DISCOVERY AND ANALYSIS OF ONCOGENIC ALTERATIONS IN TRIPLE-NEGATIVE BREAST CANCER

By

Timothy MacKenzie Shaver

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Biochemistry

August 11, 2017

Nashville, Tennessee

Approved:

Jennifer A. Pietenpol, Ph.D.

David Cortez, Ph.D.

Scott Hiebert, Ph.D.

Stephen Fesik, Ph.D.

William Tansey, Ph.D.

To David, Ruaire, Molly, and Ethan

May the world we leave for you
be better than the one left for us

ACKNOWLEDGEMENTS

I owe a debt of gratitude to many in Vanderbilt, Nashville, and beyond for their professional and personal contributions to my graduate education and my path here. I would like to devote a few pages to a set of acknowledgements that cannot hope to be comprehensive.

First and foremost, I would like to thank my mentor, Dr. Jennifer Pietenpol. Her unrivaled enthusiasm for science and infectious positivity were evident from our first conversation during the IGP poster sessions, and I have told many since that my primary reason for joining her lab was the drive I felt to be productive and impactful by association. Despite her wide array of professional and administrative responsibilities, Jennifer has never failed to be a devoted mentor and fierce advocate for my work and training. In all our interactions, Jennifer has provided a blend of scientific feedback, career guidance, and leadership mentoring that embodies her greatest strength as an advisor: leading by example. Her unfailingly optimistic approach to mentorship – focusing on individual strengths instead of personality differences to harness the most productivity from a team – is a lesson I hope to carry forward.

Another mentor whose impact was essential in my professional development is Dr. Hoang Nguyen at Baylor College of Medicine. I joined Dr. Nguyen's lab as a research technician immediately following my graduation from Rice, and my experience there laid a foundation of

technical knowledge, a critical approach to data, and an enthusiasm for the thrill of discovery that I continue to rely on. As an early hire, I benefitted from Hoang's personal training in the practice and communication of science; both were invaluable preparation for my own thesis work, and her encouragement to seek out and apply myself to graduate training was instrumental in bolstering my work ethic for the journey to a doctorate.

I also thank my committee for their guidance: Dave Cortez, Steve Fesik, Scott Hiebert, Bill Tansey, and (former member) Christine Eischen. Dr. Cortez enjoyed double duty as committee chair and a frequent audience member during joint lab meeting presentations, but always provided fresh insight and suggestions that shaped key directions of the project. I am grateful to Dr. Tansey, who joined the committee only recently but has immediately provided valuable feedback.

In the Pietenpol lab, I owe special gratitude to Hailing Jin, my scientific collaborator and a talented experimentalist whose skill at her craft is evidenced by much of the high-quality data presented throughout the following dissertation. I also owe immense acknowledgement to The Honorable Kim Johnson, whose diligence as a lab manager and unparalleled negotiating skills constitute just a fraction of her importance to the Pietenpol lab. Kim is a consummate professional who blends an impeccable work ethic with a healthy sense of humor. From his time as my rotation mentor to our joint manuscript authorship, Brian Lehmann has demonstrated the ability of a keen imagination and intellect to distill

meaningful insights from complex data; his enthusiasm and fearless approach to hypothesis generation continue to provide invaluable motivation.

My fellow trainees in the Pietenpol lab, Scott Beeler, Clayton Marshall, Gabriela Santos, and Johanna Schafer, have been key collaborators and colleagues, providing their expertise, support, and sometimes commiseration. Scott encouraged me to sign up for an online coding class with him during our first year, and in the process reshaped my graduate and career trajectory. Qing Zhang's assistance has kept the tissue culture process as painless as possible, and her formidable skill at the ornament competition has provided a level to aspire to. Former members, including Chris Pendleton, Lucy Tang, Deb Mays, and Taylor Sundby (JSA), all lent guidance, support, or a friendly ear during their time in the lab. My summer trainees and rotation students, including Veronica Gordillo, Verra Ngwa, and Spencer Lea, have each provided both valuable contributions to our projects and valuable experience in mentorship.

I am grateful to others at Vanderbilt for their guidance and support: John York, Jen Smith, Robert Dortch, Patty Mueller, and Lindsay Meyers in Biochemistry; Lynnette Stanton-Williams in the VICC; Yu Shyr and Chung-I Li in CQS; Sandra Zinkel and Christine Lovly for their feedback in joint lab meetings; Bryan Venters, Tom Stricker, Emily Hodges, Quanhu Sheng, and Olivia Koues for their sequencing knowledge and assistance; and Ferrin Wheeler for her cytogenetics expertise.

I have been fortunate to receive financial support throughout my graduate career, and would like to acknowledge the Vanderbilt Interdisciplinary Graduate Program; Jim Patton and the Cellular, Biochemical, and Molecular Sciences Training Grant (T32GM008554); Mark de Caestecker and the Vanderbilt/Howard Hughes Certificate Program in Molecular Medicine (MIG56006779); the Department of Defense Breast Cancer Research Program, for an IDEA Award to Dr. Pietenpol (W81XWH-13-1-0287); and the NCI, for a Ruth L. Kirchstein National Research Service Award (F31CA183531).

I have relied on the support of family, friends, and loved ones throughout my time in graduate school. In particular, thanks to my mother for providing an example of the power of unconditional support and the value of maintaining a lighthearted outlook on life, to my father for providing a role model of a committed professional, and to Aaron and Ciara for home-cooked meals and a home away from home.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

aa	Amino acid
AR	Androgen receptor
BL1	Basal-like 1 (TNBC subtype)
BL2	Basal-like 2 (TNBC subtype)
BLCA	Bladder urothelial carcinoma
bp	Base pair(s)
BRCA	Breast invasive carcinoma
CD	Cell death (siRNA control)
cDNA	Complementary DNA
CDS	Coding DNA sequence
CGH	Comparative genomic hybridization
COADREAD	Colorectal adenocarcinoma
d	Day(s)
DAPI	4’,6-Diamidino-2-Phenylindole
DNA	Deoxyribonucleic acid
DNA-seq	DNA sequencing
EGFR	Epidermal growth factor receptor
ER	Estrogen receptor
FGFR	Fibroblast growth factor receptor
FISH	Fluorescence/fluorescent *in situ* hybridization
FPKM	Fragments per kilobase per million

FSC-A	Forward scatter pulse area
GB	Gigabyte
GBM	Glioblastoma multiforme
GDAC	Genome Data Analysis Center (Broad Institute)
GOF	Gain of function/gain-of-function
gRNA	Guide RNA
h	Hour(s)
HDR	Homology-directed repair
HER2	Human epidermal growth factor receptor 2
HNSC	Head and neck squamous cell carcinoma
HR	Hormone receptor
HRD	Homologous recombination deficiency
IC_{50}	Half-maximal inhibitory concentration
IHC	Immunohistochemistry
IL-3	Interleukin 3
IM	Immunomodulatory (TNBC subtype)
kDa	Kilodalton
KICH	Kidney chromophobe renal cell carcinoma
KIRC	Kidney renal clear cell carcinoma
LAR	Luminal androgen receptor (TNBC subtype)
LGG	Brain lower grade glioma
LOH	Loss of heterozygosity
LUAD	Lung adenocarcinoma

LUSC	Lung squamous cell carcinoma
m	Minute(s)
M	Mesenchymal (TNBC subtype)
MATCH	Molecular Analysis of Therapy Choice (NCI)
MEFs	Mouse embryonic fibroblasts
miRNA	MicroRNA
mRNA	Messenger RNA
MSL	Mesenchymal stem-like (TNBC subtype)
mTOR	Mechanistic/mammalian target of rapamycin
NCBI	National Center for Biotechnology Information (NIH)
NCI	National Cancer Institute
NGS	Next-generation sequencing
NHEJ	Non-homologous end joining
NIH	National Institutes of Health
NT	Non-targeting
nt	Nucleotide(s)
p53	Tumor protein p53, encoded by *TP53*
PAM	Protospacer adjacent motif
PARP	Poly (ADP-ribose) polymerase
pCR	Pathological complete response
PFA	Paraformaldehyde
PI3K	Phosphoinositide 3-kinase
PR	Progesterone receptor

PRAD	Prostate adenocarcinoma
PTM	Post-translational modification
qPCR	Quantitative polymerase chain reaction
REML	Restricted maximum likelihood
RNA	Ribonucleic acid
RNA-seq	RNA sequencing
RPKM	Reads per kilobase per million
SD	Standard deviation
SEM	Standard error of the mean
SP	Signal peptide
STA	Segmental Transcript Analysis
SV40	Simian virus 40
TA	Transactivation
TCGA	The Cancer Genome Atlas (NCI)
THCA	Thyroid carcinoma
TM	Transmembrane
TNBC	Triple-negative breast cancer
TPM	Transcripts per million
UCEC	Uterine corpus endometrial carcinoma
UTR	Untranslated region
WGS	Whole-genome sequencing
WT	Wild-type
WXS	Whole-exome sequencing

CHAPTER 1

INTRODUCTION

Excluding non-melanoma skin malignancies, breast cancer is the most commonly diagnosed cancer in women and the second-leading cause of cancer death in the same group (Jemal et al., 2017). Over the past half-century, research advances have led to the understanding that breast cancer is composed of numerous subtypes with distinct molecular and clinical features. Triple-negative breast cancer (TNBC), a category encompassing 15-20% of breast cancer cases (Dent et al., 2007; Howlader et al., 2014; Plasilova et al., 2016), poses a significant clinical challenge; in contrast to hormone and cell surface receptor-driven forms of the disease, which benefit from first-line treatment and often prolonged disease control with targeted agents, the standard of care for TNBC remains cytotoxic chemotherapy. In addition, the presence of metastatic TNBC is associated with a worse prognosis, including more frequent relapse and a much shorter median time from relapse to death (Dent et al., 2007; Hudis & Gianni, 2011; Voduc et al., 2010).

In this chapter, the clinical, molecular, and genomic features of TNBC will be reviewed, with particular emphasis on two mechanistically linked oncogenic alterations that are frequent in the disease: mutations in the p53 tumor suppressor and gene rearrangements.

Triple-Negative Breast Cancer

Clinical Definition and Significance

For over fifty years, clinicians and scientists have understood that a majority of breast cancers occur due to derangement of the normal hormone signaling pathways governing mammary gland development and self-renewal, namely estrogen and progesterone (Folca et al., 1961; Glascock & Hoekstra, 1959). New breast cancer diagnoses are routinely screened for the presence of estrogen receptor (ER) and progesterone receptor (PR) by immunohistochemistry; 70-85% of these cases (varying by race) will be hormone receptor-positive (HR+) (DeSantis et al., 2016; Howlader et al., 2014).

An additional subtype of breast cancer is determined by the overexpression of human epidermal growth factor receptor 2 (HER2/neu). Originally identified in 1981 (Shih et al., 1981), HER2 was found to be amplified in human breast cancer cell lines and capable of a potent transforming effect by the end of that decade (Di Fiore et al., 1987; King et al., 1985; Moasser, 2007). Although HER2 exhibits homology to epidermal growth factor receptor (EGFR/HER1) (Schechter et al., 1984, 1985) and also displays tyrosine kinase activity, no direct activating ligand has been found for HER2; instead, it functions in heterodimers with other human epidermal growth factor receptor family members, notably EGFR/HER1 and HER3 in the context of breast cancer (Baselga & Swain, 2009; Iqbal &

Iqbal, 2014; Rubin & Yarden, 2001). HER2 is often overexpressed due to gene amplification; as a result, clinical screening strategies for HER2+ breast cancers include immunohistochemistry in concert with *in situ* hybridization techniques (Baselga & Swain, 2009; Iqbal & Iqbal, 2014; Rubin & Yarden, 2001).

Hormone receptor and HER2 positivity are not mutually exclusive; approximately 10% of breast cancer cases are dual HR+/HER2+, while approximately 5% are HER2+ alone (DeSantis et al., 2016; Howlader et al., 2014). In combination with the 70-75% of HR+/HER2- cases, these clinical categories account for ~85% of breast cancer diagnoses among all races.

The remaining ~15% of breast cancers are, as a result, clinically defined by the process of exclusion: tumors lacking ER, PR, and HER2 expression and/or amplification are termed triple-negative breast cancer. Despite the presence of some rare, nearly uniformly HR and HER2 negative histologic subtypes – *e.g.*, spindle cell carcinoma – TNBC cases cannot be distinguished on the mere basis of histology. In one assessment of 781 TNBC patients treated at the European Institute of Oncology, 89% were classified as ductal, 3.7% as apocrine, 2.3% as lobular, 1.2% as adenoid cystic, 1.2% as metaplastic, 1.1% as papillary, 0.5% as medullary, and 1% a mix of rare types, including mucinous, cribriform, mixed, and micropapillary carcinoma (Montagna et al., 2013).

Despite encompassing a heterogeneous range of tumor types, TNBCs are on average associated with the poorest outcomes among

breast cancer subtypes, with a younger age of diagnosis, increased rate of relapse, and significantly shorter latency between relapse and death (Dent et al., 2007; Haffty et al., 2006; Hudis & Gianni, 2011; Morris et al., 2007; Plasilova et al., 2016; Voduc et al., 2010). In addition, TNBC disproportionately affects minority populations in the US: while TNBC occurs in approximately 10% of cases among the non-Hispanic white population, 20-25% of breast cancer cases in non-Hispanic black women are triple-negative (DeSantis et al., 2016; Howlader et al., 2014). Due to the comparatively worse prognosis associated with TNBC, disparities in breast cancer outcomes on the basis of race are therefore in part due to differences in subtype frequency.

Molecular Subtypes

As technological advances have increasingly enabled molecular and genomic analysis of tumors, gene expression characterization has joined histopathologic assessment as a clinically relevant indicator for both prognosis and treatment determination. In a landmark 2000 study, Perou and colleagues defined ‘intrinsic subtypes’ of breast cancer based on unsupervised hierarchical clustering of gene expression signatures from 84 surgical samples (Perou et al., 2000). Importantly, these intrinsic subtypes recapitulate the known importance of receptor signaling: the luminal A, luminal B, and ‘normal-like’ subtypes are largely HR+ and associated with good to intermediate clinical outcomes, while the presence of HER2 (in a

subset of luminal B and the HER2-positive subtype) is associated with poor prognosis (Dai et al., 2015).

The basal intrinsic subtype – along with the additional claudin-low subtype identified in 2007 (Herschkowitz et al., 2007; Prat et al., 2010) – has been traditionally used synonymously with TNBC and is similarly associated with poor prognosis (Dai et al., 2015; Eroles et al., 2012). However, numerous analyses have identified that only 50-80% of TNBCs fall into the basal intrinsic subtype, necessitating a more refined assessment of transcriptional heterogeneity within TNBC (Bastien et al., 2012; Bertucci et al., 2008; de Ruijter et al., 2011; Lehmann et al., 2016; Morris et al., 2007; Rakha et al., 2009).

In 2011, the Pietenpol lab conducted an analysis of 2300 gene expression signatures collated from 14 individual studies (Lehmann et al., 2011). By using a bimodal expression filter to exclude ER+ and HER2+ cases (Carmeci et al., 1997; Press et al., 2008), 386 TNBC gene expression profiles were identified; after k-means clustering of the samples on the most differentially expressed genes, six stable clusters were identified and were termed basal-like 1 (BL1), basal-like 2 (BL2), immunomodulatory (IM), mesenchymal (M), mesenchymal stem-like (MSL), and luminal androgen receptor (LAR) on the basis of their expression signatures. Gene expression varies widely among these subtypes; as an example, LAR tumors exhibit the activation of hormonally regulated signaling pathways that more closely mimic estrogen receptor-driven

tumors than the other, more traditionally ‘basal’ TNBC molecular subtypes. More recent analysis has identified that the identification of the IM and MSL subtypes was confounded by the presence of immune-infiltrating cells and stromal cells, respectively; as a result, the major TNBC subtypes have been refined to BL1, BL2, M, and LAR (Lehmann et al., 2016).

Standard of Care Differs by Receptor Status

The process of exclusion by which TNBC is defined – lacking hormone receptor expression and HER2 amplification – has placed women diagnosed with the disease at a distinct disadvantage from the standpoint of targeted therapy. For patients with HR+ disease, anti-estrogenic agents such as tamoxifen have found clinical use and efficacy since the early 1970s (Cole et al., 1971). In numerous large-scale trials, tamoxifen has shown efficacy as an adjuvant therapy for breast cancers in both pre- and post-menopausal women (Early Breast Cancer Trialists’ Collaborative Group, 1988, 1998; Nolvadex Adjuvant Trial Organisation, 1983), and has been used as an agent for breast cancer prevention in women at increased risk due to age, history of contralateral breast cancer, and other factors (Cuzick & Baum, 1985; Fisher et al., 1998, 2005). Partly owing to the side effects of long-term tamoxifen therapy, more modern agents such as aromatase inhibitors, which inhibit ER upstream by blocking estrogen production, are now FDA-approved and in wide clinical use (Chumsri et al., 2011).

The development of targeted therapy has been of especially pronounced benefit to patients with HER2+ tumors. Trastuzumab, widely known by its brand name, Herceptin, was initially developed as a murine monoclonal antibody to cells overexpressing HER2 (Shepard et al., 1991). Subsequent humanization of the antibody and its evaluation in numerous clinical trials have validated its use as a highly effective single-agent and combination treatment for HER2+ breast cancers (Arteaga, 2003; Baselga et al., 1996; Carter et al., 1992; Cobleigh et al., 1999; Slamon et al., 2001; Vogel et al., 2002). The impact of trastuzumab in reshaping the landscape of outcomes in breast cancer subtypes has been considerable; while HER2+ patients once experienced a poor prognosis similar to TNBC patients, treatment with trastuzumab in this population extends disease-free survival and overall survival by as much as 50% (Joensuu et al., 2006; Piccart-Gebhart et al., 2005; Romond et al., 2005).

In contrast to receptor-driven forms of breast cancer, for TNBC, cytotoxic chemotherapy remains the standard of care in both the neoadjuvant and adjuvant settings (Isakoff, 2010). TNBC patients as a whole show significant clinical benefit from chemotherapy compared to HR+ patients, in large part due to the elevated mitotic count characteristic of many TNBC tumors (Rakha et al., 2007). However, responses differ between TNBC subtypes; among patients treated with a variety of neoadjuvant chemotherapy regimens, including the widely used sequential taxane and anthracycline, BL1 tumors exhibited a significantly higher

pathological complete response (pCR) compared to LAR and BL2 when examined retrospectively (Lehmann et al., 2016; Masuda et al., 2013). These findings emphasize the need to identify targetable alterations in patients whose tumors have molecular features associated with poor prognosis after mainstay chemotherapy regimens.

Potentially Targetable Features of TNBC

Numerous groups have focused on genome and transcriptome sequencing of TNBC clinical specimens and cell lines in the hope of identifying recurrent oncogenic alterations that may be amenable to targeted therapy. Instead, however, these sequencing efforts have revealed TNBC as a highly heterogeneous category of cancers encompassing a wide spectrum of mutations and copy number alterations (Banerji et al., 2012; Bianchini et al., 2016; Cancer Genome Atlas, 2012b; Curtis et al., 2012; Shah et al., 2012). The only alteration present in a majority of TNBC cases is mutation of the p53 tumor suppressor, to be discussed below. Current efforts in targeted therapy development for TNBC focus on features present in a minority of TNBC cases, including homologous recombination deficiency; androgen receptor signaling; dependence on select signaling pathways, including PI3K and Ras/MAPK; and immune evasion.

Homologous Recombination Deficiency

Deficiencies in homologous recombination are exemplified by mutations in the *BRCA1* and *BRCA2* breast cancer susceptibility genes. The BRCA1 and BRCA2 proteins are involved in the homologous repair arm of the DNA damage response, wherein sister chromatids are used as a template to repair double-strand breaks. Cells deficient in homologous repair must instead rely on error-prone alternatives such as non-homologous end joining; thus, mutations in *BRCA1* and *BRCA2* are associated with collateral mutations and chromosomal abnormalities resulting from inappropriate double-strand break repair (Atchley et al., 2008; Venkitaraman, 2009).

When inherited as germline, *BRCA1* and *2* mutations confer an approximately five-times higher risk of breast cancer incidence, and the tumors occurring in these mutation carriers are disproportionately triple-negative (Antoniou et al., 2003; Chen & Parmigiani, 2007). However, somatic mutations in the *BRCA* genes or their inactivation through other means also occur in TNBC and other cancers. Owing to their DNA repair defect, cells with deficient BRCA proteins or other alterations leading to homologous recombination deficiency (HRD) are unusually sensitive to DNA damaging agents (Turner et al., 2005); further studies have demonstrated that the inhibition of alternate DNA repair enzymes such as poly (ADP-ribose) polymerase (PARP) can be largely tolerated by normal cells with intact homologous repair machinery, but proves lethal to cells

with HRD such as *BRCA* mutants (Turner et al., 2008). This 'synthetic lethal' vulnerability has led to clinical trials assessing the efficacy of PARP inhibitors in patients with potential HRD, but many of the results to date have been inconclusive or mixed (Bianchini et al., 2016; Gelmon et al., 2011; Sinha, 2014). Since a number of variants detected by screening of the *BRCA* genes are of unknown significance and since deficiencies in other proteins can produce a 'BRCA-like' homologous repair-deficient phenotype, a number of ongoing efforts are focused on the development and implementation of assays to detect HRD within tumors (Abkevich et al., 2012; Bianchini et al., 2016; Birkbak et al., 2012; Popova et al., 2012). Despite the promise of the detection and exploitation of HRD in TNBC, such cases are likely to constitute only a minority of the disease.

Androgen Receptor Signaling

The LAR TNBC subtype, representing ~15% of TNBC (Lehmann et al., 2011, 2014), exhibits expression of the androgen receptor detectable by immunohistochemistry and an associated hormonally regulated gene expression signature; in both *in vitro* and *in vivo* assays, LAR cell lines respond to anti-androgen therapy (Lehmann et al., 2011, 2014). Clinical trials based on these findings have included screening for AR by IHC and treatment of AR+ tumors with the anti-androgens bicalutamide and enzalutamide. Results from phase I and II studies have demonstrated clinical benefit as measured by partial or complete response for a minority

of patients, even in the single-agent setting (Gucalp et al., 2013; Schwartzberg et al., 2017). Ongoing trials for patients with LAR-subtype tumors involve assessment of anti-androgens in combination with PI3K inhibitors, as discussed below.

Growth Factor Signaling Pathways

PIK3CA, encoding the alpha subunit of the phosphatidylinositol-4,5-bisphosphate 3-kinase (PI3K), is the second most commonly mutated gene in TNBC after *TP53* (Lehmann et al., 2014; Shah et al., 2012). An additional 10% of patients exhibit loss of *PTEN*, a negative regulator of the PI3K pathway whose loss is almost mutually exclusive with *PIK3CA* mutations (Shah et al., 2012). The net effect of *PIK3CA* hotspot mutations and *PTEN* loss is hyperactivation of the PI3K/AKT/mTOR pathway, which regulates cell growth, survival, and metabolism (Liu et al., 2009). *PIK3CA* mutations are significantly enriched in LAR compared to other TNBC subtypes, and preclinical studies have validated the efficacy of dual AR/PI3K inhibition (Gonzalez-Angulo et al., 2009; Lehmann et al., 2011, 2014). An investigator-initiated, randomized Phase IB/II clinical trial from Vanderbilt-Ingram Cancer Center is underway to evaluate the tolerance and efficacy of dual inhibition in patients (NCT02457910).

Immunotherapy

TNBC has emerged as an area of particular focus in the emerging field of immunotherapy. Due to its increased mutational load and chromosomal instability, TNBC cases tend to harbor a higher frequency of 'neo-antigens,' or *de novo* peptide sequences emerging from genome alterations that can trigger immune reactivity (Brown et al., 2014). As supporting evidence for increased immune reactivity of TNBC compared to other breast cancer subtypes, numerous analyses have shown an increased frequency of tumor-infiltrating lymphocytes in TNBC tumors, along with increased protein and mRNA expression of PD-L1, the immune checkpoint inhibitor whose immune-suppressive mechanism is a key target of numerous promising pharmaceutical therapies (Ali et al., 2015; Bianchini et al., 2016; Loi et al., 2013; Mittendorf et al., 2014; Sabatier et al., 2015; Wimberly et al., 2015). A number of trials are ongoing to test the efficacy of immune checkpoint inhibitors with or without chemotherapy, with the goal of refining prognostic biomarkers for patient selection (reviewed in Bianchini et al., 2016).

Challenges of Heterogeneity

Despite the rapid expansion of our knowledge regarding targetable alterations in TNBC, the heterogeneity of the disease still leaves a sizeable fraction with a poor response to chemotherapy and no obvious targeted therapies to evaluate. In a comprehensive 2012 genomic and

transcriptional analysis of TNBC by Aparicio and colleagues, the authors estimated that “20% of cases contained examples of potentially ‘clinically actionable’ somatic aberrations;” extending their categorization to commonly mutated tumor suppressors, growth factor pathways, and recurrently mutated cell structural genes, they noted that “12% of the cases did not contain somatic aberrations in any of the frequent drivers or cytoskeletal genes” (Shah et al., 2012). A comprehensive 2016 review by Vanderbilt investigators fielded a more optimistic but still sobering estimate that at least 10% of residual TNBC tumors after neoadjuvant therapy lack “alterations in pathways that can be targeted with agents currently under clinical investigation” (Balko et al., 2014; Bianchini et al., 2016).

Despite the challenges stemming from the molecular and mutational heterogeneity of the disease, several aspects of TNBC biology hold the potential to expand the spectrum of actionable alterations upon further investigation. In the remaining portion of the Introduction, two such research areas will be described, which are the focus of the dissertation research described herein: mutations in the p53 tumor suppressor and gene rearrangements.

Tumor Protein p53

Identification as a Tumor Protein

The discovery of p53 has its origins with Simian virus 40 (SV40), which was of research interest in the late 1970s due to its ability to transform cells and induce tumors in animals. In March 1979, Lane and Crawford presented SV40 T antigen immunoprecipitation of a putative host cell protein in the 50-55 kDa range (Lane & Crawford, 1979). Two months later, Linzer and Levine demonstrated that SV40 antisera could precipitate a protein of similar size from carcinoma cell lines uninfected with SV40, confirming its identity as a host cell protein (Hainaut & Wiman, 2009; Linzer & Levine, 1979). Additional groups published similar findings the same year, describing putatively cell-coded proteins in the same molecular weight range that precipitated with the T antigen (Kress et al., 1979; Levine & Oren, 2009; Melero et al., 1979; Smith et al., 1979).

In similarly timed discoveries, two groups reported the antigenicity of a p53 or p50 protein, respectively, in transformed cells (DeLeo et al., 1979; Rotter et al., 1980). Rotter *et al*. additionally noted that this antigenicity was correlated with protein abundance. These data, in concert with the SV40 findings, led to conclusions that p53 functioned as a transforming protein "hijacked" by viral machinery. Attempts were made over the next half-decade to identify the p53 coding sequence, which led to cDNA isolation and genomic cloning by numerous groups (Chumakov et al., 1982; Harlow

et al., 1985; Leppard et al., 1983; Levine & Oren, 2009; Matlashewski et al., 1984; Oren & Levine, 1983; Pennica et al., 1984; Wolf et al., 1985; Zakut-Houri et al., 1985). In a twist of fate that bears special importance for the recent history of p53 biology, many of these cloning attempts were made in tumor cells, where the protein was abundant. The resulting expression constructs were evaluated for their transforming potential, and were found to cooperate with established oncogenes in cellular transformation and to immortalize cells when expressed individually (Eliyahu et al., 1984; Jenkins et al., 1984; Levine & Oren, 2009; Parada et al., 1984). As further evidence, these expression constructs were shown to augment the tumorigenicity of already-transformed and p53-null cells (Eliyahu et al., 1985; Wolf et al., 1984). As Levine and Oren (2009) note, “by the mid[-]1980s p53 was generally acknowledged as an oncogene.” Despite this misconception, there were numerous hints of the complex biology yet to be uncovered and the critical flaw in the studies to that point in time, including a 1985 publication observing that the immortalizing effect of p53 could be enhanced by mutagenesis of the expression construct (Jenkins et al., 1985).

p53 as a Tumor Suppressor

By the late 1980s, a body of evidence had formed to suggest that the earlier interpretation of p53 as an oncogene – and the reagents used in the experiments leading to that conclusion – were flawed (Levine & Oren,

2009). Most importantly, advances in technology led to the formal establishment of the mouse *Trp53* coding sequence, and comparative analysis revealed that many tumor-derived p53 cDNAs – those which had a transforming effect when exogenously expressed – harbored mutations of the wild-type sequence (Eliyahu et al., 1988; Finlay et al., 1988; Halevy et al., 1991). The Levine and Oren labs demonstrated that, in direct contrast to the earlier findings of p53 as an enhancer of transformation, wild-type p53 in fact suppresses transformation by the ras oncoprotein (Eliyahu et al., 1989; Finlay et al., 1989).

At the same time, the Vogelstein laboratory provided incontrovertible evidence that p53 was a tumor suppressor: a 1989 study demonstrated that in two colorectal carcinomas exhibiting deletions of chromosome 17p (on which *TP53* resides), the remaining allele was the sole mRNA expressed and harbored missense mutations (Baker et al., 1989). This phenomenon, known as loss of heterozygosity (LOH), was in perfect agreement with the 'two-hit' hypothesis of tumor suppressors formulated 30 years prior by Nordling (Nordling, 1953) and refined by Knudson in the early 1970s (Knudson Jr., 1971). In 1990, germline p53 mutations were detected in six out of six analyzed families harboring the familial cancer predisposition Li-Fraumeni syndrome (Malkin et al., 1990; Srivastava et al., 1990). A transgenic mouse model with homozygous, functionally null p53 was found to be predisposed to tumor formation (Donehower et al., 1992), and further work by Vogelstein, Harris and colleagues provided a survey of

somatic mutational patterns in p53 across human cancers, providing evidence for a link between specific environmental mutagens and the type of mutation and codon affected (Hollstein et al., 1991).

Guardian of the Genome: Mechanism and Regulation

Evidence that p53 functioned as a DNA damage response protein was published soon after its establishment as a tumor suppressor. Numerous studies, following the original observation by the Kastan laboratory, documented the accumulation of p53 protein in response to DNA damage (Hainaut & Wiman, 2009; Hall et al., 1993; Kastan et al., 1992). Further, cells with inactivated p53 were found to be deficient in damage-induced cell cycle checkpoint activity (Baker et al., 1990; Diller et al., 1990; Kastan et al., 1992; Mercer et al., 1990; Michalovitz et al., 1990). Additional links were established between p53 and apoptosis induced by a variety of DNA damaging agents, including ionizing radiation and several chemotherapeutic agents (Lowe et al., 1993a, 1993b; Yonish-Rouach et al., 1991).

The role of p53 as a transcription factor for the basis of its tumor-suppressive activities was identified by numerous groups in the early 1990s, with many noting that mutant forms of p53 lacked similar transcriptional activation potential (Bargonetti et al., 1991; El-Deiry et al., 1992; Farmer et al., 1992; Fields & Jang, 1990; Funk et al., 1992; Kern et al., 1991; Pietenpol et al., 1994; Raycroft et al., 1990; Zambetti et al., 1992;

Zauberman et al., 1993). In addition to a consensus binding site mediated by a central DNA-binding core domain, these studies identified a discrete, N-terminal sequence-specific transactivation (TA) domain from amino acids 1-42, which was later expanded to include a second TA subdomain from residues 43-63 (Zhu et al., 1998). In addition to a regulatory domain that is a target of numerous post-translational modifications (PTMs), the C terminus contains a oligomerization domain that mediates the formation of transcriptionally active p53 tetramers (Chène, 2001; Iwabuchi et al., 1993; Pavletich et al., 1993; Pietenpol et al., 1994; Stürzbecher et al., 1992; Wang et al., 1993).

The first target genes of p53 to be identified were *CDKN1A*, encoding the p21 cyclin-dependent kinase inhibitor, and *MDM2*, or mouse double-minute 2 homolog. The latter was demonstrated at the time to be an antagonist in a negative feedback loop with p53 and is the key negative regulator of p53 protein stability through its E3 ligase activity (Barak et al., 1993; el-Deiry et al., 1993; Hainaut & Wiman, 2009; Wu et al., 1993). These features of p53 activity and regulation – induction of cell cycle arrest and carefully regulated, transient activation through negative feedback loops – persist to this day as core characteristics of wild-type p53's role as a tumor suppressor.

The function and regulation of p53 continue to be an active area of research, and published studies on the subject number in the tens of thousands to date. Beyond cell cycle arrest, seminal findings of the

downstream effects of p53 activation include induction of apoptosis (Lowe et al., 1993a, 1993b; Martins et al., 2006; Miyashita & Reed, 1995; Shaw et al., 1992; Yonish-Rouach et al., 1991) and cellular senescence (Flores & Blasco, 2009; Serrano et al., 1997; Tyner et al., 2002; Ventura et al., 2007; Wang et al., 1998; Xue et al., 2007). Numerous studies have also demonstrated (somewhat controversially) non-transcriptional tumor suppressive roles for p53, including interaction with the Bcl2 family of apoptosis regulatory proteins and thus an indirect role in cell death-related mitochondrial permeabilization (Green & Kroemer, 2009; Marchenko et al., 2000; Vaseva & Moll, 2009).

Like its downstream targets, the upstream activators and regulators of p53 continue to be an active area of research, but a large volume of studies have recognized that a variety of cellular stresses, including oxidative stress, oncogene activation, and DNA damage, lead to post-translational modifications of p53 and its association with regulatory cofactors, primarily at the N and C termini of the protein (Canman et al., 1998; Hainaut & Wiman, 2009; Kamijo et al., 1998; Pomerantz et al., 1998; Shieh et al., 2000; Siliciano et al., 1997; Zindy et al., 1998). The net effect of many of these modifications and protein-protein interactions is to increase the stability of the p53 protein and shield it from ubiquitination and degradation promoted by its own gene targets, including MDM2 and other E3 ligases, such as MDMX. Canonical examples of upstream p53 activators include the classical DNA damage response, exemplified by the

ATM-Chk2 and ATR-Chk1 pathways that respond to DNA damage by initiating a kinase cascade that includes N-terminal phosphorylation of p53 and its regulators, notably MDMX (Chehab et al., 2000; Chen et al., 2005; Lavin & Kozlov, 2007; Shieh et al., 2000; Smith et al., 2010). Later research has called into question the physiological relevance of ATR-Chk1 activation in p53 regulation, given temporal differences in their activity and regulatory roles (Agarwal et al., 1995; Jazayeri et al., 2006; Walker et al., 2009).

An additional body of work has been devoted to p53's role in cellular metabolism, both as a sensor of metabolic stress and a downstream regulator of metabolic pathways (reviewed in Vousden & Ryan, 2009). An example of p53's role in modulating cancer-relevant metabolism is its ability to down-regulate the glucose transporters GLUT1 and GLUT4 and activate transcription of TP53-inducible glycolysis and apoptosis regulator (*TIGAR*), encoding an enzyme that degrades the glycolytic substrate fructose-2,6-bisphosphate (Bensaad et al., 2006; Schwartzenberg-Bar-Yoseph et al., 2004).

The net effect of p53's activity and its role as a response node to a variety of upstream stresses earned p53 the title "guardian of the genome" by its co-discoverer, Sir David Lane, less than a decade after the p53 research community ascribed oncogenic roles to the protein based on unknowingly mutant expression constructs (Lane, 1992). As our understanding of the biology of p53 has evolved, it is now thought to function as a critical barrier to the early stages of tumor formation,

explaining its extremely high frequency of mutation and inactivation across all human cancers (Bartkova et al., 2005; Gorgoulis et al., 2005; Hainaut & Wiman, 2009; Halazonetis et al., 2008).

The Most Commonly Mutated Gene in Cancer

Although the high frequency of p53 mutations has been known for decades, recent cancer sequencing efforts have permitted a precise survey of mutational rates across cancer types. The Pan-Cancer effort of The Cancer Genome Atlas (TCGA), which analyzed ~3,300 tumors across 12 cancer types, identified a 42% nonsynonymous mutation rate for *TP53*, far ahead of the second-most mutated gene, *PIK3CA*, which exhibited a rate of 18% (Kandoth et al., 2013). *TP53* mutations are especially common in certain tumor types, occurring in 79% of lung squamous cell carcinoma cases and 95% of high-grade serous ovarian adenocarcinomas (Cancer Genome Atlas Research, 2011, 2012; Kandoth et al., 2013). Among breast cancers, p53 mutation rate is significantly enriched in specific subtypes; in 1,420 tumors from the METABRIC (Molecular Taxonomy of Breast Cancer International Consortium) cohort (Curtis et al., 2012), *TP53* mutations occurred in 65% of basal-like cancers and 53% of HER2-enriched but only 9.3% of the largely HR+ Luminal A subtype (Kandoth et al., 2013; Silwal-Pandit et al., 2014).

TP53 is unique among tumor suppressor genes in its mutational pattern. Rather than a broad distribution of nonsense, frameshift, and

missense mutations indicating selection for general loss of function, over 80% of mutations in *TP53* are missense, and many of these fall into discrete 'hotspots' within the core DNA binding domain (Bouaoun et al., 2016). Crystal structures of the p53 DNA binding domain in complex with DNA illustrate that these hotspot mutations fall into two main classes: contact mutants, which affect the amino acids directly in contact with the target DNA, and structural mutants, which occur in the hydrophobic core of the DNA binding domain and alter its stability and, indirectly, the conformation of the DNA contact residues (Cho et al., 1994). Even in the case of the structural mutants, the net effect is not often protein instability leading to degradation; instead, most p53 hotspot missense mutants accumulate to levels much higher than the wild-type p53 protein, which is ordinarily kept in check through its negative-regulatory feedback loop with its E3 ligase gene targets (Fig. 1).

Despite the unique hotspot mutation pattern, there are a sizeable number of nonsense, frameshift, and non-recurrent missense mutations in *TP53* that occur across all cancer types (Bouaoun et al., 2016; Kandoth et al., 2013). Functional inactivation of p53 signaling also extends beyond nonsynonymous mutations in the *TP53* locus itself; in soft tissue sarcomas, for example, gene amplification of the p53 negative regulator *MDM2* occurs at a rate similar to *TP53* mutations (Leach et al., 1993).

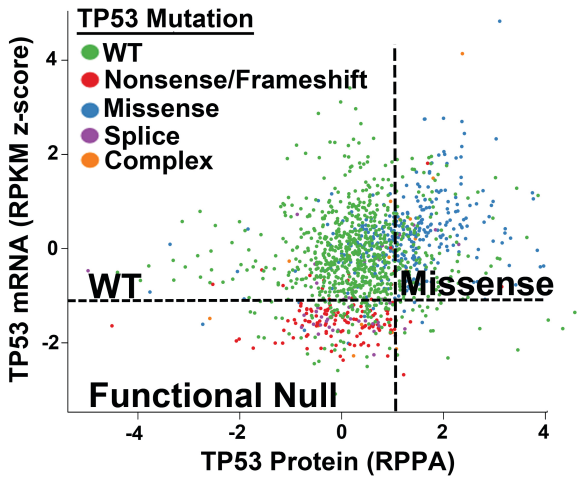


Figure 1: TP53 mutations result in differential mRNA and protein levels.
mRNA expression values (z-score) and reverse phase protein array (RPPA) protein abundance estimates from the TCGA breast cancer dataset were plotted and colored according to mutation category, listed in the legend at top left. Figure generated in collaboration with Brian Lehmann.

p53 as an Oncogene: Mutant Gain of Function

Although there is clear selection in many cancers for general p53 loss of function, the prevalence of *TP53* missense mutations – and the recurrently observed abnormal accumulation of p53 missense mutant proteins – has long intrigued the field. Building on the early experiments attributing transformation-enhancing potential to the misclassified p53 expression vectors, which contained mutant p53 cDNA, Levine and colleagues reported that mutant p53 proteins conferred “new or additional” phenotypes on cells lacking endogenous p53 protein, including enhanced tumor formation and clonogenic potential (Dittmer et al., 1993). These investigators were the first to use the term “gain of function” to describe this phenomenon (Soussi & Wiman, 2015).

The gain-of-function hypothesis received strong experimental support in 2004, when the Lozano and Jacks labs published simultaneous reports of mouse models expressing p53 hotspot mutants from the endogenous locus (Lang et al., 2004; Olive et al., 2004). Corroborating early findings, the Lozano group observed that mouse embryonic fibroblasts (MEFs) from homozygous p53 R172H mice (corresponding to the human R175H hotspot) exhibited increased transformation in a Ras-driven focus forming assay compared to homozygous p53-null MEFs. In their particular model, the overall survival and tumor spectrum of heterozygous R172H/+ mice were similar to those of heterozygous +/- (Lang et al., 2004).

In contrast, the Jacks group generated R172H and R270H mice (human R175H and R273H, respectively) in a different inbred strain and noted similar survival but significantly different tumor spectra from +/- mice in both the R172H/+ and R270H/+ genotypes (Olive et al., 2004). R172H/+ mice more frequently formed osteosarcomas and displayed more frequent metastasis, whereas R270H/+ mice developed carcinomas and B cell lymphomas with a significantly elevated frequency. The missense mutant-expressing strains exhibited loss of heterozygosity in 6/9 and 4/10 tumors analyzed, respectively, confirming their relevance to the p53 mutational selection observed in human cancers.

A later mouse model from the Xu group examined the effect of the R248W and R273H hotspot mutations introduced into a humanized p53 knock-in allele in mice (Song et al., 2007). As in both the Lozano and Jacks models, the survival of heterozygous mutant mice was similar to that of heterozygous nulls, but the spectrum of tumors observed differed significantly between the two genotypes. The mutant-expressing mice had a higher rate of interchromosomal translocations and a defect in the G2/M checkpoint, which was attributed to disruption of ATM activation by mutant interaction with the Mre11-Rad50-NBS1 (MRN) double-stranded break-binding complex.

An additional and convincing set of *in vivo* evidence for mutant p53 gain of function is found in a large cohort of French Li-Fraumeni families. Comparing individuals harboring germline p53 missense mutations versus

those harboring null mutations, the average age of tumor onset was significantly lower: 22.6 versus 37.5 years, respectively (Bougeard et al., 2008; Zerdoumi et al., 2013).

Gain-of-function phenotypes and mechanisms have been extensively documented in cell culture, with relevant publications now numbering in the hundreds (reviewed in Muller & Vousden, 2014). Making use of both exogenous overexpression and knockdown/knockout of endogenous mutants in 2D culture, 3D culture, and xenografts, research groups have documented a variety of gain-of-function phenotypes conferred by missense mutant p53, including enhanced invasion; altered migration potential; enhanced proliferation; cell cycle checkpoint inhibition; drug resistance; inhibition of apoptosis; anchorage-independent growth; increased colony formation; genomic instability; impaired cell polarity; cell fate dedifferentiation; enhanced growth in xenograft; enhanced metastasis; epithelial to mesenchymal transition; polyploidy; increased angiogenesis; and increased cell survival (Brosh & Rotter, 2009; Muller & Vousden, 2013, 2014; Oren & Rotter, 2010; Soussi & Wiman, 2015). The mechanisms postulated to explain these phenotypes are similarly varied, ranging from dominant-negative activity towards wild-type p53; inactivation of the p53 family members p63 and p73; aberrant recruitment of transcription factors and histone modifiers; *de novo* transcriptional activity, including a switch from activation to repression of wild-type p53 targets and vice versa; and inhibition of or cooperation with a variety of cancer-related pathways

processes, including the DNA damage response, autophagy, and growth factor signaling pathways (Brosh & Rotter, 2009; Muller & Vousden, 2014).

Complications in interpreting the large body of work surrounding gain-of-function mechanisms include differing phenotypes not only between distinct missense mutants, but also between the same missense mutant expressed in different cellular or tissue contexts. In Chapter 3, we demonstrate that this variability in gain of function even extends to isogenic clones expressing identical, endogenous mutant p53 alleles.

Therapeutic Targeting of Mutant p53

Therapeutic strategies under preclinical and clinical development to target mutant p53 fall into three main categories: restoration of wild-type activity, degradation of mutant protein, and inhibition of downstream pathways and/or protein-interacting partners (Muller & Vousden, 2014).

Restoring Wild-Type Activity

Many of the strategies to restore wild-type function to mutant p53 proteins are based on a mechanism of inducing proper folding and conformation of the DNA binding domain, thus making these approaches dependent on the presence of a destabilizing, structural mutant rather than a DNA contact mutant. The most advanced compound in this category, PRIMA-1, is thought to covalently bind to both wild-type and mutant p53, restoring wild-type biochemical activity and inducing apoptosis through a

mechanism that is still not fully understood (Lambert et al., 2009). APR-246, a PRIMA-1 analogue, completed a first-in-human safety trial in 2012 (Lehmann & Pietenpol, 2012; Lehmann et al., 2012), and is now being evaluated in a phase II clinical trial in ovarian cancer patients (NCT02098343).

Other mutant p53 stabilizing approaches target the structural deficiencies of particular missense mutants: the Fersht group identified a carbazole called PhiKan083 that partially rescues the destabilization of the Y220C DNA binding domain by binding to a Y220C-specific structural cavity (Boeckler et al., 2008; Liu et al., 2013). A similar compound, NSC319726, displays pronounced efficacy in cells harboring the R175H missense mutant (Yu et al., 2012). Despite the narrow specificity of these compounds, the potential patient population for a compound targeting a p53 hotspot mutant is large: as Soussi and Wiman note, among the top 15 most prevalent individual missense mutations among all mutants in the TCGA Pan-Cancer data, p53 R175H ranked fourth behind three PIK3CA hotspot mutants, and six other p53 missense mutants were in the top 15 (Kandoth et al., 2013; Soussi & Wiman, 2015).

Degrading Mutant p53

Evidence for the efficacy of mutant p53 degradation comes from a number of *in vitro* studies, many in triple-negative breast cancer cells, that have demonstrated that knockdown or loss of mutant p53 either sensitizes

cells to exogenous stress or induces outright cell cycle arrest or apoptosis (Braicu et al., 2013; Hui et al., 2006; Lim et al., 2009; Muller & Vousden, 2014; Vakifahmetoglu-Norberg et al., 2013; Zhu et al., 2011, 2013). As particularly convincing *in vivo* corroboration, the Moll group leveraged a mouse model expressing a floxed R248Q p53 missense mutant to demonstrate that loss of mutant p53 in already-established tumors induced tumor regression by apoptosis and stagnation, significantly extending animal survival (Alexandrova et al., 2015).

Several strategies to induce mutant p53 degradation have been demonstrated preclinically, including the use of HDAC inhibitors, which disrupt a HDAC6-Hsp70/90 interaction that has been shown to stabilize mutant p53 (Li et al., 2011a, 2011b). In their 2015 study, the Moll group demonstrated that Hsp90 and HDAC inhibition, either alone or in combination, significantly increased the survival of tumors with mutant, but not null, p53 alleles (Alexandrova et al., 2015). However, additional studies have demonstrated that HDAC inhibition decreases levels of both wild-type and mutant p53 *in vitro*, creating potentially deleterious side effects that must be approached with caution (Muller & Vousden, 2014; Yan et al., 2013). Several groups have also examined the role of glucose restriction-induced autophagy in promoting mutant p53 degradation but stabilizing wild-type p53 (Rodriguez et al., 2012), though the precise mechanism remains in debate (Vakifahmetoglu-Norberg et al., 2013).

Inhibiting Downstream Pathways and Interactions

Among the numerous mechanisms identified for mutant p53 gain of function, several have received focus for potential pharmaceutical inhibition. One example is the inhibition and/or sequestration of p53 family members p63 and p73 through interactions with the mutant p53 DNA binding domain (Di Como et al., 1999; Gaiddon et al., 2001). In this scenario, mutant p53 is thought to prevent these family members, especially p73, from transactivating downstream tumor suppressor pathways, many of which overlap with wild-type p53 (Li & Prives, 2007). To this end, putative small-molecule inhibitors of the mutant p53-p73 interaction, including RETRA, have displayed efficacy in suppressing *in vitro* and xenograft growth of mutant p53-expressing cells (Di Agostino et al., 2008; Kravchenko et al., 2008).

The reported interaction of mutant p53 with well-established, pharmaceutically actionable cellular pathways yields additional opportunities to inhibit mutant gain of function. In particular, the mevalonate pathway of cholesterol synthesis, which is the target of the widely prescribed cholesterol-lowering statin class of drugs, was found by the Prives group to be upregulated by mutant p53 in breast cancer cells (Freed-Pastor et al., 2012); in a 2016 study, the Iwakuma group reported that statin treatment suppressed mutant p53-expressing cancer cell growth by inhibiting mutant p53 interaction with the Hsp family member DNAJA1 (Parrales et al., 2016). Additional, potentially targetable mutant p53-

implicated cellular functions include TGF beta signaling, integrin recycling and the EGFR receptor tyrosine kinase/MAP kinase pathway (Adorno et al., 2009; Muller et al., 2009; Muller & Vousden, 2014; Sauer et al., 2010; Wang et al., 2013a).

The development of strategies to target mutant p53 gain of function activity based on published studies is complicated not only by the sheer volume of reported mechanisms, but also the inability to reproduce both phenotypes and pharmaceutical efficacy when assessing mutant p53 activity in a different cell context. A number of reports on mutant p53 gain of function activity in the literature also involve overexpression of mutant protein, which may produce an artificial readout of mutant p53 transcriptional or protein interaction capacity compared to physiological expression levels. To address these confounding effects, in Chapters 2 and 3, we describe the generation of an isogenic cell line panel in which endogenous wild-type, null, and missense mutant p53 alleles (including loss of heterozygosity) were generated in identical cell line backgrounds through CRISPR/Cas-mediated genome editing.

Gene Rearrangements

Historical Importance of Chromosomal Abnormalities

The link between chromosomal abnormalities and cancer, while long postulated, received experimental support in 1960, when Nowell and Hungerford published their findings of “a minute chromosome” – now known as the Philadelphia chromosome – in seven patients with chronic granulocytic/myelogenous leukemia (Nowell & Hungerford, 1960a, 1960b). Based on the technology at the time, these investigators were able to attribute some of the aberrant genomic material to chromosome 21 or 22, which was refined by Rowley over a decade later to a translocation between the long arms of chromosomes 9 and 22 (Rowley, 1973). In 1982, de Klein *et al*. narrowed the Philadelphia chromosome to a specific gene on chromosome 9: *ABL1*, a homologue of the transforming Abelson murine leukemia virus (de Klein et al., 1982). A decade later, identification and molecular characterization of similar abnormalities had grown rapidly, with over 50 chromosomal rearrangements identified primarily in hematopoietic malignancies (Rabbitts, 1994).

The net effect of these rearrangements was recognized at the time to consist of two main outcomes: activation of a proto-oncogene through gene locus perturbation or swapping, and the formation of a tumor-specific, chimeric protein through gene fusion (Rabbitts, 1994). In addition to the creation of kinases with increased activity, as in *BCR1-ABL*, many chimeric

proteins also were found to harness transcription factors; a canonical example is the *PML-RARA* fusion, which involves the DNA-binding retinoic acid receptor (Alcalay et al., 1991; Borrow et al., 1990; de Thé et al., 1990).

The *PML-RARA* and *BCR-ABL1* examples additionally illustrate the potential efficacy of targeting chimeric proteins. Patients with acute promyelocytic leukemia, which universally exhibit the *PML-RARA* gene fusion, are uniquely sensitive to treatment with all-*trans* retinoic acid (ATRA). This therapy dissociates transcriptional cofactors from the retinoic acid receptor and induces differentiation of leukemic promyelocytes (Huang et al., 1988; Warrell et al., 1991). Interestingly, clinical observations of the efficacy of ATRA treatment actually preceded the identification of *RARA* as a partner in the chromosomal rearrangement. In the case of *BCR-ABL1*, however, the development of a pharmaceutical inhibitor lagged behind molecular characterization of the translocation by decades. Imatinib (widely known by its trade name, Gleevec) was identified in the 1990s by a high-throughput kinase inhibitor screen at Novartis and proved to be a molecularly and clinically effective inhibitor of the BCR-ABL fusion kinase (Druker et al., 2001; Schindler et al., 2000).

Evolution in Detection Strategies

Our understanding of the number and complexity of gene rearrangements in human cancer has increased as a function of the sensitivity of our detection strategies (Mitelman et al., 2007). Initial

molecular characterization of chromosomal abnormalities leveraged chromosomal banding techniques, which could pinpoint individual regions of chromosomes but often lacked the resolution to detect changes at the gene locus level (Caspersson et al., 1970; Mitelman et al., 2007).

The advent of nucleotide probes, including fluorescence *in situ* hybridization (FISH) and array-based strategies such as comparative genomic hybridization (CGH), improved the resolution of cytogenetic analysis substantially (Kearney & Horsley, 2005; Pinkel & Albertson, 2005; Speicher & Carter, 2005). By 2007, at the dawn of the era of next-generation sequencing, 337 genes had been reported to be involved in chromosomal rearrangements (Mitelman et al., 2007).

The introduction of next-generation sequencing (NGS) technologies in the mid-2000s heralded a critical technological advance: in contrast to hybridization techniques, which required prior probe design, NGS facilitated unbiased gene fusion detection. By comparing RNA or DNA sequences to the reference genome, chimeric transcripts or genomic sequences can be identified at nucleotide resolution without prior knowledge of specific chromosomal rearrangements (Annala et al., 2013; Davare & Tognon, 2015; Mertens et al., 2015). It should be noted that any form of genomic instability leading to chromosomal breakpoints and rearrangement – including copy number-altering gene amplification and deletion, rather than the balanced translocations seen classically in hematological malignancies – has the potential to produce hybrid sequences detectable at the RNA or

DNA level; thus, not all chimeric transcripts will represent actionable gene fusion events. However, proper sequence annotation, consideration of protein domain loss or retention, and integrative analysis with copy number data are effective methods of prioritizing novel hybrid sequences for further biological validation.

Numerous algorithms and computational approaches have been developed to identify gene rearrangements using RNA-seq and WXS (whole-exome sequencing) or WGS (whole-genome sequencing) data, but a number of difficulties remain, due to both the limitations of the sequencing technology and compromises made by fusion detection algorithms (Davare & Tognon, 2015). These algorithms are designed to identify discordant or breakpoint-spanning reads in which part of the sequence maps elsewhere in the genome, but this approach is confounded by homologous regions and highly repetitive sequences throughout the genome. As a result, many gene fusion detection pipelines integrate filters to reduce false positives from homologous or repetitive regions, such as restricting potential rearrangement partners to annotated gene loci (Annala et al., 2013). While these filters are effective in enriching for existing gene fusions, they limit our ability to detect more novel events, such as rearrangements involving non-coding genomic regions. In Chapter 4, we describe discovery of both classical protein-coding gene fusions and novel rearrangements of promoters and regulatory regions with non-coding DNA using a prediction algorithm that prioritizes putative gene rearrangements

for analysis based on population-level comparisons; our approach substantially broadens the scope of potential rearrangement partners and significantly reduces false discovery and false positive events (Shaver et al., 2016).

Current Knowledge and Clinical Impact

The rapid expansion in tumor sequencing at the international, national, and institutional levels has enabled researchers to catalogue an expanse of both recurrent and low-frequency gene rearrangements. Although it is difficult to ascertain whether an individual gene rearrangement acts as a 'driver' event or is simply a collateral result of genomic instability, focusing on known oncogenic driver categories – including in-frame kinase protooncogene fusions in which the kinase domain is retained – allows an impression of the potential clinical impact of gene rearrangement detection. Analyzing nearly 7,000 tumors from TCGA, Lengauer and colleagues identified a 3% frequency of oncogenic, recurrent kinase fusions, many of which can be targeted by pharmaceutical agents in current clinical use or development (Stransky et al., 2014). Kinase fusion frequency varied by cancer type in this analysis, with a 13% rate in thyroid cancers but none detected in 66 and 529 cases of kidney chromophobe and clear cell carcinomas, respectively. In both their analysis and our studies in Chapter 4, many fusions typically associated with an individual cancer type can be identified at lower frequency in other malignancies,

suggesting that a tissue-agnostic approach to gene fusion detection may allow the repurposing of existing treatment strategies.

Despite our recent ability to identify and validate gene rearrangements as potential oncogenic alterations, particularly in carcinomas, the number of pharmaceuticals specifically approved by the FDA for gene fusions is low; other than imatinib and newer-generation *BCR-ABL1* therapeutics, only crizotinib and ceritinib – both targeting *ALK* fusions in non-small cell lung cancer – have received approval (Mertens et al., 2015). New therapeutics are on the horizon: larotrectinib, an inhibitor of *TRK* family fusions that occur in a wide variety of tumor types, has shown antitumor efficacy in clinical trials and has recently been submitted to the FDA for approval (Hyman et al., 2017).

Further, many of the recurrently rearranged genes recently identified in solid tumors, including *BRAF*, *EZH2*, *FGFR3*, *MLL*, *RET*, and *ROS1,* represent kinases, growth factor signaling pathways, and chromatin modifiers that are amenable to modulation by existing therapeutics (Kandoth et al., 2013; Mertens et al., 2015; Stransky et al., 2014). Rather than requiring the development of a *BRAF* fusion-specific inhibitor, for instance, patients harboring this rearrangement would likely respond well to existing inhibitors targeting the downstream MAP kinase pathway. In Chapter 4, we identify and biologically validate two previously unreported gene fusions in TNBC that involve proteins targetable by inhibitors already in use or under development (Schlegel et al., 2013; Shaw et al., 2013).

It is also important to note that gene rearrangements can indicate oncogenic dependence on a particular pathway without producing a chimeric protein or even chimeric mRNA. One of the earliest recurrent translocations to be discovered, *IGH-MYC*, drives expression of the complete Myc oncoprotein by juxtaposition with regulatory elements of the immunoglobulin heavy chain gene (Rabbitts & Boehm, 1991). *CD274*, encoding the PD-L1 checkpoint protein that is a major focus of current immunotherapy clinical trials, is upregulated in many cancers due to structural variants disrupting the 3' negative-regulatory region (Kataoka et al., 2016; Shaver et al., 2016). In Chapter 4, we describe additional examples of 5' and 3' regulatory disruption by rearrangement with both coding and non-coding regions of the genome; in one example, we document that upregulation of a complete *HRAS* oncogene-encoding transcript occurs due to rearrangement upstream of its transcriptional start site. In these examples, such structural variants are highly suggestive of dependence on the upregulated oncogenes and potential amenability to targeting of the affected pathway, rather than a specific chimeric protein.

CHAPTER 2

MATERIALS AND METHODS

Cell Culture and *in vitro* Experiments

Cell Culture

CAL51 cells (DSMZ, September 2013), CAL51 isogenic clones, and 293FT cells (Life Technologies) were cultured in DMEM (Gibco 11965) with 5% (v/v) FBS (Gemini). SUM185PE cells (Asterand, March 2010) were cultured in Ham's F-12 (Gibco 11765) with 5% (v/v) FBS (Gemini), 1 mg/mL hydrocortisone (Sigma-Aldrich), and 5 µg/mL insulin (Novo Nordisk). MCF10A cells (ATCC, June 2012) were cultured in DMEM:F12 (Gibco 11330) with 100 ng/mL cholera toxin (Sigma-Aldrich), 500 ng/mL hydrocortisone (Sigma-Aldrich), 20 ng/mL human epidermal growth factor (Life Technologies/Thermo Fisher), and 10 µg/mL insulin (Novo Nordisk). Ba/F3 cells (provided by Dr. Christine Lovly, Vanderbilt University, November 2014) were cultured in RPMI + GlutaMAX (Gibco 61870) with 1 ng/mL IL-3 (Life Technologies) and 5% (v/v) FBS (Gemini).

All cell lines were maintained in 100 U/mL penicillin and 100 µg/mL streptomycin (Gemini) and tested negative for mycoplasma (Lonza). Cell Line Genetics performed positive short tandem repeat DNA fingerprinting analysis on CAL51 and all initial isogenic clones (WT A1, R175H A1 and

A2, R273H A1 and A2, Null A1 and A2) in March 2016 and on SUM185PE in March 2011. We also verified CAL51 (including isogenic mutant status) and SUM185PE by manual identification of unique variants in NGS data. DNA fingerprinting analysis was not performed on MCF10A or Ba/F3 cells, but the Ba/F3 cell line displayed the previously published IL-3 dependence phenotype (Palacios & Steinmetz, 1985).

CRISPR/Cas-Mediated Genome Editing

Genome editing of CAL51 using the CRISPR/Cas9 system was performed in close adherence to the Zhang lab's protocol (Ran et al., 2013). Guide RNAs (gRNAs) were designed in the Benchling web tool; we selected 20-nt gRNAs (3-5 per desired mutation site) with the highest target specificity score that cut within at least 10 nt of the desired mutant base for screening. Complementary guide oligos were ordered (Invitrogen) with an initial 5' guanine (if not already present) and BbsI overhangs for cloning into pSpCas9(BB)-2A-GFP (PX458) (Addgene #48138). Guide oligos were resuspended at 100 μM and phosphorylated and annealed by incubation with T4 PNK and ligation buffer (New England BioLabs). Cloning into PX458 was conducted by incubating the plasmid and diluted oligo duplexes with Tango buffer (Fermentas/Thermo Scientific), DTT (Fermentas/Thermo Scientific), ATP (New England BioLabs), FastDigest BbsI (Fermentas/Thermo Scientific), and T7 ligase (New England BioLabs). The ligation products were subsequently treated with PlasmidSafe ATP-

dependent DNase (Epicentre) to digest residual linear DNA. Cloned plasmids were transformed into DH5α competent cells (Vanderbilt Molecular Biology Core Facility) and selected on LB agar plates containing 100 μg/mL ampicillin (Vanderbilt Molecular Biology Core Facility). Plasmid DNA was isolated from the resulting colonies using a QIAprep Spin Miniprep Kit (Qiagen) and sequence verified by Sanger sequencing (GenHunter).

To assess gRNA activity, cloned plasmids were transfected into 293FT cells using Lipofectamine 3000 (Thermo Fisher). After 72 h, DNA was isolated from the transfected cells using QuickExtract DNA extraction solution (Epicentre) per manufacturer's recommendations. Genomic DNA was amplified with primers designed to generate an asymmetric 300-500 bp PCR product around the gRNA cut site (Table 1) using Platinum Taq DNA Polymerase High Fidelity (Thermo Fisher). PCR products were purified using a QIAQuick PCR Purification Kit (Qiagen) and annealed by gradual reduction in temperature from 95 °C, followed by SURVEYOR digestion using SURVEYOR nuclease S and enhancer S with supplemental magnesium chloride. SURVEYOR digestion products were run on a 5% polyacrylamide TBE gel and visualized using SYBR Gold (Life Technologies). Guide RNAs producing a prominent asymmetric digestion product were selected for further use (Table 2). A negative control (no transfected plasmid) was included to rule out digestion products generated by endogenous SNP mismatches.

Associated Clones	Direction	Sequence
175A1-2, Null A1-A2	Forward	CACTTGTGCCCTGACTTTCA
175A1-2, Null A1-A2	Reverse	TTGCACATCTCATGGGGTTA
273A1-2, all B set	Forward	CCTCTGCTTGCCTCTGACCCCT
273A1-2, all B set	Reverse	TGCACCCTTGGTCTCCTCCACC

Table 1: Primer sequences for SURVEYOR gRNA activity assessment and clone genotyping.
The indicated primer pairs were designed to amplify a 300-500 nt amplification product from genomic DNA that produces asymmetric digestion products when examined for guide RNA activity by the SURVEYOR assay or for successful HDR incorporation by restriction digest screening. Primers were also used for genotyping of original single-cell clones and allele cloning, if necessary.

Associated Clones	Sequence
175A1-2, Null A1	TATCTGAGCAGCGCTCATGG
Null A2	CACCGTCCTCAGCATCTTATCCGAG
273A1-2, all B set	CACCGTGCGTGTTTGTGCCTGTCC

Table 2: Guide RNA sequences used to generate isogenic cell line panel.
Guide RNA oligos were cloned into pSpCas9(BB)-2A-GFP as described in Chapter 2 and demonstrated successful nuclease activity by SURVEYOR screening. The sequences listed above were used to generate the indicated clones.

Once the gRNA was selected for a specific mutant site, homology-directed repair (HDR) templates were designed using the Benchling web tool. HDR templates were designed to contain mutant bases corresponding to clinically observed hotspot mutations, along with a conservative mutation in the protospacer adjacent motif (PAM) to prevent CRISPR/Cas-mediated cleavage of the HDR-recombined alleles and ~75 bp of flanking homologous sequence on each side (Table 3). When the engineering of a conservative PAM mutant was not possible due to codon position, 2 additional conservative mutations were engineered into the gRNA target region to reduce complementarity. HDR templates were ordered as Ultramer oligos (IDT).

CRISPR/Cas-mediated editing of CAL51 cells and clones was conducted by incubating 2.5 µg of cloned PX458 plasmid, 5 µL of P3000 reagent (Thermo Fisher), 7.5 µL of Lipofectamine 3000 (Thermo Fisher), and (if attempting knock-in mutagenesis rather than frameshift knockout) 5 µL of 10 µM single-stranded HDR template in 250 µL of Opti-MEM (Gibco); the resulting DNA-lipid complexes were added to one well of a sub-confluent 6-well plate, with the quantity of wells scaled as necessary. After 48 h, transfected cells were trypsinized, washed in PBS, resuspended in 4% (v/v) FBS (Gemini) in PBS, and stained with 0.5 µg/mL propidium iodide (Sigma-Aldrich) for live/dead discrimination. GFP-positive live cells were gated using lipofectamine-only negative control cells with and without propidium iodide along with transfected cells lacking propidium iodide.

Associated Clones	Sequence
R175H A1-2	CACCCCCGCCCGGCACCCGCGTCCGCGCCATGGCCATC TACAAGCAGTCACAGCACATGACGGAGGTTGTGAGGC**A**C TGCCCaCACCATGAGCGCTGCTCAGATAGCGATGGTGAG CAGCTGGGGCTGGAGAGACGACAGGGCTGGTTGCC
R273H A1-2, R273H B1-10	TCCTTACTGCCTCTTGCTTCTCTTTTCCTATCCTGAGTAGT GGTAATCTACTGGGACGGAACAGCTTTGAGGTGC**A**TGTT TGcGCtTGTCCTGGGAGAGACCGGCGCACAGAGGAAGAG AATCTCCGCAAGAAAGGGGAGCCTCACCACGA

Table 3: Homology-directed repair templates used to generate isogenic cell line panel.

The indicated homology-directed repair templates were cotransfected with CRISPR/Cas and guide RNA-encoding plasmids to generate knock-in mutant cell lines, as described in Chapter 2. Bolded nucleotides indicate the clinically observed missense mutation and lower-case nucleotides depict conservative mutations designed to disrupt the PAM site (R175H) or reduce guide RNA complementarity (R273H).

Fluorescence-activated cell sorting was performed by the Vanderbilt Flow Cytometry Shared Resource. GFP+, PI- cells were sorted into 96-well plates containing complete culture medium for isolation of single cell clones.

Clonal populations were expanded 21-28 d after sorting and DNA was isolated using QuickExtract DNA extraction solution (Epicentre). Genomic DNA was amplified with primers designed to generate an asymmetric 300-500 bp PCR product around the gRNA cut site (Table 1), using Platinum Taq DNA Polymerase High Fidelity (Thermo Fisher). PCR products were purified using a QIAQuick PCR Purification Kit (Qiagen) and screened using restriction enzyme cut sites present only in HDR-recombined sequences (BtsI for R175H, NlaIII for R273H). Cells passing restriction digest checks were further verified by Sanger sequencing. When mixed traces were present due to heterozygous frameshift alleles, allele cloning was conducted by amplifying genomic DNA using primers from Table 1 with EcoRI or BamHI restriction site overhangs, followed by digestion and cloning into the pUC19 vector (New England BioLabs) for transformation. Once bacterial colonies were selected and expanded, DNA isolation was performed using a QIAprep Spin Miniprep Kit (Qiagen) and individual clones were Sanger sequenced (GenHunter) to identify single-trace alleles (Ran et al., 2013).

For CRISPR/Cas-mediated generation of knockout cells only, the resulting colonies were screened by immunoblot for a lack of p53 protein,

then confirmed by Sanger sequencing (GenHunter) and (if necessary) allele cloning, as described above.

Metabolic Rate Analysis

CAL51 cells or its isogenic clones were plated in a 96-well format and analyzed while subconfluent. Cells were assessed with alamarBlue reagent (Thermo Fisher) per the manufacturer's protocol. Blank wells were included to subtract background fluorescence. Following metabolic analysis, wells were washed with PBS, fixed with 4% paraformaldehyde (VWR) in PBS for 10 minutes, and stained with 0.5 μg/mL DAPI (Thermo Fisher) in PBS for 10 minutes, followed by additional washing. Nuclei in each well were imaged at low magnification and counted with ImageJ (NIH) using a custom macro; briefly, images were threshold processed using default settings, inverted, and processed with the watershed algorithm to resolve nuclei in close proximity, then particles were analyzed with minimum size of 15 pixels and circularity between 0.30 and 1.00. Alamar fluorescence values were divided by nuclei counts for each well to yield a per-cell metabolic rate estimate, and the presented values were normalized to the average of CAL51.

Cell Size Estimation

One million cells from CAL51 or its isogenic clones were prepared for flow cytometry cell cycle analysis by trypsinization and washing with PBS,

followed by fixation in 70% ethanol. Cells were stained in a solution consisting of 0.1% (v/v) Triton X-100 (Sigma-Aldrich) in PBS with 0.2 mg/mL DNAse-free RNAse A (Sigma-Aldrich) and 20 µg/mL propidium iodide (Sigma-Aldrich). Samples were analyzed at the Vanderbilt Flow Cytometry Shared Resource on a 3-laser LSRII (BD). Single cells were selected by pulse processing and cell cycle stages were gated by propidium iodide intensity. Cell size data are presented as forward scatter pulse area (FSC-A) of single cells in the 2N (or equivalent) phase of the cell cycle.

Xenograft Tumor Growth

Five million cells of CAL51 or its isogenic clones were resuspended in 200 µL of PBS or a 1:1 mixture of PBS and standard formulation Matrigel (BD). Cell suspensions were injected subcutaneously into the flanks of NSG mice (Jackson) or Crl:NU(NCr)-Foxn1nu athymic nude mice (Charles River) and observed for the indicated time periods. Once tumors had reached maximum size as defined by animal protocols, mice were euthanized and the tumor volume was estimated by xenograft measurement using the formula (length * width * width)/2. All experiments were conducted under the approval and guidelines of the Vanderbilt Institutional Animal Care and Use Committee.

Metaphase Spreads

Sub-confluent cells (~75%) were treated in a 10-cm culture dish for the indicated durations with 0.5 µg/mL KaryoMAX colcemid (Thermo Fisher). Media was removed and reserved, cells were trypsinized and resuspended in the reserved media. After centrifugation, cells were gently resuspended and combined with 5 mL of 0.075M potassium chloride (Fisher) while vortexing at low speeds. Cells were incubated at 37 °C with periodic mixing. After centrifugation, cells were gently resuspended and combined with 10 mL of a pre-ice chilled 3:1 methanol:glacial acetic acid fixative solution (EMD) while vortexing at low speeds.

After storage at 4 °C, cells were centrifuged and resuspended in 1 mL of fixative solution. Superfrost Plus microscope slides (Fisher) were pre-chilled at -20 °C and then placed on bench top to form condensation. Cell suspensions were dropped from a height of ~6 in onto slightly tilted slides, followed by overnight air drying, mounting with ProLong Gold Antifade Reagent with DAPI (Invitrogen), and coverslipping. Metaphase spreads were imaged using fluorescence microscopy with an oil-immersion 100x objective, and individual chromosomes within each nucleus were counted using ImageJ (NIH).

p53 Isogenics Drug Response

CAL51 cells and isogenic clones were treated with the indicated doses of doxorubicin (Pfizer) or an equivalent volume of PBS, (-)-Nutlin-3

(Johnston Lab, Vanderbilt University) or an equivalent volume of DMSO, and lovastatin (Cayman Chemical), atorvastatin (Cayman Chemical), or an equivalent volume of DMSO. Following the indicated duration of treatment, cells were lysed and immunoblotted as described below.

Cloning and Generation of Stable Cell Lines

The TMEM87B-MERTK expression construct, corresponding to amino acids 1-55 of TMEM87B (NM_032824) and amino acids 433-1000 of MERTK (NM_006343), was synthesized in a pMK-RQ-Bb vector using GeneOptimizer and GeneArt (Life Technologies) and cloned into pBABE-puro (Morgenstern & Land, 1990) (Addgene) by EcoRI/SalI restriction digest and ligation (New England BioLabs). The complete pBABE-puro-TMEM87B-MERTK expression construct sequence is available upon request. To generate stably transfected cell lines, pBABE-puro-TMEM87B-MERTK or pBABE-puro empty vector retroviruses were packaged in Phoenix cells (Orbigen) and Ba/F3 and MCF10A cells were transduced with virus for two 24 h intervals with 8 μg/mL polybrene (Sigma-Aldrich), then selected and maintained with 1.5 μg/mL (Ba/F3) or 0.5 μg/mL (MCF10A) puromycin (Sigma-Aldrich).

IC_{50} Determination

SUM185PE cells were seeded in quadruplicate (8000 cells/well) in 96-well plates. After overnight attachment, growth medium was replaced

with medium (control) or medium containing half-log serial dilutions of the FGFR inhibitor PD173074 (Selleckchem). Viability was assessed at 72 h by incubating cells with alamarBlue (Invitrogen). Half-maximal inhibitory concentration (IC_{50}) values were determined after normalization to untreated wells and double-log transformation of dose response curves to fit a linear regression predicting a concentration producing 50% viability.

Ba/F3 IL-3 Withdrawal

Stably transfected Ba/F3 cells were seeded in 6-well plates (40,000 cells/well) in growth medium ± 1 ng/mL IL-3. Live cell counts were manually determined at three sequential 24 h timepoints by visual inspection using a hemocytometer and Trypan blue (Bio-Rad).

Immunoblotting

All cells were lysed in RIPA buffer supplemented with phosphatase inhibitors (50 mM sodium fluoride, 0.2 mM sodium vanadate, 10 mM p-nitrophenyl phosphate) and protease inhibitors (10 mg/mL antipain, 10 mg/mL leupeptin, 10 mg/mL pepstatin, 10 mg/mL chymostatin [Sigma-Aldrich], 200 mg/mL 4-(2-aminoethyl)-benzenesulfonylfluoride [Millipore]). Protein quantification was conducted using the DC Protein Assay (Bio-Rad). Lysates were separated on polyacrylamide gels and transferred to polyvinyl difluoride membranes (Millipore). In Chapter 3, cells were lysed under the indicated conditions and/or at the indicated timepoints after

treatment. Immunoblotting was performed using p53 DO1 (1:1000, Santa Cruz sc-126), GAPDH MAB374 (1:1000, Millipore), MDM2 3G9 (1:2000, Millipore), and p21 DCS60 (1:2000, Cell Signaling Technologies).

In Chapter 4, SUM185PE cells were lysed 72 h after siRNA transfection. Ba/F3 cells were grown in suspension and incubated for 90 m in the indicated media before lysis. MCF10A cells seeded in 10 cm plates were incubated for 180 m in the indicated growth media before lysis. Immunoblotting was performed using FGFR3 B-9 (1:200, Santa Cruz), GAPDH MAB374 (1:1000, Millipore), MERTK D21F11 (1:1000, Cell Signaling), phospho-Akt Ser473 D9E (1:2000, Cell Signaling), total Akt (1:1000, Cell Signaling 9272), phospho-Erk1/2 Thr202/Tyr204 D.13.14.4E (1:2000, Cell Signaling), and total Erk1/2 3A7 (1:2000, Cell Signaling).

RNA and DNA Sequencing and Hybridization

CAL51 and Isogenic Clones RNA Sequencing

Total RNA was isolated from CAL51 and its isogenic clones using the Aurum Total RNA Mini kit (Bio-Rad). Five hundred ng of RNA was submitted to the Vanderbilt Technologies for Advanced Genomics (VANTAGE) core for polyA-selected, stranded mRNA library preparation. RNA and library quantity were confirmed by Qubit (Life Technologies) and qPCR, and library size and quality were assessed by Bioanalyzer (Agilent). RNA-seq was conducted at paired-end 75 bp on the Illumina HiSeq 3000.

Demultiplexed raw sequence files were trimmed to remove adapter sequences using the Flexbar v2.5 utility (Dodt et al., 2012) and were aligned to hg19 using the STAR aligner, v020201 (Dobin et al., 2013) and GENCODE (Harrow et al., 2012) v25lift37 comprehensive gene annotations. Cufflinks v2.2.1 (Roberts et al., 2011) was used to assemble transcripts and quantify transcript abundance. FPKM estimates from Cufflinks were converted to TPM estimates (Li et al., 2010) using a custom script.

For differential gene expression analysis, STAR-aligned reads were assigned to GENCODE v25lift37comprehensive genes using featureCounts v1.5.0-p3 (Liao et al., 2014) and assessed using DESeq2 v1.14.1 (Love et al., 2014) using default parameters. Genes with an adjusted p-value < 0.1 were considered significantly differentially expressed.

Array Copy Number Estimation

Whole-genome copy number analysis was performed using the CytoScan HD SNP microarray platform (Affymetrix), which uses over 743,000 SNP probes and approximately 1,953,000 non-polymorphic copy number probes with a median spacing of 880 bp. Briefly, 20 ng of whole genomic DNA isolated from cultured cell lines was digested with NspI, ligated with NspI adapter primers, and amplified using Platinum Taq (Thermo Fisher) with a GeneAmp PCR System 9700 (Thermo Fisher).

PCR products were purified and fragmented, labeled with biotin, and hybridized to the microarray chip. Chips were washed, stained, and scanned on an Affymetrix scanner. Data was analyzed using proprietary software (Chromosome Analysis Suite, Affymetrix) to detect copy number gains and losses as well as regions of homozygosity. Data analysis was conducted based on the GRCh37/hg19 genome assembly.

SUM185PE RNA Sequencing

Total RNA was isolated from SUM185PE cells using RNeasy (Qiagen). RNA quality was assessed by NanoDrop (Thermo) and Bioanalyzer (Agilent). Two μg of total RNA were used for the TruSeq Stranded Total RNA Library Prep Kit (Illumina). Libraries were quantified by Qubit (Life Technologies) and qPCR, and library size and quality were assessed by Bioanalyzer (Agilent). The constructed RNA-seq library was sequenced at the Vanderbilt Technologies for Advanced Genomics (VANTAGE) core on an Illumina HiSeq 2500 using a paired-end 100-bp protocol. Reads were de-multiplexed and trimmed using SeqPrep. FASTQ files are deposited in the NCBI Read Sequence Archive under accession number SRP077076.

SUM185PE RNA-seq reads were aligned using the TCGA RNA-seq v2 pipeline (cghub.ucsc.edu/docs/tcga/UNC_mRNAseq_summary.pdf, 7/31/2013 revision) to ensure compatibility with TCGA data. Reference data and custom scripts for exon-level expression quantification were

downloaded from the UNC database as referenced in the protocol. Alignment was conducted using the default TCGA workflow and MapSplice v12_07, RSEM v1.1.13, and UBU v1.2. Cell line RNA-seq data were grouped with the TCGA BRCA dataset for STA input and processed using the STA discovery pipeline described below.

Computational and Statistical Methods

TCGA Data Acquisition

To generate input for the Segmental Transcript Analysis discovery pipeline, we collated RNASeq Version 2 data from 5461 tumors across 14 studies from TCGA (Table 4). Per TCGA practice, the colon adenocarcinoma (COAD) and rectum adenocarcinoma (READ) data sets were combined into a single COADREAD group for STA analysis(Cancer Genome Atlas, 2012a). Exon-level expression data for each study were obtained from the Broad GDAC Firehose stddata__2013_12_10 run (doi:10.7908/C1QZ294H). Aligned RNA-seq files and portions of whole-genome sequencing (WGS) files were downloaded from the Cancer Genomics Hub (CGHub) using the gtdownload and GTFuse (Annai Systems) utilities. When two or more RNA-seq or WGS BAM files were present for a given sample, the largest file was selected to download.

TCGA Studies and Files Analyzed

Study ID	Cancer Type	RNA-seq files Analyzed	WGS files analyzed
BLCA	Bladder urothelial carcinoma	211	113
BRCA	Breast invasive carcinoma	994	113
COADREAD	Colorectal adenocarcinoma	333	22
GBM	Glioblastoma multiforme	166	31
HNSC	Head and neck squamous cell carcinoma	425	155
KICH	Kidney chromophobe	66	44
KIRC	Kidney renal clear cell carcinoma	507	34
LGG	Brain lower grade glioma	306	71
LUAD	Lung adenocarcinoma	491	145
LUSC	Lung squamous cell carcinoma	482	49
PRAD	Prostate adenocarcinoma	256	116
SKCM	Skin cutaneous melanoma	356	138
THCA	Thyroid carcinoma	498	119
UCEC	Uterine corpus endometrial carcinoma	370	114
	Total	5461	1264

Table 4: Cancer types and sample numbers analyzed from TCGA.
List of 14 cancer types from TCGA, corresponding study IDs, and the number of RNA-seq and whole-genome sequencing files analyzed per cancer type.

TCGA Breast Cancer Subtype Assignment

Since many TCGA BRCA samples lack complete clinical ER, PR, and/or HER2 annotation, we integrated clinical metadata with mRNA expression levels to assign breast cancer subtypes for the entire cohort. Our previous work has shown that TNBC cases can be reliably identified through ER, PR, and HER2 mRNA expression analysis, and additional studies have shown that ER and HER2 mRNA expression correlates with IHC and FISH analyses, respectively (Carmeci et al., 1997; Lehmann et al., 2011; Press et al., 2008). We obtained RNA-seq Version 2 genes normalized and reverse phase protein array (RPPA) data for the BRCA dataset from the Broad GDAC Firehose stddata__2013_06_06 run (doi:10.7908/C1N58KHG). The mRNA expression distributions of ER, PR, and HER2 were analyzed and cutoffs for ER, PR, and HER2 positivity (1000, 140, and 250000, respectively) were selected to differentiate bimodal states of expression (Fig. 2A-C). ER, PR, and HER2 calls were manually reviewed to test concordance with RPPA protein expression and assess sample association by principal component analysis of global gene expression. A small number of samples were manually curated by this method. Overall, 173 of the 994 breast cancer samples analyzed (17.5%) were annotated as TNBC.

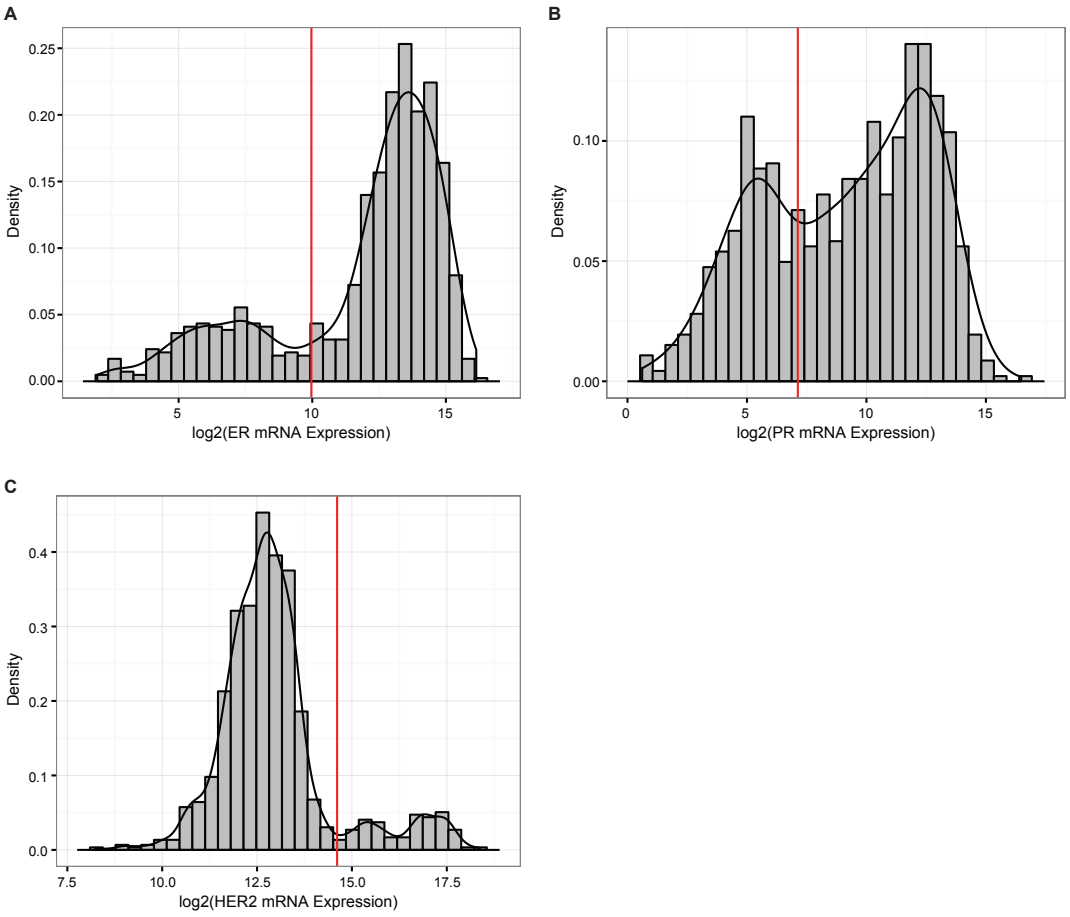


Figure 2: Discrete mRNA expression cutoffs differentiate breast cancer subtypes. Log-transformed mRNA expression values (TPM) of ER (A), PR (B), and HER2 (C) plotted as histograms of density over each interval. Vertical red lines indicate expression level chosen to represent a positive subtype call.

Data Acquisition from the Aparicio Group

RNA-seq data for 80 TNBC clinical specimens from the Aparicio group (Shah et al., 2012) were downloaded from the European Genome-phenome Archive as EGAS00001000132 on 5/18/2014. Aligned BAM files were converted to fastq using bedtools v2.17.0 and were aligned using the TCGA RNA-seq v2 pipeline, as described above, and processed using the STA discovery pipeline, as described below.

Transcript and Protein Annotation

Gene locus and exon data were determined by GENCODE and RefGene annotations as described below in the STA discovery pipeline. For exon-level expression plots, the exon order is based on sequential order of exons in the RefGene annotation. The presence of isoforms may cause deviation from the normal exon numbering scheme. For hybrid transcript frame calls, the RefGene annotation was processed to generate starting and ending frame values of 0, 1, 2 (CDS) or -1 (UTR) for each exon. Due to the potential for multiple reading frames at a given coordinate, a hybrid transcript between two exon boundaries featuring any concordance of starting and ending CDS frame was annotated as in-frame. A hybrid transcript between two exon boundaries with non-overlapping CDS frames was annotated as out-of-frame. An exon boundary with exclusively UTR status was annotated accordingly.

The domains and protein features depicted in schematics were obtained from UniProtKB (Magrane & Consortium, 2011). Features were exclusively selected from annotations with “Reviewed” status, and all coding features are to scale.

Segmental Transcript Analysis Algorithm

The median-normalized, exon-level RPKM vector for sample *i* in gene *k* is denoted as $\widetilde{\boldsymbol{RPKM}}_{ik}$. The Fscore for sample *i* in gene *k* is defined as

$$Fscore_{ik} = \frac{d_{ik}^{G} D_i}{\underset{\forall i}{stdev}\, d_{ik}^{G} D_i}$$

where d_{ik}^{G} is the geometric mean of $d(\widetilde{\boldsymbol{RPKM}}_{ik}, \widetilde{\boldsymbol{RPKM}}_{i'k})$, $\forall i \neq i'$ and $d(\mathbf{x}, \mathbf{y})$ is a distance measurement (i.e. Euclidean distance) between **x** and **y**. D_i is the biggest difference of RPKM with lag 1 in sample *i* and $stdev_{\forall i}\, d_{ik}^{G} D_i$ is the standard deviation of $d_{ik}^{G} D_i$ for gene *k*. For ranking Fscore across genes, the normalized Fscore is defined as

$$Gscore_{ik} = \frac{Fscore_{ik} - \min\limits_{\forall i, \forall k} Fscore_{ik}}{\max\limits_{\forall i, \forall k} Fscore_{ik} - \min\limits_{\forall i, \forall k} Fscore_{ik}}.$$

For ranking Gscore between samples, the Segmental Transcript Analysis score is defined as

$$STAscore_{ik} = \log_2\left(\frac{Gscore_{ik} - \underset{\forall k}{mean}\, Gscore_{ik}}{\underset{\forall k}{stdev}\, Gscore_{ik}} + 2\right),$$

where $mean_{\forall k}\, Gscore_{ik}$ is the mean of Gscore for all genes in sample *i* and $stdev_{\forall k}\, Gscore_{ik}$ is the standard deviation of Gscore for all genes in sample *i*.

Corresponding R code is included as a supplemental file to Shaver, Lehmann, ***et al.***, 2016 (Shaver et al., 2016).

STA Discovery Pipeline

Exon-level RPKM expression values for all samples in a given TCGA study were combined and divided into individual protein-coding loci based on a Gencode V19 annotation file (hg19.wgEncodeGencodeBasicV19), filtered by HGNC protein-coding gene categorization and excluding alternate haplotypes (Harrow et al., 2012). To maximize STA sensitivity and accuracy for the present analysis, we filtered out genes with four or fewer exons as well as those with an expression level of 10 RPKM or less across all samples within a study.

The resulting matrices of exon-level expression values for each gene were processed through the STA algorithm by a custom R script (Shaver et

al., 2016) to derive an STA score for each gene-sample combination. For this study, we flagged gene-sample combinations with STA scores of 2 or greater, indicating 2 or greater standard deviations above the population mean, for further sequence-level analysis. In addition, 14 genes were investigated by default in a subset of studies to evaluate the ability of STA to detect previously identified rearrangements.

For samples with transcripts passing the STA score threshold, the corresponding RNA-seq BAM file was downloaded and STA-flagged transcript regions were converted to SAM format using samtools (Li et al., 2009). Reads aligned within these regions were analyzed for the presence of the ZF:Z tag, which indicates a “fusion alignment” by MapSplice, the RNA-seq aligner used in the TCGA RNA-seq v2 workflow (Wang et al., 2010). All gene-sample combinations passing the STA score threshold were annotated with the location of any potential rearrangement partners and the number of supporting ZF:Z-tagged reads. Alignment to multiple locations within a putative rearrangement partner was noted for downstream analysis. The nucleotide position of each RNA breakpoint was annotated for the presence and coding frame/UTR status of any exon boundaries at that location, based on an hg19 refGene.txt file packaged with Integrative Genomics Viewer (Robinson et al., 2011).

To prioritize targets for validation at the whole-genome sequencing level, we filtered STA-flagged transcripts to include those with two or more supporting MapSplice fusion alignments and at least one of the two RNA

breakpoints occurring at an exon boundary. Potential intragenic rearrangements were beyond the scope of this study and were excluded. Transcripts were additionally filtered to exclude rearrangements with the mitochondrial genome and human leukocyte antigen (HLA) genes, which presented difficulty for downstream DNA-level analysis. Because of the lack of strand specificity in the original RNA-seq alignment, putative rearrangements were also assessed for the possibility of alignment to coding loci and exon boundaries on the opposite strand.

For transcripts passing these filters, corresponding portions of the WGS files were downloaded and reads were analyzed nearby the putative rearrangement partners. For rearrangements occurring in protein-coding genes, we assessed the GENCODE-annotated gene locus coordinates; for non-coding rearrangement partners, we assessed a region of DNA encompassing the RNA breakpoint ± 100 kilobases. Reads within these DNA regions were analyzed for the presence of discordant read pairs that mapped to both of the putative RNA partners, including an additional adjacent 50 kilobase “buffer” region to account for potential upstream DNA breakpoints. For any discordant reads successfully identified, all reads within 5 kilobases featuring soft-clipped regions were trimmed to exclude bases with Phred scores less than 30 using Flexbar v2.5 (Dodt et al., 2012) and then realigned to the putative rearrangement partner regions. For this alignment step, we used regions ± 5 kilobases of the discordant read coordinates as the reference genome and ran YAHA version 0.1.82 with

the options –BP 1 –X 15 –M 20(Faust & Hall, 2012). Any resulting alignments with segments mapping to a single region in each rearrangement partner, as determined by the YAHA output tags YP:i:2 and YS:i:0, were considered breakpoint-spanning reads. A number of "false positive" putative rearrangements arising from gene homology issues were filtered out by restricting alignment to these specific output tags (data not shown).

RNA and DNA breakpoints were analyzed to ensure logical orientation based on the assumed strand of the rearrangement partners. Rearrangements with at least two ZF:Z tagged RNA-seq reads, two discordant WGS read pairs, and one breakpoint-spanning WGS read were considered DNA-validated. A subset of rearrangements, including those with DNA breakpoints upstream of the affected protein-coding gene, were manually curated and validated. KANSL1-ARL17B and TFG-GPR128 fusions, both of which have been documented as structural variants in healthy individuals (Chase et al., 2010; Stefansson et al., 2005), passed the STA score threshold in numerous samples and were validated at the DNA level but were removed from the final list.

STA score distribution for previous TCGA fusion datasets

Fusion transcript partners from Table S1 from Yoshihara *et al*. (Yoshihara et al., 2014) and Supplementary Data 2 from Stransky *et al*. (Stransky et al., 2014) were linked to STA scores calculated for all TCGA

studies analyzed. Any previously identified fusion transcripts without a corresponding STA score due to differences in samples processed were excluded from density plotting and statistical analysis. When STA scores were available for both fusion partners, the higher score was used to represent the fusion transcript.

Statistical Analysis

All analyses and graphical representations were performed using R v.3.2.0-3.3.2 (R Core Team, 2016). Metabolic rates were plotted as the mean ± SD of three independent experiments, and p-values were calculated using paired, two-sided Student's t tests of the pre-normalized alamarBlue:cell count ratios for each isogenic clone versus CAL51. Cell size data were calculated as the mean ± SD of FSC-A values for 10,000 single-cell, 2N flow cytometry events, and p-values were calculated using unpaired, two-sided Student's t tests of the values for each isogenic clone versus CAL51.

The correlation of differentially expressed genes and changes in chromosomal copy number was performed by collation of RNA-seq count data for the initial isogenic cell line panel (as described above), followed by normalized log transformation using DESeq2 v1.14.1 (Love et al., 2014) with default parameters. Differences in normalized log values and log-scale chromosomal copy number values were generated and plotted as described. A linear model was fit to the data using ggplot2 (Wickham,

2009) default parameters, and the Pearson coefficient was calculated using the base R function.

The significance of the difference in frequency of aneuploidy in the validation isogenic cell line panel was assessed by combining all chromosomal counts from wild-type clonal populations and determining outlier boundaries as determined by smaller than Q1 – c * IQD or larger than Q3 * c *IQD, where Q1 and Q3 are the lower and upper quartiles, IQD is the interquartile range, and c=1.5. We then calculated the proportion of abnormal (outlier) cells in the 20 chromosomal counts from each individual clone and conducted ordinary least squares linear regression using the proportions of abnormal cells across each genotype. Significance is reported as p-values for Wald tests of each coefficient being equal to zero as well the p-value from an analysis of variance testing if the proportion of abnormal cells in a clone is significantly associated with the genotype.

Post-siRNA viability ratios were plotted as the mean ± SEM of four independent experiments, and p-values were calculated using the Wilcoxon rank-sum test with Bonferroni multiple comparison adjustment. PD173074 IC_{50} values were plotted as the mean ± SEM of three independent experiments. Ba/F3 IL-3 withdrawal data were plotted as the mean ± SD of two conditions with three independent experiments across three days, and p-values were calculated using the restricted maximum likelihood (REML)-based mixed effects model. For comparison of STA calls to Yoshihara *et al.* and Stransky *et al.*, the distribution of STA scores for all genes and

samples queried were compared to those for previously identified fusion transcripts using two-sided, two-sample Kolmogorov-Smirnov tests.

Statistical analysis of aneuploidy frequency in the validation set of isogenic clones is being conducted with Yu Shyr, Director of the Center for Quantitative Sciences; significance calculations will be included in the defense and final version of this document.

CHAPTER 3

MUTANT p53 IN AN ISOGENIC TRIPLE-NEGATIVE BREAST CANCER MODEL: EVOLUTION OF DISTINCT GAIN-OF-FUNCTION PHENOTYPES

Mutant p53 gain of function (GOF) has been extensively documented. Despite numerous ongoing preclinical and clinical efforts to therapeutically target p53 missense mutants, our understanding of the GOF phenomenon is confounded by variability in the phenotypic reproducibility and mechanisms – not only between individual missense mutants, but also the same missense mutant expressed in different contexts. The effect of differing biological states in explaining these discrepancies and the mechanistic link between the numerous reported GOF phenotypes remain unclear. To begin to address these research questions in a controlled setting and to potentially identify new therapeutic targets for triple-negative breast cancer (TNBC), in which *TP53* mutations constitute the only alteration to occur in a majority of cases, we leveraged CRISPR/Cas methodology to generate an TNBC cell line isogenic model system expressing p53 wild-type, null, and mutant alleles from the endogenous promoter. Isogenic populations expressing the R175H and R273H hotspot mutants exhibited previously published GOF phenotypes, including aneuploidy, elevated metabolism, and enhanced growth in xenograft, but

with the acquisition of distinct phenotypes occurring due to clonal variability. The missense mutant populations also exhibited significantly different gene expression profiles from both wild-type and null. These phenotypic and transcriptional changes occurred in the absence of constitutive mutant stabilization, providing evidence that mutant p53 can confer GOF phenotypes even while under negative regulation by E3 ligases such as MDM2. Our results demonstrate the difficulty of reproducing GOF phenotypes and mechanisms, even when leveraging endogenous alleles in an isogenic setting, and emphasize the challenges of clonal heterogeneity for both basic scientists and clinicians.

Introduction

Mutant p53 gain of function (GOF) has been documented in hundreds of studies since the term was first coined two and a half decades ago (Dittmer et al., 1993; Muller & Vousden, 2014). Hotspot missense mutant p53 proteins confer a variety of oncogenic phenotypes, including enhanced proliferation and survival; increased migration and invasion potential; enhanced anchorage-independent growth and colony formation; polyploidy, and larger, more metastatic growth in xenograft (Brosh & Rotter, 2009; Muller & Vousden, 2013; Oren & Rotter, 2010; Soussi & Wiman, 2015). The mechanisms reported to underlie these phenotypes are similarly varied, and include sequestration of the p53 family members p63 and p73; a dominant-negative effect over wild-type p53; altered

transcriptional activity, and aberrant protein interactions, all of which lead to inhibition or enhancement of cancer-related cellular pathways and processes (Brosh & Rotter, 2009; Muller & Vousden, 2014).

Despite consistency in the documentation of mutant p53 GOF at large, there is discrepancy in the phenotypic and mechanistic characteristics of not only distinct hotspot mutants, but also the same missense mutant expressed in different cellular or tissue contexts (Muller & Vousden, 2014). Although a number of GOF reports have evaluated the knockdown or destabilization of endogenous hotspot mutants (Alexandrova et al., 2015; Braicu et al., 2013; Hui et al., 2006; Li et al., 2011a; Lim et al., 2009; Muller & Vousden, 2014; Rodriguez et al., 2012; Vakifahmetoglu-Norberg et al., 2013; Zhu et al., 2011, 2013), a majority of GOF studies involve exogenous overexpression of mutant p53 protein, which may produce an artificial readout of mutant p53 activity, especially when evaluating transcriptional or protein interaction capacity. Most murine mouse models (Donehower & Lozano, 2009) and a limited number of *in vitro* studies (Sur et al., 2009) have leveraged isogenic evaluation of mutant p53 in the endogenous allele compared to wild-type and null, but despite their considerable advantages, *in vivo* models lack the temporal resolution and capacity for experimental manipulation afforded by cell lines.

Better understanding of mutant p53 GOF is of distinct importance for the treatment of triple-negative breast cancer (TNBC), a heterogeneous category of cancers in which *TP53* mutations constitute the only alteration

to occur in a majority of tumors (Banerji et al., 2012; Bianchini et al., 2016; Cancer Genome Atlas, 2012b; Curtis et al., 2012; Shah et al., 2012; Walerych et al., 2012). A significant fraction of TNBC cases lack alterations in any clinically actionable proteins and pathways (Bianchini et al., 2016; Shah et al., 2012); our goal is to expand our mechanistic understanding of mutant p53 GOF in the context of TNBC with the potential of elucidating new targets for therapeutic development.

Accordingly, we herein report the CRISPR/Cas-mediated creation of an isogenic TNBC cell line panel expressing wild-type, null, and missense mutant p53 from the endogenous allele in identical cell line backgrounds. Missense mutant p53-expressing cells recapitulate a variety of known GOF phenotypes in comparison to wild-type and null isogenics, but with a considerable amount of variability between biological replicates. Besides the mechanistic basis of mutant p53 GOF, our findings have relevance to further understanding of mutant protein regulation and clonal dynamics during the process of tumorigenesis.

Results

CRISPR/Cas-Mediated Editing of Endogenous p53 Allele

To model mutant p53 GOF in TNBC, we leveraged CRISPR/Cas-mediated genome editing to compare wild-type, missense mutant, and frameshift null p53 alleles in an identical cell line background. Given that

the restoration of wild-type p53 activity in p53 mutant or null cells can induce rapid cell cycle arrest or apoptosis (Kastan, 2007; Martins et al., 2006; Ventura et al., 2007; Wang et al., 2011b; Xue et al., 2007), we selected CAL51, a p53 wild-type cell line isolated from a malignant TNBC pleural effusion, as our parental population for mutant clonal derivation (Gioanni et al., 1990). CAL51 has the additional advantage of a unique, near-diploid karyotype, which facilitates modeling of the biallelic states found endogenously in tumors, such as loss of heterozygosity (LOH) (Baker et al., 1989).

In order to prioritize p53 missense mutants for modeling, we evaluated *TP53* mutational frequency in TNBC datasets from The Cancer Genome Atlas and the Aparicio group (Cancer Genome Atlas, 2012b; Shah et al., 2012). 'Hotspot' missense mutants in TNBC largely overlap with those in human cancer at large, with R175H, Y220C, R273H, I195T, and R248W representing the five most frequent amino acid changes (Fig. 3). For our initial homology-directed repair templates to knock in by leveraging CRISPR/Cas nuclease activity (Ran et al., 2013), we chose R175H and R273H, the two most common p53 missense mutants and the #4 and #8 most frequent missense mutations among all genes, respectively, in the TCGA Pan-Cancer analysis (Kandoth et al., 2013).

After derivation of a p53 wild-type, single-cell clone from the heterogeneous CAL51 population to ensure an isogenic background, we transfected this parental clone (WT A1) with plasmids encoding Cas9 and

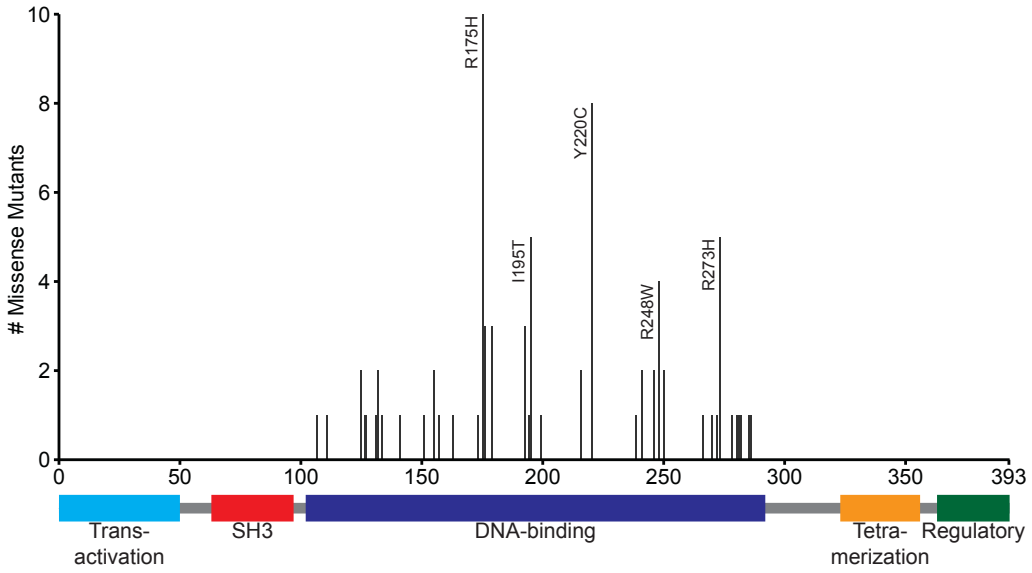


Figure 3: p53 missense mutants in triple-negative breast cancer include recurrent cancer-wide hotspots.
Mutation annotation date of TNBC cases from TCGA and Shah *et al*., 2012 were collated and assessed for the frequency of individual p53 missense mutants (n=49 total). The resulting counts were plotted as a histogram, where the x axis depicts amino acid number and associated domains for full-length p53 protein.

TP53-targeting guide RNAs along with homology-directed repair templates containing the clinically observed R175H and R273H missense mutations (Table 2 and Table 3) (Ran et al., 2013). The resulting single-cell clones exhibited a wide range of heterozygous and homozygous p53 genotypes, including wild-type, missense mutant, and frameshift null alleles. The ability of these clones to transactivate canonical p53 gene targets in response to doxorubicin treatment was largely in accordance with their expected function based on genotype (Fig. 4); notably, heterozygous wild-type/R175H and wild-type/R273H clones were capable of promoting p21 and MDM2 protein expression at a level similar to homozygous wild-type clones. This finding has relevance to the ongoing debate over the biochemical circumstances and stoichiometric ratio necessary for endogenous mutant p53 to exert a dominant-negative effect over wild-type that was originally noted in overexpression models (Chan et al., 2004; Milner & Medcalf, 1991; Sabapathy, 2015; Sun et al., 1993).

After screening, we were successful in identifying two clones for each hotspot mutant with a heterozygous missense/frameshift null genotype, mimicking the loss of heterozygosity commonly observed across cancer (Alexandrova et al., 2017; Baker et al., 1989) (Table 5). As controls for p53 loss of function, we included two additional clones harboring homozygous frameshift mutations near the frequent *TP53* R196X nonsense mutation site (Fig. 5A, Table 5).

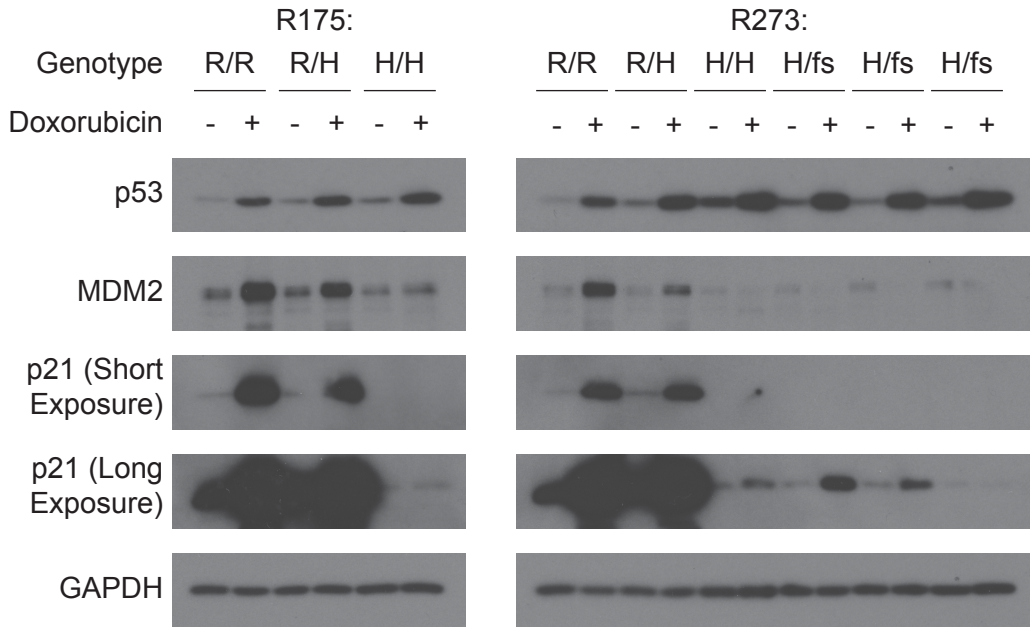


Figure 4: CRISPR/Cas-generated missense mutant p53 proteins are deficient in transactivating canonical target genes.
The diploid, p53 wild-type TNBC cell line CAL51 was targeted with a plasmid encoding *S. pyogenes*-derived Cas9 protein and homology-directed repair templates encoding the R175H or R273H missense base pair substitutions. Resulting single cell colonies were screened by restriction digest and Sanger sequencing for successful incorporation of the base pair substitution, and a variety of clones expressing combinations of the wild-type allele (R), mutant allele (H), and frameshift null allele (fs) were identified. 'R/H' notation represents the status of each allele. Immunoblotting for the p53 protein and its canonical targets MDM2 and p21 was conducted on lysates from cells treated with PBS (-) or 0.2 μM doxorubicin (+) for 24 h. GAPDH was included as a loading control.

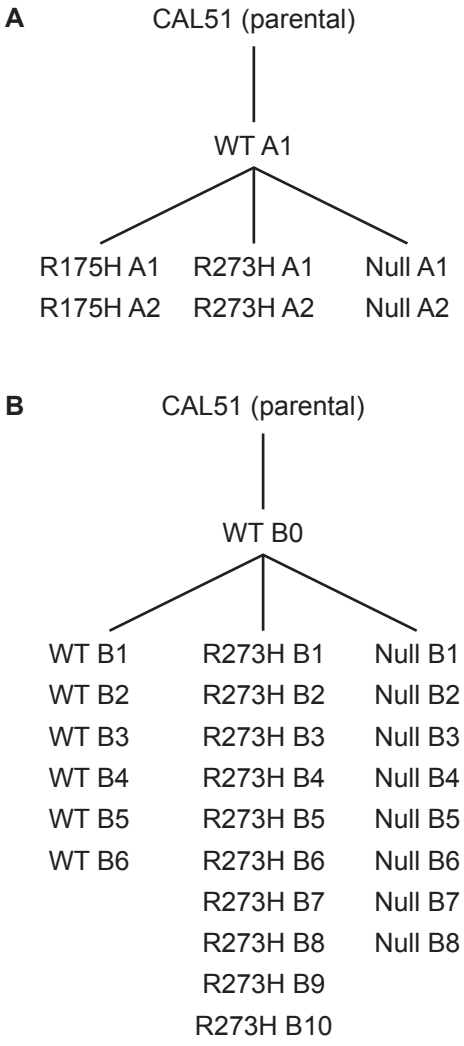


Figure 5: Isogenic cell line panels derived from single-cell CAL51 clones.
A-B, schematics depicting single-cell clonal derivation process for the initial (A) and validation (B) isogenic cell line panels. Genotypes of each clone are as indicated and distinct clone numbers represent independently derived biological replicates. WT: homozygous wild-type, R273H: heterozygous R273H/frameshift null, Null: homozygous frameshift null.

Clone ID	Protein present	PAM site/gRNA mutation	Missense mutation	Heterozygous	Second allele, if present
WT A1	+	+	-	-	
R175H A1	+	+	+	+	1 nt insertion
R175H A2	+	+	+	+	1 nt insertion
R273H A1	+	+	+	+	5 nt deletion
R273H A2	+	+	+	+	1 nt deletion
Null A1	-				
Null A2	-				
WT B1	+	+	-	-	
WT B2	+	+	-	-	
WT B3	+	+	-	-	
WT B4	+	+	-	-	
WT B5	+	+	-	-	
WT B6	+	+	-	-	
R273H B1	+	+	+	+	1 nt deletion
R273H B2	+	+	+	+	1 nt deletion
R273H B3	+	+	+	+	2 nt deletion
R273H B4	+	+	+	+	4 nt deletion
R273H B5	+	+	+	+	1 nt deletion
R273H B6	+	+	+	+	1 nt insertion
R273H B7	+	+	+	+	5 nt deletion
R273H B8	+	+	+	+	1 nt insertion
R273H B9	+	+	+	+	1 nt deletion
R273H B10	+	+	+	+	1 nt deletion
Null B1	-				
Null B2	-				
Null B3	-				
Null B4	-				
Null B5	-				
Null B6	-				
Null B7	-				
Null B8	-				

Table 5: Isogenic cell line panel genotypes by clone.
The indicated clones are annotated with whether p53 protein was detectable by immunoblot, whether the conservative PAM site mutation (R175H) or guide RNA complementarity mutations (R273H) were detected, whether the associated missense mutant was detected, whether heterozygous genotype was exhibited by mixed traces on Sanger sequencing, and, if heterozygosity was present, the genotype of the second allele as determined by allele cloning.

Missense Mutant-Expressing Isogenic Populations Exhibit GOF Phenotypes

We assessed our initial isogenic cell line panel for the acquisition of published mutant p53 GOF phenotypes (Brosh & Rotter, 2009; Muller & Vousden, 2014; Oren & Rotter, 2010; Soussi & Wiman, 2015). After more than six months in culture, one of two R175H (R175H A1) and both R273H missense mutant clones exhibited significantly elevated metabolic rates compared to their wild-type and null isogenic counterparts (Fig. 6). Assessed by flow cytometry, the same three missense mutant clones displayed significant and notable increases in cell size (Fig. 7).

Comparing growth in xenograft among the isogenic populations, we observed larger tumor growth in three out of four missense mutants compared to both wild-type and null isogenics. In both the athymic nude and NSG immunodeficient mouse models, one of two R175H (R175H A2) and both R273H missense mutant clones formed larger tumors on average (Fig. 8), although these differences lacked statistical significance due to limited sample number. Notably, the R175H clone that exhibited increased metabolic rate and cell size (R175H A1) formed tumors of equivalent size to the parental CAL51 line.

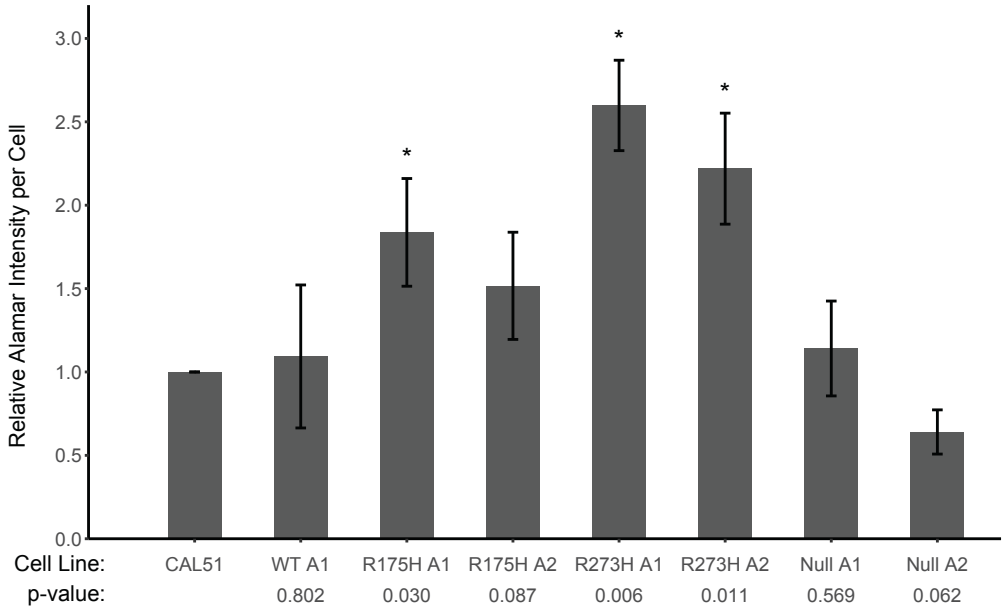


Figure 6: Missense mutant p53-expressing cells exhibit elevated metabolic reduction rates compared to wild-type and null isogenic populations.
The indicated cell line (CAL51) or isogenic clones were plated in a 96-well format. Metabolic reduction rate was assessed at 48 h using alamarBlue fluorescence, followed by washing with PBS, fixation for 10 m with 4% PFA, and staining with DAPI. Nuclei in each well were imaged at low magnification using fluorescence microscopy, then counted using ImageJ. AlamarBlue values were divided by nuclei counts to yield per-cell metabolic rate estimates, then normalized to CAL51. Values are depicted as mean +/- standard deviation of three independent experiments. P-values represent paired, two-sided Student's t tests conducted on the pre-normalized alamar:cell count ratios for each isogenic clone compared to CAL51. Clones with p-values < 0.05 are indicated with asterisks.

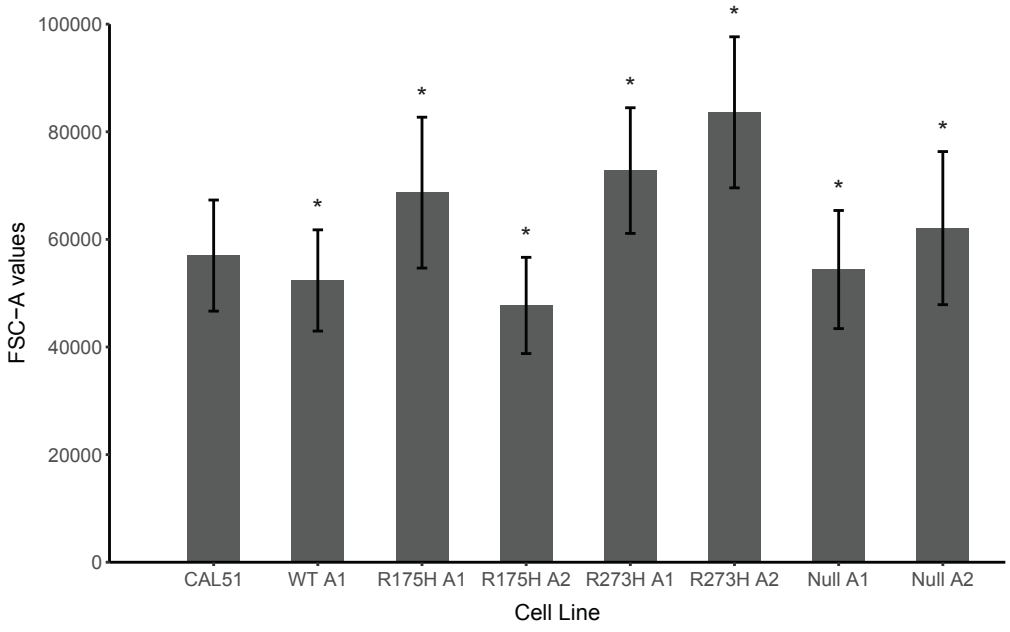


Figure 7: Missense mutant p53-expressing cells are larger in size compared to wild-type and null isogenic populations.
The indicated cell line (CAL51) or isogenic clones were grown to similar levels of sub-confluency, then trypsinized and analyzed by flow cytometry. Values depicted are the mean of forward scatter pulse area (FSC-A) for 10,000 events +/- standard deviation. Asterisks represent p-values < $1x10^{-12}$ for unpaired, two-sided Student's t tests conducted on FSC-A values for each isogenic clone compared to CAL51.

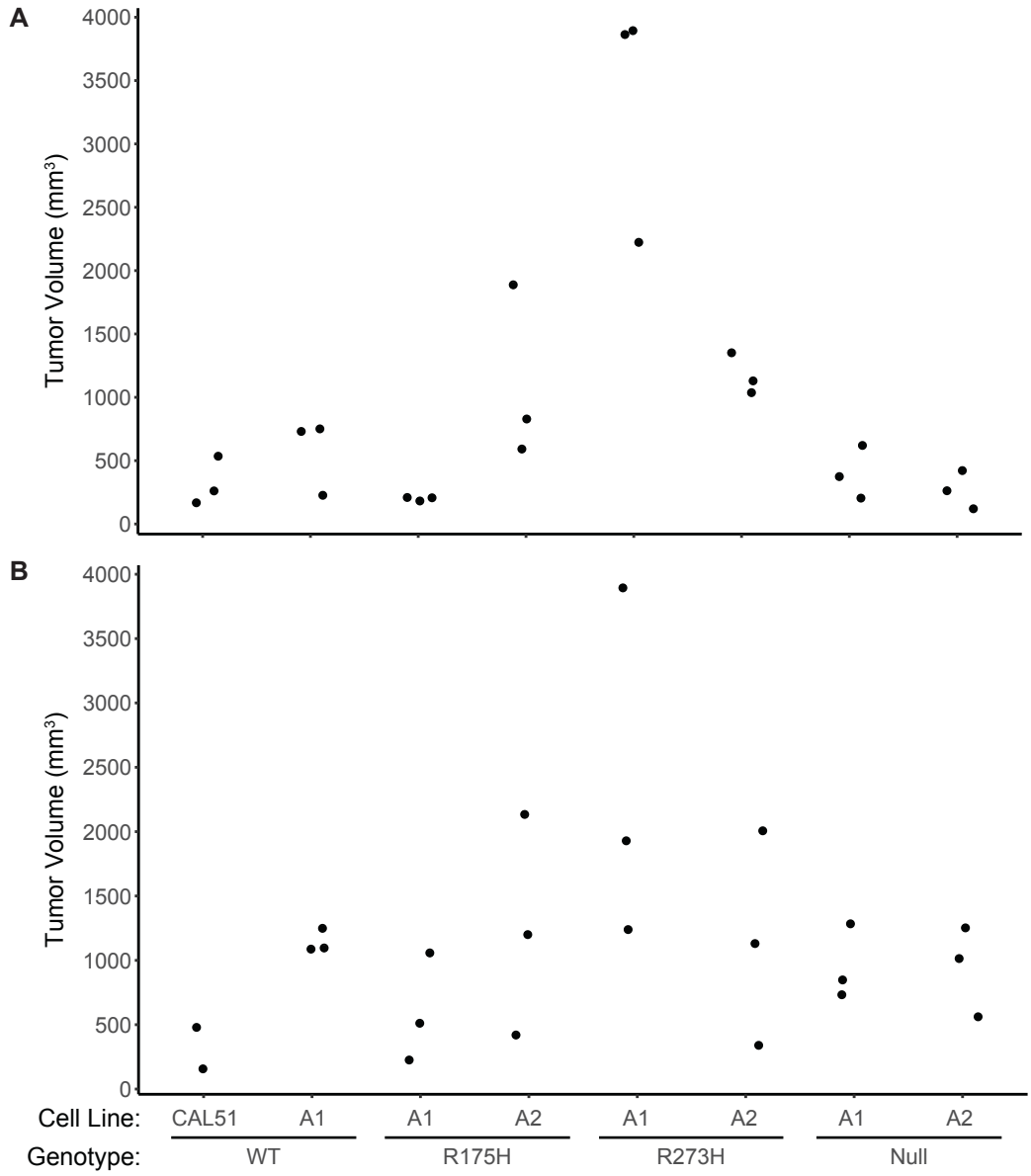


Figure 8: Missense mutant p53 expression is associated with higher xenograft tumor size compared to wild-type and null isogenic populations.
A-B, Five million cells of the indicated cell line (CAL51) or isogenic clone were injected into the flanks of immune-deficient mice, A. Cells were dissolved in 200 µL of PBS, injected into NSG mice, and grown for 40 days. B, Cells were dissolved in 100 µL of PBS and 100 µL of standard formulation Matrigel, injected into athymic nude mice, and grown for 25 days. Data points represent individual tumors; volume was estimated as (length * width^2)/2.

Missense Mutant-Expressing Isogenic Populations Exhibit a Higher Frequency of Aneuploidy and Unique Changes in Gene Expression

To investigate potential mechanisms contributing to the acquisition of these GOF phenotypes, we performed metaphase spreads to assess the ploidy of our initial isogenic cell line panel. Strikingly, one of two R175H (R175H A1) and both R273H missense mutant clones contained more than twice the number of chromosomes of CAL51 and the parental WT A1 clone. In contrast, the chromosomal counts of both null clones were similar to wild-type (Fig. 9).

We conducted RNA-seq and differential gene expression analysis on the initial isogenics (Love et al., 2014). Pair-wise analysis for differentially expressed genes between genotypes indicated a large number of changes associated with p53 loss of function (null versus wild-type and missense mutant versus wild-type), as expected. Interestingly, however, the R175H missense mutants exhibited over 140 differentially expressed genes compared to R273H (Fig. 10A-B). In addition, while there were minimal gene expression changes between the R175H and null clones (Fig. 10A), over 340 genes were significantly differentially expressed between R273H and null (Fig. 10B), implying that the mechanism underlying gene expression changes in these missense mutants extended beyond mere p53 loss of function. Further, when comparing R273H missense mutants to wild-type, null, and R175H, 71 genes were significantly differentially

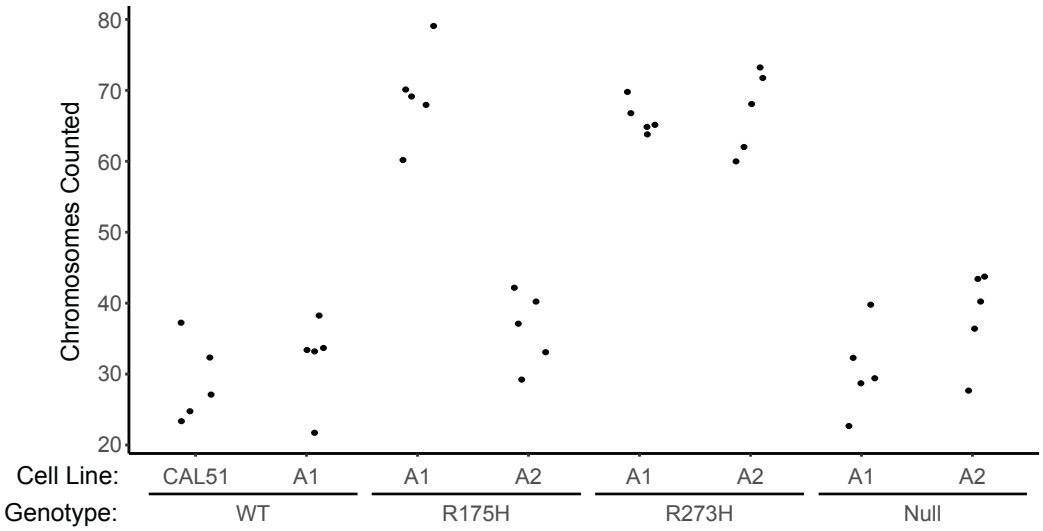


Figure 9: Isogenic cell line panel exhibits p53 missense mutant-specific aneuploidy compared to parental cells.
Cultures of the indicated cell line (CAL51) or isogenic clones were arrested in metaphase by addition of 0.5 µg/mL colcemid for 2.5 h, followed by trypsinization, treatment with 0.075 M KCl, and fixation with methanol and glacial acetic acid. The resulting cell suspension was dropped onto slides and stained with DAPI, then imaged under high magnification fluorescent microscopy. Metaphase spreads were identified and individual chromosomes were counted manually. Data points (n=5 per line) represent quantifications of individual nuclei.

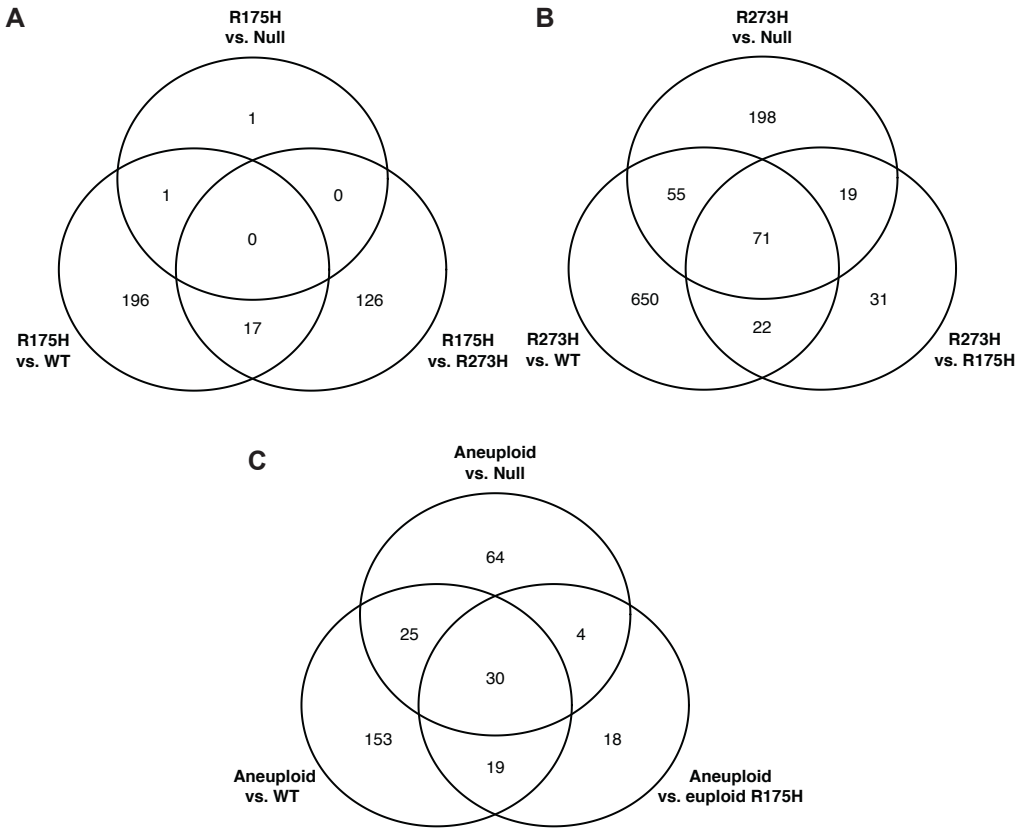


Figure 10: R273H missense mutant-expressing cell lines and aneuploid cell lines display unique gene expression patterns in comparison to other clones.
A-C, Differential gene expression analysis was conducted on RNA-seq data from isogenic cell lines using DESeq2. Pairwise comparisons were conducted between each of the indicated cell line categories and the resulting significantly different genes (adjusted p-value < 0.1) were assessed for overlaps as depicted by set diagrams. WT: WT A1. R175H: R175H A1 and A2. R273H: R273H A1 and A2. Null: Null A1 and A2. Aneuploid: R175H A1 and R273H A1 and A2.

expressed in comparison to all three groups, representing a unique, missense mutant-specific pattern of gene expression changes (Fig. 10B).

To assess whether the commonality in R273H gene expression status was due to their aneuploid status, we conducted an additional pairwise analysis of all three aneuploid clones (R175H A1, R273H A1, and R273H A2) versus wild-type, null, and the non-aneuploid missense clone (R175H A2). We identified a unique expression pattern similar to that of the R273H clones (Fig. 10C), suggesting that missense mutant-specific changes in gene expression can in part be attributed to the acquisition of aneuploidy, even when driven by different hotspot mutant proteins. The 30 genes whose expression was unique to the aneuploid clones included a number of cancer-related pathways and proteins (Fig. 11). In particular, we observed significantly elevated *MYC* expression in the aneuploid clones, along with decreased expression of the tumor suppressor *NKX3-1* and Wnt inhibitory genes *FRZB*, *WIF1*, and *RNF43*.

To examine the specific chromosomal abnormalities present in the missense mutant isogenics, we conducted whole-genome cytogenomic hybridization analysis using the high-resolution Affymetrix CytoScan HD array. While the parental WT A1 and both null isogenics exhibited largely normal chromosomal content, all four missense mutants – including, to a limited extent, the non-chromosomal count altered R175H A2 – displayed copy number gain and loss across multiple chromosomes, with both R273H clonal populations experiencing notable copy number changes for nearly

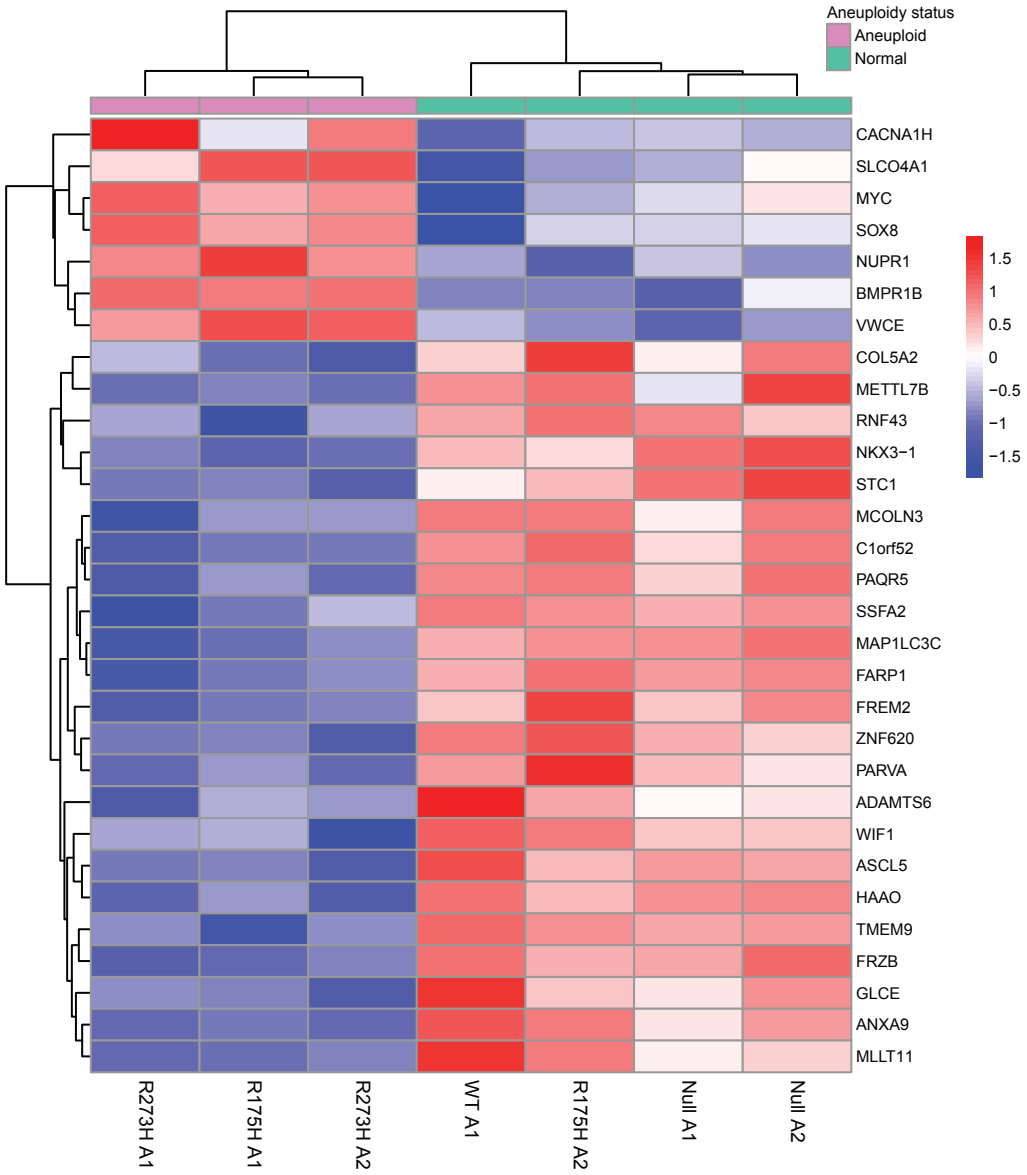


Figure 11: Aneuploid missense mutant p53-expressing cell lines display unique patterns of increased and decreased gene expression in comparison to other isogenic populations.

Differential gene expression analysis was conducted on RNA-seq data from isogenic cell lines, as in Fig. X. Thirty genes were significantly different (adjusted p-value < 0.1) between aneuploid and non-aneuploid isogenic clones, as indicated by the legend at the top. The figure displays a row- and column-clustered heat map representation of row-normalized gene expression estimates for each gene, listed at the right. Clone IDs are listed under each column. Color scale included for reference.

every chromosome (Fig. 12). For each missense mutant, copy number changes occurred across whole chromosomes, implying that the alterations occurred as a result of nondisjunction rather than focal rearrangement.

Given the somewhat consistent pattern of chromosomal copy number changes in the aneuploid clonal populations, we evaluated whether the gene expression changes seen in Fig. 10C and Fig. 11 could be attributed simply to copy number gain and loss at the chromosomal level. However, pairwise comparisons of gene expression changes versus changes in copy number of the associated chromosome did not reveal any significant correlation (Fig. 13), suggesting alterations in transcription by other means, including indirect effects of the gain and loss of regulatory factors at the copy number level or more complex transcriptional perturbation by mutant p53.

In order to evaluate the reproducibility of mutant-specific aneuploidy and control for clonal variation and sampling bias, we engineered a second, larger set of isogenic clones. Given the acquisition of aneuploidy in both initial R273H populations, in addition to their unique gene expression signature and more prominent chromosomal copy-number changes, we focused our analysis on new wild-type, null, and heterozygous R273H/frameshift clones (n=6, 8, and 10, respectively) (Fig. 5B, Table 5). After several months in culture, we conducted metaphase spreads and chromosomal counts on all 24 new validation clones. Although there was considerable heterogeneity within clonal populations compared to the

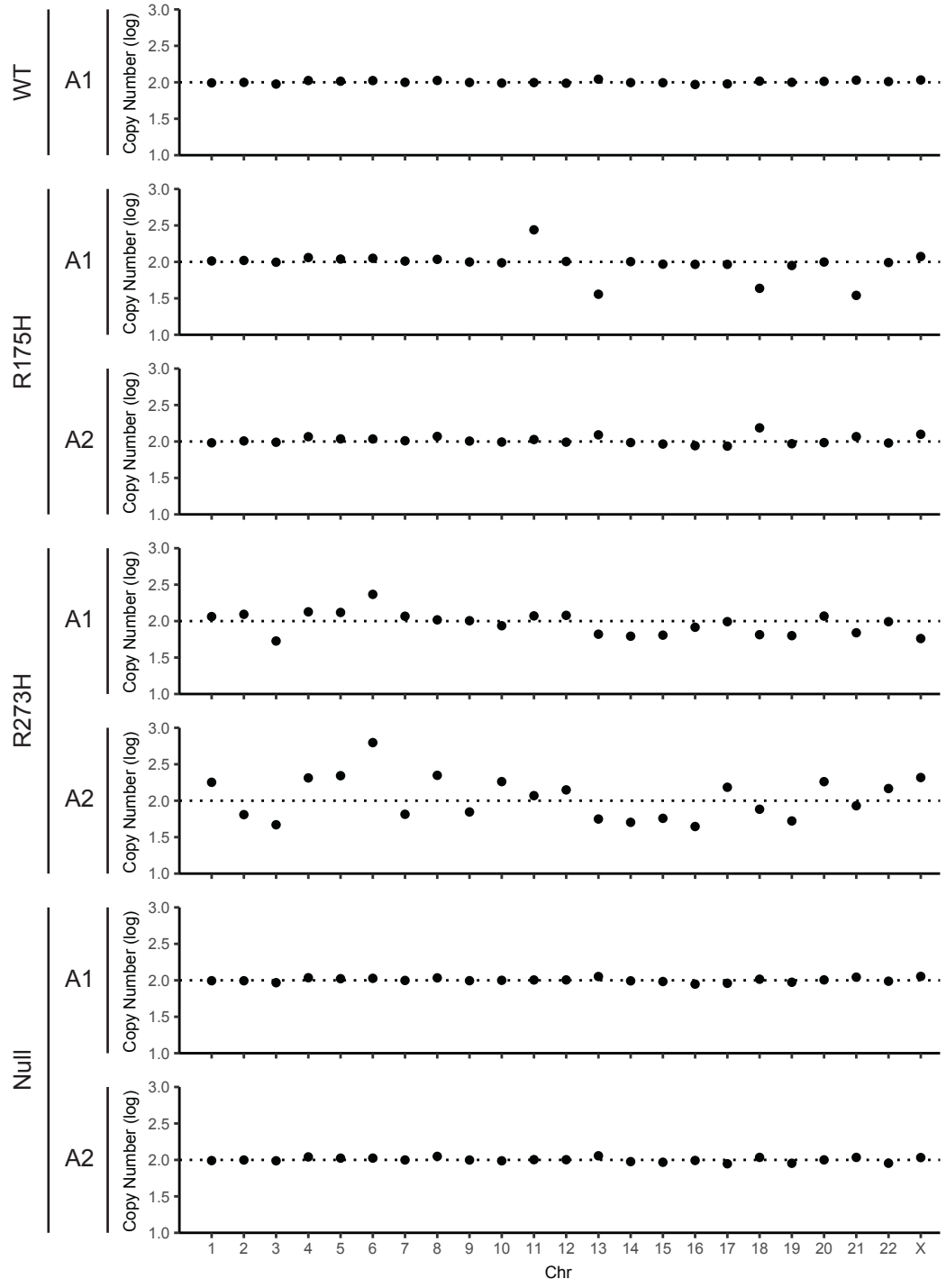


Figure 12: p53 missense mutant cell lines exhibit chromosomal copy number loss and gain compared to wild-type and null isogenic populations.
DNA from the indicated isogenic clones was isolated and analyzed on the Affymetrix CytoScan HD Array using Affymetrix Chromosome Analysis Suite Software. Median copy number state values for each chromosome were plotted as log-scale values, where 2 (dotted line) represents normal copy number by *in silico* comparative analysis. Figure generated in collaboration with Ferrin Wheeler.

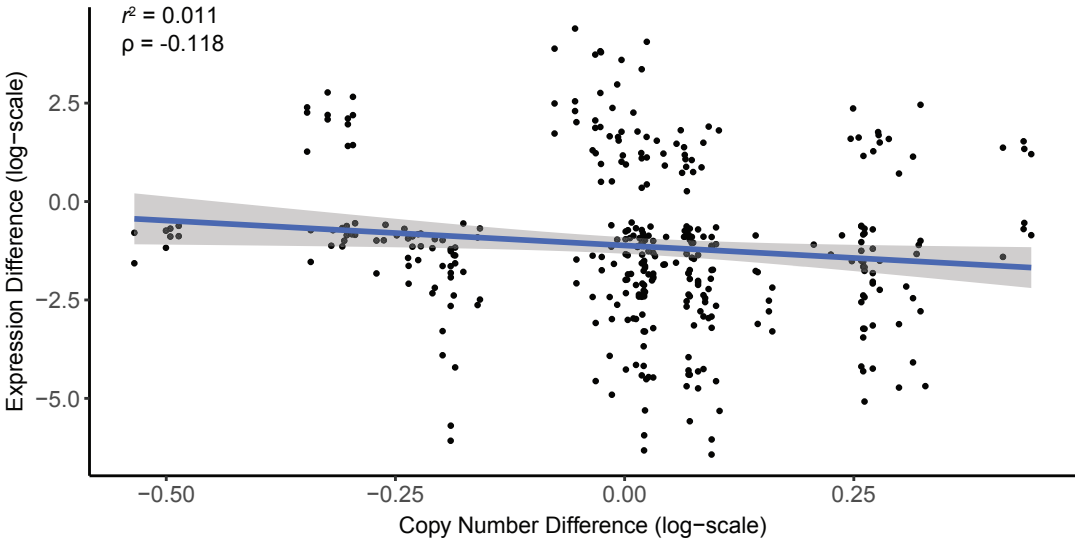


Figure 13: Aneuploid-specific gene expression changes do not correlate with changes in chromosomal copy number.
For each of the uniquely differentially expressed genes in Fig. X, a pairwise comparison was conducted between each aneuploid isogenic line from the initial set (R175H A1, R273H A1, and R273H A2) and each non-aneuploid isogenic line (WT A1, R175H A2, Null A1, Null A2). The y-axis value represents the difference in regularized log-transformed expression estimates (DESeq2) for each comparison, and the x-axis value represents the log-scale copy number difference for the chromosome containing the gene locus. The blue line represents a linear model fit to the data with a shaded 95% confidence interval; the adjusted r^2 value is depicted at top left along with the Pearson correlation coefficient (ρ) for the data.

initial isogenic set, likely owing to shorter time in culture, the extent of aneuploidy differed notably: the proportion of aneuploid cells in a clone was significantly associated with the overall genotype (p=0.008), and was significantly higher for R273H than WT (p=0.003) (Fig. 14). These data are in keeping with previous reports of mutant GOF; in addition to defects in DNA damage response resulting from p53 loss of function alone, missense mutant p53 protein has been shown to disrupt the mitotic spindle checkpoint, thereby promoting aneuploidy (Gualberto et al., 1998; Hanel & Moll, 2012; Oren & Rotter, 2010).

Knock-In Mutant p53 Protein Lacks Constitutive Stabilization and is Insensitive to Statin-Induced Degradation

Multiple publications have identified constitutive stabilization of mutant p53 by stress signals or inhibition of negative regulators as a prerequisite for GOF (Alexandrova et al., 2017; Muller & Vousden, 2014; Suh et al., 2011; Terzian et al., 2008). Surprisingly, however, our initial isogenic cell line panel displayed p53 protein accumulation in response to doxorubicin treatment for the wild-type, R175H, and R273H clonal populations, even after over twelve months in culture (Fig. 15). A similar level of protein accumulation occurred after treatment with the MDM2-p53 interaction inhibitor (-)-Nutlin-3 (Davis & Johnston, 2011; Vassilev et al., 2004), implying that mutant p53 protein in our CAL51 isogenic populations is subject to MDM2-mediated regulation, despite a lack of MDM2

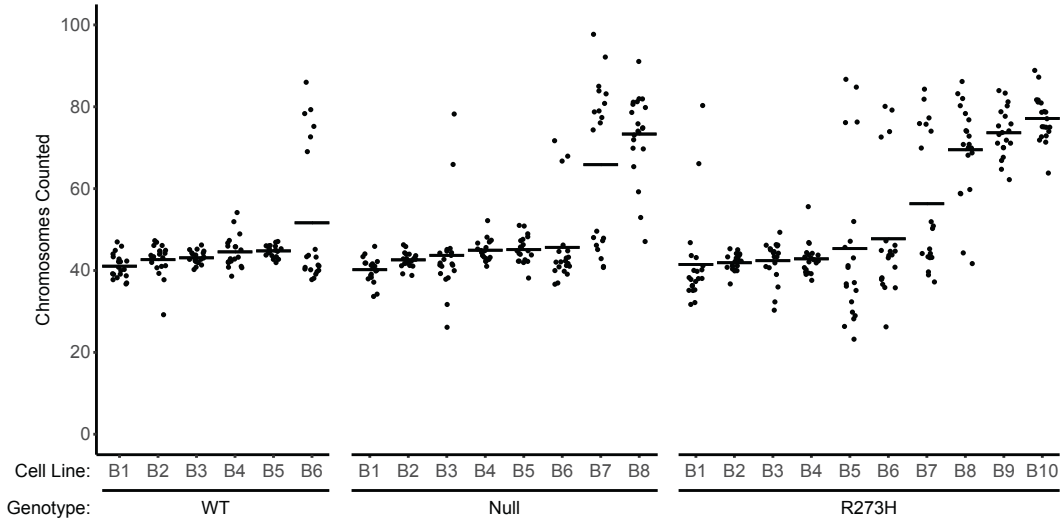


Figure 14: Missense mutant p53 isogenic cell lines exhibit a higher rate of aneuploidy than wild-type.

Cultures of the indicated isogenic clones were arrested in metaphase by addition of 0.5 µg/mL colcemid for 1.5 h, followed by trypsinization, treatment with 0.075 M KCl, and fixation with methanol and glacial acetic acid. The resulting cell suspension was dropped onto slides and stained with DAPI, then imaged under high magnification fluorescent microscopy. Metaphase spreads were identified and individual chromosomes were counted manually. Data points (n=20 per line) represent quantifications of individual nuclei, and horizontal bars depict the mean chromosome count. To determine significance, an aneuploid count threshold was defined by grouping chromosomal counts for all WT clones (n=120) and determining box plot outliers greater than Q3 * 1.5 * IQD, where Q3 is the upper quartile and IQD is the interquartile range. We calculated the proportion of outlier cells in the counts from each individual clone and conducted ordinary least squares regression using the proportions of abnormal cells across each genotype. The proportion of abnormal cells in a clone was significantly higher for R273H than WT (p=0.003), and the proportion of abnormal cells in a clone was significantly associated with the genotype by analysis of variance (p=0.008). Statistical interpretation was conducted in collaboration with Liping Du and Yu Shyr in the Center for Quantiative Sciences.

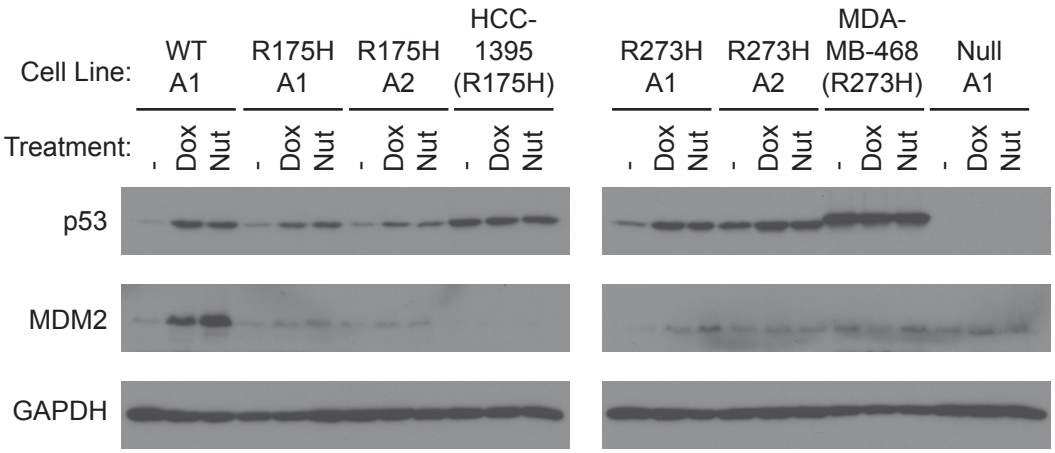


Figure 15: Isogenic p53 mutants accumulate after treatment with p53-MDM2 inhibitor Nutlin.

Immunoblot of lysates from CAL51, the indicated isogenic clones, or commercial cell lines with p53 genotype indicated in parentheses. Cells were lysed at baseline (-) or following 6 hours of treatment with 2 μM doxorubicin (Dox) or 10 μM (-)-Nutlin-3 (Nut). Blots were probed for the indicated proteins. GAPDH was included as a loading control. (-)-Nutlin-3 was provided by the Johnston Lab, Department of Chemistry, Vanderbilt University.

transactivation in response to stress (Fig. 15). In contrast, commercial TNBC cell lines expressing R175H and R273H missense mutant p53 displayed high levels of p53 protein that were unchanged in response to doxorubicin or Nutlin, despite the R273H-expressing MDA-MB-468 cell line exhibiting a similar baseline level of MDM2 protein (Fig. 15).

Previous studies have identified that mutant GOF includes modulation of the mevalonate pathway in the TNBC cell lines MDA-MB-231 and MDA-MB-468 as well as the non-transformed immortalized human mammary epithelial cell line MCF10A (Freed-Pastor et al., 2012). A later publication demonstrated that treatment with the statin class of mevalonate-5-phosphate inhibitors could degrade structural mutants such as R175H, but not contact mutants such as R273H, in a variety of cell lines (Parrales et al., 2016). We evaluated a similar dose range of atorvastatin and lovastatin in our isogenic cell line panel and confirmed no change in wild-type or R273H p53 protein levels (Fig. 16) in response to statin treatment. One R175H clonal population, R175H A2, showed no reduction in mutant p53 protein at any dose tested, while the other clonal population, R175H A1, exhibited a reduction in mutant p53 protein at the highest dose levels but lacked the dose-response relationship observed by Parrales *et al*. (Fig. 16).

The lack of constitutive stabilization of our mutant p53 protein and our inability to reproduce a dose response to statin-induced R175H degradation highlight the confounding effect of differing cellular and

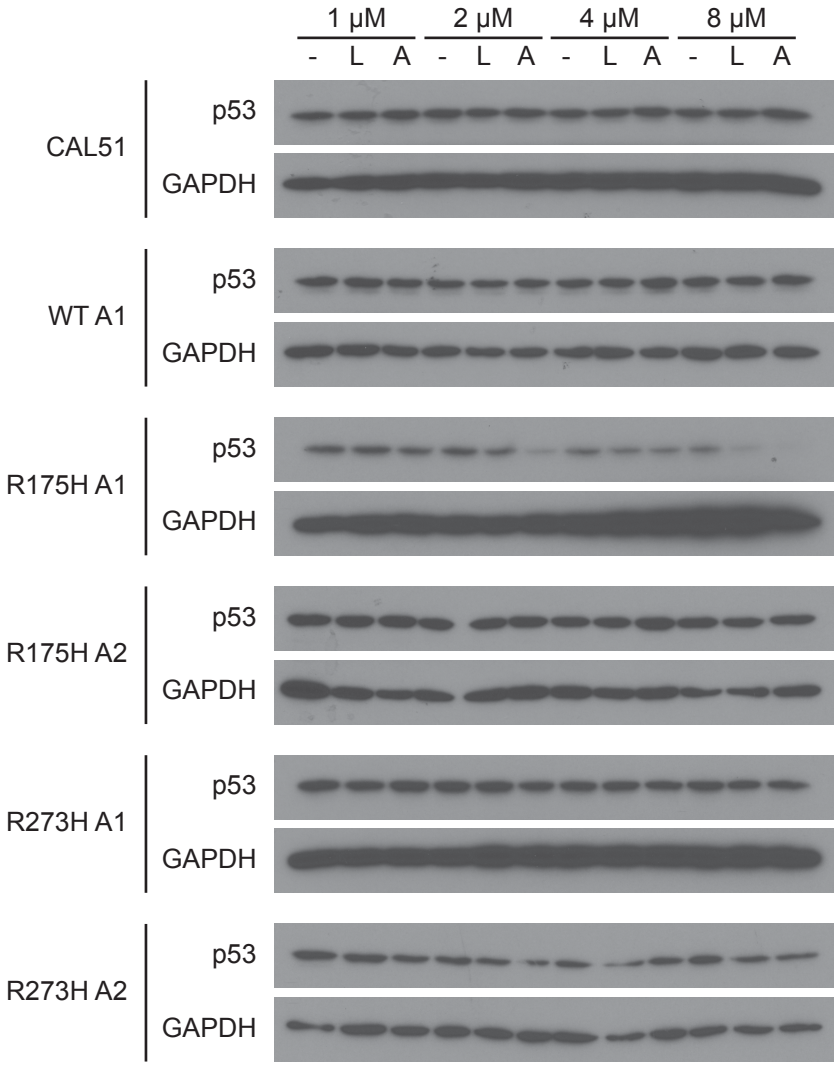


Figure 16: Wild-type and mutant p53 protein levels are unaffected by statin treatment
Cultures of the indicated cell line (CAL51) or isogenic clones were treated at ~40-50% confluency with the indicated concentrations of lorvastatin (L), atorvastatin (A), or an equivalent dilution of DMSO (-) in complete culture medium. Following 24 h of treatment, lysates were obtained and immunoblotted for total p53 protein. GAPDH was included as a loading control.

biochemical contexts. Our results occurred even after long-term growth with endogenously expressed p53; thus, findings from other GOF modeling conditions, including transient overexpression of mutant p53, are unlikely to be fully reproducible in differing biological settings, such as tumors in which p53 mutation was an early and selection pressure-defining event.

Discussion

As a result of the research described in this chapter, we demonstrate that mutant p53 cell lines expressing both the R175H and R273H missense alleles exhibit previously published GOF phenotypes, with the endogenous mutant p53 protein in our isogenic lines remaining subject to negative regulation by MDM2 despite failing to transactivate this E3 ligase in response to stress. Variability in the acquisition of GOF phenotypes manifested itself even among the two biological replicates per genotype in our initial cell line panel, with R175H A1 acquiring aneuploidy, elevated metabolism, and increased cell size, but not enhanced growth in xenograft tumors, whereas R175H A2 generated larger xenograft tumors while displaying only a slightly elevated metabolic rate and an unaltered chromosomal count and cell size. The missense mutant-specific gene expression changes we observed demonstrate that the transcriptional effects of p53 missense mutation extend beyond the mere p53 loss of function modeled by the null isogenic populations, and cannot be solely attributed to copy number loss and gain.

In our larger validation panel, we demonstrated that mutant p53 GOF (and, to a lesser extent, p53 loss of function) was associated with a higher frequency but not uniform acquisition of aneuploidy. These results demonstrate the difficulty of reproducing GOF phenotypes and mechanisms, even when leveraging endogenous alleles in an isogenic setting, and emphasize the challenges of clonal heterogeneity for both basic scientists and clinicians. Particularly in the context of CRISPR knock-in mutants, which leverage single-cell cloning and screening due to the inefficient homology-directed repair process (Ran et al., 2013), care must be taken to control for clonal variability by assessing adequate numbers of biological replicates. At present, we are extending our observations of the various GOF phenotypes we noted in our initial isogenic cell line panel to the larger validation set, with the goal of determining which GOF phenotypes and gene expression changes are statistically and mechanistically linked.

Even given the limited number of initial clones, our finding that missense isogenic populations exhibited previously published GOF phenotypes, despite the mutant p53 protein remaining subject to MDM2 regulation, calls into question whether constitutive mutant p53 stabilization is in fact a prerequisite for GOF, as suggested by previous studies (Alexandrova et al., 2017; Muller & Vousden, 2014; Suh et al., 2011; Terzian et al., 2008). The mutant protein regulation we observed closely mimics that seen in the normal tissues of knock-in mice (Terzian et al.,

2008), a model system in which tumors exhibit constitutive stabilization. Our results suggest that GOF phenotypes might be conferred by mutant p53 at an earlier stage than currently recognized, which has important implications for our understanding of tumor evolution and the selection pressures resulting from p53 mutational status.

CHAPTER 4

DIVERSE, BIOLOGICALLY RELEVANT, AND TARGETABLE GENE REARRANGEMENTS IN TRIPLE-NEGATIVE BREAST CANCER AND OTHER MALIGNANCIES

Triple-negative breast cancer (TNBC) and other molecularly heterogeneous malignancies present a significant clinical challenge due to a lack of high-frequency “driver” alterations amenable to therapeutic intervention. These cancers often exhibit genomic instability, resulting in chromosomal rearrangements that impact the structure and expression of protein-coding genes. However, identification of these rearrangements remains technically challenging. Using a newly developed approach that quantitatively predicts gene rearrangements in tumor-derived genetic material, we identified and characterized a novel oncogenic fusion involving the MER proto-oncogene tyrosine kinase (*MERTK*) and discovered a clinical occurrence and cell line model of the targetable *FGFR3-TACC3* fusion in TNBC.

Expanding our analysis to other malignancies, we identified a diverse array of novel and known hybrid transcripts, including rearrangements between non-coding regions and clinically relevant genes such as *ALK*, *CSF1R*, and *CD274*/*PD-L1*. The over 1000 genetic alterations we identified highlight the importance of considering non-coding gene rearrangement

partners, and the targetable gene fusions identified in TNBC demonstrate the need to advance gene fusion detection for molecularly heterogeneous cancers.

Introduction

Despite advances in precision medicine, the treatment of many molecularly heterogeneous cancers remains challenging due to a lack of recurrent alterations amenable to therapeutic intervention. To make significant clinical advances against these recalcitrant cancers, integrative genomic and molecular analyses are required to understand their complexity and to identify targetable features arising from lower-frequency genetic events. Our laboratory has identified six molecular subtypes of one such heterogeneous disease, triple-negative breast cancer (TNBC), with each subtype displaying unique ontologies and differential response to standard-of-care chemotherapy (Lehmann et al., 2011; Masuda et al., 2013). Ongoing genomic analysis of TNBC has identified a low frequency and widely varying clonality of therapeutically actionable alterations across subtypes, including mutations in *PIK3CA* and *BRAF*, amplification of *EGFR*, loss of *PTEN*, and expression of androgen receptor (*AR*) (Cancer Genome Atlas, 2012b; Lehmann et al., 2014; Shah et al., 2012). However, a significant proportion of TNBC cases lack somatic alterations of any established “driver” gene, highlighting the importance of continued genomic analysis and discovery integrated with molecular profiling of the tumor

microenvironment (Shah et al., 2012).

A defining feature of TNBC and other clinically challenging, molecularly heterogeneous cancers such as ovarian carcinoma and lung squamous cell carcinoma is copy number alteration (CNA), which is frequently accompanied by mutation or inactivation of the p53 tumor suppressor (Cancer Genome Atlas, 2012b; Cancer Genome Atlas Research, 2011, 2012; Hu et al., 2009). The genomic instability that gives rise to CNA can also result in chromosomal rearrangements and gene fusions, which have long been recognized as oncogenic drivers and effective drug targets in a variety of hematological and solid malignancies (Rabbitts, 1994). In recent years, the advent of next-generation sequencing (NGS) has enabled the identification of a number of gene fusion events in epithelial cancers, with varying frequency and therapeutic relevance (Shaw et al., 2013; Soda et al., 2007; Tomlins et al., 2005; Wang et al., 2011a; Williams et al., 2013).

In order to broaden our understanding of the somatic alterations underlying TNBC, we sought to identify known and novel gene rearrangements impacting the structure and/or expression level of protein-coding transcripts. While a number of computational approaches have been developed to identify hybrid RNA and DNA sequences using NGS data, numerous technical hurdles complicate accurate and efficient gene rearrangement detection (Davare & Tognon, 2015). For instance, widespread regions of sequence homology and current limitations in read

length result in sometimes-ambiguous alignment and a large number of false positives. To combat this, many detection methodologies employ filters designed to enrich for biologically relevant gene rearrangements, such as restricting both potential fusion partners to protein-coding loci (Annala et al., 2013). While these filters enrich for currently known gene fusions, they are not effective for the discovery of less canonical events, such as rearrangements between coding and non-coding regions of the genome. We developed a new algorithm, Segmental Transcript Analysis (STA), which uses exon-level expression estimates to rank a population of samples based on their likelihood of harboring a rearrangement in a given gene. We identified a number of known and novel rearrangements involving functionally diverse gene partners in TNBC, and expanded our analysis and discovery to a wider, multi-cancer cohort from The Cancer Genome Atlas (TCGA). The DNA-validated rearrangements that we identified include non-coding portions of the genome acting as both 5' and 3' partners in clinically relevant hybrid transcripts.

Results

Gene Rearrangement Prediction by STA

In order to prioritize downstream computation and minimize filters restricting the genomic location of putative rearrangement partners, we developed an algorithm known as Segmental Transcript Analysis (STA)

that uses a distance matrix approach to generate aberrant transcript scores for a population of samples. Based on exon-level expression values across a gene, STA is used to assign a normalized score for each sample by quantifying deviation from the population in both magnitude and directionality of expression. This approach allows the detection of different structural classes of rearrangements – general up- or down-regulation of a transcript might accompany promoter or untranslated region (UTR) swapping, for instance, while abrupt gain or loss of expression from one exon to the next could result from a breakpoint within the intervening DNA. While past discovery approaches using microarray datasets have leveraged exon-level expression comparison in individual transcripts (Giacomini et al., 2013), the novel population-based comparison method employed in STA effectively controls for confounding issues such as alternative splicing and normalization artifacts that cause uneven exon-level expression values across genes and facilitates detection of modest but biologically relevant changes in transcript levels.

To evaluate the utility of STA as a prediction tool for both known and novel gene rearrangements, we developed a vertically integrated NGS analysis pipeline (Fig. 17). Briefly, exon-level expression data were collated for a given tumor type and STA scores were calculated for each sample on a per-gene basis (Fig. 17A-B). RNA-seq data for each aberrant transcript passing a defined STA score threshold were assessed for evidence of rearrangement between the candidate gene and another genomic region

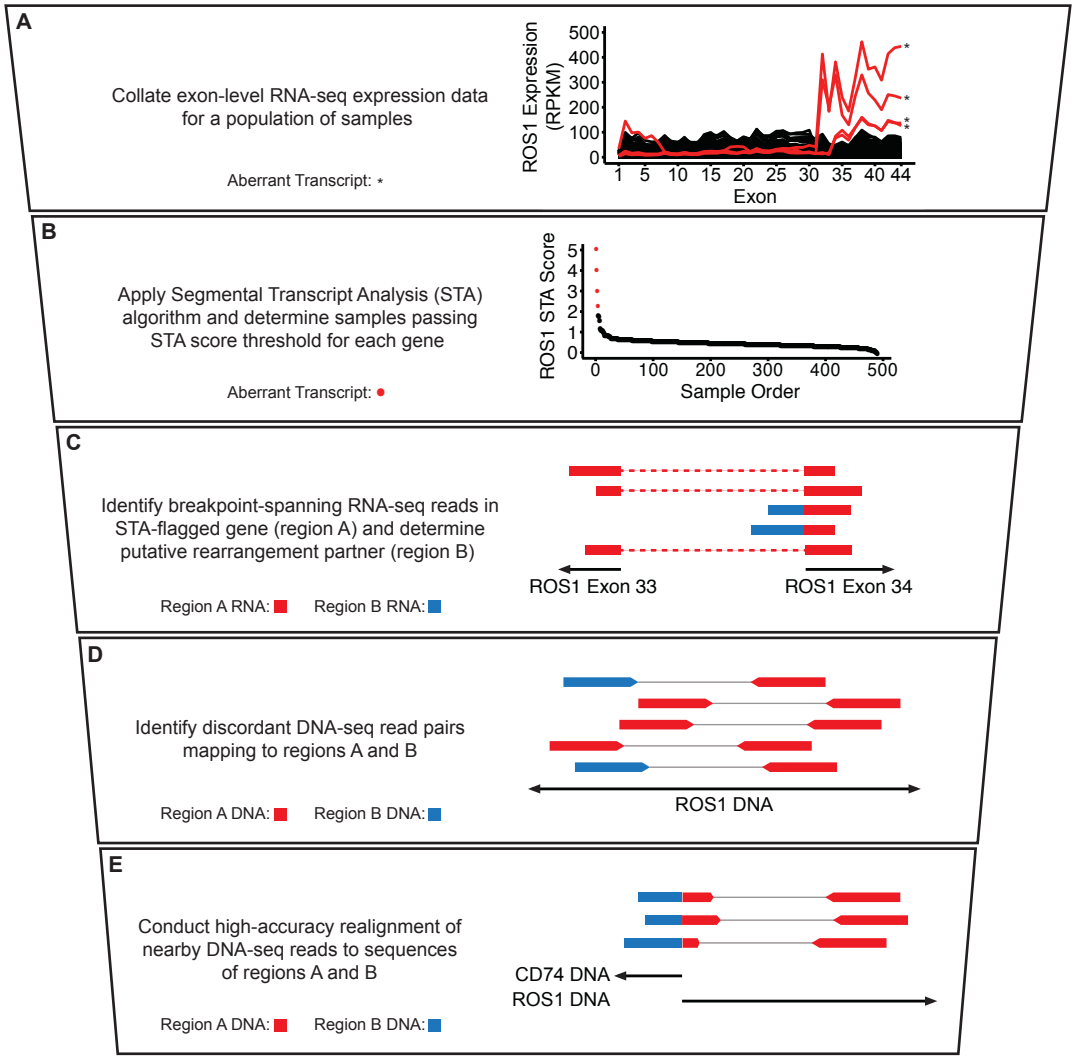


Figure 17: Quantitative prediction by STA facilitates an integrated fusion detection pipeline.

A-E, Stepwise description of STA discovery pipeline with accompanying schematics and example data. A, Exon-level expression values for a population of samples plotted as continuous lines. Samples passing STA score threshold (B) are plotted in red and denoted by asterisks; samples below the threshold are plotted in black. B, ROS1 STA scores for each sample plotted in descending order. Samples with an STA score of 2 or above are plotted in red; samples with an STA score below 2 are plotted in black. C, Schematic of RNA-seq reads. Dotted lines denote continuous read segments. D, Schematic of discordant whole-genome sequencing (WGS) read pairs. Thin lines represent denote read pairs. E, Schematic of breakpoint-spanning WGS reads identified after realignment. In C-E, colors indicate alignment location of individual read segments, as depicted in the description at left. Schematics are not to scale.

(Fig. 17C), followed by analysis and realignment of nearby whole-genome sequencing reads (WGS) to identify a DNA breakpoint with unique sequence spanning both rearrangement partners (Fig. 17D-E) (Faust & Hall, 2012). Samples displaying a discrete breakpoint upon realignment were classified as DNA-validated rearrangements. While reliant upon the availability and adequate sequencing depth of WGS data, this approach provides independent structural evidence for the validity of any detected rearrangements. As a consequence, rearrangements that might be omitted in an RNA sequencing-only approach due to concerns about homology and false positivity, such as hybrid transcripts between coding and non-coding regions of the genome, can be identified with greater confidence.

Using test RNA-seq and WGS data from TCGA, we confirmed the ability of STA to identify known gene fusions across multiple cancer types, including *CD74-ROS1* in lung adenocarcinoma, *NFASC-NTRK1* in glioblastoma, and *TMPRSS2-ETV4* in prostate adenocarcinoma (Fig. 17 and Fig. 18A-B) (Kim et al., 2014; Rikova et al., 2007; Tomlins et al., 2006). The algorithm also reliably identified therapeutically actionable anaplastic lymphoma kinase (*ALK*) fusions in lung adenocarcinoma, although DNA validation was not available in all cases due to incomplete availability of WGS data (Fig. 18C) (Soda et al., 2007). Numerous aberrant transcripts were detected for the RET receptor tyrosine kinase, which undergoes rearrangement in 10-20% of sporadic papillary thyroid cancers (Fig. 18D) (Grieco et al., 1990). When possible, we obtained DNA validation for *RET*

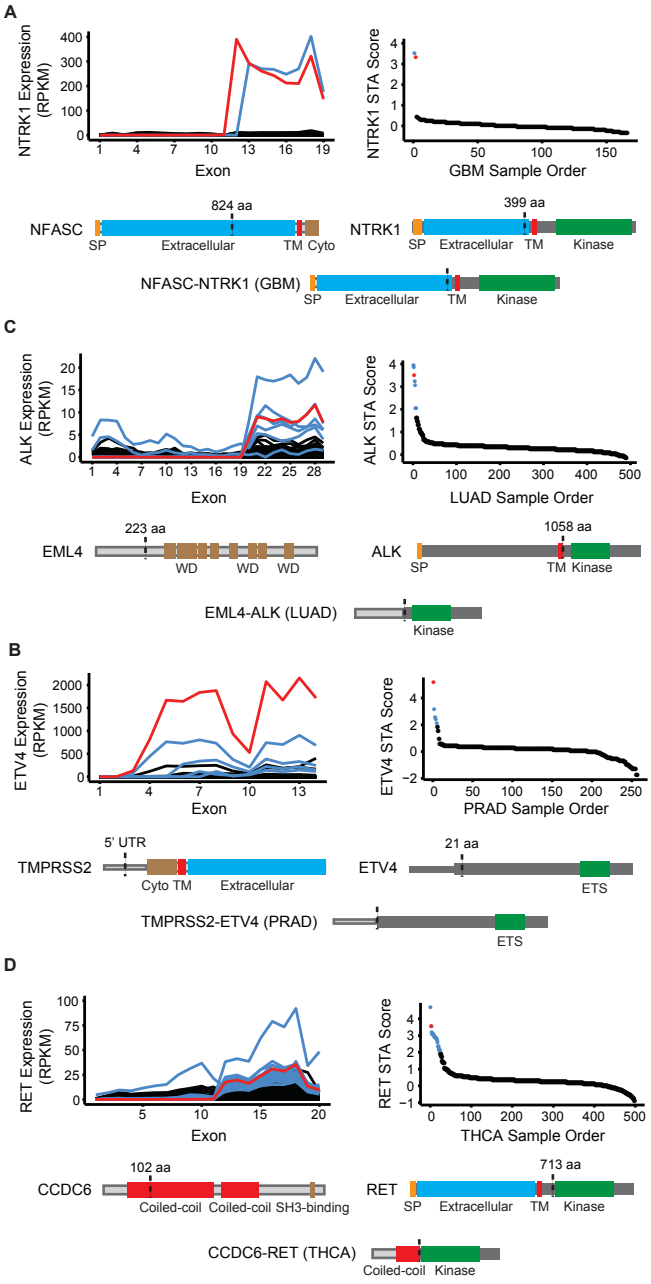


Figure 18: STA reliably predicts known gene fusions in multiple cancer types.
A-D, Four examples of STA-predicted rearrangements from multiple tumor types in TCGA. Each panel features an exon-level expression diagram and STA score plot for the gene and cancer type analyzed. Red indicates the representative DNA-validated rearrangement that is depicted at the bottom of each panel as a schematic of the resulting aberrant protein. Blue indicates additional aberrant transcripts meeting STA score threshold. Black indicates population background below threshold. Protein features and untranslated junctions (UTRs) are labeled and dotted lines indicate hybrid transcript regions. aa: amino acid; Cyto: cytoplasmic domain; GBM: glioblastoma multiforme; LUAD: lung adenocarcinoma; PRAD: prostate adenocarcinoma; SP: signal peptide; THCA: thyroid carcinoma; TM: transmembrane.

fusions in these samples, including known rearrangements with *CCDC6*, *ERC1*, and *NCOA4* (Shaver et al., 2016). To evaluate the ability of STA to predict fusions previously identified in TCGA RNA-seq, we compared the STA score distribution of fusion transcripts across all genes (Yoshihara et al., 2014) (Fig. 19A) and recurrently fused kinases (Stransky et al., 2014) (Fig. 19B) to the distribution for all genes and samples evaluated. In both cases, STA scores were significantly higher in the rearrangement-harboring transcripts ($p < 2.2x10^{-16}$).

Due to the ability of STA to detect aberrant loss of expression in addition to gain, we hypothesized that inactivation of tumor suppressors would constitute a substantial portion of STA-predicted rearrangements. Indeed, the algorithm displayed a robust ability to identify intragenic loss of expression of known tumor suppressors. In a lung squamous cell carcinoma sample, we identified a rearrangement resulting in early truncation of the SWI/SNF subunit *ARID1A* (Fig. 20A) (Jones et al., 2010). We additionally validated rearrangements with non-coding DNA in bladder and endometrial carcinoma resulting in the loss of functional domains of the PI3K negative regulator PTEN and the p300 histone acetyltransferase, respectively (Fig. 20B-C) (Gayther et al., 2000; Lee et al., 1999). In a kidney chromophobe tumor, we identified a rearrangement between a non-coding portion of chromosome 5 and the gene encoding the miRNA-processing enzyme DROSHA that results in its early truncation (Fig. 20D) (Lee et al., 2003). Interestingly, inactivating *DROSHA* mutations have

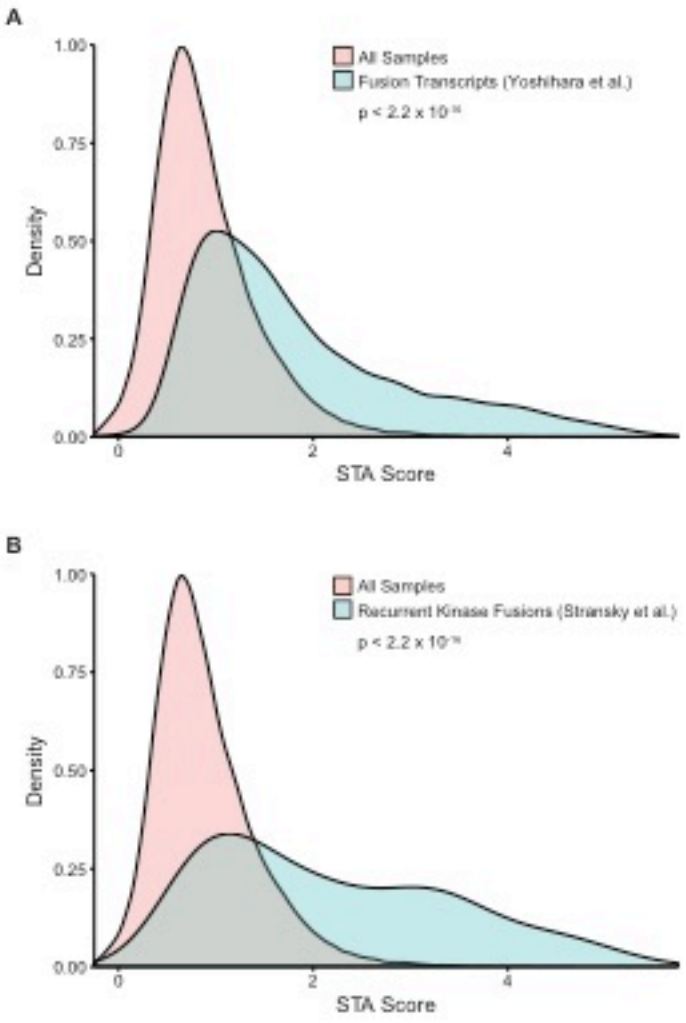


Figure 19: STA significantly enriches for known fusion transcripts in TCGA data.
A-B, Density plots comparing the STA score distribution for all genes and samples evaluated (red) to fusion transcripts identified by two previous TCGA fusion discovery studies (green). The corresponding reference is indicated in the legend. Score distributions were compared using two-sample, two-sided Kolmogorov-Smirnov tests and the resulting p-value is displayed for each panel.

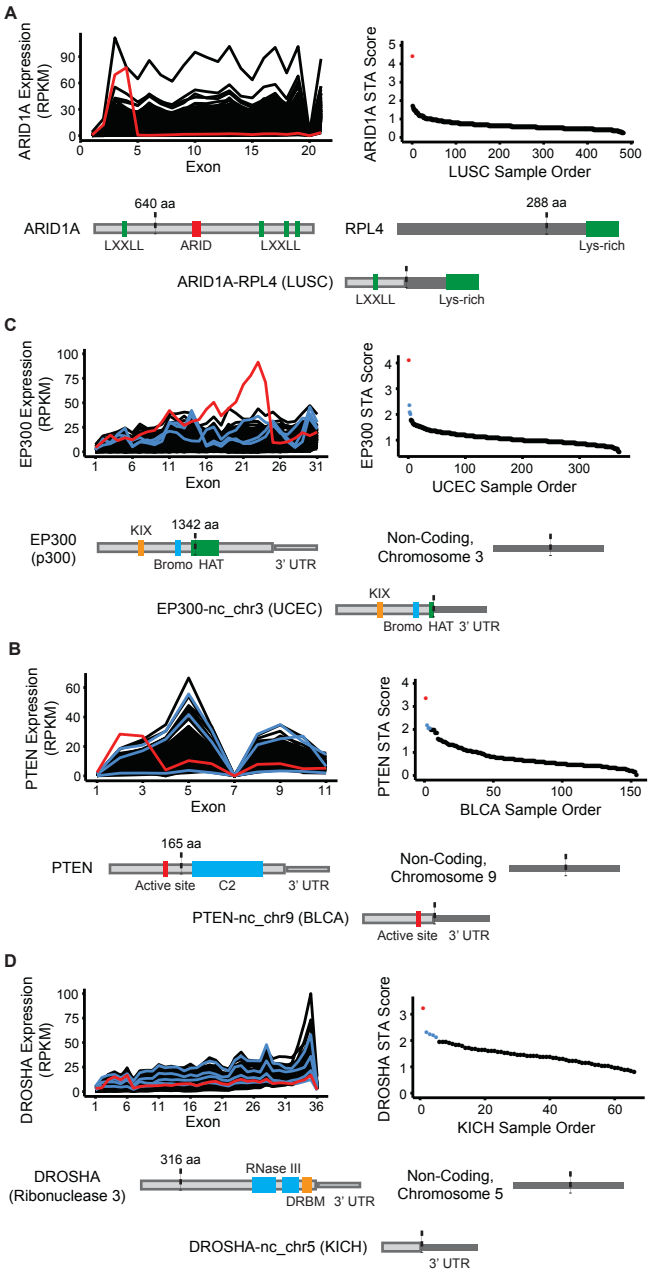


Figure 20: STA predicts tumor suppressor truncation based on intragenic loss of expression.

A-D, Four examples of STA-predicted rearrangements from multiple tumor types in TCGA. Each panel features an exon-level expression diagram and STA score plot for the gene and cancer type analyzed. Red indicates the representative DNA-validated rearrangement that is depicted at the bottom of each panel as a schematic of the resulting aberrant protein. Blue indicates additional aberrant transcripts meeting STA score threshold. Black indicates population background below threshold. Protein features and untranslated regions (UTRs) are labeled and dotted lines indicate hybrid transcript junctions. aa: amino acid; BLCA: bladder urothelial carcinoma; KICH: kidney chromophobe; LUSC: lung squamous cell carcinoma; UCEC: uterine corpus endometrial carcinoma.

recently been reported in over 10% of cases of Wilms tumor, a pediatric kidney cancer (Torrezan et al., 2014).

A Novel Oncogenic Kinase Fusion in TNBC

To discover known and novel gene rearrangements in clinical TNBC cases, we used NGS data from TCGA and performed STA prediction on 173 TNBC tumors. We identified and validated at the DNA level a previously uncharacterized fusion involving the MER proto-oncogene tyrosine kinase (*MERTK*) (Graham et al., 1994). In this rearrangement, a nearby gene encoding the transmembrane protein TMEM87B acts as 5' partner, breaking shortly after its signal peptide and fusing with the late extracellular domain-coding portion of *MERTK* (Fig. 21A). The resulting fusion transcript displays increased expression and retains the full transmembrane and intracellular kinase domains of MERTK, which is overexpressed or ectopically expressed in numerous cancers (Fig. 21B) (Cummings et al., 2013).

In order to assess if this truncated form of MERTK retains its ability to activate the oncogenic MAPK/Erk and Akt signaling pathways (Schlegel et al., 2013), we engineered a retroviral expression construct encoding the tumor-derived *TMEM87B-MERTK* fusion. In the IL3-dependent Ba/F3 mouse lymphocyte cell line, stable expression of TMEM87B-MERTK led to constitutively elevated levels of phospho-Akt and retention of robust Erk and Akt signaling even after serum starvation and withdrawal of IL3 (Fig.

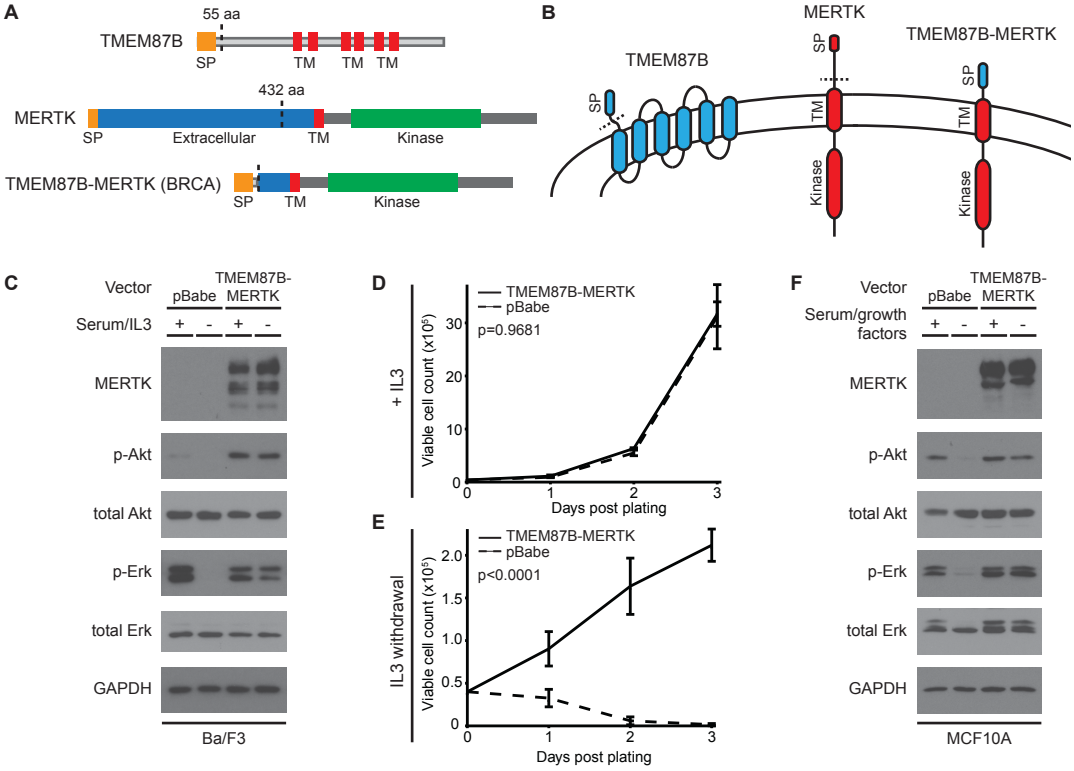


Figure 21: The TMEM87B-MERTK gene fusion in TNBC promotes constitutive oncogenic signaling and cell survival.

A, Diagram of the TMEM87B and MERTK proteins and the DNA-validated gene fusion protein product. Protein features are labeled. B, Protein schematics indicating membrane topology (not to scale). Colors indicate protein sequences as indicated. In A and B, dotted lines represent protein regions encoded by the gene fusion transcript. C, Immunoblot analysis of the indicated proteins from Ba/F3 cells transfected with an empty vector or one expressing the TMEM87B-MERTK fusion gene. Cell lines were grown in the continuous presence of 5% FBS and 1 ng/mL IL3 (+) or switched to 0.5% FBS and no IL3 (-) for 90 min. D-E, Graphs depicting growth curves of Ba/F3 cells transfected with the TMEM87B-MERTK fusion gene (solid line) or empty vector (dotted line). Cells were grown in the continuous presence of 1 ng/mL IL3 (D) or switched to no-IL3 media at day 0 (E) and viable cell counts were obtained by hemocytometer with trypan blue exclusion at the indicated timepoints. Error bars represent standard deviation of three replicates and p-values comparing the two conditions are specified at top left. F, Immunoblot analysis of the indicated proteins from MCF10A cells transfected with constructs identical to C. Cells were grown in complete growth media with 2.5% horse serum (+) or switched to base media with 0.5% horse serum and no growth factor additives for 180 min (-). aa: amino acid; BRCA: breast invasive carcinoma; SP: signal peptide; TM: transmembrane.

21C) (Palacios & Steinmetz, 1985). Accordingly, while the TMEM87B-MERTK and empty vector control cells grew similarly in the presence of IL3, the fusion protein conferred a clear survival advantage after IL3 withdrawal (Fig. 21D-E). Whereas the control cells died by day 3 after IL3 withdrawal, the TMEM87B-MERTK-expressing cells proliferated under the same conditions (Fig. 21E) and could be cultured in the absence of IL3 for at least one month (data not shown). These results are consistent with the survival-promoting role of full-length MERTK in melanoma, glioblastoma, and other cancers (Brandao et al., 2013; Schlegel et al., 2013; Wang et al., 2013b). To verify that the fusion protein-modulated signaling could be replicated in breast-derived, basal epithelial cells, we expressed TMEM87B-MERTK in immortalized MCF10A cells and observed similar activation of Erk and Akt after serum starvation and growth factor withdrawal (Fig. 21F).

Of note, we identified an identical *TMEM87B-MERTK* fusion in the lung squamous cell carcinoma RNA-seq data set from TCGA, along with a *BCL2L11-MERTK* RNA transcript with the same breakpoint in bladder carcinoma, but WGS data were not available for DNA validation (data not shown). An independent RNA-seq fusion analysis of TCGA samples corroborated the *TMEM87B-MERTK* fusion in TNBC and identified identical rearrangements in cervical carcinoma and lung adenocarcinoma (Yoshihara et al., 2014), demonstrating selection for a recurrently truncated form of MERTK in multiple cancer types.

FGFR3-TACC3 Gene Fusion is a Targetable Driver Alteration in TNBC

In order to identify and evaluate an endogenous model of oncogenic gene fusions in TNBC, we expanded our analysis to RNA-seq data from 80 additional TNBC tumors (Shah et al., 2012) and 28 TNBC cell line models. In both a tumor specimen and the SUM185PE cell line, we discovered *FGFR3-TACC3* fusions similar to the oncogenic rearrangements recently observed in glioblastoma and bladder carcinoma, which result in the fusion of the FGFR3 kinase domain to the coiled-coil domain of TACC3 (Fig. 22A-B) (Singh et al., 2012; Williams et al., 2013).

To determine if *FGFR3-TACC3* is a targetable 'driver alteration' in TNBC, we conducted knockdown and pharmaceutical inhibition of the fusion protein in SUM185PE cells. Immunoblotting for FGFR3 in cell lysates produced distinct bands consistent with the predicted molecular weights of the wild-type and fused proteins (Fig. 22C). Two siRNAs targeting FGFR3 (Fig. 22B, siRNA #1-2) reduced expression of both forms of the protein and decreased cell viability after 72 h to levels comparable to a cell-death control (Fig. 22C-D).

To verify that the viability decrease was due to the loss of fused FGFR3 rather than wild-type, we assessed two siRNAs targeting the fused portion of TACC3 (Fig. 22B, siRNA #3-4) and two siRNAs targeting sequences outside the recombined region (Fig. 22B, siRNA #5-6). The two TACC3 siRNAs targeting a portion of the transcript contained within the fusion decreased expression of the resulting hybrid protein and reduced

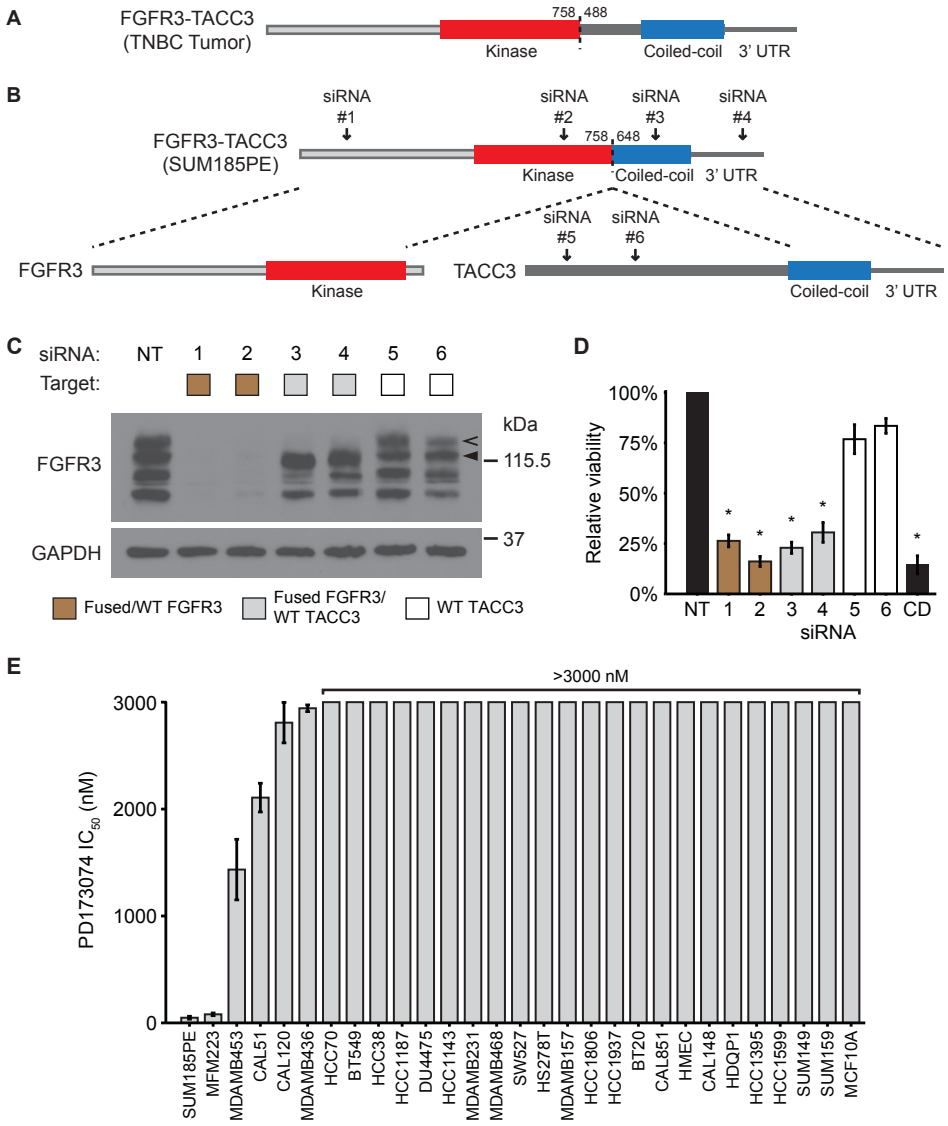


Figure 22: The FGFR3-TACC3 gene fusion is a targetable driver alteration in TNBC. A-B, diagram of the protein products of the FGFR3-TACC3 gene fusions found in a tumor sample from TNBC patient TTR0001024 (Shah *et al*.) (A) and the SUM185PE TNBC cell line (B). Protein features are labeled and dotted lines outline protein regions encoded by the gene fusion transcript. Numbers indicate amino acid position in the wild-type proteins and arrows indicate targeting locations of the siRNAs used in the experiments. C-D, Immunoblot analysis of the indicated proteins from SUM185PE lysate (C) and relative viability of the cells (D) after 72-hr treatment with the indicated siRNAs (depicted in B), a non-targeting control (NT), or a cell death-inducing positive control (CD). In C, the legend indicates proteins expected to undergo knockdown based on siRNA target location. Wild-type (WT) and fused forms of FGFR3 are denoted by a filled and hollow arrow, respectively. In D, viability as assessed by alamarBlue is normalized to the non-targeting control. Error bars represent standard error of the mean of four independent experiments. Asterisks indicate $p < 0.001$ when compared to NT control. E, Half-maximal inhibitory concentrations (IC50) of the FGFR inhibitor PD173074 for the indicated cell lines as assessed by alamarBlue assay. Error bars represent standard error of the mean of three independent experiments.

viability to a level similar to FGFR3 knockdown, whereas addition of the siRNAs targeting a portion of TACC3 outside the fused region did not significantly decrease viability or fusion protein expression (Fig. 22C-D).

Additionally, SUM185PE cells displayed pronounced sensitivity to the FGFR inhibitor PD173074 (IC_{50} = 48 ± 13 nM), whereas the majority of other cell lines tested displayed micromolar or greater half-maximal inhibitory concentrations (Fig. 22E) (Mohammadi et al., 1998). The only other line with similar sensitivity, MFM223, harbors a previously reported amplification of *FGFR2* (Turner et al., 2010).

Novel and Non-Canonical Rearrangements in TNBC

Additional STA predictions from the TCGA TNBC clinical data set led to the identification and DNA validation of a structurally and functionally diverse array of gene rearrangements. In one sample, dual 5' UTR breakpoints result in promoter swapping between the myosin heavy chain gene *MYH9* and the histone modifier gene *NFYC*, which was recently identified as an oncogene in choroid plexus carcinoma (Tong et al., 2015). The resulting fusion transcript is highly expressed and retains the entire NFYC open reading frame (Fig. 23A). In another rearrangement, the 55 kDa isoform of the transmembrane glycoprotein neuroplastin is fused to the C-terminus of the cilia-associated transcript CLUAP1, leading to a transcript encoding the signal peptide and a single extracellular Ig-like domain from neuroplastin (Fig. 23B). Importantly, a small portion of the

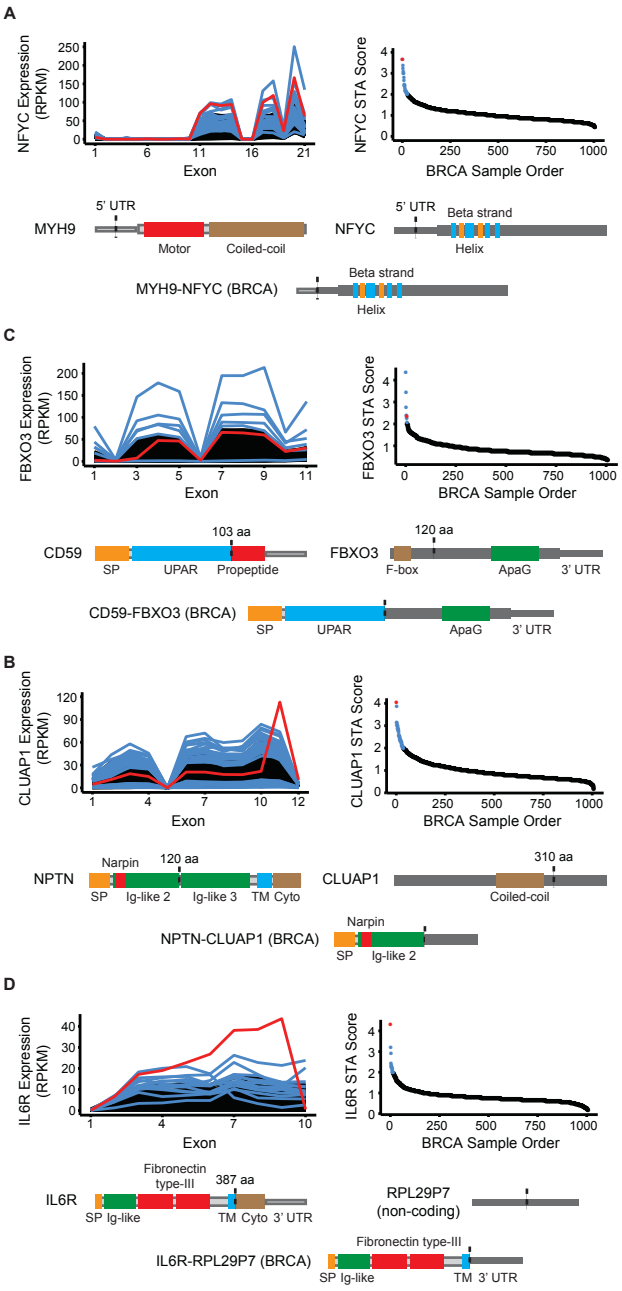


Figure 23: Triple-negative breast cancers harbor a functionally diverse array of gene rearrangements.

A-D, Four examples of STA-predicted rearrangements in triple-negative breast cancers from TCGA. Each panel features an exon-level expression diagram and STA score plot for the gene and cancer type analyzed. Red indicates the representative DNA-validated rearrangement that is depicted at the bottom of each panel as a schematic of the resulting aberrant protein. Blue indicates additional aberrant transcripts meeting STA score threshold. Black indicates background population below threshold. Protein features and untranslated regions (UTRs) are labeled and dotted lines indicate hybrid transcript junctions. aa: amino acid; BRCA: breast invasive carcinoma; Cyto: cytoplasmic domain; nt: nucleotide; SP: signal peptide; TM: transmembrane.

retained Ig-like domain was previously demonstrated to be sufficient for the FGFR1 activation exhibited by the full-length protein (Owczarek et al., 2010), implying that the *NPTN-CLUAP1* gene fusion may lead to the secretion of a paracrine FGFR1 activator.

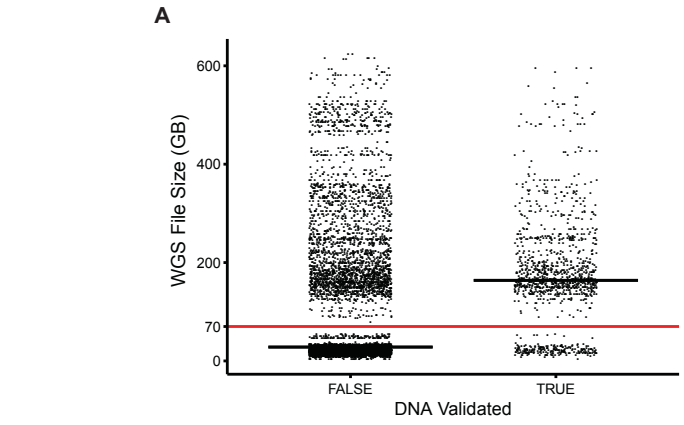
We also identified rearrangements in TNBC involving immune-related proteins. In one case, *FBXO3* undergoes rearrangement with the gene encoding the membrane attack complex inhibitor CD59*,* with resultant expression of a transcript encoding the functional domains of CD59. Interestingly, the RNA breakpoint for *CD59* occurs at a non-annotated splice site within the coding sequence that precisely mimics the cleavage site of the mature protein from a glycosylphosphatidylinositol (GPI) anchor addition signal (Fig. 23C) (Sugita et al., 1993). The consequences of GPI anchor loss and the retention of a portion of FBXO at the C-terminus were not assessed; however, soluble forms of CD59 have been previously noted to retain their complement-mediated cytotoxicity-suppressive function(Brasoveanu et al., 1997).

In a final example, the gene encoding the interleukin 6 receptor (IL6R) breaks at the junction between its transmembrane and cytoplasmic domains and undergoes rearrangement with the non-coding pseudogene *RPL29P7* (Fig. 23D). Intriguingly, previous studies have demonstrated that the cytoplasmic domain of IL6R is dispensable for its interaction with the gp130 transactivator (Jones et al., 2001), and the resulting fusion transcript is highly expressed. We speculate that the IL6R-RPL29P7 fusion protein

retains the IL6-binding and transactivation capacity of wild-type IL6R, and is overexpressed due to loss of the negative-regulatory IL6R 3' UTR. This rearrangement illustrates a mechanism by which 3' hybrid transcript formation with non-coding regions of the genome can lead to increased expression of tumor-promoting genes, similar to the increased expression resulting from 3' UTR loss in the *FGFR3-TACC3* gene fusion (Parker et al., 2013) and confirming findings from previous gene fusion discovery efforts in breast cancer tumors and cell lines (Asmann et al., 2012; Edgren et al., 2011).

Functionally and Structurally Diverse Rearrangements Across Cancer

Given the ability of STA to identify novel classes of rearrangements in TNBC, we broadened our discovery efforts using 14 additional NGS datasets from TCGA. In total, we analyzed 5461 tumor samples with RNA-seq across 14 cancer types, of which 1264 (23%) had accompanying WGS available (Table 4). We validated 1178 gene rearrangements at the RNA and DNA levels (included as a supplemental table to Shaver et al., 2016). Of note, our attempts to validate newly discovered rearrangements at the DNA level were confounded in part by variable WGS availability and sequencing depth across TCGA studies. We found a clear correlation between both the validation rate and frequency of rearrangements detected per sample and WGS file size (Figs. 24 and 25).



B

DNA Validation Rate by File Size

WGS File Size	Total Analyzed	Total Validated	Validation Rate
< 70 GB	4386	224	5.1%
≥ 70 GB	3677	954	25.9%

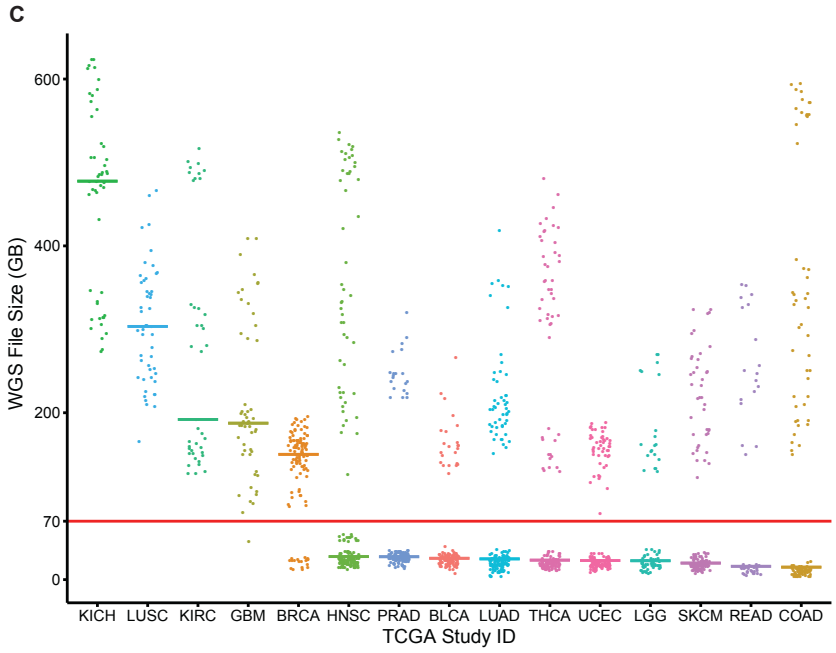
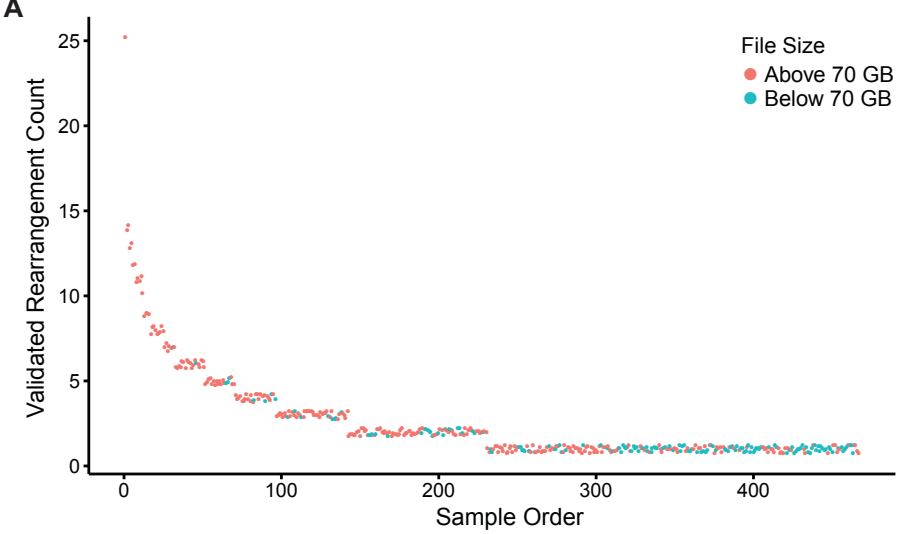


Figure 24: Low WGS depth is associated with a decreased rate of DNA validation.
A, DNA validation status of gene rearrangements predicted by STA and identified in RNA-seq data. Each point denotes the maximum whole-genome sequencing (WGS) gigabyte (GB) file size available for an individual rearrangement. Black horizontal lines indicate the median value. B, Total number of rearrangements analyzed, total number of rearrangements validated, and corresponding validation rate for samples with maximum WGS file size in each category. C, Largest WGS file size available for each sample in the TCGA studies analyzed. Studies are arranged in descending order of median file size (indicated by horizontal lines). In B and C, the red horizontal line indicates a 70 GB threshold for illustrative purposes.

A



B

Rearrangement Count vs File Size per Sample

WGS File Size	Total Samples	# Samples with ≥5 Rearrangements	Frequency
< 70 GB	151	5	3.3%
≥ 70 GB	316	65	20.6%

Figure 25: The frequency of DNA-validated rearrangements per sample is correlated with WGS depth.
A, Number of DNA-validated rearrangements per TCGA sample. Data points are ordered by descending rearrangement count and colored by the maximum whole-genome sequencing (WGS) file size available for each sample, as indicated in the legend. B, Total number of samples with maximum WGS file size in each category, the number of samples with five or more rearrangements, and the corresponding frequency.

Fusions between two protein-coding genes constituted 40% of the validated rearrangements. Among the remaining rearrangements of coding genes with non-coding DNA, a majority featured the protein-coding gene as the 5' partner, as inferred by RNA breakpoints. For 3% of total rearrangements detected, however, non-coding regions of the genome acted as 5' hybrid transcript partners and caused deregulated expression of coding genes (Fig. 26A). To characterize the function of genes undergoing each rearrangement type, we classified rearrangement partners according to a previously published oncogene and tumor suppressor prediction method (Davoli et al., 2013). The enrichment pattern of genes in these categories was consistent with tumor-promoting gain or loss of function (Fig. 26B-C). Coding genes acting as the 3' transcript partner in out-of-frame fusions included a higher proportion of tumor suppressors, for example, whereas in-frame fusions and rearrangement at UTRs included more oncogenes (Fig. 26C).

Across tumor types, we noticed a trend of tumor-promoting gene overexpression by a structurally diverse set of gene rearrangements. In a thyroid carcinoma sample, we identified a rearrangement between a 5' portion of the long non-coding RNA (lncRNA) *MALAT1* and the recurrently fused *ALK* (Soda et al., 2007), leading to extremely high expression of a transcript retaining the ALK kinase domain (Fig. 27A). Of note, the *MALAT1-ALK* rearrangement occurs upstream of *ALK* exon 16 rather than the most common exon 19 and 20 breakpoints, but numerous in-frame

A

Rearrangements Detected by Category

Category	Occurrence	Frequency
Gene to Gene, Coding	**357**	**30%**
CDS to CDS, In-Frame	187	16%
CDS to CDS, Out-of-Frame	170	14%
Gene to Gene, UTR	**112**	**10%**
UTR to CDS	58	5%
CDS to UTR	39	3%
UTR to UTR	15	1%
Non-Coding	**633**	**54%**
CDS/UTR to Non-Coding	593	50%
Non-Coding to CDS/UTR	40	3%
Other	**76**	**6%**
Total	**1178**	
3' DNA Breakpoint Upstream of Gene	**80**	**7%**

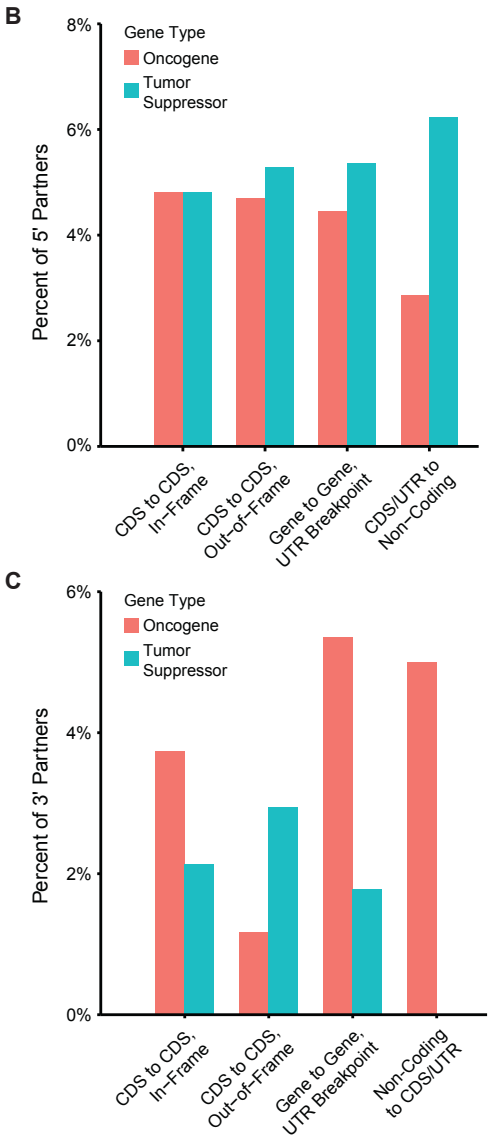


Figure 26: Rearrangements predicted by STA span a diverse array of structural and functional categories.

A, Structural categorization of DNA-validated rearrangements identified by the STA discovery pipeline. Rearrangements are described in order of 5' and 3' partners. CDS or UTR status for coding genes denotes a hybrid transcript junction occurring at an annotated exon boundary. B-C, Functional categorization of coding genes serving as hybrid transcript partners after gene rearrangement. Rearrangements are divided into structural categories, as described in a. The displayed value represents the percent of all 5' (B) and 3' (C) transcript partners in each category annotated as oncogenes or tumor suppressors, as denoted in the legend.

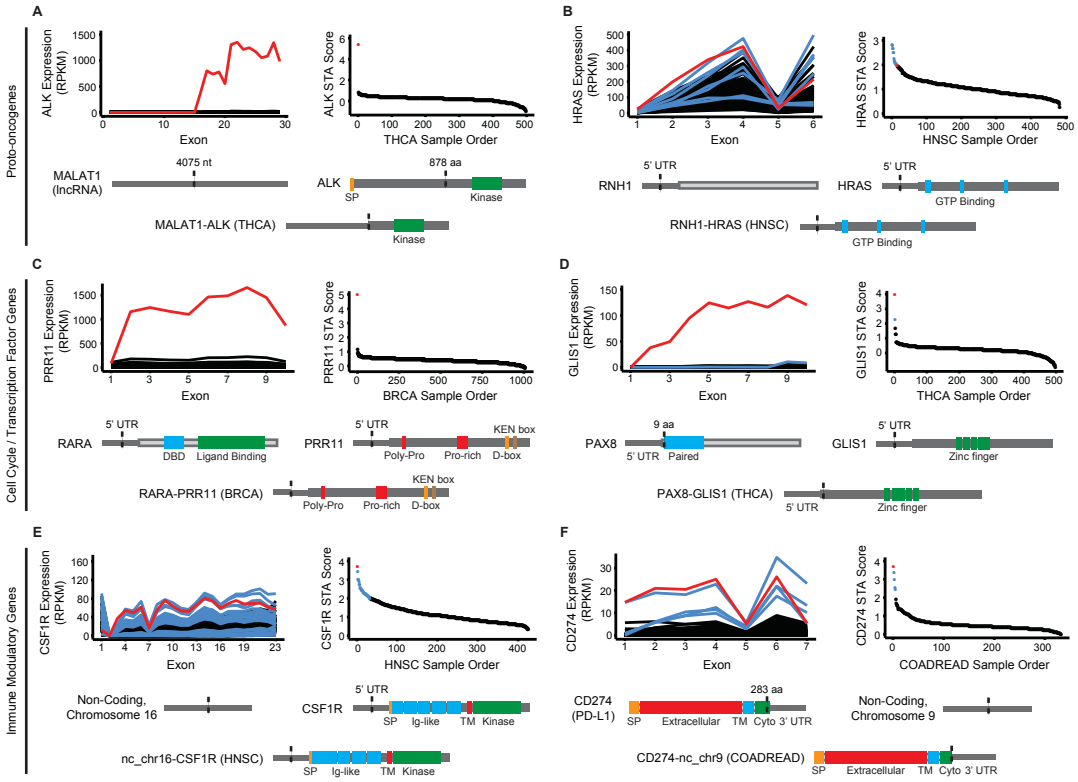


Figure 27: Overexpression of oncogenic transcripts across cancer types results from gene rearrangement with coding and non-coding DNA.
A-F, Six examples of STA-predicted rearrangements from additional tumor types in TCGA, representing the categories described at left. Each panel features an exon-level expression diagram and STA score plot for the gene and cancer type analyzed. Red indicates the representative DNA-validated rearrangement that is depicted at the bottom of each panel as a schematic of the resulting aberrant protein. Blue indicates additional aberrant transcripts meeting STA score threshold. Black indicates background population below threshold. Protein features and untranslated regions (UTRs) are labeled and dotted lines indicate hybrid transcript junctions. aa: amino acid; BRCA: breast invasive carcinoma; COADREAD: colorectal carcinoma; Cyto: cytoplasmic domain; HNSC: head and neck squamous cell carcinoma; nt: nucleotide; SP: signal peptide; THCA: thyroid carcinoma; TM: transmembrane.

methionines in the exon 16-19 region could allow translation initiation after the non-coding, 5' *MALAT1* portion of the hybrid transcript (data not shown).

We also identified abnormally high expression of the MAPK activator *HRAS* in a head and neck squamous cell carcinoma sample. WGS analysis revealed a rearrangement between *RNH1* and a DNA breakpoint upstream of the *HRAS* transcriptional start site. The 5' UTR of *RNH1* is subsequently spliced to exon 2 of *HRAS*, resulting in elevated transcription of a complete HRAS open reading frame (Fig. 27B). We noted a similar overexpression of embryonic stem cell-expressed Ras (*ERAS*) by the *PQBP1* promoter in lung squamous cell carcinoma, but the DNA breakpoint for that sample occurred within the first intron of *ERAS*. Although *KRAS* fusions have been described in metastatic prostate cancer (Wang et al., 2011a), we believe these are the first DNA-validated fusions involving *HRAS* and *ERAS* to be reported. A similar rearrangement leading to overexpression of the entire HRAS coding sequence was identified in a murine Moloney leukemia virus-induced cell line and showed the ability to transform NIH 3T3 fibroblasts (Ihle et al., 1989).

DNA breakpoints upstream of the transcriptional start site, as seen in *RNH1-HRAS*, can lead to extreme overexpression that is easily detectable by STA. In an estrogen receptor (ER)-positive breast cancer sample, the ER-responsive gene *RARA* displayed rearrangement with a region upstream of *PRR11*, leading to more than 10-fold increase of PRR11

transcript compared to the population average (Fig. 27C). While *PRR11* is a relatively understudied cell cycle progression gene with normally periodic expression, it is overexpressed in breast and lung cancer and PRR11 knockdown inhibits cancer cell proliferation (Ji et al., 2013; Zhou et al., 2014). A similar rearrangement and consequent increase in transcript levels occurred in a thyroid carcinoma, where *PAX8*, the 5' partner in the recurrent *PAX8-PPARG* thyroid fusion, is spliced to the 5' UTR of *GLIS1*, a transcription factor whose expression was previously correlated with Wnt pathway activation and epithelial to mesenchymal transition in a mouse model of breast cancer (Fig. 27D) (Kroll et al., 2000; Vadnais et al., 2014).

The non-canonical rearrangements identified using STA also include clinically relevant immune-modulatory genes. In a head and neck squamous cell carcinoma sample, we identified a hybrid transcript in which a non-coding portion of chromosome 16 is fused to the 5' UTR of the gene encoding the colony stimulating factor 1 receptor (CSF1R), leading to overexpression of a complete CSF1R-coding transcript (Fig. 27E) (Zhu et al., 2014). Additionally, a colorectal cancer sample with abnormally high expression of *CD274*, which encodes the T-cell suppressor PD-L1, harbors a rearrangement between the second-to-last exon of *CD274* and a non-coding region of chromosome 9 (Fig. 27F) (Iwai et al., 2002). The resulting hybrid transcript encodes a near-complete copy of PD-L1; as in the previous *IL6R-RPL29P7* rearrangement, we speculate that loss of the

endogenous negative-regulatory 3' UTR of CD274 leads to the increase in transcript levels.

Discussion

The wide spectrum of rearrangements identified using STA demonstrates the ability of oncogenic selection to exploit the modular architecture of the human genome to ensure tumor proliferation and survival. While some novel rearrangements emerging from our analysis consisted of classic receptor tyrosine kinase fusions, such as *TMEM87B-MERTK* in TNBC (Fig. 21), we also observed overexpression of entire coding transcripts resulting from promoter and UTR swapping, including oncogenic Ras family members (Fig. 27B).

Importantly, our ability to assess rearrangements with non-coding regions led to the detection of not only inactivating truncations in tumor suppressors, but also gain-of-function events. For instance, the gene encoding the tumor suppressor PD-L1, which has emerged as a prime target in the rapidly advancing field of cancer immunotherapy, underwent rearrangement and overexpression of a transcript harboring a non-coding sequence in place of its endogenous, negative-regulatory 3' UTR (Fig. 27F) (Pardoll, 2012).

We observed similar rearrangements multiple times in our analyses, including the *IL6R-RPL29P7* fusion in TNBC (Fig. 23D), and we hypothesize that overexpression by this mechanism may explain many of

the rearrangements occurring between oncogenes and non-coding regions of the genome (Fig. 26B). The *MALAT1-ALK* lncRNA-gene fusion and the increased expression of tumor-associated macrophage drug target CSF1R by a non-coding 5' hybrid transcript partner additionally demonstrate the clinical relevance of gene rearrangements involving non-coding regions (Fig. 27A,E) (Zhu et al., 2014). Further consideration of these non-canonical events in future gene rearrangement discovery will be critical for our understanding and treatment of TNBC and other molecularly heterogeneous cancers.

The targetable gene fusions we identified across many cancer types demonstrate the need to rapidly advance gene fusion detection for molecularly heterogeneous cancers. Although the frequency of individual gene rearrangements in these cancers may be low, the two rearrangements we validated with biological relevance for TNBC involve specific molecular targets for therapies already in clinical investigation (Shaw et al., 2013) or development (Schlegel et al., 2013). If readily detectable, several of the gene fusions we identified would provide an immediate opportunity for patient alignment to targeted therapy and would serve as biomarkers for patient selection in basket trials such as the NCI-MATCH (Molecular Analysis of Therapy Choice) Program.

Our results provide evidence for the effectiveness of STA as a quantitative prediction tool for both known and novel gene rearrangements. Although our analysis made use of RNA and DNA sequencing files already

processed by the TCGA Research Network, a pipeline based on focused assembly of STA-predicted transcripts could increase the sensitivity, efficiency, and breadth of rearrangement detection, accelerating discovery of diagnostic and prognostic markers. As tumor sequencing efforts continue, we look forward to optimizing and expanding the use of STA to comprehensively catalogue gene rearrangements across cancer.

The novel approach employed in STA enabled the discovery and analysis of known and novel gene rearrangements on a genome-wide scale, which has clear relevance to the development and repurposing of targeted therapies. Moreover, analysis of the rearrangements we identified provides insight into the multi-modular architecture of proteins and the diverse functional and regulatory domains selected for and against during tumorigenesis. Continued advances in detection methods such as STA will be critical for the treatment of diseases such as TNBC and the ability of the field to apply mechanistic insight from “exceptional responders” to individual tumor genomes containing unique mutations and rearrangements.

CHAPTER 5

CONCLUSIONS AND FUTURE DIRECTIONS

The work presented herein was conducted with the goal of expanding our knowledge of the molecular basis of triple-negative breast cancer and identifying or validating potentially targetable features for further basic or preclinical study. Numerous, comprehensive prior studies have focused on the rates of nonsynonymous mutations and copy number changes, revealing TNBC to be a highly heterogeneous collection of diseases with a wide spectrum of low-frequency alterations in targetable oncogenes (Banerji et al., 2012; Bianchini et al., 2016; Cancer Genome Atlas, 2012b; Curtis et al., 2012; Shah et al., 2012). With this in mind, we chose to focus on two features that do occur in a majority of TNBC: alterations in the p53 tumor suppressor and chromosomal instability.

Previous studies and reviews have identified mutant p53 as a potentially underappreciated oncogenic driver in TNBC (Walerych et al., 2012), and two groups have shown that loss of mutant p53 can induce cell cycle arrest or apoptosis specifically in triple-negative breast cancer cells (Braicu et al., 2013; Lim et al., 2009). To better understand the mechanism underlying the potential dependence of TNBC cells on mutant p53 for survival, we leveraged the then-nascent but now widely popular CRISPR/Cas genome editing methodology (Ran et al., 2013). In Chapter 3,

I described the engineering of an isogenic TNBC cell line panel expressing wild-type, missense mutant, and frameshift null p53 from the endogenous allele. Our goal was to enable comprehensive analysis of the biochemical and transcriptional impact of mutant p53 by comparing cell populations that differed only by individual nucleotides. In practice, we observed a rapid divergence at the phenotypic, transcriptional, and genomic levels that was driven by mutant p53. Our results provided unexpected insight regarding the variability of GOF phenotype acquisition, even in an isogenic context. To follow, I describe the implications of these findings and our ongoing efforts to leverage this unique model system to provide meaningful data for the p53 field.

Considering the extent of chromosomal instability and copy number changes present in many TNBCs (de Ruijter et al., 2011; Kwei et al., 2010), gene fusions and clinically actionable chromosomal rearrangements in general have been infrequently reported (Pareja et al., 2016). In Chapter 4, I presented our application of a new conceptual advance – quantitative rearrangement prediction by population-based, exon-level expression comparison – to take a ‘second look’ at the rapidly expanding volume of publicly available RNA and DNA sequencing datasets for TNBC and other cancers. In TNBC, we reported and biologically validated two kinase fusions that are targetable with inhibitors already in clinical investigation (Shaw et al., 2013) or development (Schlegel et al., 2013), demonstrating that advancing new capabilities for TNBC gene fusion detection has the

potential to expand clinicians' repertoire of targeted therapies for a disease where cytotoxic chemotherapy remains the standard of care (Isakoff, 2010). In this final chapter, I discuss possibilities for the future refinement and implementation of our Segmental Transcript Analysis algorithm and discuss the impact of the structurally and functionally diverse gene rearrangements we detected across TNBC and cancer at large.

Utility of the Mutant p53 Isogenic Model System

Our engineering of a mutant p53 isogenic cell line panel, reported in Chapters 2 and 3, successfully yielded a model system in which to conduct carefully controlled comparisons of numerous p53 genotypes. In particular, the ability to assess the impact of mutant p53 expressed in an endogenous mutant/null heterozygous genotype afforded confidence in the physiological relevance of our findings, especially as they relate to mutant protein regulation. Our isogenic cell line panel also represents an excellent model system to study numerous aspects of p53 biology outside the scope of our research; we look forward to sharing these cell lines with our colleagues in various biomedical research fields.

Owing to our use of 'first-generation' protocols and methodologies (Ran et al., 2013), CRISPR/Cas-mediated knock-in mutagenesis was somewhat inefficient, especially at the single-cell clone screening stage, where our frequency of identifying mutant versus frameshift alleles ranged from 1-10%. However, recent advances aimed at increasing the efficiency

of CRISPR/Cas-mediated homology-directed repair, including inhibition of non-homologous end joining and better modeling of recombination efficiency, should improve the feasibility of knock-in mutation for groups seeking to model alternate p53 alleles or other proteins in their model systems (Chu et al., 2015; Maruyama et al., 2015; Paquet et al., 2016).

New Insight into Mutant Gain of Function

Our isogenic cell line panel was successful in demonstrating the substantial oncogenic potential of mutant p53: despite differing only by a single codon in an otherwise identical cell background, our clonal R175H and R273H-expressing populations rapidly acquired phenotypes in keeping with previously reported mutant GOF (Brosh & Rotter, 2009; Muller & Vousden, 2014; Oren & Rotter, 2010; Soussi & Wiman, 2015). Our unique endogenous expression system also allowed us to confirm and build upon early findings in the field related to the stoichiometry required for mutant p53 to exert a dominant-negative effect over wild-type (Sun et al., 1993); in Figure 4 (Chapter 3, page 76), we demonstrated that heterozygous wild-type/missense mutant cell lines can stimulate the protein expression of p21 and MDM2 to a level similar to that of homozygous wild-type. This finding must be interpreted in a context that also has relevance to our modeling of mutant GOF: despite being expressed in an established tumor cell line, missense mutant p53 in our isogenic populations remains subject to MDM2-mediated negative regulation (Fig. 15). As such, this mutant protein

lacks the constitutive stabilization that is observed in commercial, tumor-derived cell lines grown under the same experimental conditions. While this may partially explain the lack of a dominant-negative effect in Figure 4, it also has important implications for the field's understanding of the biochemical state permitting acquisition of GOF phenotypes. Based in large part on observations in mutant mouse models, multiple groups have concluded that constitutive stabilization is a prerequisite for mutant p53 GOF (Alexandrova et al., 2017; Muller & Vousden, 2014; Suh et al., 2011; Terzian et al., 2008). Our findings in Chapter 3 provide direct, contravening evidence to this conclusion.

It should be noted that our model represents a rare exception: accumulation of mutant p53 protein in tumor cells has been noted since the earliest studies of p53 (Rotter et al., 1980), and is frequent in tumor specimens to the extent that p53 positivity by immunohistochemistry is a reliable surrogate for *TP53* missense mutation status (Yemelyanova et al., 2011). The fact that mutant destabilization is considered a prerequisite to GOF speaks to the prevalence of this phenotype, and therefore to a tremendous selection pressure for mutant stabilization during tumorigenesis, perhaps to enhance GOF that might rely on sequestration or overwhelming stoichiometry (Brosh & Rotter, 2009).

An additional key set of insights from our work in Chapter 3 relates to the variable acquisition of distinct GOF phenotypes among the missense mutant p53-expressing clones. In our initial isogenic cell line panel, one of

two R175H clonal populations acquired aneuploidy and an elevated metabolic rate (Fig. 6 and Fig. 9), but exhibited no difference in xenograft tumor growth compared to clonal lines expressing wild-type p53 protein (Fig. 8, page 81). In contrast, the other R175H clonal population displayed the ability to form consistently larger tumors in xenograft with an unaltered chromosomal count and no significant increase in metabolism. Even in this limited number of biological replicates, we can observe that GOF phenotypes are not universally reproducible, even in an identical cell background. This variability in acquiring GOF is especially well demonstrated by our assessment of chromosomal count in the larger validation set of isogenic cell lines: in Figure 14 (Chapter 3, page 91), we observed evidence for aneuploidy promoted by mutant p53 (and, to a lesser extent, p53 loss of function), but this GOF phenotype manifested itself as an increased predisposition rather than a uniform cause and effect. We can conclude that the presence of missense mutant p53 enables, but does not ensure, transition to an oncogenic cell state. This finding has relevance not only to our understanding of mutant p53 and the difficulty of replicating GOF phenotypes and mechanisms in differing cellular contexts, but also to the practicalities of controlling for clonal variation in similar experiments.

Ongoing Isogenic Analysis

We are actively pursuing the experimental opportunities afforded to us by our recently expanded set of isogenic cell lines. Our preliminary analysis in Figure 14 (Chapter 3, page 91) demonstrated that the acquisition of aneuploidy occurred in a majority, but not all, of the 10 new R273H-expressing isogenic populations; in ongoing experiments, we are assessing the additional GOF phenotypes reported in Chapter 3 in order to examine whether particular cellular states are acquired independently or in association with one another. Through this observational analysis, we will gain insight to potential mechanistic relationships for further study.

Additional ongoing studies include *in vitro* and xenograft clonal competition assays to examine the temporal and spatial growth dynamics of wild-type, mutant, and null isogenic clones when observed as a combined population. Given the Moll group's findings of an enrichment for sporadic p53 mutant protein-retaining cells at the periphery of mouse tumors after tamoxifen-induced p53 ablation (Alexandrova et al., 2015), we hypothesize that mutant cells will manifest a growth advantage, particularly in the nutrient-challenged xenograft setting. We also look forward to the opportunity to correlate relative growth advantage with aneuploidy status and other GOF phenotypes, as described above.

A final experimental priority is assessing the effect of mutant p53 protein loss on the various GOF phenotypes we have reported. Determining whether individual phenotypes persist after mutant protein loss

or revert to a ‘null-like’ state will yield insight into which phenotypes originate from persistent, mutant p53-permitted cell states versus an active transcriptional or biochemical role of the mutant protein. We are beginning our assessment of both shRNA-mediated knockdown and CRISPR-mediated knockout in each of our isogenic populations.

Non-Coding DNA and Regulatory Rearrangements

Some of the key contributions from our work in Chapter 4 include the identification of highly prevalent rearrangements between protein-coding gene loci and non-coding regions of the genome (Fig. 26 and Fig. 27), as well as rearrangements of 5’ and 3’ regulatory regions leading to the overexpression of clinically relevant oncogenes such as *HRAS* and *CD274/PD-L1* (Fig. 27B and F). These findings highlight the importance of considering a broad array of structural and functional categories when implementing gene rearrangement discovery for both preclinical and clinical applications.

As I described in Chapter 1, a number of existing gene fusion detection pipelines restrict potential gene rearrangement partners to protein-coding loci, which can be effective in reducing false positive detections resulting from homologous or repetitive regions of the genome while still enriching for known oncogenic fusions (Annala et al., 2013). However, our finding that non-coding regions of the genome can function as both 5’ and 3’ partners in rearrangements involving clinically relevant

genes suggests that a full understanding of rearrangement-driven gene dysregulation and truncation requires a more unbiased detection approach, which can be aided by tools such as STA.

Impact of Rearrangement Discovery for TNBC

Our results in Chapter 4 demonstrate that gene rearrangements are indeed prevalent in TNBC. As I cautioned in Chapter 1, any form of genomic instability has the ability to produce hybrid sequences detectable at the RNA or DNA level, and not all chimeric transcripts will represent actionable gene fusions. However, a number of the gene rearrangements we identified in TNBC are potentially functionally and clinically relevant (Fig. 23).

In a pair of specific examples, we identified a clinical sample and TNBC cell line harboring the recurrent *FGFR3-TACC3* kinase fusion (Fig. 22), which constituted the first report of this targetable gene fusion in TNBC. Leveraging both genetic and pharmacologic inhibition, we demonstrated oncogenic dependence of the SUM185PE cell line on the protein product of its endogenous *FGFR3-TACC3* rearrangement. We additionally identified and biologically validated a previously uncharacterized fusion in the protooncogene tyrosine kinase-encoding *MERTK* that confers constitutive oncogenic signaling and cell survival (Fig. 21). In addition to the *TMEM87B-MERTK* rearrangement in a TNBC tumor,

we noted similar gene fusions in lung squamous cell carcinoma, lung adenocarcinoma, bladder carcinoma, and cervical carcinoma.

Like almost all oncogenic alterations in TNBC, these gene fusions constituted low-frequency events (Cancer Genome Atlas, 2012b; Curtis et al., 2012; Shah et al., 2012); however, like *TMEM87B-MERTK*, *FGFR3-TACC3* fusions have been detected in multiple cancers, first in glioblastoma and bladder carcinoma and recently in non-small cell lung cancer, cervical carcinoma, and head and neck squamous cell carcinoma (Costa et al., 2016; Singh et al., 2012; Williams et al., 2013). FGFR inhibitors are in active clinical evaluation for alterations including gene amplification in addition to fusion, and recent studies have reported clinical benefit using an investigational FGFR inhibitor for patients whose tumors harbor an *FGFR3-TACC3* rearrangement (Carneiro et al., 2015; Shaw et al., 2013). As detection strategies for both gene rearrangements and oncogenic alterations at large become more tissue-agnostic and pathway-focused, it is likely that otherwise low-frequency events in individual tumor types – such as *TMEM87B-MERTK*, *FGFR3-TACC3*, and other rearrangements we detected in TNBC – may constitute a sizeable enough patient population to spur the development or repurposing of targeted therapeutics.

Future Refinement and Implementation of STA

Despite our success in identifying known and novel gene rearrangements across thousands of tumors, as described in Chapter 4, there are numerous potential opportunities to improve the predictive power and utility of STA. To maximize the performance of the algorithm during our initial optimization, we imposed two limitations on transcript analysis, as outlined in Chapter 2. Because STA functions by evaluating the Euclidean distance of exon-level expression segments, it performs best when multiple segments are present in order to leverage a population background and highlight outlier behavior. As such, we limited our analysis to transcripts containing more than four exons. Reducing or eliminating this restriction would be of benefit in expanding the spectrum of possible rearrangements to be assessed, especially given that long non-coding RNAs tend to have fewer exons than protein-coding genes (Iyer et al., 2015). Possible compromises include modification of the algorithm for transcripts containing one or two exons, such as evaluating outlier behavior on the basis of overall transcript levels rather than segment distance.

An additional restriction we implemented to improve the performance of the algorithm was the exclusion of any transcript where no sample reached an exon expression level of 10 RPKM. This approach is highly useful for reducing the confounding effect of technical artifacts in lowly expressed genes, where a small increase in RNA-seq reads mapped to a given nucleotide position might manifest themselves as a seemingly abrupt

increase in exon-level expression. However, there are undoubtedly some transcripts of potential mechanistic and clinical relevance that fall below an expression level of 10 RPKM; one approach to overcome this limitation might be a normalization strategy designed to minimize these artifacts at the low end of the expression spectrum, similar to the regularized log transformation employed in the DESeq2 differential gene expression analysis utility (Love et al., 2014).

As I described in Chapters 2 and 4, our implementation of STA involved its use as a prediction tool upstream of traditional RNA-seq detection pipelines. This approach was chosen in large part because of the existence of RNA-seq alignment and fusion detection files (Wang et al., 2010) that had already been generated by TCGA research groups, facilitating the development of an automated data retrieval and analysis pipeline. However, from a predictive perspective, this is not the ideal implementation of an approach like STA; instead, the best approach would likely be a multiple signal integration strategy in which STA is employed concurrently with traditional, single-sample sequence-based fusion detection. A similar approach was successfully demonstrated by the Hall lab to improve sensitivity and potential breadth of discovery in their LUMPY framework for structural variant detection (Layer et al., 2014).

Despite opportunities for improvement, as I reported in Chapter 4, STA proved to be a powerful discovery tool in TCGA datasets, facilitating the identification of alterations that were undetected or unreported by other

comprehensive analyses. Future research directions may include application of the STA discovery pipeline in its current form to other publicly available datasets, such as the TARGET pediatric cancer sequencing initiative (Mullighan et al., 2009; Pugh et al., 2013).

In conclusion, the research presented herein has yielded both new insights into targetable features of TNBC and new avenues for research in the ever more complex fields of p53 biology and cancer genomics. It is our hope that the impact of this work will include not only our specific research findings, but also valuable new tools in the form of our isogenic mutant p53 cell line panel and our novel, population-based approach to gene rearrangement prediction and discovery.

REFERENCES

Abkevich, V., Timms, K. M., Hennessy, B. T., Potter, J., Carey, M. S., Meyer, L. A., … Lanchbury, J. S. (2012). Patterns of genomic loss of heterozygosity predict homologous recombination repair defects in epithelial ovarian cancer. *British Journal of Cancer*, *107*(10), 1776–1782. https://doi.org/10.1038/bjc.2012.451

Adorno, M., Cordenonsi, M., Montagner, M., Dupont, S., Wong, C., Hann, B., … Piccolo, S. (2009). A Mutant-p53/Smad Complex Opposes p63 to Empower TGFβ-Induced Metastasis. *Cell*, *137*(1), 87–98. https://doi.org/10.1016/j.cell.2009.01.039

Agarwal, M. L., Agarwal, A., Taylor, W. R., & Stark, G. R. (1995). p53 controls both the G2/M and the G1 cell cycle checkpoints and mediates reversible growth arrest in human fibroblasts. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(18), 8493–7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7667317

Alcalay, M., Zangrilli, D., Pandolfi, P. P., Longo, L., Mencarelli, A., Giacomucci, A., … Lo Coco, F. (1991). Translocation breakpoint of acute promyelocytic leukemia lies within the retinoic acid receptor alpha locus. *Proceedings of the National Academy of Sciences of the United States of America*, *88*(5), 1977–81. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1848017

Alexandrova, E. M., Mirza, S. A., Xu, S., Schulz-Heddergott, R., Marchenko, N. D., & Moll, U. M. (2017). p53 loss-of-heterozygosity is a necessary prerequisite for mutant p53 stabilization and gain-of-function in vivo. *Cell Death and Disease*, *8*(3), e2661. https://doi.org/10.1038/cddis.2017.80

Alexandrova, E. M., Yallowitz, A. R., Li, D., Xu, S., Schulz, R., Proia, D. A., … Moll, U. M. (2015). Improving survival by exploiting tumour dependence on stabilized mutant p53 for treatment. *Nature*, *523*(7560), 352–356. https://doi.org/10.1038/nature14430

Ali, H. R., Glont, S.-E., Blows, F. M., Provenzano, E., Dawson, S.-J., Liu, B., … Caldas, C. (2015). PD-L1 protein expression in breast cancer is rare, enriched in basal-like tumours and associated with infiltrating

lymphocytes. *Annals of Oncology*, *26*(7), 1488–1493. https://doi.org/10.1093/annonc/mdv192

Annala, M. J., Parker, B. C., Zhang, W., & Nykter, M. (2013). Fusion genes and their discovery using high throughput sequencing. *Cancer Lett*, *340*(2), 192–200. https://doi.org/10.1016/j.canlet.2013.01.011

Antoniou, A., Pharoah, P. D. P., Narod, S., Risch, H. A., Eyfjord, J. E., Hopper, J. L., … Easton, D. F. (2003). Average Risks of Breast and Ovarian Cancer Associated with BRCA1 or BRCA2 Mutations Detected in Case Series Unselected for Family History: A Combined Analysis of 22 Studies. *The American Journal of Human Genetics*, *72*(5), 1117–1130. https://doi.org/10.1086/375033

Arteaga, C. L. (2003). Trastuzumab, an appropriate first-line single-agent therapy for HER2-overexpressing metastatic breast cancer. *Breast Cancer Research : BCR*, *5*(2), 96–100. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12631388

Asmann, Y. W., Necela, B. M., Kalari, K. R., Hossain, A., Baker, T. R., Carr, J. M., … Thompson, E. A. (2012). Detection of redundant fusion transcripts as biomarkers or disease-specific therapeutic targets in breast cancer. *Cancer Res*, *72*(8), 1921–1928. https://doi.org/10.1158/0008-5472.CAN-11-3142

Atchley, D. P., Albarracin, C. T., Lopez, A., Valero, V., Amos, C. I., Gonzalez-Angulo, A. M., … Arun, B. K. (2008). Clinical and Pathologic Characteristics of Patients With *BRCA* -Positive and *BRCA* -Negative Breast Cancer. *Journal of Clinical Oncology*, *26*(26), 4282–4288. https://doi.org/10.1200/JCO.2008.16.6231

Baker, S. J., Fearon, E. R., Nigro, J. M., Hamilton, S. R., Preisinger, A. C., Jessup, J. M., … Vogelstein, B. (1989). Chromosome 17 deletions and p53 gene mutations in colorectal carcinomas. *Science (New York, N.Y.)*, *244*(4901), 217–21. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2649981

Baker, S. J., Markowitz, S., Fearon, E. R., Willson, J. K., & Vogelstein, B. (1990). Suppression of human colorectal carcinoma cell growth by wild-type p53. *Science (New York, N.Y.)*, *249*(4971), 912–5. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2144057

Balko, J. M., Giltnane, J. M., Wang, K., Schwarz, L. J., Young, C. D., Cook, R. S., … Arteaga, C. L. (2014). Molecular Profiling of the Residual Disease of Triple-Negative Breast Cancers after Neoadjuvant Chemotherapy Identifies Actionable Therapeutic Targets. *Cancer Discovery*, *4*(2), 232–245. https://doi.org/10.1158/2159-8290.CD-13-0286

Banerji, S., Cibulskis, K., Rangel-Escareno, C., Brown, K. K., Carter, S. L., Frederick, A. M., … Meyerson, M. (2012). Sequence analysis of mutations and translocations across breast cancer subtypes. *Nature*, *486*(7403), 405–409. https://doi.org/10.1038/nature11154

Barak, Y., Juven, T., Haffner, R., & Oren, M. (1993). mdm2 expression is induced by wild type p53 activity. *The EMBO Journal*, *12*(2), 461–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8440237

Bargonetti, J., Friedman, P. N., Kern, S. E., Vogelstein, B., & Prives, C. (1991). Wild-type but not mutant p53 immunopurified proteins bind to sequences adjacent to the SV40 origin of replication. *Cell*, *65*(6), 1083–91. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1646078

Bartkova, J., Ho?ej??, Z., Koed, K., Kr?mer, A., Tort, F., Zieger, K., … Bartek, J. (2005). DNA damage response as a candidate anti-cancer barrier in early human tumorigenesis. *Nature*, *434*(7035), 864–870. https://doi.org/10.1038/nature03482

Baselga, J., & Swain, S. M. (2009). Novel anticancer targets: revisiting ERBB2 and discovering ERBB3. *Nature Reviews. Cancer*, *9*(7), 463–75. https://doi.org/10.1038/nrc2656

Baselga, J., Tripathy, D., Mendelsohn, J., Baughman, S., Benz, C. C., Dantis, L., … Norton, L. (1996). Phase II study of weekly intravenous recombinant humanized anti-p185HER2 monoclonal antibody in patients with HER2/neu-overexpressing metastatic breast cancer. *Journal of Clinical Oncology*, *14*(3), 737–744. https://doi.org/10.1200/JCO.1996.14.3.737

Bastien, R. R., Rodriguez-Lescure, L., Ebbert, M. T., Prat, A., Mun?rriz, B., Rowe, L., … Mart?n, M. (2012). PAM50 Breast Cancer Subtyping by RT-qPCR and Concordance with Standard Clinical Molecular Markers. *BMC Medical Genomics*, *5*(1), 44. https://doi.org/10.1186/1755-8794-

5-44

Bensaad, K., Tsuruta, A., Selak, M. A., Vidal, M. N. C., Nakano, K., Bartrons, R., … Vousden, K. H. (2006). TIGAR, a p53-Inducible Regulator of Glycolysis and Apoptosis. *Cell*, *126*(1), 107–120. https://doi.org/10.1016/j.cell.2006.05.036

Bertucci, F., Finetti, P., Cervera, N., Esterni, B., Hermitte, F., Viens, P., & Birnbaum, D. (2008). How basal are triple-negative breast cancers? *International Journal of Cancer*, *123*(1), 236–240. https://doi.org/10.1002/ijc.23518

Bianchini, G., Balko, J. M., Mayer, I. A., Sanders, M. E., & Gianni, L. (2016). Triple-negative breast cancer: challenges and opportunities of a heterogeneous disease. *Nature Reviews Clinical Oncology*, *13*(11), 674–690. https://doi.org/10.1038/nrclinonc.2016.66

Birkbak, N. J., Wang, Z. C., Kim, J.-Y., Eklund, A. C., Li, Q., Tian, R., … Richardson, A. L. (2012). Telomeric Allelic Imbalance Indicates Defective DNA Repair and Sensitivity to DNA-Damaging Agents. *Cancer Discovery*, *2*(4), 366–375. https://doi.org/10.1158/2159-8290.CD-11-0206

Boeckler, F. M., Joerger, A. C., Jaggi, G., Rutherford, T. J., Veprintsev, D. B., & Fersht, A. R. (2008). Targeted rescue of a destabilized mutant of p53 by an in silico screened drug. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(30), 10360–5. https://doi.org/10.1073/pnas.0805326105

Borrow, J., Goddard, A. D., Sheer, D., & Solomon, E. (1990). Molecular analysis of acute promyelocytic leukemia breakpoint cluster region on chromosome 17. *Science (New York, N.Y.)*, *249*(4976), 1577–80. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2218500

Bouaoun, L., Sonkin, D., Ardin, M., Hollstein, M., Byrnes, G., Zavadil, J., & Olivier, M. (2016). TP53 Variations in Human Cancers: New Lessons from the IARC TP53 Database and Genomics Data. *Human Mutation*, *37*(9), 865–76. https://doi.org/10.1002/humu.23035

Bougeard, G., Sesboüé, R., Baert-Desurmont, S., Vasseur, S., Martin, C., Tinat, J., … Frébourg, T. (2008). Molecular basis of the Li-Fraumeni syndrome: an update from the French LFS families. *Journal of Medical*

Genetics, *45*(8), 535–8. https://doi.org/10.1136/jmg.2008.057570

Braicu, C., Pileczki, V., Irimie, A., & Berindan-Neagoe, I. (2013). p53siRNA therapy reduces cell proliferation, migration and induces apoptosis in triple negative breast cancer cells. *Molecular and Cellular Biochemistry*, *381*(1–2), 61–68. https://doi.org/10.1007/s11010-013-1688-5

Brandao, L. N., Winges, A., Christoph, S., Sather, S., Migdall-Wilson, J., Schlegel, J., … Graham, D. K. (2013). Inhibition of MerTK increases chemosensitivity and decreases oncogenic potential in T-cell acute lymphoblastic leukemia. *Blood Cancer J*, *3*, e101. https://doi.org/10.1038/bcj.2012.46

Brasoveanu, L. I., Fonsatti, E., Visintin, A., Pavlovic, M., Cattarossi, I., Colizzi, F., … Maio, M. (1997). Melanoma cells constitutively release an anchor-positive soluble form of protectin (sCD59) that retains functional activities in homologous complement-mediated cytotoxicity. *J Clin Invest*, *100*(5), 1248–1255. https://doi.org/10.1172/JCI119638

Brosh, R., & Rotter, V. (2009). When mutants gain new powers: news from the mutant p53 field. *Nature Reviews Cancer*, *9*(10), 701–13. https://doi.org/10.1038/nrc2693

Brown, S. D., Warren, R. L., Gibb, E. A., Martin, S. D., Spinelli, J. J., Nelson, B. H., & Holt, R. A. (2014). Neo-antigens predicted by tumor genome meta-analysis correlate with increased patient survival. *Genome Research*, *24*(5), 743–750. https://doi.org/10.1101/gr.165985.113

Cancer Genome Atlas, N. (2012a). Comprehensive molecular characterization of human colon and rectal cancer. *Nature*, *487*(7407), 330–337. https://doi.org/10.1038/nature11252

Cancer Genome Atlas, N. (2012b). Comprehensive molecular portraits of human breast tumours. *Nature*, *490*(7418), 61–70. https://doi.org/10.1038/nature11412

Cancer Genome Atlas Research, N. (2011). Integrated genomic analyses of ovarian carcinoma. *Nature*, *474*(7353), 609–615. https://doi.org/10.1038/nature10166

Cancer Genome Atlas Research, N. (2012). Comprehensive genomic characterization of squamous cell lung cancers. *Nature*, *489*(7417), 519–525. https://doi.org/10.1038/nature11404

Canman, C. E., Lim, D. S., Cimprich, K. A., Taya, Y., Tamai, K., Sakaguchi, K., … Siliciano, J. D. (1998). Activation of the ATM kinase by ionizing radiation and phosphorylation of p53. *Science (New York, N.Y.)*, *281*(5383), 1677–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9733515

Carmeci, C., deConinck, E. C., Lawton, T., Bloch, D. A., & Weigel, R. J. (1997). Analysis of estrogen receptor messenger RNA in breast carcinomas from archival specimens is predictive of tumor biology. *Am J Pathol*, *150*(5), 1563–1570. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9137083

Carneiro, B. A., Elvin, J. A., Kamath, S. D., Ali, S. M., Paintal, A. S., Restrepo, A., … Johnson, M. L. (2015). FGFR3-TACC3: A novel gene fusion in cervical cancer. *Gynecologic Oncology Reports*, *13*, 53–56. https://doi.org/10.1016/j.gore.2015.06.005

Carter, P., Presta, L., Gorman, C. M., Ridgway, J. B., Henner, D., Wong, W. L., … Shepard, H. M. (1992). Humanization of an anti-p185HER2 antibody for human cancer therapy. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(10), 4285–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1350088

Caspersson, T., Zech, L., & Johansson, C. (1970). Differential binding of alkylating fluorochromes in human chromosomes. *Experimental Cell Research*, *60*(3), 315–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/5422961

Chan, W. M., Siu, W. Y., Lau, A., & Poon, R. Y. C. (2004). How many mutant p53 molecules are needed to inactivate a tetramer? *Molecular and Cellular Biology*, *24*(8), 3536–51. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15060172

Chase, A., Ernst, T., Fiebig, A., Collins, A., Grand, F., Erben, P., … Cross, N. C. (2010). TFG, a target of chromosome translocations in lymphoma and soft tissue tumors, fuses to GPR128 in healthy individuals. *Haematologica*, *95*(1), 20–26. https://doi.org/10.3324/haematol.2009.011536

Chehab, N. H., Malikzay, A., Appel, M., & Halazonetis, T. D. (2000). Chk2/hCds1 functions as a DNA damage checkpoint in G(1) by stabilizing p53. *Genes & Development*, *14*(3), 278–88. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10673500

Chen, L., Gilkes, D. M., Pan, Y., Lane, W. S., & Chen, J. (2005). ATM and Chk2-dependent phosphorylation of MDMX contribute to p53 activation after DNA damage. *The EMBO Journal*, *24*(19), 3411–3422. https://doi.org/10.1038/sj.emboj.7600812

Chen, S., & Parmigiani, G. (2007). Meta-Analysis of *BRCA1* and *BRCA2* Penetrance. *Journal of Clinical Oncology*, *25*(11), 1329–1333. https://doi.org/10.1200/JCO.2006.09.1066

Chène, P. (2001). The role of tetramerization in p53 function. *Oncogene*, *20*(21), 2611–2617. https://doi.org/10.1038/sj.onc.1204373

Cho, Y., Gorina, S., Jeffrey, P. D., & Pavletich, N. P. (1994). Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations. *Science (New York, N.Y.)*, *265*(5170), 346–55. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8023157

Chu, V. T., Weber, T., Wefers, B., Wurst, W., Sander, S., Rajewsky, K., & Kühn, R. (2015). Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nature Biotechnology*, *33*(5), 543–548. https://doi.org/10.1038/nbt.3198

Chumakov, P. M., Iotsova, V. S., & Georgiev, G. P. (1982). [Isolation of a plasmid clone containing the mRNA sequence for mouse nonviral T-antigen]. *Doklady Akademii Nauk SSSR*, *267*(5), 1272–5. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6295732

Chumsri, S., Howes, T., Bao, T., Sabnis, G., & Brodie, A. (2011). Aromatase, aromatase inhibitors, and breast cancer. *The Journal of Steroid Biochemistry and Molecular Biology*, *125*(1–2), 13–22. https://doi.org/10.1016/j.jsbmb.2011.02.001

Cobleigh, M. A., Vogel, C. L., Tripathy, D., Robert, N. J., Scholl, S., Fehrenbacher, L., … Slamon, D. J. (1999). Multinational Study of the Efficacy and Safety of Humanized Anti-HER2 Monoclonal Antibody in Women Who Have HER2-Overexpressing Metastatic Breast Cancer

That Has Progressed After Chemotherapy for Metastatic Disease. *Journal of Clinical Oncology*, *17*(9), 2639–2639. https://doi.org/10.1200/JCO.1999.17.9.2639

Cole, M. P., Jones, C. T., & Todd, I. D. (1971). A new anti-oestrogenic agent in late breast cancer. An early clinical appraisal of ICI46474. *British Journal of Cancer*, *25*(2), 270–5. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/5115829

Costa, R., Carneiro, B. A., Taxter, T., Tavora, F. A., Kalyan, A., Pai, S. A., … Giles, F. J. (2016). FGFR3-TACC3 fusion in solid tumors: mini review. *Oncotarget*, *7*(34), 55924–55938. https://doi.org/10.18632/oncotarget.10482

Cummings, C. T., Deryckere, D., Earp, H. S., & Graham, D. K. (2013). Molecular pathways: MERTK signaling in cancer. *Clin Cancer Res*, *19*(19), 5275–5280. https://doi.org/10.1158/1078-0432.CCR-12-1451

Curtis, C., Shah, S. P., Chin, S.-F., Turashvili, G., Rueda, O. M., Dunning, M. J., … Aparicio, S. (2012). The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature*, *486*(7403), 346–52. https://doi.org/10.1038/nature10983

Cuzick, J., & Baum, M. (1985). Tamoxifen and contralateral breast cancer. *Lancet (London, England)*, *2*(8449), 282. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2862460

Dai, X., Li, T., Bai, Z., Yang, Y., Liu, X., Zhan, J., & Shi, B. (2015). Breast cancer intrinsic subtype classification, clinical use and future trends. *American Journal of Cancer Research*, *5*(10), 2929–43. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/26693050

Davare, M. A., & Tognon, C. E. (2015). Detecting and targetting oncogenic fusion proteins in the genomic era. *Biol Cell*, *107*(5), 111–129. https://doi.org/10.1111/boc.201400096

Davis, T. A., & Johnston, J. N. (2011). Catalytic, enantioselective synthesis of stilbene cis-diamines: A concise preparation of (−)-Nutlin-3, a potent p53/MDM2 inhibitor. *Chemical Science*, *2*(6), 1076. https://doi.org/10.1039/c1sc00061f

Davoli, T., Xu, A. W., Mengwasser, K. E., Sack, L. M., Yoon, J. C., Park, P. J., & Elledge, S. J. (2013). Cumulative haploinsufficiency and triplosensitivity drive aneuploidy patterns and shape the cancer genome. *Cell*, *155*(4), 948–962. https://doi.org/10.1016/j.cell.2013.10.011

de Klein, A., van Kessel, A. G., Grosveld, G., Bartram, C. R., Hagemeijer, A., Bootsma, D., … Stephenson, J. R. (1982). A cellular oncogene is translocated to the Philadelphia chromosome in chronic myelocytic leukaemia. *Nature*, *300*(5894), 765–7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6960256

de Ruijter, T. C., Veeck, J., de Hoon, J. P. J., van Engeland, M., & Tjan-Heijnen, V. C. (2011). Characteristics of triple-negative breast cancer. *Journal of Cancer Research and Clinical Oncology*, *137*(2), 183–92. https://doi.org/10.1007/s00432-010-0957-x

de Thé, H., Chomienne, C., Lanotte, M., Degos, L., & Dejean, A. (1990). The t(15;17) translocation of acute promyelocytic leukaemia fuses the retinoic acid receptor ? gene to a novel transcribed locus. *Nature*, *347*(6293), 558–561. https://doi.org/10.1038/347558a0

DeLeo, A. B., Jay, G., Appella, E., Dubois, G. C., Law, L. W., & Old, L. J. (1979). Detection of a transformation-related antigen in chemically induced sarcomas and other transformed cells of the mouse. *Proceedings of the National Academy of Sciences of the United States of America*, *76*(5), 2420–4. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/221923

Dent, R., Trudeau, M., Pritchard, K. I., Hanna, W. M., Kahn, H. K., Sawka, C. A., … Narod, S. A. (2007). Triple-Negative Breast Cancer: Clinical Features and Patterns of Recurrence. *Clinical Cancer Research*, *13*(15), 4429–4434. https://doi.org/10.1158/1078-0432.CCR-06-3045

DeSantis, C. E., Fedewa, S. A., Goding Sauer, A., Kramer, J. L., Smith, R. A., & Jemal, A. (2016). Breast cancer statistics, 2015: Convergence of incidence rates between black and white women. *CA: A Cancer Journal for Clinicians*, *66*(1), 31–42. https://doi.org/10.3322/caac.21320

Di Agostino, S., Cortese, G., Monti, O., Dell'Orso, S., Sacchi, A., Eisenstein, M., … Blandino, G. (2008). The disruption of the protein

complex mutantp53/p73 increases selectively the response of tumor cells to anticancer drugs. *Cell Cycle*, *7*(21), 3440–3447. https://doi.org/10.4161/cc.7.21.6995

Di Como, C. J., Gaiddon, C., & Prives, C. (1999). p73 function is inhibited by tumor-derived p53 mutants in mammalian cells. *Molecular and Cellular Biology*, *19*(2), 1438–49. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9891077

Di Fiore, P. P., Pierce, J. H., Kraus, M. H., Segatto, O., King, C. R., & Aaronson, S. A. (1987). erbB-2 is a potent oncogene when overexpressed in NIH/3T3 cells. *Science (New York, N.Y.)*, *237*(4811), 178–82. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2885917

Diller, L., Kassel, J., Nelson, C. E., Gryka, M. A., Litwak, G., Gebhardt, M., … Vogelstein, B. (1990). p53 functions as a cell cycle control protein in osteosarcomas. *Molecular and Cellular Biology*, *10*(11), 5772–81. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2233717

Dittmer, D., Pati, S., Zambetti, G., Chu, S., Teresky, A. K., Moore, M., … Levine, A. J. (1993). Gain of function mutations in p53. *Nature Genetics*, *4*(1), 42–46. https://doi.org/10.1038/ng0593-42

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., … Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, *29*(1), 15–21. https://doi.org/10.1093/bioinformatics/bts635

Dodt, M., Roehr, J. T., Ahmed, R., & Dieterich, C. (2012). FLEXBAR-Flexible Barcode and Adapter Processing for Next-Generation Sequencing Platforms. *Biology (Basel)*, *1*(3), 895–905. https://doi.org/10.3390/biology1030895

Donehower, L. A., Harvey, M., Slagle, B. L., McArthur, M. J., Montgomery, C. A., Butel, J. S., & Bradley, A. (1992). Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. *Nature*, *356*(6366), 215–221. https://doi.org/10.1038/356215a0

Donehower, L. A., & Lozano, G. (2009). 20 years studying p53 functions in genetically engineered mice. *Nature Reviews Cancer*, *9*(11), 831–841. https://doi.org/10.1038/nrc2731

Druker, B. J., Talpaz, M., Resta, D. J., Peng, B., Buchdunger, E., Ford, J. M., … Sawyers, C. L. (2001). Efficacy and Safety of a Specific Inhibitor of the BCR-ABL Tyrosine Kinase in Chronic Myeloid Leukemia. *New England Journal of Medicine*, *344*(14), 1031–1037. https://doi.org/10.1056/NEJM200104053441401

Early Breast Cancer Trialists' Collaborative Group. (1988). Effects of Adjuvant Tamoxifen and of Cytotoxic Therapy on Mortality in Early Breast Cancer. *New England Journal of Medicine*, *319*(26), 1681–1692. https://doi.org/10.1056/NEJM198812293192601

Early Breast Cancer Trialists' Collaborative Group. (1998). Tamoxifen for early breast cancer: an overview of the randomised trials. Early Breast Cancer Trialists' Collaborative Group. *Lancet (London, England)*, *351*(9114), 1451–67. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9605801

Edgren, H., Murumagi, A., Kangaspeska, S., Nicorici, D., Hongisto, V., Kleivi, K., … Kallioniemi, O. (2011). Identification of fusion genes in breast cancer by paired-end RNA-sequencing. *Genome Biol*, *12*(1), R6. https://doi.org/10.1186/gb-2011-12-1-r6

El-Deiry, W. S., Kern, S. E., Pietenpol, J. A., Kinzler, K. W., & Vogelstein, B. (1992). Definition of a consensus binding site for p53. *Nature Genetics*, *1*(1), 45–49. https://doi.org/10.1038/ng0492-45

el-Deiry, W. S., Tokino, T., Velculescu, V. E., Levy, D. B., Parsons, R., Trent, J. M., … Vogelstein, B. (1993). WAF1, a potential mediator of p53 tumor suppression. *Cell*, *75*(4), 817–25. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8242752

Eliyahu, D., Goldfinger, N., Pinhasi-Kimhi, O., Shaulsky, G., Skurnik, Y., Arai, N., … Oren, M. (1988). Meth A fibrosarcoma cells express two transforming mutant p53 species. *Oncogene*, *3*(3), 313–21. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3060794

Eliyahu, D., Michalovitz, D., Eliyahu, S., Pinhasi-Kimhi, O., & Oren, M. (1989). Wild-type p53 can inhibit oncogene-mediated focus formation. *Proceedings of the National Academy of Sciences of the United States of America*, *86*(22), 8763–7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2530586

Eliyahu, D., Michalovitz, D., & Oren, M. (1985). Overproduction of p53 antigen makes established cells highly tumorigenic. *Nature*, *316*(6024), 158–60. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3892307

Eliyahu, D., Raz, A., Gruss, P., Givol, D., & Oren, M. (1984). Participation of p53 cellular tumour antigen in transformation of normal embryonic cells. *Nature*, *312*(5995), 646–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6095116

Eroles, P., Bosch, A., Alejandro Pérez-Fidalgo, J., & Lluch, A. (2012). Molecular biology in breast cancer: Intrinsic subtypes and signaling pathways. *Cancer Treatment Reviews*, *38*(6), 698–707. https://doi.org/10.1016/j.ctrv.2011.11.005

Farmer, G., Bargonetti, J., Zhu, H., Friedman, P., Prywes, R., & Prives, C. (1992). Wild-type p53 activates transcription in vitro. *Nature*, *358*(6381), 83–86. https://doi.org/10.1038/358083a0

Faust, G. G., & Hall, I. M. (2012). YAHA: fast and flexible long-read alignment with optimal breakpoint detection. *Bioinformatics*, *28*(19), 2417–2424. https://doi.org/10.1093/bioinformatics/bts456

Fields, S., & Jang, S. K. (1990). Presence of a potent transcription activating sequence in the p53 protein. *Science (New York, N.Y.)*, *249*(4972), 1046–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2144363

Finlay, C. A., Hinds, P. W., & Levine, A. J. (1989). The p53 proto-oncogene can act as a suppressor of transformation. *Cell*, *57*(7), 1083–93. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2525423

Finlay, C. A., Hinds, P. W., Tan, T. H., Eliyahu, D., Oren, M., & Levine, A. J. (1988). Activating mutations for transformation by p53 produce a gene product that forms an hsc70-p53 complex with an altered half-life. *Molecular and Cellular Biology*, *8*(2), 531–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2832726

Fisher, B., Costantino, J. P., Wickerham, D. L., Cecchini, R. S., Cronin, W. M., Robidoux, A., … Wolmark, N. (2005). Tamoxifen for the Prevention of Breast Cancer: Current Status of the National Surgical Adjuvant Breast and Bowel Project P-1 Study. *JNCI Journal of the National Cancer Institute*, *97*(22), 1652–1662. https://doi.org/10.1093/jnci/dji372

Fisher, B., Costantino, J. P., Wickerham, D. L., Redmond, C. K., Kavanah, M., Cronin, W. M., … Wolmark, N. (1998). Tamoxifen for prevention of breast cancer: report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study. *Journal of the National Cancer Institute*, *90*(18), 1371–88. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9747868

Flores, I., & Blasco, M. A. (2009). A p53-Dependent Response Limits Epidermal Stem Cell Functionality and Organismal Size in Mice with Short Telomeres. *PLoS ONE*, *4*(3), e4934. https://doi.org/10.1371/journal.pone.0004934

Folca, P. J., Glascock, R. F., & Irvine, W. T. (1961). Studies with tritium-labelled hexoestrol in advanced breast cancer. Comparison of tissue accumulation of hexoestrol with response to bilateral adrenalectomy and oophorectomy. *Lancet (London, England)*, *2*(7206), 796–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/13893792

Freed-Pastor, W. A., Mizuno, H., Zhao, X., Langerød, A., Moon, S.-H., Rodriguez-Barrueco, R., … Prives, C. (2012). Mutant p53 Disrupts Mammary Tissue Architecture via the Mevalonate Pathway. *Cell*, *148*(1–2), 244–258. https://doi.org/10.1016/j.cell.2011.12.017

Funk, W. D., Pak, D. T., Karas, R. H., Wright, W. E., & Shay, J. W. (1992). A transcriptionally active DNA-binding site for human p53 protein complexes. *Molecular and Cellular Biology*, *12*(6), 2866–71. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1588974

Gaiddon, C., Lokshin, M., Ahn, J., Zhang, T., & Prives, C. (2001). A Subset of Tumor-Derived Mutant Forms of p53 Down-Regulate p63 and p73 through a Direct Interaction with the p53 Core Domain. *Molecular and Cellular Biology*, *21*(5), 1874–1887. https://doi.org/10.1128/MCB.21.5.1874-1887.2001

Gayther, S. A., Batley, S. J., Linger, L., Bannister, A., Thorpe, K., Chin, S. F., … Caldas, C. (2000). Mutations truncating the EP300 acetylase in human cancers. *Nat Genet*, *24*(3), 300–303. https://doi.org/10.1038/73536

Gelmon, K. A., Tischkowitz, M., Mackay, H., Swenerton, K., Robidoux, A., Tonkin, K., … Oza, A. (2011). Olaparib in patients with recurrent high-grade serous or poorly differentiated ovarian carcinoma or triple-

negative breast cancer: a phase 2, multicentre, open-label, non-randomised study. *The Lancet Oncology*, *12*(9), 852–861. https://doi.org/10.1016/S1470-2045(11)70214-5

Giacomini, C. P., Sun, S., Varma, S., Shain, A. H., Giacomini, M. M., Balagtas, J., … Pollack, J. R. (2013). Breakpoint analysis of transcriptional and genomic profiles uncovers novel gene fusions spanning multiple human cancer types. *PLoS Genet*, *9*(4), e1003464. https://doi.org/10.1371/journal.pgen.1003464

Gioanni, J., Le François, D., Zanghellini, E., Mazeau, C., Ettore, F., Lambert, J. C., … Dutrillaux, B. (1990). Establishment and characterisation of a new tumorigenic cell line with a normal karyotype derived from a human breast adenocarcinoma. *British Journal of Cancer*, *62*(1), 8–13. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2390488

Glascock, R. F., & Hoekstra, W. G. (1959). Selective accumulation of tritium-labelled hexoestrol by the reproductive organs of immature female goats and sheep. *The Biochemical Journal*, *72*, 673–82. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/13828338

Gonzalez-Angulo, A. M., Stemke-Hale, K., Palla, S. L., Carey, M., Agarwal, R., Meric-Berstam, F., … Hennessy, B. T. (2009). Androgen Receptor Levels and Association with PIK3CA Mutations and Prognosis in Breast Cancer. *Clinical Cancer Research*, *15*(7), 2472–2478. https://doi.org/10.1158/1078-0432.CCR-08-1763

Gorgoulis, V. G., Vassiliou, L.-V. F., Karakaidos, P., Zacharatos, P., Kotsinas, A., Liloglou, T., … Halazonetis, T. D. (2005). Activation of the DNA damage checkpoint and genomic instability in human precancerous lesions. *Nature*, *434*(7035), 907–913. https://doi.org/10.1038/nature03485

Graham, D. K., Dawson, T. L., Mullaney, D. L., Snodgrass, H. R., & Earp, H. S. (1994). Cloning and mRNA expression analysis of a novel human protooncogene, c-mer. *Cell Growth Differ*, *5*(6), 647–657. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8086340

Green, D. R., & Kroemer, G. (2009). Cytoplasmic functions of the tumour suppressor p53. *Nature*, *458*(7242), 1127–1130. https://doi.org/10.1038/nature07986

Grieco, M., Santoro, M., Berlingieri, M. T., Melillo, R. M., Donghi, R., Bongarzone, I., … Vecchio, G. (1990). PTC is a novel rearranged form of the ret proto-oncogene and is frequently detected in vivo in human thyroid papillary carcinomas. *Cell*, *60*(4), 557–563. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2406025

Gualberto, A., Aldape, K., Kozakiewicz, K., & Tlsty, T. D. (1998). An oncogenic form of p53 confers a dominant, gain-of-function phenotype that disrupts spindle checkpoint control. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(9), 5166–71. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9560247

Gucalp, A., Tolaney, S., Isakoff, S. J., Ingle, J. N., Liu, M. C., Carey, L. A., … Translational Breast Cancer Research Consortium (TBCRC 011). (2013). Phase II Trial of Bicalutamide in Patients with Androgen Receptor-Positive, Estrogen Receptor-Negative Metastatic Breast Cancer. *Clinical Cancer Research*, *19*(19), 5505–5512. https://doi.org/10.1158/1078-0432.CCR-12-3327

Haffty, B. G., Yang, Q., Reiss, M., Kearney, T., Higgins, S. A., Weidhaas, J., … Toppmeyer, D. (2006). Locoregional Relapse and Distant Metastasis in Conservatively Managed Triple Negative Early-Stage Breast Cancer. *Journal of Clinical Oncology*, *24*(36), 5652–5657. https://doi.org/10.1200/JCO.2006.06.5664

Hainaut, P., & Wiman, K. G. (2009). 30 years and a long way into p53 research. *The Lancet. Oncology*, *10*(9), 913–9. https://doi.org/10.1016/S1470-2045(09)70198-6

Halazonetis, T. D., Gorgoulis, V. G., & Bartek, J. (2008). An Oncogene-Induced DNA Damage Model for Cancer Development. *Science*, *319*(5868), 1352–1355. https://doi.org/10.1126/science.1140735

Halevy, O., Rodel, J., Peled, A., & Oren, M. (1991). Frequent p53 mutations in chemically induced murine fibrosarcoma. *Oncogene*, *6*(9), 1593–600. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1923526

Hall, P. A., McKee, P. H., Menage, H. D., Dover, R., & Lane, D. P. (1993). High levels of p53 protein in UV-irradiated normal human skin. *Oncogene*, *8*(1), 203–7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8093810

Hanel, W., & Moll, U. M. (2012). Links between mutant p53 and genomic instability. *Journal of Cellular Biochemistry*, *113*(2), 433–9. https://doi.org/10.1002/jcb.23400

Harlow, E., Williamson, N. M., Ralston, R., Helfman, D. M., & Adams, T. E. (1985). Molecular cloning and in vitro expression of a cDNA clone for human cellular tumor antigen p53. *Molecular and Cellular Biology*, *5*(7), 1601–10. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3894933

Harrow, J., Frankish, A., Gonzalez, J. M., Tapanari, E., Diekhans, M., Kokocinski, F., … Hubbard, T. J. (2012). GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res*, *22*(9), 1760–1774. https://doi.org/10.1101/gr.135350.111

Herschkowitz, J. I., Simin, K., Weigman, V. J., Mikaelian, I., Usary, J., Hu, Z., … Perou, C. M. (2007). Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. *Genome Biology*, *8*(5), R76. https://doi.org/10.1186/gb-2007-8-5-r76

Hollstein, M., Sidransky, D., Vogelstein, B., & Harris, C. C. (1991). p53 mutations in human cancers. *Science (New York, N.Y.)*, *253*(5015), 49–53. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1905840

Howlader, N., Altekruse, S. F., Li, C. I., Chen, V. W., Clarke, C. A., Ries, L. A. G., & Cronin, K. A. (2014). US incidence of breast cancer subtypes defined by joint hormone receptor and HER2 status. *Journal of the National Cancer Institute*, *106*(5). https://doi.org/10.1093/jnci/dju055

Hu, X., Stern, H. M., Ge, L., O'Brien, C., Haydu, L., Honchell, C. D., … Cavet, G. (2009). Genetic alterations and oncogenic pathways associated with breast cancer subtypes. *Mol Cancer Res*, *7*(4), 511–522. https://doi.org/10.1158/1541-7786.MCR-08-0107

Huang, M. E., Ye, Y. C., Chen, S. R., Chai, J. R., Lu, J. X., Zhoa, L., … Wang, Z. Y. (1988). Use of all-trans retinoic acid in the treatment of acute promyelocytic leukemia. *Blood*, *72*(2), 567–72. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3165295

Hudis, C. A., & Gianni, L. (2011). Triple-Negative Breast Cancer: An Unmet Medical Need. *The Oncologist*, *16*(Supplement 1), 1–11.

https://doi.org/10.1634/theoncologist.2011-S1-01

Hui, L., Zheng, Y., Yan, Y., Bargonetti, J., & Foster, D. A. (2006). Mutant p53 in MDA-MB-231 breast cancer cells is stabilized by elevated phospholipase D activity and contributes to survival signals generated by phospholipase D. *Oncogene*, *25*(55), 7305–7310. https://doi.org/10.1038/sj.onc.1209735

Hyman, D. M., Laetsch, T. W., Kummar, S., DuBois, S. G., Farago, A. F., Pappo, A. S., … Drilon, A. E. (2017). The efficacy of larotrectinib (LOXO-101), a selective tropomyosin receptor kinase (TRK) inhibitor, in adult and pediatric TRK fusion cancers. In *ASCO Annual Meeting*.

Ihle, J. N., Smith-White, B., Sisson, B., Parker, D., Blair, D. G., Schultz, A., … et al. (1989). Activation of the c-H-ras proto-oncogene by retrovirus insertion and chromosomal rearrangement in a Moloney leukemia virus-induced T-cell leukemia. *J Virol*, *63*(7), 2959–2966. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2542606

Iqbal, N., & Iqbal, N. (2014). Human Epidermal Growth Factor Receptor 2 (HER2) in Cancers: Overexpression and Therapeutic Implications. *Molecular Biology International*, *2014*, 1–9. https://doi.org/10.1155/2014/852748

Isakoff, S. J. (2010). Triple-negative breast cancer: role of specific chemotherapy agents. *Cancer Journal (Sudbury, Mass.)*, *16*(1), 53–61. https://doi.org/10.1097/PPO.0b013e3181d24ff7

Iwabuchi, K., Li, B., Bartel, P., & Fields, S. (1993). Use of the two-hybrid system to identify the domain of p53 involved in oligomerization. *Oncogene*, *8*(6), 1693–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8502489

Iwai, Y., Ishida, M., Tanaka, Y., Okazaki, T., Honjo, T., & Minato, N. (2002). Involvement of PD-L1 on tumor cells in the escape from host immune system and tumor immunotherapy by PD-L1 blockade. *Proc Natl Acad Sci U S A*, *99*(19), 12293–12297. https://doi.org/10.1073/pnas.192461099

Iyer, M. K., Niknafs, Y. S., Malik, R., Singhal, U., Sahu, A., Hosono, Y., … Chinnaiyan, A. M. (2015). The landscape of long noncoding RNAs in the human transcriptome. *Nature Genetics*, *47*(3), 199–208.

https://doi.org/10.1038/ng.3192

Jazayeri, A., Falck, J., Lukas, C., Bartek, J., Smith, G. C. M., Lukas, J., & Jackson, S. P. (2006). ATM- and cell cycle-dependent regulation of ATR in response to DNA double-strand breaks. *Nature Cell Biology*, *8*(1), 37–45. https://doi.org/10.1038/ncb1337

Jemal, A., Ward, E. M., Johnson, C. J., Cronin, K. A., Ma, J., Ryerson, A. B., … Weir, H. K. (2017). Annual Report to the Nation on the Status of Cancer, 1975–2014, Featuring Survival. *JNCI: Journal of the National Cancer Institute*, *109*(9). https://doi.org/10.1093/jnci/djx030

Jenkins, J. R., Rudge, K., Chumakov, P., & Currie, G. A. (1985). The cellular oncogene p53 can be activated by mutagenesis. *Nature*, *317*(6040), 816–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3903515

Jenkins, J. R., Rudge, K., & Currie, G. A. (1984). Cellular immortalization by a cDNA clone encoding the transformation-associated phosphoprotein p53. *Nature*, *312*(5995), 651–4. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6095117

Ji, Y., Xie, M., Lan, H., Zhang, Y., Long, Y., Weng, H., … Bu, Y. (2013). PRR11 is a novel gene implicated in cell cycle progression and lung cancer. *Int J Biochem Cell Biol*, *45*(3), 645–656. https://doi.org/10.1016/j.biocel.2012.12.002

Joensuu, H., Kellokumpu-Lehtinen, P.-L., Bono, P., Alanko, T., Kataja, V., Asola, R., … FinHer Study Investigators. (2006). Adjuvant Docetaxel or Vinorelbine with or without Trastuzumab for Breast Cancer. *New England Journal of Medicine*, *354*(8), 809–820. https://doi.org/10.1056/NEJMoa053028

Jones, S. A., Horiuchi, S., Topley, N., Yamamoto, N., & Fuller, G. M. (2001). The soluble interleukin 6 receptor: mechanisms of production and implications in disease. *FASEB J*, *15*(1), 43–58. https://doi.org/10.1096/fj.99-1003rev

Jones, S., Wang, T. L., Shih Ie, M., Mao, T. L., Nakayama, K., Roden, R., … Papadopoulos, N. (2010). Frequent mutations of chromatin remodeling gene ARID1A in ovarian clear cell carcinoma. *Science*, *330*(6001), 228–231. https://doi.org/10.1126/science.1196333

Kamijo, T., Weber, J. D., Zambetti, G., Zindy, F., Roussel, M. F., & Sherr, C. J. (1998). Functional and physical interactions of the ARF tumor suppressor with p53 and Mdm2. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(14), 8292–7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9653180

Kandoth, C., McLellan, M. D., Vandin, F., Ye, K., Niu, B., Lu, C., … Ding, L. (2013). Mutational landscape and significance across 12 major cancer types. *Nature*, *502*(7471), 333–9. https://doi.org/10.1038/nature12634

Kastan, M. B. (2007). Wild-Type p53: Tumors Can't Stand It. *Cell*, *128*(5), 837–840. https://doi.org/10.1016/j.cell.2007.02.022

Kastan, M. B., Zhan, Q., el-Deiry, W. S., Carrier, F., Jacks, T., Walsh, W. V, … Fornace, A. J. (1992). A mammalian cell cycle checkpoint pathway utilizing p53 and GADD45 is defective in ataxia-telangiectasia. *Cell*, *71*(4), 587–97. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1423616

Kataoka, K., Shiraishi, Y., Takeda, Y., Sakata, S., Matsumoto, M., Nagano, S., … Ogawa, S. (2016). Aberrant PD-L1 expression through 3'-UTR disruption in multiple cancers. *Nature*, *534*(7607), 402–406. https://doi.org/10.1038/nature18294

Kearney, L., & Horsley, S. W. (2005). Molecular cytogenetics in haematological malignancy: current technology and future prospects. *Chromosoma*, *114*(4), 286–294. https://doi.org/10.1007/s00412-005-0002-z

Kern, S. E., Kinzler, K. W., Bruskin, A., Jarosz, D., Friedman, P., Prives, C., & Vogelstein, B. (1991). Identification of p53 as a sequence-specific DNA-binding protein. *Science (New York, N.Y.)*, *252*(5013), 1708–11. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2047879

Kim, J., Lee, Y., Cho, H. J., Lee, Y. E., An, J., Cho, G. H., … Nam, D. H. (2014). NTRK1 fusion in glioblastoma multiforme. *PLoS One*, *9*(3), e91940. https://doi.org/10.1371/journal.pone.0091940

King, C. R., Kraus, M. H., & Aaronson, S. A. (1985). Amplification of a novel v-erbB-related gene in a human mammary carcinoma. *Science (New York, N.Y.)*, *229*(4717), 974–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2992089

Knudson Jr., A. G. (1971). Mutation and cancer: statistical study of retinoblastoma. *Proceedings of the National Academy of Sciences of the United States of America*, *68*(4), 820–3. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/5279523

Kravchenko, J. E., Ilyinskaya, G. V., Komarov, P. G., Agapova, L. S., Kochetkov, D. V., Strom, E., … Chumakov, P. M. (2008). Small-molecule RETRA suppresses mutant p53-bearing cancer cells through a p73-dependent salvage pathway. *Proceedings of the National Academy of Sciences*, *105*(17), 6302–6307. https://doi.org/10.1073/pnas.0802091105

Kress, M., May, E., Cassingena, R., & May, P. (1979). Simian virus 40-transformed cells express new species of proteins precipitable by anti-simian virus 40 tumor serum. *Journal of Virology*, *31*(2), 472–83. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/225566

Kroll, T. G., Sarraf, P., Pecciarini, L., Chen, C. J., Mueller, E., Spiegelman, B. M., & Fletcher, J. A. (2000). PAX8-PPARgamma1 fusion oncogene in human thyroid carcinoma [corrected]. *Science*, *289*(5483), 1357–1360. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10958784

Kwei, K. A., Kung, Y., Salari, K., Holcomb, I. N., & Pollack, J. R. (2010). Genomic instability in breast cancer: pathogenesis and clinical implications. *Molecular Oncology*, *4*(3), 255–66. https://doi.org/10.1016/j.molonc.2010.04.001

Lambert, J. M. R., Gorzov, P., Veprintsev, D. B., S?derqvist, M., Segerb?ck, D., Bergman, J., … Bykov, V. J. N. (2009). PRIMA-1 Reactivates Mutant p53 by Covalent Binding to the Core Domain. *Cancer Cell*, *15*(5), 376–388. https://doi.org/10.1016/j.ccr.2009.03.003

Lane, D. P. (1992). p53, guardian of the genome. *Nature*, *358*(6381), 15–16. https://doi.org/10.1038/358015a0

Lane, D. P., & Crawford, L. V. (1979). T antigen is bound to a host protein in SV40-transformed cells. *Nature*, *278*(5701), 261–3. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/218111

Lang, G. A., Iwakuma, T., Suh, Y.-A., Liu, G., Rao, V. A., Parant, J. M., … Lozano, G. (2004). Gain of Function of a p53 Hot Spot Mutation in a Mouse Model of Li-Fraumeni Syndrome. *Cell*, *119*(6), 861–872.

https://doi.org/10.1016/j.cell.2004.11.006

Lavin, M. F., & Kozlov, S. (2007). ATM Activation and DNA Damage Response. *Cell Cycle*, *6*(8), 931–942. https://doi.org/10.4161/cc.6.8.4180

Layer, R. M., Chiang, C., Quinlan, A. R., & Hall, I. M. (2014). LUMPY: a probabilistic framework for structural variant discovery. *Genome Biology*, *15*(6), R84. https://doi.org/10.1186/gb-2014-15-6-r84

Leach, F. S., Tokino, T., Meltzer, P., Burrell, M., Oliner, J. D., Smith, S., … Vogelstein, B. (1993). p53 Mutation and MDM2 amplification in human soft tissue sarcomas. *Cancer Research*, *53*(10 Suppl), 2231–4. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8387391

Lee, J. O., Yang, H., Georgescu, M. M., Di Cristofano, A., Maehama, T., Shi, Y., … Pavletich, N. P. (1999). Crystal structure of the PTEN tumor suppressor: implications for its phosphoinositide phosphatase activity and membrane association. *Cell*, *99*(3), 323–334. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10555148

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., … Kim, V. N. (2003). The nuclear RNase III Drosha initiates microRNA processing. *Nature*, *425*(6956), 415–419. https://doi.org/10.1038/nature01957

Lehmann, B. D., Bauer, J. A., Chen, X., Sanders, M. E., Chakravarthy, A. B., Shyr, Y., & Pietenpol, J. A. (2011). Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J Clin Invest*, *121*(7), 2750–2767. https://doi.org/10.1172/JCI45014

Lehmann, B. D., Bauer, J. A., Schafer, J. M., Pendleton, C. S., Tang, L., Johnson, K. C., … Pietenpol, J. A. (2014). PIK3CA mutations in androgen receptor-positive triple negative breast cancer confer sensitivity to the combination of PI3K and androgen receptor inhibitors. *Breast Cancer Res*, *16*(4), 406. https://doi.org/10.1186/s13058-014-0406-x

Lehmann, B. D., Jovanović, B., Chen, X., Estrada, M. V., Johnson, K. N., Shyr, Y., … Pietenpol, J. A. (2016). Refinement of Triple-Negative Breast Cancer Molecular Subtypes: Implications for Neoadjuvant Chemotherapy Selection. *PLOS ONE*, *11*(6), e0157368.

https://doi.org/10.1371/journal.pone.0157368

Lehmann, B. D., & Pietenpol, J. A. (2012). Targeting Mutant *p53* in Human Tumors. *Journal of Clinical Oncology*, *30*(29), 3648–3650. https://doi.org/10.1200/JCO.2012.44.0412

Lehmann, S., Bykov, V. J. N., Ali, D., Andrén, O., Cherif, H., Tidefelt, U., … Andersson, P.-O. (2012). Targeting p53 in Vivo: A First-in-Human Study With p53-Targeting Compound APR-246 in Refractory Hematologic Malignancies and Prostate Cancer. *Journal of Clinical Oncology*, *30*(29), 3633–3639. https://doi.org/10.1200/JCO.2011.40.7783

Leppard, K., Totty, N., Waterfield, M., Harlow, E., Jenkins, J., & Crawford, L. (1983). Purification and partial amino acid sequence analysis of the cellular tumour antigen, p53, from mouse SV40-transformed cells. *The EMBO Journal*, *2*(11), 1993–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6315411

Levine, A. J., & Oren, M. (2009). The first 30 years of p53: growing ever more complex. *Nature Reviews Cancer*, *9*(10), 749–758. https://doi.org/10.1038/nrc2723

Li, B., Ruotti, V., Stewart, R. M., Thomson, J. A., & Dewey, C. N. (2010). RNA-Seq gene expression estimation with read mapping uncertainty. *Bioinformatics*, *26*(4), 493–500. https://doi.org/10.1093/bioinformatics/btp692

Li, D., Marchenko, N. D., & Moll, U. M. (2011a). SAHA shows preferential cytotoxicity in mutant p53 cancer cells by destabilizing mutant p53 through inhibition of the HDAC6-Hsp90 chaperone axis. *Cell Death and Differentiation*, *18*(12), 1904–1913. https://doi.org/10.1038/cdd.2011.71

Li, D., Marchenko, N. D., Schulz, R., Fischer, V., Velasco-Hernandez, T., Talos, F., & Moll, U. M. (2011b). Functional Inactivation of Endogenous MDM2 and CHIP by HSP90 Causes Aberrant Stabilization of Mutant p53 in Human Cancer Cells. *Molecular Cancer Research*, *9*(5), 577–588. https://doi.org/10.1158/1541-7786.MCR-10-0534

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … Genome Project Data Processing, S. (2009). The Sequence

Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Li, Y., & Prives, C. (2007). Are interactions with p63 and p73 involved in mutant p53 gain of oncogenic function? *Oncogene*, *26*(15), 2220–2225. https://doi.org/10.1038/sj.onc.1210311

Liao, Y., Smyth, G. K., & Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics (Oxford, England)*, *30*(7), 923–30. https://doi.org/10.1093/bioinformatics/btt656

Lim, L. Y., Vidnovic, N., Ellisen, L. W., & Leong, C.-O. (2009). Mutant p53 mediates survival of breast cancer cells. *British Journal of Cancer*, *101*(9), 1606–1612. https://doi.org/10.1038/sj.bjc.6605335

Linzer, D. I., & Levine, A. J. (1979). Characterization of a 54K dalton cellular SV40 tumor antigen present in SV40-transformed cells and uninfected embryonal carcinoma cells. *Cell*, *17*(1), 43–52. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/222475

Liu, P., Cheng, H., Roberts, T. M., & Zhao, J. J. (2009). Targeting the phosphoinositide 3-kinase pathway in cancer. *Nature Reviews. Drug Discovery*, *8*(8), 627–44. https://doi.org/10.1038/nrd2926

Liu, X., Wilcken, R., Joerger, A. C., Chuckowree, I. S., Amin, J., Spencer, J., & Fersht, A. R. (2013). Small molecule induced reactivation of mutant p53 in cancer cells. *Nucleic Acids Research*, *41*(12), 6034–6044. https://doi.org/10.1093/nar/gkt305

Loi, S., Sirtaine, N., Piette, F., Salgado, R., Viale, G., Van Eenoo, F., … Sotiriou, C. (2013). Prognostic and Predictive Value of Tumor-Infiltrating Lymphocytes in a Phase III Randomized Adjuvant Breast Cancer Trial in Node-Positive Breast Cancer Comparing the Addition of Docetaxel to Doxorubicin With Doxorubicin-Based Chemotherapy: BIG 02-98. *Journal of Clinical Oncology*, *31*(7), 860–867. https://doi.org/10.1200/JCO.2011.41.0902

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550. https://doi.org/10.1186/s13059-014-0550-8

Lowe, S. W., Ruley, H. E., Jacks, T., & Housman, D. E. (1993a). p53-dependent apoptosis modulates the cytotoxicity of anticancer agents. *Cell*, *74*(6), 957–67. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8402885

Lowe, S. W., Schmitt, E. M., Smith, S. W., Osborne, B. A., & Jacks, T. (1993b). p53 is required for radiation-induced apoptosis in mouse thymocytes. *Nature*, *362*(6423), 847–849. https://doi.org/10.1038/362847a0

Magrane, M., & Consortium, U. (2011). UniProt Knowledgebase: a hub of integrated protein data. *Database (Oxford)*, *2011*, bar009. https://doi.org/10.1093/database/bar009

Malkin, D., Li, F. P., Strong, L. C., Fraumeni, J. F., Nelson, C. E., Kim, D. H., … Tainsky, M. A. (1990). Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. *Science (New York, N.Y.)*, *250*(4985), 1233–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1978757

Marchenko, N. D., Zaika, A., & Moll, U. M. (2000). Death signal-induced localization of p53 protein to mitochondria. A potential role in apoptotic signaling. *The Journal of Biological Chemistry*, *275*(21), 16202–12. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10821866

Martins, C. P., Brown-Swigart, L., & Evan, G. I. (2006). Modeling the Therapeutic Efficacy of p53 Restoration in Tumors. *Cell*, *127*(7), 1323–1334. https://doi.org/10.1016/j.cell.2006.12.007

Maruyama, T., Dougan, S. K., Truttmann, M. C., Bilate, A. M., Ingram, J. R., & Ploegh, H. L. (2015). (SI) Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nature Biotechnology*, (October 2014), 1–8. https://doi.org/10.1038/nbt.3190

Masuda, H., Baggerly, K. A., Wang, Y., Zhang, Y., Gonzalez-Angulo, A. M., Meric-Bernstam, F., … Ueno, N. T. (2013). Differential response to neoadjuvant chemotherapy among 7 triple-negative breast cancer molecular subtypes. *Clin Cancer Res*, *19*(19), 5533–5540. https://doi.org/10.1158/1078-0432.CCR-13-0799

Matlashewski, G., Lamb, P., Pim, D., Peacock, J., Crawford, L., &

Benchimol, S. (1984). Isolation and characterization of a human p53 cDNA clone: expression of the human p53 gene. *The EMBO Journal*, *3*(13), 3257–62. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6396087

Melero, J. A., Stitt, D. T., Mangel, W. F., & Carroll, R. B. (1979). Identification of new polypeptide species (48-55K) immunoprecipitable by antiserum to purified large T antigen and present in SV40-infected and -transformed cells. *Virology*, *93*(2), 466–80. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/222051

Mercer, W. E., Shields, M. T., Amin, M., Sauve, G. J., Appella, E., Romano, J. W., & Ullrich, S. J. (1990). Negative growth regulation in a glioblastoma tumor cell line that conditionally expresses human wild-type p53. *Proceedings of the National Academy of Sciences of the United States of America*, *87*(16), 6166–70. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2143581

Mertens, F., Johansson, B., Fioretos, T., & Mitelman, F. (2015). The emerging complexity of gene fusions in cancer. *Nat Rev Cancer*, *15*(6), 371–381. https://doi.org/10.1038/nrc3947

Michalovitz, D., Halevy, O., & Oren, M. (1990). Conditional inhibition of transformation and of cell proliferation by a temperature-sensitive mutant of p53. *Cell*, *62*(4), 671–80. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2143698

Milner, J., & Medcalf, E. A. (1991). Cotranslation of activated mutant p53 with wild type drives the wild-type p53 protein into the mutant conformation. *Cell*, *65*(5), 765–74. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2040013

Mitelman, F., Johansson, B., & Mertens, F. (2007). The impact of translocations and gene fusions on cancer causation. *Nature Reviews Cancer*, *7*(4), 233–245. https://doi.org/10.1038/nrc2091

Mittendorf, E. A., Philips, A. V., Meric-Bernstam, F., Qiao, N., Wu, Y., Harrington, S., … Alatrash, G. (2014). PD-L1 Expression in Triple-Negative Breast Cancer. *Cancer Immunology Research*, *2*(4), 361–370. https://doi.org/10.1158/2326-6066.CIR-13-0127

Miyashita, T., & Reed, J. C. (1995). Tumor suppressor p53 is a direct

transcriptional activator of the human bax gene. *Cell*, *80*(2), 293–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7834749

Moasser, M. M. (2007). The oncogene HER2: its signaling and transforming functions and its role in human cancer pathogenesis. *Oncogene*, *26*(45), 6469–6487. https://doi.org/10.1038/sj.onc.1210477

Mohammadi, M., Froum, S., Hamby, J. M., Schroeder, M. C., Panek, R. L., Lu, G. H., … Hubbard, S. R. (1998). Crystal structure of an angiogenesis inhibitor bound to the FGF receptor tyrosine kinase domain. *EMBO J*, *17*(20), 5896–5904. https://doi.org/10.1093/emboj/17.20.5896

Montagna, E., Maisonneuve, P., Rotmensz, N., Cancello, G., Iorfida, M., Balduzzi, A., … Colleoni, M. (2013). Heterogeneity of Triple-Negative Breast Cancer: Histologic Subtyping to Inform the Outcome. *Clinical Breast Cancer*, *13*(1), 31–39. https://doi.org/10.1016/j.clbc.2012.09.002

Morgenstern, J. P., & Land, H. (1990). Advanced mammalian gene transfer: high titre retroviral vectors with multiple drug selection markers and a complementary helper-free packaging cell line. *Nucleic Acids Res*, *18*(12), 3587–3596. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2194165

Morris, G. J., Naidu, S., Topham, A. K., Guiles, F., Xu, Y., McCue, P., … Mitchell, E. P. (2007). Differences in breast carcinoma characteristics in newly diagnosed African-American and Caucasian patients: a single-institution compilation compared with the National Cancer Institute's Surveillance, Epidemiology, and End Results database. *Cancer*, *110*(4), 876–84. https://doi.org/10.1002/cncr.22836

Muller, P. A. J., Caswell, P. T., Doyle, B., Iwanicki, M. P., Tan, E. H., Karim, S., … Vousden, K. H. (2009). Mutant p53 Drives Invasion by Promoting Integrin Recycling. *Cell*, *139*(7), 1327–1341. https://doi.org/10.1016/j.cell.2009.11.026

Muller, P. A. J., & Vousden, K. H. (2013). P53 Mutations in Cancer. *Nature Cell Biology*, *15*(1), 2–8. https://doi.org/10.1038/ncb2641

Muller, P. A. J., & Vousden, K. H. (2014). Mutant p53 in Cancer: New Functions and Therapeutic Opportunities. *Cancer Cell*, *25*(3), 304–

317. https://doi.org/10.1016/j.ccr.2014.01.021

Mullighan, C. G., Su, X., Zhang, J., Radtke, I., Phillips, L. A. A., Miller, C. B., … Children's Oncology Group. (2009). Deletion of *IKZF1* and Prognosis in Acute Lymphoblastic Leukemia. *New England Journal of Medicine*, *360*(5), 470–480. https://doi.org/10.1056/NEJMoa0808253

Nolvadex Adjuvant Trial Organisation. (1983). Controlled trial of tamoxifen as adjuvant agent in management of early breast cancer. Interim analysis at four years by Nolvadex Adjuvant Trial Organisation. *Lancet (London, England)*, *1*(8319), 257–61. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6130291

Nordling, C. O. (1953). A new theory on cancer-inducing mechanism. *British Journal of Cancer*, *7*(1), 68–72. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/13051507

Nowell, P. C., & Hungerford, D. A. (1960a). A minute chromosome in human chronic granulocytic leukemia. *Science*, *132*(1497).

Nowell, P. C., & Hungerford, D. A. (1960b). Chromosome studies on normal and leukemic human leukocytes. *Journal of the National Cancer Institute*, *25*, 85–109. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14427847

Olive, K. P., Tuveson, D. A., Ruhe, Z. C., Yin, B., Willis, N. A., Bronson, R. T., … Jacks, T. (2004). Mutant p53 Gain of Function in Two Mouse Models of Li-Fraumeni Syndrome. *Cell*, *119*(6), 847–860. https://doi.org/10.1016/j.cell.2004.11.004

Oren, M., & Levine, A. J. (1983). Molecular cloning of a cDNA specific for the murine p53 cellular tumor antigen. *Proceedings of the National Academy of Sciences of the United States of America*, *80*(1), 56–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6296874

Oren, M., & Rotter, V. (2010). Mutant p53 gain-of-function in cancer. *Cold Spring Harbor Perspectives in Biology*, *2*(2), a001107. https://doi.org/10.1101/cshperspect.a001107

Owczarek, S., Kiryushko, D., Larsen, M. H., Kastrup, J. S., Gajhede, M., Sandi, C., … Soroka, V. (2010). Neuroplastin-55 binds to and signals

through the fibroblast growth factor receptor. *FASEB J*, *24*(4), 1139–1150. https://doi.org/10.1096/fj.09-140509

Palacios, R., & Steinmetz, M. (1985). Il-3-dependent mouse clones that express B-220 surface antigen, contain Ig genes in germ-line configuration, and generate B lymphocytes in vivo. *Cell*, *41*(3), 727–734. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3924409

Paquet, D., Kwart, D., Chen, A., Sproul, A., Jacob, S., Teo, S., … Tessier-Lavigne, M. (2016). Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature*, *533*(7601), 125–129. https://doi.org/10.1038/nature17664

Parada, L. F., Land, H., Weinberg, R. A., Wolf, D., & Rotter, V. (1984). Cooperation between gene encoding p53 tumour antigen and ras in cellular transformation. *Nature*, *312*(5995), 649–51. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6390217

Pardoll, D. M. (2012). The blockade of immune checkpoints in cancer immunotherapy. *Nat Rev Cancer*, *12*(4), 252–264. https://doi.org/10.1038/nrc3239

Pareja, F., Geyer, F. C., Marchiò, C., Burke, K. A., Weigelt, B., & Reis-Filho, J. S. (2016). Triple-negative breast cancer: the importance of molecular and histologic subtyping, and recognition of low-grade variants. *Npj Breast Cancer*, *2*, 16036. Retrieved from http://dx.doi.org/10.1038/npjbcancer.2016.36

Parker, B. C., Annala, M. J., Cogdell, D. E., Granberg, K. J., Sun, Y., Ji, P., … Zhang, W. (2013). The tumorigenic FGFR3-TACC3 gene fusion escapes miR-99a regulation in glioblastoma. *J Clin Invest*, *123*(2), 855–865. https://doi.org/10.1172/JCI67144

Parrales, A., Ranjan, A., Iyer, S., Padhye, S., Weir, S., Roy, A., & Iwakuma, T. (2016). DNAJA1 controls the fate of misfolded mutant p53 through the mevalonate pathway. *Nature Cell Biology*, *18*(11), 1233–1243. https://doi.org/10.1038/ncb3427

Pavletich, N. P., Chambers, K. A., & Pabo, C. O. (1993). The DNA-binding domain of p53 contains the four conserved regions and the major mutation hot spots. *Genes & Development*, *7*(12B), 2556–64. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8276238

Pennica, D., Goeddel, D. V, Hayflick, J. S., Reich, N. C., Anderson, C. W., & Levine, A. J. (1984). The amino acid sequence of murine p53 determined from a c-DNA clone. *Virology*, *134*(2), 477–82. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6400059

Perou, C. M., S?rlie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Rees, C. A., … Botstein, D. (2000). Molecular portraits of human breast tumours. *Nature*, *406*(6797), 747–752. https://doi.org/10.1038/35021093

Piccart-Gebhart, M. J., Procter, M., Leyland-Jones, B., Goldhirsch, A., Untch, M., Smith, I., … Herceptin Adjuvant (HERA) Trial Study Team. (2005). Trastuzumab after Adjuvant Chemotherapy in HER2-Positive Breast Cancer. *New England Journal of Medicine*, *353*(16), 1659–1672. https://doi.org/10.1056/NEJMoa052306

Pietenpol, J. A., Tokino, T., Thiagalingam, S., el-Deiry, W. S., Kinzler, K. W., & Vogelstein, B. (1994). Sequence-specific transcriptional activation is essential for growth suppression by p53. *Proceedings of the National Academy of Sciences of the United States of America*, *91*(6), 1998–2002. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8134338

Pinkel, D., & Albertson, D. G. (2005). Array comparative genomic hybridization and its applications in cancer. *Nature Genetics*, *37*(6s), S11–S17. https://doi.org/10.1038/ng1569

Plasilova, M. L., Hayse, B., Killelea, B. K., Horowitz, N. R., Chagpar, A. B., & Lannin, D. R. (2016). Features of triple-negative breast cancer: Analysis of 38,813 cases from the national cancer database. *Medicine*, *95*(35), e4614. https://doi.org/10.1097/MD.0000000000004614

Pomerantz, J., Schreiber-Agus, N., Liégeois, N. J., Silverman, A., Alland, L., Chin, L., … DePinho, R. A. (1998). The Ink4a tumor suppressor gene product, p19Arf, interacts with MDM2 and neutralizes MDM2's inhibition of p53. *Cell*, *92*(6), 713–23. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9529248

Popova, T., Manie, E., Rieunier, G., Caux-Moncoutier, V., Tirapo, C., Dubois, T., … Stern, M.-H. (2012). Ploidy and Large-Scale Genomic Instability Consistently Identify Basal-like Breast Carcinomas with BRCA1/2 Inactivation. *Cancer Research*, *72*(21), 5454–5462.

https://doi.org/10.1158/0008-5472.CAN-12-1470

Prat, A., Parker, J. S., Karginova, O., Fan, C., Livasy, C., Herschkowitz, J. I., … Perou, C. M. (2010). Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Research*, *12*(5), R68. https://doi.org/10.1186/bcr2635

Press, M. F., Finn, R. S., Cameron, D., Di Leo, A., Geyer, C. E., Villalobos, I. E., … Koehler, M. T. (2008). HER-2 gene amplification, HER-2 and epidermal growth factor receptor mRNA and protein expression, and lapatinib efficacy in women with metastatic breast cancer. *Clin Cancer Res*, *14*(23), 7861–7870. https://doi.org/10.1158/1078-0432.CCR-08-1056

Pugh, T. J., Morozova, O., Attiyeh, E. F., Asgharzadeh, S., Wei, J. S., Auclair, D., … Maris, J. M. (2013). The genetic landscape of high-risk neuroblastoma. *Nature Genetics*, *45*(3), 279–84. https://doi.org/10.1038/ng.2529

Rabbitts, T. H. (1994). Chromosomal translocations in human cancer. *Nature*, *372*(6502), 143–149. https://doi.org/10.1038/372143a0

Rabbitts, T. H., & Boehm, T. (1991). Structural and functional chimerism results from chromosomal translocation in lymphoid tumors. *Advances in Immunology*, *50*, 119–46. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1950795

R Core Team. (2016). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org/

Rakha, E. A., El-Sayed, M. E., Green, A. R., Lee, A. H. S., Robertson, J. F., & Ellis, I. O. (2007). Prognostic markers in triple-negative breast cancer. *Cancer*, *109*(1), 25–32. https://doi.org/10.1002/cncr.22381

Rakha, E. A., Elsheikh, S. E., Aleskandarany, M. A., Habashi, H. O., Green, A. R., Powe, D. G., … Ellis, I. O. (2009). Triple-Negative Breast Cancer: Distinguishing between Basal and Nonbasal Subtypes. *Clinical Cancer Research*, *15*(7), 2302–2310. https://doi.org/10.1158/1078-0432.CCR-08-2132

Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. a, & Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nature Protocols*, *8*(11), 2281–308. https://doi.org/10.1038/nprot.2013.143

Raycroft, L., Wu, H. Y., & Lozano, G. (1990). Transcriptional activation by wild-type but not transforming mutants of the p53 anti-oncogene. *Science (New York, N.Y.)*, *249*(4972), 1049–51. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2144364

Rikova, K., Guo, A., Zeng, Q., Possemato, A., Yu, J., Haack, H., … Comb, M. J. (2007). Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. *Cell*, *131*(6), 1190–1203. https://doi.org/10.1016/j.cell.2007.11.025

Roberts, A., Pimentel, H., Trapnell, C., & Pachter, L. (2011). Identification of novel transcripts in annotated genomes using RNA-Seq. *Bioinformatics*, *27*(17), 2325–2329. https://doi.org/10.1093/bioinformatics/btr355

Robinson, J. T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., & Mesirov, J. P. (2011). Integrative genomics viewer. *Nat Biotechnol*, *29*(1), 24–26. https://doi.org/10.1038/nbt.1754

Rodriguez, O. C., Choudhury, S., Kolukula, V., Vietsch, E. E., Catania, J., Preet, A., … Avantaggiati, M. L. (2012). Dietary downregulation of mutant p53 levels via glucose restriction. *Cell Cycle*, *11*(23), 4436–4446. https://doi.org/10.4161/cc.22778

Romond, E. H., Perez, E. A., Bryant, J., Suman, V. J., Geyer, C. E., Davidson, N. E., … Wolmark, N. (2005). Trastuzumab plus Adjuvant Chemotherapy for Operable HER2-Positive Breast Cancer. *New England Journal of Medicine*, *353*(16), 1673–1684. https://doi.org/10.1056/NEJMoa052122

Rotter, V., Witte, O. N., Coffman, R., & Baltimore, D. (1980). Abelson murine leukemia virus-induced tumors elicit antibodies against a host cell protein, P50. *Journal of Virology*, *36*(2), 547–55. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6159484

Rowley, J. D. (1973). Letter: A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining. *Nature*, *243*(5405), 290–3. Retrieved from

http://www.ncbi.nlm.nih.gov/pubmed/4126434

Rubin, I., & Yarden, Y. (2001). The basic biology of HER2. *Annals of Oncology : Official Journal of the European Society for Medical Oncology*, *12 Suppl 1*, S3-8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11521719

Sabapathy, K. (2015). The Contrived Mutant p53 Oncogene - Beyond Loss of Functions. *Frontiers in Oncology*, *5*, 276. https://doi.org/10.3389/fonc.2015.00276

Sabatier, R., Finetti, P., Mamessier, E., Adelaide, J., Chaffanet, M., Ali, H. R., … Bertucci, F. (2015). Prognostic and predictive value of PDL1 expression in breast cancer. *Oncotarget*, *6*(7), 5449–5464. https://doi.org/10.18632/oncotarget.3216

Sauer, L., Gitenay, D., Vo, C., & Baron, V. T. (2010). Mutant p53 initiates a feedback loop that involves Egr-1/EGF receptor/ERK in prostate cancer cells. *Oncogene*, *29*(18), 2628–2637. https://doi.org/10.1038/onc.2010.24

Schechter, A. L., Hung, M. C., Vaidyanathan, L., Weinberg, R. A., Yang-Feng, T. L., Francke, U., … Coussens, L. (1985). The neu gene: an erbB-homologous gene distinct from and unlinked to the gene encoding the EGF receptor. *Science (New York, N.Y.)*, *229*(4717), 976–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2992090

Schechter, A. L., Stern, D. F., Vaidyanathan, L., Decker, S. J., Drebin, J. A., Greene, M. I., & Weinberg, R. A. (1984). The neu oncogene: an erb-B-related gene encoding a 185,000-Mr tumour antigen. *Nature*, *312*(5994), 513–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6095109

Schindler, T., Bornmann, W., Pellicena, P., Miller, W. T., Clarkson, B., & Kuriyan, J. (2000). Structural mechanism for STI-571 inhibition of abelson tyrosine kinase. *Science (New York, N.Y.)*, *289*(5486), 1938–42. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10988075

Schlegel, J., Sambade, M. J., Sather, S., Moschos, S. J., Tan, A. C., Winges, A., … Graham, D. K. (2013). MERTK receptor tyrosine kinase is a therapeutic target in melanoma. *J Clin Invest*, *123*(5), 2257–2267. https://doi.org/10.1172/JCI67816

Schwartzberg, L. S., Yardley, D. A., Elias, A. D., Patel, M., LoRusso, P., Burris, H. A., … Traina, T. A. (2017). A Phase I/Ib Study of Enzalutamide Alone and in Combination with Endocrine Therapies in Women with Advanced Breast Cancer. *Clinical Cancer Research*. https://doi.org/10.1158/1078-0432.CCR-16-2339

Schwartzenberg-Bar-Yoseph, F., Armoni, M., & Karnieli, E. (2004). The tumor suppressor p53 down-regulates glucose transporters GLUT1 and GLUT4 gene expression. *Cancer Research*, *64*(7), 2627–33. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15059920

Serrano, M., Lin, A. W., McCurrach, M. E., Beach, D., & Lowe, S. W. (1997). Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16INK4a. *Cell*, *88*(5), 593–602. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9054499

Shah, S. P., Roth, A., Goya, R., Oloumi, A., Ha, G., Zhao, Y., … Aparicio, S. (2012). The clonal and mutational evolution spectrum of primary triple-negative breast cancers. *Nature*, *486*(7403), 395–399. https://doi.org/10.1038/nature10933

Shaver, T. M., Lehmann, B. D., Beeler, J. S., Li, C. I., Li, Z., Jin, H., … Pietenpol, J. A. (2016). Diverse, biologically relevant, and targetable gene rearrangements in triple-negative breast cancer and other malignancies. *Cancer Research*, *76*(16), 4850–4860. https://doi.org/10.1158/0008-5472.CAN-16-0058

Shaw, A. T., Hsu, P. P., Awad, M. M., & Engelman, J. A. (2013). Tyrosine kinase gene rearrangements in epithelial malignancies. *Nat Rev Cancer*, *13*(11), 772–787. https://doi.org/10.1038/nrc3612

Shaw, P., Bovey, R., Tardy, S., Sahli, R., Sordat, B., & Costa, J. (1992). Induction of apoptosis by wild-type p53 in a human colon tumor-derived cell line. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(10), 4495–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1584781

Shepard, H. M., Lewis, G. D., Sarup, J. C., Fendly, B. M., Maneval, D., Mordenti, J., … Ullrich, A. (1991). Monoclonal antibody therapy of human cancer: taking the HER2 protooncogene to the clinic. *Journal of Clinical Immunology*, *11*(3), 117–27. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1679763

Shieh, S. Y., Ahn, J., Tamai, K., Taya, Y., & Prives, C. (2000). The human homologs of checkpoint kinases Chk1 and Cds1 (Chk2) phosphorylate p53 at multiple DNA damage-inducible sites. *Genes & Development*, *14*(3), 289–300. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10673501

Shih, C., Padhy, L. C., Murray, M., & Weinberg, R. A. (1981). Transforming genes of carcinomas and neuroblastomas introduced into mouse fibroblasts. *Nature*, *290*(5803), 261–4. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7207618

Siliciano, J. D., Canman, C. E., Taya, Y., Sakaguchi, K., Appella, E., & Kastan, M. B. (1997). DNA damage induces phosphorylation of the amino terminus of p53. *Genes & Development*, *11*(24), 3471–81. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9407038

Silwal-Pandit, L., Vollan, H. K. M., Chin, S.-F., Rueda, O. M., McKinney, S., Osako, T., … Langerod, A. (2014). TP53 Mutation Spectrum in Breast Cancer Is Subtype Specific and Has Distinct Prognostic Relevance. *Clinical Cancer Research*, *20*(13), 3569–3580. https://doi.org/10.1158/1078-0432.CCR-13-2943

Singh, D., Chan, J. M., Zoppoli, P., Niola, F., Sullivan, R., Castano, A., … Iavarone, A. (2012). Transforming fusions of FGFR and TACC genes in human glioblastoma. *Science*, *337*(6099), 1231–1235. https://doi.org/10.1126/science.1220834

Sinha, G. (2014). Downfall of Iniparib: A PARP Inhibitor That Doesn't Inhibit PARP After All. *JNCI Journal of the National Cancer Institute*, *106*(1), djt447-djt447. https://doi.org/10.1093/jnci/djt447

Slamon, D. J., Leyland-Jones, B., Shak, S., Fuchs, H., Paton, V., Bajamonde, A., … Norton, L. (2001). Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *The New England Journal of Medicine*, *344*(11), 783–92. https://doi.org/10.1056/NEJM200103153441101

Smith, A. E., Smith, R., & Paucha, E. (1979). Characterization of different tumor antigens present in cells transformed by simian virus 40. *Cell*, *18*(2), 335–46. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/227604

Smith, J., Mun Tho, L., Xu, N., & A. Gillespie, D. (2010). The ATM-Chk2 and ATR-Chk1 Pathways in DNA Damage Signaling and Cancer. In *Advances in cancer research* (Vol. 108, pp. 73–112). https://doi.org/10.1016/B978-0-12-380888-2.00003-0

Soda, M., Choi, Y. L., Enomoto, M., Takada, S., Yamashita, Y., Ishikawa, S., … Mano, H. (2007). Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. *Nature*, *448*(7153), 561–566. https://doi.org/10.1038/nature05945

Song, H., Hollstein, M., & Xu, Y. (2007). p53 gain-of-function cancer mutants induce genetic instability by inactivating ATM. *Nature Cell Biology*, *9*(5), 573–580. https://doi.org/10.1038/ncb1571

Soussi, T., & Wiman, K. G. (2015). TP53: an oncogene in disguise. *Cell Death and Differentiation*, *22*(8), 1239–49. https://doi.org/10.1038/cdd.2015.53

Speicher, M. R., & Carter, N. P. (2005). The new cytogenetics: blurring the boundaries with molecular biology. *Nature Reviews Genetics*, *6*(10), 782–792. https://doi.org/10.1038/nrg1692

Srivastava, S., Zou, Z. Q., Pirollo, K., Blattner, W., & Chang, E. H. (1990). Germ-line transmission of a mutated p53 gene in a cancer-prone family with Li-Fraumeni syndrome. *Nature*, *348*(6303), 747–9. https://doi.org/10.1038/348747a0

Stefansson, H., Helgason, A., Thorleifsson, G., Steinthorsdottir, V., Masson, G., Barnard, J., … Stefansson, K. (2005). A common inversion under selection in Europeans. *Nat Genet*, *37*(2), 129–137. https://doi.org/10.1038/ng1508

Stransky, N., Cerami, E., Schalm, S., Kim, J. L., & Lengauer, C. (2014). The landscape of kinase fusions in cancer. *Nat Commun*, *5*, 4846. https://doi.org/10.1038/ncomms5846

Stürzbecher, H. W., Brain, R., Addison, C., Rudge, K., Remm, M., Grimaldi, M., … Jenkins, J. R. (1992). A C-terminal alpha-helix plus basic region motif is the major structural determinant of p53 tetramerization. *Oncogene*, *7*(8), 1513–23. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1321401

Sugita, Y., Nakano, Y., Oda, E., Noda, K., Tobe, T., Miura, N. H., & Tomita, M. (1993). Determination of carboxyl-terminal residue and disulfide bonds of MACIF (CD59), a glycosyl-phosphatidylinositol-anchored membrane protein. *J Biochem*, *114*(4), 473–477. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8276756

Suh, Y.-A., Post, S. M., Elizondo-Fraire, A. C., Maccio, D. R., Jackson, J. G., El-Naggar, A. K., … Lozano, G. (2011). Multiple Stress Signals Activate Mutant p53 In Vivo. *Cancer Research*, *71*(23), 7168–7175. https://doi.org/10.1158/0008-5472.CAN-11-0459

Sun, Y., Dong, Z., Nakamura, K., & Colburn, N. H. (1993). Dosage-dependent dominance over wild-type p53 of a mutant p53 isolated from nasopharyngeal carcinoma. *FASEB Journal : Official Publication of the Federation of American Societies for Experimental Biology*, *7*(10), 944–50. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8344492

Sur, S., Pagliarini, R., Bunz, F., Rago, C., Diaz, L. A., Kinzler, K. W., … Papadopoulos, N. (2009). A panel of isogenic human cancer cells suggests a therapeutic approach for cancers with inactivated p53. *Proceedings of the National Academy of Sciences*, *106*(10), 3964–3969. https://doi.org/10.1073/pnas.0813333106

Terzian, T., Suh, Y.-A., Iwakuma, T., Post, S. M., Neumann, M., Lang, G. a, … Lozano, G. (2008). The inherent instability of mutant p53 is alleviated by Mdm2 or p16INK4a loss. *Genes & Development*, *22*(10), 1337–44. https://doi.org/10.1101/gad.1662908

Tomlins, S. A., Mehra, R., Rhodes, D. R., Smith, L. R., Roulston, D., Helgeson, B. E., … Chinnaiyan, A. M. (2006). TMPRSS2:ETV4 gene fusions define a third molecular subtype of prostate cancer. *Cancer Res*, *66*(7), 3396–3400. https://doi.org/10.1158/0008-5472.CAN-06-0168

Tomlins, S. A., Rhodes, D. R., Perner, S., Dhanasekaran, S. M., Mehra, R., Sun, X. W., … Chinnaiyan, A. M. (2005). Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science*, *310*(5748), 644–648. https://doi.org/10.1126/science.1117679

Tong, Y., Merino, D., Nimmervoll, B., Gupta, K., Wang, Y. D., Finkelstein,

D., … Gilbertson, R. J. (2015). Cross-Species Genomics Identifies TAF12, NFYC, and RAD54L as Choroid Plexus Carcinoma Oncogenes. *Cancer Cell*, *27*(5), 712–727. https://doi.org/10.1016/j.ccell.2015.04.005

Torrezan, G. T., Ferreira, E. N., Nakahata, A. M., Barros, B. D., Castro, M. T., Correa, B. R., … Carraro, D. M. (2014). Recurrent somatic mutation in DROSHA induces microRNA profile changes in Wilms tumour. *Nat Commun*, *5*, 4039. https://doi.org/10.1038/ncomms5039

Turner, N. C., Lord, C. J., Iorns, E., Brough, R., Swift, S., Elliott, R., … Ashworth, A. (2008). A synthetic lethal siRNA screen identifying genes mediating sensitivity to a PARP inhibitor. *The EMBO Journal*, *27*(9), 1368–1377. https://doi.org/10.1038/emboj.2008.61

Turner, N., Lambros, M. B., Horlings, H. M., Pearson, A., Sharpe, R., Natrajan, R., … Reis-Filho, J. S. (2010). Integrative molecular profiling of triple negative breast cancers identifies amplicon drivers and potential therapeutic targets. *Oncogene*, *29*(14), 2013–2023. https://doi.org/10.1038/onc.2009.489

Turner, N., Tutt, A., & Ashworth, A. (2005). Targeting the DNA repair defect of BRCA tumours. *Current Opinion in Pharmacology*, *5*(4), 388–393. https://doi.org/10.1016/j.coph.2005.03.006

Tyner, S. D., Venkatachalam, S., Choi, J., Jones, S., Ghebranious, N., Igelmann, H., … Donehower, L. A. (2002). p53 mutant mice that display early ageing-associated phenotypes. *Nature*, *415*(6867), 45–53. https://doi.org/10.1038/415045a

Vadnais, C., Shooshtarizadeh, P., Rajadurai, C. V, Lesurf, R., Hulea, L., Davoudi, S., … Nepveu, A. (2014). Autocrine Activation of the Wnt/beta-Catenin Pathway by CUX1 and GLIS1 in Breast Cancers. *Biol Open*, *3*(10), 937–946. https://doi.org/10.1242/bio.20148193

Vakifahmetoglu-Norberg, H., Kim, M., Xia, H. -g., Iwanicki, M. P., Ofengeim, D., Coloff, J. L., … Yuan, J. (2013). Chaperone-mediated autophagy degrades mutant p53. *Genes & Development*, *27*(15), 1718–1730. https://doi.org/10.1101/gad.220897.113

Vaseva, A. V., & Moll, U. M. (2009). The mitochondrial p53 pathway. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, *1787*(5), 414–

420. https://doi.org/10.1016/j.bbabio.2008.10.005

Vassilev, L. T., Vu, B. T., Graves, B., Carvajal, D., Podlaski, F., Filipovic, Z., … Liu, E. A. (2004). In Vivo Activation of the p53 Pathway by Small-Molecule Antagonists of MDM2. *Science*, *303*(5659), 844–848. https://doi.org/10.1126/science.1092472

Venkitaraman, A. R. (2009). Linking the Cellular Functions of *BRCA* Genes to Cancer Pathogenesis and Treatment. *Annual Review of Pathology: Mechanisms of Disease*, *4*(1), 461–487. https://doi.org/10.1146/annurev.pathol.3.121806.151422

Ventura, A., Kirsch, D. G., McLaughlin, M. E., Tuveson, D. A., Grimm, J., Lintault, L., … Jacks, T. (2007). Restoration of p53 function leads to tumour regression in vivo. *Nature*, *445*(7128), 661–665. https://doi.org/10.1038/nature05541

Voduc, K. D., Cheang, M. C. U., Tyldesley, S., Gelmon, K., Nielsen, T. O., & Kennecke, H. (2010). Breast Cancer Subtypes and the Risk of Local and Regional Relapse. *Journal of Clinical Oncology*, *28*(10), 1684–1691. https://doi.org/10.1200/JCO.2009.24.9284

Vogel, C. L., Cobleigh, M. A., Tripathy, D., Gutheil, J. C., Harris, L. N., Fehrenbacher, L., … Press, M. (2002). Efficacy and Safety of Trastuzumab as a Single Agent in First-Line Treatment of *HER2* -Overexpressing Metastatic Breast Cancer. *Journal of Clinical Oncology*, *20*(3), 719–726. https://doi.org/10.1200/JCO.2002.20.3.719

Vousden, K. H., & Ryan, K. M. (2009). p53 and metabolism. *Nature Reviews Cancer*, *9*(10), 691–700. https://doi.org/10.1038/nrc2715

Walerych, D., Napoli, M., Collavin, L., & Del Sal, G. (2012). The rebel angel: mutant p53 as the driving oncogene in breast cancer. *Carcinogenesis*, *33*(11), 2007–17. https://doi.org/10.1093/carcin/bgs232

Walker, M., Black, E. J., Oehler, V., Gillespie, D. A., & Scott, M. T. (2009). Chk1 C-terminal regulatory phosphorylation mediates checkpoint activation by de-repression of Chk1 catalytic activity. *Oncogene*, *28*(24), 2314–2323. https://doi.org/10.1038/onc.2009.102

Wang, K., Singh, D., Zeng, Z., Coleman, S. J., Huang, Y., Savich, G. L., … Liu, J. (2010). MapSplice: accurate mapping of RNA-seq reads for splice junction discovery. *Nucleic Acids Res*, *38*(18), e178. https://doi.org/10.1093/nar/gkq622

Wang, W., Cheng, B., Miao, L., Mei, Y., & Wu, M. (2013a). Mutant p53-R273H gains new function in sustained activation of EGFR signaling via suppressing miR-27a expression. *Cell Death and Disease*, *4*(4), e574. https://doi.org/10.1038/cddis.2013.97

Wang, X. S., Shankar, S., Dhanasekaran, S. M., Ateeq, B., Sasaki, A. T., Jing, X., … Chinnaiyan, A. M. (2011a). Characterization of KRAS rearrangements in metastatic prostate cancer. *Cancer Discov*, *1*(1), 35–43. https://doi.org/10.1158/2159-8274.CD-10-0022

Wang, Y., Blandino, G., Oren, M., & Givol, D. (1998). Induced p53 expression in lung cancer cell line promotes cell senescence and differentially modifies the cytotoxicity of anti-cancer drugs. *Oncogene*, *17*(15), 1923–1930. https://doi.org/10.1038/sj.onc.1202113

Wang, Y., Moncayo, G., Morin Jr., P., Xue, G., Grzmil, M., Lino, M. M., … Hemmings, B. A. (2013b). Mer receptor tyrosine kinase promotes invasion and survival in glioblastoma multiforme. *Oncogene*, *32*(7), 872–882. https://doi.org/10.1038/onc.2012.104

Wang, Y., Reed, M., Wang, P., Stenger, J. E., Mayr, G., Anderson, M. E., … Tegtmeyer, P. (1993). p53 domains: identification and characterization of two autonomous DNA-binding regions. *Genes & Development*, *7*(12B), 2575–86. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8276240

Wang, Y., Suh, Y.-A., Fuller, M. Y., Jackson, J. G., Xiong, S., Terzian, T., … Lozano, G. (2011b). Restoring expression of wild-type p53 suppresses tumor growth but does not cause tumor regression in mice with a p53 missense mutation. *Journal of Clinical Investigation*, *121*(3), 893–904. https://doi.org/10.1172/JCI44504

Warrell, R. P., Frankel, S. R., Miller, W. H., Scheinberg, D. A., Itri, L. M., Hittelman, W. N., … Dmitrovsky, E. (1991). Differentiation Therapy of Acute Promyelocytic Leukemia with Tretinoin (All-trans-Retinoic Acid). *New England Journal of Medicine*, *324*(20), 1385–1393. https://doi.org/10.1056/NEJM199105163242002

Wickham, H. (2009). *ggplot2: elegant graphics for data analysis*. Springer New York.

Williams, S. V, Hurst, C. D., & Knowles, M. A. (2013). Oncogenic FGFR3 gene fusions in bladder cancer. *Hum Mol Genet*, *22*(4), 795–803. https://doi.org/10.1093/hmg/dds486

Wimberly, H., Brown, J. R., Schalper, K., Haack, H., Silver, M. R., Nixon, C., … Rimm, D. L. (2015). PD-L1 Expression Correlates with Tumor-Infiltrating Lymphocytes and Response to Neoadjuvant Chemotherapy in Breast Cancer. *Cancer Immunology Research*, *3*(4), 326–332. https://doi.org/10.1158/2326-6066.CIR-14-0133

Wolf, D., Harris, N., Goldfinger, N., & Rotter, V. (1985). Isolation of a full-length mouse cDNA clone coding for an immunologically distinct p53 molecule. *Molecular and Cellular Biology*, *5*(1), 127–32. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2580227

Wolf, D., Harris, N., & Rotter, V. (1984). Reconstitution of p53 expression in a nonproducer Ab-MuLV-transformed cell line by transfection of a functional p53 gene. *Cell*, *38*(1), 119–26. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6088057

Wu, X., Bayle, J. H., Olson, D., & Levine, A. J. (1993). The p53-mdm-2 autoregulatory feedback loop. *Genes & Development*, *7*(7A), 1126–32. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8319905

Xue, W., Zender, L., Miething, C., Dickins, R. A., Hernando, E., Krizhanovsky, V., … Lowe, S. W. (2007). Senescence and tumour clearance is triggered by p53 restoration in murine liver carcinomas. *Nature*, *445*(7128), 656–660. https://doi.org/10.1038/nature05529

Yan, W., Liu, S., Xu, E., Zhang, J., Zhang, Y., Chen, X., & Chen, X. (2013). Histone deacetylase inhibitors suppress mutant p53 transcription via histone deacetylase 8. *Oncogene*, *32*(5), 599–609. https://doi.org/10.1038/onc.2012.81

Yemelyanova, A., Vang, R., Kshirsagar, M., Lu, D., Marks, M. A., Shih, I. M., & Kurman, R. J. (2011). Immunohistochemical staining patterns of p53 can serve as a surrogate marker for TP53 mutations in ovarian carcinoma: an immunohistochemical and nucleotide sequencing analysis. *Modern Pathology*, *24*(9), 1248–1253.

https://doi.org/10.1038/modpathol.2011.85

Yonish-Rouach, E., Resnftzky, D., Lotem, J., Sachs, L., Kimchi, A., & Oren, M. (1991). Wild-type p53 induces apoptosis of myeloid leukaemic cells that is inhibited by interleukin-6. *Nature*, *352*(6333), 345–347. https://doi.org/10.1038/352345a0

Yoshihara, K., Wang, Q., Torres-Garcia, W., Zheng, S., Vegesna, R., Kim, H., & Verhaak, R. G. (2014). The landscape and therapeutic relevance of cancer-associated transcript fusions. *Oncogene*. https://doi.org/10.1038/onc.2014.406

Yu, X., Vazquez, A., Levine, A., & Carpizo, D. (2012). Allele-Specific p53 Mutant Reactivation. *Cancer Cell*, *21*(5), 614–625. https://doi.org/10.1016/j.ccr.2012.03.042

Zakut-Houri, R., Bienz-Tadmor, B., Givol, D., & Oren, M. (1985). Human p53 cellular tumor antigen: cDNA sequence and expression in COS cells. *The EMBO Journal*, *4*(5), 1251–5. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/4006916

Zambetti, G. P., Bargonetti, J., Walker, K., Prives, C., & Levine, A. J. (1992). Wild-type p53 mediates positive regulation of gene expression through a specific DNA sequence element. *Genes & Development*, *6*(7), 1143–52. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1628822

Zauberman, A., Barak, Y., Ragimov, N., Levy, N., & Oren, M. (1993). Sequence-specific DNA binding by p53: identification of target sites and lack of binding to p53 - MDM2 complexes. *The EMBO Journal*, *12*(7), 2799–808. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8334996

Zerdoumi, Y., Aury-Landas, J., Bonaïti-Pellié, C., Derambure, C., Sesboüé, R., Renaux-Petel, M., … Flaman, J.-M. (2013). Drastic effect of germline TP53 missense mutations in Li-Fraumeni patients. *Human Mutation*, *34*(3), 453–61. https://doi.org/10.1002/humu.22254

Zhou, F., Liu, H., Zhang, X., Shen, Y., Zheng, D., Zhang, A., … Li, H. (2014). Proline-rich protein 11 regulates epithelial-to-mesenchymal transition to promote breast cancer cell invasion. *Int J Clin Exp Pathol*, *7*(12), 8692–8699. Retrieved from

http://www.ncbi.nlm.nih.gov/pubmed/25674234

Zhu, H.-B., Yang, K., Xie, Y.-Q., Lin, Y.-W., Mao, Q.-Q., & Xie, L.-P. (2013). Silencing of mutant p53 by siRNA induces cell cycle arrest and apoptosis in human bladder cancer cells. *World Journal of Surgical Oncology*, *11*(1), 22. https://doi.org/10.1186/1477-7819-11-22

Zhu, H., Mao, Q., Lin, Y., Yang, K., & Xie, L. (2011). RNA interference targeting mutant p53 inhibits growth and induces apoptosis in DU145 human prostate cancer cells. *Medical Oncology*, *28*(S1), 381–387. https://doi.org/10.1007/s12032-010-9679-9

Zhu, J., Zhou, W., Jiang, J., & Chen, X. (1998). Identification of a novel p53 functional domain that is necessary for mediating apoptosis. *The Journal of Biological Chemistry*, *273*(21), 13030–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9582339

Zhu, Y., Knolhoff, B. L., Meyer, M. A., Nywening, T. M., West, B. L., Luo, J., … DeNardo, D. G. (2014). CSF1/CSF1R blockade reprograms tumor-infiltrating macrophages and improves response to T-cell checkpoint immunotherapy in pancreatic cancer models. *Cancer Res*, *74*(18), 5057–5069. https://doi.org/10.1158/0008-5472.CAN-13-3723

Zindy, F., Eischen, C. M., Randle, D. H., Kamijo, T., Cleveland, J. L., Sherr, C. J., & Roussel, M. F. (1998). Myc signaling via the ARF tumor suppressor regulates p53-dependent apoptosis and immortalization. *Genes & Development*, *12*(15), 2424–33. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9694806