

INVESTIGATING THE PROFILE OF ABERRANTLY EXPRESSED MEIOTIC
GENES IN ALT POSITIVE AND NEGATIVE CANCER CELL LINES

by
Zach Perdun

A thesis submitted to Johns Hopkins University in conformity with the requirements for the
degree of Master of Science

Baltimore, Maryland

June 2017

Abstract

Meiosis is a unique form of cellular division that requires the use of genes normally not expressed in other tissues. Recently, it has been discovered that many of these meiotic genes are being expressed in cancer cell lines. Some of the cell lines identified follow the Alternative Lengthening of Telomeres (ALT) pathway as a means to extend their telomeres in the absence of telomerase, successfully overcoming replicative mortality. A group of meiotic genes identified play a role in mechanisms such as homologous recombination and DNA damage repair – processes essential for the ALT pathway. In an effort to better understand which meiotic genes are necessary and unique to the ALT mechanism, we surveyed ALT positive and ALT negative cancer cell lines for specific gene expression. After identifying genes of interest, we introduced the Auxin Induced Degradation (AID) system into the ALT positive U2OS pediatric osteosarcoma cell line as a method to conditionally deplete candidate proteins of interest during specific stages of the cell cycle.

Acknowledgements

I would like to acknowledge the department of Biochemistry and Molecular Biology. My gratitude extends to the outstanding professors and researchers that make this department what it is. I would especially like to thank Grace Hwang for her instruction and guidance in the lab, as well as Dr. Marina Pryzhkova for her instruction in various lab techniques, especially cell culture, PCR primer development, and CRISPR transfection. These two individuals have instilled in me a lab skillset that will aid me in my future pursuits as a researcher, and I would not have the experience I do without them.

Lastly, I would like to thank Dr. Philip Jordan for his mentorship and guidance during my two years here at The Johns Hopkins Bloomberg School of Public Health. He has had a lasting positive impact on my future in research and medicine, and I would not be the person I am today without him.

Table of Contents

Introduction	1-18
Replicative Immortality and Telomere Function in Cancer	1-2
Misexpression of Meiotic Genes in Cancer	2-5
DNA Damage Responses and Cancer	5-7
The ALT Phenotype and Homologous Recombination	7-11
Meiotic misexpression in the ALT pathway	11-12
Auxin Induced Degradation (AID)	12-15
Materials and Methods	19-36
Cell Lines for Protein Isolates	19
Cell Culture Conditions	20
Protein Isolation and quantification	20-23
Western Blot Analysis	23-25
RT-PCR and QPCR Procedure	26
UCSC Cancer Browser Screening	26-27
Plasmids used in transfection	27-31
Transfection of the AID system	31-32
Single Cell Cloning and Selection	32-36
Results	36-65
RT-PCR and Q-PCR DATA	36-47
Western Blot Analysis	47-57
Analysis of successful transfection of AID system	58-63
UCSC Cancer Browser Results	64
Discussion	65-68
Future Direction	68-71
References	72-77

List of Figures

Figure 1: The Shelterin Complex	15
Figure 2: Conservative DNA synthesis during BIR	16
Figure 3: Auxin Induced Degradation (AID) system	17
Figure 4: Mechanism of cationic-lipid mediated delivery	18
Figure 5: Serial dilution protocol used for protein gradient experiment	22
Figure 6: Results for protein gradient experiment	23
Figure 7: Donor construct for insertion of osTIR1 into the AAVS1 safe harbor locus	30
Figure 8: Donor construct for insertion of SMC5-mAID	30
Figure 9: Donor construct for insertion of SMC6-mAID	31
Figure 10: RT-PCR data for a group of meiotic genes corresponding to Table 5	41
Figure 11: RT-PCR data for a group of meiotic genes corresponding to Table 6	42
Figure 12: RT-PCR results for the osteosarcoma cell lines U2OS and MG63	43
Figure 13: RT-PCR results for the osteosarcoma cell lines SAOS2 and SJSA1	44
Figure 14: Q-PCR data for a group of meiotic genes surveyed for in U2OS and SAOS2	45
Figure 15: Q-PCR data for a group of meiotic genes surveyed for in SJSA1 and MG63	46
Figure 16: Western blot analysis of DMC1 (1)	49
Figure 17: Western blot analysis of DMC1 (2)	50
Figure 18: Western blot analysis of HFM1 (1)	50
Figure 19: Western blot analysis of HFM1 (2)	51
Figure 20: Western blot analysis of HFM1 (3)	51
Figure 21: Western blot analysis of HOP2	52
Figure 22: Western blot analysis of HORMAD2 (1)	52
Figure 23: Western blot analysis of HORMAD2 (2)	53
Figure 24: Western blot analysis of NSE2	53
Figure 25: Western blot analysis of NSE4a	54
Figure 26: Western blot analysis of REC8	54
Figure 27: Western blot analysis of SMC5	55
Figure 28: Western blot analysis of SMC6 (1)	55
Figure 29: Western blot analysis of SMC6 (2)	56
Figure 30: Western blot analysis of SYCE2 (1)	56
Figure 31: Western blot analysis of SYCE2 (2)	57
Figure 32: Confirming osTIR1 transfection in the U2OS Het. Pop. and single cell clones	61
Figure 33: Confirming osTIR1 transfection in the SAOS1 Het. Pop.	62
Figure 34: Assessing SMC5-mAID transfection in U2OS-osTIR1(clone 3)	62
Figure 35: Assessing 3 variants of SMC6-mAID transfection in U2OS-osTIR1(clone 3)	63
Figure 36: Meiotic gene expression stratified by ATRX expression in lower grade glioma	64

List of Tables

Table 1: Cell lines used in protein analysis	19
Table 2: List of antibodies used in western blot procedure and analysis via ECL	25
Table 3: Meiotic genes of interest used in QPCR analysis	26
Table 4: Meiotic Genes of Interest used in UCSC Cancer Browser screening	27
Table 5: Primer sequence and expected product band size for Figure 10	39-40
Table 6: Primer sequence and expected product band size for Figure 11	41
Table 7: Meiotic genes of interest used in RT-PCR and Q-PCR analyses	42
Table 8: Genes not expressed in 2 ALT (+), 2 ALT (-) or any of the 4 cell lines assessed	47
Table 9: Genes assayed for expression in western blots	49

List of Abbreviations

Structures, Processes: AAVS1, Adeno-Associated Virus Integration Site 1; AiCNA, Allelic-imbalanced Copy Number Aberrations; AID, Auxin Induced Degradation; mAID, Mini-AID; ALT, Alternative Lengthening of Telomeres; APBs, ALT-Associated Promyelocytic Leukaemia Nuclear Bodies; CRISPR, Clustered regularly interspaced short palindromic repeats; CMV, Cytomegalovirus; Ct, Average Threshold Cycle; CTL, Cytotoxic T-Lymphocyte Response; DSB, Double-Stranded Break; cDNA, Complementary DNA; LUAD, Lung Adenocarcinoma; MMR, Mismatch Repair; NER, Nucleotide Excision Repair; NHEJ, Non-Homologous End-Joining; PCR, Polymerase Chain Reaction; PVDF, Polyvinylidene Difluoride; Q-PCR, Quantitative RT-PCR; RT-PCR, Reverse Transcription PCR; gRNA, Guide RNA; SSB, Single-Stranded Break; SUMO, Small Ubiquitin-Like Modifier; TNBSs, Triple-Negative Breast Cancers; UCSC, University of California Santa Cruz; WT, Wild Type

Solutions: BCA, Bicinchoninic Acid; BSA, Bovine Serum Albumin; BSA-T, BSA in PBS-T; CGM, Complete Growth Media; DMEM, Dulbecco's Modified Eagle Medium; ECL, Enhanced Chemi-Luminescence; EDTA, Ethylenediaminetetraacetic Acid; FBS, Fetal Bovine Serum; G418, Gentamicin; IAA, Indole-3-Acetic Acid; MMS, methanesulfonate, PBS, Phosphate Buffered Saline; PBS-T, PBS+0.2% tween 20; RIPA, Radioimmunoprecipitation

Genes/Proteins: ATM, Ataxia-Telangiectasia Mutated; ATR, Ataxia Telangiectasia Related; ATRX, ATP-dependent helicase ATRX; BAGE, B-Melonama Antigen; BRCA2, breast cancer protein 2; CCNB1IP1, Cyclin B1 Interacting Protein 1; CCR5, Chemokine Receptor 5; CHK2, Checkpoint kinase 2; COUP-TF2, COUP transcription factor 2; DAXX, Death-associated protein

6; DMC1, Disrupted Meiotic cDNA 1; eSpCAS9, Enhanced Specificity *Streptococcus pyogenes* Cas9; FKBP6, FK506 Binding Protein 6; GAGE, G-Melonama Antigen; H2AX, Histone 2AX; γ H2AX, Phosphorylated H2AX; HFM1, Helicase Family Member 1; HOP2, Homologous-Pairing Protein 2; HORMAD1, HORMA Domain Protein 1; HORMAD2, HORMA Domain Protein 2; HTERT, Human Telomerase Reverse Transcriptase; MAGE-1, Melanoma Antigen-1; MAJIN, Membrane Anchored Junction Protein; MDC1, mediator of DNA damage checkpoint 1; MDM2, mouse double minute 2 homolog; MEI1, Meiotic Double-Stranded Break Formation Protein 1; MEIKIN, Meiotic kinetochore Factor; MEIOB, Meiosis Specific With OB Domains; MMS21, methyl-methanesulfonate protein 21; MND1, Meiotic Nuclear Divisions 1; MRE11, meiotic recombination 11; MRN complex, MRE11-RAD50-NBS1 complex; MSH4, MutS Homolog 4; MSH5, MutS Homolog 5; MSN1, Multicopy suppressor of SNF1 mutation; NBS1, Nijmegen breakage syndrome protein 1; NSE2, Non-SMC Element 2; NSE4a, Non-SMC Element 4a; NuRD, Nucleosome Remodeling Deacetylase; P53, Tumor Protein 53; PCNA, Proliferating Cell Nuclear Antigen; PML, Promyelocytic Leukaemia; POLD3, Non-Essential Subunit of DNA Polymerase Delta; POT1, Protection of Telomeres 1; RAD, radiation; RAD21L1, RAD21 Cohesin Complex Component Like 1; RAP1, Ras-related protein 1; REC8, Recombination Protein 8; RFC, Replication Factor C; RNF212, Ring Finger Protein 212; SC, Synaptonemal Complex; SCF Complex, SKP, CULLIN, and F-BOX; SMC3, Structural Maintenance of Chromosomes 3; SMC5/6, Structural Maintenance of Chromosomes Complex 5/6; SPATA22, Spermatogenesis Associated 22; STAG3, Stromal Antigen 3; SYCE1, Synaptonemal Complex Central Element Protein 1; SYCE2, Synaptonemal Complex Central Element Protein 2; SYCE3, Synaptonemal Complex Central Element Protein 3; SYCP1, Synaptonemal Complex Protein 1; SYCP2, Synaptonemal Complex Protein 2; SYCP2L,

Synaptonemal Complex Protein 2 Like; SYCP3, Synaptonemal Complex Protein 3; TERB1, Telomere Repeat Binding Bouquet Formation Protein 1; TERB2, Telomere Repeat Binding Bouquet Formation Protein 2; TEX11, Testis Expressed Gene 11; TEX12, Testis Expressed Gene 12; TIN2, TRF1 Interacting Protein 2; TIR1, Transport Inhibitor Response 1; osTIR1, *O. sativa* TIR1; TPP1, POT1 and TIN2 Interacting Protein; TRF1, Telomeric-Repeat Binding Factor 1; TRF2, Telomeric-Repeat Binding Factor 2; XRCC2, X-Ray Repair Cross Complementing 2; ZNF827, Zinc Finger Protein 827

Introduction

Replicative Immortality and Telomere Function in Cancer

One of the major hallmarks of cancer is an ability to replicate indefinitely, thereby successfully overcoming replicative mortality that most somatic cells eventually succumb to. Characterizing the processes that facilitate a somatic cell to become cancerous has proved a daunting task for researchers; however, the functionality of the telomeres plays a crucial role in determining if a cell will continue to divide. As cells undergo multiple rounds of DNA replication, they lose their telomeric DNA as a result of incomplete DNA synthesis. This functions as an endogenous mitotic clock, which eventually leads a cell to enter into replicative senescence and reach what is called its “Hayflick Limit” (Hayflick and Moorhead, 1961; Shay and Wright, 2011).

Telomeres contain stretches of G-rich tandemly repeated sequences of DNA located on the ends of every chromosome (Pickett and Reddel, 2015). Because telomeric DNA contains repetitive sequences, they have the capacity to form various secondary structures that physically prevent complete DNA replication. As a result, each round of replication leaves about 50-200 bp of DNA un-replicated at the 3’ end, creating what is known as the end replication problem (Levy *et al.*, 1991). The majority of cancer cells contain a mutation that results in the upregulation of the gene coding for telomerase. This is a specialized enzyme that has reverse transcriptase activity accomplished through the catalytic subunit called human telomerase reverse transcriptase (HTERT). However, there are a small subset of cancers (10-15%) that utilize the Alternative Lengthening of Telomeres (ALT) pathway to extend telomeric DNA in the absence of telomerase.

The shelterin protein complex plays an important role in inhibiting initiation of DNA damage responses at the telomere by protecting single-stranded DNA ends from being detected as DNA damage. The shelterin complex consists of six core proteins – telomeric-repeat binding factor 1 (TRF1), TRF2, TRF1 interacting protein 2 (TIN2), protection of telomeres 1 (POT1), the POT1 and TIN2 interacting protein (TPP1) and the transcriptional repressor/activator protein (RAP1) (Figure 1; Deng *et al.*, 2008). When telomeres become shorter due to multiple cycles of replication, it is more likely that they will elicit a DNA damage response as a result of ineffective shelterin functioning. DNA damage responses can result in a variety of events such as inaccurate non-homologous end joining, anaphase bridges, aneuploidy, and reactivation of telomeric DNA replication enzymes.

Misexpression of Meiotic Genes in Cancer

Recently, researchers have demonstrated that multiple genes, once thought to only play a role in meiosis, are actually upregulated in cancer. Furthermore, many of these genes may play an important role in the processes necessary for cancer cells to overcome replicative mortality or avoid cellular senescence. To date, there have been more than 200 genes identified whose expression is restricted to germ cells but often reactivated and aberrantly expressed in tumor cells (Almeida *et al.*, 2009). Thus, it is apparent that the processes involved in gametogenesis and tumorigenesis have various overlaps and similarities. Furthermore, the correlation between epigenetic alterations and aberrant genetic expression patterns in cancer could provide insight on how these genes are being controlled (Wang *et al.*, 2016). Identification of these genes is advantageous because they offer an attractive target for therapy due to their lack of expression in tissues other than the gonads.

Of the meiotic specific genes found to be aberrantly expressed in cancer, they can be broken down into two useful groups; those found on the X chromosome and those found on the autosomes. Scientists have identified that those misexpressed meiotic genes present on the X chromosome are most advantageous for the development of experimental cancer vaccines. The first antigen capable of inducing a cytotoxic T-lymphocyte response (CTL) in cancer patients was termed melanoma antigen-1 (MAGE-1), later discovered to be part of a multigene family. Using T-cell epitope cloning methods, researchers have identified two additional antigen gene families: BAGE and GAGE. The gene families MAGE, BAGE, and CAGE are activated in a wide range of cancers and have been mapped to the X chromosome (Scanlan *et al.*, 2002). Many of the gene families identified on the X chromosome function as transcription factors or other regulatory proteins; however, the majority of them have yet to have their function understood. Additionally, some of the non-X chromosome gene families aberrantly expressed in cancer function as components that aid in the structural integrity of DNA and other cellular components through gametogenesis, while others function as regulatory proteins (Simpson *et al.*, 2005).

Some of the meiotic specific genes identified as being aberrantly expressed in tumor cells are involved in processes that promote meiotic recombination. During meiosis, proteins induce DNA double strand breaks to facilitate homologous recombination, which is essential for increasing genetic diversity among gametes and facilitating proper segregation of homologous chromosomes during the first meiotic division. A protein structure called the synaptonemal complex (SC) forms between two homologous chromosomes and mediates the chromosome pairing and recombination (Baudat *et al.*, 2013). Accurate recombination and synapsis between homologs is essential to prevent uneven segregation of chromosomes resulting in aneuploidy. One hallmark of cancer cells is their lack of a normal karyotype. Cancer cells tend to have

numerous, fragmented chromosomes that fail to segregate properly during multiple rounds of division. Genes whose protein products make up the SC, such as *SYCP1*, or those that function in homologous chromosome pairing and recombination during SC formation, such as *SPO11*, have been identified as being expressed in some cancer lines. *SPO11* is a component of a type IIB topoisomerase that forms DNA double stranded breaks (DSBs) during meiosis. These DSBs are essential for the initiation of homologous recombination. *SYCP1* proteins form the transverse filaments of the SC, which bridges the two homologous chromosomes together (Simpson *et al.*, 2005). Thus, meiotic proteins involved in chromosome structure, recombination and segregation are of interest in cancer, as they may be aberrantly expressed in response to, or even a source of, these malfunctions.

One meiotic gene of interest that is aberrantly expressed in some cancer lines is *MEIOB*, which is also essential for meiotic recombination and exhibits 3'-5' exonuclease activity. *MEIOB* functions by localizing to meiotic chromatin and forming a complex with *SPATA22*, which is thought to regulate its nuclease activity, and enhance the specificity of the substrate that *MEIOB* binds to (Luo *et al.*, 2013). It has been demonstrated that the *MEIOB-SPATA22* complex plays a critical role during the development of cancer, especially lung adenocarcinoma (LUAD) (Wang *et al.*, 2016). Another example is the HORMA domain protein 1 (*HORMAD1*), which is expressed in triple-negative breast cancers (TNBSs). *HORMAD1* is essential for normal SC formation and the recruitment of *ATR* to un-synapsed chromatin during meiosis. In DSB repair, DNA end resection results in single-stranded 3' overhangs bound by *RAD51* and *DMC1* recombinases. These proteins facilitate homology search to find a primer template for new DNA synthesis to occur between homologs or sister chromatids (Daniel *et al.*, 2011).

HORMAD1 is involved in inhibiting RAD51-dependent sister chromatid recombination in favor of DMC1-mediated recombination with the homologous chromosome during meiosis. The inhibition of RAD51 function in cancer cells promotes alternative forms of DNA repair, resulting in the generation of Allelic-imbalanced Copy Number Aberrations (AiCNA) (Watkins *et al.*, 2015). Furthermore, DMC1 has been indicated as having increased expression in certain cancer lines (Kalejs *et al.*, 2006). Thus, increased DMC1-mediated recombination in cancer may result in further aberrations such as translocations, insertions, or deletions. Taken together, meiotic specific proteins such as DMC1, HORMAD1, and SPO11 that are involved in the initiation of cross over events and successful resolution via HR play an important role in the genomic instability of cancer cells. Researchers are still working to better understand the function of these meiotic specific HR genes when they are aberrantly expressed in cancer cell lines.

DNA Damage Responses and Cancer

After multiple cell divisions telomere functioning can become impaired, and the canonical DNA damage response pathway is activated. This involves p53, which facilitates a cascade of events leading to apoptosis or replicative senescence (Deng *et al.*, 2008). One of the main proteins involved in helping to control p53 concentrations in the cell is MDM2, which binds to p53 and facilitates its degradation (Momand *et al.*, 1999). When the cell senses DNA damage, MDM2 is inhibited, thus concentrations of p53 in the cell rise. This results in the activation of various downstream signaling pathways that ultimately lead to apoptosis or cellular senescence.

Continuously dividing somatic cells succumb to various sources of DNA damage throughout their lifespan. Some DNA damage is due to environmental factors, such as UV radiation, while others are intrinsic, such as errors in DNA replication. Sometimes these genetic insults result in single base pair mutations (Tomasetti *et al.*, 2017). Cells have developed ways to detect and correct these errors, such as nucleotide excision repair (NER) and mismatch repair (MMR). However, sometimes DNA damage can result in single and double stranded breaks (SSB and DSB, respectively), which require a more specific method of detection and correction. The initial steps of detection involve recognition of free DNA ends by the MRE11-RAD50-NBS1 complex (MRN complex), which activates ataxia-telangiectasia mutated (ATM). Activated ATM will phosphorylate substrates such as CHK2, p53, and H2AX. Phosphorylated H2AX (γ H2AX) is recognized by MDC1, which spreads activated ATM and γ H2AX over a large area of the chromatin (Marechal and Xou, 2013). This process creates a platform for subsequent reactions to take place to fix the SSB or DSB. The most severe form of DNA damage is the DSB, and it is an essential part of the ALT mechanism.

Two major ways cells repair DSBs is through homologous recombination (HR) and non-homologous end-joining (NHEJ). In HR, DNA is first resected to create a 3' single-stranded overhang before a homologous chromosome is used as a DNA template to synthesize the missing nucleotides, resulting in more accurate DSB resolution. NHEJ, however, is more error prone and involves DNA ligase IV functioning to re-ligate the two free DNA ends together without resection or a homologous template (Polo and Jackson, 2011). If nucleotides are lost during DSB formation they are not accounted for during NHEJ, which could result in point mutations. Both HR and NHEJ indicate some level of genomic instability occurring in the cell, and proliferating cells demonstrate the highest propensity for HR (Bishop and Schiestl, 2002).

In meiosis, HR facilitates exchange of genetic information between maternal and paternal alleles, thus generating genetic diversity. Furthermore, it facilitates accurate segregation of homologous chromosomes during meiosis I by forming chiasmata, thus ensuring that aneuploidy does not occur (Filippo *et al.*, 2008). In cancer, HR machinery is highly active as the DNA of cancer cells is fragmented, full of damage, and cells are rapidly dividing. In rapidly dividing cell populations, such as those found in the intestine and uterine epithelia, HR could be a source of genetic alterations and potentially result in a loss of heterozygosity or cause aberrant genomic rearrangements that may eventually lead to carcinogenesis (Bishop and Schiestl, 2002).

If DNA damage occurs and is not successfully fixed, there are a number of problems that could result. If the damage incurred results in a non-functional tumor suppressor protein, for example, the cell could lack the ability to turn off mitogenic signals resulting in continued cell divisions. Likewise, if the damage results in an overactive oncogenic protein, this cell could continue proliferating as the signal produces positive feedback on mitogenic pathways.

Furthermore, if the damage occurs at the telomeres, the cell may upregulate telomerase or the ALT pathway. This could potentially enhance telomere elongation and prepare the cell for multiple cellular divisions. Many cancer promoting genes are affected as a result of DNA damaging events, resulting in faulty mitogenic pathways or the inhibition of apoptosis or cellular senescence.

The ALT Phenotype and Homologous Recombination

The ALT phenotype is characterized by the presence of large specialized ALT-associated promyelocytic leukaemia (PML) nuclear bodies (APBs), which co-localize with telomeric DNA and other telomere associated proteins (Muntoni and Reddel, 2005). APBs contain proteins that

are involved in DNA recombination and DNA repair, which is a vital part of the ALT mechanism. Furthermore, APBs contain telomere specific proteins such as TRF1 and TRF2, which are part of the shelterin complex. This complex helps to maintain the structural integrity of the telomere by preventing the telomere from eliciting a DNA damage response or illegitimate nucleolytic degradation (Deng *et al.*, 2008). In addition, the ALT phenotype can be identified by large amounts of extrachromosomal telomeric DNA, which can exist in various forms, including predominantly double-stranded telomeric circles (t-circles), partially single-stranded circles (C-circles or G-circles), linear double-stranded DNA, and “t-complex” DNA that contain abnormal, branched structures (Cesare and Reddel, 2010).

The processes that facilitate ALT are usually results in telomeric DNA that is not sequentially or structurally the same as normal telomeres. Normal telomeric DNA contain stretches of G-rich tandemly repeated sequences. These sequences provide the telomeric DNA with the capacity to form various secondary structures, such as G-quadruplexes and T-loops, as well as other hypothetical structures such as triple helices, four-way junctions, and D-loops (Gilson and Geli, 2007). These structures form physical barriers for replication machinery to combat, thus replication at the telomeres is incomplete, resulting in about 50-200 bp of DNA unreplicated at the 3' end during each round of cellular division. Because telomeric DNA contains a repetitive sequence, when ALT positive cells extend their telomeres using HR, the invading strand may not always match up perfectly as a template to allow for the canonical TTAGGG sequence to be repeated accurately. This process results in telomeric DNA becoming interspersed with variable, non-canonical sequences, which is a hallmark of the ALT phenotype. As a result, telomeric binding proteins, like those in the shelterin complex, are not able to bind efficiently to telomeres and therefore lose their function (Bechter *et al.*, 2004).

The HR repair mechanism is essential for the ALT pathway. A collection of core genes, known as the *RAD52* epistasis group, is required for HR. The protein products of this group in humans includes the MRN complex – consisting of MRE11, RAD50, and NBS1 – BRCA2, RAD52, RAD54, RAD54B, RAD51B-RAD51C complex, RAD51D-XRCC2 complex, and the RAD51C-XRCC3 complex (Filippo et al., 2008). In addition, the expression of a variety of other accessory genes are required for ALT – some are also expressed in meiosis. Whether recombination is occurring in meiosis or as part of the ALT pathway, the process starts with a DSB followed by end resection and strand invasion on a DNA template – a homologous chromosome in the case of homologous recombination.

In the ALT mechanism, it is thought that a recombination-dependent replication process called Break Induced Repair (BIR) is necessary. This mechanism is used to repair broken chromosomes when a single-stranded overhang is present in DNA (Kraus *et al.*, 2001). Single-stranded overhangs are common in stalled replication forks, which frequently occur at telomeres due to the abnormal secondary structures present. When a fork collapses in a telomere, it is unlikely to be resolved by incoming or dormant forks, since it is thought that human telomeres lack replication origins. In BIR, a D-loop is formed when the 3'-end of a single-stranded DNA invades a double-stranded homologous DNA segment. This invading strand then serves as a primer for initiation of DNA replication. The non-essential subunit of DNA polymerase delta, POLD3, is required for BIR and leads to conservative DNA replication (Figure 2; Roumelioti et al., 2016). Furthermore, the initial sensor of telomere damage that establishes DNA polymerase delta (through POLD3) is the proliferating cell nuclear antigen (PCNA), which is loaded by replication factor C (RFC). BIR at the telomeres is thought to be independent of ATM, ATR, or RAD51, but requires the RFC-PCNA-Pol delta axis (Dilley *et al.*, 2016).

The SMC5/6 complex promotes the repair of DNA DSBs through using HR, and in ALT positive cells, it is required for telomeric DNA to be included in APBs (Amorim *et al.*, 2016). Furthermore, the SUMO ligase MMS21, part of the SMC5/6 complex, is required for APB formation through stimulating the SUMOylation of various subunits of the shelterin complex (Potts & Yu, 2007). Another important meiotic protein complex thought to be misexpressed and functioning in the ALT pathway is the ATP-dependent helicase protein ATRX and its H3.3-specific histone chaperone DAXX. This complex plays an important role in PML bodies, and most ALT positive cancer cells in humans have a mutated, non-functional ATRX-DAXX protein complex (Heaphy *et al.*, 2011). It is not apparent that ATRX possess G-quadruplex DNA (G4-DNA) unwinding activity, and it may overcome this obstacle indirectly by facilitating the histone H3.3 deposition so that DNA is maintained in a B-form conformation, or by promoting a fork bypass through template switching (Amorim *et al.*, 2016). Thus, it has been suggested that a potential consequence of losing ATRX functioning is a higher frequency of G-quadruplexes, which results in more DNA damage at the telomeres. The presence of G-quadruplexes can promote the ALT phenotype by presenting a barrier to the replication fork, thus causing fork stalling, collapse, and subsequent activation of the HR mechanism as a means to repair the damaged DNA (Amorim *et al.*, 2016).

The heterogeneity within telomeric DNA, caused by BIR or other repair mechanisms, can create high-affinity binding sites for a group of nuclear hormone receptors (Pickett and Reddel, 2015). A zinc finger protein known as ZNF827 has been demonstrated as being recruited to telomeres in ALT positive cells by these nuclear hormone receptors such as TRF and COUP-TF2 (Conomos *et al.*, 2014). Once ZNF827 has been recruited to the telomeres, it helps to recruit the Nucleosome Remodeling Deacetylase (NuRD) complex in a sequence specific manner (Laubert

and Rauchman, 2006; Conomos *et al.*, 2014). In *Drosophila*, the NuRD complex, containing the subunit MI-2, functions to promote chromosome condensation during meiosis in oocytes (Nikalayevich and Ohkura, 2015). The NuRD complex contains nucleosome-remodeling and histone-deacetylation functions. The nucleosome remodeling function can displace shelterin, while the histone-deacetylation function (in conjunction with ZNF827) may play a role in countering histone demethylation by compacting telomeric chromatin. Once the NuRD-ZNF827 complex is established, it can recruit proteins involved in DDR and HR (Conomos *et al.*, 2014).

Meiotic misexpression in the ALT pathway

One of the meiotic specific complexes that is of particular interest in interacting in the ALT pathway is HOP2-MND1. This protein stimulated D-loop formation through interacting with the two recombinases RAD51 and DMC1, thus ensuring recombination between homologous chromosomes is favored over recombination between sister chromatids during meiosis (Cho *et al.*, 2014). Because HOP2-MND1 favors specific DNA molecules, it is possible that it increases the variety of templates available during HR for the ALT pathway to work. ALT positive cells have characteristics that are unique to this pathway, thus it may contain specific features that allow interaction between RAD51, HOP2, and MND1 (Arnoult & Karlseder, 2014). These interactions increase the sequence heterogeneity, which is conducive to the ALT phenotype. Variation in telomeric DNA sequences of ALT positive cells has been shown to be important for its interaction with many of the proteins involved in the ALT (Lee *et al.*, 2014). In addition, HOP2-MND1 has been demonstrated as localizing to APBs on the telomeres of ALT positive cells, and knocking down HOP2-MND1 in these cells results in a significant reduction in APB formation (Cho *et al.*, 2014).

Proteins involved in HR play an important role in facilitating their respective function in the ALT pathway. In meiosis, the protein disrupted meiotic cDNA 1 (DMC1) forms right-handed helical filaments on ssDNA in an ATP-dependent manner, catalyzing the pairing of homologous DNA and facilitating strand exchange reactions within these nucleoprotein filaments (Filippo *et al.*, 2008). Likewise, the ATP dependent DNA helicase homolog HFM1 is expressed during meiosis and helps to facilitate successful cross-over formation during HR, which leads to the complete synapsis of homologous chromosomes (Pu *et al.*, 2016). Proteins like these, which are known to play an important role in meiosis, are of interest in their interaction with the ALT mechanism and their propagation in cancer in general.

Auxin Induced Degradation (AID)

The Auxin Induced Degradation (AID) mechanism is a tool used to quickly, conditionally, and reversibly deplete a protein of interest *in vitro*. This system is a plant specific mechanism and includes two parts: 1 – the SCF complex, which is an E3 ubiquitin ligase and consists of three proteins – SKP, CULLIN, and a variable F-BOX, and 2 – target proteins that harbor a specific domain/motif known as the auxin-inducible degron. In plants, this system is utilized as a method of depleting members of the AUX/AII family of transcription repressors (which all share four conserved domain/motifs) in the presence of indole-3-acetic acid (IAA or auxin) (Gray *et al.*, 2001). Researchers have been able to harvest one of the AUX/AII F-BOX domains of the plant SCF E3 ubiquitin ligase called TIR1 and transplant it into other eukaryotes, as they lack the auxin response but still utilize the SCF degradation pathway. Additionally, researchers have identified and isolated the domain/motif of the AUX/AII protein that the TIR1 protein binds to in the presence of auxin (Nishimura *et al.*, 2009). Thus, this domain/motif can be

incorporated into a protein of interest by inserting it into the 3' or 5' end of the coding sequence for the protein; thus, in the presence of auxin, this protein will be ubiquitinated by the SCF-TIR1 and subsequently degraded via the proteasome (Figure 3; Natsume *et al.*, 2016).

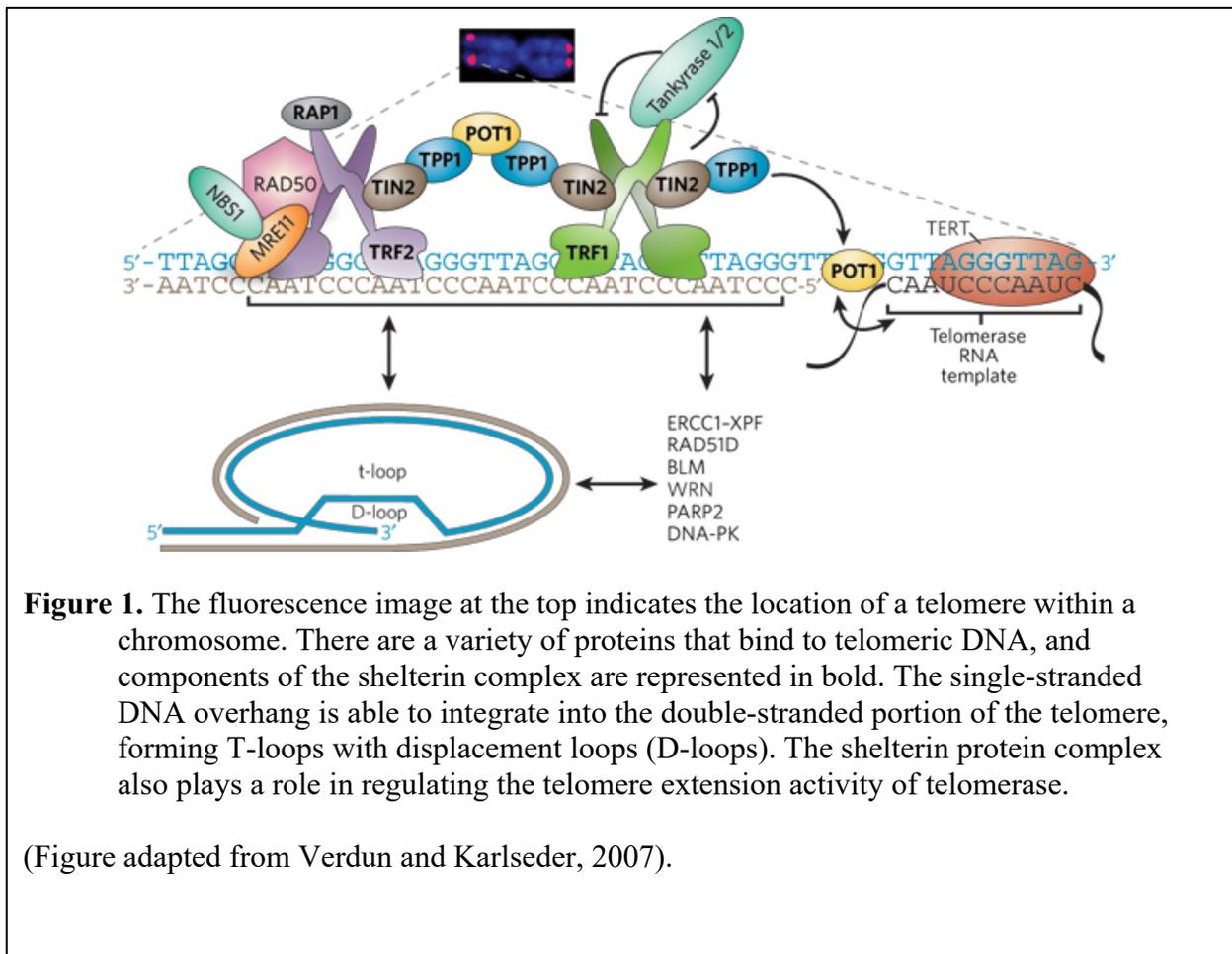
Auxin shows no adverse effects on cells *in vitro*, so utilizing these mechanisms as a means to conditionally deplete a protein of interest in eukaryotic cell systems is advantageous. This mechanism is more efficient and specific than traditional RNAi silencing and results in less off target effects (Jackson and Linsley, 2010). Researchers have also identified ways to enhance the interactions between the TIR1 protein, the SCF complex, and the AID-tagged protein of interest (Figure 3). The TIR1 protein derived from the plant *O. sativa* (osTIR1), which grows in a warmer environment, has been identified as working optimally around 37°C. This thermostable TIR1 works better for transfection of cells in culture because they grow best at 37°C. (Nishimura *et al.*, 2009). Furthermore, a new mini-AID (mAID) tag is being used, which is about 47 amino acids long. The smaller tag on the protein decreases the likelihood that it will physically or functionally interfere with the normal role the protein of interest plays in the cell (Brosh *et al.*, 2016). In addition to enhancing the interactions between the proteins themselves, scientists have also enhanced the specificity of CRISPR to function in successfully inserting the osTIR1 and the mAID tag in their respective locations in the genome. A specific kind of CAS9 called enhanced specificity *Streptococcus pyogenes* Cas9 (eSpCAS9) has been modified to decrease off-target effects by attenuating mismatches between sgRNA and target DNA so they would be less energetically favorable (Slaymaker *et al.*, 2016).

There are various methods of inserting the TIR1 protein into the genome of the host cell using CRISPR technology. Likewise, there are many locations within the genome for incorporation. One of the most successful and efficient locations in the genome to insert a

constitutively expressed TIR1 gene is in the adeno-associated virus integration site 1 (AAVS1) safe harbor locus. This site is located on chromosome 19 (position 19q13.42), and incorporation here will not interfere with the expression of other endogenous genes. Other loci that have been targeted for transgene addition are the chemokine (CC motif) receptor 5 (CCR5) and the human orthologue of the mouse ROSA26 locus; however, there are shortcomings to each site that may differ by the cell line being used (Sadelain *et al.*, 2012). For example, if random integration occurs within or near a cancer-related gene it could result in oncogenesis or cells entering senescence, depending on what gene is interfered with. Integration within an oncogene could result in variable transgene expression or insertional oncogenesis (Sadelain *et al.*, 2012). These variables have the potential to confound results in experiments conducted using these transgenic cells.

Two major methods have been employed to integrate the CRISPR/CAS9 machinery and template DNA into the cell – electroporation and cationic lipid mediated transfection. Electroporation may result in a lower yield of successfully transfected cells as well as fragmented template DNA to be inserted. Cationic lipid mediated transfection is one of the most efficient methods, and constitutes incubating the CRISPR/CAS9 machinery and template DNA with liposomes, which subsequently form DNA-liposome complex. Cationic lipids consist of a positively charged head group bound to one or two hydrocarbon chains, which forms a micelle in solution. The positively charged head group binds to the negatively charged backbone of the nucleic acid, facilitating DNA condensation. This positively charged surface of the lipids also mediates fusion of the liposome-DNA complex with the negatively charged cell membrane, which subsequently enters the cell through endocytosis (Figure 4; ThermoFisher Scientific; Liang *et al.*, 2015). This method has proven to be more efficient, simple to perform, work on a

wide variety of eukaryotic cells, be less toxic to cells, and have higher yields of successful transfection than other methods (ThermoFisher Scientific).



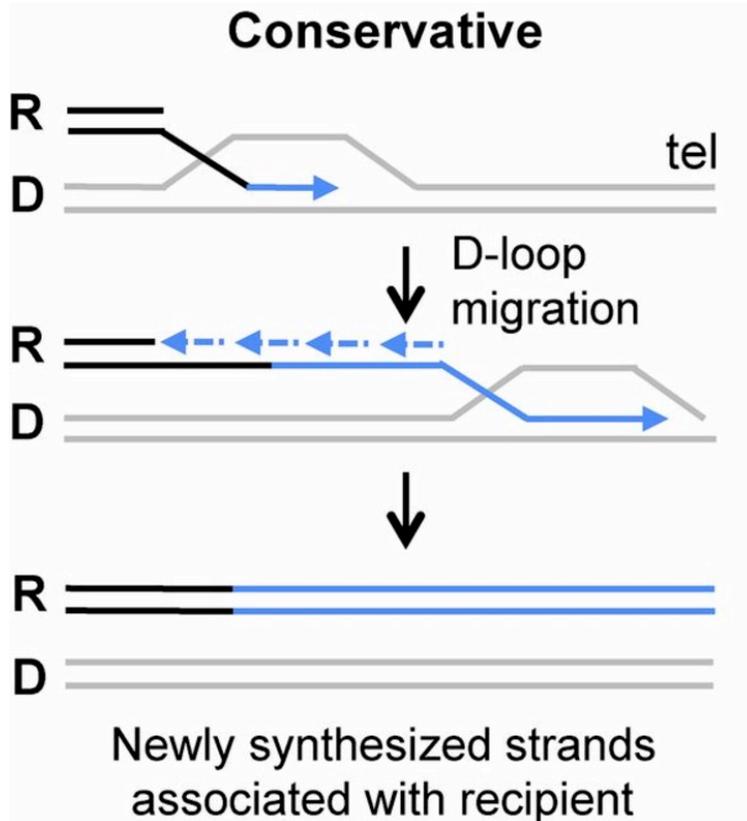


Figure 2. Model for conservative DNA synthesis during BIR. If the D-loop migrates towards the telomere (tel) and synthesis of the lagging strand is initiated on the displaced, nascent strand, then both newly synthesized strands (in blue) will segregate with the recipient chromosome.

Note: R-recipient strand; D-donor strand; tel-telomere end

(Figure adapted from Donnianni and Symington, 2013).

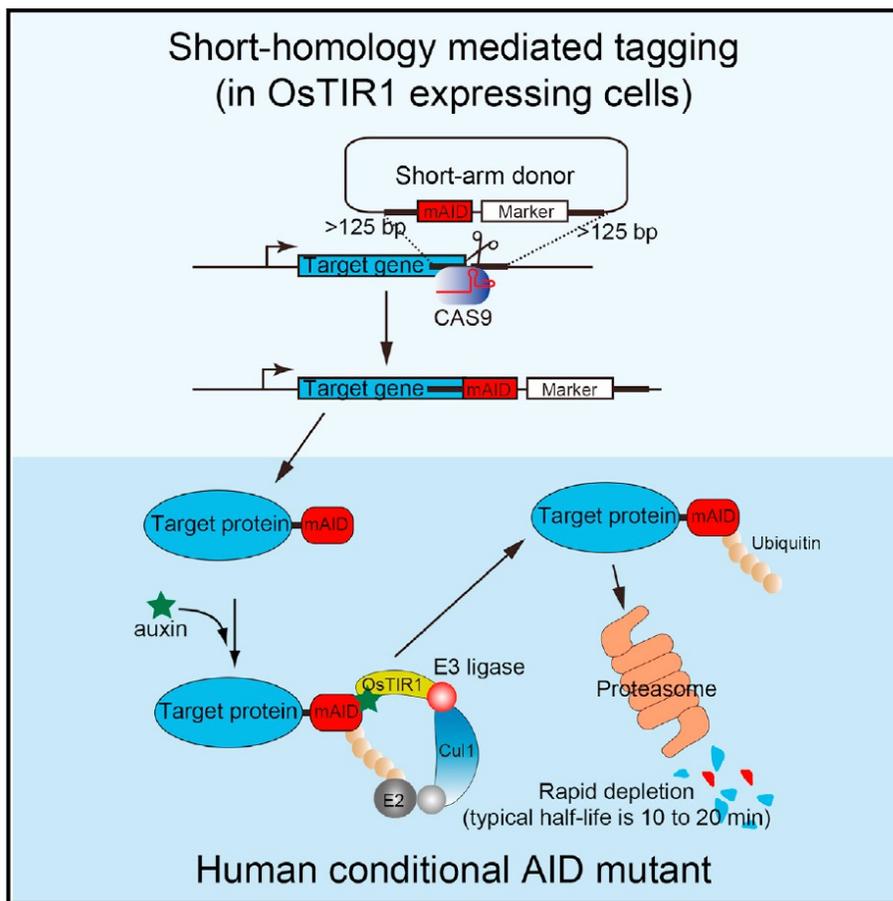


Figure 3. In this cell line, a TIR1 protein derived from the plant *O. sativa* (osTIR1) is being constitutively expressed. Using CRISPR and homology directed repair mechanisms, a mini-version of the AID tag (mAID) is successfully inserted on the 3' end of coding sequence for the target protein of interest. After transcription and translation, the target protein has the mAID tag on it and is recognized by the osTIR1 F-BOX domain of the SCF ubiquitin ligase and subsequently degraded via the proteasome.

(Figure adapted from Natsume, T. *et al.*, 2016).

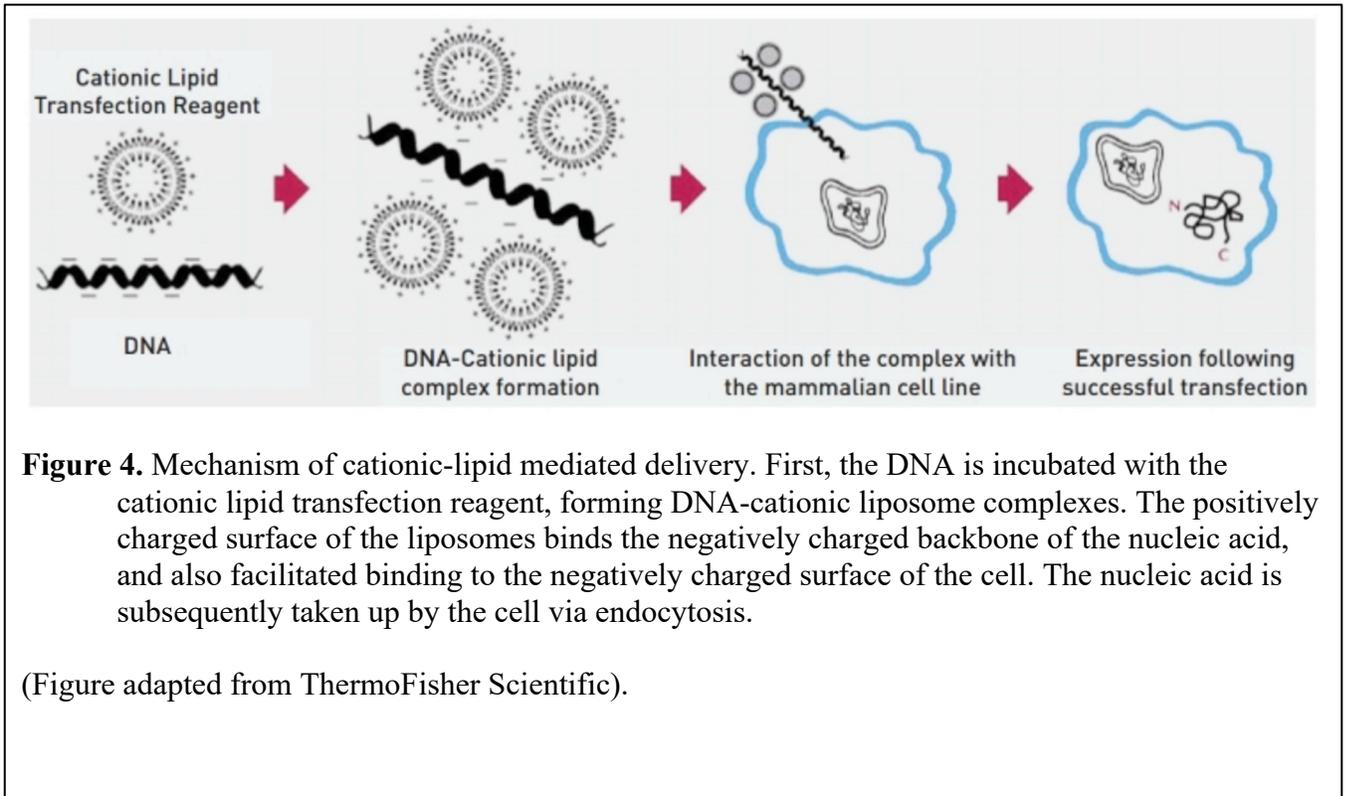


Figure 4. Mechanism of cationic-lipid mediated delivery. First, the DNA is incubated with the cationic lipid transfection reagent, forming DNA-cationic liposome complexes. The positively charged surface of the liposomes binds the negatively charged backbone of the nucleic acid, and also facilitated binding to the negatively charged surface of the cell. The nucleic acid is subsequently taken up by the cell via endocytosis.

(Figure adapted from ThermoFisher Scientific).

Materials and Methods

Cell Lines for Protein Isolates

Protein isolates for this project were provided by Dr. Alan Meeker's lab. 23 different ALT positive and ALT negative cell lines derived from various tissue were used analyzed. In Table 1 below, the cell lines used in our protein isolate analysis are presented, organized by tissue origin and ALT phenotype:

Table 1. Cell lines used in protein analysis

	ALT Positive	ALT Negative	Non-Immortalized
SV-Immortalized (Fibroblast)	<ul style="list-style-type: none"> • GM847 	<ul style="list-style-type: none"> • VA13 	
hTERT-Immortalized		<ul style="list-style-type: none"> • BJ-TERT 	
Osteosarcoma	<ul style="list-style-type: none"> • U2OS • SAOS2 	<ul style="list-style-type: none"> • MG63 • SJSA1 	
Pediatric Glioma	<ul style="list-style-type: none"> • 1118 • SJ-GBM2 	<ul style="list-style-type: none"> • KNS42 • SF188 • UW479 	
Normal Breast Stroma			<ul style="list-style-type: none"> • 1 Patient-Derived Sample
Prostate Mesenchymal			<ul style="list-style-type: none"> • 3 Patient Derived Samples (12493, 13407, 12493)
ATRX Isogenics		<ul style="list-style-type: none"> • U251 • U251 ATRX-/- M • ATRX -/- Q • MOG • MOG ATRX-/- A • MOG ATRX-/- I 	

Cell Culture Conditions

In addition to obtaining and analyzing protein isolates from these cell lines, we also obtained frozen cell suspensions of the U2OS and SAOS2 cell lines from Dr. Alan Meeker's lab for culture. Both U2OS and SAOS2 cells were cultured in complete growth media (CGM) – DMEM supplemented with 10% heat inactivated FBS and 1% pen/strep – and incubated at 37°C, 5% CO² in filtered cell culture vessels. U2OS cells were passaged at a sub-cultivation ratio of 1:3 every 3 days while SAOS2 cells were passaged at a sub-cultivation ratio of 1:5 every 5 days; at this point the cells reached sub-confluence. To passage, cells were first washed twice with 1 X PBS and then incubated in 0.25% Trypsin-EDTA for 1 minute at 37°C to facilitate detachment of cells from the cell culture vessel surface. The Trypsin-EDTA was inactivated by adding 2 volumes CGM. After quenching, the cells were centrifuged at 200xg for 3 minutes. U2OS cells were re-suspended in 6mL CGM, with 2mL being plated in a cell culture vessel with equal surface area as the original, resulting in a subcultivation ratio of 1:3. Likewise, SAOS2 cells were re-suspended in 10mL CGM, with 2mL being plated in a cell culture vessel with equal surface area as the original, resulting in a subcultivation ratio of 1:5.

Protein Isolation and quantification

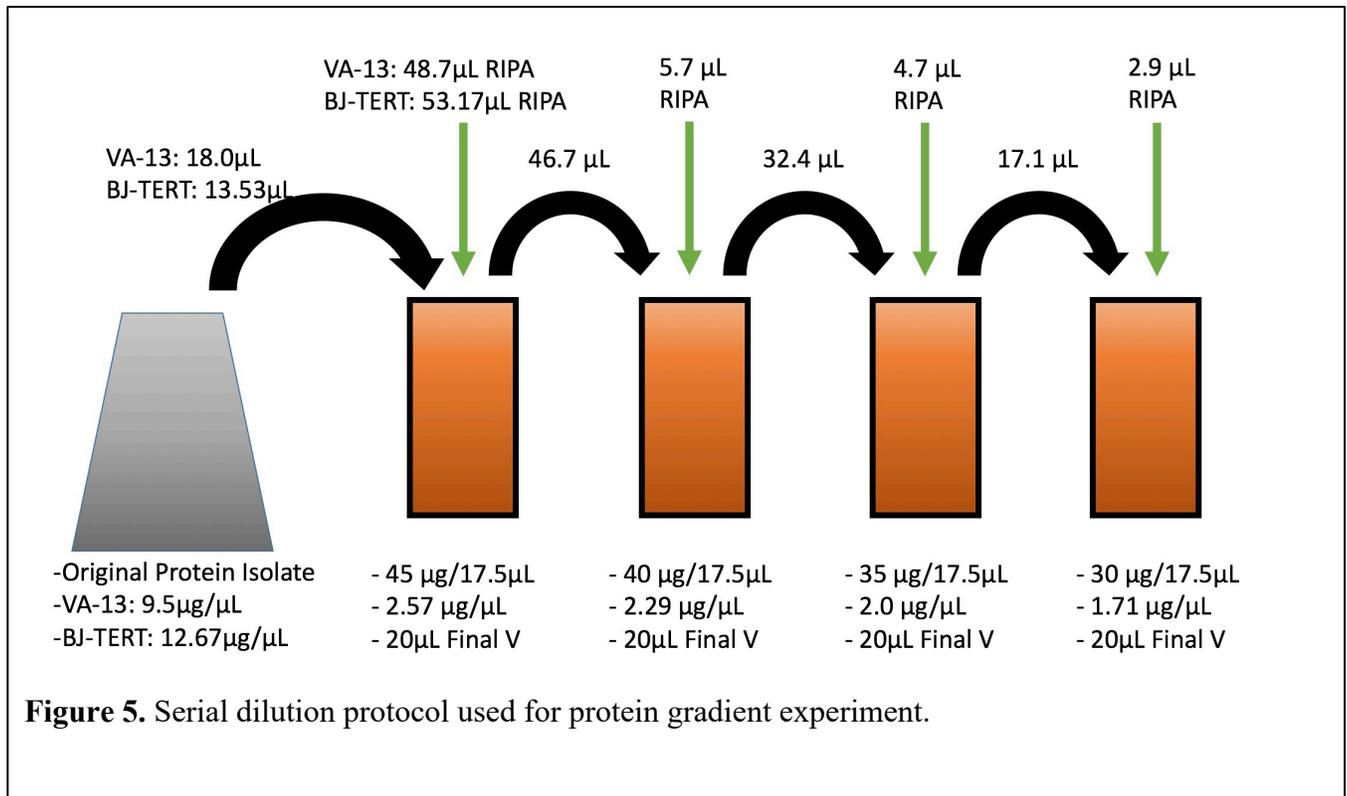
For western blot analysis, cells were taken from one T-75 culture flask at 80% confluence. Cells were collected using trypsin-EDTA detachment methods previously described. Cells were centrifuged at 200xg for 3 minutes followed by aspiration of the supernatant. Following, cells were re-suspended in 1X PBS and subsequently centrifuged again at 200xg for 3 minutes. This wash step was performed one more time before re-suspending cells in RIPA buffer (Santa Cruz) supplemented with 20 µL/1mL Protease Inhibitor (Roche). Cells were left rocking

at 4°C for 20 minutes. The solution was vortexed three times for 5 seconds each and left rocking at 4°C for an additional 30 minutes. To disassociate proteins from DNA and other cellular components, lysates were sonicated (Bioruptor sonication system) at high intensity for 5 minutes with 30 second on/off intervals. Finally, cell lysates were centrifuged at 10,000rpm for 10 minutes and the supernatant was saved as the protein isolate.

The concentration of protein in each sample was quantified using the bicinchoninic acid (BCA) assay and normalized to allow for band intensity to be compared on a western blot. In the BCA assay, peptide bonds reduce Cu^{2+} ions to Cu^+ in a temperature dependent manner. Two molecules of bicinchoninic acid subsequently chelate with each Cu^+ ion, which forms a solution that strongly absorbs light at 562nm (Thermo Scientific). Subsequently, protein samples were compared against a standard curve using bovine serum albumin (BSA) as the standard. The stock concentration of BSA was $2000\mu\text{g}/\mu\text{L}$ and the standard curve was built by performing a serial dilution down to $125\mu\text{g}/\mu\text{L}$. After the concentration of each protein sample was determined, samples were subsequently diluted to the same concentration ($2.3\mu\text{g}/\mu\text{L}$). $20\mu\text{L}$ of each quantized sample was added to $20\mu\text{L}$ of 2x running buffer supplemented 1:20 with β -mercaptaethanol and incubated at 95°C for 5 minutes.

The optimal protein concentration of $2.3\mu\text{g}/\mu\text{L}$ was determined by running a protein gradient concentration experiment. At a concentration of $2.3\mu\text{g}/\mu\text{L}$, $40\mu\text{g}$ of protein is present in $17.5\mu\text{L}$. After adding equal amounts of running buffer, the volume of $35\mu\text{L}$ contained $40\mu\text{g}$ of protein. Protein isolates of BJ-TERT and VA-13 underwent a serial dilution, so that 45, 40, 35, and $30\mu\text{g}$ were added to separate wells of a 4-5% gradient polyacrylamide gel (figure 5). The gel was subsequently electrophoresed at 100V to separate proteins based on size. The gel was then transferred to a PVDF membrane using the TransBlot Turbo system and HFM1 was stained for

along. HFM1 should be expressed at very low concentrations in these cells, since it is a meiotic specific protein. Thus, by doing this protein concentration experiment and staining for HFM1, other proteins that may be expressed at higher concentrations should also be detected. The results indicated that 2.3 $\mu\text{g}/\mu\text{L}$, or 40 $\mu\text{g}/\text{well}$, was the optimal protein concentration to use (figure 6).



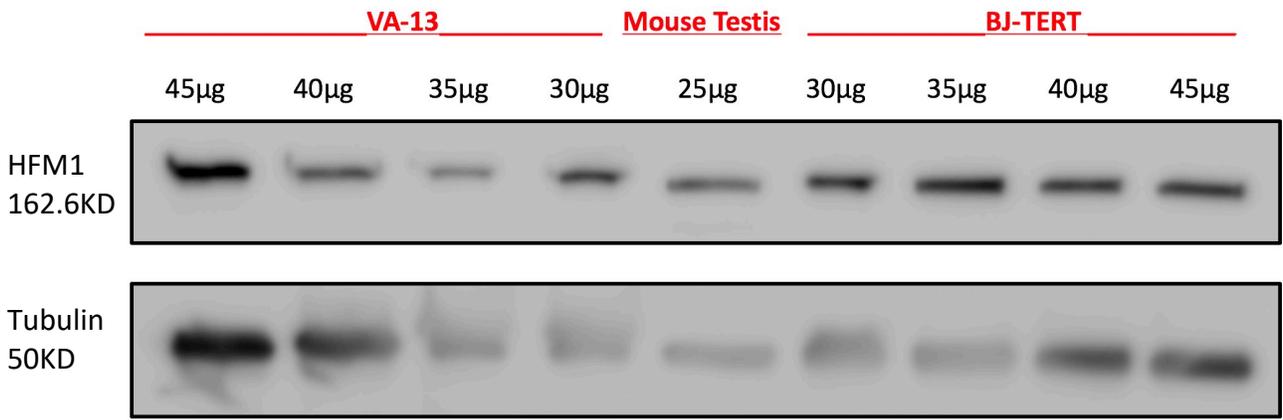


Figure 6. Results for protein gradient experiment. The top lane is HFM1, while the bottom lane is Tubulin. The middle lane, labeled mouse testis (25µg), was used as a control. 40µg per well was chosen because it was the lowest concentration that showed a clear result for both BJ-TERT and VA-13.

Western Blot Analysis

Protein samples were added to a 4-15% polyacrylamide gradient, pre-cast SDS-PAGE gel at a quantity of 40µg/well and subsequently electrophoresed at 100V to separate proteins based on size. Separated proteins were then transferred from the gel to a PVDF membrane using the TransBlot turbo system. After transfer, PVDF membranes were blocked in sterile filtered 3% bovine serum albumin (BSA) in 0.2% Tween20-PBS (PBS-T) (BSA-T) overnight at 4°C. Both primary and secondary antibodies were diluted in 3% BSA-T to their appropriate concentration (Table 2). The membrane was incubated in primary antibody for 2 hours at room temperature (RT) and then washed 2 times 15 minutes each with 0.2% PBS-Tween20 (PBS-T) at RT. The membrane was then incubated in its respective secondary antibody conjugated to horse radish

peroxidase for 1 hour at RT and subsequently washed 1 time 15 minutes followed by 2 times 5 minutes in PBS-T. Finally, one last 5-minute wash in 1 X PBS was performed before immediately imaging using enhanced chemi-luminescence (ECL).

Membranes were incubated in the 1:1 mixture of clarity ECL substrates (Bio-Rad) for 2 minutes in the dark at RT before being transferred to a plastic paper cover slip and imaged using the Syngene SR5 system. The optimal exposure time was determined by setting the system to take images at various time points, such as 10, 20, 40, 60, and 90 minutes of exposure time. The images were then compared to determine the optimal exposure time for each western blot. This time varied between a few minutes to up to over an hour depending on the sensitivity of the antibody, specificity of the antibody, or prevalence of the specific protein of interest in the cell.

1° or 2°	Antigen Specificity	Supplier	Protein Size (kDa)	Dilution Used
Primary Antibody	Mouse α DMC1	Thermo: MA1-20220	37.8	1:1000
	Mouse α HFM1	Santa Cruz: SC-514597	162.6	1:500
	Rabbit α HOP2	Novus: NBP1-92301	24.7	1:200
	Rabbit α HORMAD2	Abcam: AB106256	34	1:500
	Mouse α NSE2	Abnova: H00286053-B01	27.9	1:200
	Rabbit α NSE4a	Sigma: HPA037459	43.7	1:100
	Rabbit α REC8	Abcam: AB38372	67.4	1:5000
	Mouse α Tubulin	Sigma: T9026	50	1:5000
	Rabbit α osTIR1	Kanemaki	74	1:2000
	Rabbit α SMC5	Novus: 100-469	129	1:500
	Rabbit α SMC6	Abcam: AB155495	126	1:500
	Goat α SYCE2	Santa Cruz: SC240935	19.5	1:100
Secondary Antibody	Mouse α HRP	Invitrogen: R21455	N/A	1:5000
	Rabbit α HRP	Invitrogen: A10533	N/A	1:5000
	Goat α HRP	Invitrogen: R21459	N/A	1:5000

Table 2. List of antibodies used in western blot procedure and analysis via ECL.

RT-PCR and QPCR Procedure

Reverse transcription PCR (RT-PCR) and Quantitative RT-PCR (Q-PCR) analysis of a group of meiotic genes of interest (Table 3) was performed on four different osteosarcoma cell lines. Each tissue type has a unique profile; therefore, we chose to analyze only osteosarcoma cells to avoid discrepancies between tissue types. We analyzed two ALT positive cell lines (U2OS and SAOS2) and two ALT negative (telomerase positive) cell lines (SJSA1 and MG63).

Table 3. Meiotic genes of interest used in QPCR analysis

1) <i>SPO11</i>	10) <i>HORMAD1</i>	19) <i>FKBP6</i>	28) <i>MEIOB</i>
2) <i>DMC1</i>	11) <i>HORMAD2</i>	20) <i>HFMI</i>	29) <i>SPATA22</i>
3) <i>MEI1</i>	12) <i>SYCP1</i>	21) <i>MSH5</i>	30) <i>TERB1</i>
4) <i>MND1</i>	13) <i>SYCP2</i>	22) <i>MSH4</i>	31) <i>MEIKIN</i>
5) <i>RAD21L1</i>	14) <i>SYCP3</i>	23) <i>TEX11</i>	32) <i>TERB2</i>
6) <i>SMC3</i>	15) <i>SYCE1</i>	24) <i>TEX12</i>	33) <i>MAJIN</i>
7) <i>REC8</i>	16) <i>SYCE2</i>	25) <i>RNF212</i>	
8) <i>STAG3</i>	17) <i>SYCE3</i>	26) <i>CCNB1IP1</i>	
9) <i>HOP2</i>	18) <i>SYCP2L</i>	27) <i>MSNI</i>	

UCSC Cancer Browser Screening

An analysis of the levels of expression of a group of meiotic genes of interest (Table 4) in lower grade glioma was performed using the UCSC Cancer Browser data. This analysis was performed using data from the TCGA Brain Lower Grade Glioma dataset. The data presented is

stratified by ATRX expression levels, given that ATRX is lost in most of the cancer cell lines that demonstrate the ALT phenotype.

Table 4. Meiotic Genes of Interest used in UCSC Cancer Browser screening

<i>SPO11</i>	<i>HORMAD1</i>	<i>MSH4</i>
<i>MEI11</i>	<i>HORMAD2</i>	<i>TEX11</i>
<i>DMC1</i>	<i>SYCP1</i>	<i>TEX12</i>
<i>MND1</i>	<i>SYCP2</i>	<i>RNF212</i>
<i>RAD21L1</i>	<i>SYCP3</i>	<i>CCNB1LP1</i>
<i>SMC3</i>	<i>SYCE1</i>	<i>MNS1</i>
<i>REC8</i>	<i>FKBP6</i>	<i>SPATA22</i>
<i>STAG3</i>	<i>HFM1</i>	
<i>PSMC3IP</i>	<i>MSH5</i>	

Plasmids used in transfection

Plasmids used in the transfection process were maintained in *E. coli*, which were subsequently expanded and the DNA extracted. The *E. coli* was grown and harvested for a plasmid midi-preparation to isolate the plasmid DNA for transfection. Both the donor DNA plasmids and the gRNA/CAS9 plasmids contained sequences necessary for replication of the plasmid within *E. coli*, such as a high-copy-number ColE1/pMB1/pBR322/pUC origin of replication. Furthermore, a lac operator was present, which allows the lac repressor to bind and inhibit transcription of the plasmid within the *E. coli*. A gene coding for a β -lactamase is present

in each plasmid, which confers resistance to ampicillin. This gene allows for cells that contain the construct to survive in the presence of ampicillin, therefore only bacteria containing the plasmid will be able to grow while those that do not will die.

The ALT positive U2OS and SAOS2 cells were chosen for transfection. Both cell lines were transfected with a plasmid containing a coding sequence for the F-box protein osTIR1 (osTIR1-CMV-Puro) (pMK232; addgene ID: 72834; lab plasmid number: 214). This plasmid contains homology arms for insertion into the AAVS1 safe harbor locus between exon 1 and exon 2 and is inserted via the AAVS1 T2 CRISPR plasmid in pX330 (pX330; addgene ID:72833; lab plasmid number: 213). Once inserted, the osTIR1 gene is constitutively expressed using a cytomegalovirus (CMV) promoter as well as a CMV enhancer. A puromycin resistance cassette coding for puromycin N-acetyltransferase is also part of this plasmid, which allows for negative selection of successfully transfected cells (Figure 7). During transfection of this plasmid, the gRNA recognizes and binds to a 14bp segment in the AAVS1 locus between nucleotides 5482 and 5495, which is subsequently cut by CAS9 at nucleotide 5492. The new osTIR1 DNA donor construct is then inserted via homologous recombination.

After single cell cloning, U2OS-osTIR1 clone 3 cells were chosen for subsequent transfection of the SMC5-mAID construct. This construct was created using a plasmid containing mAID-Hygro (pMK287; addgene ID: 72825; lab plasmid number: 207). The donor plasmid for the SMC5-mAID targeting contains a coding sequence for the mAID protein located at the 3' end of the coding sequence for the SMC5 protein, adjacent to exon 25. Two guide RNA's specific for the C-terminus of the SMC5 locus, both designed using the pX330 plasmid (lab plasmid numbers: 221 and 222), were used for successful integration of the donor sequence. The endogenous promoter for SMC5 is used to drive transcription of the SMC5-mAID protein.

This donor plasmid also contains a coding sequence for an aminoglycoside phosphotransferase, which inactivates hygromycin and thus confers resistance. The expression of this cassette is driven by the mouse phosphoglycerate kinase 1 promoter, and the coding sequence is downstream of the coding sequence for the SMC5-mAID protein (Figure 8).

In addition to transfection of the SMC5-mAID, SMC6-mAID was also transfected into U2OS-osTIR1 Clone 3 cells. The donor plasmid used for this transfection was a kind gift from Dr. Andrew Holland's lab (lab plasmid number 227). Two guide RNA's specific for the C-terminus of the SMC6 locus, both designed using the pX330 plasmid (lab plasmid numbers: 225 and 226), were used for successful integration of the donor sequence. Three variations of the donor construct were used: an undigested, circular DNA donor construct, a single KPN1 digested linear donor construct, and a double KPN1 and FSE1 digested donor construct (Figure 9). The latter digestion resulted in a donor construct including the homology arms and what is located in between. This was isolated from the DNA remaining after digestion using a PCR cleanup kit.

The donor plasmid for the SMC6-mAID targeting contains a coding sequence for the mAID protein located at the 3' end of the coding sequence for the SMC6 protein. Directly adjacent to the mAID coding region is a FLAG tag coding sequence, which is subsequently bound to the SMC6-mAID protein after it is transcribed and translated. Downstream of the coding sequence for the SMC6-mAID protein, a coding sequence for an aminoglycoside phosphotransferase is present, which confers resistance to neomycin, kanamycin, and Geneticin (G418). Past the 3' homology arm for SMC6 is a coding sequence for another aminoglycoside phosphotransferase that confers resistance to Hygromycin. Past the origin of replication that is used to replicate this plasmid in *E. coli* is a coding region for β -lactamase, which confers

resistance to ampicillin and allows for negative selection during expansion of *E. coli* containing the plasmid (Figure 9).

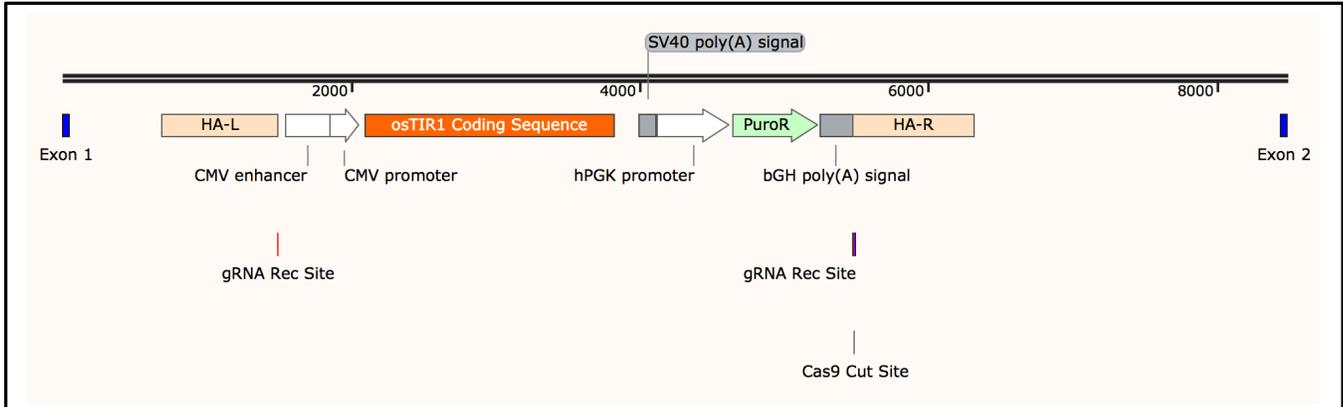


Figure 7. Donor construct for insertion of osTIR1 into the AAVS1 safe harbor locus. This construct includes the coding sequence for the F-Box protein osTIR1. Additionally, a puromycin resistance cassette is located within the sequence to be inserted, thus allowing for negative selection of successfully transfected cells.

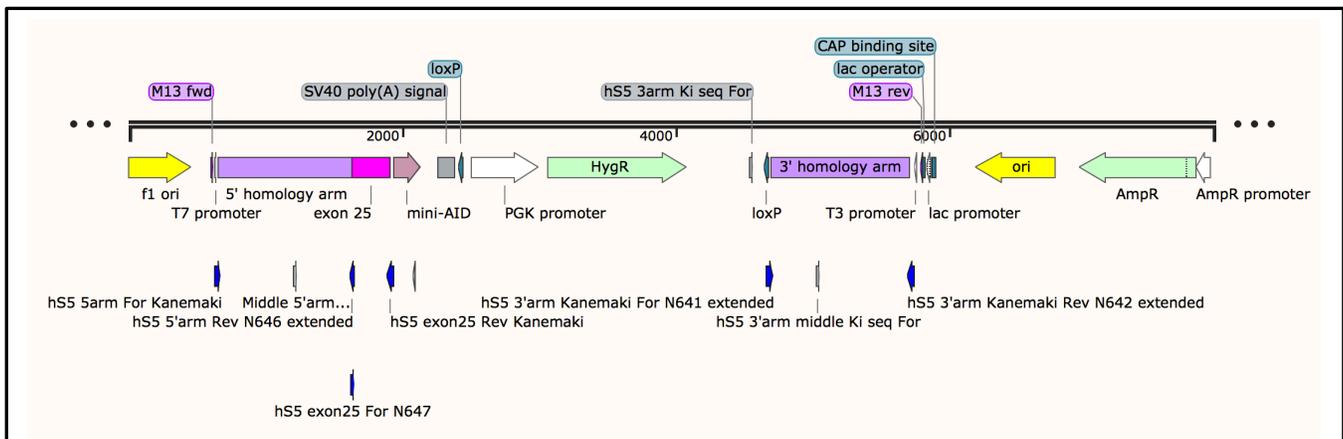


Figure 8. Donor construct for insertion of SMC5-mAID. This construct includes a coding sequence for the mAID protein located on the 3' end of the coding sequence for the SMC5 protein, directly adjacent to exon 25. This construct includes a hygromycin resistance cassette, which allows for negative selection of successfully transfected cells.

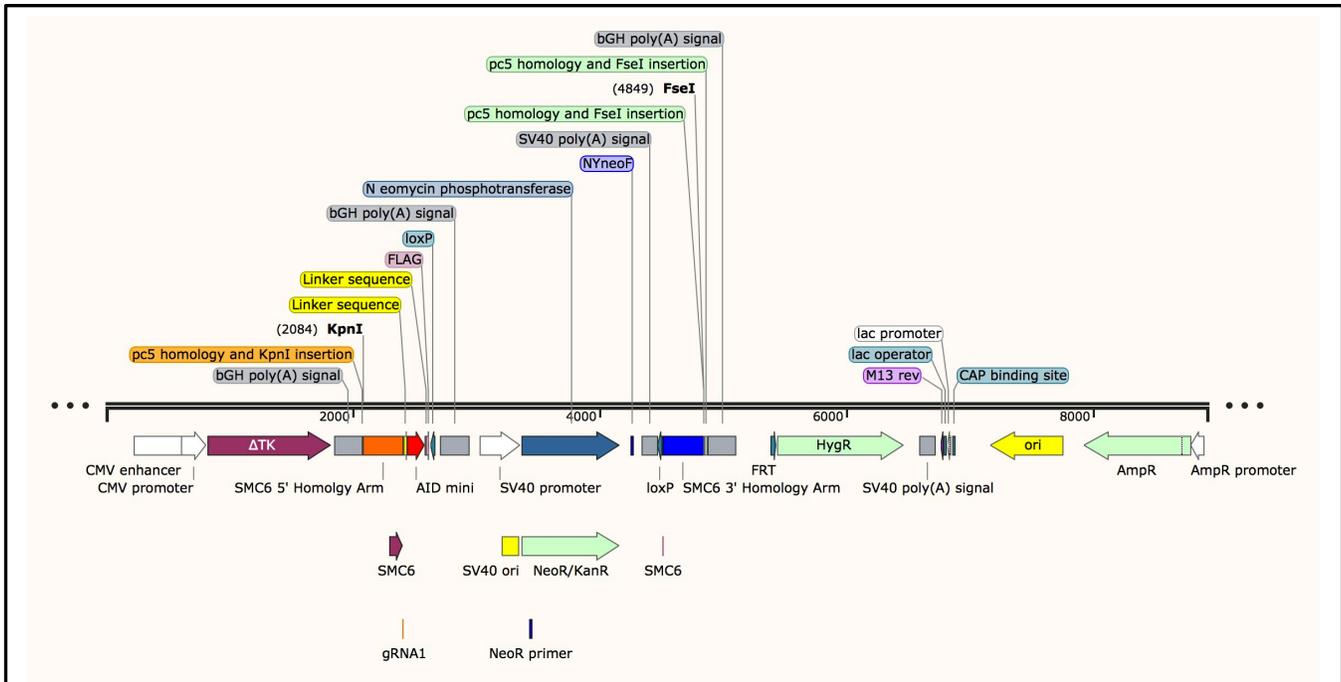


Figure 9. Donor construct for insertion of SMC6-mAID. This construct includes a coding sequence for the mAID protein located on the 3' end of the coding sequence for the SMC5 protein. This construct includes a neomycin resistance cassette, which allows for negative selection of successfully transfected cells using Geneticin (G418).

Transfection of the AID system

Cells were plated on 6-well plates 2 days before transfection, and allowed to grow to 60-70% confluence. Immediately before cells were transfected, media was aspirated from the cells and fresh CGM was added. Cells were transfected using the LipoJet in Vitro DNA and siRNA Transfection Kit (Ver. II) (SignaGen Laboratories). The liposomes used in this transfection kit are formulated from novel fluorinated cationic lipids, and this is one of the most efficient transfection kits on the market. A working solution of the LipoJet transfection buffer was prepared by diluting the 5X concentrated stock solution to 1X. For preparation of the LipoJet-

DNA complexes, a total of 1ug of DNA was added to 200uL of the 1X LipoJet transfection buffer for each well of the 6-well plate to be transfected. This 1ug of DNA included 0.3ug of the Cas9 and guide RNA(s) (gRNA) and 0.7ug of the donor DNA.

This mixture was vortexed briefly before 4uL of the LipoJet reagent, including the liposomes, was added. This solution was then incubated at room temperature for 10 minutes to allow the LipoJet-DNA complexes to form. 200uL of this solution was added to each well of the 6-well plate and swirled gently to homogenize the solution. Cells were then incubated at 37°C, 5% CO² for 24 hours. After 24 hours, media containing the transfection reagent was aspirated and fresh CGM was added. Subsequently, the appropriate antibiotics were administered to the media to negatively select for cells that were successfully transfected.

Single Cell Cloning and Selection

After cells were transfected with osTIR1 in 6-well plates, antibiotics were introduced into the media for negative selection. Cells were incubated in CGM supplemented with 1µg/mL of puromycin, and this antibiotic containing media was refreshed every other day. Puromycin was used as a selective antibiotic because the osTIR1 construct contains a gene coding for puromycin N-acetyltransferase, which inactivates puromycin within the cell. Thus, cells containing this construct will survive in the presence of puromycin, while those that do not will die. The optimal concentration of puromycin was determined by incubating wild type (WT) cells not transfected with the osTIR1 construct in 0.5µg/mL, 1µg/mL, and 1.5µg/mL puromycin until there was complete cell death. 1µg/mL puromycin was determined as being completely toxic to WT cells and there was complete cell death after 5 days of incubation, while 1.5µg/mL puromycin resulted in complete cell death after just 2 days. Thus, 1.5µg/mL puromycin was too toxic, and may result

in a lower yield of osTIR1 transfected cells during selection. As a control, WT cells were incubated with 1 µg/mL puromycin alongside osTIR1 transfected cells undergoing selection. After 5 days of incubation with 1 µg/mL puromycin, cells successfully transfected with the osTIR1 construct remained, while those not transfected were killed off. Cells were then transferred to a T-25 filtered culture flask and incubated with 1 µg/mL puromycin for an additional 3 days to be sure there were no un-transfected cells still present.

After cells transfected with osTIR1 were selected for using puromycin, single cell clones were obtained. Cells were collected using trypsin-EDTA detachment methods previously described. Cells were then collected by spinning them down at 200xg for 3 min and re-suspended in 5mL CGM without antibiotics. The concentration of cells was calculated using a hemocytometer and an aliquot of the cell suspension was prepared at a concentration of 1cell/100µL. This suspension was then plated into a 96-well plate using a multichannel pipette. Afterwards, the plates were incubated at 37°C, 5% CO² for 3-5 hours to allow cells to settle to the bottom of the plate. Each well of the 96-well plate was analyzed, and wells that had only one cell successfully plated in them were marked. The cell concentration of 1cell/100µL was used because each well of a 96-well plate holds 200µL. Thus, cells were plated at 0.5 cells/well, which provided the best chance that out of 96 wells, some wells would be successfully plated with only one cell in them.

Cells successfully plated at one cell per well were allowed to proliferate until they became sub-confluent. Cells were then collected using trypsin-EDTA detachment methods previously described. After quenching the trypsin, cells were immediately transferred to one well of a 24-well plate, and allowed to proliferate until they became sub-confluent. This process continued, transferring cells next to a 12-well plate, 6-well plate, and finally a T-25 filtered cell

culture flask. Interestingly, the SAOS2 cell line was not proliferative after being plated as single cell clones and cells would enter senescence after a few rounds of division, eventually undergoing apoptosis. It appears these cells need to be in groups to thrive, possibly due to growth factors secreted from neighboring cells. Thus, U2OS cells were used for single cell cloning and subsequent transfections. U2OS-osTIR1 single cell clones were harvested and their protein extracted following the methods previously described.

During expansion, the morphology and rate of proliferation were observed and compared to the WT U2OS cell line. Transfected U2OS single cell clones that most closely resembled the WT cell line during expansion and most robustly expressed the osTIR1 protein were chosen for subsequent transfection of mAID. U2OS Clone 3 was picked for subsequent transfection of the SMC5-mAID construct. Cells successfully transfected with the SMC5-mAID construct were selected for using hygromycin as a selective antibiotic, and the antibiotic containing media was refreshed every other day. The SMC5-mAID construct contains a coding region for an aminoglycoside phosphotransferase. This protein is an enzyme that catalyzes the transfer of a phosphate from ATP to an aminoglycoside such as hygromycin. Thus, this enzyme inactivates hygromycin, and cells successfully transfected will confer resistance to this antibiotic. The optimal concentration of hygromycin was determined by incubating U2OS-osTIR1 Clone 3 cells not transfected with the SMC5-mAID construct in 150 μ g/mL, 200 μ g/mL, and 250 μ g/mL hygromycin until there was complete cell death.

200 μ g/mL hygromycin was determined as being sufficiently toxic to U2OS-osTIR1 Clone 3 cells and there was complete cell death after 9 days of incubation, while 250 μ g/mL hygromycin resulted in complete cell death after just 6 days. Additionally, 150 μ g/mL hygromycin resulted in complete cell death after 11 days of incubation. As a control, U2OS-

osTIR1 Clone 3 cells were incubated with 200µg/mL hygromycin alongside SMC5-mAID transfected U2OS-osTIR1 Clone 3 cells undergoing selection. After 9 days of incubation with 200µg/mL hygromycin, cells successfully transfected with the osTIR1 construct remained, while those not transfected were killed off. Cells were then transferred to a T-25 filtered culture flask and incubated with 200µg/mL hygromycin for an additional 3 days to be sure there were no un-transfected cells still present. After successfully transfected cells were selected for using hygromycin, single cell clones were obtained using the same method described previously.

Successfully transfected U2OS-osTIR1 clone 3 cells were also used to transfect the SMC6-mAID construct. U2OS-osTIR1 cells successfully transfected with the SMC6-mAID construct were selected for using Genticin (G418). The SMC6-mAID construct contains a coding sequence for an aminoglycoside phosphotransferase, which inactivates G418. This aminoglycoside phosphotransferase also confers resistance to kanamycin and neomycin, and it is called a neo cassette. The optimal concentration of G418 was determined by incubating U2OS-osTIR1 Clone 3 cells not transfected with the SMC6-mAID construct in 200µg/mL, 400µg/mL, and 600µg/mL G418 until there was complete cell death. In wells containing 400µg/mL G418, cells died after 10 days, while cells incubated with 600µg/mL G418 died after 9 days.

Additionally, cells incubated with 200µg/mL G418 died after 11 days. 400µg/mL was chosen as the optimal concentration of G418 to use for selection. As a control, U2OS Clone 3 cells were incubated with 400µg/mL G418 alongside SMC6-mAID transfected U2OS Clone 3 cells undergoing selection. After 10 days of incubation with 400µg/mL G418, cells successfully transfected with the osTIR1 construct remained, while those not transfected were killed off. Cells were then transferred to a T-25 filtered culture flask and incubated with 400µg/mL G418 for an additional 3 days to be sure there were no un-transfected cells still present. After successfully

transfected cells were selected for using G418, single cell clones were obtained using the same methods described previously.

Results

RT-PCR and Q-PCR DATA

RT-PCR and Q-PCR data for the ALT (+) and ALT (-) osteosarcoma cell lines assessed is displayed below. Each tissue type has a unique profile; therefore, we chose to analyze only osteosarcoma cells to avoid discrepancies between tissue types. The gels displaying RT-PCR results using testis as a control are displayed below (Figure 10 & 11). Each lane number, primer sequence, and expected product size that correspond to the testis control gels in Figure 10 and 11 are displayed in table 5 and 6, respectively. The RT-PCR results for ALT positive cell lines (U2OS and SAOS2) and ALT negative cell lines (SJSA1 and MG63) are displayed in Figure 12 and 13, respectively. The Q-PCR data for the ALT positive cell lines (U2OS and SAOS2) and ALT negative cell lines (SJSA1 and MG63) are displayed in Figure 14 and Figure 15, respectively. The average threshold cycle (Ct) is plotted for each gene that was tested, and those genes that have no bars indicate that no mRNA was found. The threshold cycle is the cycle during the PCR reaction at which the quantity of DNA has reached a standard threshold that is the same for all samples. Thus, a higher threshold cycle corresponds to a higher number of PCR cycles required to reach the threshold, and a lower initial quantity of cDNA for the specific gene present. A lower concentration of mRNA in the cell for a specific gene of interest means there

will be less cDNA to start with during Q-PCR and it will take more PCR cycles to reach the threshold, resulting in a higher threshold cycle.

Genes that were found to not be expressed in the RT-PCR analysis correlate closely to genes not expressed in the Q-PCR analysis, with a few exceptions. In the RT-PCR results, it was observed that there was not a band corresponding to the correct band size for the amplification products for genes *SYCP2L2* and *MSH4* in the ALT (+) U2OS cell line; however, there was mRNA found for these genes in the U2OS cell line according to the Q-PCR results. Interestingly, these two genes had the highest Ct values out of all genes found to be expressed in the U2OS cell line, meaning they had the lowest level of expression. This was also true for the genes *HORMAD2*, *TEX11*, *TEX12*, and *MAJIN* in the ALT (-) cell line MG63. There was not a band corresponding to the correct band size for the amplification products for these genes; however, they were indicated as being expressed in the Q-PCR results.

There were similarities and differences in which genes were observed in the Q-PCR data as being expressed between the ALT (+) and ALT (-) cells (Table 8). There was a larger group of genes not expressed in both ALT (+) cell lines analyzed than there were in the ALT (-) cell lines; thus, there was more meiotic gene expression in ALT (-) cells. Interestingly, all genes not expressed in the ALT (-) cell lines assessed are also not expressed in the ALT (+) cell lines assessed, except for *TERBI*. The ALT (+) U2OS cell line we assessed expressed *TERBI*, while both ALT (-) cell lines we assessed do not. Within the two ALT (+) cell lines we assessed, some genes were expressed in one cell line but not expressed in the other. In the two ALT (+) cell lines we assessed, there were 5 genes that were expressed in the U2OS cell line, but not expressed in the SAOS2 cell line: *MEI1*, *HORMAD1*, *SYCE1*, *TEX12*, and *TERBI*. Furthermore, there was only one gene that was expressed in the SAOS2 cell line but not expressed in the U2OS cell line:

MEIOB. These discrepancies were also true between the two ALT (-) cell lines we assessed. There were 4 genes that were expressed in the SJSA1 cell line but not in the MG63 cell line: *RAD21L1*, *STAG3*, *SYCE3*, and *SYCP2L*. Additionally, there were 5 genes that were expressed in the MG63 cell line but not in the SJSA1 cell line: *HORMAD2*, *HFM1*, *TEX11*, *TEX12*, and *MAJIN*. There were more discrepancies between the Alt (-) cell lines than there were between the ALT (+) cell lines, and the U2OS cell line expressed almost all the genes that were expressed in the SJSA1 cell line.

The Q-PCR data provides relative levels of expression of each of the genes found to be expressed in the ALT (+) and ALT (-) cell lines we assessed. Because each round of PCR amplification results in twice as much DNA as was present in the last round, a reaction that has a threshold cycle of 2 cycles larger than another indicates there is 4 times less initial cDNA in that reaction. Thus, there is 4 times lower expression of that gene and an exponential relationship exists between the two reactions. Of the genes we found were expressed, there was varied levels of expression between all 4 cells assessed. Between the two ALT (+) cell lines, genes that were expressed in both cell lines were expressed in similar amounts, with one major exception – *SYCP2L* was expressed in much higher quantities in the SAOS2 cell line than it was in the U2OS cell line. The threshold cycle for *SYCP2L* in the SAOS2 cell line was 31.39, while the threshold cycle in the U2OS cell line was 36.82. This indicates there was roughly 30 times the amount of starting cDNA (and therefore mRNA) for the transcription of *SYCP2L* in the SAOS2 cell line compared to the U2OS cell line.

Similarly, between the two ALT (-) cell lines, genes that were expressed in both cell lines were expressed in similar amounts, with one major exception - *MEIOB* was expressed in much higher quantities in the MG63 cell line than it was in the SJSA1 cell line. The threshold cycle for

MEIOB in the MG63 cell line was 31.61, while the threshold cycle in the SJSA1 cell line was 38.07. This indicates there was roughly 42 times the amount of starting cDNA (and therefore mRNA) for the transcription of *MEIOB* in the MG63 cell line compared to the SJSA1 cell line. Additionally, *DMC1* was expressed in larger quantities in the MG63 cell line than it was in the SJSA1 cell line (about 10.5 times more). Genes that were found to be expressed in either one or both ALT (+) and ALT (-) cell lines, respectively, there were very similar levels of expression of each gene between the ALT (+) and ALT (-) phenotype. However, there were two genes that stuck out that were both expressed in higher amounts in one or both ALT (+) cell lines. *HORMAD1* was expressed about 25 time more in the ALT (+) U2OS cell line than it was expressed in both the ALT (-) MG63 and SJSA1 cell lines. Additionally, *MEIOB* was expressed about 35 times more in the ALT (+) SAOS2 cell line than the average level expression between both ALT (-) MG63 and SJSA1 cell lines.

Table 5. Primer sequence and expected product band size for Figure 10. (Note: For = forward primer sequence, Rev = reverse primer sequence)

Lane #	Gene	Primer Sequence	Product Size (bp)
1 For	<i>SPO11</i>	ACAGAGCAACACTTATGCAACC	219
1 Rev		GCACCACAGGTACAATTCCT	
2 For	<i>DMC1</i>	AGAAACATGGAATTAACGTGGCT	185
2 Rev		AAATGCAGTCAAGAATCCTGGTT	
3 For	<i>MEI1</i>	GGTCACGCAACTGGTGTCTC	163
3 Rev		GCAACGGATAGTCTGCTCCA	
4 For	<i>MND1</i>	TGTGAGAGGATCGGAACCTTCT	163
4 Rev		CACATCGGCCAATTTTAGCTTTC	
5 For	<i>RAD21L1</i>	AACCAAAGCAGACCAGAAGAAA	174
5 Rev		GAGGCTTCCAGAACTATGTTCAA	
6 For	<i>SMC3</i>	AACATAATGTGATTGTGGGCAGA	244
6 Rev		TCCTTTTGGCACCAATAACTCT	
7 For	<i>REC8</i>	TCCGCGTCTATTCTCAACAATG	172
7 Rev		GGATCTGGAGCATCTTCTAGGG	
8 For	<i>STAG3</i>	GCAAAACGACCACCGAAAACA	162
8 Rev		CCACTCATCTACCAAAGACTGC	

Table 5 (continued). Primer sequence and expected product band size for Figure 10

9 For	<i>HOP2</i>	CTCCCAGGATGTGTTCCGGG	285
9 Rev		GTGGTCAGGGCACTAGATAATTC	
10 For	<i>HORMAD1</i>	AGCAACGAATCTAGCATGTTGT	185
10 Rev		TCACCATCCTTAAAACCGGGA	
11 For	<i>HORMAD2</i>	CTGCATCACAATACACAAGGCT	164
11 Rev		AATGGCGTTCTCCATAAGAGC	
12 For	<i>SYCP1</i>	CCCTTTGCATTGTTCTGTACCA	87
12 Rev		GAAAGTGGAATCGCCTCCCA	
13 For	<i>SYCP2</i>	AGCTGCAAATACCATCAGATGAA	131
13 Rev		CTCTGGCACAGTAACTGCTTC	
14 For	<i>SYCP3</i>	TATCCAGGAAATCTGGGAAGCC	232
14 Rev		GAGCCTTGTAATGTCAACTCCA	
15 For	<i>SYCE1</i>	AGGTTGGCATTGAGGAACAG	150
15 Rev		CGCCTTGACCAGTTCTCT	
16 For	<i>SYCE2</i>	GTCGGGACTCTACTTCTCCTC	83
16 Rev		TGTTGATGTTTTCGATCAGCTCC	
17 For	<i>SYCE3</i>	ACAACATGCTGAAAATGCTGTC	228
17 Rev		GCCTTTGCTTGGTCTCATGC	
18 For	<i>SYCP 2L</i>	AGATGAACCTCTGCTAATTCGGC	221
18 Rev		TCACCAGGAATCCTAAGCTAAGT	
19 For	<i>FKBP6</i>	TTCTGTTCAAACCGAACTACGC	134
19 Rev		GAGAGCACAAAACCTGTCTGACT	
20 For	<i>HFM1</i>	TTTGCCTGCTCACCTAGTAGT	165
20 Rev		TCCCTTGTGCTTAATCGAGTCA	
21 For	<i>MSH5</i>	TGGCAGGTTCTCTACAAGACT	135
21 Rev		GAGGCTGGCGATATGGTGC	
22 For	<i>MSH4</i>	TTCAGCACTGTCTAATGGAGG	155
22 Rev		TCTATCATGGCTGTCTGTTCACT	
23 For	<i>TEX11</i>	CTGCCAGTAGTTTTGAGGTACAA	166
23 Rev		GCCTCTTTGGCCTTATCAAGTTG	
24 For	<i>TEX12</i>	ATGGCAAATCACCTGTAAAGCC	85
24 Rev		GCTGTGGACTATCTGGCACTG	
25 For	<i>RNF212</i>	CTGGGTGTTCTGTAATCGCTG	150
25 Rev		AAAACGTACGACAAGGAGCTTT	
26 For	<i>CCNB1IP1</i>	CGTGTTGGACATTAGCTCCCG	250
26 Rev		GCGCTCCATAAGTTTCTCAGAG	
27 For	<i>MNS1</i>	TGCGGCAACGTGAAGATTTG	175
27 Rev		GCAGCCTGTAGACTAATTCC	
28 For	<i>MEIOB</i>	CCTGCAACTCTAGCAACTGT	198
28 Rev		TGCAAGCACGTTAATAATCCTCC	
29 For	<i>SPATA22</i>	GAGAGGGCTAGACAAAAACAGTG	204
29 Rev		TGTTACGCCAATACTTCATGC	
30 For	<i>TERB1</i>	CCCACGTAGAAGACAACGACT	186
30 Rev		TCCACAGCCTCCGTCCTT	
31 For	<i>MEIKIN W1</i>	GGCTTTTGCTTTCGCTCGAT	121
31 Rev		GTGCCTCTGGACTIONCTCGATG	
32 For	<i>MEIKIN W2</i>	CGGCTTGTGGAAGATTGCAG	441
32 Rev		TGTGGTGCCTTCTCTATCGC	
33 For	<i>MEIKIN W3</i>	ACTCCCTGTGACTTAGCCCT	441
33 Rev		CAATCTTCGACAAGCCGTGC	

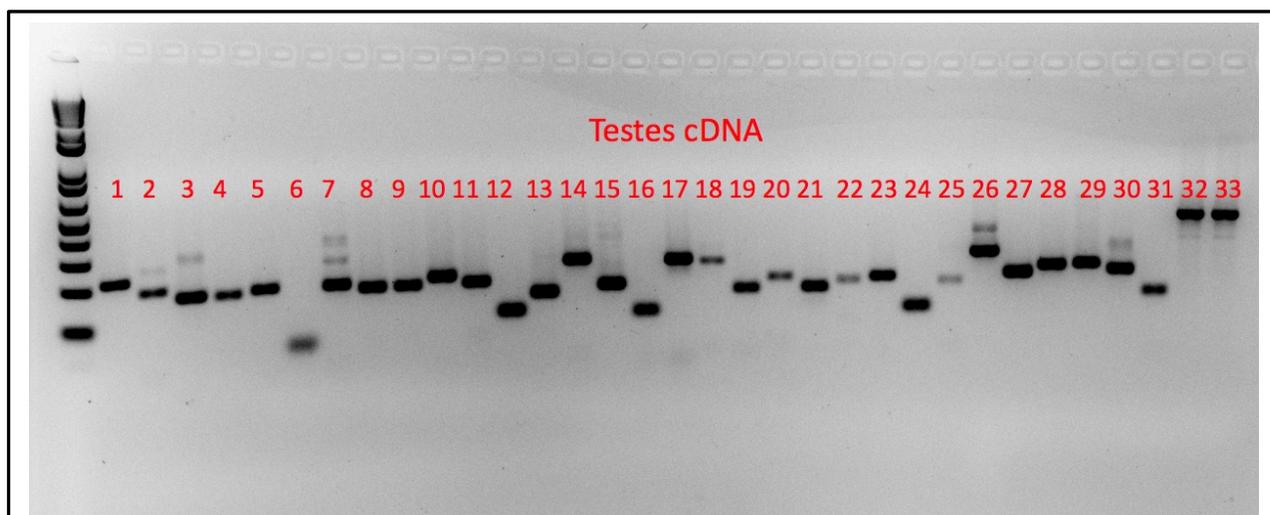


Figure 10. RT-PCR data for a group of meiotic genes corresponding to Table 5.

Table 6. Primer sequence and expected product band size for Figure 11. (Note: For = forward primer sequence, Rev = reverse primer sequence)

Lane #	Gene	Primer Sequence	Product Size (bp)
1 For	<i>SMC3</i>	AACATAATGTGATTGTGGGCAGA	244
1 Rev		TCCTTTTGGCACCAATAACTCT	
2 For	<i>SMC3 W1</i>	CTGGCCCGTGCTTTCACAT	255
2 Rev		CCTTGCTGGGTCTCGATCT	
3 For	<i>SMC3 W2</i>	CACATGCGTGGAAGTCACTG	338
3 Rev		GGCTGACTTGGTCACCTTCC	
4 For	<i>SMC3 W3</i>	GCTAGACCACTCCGTCGAA	400
4 Rev		CATAGTGAAAGCACGGCCA	
5 For	<i>TERB2 W1</i>	CTGCGTGCCTGCAAAAAGAA	180
5 Rev		AATGCTTTTCTGGGCTTTGGA	244
6 For	<i>TERB2 W2</i>	TCATTCTTCTCCTGCGTGC	160
6 Rev		TTGTGCTCTGTTGCTAGTTCAC	
7 For	<i>TERB2 W3</i>	GTGAACTAGCAACAGAGCACA	179
7 Rev		GTCATGTAGCTCCCAAGGA	
8 For	<i>MAJIN W1</i>	GAGCTGGAGGATTCTGTCCG	343
8 Rev		TCCCTATTCTGTCCAGCCCT	
9 For	<i>MAJIN W2</i>	AGGGCTGGACAGAATAGGGA	305
9 Rev		GAAGAATGGCTCGGGAGGAG	
10 For	<i>MAJIN W3</i>	TGGGAGAGAGTTTCCACCT	247
10 Rev		TCCCTATTCTGTCCAGCCCT	

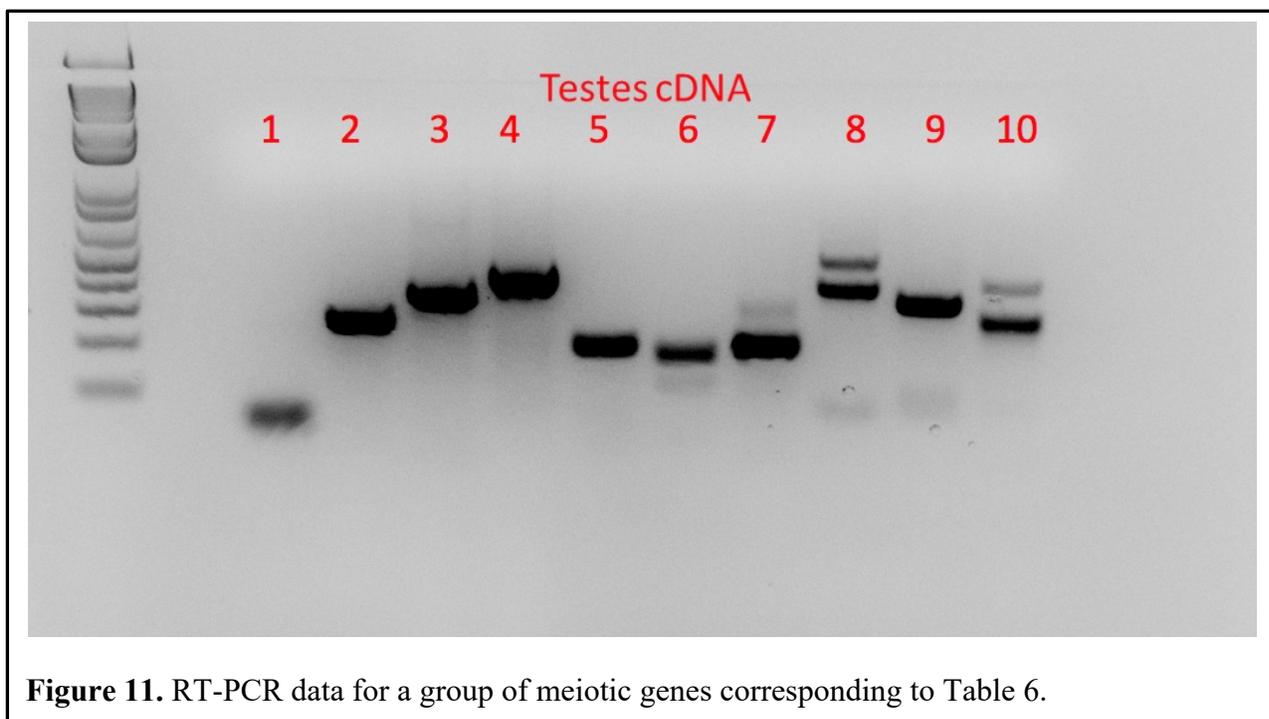


Table 7. Meiotic genes of interest used in RT-PCR and Q-PCR analyses. For U2OS, SAOS2, and SJSA1, the last two lanes that are not numbered correspond to TBP used as a control.

1) <i>SPO11</i>	10) <i>HORMAD1</i>	19) <i>FKBP6</i>	28) <i>MEIOB</i>
2) <i>DMC1</i>	11) <i>HORMAD2</i>	20) <i>HFM1</i>	29) <i>SPATA22</i>
3) <i>MEI1</i>	12) <i>SYCP1</i>	21) <i>MSH5</i>	30) <i>TERB1</i>
4) <i>MND1</i>	13) <i>SYCP2</i>	22) <i>MSH4</i>	31) <i>MEIKIN</i>
5) <i>RAD21L1</i>	14) <i>SYCP3</i>	23) <i>TEX11</i>	32) <i>TERB2</i>
6) <i>SMC3</i>	15) <i>SYCE1</i>	24) <i>TEX12</i>	33) <i>MAJIN</i>
7) <i>REC8</i>	16) <i>SYCE2</i>	25) <i>RNF212</i>	34) Water Control
8) <i>STAG3</i>	17) <i>SYCE3</i>	26) <i>CCNB1IP1</i>	
9) <i>HOP2</i>	18) <i>SYCP2L</i>	27) <i>MSNI</i>	

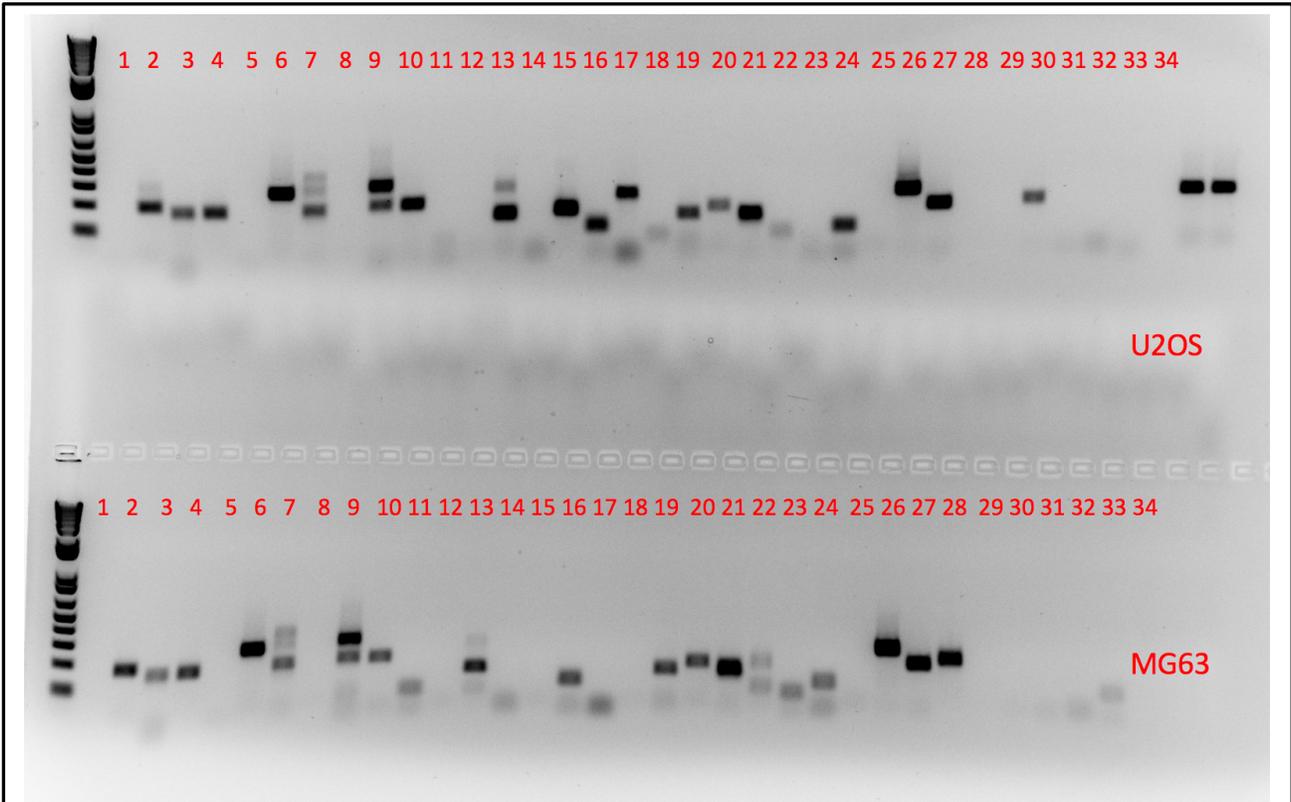


Figure 12. RT-PCR results for the osteosarcoma cell lines U2OS (ALT+) and MG63 (ALT-). Each lane corresponds to a different meiotic gene of interest surveyed for (Table 7). The numbers for the lanes align with the numbers for each gene of interest in Table 7.

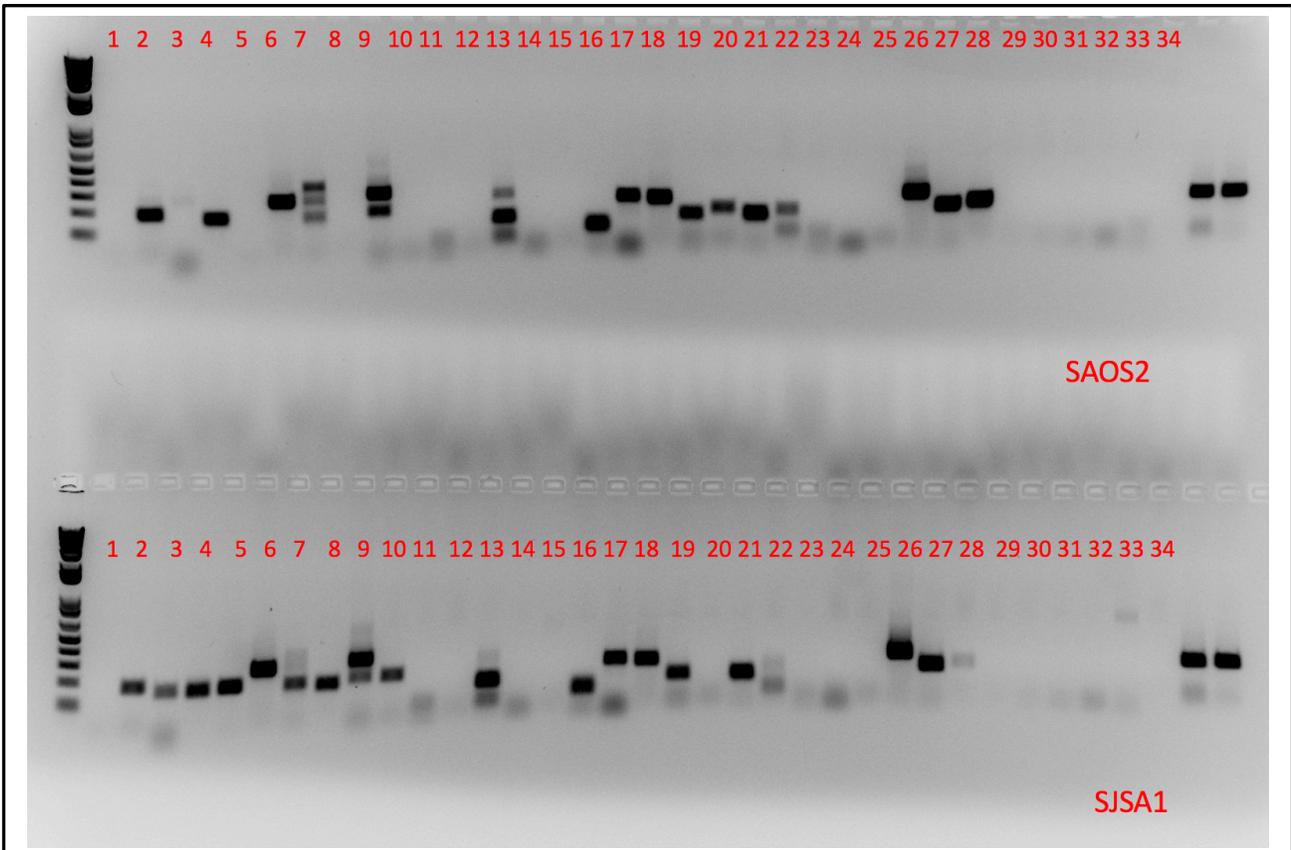


Figure 13. RT-PCR results for the osteosarcoma cell lines SAOS2 (ALT+) and SJSA1 (ALT-). Each lane represents a different meiotic gene of interest surveyed for (Table 7). The numbers for the lanes align with the numbers for each gene of interest in Table 7.

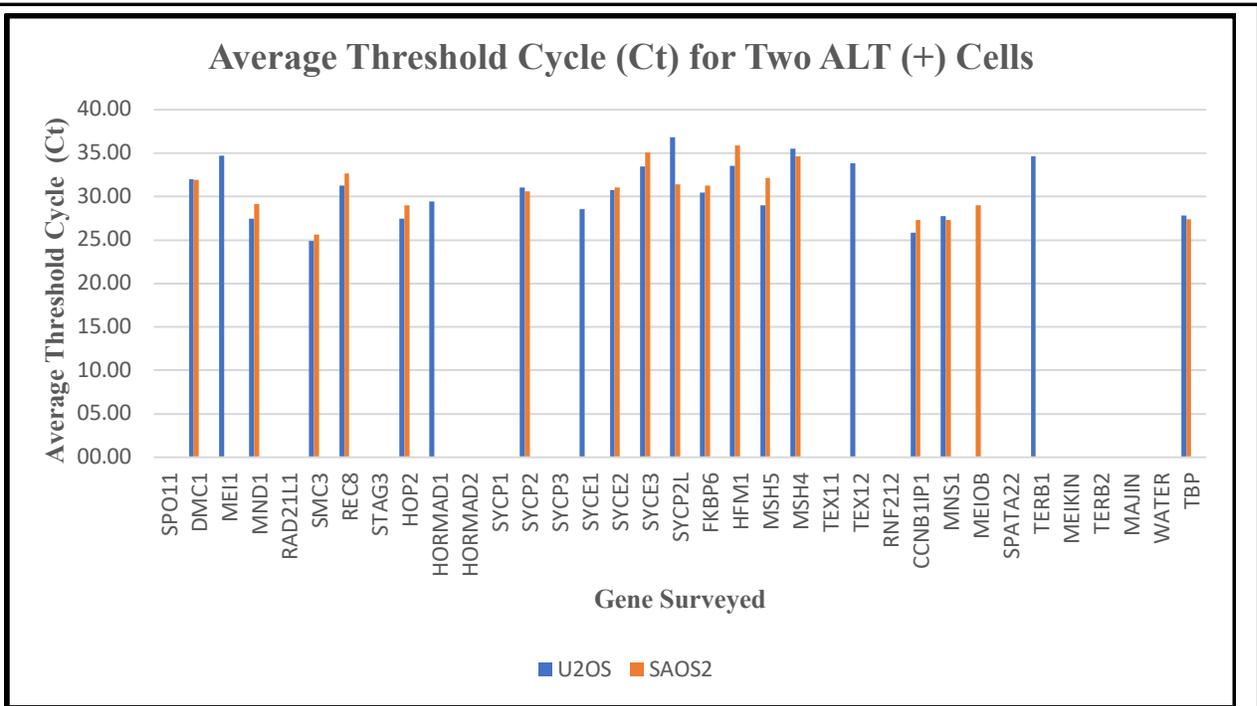


Figure 14. Q-PCR data for a group of meiotic genes surveyed for in two ALT (+) cells (U2OS and SAOS2). The average threshold cycle (Ct) is presented for each gene. The larger the average threshold, the lower concentration of mRNA for each gene was found present in each cell.

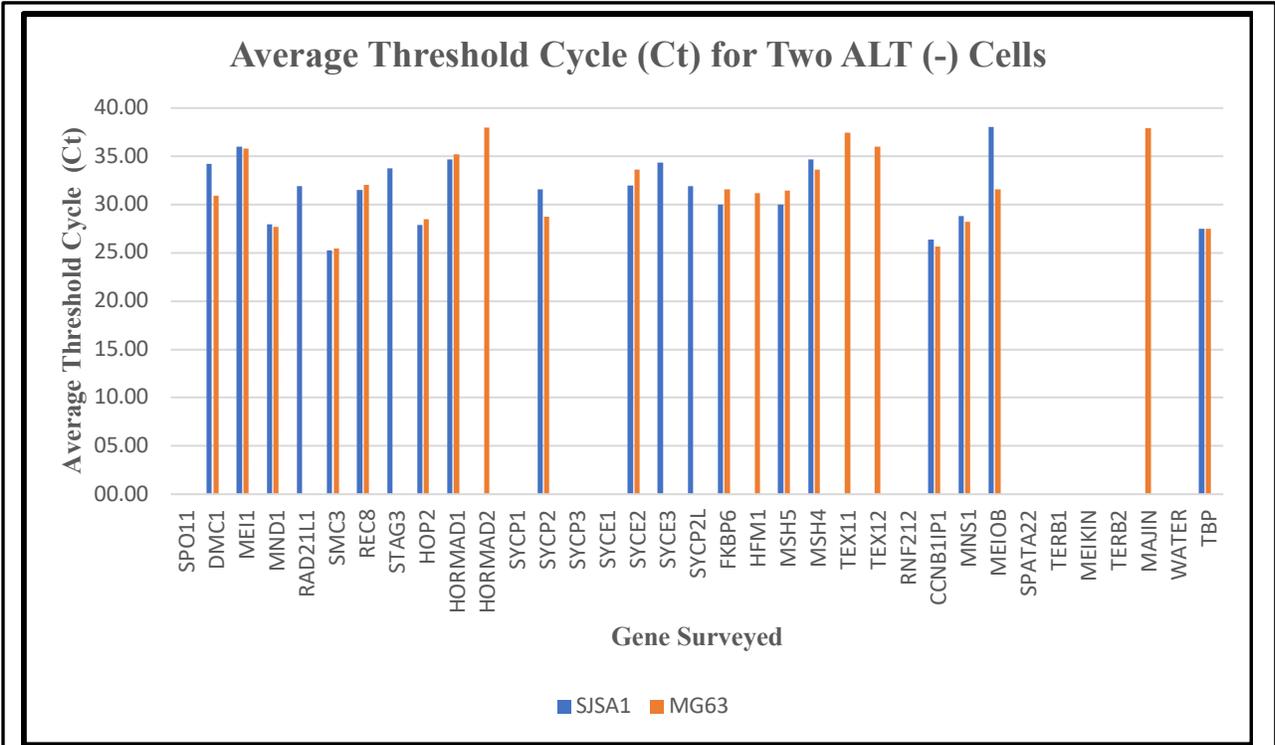


Figure 15. Q-PCR data for a group of meiotic genes surveyed for in two ALT (-) cells (SJS1 and MG63). The average threshold cycle (Ct) is presented for each gene. The larger the average threshold, the lower concentration of mRNA for each gene was found present in each cell.

Table 8. Genes not expressed in both ALT (+), both ALT (-) or not expressed in any of the 4 cell lines assessed.

*Note: Yellow – not expressed in all cell lines assessed; Green – not expressed in only ALT (+) cell lines assessed; Blue – not expressed in only ALT (-) cell lines assessed

Not Expressed in ALT (+)	Not Expressed in ALT (-)	Not Expressed in all four cell lines
<i>SPO11</i>	<i>SPO11</i>	<i>SPO11</i>
<i>SYCP3</i>	<i>SYCP3</i>	<i>SYCP3</i>
<i>RNF212</i>	<i>RNF212</i>	<i>RNF212</i>
<i>SPATA22</i>	<i>SPATA22</i>	<i>SPATA22</i>
<i>MEIKIN</i>	<i>MEIKIN</i>	<i>MEIKIN</i>
<i>TERB2</i>	<i>TERB2</i>	<i>TERB2</i>
<i>RAD21L1</i>	<i>TERB1</i>	
<i>STAG3</i>		
<i>HORMAD2</i>		
<i>SYCP1</i>		
<i>TEX11</i>		
<i>MAJIN</i>		

Western Blot Analysis

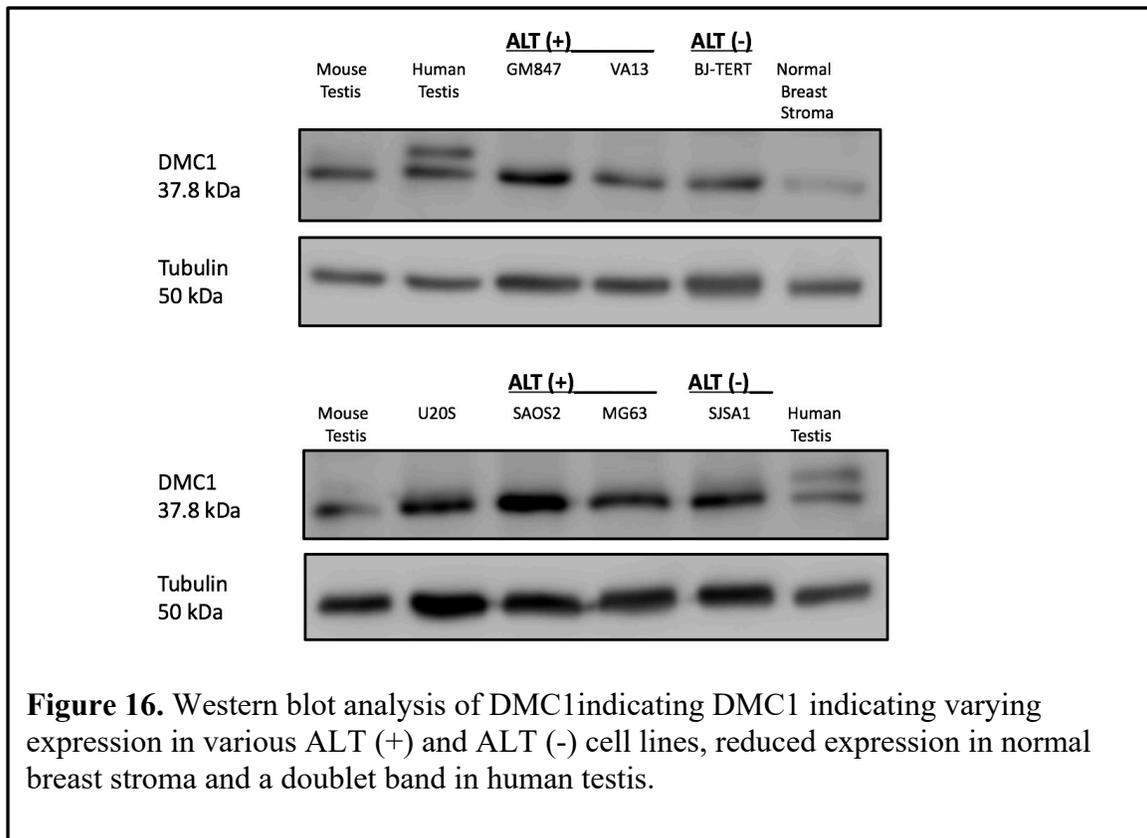
Various western blots were performed to assess the level of expression of meiotic genes of interest (Table 9). Each western blot is displayed below (Figures 16-31) in the order they appear in Table 9. Many of these genes that were originally thought to be meiosis-specific (*DMC1*, *HFMI*, *HOP2*, *SYCE2*) were identified as being expressed at various concentrations in both ALT (+) and ALT (-) cancer cell lines. There does not seem to be a preference for these genes being expression in ALT (+) versus ALT (-) cell lines, rather, there seems to be varying expression in both ALT (+) and ALT (-) cell lines.

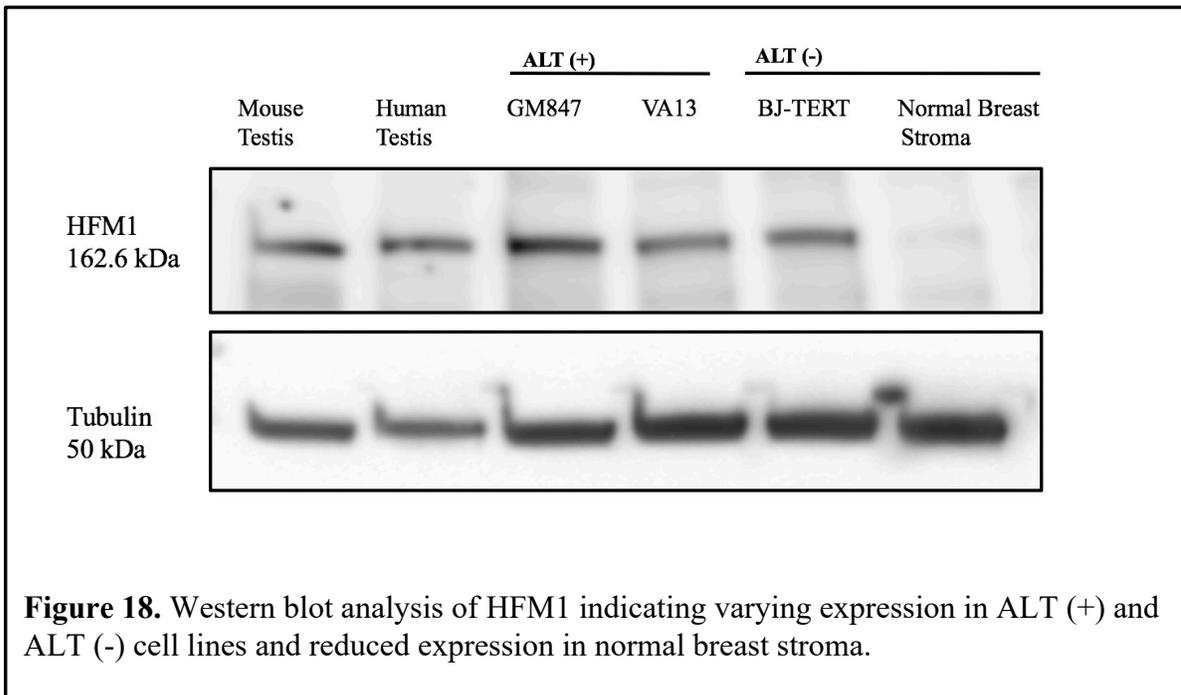
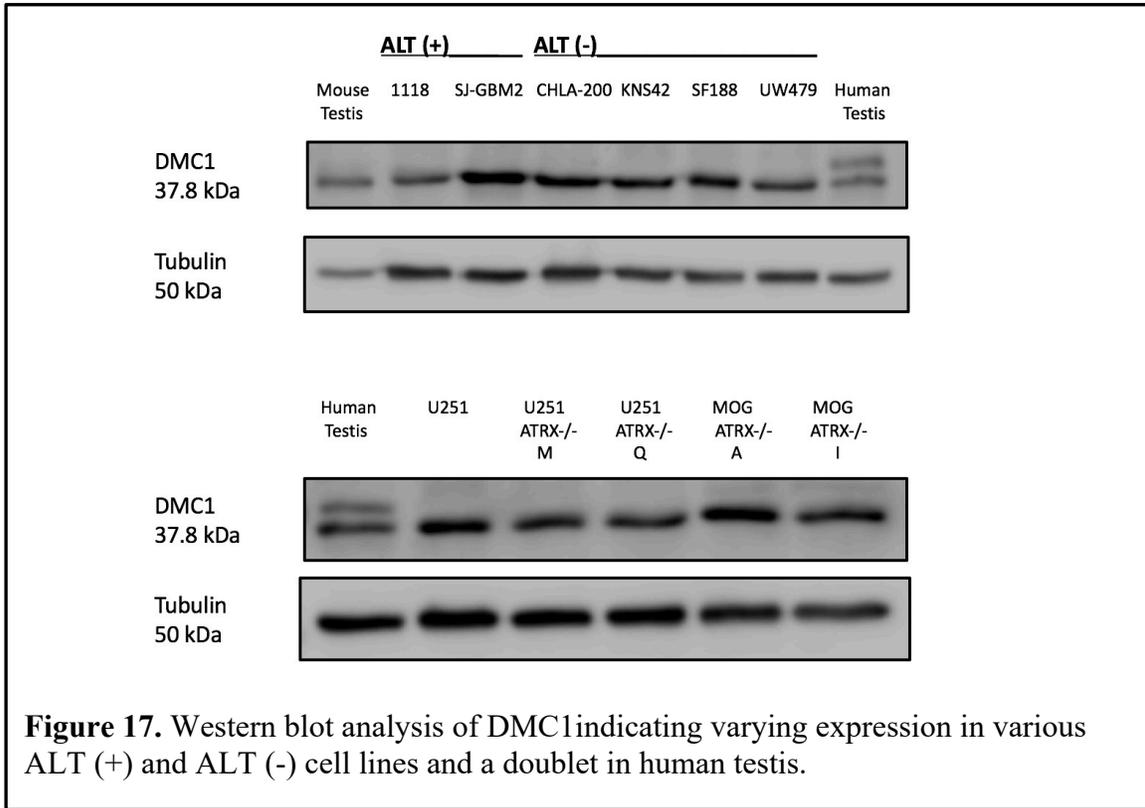
Interestingly, *HFMI* was identified as having reduced expression in the non-immortalized normal breast stromal cells. Originally, we thought this may have something to do with the non-immortalized nature of this particular cell line. Thus, we analyzed three patient

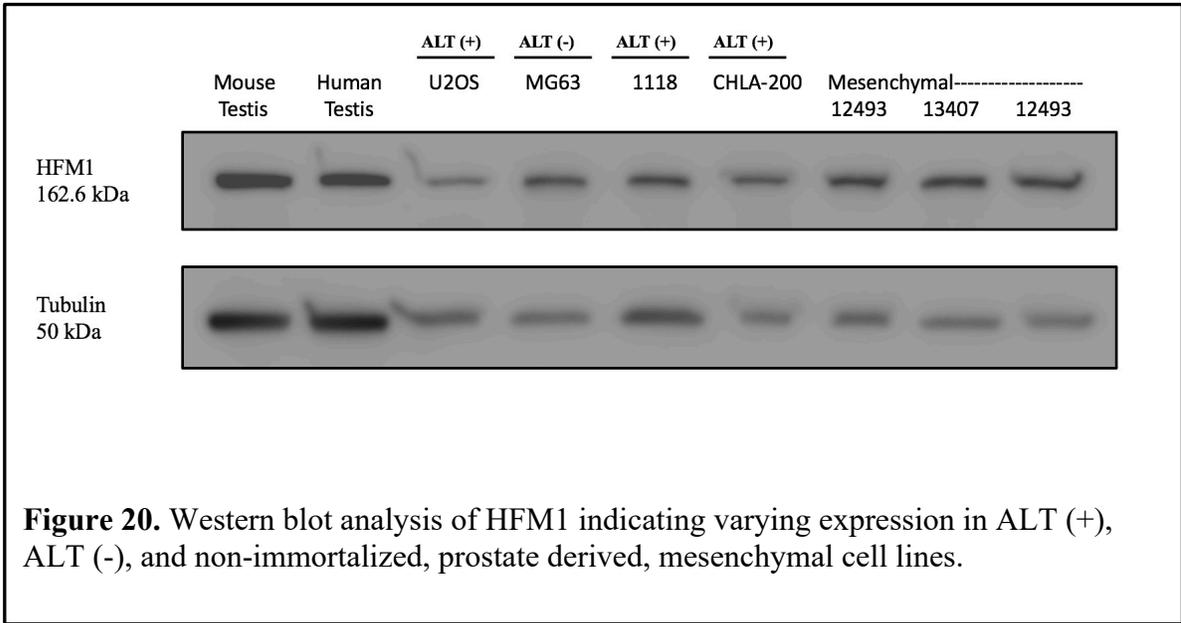
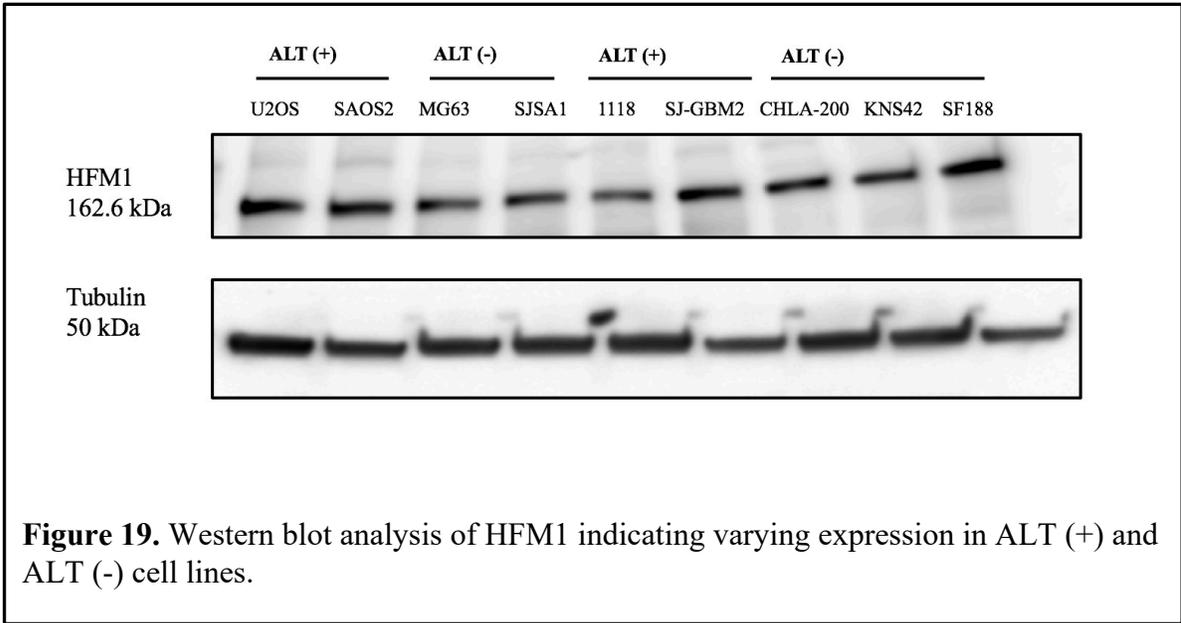
derived samples of non-immortalized prostate derived mesenchymal cells (Figure 20). This result displayed expression of *HFMI* in all three samples; therefore, we concluded that the decreased expression of *HFMI* in normal breast stroma was not due to the non-immortalized nature of this cell line. Normal breast stroma is a specialized form of tissue that regulates the proliferation, differentiation, and survival of the mammary gland (Arendt *et al.*, 2010). Therefore, the downregulation of *HFMI* in normal breast stroma may be due, in part, to the specific function of this tissue. Furthermore, *DMCI* was also identified as having reduced expression in normal breast stroma. *DMCI* is a meiosis-specific member of the recA-like gene family of recombinases that stabilizes strand exchange intermediates during homologous recombination, while *HFMI* is an ATP-dependent helicase required to form crossovers and complete synapsis of homologous chromosomes (Wang *et al.*, 2014). Both *DMCI* and *HFMI* are meiotic specific genes necessary for homologous recombination; thus, the fact that they are both downregulated in normal breast stroma alludes to a functional explanation for this.

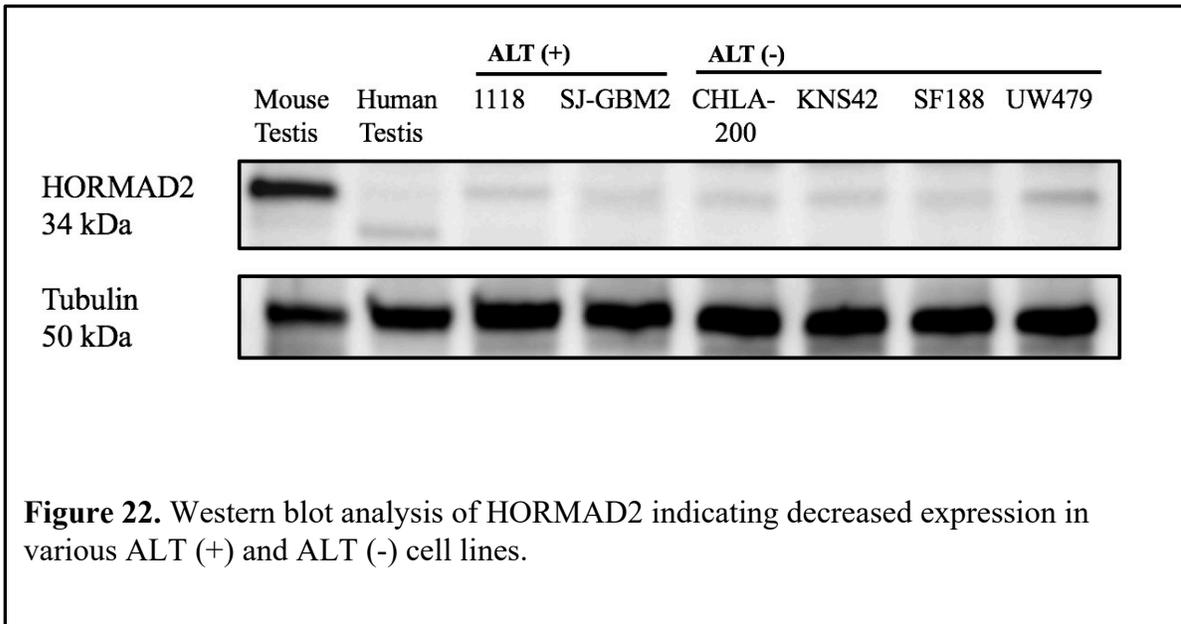
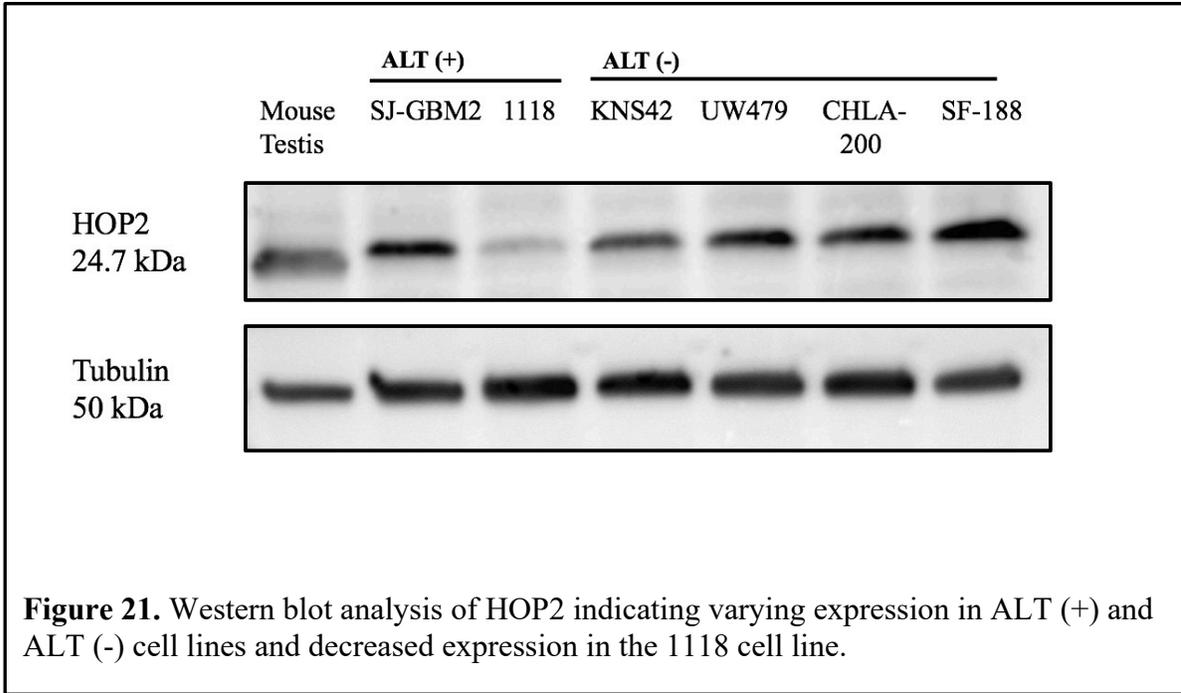
Table 9. Proteins assayed for expression in western blots

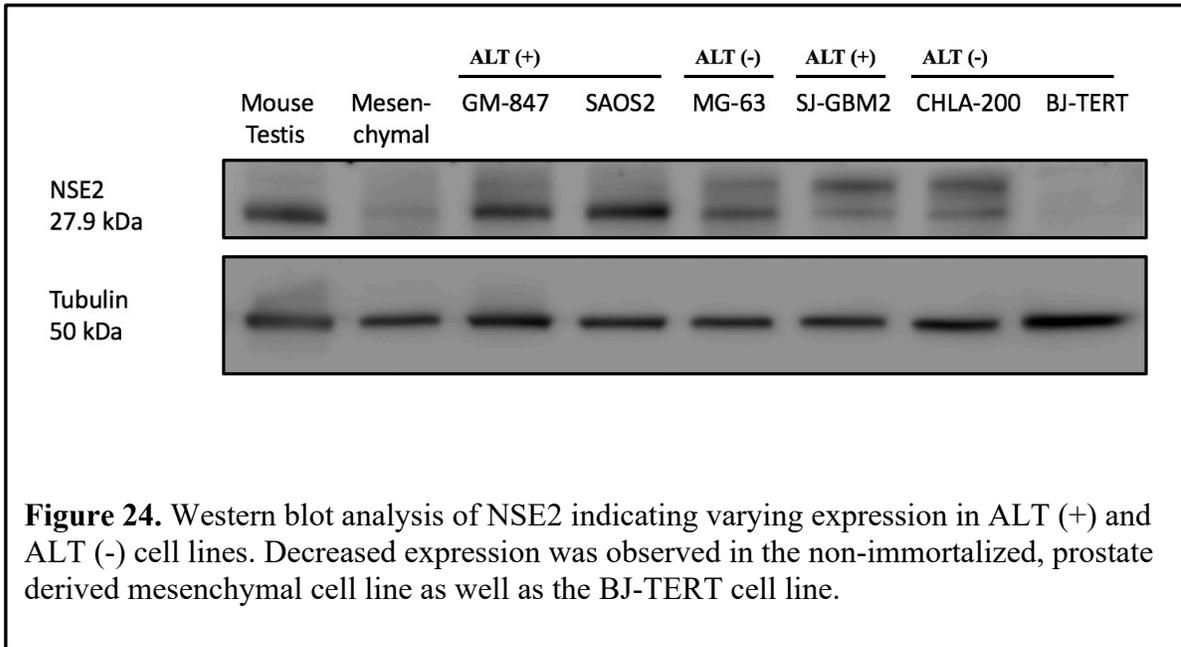
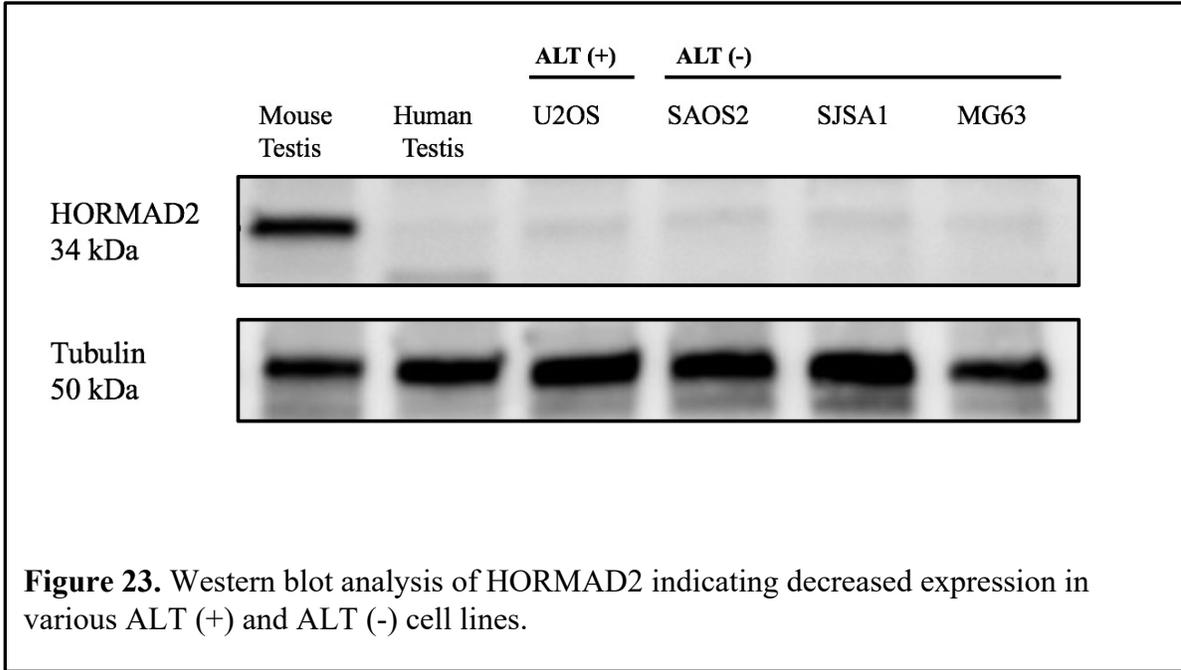
1) DMC1	3) HOP2	5) NSE2	7) REC8	9) SMC6
2) HFM1	4) HORMAD2	6) NSE4a	8) SMC5	10) SYCE2











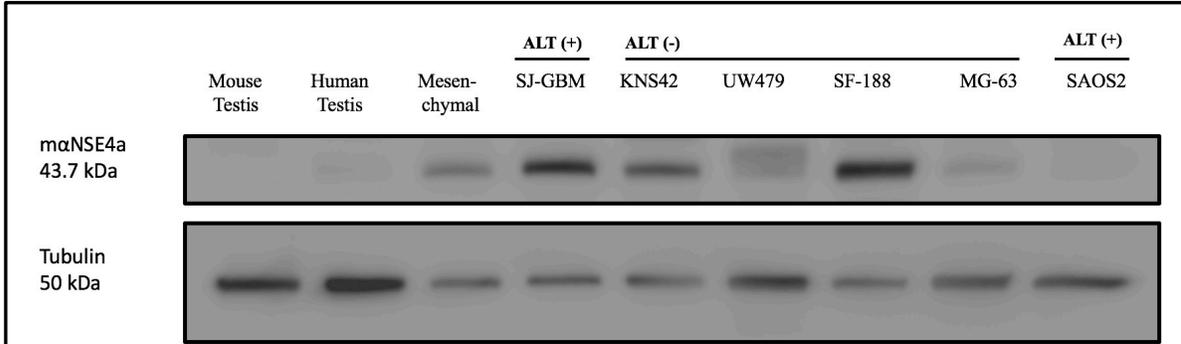


Figure 25. Western blot analysis of NSE4a indicating varying expression in ALT (+) and ALT (-) cell lines. Indicated is decreased expression in non-immortalized, prostate derived mesenchymal cells, UW479, MG63, and SAOS2.

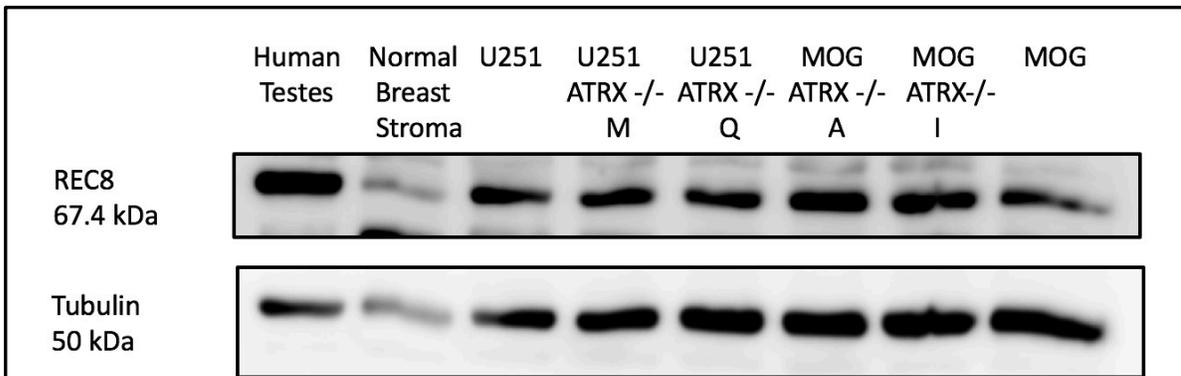
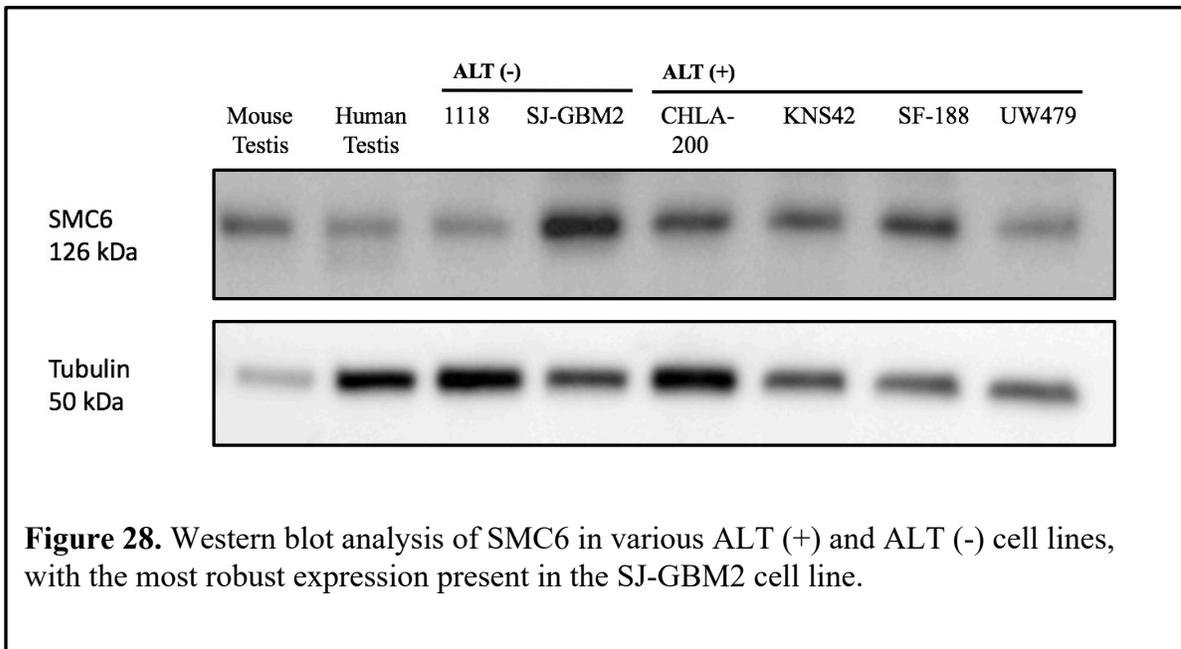
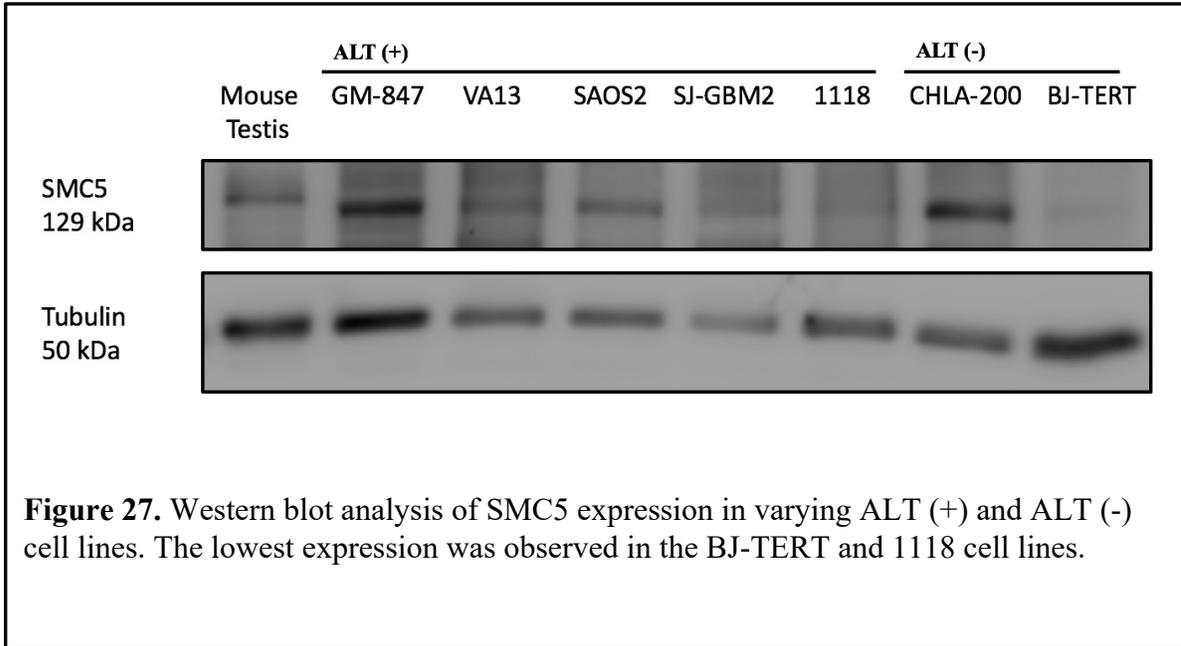
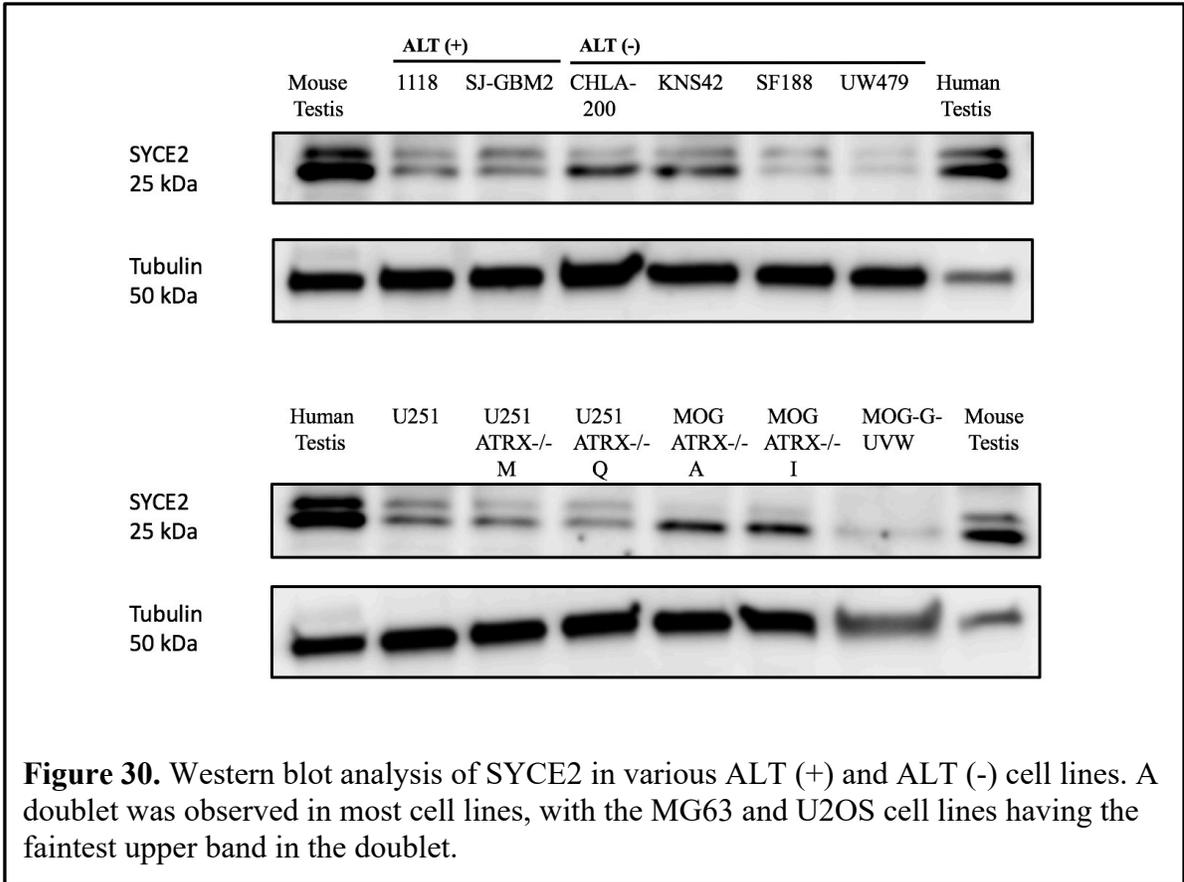
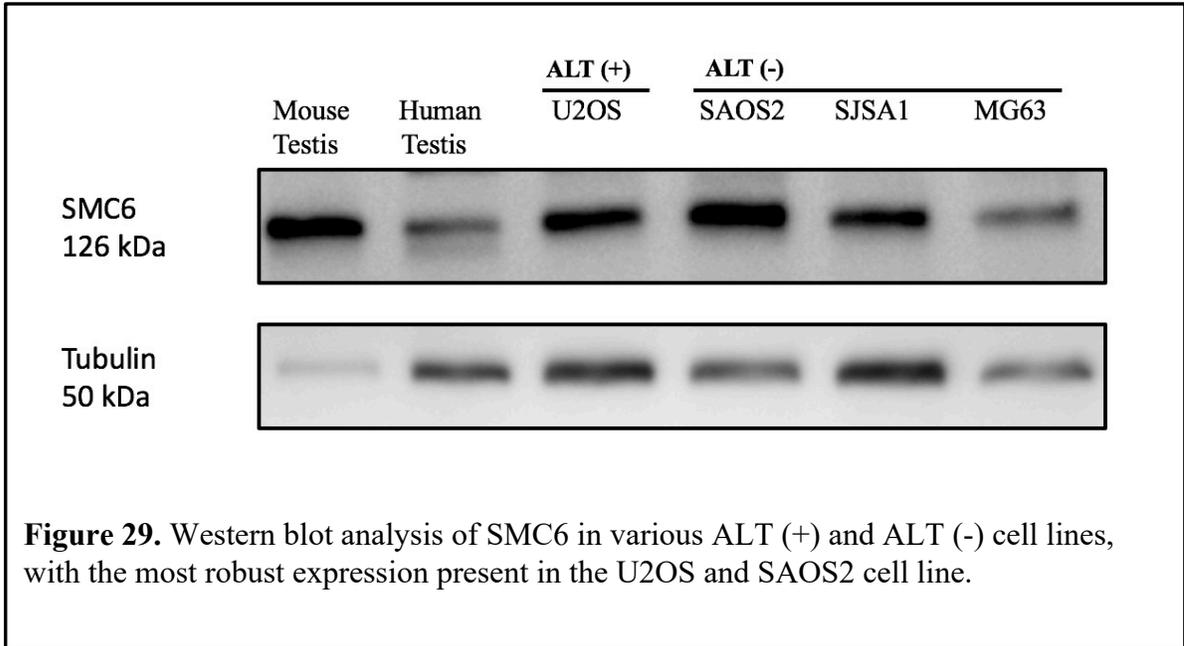


Figure 26. Western blot analysis of REC8 indicating expression in U251 and MOG ATRX isogenic cell lines as well as normal breast stromal cells.





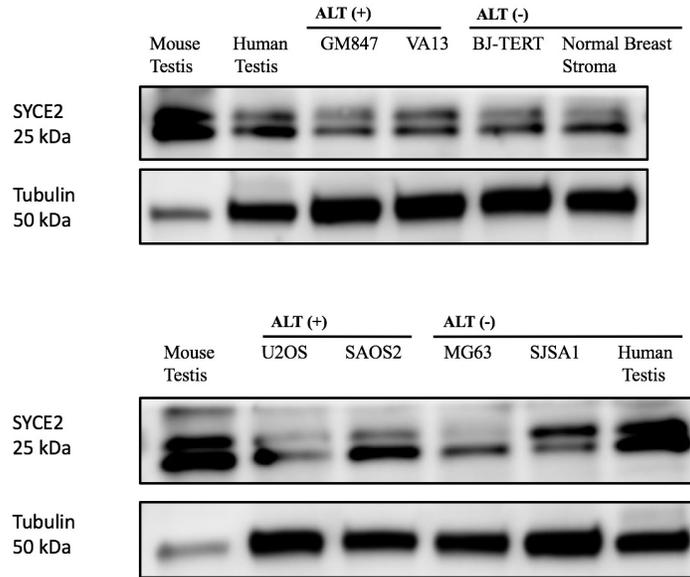


Figure 31. Western blot analysis of SYCE2 in various ALT (+) and ALT (-) cell lines. A doublet was observed in most cell lines, with the lowest expression observed in the UW479 cell line.

Analysis of successful transfection of AID system

After the U2OS and SAOS2 cell lines were transfected with osTIR1 and subsequently selected for using puromycin, the heterogeneous population was tested for successful insertion of the donor construct by harvesting the cells and running a western blot using an antibody specific for the osTIR1 protein. If the cells were successfully transfected, the *osTIR1* gene will be constitutively expressed under the CMV promoter, and the osTIR1 protein will be present in the cells. As a negative control, WT cells that were not transfected with the osTIR1 construct were harvested and ran alongside the transfected cells. Figure 32 indicates a band is present corresponding to the correct size for the osTIR1 protein in the heterogeneous population of U2OS cells, while there is no band present in the WT lane. Thus, it can be concluded that there are cells present in this heterogeneous population expressing osTIR1. In Figure 33, the same analysis for the heterogeneous population of osTIR1 transfected SAOS2 cells is presented. Likewise, there is a band present corresponding to the correct size for the osTIR1 protein, while there is no band present in the WT lane; thus, it can be concluded that there are cells present in this heterogeneous population expressing osTIR1.

After single cell cloning was performed on the heterogeneous population of osTIR1 transfected U2OS cells, clones were analyzed for successful incorporation of the osTIR1 construct. Cells were harvested and a western blot was performed using an antibody specific for the osTIR1 protein. The WT U2OS cell line that was not transfected with the osTIR1 construct was used as a negative control, while the osTIR1-transfected heterogeneous population of U2OS cells was used as a positive control (Figure 32). The results indicate a band present corresponding to the correct size for the osTIR1 protein in the osTIR1-transfected heterogeneous population of U2OS cells, while there was no band present in the WT U2OS cells not transfected

with the osTIR1 construct. Furthermore, there is a band present for Clone 3 isolated from the osTIR1-transfected heterogeneous U2OS cell population, indicating this clone was isolated from a parent cell that was successfully transfected with the osTIR1 construct.

The band intensity for Clone 3 appears to be stronger than the heterogeneous population, which is expected because the single cell clone population was derived from one cell that is expressing the osTIR1 construct. In the heterogeneous population, there may be cells that randomly incorporated the puromycin resistance portion of the osTIR1 construct; therefore, they survive during puromycin selection, but they do not express the *osTIR1* gene. Because Clone 3 comes from one single cell, the same quantity of protein will contain a larger concentration of osTIR1 in it because almost all cells are expressing the *osTIR1* gene.

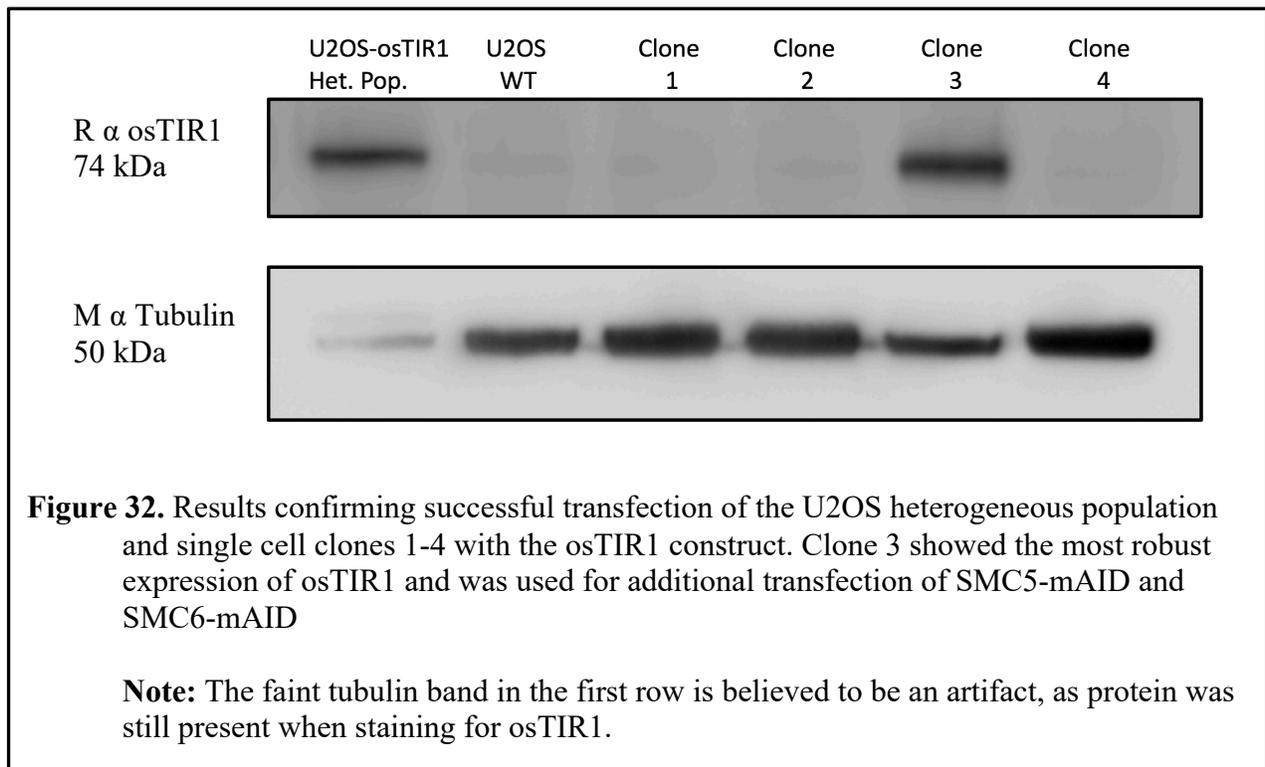
Clone 3, the isolated osTIR1-expressing U2OS cell line, was transfected with the SMC5-mAID construct and subsequently selected for using hygromycin, single cell clones were obtained. These single cell clones were analyzed for successful incorporation of the SMC5-mAID construct. Cells were harvested and a western blot performed using an antibody specific for the mAID tag. A DLD-1 cell line transfected with the osTIR1 and an SMC6-mAID constructs was used as a positive control. This DLD-1 cell line was previously confirmed as being successfully transfected with both the osTIR1 and SMC6-mAID constructs during an experiment in which auxin was added and the depletion of SMC6 was observed.

In the results presented below (Figure 34), a band appears where the expected size of the SMC6 protein (126 kDa) should be located for the DLD1 cell line, which was used as a positive control. However, there does not appear to be a band located where the expected size of the SMC5 protein (129 kDa) would be. Therefore, it does not appear that the SMC5-mAID is being expressed (or lack thereof) in a way that would allow for the antibody to specific for mAID to

bind to it and produce a band. These cells survived under hygromycin selection; thus, this construct could have been randomly inserted or inserted in a way that did inhibited expression of the SMC5-mAID gene, such as the construct being inserted in a way that obstructed the endogenous promotor for the *SMC5* gene. Furthermore, adding the tag to the 3' end of the coding sequence for the *SMC5* gene could have affected its processing or folding, resulting in suppressed expression. To better understand the details of this result, it would be advantageous to run a western blot using an antibody specific for SMC5 and see if there is expression of *SMC5* in the cell at all. Additionally, primers designed to amplify through this insert using PCR analysis should be employed to confirm successful integration of the SMC5-mAID construct on the nucleotide level.

Additionally, Clone 3, the isolated osTIR1-expressing U2OS cell line, was transfected with the three variations of the SMC6-mAID construct (undigested, single KPN1 digestion, and double KPN1 and FSE1 digestion) and subsequently selected for using G418, single cell clones were obtained. These single cell clones were analyzed for successful incorporation of the SMC6-mAID construct. Cells were harvested and a western blot performed using an antibody specific for the mAID tag. A DLD-1 cell line transfected with the osTIR1 and an SMC6-mAID constructs was used as a positive control. This DLD-1 cell line was previously confirmed as being successfully transfected with both the osTIR1 and SMC6-mAID constructs during an experiment in which auxin was added and the depletion of SMC6 was observed. In the results presented below (Figure 35), a band appears where the expected size of the SMC6 protein (126 kDa) should be located for the DLD1 cell line, which was used as a positive control. However, there does not appear to be a band located where the expected size of the SMC6 protein would be

in the undigested, single, or double digested constructs. The same explanations as described for the SMC5-mAID negative result apply to this result as well.



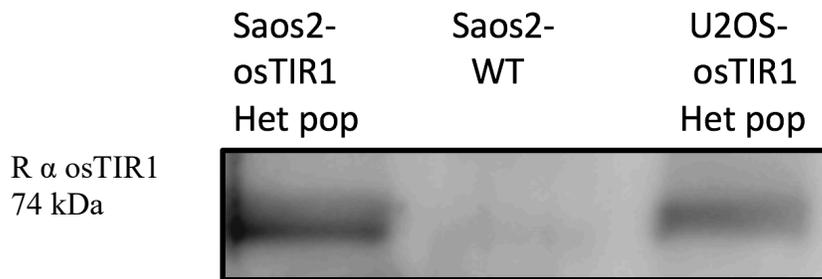


Figure 33. Results confirming successful transfection of SAOS2 heterogeneous population of cells with the osTIR1 construct. The U2OS-osTIR1 heterogeneous population was used as a positive control, while the WT SAOS2 cells were used as a negative control since they have not been transfected with the osTIR1 construct.

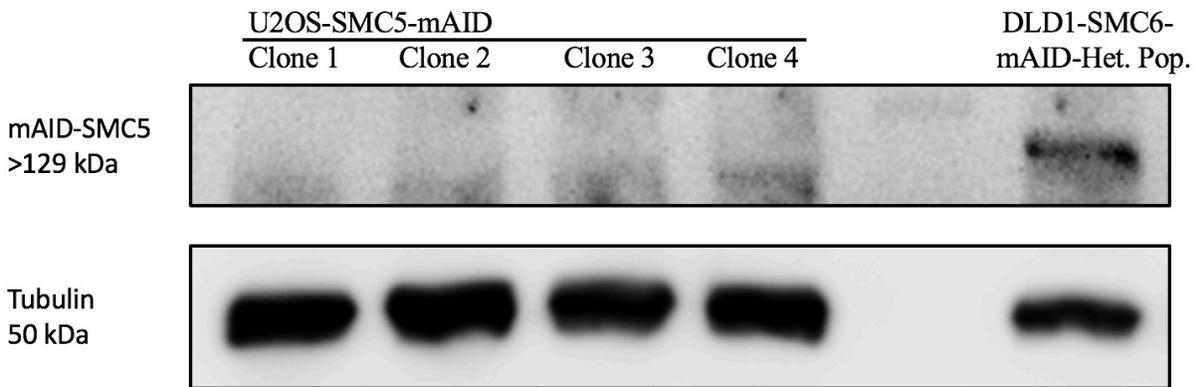


Figure 34. Western blot results (using an antibody specific for the mAID protein tag) for single cell clones of U2OS-osTIR1(clone 3)-SMC5-mAID cells. The expected band size should be a little larger than 129 kDa because the antibody specific for the mAID will bind the mAID protein, which is tagged to the SMC5 protein. The DLD1 cells that were used as a positive control indicate the mAID protein tag is present and appears around the same band size as SMC6 (126 kDa). It does not appear that the mAID protein is present where the expected band for the SMC5 protein would be.

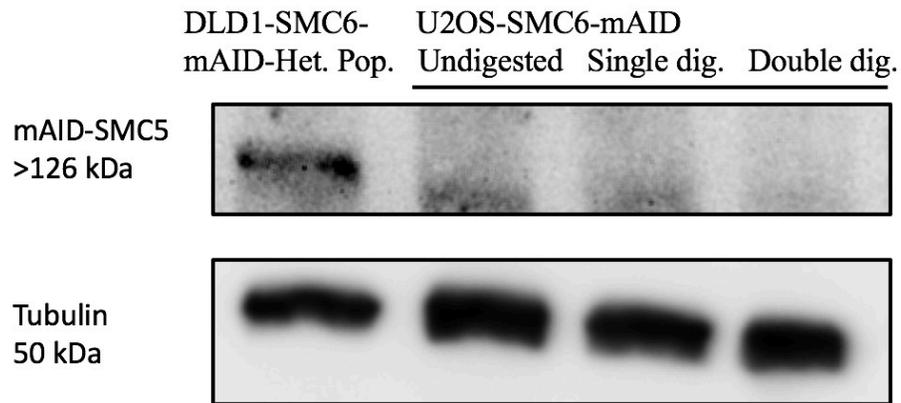


Figure 35. Western blot results (using an antibody specific for the mAID protein tag) for single cell clones of U2OS-osTIR1(clone 3)-SMC6-mAID cells. Presented are results for the undigested, single, and double digested construct. The expected band size should be a little larger than 126 kDa because the antibody specific for the mAID will bind the mAID protein, which is tagged to the SMC6 protein. The DLD1 cells that were used as a positive control indicate the mAID protein tag is present and appears around the same band size as SMC6. It does not appear that the mAID protein is present where the expected band for the SMC6 protein would be.

Discussion

In the present research, a group of meiotic genes were identified and assessed that may be playing an important role in cancer. These genes pose an opportunity for therapeutic intervention due to their lack of expression in somatic tissues, other than the gonads. Some of the processes these genes are involved in, including homologous recombination and maintaining genomic structure and integrity, have been identified as playing a role in the alternative lengthening of telomeres pathway (Deng *et al.*, 2008; Cesare and Reddel, 2010). The ALT pathway utilizes many of the same processes that are important for meiosis, and cells that follow this pathway extend their telomeres in the absence of telomerase as a method of successfully evading apoptosis and overcoming replicative mortality. In this study, expression of the group of meiotic genes identified were tested for in both ALT (+) and ALT (-) cells. These analyses helped to better understand which meiotic genes may be important for the ALT pathway.

We assessed the expression of a group of meiotic genes of interest in two ALT (+) and ALT (-) osteosarcoma cancer cell lines through RT-PCR and Q-PCR analysis. It was important to use osteosarcoma in all four cell lines because each tissue type has a unique profile. Interestingly, there were more meiotic genes found to be expressed in the ALT (-) cell lines than the ALT (+) cell lines assessed. Furthermore, all genes that were not expressed in the ALT (-) cell lines assessed were also not expressed in the ALT (+) cell lines assessed, with one exception – *TERB1* was expressed in the ATL (+) U2OS cell line but not expressed in either of the ALT (-) cell lines we assessed. In the future, it would be advantageous to assess *TERB1* expression in these various ALT (+) and ALT (-) cell lines using a western blot analysis. TERB1, a meiosis-specific protein, is a component of the MAJIN-TERB1-TERB2 complex. This complex associates with telomeric proteins and promotes telomere cap exchange by helping to attach

telomeric DNA to the inner nuclear membrane and replace the protective cap of telomeric chromosomes during meiosis (UniProt). It is possible that this protein is interacting with telomeres in the ALT (+) U2OS cell line to extend telomeres by facilitating homolog search within the cell. Additional research assessing the interaction of this complex with APB's would be advantageous, since APBs have been identified as facilitating homolog search with telomeres in ALT (+) cells (Muntoni and Reddel, 2005; Deng *et al.*, 2008).

There were differences between the two ALT (+) cell lines we assessed, and some genes were expressed in one cell line but not in the other. It is possible that genes expressed in one ALT (+) cell line but not the other are not essential to the ALT mechanism. If a gene is essential to the ALT mechanism, then it is more likely that it would be expressed in both ALT (+) cell lines. There were 5 genes expressed in the ALT (+) U2OS cell line that were not expressed in the ALT (+) SAOS2 cell line. Additionally, there was only one gene that was expressed in the SAOS2 cell line but not expressed in the U2OS cell line. It appears that the U2OS cell line expresses many genes that may not be expressed in other ALT (+) and ALT (-) cell lines. The U2OS cell line was harvested from a 14-year-old female in 1964 (ATCC). As subsequent divisions have occurred, the cell may have become more genetically heterogeneous in nature. This could explain why there are so many meiotic genes expressed in the U2OS cell line that are not expressed in other ALT (+) or ALT (-) cell lines. Focusing on those genes that are expressed in U2OS but also expressed in other ALT (+) cell lines could help better characterize which genes are essential for the ALT mechanism.

Between the genes that were expressed in both ALT (+) cell lines assessed, *SYCP2L* was expressed about 30 time more in the SAOS2 cell line than it was in the U2OS cell line. The SAOS2 cell line could rely on *SYCP2L* as a means of extending its telomeres in the absence of

telomerase more than the U2OS cell line. SYCP2L is a part of the synaptonemal complex and plays an important part in homologous chromosome pairing during meiosis. Genes involved in homologous recombination have been identified as playing a role in the ALT mechanism (Kraus *et al.*, 2001; Roumelioti *et al.*, 2016). Therefore, this gene could be upregulated to play a part in the ALT mechanism. Although this gene is expressed in much higher amounts in the SAOS2 cell line compared to the U2OS cell line, it could be playing a part in the ALT mechanism as it is still expressed in both ALT positive cell lines we assessed.

Additionally, there were differences between the two ALT (-) cell lines we assessed. There were 4 meiotic genes expressed in the SJSA1 cell line but not in the MG63 cell line, and there were 5 meiotic genes expressed in the MG63 cell line but not in the SJSA1 cell line. There were more genes that differed between the two ALT (-) cell lines than those that differed between the two ALT (+) cell lines we assessed. Many of these meiotic genes expressed in one ALT (-) cell line but not the other, such as *SYCE3*, *SYCP2L*, *TEX11*, and *TEX12*, are involved in cross-over events during meiosis. It is possible that homologous recombination is important for some cells that do not follow the ALT mechanism for telomere extension. However, it does not seem that it is essential, as these genes were expressed in one ALT (-) cell line assessed, but not the other. Additionally, the ALT (+) U2OS cell line expressed almost all the genes that were expressed in the SJSA1 cell line. Thus, more research should be done on these meiotic genes to better understand which are essential for the ALT mechanism, and which are transiently expressed between both ALT (+) and ALT (-) cell lines.

The western blot analysis performed in this study indicated that there was varied expression of a group of meiotic genes in both ALT (+) and ALT (-) cell lines. Thus, it is possible that these meiotic genes play an important part in the development of cancer in general,

and do not favor either the ALT (+) or ALT (-) mechanism of telomere elongation. Many of the meiotic genes assessed in our western blot analysis play a part in meiotic recombination.

Although recombination is an important part of the ALT mechanism, it may also play a role in the development of both ALT (+) and ALT (-) cancer. In our analysis, *HFMI* and *DMCI* both had suppressed expression in non-immortalized normal breast stroma cells. The stroma in normal breast tissue plays an important part in guiding the development of the mammary gland, and the downregulation of both *HFMI* and *DMCI* could be due to some specialized role these cells have. *HFMI* and *DMC1* are involved in meiotic recombination; thus, it is possible that these normal breast stroma cells undergo much less recombination. Cells that undergo high levels of recombination, such as gametes, do so to increase the genetic variation within their DNA. Therefore, it is possible that normal breast stroma cells maintain high levels of genome integrity to facilitate their specialized function in the development of the mammary gland.

Future Direction

In the future, more analysis should be performed on these meiotic genes to better understand which genes are involved in the ALT pathway, and which of those are important for the development of cancer in general. In this study, there were various discrepancies between which genes were identified as being expressed in ATL (+) and ALT (-) cancer cell lines. RT-PCR and Q-PCR analysis indicated that there were more variations in meiotic gene expression between the two ALT (-) cell lines assessed than was apparent between the two ALT (+) cell lines assessed. Thus, it would be advantageous to further analyze these discrepancies between

Alt (+) cell lines and ALT (-) cell lines to better understand which meiotic genes are important for the ALT pathway.

The introduction of the auxin induced degradation mechanism into the ALT (+) U2OS cell line provides a great opportunity to analyze how the cell reacts to the depletion of any gene of interest. This system allows the mAID-tagged protein of interest to be depleted rapidly and reversibly. Therefore, it would be advantageous in the future to perform experiments such as depleting SMC5 or SMC6 from the U2OS cells in the presence of DNA damage and assessing how the cell recovers in the initial stages with SMC5 or SMC6 depleted, but then re-introducing SMC5 or SMC6 later by removing auxin from the growth media and observing what changes. If SMC5/6 is playing an important role during a specific stage in the DDR, it could be observed easily with this system.

Furthermore, this system is titratable, and it is possible to partially deplete the protein of interest. The amount of control of variables this system provides make it advantageous over previous siRNA techniques of protein knock-down. Previously, it has been demonstrated that knockdown of the human SMC5/6 complex by RNA interference (RNAi) inhibits telomere HR in ALT (+) cells, resulting in telomere shortening and cells entering senescence (Potts *et al.*, 2007). However, the particular RNAi molecules used were later found to be subject to off target effects, which causes some concern with the interpretation of the data (Wu *et al.*, 2012). Furthermore, in general, RNAi does not completely knock-out the protein of interest. Thus, the AID system in the ALT (+) U2OS cell provides an opportunity to assess the validity of these findings by ensuring complete depletion of the SMC5/6 complex with no off-target effects that could possibly confound the results.

The SMC5/6 protein complex has been indicated as playing a role in the ALT mechanism. The SUMO E3 ligase MMS21 (also known as NSE2) is a component of the SMC5/6 complex, and it has been indicated as having SUMOylation properties (Chung *et al.*, 2012). NSE2 of the SMC5/6 complex has been indicated as playing a role in facilitating homologous recombination within APBs of ALT (+) cancer cells by SUMOylation components of the shelterin complex, such as RAP1, TIN2, TRF1 and TRF2 (Potts, 2009). SUMOylation of components of the shelterin complex may promote telomeres to associate with PML bodies, as proteins within PML bodies have a high affinity for SUMO. Knockdown of SMC5/6 in ALT (+) cancer cells results in decreased recombination at telomeres, shortening of telomeres, and cells entering senescence (Potts and Yu, 2007). Additionally, in fission yeast treated with the DNA damaging agent methyl methanesulfonate (MMS), there is significantly more SUMOylation of both SMC6 and NSE4 by the SUMO ligase function of MMS21. It has been demonstrated that the SMC5/6 complex localizes to telomeres in budding and fission yeast. Therefore, this suggests that SUMOylation of SMC5/6 by MMS21 could potentially increase the affinity of the SMC5/6 complex to telomeres in telomerase-positive fission yeast (Potts, 2009).

Additionally, it has been speculated that various other proteins, such as components of the NuRD-ZNF827 complex, undergo SUMOylation within the APBs (Conomos *et al.*, 2014). Thus, the SMC5/6 complex plays an important role in SUMOylation events required for the ALT mechanism and the formation and subsequent maintenance of APBs associated with telomeres in ALT positive cells. In this study, SMC5 and SMC6 were independently tagged with the mAID tag in the U2OS cell line and the AID system introduced. Further studies utilizing the AID mechanism to deplete SMC5 and SMC6 independently in the ALT (+) U2OS cell line should be performed in the future to assess how a lack of SUMOylation within the APBs affects the

accumulation of APBs and subsequently the ALT mechanism. Using the AID mechanism in human ALT (+) U2OS cells provides an opportunity to further elucidate studies that have only been conducted on yeast and expand our understanding of how the SMC5/6 complex is functioning in the ALT mechanism in humans.

Additionally, the progression of ALT (+) cell through the cell cycle may rely on the functions of the SMC5/6 complex. Therefore, it would be advantageous to conduct experiments with cell synchronization, and test the effects of cell cycle progression in the presence or absence of SMC5 or SMC6. Experiments such as synchronizing cells at the G2/M stage of the cell cycle using nocodazole, then releasing in the presence or absence of the SMC5/6 complex using the AID system could help gain a better understanding of how the SMC5/6 complex is functioning during this stage of the cell cycle. Additionally, cells could also be halted at the G1/S stage of the cell cycle using a thymidine block. SMC5/6 has been indicated as playing a role in the collapse of replication forks, and thymidine induces replication fork collapse (Roy *et al.*, 2015). At telomeres, replication forks commonly collapse due to the irregular secondary structures that form, causing single-stranded overhangs. The BIR mechanism is used to repair broken chromosomes when a single-stranded overhang is present in DNA and plays an important part in the ALT mechanism (Kraus *et al.*, 2001). Thus, the AID system in the ALT (+) U2OS cell line can be used to better understand the role of the SMC5/6 complex in BIR and the ALT mechanism during specific stages of the cell cycle.

References

- Almeida, L. G., Sakabe, J. J., deOliveira, A. R., Silva, M. C., Mundstein, A. S., Cohen, T., et al. (2009). CT database: a knowledge-base of high-throughput and curated data on cancer-testis antigens. *Nucleic Acids Research* 37(Database issue), D816–D819.
- Amorim, J. P., Santos, G., Vinagre, J., & Soares, P. (2016). The Role of ATRX in the Alternative Lengthening of Telomeres (ALT) Phenotype. *Genes*, 7(66).
- Arendt, L. M., Rudnick, J. A., Keller, P. J. & Kuperwasser, C. (2010). Stroma in Breast Development and Disease. *Seminars in Cell Development Biology* 21(1), 11-18.
- Arnoult, N., & Karlseder, J. (2014). ALT Telomeres Borrow from Meiosis to get Moving. *Cell*, 159, 11-12.
- ATCC HTB-96 U-2OS <https://www.atcc.org/Products/All/HTB-96.aspx>
- Baudat, F., Imai, Y. & de Massy, B. (2013). Meiotic recombination in mammals: localization and regulation. *Nature Reviews Genetics* 14, 794–806.
- Betcher, O. E., Shay, J. W., & Wright, W. E. (2004). The Frequency of Homologous Recombination in Human ALT Cells. *Cell Cycle*, 3(5), 547-549.
- Bishop, A. J. R. & Schiestl, R. H. (2002). Homologous Recombination and its Role in Carcinogenesis. *Journal of Biomedicine and Biotechnology* 2(2), 75-85.
- Brosh, R., Hrynyk, I., Shen, J., Waghray, A., Zheng, N. et al. (2016). A Dual Molecular Analogue Tuner for Dissecting Protein Function in Mammalian Cells. *Nature Communications* 7(11742).
- Cesare A. J., & Reddel, R. R. (2010). Alternative Lengthening of Telomeres: Models, Mechanisms and Implications. *Nature Reviews: Genetics*, 11, 319-330.

- Cho, N. W., Dilley, R. L., Lampson, M. A., & Greenberg, R. A. (2014). Interchromosomal Homology Searches Drive Directional ALT Telomere Movement and Synapsis. *Cell*, 159, 108-121.
- Chung, I., Osterwald, S., Deeg, K. I., & Rippe, K. (2012). PML Body Meets Telomere: The Beginning of an ALTERNATE Ending?. *Nucleus*, 3(3), 263-275.
- Conomos, D., Reddel, R. R., & Pickett, H. A. (2014). NuRD-ZNF827 Recruitment to Telomeres Creates a Molecular Scaffold for Homologous Recombination. *Nature Structural & Molecular Biology*, 21(9), 760-770.
- Daniel, K., Lang, J., Haches K., Fru, J., Anastassiadis, K. et al. (2011). Meiotic Homologue Alignment and its Quality Surveillance are Controlled by Mouse HORMAD1. *Nature Cell Biology* 13(5), 599-612.
- Deng, Y., Chan, S., & Chang, S. (2008). Telomere Dysfunction and Tumor Suppression-the Senescence Connection. *Nature Review Cancer*, 8(6), 450-458.
- Dilley, R. L., Verma, P., Choo, N. W., Winters, H. D., Wondisford, A. R. et al. (2016). Break-Induced Telomere Synthesis Underlies Alternative Telomere Maintenance. *Nature* 539, 54-58.
- Donnianni, R. A. & Symington, L. S. (2013). Break-induced Replication Occurs by Conservative DNA Synthesis. *Proceedings of the National Academy of Sciences of the USA* 110(33), 13475-13480.
- Filippo, J. S., Sung, P., & Klein, H. (2008). Mechanism of Eukaryotic Homologous Recombination, *Annual Review of Biochemistry*, 77, 229-257.
- Gilson, E., & Geli, V. (2007). How Telomeres are Replicated. *Nature Reviews: Molecular Cell Biology*, 8, 825-838.

- Gray, W. M., Kepinski, S., Rouse, D., Leyser, O. & Estelle, M. (2001). Auxin Regulates SCF^{TIR1}-dependent Degradation of AUX/IAA Proteins. *Nature* 414, 271-276.
- Hayflick L., & Moorhead P.S. (1961). The Serial Cultivation of Human Diploid Cell Strains. *Experimental Cell Research*, 25, 585–662.
- Heaphy, C. M., de Wilde, R. F., Jiao, Y., Klein A. P., Edil, B. H., Shi, C., et al. (2011). Altered Telomeres in Tumors with *ATRX* and *DAXX* Mutations. *Science*, 333(6041), 425.
- Jackson A. L. & Linsley, P. S. (2010). Recognizing and avoiding siRNA off-target effects for target identification and therapeutic application. *Nature Reviews Drug Discovery* 9, 57-67.
- Kalejs, M., Ivanov, A., Plakhins, G., Cragg, M. S., Emzinsh, D. et al. (2006). Upregulation of Meiosis-Specific Genes in Lymphoma Cell Lines Following Genotoxic Insult and Induction of Mitotic Catastrophe. *BMC Cancer* 6(6).
- Kraus, E., Leung, W., & Haber, J. E. (2001). Break-Induced Replication: A Review and an Example in Budding Yeast. *PNAS*, 98(15).
- Lauberth, S. M., & Rauchman, M. (2006). A Conserved 12-Amino Acid Motif in Sall1 Recruits the Nucleosome Remodeling and Deacetylase Corepressor Complex. *Journal of Biological Chemistry*, 281(33), 23922-23931.
- Lee, M., Hills, M., Conomos, D., Stutz, M. D., Dagg, R. A., Lau, L. M. S., et al. (2014). Telomere Extension by Telomerase and ALT Generate Variant Repeats by Mechanistically Distinct Processes. *Nucleic Acids Research*, 42(3), 1733-1746.
- Levy, M., Allsopp, R. C., Futcher, A. B., Greider, C. W. & Harley, C. B. (1991). Telomere End-Replication Problem and Cell Aging. *Journal of Molecular Biology*, 225, 951-960.

- Liang, X., Potter, J., Kumar, S., Zou, Y., Quintanilla, R. et al. (2015). Rapid and Highly Efficient Mammalian Cell Engineering Via Cas9 Protein Transfection. *Journal of Biotechnology* 208, 44-53.
- Luo, M. Yang, F., Leu, N. A., Lancaiche, J., Handel, M. A., et al. (2013). MEIOB exhibits single-stranded DNA-binding and exonuclease activities and is essential for meiotic recombination. *Nature Communications* 4, 2788.
- Marechal, A. & Zou, L. (2013). DNA Damage Sensing by the ATM and ATR Kinases. *Cold Spring Harbor Perspectives in Biology*, 5(9).
- Momand, J., Wu, H., & Dasgupta, G. (1999). MDM2 – Master Regulator of the p53 Tumor Suppressor Protein. *Gene* 242, 15-29.
- Muntoni, A., & Reddel, R. R. (2005). The First Molecular Details of ALT in Human Tumor Cells. *Human Molecular Genetics*, 14(2), 191-196.
- Natsume, T., Klyomitsu, T., Saga, Y. & Kanemaki, M. T. (2016). Rapid Protein Depletion in Human Cells by Auxin-Inducible Degron Tagging with Short Homology Donors. *Cell Reports* 15, 210-218.
- Nikalayevich, E. & Ohkura H. (2015). The NuRD Nucleosome Remodeling Complex and NHK-1 Kinase are Required for Chromosome Condensation in Oocytes. *Journal of Cell Science* 128(3), 566-575.
- Nishimura, K., Fukagawa, T., Takisawa, H., Kakimoto, T. & Kanemaki, M. (2009). An Auxin-based Degron System for the Rapid Depletion of Proteins in Nonplant Cells. *Nature Methods* 6(12), 917-923.

- Pickett, H. A., & Reddel, R. R. (2015). Molecular Mechanisms of Activity and Depression of Alternative Lengthening of Telomeres. *Nature Structural & Molecular Biology*, 22(11), 875-880.
- Polo, S. E. & Jackson S. P. (2011). Dynamics of DNA Damage Response Proteins at DNA Breaks: A Focus on Protein Modifications. *Genes and Development*, 25, 409-433.
- Potts, P.R. and Yu, H. (2007). The SMC5/6 complex maintains telomere length in ALT cancer cells through SUMOylation of telomere-binding proteins. *Nature Structural Molecular Biology* 14(7), 581–590.
- Potts, P. R. (2009). The Yin and Yang of the MMS21-SMC5/6 ligase complex in homologous recombination. *DNA Repair* 8, 499-506.
- Pu, D., Wang, C., Cao, J., Shen, Y., Jiang, H., Liu, J., et al. (2016). Association Analysis Between HFM1 Variation and Primary Ovarian Insufficiency in Chinese Women. *Clinical Genetics*, 89(5), 597-602.
- Roumelioti, F., Sotiriou, S. K., Katsini, V., Chiourea, M., Halazonetis, T. D. et al. (2016). Alternative Lengthening of Human Telomeres is a Conservative DNA Replication Process with Features of Beak-Induced Replication. *EMBO Reports*.
- Roy, M., Dhanaraman, T., & D'Amours, D. (2015). The Smc5-Smc6 Heterodimer Associates with DNA Through Several Independent Binding Domains. *Scientific Reports*, 5.
- Sadelain, M., Papapetrou, E. P. & Bushman, F. (2012). Safe Harbours for the Integration of New DNA in the Human Genome. *Nature Reviews Cancer* 12, 51-58.
- Scanlan, M. J., Gure, A. O., Jungbluth, A. A., Old, L. J. & Chen, Y. T. (2002). Cancer/testis antigens: an expanding family of targets for cancer immunotherapy. *Immunology Reviews* 188, 22–32.

- Shay, J. W., & Wright, W. E. (2011). Role of Telomeres and Telomerase in Cancer. *Seminars Cancer Biology*, 21(6), 349-353.
- Simpson, A. J., Caballero, O. L., Jungbluth, A., Chen, Y. T. & Old, L. J. (2005). Cancer/testis antigens, gametogenesis and cancer. *Nature Reviews Cancer*, 615–625.
- Slaymaker, I. M., Gao, L., Zetsche, B., Scott, D. A., Yan, W. X. et al. (2016). Rationally Engineered Cas9 Nucleases with Improved Specificity. *Science* 351(6268), 84-88.
- Tomasetti, C., Li, L. & Vogelstein, B. (2017). Stem Cell Division, Somatic Mutations, Cancer Etiology, and Cancer Prevention. *Science* 355, 1330-1334.
- UniProtKB-Q8NA31 (TERB1_HUMAN) <http://www.uniprot.org/uniprot/Q8NA31>
- Watkins, J., Weekes, D., Shah, V., Gazinska, P., Joshi, S., et al. (2015). Genomic complexity profiling reveals That HORMAD1 overexpression contributes to homologous recombination deficiency in triple-negative breast cancers. *Cancer Discovery* 5, 488–505.
- Wang, C., Gu, Y., Zhang, K., Xie, K., Zhu, M., Dai, N., et al. (2015). Systematic Identification of Genes with a Cancer-Testis Expression Pattern in 19 Cancer Types. *Nature Communications* 7(10499).
- Wang, J., Zhang, W., Jiang, H. & Wu, B. (2014). Mutations in *HFMI* in Recessive Primary Ovarian Insufficiency. *The New England Journal of Medicine* 370(10), 972-974.
- Wu, N., Kong, X., Ji, Z., Zeng, W., Potts, P. R. et al. (2012). Scc1 SUMOylation by Mms21 promotes sister chromatid recombination through counteracting Wapl. *Genes and Development* 26, 1473-1485.