see implementation strategies). Additionally, probands will be surveyed on their experiences and can give feedback on the toolkit materials to give recommendations on improving FACTT. Before expanding the implementation of the toolkit to other clinics, those suggestions and additional materials can be incorporated into the kit. Since we are implementing the same strategies

into 3 clinics, we can also tailor implementation and compare findings based on the different settings and clinic infrastructures, which will advise implementation and adoption for future clinics.

Implementation Strategies

The following strategies are adapted from ERIC implementation strategies and defined according to the specifications recommended in Proctor et al. [5,6]

Strategy: Capture and share local knowledge	
ERIC Definition: "Capture local knowledge from implementation sites on how implementers and clinicians made something work in their setting and then share it with other sites." **Actor:** Champions and current pilot study team **Action:** The current pilot study will be used to improve the toolkit and help overcome or anticipate barriers for implementation prior to the probands being enrolled in this study. **Targets of Action:** Implementation team for this project (see next strategy) **Temporality/Dose:** Once, in the beginning of study	**Implementation Outcome:** Effectiveness **Justification:** The lessons learned from the on-going study will help guide our implementation strategies, so the ones planned here may be adjusted or more added to help mitigate newly identified barriers and facilitators. PRISM reported that beginning with a barrier and needs assessment leads to effective implementation.[4]

Strategy: Create Implementation/Clinical Team	
ERIC Definition: "Change who serves on the clinical team, adding different disciplines and different skills to make it more likely that the clinical innovation is delivered (or is more successfully delivered)" **Actors:** The team will be led by a champion in each clinic and will also consist of genetic counselors, nurses, oncologists, and research coordination staff. Champions will be an establish oncologist in the Division that is willing and qualified to lead and oversee their team. **Action:** The team will coordinate the efforts and record who/when toolkits are given out and track and survey probands and family members. Genetic counselors and nurses will give toolkit consultations by review the materials with probands and doing the pedigree with them. Oncologists will advise study and help identify probands for team. Research will collect electronic health records, consent patients, make and disseminate surveys, and record results. **Targets of Action:** Probands receiving toolkits. **Temporality:** Team will be assembled at the start the study and will be assessed 6 months after start date. The team will meet bi-monthly to check on progress and coordinate efforts. **Dose:** Once at start of the study.	**Implementation Outcome:** Feasibility of having dedicated team and genetic counselors in clinic. Adoption (uptake of toolkits), Implementation **Justification:** One study on cascade genetic testing reported that having a dedicated team helped implementation. [6] It is also recommended in the PRISM model to have a dedicated team for Implementation and sustainability infrastructure.[4] In cancer clinics, it has been shown that having genetic counselors in the clinic helps uptake of genetic testing for patients with BRCA1/2 mutations. [7]

Strategy: Conduct Educational Trainings	
ERIC Definition: "Hold meetings targeted toward different stakeholder groups (*e.g.*, providers, administrators, other organizational stakeholders, and community, patient/consumer, and family stakeholders) to teach them about the clinical innovation" **Actors:** Clinic champion will be in charge of leading educational trainings to team. **Action:** Trainings will teach the overview and scope of the project, its goals, the items of the toolkit, how the genetic counselors and nurses should consult probands about the toolkit materials and pedigree tool, and how surveys for probands and team. **Targets of Action:** Implementation team **Temporality:** Trainings will be held after implementation team is assembled and 2 training will be held within the same week. **Dose:** There will be 2, 1 hour trainings. And a retraining effort 6 months after start of project, to re-introduce goals of project, progress, and make any adaptions.	**Implementation Outcome:** Effectiveness, Adoption, Implementation, Maintenance **Justification: Communication and preparation of the team will facilitate and standardized the dissemination of the toolkit to probands.** There was a study that reported that training staff on how to deal with family communications was needed for successful cascade testing, therefore educational trainings are necessary for this kind of dissemination. [8]

Strategy: Purposefully Reexamine the Implementation	
ERIC Definition: "Monitor progress and adjust clinical practices and implementation strategies to continuously improve the quality of care" **Actors:** Implementation team **Action:** We will check into the program every 6 months to assess how the team is doing and if there are adjustments that need to be made to strategies or protocols for implementing the toolkit. **Targets of Action:** Implementation team and probands **Temporality/Dose:** Every 6-months for the duration of the project.	**Implementation Outcome:** Adoption, Maintenance **Justification:** Continuously assessing progress will facilitate adaption and hopefully improve the adoption of the toolkit within each clinic's routine procedures, so the process can be streamlined and maintained.

Approach

Study design and Evaluation: This study will be a randomized clinical trial. Each clinic will enroll up to 120 probands into the study, 40 of which will be given the toolkit and do the online pedigree tool with a nurse while the other 40 will be given the toolkit and complete the pedigree with a genetic counselor, and 40 probands will be a baseline control group where they are given the toolkit but not consulted by a nurse or genetic counselor. To control for randomization, every other proband that is enrolled in the study will be put into each group. Therefore, both the nurses and genetic counselors will conduct consultations over the same time frame.

Implementation Outcomes: The study is also a hybrid III design, as the focus is on measuring the implementation strategies and outcomes for FACTT adoption and actors in the cancer clinics. To assess effectiveness, reach, and adoption, the probands enrolled in the study will be surveyed at the time of enrollment and 6-months after enrollment. Adoption, maintenance, and implementation will be assessed with a survey for the oncologists, nurses, genetic counselors, and staff on the implementation team every 6 months for the duration of the study to assess their perceived barriers to implementing the toolkit in their clinic and to make adjustments while the project is on-going.

We will measure effectiveness of the intervention by assessing the number of family members that undergo genetic testing for each proband. This measure will be proband-reported for the cascade testing rates of first-degree relatives, which will be defined as the number of first-degree relatives tested divided by the number of living first-degree relatives of each proband. This outcome will allow us to assess the efficacy of the toolkit to compare between the nurse and genetic counselor consultations. We will also collect each probands' demographics, cancer history data, cancer status, and specific mutation status from

their electronic health record, to assess any differences in the reach of FACTT between these demographics.

Internal and external validity: There are several barriers to consider when performing cascade genetic testing, mainly because we cannot control for diverse family relations and dynamics, communication styles, and family members having differing access to health care. These are the threats to internal validity, since the kit and delivery method are not the only factors that determine if family members are able to undergo genetic testing and preventative care. In terms of external validity, the structure of each division participating in the study may be different, but they are all within the Siteman Cancer Center and have the same resources at the Barnes Jewish hospital. Therefore, applying the findings here to other clinics and hospitals systems may be challenging depending on the internal structures of the clinics, resources available, and personnel available. For example, some rural hospitals may not have genetic counseling services or personnel to consult probands about genetic toolkits.

PRECIS-2 wheel

The following table displays the ratings considered by the PRECIS-2 Wheel.[7]

[7]Element	Rating (Scale 1-5)	Rationale
Eligibility	5	Probands selected for each group will be the same. We will not be selecting for any certain demographic for type of patient, their eligibility will be their status as a proband having potentially at-risk family members.
Recruitment	4	Patients will be recruited in the clinics and the implementation team will use health records to identify patients that are probands after the genetic testing results are received by the clinic and they will be asked to enroll in the trial.
Setting	4	Implementation of the toolkit will be done in the same clinical settings for each proband, but they will need an extra appointment to meet with the nurse or genetic counselor to be given the kit, consulted about its contents and do the online pedigree tool with the nurse of counselor.
Organization	2	Study will create a new team in the divisions and introduce genetic counselors directly to the clinics to implement the toolkit and give extra training to nurses and counselors for the toolkit and tools for dealing with family relations.
Flexibility: Delivery	2	The implementation team can be flexible by allowing adaptions and adjustments to implementation during the study. But will be strict in having every other proband enrolled be consulted about the toolkit with either a nurse of genetic counselor.
Flexibility: Adherence	4	Adherence to using the kit will be not be judged in enrolling patients for the study. Since we are measuring the efficacy of how many family members are reached and if they undergo genetic testing, we actually want to include all participates that are given the kit and complete follow-up surveys.
Follow-up	5	The follow-up is the same for all probands in the study.
Primary Outcome	2	We are measuring outcome in terms of how many probands utilize the kit and family members are undergo testing and preventative care, while also testing their trust of the nurse/genetic counselor that

		consults them about the kit to see if differing the delivery person affects this outcome and seeing how feasible it is to have a nurse versus counselor in the clinic delivering the kits.
Primary analysis	5	We will use and analyze all data from surveys collected and stats from health records and kits given to probands etc.

Implications

At the completion of the study, we will determine if there is difference in the effectiveness of the FACTT between the groups that were given no consultation, or given a consultation with a nurse or genetic counselor. Then the implementation team will write a suggested protocol for implementation based on the outcomes and evaluation of the staff on the feasibility of the implementation strategies, which would be the condensed plan for implementation that can be disseminated to other clinics. In the end, we hope that FACTT will be seamlessly adopted into the divisions, so it can benefit and reach all probands and their family members at Siteman and help patients receive preventative strategies and eventually reduce hereditary cancer deaths.

The results of this work will be presented to faculty and staff of each of the divisions by the implementation team champion during grand rounds. This will also be when the final plan for uptake and sustainability will be announced (see below). We will publish the findings in a peer-reviewed journal and the results can be presented at conferences to disseminate this work and protocol to other clinicians. A presentation designed for non-researchers/clinicians will be presented to patient advocacy groups that were consulted during the implementation and have previously helped refine the toolkit. Pamphlets can also be created to summarize the results and materials in the toolkit for advocates and the patients involved in the study.

Plan for sustainability

In order to maintain the use of FACTT in the clinics moving forward, the personnel (dedicated nurse/genetic counselor/coordination staff) that are determined to lead to the most effective outcomes, will become permanent members(s) of the clinical team in the divisions and will continue to disseminate the toolkits to probands in the most efficient manner possible. Depending on the success of the study, to

further maintain FACTT we could use more implementation strategies such as mandating the change, so that providing the toolkit and a consultation to all probands in the clinic becomes the standard of care within the divisions. In the future we could also use implementation strategies to stage a scale-up so we can implement the tool to other divisions in the Siteman Cancer Center and eventually expand it to other hospitals in our network and around the state.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA: A Cancer Journal for Clinicians.* 2018;68(1):7-30.
2. Childers CP, Childers KK, Maggard-Gibbons M, Macinko J. National Estimates of Genetic Testing in Women With a History of Breast or Ovarian Cancer. *J Clin Oncol.* 2017;35(34):3800-3806.
3. Frey MK, Kahn RM, Lipkin K, et al. Prospective feasibility trial of a novel strategy of facilitated cascade genetic testing using telephone counseling and mailed saliva kit genetic testing. *Gynecologic Oncology.* 2019;154:281.
4. Feldstein AC, Glasgow RE. A practical, robust implementation and sustainability model (PRISM) for integrating research findings into practice. *Jt Comm J Qual Patient Saf.* 2008;34(4):228-243.
5. Proctor EK, Powell BJ, McMillen JC. Implementation strategies: recommendations for specifying and reporting. *Implementation Science.* 2013;8(1):139.
6. Powell BJ, Waltz TJ, Chinman MJ, et al. A refined compilation of implementation strategies: results from the Expert Recommendations for Implementing Change (ERIC) project. *Implementation Science.* 2015;10(1):21.
7. Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. *BMJ : British Medical Journal.* 2015;350:h2147.

Motor Vehicle Crash Testing Regulations for More Inclusive Populations

The Following is a policy memo I co-authored with members of ProSPER that was published in a special issue of the Journal of Science Policy and Governance, Volume 18, Issue 04 on September 27, 2021 under https://doi.org/10.38126/JSPG180410

Hannah E. Frye[1,5*], Daphne Ko[2,4,5], Emilee N. Kotnik[3,4,5], Nathan Zelt[4,5]

Note: All authors contributed equally to this work.

[1]Washington University in St. Louis, Department of Anesthesiology, St. Louis, MO

[2]Washington University in St. Louis, Department of Molecular Microbiology, St. Louis, MO

[3]Washington University in St. Louis, Department of Obstetrics and Gynecology, St. Louis, MO

[4]Washington University in St. Louis, Division of Biological Sciences, St. Louis, MO

[5]Washington University in St. Louis, ProSPER, St. Louis, MO

[*]Corresponding author: hfrye@wustl.edu

Keywords: motor vehicle safety; sex disparities; crash test safety; diverse body types; NHTSA; crash test dummy

Executive Summary: There is a stark disparity in motor vehicle crash deaths and injuries between male and female drivers. Female drivers are 13% more likely to be killed than their male counterparts in similar motor accidents. However, vehicle safety test practices do not account for diverse body proportions when assessing safety outcomes. Vehicle crash testing standards only require testing of two variations of adult-sized crash test

dummies: a 50th percentile male and a 5th percentile female. Automotive companies are not required to test safety outcomes in crash test model's representative of average female proportions or of differing body size proportions and physiological compositions. Current crash test standards are regulated by the National Highway Traffic Safety Administration (NHTSA) under the US Department of Transportation. This memo proposes three actions for the NHTSA and the Department of Transportation to address disparities in vehicle safety outcomes: (i) update safety standard requirements to include a 50th percentile female crash test dummy, (ii) implement a federal tax incentive program for companies to include a greater diversity of vehicle occupant models, and (iii) allocate funds for research and development of virtual crash testing models. These proposed initiatives seek to raise the minimum safety requirements and prioritize wider representation of vehicle occupants to improve parity in vehicle safety outcomes.

I. Statement of the issue

In the past few decades, vehicles have become increasingly safer, with fewer crash deaths and injuries (Kahane 2013). This is due, in part, to stricter safety regulations and technological advances in automobile safety features, such as increased safety belt use and advanced air bag installation. However, there remains a stark disparity in vehicle crash deaths and injuries when comparing male and female drivers involved in similar vehicle accidents.

Men drive more miles, are involved in more vehicle crashes, and are more likely to engage in risky behavior while driving compared to women, yet women are about 13-20% more likely to be killed than men in similar motor accidents (IIHS 2021).

These disparities in vehicle crash safety outcomes are not adequately addressed in present-day safety vehicle safety testing requirements. Crash testing currently uses 50th percentile male dummies (171 lb, 5ft

9in) but only 5th percentile female dummies (108 lb, 4ft 11in) to represent adult vehicle occupants (NHTSA 2021). Additionally, female dummies are scaled-down versions of male dummies and do not account for sex differences in mass distribution, muscle and ligament strength, and bone structure and density (Linder & Svedberg 2019). Female crash test dummies do not represent average female vehicle occupants, which leads to inaccurate crash test data collection and inadequate safety features. For example, due to differences in neck musculature, females are about twice as likely to suffer from whiplash injuries than males. Whiplash protection seats, designed in response to crash test results, are more effective for males than females (Linder & Svedberg 2019).

Physiological differences between male and female occupants are not the sole factor in risk disparity. Compared to males, female drivers are more likely to be in passenger cars than heavier vehicles like SUVs and vans, and drivers in heavier vehicles are less likely to be injured (Bose 2011). Female injury risk was reduced after accounting for differences in vehicle and crash type (Brumbelow & Jermakian 2021). However, both of these studies concluded that females still have a higher risk of injury compared to males, particularly extremity injury, and that female-specific safety designs need to be incorporated.

Other factors beyond sex, such as age and weight, also lead to risk disparities. Fatality risk increases with age (Kahane 2013), and research groups have found that obese and underweight drivers have increased risk of death (Zhu 2006, Viano 2008). Current crash test dummies do not represent underweight, obese, elderly, or disabled body types (Cornell Law School, n.d.).

II. Current car safety standards

Vehicle safety tests are regulated by the NHTSA, which is a federal agency under the US Department of Transportation. The NHTSA implements laws from Congress by writing and enforcing the Federal Motor Vehicle Safety Standards (FMVSS), which sets requirements for vehicles that test for crash avoidance, crashworthiness, and post-crash survivability.

These standards are developed and supported by research conducted under the NHTSA through the Vehicle Research and Test Center. Their responsibilities include crash test dummy standardization, as well as testing and research on crash avoidance and crashworthiness. In the United States, crash tests are performed by the NHTSA and the Insurance Institute for Highway Safety, a nonprofit organization. Both organizations evaluate vehicles based on crash test performance (Consumer Reports, n.d.).

One example of regulatory success in improving car safety is the response to airbag deaths in low-speed crashes that were not expected to be fatal. Airbag deployment was designed to protect 50th percentile males in crashes. However, the force of these airbags was too strong for children and smaller women, leading to 179 deaths between 1996 and 2000 (Barry 2019). The NHTSA responded by implementing new airbag regulations, relaxing testing requirements to quickly promote lifesaving improvements (Hollowell, n.d.), and requiring advanced airbags that deploy with a force relative to the weight of the vehicle occupant. These measures significantly reduced airbag fatalities, demonstrating that regulation can mitigate vehicle injury and fatalities. Yet, since these safety measures were designed and tested primarily using 50th percentile male dummies, female drivers have higher risk of injury (Bose 2011).

Automotive companies must test several different crash conditions in order to meet FMVSS, using both front and side impacts at different speeds (NHTSA 1999). In these tests, the effectiveness of the safety equipment is determined by measurements of acceleration, force, and deflection of a representative crash test dummy (Hollowell, n.d.). After meeting a minimum requirement of FMVSS, the NHTSA gives the vehicle a rating out of five stars for its performance in various types of accidents (Consumer Reports 2014). These tests determine injury severity based on the Abbreviated Injury Scale and use the results from all types of crashes to give a vehicle an overall rating. Unlike the compliance testing of FMVSS, which uses the 5th percentile female dummy in a variety of positions, the NHSTA star ratings only procure data with the female dummy as a passenger on the driver's side. This is notable since the NHSTA star ratings are made available to the public while compliance testing remains private. Separate ratings are

used for the passenger and driver side dummies, but in testing, a 50th percentile male dummy is used as the driver while the 5th percentile female and children dummies are put in passenger positions (Consumer Reports 2014). Additionally, the FMVSS requires maximum measurements of force and acceleration, but excludes considerations for differences in bone density, mass distribution, and other physiological differences between men and women in the injury calculations (Cornell Law School, n.d.; Linder & Svedberg 2019). Current standards do not represent diverse body types since they exclude publicly available data on female drivers and don't use dummies with more inclusive body types.

While regulations for crash tests are often updated to improve vehicle safety, standards for crash test dummies have been slow to change. Efforts by the NHTSA to expand its crash test dummy repertoire in the 1980s were hindered by budget cuts and shifts in government attitudes towards regulation. The cost and time needed to develop new dummies continue to be key factors in maintaining the status quo. According to a senior research engineer at the IIHS, crash test dummies can take over twenty years to develop (Barry 2019). In support of the current standards, an NHTSA representative explained that using a wider range of crash test dummies (50th percentile male and 5th percentile female) allows for better protection for a greater range of occupants (Putka 2021).

III. Car safety research and development

A crash test dummy representing the 50th percentile female body has yet to be commercially produced or used in US automotive testing, despite there being a physical prototype model, the BioRID 50F, and a computer model called EvaRID. However, due to cost and the absence of regulatory incentive to incorporate these models in vehicular safety testing, they are not yet fully developed nor widely used in crash testing (Gendered Innovations, n.d.). Some individual automotive companies are implementing additional safety testing to improve safety measures for a wider diversity of drivers and passengers. For example, in 2019 Volvo launched the Equal Vehicles for All initiative with the aim to expand crash testing research for a greater variety of body types and has shared more than 40 years' worth of their data

and research (Gendered Innovations, n.d.). However, these additional safety measures vary in scope and efficacy across brands and vehicle models. Unless FMVSS requirements are updated to reflect updated proportions, present-day vehicle operators will continue to experience poor safety outcomes in automobile accidents.

Considering the cost, labor, and time of developing diverse crash test dummies and running multiple crash tests per dummy, some companies have turned to virtual models. As of 2021, Toyota has made its virtual human-modeling software, Total Human Model of Safety, free access, without a government incentive to do so. This system digitizes crash testing and performs computer simulations and injury analyses for vehicle collisions. The system can replicate vehicle impacts on the muscle tissue, skeleton, and internal organs of several sexes, age groups, and body types in greater detail (Autovista Group 2020). Virtual crash testing has the potential to gather crash data and inform safer vehicle designs, especially since Toyota can continue to improve the software when it is more widely available to users (Toyota Motor Corporation 2020).

IV. Policy recommendations

Vehicle safety can be improved through federal regulation and support. This memo recommends the following measures:

i. Update FMVSS to require a 50th percentile female crash test dummy in vehicle crash testing.

Update the FMVSS to include a female crash test dummy proportioned for the 50th percentile of adult American women in automotive crash testing. Additionally, this crash test dummy must be tested in driver and passenger positions.

Advantages

Requiring the inclusion of a 50th percentile female crash test dummy in vehicle safety testing will ensure that crash test data incorporates average female proportions. Testing updated female crash test dummies in both driver and passenger placements will improve representation of actual vehicle operators. Inclusion of a female crash test dummy in vehicle crash testing is a promising first step to address disparities in vehicle occupant safety outcomes.

Disadvantages

Changes in safety testing requirements will require significant administrative and financial investment by automotive companies for compliance with new guidelines. While 50th percentile female crash test dummies have already undergone significant research and development, additional research may be required to ensure standardized implementation for safety testing across companies. Finally, this requirement still only represents a limited subset of vehicle operators.

ii. Incentivize automotive companies to incorporate additional safety testing for diverse vehicle occupants

Inclusion of an updated female crash test dummy improves the representation of diverse vehicle occupants in safety testing protocols, but it still provides only a limited scope of occupants. Therefore, we propose the use of federal tax incentives to encourage automotive companies to incorporate additional diversity in vehicle safety testing. While the NHTSA or IIHS issue industry-wide safety ratings based on testing at their affiliated sites, automotive companies also conduct their own crash testing to ensure compliance with FMVSS in the United States. The NHTSA can recommend and incentivize the inclusion of additional standardized crash test dummy specifications for automotive companies to incorporate in company-specific car safety testing to represent a wider diversity of vehicle occupants. These can include models representing higher and lower size percentiles, weakened bone density or joint strength, internal or external implants, etc. In addition, virtual crash testing presents promising opportunities for cost-effective testing of diverse occupant safety outcomes.

Advantages

The utilization of a tax incentive approach for the incorporation of additional vehicle safety testing will encourage automotive companies to consider vehicle safety data from a wider diversity of vehicle occupants. The automotive industry will retain the option to consider which safety measures beyond the basic FMVSS requirements will be valued by their consumers. Incentivizing the inclusion of diverse crash test models, whether physical or virtual, may promote further development of safety models which represent a wider range of vehicle occupants to improve overall safety outcomes.

Disadvantages

An incentive model does not require additional safety testing of diverse vehicle occupancy. Safety testing will still vary between vehicle brands and models, which may not be clearly communicated to consumers. In addition, more research may be required to set national standards for diverse crash test dummies. Finally, inclusion of tax incentives for companies to incorporate improved diversity in crash test dummies may result in reduced federal tax revenue received from the automotive sector.

iii. Fund vehicle research and development for virtual crash testing and analysis

The NHTSA requests a total of $32,805,000 for Vehicle Safety Research and Analysis in their Fiscal Year 2021 budget estimate. This budget focuses on improving research for driving automation, advanced vehicle safety technologies, and crash survivability, including the development of innovative physical and virtual testing tools for crashworthiness (NHTSA 2021). We recommend that the appropriations committee approve this investment in technology to make vehicles safer.

Advantages

Virtual crash testing allows for a greater number of crash test simulations on a greater variety of vehicles and body types. This allows for more data and analyses without the costs and time of physical crash

testing and dummies. Approving this budget to fund the development and access of virtual crash test systems, like Toyota's system, would lead to more vehicle manufactures having the resources and data to improve car safety for more people.

Disadvantages

Virtual crash testing can never completely replace physical crash testing, as virtual models are based on the physical world. There will always be a need to confirm and validate virtual models with physical testing, which will still need better representative dummies to create safety features (Barry 2019).

V. Conclusions:

Despite increased safety measures in vehicle development and testing, there remain significant demographic disparities in safety outcomes following vehicular accidents. Federal agencies can help prevent unnecessary injury and fatalities by updating crash test dummy requirements to be more representative, incentivizing automotive companies to be more inclusive and extensive in their crash testing research, and by investing in virtual crash testing technology. Implementing these recommendations can help make vehicles safer for more people.

References

Autovista Group. "Toyota to Make Its Virtual Crash Test Dummy Freely Available," June 2020. https://autovistagroup.com/news-and-insights/toyota-make-its-virtual-crash-test-dummy-freely-available.

Barry, Keith. "The Crash Test Bias: How Male-Focused Testing Puts Female Drivers at Risk." Consumer Reports, October 2019. https://www.consumerreports.org/car-safety/crash-test-bias-how-male-focused-testing-puts-female-drivers-at-risk/.

Bose, Dipan, Maria Segui-Gomez, and Jeff R. Crandall. "Vulnerability of Female Drivers Involved in Motor Vehicle Crashes: An Analysis of US Population at Risk." *American Journal of Public Health* 101, no. 12 (December 2011): 2368–73. https://doi.org/10.2105/AJPH.2011.300275.

Brumbelow, Matthew L, and Jessica S Jermakian. "Injury Risks and Crashworthiness Benefits for Females and Males: Which Differences Are Physiological?" *Insurance Institute for Highway Safety*, February 2021, 18.

Consumer Reports. "NHTSA Crash Test 101: How Crash Worthiness Is Measured and How Crash Ratings Can Help You Choose Your next Car," April 2014. https://www.consumerreports.org/cro/2011/08/crash-test-101/index.htm.

Cornell Law School: Legal Information Institute. "49 CFR Part 572 - ANTHROPOMORPHIC TEST DEVICES," n.d. https://www.law.cornell.edu/cfr/text/49/part-572.

Gendered Innovations. "Inclusive Crash Test Dummies: Rethinking Standards and Reference Models," n.d. https://genderedinnovations.stanford.edu/case-studies/crash.html#tabs-2.

Hollowell, William T, Hampton C Gabler, Sheldon L Stucki, Stephen Summers, and James R Hackney. "REVIEW OF POTENTIAL TEST PROCEDURES FOR FMVSS NO. 208." Office of Vehicle Safety: National Highway Safety Administration, n.d.

IIHS. "Fatality Facts 2019: Males and Females." Insurance Institute for Highway Safety, March 2021. https://www.iihs.org/topics/fatality-statistics/detail/males-and-females#fn3.

Jehle, Dietrich, Seth Gemme, and Christopher Jehle. "Influence of Obesity on Mortality of Drivers in Severe Motor Vehicle Crashes." *The American Journal of Emergency Medicine* 30, no. 1 (January 2012): 191–95. https://doi.org/10.1016/j.ajem.2010.10.017.

Kahane, Charles J. "Injury Vulnerability & Effectiveness of Occupant Protection Technologies for Older Occupants and Women." NHTSA Technical Report. Office of Vehicle Safety: National Highway Safety Administration, May 2013. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/811766.

Lavallière, Martin, Mathieu Tremblay, Frédéric Lefebvre, Maxime Billot, and Grant A. Handrigan. "Aging, Obesity, and Motor Vehicle Collisions." *Frontiers in Sustainable Cities* 2 (July 14, 2020): 33. https://doi.org/10.3389/frsc.2020.00033.

Linder, Astrid, and Wanna Svedberg. "Review of Average Sized Male and Female Occupant Models in European Regulatory Safety Assessment Tests and European Laws: Gaps and Bridging Suggestions." *Accident Analysis & Prevention* 127 (June 2019): 156–62. https://doi.org/10.1016/j.aap.2019.02.030.

NHTSA. "FY21 NHTSA Budget Estimate." National Highway Traffic Safety Administration, 2021. https://www.nhtsa.gov/document/fy2021-nhtsa-budget-estimate.

NHTSA. “NHTSA’s Crash Test Dummies.” National Highway Traffic Safety Administration, n.d. https://www.nhtsa.gov/nhtsas-crash-test-dummies.

NHTSA. “Standard No. 208,” March 1999. https://icsw.nhtsa.gov/cars/rules/import/FMVSS/#SN208.

Putka, Sophie. “Why Are There No Crash Test Dummies That Represent Average Women?” Discover Magazine, February 16, 2021. https://www.discovermagazine.com/technology/why-are-there-no-crash-test-dummies-that-represent-average-women.

Toyota Motor Corporation. “Toyota Offers Free Access to THUMS Virtual Human Body Model Software.” Toyota Motor Corporation, June 16, 2020. https://global.toyota/en/newsroom/corporate/32665896.html.

Viano, David C., Chantal S. Parenteau, and Mark L. Edwards. “Crash Injury Risks for Obese Occupants Using a Matched-Pair Analysis.” *Traffic Injury Prevention* 9, no. 1 (February 19, 2008): 59–64. https://doi.org/10.1080/15389580701737645.

Zhu, Shankuan, Peter M. Layde, Clare E. Guse, Purushottam W. Laud, Frank Pintar, Raminder Nirula, and Stephen Hargarten. “Obesity and Risk for Death Due to Motor Vehicle Crashes.” *American Journal of Public Health* 96, no. 4 (April 2006): 734–39. https://doi.org/10.2105/AJPH.2004.0581