MULTI-OMIC ASSESSMENT OF ONCOGENIC PIK3CA MUTATIONS USING ISOGENIC CELL LINE MODELS

By

Adam Xavier Miranda

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

Cancer Biology

August 9, 2024

Nashville, Tennessee

Approved:

Vito Quaranta, M.D. (Committee Chair)

Ann Richmond, Ph.D.

Thomas Stricker, Ph.D.

Emily Hodges, Ph.D. (Co-mentor)

Ben Ho Park, M.D.,Ph.D. (Primary Mentor)

I dedicate this thesis to my loving wife. Thank you for sticking by me throughout this entire (ordeal) experience. I could not have done this without your support. I can't wait to see what the chapters of our lives after this are like.

ACKNOWLEDGMENTS

The work presented herein was made possible by financial support from the Predoctoral Biochemistry and Chemistry in Cancer T32 (T32CA009582) (A.X.M.). The American Cancer Society 131356-RSG-17-160-01-CSM as well as The Breast Cancer Research Foundation, Komen Foundation, and NIH CA214494/CA194024 (B.H.P.). We would also like to thank and acknowledge the support of The Canney Foundation, the Sage Patient Advocates, the Marcie and Ellen Foundation, The Eddie and Sandy Garcia Foundation, the support of Amy and Barry Baker, the support of John and Donna Hall, and the Vanderbilt-Ingram Cancer Center support grant (NIH CA068485) and Breast Cancer SPORE (NIH CA098131). Support from the Vanderbilt core facilities was also paramount to the completion of many of the experiments described within. Chief among these core facilities are the VANTAGE sequencing core and the Functional Genomics core. Schematic figures were also generated with the help of biorender.com. The data used for the analyses described in this dissertation were obtained from:the GTEx Portal on 05/26/24.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Introduction

1.1 *PIK3CA* mutations in breast cancer

Breast cancer is a serious public health problem and continues to be by far the most common type of cancer diagnosed in American women. In 2023 alone, it was estimated that there would be over 300,000 new breast cancer diagnoses in the United States. This number was more than double the prediction for lung cancer, which was the next most common cancer type in women (Siegel et al., 2023). Recent trends however, show that mortality rates across all cancers in both men and women have been decreasing over the past 20 years (Siegel et al., 2023). A major contributor to this trend has been the implementation and advancement of targeted therapies. Targeted therapies is a term that describes a class of therapeutic molecules that inhibit the function of specific proteins within cancer cells (Waarts et al., 2022). These therapies are the foundation of precision medicine and allow for personalized treatment plans that are specific to the mutations present in each patient's tumor. Cancer is a disease of accumulated mutations and an individual patient may have a totally different set of mutations from another with the same disease (Fearon and Vogelstein, 1990). The promise of precision medicine is for clinicians to specifically target the altered proteins present within a given patient's disease and to create a personalized treatment plan for each patient.

While targeted therapies have greatly improved patient outcomes in breast cancer, there is still room for improvement. One major area of improvement is better targeting of *PIK3CA* mutations in breast cancer. *PIK3CA* is the most commonly mutated gene in breast cancer and mutations in this gene are found in over one third of all breast cancers (Cerami et al., 2012; Gao et al., 2013; Bachman et al., 2004). This high rate of mutation would make *PIK3CA*-mutant breast cancer, specifically, the third most common type of cancer

diagnosed in women, behind only breast cancer as a whole and lung cancers (Siegel et al., 2023).

The location and frequency of *PIK3CA* mutations in breast cancer is worth highlighting. Mutational analyses of breast tumors and breast tumor cell lines has shown that most mutations cluster in ”hotspot” regions in exons 9 and 20, corresponding to the helical and catalytic domains of *PIK3CA*, respectively (Bachman et al., 2004; Saal et al., 2005). In fact, Bachman's 2004 paper in Cancer Biology & Therapy expanded upon *PIK3CA* mutational status in breast cancers from a previous study examining the frequency of *PIK3CA* mutations in colorectal and other human cancers, where the sample size of breast cancers was much smaller (Samuels et al., 2004). From this study, it was thought that *PIK3CA* mutations in breast cancer were an uncommon event. They also observed more mutations in the helical (47%) versus kinetic (33%) domain of *PIK3CA* across various cancers. With a larger sample size, specifically that of breast cancer tissue alone, Saal et al. demonstrated that *PIK3CA* mutations are present in a quarter of human breast cancers. In fact, using sequencing data from forty-one primary breast cancers, Saal et al. identfied nine that contain somatic missense mutations. Of the nine, two mutations were found in exon 1, one in exon 9, and the remaining within exon 20. Confirming their somatic origin, DNA from adjacent normal mammary or lymphoid tissue for each sample where a mutation was detected revealed the wild-type sequence. The somatic nature of these mutations shows that these mutations are acquired after birth and that *PIK3CA* mutations do not represent a germline risk variant that presents heritable risk. In addition to primary breast tissue samples, sequencing data from that of 12 well-established, commonly used human breast cancer cell lines showed that 33%, or four of the twelve, harbor *PIK3CA* mutations, half of which were in exon 9, the other in exon 20 (Bachman et al., 2004). These findings contrast that of Samuels et al. where their sample tissue of origin, which was mostly colorectal, did not implicate an oncogenic role for *PIK3CA* in breast cancers. As such, the mutational profiling of *PIK3CA* in primary breast tissue and breast cancer cell lines in Bachman et al.

illuminated *PIK3CA* as the most mutated oncogene in breast cancer. Furthermore, the positions of the mutations within *PIK3CA* were predominantly within exon 20, or the kinase domain, as opposed to exon 9 for colorectal cancer in Samuels et al. These findings are further supported by the sequencing of breast cancer cases collected within cBioPortal, which contains a collection of many clinical studies and sequencing data (Cerami et al., 2012; Gao et al., 2013). The cBioPortal maintains that *PIK3CA* mutations are more common (36% in all breast cancer samples) than *TP53* mutations (34% in all breast cancer samples) which is another gene that is commonly mutated in breast cancer but occasionally reported as the most frequently mutated gene in breast cancer.

Saal's 2005 paper in Cancer Research corroborated the findings of Bachman et. al with a mutational analysis of *PIK3CA* in two large cohorts of breast tumors, totaling 295 breast tumor samples, and a diverse range of breast tumor cell lines; Saal et al observed a *PIK3CA* mutation rate of 26% in primary breast tissues, or 77 of 292, and 28% of breast cancer cell lines, or 14 of 50. These *PIK3CA* mutations rates are very similar to the those observed for breast tumors and cell lines in Bachman et al. When they looked more closely at the positions of *PIK3CA* mutations, they also found that those of exon 20 predominate over exon 9. Beyond the exon 20 and exon9 "hotspots", Saal et al. observed a third "hotspot" in exon 7. Together, the findings of both Bachman et al. and Saal et al. provided strong evidence of a functionally important for *PIK3CA* in breast cancer tumor growth. Thus, their research laid the incentive to further dive into *PIK3CA* function and how that changes depending on the mutations present within the gene in breast cancers, specifically.

1.1.1 The role of PI3K in biochemical signaling and cancer

PIK3CA is a member of the family of lipid kinases, called Phosphatidylinosoitol 3-Kinases (PI3Ks), that catalyze the transfer of the gamma phosphate group on ATP to the 3'-hydroxyl group of phosphoinositides. This reaction triggers a cascade of intracellular signaling events that range from regulation of cell survival, proliferation, and motility to vesicle

trafficking and structural and mechanical modeling of the cell cytoskeleton (Ghigo et al., 2015; Engelman et al., 2006). Given the role of the PI3K pathway in a diverse range of cellular processes, it has become evident that understanding PI3Ks and their role in eukayortic development is essential to understand cellular disease states, especially in cancer. To start, PI3Ks are grouped into three classes (I-III) based on their structure and substrate specificity. In mammals, class I PI3Ks are further subdivided into subclasses IA and IB due to differing modes of regulation; class IA kinases bind to or are activated by growth factor receptor tyrosine kinases (RTKs), whereas G-protein-coupled receptors for class IB kinases (Table 1.1). Class IA PI3Ks are composed of a p110 catalyic subunit and a p85 regulatory subunit. Three genes - *PIK3CA*, *PIK3CB*, and *PIK3CD* - encode p110α, p110β and p110δ, respectively, three highly homologous p110 subunit isoforms. These class IA catalytic isoforms couple with any of five isoforms of the p85 regulatory subunit: p85α, p85β, p55α, p50α and p55γ. p55α and p50α are splicing variants of p85α that arise through alternative transcription-initiation sites and are encoded by *PIK3R1*; *PIK3R2* for p85β; and finally *PIK3R3* for p55γ (Thorpe et al., 2015; Fox et al., 2020). Of the five, p85α, which acts as a tumor suppressor by regulating and stabilizing p110α, is the most abundantly expressed isoform in most normal tissues and its expression is often reduced in cancer (Ueki et al., 2002; Shekar et al., 2005; Taniguchi et al., 2010). *PIK3R1*, while implicated in breast cancer and other cancers like endometrial and pancreatic, occurs at a lower frequency than mutations in *PIK3CA* (Koboldt et al., 2012; Jaiswal et al., 2009; Urick et al., 2011). Finally, class IB PI3Ks are heterodimers consisting of regulatory subunits, p101 and p87, encoded by *PIK3R5* and *PIK3R6*, respectively, and a p110γ catalytic subunit, encoded by *PIK3CG* (Thorpe et al., 2015; Fox et al., 2020).

Class I PI3 kinases			
Gene Name	**Protein Name**	**Subclass**	**Stimulating Receptor Class**
PIK3CA	p110α	IA	RTK (Receptor Tyrosine Kinase)
PIK3CB	p110β	IA	RTK
PIK3CD	p110δ	IA	RTK
PIK3CG	p110γ	IB	GPCR (G-protein Coupled Receptor)

Table 1.1: **PI3K Summary Table.** This table summarizes the different class I PI3K catalytic subunits.

It's important to note that p110α, which is encoded by the *PIK3CA* gene, is one of the most frequently mutated oncogenes in cancer (Gymnopoulos et al., 2007; Samuels et al., 2004). first revealed the high frequency of somatic mutations in the *PIK3CA* gene in various human cancers, including those of the brain, breast, colon, gastric, and lung (Samuels et al., 2004). The high frequency of mutation in this particular gene across cancers shows that this gene has distinct capacity to influence the hallmarks and onset of cancer compared to the other classes of this kinase and other subunits of the PI3K complexes.

To further expand upon p110α of class I PI3Ks, this catalyic subunit specificlaly phosphorylates the membrane phospholipid, phosphatidylinositol 4,5-bisphosphate(PIP2). While this singular chemical reaction may seem rather insignificant in isolation, it actually sits at the hub of a crucial signal transduction cascade. Activation of the PI3K complex is stimulated by multiple growth factor receptor tyrosine kinases, and the phosphorylated form of PIP2 (PIP3) is cleaved to produce the secondary signaling molecules: inositol 1,4,5-trisphosphate(IP3) and diacyl glycerol. Both secondary signalers are key stimulants to a diverse array of cellular functions including cell proliferation and survival (Nussinov et al., 2021; Yang et al., 2019). The frequency of which *PIK3CA* and/or its direct antagonist, PTEN, a phosphatase that removes phosphorylation from PIP3, are mutated across all can-

cer types (over 1 in 6 across all cancer types), speaks to how critical this reaction is to the promotion of cell growth, which is the core hallmark of cancer (Cerami et al., 2012; Gao et al., 2013).

1.1.1.1 AKT and its associated substrates

To specifically illustrate the mechanisms of the role PI3K signaling plays in cell growth and proliferation, a description of the downstream actors in the pathway is provided. The next major actor in the PI3K biochemical signaling cascade is protein kinase B, also better known as AKT. AKT is a family of three serine/threonine protein kinase isoforms (AKT1, AKT2, and AKT3). AKT is activated by interaction with the cleaved PIP3 molecule in addition to phosphorylation at the threonine-308 residue by phosphoinositide-dependent kinase 1 (Nitulescu et al., 2018). The activation of AKT by the product of the PI3K reaction is what makes AKT a downstream actor in the PI3K signaling pathway. Activated AKT has a number of key substrates that promote downstream promotion of cell growth and oncogenesis.

The first AKT substrate to play a key role in the development of cancer is BAD. BAD is a member of the bcl-2 family of proteins that forms heterodimers with the proteins Bcl-XL and Bcl-2. This activity of BAD allows for cells to undergo apoptosis in the presence of DNA damage. AKT however, interferes with this function of BAD by phosphorylating the protein at the serine-136 residue. This phosphorylation of BAD prevents the protein from complexing with the Bcl-XL and Bcl-2 proteins. Thus the activity of AKT is able to support a cellular phenotype that is resistant to apoptosis associated with DNA damage (Hayakawa et al., 2000; Nitulescu et al., 2018). This particular action of AKT can make cells resistant to the DNA damage induced by platinum chemotherapy drugs as well (Dasari and Tchounwou, 2014). Thus, the interaction of AKT and BAD plays influences both the ”Genome instability and mutation” and ”Resisting cell death” hallmarks of cancer (Hanahan, 2022).

Another critical AKT substrate with implications in oncogenesis are the FOXO family

of transcription factors (TFs). The FOXO, forkhead box class O, TF family is comprised of four member proteins (FOXO1, FOXO3, FOXO4, and FOXO6) each of which are expressed at different levels across tissue types in the body. FOXO TFs are responsible for binding DNA and promoting the expression of key genes that modulate a cells progress through the cell cycle. Key factors that are increased in expression by the DNA-binding activity of FOXO family TFs are p27 and p21. Both of these proteins inhibit members of the cyclin-dependent kinase (CDK) family of proteins and slow a cell's progress through the cell cycle (Nitulescu et al., 2018; Wang et al., 2014). Other key factors influenced by the activity of the FOXO family of TFs is are the TRAIL and FAS ligand proteins which can promote the initiation of the extrinsic apoptosis pathway (Wang et al., 2014; Davidson et al., 2024; Pommier et al., 2004). Phosphorylation of FOXO proteins by AKT prevents the nuclear translocation of the TFs and thus their ability to bind DNA and influence the gene expression program of the cell (Wang et al., 2014). The role that AKT plays in the inhibition of FOXO TF activity collectively promotes progression of a cell through the cell cycle and limits the cell's potential to activate extrinsic apoptosis.

The next key AKT substrate is TSC2 (tuberosclerosis complex 2). TSC2 is a member of the tuberosclerosis complex which is a heterodimer of the TSC1 and TSC2 proteins. Mutations in either of these genes contributes to the onset of tuberosclerosis, the namesake of the complex (Crino et al., 2006). AKT phosphorylates TSC2 at a number of different residues but the serine-939 and threonine-1462 sites have been most implicated with influence on the downstream signaling of the pathway. Phosphorylation of TSC2 by AKT prevents the function of the TSC complex to inhibit the Rheb-GTPase. This allows Rheb-GTP to activate the MTORC1 complex (Huang and Manning, 2009). Activity of the MTORC1 complex has a variety of key substrates of its own that promote growth and proliferation of cells.

1.1.1.2 mTORC1 and its associated substrates

The mTORC1 complex is one of two major complexes to feature the mammalian target of rapamycin (mTOR) serine/threonine kinase as its key catalytic constituent. What differentiates the mTORC1 complex from its sister complex mTORC2 is the scaffolding protein known as the regulatory associated protein of mTOR, or RAPTOR. the mTORC2 complex replaces RAPTOR with a different scaffolding protein called the rapamycin-insensitive companion of mTOR, or RICTOR. These different scaffolding proteins change the binding partners of the different mTOR complexes. It is the difference in these associated protein partners that dictate different downstream substrates of the mTOR kinase (Tian et al., 2019). While mTORC2 activation has some cross talk with the activity of mTORC1, they generally have a negative feedback relationship in which mTORC1 activity indirectly inhibits mTORC2 activity. mTORC2 however, can be activated from biochemical pathways independent of PI3K activity. This is critical in that a key substrate of mTORC2 is AKT. This means that mTORC2 activity can positively upregulate the activity of mTORC1 (Szwed et al., 2021).

The primary substrate of the mTORC1 complex to play a major role in cancer is 4EBP1. 4EBP1, or 4E binding protein 1, binds to the translation initiation complex, eukaryotic initiation factor 4F (eIF4F), particularly at the subunit eIF4E. eIF4E binds to the 5' cap on mature mRNA molecules and recruits the other members of the eIF4F complex to promote ribsomal binding and translation of the bound mRNA molecule. However, the interaction between 4EBP1 and eIF4E prevents the recruitment of the other members of the eIF4F complex, particularly eIF4G. In summary, 4EBP1 is an impediment to the translation of many mRNA molecules (Ma and Blenis, 2009). Phosphorylation of 4EBP1 by the mTORC1 complex leads to dissociation of the 4EBP1/eIF4E binding and allows for increased translation of mRNA. 4EBP1 preferentially binds to specific mRNA molecules and this function of mTORC1 allows for increased translation of those mRNA molecules. mRNA that undergo increased translation in the presence of phosphorylated 4EBP1 include

VEGF-A (a regulator of angiogenesis), Bcl-2 (an inhibitor of apoptosis), and HIF1α (a hypoxia associated TF) (Qin et al., 2016). All told, the phosphorylation of 4EBP1 promotes oncogenesis through the indirect promotion of angiogenesis and inhibition of apoptosis. It is important to keep in mind that activity of the mTORC1 complex is necessary for the normal growth and proliferation of cells, but that over activation of this complex can lead to the aberrant increased translation of specific proteins.

Another regulatory substrate of the mTORC1 complex is S6K. S6K, or ribosomal S6 Kinase, is a phosphorylated substrate of the mTORC1 complex that has been implicated in cancer progression and particularly that of ER-positive breast cancer progression (Tian et al., 2019). Phosphorylation of S6K by mTORC1 promotes activity of the kinase and a key substrate of S6K is ERα. Phosphorylation of ERα by S6K creates a unique feed-forward feedback loop in which ERα increases transcription of the S6K gene *RPS6KB1*. This feedback loop further supports the gene expression profiles promoted by the ERα TF which regulates the expression of many oncogenic genes (Sridharan and Basu, 2020).

1.2 State of PI3K inhibition for breast cancer in the clinic and beyond

1.2.1 Currently approved therapies

1.2.1.1 Alpelisib

Despite the crucial role *PIK3CA* plays in breast cancer, and the high prevalence of mutations, there are currently only two approved targeted therapies for PIK3CA mutant breast cancer. The first of these is Alpelisib. Alpelisib is an isoform specific inhibitor that specifically targets the p110α subunit encoded by *PIK3CA*. This drug unfortunately has many key shortcomings that limit its efficacy in treating *PIK3CA* mutant breast cancer. In the phase III clinical trial, SOLAR-1, in which alpelisib was approved for use in *PIK3CA*-mutant breast cancers, the overall response rate to alpelisib was 26.6% across all patients in the treatment arm. While any improvement to a patient's condition should be considered a positive, this overall response rate pales in comparison to the number of patients that ex-

perienced adverse events while taking alpelisib. Of the patients in the treatment arm of the study, 76% experienced a grade 3 or greater adverse event. These grade 3 events included hyperglycemia (64.8%), diarrhea (59.5%), and rash (36.3%). This is more than double the rate of patients in placebo group (35.5%). Of significant note, 25% of the patients in the treatment arm withdrew from treatment due to the adverse events associated with taking alpelisib (André et al., 2021). The most common adverse event impacting patients that withdrew from the trial was hyperglycemia. While alpelisib is an isoform specific inhibitor of the p110α subunit, it is not a mutation-specific inhibitor and it inhibits the function of wild type and mutant PI3K alike. PI3K plays a key role in regulating glucose metabolism, in addition to the role it plays in cell proliferation, and inhibiting its normal functions leads to the metabolic disruption experienced by many of the patients in this trial (Hoxhaj and Manning, 2019). This indiscriminate inhibition of PI3K function disrupts the function of normal cells as much as cancerous ones which leads to high levels of toxicity in patients.

1.2.1.2 Capivasertib

The other clinically approved therapy for *PIK3CA*-mutant breast cancer is capivasertib. Capivasertib also known as TRUQAP, is an AKT inhibitor produced by AstraZeneca. AKT is a downstream kinase actor in the PI3K signaling pathway with many substrates that affect cellular proliferative capacity. AKT has a strong affinity for the secondary signaling molecule IP_3 produced by the function of the PI3K complex(Miao et al., 2010; Manning and Cantley, 2007). AKT inhibition using capivasertib has been shown to reduce signaling in all three isoforms of AKT in preclinical models (Davies et al., 2012). Despite targeting a different protein complex, capivasertib suffers from many of the same shortcomings that affect alpelisib. In the phase II clinical trial, FAKTION, patients taking capivasertib saw an increase in progression-free survival and overall survival by about 6 months over the placebo (10 months to 4 month progression-free and 29 months to 23 months overall). The difference in progression-free survival was the same in patients that had demonstrated

perturbation in the PI3K pathway, but the overall survival was improved for this subset of patients (39 months to 20 moths overall survival). In the FAKTION trial, 81% of patients in the treatment arm experienced hyperglycemia with a third of those cases being recorded as grade 3 (Howell et al., 2022). Objective response rate to capivasertib treatment was 22.9% in the total population of patients within the treatment arm of the CAPItello-291 phase III clinical trial. This response rate translated to an median increase in progression free survival of 3.5 months. The objective response rate was slightly higher at 28.8% in the subset of patients that specifically had a gene mutation in a gene associated with the AKT pathway, which led to an increased median progression-free survival of about 4 months. 39.2% of patients in the treatment arm of the CAPItello-291 trial experienced a grade 3 adverse event and 34.9% of patients required a dose interruption due to adverse events. Common adverse events experienced by patients taking capivasertib include rash, diarrhea, and hyperglycemia which are all adverse events that were also experienced by patients in the clinical trial for alpelisib (Turner et al., 2023).

Taken together, the data surrounding these two targeted therapies paint an harsh picture for the current state of treatment for *PIK3CA*-mutant breast cancer. Both of these therapies target major actors in this key oncogenic pathway, but the rate at which patients see positive response to the drug and experience adverse reactions to the treatment are very similar. These treatments present clinicians with a difficult decision in which they must weigh a patients ability to cope with potential adverse effects to their potential to see clinical improvement with one of these drugs. All of these decisions get further complicated by the interaction between these drugs and other diseases like diabetes (Ekanayake et al., 2022). Patients with preexisting type I diabetes and uncontrolled type II diabetes were excluded from both the SOLAR-1 trial for alpelisib and the CAPItello-291 trial for capivaseritib due to anticipated complications involving these therapies and hyperglycemia (Ge et al., 2023; Turner et al., 2023).

1.2.2 Mutation-specific therapies

Interestingly, *PIK3CA* mutations in breast cancer are highly dichotomous with over 40% of these mutations being either E545K or H1047R specifically (Cerami et al., 2012; Gao et al., 2013). These two mutations occur in different domains of the encoded protein and may have a differential impact on the function of PI3K. The E545K mutation occurs within the helical activation domain of the subunit, while the H1047R mutation is within the C-terminus domain of the protein. Mutations at both of these residues show an increase in enzymatic activity, however there are also studies that demonstrate that the downstream effects of these mutations are quite different (Gabelli et al., 2010; Hart et al., 2015). In one of these studies, the researchers demonstrated that the *PIK3CA* mutant genotype of breast cells impacted how breast cells responded to the targeted treatment lapatinib, which is an inhibitor of growth factor receptors. It was found that the E545K mutant cells were sensitive to the treatment, but that H1047R mutant cells continued to activate downstream AKT signaling (Garay et al., 2021). Findings like this suggest that additional study into the function and role of these mutations is necessary to understand mechanisms of resistance to existing therapies and uncover new less toxic means of treating cancers with these mutations.

1.2.2.1 LOXO-783

There is currently one mutation-specific targeted therapy of note that is in phase I clinical trials. This inhibitor is currently known as LOXO-783 and is produced by Loxo Oncology at Eli Lilly (Klippel et al., 2021). This inhibitor is a targeted inhibitor of H1047R *PIK3CA*-mutant PI3K complex specifically (Klippel et al., 2021; Puca et al., 2022). Preclinical models testing this inhibitor show very specific inhibition of the H1047R mutant PI3K with low selectivity in cell models without *PIK3CA* mutations. This molecule also shows high penetrance to the central nervous system in *in vivo* models (Klippel et al., 2021). These factors make this inhibitor an especially strong candidate to improve the standard

of care for patients with H1047R *PIK3CA*-mutant disease. The promise of this inhibitor is that it could be a less toxic alternative to the currently approved therapies that which are not mutation-specific. The CNS penetrance of this inhibitor also provides hope that this molecule could be effective in the treatment of metastatic disease as well. The phase I clinical trial for LOXO-783 called PIKASSO-01 has an estimated completion date of May 2025 and there are no associated publications for this trial yet (Chau and Loxo Oncology, 2024).

1.2.2.2 STX-478

STX-478 is another small molecule inhibitor of PI3Kα produced by Scorpion Therapeutics that is hailed as another mutation-specific therapeutic of H1047R *PIK3CA*-mutant breast cancer. This molecule, like LOXO-783, is currently in a phase I clinical trial. STX-478 binds a non-mutated allosteric site within the p110α protein. This unique binding site is distinct from the binding targets of other similar inhibitors and is supposedly inaccessible in the WT version of the protein. Despite demonstrating specific efficacy in H1047 residue mutant cell line models and *in vitro* enzyme inhibition assays, STX-478 showed in preclinical *in vivo* models to have efficacy in slowing tumor growth of tumors containing other *PIK3CA* mutations. STX-478 was effective in slowing the growth of PDX models containing hotspot E545K mutations and the less common H1047L and E542K mutations (Buckbinder et al., 2023). While these results are promising for the wider applicability of this therapeutic, they do call into question the actual mutation-specificity of the molecule. The binding pocket is seemingly only accessible in the H1047 residue mutants per some assays, but other assays show efficacy in E545 and E542 residue mutants as well. Results from the ongoing clinical trials will hopefully clarify which mutant versions of the p110α protein are most susceptible to this means of inhibition.

1.2.2.3 RLY-2608

RLY-2608 is a particularly encouraging mutation-specific inhibitor. RLY-2608 has unique binding affinity for an allosteric site that is only present in the H1047R mutant version of the p110α protein. This binding pocket was discovered using CRYO-EM imaging of the mutant protein and this inhibitor was designed to fit that pocket. The most encouraging aspect of this inhibitor is that its inhibitory activity was effective in preclinical *in vivo* models and did not induce hyperglycemia in these models. Hyperglycemia is a common treatment limiting adverse effect that is reported in many of the clinical trials described with regards to PI3K inhibition. Reducing this specific side effect could be a major step in improving the quality of care for patients (Varkaris et al., 2024). The phase I clinical trial (NCT05216432) has not yet concluded or published results, but a phase II trial has already begun enrollment (NCT05216432) (Belli et al., 2023).

1.2.2.4 Invasolisib

Invasolisib is a PI3Kα specific inhibitor with mutation-specific effect in preclinical models. In cell line models and mouse *in vivo* models, invasolisib showed specific inhibition of multiple mutant variants of PI3Kα with minimal disruption to the WT version of the protein complex. Invasolisib works by marking PI3Kα proteins for degradation. Results of the phase I clinical trial for use in breast cancer (NCT03006172) have yet to be published, but a phase III trial is currently underway (INAVO-121) (Song et al., 2022; Juric et al., 2023).

1.2.3 Other PI3K inhibitors of note in breast cancer treatment

1.2.3.1 Buparlisib

Aside from these currently approved or in trial therapies, there are a number of other notable PI3K inhibitors. Among these is buparlisib. Buparlisib is a pan-PI3K inhibitor that inhibits not only the α isoform like alpelisib, but also inhibits β, γ, and δ isoforms as well (Fuso et al., 2022). The results of the buparlisib clinical trials showed that treatment with this molecule had only modest effects on the improvement of patients' condition while also

being associated with similar adverse events to other PI3K inhibitors. In the BELLE-2 phase III clinical trial, hyperglycemia was reported as an adverse effect in 43.1% of patients in the treatment arm of the study and 13.2% of patients withdrew from the study due to adverse effects(Fuso et al., 2022; Baselga et al., 2017). Buparlisib has high penetrance of the blood-brain barrier and an increase in depression and anxiety were reported in 27% and 23% of patients respectively in the BELLE-2 trial. Suicidal ideation was also reported in other trials for this drug (Fuso et al., 2022). The frequency of these toxic adverse events in combination with low overall response rates (11.8% in the BELLE-2 trial, which was only 4% higher than the overall response of patients receiving standard of care in the placebo arm of the study) has ultimately prevented buparlisib from receiving full FDA approval in the clinic (Baselga et al., 2017).

1.2.3.2 Pictilisib

Pictilisib is another PI3K inhibitor that was evaluated in clinical trials for use in breast cancer treatment. Pictilisib inhibits both the α and δ isoforms of the PI3K complex, but shows little capacity to inhibit the β and γ isoforms. This limited isoform inhibition is a major distinguishing characteristic of pictilisib from buparlisib which inhibits all four of the major isoforms of the PI3K complex (Fuso et al., 2022). Currently pictilisib has been discontinued due to poor results from the PEGGY phase II clinical trial. In the PEGGY clinical trial, 70.3% of patients in the treatment arm discontinued treatment. Of the patients that discontinued treatment, 56.3% of the patients had progression of disease and needed to change treatment and 23.4% of the discontinuing patients withdrew due to adverse effects associated with treatment (Vuylsteke et al., 2016). These results depict pictilisib as a toxic therapeutic with low efficacy.

1.2.3.3 Eganelisib

Eganelisib is a unique PI3K inhibitor in that it is an isoform specific inhibitor that exclusively targets the gamma isoform of PI3K. The gamma isoform is distinct from the other

class I PI3K complexes in that it is activated by G-protein coupled receptors instead of receptor tyrosine kinases. The gamma catalytic subunit is also rarely mutated in breast cancers and mutations in the *PIK3CG* gene occur in roughly 1.5% of breast cancer patients (Cerami et al., 2012; Gao et al., 2013). The goal of targeting PI3Kγ was not to affect the cancer cells themselves, but to rather reactivate myeloid cells in the tumor microenvironment to act against the tumor. Eganelisib was evaluated for use in solid cancers in combination with nivolumab in the dose escalation phase I MARIO-1 clinical trial (This trial enrolled 29 triple negative breast cancer patients). In the monotherapy portion of the trial there were no dose-limiting toxicities associated with eganelisib treatment alone. Only three patients in this arm of the trial withdrew from treatment due to treatment related adverse events with the most commonly reported event being increased hepatic aminotransferases. It is also important to note that one patient achieved partial response with eganelisib monotherapy (this patient had peritoneal mesothelioma). In the combination therapy dose escalation arm with nivolumab there were a few dose-limiting toxicity events reported at some of the higher doses which included maculopapular rash and increased hepatic aminotransferases. Of the patients in this arm of the trial 2 reported a partial response of their disease (these patients had adrenal cell carcinoma and gallbladder carcinoma). Overall the results of the MARIO-1 trial depict eganelisib as a therapy with low toxicity (Hong et al., 2023). The ongoing MARIO-3 phase II trial will further evaluate the effects of this inhibitor on triple negative breast cancer more specifically.

1.2.3.4 Copanlisib

Copanlisib is a PI3K inhibitor with specific affinity for the alpha and delta isoforms of the complex. Copanlisib was evaluated in a unique tumor agnostic phase II trial for patients with known *PIK3CA* mutations (NCI-MATCH). The overall response rate to copanlisib treatment was 16%. Of the 22 patients evaluated for efficacy of treatment, 4 reported a partial response. The two patients in this arm of the study with breast cancer reported pro-

gressive disease and discontinued treatment. Thirty patients were included in the arm of the study that evaluated the toxicity of copanlisib treatment. 63% of the patients in this arm of the study reported with grade 3 or higher hyperglycemia with rash and hypertension being the other most commonly reported grade 3 or higher adverse events. Progression of disease was however the most common reason for discontinuing treatment (68%) (Damodaran et al., 2022; Markham, 2017).

1.2.3.5 MEN1611

MEN1611 is a delta-sparing inhibitor of PI3K activity. PI3Kδ is primarily expressed in immune cells and particularly B cells. The principle behind the delta-sparing nature of MEN1611 is that it will inhibit the growth of cancer cells with its inhibition of the alpha and beta isoforms, but will not disrupt B cell function which could improve immune response compared to pan-PI3K inhibitors. MEN1611 showed some mutation-specificity in its selectivity in cell line models and was effective in slowing tumor growth in xenograft models. It is important to note however, that the xenograft models used to evaluate MEN1611 did not harbor the common hotspot *PIK3CA* mutations, but instead featured a variety of less common mutations (K111E, E545A, and C420R) (Fiascarelli et al., 2023). This is a bit of an odd choice given the cell line models all featured the prominent H1047R or E545K mutations. The B-PRECISE-01 phase I clinical trial is currently underway. Early results show 36.5% of patients exhibited partial or complete response which is very encouraging. 51.6% of patients however experienced treatment interruptions due to adverse events with hyperglycemia being the most common and 14.5% of patients were forced to discontinue treatment due to adverse effects (Piccart et al., 2023).

1.2.3.6 CYH-33

CYH-33 is a PI3Kα specific inhibitor. CYH-33 is being developed by Haihe Biopharma and is currently in phase I clinical trial (NCT04856371)for use in hormone receptor positive breast cancer. CYH-33 profiles similarly in terms of its inhibition to alpelisib, however in

preclinical models it did demonstrate some unique properties with regards to immune cell recruitment to the tumor site of xenograft models in mice (Sun et al., 2021).

Summary of Inhibitors		
Inhibitor Name	**Inhibitor Target**	**Clinical Approval**
Alpelisib	PI3K (alpha)	Approved
Capivasertib	AKT	Approved
LOXO-783	H1047R mutant PI3K alpha	In Trial (Phase I)
STX-478	H1047R mutant PI3K alpha (possibly other residues)	In Trial (Phase I)
RLY-2608	H1047R mutant PI3K alpha	In Trial (Phase II)
Invasolisib	mutant PI3K alpha	In Trial (Phase III)
Buparlisib	PI3K (alpha, beta, gamma, and delta)	Not approved (Phase III)
Pictilisib	PI3K (alpha and delta)	Not approved (Phase II)
Eganelisib	PI3K (gamma)	In trial (Phase II)
Copanlisib	PI3K (alpha and delta)	In trial (Phase I) for breast cancer approved for use in follicular lymphoma
MEN1611	PI3K (alpha, beta, and gamma)	In trial (Phase I)
CYH-33	PI3K (alpha)	In trial (Phase I)

Table 1.2: **PI3K Inhibitors in breast cancer summary table.** This table summarizes the discussed targeted therapies used for treatment of *PIK3CA*-mutant breast cancer.

There are a number of other PI3K inhibitors that are being evaluated for use in other cancer types and are not currently being tested for efficacy in breast cancer specifically. Mutations in PIK3CA are common across a variety of cancer types, especially epithelial cancers like lung and colorectal cancer types (Belli et al., 2023; Sirico et al., 2023). A

summary of these other inhibitors that are in ongoing clinical trials can be found in Table 1.3.

Summary of Inhibitors			
Inhibitor Name	**Inhibitor Target**	**Cancer Type**	**Clinical Approval**
Taselisib	PI3K (alpha and gamma)	Refractory cancers, lymphoma, and multiple Myeloma	In trial (Phase II)
GSK2636771	PI3K (beta)	Solid Tumors broadly	In trial (Phase II)
Duvelisib	PI3K (gamma and delta)	Melanoma and Head&Neck	In trial (Phase II)
IOA-244	PI3K (delta)	Solid Tumors broadly	In trial (Phase I)
Linperlisib	PI3K (delta)	Thymic Cancer	In trial (Phase II)
TQ-B3525	PI3K (alpha and delta)	Follicular Lymphoma	In trial (Phase II)
Pilaralisib	PI3K (all isoforms)	Not in Trial	Not in Trial
Sonolisib	PI3K (all isoforms)	Not in Trial	Not in Trial
Sonolisib	PI3K (alpha)	Not in Trial	Not in Trial
SAR260301	PI3K (beta)	Not in Trial	Not in Trial
AZD8154	PI3K (gamma and delta)	Not in Trial	Not in Trial
AMG319	PI3K (delta)	Not in Trial	Not in Trial
Idelasib	PI3K (delta)	Lymphoma	Not in Trial
Umbralisib	PI3K (delta)	Lymphoma	Not in Trial
Nemiralisib	PI3K (delta)	Not in Trial	Not in Trial
Leniolisib	PI3K (delta)	Not in Trial	Not in Trial
Parsaclisib	PI3K (delta)	Not in Trial	Not in Trial
Seletalisib	PI3K (delta)	Not in Trial	Not in Trial

Table 1.3: **PI3K Inhibitors in other cancers summary table.** This table summarizes PI3K targeted therapies used for treatment of cancer types other than breast cancer. Adapted from Belli et al. (2023)

Summary of Inhibitors			
Inhibitor Name	**Inhibitor Target**	**Cancer Type**	**Clinical Approval**
Zandelisib	PI3K (delta)	Not in Trial	Not in Trial
SHC014748	PI3K (delta)	Not in Trial	Not in Trial
KA2237	Unknown specificity	Not in Trial	Not in Trial
HMPL-689	Unknown specificity	Not in Trial	Not in Trial

Table 1.3: **Continued**

1.3 Isogenic Cell Line Models

One of the biggest challenges in cancer biology is determining which mutations in a given tumor are contributing to the tumor's growth and which mutations offer a therapeutic vulnerability. To understand the role of individual mutations, isogenic cell line models are an important tool for studying single mutations in near isolation. Isogenic cell line models are model systems created using genome editing to specifically introduce a mutation of interest into an immortalized cell line.

The origin of using isogenic cell line models begins with what are considered classic genetic knockout experiments. Knockout experiments first became prominent in the early 1990s and were used to characterize the function of genes on specific processes and phenotypes (Nelson, 2015). In these experiments, specific genes would be mutated, typically within the embryos of model organisms and the impact of these modifications on various processes would be evaluated. In one of the first knockout experiments, The gene α*CAMKII* was disrupted in a mouse model. Mice deficient in CAMKII exhibited reduced memory and the functional link between this enzyme and neurological development was established Silva et al. (1992).

The advent of more advanced genetic modification tools has allowed for the more spe-

cific modification of genes. Older genetic modifications could disrupt the gene sequence or induce random mutations into the gene sequences, but lacked the specificity to induce single base pair alterations. Many cancer mutations involve single base pair substitutions in key genes (Bacolla et al., 2014). To properly study these mutations, newer techniques like recombinant AAV and CRISPR-Cas9 editing must be used to introduce the specific mutation with minimal off target effects. AAV editing, which stands for adeno-associated virus editing, uses adeno viruses to deliver recombinant single stranded DNA molecules with the desired sequence. These molecules are incorporated into the genome using the cell's native homology directed repair mechanisms (Bijlani et al., 2021). CRISPR-Cas9 editing, which stands for clustered randomly interspersed palindromic repeats, is the technology that received the Nobel prize for chemistry in 2020. CRISPR-Cas9 gene editing is actually more focused on the Cas9 part of the name. Cas9 is a bacterial nuclease enzyme that cuts DNA specifically at the sequence specified within its partner guide RNA (gRNA). By creating custom guide RNA molecules, genomic edits can be introduced at specific sites in the genome (Jinek et al., 2012; Cong et al., 2013). Specific knock-in edits can be introduced using homology directed repair with an accompanying template molecule.

An example of an isogenic cell line model is our MCF-10A model for *PIK3CA* mutations. This model was created using AAV to independently introduce the E545K and H1047R mutations into the genome of the MCF-10A cell line (Gustin et al., 2009) (Figure 1.1). MCF-10A is a nontumorigenic, spontaneously immortalized cell line derived from noncancerous breast epithelium (Puleo and Polyak, 2011; Soule et al., 1990). The key advantage of this model is that the MCF-10A parental cell line is largely unmutated. This allows users of this model to study the effects of the *PIK3CA* mutations in a breast cell in near isolation without other cancerous mutations confounding the observations. In making comparisons between the parental MCF-10A cell line and the versions with each of the *PIK3CA* mutations, we can better understand the distinct effects that either of the mutations can induce in breast cells. This particular isogenic model has been used by many

researchers to study these mutations for more discovery based purposes (Hart et al., 2015; Wu et al., 2014) to even testing the efficacy of new targeted therapies like STX-478 (Buckbinder et al., 2023).

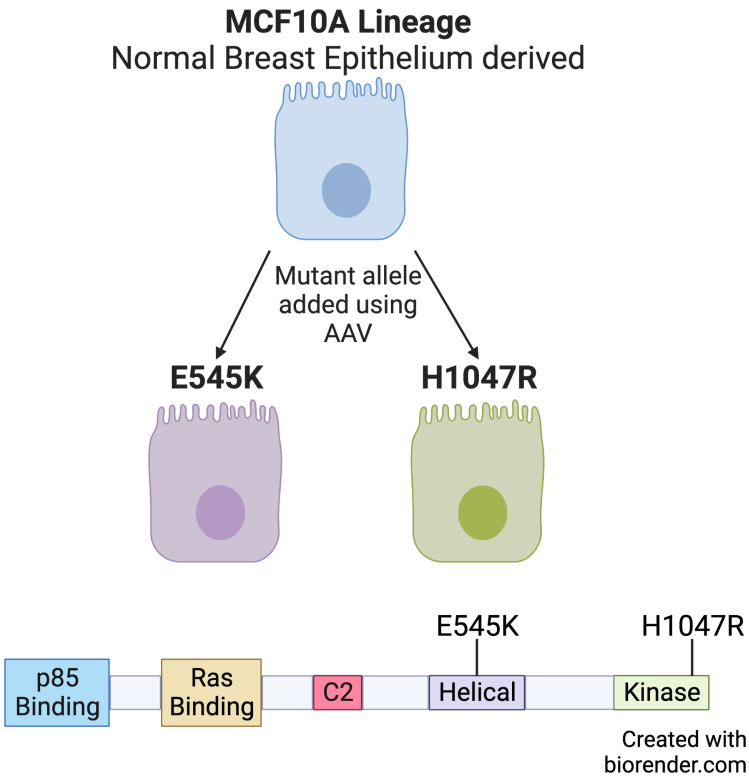

Figure 1.1: **MCF-10A *PIK3CA* Isogenic Cell Line Schematic**. This diagram shows the cell line model described in Gustin et al.(Gustin et al., 2009) as well as the relative locations and associated protein domains of the breast cancer hotspot *PIK3CA* mutations

.

1.4 Sequencing for molecular Profiling and Epigenetics in Cancer

1.4.1 DNA Sequencing

The advent of genomic sequencing was perhaps the most pivotal advancement for the field of precision medicine for cancer biology. The completion of the human genome project in 2004 established a draft of the healthy human genome (Lander et al., 2001; Hood and Rowen, 2013). This and later versions of the human genome provided researchers with a reference genome to use in comparison with cancer mutated genomes. Comparisons be-

tween cancer mutated genomes and reference human genomes have allowed for the identification of driver cancer mutations as well as biomarkers that increase risk for disease. Identification of these biomarkers has provided the basis for the field of precision oncology, which seeks to identify patient-specific treatment plans based on biomarkers present within a patient's disease.

There are a number of array-based assays that are used in the clinic to determine the presence of specific genetic biomarkers within a patient's tumor. These assays can detect the presence of specific known and actionable mutations within a tumor (Tortajada-Genaro et al., 2023). The shortcoming of these assays is that they can only inform the state of biomarkers on the array, but they do not provide any information beyond that.

Whole exome sequencing (WES), which was first used in breast and colorectal cancers in 2007, is now a standard clinical practice for many cancer patients (Hodges et al., 2007; Wheeler and Wang, 2013; Wood et al., 2007). WES allows clinicians to identify all gene coding mutations present within a tumor. Identifying key mutations within a tumor is crucial for many treatment decisions, as the presence of certain mutations may prompt a clinician to opt for a specific targeted therapy (Menzel et al., 2023). Whole genome sequencing (WGS) is also becoming more common as the cost for sequencing the whole genome has decreased. The advantage provided by WGS over WES is that WGS can identify mutations outside of genes and chromosomal rearrangements that WES cannot detect. This also means that there is more utility in WGS as a discovery tool to uncover new mutations for further study.

WES remains the preferred assay in the clinic as there are not enough actionable extragenic mutations to currently justify the added cost of WGS.

1.4.2 Gene Expression

Beyond the analysis of tumor genotypes, alternative sequencing techniques can uncover more insights about the molecular state of cancer cells. RNA sequencing (RNA-seq) is an-

other sequencing technique that is commonly used to make treatment decisions for cancer patients (Buzdin et al., 2019; Hedenfalk et al., 2020). RNA-seq is used to assess the relative levels of distinct mRNA molecules present within a tissue. With this information scientists and clinicians can determine which genes are being actively transcribed in a tumor. The profile of expressed genes in a given cell can inform a lot about the identity and state of cells. Changes in gene expression throughout development can determine cell fate decisions and can inform things like responses to stress (Kuo et al., 1992). Gene expression can be very informative for the development of normal cells, but can also provide a lot of insight into how cancer cells have begun to differ from the normal cells and how they might respond to various treatments. In comparing the gene expression profile of cells in a diseased or cancerous state to the normal/healthy gene expression profile, we can identify gene expression biomarkers that can indicate the presence of disease, and/or the mechanisms by which diseased cells have changed.

There are a number of expression-based biomarkers that can inform treatment decisions for patients. Identifying key overexpressed genes like HER2 in breast cancer can be crucial to treatment decisions as there are many targeted therapies, like trastuzumab, that can be highly effective for HER2 overexpressed cancers even when genetic modification is not present (Baselga, 2001; Buzdin et al., 2019; Perou et al., 2000). This is a key distinction between mutational profiling and gene expression assays. The biomarkers and targets of therapy uncovered by these different approaches can be totally different.

Despite the benefits of gene expression analysis in cancer, RNA sequencing is still not a commonly used tool in the clinic. The number of actionable expression biomarkers is still too low to justify what may be an increased cost to a patient depending on their insurance coverage for these types of analyses. Even limited microarray and panel based assays can be costly to a patient(Gordon et al., 2020; Ergin et al., 2022). An example of a gene expression test that has gained traction for clinical use in breast cancer is Oncotype DX. Oncotype DX is a limited gene expression assay which evaluates the RNA expression

level of 22 different genes. Assaying the expression of just this selection of genes creates a well validated prediction of a given patient's breast cancer subtype (Luminal A, Luminal B, Her2+, or Triple Negative) (McVeigh and Kerin, 2017). This quick classification can help clinicians make key treatment decisions in cases where a patient's subtype can be a crucial factor. For example, the use of aromatase inhibitors for the ER+ subtypes or trastuzumab for HER2 positive disease (Baselga, 2001; Xie et al., 2022). Both of these therapies have high efficacy in breast cancers of these specific subtypes and tests like Oncotype DX can help patients be connected to these treatments early. In the future, the hope is that access to these tests would be more accessible and affordable to all patients. Increased accessibility and use in the clinic would also improve the incentives for the development of more and similar gene expression based tests to be used in other cancer types as well.

Currently there are many consortium based initiatives that are assaying the gene expression of many tumor samples to create a resource for future study that will aid in the discovery of more actionable expression biomarkers. The largest of these resources is the Cancer Genome Atlas (TCGA), which is a joint project between the National Cancer Institute (NCI) and the National Human Genome Research Institute (NHGRI) (Grossman et al., 2016). This project, to date, has profiled the expression of over 25,000 patient samples across over 60 different primary sites. This publicly available resource allows for the robust study of gene expression across a number of tumor types with diverse mutational backgrounds. Large scale initiatives like TCGA will continue to push the field forward and incentivize increased expression profiling for all cancer patients. Recent efforts in breast cancer research specifically have used the larger number of patient RNA-seq samples (1,095) available in the TCGA-BRCA resource to further refine and improve upon the PAM50 classification of breast cancer subtypes, which were originally defined using data from 189 patients (Parker et al., 2009; Thennavan et al., 2021).

1.4.3 Chromatin Accessibility

Chromatin accessibility (ChrAcc) is an important epigenetic marker for transcriptional activity and gene regulation. This means that it not only denotes active genes in a cell or tissue, but noncoding enhancer elements that can influence the expression of genes. ChrAcc can also identify genes that may not be actively expressed, but may exist in poised state, prepared to be expressed should a key stimuli impact the cells. ChrAcc is not currently used in the clinic to make treatment decisions as there are very few, if any, targeted treatments that are poised to take advantage of the information provided by ChrAcc. Evaluation of chromatin accessibility can still be a valuable resource in better understanding the mechanisms underlying shifts in gene expression undergone by cancer cells as cells mutate and evolve (Corces et al., 2018).

For comparison, only 410 tumor samples have been assayed for chromatin accessibility in TCGA compared to the over 25,000 that have been assayed for gene expression (Grossman et al., 2016; Corces et al., 2018). This shows just how far this research has to go before clinical implementation becomes the norm. Despite the limitations in available data for chromatin accessibility in cancer patient samples, many studies show the promise of chromatin accessibility as a prognostic marker. One such study, Liu (2020), demonstrates a link between breast cancer patient survival and the accessibility of a particular locus within the promoter of *ESR1* (Liu, 2020).

There has also been some interesting work in which ChrAcc and gene expression have been combined to define and project the trajectory of Alzheimer's disease (Morabito et al., 2021). This study identified a number of relevant candidate cis-regulatory elements with high correlation to the onset of Alzheimer's disease. The identification of these elements implicates the transcription factors FOSL2 and NRF1 as key regulatory proteins that promote Alzheimer's onset when active in astrocytes and oligodendrocytes respectively (Morabito et al., 2021). In making the connection between these accessible gene regulatory regions and regulatory transcription factors new therapeutic targets have been identified. In

the future, similar studies could be used to predict clonal expansion, drug resistance and response to specific therapeutic strategies in the context of cancer.

1.5 Goal of This Dissertation

In this dissertation, I will demonstrate the utility of combining the discovery tools of isogenic cell lines and molecular profiling by sequencing. This dissertation outlines findings made through the application of multiple sequencing techniques used on isogenic cell line models of hotspot *PIK3CA* breast cancer mutations. The current state of targeted treatment for *PIK3CA* mutant breast cancer is lacking and new discoveries are needed to improve upon this. Our findings elucidate many distinct differences in cell signaling caused by these mutations, and also highlight potential mutation-preferential targets of treatment that can be used to improve upon current treatment paradigms for *PIK3CA*-mutant breast cancer. The results of this work build upon our understanding of these independent mutations and provide a framework for future study of oncogenic mutations using functional genomic analyses.

Data presented within the supporting appendix of this dissertation reflects work performed on an other project during my time in graduate school. While distinct from the primary story presented within, this project supports the overall goal and theme of the dissertation and my graduate studies. In this appendix we evaluate chromatin interactions with respect to the E545K *PIK3CA* mutation in a different isogenic breast cell line model.

CHAPTER 2

Hotspot *PIK3CA* mutations exhibit distinct impacts on breast cells

2.1 Hotspot mutations drive distinct response to PI3K inhibition

The dichotomous hotspot mutations of *PIK3CA* impact different domains of the encoded protein, but have seldom been evaluated in the context of having unique impacts on downstream signaling. A few studies highlight some key differences and observations, but the fact remains that clinically, all *PIK3CA* mutations in breast cancer are evaluated and treated as if they are the same (André et al., 2021; Hart et al., 2015; Janku et al., 2013). In the SOLAR-1 clinical trial, in which alpelisib was approved for use in *PIK3CA*-mutant breast cancer, there were no analyses performed to determine whether either hotspot mutation presented as a better responder to treatment (André et al., 2021).

In an MCF-10A isogenic model for these mutations, previously developed in the Park lab (Gustin et al., 2009), we evaluated the response of breast cells bearing either the E545K mutation or the H1047R mutation to alpelisib (Figure 2.1). The results of this assay show that E545K mutant cells demonstrate increased sensitivity to treatment with alpelisib compared to the response by H1047R mutant cells. These results are somewhat consistent with the findings of Janku et al.(Janku et al., 2013). Janku et al found that breast cancer patients with H1047R mutant disease were more sensitive to treatment with targeted therapies of PI3K pathway family members. While this study found that H1047R mutant disease were more susceptible to targeted therapies, this study was based on the use of other inhibitors of other pathway members and not direct targeting of the *PIK3CA* product using alpelisib. Ultimately, these results demonstrate that the *PIK3CA* mutant genotype of breast cancer can be a key predictive factor in drug response. The results of this experiment and corroborating clinical data prompted us to further investigate the effects of these mutations and how differences between them could be used to better understand their role on drug

resistance.

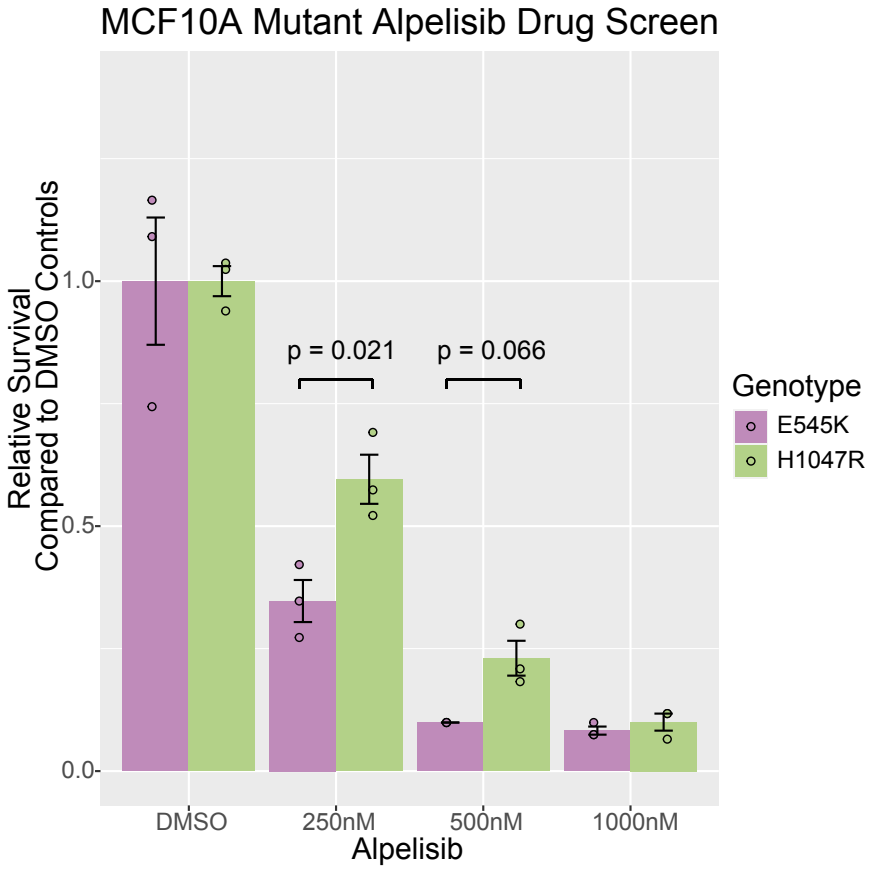

Figure 2.1: **Alpelisib drug screen in MCF-10A isogenic *PIK3CA* mutant cells.** This bar plot shows relative survival of the mutant cell lines in response to differing concentrations of alpelisib.Significance calculated using an ANOVA with a post-hoc Sidak's test

2.2 *PIK3CA* Mutation-induced Changes to Gene Expression

To expand upon the differential findings of drug resistance, we performed RNA-seq across isogenic cell line models to identify differences in gene expression induced by each of these mutations. Gene expression analysis via RNA-seq, can be used to uncover nuanced differences between cell types within a tissue as the expressed genes within a cell are critical to how a cell functions and the role it plays (Cheng et al., 2023; Momeni et al., 2023). In using these tools, we seek to understand how the independent hotspot *PIK3CA* mutations are influencing the function and role of breast cells.

2.2.1 RNA-sequencing of MCF-10A Isogenic Model

For these analyses, we once again made use of the MCF-10A isogenic model described before (Gustin et al., 2009). The primary advantage of this model for this purpose is the dearth of other mutations that could confound our observed effect of the *PIK3CA* mutations. RNA-sequencing was performed across each of the three cell lines in the model representing each of the genotypes of PIK3CA (WT, E545K, and H1047R). To demonstrate the differences in gene expression induced by the mutants, we focused our primary analysis on the direct comparison of the E545K and H1047R cell lines (Figure 2.2). This work can be found in the preprint Miranda et al.(Miranda et al., 2024).

From our analysis we identified 1271 genes exhibiting differential expression between the mutant cell lines. These differentially expressed genes (DEGs) highlight distinct patterns of expression between these mutant cell lines and highlight changes in key oncogenic pathways, which reveal differing modes of carcinogenesis. Differential expression was determined using DESeq2(Love et al., 2014). Firstly in Figure 2.2A, we highlight the expression patterns of the DEGs across the three cell lines in the model. A graphical representation of the DEGs and their relative expression in the mutant cell lines is represented in Figure 2.2B. What becomes apparent from this analysis, is how many of the genes with increased expression in the E545K cell line share increased expression with the WT cell line, while genes with increased expression in the H1047R mutant cells do not. This shows a clear deviation in the biochemical signaling of cells with the H1047R genotype.

To better ascertain the function and role the DEGs may be playing in terms of their impact on cell function, we performed Gene Set Enrichment Analysis (GSEA) of the DEGs using the MSigDB Hallmark pathway collection (Figure 2.2C). The results of this analysis show enrichment of genes within different pathways that can promote carcinogenesis. Within the E545K mutant cells we see enrichment of genes within the apoptotic pathway and the Epithelial-mesenchymal transition pathway. Both of these pathways promote survival and adaptation of cells as they progress toward aggressive and malignant expansion.

In the H1047R mutant cells, enrichment of genes associated with estrogen signaling is shown. Disruption of estrogen signaling is common in many breast cancers and is a key factor in treatment response and resistance (Clusan et al., 2023). These findings support the idea of distinct modes of tumorigenesis defined by these mutations and provide further evidence for the clinical need to evaluate these mutations independently.

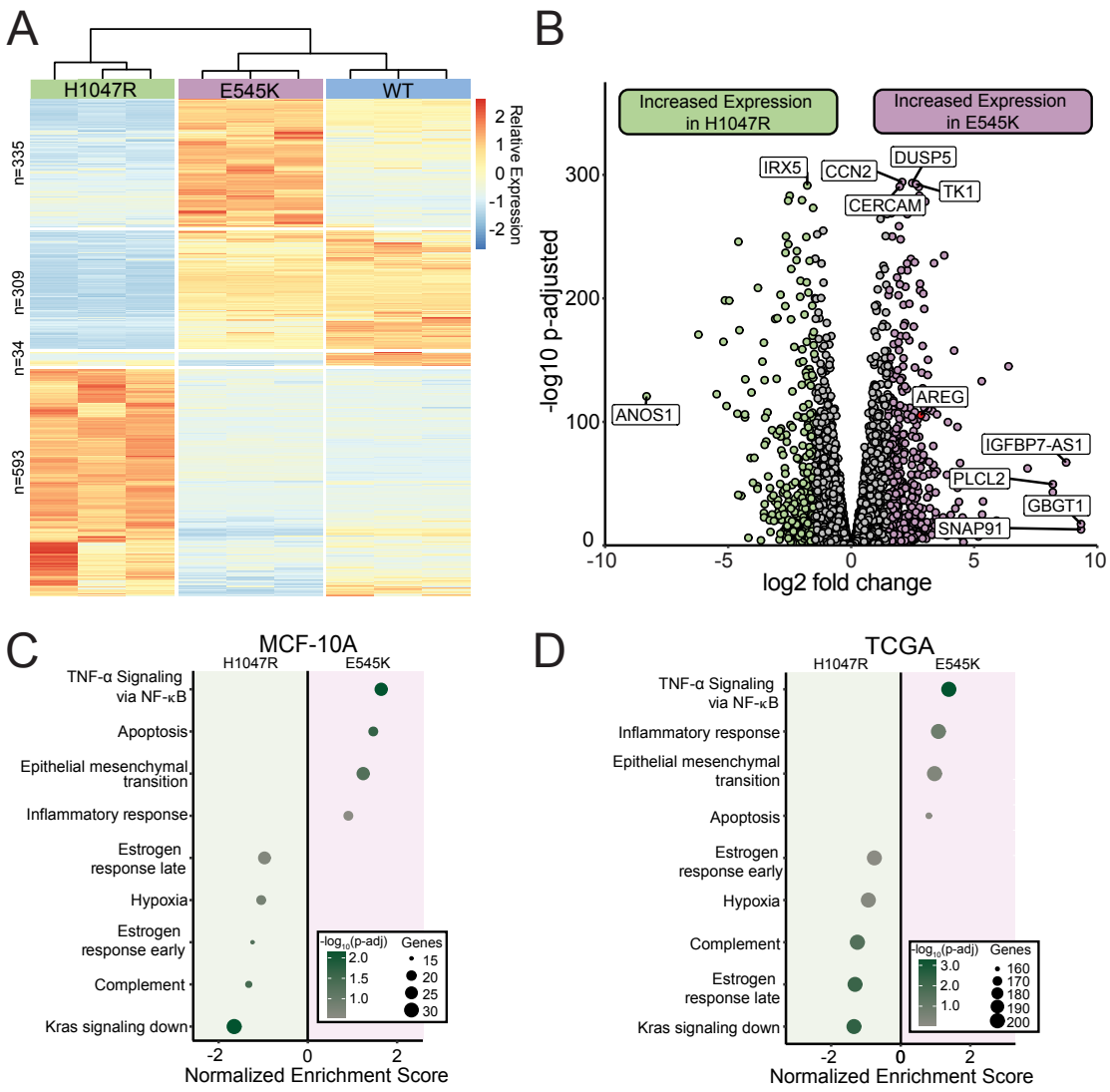

Figure 2.2: **Comparison of RNA sequencing in *PIK3CA* mutant cell lines in the MCF-10A model. RNA-seq captures distinct gene expression differences induced by *PIK3CA* hotspot mutations in isogenic cell line models which are reflected in TCGA patient samples.** (A) Heatmap of normalized counts for 1271 differentially expressed genes. Hierarchical clustering of these genes reveals that E545K cells bear more similarity to WT than H1047R. (B) Volcano plot of differential expression between mutant isogenic cell lines. Differentially expressed genes (DEGs) were defined by the criteria: fold change $> |1.5|$, Padj < 0.05. (C, D) Dot plot showing results from GSEA pathway enrichment analyses using the hallmark gene sets for (C) MCF-10A DEGs and (D) expression data from TCGA-BRCA samples. Pathways shown in panel D are those that are significantly enriched, shared and concordant with those identified as significant in the MCF-10A cell line.

2.2.2 Independent Validation in TCGA Patient Samples

To validate the gene expression profiles from our cell line model and to support the clinical relevance of these findings, we performed an independent validation of our GSEA findings using the RNA-seq samples of actual breast cancer patients with the E545K and H1047R mutations of *PIK3CA*. The TCGA-BRCA database has 72 RNA-seq patient samples with the E545K genotype and 135 samples with the H1047R genotype. In performing GSEA on these samples, we find that all 14 of the pathways found to be significantly enriched in our cell line model are enriched within comparison of the TCGA samples. Shown in Figure 2.2D it can be observed that 11 of those 14 pathways show concordant behavior with the cell line model. These results bolster our findings and demonstrate the parallels in gene expression between the isogenic cell line model and actual patient samples in the TCGA database.

2.3 Mutation-induced Changes to Chromatin Accessibility (ChrAcc)

2.3.1 Assessing ChrAcc in the MCF-10A Isogenic cell line model

To improve our understanding of the underlying mechanisms driving the differences in gene expression observed between the *PIK3CA*-mutant cell lines in our model, we also performed ATAC-seq across all of the cell lines to assess ChrAcc differences induced by *PIK3CA* genotype. From our analysis, we made several key findings. Firstly, we identified 8672 differentially accessible regions between the E545K mutant cells and the H1047R mutant cells (Figure 2.3A). From these regions, we identified two clusters that most closely associate with each *PIK3CA* genotype. The first cluster highlighted in purple in Figure 2.3A is designated as the E545K-preferred cluster and contains 1125 regions. The second cluster highlighted in green is designated as the H1047R-preferred cluster and contains 2668 regions. It is important to note that both of these clusters show high enrichment of non-gene-coding and non-promoter annotations (over 70% for both clusters). This suggests that these regions of altered accessibility are likely enhancers and that these sites will likely

have influence on the expression of multiple genes beyond whichever happens to in the closest proximity.

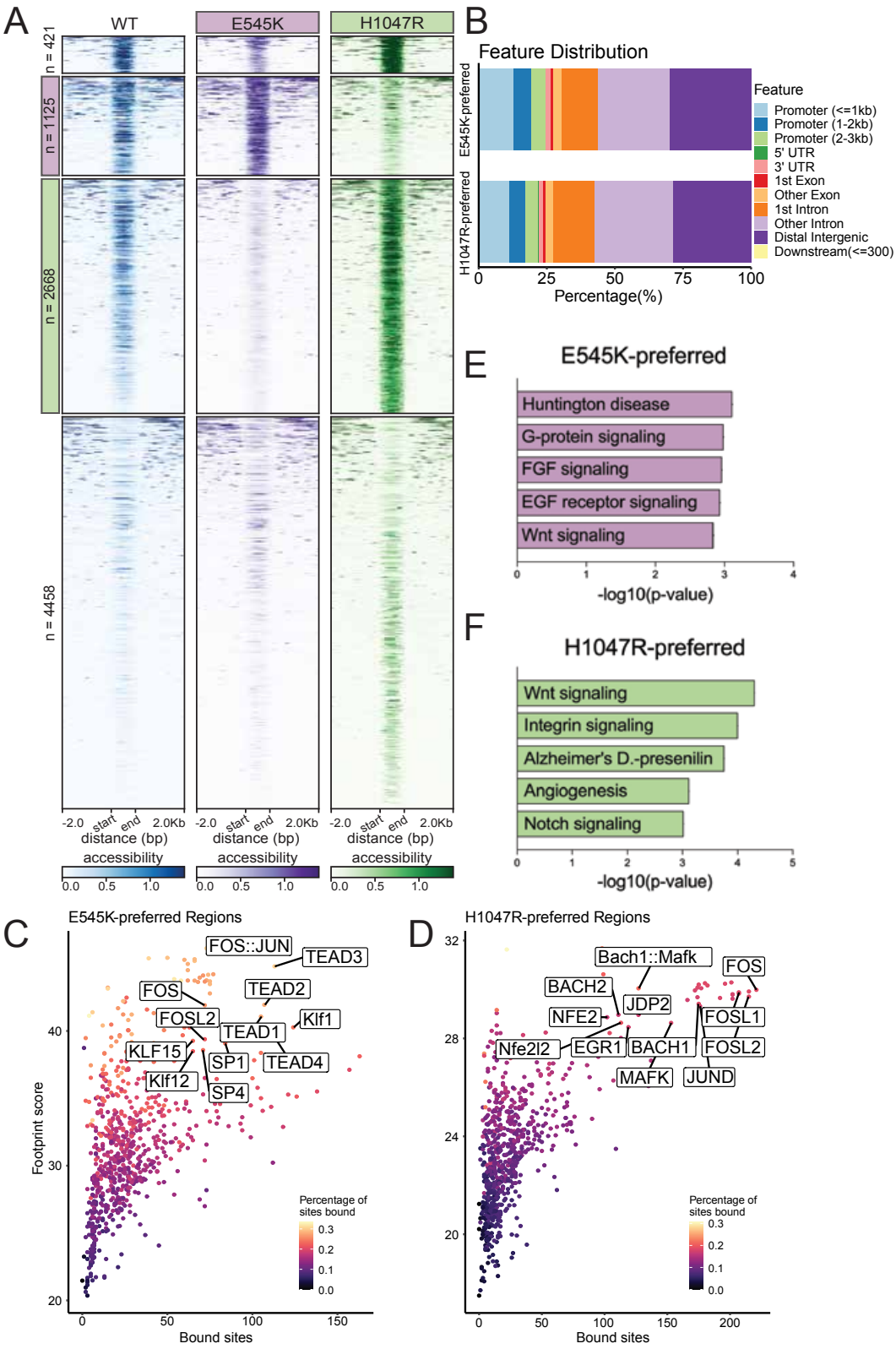

Figure 2.3: **Comparison of ChrAcc measured by ATAC sequencing in *PIK3CA* mutant cell lines in the MCF-10A model.** (A) Heatmap of accessibility at 16660 peaks exhibiting differential accessibility between the mutant isogenic cell lines.

Figure 2.3: **Continued** Peaks are divided based on k-means clustering. The second cluster highlighted in pink has been designated as E545K-preferred. The third cluster highlighted in green has been designated as H1047R-preferred. (B) Distribution of genomic feature annotations of regions within the E545K-preferred and H1047R-preferred clusters. (C) Scatter plots of TOBIAS transcription factor footprinting of accessibility in the E545K-preferred regions. (D) Scatter plots of TOBIAS transcription factor footprinting of accessibility in the H1047R-preferred regions. (E) Bar plot of pathway enrichment from the PANTHER pathway database analysis performed on genes uniquely annotated to the E545K-preferred cluster. (F) Bar plot of pathway enrichment from the PANTHER pathway database analysis performed on genes uniquely annotated to the H1047R-preferred cluster regions.

Further analysis of these regions reveals a strong change in the transcription factor (TF) footprinting in these regions. TF footprinting is a unique analysis to the assessment of ChrAcc as it allows for the measurement of TF binding activity through changes in accessibility specifically at the TF binding motif sequences. What is observed within our comparison of footprinting in the mutant-preferred regions, is that there is unique enrichment for binding of members of the TEAD TF family in the E545K-preferred regions (Figure 2.3C). Activity of the TEAD TF family has a strong association with canonical PI3K/AKT signaling and has also been shown to be a promoter of the epithelial to mesenchymal transition (EMT) (Borreguero-Muñoz et al., 2019; Chi et al., 2022). While this unique enrichment of the TEAD family is observed in the E545K-preferred regions, the H1047R regions have a high association with multiple AP-1 TF family member proteins (Figure 2.3D). The AP-1 family TFs are considered pioneer TFs as they are able to interact with chromatin remodeling proteins to promote a proliferative gene expression program (Vierbuchen et al., 2017; Shaulian and Karin, 2001). Increased activity of the AP-1 family members is also typically associated with increased activation of the MAPK cascade and not PI3K/AKT signaling, which makes its high enrichment within regions associated with this specific *PIK3CA* genotype very curious. The results of our TF footprinting are also consistent with the enrichment of total TF motifs found within these regions as well (Figure 2.4). This suggests that we are not just identifying changes in the accessibility at these

binding sites, but that the regions we are defining are quite distinct in sequence and overall potential for gene regulatory activity.

To further characterize the regions within the mutation-preferred clusters, gene annotation was performed so that gene pathway enrichment analyses could be performed. Genes were annotated using a nearest neighbor approach. While this approach is a bit limited in terms of the number of genes included in the analysis, there is still strong correlation between the proximity of an enhancer element and the TSS of a gene as far as regulatory activity. The results of this analysis once again revealed a strong distinction in the gene regulatory environment of breast cells by *PIK3CA* genotype. For regions in the E545K-preferred cluster there is high enrichment of annotated genes with functions related to FGF signaling and EGF receptor signaling (Figure 2.3E). Both of these pathways are associated with the activity of growth factor receptor tyrosine kinases that are directly associated with the activity of the PI3K complex (Nussinov et al., 2021; Yang et al., 2019). Interestingly, the annotated genes in the H1047R-preferred cluster do not show the same enrichment for these pathways. The annotated genes associated with this cluster instead show enrichment for both the Notch and Wnt signaling pathways (Figure 2.3F). While both of these are associated with tumor growth in a variety of cancer types, neither pathway is associated with canonical PI3K activity (Zhan et al., 2016; Zhou et al., 2022).

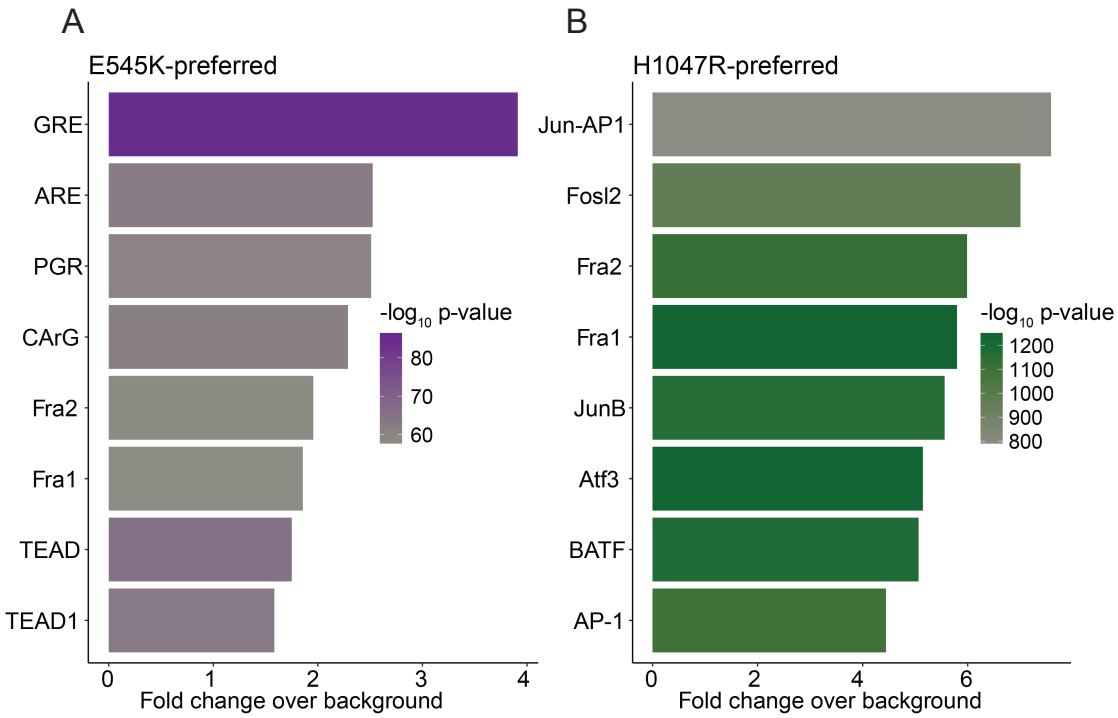

Figure 2.4: **Transcription factor motif enrichment in mutation-preferred clusters.** Bar plots showing the results of HOMER TF motif enrichment analysis within the (A) E545K-preferred and (B) H1047R-preferred cluster regions.

In conclusion, our analysis of ChrAcc in our isogenic model of *PIK3CA* mutations in breast cancer further characterizes the differences in cell signaling and gene regulation that was first described by the RNA-seq experiments in the previous section. Mutation dependent changes in ChrAcc characterize differential activity of the PI3K complex that drive distinct tumorigenic gene expression profiles that are mediated by different TF families.

2.4 Chapter 2 Materials and Methods

These methods are published in Miranda et al. 2024 on Biorxiv (Miranda et al., 2024). A version of these methods can also be found where the data is stored at the Gene Expression Omnibus (GEO) with the accession number: GSE247822

2.4.1 Cell Culture

MCF-10A parental cell lines were purchased from American Type Culture Collection (ATCC). MCF-10A cell line knock-ins were generated as previously described [14]. All cell lines were grown in 5% CO_2 at 37^{o}C with 1% Penicillin/Streptomycin in respective media conditions. Parental MCF-10A cell lines were cultured in DMEM/F12 (1:1) supplemented with 5% horse serum, 20 ng/ml epidermal growth factor (EGF), 10 µg/ml insulin (Roche), 0.5 µg/mL hydrocortisone (Sigma), and 100 ng/ml cholera toxin (Sigma). Knock-in cell lines were maintained in MCF-10A media in the absence of EGF. For all sequencing assays, cells were transferred to assay media 24 hours prior to sample collection. Assay media contains phenol red-free DMEM/F12 (1:1) supplemented with 1% charcoal-dextran stripped FBS (Fisher), 0.2 ng/ml EGF, 10 µg/ml insulin, 0.5 µg/mL hydrocortisone, and 100 ng/ml cholera toxin.

2.4.2 Alpelisib Drug Screen

Cells were plated in their respective maintenance media conditions at a density of 50,000 cells/well in 12-well plates. 24 hours following seeding, cells were treated with 250nM, 500nM, 1000nM of alpelisib (SelleckChem, S2814), or an equivalent volume of DMSO as a negative control. This experiment was performed with three replicates for each cell line and each treatment condition. Cells were counted and viability was measured 72 hours following treatment using a Vi-CELL BLU cell viability analyzer (Beckman Coulter).

2.4.3 RNA-Seq

RNA was isolated and prepared using the Qiagen RNeasy kit. Libraries were prepared by the Vanderbilt Technologies for Advanced Genomics (VANTAGE) Core using the Illumina Ribo-Zero Plus rRNA Depletion Kit. Each library was sequenced on an Illumina NovaSeq, PE150, at a requested depth of 50 million reads. All code and the specific parameters used in all data analyses can be found at:
(https://github.com/adamxmiranda/PIK3CA). All sequencing library reads were trimmed

of adapters and assessed for quality using the Trim Galore! (version 0.4.0) Wrapper of Cutadapt and FastQC (Andrews, 2010; Martin, 2011). Trimmed reads were mapped to the human genome assembly hg38 using the Spliced Transcripts Alignment to a Reference (STAR) aligner (version 2.5.4b) (Dobin et al., 2013). Mapped reads were sorted and filtered for a mapping quality score over 30 using the SAMtools package (version 1.5) (Danecek et al., 2021). Reads were counted to gene transcripts using featureCounts (version 2.0.0) to version 32 of the GENCODE transcripts (Frankish et al., 2019; Liao et al., 2014). A summary of these processing steps can be found in Table 2.1. Differential gene expression was identified between conditions using the DESeq2 package (Love et al., 2014). Pathway analysis was performed using the fgsea package on the Hallmark gene set from MsigDB (Liberzon et al., 2015; Korotkevich et al., 2021).

RNA-seq			
Sample ID	**Sequenced Reads**	**Reads Counted to Transcripts**	**Percent Counted**
MCF10AWT_rep1	74,202,057	68,066,157	91.7308
MCF10AWT_rep2	52,390,047	48,022,937	91.6642
MCF10AWT_rep3	51,651,104	47,545,947	92.0521
MCF10AE545K_rep1	73,874,007	67,122,851	90.8613
MCF10AE545K_rep2	57,313,955	51,485,649	89.8309
MCF10AE545K_rep3	52,917,763	47,766,995	90.2665
MCF10AH1047R_rep1	206,598,621	188,546,753	91.2623
MCF10AH1047R_rep2	332,803,463	302,986,459	91.0407
MCF10AH1047R_rep3	299,799,005	259,056,300	86.4100

Table 2.1: **RNA-seq Summary Table.** This table summarizes the preprocessing of RNA-seq reads prior to differential expression analyses.

2.4.4 ATAC-Seq

Nuclei were isolated and ATAC-seq libraries were prepared using previously published methods (Buenrostro et al., 2013; Barnett et al., 2020). Libraries were sequenced by the VANTAGE Core on an Illumina NovaSeq PE150, at a requested depth of 50 million reads. Reads from the ATAC-seq libraries were trimmed using the same process described in the RNA-seq section. All code and specific parameters used in all data analyses can be found at:

(https://github.com/adamxmiranda/PIK3CA). Trimmed reads were mapped to the human genome assembly hg38 using the BBTools (version 38.69) package and Burrows-Wheeler Aligner (version 0.7.17) (Bushnell et al., 2017; Doench et al., 2016). Quality filtering was performed on the mapped reads using SAMtools (Danecek et al., 2021). A summary of the sequencing preprocessing can be found in Table 2.2. Peaks of accessibility were called using Genrich (version 0.6.1) and differential accessibility was determined using DESeq2 (version 1.34.0) (Love et al., 2014). Accessible regions were clustered using k-means clustering. Gene annotation and pathway enrichment was performed using GREAT (version 4.0.4) (Tanigawa et al., 2022; McLean et al., 2010). The gene annotation parameters used for GREAT were the default parameters of 5kb upstream of the transcription start site (TSS), 1kb downstream of the TSS, or up to 1000kb in either direction for distal regions. Pathway enrichment was performed on uniquely annotated genes using the Enrichr web browser tool and the PANTHER database (Chen et al., 2013; Kuleshov et al., 2016; Mi et al., 2013). Motif enrichment and transcription factor (TF) footprinting were performed using HOMER (version 4.10) and TOBIAS (version 0.13.3), respectively, to identify TF potentially binding to identified accessible peak clusters (Bentsen et al., 2020; Heinz et al., 2010).

ATAC-seq				
Sample ID	**Sequenced Reads**	**Mapped Reads**	**MapQ Filter**	**Filtered Reads**
MCF10AWT_rep1	183,619,704	180,934,051	157,598,707	111,161,682
MCF10AWT_rep2	150,161,177	141,922,792	123,114,133	94,889,407
MCF10AE545K_rep1	117,963,979	112,973,301	99,514,493	77,325,492
MCF10AE545K_rep2	170,250,130	159,992,507	139,707,287	77,564,129
MCF10AH1047R_rep1	230,336,311	212,474,650	183,344,674	122,127,880
MCF10AH1047R_rep2	127,442,609	120,517,187	101,337,602	54,625,387

Table 2.2: **ATAC-seq Summary Table.** This table summarizes the preprocessing of ATAC-seq reads prior to peak calling and differential accessibility analyses.

CHAPTER 3

Therapeutic Target Identification

3.1 CRISPR KO Screens as a Discovery Tool

Genetic knockout (KO) or loss of function experiments are one of the most fundamental experiments in modern genetics. In disrupting the function of a single gene and observing the effects on phenotype, scientists for generations have been able to functionally characterize the role of individual genes in myriad contexts. These methods are so key to genetic research that two different approaches to performing these experiments have won Nobel prizes in 2007 and 2020 (Valancius and Smithies, 1991; Jinek et al., 2012). The next phase of these studies is to increase the scale at which they can be performed. Classically, a single gene knockout could be performed at a time. This would only allow for the functional evaluation of a single gene or genetic locus. Newer techniques, however, have allowed for entire genomes to be evaluated in a single experiment. These experiments have revolutionized not only the efficiency of gene KO experiments, but have also allowed for more interesting and complex experiments to be developed.

Currently, one of the most powerful tools to perform whole genome KO experiments is the whole genome CRISPR KO screen. These experiments make use of a library of plasmids that each encode CRISPR-Cas9 gene editing proteins as well as a targeted guide RNA that can make specific edits to the genome. This library of plasmids can be transduced into a population of cells using lentiviral vectors at a set concentration such that individual cells receive a single plasmid and have only one gene knockout. This strategy creates a population of cells with individual gene KOs that can be studied in a variety of ways (Shalem et al., 2014).

In cancer research specifically, CRISPR KO screens can be used to uncover mechanisms of drug resistance, targets of therapy, and even mechanisms governing metastasis in

specific cell models. The utility of this type of screening is that it can greatly expand the number of genes tested and identified with relevance to a variety of cancer related mechanisms (He et al., 2021).

3.2 Pruning gene targets for Use in a Select CRISPR KO Screen

A weakness and limitation of whole genome CRISPR KO screens is the amount of resources and cells necessary to screen every gene in the genome. To perform a whole genome screen properly using the Brunello library, which is considered the standard for human genome knockout screens, cells on the scale of hundreds of millions are needed to achieve the proper level of viral transduction (Sanson et al., 2018; Joung et al., 2017; Zhang et al., 2014). Therefore, a great strategy to circumvent this weakness is to limit the number of genes being targeted and thus the number of guides in the library pool being screened. Reducing gene targets not only reduces the number of cells needed for screening, but can also reduce the noise in screening results by focusing the assay only on genes of specific interest to the study at hand.

To prune gene targets for a select CRISPR KO screen in our MCF-10A model of *PIK3CA* mutations, we identified genes based on two different categories. Genes selected for targeting in our KO screen needed to exhibit relevance based on both our RNA-seq assay as well as our ATAC-seq assay. These genes would represent those that not only exhibited differentially expressed genes within our system, but specifically those that were near regions of dynamic regulation. Our final selection criteria were as follows:

1. Genes must exhibit significantly increased RNA expression (log2 fold change greater than 1.5 and p-adjusted value less than 0.05) specifically in one of the *PIK3CA*-mutant cell lines compared to wild type parental line.

2. Selected genes must also be annotated to a region that demonstrated significantly increased accessibility in either mutant cell line compared to the parental cell line.

These selection criteria provided a list of 312 genes, which was further reduced to 280

by availability of targeting guides within the Brunello library (Figure 3.1).

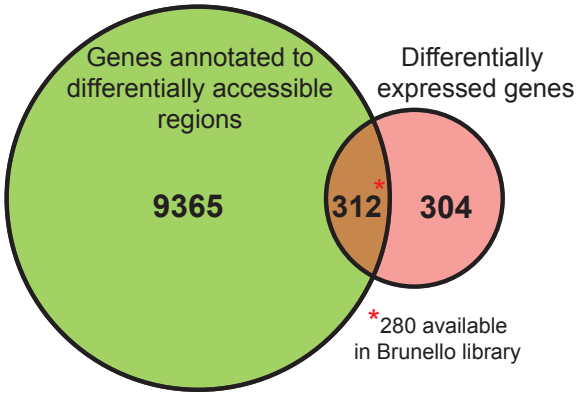

Figure 3.1: **CRISPR KO screen target pruning** Venn Diagram showing the breakdown of genes selected to be included in the select CRISPR KO screen

3.3 Select CRISPR KO Screen in MCF-10A Model

In performing our select CRISPR KO screen, we were able to find many key gene targets with *PIK3CA* mutation-specific effects on the survival/proliferation of cells within our MCF-10A model. For this screen, we focused our initial analysis on the differential drop out of gene targeting guides between each of the *PIK3CA*-mutant cell lines compared to the parental cell line independently. This initial analysis effectively identified genes that when knocked out would induce a specific decrease in the proliferation/survival of *PIK3CA*-mutant cells. These genes present with a negative Z-score per our analysis. However, to identify mutation-specific activity, we compared the Z-scores from the E545K-WT compar-

ison and the H1047R-WT comparison to identify genes that exhibited discordant activity. From this comparison, we identified 36 genes with a z-score difference greater than +/- 1.5. These genes are highlighted in red in Figure 3.2. The five genes with the greatest z-score difference are labeled. These genes represent those with a genotype-specific negative effect on proliferation/survival.

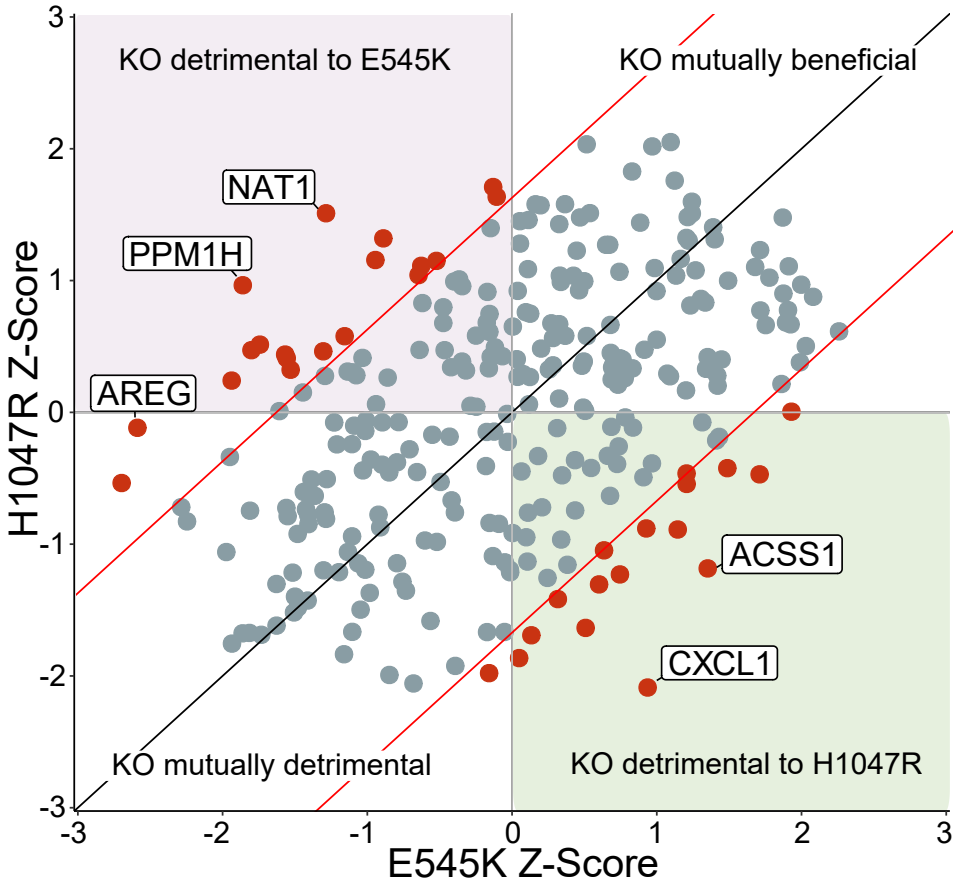

Figure 3.2: **CRISPR KO screen results** Wolverine plot shows the breakdown of gene Z-scores from each mutant cell line relative to the parental cell line. Significant hits, shown in red, were defined by those with an absolute value Z-score difference >1.5.

To further characterize these top five hits for prioritization in further analyses, we analyzed the expression of these genes in both our cell line model as well as the TCGA-BRCA patient samples we analyzed previously. In our cell line model, four of the genes exhibited differential expression between the two mutant cell lines (Figure3.3A). Each of the genes

already demonstrated differential expression in one of the mutants relative to the parental *PIK3CA* WT line. Of the genes that showed differential expression, three of them showed concordance with the genotype of *PIK3CA* for which they exhibited specific selectivity in the CRISPR screen. These genes were AREG, CXCL1, and ACSS1. AREG and CXCL1 both encode protein products that are secreted factors that bind cell membrane bound receptors and can promote signaling through a variety of pathways (Lindzen et al., 2021; Wang et al., 2017). ACSS1 meanwhile encodes a subunit of the acetyl-CoA synthetase complex (Zhang et al., 2016). Each of these genes have been independently linked to tumorigenic mechanisms in multiple cancer types which would make any of them a great target for further study.

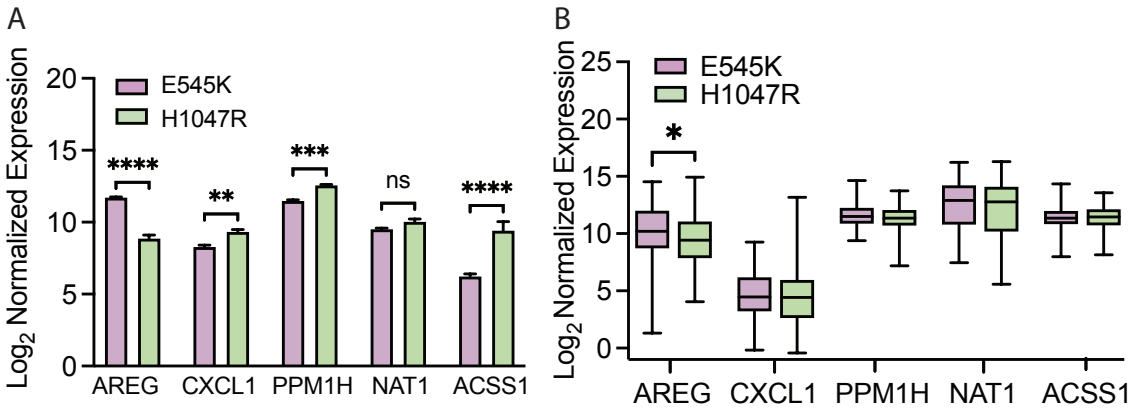

Figure 3.3: **Expression of CRISPR KO Screen Hits** (A) Expression of the top five CRISPR Screen hits in our MCF-10A mutant cell lines. Significance calculated using an ANOVA with a post-hoc Fisher's LSD test. * = p-value 0.05-0.0332, ** = p-value .0332-0.0021, *** = p-value 0.0021-0.0002, ****= p-value $<$0.0002. (B) Box and whisker plot showing the expression of the top five CRISPR screen hit genes in patient samples from TCGA-BRCA. Significance calculated using an ANOVA with a post-hoc Fisher's LSD test. Significance threshold is 0.05.

However, analysis of the expression of these genes in the TCGA-BRCA patient samples yields only one gene with differential expression between samples based on *PIK3CA* genotype (Figure 3.3B). AREG exhibits significantly increased expression in the E545K

mutant samples compared to the H1047R mutant samples. This result is consistent with what is observed in our MCF-10A cell model. Based on these findings we chose to focus our further validation experiments on AREG specifically.

3.4 Chapter 3 Materials and Methods

These methods are published in Miranda et al. 2024 on Biorxiv (Miranda et al., 2024). A version of these methods can also be found where the data is stored at the Gene Expression Omnibus (GEO) with the accession number: GSE247822.

3.4.1 Gene Target Pruning

Differential expression was determined using DESeq2 at a threshold of (log2 fold change greater than 1.5 and p-adjusted value less than 0.05) (Love et al., 2014). Differential accessibility was assessed using DESeq2 and nearest neighbor annotation for differentially accessible regions was performed using ChIPseeker (version 1.30.3) with the default annotation conditions of +/- 3kb from the TSS (Love et al., 2014; Wang et al., 2022). The Brunello whole genome sgRNA library was modified for the selected 280 genes and prepared by the Vanderbilt Functional Genomics core in the lentiCRISPRv2 plasmid background (Doench et al., 2016; Sanjana et al., 2014).

3.4.2 CRISPR KO Screen

MCF-10A cells were cultured in the maintenance media conditions at a density of 500,000 cells/well in a 6-well plate. 24 hours after seeding, cells were infected with viral supernatant in maintenance media containing 5 µg/mL polybrene. 24 hours post-infection cells were placed in selection media containing 1 µg/mL of puromycin and maintained for two weeks. Following two weeks of selection, libraries were prepared and sequenced following the protocol described in (Sanjana et al., 2014). A graphical representation of this strategy can be found in Figure 3.4. Libraries were sequenced by the VANTAGE core and analysis was performed using the maximum likelihood estimation (MLE) algorithm within the

MAGeCK software package(version 0.5.9.5) (Li et al., 2014, 2015).

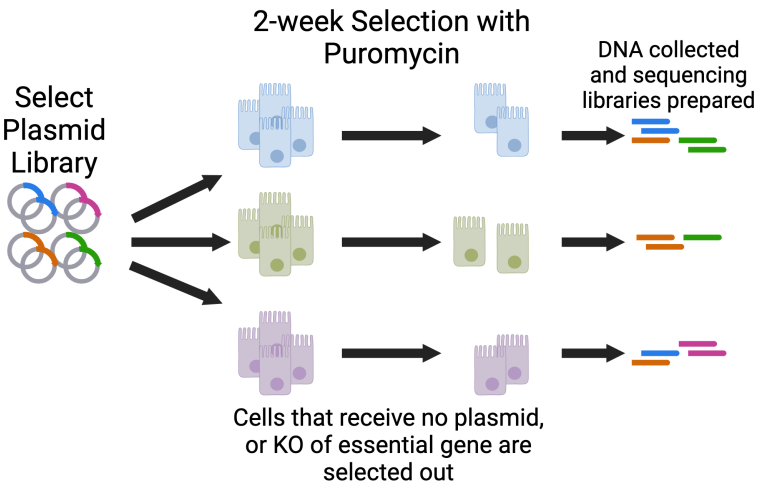

Figure 3.4: **Schematic of CRISPR Screen Strategy** This flowchart shows the experimental set up for the CRISPR KO screen experiment.

CHAPTER 4

Amphiregulin as a mutation-preferential target of therapy in E545K-mutant PIK3CA Breast Cancer

4.1 Introduction to amphiregulin

4.1.1 The function of amphiregulin

AREG, or amphiregulin, is a growth factor ligand with binding affinity for both EGFR and HER2. Signaling through either of these receptors can promote activation of the PI3Kα complex (Lindzen et al., 2021; Macdonald-Obermann and Pike, 2014). The near direct interaction of amphiregulin with the mutated product of *PIK3CA* made it a particularly interesting target of continued study.

AREG is one of the seven ligand agonists of the EGF receptor (Macdonald-Obermann and Pike, 2014). AREG was first characterized as a growth modulating protein in the MCF-7 breast cancer cell line. The name amphiregulin stems from the characterization that it can promote growth in some cancer cell lines and inhibit growth in others (Shoyab et al., 1988). Expression of AREG is a key to proper development of the mammary duct and the expression of this gene is closely associated with signaling through the estrogen receptor (McBryan et al., 2008). Newly translated AREG exists as an inactive membrane-associated precursor that is cleaved by an ADAM metalloproteinase into the canonical extracellular ligand molecule and an intracellular molecule. The intracellular signaling molecule is much less characrerized but has been correlated with PI3K signaling, MAPK signaling, and chromatin remodeling (McBryan et al., 2008; Seefried et al., 2022).

4.1.2 Amphiregulin in breast cancer

Amphiregulin has long been established as an important growth factor in early mammary duct development (Luetteke et al., 1999; McBryan et al., 2008). However, in mature breast tissue amphiregulin expression decreases and modulation of cell proliferation is largely

regulated by signaling by the EGF ligand (McBryan et al., 2008; Lee et al., 2007). While these ligands bind to the same receptor proteins, the downstream signaling pathways can diverge quite distinctly based on the activating ligand (Wilson et al., 2009). This shift in growth factor signaling marks the shift from developing prepubescent breast tissue to mature breast tissue, which undergoes highly controlled periods of cell proliferation.

The shift to EGF mediated proliferative signaling however, becomes reversed in early neoplasias of breast tissues. A histopathological study of noncancerous breast neoplasias showed that neoplastic epithelial cells show a distinct increase in the gene expression of AREG that is accompanied by a decrease in the expression of EGF (Lee et al., 2007). These data show a reversion of these neoplastic cells to an earlier developmental stage that promotes more sustained cell division. An increase in AREG expression has been shown to be a key biomarker of oncogenesis in multiple epithelial tissue types including not only breast but also lung, colon, liver, ovary and prostate (Busser et al., 2011).

In cancers with demonstrated AREG overexpression, cells proliferate through a self-sufficient autocrine, juxtacrine, and paracrine loop. In this loop, cells sustain proliferative and oncogenic gene expression profiles and phenotypes through increased signaling events involving both the intracellular and extracellular cleaved domains of AREG (Bolitho et al., 2021; Busser et al., 2011). These signaling events not only promote the growth of cells with increased AREG expresssion, but also create a proliferative microenvironment where other nearby cells can also receive proliferative signals from extracellular AREG. This growth factor promoted loop differs from the growth factor independence that many cancers use as a proliferative signaling mechanism as disruption of AREG can reduce the proliferation of all cells (Bolitho et al., 2021).

The oncogenic signals in AREG overexpressing cells are characterized by three primary elements: increased activation of PI3K pathway member proteins, increased telomerase activity, and epithelial-mesenchymal transition phenotypes (Matsui et al., 2000; McBryan et al., 2008; Bolitho et al., 2021). Activation of PI3K signaling by AREG particularly

enhances the phosphorylation of AKT which leads to increased phosphorylation of the apoptotic proteins Bad and caspase-9. These changes in cell signaling not only promote proliferation of the cell, but make the cell resistant to apoptosis as well (Holbro et al., 2003).

AREG overexpression is also correlated to increased telomerase activity (Matsui et al., 2000; McBryan et al., 2008). Telomerase is an enzyme that extends the telomere ends of chromosomes. As cells divide, these telomere ends can degrade. Degradation of these telomere ends limits the capacity of a cell to divide as these structures are necessary for the proper segregation of chromosomes during cell division. Increased activation of telomerase is common in many cancers as it increases the replicative cycles a cell can undertake (Kim et al., 1994). This increase in telomerase activity in AREG overexpressing cells sustains the proliferative microenvironment and allows for increased proliferative potential for all cells in the microenvironment (Busser et al., 2011).

The third oncogenic element of AREG overexpression is the promotion of epithelial-mesenchymal transition phenotypes. Epithelial-mesenchymal transition, or EMT, describes a change in the characteristics of an epithelial cell to acquire traits more like that of a mesenchymal cell. Those traits are primarily a lack of dependence on cell-cell adhesion and motility (Nisticò et al., 2012). The acquisition of these traits is correlated to a cancer cell's ability to metastasize and form new tumors in tissues beyond the tumor origin site. One of the key mechanisms through which AREG promotes these changes in cells is through the redistribution of E-cadherin proteins (Chung et al., 2005; Wilson et al., 2009). E-cadherins are important proteins for the establishment of cell-cell junctions in epithelial tissues. These proteins bind to E-cadherins on the surfaces of other epithelial cells to orient and organize the cells. Treatment of cells with extracellular AREG leads to intracellular localization of E-cadherin proteins and loss of E-cadherin mediated junctions (Chung et al., 2005). The loss of these junctions allows for disorder of the epithelial tissue and a loss of contact-dependent signaling that would slow the proliferation of these AREG overexpressing cells. Treatment of extracellular AREG has also been shown to increase the motility invasiveness

of breast epithelial and breast cancer cells in both 2D and 3D culture conditions (Silvy et al., 2001; Willmarth and Ethier, 2006).

In summary, amphiregulin is a critical growth factor ligand and intracellular signaling molecule with the capacity to promote a proliferative gene expression profile within the AREG activated cell and in other nearby cells in the microenvironment. Amphiregulin is found to be overexpressed in a wide variety of tumor types, primarily epithelial ones, and promotes a number of key oncogenic hallmarks beyond increased proliferation. The hallmarks of cancer associated with amphiregulin overexpression are: sustaining proliferative signaling, enabling replicative immortality, nonmutational epigentic reprogramming, unlocking phenotypic plasticity, and resisting cell death (Hanahan, 2022).

4.1.3 Amphiregulin in our model

In our MCF-10A isogenic cell line model we observe a number of key phenotypes with regards to AREG. Firstly, both PIK3CA mutant cell lines exhibit increased expression of AREG over WT cell lines (Figure 4.1). We also observe a region of differential accessibility annotated to the AREG gene. This region is discussed at length in section 4.3. Lastly, our observations culminate with selective essentiality of AREG in our mutant cell lines that is not present within the WT cells.

Increased amphiregulin expression in our mutant cell line models describes a similar autocrine positive feedback loop described in previous studies (McBryan et al., 2008). Increased PI3K signaling in mutant cells is further amplified by increased production of amphiregulin that also activates PI3K through growth factor receptors. Amphiregulin exhibits increased expression particularly in E545K mutant cells (Figure 2.2B and 4.1) and E545K mutant patient breast cancer samples (Figure 3.3B). This pattern of expression corresponds to a mutation-preferential selective effect observed in our select CRISPR KO screen (Figure 3.2). Genetic knockout of the *AREG* gene in the CRISPR KO screen led to a decrease in the proliferation and survivability of both mutant cell lines, but had a much more selective

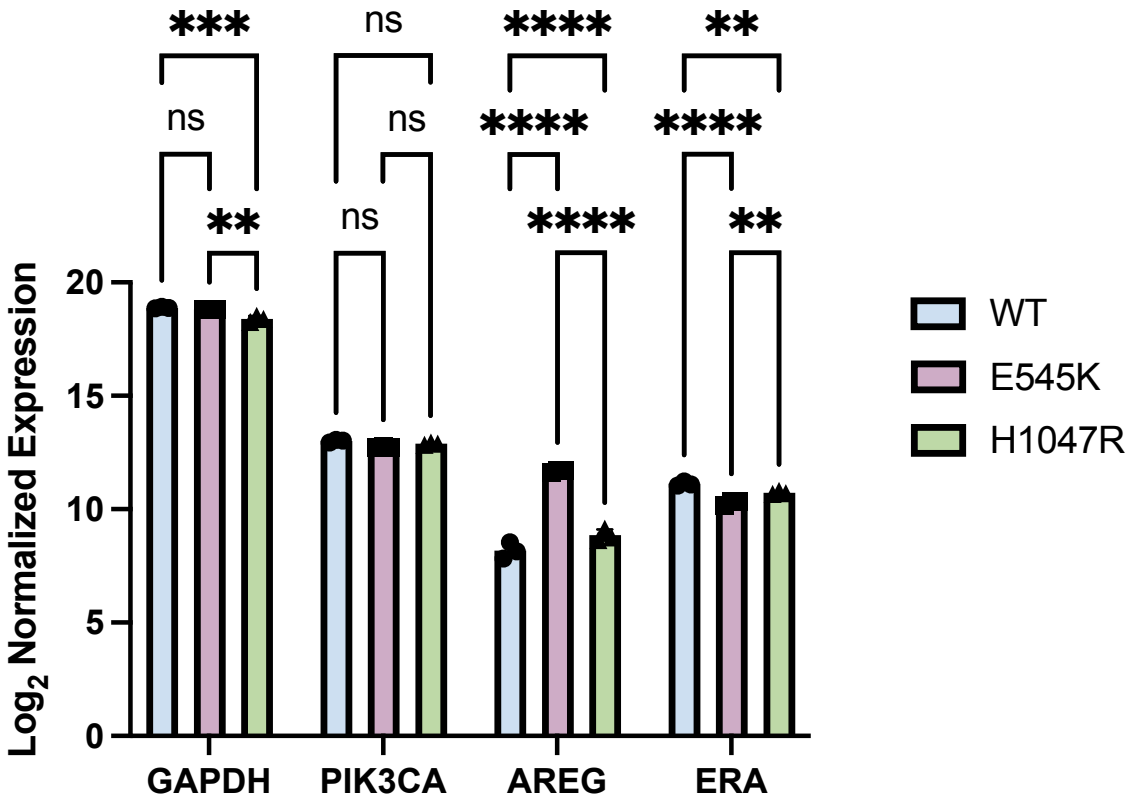

Figure 4.1: **Expression of key genes across the isogenic cell lines.** Bar plot showing the normalized expression of genes across all three cell lines in our isogenic model. The significance tests shown are based on a two-way ANOVA and a post-hoc Tukey test. These are not representative of differential expression analyses shown in other figures. * = p-value 0.05-0.0332, ** = p-value .0332-0.0021, *** = p-value 0.0021-0.0002, ****= p-value $<$0.0002.

effect on the E545K mutant cells. This implies a selective dependence on amphiregulin unique to cells bearing the E545K genotype.

4.2 Inhibition of amphiregulin in MCF-10A Model

To further to interrogate the selective dependence on amphiregulin in our cell model, we performed a knockdown of the gene in our cell line model using three different siRNA molecules. The goal of this experiment was to observe the impact on survival/proliferation of cells in our model with only a reduction of the viable amphiregulin molecules in our system. The results of this assay recapitulates and supports the findings of our CRISPR KO

screen. Inhibition of AREG by siRNA did induce a universal decrease in the expression of AREG, but there was only a significant decline in proliferation/survival of E545K mutant cells (Figure 4.2). Not only did these results support the results of our previous experiment, but it further reinforces the findings because the essentiality phenotype observed in E545K cells is sustained even when the reduction of AREG in the system is incomplete. This highlights the importance AREG to the E545K mutant cells and their promotion of their proliferative signaling.

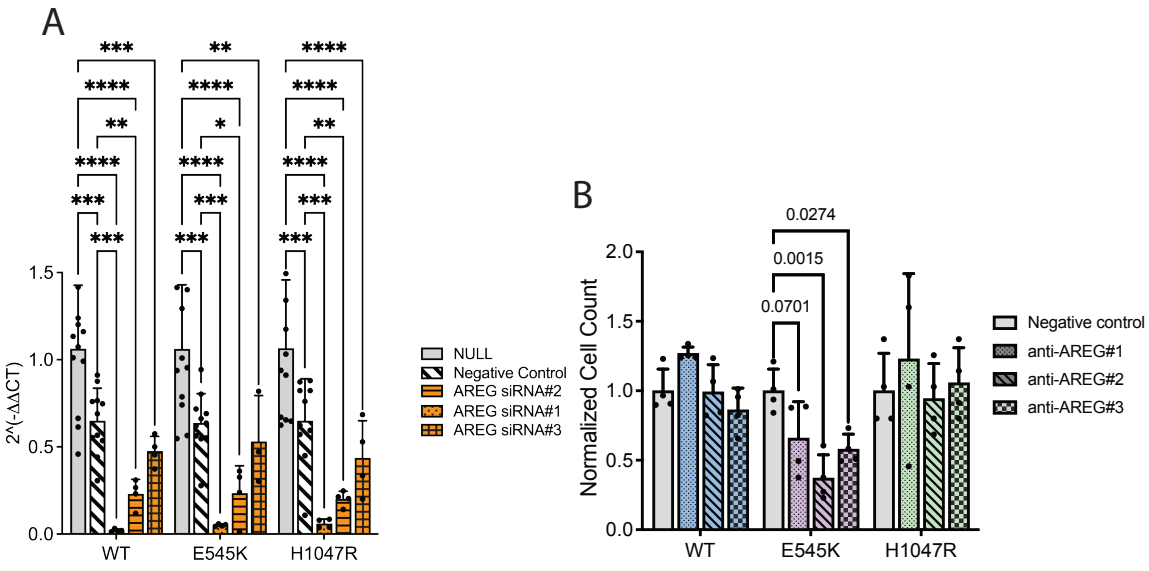

Figure 4.2: **SiRNA knockdown of AREG** (A) Bar plot showing the change in expression of AREG following the introduction of AREG inhibiting siRNA molecules. Outliers were removed using the ROUT method with a Q threshold of 1%. Significance calculated using an ANOVA with a post-hoc Fisher's LSD test. * = p-value 0.05-0.0332, ** = p-value .0332-0.0021, *** = p-value 0.0021-0.0002, ****= p-value <0.0002.(B) Bar plot showing relative cell count as a proxy for proliferation/survivability following the introduction of AREG inhibiting siRNA molecules. Significance calculated using an ANOVA with a post-hoc Tukey's test.

Additional support for the relationship between AREG and E545K mutant breast cells was provided in a subsequent experiment. AREG was also inhibited using a neutralizing antibody introduced to the media. As an extracellular signaling molecule, AREG can be inhibited by the addition of antibody to the media without the need for a delivery mechanism to get the antibody inside the cells. This further improves the therapeutic potential of AREG as a targeted therapy could be much simpler in design for this protein (Alves et al., 2016; Zhou et al., 2022). A lot of drug design can be focused on how to get the treatment molecule into the target cell. AREG being an extracellular protein would really streamline much of that design process.

The results of our assay can be seen in Figure 4.3. Once again we demonstrate a sensitivity to the perturbation of AREG within our E545K mutant cell line. This observation, however, is accompanied by an observed reduction in the proliferation/ survival of H1047R mutant cells. This is to be expected to a degree as the CRISPR KO screen results demonstrated that genetic knockout of AREG has a slight selective effect on H1047R cell lines too. Most importantly however, is that there is no significant reduction in the viability and growth of the WT parental cell line. This reinforces AREG as a potential target with strong selectivity of *PIK3CA* mutant cells with minimal predicted toxicity to healthy cells, which is a major weakness in the current treatment of *PIK3CA* mutant breast cancers.

4.3 Identification of a putative enhancer region of *AREG*

4.3.1 ATAC-seq identifies putative enhancer region upstream of *AREG* transcription start site

ATAC-seq, as described earlier, identifies regions of chromatin accessible to protein binding. These regions of accessible chromatin are typically associated with sites of active transcription and gene start sites, but also associate with TF binding sites in intergenic loci as well. Accessible TF binding sites within intergenic regions of chromatin are considered enhancer regions as they may influence the transcription level of target genes.

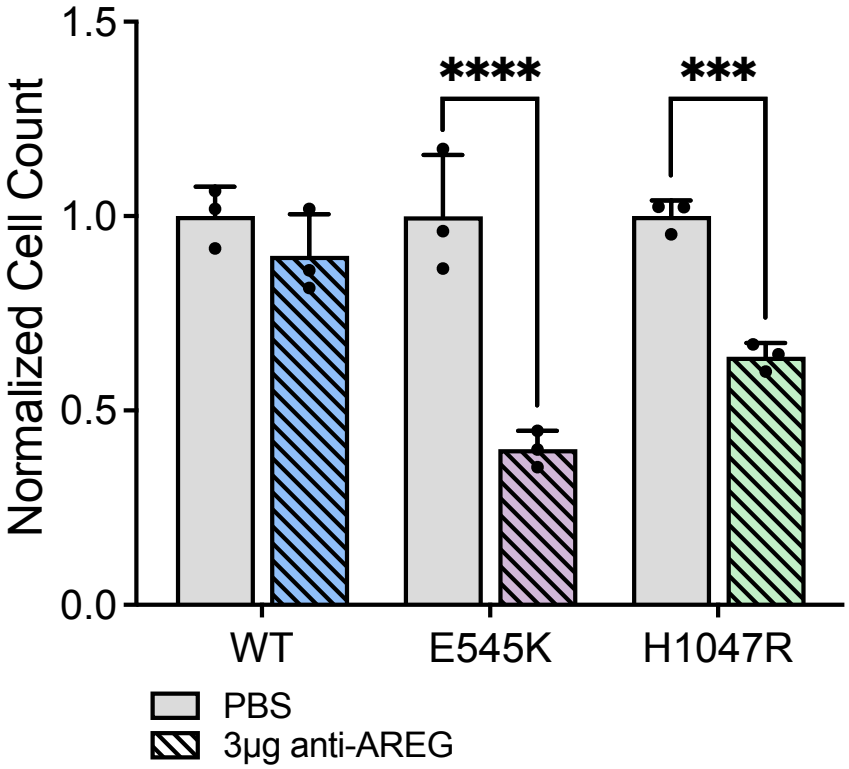

Figure 4.3: **Inhibition of AREG using a neutralizing antibody** Bar plot showing relative cell count as a proxy for proliferation/survivability following the introduction of 3µg of AREG neutralizing antibody; (RandD Systems, MAB262-SP). Significance calculated using an ANOVA with a post-hoc Sidak's test. * = p-value 0.05-0.0332, ** = p-value .0332-0.0021, *** = p-value 0.0021-0.0002, ****= p-value $<$0.0002.

One aspect of the criteria for the selection of gene targets within our CRISPR KO screen was that selected genes must also be annotated to a region that demonstrated significantly increased accessibility in either mutant cell line compared to the parental cell line. This means that there was at least one putative enhancer region annotated to AREG. Our gene of interest AREG was annotated to a single putative enhancer region at the chr4:74435384-74435596 locus. This region is almost 10 kilobases from the transcription start site of the AREG gene and is called as uniquely accessible in genes with the E545K genotype. A graphical representation of this locus can be found in Figure 4.4. This unique accessibility profile corresponds to the unique regulation of AREG and its relationship to breast cells with the E545K genotype. These results suggest that the increased expression of the AREG

gene observed in E545K cells is promoted by TF binding activity at this uniquely accessible peak.

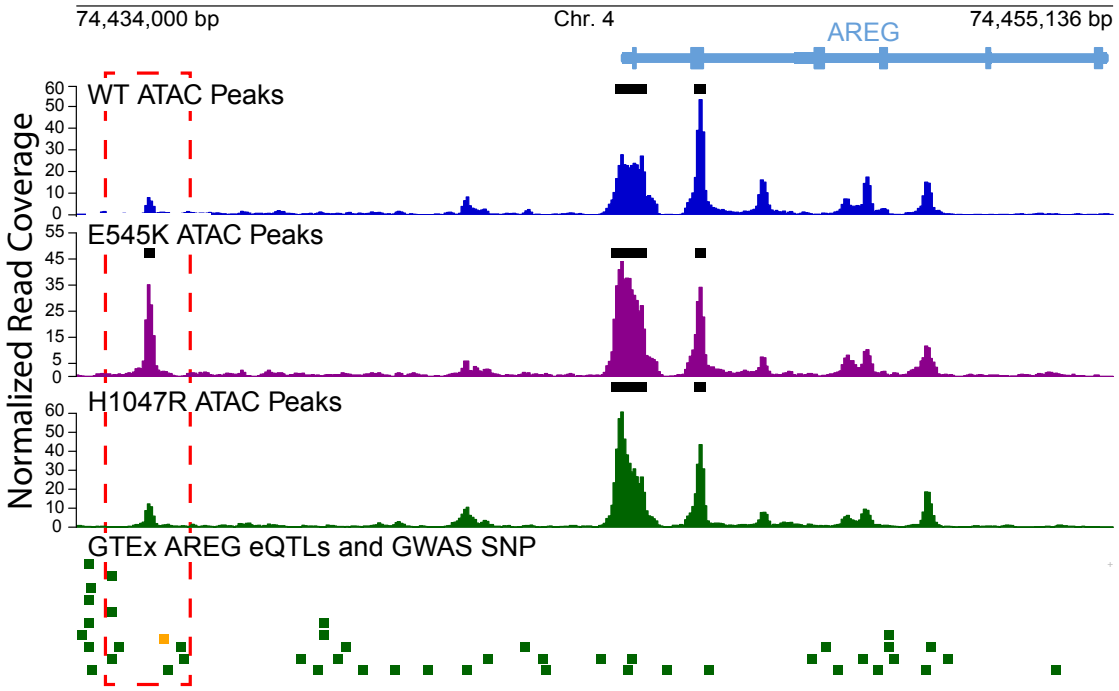

Figure 4.4: **Putative enhancer of AREG locus** Genomic tracks showing the ATAC-seq data across the isogenic cell lines alongside key SNPs at the AREG gene locus. The red box highlights the differentially accessible region annotated to the AREG gene. The green bars designate GTEx AREG eQTLs. The gold bar designates a GWAS SNP associated with breast cancer.

At this locus, we also observe 5 different eQTLs (expression quantitative trait loci) predicted to affect AREG expression within 1 kilobase of the called peak region (Lonsdale et al., 2013). eQTLs are sites of genetic variability that have been shown to impact the expression of genes depending on the genotype at these sites (Carithers et al., 2015; Aguet et al., 2020). A full list of the eQTLs can be found in Table 4.1. The clustering of AREG eQTLs at this locus reinforces our association between activity at this locus and the reg-

ulation of AREG expression specifically. We also identified a SNP (rs28570600) around this locus that has been associated with breast cancer risk by GWAS (Zhang et al., 2020). This is another layer of evidence linking the significance of this region to a potential role in breast cancer.

SNP ID	Coordinates	P-Value	gDNA Beta Score	Tissue Identified In
rs55738416	chr4:74435873	1.20E-08	0.19	Cells-Cultured Fibroblasts
rs1845569	chr4:74434878	4.80E-08	0.26	Testis
rs746861450	chr4:74434727	5.20E-07	0.44	Heart-Left Ventricle
rs746861450	chr4:74434727	6.40E-07	0.48	Nerve-Tibial
rs746861450	chr4:74434727	1.10E-06	0.41	Adipose-Subcutaneous
rs746861450	chr4:74434727	2.80E-06	0.22	Whole Blood
rs972076808	chr4:74436186	3.30E-06	0.45	Cells-Cultured Fibroblasts
rs1005463662	chr4:74436131	2.30E-05	0.15	Cells-Cultured Fibroblasts

Table 4.1: **eQTL Summary Table** This table summarizes eQTLs of AREG near the putative enhancer region.

4.3.2 CRISPR-mediated deletion of putative enhancer of AREG

Much like the concept behind the gene knockout experiments discussed before, we performed a genetic knockout of the putative enhancer region to determine its effect on AREG expression and cell growth and proliferation. To execute this experiment we designed guide RNAs that would target part of the called accessible peak and some of the eQTLs. The goal of this strategy is to excise the region we suspect to have regulatory influence over AREG. A combination of these guides complexed to Cas9 protein was introduced to the cells in

our isogenic model.

Results of this assay are shown in Figure 4.5. From these results, it can be seen that deletion of this enhancer region induces a significant decrease in AREG expression across each of the cell lines in our model (Figure 4.5A). This demonstrates a regulatory relationship between the putative enhancer region and the expression of AREG in breast cells. More interestingly, the deletion of this region had a less universal effect on proliferation/-survival. Deletion of the putative AREG enhancer induced a specific reduction in proliferation and survival unique to the E545K cells in our model (Figure 4.5B). This result was once again consistent with the relationship we have characterized between cells with the E545K genotype and dependence on AREG expression and activity. Whether perturbed at the genetic, RNA, protein, or even regulatory level; perturbation of AREG has a specific effect on PIK3CA mutant cells particularly of the E545K genotype. This mutation-specific effect drives home the conclusion that our analysis of functional genomic data can identify potential therapeutic targets with mutation-specific selection and limited toxicity to non-mutant cells.

4.3.3 Synthetic lethality of AREG loss in E545K

This relationship between the perturbation of AREG (and/or its putative enhancer) and the mutation-specific decline in survival of the E545K mutant cells describes a synthetic lethal relationship in which loss of this genetic region or molecule is incompatible with growth and survival of these specific mutant cells. This synthetical lethal relationship highlights a potential therapeutic vulnerability and there have been a number of recent advances in the research of cancer therapeutics designed to take advantage of these synthetic lethal relationships.

The biggest success story for synthetic lethal cancer therapeutics is the inhibition of PARP in BRCA1/2 mutant breast cancer. BRCA1 and BRCA2 are both commonly mutated genes in breast cancer that encode proteins that play critical roles in DNA damage

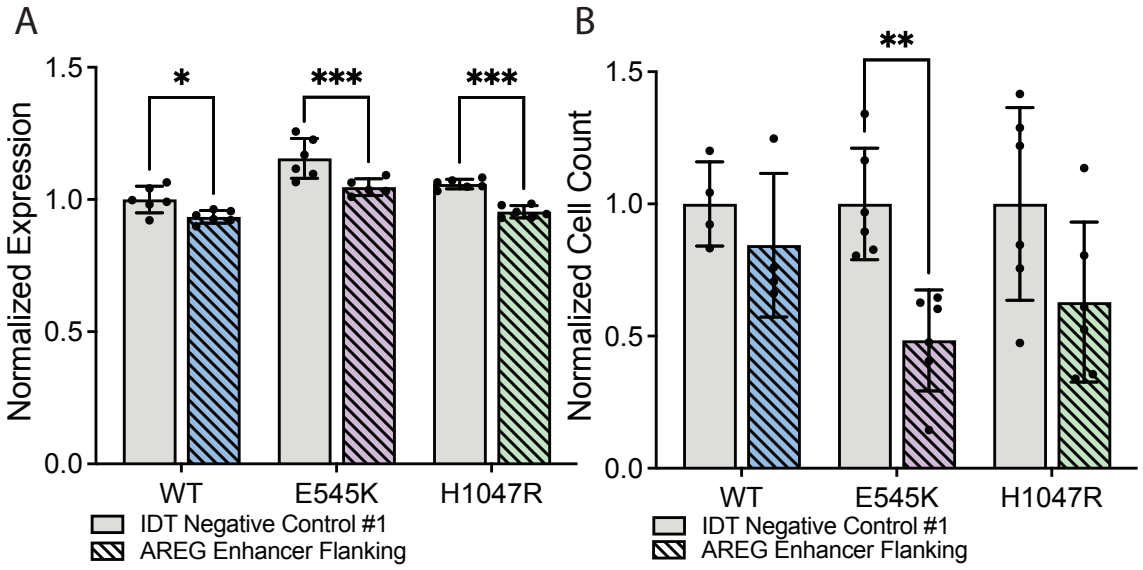

Figure 4.5: **CRISPR-mediated deletion of putative AREG Enhancer** (A) Bar plot showing the change in expression of AREG by qPCR following the introduction of AREG enhancer flanking RNPs. (B) Bar plot showing relative cell count as a proxy for proliferation/-survivability following the introduction of AREG enhancer flanking RNPs. Significance of A and B calculated using an ANOVA with a post-hoc Sidak's test. * = p-value 0.05-0.0332, ** = p-value .0332-0.0021, *** = p-value 0.0021-0.0002, ****= p-value $<$0.0002.

repair (Prakash et al., 2015; Kuchenbaecker et al., 2017). PARP also encodes a DNA damage repair protein, but it is not as frequently mutated as the BRCA proteins. The synthetic lethal relationship of these molecules is such that when PARP is inhibited, in addition to the loss of function in the BRCA genes from mutation, the cell is unable to resolve multiple forms of DNA damage and ultimately succumbs from the accumulation of too many breaks in the DNA. Inhibition of PARP has been highly successful in the clinic with a significant increase in progression-free survival reported in the SOLO2 phase III clinical trial of olaparib (Poveda et al., 2021). Loss of PARP makes BRCA mutant cells incompatible

with survival. This is similar to our demonstrated relationship between AREG and *PIK3CA* mutant cells.

The success of PARP inhibition has been an inspiration for additional research into targeting synthetic lethal relationships in cancer. One of the major initiatives to accomplish this is the Cancer Dependency Map (DepMap). This initiative synthesizes the findings from many CRISPR KO screens across cancer cell lines to identify gene dependencies across a number of cancer types and mutational backgrounds (Setton et al., 2021; Pacini et al., 2021). In fact, when analyzing data from DepMap with relation to PIK3CA mutations and AREG, we see specific essentiality of AREG specifically in breast cancer cell lines with E545K mutant *PIK3CA* (Figure 4.6).

4.4 Chapter 4 Materials and Methods

These methods are published in Miranda et al. 2024 on Biorxiv (Miranda et al., 2024). A version of these methods can also be found where the data is stored at the Gene Expression Omnibus (GEO) with the accession number: GSE247822.

4.4.1 siRNA Inhibition of AREG

Cells were plated in their respective maintenance media at a density of 50,000 cells/well in a 12-well plate. 24 hours following seeding, cells were treated with three different commercially validated AREG targeting siRNA (Ambion), a negative control siRNA (Invitrogen, 4390843), or a null transfection condition using the Lipofectamine RNAiMAX Transfection Reagent (Invitrogen, 13778100) at a concentration of 10 pmol siRNA per well. 24 hours post-transfection, RNA was prepared from cells using the Qiagen RNeasy kit. Four replicates were prepared from each cell line and each treatment condition. RNA was converted to cDNA using the iScript cDNA Synthesis Kit (Bio-Rad, 1708890). qPCR was performed using the AREG and ACTB qPCR primer sets for each sample with the SYBR Green PCR Master Mix (Applied Biosystems, 4309155). Expression of AREG was calculated relative to the expression of housekeeping gene ACTB.

To assess the impact on survival and proliferation, cells were plated in their respective maintenance media at a density of 30,000 cells/well in a 24-well plate. Cells were treated with one of three different commercially validated AREG targeting siRNA (Ambion) or a negative control siRNA (Invitrogen, 4390843) using the Lipofectamine RNAiMAX Transfection Reagent (Invitrogen, 13778100) at a concentration of 5 pmol siRNA per well. Cell counts and viability were measured 24 hours following treatment using a Vi-CELL BLU cell viability analyzer (Beckman Coulter). This experiment was performed with four replicates in each cell line and each treatment condition.

4.4.2 Anti-AREG Antibody Assay

Cells were plated in their respective maintenance media conditions at a density of 50,000 cells/well in 12-well plates. 24 hours following seeding, cells were treated with 1, 3, or 5µg of AREG neutralizing antibody (RandD Systems, MAB262-SP) or an equivalent volume of PBS. This experiment was performed with three replicates for each cell line and each treatment condition. Cells were counted and viability was measured 72 hours following treatment using a Vi-CELL BLU cell viability analyzer (Beckman Coulter).

4.4.3 Deletion of putative AREG Enhancer

Two sgRNAs were designed targeting the locus, chr4:74435384-74435596, which is annotated to the AREG gene as well as nearby AREG eQTLs identified from the GTEx Portal. The guide RNAs were purchased as sgRNAs from IDT with custom targeting sequences. Cells were plated in their respective maintenance media conditions at a density of 50,000 cells/well in a 12-well plate. Transfection mixtures were prepared and added to each of the cell lines according to the guidelines described in the Lipofectamine CRISPRMAX Transfection Reagent Kit (Invitrogen, CMAX00001) using the provided reagents and the Alt-R S.p. HiFi Cas9 Nuclease V3 (IDT, 1081060) in Opti-MEM Reduced Serum Media (ThermoFisher, 31985062). Separate mixtures were prepared containing either a 1:1 mixture of the designed AREG enhancer flanking guides or a control mixture of Alt-R®

CRISPR-Cas9 Negative Control crRNA 1 (IDT, 1072544) duplexed to Alt-R® CRISPR-Cas9 tracrRNA DNA and RNA were collected from separate experimental replicates 24 hours after transfection (at least 4 technical replicates were prepared for each cell line in each treatment condition). Cell counts and viability were measured 72 hours following transfection using a Vi-CELL BLU cell viability analyzer (Beckman Coulter).

DNA was collected using the Wizard SV 96 Genomic DNA Purification System (Promega, A2370). PCR was performed using Phusion High-Fidelity PCR Master Mix (NEB, M0531) and the enhancer deletion validation primer set. PCR mix and thermocycler conditions were set according to the Phusion Master Mix protocol provided by NEB. PCR products were measured and visualized using a D5000 ScreenTape System (Agilent, 5067).

RNA was isolated from cells using the RNeasy Plus Mini Kit (Qiagen, 74134). RNA was converted to cDNA using the iScript cDNA Synthesis Kit (Bio-Rad, 1708890). qPCR was performed using the AREG and ACTB qPCR primer sets for each sample with the SYBR Green PCR Master Mix (Applied Biosystems, 4309155). Expression of AREG was calculated relative to the expression of housekeeping gene ACTB.

A full summary of the oligos used in these assays can be found in Table 4.2.

Oligo Name	Oligo Type	Sequence
AREG Enhancer Upstream	sgRNA	5'mC*mC*mC*rA rA rU rG rA rA rG rA rG rC rA rC rG rG rU rC rA rG rU rU rU rU rA rG rA rG rC rU rA rG rA rA rA rR rA rG rC rA rA rG rU rU rA rA rA rA rA rA rG rG rC rU rA rG rU rC rC rG rU rU rA rU rC rA rA rC rU rU rG rA rA rA rA rA rG rU rG rG rC rA rC rC rG rA rG rU rC rG rG rU rG rC mU*mU*mU*rU-3'
AREG Enhancer Downstream	sgRNA	5'mG*mU*mU*rU rA rC rA rA rC rG rA rC rC rU rG rC rU rC rA rU rG rU rU rU rU rA rG rA rG rC rU rA rG rA rA rA rU rA rG rC rA rA rG rU rU rA rA rA rA rA rG rG rC rU rA rG rU rC rC rG rU rU rA rU rC rA rA rC rU rU rG rA rA rA rA rA rG rU rG rG rC rA rC rC rG rA rG rU rC rG rG rU rGr-CmU*mU*mU*rU-3'

Table 4.2: **Oligo Table** This table contains the sequences of all oligos used in the work described in Chapter 4.

AREG Enhancer Deletion Validation F	PCR Primer	5'-GGAGGCATTTTGTTTCCTGG-3'
AREG Enhancer Deletion Validation R	PCR Primer	5'-GGGGGAAAGGTTAACCAACT-3'
AREG qPCR F	qPCR Primer	5'-TTCCAACACCCGCTCGTTTTGG-3'
AREG qPCR R	qPCR Primer	5'-AGTAGTCATAGTCGGCTCCC-3'
ACTB qPCR F	qPCR Primer	5'-GTTGTCGACGACGAGCG-3'
ACTB qPCR R	qPCR Primer	5'-GCACAGAGCCTCGCCTT-3'
Silencer Select Negative Control #1 siRNA	siRNA	Unlisted (Catalog #: Invitrogen, 493084)
AREG siRNA #1 (Ambion, s1549)	siRNA	GACAAUACGUCAGGAAAUAtt
AREG siRNA #2 (Ambion, 10332)	siRNA	GGGAGGGCAAAAAUGGAAAAtt
AREG siRNA #3 (Ambion, 147186)	siRNA	GCCGACUAUGACUACUCAGtt

Table 4.2: **Continued**

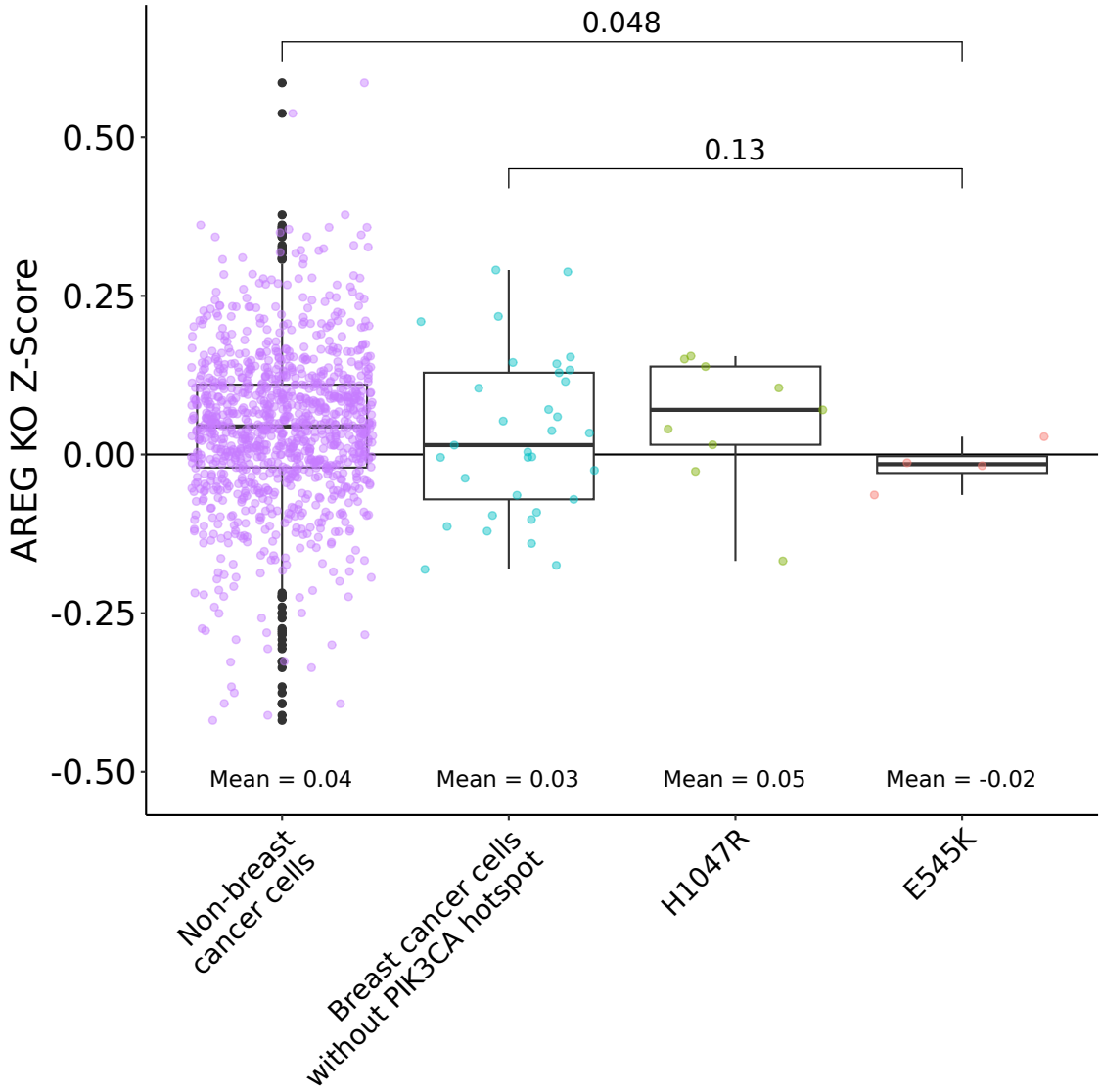

Figure 4.6: **DepMap essentiality data of AREG** Box and whisker plots showing the Z-score of AREG KO in all cell lines from the DepMap study stratified by cancer type and PIK3CA mutation. p-values calculated using nonparametric t-test.

CHAPTER 5

Conclusions for the *PIK3CA* story

5.1 Conclusions on the differential gene regulation of PIK3CA mutant cells

Our assessment of gene expression and gene regulation within our isogenic cell line model of *PIK3CA* hotspot mutations uncovered divergent tumorigenic mechanisms that underscore strong differences in biochemical signaling. From our RNA sequencing we identify a gene expression profile in the E545K cells that appears to align with what we would expect for cells with overly active PI3K signaling. The H1047R cells however show enrichment of genes associated with other tumorigenic pathways, like estrogen signaling and K-Ras, that do not canonically overlap with the activity of PI3K.

These distinct profiles of gene regulation are further supported by our findings in the ATAC-seq data. Our ATAC-seq data shows unique profiles of accessible chromatin dependent on the genotype of *PIK3CA*. These accessible regions are also bound by distinct groups of transcription factors. E545K cells show increased accessibility at and binding of TEAD transcription factors. Members of the TEAD transcription factor family are canonically activated by PI3K signaling. This contrasts with the accessible regions found in H1047R cells. H1047R-preferred regions are primarily associated with AP-1 transcription factors, which are not canonically activated by PI3K signaling.

Together these results demonstrate a model in which the genotype of PIK3CA mutation has a distinct impact on the tumorigenic signaling of breast cells. Despite these mutations occurring within the same gene, the impact on cell signaling is quite different. The E545K mutation promotes what seems to be canonical signaling of PI3K, while the H1047R mutation promotes the sister MAPK pathway of proliferative signaling. These distinct expression profiles are also replicated within the gene expression of actual breast cancer samples. A visual representation of this proposed model can be found in Figure 5.1.

5.2 Conclusions on AREG

In short, our findings on AREG identify it as a gene target of particular interest for breast cancers with E545K *PIK3CA* mutations. AREG not only presents with increased expression unique to E545K mutant cells, but it also presents with a unique putative enhancer that is upstream of the TSS. These observations, with regards to the expression, are also supported by data from actual breast cancer samples from the TCGA-BRCA database, in which samples with the E545K mutation exhibited increased expression of AREG compared to samples with the H1047R mutation. We also demonstrate that that perturbations of AREG at multiple molecular levels have mutation-specific impacts on the proliferation/survival of *PIK3CA* mutant cells. The results of these assays suggest a near mutation-specific oncogene dependence on the presence and function of AREG to E545K mutant cells and breast cancers. Per the biochemical role of AREG as a growth factor receptor, and by consequence PI3K, agonist; a model of a unique positive feedback loop can be proposed. In this model, E545K mutant PI3K complex promotes increased expression of AREG. This increased production of AREG proteins further promotes cell growth signaling through the PI3K pathway as well as even more AREG. Perturbation of AREG however, disrupts this cycle and can slow this proliferative signaling.

All of these factors point to AREG as a strong candidate for future study as a therapeutic target with a synthetic lethal relationship with mutant *PIK3CA*. From our experimental validation, multiple points of vulnerability were identified. Perturbation of AREG at the DNA, RNA, and protein levels all had deleterious effects on the *PIK3CA* mutant cells with minimal effect on the WT parental line. Firstly, this suggests minimal toxicity in targeting AREG, which is the primary weakness of current treatments for PIK3CA mutant cancers. Secondly, this means that a variety of treatment modalities could be explored in the drug development phase. An inhibitor of AREG function directly could be just as viable or worth pursuit as an inhibitor that blocks a certain TF from the AREG enhancer site.

Data presented in Appendix A further supports AREG as a target and also implicates

other nearby growth factor ligand genes as potential agonists of the feedback loop model.

5.3 Impact on the field

The current paradigm in the field of cancer precision medicine is to identify an oncogenic protein and to develop a specific inhibitor that neutralizes the function of that protein. The research presented within this dissertation challenges this paradigm as a best practice for all oncogenic proteins. In the case of the *PIK3CA* gene and the encoded p110α subunit of PI3K in breast cancer, inhibition of this specific protein and complex leads to serious toxicity in patients. Through our functional genomic analyses of isogenic cell lines, we identified alternative targets of treatment with specific effects on mutant cells. Our approach also provides a broadly applicable framework for identifying mutation-specific targets of treatment using other isogenic cell line models as well.

This research also challenges another paradigm that all oncogenic mutations within a gene have the same effect on the phenotype of a cell. In previous studies, the two hotspot mutations of *PIK3CA* were treated as if they would have identical impacts on cell phenotypes. This was perhaps most importantly the case in the SOLAR-1 clinical trial of alpelisib in which the differential response of patients with either hotspot mutation was not even considered as part of the study (André et al., 2021). Not only does our research demonstrate that the different mutations in *PIK3CA* influence response to alpelisib, but that these mutations drive oncogenic phenotypes through distinct pathways and gene regulatory mechanisms (Figure 5.1). Not only do these findings demand further evaluation of clinical response to PI3K inhibitors based on *PIK3CA* genotype, but they also suggest that other common oncogenes with multiple hotspot mutations need to be reevaluated in a mutation-specific manner.

5.4 Limitations of this study

While this study does make a number of impactful observations, there are some considerations that must be taken into account when evaluating the discussed results.

The first of these limitations is that each of the described assays is performed in a single isogenic cell line model. When using a single model, it can be difficult to make a distinction in experimental results being due to the intended biological variable or an idiosyncracy of the particular model. Beyond this being a single cell lineage model, the mutant cell lines are both single clones which may be prone to peculiarities after being bottlenecked in a single cell dilution during development. To offset this limitation of the study, we supported our findings with patient data when possible. This support using patient data is largely used with regards to the gene expression and RNA-seq assays. The cancer genome atlas (TCGA, specifically the TCGA-BRCA) contains data from RNA-seq performed on actual patient breast cancer samples. Using this data, we were able to show that expression patterns from our cell line model are consistent with what is seen in patient samples (Figure 2.2C, 2.2D, 3.3B). The ultimate goal of our study is to find vulnerabilities that will be effective in patients, so being able to verify our results with actual patient data is encouraging for future applications of the presented findings.

A limitation in our approach to identifying candidate genes for our CRISPR KO screen was that we only relied upon gene expression data from RNA-seq and ChrAcc data from ATAC-seq. While identifying genes from these assays yielded a fine number of candidates to evaluate with CRISPR KO screen, there were a number of other assays or epigenetic marks that could have been considered in this selection process. Among these potential assays was ChIP-seq for histone marks. The various posttranslational modifications to histone proteins have strong correlations to the mechanisms of gene expression (Dong and Weng, 2013). ChIP-seq, and the more modern CUT&RUN assay, are assays that can be used to assess the presence of specific histone modifications throughout the genome (O'Geen et al., 2011; Meers et al., 2019). Using the presence of specific histone markers could have been another layer of data that could have further refined our gene selection and helped to better characterize the molecular mechanisms defining the distinctions between our mutant cell lines. Our choice to focus on ChrAcc however, stems from the nature of the peaks this data

provides. Our ultimate goal from these assays was to identify a select list of genes to use for our CRISPR KO library and ATAC-seq peaks are concise and punctuated peaks of less than 1000 base pairs. The concise nature of these peaks makes it easy to connect these peaks with specific genes they may have association with. Histone modifications however, produce peaks that occupy larger swaths of the genome (˜100kb) (Ma and Zhang, 2020). These large peaks can make it hard to make connections between these putative regulatory regions and genes of influence. These larger regions may also contain many transcription factor binding sites, which would also cloud which transcription factors are playing a role in regulating the expression of the cells. ATAC-seq peaks, by nature of their relative size contain fewer TF binding sites which makes it far easier to infer which TFs are actively playing a role in gene expression. While ATAC-seq has certain advantages over ChIP-seq, there is still a degree of assumption that is made when connecting enhancer loci to their target gene promoters when using either of these techniques. In our study, we use a ”nearest neighbor” approach to determine enhancer-gene relationships. This approach makes the assumption that the gene promoter closest to the accessible site is the target of that putative enhancer. While there is a strong correlation between the base pair distance of an enhancer and its target gene/genes, this is not always the case (Qin et al., 2022). Only assays of chromatin conformation can truly provide evidence linking the interaction between an enhancer element and a gene promoter (Pope et al., 2014). In the following Appendix A, we show data from our own chromatin conformation assays in a similar isogenic cell line model.

The results from our select CRISPR KO screen also reflects only a single trial for this experiment. Additional trials may yield some variation in the results for genes to meet the defined significance threshold, and additional repetition could yield more true hits of interest that were labeled as false negatives in this single trial. While we would have liked to have done additional trials of this assay, a change in leadership at the Functional Genomics Core at that time led to our select library of 280 genes being lost. Despite the singleton nature of this screen experiment, our accompanying experimental validation of AREG

inhibition as well as our findings from the clinical samples in TCGA and METABRIC databases (Miranda et al., 2024) provides confidence that AREG, our primary hit, indeed has an important biological function in this system.

A shortcoming of our validation experiments for AREG inhibition is that they were all performed in the same noncancer derived model. Our MCF-10A model is almost entirely wild type, which is a great advantage when we want to study the function of the *PIK3CA* mutations in near isolation. However, a key difference between our model and cancer derived models is the presence of other mutations. While we may have identified selective hits from our CRISPR KO screen with specific function with regards to these mutations in isolation, it is unlikely that an actual breast cancer would have only a single mutation (this may be the case more commonly in a few other cancer types, but would be exceedingly rare in breast cancer (Chalmers et al., 2017)). So what we are missing in our model, is the potential for the epistasis of other mutant genes to compensate for a factor like AREG inhibition. In also doing these experiments in a cancer-derived model it may become evident that AREG inhibition alone is insufficient to produce the same effect in the presence of other oncogenic mutations. This type of information would be greatly informative and may preempt us to try AREG inhibition in combination with another therapeutic strategy when working in more complex model systems.

We also never validated the protein levels of AREG in our system. All of our validation experiments were predicated on using the RNA expression level of AREG as our measure for AREG levels in the cell. While RNA expression does have a high correlation to the protein level of a given molecule in the cell, this relationship is not always a consistent ratio. See our discussion of 4EBP1 in section 1.1.1.2 for an example of a protein interaction creating a rate limiting step prior to translation. However, while the protein levels of AREG were never evaluated specifically using a western blot or ELISA assay, our results demonstrate that perturbation of AREG at the RNA level via siRNA and perturbation of a gene regulatory region were sufficient in producing significant results similar to that of the

neutralizing antibody with regards to cell survival. These results provide evidence that the rate limiting step in AREG protein production is likely at the RNA expression level and not at the protein translation step. This case however, should not be taken as the rule and future validation of other screen hits should keep this in mind when evaluating the results of validation experiments.

Another limitation of this study is that all of the described assays are performed in a 2D *in vitro* culture system. Cancer is a complicated disease and factors like the three-dimensional nature of a tumor or the presence of other cell types can greatly influence not only the gene expression profiles of our cell lines, but they can also influence how cells respond to perturbations like AREG inhibition. Future experiments need to address this limitation and study the impact of these observations in more complex models like a 3D culture, or an *in vivo* mouse model. These models can better mimic these other variables present in real disease that could influence the interpretation of our results. In doing these experiments, it may come to light that the presence of other cell types can sustain E545K mutant cells even when AREG is inhibited. This is an observation that would be impossible to make in our current model, but could be highly relevant in the future development of tareted therapies or advancement of AREG inhibition. Were this the case, we may need to consider a combinatorial therapeutic strategy in addition to AREG inhibition to see more promising results in a more nuanced model system.

5.5 Future directions

There are a number of promising future directions to further the findings of this dissertation. The foremost of these directions should be the further validation of AREG as a therapeutic target of treatment in E545K *PIK3CA* mutant breast cancer. For the reasons outlined above, AREG profiles as a particularly interesting therapeutic target that could have real clinical applications. A potential specific inhibitor for E545K mutant *PIK3CA* would be of particular clinical value as the mutation-specific inhibitors that are currently in trial both

target the H1047R variant. As an accompaniment to the *in vitro* testing of AREG inhibition in our cell line models, validation of the impacts of AREG inhibition in pre-clinical *in vivo* models would go a long way to supporting this target as being worthy of investment in drug design and discovery. A specific experiment to address this would be testing the impact of AREG inhibition on the growth of *PIK3CA* mutant patient-derived xenograft tumors in mice. Patient-derived systems would be a good addition to 3D cultures and/or *in vivo* models. Patient-derived models are created from patient tumor tissue collected in the clinic. In the limitations section, we discussed how the presence of other mutations would be a necessary factor to evaluate the efficacy of AREG inhibition. Patient-derived models have an advantage over immortalized cell line models in that the mutations present within the patient-derived tissue are more representative of mutations within actual breast cancers. Many breast cancer immortalized cell line models like MCF-7 are genetically unstable and may have mutations that would not be representative of clinical samples.

The select CRISPR KO screen highlighted in this dissertation uncovered a number of potential gene targets with mutation-specific effects on the survival/proliferation of cells. While *AREG* profiled as the most promising among these genes and was further validated in this study, there are a number of other hits from this screen that could also prove to be promising. The screen hit that I would choose to follow up on next would be *CXCL1*. This gene also encodes an extracellular signaling molecule that can promote similar signaling pathways that AREG does (Wang et al., 2017). Knock out of the *CXCL1* gene showed a specific detrimental effect on H1047R cells in our select CRISPR KO screen which is opposite the effect demonstrated by knock out of the *AREG* gene.

One thing to consider when targeting CXCL1 in more complicated *in vivo* models is that CXCL1 primarily binds the CXCR2 receptor. CXCR2 is highly promiscuous with high affinity for seven different ligands. One of these other ligands is CXCL2 (Korbecki et al., 2022). According to the GTEx portal, the CXCL2 gene is expressed at a comparable level to CXCL1 in breast tissue (Lonsdale et al., 2013). This could create a scenario in which

the presence of CXCL2 in an in vivo model could compensate for CXCL1 inhibition. This is a consideration for the testing of CXCL1 inhibition, but the experiments will need to be performed to see if any of this bears out.

Ultimately, identifying a viable target of treatment for both breast cancer hotspot mutations of *PIK3CA* would really demonstrate the strength of this experimental framework and provide improved treatment to a larger population of patients.

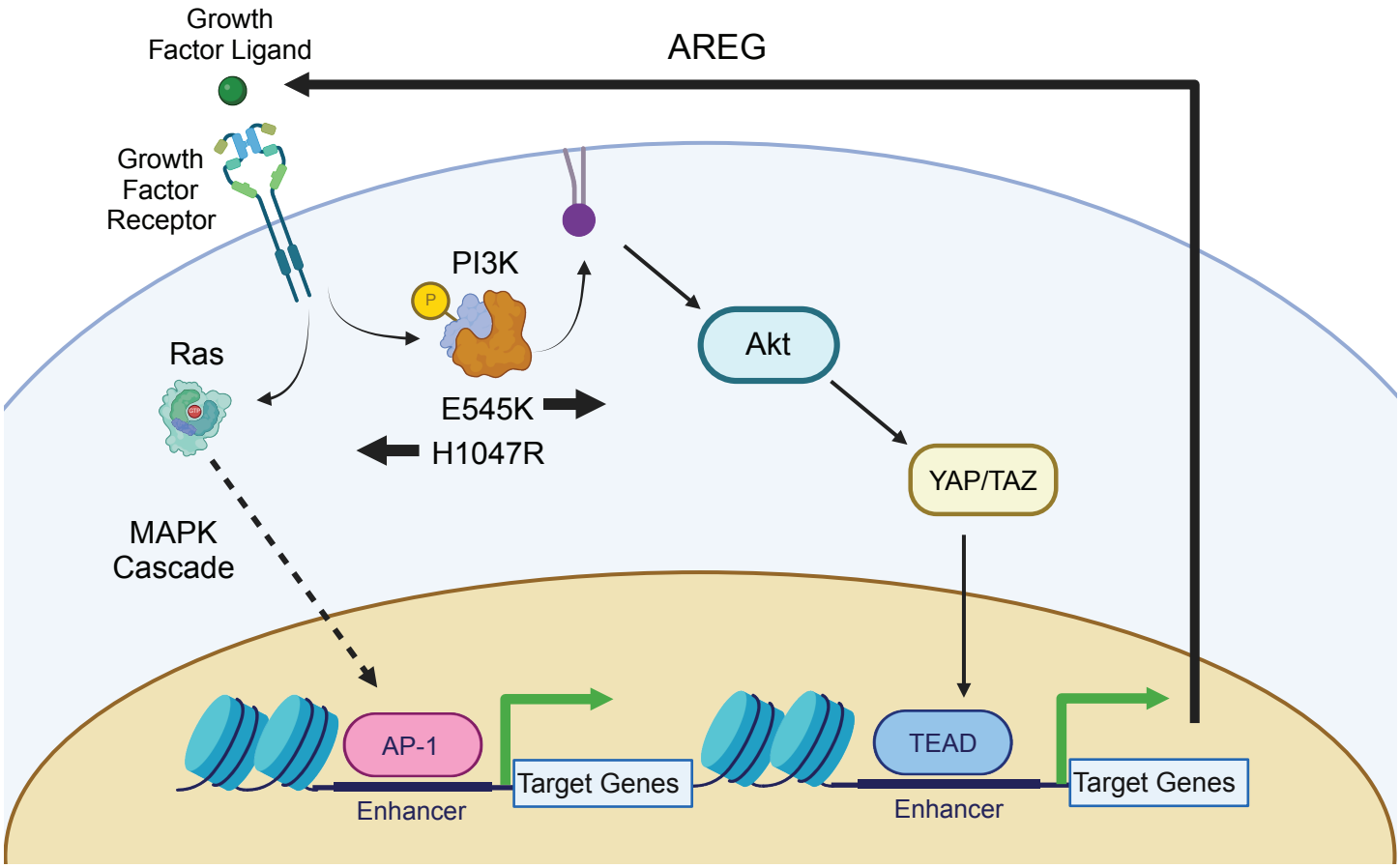

Figure 5.1: **Pathway Diagram of Proposed Feedback Loop** This pathway Diagram illustrates the proposed pathway of action for AREG in our *PIK3CA* mutant cells in relation to other differentially regulated genes.

Appendix A

Chromatin Conformation Analysis of *PIK3CA* Mutant Breast Cells

A.1 The Role of Chromatin Conformation in Gene Expression

The organization of chromatin within the nucleus is an important mechanism for the regulation of gene expression in a given cell (Pope et al., 2014). Many models of gene expression visualize DNA in one or two dimensions in which accessible promoters are bound by TFs and RNA polymerase and gene expression is increased. These models often ignore or oversimplify the fact that the nucleic acids and proteins of chromatin are three-dimensional molecules that need to coordinate and interact within three-dimensional space.

The best known demonstration of chromatin conformation in gene regulation is the enhancer-gene promoter looping with regards to transcription initiation. The canonical model of transcription initiation demonstrates that a TF-bound distal enhancer region must be brought into contact with a TF-bound promoter to either recruit RNA polymerase and/or to initiate the transcription process. This looping event is mediated by a number of specialized transcription factors and shows how chromatin conformation is key to the regulation of gene expression. Enhancer elements need to be able to co-localize and interact with gene targets to promote proper expression (Kagey et al., 2010; Valton and Dekker, 2016).

Outside of enhancer-gene promoter looping there are additional classifications of regulation relevant to the regulation of gene expression. The first of these are ”associated domains”. There are number of different kinds of associated domains, but the two main types are topologically associated domains (TADs) and lamina associated domains (LADs). TADs are a bit of a catch-all term as far as associated domains go and generally refer to regions of chromatin with high levels of interaction as measured by Hi-C. These regions are typically separated, or punctuated by binding sites for the transcription factor CTCF, which serves as an anchor for the looping of larger chromatin domains. TADs typically define

the range at which enhancers can interact with target genes as gene-enhancer interactions are greatly increased between elements within a TAD (Valton and Dekker, 2016). LADs are a slightly more unique associated domain as they refer to regions of the genome that congregate at, and associate with, the nuclear membrane. These regions exist at the periphery of the nucleus and are identified through the association of Lamin A with chromatin. These regions have some unique gene expression interactions due to their association with the proteins of the nuclear membrane (Lochs et al., 2019).

Another major classification of chromatin conformation are A/B compartments. This classification defines different TADs or groups of TADs. A compartments are classified as regions of active transcription. A compartments are correlated with increased expression of included genes and higher levels of intra-TAD interaction. B compartments are just the opposite and are typically considered regions of heterochromatin. These are classified by low gene expression and low intra-TAD interactions. These compartments are representative of a phenomenon known as transcriptional hubs. Transcriptional hubs describe foci within the nucleus with high levels of gene transcription.

A.2 Chromatin Conformation in Cancer

In cancer, many aspects of the regulation of gene expression become disrupted and dysregulated. Conformation of chromatin is no exception to these changes; however, it remains an understudied aspect of gene regulation in cancer. Part of this lack of study can be attributed to how dynamic and difficult to assay these interactions can be. Currently, the best assays to observe and measure the three-dimensional state of chromatin are resource and tissue intensive and direct clinical relevance is hard to parse. There are a few studies that do demonstrate the potential of this research and provide hope for furthering our understanding of how this key element of gene regulation responds to tumorigenic perturbations (Valton and Dekker, 2016).

An example of chromatin conformation playing a role in cancer comes from Flavahan

et al (Flavahan et al., 2016). In this study, they identify a mutation in an insulator region near the *PDGFRA* gene that occurs in glioma. This insulator region contains a binding site for CTCF which separates two different TAD regions. The mutation in this site disrupts the recognition sequence required for TAD binding and ultimately reduces binding of the protein. Disruption of this binding site removes the boundary between the once separated TADs and allows for *PDGFRA* to be upregulated by other more distal enhancers these glioma cells. The critical finding of this study is the presence of a noncoding mutation with influence on the expression of a major oncogene for this tumor type. It is often difficult to identify noncoding mutations with specific relevance to cancer, and not only did the study of chromatin conformation identify this mutation, but it also provided a mechanism for how this mutation influences the cancer cells.

A.3 PIK3CA Mutation Effect on Chromatin Conformation

To study the effects of a single mutation on chromatin conformation in a breast cancer context, we made use of another isogenic cell line model of *PIK3CA* mutations. This model was developed in the MCF7 breast cancer lineage. The MCF7 cell line already carries an E545K mutation in the PIK3CA gene (Horwitz et al., 1975; Lee et al., 2015). This mutation was corrected to the WT genotype using AAV-mediated recombination. Initial characterization of these cell lines demonstrated that the cells with the WT version of the PIK3CA gene grew at a slower pace and had different response to drugs (Beaver et al., 2013). Specifically, the corrected cell lines had a reduced response to the PI3K inhibitor GDC-0941.

Using this isogenic cell line model, we performed bead capture HiC, also known as tethered chromosome capture, in triplicate across the two cell lines. This method allowed for us to assess chromatin interactions anchored by proteins across the genome. Chromatin interactions anchored by proteins are more likely to be stable biologically relevant interactions.

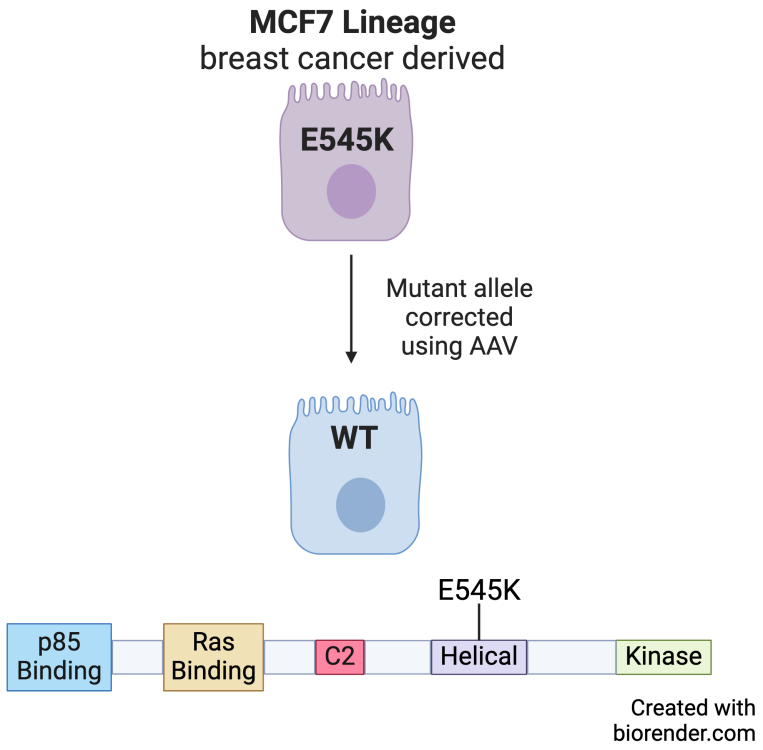

Figure A.1: **Schematic of MCF7 Isogenic Model** This schematic illustrates the isogenic model developed in the MCF7 lineage. This model was used for the chromatin conformation experiments.

From our analysis of this data, we can conclude that there is significant chromatin rearrangement induced by genotype of the PIK3CA gene. In Figure A.2 we show the AREG gene locus, which was shown to be of interest from our previous experiments involving the MCF10A isogenic cell line model. In this figure, it can be appreciated that there are many more distal interactions present within the PIK3CA mutant cell line that are not present within the parental cell line. These interactions also overlap the nearby genes EREG and EPGN. These genes encode epiregulin and epithelial mitogen, respectively. Both of these gene products, like AREG, are Growth factor ligands and serve similar functions (Abud et al., 2021). Increased interactions correlates with compartment A classification of associated loci, which correlates to increased expression of genes within this associating domain (Pope et al., 2014).

These results, in conjunction with our findings from our experiments in the MCF-10A isogenic model, point to this region being particularly sensitive to the downstream biochemical signaling induced by the E545K mutant PI3K complex. These findings also reemphasize the importance of this locus in a cancer background which was a key limitation mentioned in the limitations section 5.4. Additionally, the genes clustered at this locus share common functions, which further supports the positive feedback loop model posited earlier. Altered regulation at this locus could influence much more than the expression of AREG, but any of these other growth factor ligand genes. Any of these genes could promote the feedback loop we observed in the MCF-10A isogenic model.

A.4 Future Directions of Appendix A

The initial goal of the chromatin conformation project was to generate a "four-dimensional" map of chromatin to understand the evolving chromatin dynamics of cancer cells as they acquired mutations. This goal shifted as the main goal of my research limited its scope to focus exclusively on the *PIK3CA* mutations. With this new focus it was my intention to integrate the chromatin conformation data into the main project to support the findings of that study. However, a number of key issues arose which ultimately led to this data being left on the cutting room floor.

Firstly, there is only quality chromatin conformation data from the MCF7 cell line model. Attempts to generate libraries from the MCF-10A lineage cell lines were unsuccessful. Further troubleshooting and optimization of the Bead Capture HiC protocol to successfully generate libraries from the MCF-10A cell lines is needed and should greatly enrich these data.

Secondly, there is not an H1047R mutant analog of the MCF7 cell line lineage. This means that there are no means to look for mutation-specific chromatin interactions from this model. We attempted to generate an isogenic H1047R mutant in the MCF7 cell line, but were ultimately unsuccessful. To generate this cell line, we attempted CRISPR-Cas9

mediated genetic modification on the ”corrected” MCF7 cell line. We were unable to generate any positively modified cells and the cells were not very tolerant of the modification and single cell dilution process. I attribute this to the general frailty of the corrected MCF7 cells which do not grow as well in culture. I think a better strategy may be to try and create the H1047R cell line from the parental MCF7 cells. There are a number of obstacles to consider here as there are two E545K mutant alleles that would need to be corrected in addition to adding the H1047R mutation to a *PIK3CA* allele. This is why we did not initially choose this method, but given the difficulty we experienced with the corrected cells it could be a good alternative. Even doing a multi-step staged strategy could yield better results than working from fully correcting the E545K mutations as a first step.

We also attempted to use the T47D cell line as a proxy for the H1047R comparison as this cell line is derived from the same subtype of breast cancer as the MCF7 cell line and natively includes an H1047R mutation of *PIK3CA*. This however introduces a number of confounding variables as the T47D cell line also has a TP53 loss of function mutation that may also affect chromatin conformation.

Ultimately, there is still a lot of interesting data and findings that can be gleaned from the chromatin conformation data that has already been generated, and diving deeper into this data is the most attainable future direction for this project.

A.5 Appendix A Materials and Methods

A.5.1 Cell Culture

Parental MCF7 cells were purchased from ATCC. The *PIK3CA* corrected cell line was generated as described in Beaver et al. (2013). Both cell lines were grown in 5%CO_2 cultured at 37^{o}C. Culture media for these cells was DMEM supplemented with 10% Fetal bovine serum and 1% Penicillin/Streptomycin.

A.5.2 Bead Capture HiC

A.5.2.1 Chromatin Isolation

Cells were grown to 80% confluence in a T75 flask prior to collection. Cells were trypsinized using TypLE Express Enzyme (ThermoFisher, 12604013). The reaction was stopped using culture media and cells were spun down. Supernatant was removed and cells were resuspended in PBS (25mL) and rocked at 4°C to reduce clumping. formaldehyde was added to the suspended cells to crosslink the chromatin at a concentration of 1% (1.625mL of 16% formaldehyde solution; ThermoFisher, 28906). The crosslinking reaction was carried out with rocking for 10 minutes and stopped with the addition of .23g of glycine (Research Products International, G36050). Samples were spun down and washed with PBS. Samples were spun down and cell membranes were lysed with NP-40 lysis buffer (Invitrogen, FNN0021) and incubated on ice for 15 minutes. Samples were spun again and the cellular debris was removed. Nuclear membranes were then lysed using SDS lysis buffer (Millipore, 20-163).

A.5.2.2 Proximity Ligation

Chromatin was diluted with water to a volume of 400µL and incubated at 65°C for five minutes. 105µL of Iodoacetyl PEG-2 Biotin(ThermoFisher, 21334) was added to the samples and rocked for 1 hour to biotinylate DNA bound proteins. Following this incubation TritonX-100 was added to each sample. samples were then incubated for 10 minutes on ice and then 10 minutes at 37°C. Chromatin was then digested using 400 units of HindIII-HF (NEB, R3104) in CutSmart (NEB, B7204) buffer overnight at 37°C. Digested chromatin was bound to MyOne Streptavidin T1 Dynabeads (Invitrogen, 65601). Beads were washed with PBS-Tween20 prior to addition to the samples. Samples and beads were incubated together for an hour at RT. Excess streptavidin was bound using a 1:1 mixture of the Iodoacetyl biotin and beta-mercapto ethanol. DNA cut site ends were repaired using the NEBNext End Repair module (NEB, E6050). DNA ends were proximity ligated us-

ing Blunt/TA Ligase (NEB, M0367) and incubated overnight. DNA was eluted from the beads through incubations with RNase A and Proteinase K. DNA was purified using the QIAamp DNA Blood Mini Kit (Qiagen, 51104). Libraries were prepared and sequenced by the VANTAGE core.

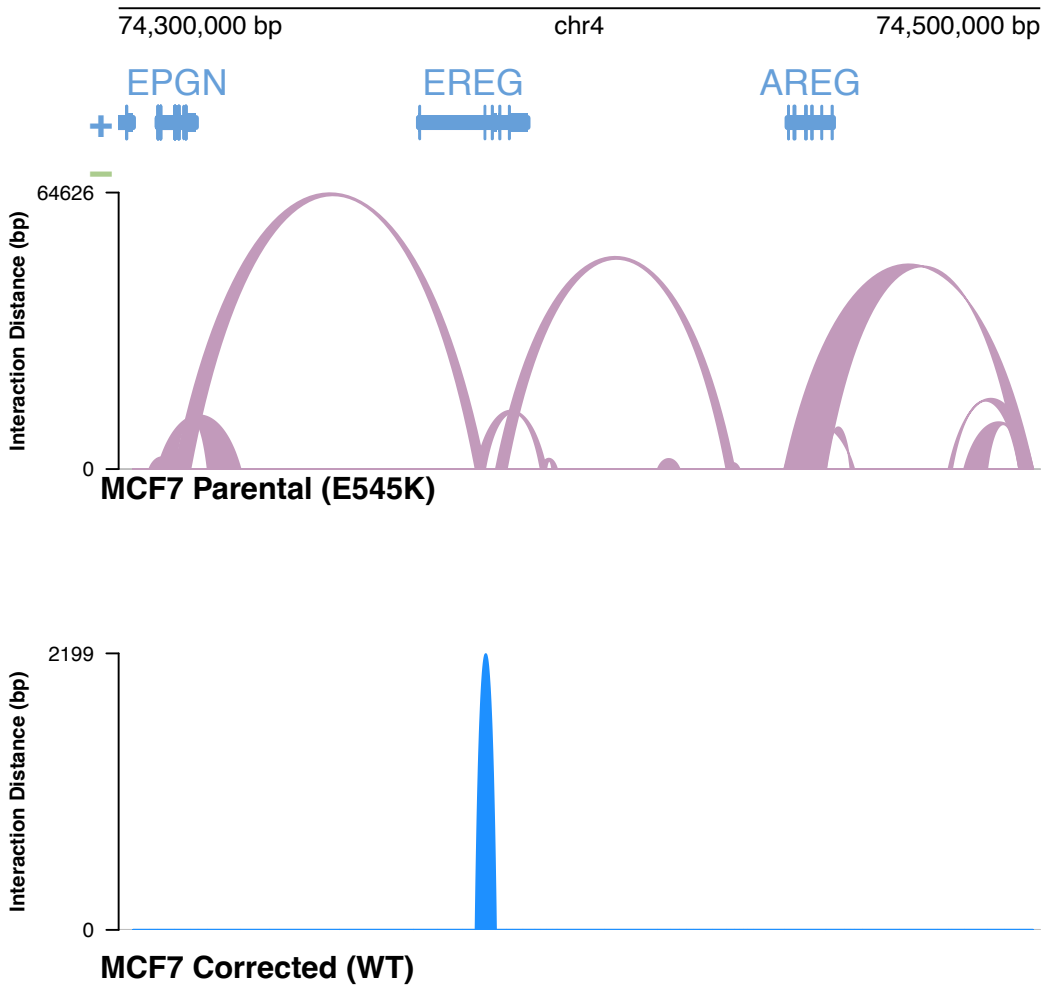

Figure A.2: **Genomic Tracks of Chromatin Interactions at the AREG gene locus** This schematic shows arches between genomic loci shown to interact through protein mediated chromatin loops identified by bead capture HiC.

References

Abud, H. E., Chan, W. H., and Jardé, T. (2021). Source and impact of the egf family of ligands on intestinal stem cells. *Frontiers in Cell and Developmental Biology*, 9:685665.

Aguet, F., Barbeira, A. N., Bonazzola, R., Brown, A., Castel, S. E., Jo, B., Kasela, S., Kim-Hellmuth, S., Liang, Y., Oliva, M., Flynn, E. D., Parsana, P., Fresard, L., Gamazon, E. R., Hamel, A. R., He, Y., Hormozdiari, F., Mohammadi, P., Muñoz-Aguirre, M., Park, Y. S., Saha, A., Segrè, A. V., Strober, B. J., Wen, X., Wucher, V., Ardlie, K. G., Battle, A., Brown, C. D., Cox, N., Das, S., Dermitzakis, E. T., Engelhardt, B. E., Garrido-Martín, D., Gay, N. R., Getz, G. A., Guigó, R., Handsaker, R. E., Hoffman, P. J., Im, H. K., Kashin, S., Kwong, A., Lappalainen, T., Li, X., MacArthur, D. G., Montgomery, S. B., Rouhana, J. M., Stephens, M., Stranger, B. E., Todres, E., Viñuela, A., Wang, G., Zou, Y., Anand, S., Gabriel, S., Graubert, A., Hadley, K., Huang, K. H., Meier, S. R., Nedzel, J. L., Nguyen, D. T., Balliu, B., Conrad, D. F., Cotter, D. J., deGoede, O. M., Einson, J., Eskin, E., Eulalio, T. Y., Ferraro, N. M., Gloudemans, M. J., Hou, L., Kellis, M., Li, X., Mangul, S., Nachun, D. C., Nobel, A. B., Park, Y., Rao, A. S., Reverter, F., Sabatti, C., Skol, A. D., Teran, N. A., Wright, F., Ferreira, P. G., Li, G., Melé, M., Yeger-Lotem, E., Barcus, M. E., Bradbury, D., Krubit, T., McLean, J. A., Qi, L., Robinson, K., Roche, N. V., Smith, A. M., Sobin, L., Tabor, D. E., Undale, A., Bridge, J., Brigham, L. E., Foster, B. A., Gillard, B. M., Hasz, R., Hunter, M., Johns, C., Johnson, M., Karasik, E., Kopen, G., Leinweber, W. F., McDonald, A., Moser, M. T., Myer, K., Ramsey, K. D., Roe, B., Shad, S., Thomas, J. A., Walters, G., Washington, M., Wheeler, J., Jewell, S. D., Rohrer, D. C., Valley, D. R., Davis, D. A., Mash, D. C., Branton, P. A., Sobin, L., Barker, L. K., Gardiner, H. M., Mosavel, M., Siminoff, L. A., Flicek, P., Haeussler, M., Juettemann, T., Kent, W. J., Lee, C. M., Powell, C. C., Rosenbloom, K. R., Ruffier, M., Sheppard, D., Taylor, K., Trevanion, S. J., Zerbino, D. R., Abell, N. S., Akey, J., Chen, L., Demanelis, K., Doherty, J. A., Feinberg, A. P., Hansen, K. D., Hickey, P. F., Hou, L., Jasmine, F., Jiang, L., Kaul, R., Kellis, M., Kibriya, M. G., Li, J. B., Li, Q., Lin, S., Linder, S. E., Pierce, B. L., Rizzardi, L. F., Smith, K. S., Snyder, M., Stamatoyannopoulos, J., Tang, H., Wang, M., Branton, P. A., Carithers, L. J., Guan, P., Koester, S. E., Little, A. R., Moore, H. M., Nierras, C. R., Rao, A. K., Vaught, J. B., and Volpi, S. (2020). The gtex consortium atlas of genetic regulatory effects across human tissues. *Science*, 369:1318–1330.

Alves, A. C., Ribeiro, D., Nunes, C., and Reis, S. (2016). Biophysics in cancer: The relevance of drug-membrane interaction studies. *BBA Biomembranes*, 1858:2231–2244.

Andrews, S. (2010). Fastqc. *Babraham Bioinformatics*.

André, F., Ciruelos, E. M., Juric, D., Loibl, S., Campone, M., Mayer, I. A., Rubovszky, G., Yamashita, T., Kaufman, B., Lu, Y. S., Inoue, K., Pápai, Z., Takahashi, M., Ghaznawi, F., Mills, D., Kaper, M., Miller, M., Conte, P. F., Iwata, H., and Rugo, H. S. (2021). Alpelisib plus fulvestrant for pik3ca-mutated, hormone receptor-positive, human epidermal growth factor receptor-2-negative advanced breast cancer: final overall survival

results from solar-1. *Annals of Oncology : Official Journal of the European Society for Medical Oncology*, 32:208–217.

Bachman, K. E., Argani, P., Samuels, Y., Silliman, N., Ptak, J., Szabo, S., Konishi, H., Karakas, B., Blair, B. G., Lin, C., Peters, B. A., Velculescu, V. E., and Park, B. H. (2004). The pik3ca gene is mutated with high frequency in human breast cancers. *Cancer Biology and Therapy*, 3:772–775.

Bacolla, A., Cooper, D. N., and Vasquez, K. M. (2014). Mechanisms of base substitution mutagenesis in cancer genomes. *Genes (Basel)*, 5:108–146.

Barnett, K. R., Decato, B. E., Scott, T. J., Hansen, T. J., Chen, B., Attalla, J., Smith, A. D., and Hodges, E. (2020). Atac-me captures prolonged dna methylation of dynamic chromatin accessibility loci during cell fate transitions. *Molecular Cell*, 77:1350–1364.

Baselga, J. (2001). Clinical trials of herceptin® (trastuzumab). *European Journal of Cancer*, 37:18–24.

Baselga, J., Im, S. A., Iwata, H., Cortés, J., Laurentiis, M. D., Jiang, Z., Arteaga, C. L., Jonat, W., Clemons, M., Ito, Y., Awada, A., Chia, S., Jagiełło-Gruszfeld, A., Pistilli, B., Tseng, L. M., Hurvitz, S., Masuda, N., Takahashi, M., Vuylsteke, P., Hachemi, S., Dharan, B., Tomaso, E. D., Urban, P., Massacesi, C., and Campone, M. (2017). Buparlisib plus fulvestrant versus placebo plus fulvestrant in postmenopausal, hormone receptor-positive, her2-negative, advanced breast cancer (belle-2): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet Oncology*, 18:904–916.

Beaver, J. A., Gustin, J. P., Yi, K. H., Rajpurohit, A., Thomas, M., Gilbert, S. F., Rosen, D. M., Park, B. H., and Lauring, J. (2013). Pik3ca and akt1 mutations have distinct effects on sensitivity to targeted pathway inhibitors in an isogenic luminal breast cancer model system. *Clinical Cancer Research*, 19:5413–5422.

Belli, C., Repetto, M., Anand, S., Porta, C., Subbiah, V., and Curigliano, G. (2023). The emerging role of pi3k inhibitors for solid tumour treatment and beyond. *British Journal of Cancer*, 128:2150.

Bentsen, M., Goymann, P., Schultheis, H., Klee, K., Petrova, A., Wiegandt, R., Fust, A., Preussner, J., Kuenne, C., Braun, T., Kim, J., and Looso, M. (2020). Atac-seq footprinting unravels kinetics of transcription factor binding during zygotic genome activation. *Nature Communications*, 11:4267.

Bijlani, S., Pang, K. M., Sivanandam, V., Singh, A., and Chatterjee, S. (2021). The role of recombinant aav in precise genome editing. *Frontiers in Genome Editing*, 3.

Bolitho, C., Moscova, M., Baxter, R. C., and Marsh, D. J. (2021). Amphiregulin increases migration and proliferation of epithelial ovarian cancer cells by inducing its own expression via pi3-kinase signaling. *Molecular and Cellular Endocrinology*, 533:111338.

Borreguero-Muñoz, N., Fletcher, G. C., Aguilar-Aragon, M., Elbediwy, A., Vincent-Mistiaen, Z. I., and Thompson, B. J. (2019). The hippo pathway integrates pi3k–akt signals with mechanical and polarity cues to control tissue growth. *PLOS Biology*, 17:e3000509.

Buckbinder, L., Jean, D. J. S., Tieu, T., Ladd, B., Hilbert, B., Wang, W., Alltucker, J. T., Manimala, S., Kryukov, G. V., Brooijmans, N., Dowdell, G., Jonsson, P., Huff, M., Guzman-Perez, A., Jackson, E. L., Goncalves, M. D., and Stuart, D. D. (2023). Stx-478, a mutant-selective, allosteric pi3kα inhibitor spares metabolic dysfunction and improves therapeutic response in pi3kα-mutant xenografts. *Cancer Discovery*, 13:2432–2447.

Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y., and Greenleaf, W. J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, dna-binding proteins and nucleosome position. *Nature Methods*, 10:1213–1218.

Bushnell, B., Rood, J., and Singer, E. (2017). Bbmerge – accurate paired shotgun read merging via overlap. *PLOS ONE*, 12:0185056.

Busser, B., Sancey, L., Brambilla, E., Coll, J. L., and Hurbin, A. (2011). The multiple roles of amphiregulin in human cancer. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*, 1816:119–131.

Buzdin, A., Sorokin, M., Garazha, A., Glusker, A., Aleshin, A., Poddubskaya, E., Sekacheva, M., Kim, E., Gaifullin, N., Giese, A., Seryakov, A., Rumiantsev, P., Moshkovskii, S., and Moiseev, A. (2019). Rna sequencing for research and diagnostics in clinical oncology. *Seminars in Cancer Biology*, 60:311–323.

Carithers, L. J., Ardlie, K., Barcus, M., Branton, P. A., Britton, A., Buia, S. A., Compton, C. C., Deluca, D. S., Peter-Demchok, J., Gelfand, E. T., Guan, P., Korzeniewski, G. E., Lockhart, N. C., Rabiner, C. A., Rao, A. K., Robinson, K. L., Roche, N. V., Sawyer, S. J., Segrè, A. V., Shive, C. E., Smith, A. M., Sobin, L. H., Undale, A. H., Valentino, K. M., Vaught, J., Young, T. R., Moore, H. M., Barker, L., Basile, M., Battle, A., Boyer, J., Bradbury, D., Bridge, J. P., Brown, A., Burges, R., Choi, C., Colantuoni, D., Cox, N., Dermitzakis, E. T., Derr, L. K., Dinsmore, M. J., Erickson, K., Fleming, J., Flutre, T., Foster, B. A., Gamazon, E. R., Getz, G., Gillard, B. M., Guigó, R., Hambright, K. W., Hariharan, P., Hasz, R., Im, H. K., Jewell, S., Karasik, E., Kellis, M., Kheradpour, P., Koester, S., Koller, D., Konkashbaev, A., Lappalainen, T., Little, R., Liu, J., Lo, E., Lonsdale, J. T., Lu, C., MacArthur, D. G., Magazine, H., Maller, J. B., Marcus, Y., Mash, D. C., McCarthy, M. I., McLean, J., Mestichelli, B., Miklos, M., Monlong, J., Mosavel, M., Moser, M. T., Mostafavi, S., Nicolae, D. L., Pritchard, J., Qi, L., Ramsey, K., Rivas, M. A., Robles, B. E., Rohrer, D. C., Salvatore, M., Sammeth, M., Seleski, J., Shad, S., Siminoff, L. A., Stephens, M., Struewing, J., Sullivan, T., Sullivan, S., Syron, J., Tabor, D., Taherian, M., Tejada, J., Temple, G. F., Thomas, J. A., Thomson, A. W., Tidwell, D., Traino, H. M., Tu, Z., Valley, D. R., Volpi, S., Walters, G. D., Ward, L. D., Wen, X., Winckler, W., Wu, S., Zhu, J., Abdallah, A., Addington, A., Anderson, J. M., Bender, P. K., Cosentino, M., Diaz-Mayoral, N., Engel, T., Garci, F., Green, A., Hammond, T.,

Jaffe, K., Keen, J., Kennedy, M., Kigonya, P., Lander, B., Nampally, S., Ny, C., Robb, J., Santhanum, V., Sharopova, N., Singh, S., Soria, C., Sturcke, A., Sukari, S., Thomson, E. J., Tomaszewski, M., Trowbridge, C., Udoye, F., Vanscoy, D., Vatanian, N., Wilder, E. L., and Williams, P. (2015). A novel approach to high-quality postmortem tissue procurement: The gtex project. *Biopreservation and Biobanking*, 13:311–317.

Cerami, E., Gao, J., Dogrusoz, U., Gross, B. E., Sumer, S. O., Aksoy, B. A., Jacobsen, A., Byrne, C. J., Heuer, M. L., Larsson, E., Antipin, Y., Reva, B., Goldberg, A. P., Sander, C., and Schultz, N. (2012). The cbio cancer genomics portal: An open platform for exploring multidimensional cancer genomics data. *Cancer Discovery*, 2:401–404.

Chalmers, Z. R., Connelly, C. F., Fabrizio, D., Gay, L., Ali, S. M., Ennis, R., Schrock, A., Campbell, B., Shlien, A., Chmielecki, J., Huang, F., He, Y., Sun, J., Tabori, U., Kennedy, M., Lieber, D. S., Roels, S., White, J., Otto, G. A., Ross, J. S., Garraway, L., Miller, V. A., Stephens, P. J., and Frampton, G. M. (2017). Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. *Genome Medicine*, 9:1–14.

Chau, V. and Loxo Oncology, I. (2024). A study of loxo-783 in patients with breast cancer other solid tumors (pikasso-01). *ClinicalTrials.gov*, NCT05307705.

Chen, E. Y., Tan, C. M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G. V., Clark, N. R., and Ma'ayan, A. (2013). Enrichr: Interactive and collaborative html5 gene list enrichment analysis tool. *BMC Bioinformatics*, 14:1–14.

Cheng, C., Chen, W., Jin, H., and Chen, X. (2023). A review of single-cell rna-seq annotation, integration, and cell–cell communication. *Cells*, 12:1970.

Chi, M., Liu, J., Mei, C., Shi, Y., Liu, N., Jiang, X., Liu, C., Xue, N., Hong, H., Xie, J., Sun, X., Yin, B., Meng, X., and Wang, B. (2022). Tead4 functions as a prognostic biomarker and triggers emt via pi3k/akt pathway in bladder cancer. *Journal of Experimental and Clinical Cancer Research*, 41:1–20.

Chung, E., Graves-Deal, R., Franklin, J. L., and Coffey, R. J. (2005). Differential effects of amphiregulin and tgf-a on the morphology of mdck cells. *Experimental Cell Research*, 309:149–160.

Clusan, L., Ferrière, F., Flouriot, G., and Pakdel, F. (2023). A basic review on estrogen receptor signaling pathways in breast cancer. *International Journal of Molecular Sciences*, 24:6834.

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A., and Zhang, F. (2013). Multiplex genome engineering using crispr/cas systems. *Science*, 339:819–823.

Corces, M. R., Granja, J. M., Shams, S., Louie, B. H., Seoane, J. A., Zhou, W., Silva, T. C., Groeneveld, C., Wong, C. K., Cho, S. W., Satpathy, A. T., Mumbach, M. R., Hoadley, K. A., Robertson, A. G., Sheffield, N. C., Felau, I., Castro, M. A., Berman,

B. P., Staudt, L. M., Zenklusen, J. C., Laird, P. W., Curtis, C., Greenleaf, W. J., and Chang, H. Y. (2018). The chromatin accessibility landscape of primary human cancers. *Science*, 362:1898.

Crino, P. B., Nathanson, K. L., and Henske, E. P. (2006). The tuberous sclerosis complex. *New England Journal of Medicine*, 355:1345–56.

Damodaran, S., Zhao, F., Deming, D. A., Mitchell, E. P., Wright, J. J., Gray, R. J., Wang, V., McShane, L. M., Rubinstein, L. V., Patton, D. R., Williams, P. M., Hamilton, S. R., Suga, J. M., Conley, B. A., Arteaga, C. L., Harris, L. N., O'Dwyer, P. J., Chen, A. P., and Flaherty, K. T. (2022). Phase ii study of copanlisib in patients with tumors with pik3ca mutations: Results from the nci-match ecog-acrin trial (eay131) subprotocol z1f. *Journal of Clinical Oncology*, 40:1552–1561.

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., and Li, H. (2021). Twelve years of samtools and bcftools. *GigaScience*, 10:1–4.

Dasari, S. and Tchounwou, P. B. (2014). Cisplatin in cancer therapy: molecular mechanisms of action. *European Journal of Pharmacology*, 740:364.

Davidson, B. A., Miranda, A. X., Reed, S. C., Bergman, R. E., Kemp, J. D., Reddy, A. P., Pantone, M. V., Fox, E. K., Dorand, R. D., Hurley, P. J., Croessmann, S., and Park, B. H. (2024). An in vitro crispr screen of cell-free dna identifies apoptosis as the primary mediator of cell-free dna release. *Communications Biology 2024 7:1*, 7:1–15.

Davies, B. R., Greenwood, H., Dudley, P., Crafter, C., Yu, D. H., Zhang, J., Li, J., Gao, B., Ji, Q., Maynard, J., Ricketts, S. A., Cross, D., Cosulich, S., Chresta, C. C., Page, K., Yates, J., Lane, C., Watson, R., Luke, R., Ogilvie, D., and Pass, M. (2012). Preclinical pharmacology of azd5363, an inhibitor of akt: Pharmacodynamics, antitumor activity, and correlation of monotherapy activity with genetic background. *Molecular Cancer Therapeutics*, 11:873–887.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T. R. (2013). Sequence analysis star: ultrafast universal rna-seq aligner. *Bioinformatics*, 29:15–21.

Doench, J. G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E. W., Donovan, K. F., Smith, I., Tothova, Z., Wilen, C., Orchard, R., Virgin, H. W., Listgarten, J., and Root, D. E. (2016). Optimized sgrna design to maximize activity and minimize off-target effects of crispr-cas9. *Nature Biotechnology*, 34:184–191.

Dong, X. and Weng, Z. (2013). The correlation between histone modifications and gene expression. *Epigenomics*, 5:113.

Ekanayake, P. S., Gerwer, J., and McCowen, K. (2022). Alpelisib - induced hyperglycemia. *Acta Endocrinologica (Bucharest)*, 18:115–117.

Engelman, J. A., Luo, J., and Cantley, L. C. (2006). The evolution of phosphatidylinositol 3-kinases as regulators of growth and metabolism. *Nature Reviews Genetics*, 7(8):606–619.

Ergin, S., Kherad, N., and Alagoz, M. (2022). Rna sequencing and its applications in cancer and rare diseases. *Molecular Biology Reports*, 49:2325–2333.

Fearon, E. R. and Vogelstein, B. (1990). A genetic model for colorectal tumorigenesis. *Cell*, 61:759–767.

Fiascarelli, A., Merlino, G., Capano, S., Talucci, S., Bisignano, D., Bressan, A., Bellarosa, D., Carrisi, C., Paoli, A., Bigioni, M., Tunici, P., Irrissuto, C., Salerno, M., Arribas, J., de Stanchina, E., Scaltriti, M., and Binaschi, M. (2023). Antitumor activity of the pi3k δ-sparing inhibitor men1611 in pik3ca mutated, trastuzumab-resistant her2breast cancer. *Breast Cancer Research and Treatment*, 199:13.

Flavahan, W. A., Drier, Y., Liau, B. B., Gillespie, S. M., Venteicher, A. S., Stemmer-Rachamimov, A. O., Suvà, M. L., and Bernstein, B. E. (2016). Insulator dysfunction and oncogene activation in idh mutant gliomas. *Nature*, 529:110–114.

Fox, M., Mott, H. R., and Owen, D. (2020). Class IA PI3K regulatory subunits: p110-independent roles and structures. *Biochemical Society Transactions*, 48(4):1397.

Frankish, A., Diekhans, M., Ferreira, A. M., Johnson, R., Jungreis, I., Loveland, J., Mudge, J. M., Sisu, C., Wright, J., Armstrong, J., Barnes, I., Berry, A., Bignell, A., Sala, S. C., Chrast, J., Cunningham, F., Domenico, T. D., Donaldson, S., Fiddes, I. T., Girón, C. G., Gonzalez, J. M., Grego, T., Hardy, M., Hourlier, T., Hunt, T., Izuogu, O. G., Lagarde, J., Martin, F. J., Martínez, L., Mohanan, S., Muir, P., Navarro, F. C., Parker, A., Pei, B., Pozo, F., Ruffier, M., Schmitt, B. M., Stapleton, E., Suner, M. M., Sycheva, I., Uszczynska-Ratajczak, B., Xu, J., Yates, A., Zerbino, D., Zhang, Y., Aken, B., Choudhary, J. S., Gerstein, M., Guigó, R., Hubbard, T. J., Kellis, M., Paten, B., Reymond, A., Tress, M. L., and Flicek, P. (2019). Gencode reference annotation for the human and mouse genomes. *Nucleic Acids Research*, 47:D766–D773.

Fuso, P., Muratore, M., D'angelo, T., Paris, I., Carbognin, L., Tiberi, G., Pavese, F., Duranti, S., Orlandi, A., Tortora, G., Scambia, G., and Fabi, A. (2022). Pi3k inhibitors in advanced breast cancer: The past, the present, new challenges and future perspectives. *Cancers (Basel)*, 14:2161.

Gabelli, S., Huang, C.-H., Mandelker, D., Schmidt-Kittler, O., Vogelstein, B., and Amzel, L. M. (2010). Structural effects of oncogenic pi3kα mutations. *Current Topics in Microbiology Immunology*, 347:43–53.

Gao, J., Aksoy, B. A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S. O., Sun, Y., Jacobsen, A., Sinha, R., Larsson, E., Cerami, E., Sander, C., and Schultz, N. (2013). Integrative analysis of complex cancer genomics and clinical profiles using the cbioportal. *Science Signaling*, 6:269.

Garay, J. P., Smith, R., Devlin, K., Hollern, D. P., Liby, T., Liu, M., Boddapati, S., Watson, S. S., Esch, A., Zheng, T., Thompson, W., Babcock, D., Kwon, S., Chin, K., Heiser, L., Gray, J. W., and Korkola, J. E. (2021). Sensitivity to targeted therapy differs between her2-amplified breast cancer cells harboring kinase and helical domain mutations in pik3ca. *Breast Cancer Research*, 23:81.

Ge, X., Behrendt, C. E., Yost, S. E., Patel, N., Samoa, R., Stewart, D., Sedrak, M., Lavasani, S., Waisman, J., Yuan, Y., and Mortimer, J. (2023). Predicting hyperglycemia among patients receiving alpelisib plus fulvestrant for metastatic breast cancer. *The Oncologist*, 28:e488.

Ghigo, A., Morello, F., Perino, A., and Hirsch, E. (2015). Phosphoinositide 3-kinases in health and disease. *Subcellular Biochemistry*, 58:183–213.

Gordon, L. G., White, N. M., Elliott, T. M., Nones, K., Beckhouse, A. G., Rodriguez-Acevedo, A. J., Webb, P. M., Lee, X. J., Graves, N., and Schofield, D. J. (2020). Estimating the costs of genomic sequencing in cancer control. *BMC Health Services Research*, 20.

Grossman, R. L., Heath, A. P., Ferretti, V., Varmus, H. E., Lowy, D. R., Kibbe, W. A., and Staudt, L. M. (2016). Toward a shared vision for cancer genomic data. *New England Journal of Medicine*, 375:1109–1112.

Gustin, J. P., Karakas, B., Weiss, M. B., Abukhdeir, A. M., Lauring, J., Garay, J. P., Cosgrove, D., Tamaki, A., Konishi, H., Konishi, Y., Mohseni, M., Wang, G., Rosen, D. M., Denmeade, S. R., Higgins, M. J., Vitolo, M. I., Bachman, K. E., and Park, B. H. (2009). Knockin of mutant pik3ca activates multiple oncogenic pathways. *Proceedings of the National Academy of Sciences of the United States of America*, 106:2835–2840.

Gymnopoulos, M., Elsliger, M. A., and Vogt, P. K. (2007). Rare cancer-specific mutations in pik3ca show gain of function. *Proceedings of the National Academy of Sciences of the United States of America*, 104(13):5569.

Hanahan, D. (2022). Hallmarks of cancer: New dimensions. *Cancer discovery*, 12:31–46.

Hart, J. R., Zhang, Y., Liao, L., Ueno, L., Du, L., Jonkers, M., Yates, J. R., and Vogt, P. K. (2015). The butterfly effect in cancer: A single base mutation can remodel the cell. *Proceedings of the National Academy of Sciences of the United States of America*, 112:1131–1136.

Hayakawa, J., Ohmichi, M., Kurachi, H., Kanda, Y., Hisamoto, K., Nishio, Y., Adachi, K., Tasaka, K., Kanzaki, T., and Murata, Y. (2000). Inhibition of bad phosphorylation either at serine 112 via extracellular signal-regulated protein kinase cascade or at serine 136 via akt cascade sensitizes human ovarian cancer cells to cisplatin. *Cancer Research*, 60:5988–5994.

He, C., Han, S., Chang, Y., Wu, M., Zhao, Y., Chen, C., and Chu, X. (2021). Crispr screen in cancer: status quo and future perspectives. *American Journal of Cancer Research*, 11:1031.

Hedenfalk, I. A., Garzia, L., Liu, B., Li, X., Wang, Y., Mashock, M., Tong, Z., Mu, X., Chen, H., Zhou, X., Zhang, H., and Zhao, G. (2020). Changing technologies of rna sequencing and their applications in clinical oncology. *Frontiers in Oncology*, 1:447.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., Cheng, J. X., Murre, C., Singh, H., and Glass, C. K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and b cell identities. *Molecular cell*, 38:576–589.

Hodges, E., Xuan, Z., Balija, V., Kramer, M., Molla, M. N., Smith, S. W., Middle, C. M., Rodesch, M. J., Albert, T. J., Hannon, G. J., and McCombie, W. R. (2007). Genome-wide in situ exon capture for selective resequencing. *Nature Genetics 2007 39:12*, 39:1522–1527.

Holbro, T., Civenni, G., and Hynes, N. E. (2003). The erbb receptors and their role in cancer progression. *Experimental Cell Research*, 284:99–110.

Hong, D. S., Postow, M., Chmielowski, B., Sullivan, R., Patnaik, A., Cohen, E. E., Shapiro, G., Steuer, C., Gutierrez, M., Yeckes-Rodin, H., Ilaria, R., O'Connell, B., Peng, J., Peng, G., Zizlsperger, N., Tolcher, A., and Wolchok, J. D. (2023). Eganelisib, a first-in-class pi3k-[... formula ...] inhibitor, in patients with advanced solid tumors: Results of the phase 1/1b mario-1 trial. *Clinical cancer research : an official journal of the American Association for Cancer Research*, 29:2210.

Hood, L. and Rowen, L. (2013). The human genome project: Big science transforms biology and medicine. *Genome Medicine*, 5:1–8.

Horwitz, K. B., Costlow, M. E., and McGuire, W. L. (1975). Mcf-7: A human breast cancer cell line with estrogen, androgen, progesterone, and glucocorticoid receptors. *Steroids*, 26:785–795.

Howell, S. J., Casbard, A., Carucci, M., Ingarfield, K., Butler, R., Morgan, S., Meissner, M., Bale, C., Bezecny, P., Moon, S., Twelves, C., Venkitaraman, R., Waters, S., de Bruin, E. C., Schiavon, G., Foxley, A., and Jones, R. H. (2022). Fulvestrant plus capivasertib versus placebo after relapse or progression on an aromatase inhibitor in metastatic, oestrogen receptor-positive, her2-negative breast cancer (faktion): overall survival, updated progression-free survival, and expanded biomarker analysis from a randomised, phase 2 trial. *The Lancet. Oncology*, 23:851.

Hoxhaj, G. and Manning, B. D. (2019). The pi3k–akt network at the interface of oncogenic signalling and cancer metabolism. *Nature Reviews Cancer*, 20:74–88.

Huang, J. and Manning, B. D. (2009). A complex interplay between akt, tsc2, and the two mtor complexes. *Biochemical Society transactions*, 37:217.

Jaiswal, B. S., Janakiraman, V., Kljavin, N. M., Chaudhuri, S., Stern, H. M., Wang, W., Kan, Z., Dbouk, H. A., Peters, B. A., Waring, P., Dela Vega, T., Kenski, D. M., Bowman, K. K., Lorenzo, M., Li, H., Wu, J., Modrusan, Z., Stinson, J., Eby, M., Yue, P., Kaminker,

J. S., de Sauvage, F. J., Backer, J. M., and Seshagiri, S. (2009). Somatic mutations in p85α promote tumorigenesis through class IA PI3K activation. *Cancer Cell*, 16(6):463.

Janku, F., Wheler, J. J., Naing, A., Falchook, G. S., Hong, D. S., Stepanek, V. M., Fu, S., Piha-Paul, S. A., Lee, J. J., Luthra, R., Tsimberidou, A. M., and Kurzrock, R. (2013). Pik3ca mutation h1047r is associated with response to pi3k/akt/mtor signaling pathway inhibitors in early phase clinical trials. *Cancer research*, 73:276.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012). A programmable dual-rna-guided dna endonuclease in adaptive bacterial immunity. *Science*, 337:816–821.

Joung, J., Konermann, S., Gootenberg, J. S., Abudayyeh, O. O., Platt, R. J., Brigham, M. D., Sanjana, N. E., and Zhang, F. (2017). Genome-scale crispr-cas9 knockout and transcriptional activation screening. *Nature Protocols 2017 12:4*, 12:828–863.

Juric, D., Kalinsky, K., Im, S.-A., Ciruelos, E. M., Bianchini, G., Barrios, C. H., Jacot, W., Schmid, P., Loi, S., Rugo, H. S., Craine, V., Hutchinson, K. E., Flechais, A., Thanopoulou, E., and Harbeck, N. (2023). Inavo121: Phase iii study of inavolisib (inavo) + fulvestrant (ful) vs. alpelisib (alp) + ful in patients (pts) with hormone receptor-positive, her2-negative (hr+, her2–) pik3ca-mutated (mut) locally advanced or metastatic breast cancer (la/mbc). *Journal of Clinical Oncology*, 41:TPS1123–TPS1123.

Kagey, M. H., Newman, J. J., Bilodeau, S., Zhan, Y., Orlando, D. A., Berkum, N. L. V., Ebmeier, C. C., Goossens, J., Rahl, P. B., Levine, S. S., Taatjes, D. J., Dekker, J., and Young, R. A. (2010). Mediator and cohesin connect gene expression and chromatin architecture. *Nature*, 467:430–435.

Kim, N. W., Piatyszek, M. A., Prowse, K. R., Harley, C. B., West, M. D., Ho, P. L., Coviello, G. M., Wright, W. E., Weinrich, S. L., and Shay, J. W. (1994). Specific association of human telomerase activity with immortal cells and cancer. *Science*, 266:2011–2015.

Klippel, A., Wang, R., Puca, L., Faber, A. L., Shen, W., Bhagwat, S. V., Karukurichi, K., Zhang, F. F., Perez, C., Rama-Garda, R., Ramos, A., Zheng, Y., Bonday, Z., Thomas, J., Brooks, H. B., Kindler, L. J., Bogner, S. M., Zolfaghari, P., II, M. H., Callies, S., Mattioni, B., Lebrun, L., Durbin, J., Anderson, E., Christopher G. Mayne, E. A. K., Kolakowski, G., Andrews, S. W., and Brandhuber, B. J. (2021). Preclinical characterization of loxo-783 (lox-22783), a highly potent, mutant-selective and brain-penetrant allosteric pi3kα h1047r inhibitor. *AACR-NCI-EORTC Virtual International Conference On Molecular Targets And Cancer Therapeutics; Conference presentation of unpublished data.*

Koboldt, D. C., Fulton, R. S., McLellan, M. D., Schmidt, H., Kalicki-Veizer, J., McMichael, J. F., Fulton, L. L., Dooling, D. J., Ding, L., Mardis, E. R., Wilson, R. K., Ally, A., Balasundaram, M., Butterfield, Y. S., Carlsen, R., Carter, C., Chu, A., Chuah, E., Chun, H. J. E., Coope, R. J., Dhalla, N., Guin, R., Hirst, C., Hirst, M., Holt, R. A., Lee, D., Li, H. I., Mayo, M., Moore, R. A., Mungall, A. J., Pleasance, E., Robertson,

A. G., Schein, J. E., Shafiei, A., Sipahimalani, P., Slobodan, J. R., Stoll, D., Tam, A., Thiessen, N., Varhol, R. J., Wye, N., Zeng, T., Zhao, Y., Birol, I., Jones, S. J., Marra, M. A., Cherniack, A. D., Saksena, G., Onofrio, R. C., Pho, N. H., Carter, S. L., Schumacher, S. E., Tabak, B., Hernandez, B., Gentry, J., Nguyen, H., Crenshaw, A., Ardlie, K., Beroukhim, R., Winckler, W., Getz, G., Gabriel, S. B., Meyerson, M., Chin, L., Kucherlapati, R., Hoadley, K. A., Auman, J. T., Fan, C., Turman, Y. J., Shi, Y., Li, L., Topal, M. D., He, X., Chao, H. H., Prat, A., Silva, G. O., Iglesia, M. D., Zhao, W., Usary, J., Berg, J. S., Adams, M., Booker, J., Wu, J., Gulabani, A., Bodenheimer, T., Hoyle, A. P., Simons, J. V., Soloway, M. G., Mose, L. E., Jefferys, S. R., Balu, S., Parker, J. S., Hayes, D. N., Perou, C. M., Malik, S., Mahurkar, S., Shen, H., Weisenberger, D. J., Triche, T., Lai, P. H., Bootwalla, M. S., Maglinte, D. T., Berman, B. P., Van Den Berg, D. J., Baylin, S. B., Laird, P. W., Creighton, C. J., Donehower, L. A., Noble, M., Voet, D., Gehlenborg, N., Di Cara, D., Zhang, J., Zhang, H., Wu, C. J., Yingchun Liu, S., Lawrence, M. S., Zou, L., Sivachenko, A., Lin, P., Stojanov, P., Jing, R., Cho, J., Sinha, R., Park, R. W., Nazaire, M. D., Robinson, J., Thorvaldsdottir, H., Mesirov, J., Park, P. J., Reynolds, S., Kreisberg, R. B., Bernard, B., Bressler, R., Erkkila, T., Lin, J., Thorsson, V., Zhang, W., Shmulevich, I., Ciriello, G., Weinhold, N., Schultz, N., Gao, J., Cerami, E., Gross, B., Jacobsen, A., Sinha, R., Aksoy, B. A., Antipin, Y., Reva, B., Shen, R., Taylor, B. S., Ladanyi, M., Sander, C., Anur, P., Spellman, P. T., Lu, Y., Liu, W., Verhaak, R. R., Mills, G. B., Akbani, R., Zhang, N., Broom, B. M., Casasent, T. D., Wakefield, C., Unruh, A. K., Baggerly, K., Coombes, K., Weinstein, J. N., Haussler, D., Benz, C. C., Stuart, J. M., Benz, S. C., Zhu, J., Szeto, C. C., Scott, G. K., Yau, C., Paull, E. O., Carlin, D., Wong, C., Sokolov, A., Thusberg, J., Mooney, S., Ng, S., Goldstein, T. C., Ellrott, K., Grifford, M., Wilks, C., Ma, S., Craft, B., Yan, C., Hu, Y., Meerzaman, D., Gastier-Foster, J. M., Bowen, J., Ramirez, N. C., Black, A. D., Pyatt, R. E., White, P., Zmuda, E. J., Frick, J., Lichtenberg, T. M., Brookens, R., George, M. M., Gerken, M. A., Harper, H. A., Leraas, K. M., Wise, L. J., Tabler, T. R., McAllister, C., Barr, T., Hart-Kothari, M., Tarvin, K., Saller, C., Sandusky, G., Mitchell, C., Iacocca, M. V., Brown, J., Rabeno, B., Czerwinski, C., Petrelli, N., Dolzhansky, O., Abramov, M., Voronina, O., Potapova, O., Marks, J. R., Suchorska, W. M., Murawa, D., Kycler, W., Ibbs, M., Korski, K., Spychała, A., Murawa, P., Brzeziński, J. J., Perz, H., Łaźniak, R., Teresiak, M., Tatka, H., Leporowska, E., Bogusz-Czerniewicz, M., Malicki, J., Mackiewicz, A., Wiznerowicz, M., Van Le, X., Kohl, B., Viet Tien, N., Thorp, R., Van Bang, N., Sussman, H., Phu, B. D., Hajek, R., Hung, N. P., Phuong, T. V. T., Thang, H. Q., Khan, K. Z., Penny, R., Mallery, D., Curley, E., Shelton, C., Yena, P., Ingle, J. N., Couch, F. J., Lingle, W. L., King, T. A., Gonzalez-Angulo, A. M., Dyer, M. D., Liu, S., Meng, X., Patangan, M., Waldman, F., Stöppler, H., Rathmell, W. K., Thorne, L., Huang, M., Boice, L., Hill, A., Morrison, C., Gaudioso, C., Bshara, W., Daily, K., Egea, S. C., Pegram, M. D., Gomez-Fernandez, C., Dhir, R., Bhargava, R., Brufsky, A., Shriver, C. D., Hooke, J. A., Campbell, J. L., Mural, R. J., Hu, H., Somiari, S., Larson, C., Deyarmin, B., Kvecher, L., Kovatich, A. J., Ellis, M. J., Stricker, T., White, K., Olopade, O., Luo, C., Chen, Y., Bose, R., Chang, L. W., Beck, A. H., Pihl, T., Jensen, M., Sfeir, R., Kahn, A., Chu, A., Kothiyal, P., Wang, Z., Snyder, E., Pontius, J., Ayala, B., Backus, M., Walton, J., Baboud, J., Berton, D., Nicholls, M., Srinivasan, D., Raman, R., Girshik, S., Kigonya, P., Alonso, S., Sanbhadti, R., Barletta, S., Pot, D., Sheth, M., Demchok, J. A., Shaw, K. R., Yang, L., Eley, G.,

Ferguson, M. L., Tarnuzzer, R. W., Zhang, J., Dillon, L. A., Buetow, K., Fielding, P., Ozenberger, B. A., Guyer, M. S., Sofia, H. J., and Palchik, J. D. (2012). Comprehensive molecular portraits of human breast tumours. *Nature*, 490(7418):61–70.

Korbecki, J., Kupnicka, P., Chlubek, M., Goracy, J., Gutowska, I., and Baranowska-Bosiacka, I. (2022). Cxcr2 receptor: Regulation of expression, signal transduction, and involvement in cancer. *International Journal of Molecular Sciences*, 23.

Korotkevich, G., Sukhov, V., Budin, N., Shpak, B., Artyomov, M. N., and Sergushichev, A. (2021). Fast gene set enrichment analysis. *Biorxiv*.

Kuchenbaecker, K. B., Hopper, J. L., Barnes, D. R., Phillips, K. A., Mooij, T. M., Roos-Blom, M. J., Jervis, S., van Leeuwen, F. E., Milne, R. L., Andrieu, N., Goldgar, D. E., Terry, M. B., Rookus, M. A., Easton, D. F., Antoniou, A. C., McGuffog, L., Evans, D. G., Barrowdale, D., Frost, D., Adlard, J., Ong, K. R., Izatt, L., Tischkowitz, M., Eeles, R., Davidson, R., Hodgson, S., Ellis, S., Nogues, C., Lasset, C., Stoppa-Lyonnet, D., Fricker, J. P., Faivre, L., Berthet, P., Hooning, M. J., van der Kolk, L. E., Kets, C. M., Adank, M. A., John, E. M., Chung, W. K., Andrulis, I. L., Southey, M., Daly, M. B., Buys, S. S., Osorio, A., Engel, C., Kast, K., Schmutzler, R. K., Caldes, T., Jakubowska, A., Simard, J., Friedlander, M. L., McLachlan, S. A., Machackova, E., Foretova, L., Tan, Y. Y., Singer, C. F., Olah, E., Gerdes, A. M., Arver, B., and Olsson, H. (2017). Risks of breast, ovarian, and contralateral breast cancer for brca1 and brca2 mutation carriers. *JAMA*, 317:2402–2416.

Kuleshov, M. V., Jones, M. R., Rouillard, A. D., Fernandez, N. F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S. L., Jagodnik, K. M., Lachmann, A., McDermott, M. G., Monteiro, C. D., Gundersen, G. W., and Maayan, A. (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic acids research*, 44:W90–W97.

Kuo, C. J., Conley, P. B., Chen, L., Sladek, F. M., Darnell, J. E., and Crabtree, G. R. (1992). A transcriptional hierarchy involved in mammalian cell-type specification. *Nature*, 355:457–461.

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., Fitzhugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., Levine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J. P., Miranda, C., Morris, W., Naylor, J., Raymond, C., Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Wyman, D., Rogers, J., Sulston, J., Ainscough, R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter, N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A., Jones, M., Lloyd, C., McMurray, A., Matthews, L., Mercer, S., Milne, S., Mullikin, J. C., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Waterston, R. H., Wilson, R. K., Hillier, L. W., McPherson, J. D., Marra, M. A., Mardis, E. R., Fulton, L. A., Chinwalla, A. T., Pepin, K. H., Gish, W. R., Chissoe, S. L., Wendl, M. C., Delehaunty, K. D., Miner, T. L., Delehaunty, A., Kramer, J. B., Cook, L. L., Fulton, R. S., Johnson, D. L., Minx, P. J., Clifton, S. W., Hawkins, T., Branscomb, E., Predki, P., Richardson,

P., Wenning, S., Slezak, T., Doggett, N., Cheng, J. F., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., Gibbs, R. A., Muzny, D. M., Scherer, S. E., Bouck, J. B., Sodergren, E. J., Worley, K. C., Rives, C. M., Gorrell, J. H., Metzker, M. L., Naylor, S. L., Kucherlapati, R. S., Nelson, D. L., Weinstock, G. M., Sakaki, Y., Fujiyama, A., Hattori, M., Yada, T., Toyoda, A., Itoh, T., Kawagoe, C., Watanabe, H., Totoki, Y., Taylor, T., Weissenbach, J., Heilig, R., Saurin, W., Artiguenave, F., Brottier, P., Bruls, T., Pelletier, E., Robert, C., Wincker, P., Rosenthal, A., Platzer, M., Nyakatura, G., Taudien, S., Rump, A., Smith, D. R., Doucette-Stamm, L., Rubenfield, M., Weinstock, K., Hong, M. L., Dubois, J., Yang, H., Yu, J., Wang, J., Huang, G., Gu, J., Hood, L., Rowen, L., Madan, A., Qin, S., Davis, R. W., Federspiel, N. A., Abola, A. P., Proctor, M. J., Roe, B. A., Chen, F., Pan, H., Ramser, J., Lehrach, H., Reinhardt, R., McCombie, W. R., Bastide, M. D. L., Dedhia, N., Blöcker, H., Hornischer, K., Nordsiek, G., Agarwala, R., Aravind, L., Bailey, J. A., Bateman, A., Batzoglou, S., Birney, E., Bork, P., Brown, D. G., Burge, C. B., Cerutti, L., Chen, H. C., Church, D., Clamp, M., Copley, R. R., Doerks, T., Eddy, S. R., Eichler, E. E., Furey, T. S., Galagan, J., Gilbert, J. G., Harmon, C., Hayashizaki, Y., Haussler, D., Hermjakob, H., Hokamp, K., Jang, W., Johnson, L. S., Jones, T. A., Kasif, S., Kaspryzk, A., Kennedy, S., Kent, W. J., Kitts, P., Koonin, E. V., Korf, I., Kulp, D., Lancet, D., Lowe, T. M., McLysaght, A., Mikkelsen, T., Moran, J. V., Mulder, N., Pollara, V. J., Ponting, C. P., Schuler, G., Schultz, J., Slater, G., Smit, A. F., Stupka, E., Szustakowki, J., Thierry-Mieg, D., Thierry-Mieg, J., Wagner, L., Wallis, J., Wheeler, R., Williams, A., Wolf, Y. I., Wolfe, K. H., Yang, S. P., Yeh, R. F., Collins, F., Guyer, M. S., Peterson, J., Felsenfeld, A., Wetterstrand, K. A., Myers, R. M., Schmutz, J., Dickson, M., Grimwood, J., Cox, D. R., Olson, M. V., Kaul, R., Raymond, C., Shimizu, N., Kawasaki, K., Minoshima, S., Evans, G. A., Athanasiou, M., Schultz, R., Patrinos, A., and Morgan, M. J. (2001). Initial sequencing and analysis of the human genome. *Nature 2001 409:6822*, 409:860–921.

Lee, A. V., Oesterreich, S., and Davidson, N. E. (2015). Mcf-7 cells—changing the course of breast cancer research and care for 45 years. *JNCI: Journal of the National Cancer Institute*, 107:73.

Lee, S., Medina, D., Tsimelzon, A., Mohsin, S. K., Mao, S., Wu, Y., and Allred, D. C. (2007). Alterations of gene expression in the development of early hyperplastic precursors of breast cancer. *American Journal of Pathology*, 171:252–262.

Li, W., Köster, J., Xu, H., Chen, C. H., Xiao, T., Liu, J. S., Brown, M., and Liu, X. S. (2015). Quality control, modeling, and visualization of crispr screens with mageck-vispr. *Genome Biology*, 16:1–13.

Li, W., Xu, H., Xiao, T., Cong, L., Love, M. I., Zhang, F., Irizarry, R. A., Liu, J. S., Brown, M., and Liu, X. S. (2014). Mageck enables robust identification of essential genes from genome-scale crispr/cas9 knockout screens. *Genome biology*, 15:554.

Liao, Y., Smyth, G. K., and Shi, W. (2014). Featurecounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, 30:923–930.

Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J. P., and Tamayo, P. (2015). The molecular signatures database hallmark gene set collection. *Cell Systems*, 1:417–425.

Lindzen, M., Ghosh, S., Noronha, A., Drago, D., Nataraj, N. B., Leitner, O., Carvalho, S., Zmora, E., Sapoznik, S., Shany, K. B., Levanon, K., Aderka, D., Ramírez, B. S., Dahlhoff, M., McNeish, I., and Yarden, Y. (2021). Targeting autocrine amphiregulin robustly and reproducibly inhibits ovarian cancer in a syngeneic model: roles for wildtype p53. *Oncogene*, 40:3665–3679.

Liu, Y. (2020). Clinical implications of chromatin accessibility in human cancers. *Oncotarget*, 11:1666–1678.

Lochs, S. J., Kefalopoulou, S., and Kind, J. (2019). Lamina associated domains and gene regulation in development and cancer. *Cells*, 8:271.

Lonsdale, J., Thomas, J., Salvatore, M., Phillips, R., Lo, E., Shad, S., Hasz, R., Walters, G., Garcia, F., Young, N., Foster, B., Moser, M., Karasik, E., Gillard, B., Ramsey, K., Sullivan, S., Bridge, J., Magazine, H., Syron, J., Fleming, J., Siminoff, L., Traino, H., Mosavel, M., Barker, L., Jewell, S., Rohrer, D., Maxim, D., Filkins, D., Harbach, P., Cortadillo, E., Berghuis, B., Turner, L., Hudson, E., Feenstra, K., Sobin, L., Robb, J., Branton, P., Korzeniewski, G., Shive, C., Tabor, D., Qi, L., Groch, K., Nampally, S., Buia, S., Zimmerman, A., Smith, A., Burges, R., Robinson, K., Valentino, K., Bradbury, D., Cosentino, M., Diaz-Mayoral, N., Kennedy, M., Engel, T., Williams, P., Erickson, K., Ardlie, K., Winckler, W., Getz, G., DeLuca, D., MacArthur, D., Kellis, M., Thomson, A., Young, T., Gelfand, E., Donovan, M., Meng, Y., Grant, G., Mash, D., Marcus, Y., Basile, M., Liu, J., Zhu, J., Tu, Z., Cox, N. J., Nicolae, D. L., Gamazon, E. R., Im, H. K., Konkashbaev, A., Pritchard, J., Stevens, M., Flutre, T., Wen, X., Dermitzakis, E. T., Lappalainen, T., Guigo, R., Monlong, J., Sammeth, M., Koller, D., Battle, A., Mostafavi, S., McCarthy, M., Rivas, M., Maller, J., Rusyn, I., Nobel, A., Wright, F., Shabalin, A., Feolo, M., Sharopova, N., Sturcke, A., Paschal, J., Anderson, J. M., Wilder, E. L., Derr, L. K., Green, E. D., Struewing, J. P., Temple, G., Volpi, S., Boyer, J. T., Thomson, E. J., Guyer, M. S., Ng, C., Abdallah, A., Colantuoni, D., Insel, T. R., Koester, S. E., Little, A. R., Bender, P. K., Lehner, T., Yao, Y., Compton, C. C., Vaught, J. B., Sawyer, S., Lockhart, N. C., Demchok, J., and Moore, H. F. (2013). The genotype-tissue expression (gtex) project. *Nature Genetics*, 45:580–585.

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for rna-seq data with deseq2. *Genome Biology*, 15:550.

Luetteke, N. C., Qiu, T. H., Fenton, S. E., Troyer, K. L., Riedel, R. F., Chang, A., and Lee, D. C. (1999). Targeted inactivation of the egf and amphiregulin genes reveals distinct roles for egf receptor ligands in mouse mammary gland development. *Development (Cambridge, England)*, 126:2739–2750.

Ma, S. and Zhang, Y. (2020). Profiling chromatin regulatory landscape: insights into the development of chip-seq and atac-seq. *Molecular Biomedicine*, 1:1–13.

Ma, X. M. and Blenis, J. (2009). Molecular mechanisms of mtor-mediated translational control. *Nature Reviews Molecular Cell Biology 2009 10:5*, 10:307–318.

Macdonald-Obermann, J. L. and Pike, L. J. (2014). Different epidermal growth factor (egf) receptor ligands show distinct kinetics and biased or partial agonism for homodimer and heterodimer formation. *The Journal of Biological Chemistry*, 289:26178.

Manning, B. D. and Cantley, L. C. (2007). Akt/pkb signaling: Navigating downstream. *Cell*, 129:1261.

Markham, A. (2017). Copanlisib: First global approval. *Drugs*, 77:2057–2062.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, 17:10–12.

Matsui, M., Miyasaka, J., Hamada, K., Ogawa, Y., Hiramoto, M., Fujimori, R., and Aioi, A. (2000). Influence of aging and cell senescence on telomerase activity in keratinocytes. *Journal of Dermatological Science*, 22:80–87.

McBryan, J., Howlin, J., Napoletano, S., and Martin, F. (2008). Amphiregulin: Role in mammary gland development and breast cancer. *Journal of Mammary Gland Biology and Neoplasia*, 13:159–169.

McLean, C. Y., Bristor, D., Hiller, M., Clarke, S. L., Schaar, B. T., Lowe, C. B., Wenger, A. M., and Bejerano, G. (2010). Great improves functional interpretation of cis-regulatory regions. *Nature biotechnology*, 28:495–501.

McVeigh, T. P. and Kerin, M. J. (2017). Clinical use of the oncotype dx genomic test to guide treatment decisions for patients with invasive breast cancer. *Breast Cancer : Targets and Therapy*, 9:393.

Meers, M. P., Bryson, T. D., Henikoff, J. G., and Henikoff, S. (2019). Improved cut&run chromatin profiling tools. *eLife*, 8.

Menzel, M., Ossowski, S., Kral, S., Metzger, P., Horak, P., Marienfeld, R., Boerries, M., Wolter, S., Ball, M., Neumann, O., Armeanu-Ebinger, S., Schroeder, C., Matysiak, U., Goldschmid, H., Schipperges, V., Fürstberger, A., Allgäuer, M., Eberhardt, T., Niewöhner, J., Blaumeiser, A., Ploeger, C., Haack, T. B., Tay, T. K. Y., Kelemen, O., Pauli, T., Kirchner, M., Kluck, K., Ott, A., Renner, M., Admard, J., Gschwind, A., Lassmann, S., Kestler, H., Fend, F., Illert, A. L., Werner, M., Möller, P., Seufferlein, T. T. W., Malek, N., Schirmacher, P., Fröhling, S., Kazdal, D., Budczies, J., and Stenzinger, A. (2023). Multicentric pilot study to standardize clinical whole exome sequencing (wes) for cancer patients. *npj Precision Oncology*, 7:1–11.

Mi, H., Muruganujan, A., and Thomas, P. D. (2013). Panther in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. *Nucleic acids research*, 41:377–386.

Miao, B., Skidan, I., Yang, J., Lugovskoy, A., Reibarkh, M., Long, K., Brazell, T., Durugkar, K. A., Maki, J., Ramana, C. V., Schaffhausen, B., Wagner, G., Torchilin, V., Yuan, J., and Degterev, A. (2010). Small molecule inhibition of phosphatidylinositol-3,4,5-triphosphate (pip3) binding to pleckstrin homology domains. *Proceedings of the National Academy of Sciences of the United States of America*, 107:20126–20131.

Miranda, A. X., Kemp, J., Davidson, B., Bellomo, S. E., Agan, V., Manoni, A., Marchiò, C., Croessmann, S., Park, B. H., and Hodges, E. (2024). Genomic dissection and mutation-specific target discovery for breast cancer pik3ca hotspot mutations. *BioRxiv*.

Momeni, K., Ghorbian, S., Ahmadpour, E., and Sharifi, R. (2023). Unraveling the complexity: understanding the deconvolutions of rna-seq data. *Translational Medicine Communications*, 8:1–8.

Morabito, S., Miyoshi, E., Michael, N., Shahin, S., Martini, A. C., Head, E., Silva, J., Leavy, K., Perez-Rosendahl, M., and Swarup, V. (2021). Single-nucleus chromatin accessibility and transcriptomic characterization of alzheimer's disease. *Nature Genetics*, 53:1143–1155.

Nelson, N. C. (2015). A knockout experiment: Disciplinary divides and experimental skill in animal behaviour genetics. *Medical History*, 59:465–485.

Nisticò, P., Bissell, M. J., and Radisky, D. C. (2012). Epithelial-mesenchymal transition: General principles and pathological relevance with special emphasis on the role of matrix metalloproteinases. *Cold Spring Harbor Perspectives in Biology*, 4.

Nitulescu, G. M., Venter, M. V. D., Nitulescu, G., Ungurianu, A., Juzenas, P., Peng, Q., Olaru, O. T., Grădinaru, D., Tsatsakis, A., Tsoukalas, D., Spandidos, D. A., and Margina, D. (2018). The akt pathway in oncology therapy and beyond (review). *International Journal of Oncology*, 53:2319.

Nussinov, R., Zhang, M., Tsai, C. J., and Jang, H. (2021). Phosphorylation and driver mutations in pi3kα and pten autoinhibition. *Molecular Cancer Research*, 19:543–548.

O'Geen, H., Echipare, L., and Farnham, P. J. (2011). Using chip-seq technology to generate high-resolution profiles of histone modifications. *Methods in Molecular Biology (Clifton, N.J.)*, 791:265.

Pacini, C., Dempster, J. M., Boyle, I., Gonçalves, E., Najgebauer, H., Karakoc, E., van der Meer, D., Barthorpe, A., Lightfoot, H., Jaaks, P., McFarland, J. M., Garnett, M. J., Tsherniak, A., and Iorio, F. (2021). Integrated cross-study datasets of genetic dependencies in cancer. *Nature Communications*, 12:1–14.

Parker, J. S., Mullins, M., Cheang, M. C. U., Leung, S., Voduc, D., Vickery, T., Davies, S., Fauron, C., He, X., Hu, Z., Quackenbush, J. F., Stijleman, I. J., Palazzo, J., Marron, J. S., Nobel, A. B., Mardis, E., Nielsen, T. O., Ellis, M. J., Perou, C. M., and Bernard, P. S. (2009). Supervised risk predictor of breast cancer based on intrinsic subtypes. *Journal of Clinical Oncology*, 27:1160–1167.

Perou, C. M., Sørile, T., Eisen, M. B., Rijn, M. V. D., Jeffrey, S. S., Ress, C. A., Pollack, J. R., Ross, D. T., Johnsen, H., Akslen, L. A., Øysteln Fluge, Pergammenschlkov, A., Williams, C., Zhu, S. X., Lønning, P. E., Børresen-Dale, A. L., Brown, P. O., and Botstein, D. (2000). Molecular portraits of human breast tumours. *Nature*, 406:747–752.

Piccart, M., Hennequin, A., Borrego, M. R., de Romani, S. E., Williams, A., Rodríguez, B. J., Conte, G. D., Howell, S. J., Palleschi, M., Simonelli, M., Duhoux, F. P., Tosi, D., de Speville Uribe, B. D., Gilarranz, Y. J., Tassone, P., Curigliano, G., Waters, S., Aftimos, P., Wildiers, H., Scartoni, S., Vallespir, B. P., Shankaraiah, R. C., Grzegorzewski, K., and Habboubi, N. (2023). Abstract pd18-05: Men1611, a pi3k inhibitor, combined with trastuzumab ± fulvestrant for her2+/pik3ca mutant advanced or metastatic breast cancer: updated safety and efficacy results from the ongoing phase 1b study (b-precise-01). *Cancer Research*, 83:PD18–05.

Pommier, Y., Sordet, O., Antony, S., Hayward, R. L., and Kohn, K. W. (2004). Apoptosis defects and chemotherapy resistance: molecular interaction maps and networks. *Oncogene*, 23:2934–2949.

Pope, B. D., Ryba, T., Dileep, V., Yue, F., Wu, W., Denas, O., Vera, D. L., Wang, Y., Hansen, R. S., Canfield, T. K., Thurman, R. E., Cheng, Y., Gülsoy, G., Dennis, J. H., Snyder, M. P., Stamatoyannopoulos, J. A., Taylor, J., Hardison, R. C., Kahveci, T., Ren, B., and Gilbert, D. M. (2014). Topologically associating domains are stable units of replication-timing regulation. *Nature*, 515:402–405.

Poveda, A., Floquet, A., Ledermann, J. A., Asher, R., Penson, R. T., Oza, A. M., Korach, J., Huzarski, T., Pignata, S., Friedlander, M., Baldoni, A., Park-Simon, T. W., Tamura, K., Sonke, G. S., Lisyanskaya, A., Kim, J. H., Filho, E. A., Milenkova, T., Lowe, E. S., Rowe, P., Vergote, I., Pujade-Lauraine, E., Byrski, T., Pautier, P., Harter, P., Colombo, N., Scambia, G., Nicoletto, M., Nussey, F., Clamp, A., Penson, R., Velasco, A. P., Rodrigues, M., Lotz, J. P., Selle, F., Ray-Coquard, I., Provencher, D., Aparicio, A. P., Boixader, L. V., Scott, C., Yunokawa, M., Medioni, J., Pécuchet, N., Dubot, C., Rouge, T. D. L. M., Kaminsky, M. C., Weber, B., Lortholary, A., Parkinson, C., Ledermann, J., Williams, S., Banerjee, S., Cosin, J., Hoffman, J., Plante, M., Covens, A., Sonke, G., Joly, F., Hirte, H., Amit, A., Matsumoto, K., Tjulandin, S., Kim, J. H., Gladieff, L., Sabbatini, R., O'Malley, D., Timmins, P., Kredentser, D., Milagro, N. L., Ginesta, M. P. B., Martorell, A. T., Lista, A. G. D. L., González, B. O., Mileshkin, L., Mandai, M., Boere, I., Ottevanger, P., Nam, J. H., Filho, E., Hamizi, S., Cognetti, F., Warshal, D., Dickson-Michelson, E., Kamelle, S., McKenzie, N., Rodriguez, G., Armstrong, D., Chalas, E., Celano, P., Behbakht, K., Davidson, S., Welch, S., Helpman, L., Fishman, A., Bruchim, I., Sikorska, M., Słowińska, A., Rogowski, W., Bidziński, M., Śpiewankiewicz, B., Herraez, A. C., Fernández, C. M., Gropp-Meier, M., Saito, T., Takehara, K., Enomoto, T., Watari, H., Choi, C. H., Kim, B. G., Kim, J. W., and Hegg, R. (2021). Olaparib tablets as maintenance therapy in patients with platinum-sensitive relapsed ovarian cancer and a brca1/2 mutation (solo2/engot-ov21): a final analysis of a double-blind, randomised, placebo-controlled, phase 3 trial. *The Lancet Oncology*, 22:620–631.

Prakash, R., Zhang, Y., Feng, W., and Jasin, M. (2015). Homologous recombination and human health: The roles of brca1, brca2, and associated proteins. *Cold Spring Harbor Perspectives in Biology*, 7:a016600.

Puca, L., Dowless, M., Perez, C., Ortiz-Ruiz, M. J., Donoho, G., Capen, A., Huber, L., Bogner, S., Fei, D., Manro, J. R., Yu, C. P., Xu, W. G., Wang, R., Chen, S., Hicks, M., Zolfaghari, P., Faber, A., Gilmour, R., Ramstetter, M., Chang, M., Lallena, M. J., Gong, X., Hyman, D. M., Smyth, L. M., Brandhuber, B. J., Taylor, B. S., and Klippel, A. (2022). Loxo-783: A potent, highly mutant selective and brain-penetrant allosteric pi3kα h1047r inhibitor in combination with standard of care (soc) treatments in preclinical pi3kα h1047r-mutant breast cancer models. *San Antonio Breast Cancer Symposium; Poster presentation of unpublished data*.

Puleo, J. and Polyak, K. (2011). The mcf10 model of breast tumor progression. *Cancer Research*, 81:4183–4185.

Qin, T., Lee, C., Li, S., Cavalcante, R. G., Orchard, P., Yao, H., Zhang, H., Wang, S., Patil, S., Boyle, A. P., and Sartor, M. A. (2022). Comprehensive enhancer-target gene assignments improve gene set level interpretation of genome-wide regulatory data. *Genome Biology*, 23:1–30.

Qin, X., Jiang, B., and Zhang, Y. (2016). 4e-bp1, a multifactor regulated multifunctional protein. *Cell Cycle*, 15:781.

Saal, L. H., Holm, K., Maurer, M., Memeo, L., Su, T., Wang, X., Yu, J. S., Malmström, P. O., Mansukhani, M., Enoksson, J., Hibshoosh, H., Åke Borg, and Parsons, R. (2005). Pik3ca mutations correlate with hormone receptors, node metastasis, and erbb2, and are mutually exclusive with pten loss in human breast carcinoma. *Cancer Research*, 65:2554–2559.

Samuels, Y., Wang, Z., Bardelli, A., Silliman, N., Ptak, J., Szabo, S., Yan, H., Gazdar, A., Powell, S. M., Riggins, G. J., Willson, J. K., Markowitz, S., Kinzler, K. W., Vogelstein, B., and Velculescu, V. E. (2004). High frequency of mutations of the pik3ca gene in human cancers. *Science (New York, N.Y.)*, 304(5670):554.

Sanjana, N. E., Shalem, O., and Zhang, F. (2014). Improved vectors and genome-wide libraries for crispr screening. *Nature methods*, 11:783–784.

Sanson, K. R., Hanna, R. E., Hegde, M., Donovan, K. F., Strand, C., Sullender, M. E., Vaimberg, E. W., Goodale, A., Root, D. E., Piccioni, F., and Doench, J. G. (2018). Optimized libraries for crispr-cas9 genetic screens with multiple modalities. *Nature Communications*, 9:1–15.

Seefried, F., Haller, L., Fukuda, S., Thongmao, A., Schneider, N., Utikal, J., Higashiyama, S., Bosserhoff, A. K., and Kuphal, S. (2022). Nuclear areg affects a low-proliferative phenotype and contributes to drug resistance of melanoma. *International Journal of Cancer*, 151:2244–2264.

Setton, J., Zinda, M., Riaz, N., Durocher, D., Zimmermann, M., Koehler, M., Reis-Filho, J. S., and Powell, S. N. (2021). Synthetic lethality in cancer therapeutics: The next generation. *Cancer Discovery*, 11:1626–1635.

Shalem, O., Sanjana, N. E., Hartenian, E., Shi, X., Scott, D. A., Mikkelsen, T. S., Heckl, D., Ebert, B. L., Root, D. E., Doench, J. G., and Zhang, F. (2014). Genome-scale crispr-cas9 knockout screening in human cells. *Science (American Association for the Advancement of Science)*, 343:84–87.

Shaulian, E. and Karin, M. (2001). Ap-1 in cell proliferation and survival. *Oncogene*, 20:2390–2400.

Shekar, S. C., Wu, H., Fu, Z., Yip, S. C., Nagajyothi, Cahill, S. M., Girvin, M. E., and Backer, J. M. (2005). Mechanism of constitutive phosphoinositide 3-kinase activation by oncogenic mutants of the p85 regulatory subunit. *Journal of Biological Chemistry*, 280(30):27850–27855.

Shoyab, M., McDonald, V. L., Bradley, J. G., and Todaro, G. J. (1988). Amphiregulin: A bifunctional growth-modulating glycoprotein produced by the phorbol 12-myristate 13-acetate-treated human breast adenocarcinoma cell line mcf-7. *Proceedings of the National Academy of Sciences of the United States of America*, 85:6528–6532.

Siegel, R. L., Miller, K. D., Sandeep, N., Mbbs, W., Ahmedin, ., Dvm, J., and Siegel, R. L. (2023). Cancer statistics, 2023. *CA: A Cancer Journal for Clinicians*, 73:17–48.

Silva, A. J., Paylor, R., Wehner, J. M., and Tonegawa, S. (1992). Impaired spatial learning in alpha-calcium-calmodulin kinase ii mutant mice. *Science*, 257:206–212.

Silvy, M., Giusti, C., Martin, P. M., and Berthois, Y. (2001). Differential regulation of cell proliferation and protease secretion by epidermal growth factor and amphiregulin in tumoral versus normal breast epithelial cells. *British Journal of Cancer*, 84:936.

Sirico, M., D'Angelo, A., Gianni, C., Casadei, C., Merloni, F., and Giorgi, U. D. (2023). Current state and future challenges for pi3k inhibitors in cancer therapy. *Cancers*, 15.

Song, K. W., Edgar, K. A., Hanan, E. J., Hafner, M., Oeh, J., Merchant, M., Sampath, D., Nannini, M. A., Hong, R., Phu, L., Forrest, W. F., Stawiski, E., Schmidt, S., Endres, N., Guan, J., Wallin, J. J., Cheong, J., Plise, E. G., Phillips, G. D. L., Salphati, L., Heffron, T. P., Olivero, A. G., Malek, S., Staben, S. T., Kirkpatrick, D. S., Dey, A., and Friedman, L. S. (2022). Rtk-dependent inducible degradation of mutant pi3kα drives gdc-0077 (inavolisib) efficacy. *Cancer Discovery*, 12:204.

Soule, H. D., Maloney, T. M., Wolman, S. R., Peterson, Ward D., J., Brenz, R., McGrath, C. M., Russo, J., Pauley, R. J., Jones, R. F., and Brooks, S. C. (1990). Isolation and characterization of a spontaneously immortalized human breast epithelial cell line, mcf-101. *Cancer Research*, 50(18):6075–6086.

Sridharan, S. and Basu, A. (2020). Distinct roles of mtor targets s6k1 and s6k2 in breast cancer. *International Journal of Molecular Sciences*, 21:1199.

Sun, P., Zhang, X., Wang, R.-J., Ma, Q.-Y., Xu, L., Wang, Y., Liao, H.-P., Wang, H.-L., Hu, L.-D., Kong, X., Ding, J., and Meng, L.-H. (2021). Pi3kα inhibitor cyh33 triggers antitumor immunity in murine breast cancer by activating cd8 + t cells and promoting fatty acid metabolism. *Journal of Immunotherapy in Cancer*, 9:3093.

Szwed, A., Kim, E., and Jacinto, E. (2021). Regulation and metabolic functions of mtorc1 and mtorc2. *Physiological Reviews*, 101:1371.

Tanigawa, Y., Dyer, E. S., and Bejerano, G. (2022). Whichtf is functionally important in your open chromatin data? *PLoS Computational Piology*, 18.

Taniguchi, C. M., Winnay, J., Kondo, T., Bronson, R. T., Guimaraes, A. R., Alemán, J. O., Luo, J., Stephanopoulos, G., Weissleder, R., Cantley, L. C., and Kahn, C. R. (2010). The phosphoinositide 3-kinase regulatory subunit p85alpha can exert tumor suppressor properties through negative regulation of growth factor signaling. *Cancer Research*, 70(13):5305–5315.

Thennavan, A., Beca, F., Xia, Y., Garcia-Recio, S., Allison, K., Collins, L. C., Tse, G. M., Chen, Y. Y., Schnitt, S. J., Hoadley, K. A., Beck, A., and Perou, C. M. (2021). Molecular analysis of tcga breast cancer histologic types. *Cell Genomics*, 1.

Thorpe, L. M., Yuzugullu, H., and Zhao, J. J. (2015). PI3K in cancer: divergent roles of isoforms, modes of activation, and therapeutic targeting. *Nature reviews. Cancer*, 15(1):7.

Tian, T., Li, X., and Zhang, J. (2019). mtor signaling in cancer and mtor inhibitors in solid tumor targeting therapy. *International Journal of Molecular Sciences*, 20.

Tortajada-Genaro, L. A., Lazaro, A., Martorell, S., and Maquieira, A. (2023). Nucleotide-selective amplification and array-based detection for identifying multiple somatic mutations. *Analytica Chimica Acta*, 1265:341343.

Turner, N. C., Oliveira, M., Howell, S. J., Dalenc, F., Cortes, J., Moreno, H. L. G., Hu, X., Jhaveri, K., Krivorotko, P., Loibl, S., Murillo, S. M., Okera, M., Park, Y. H., Sohn, J., Toi, M., Tokunaga, E., Yousef, S., Zhukova, L., de Bruin, E. C., Grinsted, L., Schiavon, G., Foxley, A., and Rugo, H. S. (2023). Capivasertib in hormone receptor–positive advanced breast cancer. *New England Journal of Medicine*, 388:2058–2070.

Ueki, K., Fruman, D. A., Brachmann, S. M., Tseng, Y.-H., Cantley, L. C., and Kahn, C. R. (2002). Molecular balance between the regulatory and catalytic subunits of phosphoinositide 3-kinase regulates cell signaling and survival. *Molecular and Cellular Biology*, 22(3):965.

Urick, M. E., Rudd, M. L., Godwin, A. K., Sgroi, D., Merino, M., and Bell, D. W. (2011). Pik3r1 (p85α) is somatically mutated at high frequency in primary endometrial cancer. *Cancer research*, 71(12):4061–4067.

Valancius, V. and Smithies, O. (1991). Testing an "in-out" targeting procedure for making subtle genomic modifications in mouse embryonic stem cells. *Molecular and Cellular Biology*, 11:1402.

Valton, A. L. and Dekker, J. (2016). Tad disruption as oncogenic driver. *Current Opinion in Genetics and Development*, 36:34–40.

Varkaris, A., Pazolli, E., Gunaydin, H., Wang, Q., Pierce, L., Boezio, A. A., Bulku, A., DiPietro, L., Fridrich, C., Frost, A., Giordanetto, F., Hamilton, E. P., Harris, K., Holliday, M., Hunter, T. L., Iskandar, A., Ji, Y., Larivée, A., LaRochelle, J. R., Lescarbeau, A., Llambi, F., Lormil, B., Mader, M. M., Mar, B. G., Martin, I., McLean, T. H., Michelsen, K., Pechersky, Y., Puente-Poushnejad, E., Raynor, K., Rogala, D., Samadani, R., Schram, A. M., Shortsleeves, K., Swaminathan, S., Tajmir, S., Tan, G., Tang, Y., Valverde, R., Wehrenberg, B., Wilbur, J., Williams, B. R., Zeng, H., Zhang, H., Walters, W. P., Wolf, B. B., Shaw, D. E., Bergstrom, D. A., Watters, J., Fraser, J. S., Fortin, P. D., and Kipp, D. R. (2024). Discovery and clinical proof-of-concept of rly-2608, a first-in-class mutant-selective allosteric pi3kα inhibitor that decouples antitumor activity from hyperinsulinemia. *Cancer Discovery*, 14:240–257.

Vierbuchen, T., Ling, E., Cowley, C. J., Couch, C. H., Wang, X., Harmin, D. A., Roberts, C. W., and Greenberg, M. E. (2017). Tead4 functions as a prognostic biomarker and triggers emt via pi3k/akt pathway in bladder cancer. *Molecular Cell*, 41:1–20.

Vuylsteke, P., Huizing, M., Petrakova, K., Roylance, R., Laing, R., Chan, S., Abell, F., Gendreau, S., Rooney, I., Apt, D., Zhou, J., Singel, S., and Fehrenbacher, L. (2016). Pictilisib pi3kinase inhibitor (a phosphatidylinositol 3-kinase [pi3k] inhibitor) plus paclitaxel for the treatment of hormone receptor-positive, her2-negative, locally recurrent, or metastatic breast cancer: Interim analysis of the multicentre, placebo-controlled, phase ii randomised peggy study. *Annals of Oncology*, 27:2059–2066.

Waarts, M. R., Stonestrom, A. J., Park, Y. C., and Levine, R. L. (2022). Targeting mutations in cancer. *The Journal of Clinical Investigation*, 132.

Wang, Q., Li, M., Wu, T., Zhan, L., Li, L., Chen, M., Xie, W., Xie, Z., Hu, E., Xu, S., and Yu, G. (2022). Exploring epigenomic datasets by chipseeker. *Current Protocols*, 2.

Wang, Y., Liu, J., Jiang, Q., Deng, J., Xu, F., Chen, X., Cheng, F., Zhang, Y., Yao, Y., Xia, Z., Xu, X., Su, X., Huang, M., Dai, L., Yang, Y., Zhang, S., Yu, D., Zhao, R. C., Wei, Y., and Deng, H. (2017). Human adipose-derived mesenchymal stem cell-secreted cxcl1 and cxcl8 facilitate breast tumor growth by promoting angiogenesis. *Stem Cells*, 35:2060–2070.

Wang, Y., Zhou, Y., and Graves, D. T. (2014). Foxo transcription factors: Their clinical significance and regulation. *BioMed Research International*, 2014.

Wheeler, D. A. and Wang, L. (2013). From human genome to cancer genome: The first decade. *Genome Research*, 23:1054–1062.

Willmarth, N. E. and Ethier, S. P. (2006). Autocrine and juxtacrine effects of amphiregulin on the proliferative, invasive, and migratory properties of normal and neoplastic human mammary epithelial cells *. *Journal of Biological Chemistry*, 281:37728–37737.

Wilson, K. J., Gilmore, J. L., Foley, J., Lemmon, M. A., and Riese, D. J. (2009). Functional selectivity of egf family peptide growth factors: Implications for cancer. *Pharmacology and Therapeutics*, 122:1–8.

Wood, L. D., Parsons, D. W., Jones, S., Lin, J., Sjöblom, T., Leary, R. J., Shen, D., Boca, S. M., Barber, T., Ptak, J., Silliman, N., Szabo, S., Dezso, Z., Ustyanksky, V., Nikolskaya, T., Nikolsky, Y., Karchin, R., Wilson, P. A., Kaminker, J. S., Zhang, Z., Croshaw, R., Willis, J., Dawson, D., Shipitsin, M., Willson, J. K., Sukumar, S., Polyak, K., Ben, H. P., Pethiyagoda, C. L., Pant, P. V., Ballinger, D. G., Sparks, A. B., Hartigan, J., Smith, D. R., Suh, E., Papadopoulos, N., Buckhaults, P., Markowitz, S. D., Parmigiani, G., Kinzler, K. W., Velculescu, V. E., and Vogelstein, B. (2007). The genomic landscapes of human breast and colorectal cancers. *Science*, 318:1108–1113.

Wu, X., Renuse, S., Sahasrabuddhe, N. A., Zahari, M. S., Chaerkady, R., Kim, M. S., Nirujogi, R. S., Mohseni, M., Kumar, P., Raju, R., Zhong, J., Yang, J., Neiswinger, J., Jeong, J. S., Newman, R., Powers, M. A., Somani, B. L., Gabrielson, E., Sukumar, S., Stearns, V., Qian, J., Zhu, H., Vogelstein, B., Park, B. H., and Pandey, A. (2014). Activation of diverse signalling pathways by oncogenic pik3ca mutations. *Nature Communications*, 5.

Xie, Y., Yang, L., Wu, Y., Zheng, H., and Gou, Q. (2022). Adjuvant endocrine therapy in patients with estrogen receptor-low positive breast cancer: A prospective cohort study. *The Breast : Official Journal of the European Society of Mastology*, 66:89.

Yang, J., Nie, J., Ma, X., Wei, Y., Peng, Y., and Wei, X. (2019). Targeting pi3k in cancer: mechanisms and advances in clinical trials. *Molecular Cancer*, 18:1–28.

Zhan, T., Rindtorff, N., and Boutros, M. (2016). Wnt signaling in cancer. *Oncogene*, 36:1461–1473.

Zhang, F., Wen, Y., and Guo, X. (2014). Crispr/cas9 for genome editing: progress, implications and challenges. *Human Molecular Genetics*, 23:R40–R46.

Zhang, H., Ahearn, T. U., Lecarpentier, J., Barnes, D., Beesley, J., Qi, G., Jiang, X., O'Mara, T. A., Zhao, N., Bolla, M. K., Dunning, A. M., Dennis, J., Wang, Q., Ful, Z. A., Aittomäki, K., Andrulis, I. L., Anton-Culver, H., Arndt, V., Aronson, K. J., Arun, B. K., Auer, P. L., Azzollini, J., Barrowdale, D., Becher, H., Beckmann, M. W., Behrens, S., Benitez, J., Bermisheva, M., Bialkowska, K., Blanco, A., Blomqvist, C., Bogdanova, N. V., Bojesen, S. E., Bonanni, B., Bondavalli, D., Borg, A., Brauch, H., Brenner, H., Briceno, I., Broeks, A., Brucker, S. Y., Brüning, T., Burwinkel, B., Buys, S. S., Byers, H., Caldés, T., Caligo, M. A., Calvello, M., Campa, D., Castelao, J. E., Chang-Claude, J., Chanock, S. J., Christiaens, M., Christiansen, H., Chung, W. K., Claes, K. B., Clarke, C. L., Cornelissen, S., Couch, F. J., Cox, A., Cross, S. S., Czene, K., Daly, M. B., Devilee, P., Diez, O., Domchek, S. M., Dörk, T., Dwek, M., Eccles, D. M., Ekici,

A. B., Evans, D. G., Fasching, P. A., Figueroa, J., Foretova, L., Fostira, F., Friedman, E., Frost, D., Gago-Dominguez, M., Gapstur, S. M., Garber, J., García-Sáenz, J. A., Gaudet, M. M., Gayther, S. A., Giles, G. G., Godwin, A. K., Goldberg, M. S., Goldgar, D. E., González-Neira, A., Greene, M. H., Gronwald, J., Guénel, P., Häberle, L., Hahnen, E., Haiman, C. A., Hake, C. R., Hall, P., Hamann, U., Harkness, E. F., Heemskerk-Gerritsen, B. A., Hillemanns, P., Hogervorst, F. B., Holleczek, B., Hollestelle, A., Hooning, M. J., Hoover, R. N., Hopper, J. L., Howell, A., Huebner, H., Hulick, P. J., Imyanitov, E. N., Isaacs, C., Izatt, L., Jager, A., Jakimovska, M., Jakubowska, A., James, P., Janavicius, R., Janni, W., John, E. M., Jones, M. E., Jung, A., Kaaks, R., Kapoor, P. M., Karlan, B. Y., Keeman, R., Khan, S., Khusnutdinova, E., Kitahara, C. M., Ko, Y. D., Konstantopoulou, I., Koppert, L. B., Koutros, S., Kristensen, V. N., Laenkholm, A. V., Lambrechts, D., Larsson, S. C., Laurent-Puig, P., Lazaro, C., Lazarova, E., Lejbkowicz, F., Leslie, G., Lesueur, F., Lindblom, A., Lissowska, J., Lo, W. Y., Loud, J. T., Lubinski, J., Lukomska, A., MacInnis, R. J., Mannermaa, A., Manoochehri, M., Manoukian, S., Margolin, S., Martinez, M. E., Matricardi, L., McGuffog, L., McLean, C., Mebirouk, N., Meindl, A., Menon, U., Miller, A., Mingazheva, E., Montagna, M., Mulligan, A. M., Mulot, C., Muranen, T. A., Nathanson, K. L., Neuhausen, S. L., Nevanlinna, H., Neven, P., Newman, W. G., Nielsen, F. C., Nikitina-Zake, L., Nodora, J., Offit, K., Olah, E., Olopade, O. I., Olsson, H., Orr, N., Papi, L., Papp, J., Park-Simon, T. W., Parsons, M. T., Peissel, B., Peixoto, A., Peshkin, B., Peterlongo, P., Peto, J., Phillips, K. A., Piedmonte, M., Plaseska-Karanfilska, D., Prajzendanc, K., Prentice, R., Prokofyeva, D., Rack, B., Radice, P., Ramus, S. J., Rantala, J., Rashid, M. U., Rennert, G., Rennert, H. S., Risch, H. A., Romero, A., Rookus, M. A., Rübner, M., Rüdiger, T., Saloustros, E., Sampson, S., Sandler, D. P., Sawyer, E. J., Scheuner, M. T., Schmutzler, R. K., Schneeweiss, A., Schoemaker, M. J., Schöttker, B., Schürmann, P., Senter, L., Sharma, P., Sherman, M. E., Shu, X. O., Singer, C. F., Smichkoska, S., Soucy, P., Southey, M. C., Spinelli, J. J., Stone, J., Stoppa-Lyonnet, D., Swerdlow, A. J., Szabo, C. I., Tamimi, R. M., Tapper, W. J., Taylor, J. A., Teixeira, M. R., Terry, M. B., Thomassen, M., Thull, D. L., Tischkowitz, M., Toland, A. E., Tollenaar, R. A., Tomlinson, I., Torres, D., Troester, M. A., Truong, T., Tung, N., Untch, M., Vachon, C. M., van den Ouweland, A. M., van der Kolk, L. E., van Veen, E. M., vanRensburg, E. J., Vega, A., Wappenschmidt, B., Weinberg, C. R., Weitzel, J. N., Wildiers, H., Winqvist, R., Wolk, A., Yang, X. R., Yannoukakos, D., Zheng, W., Zorn, K. K., Milne, R. L., Kraft, P., Simard, J., Pharoah, P. D., Michailidou, K., Antoniou, A. C., Schmidt, M. K., Chenevix-Trench, G., Easton, D. F., Chatterjee, N., and García-Closas, M. (2020). Genome-wide association study identifies 32 novel breast cancer susceptibility loci from overall and subtype-specific analyses. *Nature Genetics*, 52:572–581.

Zhang, J., Baddoo, M., Han, C., Strong, M. J., Cvitanovic, J., Moroz, K., Dash, S., Flemington, E. K., Wu, T., Zhang, J., Baddoo, M., Han, C., Strong, M. J., Cvitanovic, J., Moroz, K., Dash, S., Flemington, E. K., and Wu, T. (2016). Gene network analysis reveals a novel 22-gene signature of carbon metabolism in hepatocellular carcinoma. *Oncotarget*, 7:49232–49245.

Zhou, B., Lin, W., Long, Y., Yang, Y., Zhang, H., Wu, K., and Chu, Q. (2022). Notch

signaling pathway: architecture, disease, and therapeutics. *Signal Transduction and Targeted Therapy*, 7:95.