CTCF Loss Potentiates *TP53* Mediated Gene Transcription in Breast Tissue

by

Meredyth Elisseou, Experimental Medicine

McGill University, Montreal

December 2022

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree

of

Master of Science in Experimental Medicine

Supervisor: Dr. Michael Witcher, Department of Oncology

Table of Contents

1. Abstract

A single copy loss of CTCF is found in about 50% of breast cancer patients. Based on clinical TCGA data we hypothesized that the loss of CTCF may potentiate *TP53* target gene expression in patients. Using MCF10A cells as a model, we deleted a single copy of CTCF using CRISPR/Cas9. We found, using qPCR and RNA-seq, that cells expressing low CTCF displayed an enhanced *TP53* response after exposure to chemotherapeutics. Using ATAC-seq, we aimed to explore whether changes in chromatin accessibility at *TP53* target genes within MCF10A $CTCF^{+/-}$ cells compared to control cells following the induction of DNA damage (6μM cisplatin for 8h). We discovered that accessibility of a subset of transcription start sites is associated with heightened gene expression in $CTCF^{+/-}$ compared to the control. Interestingly, for 4 *TP53* associated genes, there is increased accessibility on both transcriptional start sites and termination sites following the induction of DNA damage. The importance of chromatin accessibility at these two regions is still under investigation. Additionally, accessible gene regions at both sites also appear to have greater enrichment within topologically associated domains within $CTCF^{+/-}$ cells compared to the controls. We propose that the increased accessibility of *TP53* target genes following damage represents a mechanism enhancing the efficacy of the *TP53*-regulated DNA damage response in MCF10A cell lines.

2. Le Résumé

La perte d'une seule copie du gène CTCF est retrouvée chez environ 50% des patientes atteintes d'un cancer du sein. Sur la base des données cliniques du TCGA, nous avons émis l'hypothèse que la perte de CTCF pourrait potentialiser l'expression des gènes cible de *TP53* chez les patientes. En utilisant les cellules MCF10A comme modèle, nous avons supprimé une seule copie du gène CTCF à l'aide de la technologie CRISPR/Cas9. En utilisant les techniques de qPCR et d'ARN-seq, nous avons découvert que les cellules portant un faible niveau de CTCF affichaient une réponse au gène *TP53* améliorée après exposition à la chimiothérapie. À l'aide de l'ATAC-seq, nous avons cherché à déterminer si l'induction élevée de la transcription des gènes cible de *TP53* était associée à des modifications de la structure ouverte ou fermée de la chromatine. Plus précisément, nous nous sommes intéressés à la comparaison de l'accessibilité de la chromatine aux gènes cibles de *TP53* dans les cellules MCF10A CTCF$^{+/-}$ par rapport aux cellules témoins après l'induction de dommages à l'ADN (cisplatine 6μM pendant 8h). Nous avons découvert que l'accessibilité du site d'initiation de la transcription est associée à une expression génique accrue dans les cellules CTCF$^{+/-}$ par rapport aux cellules contrôle. De plus, pour un sous-ensemble de gènes répondant à *TP53*, il existe une accessibilité accrue sur les sites d'initiation et les sites de terminaison de la transcription à la suite de l'induction de dommages à l'ADN. L'importance de l'accessibilité de la chromatine dans ces deux régions est encore à l'étude. Par ailleurs, les régions géniques accessibles sur les deux sites semblent également avoir un plus grand enrichissement en TADS dans les cellules CTCF$^{+/-}$ par rapport aux cellules contrôle. Ainsi, nous proposons que l'accessibilité accrue des gènes cibles de *TP53* à la suite de dommages à l'ADN représente un mécanisme améliorant l'efficacité de la réponse aux dommages à l'ADN régulée par *TP53*.

3. Acknowledgements

First, I would like to extend my thanks to Dr. Michael Witcher for his supervision and guidance. His considerable knowledge in the field of epigenetics, editorial assistance, and mentoring have been an essential asset in the completion of this project. I greatly appreciated his prompts to participate in seminars and journal clubs so that I could learn from and interact with peers and professionals in medical research.

Next, I would like to thank the members of the Witcher lab for contributing to my positive experience. I would like to extend a special thanks to Tiejun Zhao for ensuring I had access to all the materials I needed for this project as well as the advice he provided. I would also like to thank a previous member of the lab, Cheng Kit Wong, for training me and all his assistance. I would also like to thank Benjamin Lebeau for his advice in experimental planning and additional training as well as his contributions to the bioinformatic analysis of my project. In addition, I would like to thank Korin Sahinyan for her expertise and advice given while I participated in implementing the ATAC-seq protocol in my lab. Thank you to all other members of the lab for the productive discussions, ideas, support, and positive work environment.

My committee members, which included Dr. Kostas Pantopoulos, Dr. Josie Ursini-Siegel, and Dr. Volker Blank also deserve a special thanks for their helpful advice and recommendations. Their contributions to my committee meeting were exceptionally helpful.

Finally, I would like to thank my family and friends who supported me and gave me encouragement.

4. Statement of Contribution of Author

This thesis is completely written by Meredyth Laurel Georgia Elisseou (the author) with revisions provided by Dr. Michael Witcher. All experiments described were completed by the author under the advice and supervision of Dr. Michael Witcher.

Previously completed RNA-seq in our lab was carried out by Cheng Kit Wong. The processing of data for RNA-seq and ATAC-seq, including extraction and normalization of read counts from raw sequencing files (fastq files) and generating Log_2 fold change of differentially expressed genes for all samples, was done by Benjamin Lebeau. Analysis of the data was completed by the author and Benjamin Lebeau.

MCF10A cell lines were generated in our lab by previous members.

5. List of Figures

6. List of Symbols, Nomenclature, or Abbreviations

- Transcription Start Site (TSS)
- Transcription Termination Site (TTS)
- Topologically Associating Domain (TAD)
- Assay for Transposase Accessible Chromatin (ATAC-seq)
- Copy Number Loss (CNL)
- Double Strand Break (DSB)
- Homologous Recombination (HR)
- Non-homologous end joining (NHEJ)
- CCCTC-binding factor (CTCF)
- The Cancer Genome Atlas (TCGA)
- BBC3; PUMA
- $CTCF^{+/-}$ #1; KO1; #1
- $CTCF^{+/-}$ #2; KO2; #2
- Wild type; WT; Ctl; Control
- *TP53* (Human P53); Trp53 (Murine p53)

7. Introduction

7.1 Breast Cancer

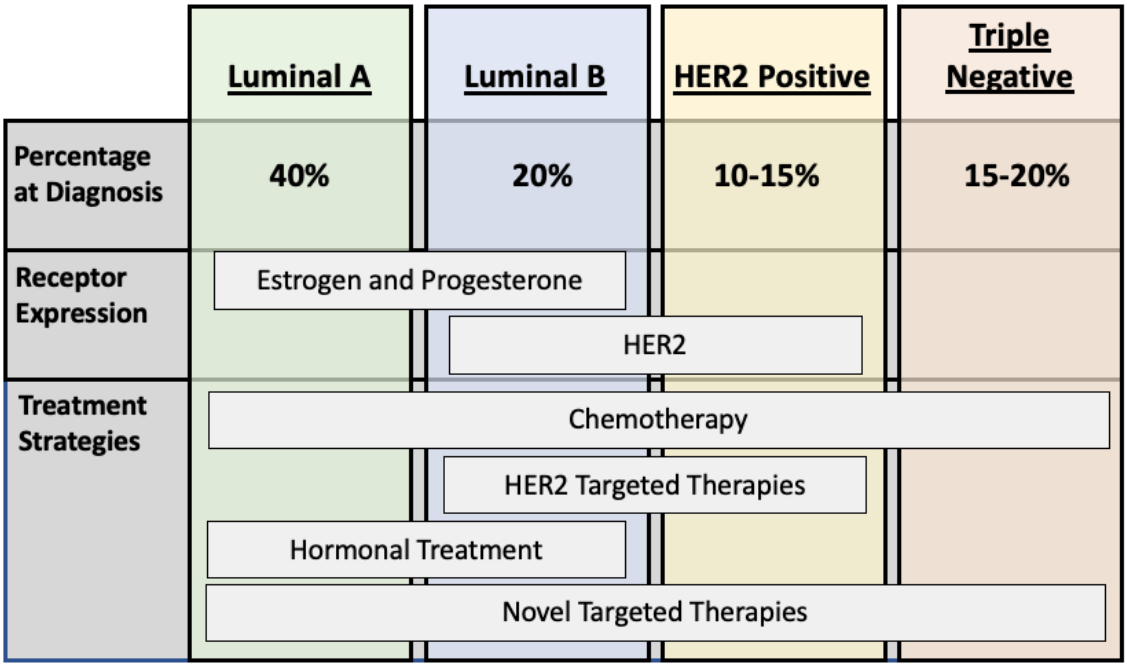


Figure I. Breast Cancer Molecular Subtypes. Breast Cancer can be categorized into four types: Luminal A, Luminal B, HER2 Positive, and Triple Negative.

Currently, breast cancer is the most common cancer affecting women globally and it represents almost a quarter of all cancers diagnosed in women (Ferlay et al. 2015). The prevalence of breast cancer appears to be the result of interactions between multiple genes and epigenetic factors that may be influenced by the environment (Kaaks et al. 2005; Wu et al. 2015).

Breast cancer is classified into subtypes based on the expression of predictive molecular biomarkers that dictate the type of treatment administered. The luminal A subtype is the most common cohort of breast cancer patients and is associated with the best prognosis among all subtypes. This subtype expresses high levels of estrogen receptor (ER) and/or progesterone

receptor (PR). Luminal A breast cancers usually respond well to anti-estrogens and the National Cancer Institute reports a 5-year relative survival that is superior to other breast cancer subtypes (94.4%). In a study reviewing relapse rates, Ignatov et al. (2018) found that the local and regional rate was 1.5% and 0.7%, respectively.

Luminal B subtype is characterized by ESR1+ status, a high proliferative index, and commonly, HER2 amplification. Luminal B patients generally relapse sooner than Luminal A after therapy and show a poorer prognosis than those patients with a Luminal A subtype; the 5-year relative survival is reported to be 90.7% (The National Cancer Institute; Howlader et al, 2018). Ignatov et al. (2018) found that the local and regional recurrence rate was 2.9% and 1.5%, respectively.

The third subtype of breast cancer is HER2-enriched, and this subtype has a slightly worse prognosis than Luminal subtypes; the 5-year relative survival is reported to be 84.8% (The National Cancer Institute; Seung et al., 2020). The HER2-enriched subtype is characterized by a highly proliferative tumor with overexpression of the HER2 receptor due to amplification of the gene. The National Cancer institute reports the 5-year relative survival percentage as 84.8%. Additionally, Ignatov et al. (2018) found that the local and regional recurrence rate was 7.5% and 3.4%, respectively. HER2+ tumors are treated with antibody-based therapies that prohibit signaling through the HER2 cell surface receptor.

Breast cancers that lack ER/PR and HER2 expression are classified as triple-negative breast cancers, which is the subtype with the worst prognosis. The 5-year relative survival percentage is 77.1% for this subtype (The National Cancer Institute). The investigation by Ignatov et al. (2018) found that the local and regional relapse rate was 7.6% and 3.3%, respectively. The standard of treatment for triple negative breast cancer patients is chemotherapy

and radiotherapy. Regardless of subtype, the stage of the cancer when it is diagnosed is a major determining factor in survival.

7.2 Haploinsufficient Tumor Suppressor Genes

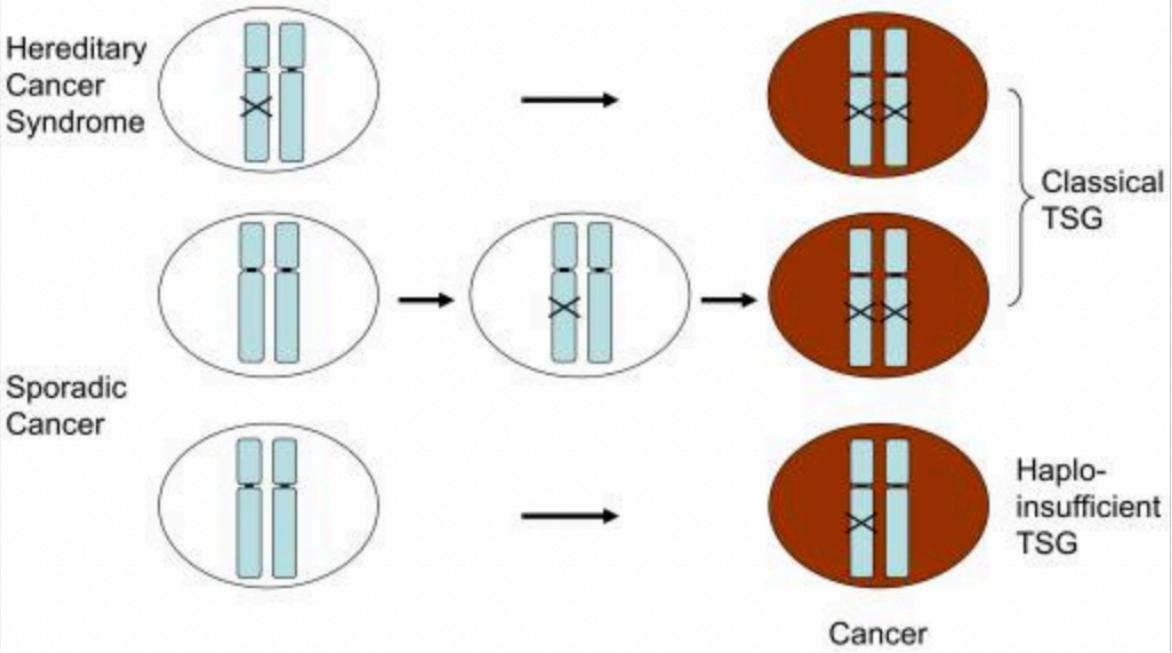


Figure II. Breast cancer tumor development caused by classical and haploinsufficient tumor suppressor genes. Haploinsufficient tumor suppressor genes (TSG) retain one unmuted wild-type copy of the gene (Inoue et al. 2017).

Cancer is an exceedingly complicated genetic disease which is often triggered by alterations in gene expression and diverse DNA mutations. From the perspective of cancer, there are two primary classes of genes that impact tumor growth; tumor suppressors and oncogenes (Alberts et al. 2002; Inoue et al. 2017). The most prevalent early models have indicated that a complete loss of a tumor suppressor gene (two copies) must be achieved to impact pro-cancer pathways (Knudson, 1971). Despite this long-held belief, in recent years, more research has demonstrated that in some instances, only one allele of a tumor suppressor gene needs to be lost

to support oncogenesis; this is known as haploinsufficiency. This phenomenon was first described by Fero et al. in 1998, when the cyclin dependent kinase inhibitor, p27, was described as a haploinsufficient tumor suppressor. The tumors that result from $p27^{+/-}$ tumors in mice retained one unmutated wild type allele, demonstrating that the loss of a single allele in these tumor suppressor genes can indeed contribute to tumorigenesis (Fero et al. 1998; Quon & Berns, 2001). Understanding the consequences of haploinsufficiency in tumor suppressing genes is thus essential to our ability to approach and target mechanisms that result from tumor suppressor haploinsufficiency.

Although haploinsufficiency in tumor suppressor genes has been mainly described through the lens of mouse models, copy number loss has also been investigated in human tissue. Changes in the copy number of genes such as HER2 have been extensively documented in breast cancer and are present in model cell lines (Chi et al. 2018; Janiszewska et al, 2021). This prompted efforts in searching for the identity of the tumor suppressor gene by comparative genomic hybridization techniques to look for overlapping regions of deletions. These studies identified the q-arm of chromosome 16 as a region commonly deleted in many cancers (Rakha et al. 2006; Mason et al. 2000; Massenkeil et al. 1995). The Cancer Genome Atlas (TCGA) shows that the epigenetic regulatory protein CTCF (CCCTC-binding factor) undergoes copy number loss (CNL) in over 50% of breast tumors (Kemp et al. 2014). Hemizygous deletions at chromosomal location 16q22.1, that encompasses CTCF, frequently occur in many human cancers beyond breast cancer (Filippova et al., 1998). CTCF is critical for chromatin organization and acts as a master regulator of the genome. Researchers demonstrate that loss of a single copy of CTCF triggers epigenetic changes and markedly increases cancer development and progression in mice and humans. A 2014 study on a $CTCF^{+/-}$ mouse model has shown an

increased rate of spontaneous lymphoma development in the $CTCF^{+/-}$ mice as compared to WT mice (Kemp et al. 2014). This has demonstrated CTCF's role as a haploinsufficient tumor suppressor gene.

7.3 CTCF

CCCTC-binding factor (CTCF) is a multifunctional epigenetic regulatory protein (Filippova et al. 2002) consisting of a Zinc-finger domain comprising a disordered N-terminal domain, an 11 DNA-binding Zinc-finger central domain, and a C-terminal that may be phosphorylated as a modulator of activity. Evidence regarding the function of the C-terminal has been contradictory, either supporting (Saldaña-Meyer et al. 2014) or denying (Xiao et al. 2011) its mediating role in CTCF-cohesion interactions. CTCF was first characterised as a repressive transcriptional factor of c-Myc gene expression (Lobanenkov et al. 1990), followed by subsequent characterizations of its other functions in regulating the genome. Apart from its role as a transcription factor (Lobanenkov et al. 1990; Peña-Hernández et al. 2015), CTCF has also been described to act in the formation of chromatin boundaries (Witcher & Emerson, 2009), acting as an enhancer insulator element (Hark et al. 2000), playing a role in three-dimensional genome organization (Tang et al. 2015), and to facilitate the repair of DNA double-strand breaks via homologous recombination (Hilmi et al. 2017; Lang et al. 2017; Hwang et al. 2019). Many of its disparate functions have been ascribed to its ability to facilitate long range interactions and are often associated with topologically associating domains (TADs) (Bell et al. 1999; Wutz et al. 2017).

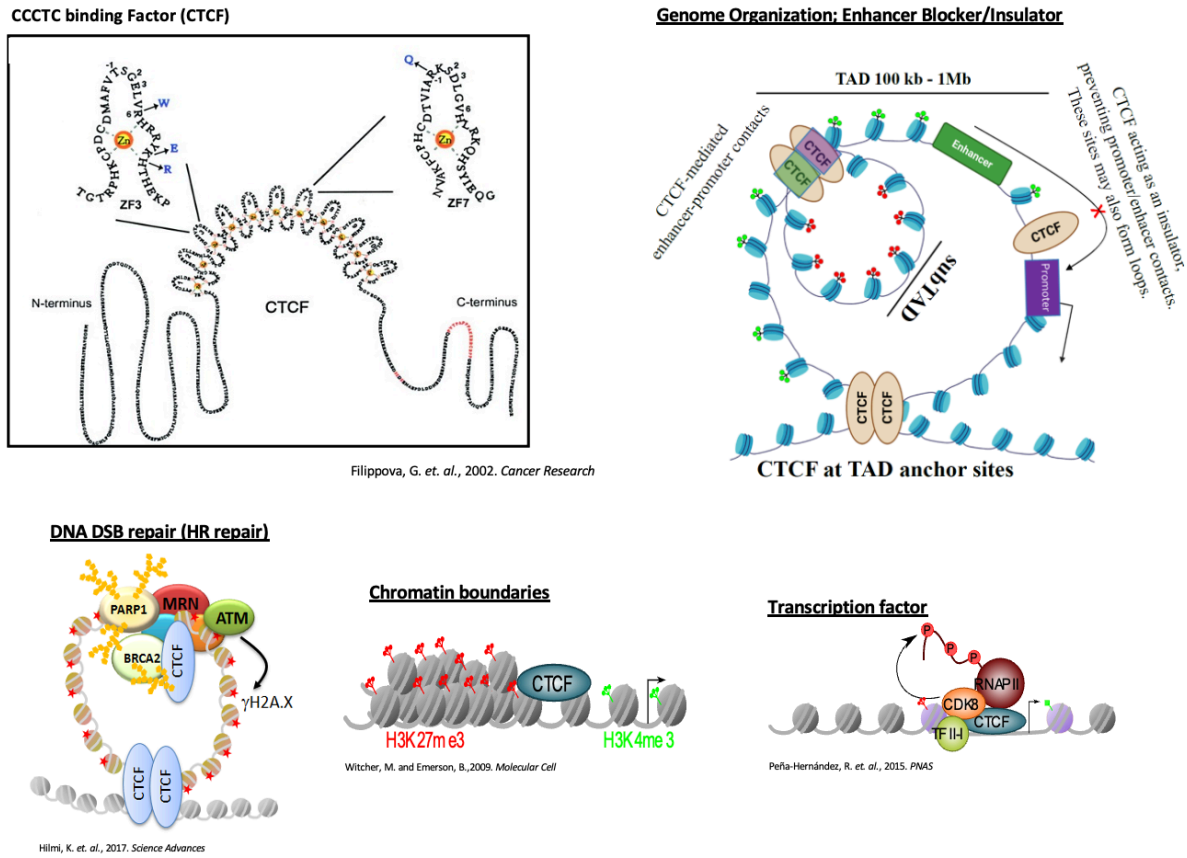


Figure III. CTCF is a multifunctional regulatory protein. Some of the related functions of CTCF include genome organization, acting as an enhancer blocker, maintaining chromatin boundaries, acting as a transcription factor, and contributing to DNA DSB repair. (Dr. Michael Witcher lab; Wong. 2021).

7.3.1 Transcription Factor

CTCF is a well described gene regulator through its role as a transcription factor. It was first identified as a transcription factor by Lobanenkov et al. in 1990, where they found that the protein binds to motifs upstream of the c-Myc gene and was essential for its transcriptional regulation. Later research by Peña-Hernández in 2015 found that the transcription factor, TFII-I's role in regulating gene expression is mediated by CTCF. This mediation resulted in the promotion of CDK8 recruitment and Polymerase II phosphorylation. In addition, CTCF is also

known to be found at RNA polymerase II stalling and termination sites (Egloff et al. 2009). In addition, around 20% of CTCF binding sites are within 2kb of transcription start sites (TSS) (Cuddapah et al. 2009; Kim et al. 2007). This may indicate a role where CTCF participates in the regulation of transcription, particularly at the 5' end of a gene.

7.3.2 Genome Organization & Chromatin Boundaries

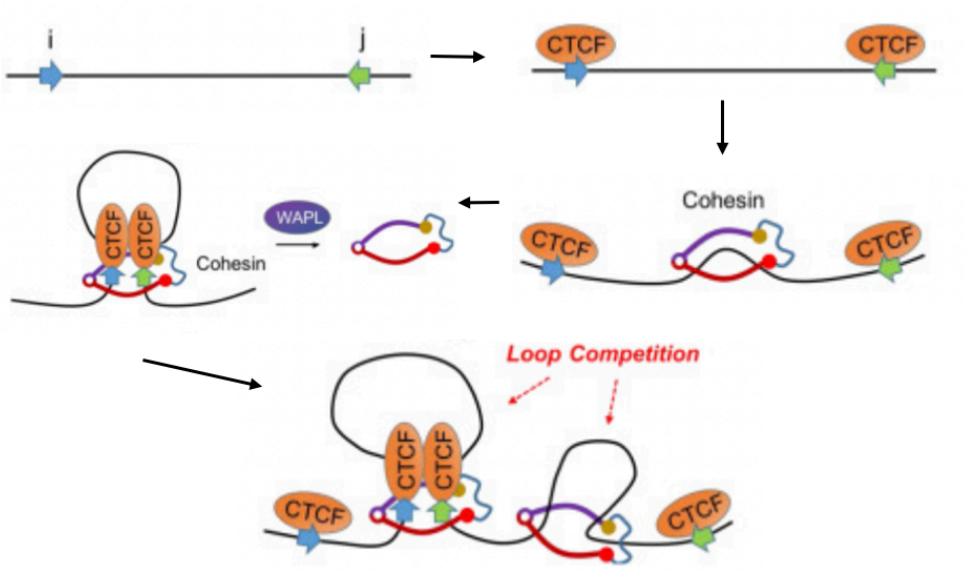


Figure IV. Loop extrusion model. Cohesin is loaded between CTCF, and the loop forms progressively as cohesin translocated along the chromatin fiber until it reaches CTCF (Xi & Beer, 2021).

Three-dimensional studies on chromatin structure have found that chromosomes are organized in loops that can be 100kb to 1Mb in size; these loops are known as topologically associating domains (TADs). Notably, a prominent feature of TADs is the presence of CTCF and cohesin binding at their boundaries, acting to anchor the large chromatin loop. Both CTCF and its cohesion binding partner are important for the formation of TADs (Luo et al. 2018; Narendra et al. 2015). Recent research has shown that the N-terminus of CTCF directly interacts with

subunits of the human cohesin complex in order to act as an anchor for these loops (Li et al. 2020). The loop extrusion model was proposed wherein cohesin translocation forms a chromatin loop until the formation is stopped by CTCF, which acts as an anchor at the boundaries of TADs (Xi & Beers, 2021; Fudenburg et al. 2016; Sanborn et al. 2015).

7.3.3 Chromatin Boundaries

Chromatin is generally described as being in one of two states: transcriptionally active euchromatin (open) or transcriptionally inactive heterochromatin (closed). In some instances, heterochromatin can aberrantly spread across the genome, leading to the inactivation of genes. Thus, boundaries to prevent this spread and preserve gene expression are necessary. Research by Witcher & Emerson (2009) found that CTCF may play an essential role in establishing tumor suppressor genes in chromosomal domains by forming boundaries. In addition, other research has shown that CTCF can block heterochromatin from spreading (Ling et al. 2006). Later studies have found that CTCF binding sites were often found at the boundaries of domains containing H3K27me3 marks, possibly to prevent the spreading of these marks (Barski et al. 2007; Cuddapah et al. 2009). Thus, CTCF can also act as a boundary between active euchromatin regions and inactive heterochromatin regions.

7.3.4 DNA Double Strand Break Repair

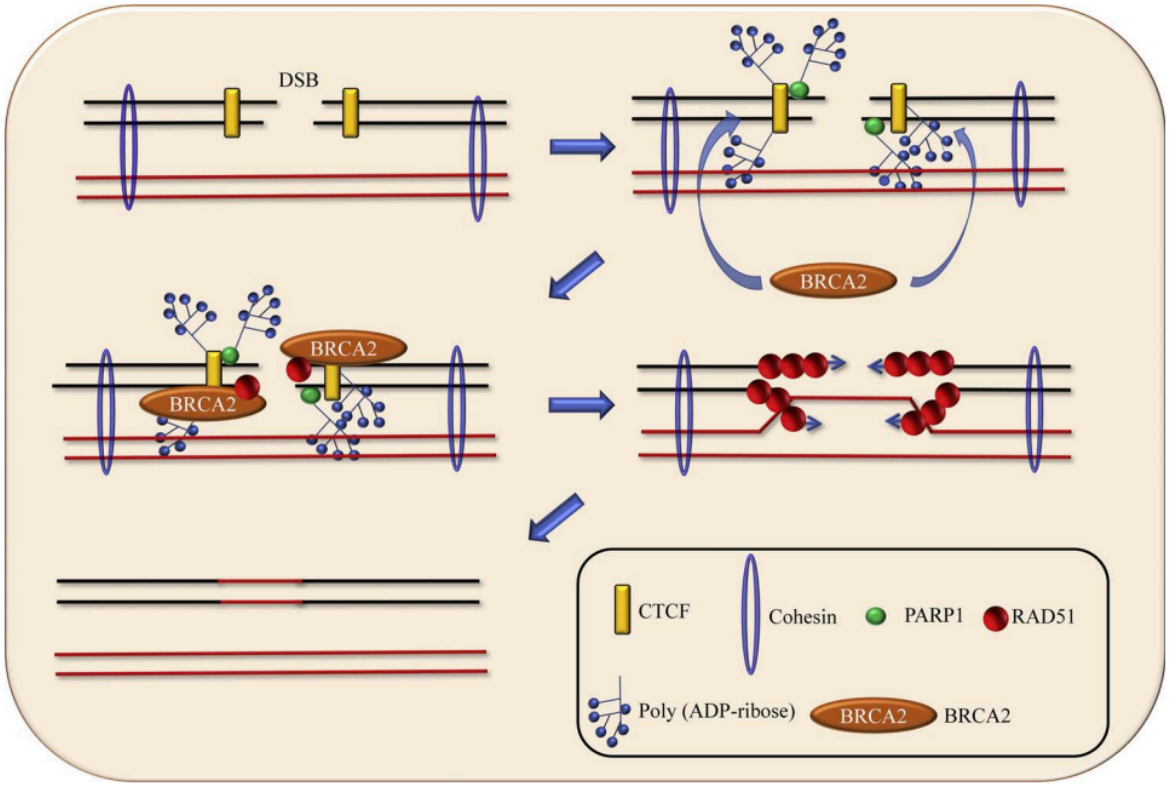


Figure V. CTCF recruits BRCA2 to DSBs during HR. During DSB repair, CTCF recruits BRCA2 to DSBs in a PARlation dependent manner. BRCA2 in turn recruits RAD 51, which form filaments to allow strand invasion and homologous recombination. (Tanwar et al. 2019).

When cells are exposed to DNA damaging conditions, whether they are chemical or physical, a double-strand break (DSB) may occur. These breaks may be repaired by either non-homologous end joining (NHEJ) or by homologous recombination (HR) (Altieri et al. 2008). For HR repair, CTCF has been identified as having a role in the facilitation of double-strand break response (Hilmi et al. 2017; Han et al. 2017; Natale et al. 2017; Lang et al. 2017; Hwang et al. 2019). Recent research indicates that CTCFs contribution to the mediation of long-range interactions likely creates a domain for repair within a loop (Natale et al. 2017). These researchers also found that CTCF depletion led to a reduction in the number of γH2AX nano-

foci. In addition, research by Hilmi et al (2017) has shown a direct role in HR for CTCF, which recruits BRCA2 to double-strand breaks; once BRCA2 is present, it recruits RAD51, which facilitates strand invasion for HR (Liu et al. 2010; Baumann et al. 1996). Notably, Hilmi et al. (2017) found that the recruitment of BRCA2 is dependent on the PARlation of CTCF to facilitate HR.

7.3.5 Insulator

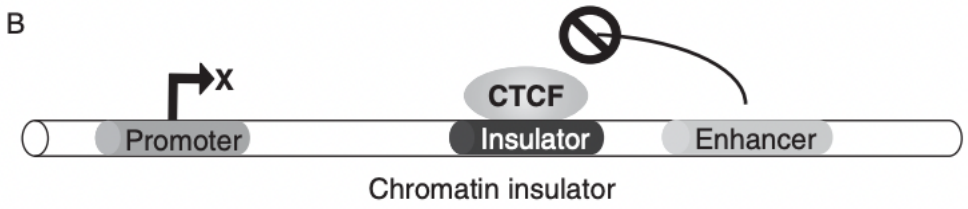


Figure VI. CTCF binds to insulators and functions as a chromatin insulator protein. CTCF prevents interaction between a promoter and nearby enhancers or silencers (Bell and Felsenfeld, 2000; Bell et al., 1999, 2001; Hark et al., 2000; Kanduri et al., 2000; Fillipova et al. 2008).

An insulator is defined as a DNA element, generally bound by CTCF, that is located between an enhancer and gene promoter; this prevents interactions between the two, thereby prohibiting transcriptional activation. DNA sequences from drosophila scs and scs' (insulator) elements were the first to be described as insulators by Udvardy et al. (1985). Later research by Bell et al (1999) has demonstrated that CTCF is able to direct enhancer insulating in vertebrates. CTCF was thought to act as an enhancer blocker protein when insulator elements were found upstream of ß-globin locus in chicken; additionally, they identified a CTCF site that plays a role in the insulator elements activity. CTCF's role as an insulator/blocking element was further solidified by research by Hark et al. (2000) where they determined that Igf2 imprinting requires a

blocking element between the gene and enhancer; they further argued that this element is likely mediated by CTCF binding to the imprinting control region.

7.3.6. Loss of CTCF

While most tumors suppressors necessitate a loss of both alleles to promote tumor initiation, loss of a single copy of CTCF triggers epigenetic changes and markedly increases lymphoma development and progression in mice (Kemp et al. 2014). This study also analyzed TCGA data and found CTCF to be deleted in ~50% or more of all human breast tumors. However, CTCF heterozygous mice did not display clear effects on solid tissue. A study by Moore et al. (2012) demonstrated that a complete loss of CTCF to be embryonic lethal, while $CTCF^{+/-}$ and $^{+/+}$ mice appeared to have normal development. This demonstrated that at least one copy of CTCF is necessary for development. Additionally, other clinical studies suggest that loss of CTCF is associated with higher grade tumors (Akhtar et al. 2020; Kemp et al. 2014). Mechanistically, the oncogenic impact of CTCF CNL remains obscure, however, CTCF has been demonstrated as a haploinsufficient tumour suppressor.

Contrasting these data, previous studies have indicated that knockdown of CTCF in breast cancer cell lines is associated with reduced cell proliferation (Lee et al. 2017; Mustafa et al. 2015). Thus, there appears a gradient where low CTCF may be oncogenic, but a complete loss is detrimental.

7.4 *TP53*

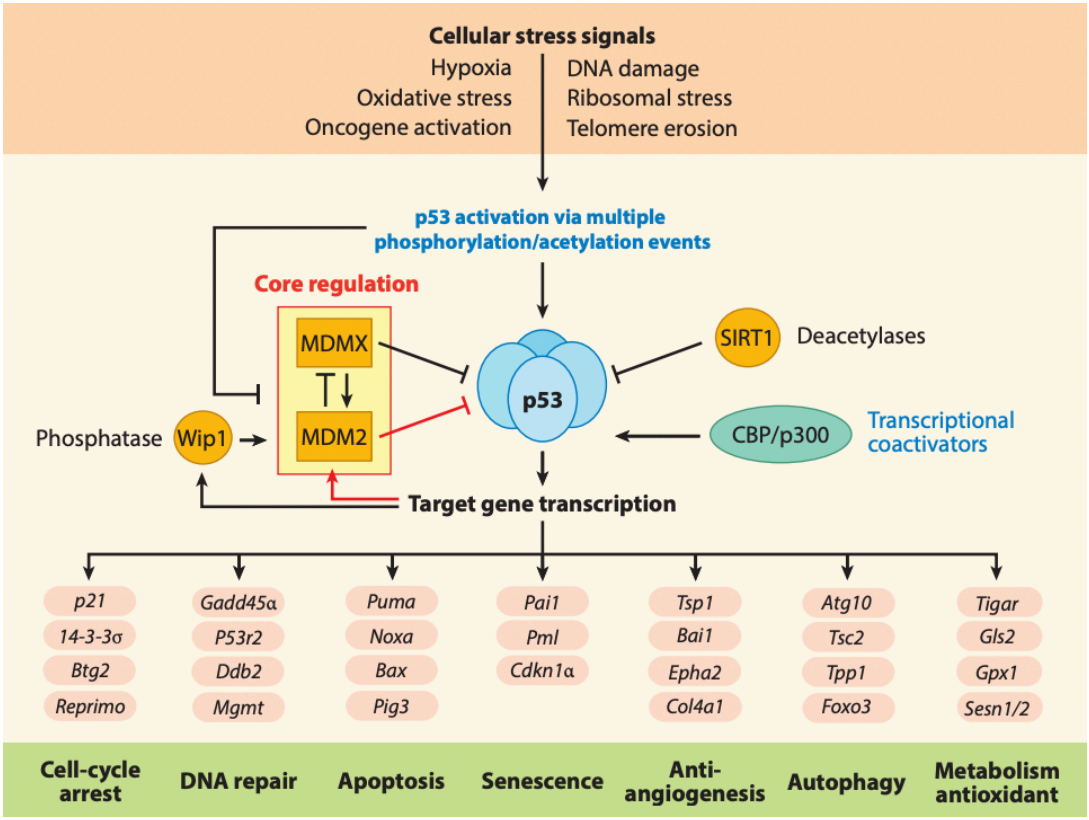


Figure VII. Overview of *TP53* associated pathways. *TP53* is activated following stress, including the induction of DNA damage. Genes associated with cell-cycle progression are downregulated while genes associated with DNA damage response are upregulated (Joerger et al. 2016).

TP53 is highly conserved and is the most frequently mutated gene across all cancer types (Sabapathy & Lane, 2019; Lawrence et al. 2014). It is commonly described as the "guardian of the genome". It was first described by several research groups in 1979 (Lane & Crawford, 1979; DeLeo et al. 1979; Kress et al. 1979) who were investigating Simian Virus 40 (SV40) cancer cells. The researchers found indicators of a protein at 53kDa that was a target of SV40. Additional research by Jörnvall et al. in 1982 found that Trp53 is highly conserved, which alluded to its essential role in cellular function. Later, several research groups transfected Trp53

into cells with the Ras oncogene and it was found to promote cellular transformation, thus Trp53 was thought to function as a wild-type oncogene (Eliyahu et al., 1984; Jenkins et al., 1984; Parada et al. 1984; Eliyahu et al. 1985). His finding turned out to be erroneous and what the researchers considered to be a wild-type Trp53 oncogene was later found to be a mutant protein. *TP53* was later re-classified as a tumor suppressor gene when *TP53* deletions were described in tumors from colorectal cancers (Baker et al. 1989). As a tumor suppressor protein, *TP53* was found to be activated in response to hypoxia, DNA damage, and loss of normal cell contacts (Fridman and Lowe, 2003).

7.4.1 Regulation

The transcription of several genes is regulated by *TP53* (Riley et al. 2008); *TP53* is known to influence the upregulation and downregulation on genes involved in a broad range of biological activities. The transactivation domain of *TP53* was initially characterized by three separate groups in 1990 (Fields & Jang. 1990; O'Rourke et al. 1990; Raycroft et al. 1990). *TP53* functions as a transcription factor through its transactivation domain as a homo-tetramer (Friedman et al. 1993). Following the induction of stress, *TP53* binds as a tetramer to a variety of targets which contain two half-site motifs; this leads to the transcription and expression of *TP53* associated genes. (Friedman et al. 1993; McLure & Lee. 1998; Nagaich et al. 1999). Typically, the binding site of *TP53* is found on the promoter region within 400 base pairs of the transcriptional start site, although this is not always the case (Tarasov et al. 2007). Notably, *TP53* mediated transcription is preceded by chromatin modifications to create a more accessible chromatin region at the binding site. *TP53* has been demonstrated to participate in the recruitment of several histone acetyl transferases (Barlev et al. 2001; Grossman. 2001).

TP53 functions as a transcription factor that primarily upregulates many genes in response to cell stress. This is a fast response, regulated by posttranslational modifications. As many as 36 amino acids of *TP53* have been found to be modified in several studies (Kruse & Gu. 2008). Researchers found that the human papilloma virus E6 protein induced the ubiquitination of *TP53* to avoid apoptosis (Everett et al. 1997; Scheffner et al. 1993). The E3 ubiquitin ligase, MDM2, is considered a major regulator of *TP53*. Activation of *TP53* by phosphorylation leads to an upregulation of MDM2, which in turn, feedbacks to downregulate *TP53*. MDM2 acts by targeting *TP53*s N-terminal to prevent transcription, and then by ubiquitinating *TP53*, leading to its degradation (Haput et al. 1997; Honda et al. 2000; Kubutat et al. 1997; Oliner et al. 1993). Notably, MDM2 binding alone is insufficient to supress *TP53* (Itahana et al. 2007). The MDM4 protein lacks the ubiquitin ligase activity found in MDM2. MDM4 is known to physically interact with MDM2, forming a stable and efficient heterodimer *in vitro* due to a RING domain (Tanimura et al. 1999; Wade et al. 2010). A previous study by Xiong et al (2006) which investigated the central nervous system of mice where MDM2 or MDM4 was not expressed. The results indicated that MDM2 and MDM4 appear to work in synergy to regulate Trp53.

Additionally, some studies indicate a direct role of CTCF in the regulation of *TP53*. A study by Soto-Reyes & Recillas-Targa (2010) demonstrated that CTCF binds to the promoter region of TP53, thus causing a more accessible chromatin conformation. The loss of CTCF lead to an increase in repressive histone marks in the upstream regulatory region of the sequence. Thus, CTCF was demonstrated to protect *TP53* from repressive marks. Upon the induction of damage, Wrap53 RNA was found to be involved in the induction of *TP53* expression in a study by Mahmoudi et al. (2009). Saldana-Meyer et al. (2014) sought to investigate whether these two pathways were associated. They determined that a partial loss of CTCF resulted in depleted

TP53 mRNA and Wrap53 levels. This suggested a CTCF-mediated regulation of *TP53* and Wrap53 following the induction of DNA damage.

7.4.2 Cell Cycle Arrest

After the induction of DNA damage, Trp53-mediated transcriptional activity is important to mediate cell cycle arrest that allows the cell time to repair damage (Chen, 2016). Some researchers hypothesize that the key mechanism of Trp53-mediated arrest is through transcriptional downregulation of many cell cycle genes (Engeland, 2018). This downregulation is thought to be indirect, as RNA-seq and ChIP-seq revealed that very few *TP53* downregulated genes have *TP53* binding (Fischer et al. 2014).

Alternatively, researchers have found that Trp53-induced G1/S arrest occurs when *TP53* upregulates the transcription of the cyclin-dependent kinase inhibitor, CDKN1A/p21 (el-Diery et al. 1994; Deng et al. 1995). As a result, CDKN1A inhibits cyclin-dependent kinase complexes that would otherwise phosphorylate p107 and p130 (Farkas et al. 2002). In addition, this inhibition of cyclin-kinase complexes may also cause G2/M arrest due to the upregulation of GADD45a which binds to CDK1, thus dissociating the cyclin-CDK1 complex (Zhan et al. 1999). Overall, the induction of cell cycle arrest is an essential mechanism which acts as a barrier to uncontrolled proliferation in addition to allowing the cell to facilitate DNA repair mechanisms.

In relation to CTCF, its association with *TP53* in regards to cell cycle progression has not been fully described in academic literature. CTCF is a repressive transcriptional factor of c-Myc gene expression (Lobanenkov et al. 1990). A study by Rasko et al. (2001) found that CTCF expression in several tumour lines led to growth arrest without apoptosis, however this could not be described only by the repression of c-Myc. Interestingly, a study by Sherr (1998) found that the $p19^{ARF}$ (ARF) promotor binding to CTCF led to increased transcriptional activity. One

mechanism by which *TP53* can be stabilized is by the inhibition of MDM2 activity via its association with ARF (Matheu et al. 2008). Although it may be indirect, CTCFs association with ARF may indicate a role in the cell cycle regulation by *TP53*.

7.4.3 Senescence

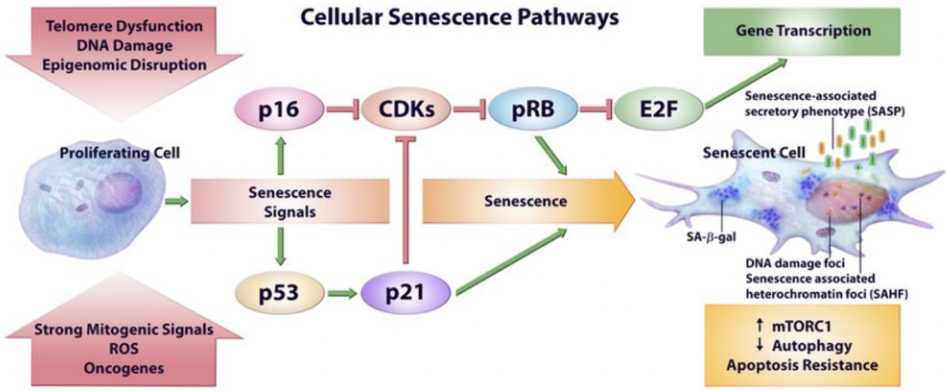


Figure VIII. An overview of cellular senescence. Various stressors can induce cellular senescence including DNA damage. The damage response pathway activates *TP53* which contributes to directly to the *TP53*/p21^{cip1} pathways or indirectly with the p16^{INK4a}/Rb pathway (Sultana et al. 2018).

In the simplest terms, senescence can be described as the halting or exhaustion of cell proliferative capacity; this is generally an irreversible process. A variety of factors can cause senescence including cell aging, reactive oxygen species, diminished telomere length, or DNA damage (Chandrasekaran et al. 2016; van Deursen. 2014; Takahashi et al. 2017). In the case of DNA damage, it causes a stress-induced premature senescence that may play a role in various disease states (Kritsilis et al. 2018).

The DNA damage response can directly cause the activation of *TP53* (Vanentine et al. 2011); One way this occurs is when ATM/ATR phosphorylates MDM2 and *TP53* to activate

TP53/p21^{cip1} pathway (Hu et al. 2012). As a CDK inhibitor, p21^{cip1} is necessary for TP53 mediated G1/S and G2/M cell cycle arrest checkpoints (Rufini et al. 2013). Importantly, p21^{cip1} is known to bind to many pro-apoptotic agents in order to promote cell senescence. A study by Zhang et al. (2011) showed that doxorubicin-induced damage caused *TP53* to inhibit p21^{cip1} in colorectal cancer cells.

Additionally, the maintenance of cellular senescence can be attributed to the p16^{INK4a}/Rb pathway (Rayess et al. 2012). Observations found that after the cells became senescent, *TP53* was depleted, however p16 levels remained high (Dolan et al. 2015); this suggests the importance of p16 in maintaining senescence. The activation of p16 is responsible for separating senescence dominated by *TP53*, which is reversible, and the p16^{INK4a}/Rb induced irreversible pathway (Helmbold et al. 2009).

7.4.4 Apoptosis

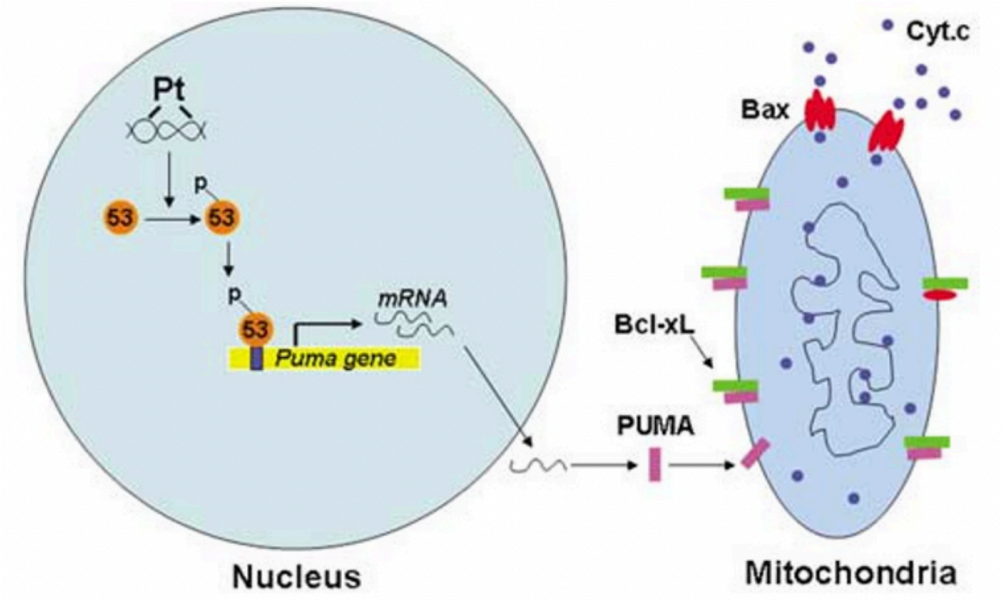


Figure IX. Schematic of *TP53* and PUMA mediated apoptosis. Following Cisplatin treatment of renal cells, activated *TP53* upregulates PUMA transcription, thus leading to the induction of apoptosis (Jiang et al. 2006).

Apoptosis is induced by either death receptor pathways or by mitochondrial pathways, which are mediated by BCL-2 proteins; these can either have pro-apoptotic or anti-apoptotic properties. *TP53* participates in transcription independent apoptosis in the mitochondria (Speidel, 2010).

The BCL-2 binding component 3 (BBC3), most well known as PUMA, was found to be a downstream target of *TP53* by several independent research groups (Han et al. 2001; Yu et al. 2001). Lys120 acetylation of *TP53* is a post-translational modification which mediates H4 acetylation at the PUMA promoter; thus, following DNA damage, PUMA expression is upregulated (Tang et al. 2006). PUMA activates the BAX and BAK effectors directly, causing mitochondrial outer membrane permeabilization and thus releasing apoptotic molecules such as cytochrome c into the cytoplasm. Cytochrome c will then bind to apoptotic protease activating

factor 1, thus causing apoptosis in cells (Green. 2018). In another study, Gomes and Espinosa (2010) demonstrated that upon treatment with the DNA damaging agent 5-Fluorouracil, CTCF binding is lost from PUMA, leading to its transcriptional upregulation. Thus, loss of CTCF may represent a *TP53*-independent mechanism by which PUMA is regulated.

7.4.5 DNA Repair

The absence of *TP53* is found in many cancer cases due to the lost ability to halt proliferation; thus, errors in the genome continue to be replicated with limited opportunity for repair. As a tumor suppressor, *TP53* also physically interacts with proteins to promote DNA damage repair processes, such as homologous recombination. Several studies have indicated a direct interaction between *TP53* and RAD51 (Arias-Lopez et al. 2006; Fong et al. 2011). Although *TP53*'s interaction with RAD51 is weak, Hine et al. (2014) found that its regulation of RAD51 expression is direct. *TP53* was also demonstrated to bind unphosphorylated RPA, forming a complex; upon induction of double strand breaks, RPA is phosphorylated, and the *TP53*-RPA complex is dissociated. Next, the phosphorylated RPA binds to ssDNA formed after end resection which has two purposes: blocking degradation and allowing recruitment for strand invasion (Serrano et al. 2013), which is critical for homologous recombination repair.

7.5 Chromatin Accessibility and Gene Expression

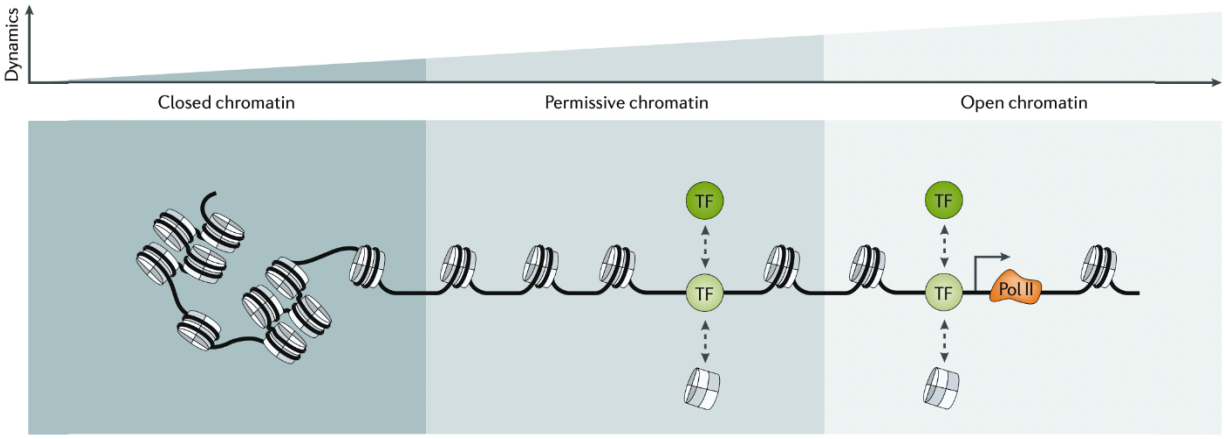


Figure X. Various states of chromatin accessibility dynamics on the genome. A nucleosome represents a unit of chromatin where DNA is wound around 8 histones to form one unit. Tightly packed nucleosomes are associated with closed chromatin with minimal transcriptional activity. Loosely packed nucleosomes are associated with open chromatin and generally have higher transcriptional activity (Klemm et al. 2019).

Chromatin accessibility is non-uniform and dictates the extent to which transcription factors and other nuclear elements can physically contact regions of the chromatin. Topological organization of the chromatin as well as chromatin binding factors can affect the degree of accessibility. Often post translational modifications of the nucleosome such as methylation and acetylation can indicate the organization of the chromatin (Allis & Jenuwein. 2016). Heterochromatin has low accessibility due to its condensed nucleosomes and generally has low transcriptional activity. Euchromatin has high accessibility with loosely packed nucleosomes and is often a region of high transcriptional activity. Although only 3% of the genome is accessible, around 90% of bound transcription factors are found in these regions (Thurman et al. 2012), thus the accessibility of the genome largely dictates the transcriptional activity.

CTCF binding occupancy is known to contribute to the regulation of transcriptional activity, however its association with changes in accessibility have not been fully described in the literature. Interestingly, in a study by Xu et al. (2021) it was found that a depletion of CTCF affected chromatin interactions and accessibility globally. The researchers theorized that CTCF's association with chromatin accessibility played a role in transcriptional regulation.

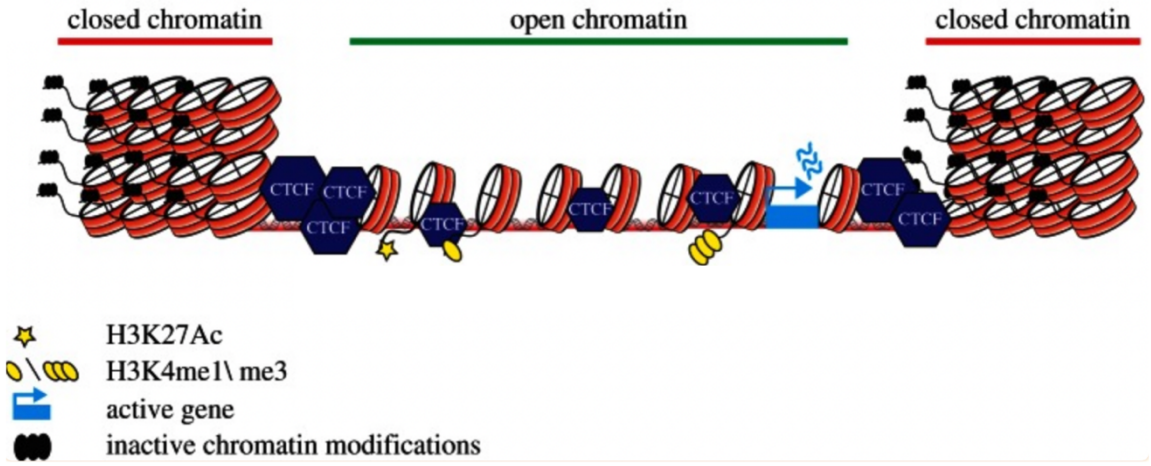


Figure XI. A versatile role for CTCF. CTCF binding sites can be found at boundaries between active and inactive chromatin regions (Holwerda & de Laat. 2013).

7.5 Assay for Transposase-Accessible Chromatin

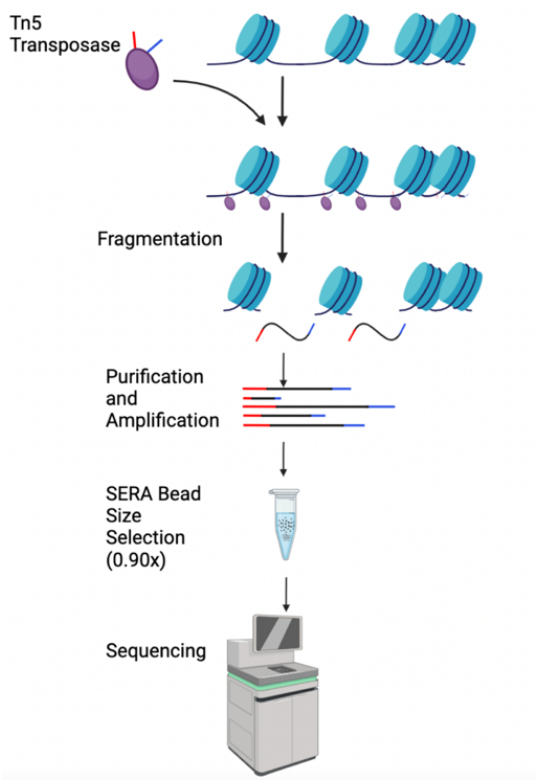


Figure XII. An assay for transposase accessible chromatin (ATAC-seq). Tn5 transposase flanks nucleosomes on open regions of chromatin. Cut fragments are purified and amplified for library preparation with custom barcode adapters. 0.09x bead size selection is used to remove fragments >200bp in order to eliminate the presence of primer dimers in sequencing. Visuals designed using BioRender.

One method used to describe the epigenetic landscape is ATAC-seq. The assay for transposase accessible chromatin (ATAC-seq) is a simple protocol, which only requires 50,000 cells and utilizes a Tn5 transposase to insert illumina adapters into regions of accessible chromatin. This technique was first described in 2013 by Buenrostro et al. in a study investigating human T cells. ATAC-seq can be utilized to profile regions of accessible chromatin, identifying nucleosome-bound positions, and can even identify transcription factor

"footprinting"; the information obtained from sequencing depends on the number of sequencing reads obtained. In addition, ATAC seq may also identify nucleosome position by utilizing fragments that represent nucleosome monomers. For the purposes of this project, at least 50 million reads per pair is sufficient to obtain data describing regions of accessible chromatin.

7.6 Aim

Based on clinical data available from 1217 breast tumors obtained from the TCGA, we looked at 116 classical *TP53* target genes (Fischer, 2017). Our review found that target genes that are positively correlated with high levels of *TP53* expression, tend to be negatively correlated with CTCF expression (Figure S1). This suggests an inverse correlation between CTCF levels and *TP53* target gene expression. We hypothesize that CTCF may play a role in the negative regulation of *TP53* target genes. A previous study indicates a repressive role for CTCF in regulating the classical *TP53* target gene BBC3/PUMA, but little is known about a role for CTCF in modulating the *TP53* response beyond this (Gomes and Espinosa, 2010).

As single copy loss of CTCF is found in about 50% of breast cancer patients, it is possible that the loss of CTCF results in deregulation of *TP53* target genes in these patients, possibly potentiating *TP53*-mediated response. Thus, loss of CTCF could potentially predict enhanced responses to chemotherapeutics in *TP53* WT tumors. In this project, we will explore the effects of CTCF copy number loss on chromatin accessibility and *TP53*-mediated transcription. We hypothesize that changes in chromatin accessibility after CTCF CNL may impact the *TP53*-mediated transcriptional response.

8. Methods

8.1 Cell Culture

Immortalized MCF10A cells were utilized to avoid the confounding factors that would need to be considered with tumor cells. MCF10A cells were cultured in DMEM/F12 50/50 mix media (Wisent, cat# 319-085-CL) supplemented with 2% horse serum (Wisent, cat# 065150), 0.5µg/mL hydrocortisone (Sigma, cat# H0888-1G), 0.02µg/mL epidermal growth factor (Wisent, cat# 511-110-UM), 0.01mg/mL insulin (Wisent, cat# H511-016-U6) and 0.1µg/mL cholera toxin (Sigma, cat# C8052-2MG). MCF10A, cell cultures were maintained at 1:6 and discarded once they reach passage 4. 67NR cells were cultured in DMEM (Wisent, cat#319-005-CL) supplemented with 10% fetal bovine serum (Gibco, cat# 12483-020). For CTCF single copy loss cell lines previously generated by CRISPR/Cas9-mediated gene knockdown, single allele knockdown was confirmed by sanger sequencing and lower protein expression of CTCF was confirmed by western blot (Figure S2).

All cell lines were cultured at 37°C incubator with 5% CO_2. For treatment of cells with chemotherapeutic agents, cells were treated with 6µM of cisplatin (Jewish General Hospital, Montreal) for durations indicated.

8.2 RNA Isolation

After treatment, media was aspirated from each well of 6 well plate. 350µL of lysis buffer (Sigma Aldrich, cat# L8285-350mL) containing 10µL/mL beta-mercaptoethanol (Sigma Aldrich, cat# M3148-2mL) was added to each well. RNA was isolated according to protocol from Aurum™ Total RNA Mini Kit (Biorad, cat# 732-6820) and eluted in 40µL of nuclease-free water. Concentration of purified RNA was measured using nanodrop and stored at -80°C.

8.3 Polymerase Chain Reaction (PCR)

PCR was performed following the protocol from Advantech 2X Hot-Start PCR MasterMix, With Dye (Diamed, cat# AD100-12102). 50 to 100ng of DNA template was used for each reaction.

8.4 Reverse Transcription Quantitative Polymerase Chain Reaction (RT-qPCR)

Reverse transcription was performed according to protocol using Advantech 5X Reverse Transcription Mastermix (Advantech, cat# AD100-31401). Briefly, a total of 100ng of RNA is added to 4µL of 5X All-In-One RT Mastermix and topped up to 20µL per reaction with nuclease-free water. cDNA synthesis is carried out in a PCR machine (Biorad, T100 Thermal Cycler) using the following steps: 25°C for 10mins, 50°C for 60mins, 85°C for 5mins and finally hold at 4°C. Resulting cDNA is diluted 10x in nuclease-free water and then stored at -20°C until usage. GoTaq® qPCR Master Mix (Promega, cat# A6001) was used for qPCR. 5µL of master mix was added to 1µL of 0.5µM forward and reverse primer mix, and then 2µL of cDNA was added and topped up with 2µL of nuclease-free water to 10µL total volume per reaction. qPCR was performed in a qPCR machine (Applied Biosystems, QuantStudio 3) using the following steps: heated lid at 105°C, 50°C for 2mins, 95°C for 2mins, 40 cycles of 95°C for 15secs and 60°C for 1min, followed by melt curve stage of 95°C for 15secs, 60°C for 1min and 95°C for 15secs. Results were analysed using QuantStudio™ Design & Analysis Desktop Software v1.5.1 (Thermo Fisher Scientific).

qPCR Primers for MCF10A & 67NR:

Target	**Forward sequence**	**Reverse sequence**
B-actin	AGGCACCAGGGCGTGAT	GCCCACATAGGAATCCTTCTGAC
GAPDH	ACAGTCAGCCGCATCTTCTT	ACGACCAAATCCGTTGACTC
RPL4	GCTCTGGCCAGGGTGCTTTTG	ATGGCGTATCGTTTTTGGGTTGT

RPLP0	TTAAACCCTGCGTGGCAATCC	CCACATTCCCCCGGATATGA
18S	GCTTAATTTGACTCAACACGGGA	AGCTATCAATCTGTCAATCCTGTC
BBC3	GCAGGCACCTAATTGGGCT	ATCATGGGACTCCTGCCCTTA
BAX	GGTTGTCGCCCTTTTCTACT	AAGTCCAATGTCCAGCCCAT
CDKN1A	GACTCTCAGGGTCGAAAACG	GGATTAGGGCTTCCTCTTGG

8.5 Western Blotting

Cells are cultured in 10 cm plastic dishes until 80% confluent. Treated and untreated cells were collected by first aspirating the media, following by addition of ice-cold PBS. A cell scraper is used to scrape the cells off the dish and the cell suspension is transferred into a 1.5mL tube. The cells are centrifuged at 2,500 rpm for 5 minutes at 4°C. The supernatant was removed, and cell pellet used for lysis immediately, or flash frozen with liquid nitrogen and kept at -80°C until required. Cell pellet was lysed using 100µL of lysis buffer [10% glycerol, 0.5% NP-40, 0.5% Triton X-100, 420mM NaCl, 20mM Tris (pH 7.5), 2mM MgCl2, 1mM EDTA, 1mM dithiothreitol, 2mM phenylmethylsulfonyl fluoride (PMSF), 1mM P8340 Cocktail inhibitor (Roche), 1mM bis-glycerol phosphate, and 1mM NaF] on ice for 15 minutes, with agitation at intervals. Cell suspension is then centrifuged at 13,000 rpm for 15 minutes at 4°C. The resulting supernatant was transferred to a new 1.5mL tube. To measure protein concentration, samples were diluted 20x with distilled water (2µL sample in 38µL water). Then, 10µL of diluted protein was added to 200µL of Bradford reagent (Thermo Fisher, cat# 1856209) and mixed in a 96-well plate. After 5mins, the 96-well plate was read using plate reader at an absorbance of 595nm. Protein standards were prepared at 500µg/mL, 250µg/mL, 125µg/mL, 62.5µg/mL, 31.25µg/mL and 0µg/mL with 2mg/mL of BSA (Thermo Scientific, cat# 23209) diluted with distilled water.

A total of 25µg of proteins were added to 6x loading buffer and loaded onto an 8% gel. The gel was run at 100V for 1hr and transferred to a nitrocellulose membrane (Pall, cat# 66485) at 100V for 1hr at 4°C. After the transfer, the blot was blocked with 5% skimmed milk/BSA in Tris-Buffered Saline in 0.1% Tween-20 (0.1% TBST) [20mM Tris, 137mM NaCl, and 0.1% Tween-20] for at least 1 hour at room temperature, or overnight at 4°C. Primary antibody in 5% skimmed milk/BSA was added for overnight at 4°C. The blot was then washed with 0.1% TBST 3 times for 10 minutes each wash. Secondary antibody in 5% skimmed milk/BSA was added to the blot for 1 hour at room temperature. The blot was washed again with 3 times of 0.1% TBST for 10mins each wash and incubated with ECL substrate (Biorad, cat# 170-5061) for film exposure.

Antibodies for Western Blotting:

Target	**Source**	**Dilution factor**
p53 (DO-7)	Cell Signalling #48818	1:10000
Phosphorylated-p53 (Ser15)	Cell Signalling #9284	1:4000
MDM2	Cell Signalling #86934	1:2000
β-actin	Sigma Aldrich #A2228	1:5000
GAPDH (14C10)	Cell Signalling #2118	1:5000
CTCF	BD Biosciences #612149	1:1000

8.6 Assay for Transposase-Accessible Chromatin Sequencing

ATAC-RSB (Resuspension Buffer)

Reagent	**Final Concentration**	**Volume for 50 mL**

1M Tris-HCl pH 7.4	10 mM	500 µL
5M NaCl	10 mM	100 µL
1M $MgCl_2$	3 mM	150 µL
Sterile Water	N/A	49.25 mL

The protocol in 8.6 and 8.6.1 are based on Corces et al. 2017; Buenrostro et al. 2015; and utilizes an existing protocol from the Howard Chang Lab. After ensuring the MCF10A cells are ideally 90-95% viable, the media was aspirated and washed with 5ml D-PBS (Wisent, 311-425-CL) which was aspirated and 2ml TRYPSIN/EDTA (Wisent, 325-043-EL) was added. The cells were placed in an incubator until the cells began to detach. Plates were removed from the incubator and 4ml of supplemented DMEM/F12 50/50 mix media (Wisent, cat# 319-085-CL) was added. 6ml of cells in media were transferred to a tube and centrifuged at 1200 RPM for 4 minutes. The supernatant was aspirated, and the cell pellet was resuspended in 6 mL of the supplemented DMEM/F12. Cells were counted and 50,000 viable cells were transferred to a 1.5 mL Eppendorf. For best results, the cells were used fresh and were not stored before proceeding to the transposition reaction.

Before starting the transposition, a 50 mL stock of the ATAC resuspension buffer (ATAC-RSB) was prepared. 500 µL 1M Tris-HCl pH 7.4 (10 mM final), 100 µL 5M NaCl (10mM final), 150 µL 1M $MgCl_2$ (3mM final), and 49.25mL sterile water was combined to make the ATAC-RSB and was immediately placed on ice. Next, the ATAC-RSB-1 lysis buffer was prepared. 50 µL (per sample) ATAC-RSB was supplemented with 0.1% NP40 (IGEPAL® CO-630 at 10% stock concentration; Sigma-Aldrich, 68412-54-4), 0.1% Tween-20 (at 10% stock concentration; Sigma-Aldrich, 9005-64-5), and 0.01% Digitonin (Diluted 1:1 with water to make

a 1% stock solution; Promega, G9441) and was put on ice. Next, the ATAC-RSB-2 was prepared with 1ml ATAC-RSB (per sample) containing 0.1% Tween-20 (at 10% stock concentration; Sigma-Aldrich, 9005-64-5) and was put on ice.

The cells (50,000) were pelleted at 500 RCF at 4°C for 5 minutes in a fixed angle centrifuge. The supernatant was aspirated in multiple steps and with great caution to avoid the small cell pellet, using a p200 pipet. 50 µL cold ATAC-RSB-1 was added per sample and pipetted up and down 3 times. It was then incubated on ice for 3 minutes. The lysis was then washed out with 1 mL cold ATAC-RSB-2 per sample and the tube was inverted 3 times. The nuclei were pelleted at 500 RCF for 10 minutes at 4°C in a centrifuge. All supernatant was carefully aspirated to avoid the small cell pellet, using multiple pipetting steps. The cells were resuspended in the transposition mixture; 25 µL 2x TD buffer (Illumina, 20034210), 2.5 µL transposase (100nM final, Illumina, 20034210), 16.5 µL D-PBS (Wisent, 311-425-CL), 0.5 µL 1% digitonin (Promega, G9441), 0.5 µL 10% Tween-20 (Sigma-Aldrich, 9005-64-5), and 5 µL sterile water. The transposition mixture was pipetted up and down 6 times. The transposition reaction was incubated at 37°C for 24 minutes. The Eppendorf was agitated every 4 minutes. Following the reaction, the samples were immediately put on ice.

8.6.1 Library Preparation

Purification

Following transposition, a DNA purification reaction was performed using the Zymo DNA Clean and Concentrator-5 Kit (D4014). Following the kits instruction, in a 1.5 mL tube, add 5 volumes (~250 µL) of DNA binding buffer to each volume of DNA sample. Mix by vortexing briefly. Transfer the mixture to the provided Zymo-Spin Column in a collection tube.

Centrifuge for 30 seconds at 12,000 RPM and discard flowthrough. Add 200 µL DNA wash buffer to the column. Centrifuge at 12,000 RPM for 30 second. The wash step was repeated. Transfer the column to a 1.5 mL tube and elute the DNA in 12 µL of elution buffer for 2 minutes. It was centrifuged for 1 minute at 12,000 RPM. The elution step was repeated. The eluted DNA (24 µL) was stored at -20°C until ready for amplification or used immediately.

Pre-Amplification

The PCR pre-amplification mixture was prepared using 2.5 µL of the 25 µM Primer Ad1 (IDT, Custom Illumina Adapters), 2.5 µL of the 25 µM Primer Ad2.x (A unique custom Ad2.x is used for each sample, IDT), 25 µL 2X NEBNext Master Mix (New England BioLabs, M0541S), and 20 µL of the transposed sample. The samples are amplified by PCR under the following conditions: 72°C for 5 minutes, 98°C for 30 seconds, then 5 cycles of 98°C for 10 seconds, 63°C for 30 seconds, 72°C for 1 minute, and put on hold at 4°C.

qPCR Amplification to Determine Additional Cycles

According to Buenrostro et al. in 2015, In order to reduce size bias and GC bias in PCR, the number of additional PCR cycles was determined with qPCR to avoid stopping amplification prior to saturation. Using. the pre-amplified mixture, a qPCR amplification mixture was prepared with 4 µL sterile water, 0.5 µL 25µM Primer Ad1 (IDT, custom adapters), 0.5 µL 25µM Primer Ad2.x (IDT, custom adapters), 2x GoTaq PCR mix (Promega, A600A), and 2 µL of the pre-amplified sample. The qPCR was run under the following cycling conditions: 72°C for 5 minutes, 98°C for 30 seconds; then 40 cycles of 98°C for 10 seconds, 63°C for 30 seconds, 72°C for 1 minute, and put on hold at 4°C. After qPCR amplification, manually assess the amplification profiles and determine the requirednumber of additional cycles to amplify. To calculate the additional number of cycles needed (N), linear Rn versus cycle was plotted and

used to determine the cycle number that corresponds to ¼ of the maximum fluorescent intensity (Buenrostro et al. 2015).

PCR Amplification

The remaining PCR pre-amplification mixture were amplified by PCR under the following conditions: 98°C for 30 seconds, then N cycles of 98°C for 10 seconds, 63°C for 30 seconds, 72°C for 1 minute, and put on hold at 4°C. N is the number of cycles determined by qPCR.

Post-Amplification Purification

Following amplification, a DNA purification reaction was performed using the Zymo DNA Clean and Concentrator-5 Kit (D4014). Following the kits instruction, in a 1.5 mL tube, add 5 volumes of DNA binding buffer to each volume of DNA sample. Mix by vortexing briefly. Transfer the mixture to the provided Zymo-Spin Column in a collection tube. Centrifuge for 30 seconds at 12,000 RPM and discard flowthrough. Add 200 µL DNA wash buffer to the column. Centrifuge at 12,000 RPM for 30 second. The wash step was repeated. Transfer the column to a 1.5 mL tube and elute the DNA in 12 µL of nuclease-free water for 2 minutes. It was centrifuged for 1 minute at 12,000 RPM. The elution step was repeated. The eluted DNA was stored at –20°C.

Library Preparation Barcode Adapters

Primer Name	Sequence
Ad1_noMX	AATGATACGGCGACCACCGAGATCTACACTCGTCGGCAGCGTCAGATGTG
Ad2.1_TAAGGCGA	CAAGCAGAAGACGGCATACGAGATTCGCCTTAGTCTCGTGGGCTCGGAGATGT
Ad2.2_CGTACTAG	CAAGCAGAAGACGGCATACGAGATCTAGTACGGTCTCGTGGGCTCGGAGATGT
Ad2.3_AGGCAGAA	CAAGCAGAAGACGGCATACGAGATTTCTGCCTGTCTCGTGGGCTCGGAGATGT

Ad2.4_TCCTGAGC	CAAGCAGAAGACGGCATACGAGATGCTCAGGAGTCTCGTGGGCTCGGAGATGT
Ad2.5_GGACTCCT	CAAGCAGAAGACGGCATACGAGATAGGAGTCCGTCTCGTGGGCTCGGAGATGT
Ad2.6_TAGGCATG	CAAGCAGAAGACGGCATACGAGATCATGCCTAGTCTCGTGGGCTCGGAGATGT
Ad2.7_CTCTCTAC	CAAGCAGAAGACGGCATACGAGATGTAGAGAGGTCTCGTGGGCTCGGAGATGT
Ad2.8_CAGAGAGG	CAAGCAGAAGACGGCATACGAGATCCTCTCTGGTCTCGTGGGCTCGGAGATGT
Ad2.9_GCTACGCT	CAAGCAGAAGACGGCATACGAGATAGCGTAGCGTCTCGTGGGCTCGGAGATGT
Ad2.10_CGAGGCTG	CAAGCAGAAGACGGCATACGAGATCAGCCTCGGTCTCGTGGGCTCGGAGATGT
Ad2.11_AAGAGGCA	CAAGCAGAAGACGGCATACGAGATTGCCTCTTGTCTCGTGGGCTCGGAGATGT
Ad2.12_GTAGAGGA	CAAGCAGAAGACGGCATACGAGATTCCTCTACGTCTCGTGGGCTCGGAGATGT
Ad2.13_GTCGTGAT	CAAGCAGAAGACGGCATACGAGATATCACGACGTCTCGTGGGCTCGGAGATGT
Ad2.14_ACCACTGT	CAAGCAGAAGACGGCATACGAGATACAGTGGTGTCTCGTGGGCTCGGAGATGT
Ad2.15_TGGATCTG	CAAGCAGAAGACGGCATACGAGATCAGATCCAGTCTCGTGGGCTCGGAGATGT
Ad2.16_CCGTTTGT	CAAGCAGAAGACGGCATACGAGATACAAACGGGTCTCGTGGGCTCGGAGATGT
Ad2.17_TGCTGGGT	CAAGCAGAAGACGGCATACGAGATACCCAGCAGTCTCGTGGGCTCGGAGATGT
Ad2.18_GAGGGGTT	CAAGCAGAAGACGGCATACGAGATAACCCCTCGTCTCGTGGGCTCGGAGATGT

8.6.2 Size Selection

The aim was to select for DNA greater than 200 bp in order to eliminate primer-dimers (<200 bp), which would interfere with sequencing. Starting with 20 µL of the amplified DNA, add 18 µL of room temperature DNA Size Selection and PCR clean up kit (SERA-Mag Select, 29343052) so the ratio of bead solution to DNA is now ~0.9 to 1. Mix well until the solution is homogenous. The solution was allowed to incubate at room temperature for 8 minutes. Following incubation, the lid was carefully opened, and the tube was placed on a magnetic

separation rack until the beads separate and the solution becomes clear. The beads have now captured the required size of DNA and the supernatant was discarded. 100 µL of freshly prepared 80% ethanol was added slowly and gently to the tube on the magnetic separation stand. The beads were given time to separate again, and the supernatant was discarded. This was repeated for a total of two wash steps. The tube lid was gently opened, and the beads were allowed to dry at room temperature for no more than 5 minutes, or until the colour began to shift. 30 µL of TLE buffer (1 mL 1M Tris-HCl pH 8.0, 20 µL 0.5M EDTA pH 8.0, and 98.8 mL nuclease-free water) was added to each sample and mixed well. It was allowed to incubate with the lid closed at 37°C for 10 minutes and then 10 minutes at room temperature. The lid was carefully opened and put back onto the magnetic separation rack until the solution became clear. The supernatant was transferred to a new tube and the beads were discarded.

8.6.3 DNA Agarose Gel

To confirm the success of the size selection, 1.2% Agrose gel was prepared using 1.2g Agrose D1-LE (Wisent. 800-015-CG) and 100 mL 0.5X TBE buffer (Wisent, 880-545-LL), which was then heated until the solution become homogenous. Once the solution had cooled to ~60°C, and 10 µL (1:10,000) GelGreen Nucleic Acid Gel Stain (Biotium, 41005) was added and carefully mixed to avoid bubbles until the solution was homogenous and was poured into a 100 mL gel mold with wells inserted and allowed to solidify. In a gel electrophoresis tray, the solidified gel was carefully placed and the 0.5X TBE buffer was added until it was submerged slightly. The well-mold was carefully removed and a 100bp DNA Ladder was added (SMOBio, DM2100) and 12-14ng of DNA was added to the wells with10X FastDigest Green Buffer (Thermo Fisher, B72). The electrophoresis tray was attached to a PowerPac™ Basic Power

Supply machine (BIO-RAD, 1645050) and allowed to run to completion. Gels were imaged using a ChemiDoc Imaging System (BIO-RAD, 12003153).

8.6.4 Picogreen DNA Concentration

Due to the small concentration of DNA in ATAC-seq, the Quan-iT PicoGreen dsDNA Assay Kit (Thermo Fisher Scientific, P7589) was utilized to measure concentration. From the kit, 1X TE buffer was prepared from the 20X stock solution. DNA standards were prepared by diluting the stock DNA that is provided in the kit in 1X TE solution (10 ng/µL, 5 ng/µL, 2 ng/µL, 1.5 ng/µL, 1.0 ng/µL, 0.75 ng/µL, 0.5 ng/µL, 0.25 ng/µL, 0.1 ng/µL and 0.05 ng/µL.). 5 uL of 1X TE was pipetted into each well of a black 384 well Assay Plate (Corning, 4514). The number of wells necessary were calculated depending on the number of samples; each sample standard was run in duplicate. 1µl of the DNA sample and 1 µL of the standard were added to the corresponding well containing 5 µL of 1X TE buffer. In a dark area, a 1:200 dilution of the stock PicoGreen was prepared in 1X TE buffer. 5 µL of the diluted PicoGreen was pipetted into each well. The amount of DNA in each sample was quantified using a EnSpire Multimode Plate Reader. This protocol was developed, and modified utilizing recommendations and advice given by Korin Sahinyan.

8.6.5 Sequencing

Samples were sequenced by The Hospital for Sick Children at The Center for Applied Genomics. Quality control Bioanalysis was performed prior to sequencing for both untreated (Figure S4) and treated (Figure S5) samples. DNA was sequenced using Illumina NovaSeq -S1 flow cell. The conditions were paired-end sequencing 2x150bp reads. We required a minimum of 50 million pairs of reads per sample and this service yielded ~ 70-88 million pairs of reads per

library. NCBI GEO Series record: GSE217698, available November 10, 2023 (Fastq and Bigwig files).

9. Results

Genes with greater accessibility on the TSS and TTS are significantly enriched within TADs

To gain better insight into the impact of CTCF haplo-insufficiency on gene expression, we performed RNA-seq analysis. We found that changes in gene expression of our $CTCF^{+/-}$ #1 and #2 clones relative to the control MCF10A cells, are strongly consistent, thus, we used $CTCF^{+/-}$ #1 as a model system to study chromatin accessibility (Figure S3). As previously described, accessibility of the transcription start site (TSS) is a good proxy to describe transcriptional activity. Thus, we aimed to investigate the accessibility of all genes on the TSS and its association with upregulated gene expression of MCF10A $CTCF^{+/-}$ #1 cells before the induction of DNA damage. We crossed our pre-treatment ATAC-seq data with previously obtained RNA-seq for MCF10A to evaluate any significant association between expression and accessibility. Untreated $CTCF^{+/-}$ #1 cells were compared to the WT on genes where the TSS was differentially accessible on all genes compared to all genes ($r=0.2580$, $p<0.0001$) where a significant but mild association was found (Figure 1a). Next, we filtered for differentially expressed genes where the TSS had increased accessibility basally in the untreated $CTCF^{+/-}$ #1 compared to the WT and found a significant correlation between increased TSS accessibility and gene expression ($r=0.4520$, $p<0.001$). In addition, proportionally more genes were upregulated when the TSS was more accessible (Figure 1b). Finally, we investigated genes where both the TSS and transcription termination site (TTS) had increased accessibility in the $CTCF^{+/-}$ #1 compared to the WT and did not find a significant

correlation. Despite the lack of correlation, most genes where both the TSS and TTS had increased accessibility were upregulated compared to genes which only had increased accessibility on the TSS (Figure 1c). As expected, accessibility of the TSS is associated with gene overexpression in the $CTCF^{+/-}$ #1 compared to the WT.

Next, we compared the relationship between gene enrichment at TAD boundaries or within TADs to chromatin accessibility at the TSS and TTS. First, we investigated the proportion of all genes located either on TAD boundaries (2757) or within TADs (9166) (Figure 1d). We filtered for genes with open chromatin on the TSS and found proportionally more genes within TADs (473) than on TAD boundaries (122) (Figure 1e). Finally, we investigated genes with open chromatin on both the TSS and TTS (Figure 1f) and found that most genes were enriched within TADs (70) compared to genes on TAD boundaries (5). As expected, the majority of genes with increased TSS activity were enriched within TADs; these finding are consistent with the literature.

Five *TP53* target genes have increased accessibility on both the TSS and TTS after the induction of DNA DSB

We next sought to induce DNA damage to investigate differential accessibility on the TSS and TTS as a representation of transcriptional activity. We chose to utilize Cisplatin to cause the induction of DNA DSB; previous members of our lab determined that 6μM treatment of MCF10A cells for 8h was sufficient to cause heightened *TP53* associated gene expression (Figure S2). This treatment causes DNA adducts that form intra- and interform crosslinks; the interstrand crosslinks stall replication forks and result in DNA DSB (Kartalou & Essigmann. 2001; Enoiu et al. 2012). We confirmed the success of our treatment wherein *TP53* associated genes tended to have increased chromatin accessibility following treatment (Figure 4a). We

investigated the accessibility of all genes on the TSS and its association with upregulated gene expression for MCF10A CTCF$^{+/-}$ (#1) cells after induction of DNA damage with 6μM Cisplatin for 8 hours. We crossed our post-treatment ATAC-seq data with previously obtained RNA-seq for MCF10A to evaluate any significant association between expression and accessibility. For all differentially expressed genes (11923), there was a significant but mild association between the accessibility of the TSS and gene expression following damage (r=0.03147; p=0.0006) (Figure 2a). The 11923 genes from figure 2a were filtered to only include genes that have upregulated gene expression (644), and a significant but mild association was found (r=0.03887; p=0.0006) (Figure 2b). Further analysis found that thirteen genes (including, but not limited to *TP53* associated genes) were found to have significantly increased accessibility on the TSS. Of the 13 genes with increased accessibility of the TSS in this group, it is worth noting that the *TP53* target genes FDXR and BBC3 had significantly increased accessibility of both the TSS and TTS ($p<0.05$). This was unexpected as both genes are described as *TP53* associated even though this analysis was not limited to only *TP53* associated genes.

As an alternative analysis, we modified the significance threshold to $p \leq 0.25$ to increase the proportion of genes with more open TSS after damage to investigate more *TP53* target genes. This change only resulted in marginal accessibility changes (Figure 2c, e). Focusing on just the upregulated genes (644), we found 57 genes (globally) with increased accessibility of the TSS in the knockdown; interestingly, we found that 5 of the 57 genes with increased TSS accessibly also had increased accessibility at the TTS. These were BBC3, CDKN1A, GDF15, FDXR, and NECTIN4. Notably, although this analysis was not exclusive to *TP53* genes and included a wide variety of genes, 4 out of these 5 are described *TP53* associated genes (Figure 2e). Notably, this is a very small subset and likely represents a very specific response.

Finally, we investigated genes with differentially accessible TSS following damage to find the proportion of genes that were found within TADs or were constant/gained/lost on TAD boundaries. Overall, when comparing genes with an open TSS for all differentially expressed genes (641) (Figure 2d) to only upregulated genes (57) (Figure 2f), there is a marginal difference. Thus, genes with an open TSS after damage are not enriched within TADs.

Accessibility of the TSS after exposure to cisplatin

Since we are interested in transcriptional activity, we next chose to review accessibility of the transcription start site (TSS), a critical regulatory region for gene activation. As mentioned previously, although ~3% of the genome is accessible during interphase, around 90% of transcription factors are found within these regions, thus accessibility is used as a proxy to determine transcriptional activity. Here, we sought to evaluate only the accessibility profile obtained from ATAC-seq without any comparison to RNA-seq. First, we sought to compare the accessibility of the TSS globally between WT and $CTCF^{+/-}$ following treatment with 6μM Cisplatin for 8h. Before comparting the Log Fold Change (LogFC) between groups, we sought to visualize the global TSS accessibility of most of the annotated protein coding genes, totalling 14254 (Figure 3a). In this heatmap, each row represents a gene, and the columns represent the following: MCF10A WT untreated, WT treated, $CTCF^{+/-}$ #1 untreated, $CTCF^{+/-}$ #1 treated, $CTCF^{+/-}$ #2 untreated, $CTCF^{+/-}$ #2 treated. The signal was determined by using the standard error of the mean (SEM) for the triplicate data which indicates the scaled read counts for the TSS of a gene; intensity of colour in the figure is related to strength of the signal, with darker red being associated with a stronger signal and thus a greater degree of accessibility. To assess how many genes were open at the TSS, we decided to label a signal that is greater than the average signal for all genes to be “open”. We determined the average SEM of the signal of all genes basal

condition to determine the average accessibility of the TSS, and values greater than the average were deemed open for the following: WT untreated (4659), WT treated (6968), CTCF$^{+/-}$ #1 untreated (4888), CTCF$^{+/-}$ #1 treated (5358), CTCF$^{+/-}$ #2 untreated (4462), CTCF$^{+/-}$ #2 treated (4919) (Figure 3a). Upon a more specific review, several key genes that were expected to be upregulated following the induction of damage, such as BBC3, were found to have increased accessibility of the TSS following treatment.

In addition, we utilized volcano plots to characterize differentially accessible TSS for untreated MCF10A WT and CTCF$^{+/-}$ (#1 and #2) cells. Here, we compare the accessibility for the untreated condition between WT and both CTCF knockdowns. First, accessibility of the TSS for the CTCF$^{+/-}$ #1 compared to the WT was investigated and it was found that there were 224 genes with significantly increased accessibility (Figure 3b). Interestingly, the *TP53* target gene, BBC3 had significantly greater accessibility at the TSS in CTCF$^{+/-}$ #1 cells compared to WT at the basal level. Accessibility of the TSS for the CTCF$^{+/-}$ #2 compared to the WT found that there were 597 genes with increased accessibility (Figure 3c). Next, the accessibility of the TSS for both CTCF$^{+/-}$ #1 and #2 compared to the WT showed an increase in accessibility in 130 genes (Figure 3d).

Next, we assessed changes in accessibility of the TSS for cells exposed to 6μM Cisplatin for 8h using MCF10A WT and CTCF$^{+/-}$ (clones #1 & #2) cells. Here, we compare the accessibility following treatment between WT and CTCF knockdowns. First, accessibility of the WT treated as compared to WT untreated cells found that there was an increase in ATAC-seq reads at 104 genes (Figure 3e). The CTCF$^{+/-}$ #1 treated as compared to #1 untreated found an increase in 39 genes (Figure 3f). CTCF$^{+/-}$ #2 treated as compared to #2 untreated found and increase in accessibility of 34 genes (Figure 3g). CTCF$^{+/-}$ #1 treated as compared to WT treated saw 204 genes with increased accessibility (Figure 3h). Finally, the accessibility of CTCF$^{+/-}$ #2

treated as compared to WT treated was assessed and found that 124 genes had increased accessibility on the knockdown compared to the control (Figure 3i). Overall, the number of genes with differences in accessibility profile between CTCF and WT status appears to be greater than the differences for treatment status.

Accessibility of the TSS of *TP53* target genes after exposure to cisplatin

Using the same methods as the previous section (Figure 3), we compared the accessibility profile of the TSS globally between WT and $CTCF^{+/-}$ following treatment (6μM Cispatin, 8h) after filtering to include only *TP53* target genes. Here, we evaluate and compare TSS accessibility for *TP53* target genes in MCF10A WT and $CTCF^{+/-}$ cells between all treatment conditions. Before comparting the Log Fold Change (LogFC) between groups, we sought to visualize the global TSS accessibility of all putative *TP53* associated genes (298). Each row visually represents the signal of a gene, and the columns represent the following: MCF10A wild-type untreated, WT treated, $CTCF^{+/-}$ #1 untreated, $CTCF^{+/-}$ #1 treated, $CTCF^{+/-}$ #2 untreated, $CTCF^{+/-}$ #2 treated (Figure 4a). The signal was determined by using the standard error of the mean (SEM) for triplicate data indicated by the scaled read counts for the TSS of a gene; intensity of colour in the figure is related to strength of the signal, with darker red indicating a stronger signal and thus a greater degree of accessibility. To assess how many genes were open at the TSS, we decided to label a signal that is greater than the average signal for all genes to be "open". We determined the average SEM of the signal of all genes basal condition to determine the average accessibility of the TSS, and values greater than the average were open for the following: WT untreated (111), WT treated (164), $CTCF^{+/-}$ #1 untreated (112), $CTCF^{+/-}$ #1 treated (135), $CTCF^{+/-}$ #2 untreated (102), $CTCF^{+/-}$ #2 treated (108) (Figure 4a). In general, there

appears to be a trend where more *TP53* target genes become more accessible at the TSS following the induction of damage.

In addition, we utilized volcano plots to characterize differentially accessible TSS for untreated MCF10A WT and CTCF$^{+/-}$ (#1 and #2) cells. Here, we compare the accessibility for the untreated condition between WT and both CTCF knockdowns for *TP53* target genes. First, basal accessibility of the TSS for the CTCF$^{+/-}$ #1 compared to the WT was investigated and it was found that there were 10 genes with significantly increased accessibility in the knockdown (Figure 4b). Interestingly, the *TP53* target gene, BBC3 had significantly greater accessibility at the TSS in CTCF$^{+/-}$ #1 cells compared to WT at the basal level. Accessibility of the TSS for the CTCF$^{+/-}$ #2 compared to the WT found that there were 6 genes with increased accessibility (Figure 4c).

Next, we assessed changes in accessibility of the TSS for the treated (6μM Cisplatin, 8h) condition for MCF10A WT and CTCF$^{+/-}$ (#1 & #2). Here, we compare the accessibility following treatment between WT and CTCF knockdowns for *TP53* target genes. First, accessibility of the WT treated as compared to WT untreated cells found that there was a significant increase in only 2 genes after treatment (Figure 4d). The CTCF$^{+/-}$ #1 treated as compared to #1 untreated did not find genes with increased accessibility following treatment (Figure 4e). Similarly, CTCF$^{+/-}$ #2 treated as compared to #2 untreated did not find a significant increase in the proportion of accessible genes following treatment (Figure 4f). CTCF$^{+/-}$ #1 treated as compared to WT treated saw 5 genes with significantly increased accessibility in the knockdown (Figure 4g). Finally, the accessibility of CTCF$^{+/-}$ #2 treated as compared to WT treated was assessed and found that 2 genes had increased accessibility in the knockdown compared to the control (Figure 4h). Again, the differences in number of genes with increased

accessibility appear to be slightly greater when comparing $CTCF^{+/-}$ and WT status. Notably, notably the sample size in this instance is very small.

Accessibility of the TSS is associated with the upregulation of gene expression

Again, we cross-referenced our ATAC-seq results with previously completed RNA sequencing using the same samples including MCF10A WT, and $CTCF^{+/-}$ cells with, or without, cisplatin exposure. By integrating these data, we could investigate whether differentially expressed genes show significant changes to chromatin accessibility surrounding the TSS (± 3kb) (Ackermann et al. 2016) after CTCF loss or upon cisplatin exposure. We compared Basal accessibility of the TSS (± 3kb) and its association with differentially expressed genes between the MCF10A WT and $CTCF^{+/-}$ (#1 and #2) untreated cells. It was found that for all genes (11923) with a base mean greater than 100, changes in TSS accessibility positively correlated significantly, but mildly with changes in gene expression. Differential gene expression before the induction of DNA damage positively corelated with TSS accessibility between the WT and $CTCF^{+/-}$ #2 ($r=0.1877$, $p<0.0001$) (Figure 5a). Similarly, differential gene expression before the induction of DNA damage positively corelated with TSS accessibility between the WT and $CTCF^{+/-}$ #1 ($r=0.2580$, $p<0.0001$) (Figure 5b). Comparing differential gene expression before DNA damage to TSS accessibility of the WT compared to both $CTCF^{+/-}$ #1 and #2 also revealed a positive correlation ($r= 6518$, $p<0.0001$) (Figure 5c). Overall, the gene accessibility and expression was demonstrated to be significantly associated with each other. Comparing the WT condition to both knockdown conditions yielded the greatest positive correlation among the groups. This data indicates that increased accessibility of the TSS is associated with increased gene expression.

Changes in CTCF binding at target genes does not strongly influence the association between accessibility and gene expression

Next, we investigated the presence of CTCF around the promoters of genes to see whether a loss at a promoter impacts the association between basal gene expression and TSS accessibility for MCF10A WT and $CTCF^{+/-}$ (#1 and #2) untreated cells. At a subset of genes, we assess the impact of CTCF presence. For genes where there was a constant presence of CTCF at the promoter of a gene (3947 total), a significant but reduced association between accessibility and gene expression for both knockdowns. For basal TSS accessibility of $CTCF^{+/-}$ compared to WT, for #1 426 genes were significantly more open in the knockdown and for #2 693 genes were significantly more open at the TSS in the knockdown (LogFC>1) (Figure 6a). For genes where there was a loss of CTCF at the promoter of a gene (1097), a significant but reduced association between accessibility and gene expression for both knockdowns. When comparing accessibility of the TSS of the $CTCF^{+/-}$ compared to WT, for #1 90 genes were significantly more open at the TSS in the knockdown and for #2 150 genes were significantly more open in the knockdown (Figure 6b). For genes where there was a gain of CTCF at the promoter of a gene (41), a significant association between accessibility and gene expression for $CTCF^{+/-}$ #1 and a reduced association in $CTCF^{+/-}$ #2 was found. When comparing accessibility between the knockdown and WT, #1 had 17 genes that were more open in the knockdown and #2 did not show any genes that were more open in the knockdown compared to WT (Figure 6c). Overall, the presence of CTCF at the promoter of a gene, whether it is lost or constant, slightly reduces the association between accessibility at the TSS and differential gene expression. The only exception is $CTCF^{+/-}$ #1 in Figure 6c, however it's important to note that the sample size in that figure was still quite small (41 genes).

In addition, our analysis revealed that the presence of CTCF proximal to a gene does not appear to be predictive of the association between gene expression and TSS accessibility when the binding of CTCF is constant (Figure 7a), or even when CTCF is lost around the gene in $CTCF^{+/-}$ cells (Figure 7b). In summary, the presence of CTCF on a promoter or around a gene does not appear to be indicative of differential TSS accessibility and gene expression of $CTCF^{+/-}$ cells compared to the WT at the basal level.

The critical *TP53* target gene, BBC3, has differential accessibility in $CTCF^{+/-}$ cells compared to the control

We next decided to review the sequencing tracks of BBC3 and three others key *TP53* associated genes to visualize the accessibility profile of these genes. First, we investigated datasets obtained from ATAC-seq, comparing cisplatin exposed cells to control MCF10A WT and $CTCF^{+/-}$ cells. ATAC-seq data may be visualized using the web-based viewer, IGV. Accessibility is visualized as signals, which represent scaled read counts of regions of chromatin to indicate regions of open chromatin. Sequencing was carried out in triplicate and ATAC-seq peaks were visualized genome-wide using IGV tracks based on BAM files. Open chromatin is confirmed based on the signal in a particular region. We viewed 4 critical genes involved in the *TP53* transcriptional response; DNA repair (GADD45a) (Zhan et al. 1999), cell cycle arrest/senescence (CDKN1A) (Brugarolas et al. 1999), core regulation of *TP53* (MDM2) (Kubbutat et al. 1997), and apoptosis (BBC3) (Speidel, 2010). This was done to query whether changes in accessibility might differ depending on the pathway. We visualized ATAC-seq genome tracks indicating chromatin accessibility before and following the induction of DNA damage by 8h 6μM Cisplatin treatment to MCF10A WT cells and $CTCF^{+/-}$ cells (clone #2). Tracks show peak intensity signal along the y-axis and genomic location along the x-axis. For

each gene, from top to bottom, the order of the tracks are as follows: WT cells without cisplatin treatment (dark blue), WT cells after 8 hours of cisplatin treatment (dark orange), $CTCF^{+/-}$ #2 cells without cisplatin treatment (light blue), and $CTCF^{+/-}$ #2 cells after 8 hours of cisplatin treatment (light orange) (Figure 8). The GADD45A gene indicates accessibility of WT and $CTCF^{+/-}$ #2 before and following the induction of damage. From observation, it appears to have greater accessibility following damage, but marginal differences are noted between the knockdown and WT (Figure 8a). The CDKN1A/p21 gene indicates accessibility of WT and $CTCF^{+/-}$ #2 before and following the induction of damage. Similarly, while damage seems to increase accessibility, little difference is noted for differing CTCF status (Figure 8b). The MDM2 gene indicates accessibility of WT and $CTCF^{+/-}$ #2 before and following the induction of damage. Again, damage increases accessibility but there is little noticeable difference between the WT and knockdown (Figure 8c). The BBC3/PUMA gene indicates accessibility of WT and $CTCF^{+/-}$ #2 before and following the induction of damage. There appears to be an increase in gene accessibility following damage. Additionally, there appears to be greater accessibility in the knockdown compared to the WT; the basal accessibility also appears to be greater in the knockdown (Figure 8d).

As BBC3 appeared quite often, we decided to investigate further. An additional review of accessibility (scaled read counts) for BBC3 at the TSS (Figure 9a) and TTS (Figure 9b) for WT and $CTCF^{+/-}$ cells before and following damage agreed with the finding that it is more accessible in the knockdown compared to WT following damage (Figure 9). Of these *TP53* targets, BBC3 was found to be the only target that is basally more accessible at the TSS in $CTCF^{+/-}$ cells before the induction of damage.

10. Figures

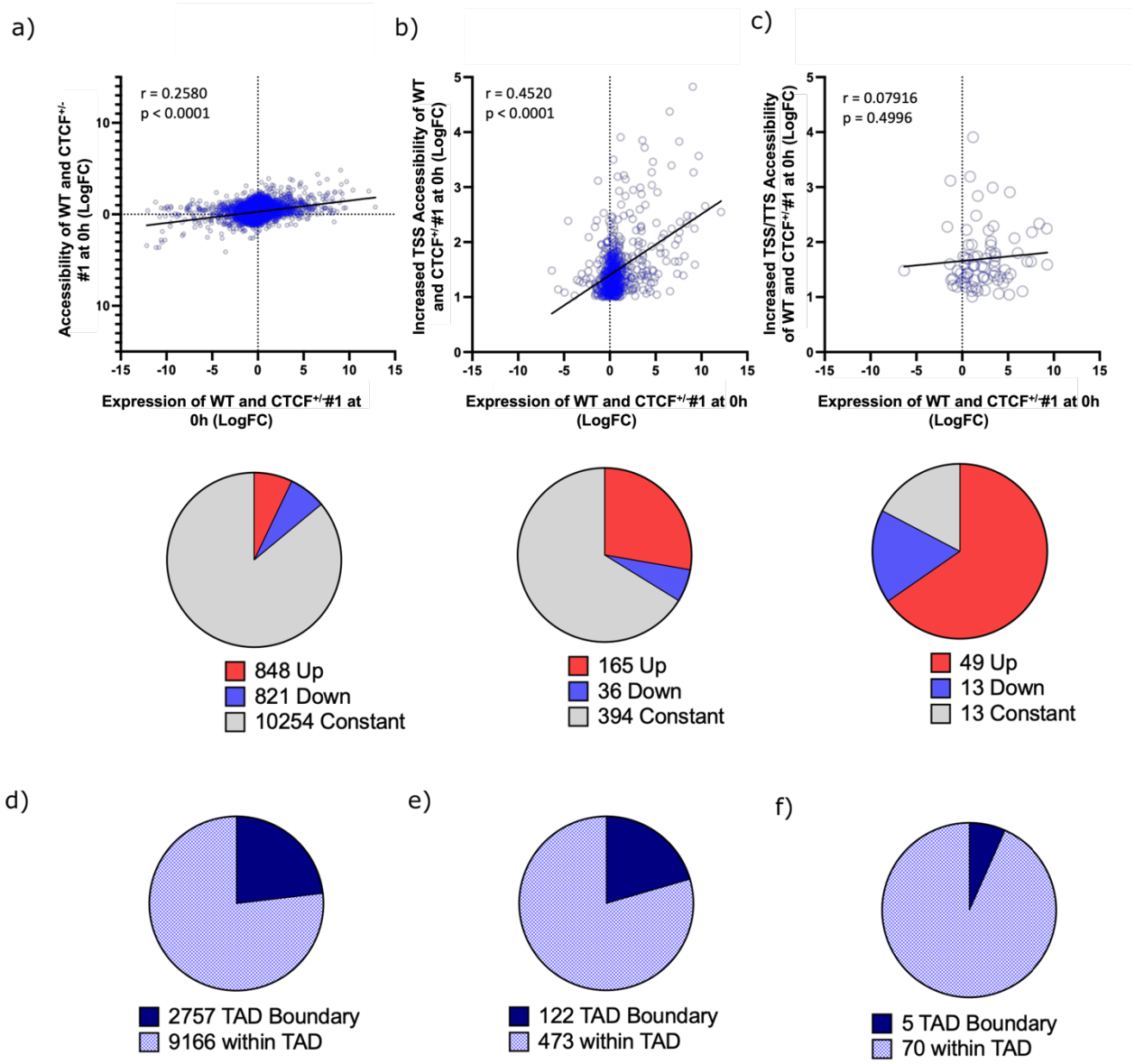


Figure 1. Basal accessibility of the TSS (± 3kb) and TTS (± 3kb) and its association with differentially expressed genes between the MCF10A WT and CTCF$^{+/-}$ (#1) untreated cells. Scatterplot (Top) of #1 cells were compared to the WT group on genes where the genome was differentially accessible and pie charts (bottom) indicating proportions of differentially expressed genes on a) all genes (848 upregulated, 821 downregulated, and 10252 constant), b) genes where the TSS had increased accessibility (165 upregulated, 36 downregulated, and 394 constant), and c) both the TSS and TTS had increased accessibility (49 upregulated, 13 downregulated, and 13 constant). Pie charts of differentially expressed genes within Topologically Associating domains

(TAD) or on TAD boundaries for d) all genes (2757 on TAD boundary, 9166 within TADs), e) genes with increased TSS accessibility (122 on TAD boundaries, 473 within TADs), or f) genes with increased accessibility on both TSS and TTS (5 on TAD boundaries, 70 within TADs). Up/Down indicates absolute LogFC (fold change) ≥ 1 and pv/padj ≤0.05.

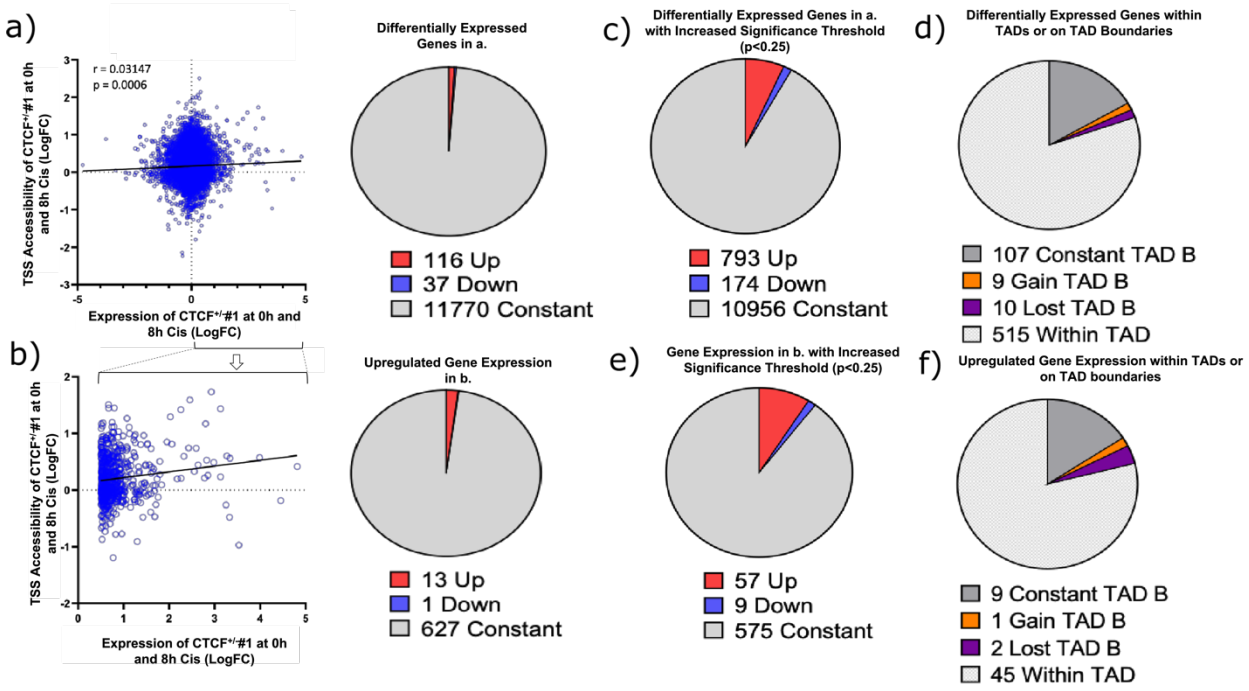


Figure 2. Accessibility of the TSS (± 3kb) and its association with differentially expressed genes of the MCF10A CTCF$^{+/-}$ (#1) treated (6μM Cisplatin, 8h) compared to non-treated cells. Scatterplot (left) differential expression of #1 following treatment were compared to the differential accessibility following treatment and pie charts (right) indicating proportions of differentially expressed genes in treated cells compared to untreated for a) all genes with increased TSS accessibility and b) filtered for genes that are differentially expressed. (LogFC > 0.5, pv ≤0.05). Pie charts (right) indicating proportions of differentially expressed genes in treated cells compared to untreated with modified thresholds of pv≤ 0.25 to investigate a greater proportion of genes with differentially accessible TSS following damage for c) all genes (793 upregulated, 174 downregulated, and 10956 constant), d) genes with TAD enrichment (107 constant on TAD boundaries, 9 gained on TAD boundaries, 10 lost on TAD boundaries, and 515 within TADs), e) genes with increased accessibility (57 upregulated, 9 downregulated, and 575 constant), and f) genes with TAD enrichment that have increased TSS accessibility (9 constant

on TAD boundaries, 1 gained on TAD boundaries, 2 lost on TAD boundaries, and 45 within TADs).

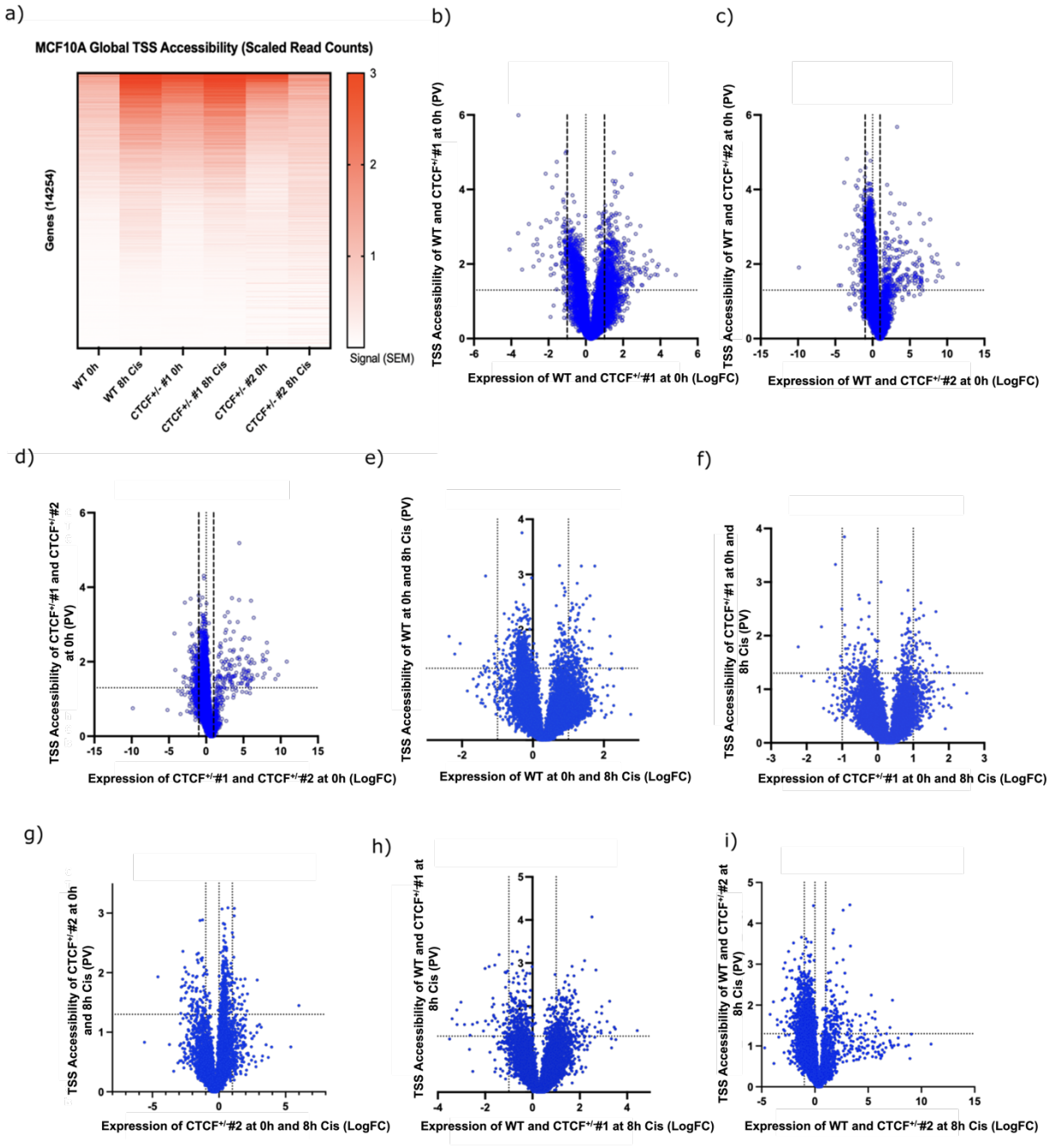


Figure 3. Global Accessibility of the TSS for MCF10A cells. a) Heatmap of global accessibility of the TSS for MCF10A cells; SEM of triplicates was utilized to determine the signal (Scaled read counts). Volcano plot characterizing differentially accessible TSS for untreated MCF10A WT and $CTCF^{+/-}$ cells (knockdown #1 & #2). Dotted lines indicate thresholds of significance (x-axis = ±1 LogFC, y-axis =1.3 Logpv). At the TSS in the untreated condition, b) #1 as compared to WT cells, c) #2 as compared to WT cells and d) #2 as compared to #1 was assessed. The accessibility of the treated (6μM Cisplatin, 8h) condition at the TSS, e) WT treated as compared to WT untreated, f) #1 treated as compared to #1 untreated, g) #2 treated as compared to #2 untreated, h) #1 treated as compared to WT treated, and i) #2 treated as compared to WT treated was assessed.

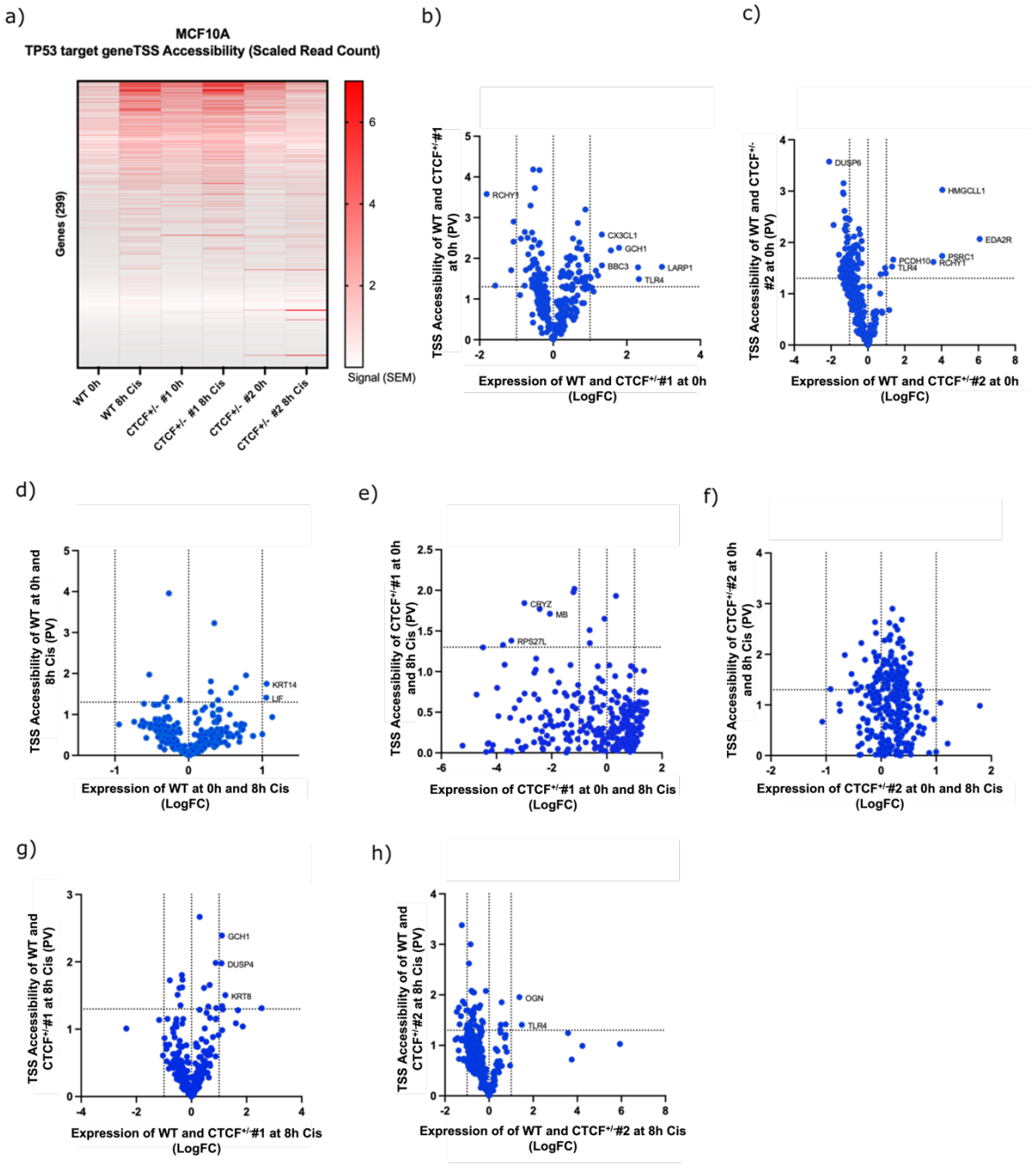


Figure 4. Global Accessibility of the TSS for *TP53* target genes for MCF10A cells. a) Heatmap of global accessibility of the TSS for MCF10A cells, filtered to include 299 *TP53* target genes; SEM of triplicates was utilized to determine the signal (Scaled read counts). Volcano plot characterizing differentially accessible TSS for untreated MCF10A WT and CTCF$^{+/-}$ cells (knockdown #1 & #2). Dotted lines indicate thresholds of significance (x-axis = ±1

LogFC, y-axis =1.3 Logpv). At the TSS in the untreated condition, b) #1 as compared to WT cells, c) #2 as compared to WT cells was assessed. The accessibility of the treated (6μM Cisplatin, 8h) condition at the TSS for, e) WT treated as compared to WT untreated, f) #1 treated as compared to #1 untreated, g) #2 treated as compared to #2 untreated, h) #1 treated as compared to WT treated, and i) #2 treated as compared to WT treated was assessed.

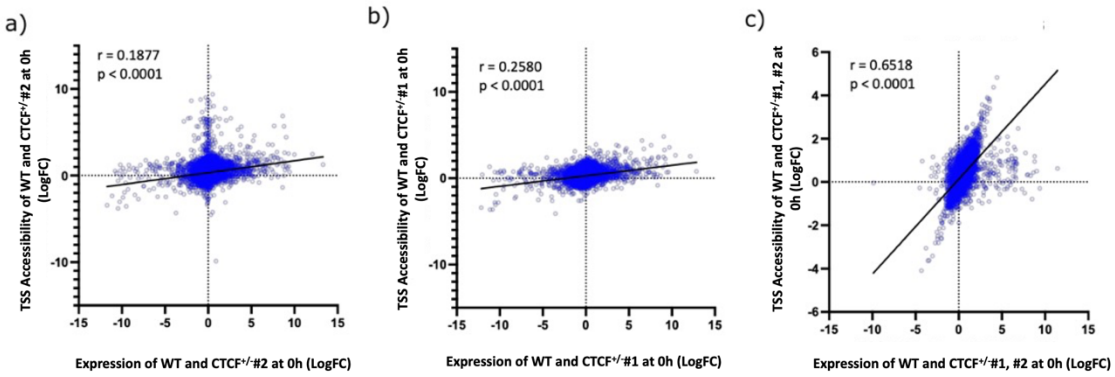


Figure 5. Basal accessibility of the TSS (± 3kb) and its association with differentially expressed genes between the MCF10A WT and CTCF$^{+/-}$ (#1 and #2) untreated cells. 11923 upregulated genes were selected from RNA-seq data by filtering for genes with a base mean greater than 100. Scatterplot (Log2 Fold-Change scale) of RNA-seq expression for triplicates of ATAC-seq WT data and a) CTCF$^{+/-}$ #2, b) CTCF$^{+/-}$ #1, c) CTCF$^{+/-}$ #1 and #2.

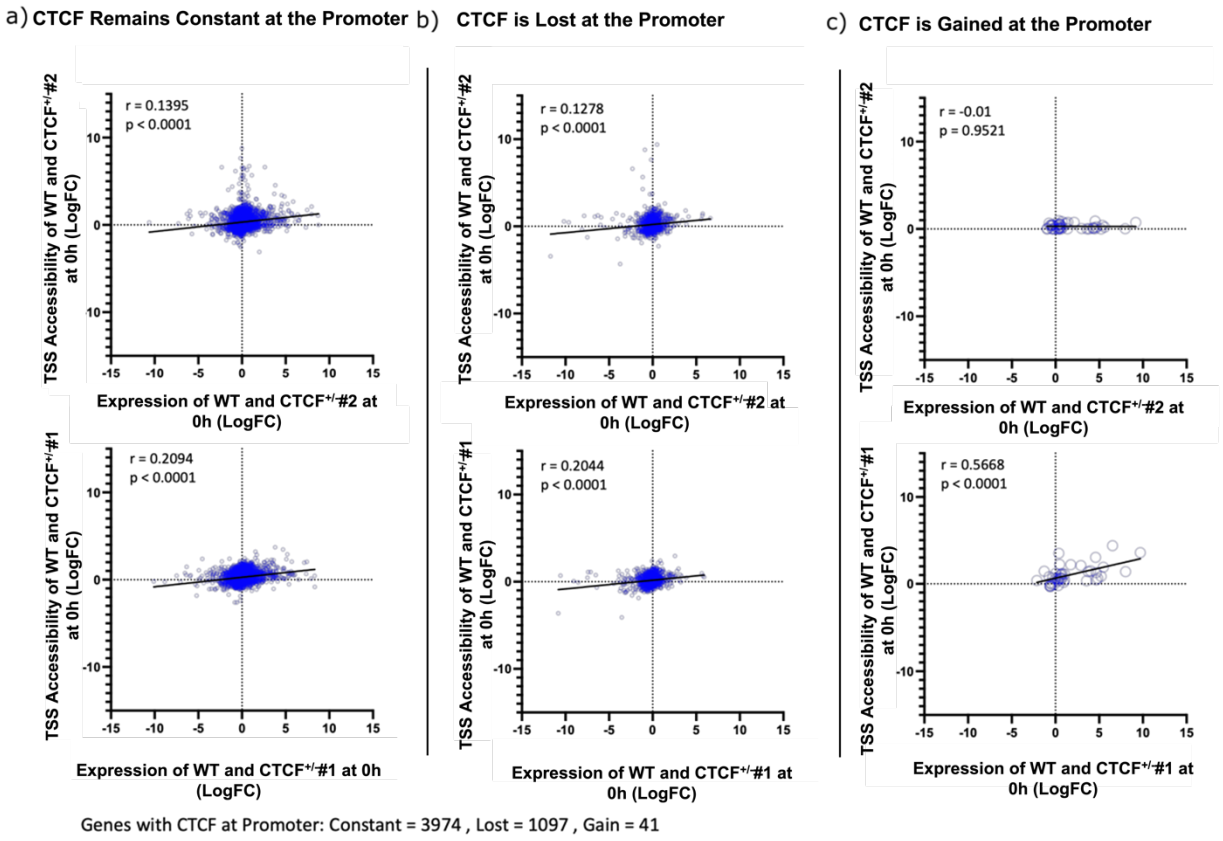


Figure 6. Basal accessibility of the TSS (± 3kb) and its association with differential CTCF presence at the promoter between the MCF10A WT and CTCF$^{+/-}$ (#1 and #2) untreated cells. #2 (top) and #1 (bottom) cells were compared to the WT group on genes where CTCF presence at the promoter of the gene was a) constant (3947 genes) b) lost at the promoter (1097 genes) or c) gained at the promoter (41 genes).

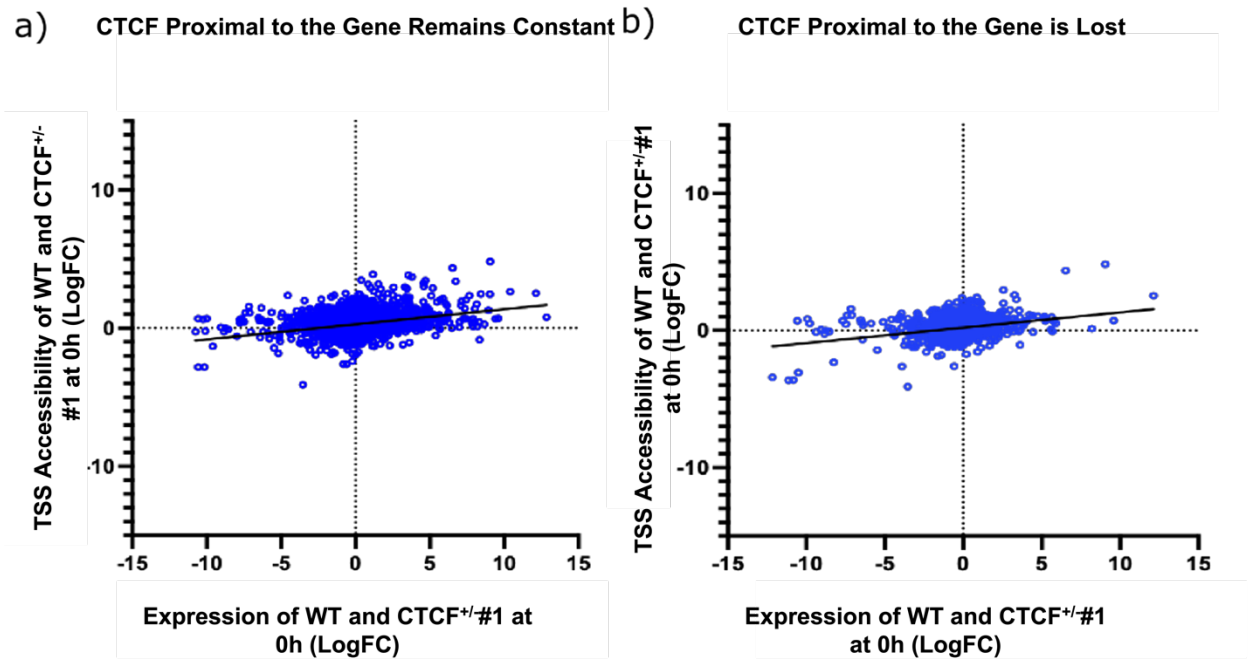
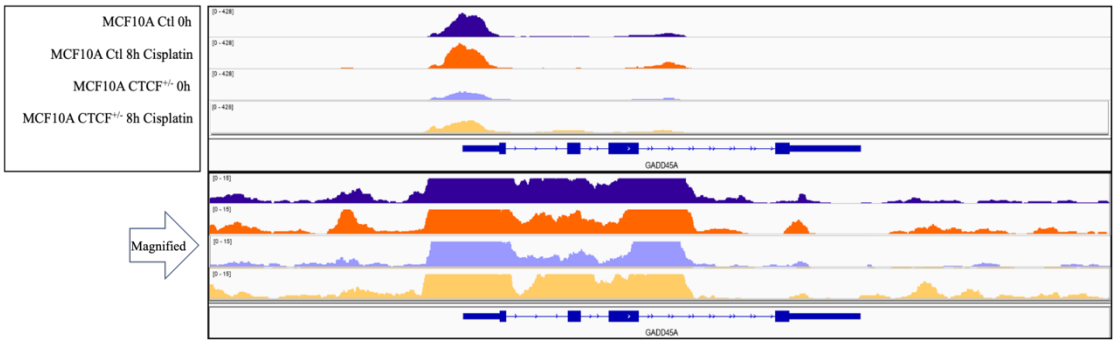


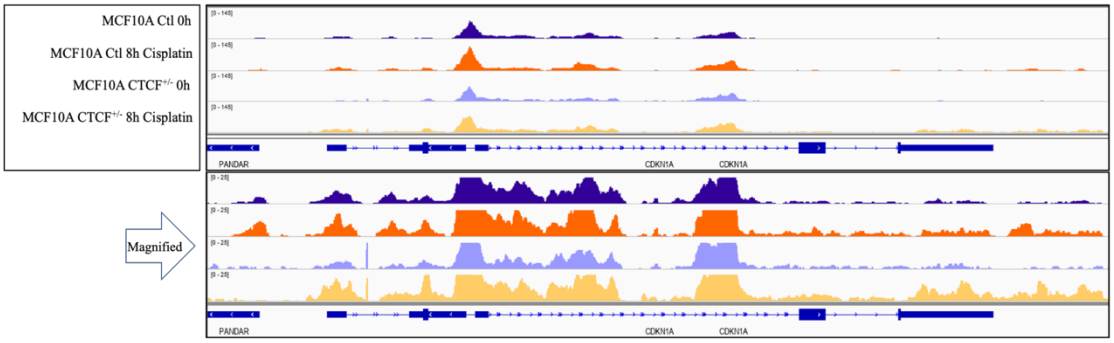
Figure 7. Basal accessibility of the TSS (± 3kb) and its association with differential CTCF presence around genes between the MCF10A WT and CTCF$^{+/-}$ (#1) untreated cells.

Scatterplot of #1 cells compared to the WT group on genes where CTCF presence around a gene was a) constant (8303 genes) ($r = 0.2321$, $p < 0.0001$) or b) lost around the gene (2225 genes) ($r = 0.2584$, $p < 0.0001$).

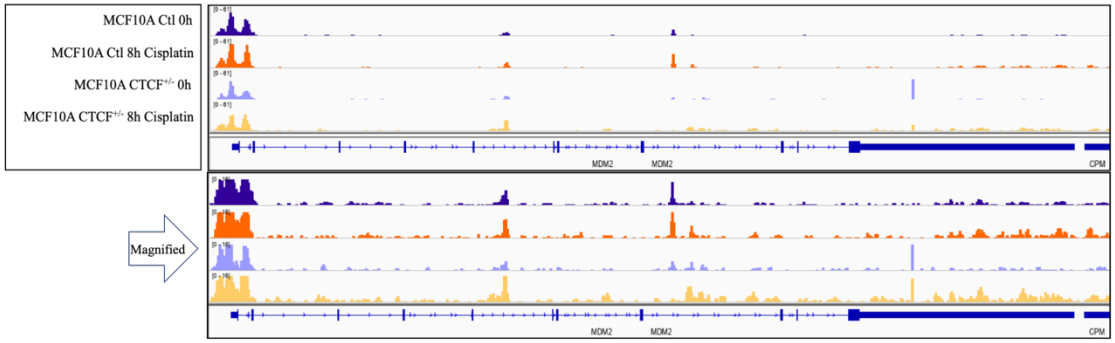
a)


MCF10A Ctl 0h
MCF10A Ctl 8h Cisplatin
MCF10A CTCF+/- 0h
MCF10A CTCF+/- 8h Cisplatin
GADD45A
Magnified
GADD45A

b)


MCF10A Ctl 0h
MCF10A Ctl 8h Cisplatin
MCF10A CTCF+/- 0h
MCF10A CTCF+/- 8h Cisplatin
PANDAR
CDKN1A
CDKN1A
Magnified
PANDAR
CDKN1A
CDKN1A

c)


MCF10A Ctl 0h
MCF10A Ctl 8h Cisplatin
MCF10A CTCF+/- 0h
MCF10A CTCF+/- 8h Cisplatin
MDM2
MDM2
CPM
Magnified
MDM2
MDM2
CPM

d)

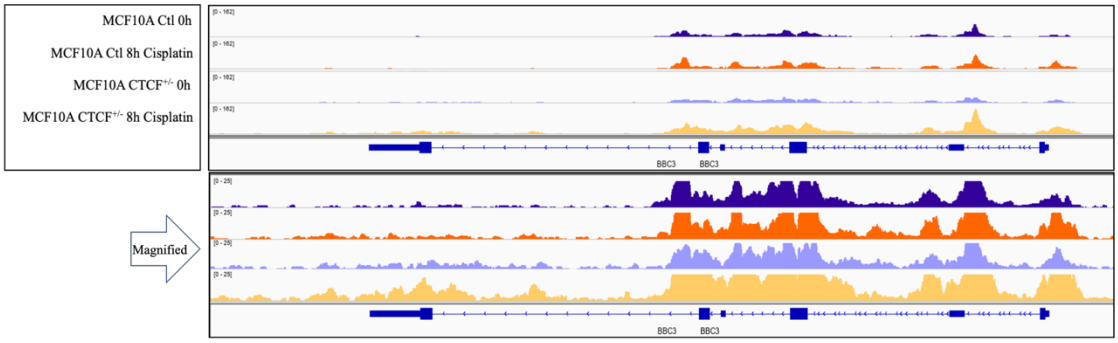

MCF10A Ctl 0h
MCF10A Ctl 8h Cisplatin
MCF10A CTCF+/- 0h
MCF10A CTCF+/- 8h Cisplatin
BBC3
BBC3
Magnified
BBC3
BBC3

Figure 8. ATAC-seq genome tracks indicating chromatin accessibility before and following the induction of DNA damage by 6µM Cisplatin treatment to MCF10A WT cells and CTCF$^{+/-}$ cells (#2). Tracks show peak intensity signal (y-axis). For each gene, from top to bottom, the order of the tracks are as follows: WT cells without cisplatin treatment, WT cells after 8 hours of cisplatin treatment, #2 cells without cisplatin treatment, and #2 cells after 8 hours of cisplatin treatment. a) The GADD45A gene is indicated (1q31.3; 67,685,201 bp – 67,688,334 bp). b) The CDKN1A/p21 gene is indicated (6q21.2; 36,676,460 bp – 36,687,339 bp). c) The MDM2 Gene is indicated (12q15; 68,808,177 bp – 68,845,544 bp). d) The BBC3/PUMA gene is indicated (19q13.32; 47,220,822 bp – 47,232,766 bp).

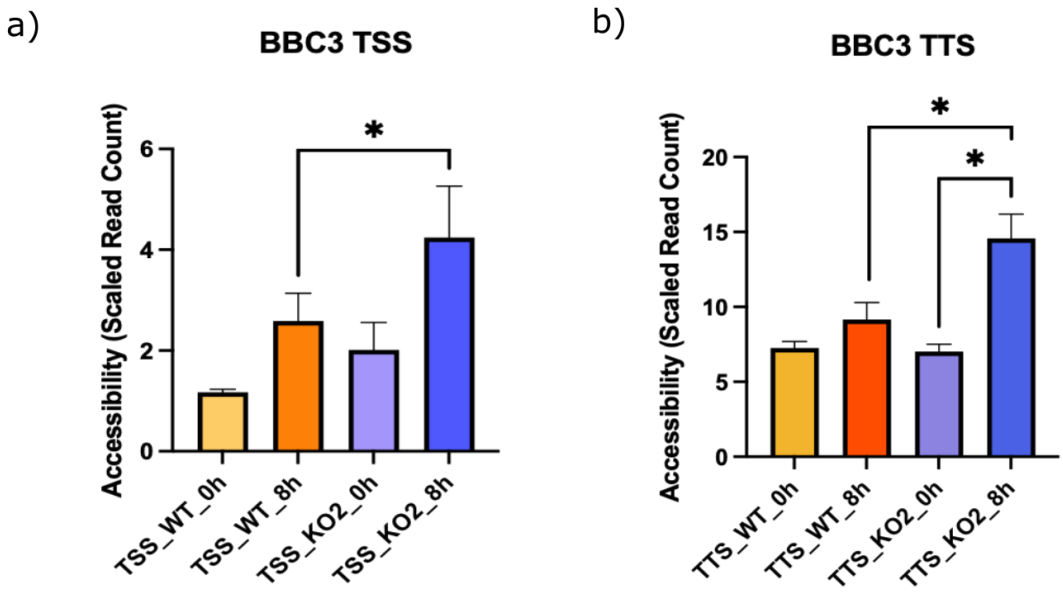


Figure 9. Changes in Accessibility of the TSS and TTS following 8h 6µM Cisplatin treatment on the BBC3 gene. The y-axis indicates the signal (scaled read counts) to describe accessibility. a) Accessibility of the TSS of the BBC3 gene for the following conditions in order: WT 0h, WT 8h Cisplatin, CTCF$^{+/-}$ 0h, and CTCF$^{+/-}$ 8h Cisplatin. b) Accessibility of the TTS of the BBC3 gene for the following conditions in order: WT 0h, WT 8h Cisplatin, CTCF$^{+/-}$ 0h, and CTCF$^{+/-}$ 8h Cisplatin.

11. Discussion

CTCF CNL results in 5'-3' accessibility in some *TP53* genes following DSB

In this thesis, we show that globally, the significant changes in accessibility of the TSS can be influenced by loss of CTCF before and after damage at a subset of genes. Among *TP53* target genes, CTCF loss impacts changes in accessibility to a small number of key targets. Additionally, we found that in cells with a single copy loss of CTCF, select *TP53* target genes have increased accessibility on both the TSS and TTS after the induction of DNA DSB when compared to the control. Of the 5 genes (BBC3, CDKN1A, GDF15, FDXR, and NECTIN4) with increased accessibility of the TSS and TTS in the $CTCF^{+/-}$ following damage, the genes BBC3, CDKN1A, FDXR and GDF15 are classical *TP53* target genes (Hwang et al. 2001; Speidel, 2010; Brugarolas et al. 1999; Khaled et al. 2012). We also demonstrated that accessibility of the TSS is positively associated with differentially expressed genes between the control and $CTCF^{+/-}$ globally. Finally, we demonstrated that genes with greater accessibility on the TSS and TTS are significantly enriched within topologically associating domains. Additionally, BBC3 is the only *TP53* target gene that has greater accessibility in the $CTCF^{+/-}$ cells compared to WT before the induction of damage. BBC3 also becomes more accessible at the TSS in $CTCF^{+/-}$ following the induction of DNA DSB. Thus, the 5'-3' transposase accessibility in $CTCF^{+/-}$ cells following damage appears to be limited to a select few *TP53* target genes. We propose that the single copy loss of CTCF may result in loss of regulation in chromatin loops, and upon the induction of DNA DSB in deficient cells, upregulated *TP53* target genes see their TSS and TTS become more accessible. We hypothesise that this increase in 5'-3' accessibility may represent a mechanism by which specific *TP53* target genes may have increased transcriptional efficacy following damage by 'circularizing' where the TSS and TTS become more accessible. We hypothesize that the loss

of CTCF may mediate this process by destabilizing chromatin loops and bringing the accessible TSS and TTS into proximity (Figure 10). We note that this hypothesized process would be limited to specific *TP53* target genes, especially, BBC3 and thus may be associated with apoptosis. Indeed, CTCF loss has been previously demonstrated to alter chromatin accessibility upon depletion (Xu et al. 2021).

Cisplatin was utilized to induce DSBs

The accumulation of DNA damage is widely considered to be a major contributor to cancer. Although damage can alter gene expression, its relationship with altered accessibility transcription has been poorly described in the literature. Regions of "open" chromatin are typically associated with higher levels of transcription. Indeed, around 90% of bound transcription factors are found in regions of accessible chromatin (Thurman et al. 2012). Thus, we induced DNA DSB damage to investigate differential accessibility on the TSS and TTS as a representation of transcriptional activity. DNA DSBs can cause mutation, deletion, and translocations in the genome if they are not repaired. Our use of Cisplatin treatment led to the induction of DNA DSB. Cisplatin treatment causes DNA adducts that form intra- and interform crosslinks; while intrastrand crosslinks are repaired by nucleotide excision repair pathways, interstrand crosslinks stall replication forks and result in DNA DSB and are thus potentially lethal (Kartalou & Essigmann. 2001; Enoiu et al. 2012). Previously completed RNA-seq with doxorubicin treated MCF10A cells indicated slightly higher differential gene expression between the WT and $CTCF^{+/-}$ cells. Additionally, western blots indicated greater magnitude of protein stabilization for *TP53* and p-*TP53* at 8h in doxorubicin treatment compared to cisplatin. Doxorubicin leads to DNA DSBs by intercalation into DNA which in turn prevents DNA and

RNA synthesis (Momparler et at. 1976). Thus, an ATAC-seq performed on MCF10A cells utilizing 500nM Doxorubicin treatment may reveal slightly stronger results.

BBC3 becomes more accessible in CTCF$^{+/-}$ cells following damage

From a review of IGV tracks for several *TP53* target genes, we assessed the accessibility of the treated and untreated WT and CTCF$^{+/-}$. The GADD45A, CDKN1A, and MDM2 genes appeared to have an increase in accessibility following the induction of DNA damage but shows marginal difference between the accessibility of the WT and CTCF$^{+/-}$ #2 cells (Figure 8). The BBC3/PUMA gene indicates an increase in gene accessibility following DNA DSB. In addition, there appears to be a greater difference in accessibility following damage for CTCF$^{+/-}$ #2 cells compared to WT. Finally, there appears to be greater accessibility in the CTCF$^{+/-}$ compared to the WT at the basal level (Figure 8). Taken together, a visual review of the tracks alone suggested that BBC3 may be more accessible in CTCF$^{+/-}$ cells compared to the control after damage induction. Indeed, In the case of BBC3, we demonstrated that, similarly to the visual IGV data, our analysis revealed that this gene becomes more accessible in the treated CTCF$^{+/-}$ cells compared to the control. Not just BBC3, but 5 key upregulated genes have increased accessibility at both the TSS and TTS in CTCF$^{+/-}$ cells. Notably, this phenomena of increased TSS and TTS accessibility in CTCF$^{+/-}$ cells has not been described in the literature. While the greater accessibility following damage in CTCF$^{+/-}$ cells is indeed interesting, this is still limited to few genes; however, of the 5 genes with more accessible TSS and TTS in CTCF$^{+/-}$ cells, most are *TP53* targets. These genes were CDKN1A, GDF15, NECTIN4, FDXR, and BBC3; NECTIN4 was the only gene without an association to *TP53.* From our analysis, BBC3 and FDXR were found to have the greatest significance.

FDXR is involved in the transfer of electrons from NADPG to cytochrome P450 in the mitochondria. Additionally, Hwang et al. (2001) indicated that Trp53 induced FDXR in colorectal carcinoma cells treated with 5-fluorouracil. Additionally, disruption of the FDXR gene showed a decrease in the sensitivity of the cells to 5-FU induced apoptosis. Thus, FDXR is regulated by *TP53* and is involved in apoptotic pathways following the induction of damage.

Finally, PUMA/BBC3 is a major *TP53* target associated with the mitochondrial regulation of apoptosis. PUMA can rapidly lead to the induction of apoptosis via BAX pathways within the mitochondria (Han et al. 2001). Additionally, BBC3 was the only gene shown to be significantly more accessible at the TSS before the induction of damage and after in $CTCF^{+/-}$ cells. This indicates that the increased accessibility in $CTCF^{+/-}$ #1 cells after damage is not due to damage alone.

5'-3' accessibility following DSB in $CTCF^{+/-}$ cells may be associated with apoptotic pathways

Notably, several of the previously described genes are associated with pro-apoptotic pathways and mitochondrial pathways. BBC3 and FDXR had the most significant p-values out of all the genes with increased accessibility of the TSS and TTS after the induction of damage in $CTCF^{+/-}$ cells. Both of those genes are localized to the mitochondria. The literature has not generally described any direct association between BBC3 and FDXR, however, their relationship to apoptosis may indicate that the increase in TSS and TTS accessibility in the knockdown cells may be associated with these pathways. BBC3 is a well described *TP53* target gene and FDXR has been demonstrated to be inducted by *TP53* and may represent a response to damage in cells with deregulated chromatin structure due to CTCF loss. Thus, the importance of the increase in accessibility of these regions will continue to be investigated. Indeed, if CTCF can be

demonstrated as a biomarker for accessibility profiles of these genes, it could function to predict efficacy of treatment.

The importance of 5'-3' accessibility due to CTCF CNL is still under investigation

CTCF is crucial for the organization of hierarchical chromatin structure. The loss of CTCF has been shown to deregulate nucleosome density and its chromatin boundaries, thus increasing accessibility within chromatin loops (Khoury et al., 2020). Additionally, CTCF has been demonstrated to be important for RNA Polymerase II mediated clustering interactions of activator proteins on chromatin (Lee et al. 2022). Overall, CTCF appears to be essential for the regulation of multiple chromatin structures that promote transcriptional activation. While loss of CTCF potentiated the activation of many genes after cisplatin, mechanistically, it is clear that open and closed chromatin conformation can explain only a subset of these changes. Thus, it would be valuable to investigate other possible causes, such as changes to enhancer-promoter interactions. It would also be worthwhile to investigate the importance of 5'-3' chromatin accessibility, and to determine whether this is indeed due to circularization of the gene to enhance cycles of RNA Polymerase II-driven transcription. To accomplish this, we would need to investigate the 3D chromatin structure on a relatively discrete genome region so Hi-C analysis would likely be ineffective. An appropriate alternative may be chromatin conformation capture (3C) that allows more precise mapping of proximal DNA-DNA interactions (Dekker et al. 2002).

We could also investigate interactions between the TSS and TTS of an individual gene moving forward by utilizing Fluorescence Resonance Energy Transfer (FRET) tags. FRET provides insight into intra- and intermolecular distances on a nanometer scale (Algar et al. 2019). FRET is traditionally used for identifying protein interactions by attaching a donor and acceptor fluorophores to the proteins by fusion or by tagged antibodies. Although, it was originally

utilized for proteins, DNA-based applications have also emerged by utilizing FRET-based hybridization probes (Didenko. 2001). By tagging the TSS and TTS of these 5 genes with increased accessibility, we could validate whether the distance between the two changes following the induction of DNA damage in the control and CTCF$^{+/-}$ cells. Alternatively, a CLOuD9 approach could be taken. CLOuD9 is a CISPR based approach that allows enforced interactions between two distinct loci (Seow et al. 2022). Thus, this technique could be used to force circularization of single genes to examine the effect on transcriptional outputs.

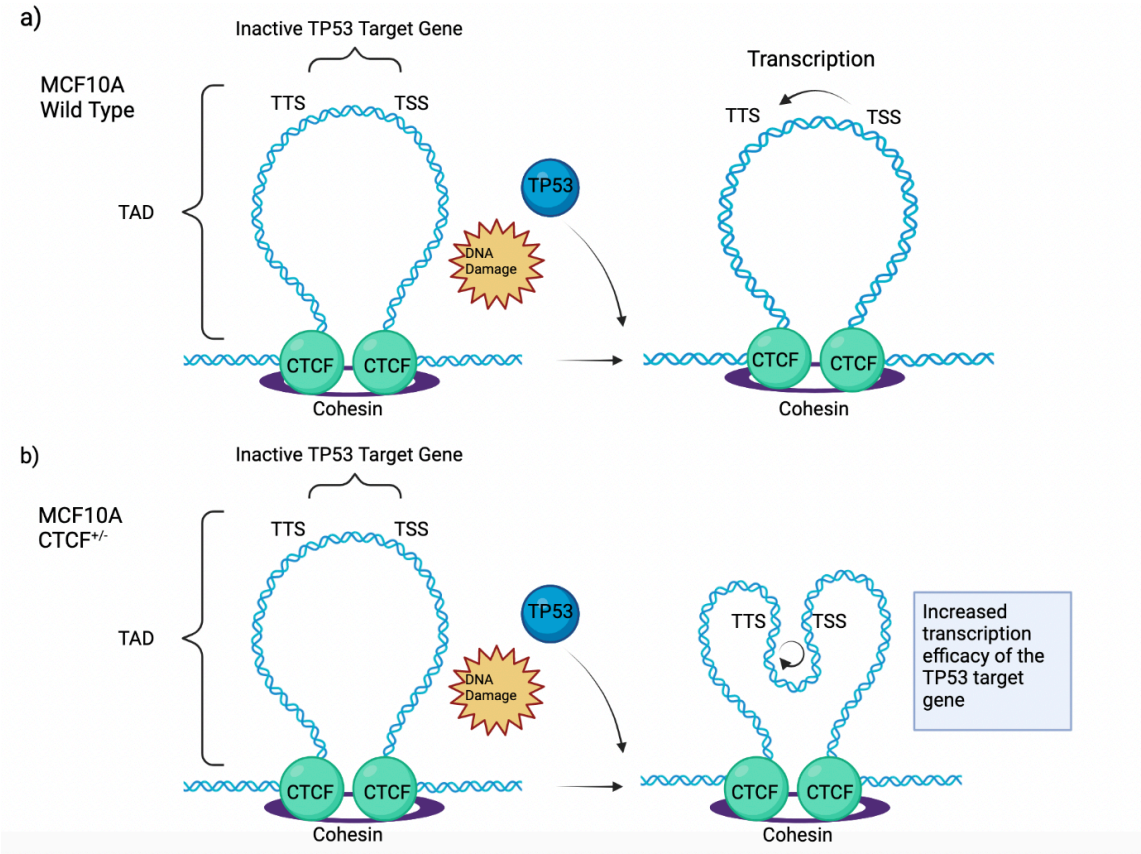


Figure 10. Hypothesized mechanism by which genes with increased accessibility of both the TSS and TTS may increase the efficacy of transcription in the case of CTCF haploinsufficient MCF10A cells. Within a topologically associating domain (TAD), following the induction of DNA damage and activation of *TP53* in a) WT cells, and b) CTCF$^{+/-}$ cells where

select inactive *TP53* target genes (such as BBC3) have greater accessibility at the TSS and TTS. The TSS and TTS are brought into proximity, possibly mediated by CTCF loss, where the transcription of the target gene becomes more efficient. Designed with BioRender.

12. Conclusion

In this project we demonstrated that globally, the changes in accessibility of the TSS can be influenced by loss of CTCF before and after damage in some genes. For *TP53* target genes globally, CTCF loss can also affect the changes in accessibility to a smaller extent. We also showed that accessibility of the transcription start site is associated with heightened gene expression in $CTCF^{+/-}$ compared to the control. Interestingly, for a subset of *TP53* response genes, there is increased accessibility on both transcriptional start sites and termination sites following the induction of DNA damage. In BBC3, we find that accessibility is also greater before the induction of damage in $CTCF^{+/-}$ cells, indicating that the difference in accessibility is not due to damage alone. Thus, the importance of chromatin accessibility at these two regions is still under investigation. We propose that the increased accessibility of *TP53* target genes in $CTCF^{+/-}$ cell following damage represents a mechanism enhancing the efficacy of the *TP53*-regulated DNA damage response. Indeed, if CTCF can be demonstrated as a biomarker for accessibility profiles of these *TP53* target genes, it could function to predict efficacy of treatment for breast cancer patients.

13. References

Arias-Lopez, C., Lazaro-Trueba, I., Kerr, P., Lord, C., Dexter, T., Iravani, M., Ashworth, A., & Silva, A. (2006). p53 modulates homologous recombination by transcriptional regulation of the RAD51 gene. *EMBO Rep. 7,* 219–224.

Alberts, B., Johnson, A., Lewis, J., et al. (2002). Molecular Biology of the Cell. 4th edition. New York: Garland Science. The Molecular Basis of Cancer-Cell Behavior. Available from: https://www.ncbi.nlm.nih.gov/books/NBK26902/

Algar, W., Hildebrandt, N., Vogel, S. *et al*. (2019). FRET as a biomolecular research tool — understanding its potential while avoiding pitfalls. *Nat Methods. 16*, 815–829.

Allis, C. & Jenuwein, T. (2016). The molecular hallmarks of epigenetic control. *Nat. Rev. Genet*. *17*, 487–500.

Altieri, F., Grillo, C., Maceroni, M., & Chichiarelli, S. (2008). DNA damage and repair: from molecular mechanisms to health implications. *Antioxid. Redox Signal*. *10*, 891–937.

Akhtar, S., et al. (2020). Association of Mutation and Low ow Expression of the CTCF Gene with Breast Cancer Progression. *SPJ. 28,* 607-614.

Baker, S., Fearon, E., Nigro, J., Hamilton, S., Preisinger, A., Jessup, J., vanTuinen, P., Ledbetter, D., Barker, D., Nakamura, Y.*, et al.* (1989). Chromosome 17 deletions and p53 gene mutations in colorectal carcinomas. *Science. 244*, 217-221.

Baumann, P., Benson, F., & West, S. (1996). Human Rad51 protein promotes ATP-dependent homologous pairing and strand transfer reactions in vitro, *Cell. 87*, 757–766.

Barlev, N., Liu, L., Chehab, N., Mansfield, K., Harris, K., Halazonetis, T., & Berger, S. (2001). Acetylation of p53 activates transcription through recruitment of coactivators/histone acetyltransferases. *Mol Cell*. *8,* 1243–1254.

Barski, A., Cuddapah, S., Cui, K., Roh, T., Schones, D., Wang, Z., Wei, G., Chepelev, I., & Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. *Cell. 129*, 823–837.

Bell, A., West, A., & Felsenfeld, G. (1999). The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell. 98*, 387–396.

Buenrostro, J., Giresi, P., Zaba, L., Chang, H., & Greenleaf, W. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods. 10(12),* 1213-1218.

Buenrostro, J., Wu, B., Chang, H., & Greenleaf, W. (2015). ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr Protoc Mol Biol*. *109,* 21.29.1-21.29.9.

Brugarolas, J., Moberg, K., Boyd, S., Taya, Y., Jacks, T., & Lees, J. (1999). Inhibition of cyclin-dependent kinase 2 by p21 is necessary for retinoblastoma protein-mediated G1 arrest after gamma-irradiation. *Proc. Natl. Acad. Sci. U.S.A. 96*, 1002-1007.

Chandrasekaran, A., Idelchik, P., & Melendez, J. (2017). Redox control of senescence and age-related disease. *Redox Biol. 11*, 91–102.

Chen, J. (2016). The Cell-Cycle Arrest and Apoptotic Functions of p53 in Tumor Initiation and Progression. *Cold Spring Harb Perspect Med. 6(3).*

Chi, C., Murphy, L., & Hu, P. (2018). Recurrent copy number alterations in young women with breast cancer. *Oncotarget. 9(14*), 11541-11558.

Corces, M., Trevino, A., Hamilton, E., Greenside, P., Sinnott-Armstrong, N., Vesuna, S., Satpathy, A., Rubin, A., Montine, K., et al. (2017). An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat Methods*. *14(10),* 959-962.

Cuddapah, S., Jothi, R., Schones, D., Roh, T., Cui, K., & Zhao, K. (2009). Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. *Genome Res*. *19*, 24-32.

Dekker, J., Rippe, K., Dekker, M., & Kleckner, N. (2002). Capturing chromosome conformation. *Science*. *295(5558),* 1306–1311.

DeLeo, A., Jay, G., Appella, E., Dubois, G., Law, L., & Old, L. (1979). Detection of a transformation-related antigen in chemically induced sarcomas and other transformed cells of the mouse. *Proc Natl Acad Sci USA*. *76*, 2420-2424.

Deng, C., Zhang, P., Harper, J., Elledge, S., & Leder, P. (1995). Mice lacking $p21^{CIP1}$/WAF1 undergo normal development, but are defective in G1 checkpoint control. *Cell. 82*, 675–684.

Didenko, V. (2001). DNA probes using fluorescence resonance energy transfer (FRET): designs and applications. *Biotechniques. 5*,1106-1116, 1118, 1120-1121.

Dolan, D., Zupanic, A., Nelson, G., Hall, P., Miwa, S., Kirkwood, T., & Shanley D. (2015). Integrated Stochastic Model of DNA Damage Repair by Non-homologous End Joining and p53/p21-Mediated Early Senescence Signalling. *PLoS Comput. Biol. 11*, e1004246.

Egloff, S., Al-Rawaf, H., O'Reilly, D., & Murphy, S. (2009). Chromatin structure is implicated in "late" elongation checkpoints on the U2 snRNA and beta-actin genes. *Mol Cell Biol*. *29*, 4002-4013.

el-Deiry, W., Harper, J., O'Connor, P., Velculescu, V., Canman, C., Jackman J, et al. (1994). WAF1/CIP1 is induced in p53-mediated G1 arrest and apoptosis. *Cancer Res. 54*, 1169–1174.

Eliyahu, D., Raz, A., Gruss, P., Givol, D., & Oren, M. (1984). Participation of p53 cellular tumour antigen in transformation of normal embryonic cells. *Nature*. *312*, 646-649.

Eliyahu, D., Michalovitz, D., & Oren, M. (1985). Overproduction of p53 antigen makes established cells highly tumorigenic. *Nature*. *316(6024),* 158-160.

Engeland, K. (2018). Cell cycle arrest through indirect transcriptional repression by p53: I have a DREAM. *Cell Death Differ. 25(1),* 114-132.

Enoiu, M., Jiricny, J., & Schärer O. (2012). Repair of cisplatin-induced DNA interstrand crosslinks by a replication-independent pathway involving transcription-coupled repair and translesion synthesis. *Nucleic Acids Res*. *40(18),* 8953-8964.

Espinosa, J., & Emerson, B. (2001). Transcriptional regulation by p53 through intrinsic DNA/chromatin binding and site-directed cofactor recruitment. *Mol Cell. 8*, 57-69.

Everett, R., Meredith, M., Orr, A., Cross, A., Kathoria, M., & Parkinson, J. (1997). A novel ubiquitin specific protease is dynamically associated with the PML nuclear domain and binds to a herpes virus regulatory protein. *EMBO J. 16*, 1519–1530.

Farkas, T., Hansen, K., Holm, K., Lukas, J., & Bartek, J. (2002). Distinct phosphorylation events regulate p130- and p107-mediated repression of E2F-4. *J Biol Chem*. *277*, 26741–26752.

Fero, M., Randel, E., Gurley, K., Roberts, J., & Kemp, C. (1998). The murine gene $p27^{Kip1}$ is haplo-insufficient for tumor suppression. *Nature*. *396,* 177–180.

Fields, S., & Jang, S. (1990). Presence of a potent transcription activating sequence in the p53 protein. *Science*. *249*, 1046–1049.

Filippova, G., Lindblom, A., Meincke, L., Klenova, E., Neiman, P., Collins, S., Doggett, N., & Lobanenkov, V. (1998). A widely expressed transcription factor with multiple DNA sequence specificity, CTCF, is localized at chromosome segment 16q22.1 within one of the smallest regions of overlap for common deletions in breast and prostate cancers. *Genes Chromosomes Cancer. 22(1),* 26-36.

Filippova, G., Qi, C., Ulmer, J., Moore, J., Ward, M., Hu, Y., Loukinov, D., Pugacheva, E., Klenova, E., Grundy, P., Feinberg, A., Cleton-Jansen, A., et al. (2002). Tumor-associated zinc finger mutations in the CTCF transcription factor selectively alter tts DNA-binding specificity. *Cancer Res. 62(1),* 48-52.

Filippova, G. (2007). Genetics and Epigenetics of the Multifunctional Protein CTCF. *Current Topics in Developmental Biology*. *80*, 337-360.

Fischer, M., Steiner, L., & Engeland, K. (2014). The transcription factor p53: not a repressor, solely an activator. *Cell Cycle*. *13*, 3037–3058.

Fischer, M. (2017). Census and evaluation of p53 target genes. *Oncogene. 36(28),* 3943-3956.

Fong, V., Osterbur, M., Capella, C., Kim, Y., Hine, C., Gorbunova, V., Seluanov, A., & Dewhurst, S. (2011). Adenoviral vector driven by a minimal Rad51 promoter is selective for p53–deficient tumor cells. *PLoS ONE*. *6,* e28714.

Friedman, P., Chen, X., Bargonetti, J., & Prives, C. (1993). The p53 protein is an unusually shaped tetramer that binds directly to DNA. *Proc Natl Acad Sci U S A. 90*, 3319-3323.

Fridman, J., & Lowe, S. (2003). Control of apoptosis by p53. *Oncogene. 22(56),* 9030-9040.

Fudenberg, G., et al. (2016). Formation of chromosomal domains by loop extrusion. *Cell Rep. 15*, 2038–2049.

Gomes, N., & Espinosa, J. (2010). Gene-specific repression of the p53 target gene PUMA via intragenic CTCF-Cohesin binding. *Genes Dev*. *24*, 1022-1034.

Green, D. (2018). Cell death: apoptosis and other means to an end. New York: Cold Spring Harbor Press.

Grossman, S. (2001). p300/CBP/p53 interaction and regulation of the p53 response. *Eur J Biochem/FEBS. 268*, 2773–2778.

Han, J., Flemington, C., Houghton, A., et al. (2001). Expression of bbc3, a proapoptotic BH3-only gene, is regulated by diverse cell death and survival signals. *Proc Natl Acad Sci USA. 20*, 11318–11323.

Han, D., Chen, Q., Shi, J., Zhang, F., & Yu, X. (2017). CTCF participates in DNA damage response via poly(ADP-ribosyl)ation, *Sci. Rep*. *7*, 43530.

Hark, A., Schoenherr, C., Katz, D., Ingram, R., Levorse, J., Tilghman, S. (2000). CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature. 6785*, 486-489.

Haupt, Y., Maya, R., Kazaz, A., & Oren, M. (1997). Mdm2 promotes the rapid degradation of p53. *Nature. 387*, 296–299.

Helmbold, H., Kömm, N., Deppert, W., & Bohn, W. (2009). Rb2/p130 is the dominating pocket protein in the p53–p21 DNA damage response pathway leading to senescence. *Oncogene. 28*, 3456–3467.

Hilmi, K., Jangal, M., Marques, M., Zhao, T., Saad, A., Zhang, C., Luo, V., Syme, A., Rejon, C., Yu, Z., Krum, A., Fabian, M., Richard, S., Alaoui-Jamali, M., et al. (2017). CTCF facilitates DNA double-strand break repair by enhancing homologous recombination repair. *Sci Adv. 3(5),* e1601898.

Hine, C., Li, H., Xie, L., Mao, Z., Seluanov, A., & Gorbunova, V. (2014). Regulation of Rad51 promoter. *Cell Cycle. 13*, 2038–2045.

Honda, R., & Yasuda, H. (2000). Activity of MDM2, a ubiquitin ligase, toward p53 or itself is dependent on the RING finger domain of the ligase. *Oncogene. 19*, 1473–1476.

Howlader, N., Cronin, K., Kurian, A., & Andridge, R. (2018). Differences in Breast Cancer Survival by Molecular Subtypes in the United States. *Cancer Epidemiol Biomarkers Prev. 27(6),* 619-626.

Holwerda, S. & de Laat, W. (2013). CTCF: the protein, the binding partners, the binding sites and their chromatin loops. *Philos Trans R Soc Lond B Biol Sci. 368(1620),* 20120369.

Hu, W., Feng, Z., & Levine, A. (2012). The Regulation of Multiple p53 Stress Responses is Mediated through MDM2. *Genes Cancer. 3*, 199–208.

Hwang, P., Bunz, F., Yu, J., Rago, C., Chan, T., Murphy, M., Kelso, G., Smith, R., Kinzler, K., & Vogelstein, B. (2001). Ferredoxin reductase affects p53-dependent, 5-fluorouracil-induced apoptosis in colorectal cancer cells. *Nat Med. 7(10),* 1111-1117.

Hwang, S., Kang, M., Baik, C., Lee, Y., Hang, N., Kim, B., Han, J., Jeong, J., Park, D., Myung, K., & Lee, J. (2019). CTCF cooperates with CtIP to drive homologous recombination repair of double-strand breaks. *Nucleic Acids Res. 47(17),* 9160-9179.

Ignatov, A., Eggemann, H., Burger, E., & Ignatov, T. (2018). Patterns of breast cancer relapse in accordance to biological subtype. *J Cancer Res Clin Oncol. 144(7),* 1347–1355.

Inoue, K., & Fry, E. (2017). Haploinsufficient tumor suppressor genes. *Adv Med Biol*. *118*, 83-122.

Itahana, K., Mao, H., Jin, A., Itahana, Y., Clegg, H., Lindstrom, M., Bhat, K., Godfrey, V., Evan, G., & Zhang, Y. (2007). Targeted inactivation of Mdm2 RING finger E3 ubiquitin ligase activity in the mouse reveals mechanistic insights into p53 regulation. *Cancer Cell. 12*, 355–366.

Janiszewska, M., Stein, S., Metzger Filho, O., Eng, J., Kingston, N., Harper, N., Rye, I., Alečković, M., Trinh, A., Murphy, K., Marangoni, E., Cristea, S., et al. (2021). The impact of tumor epithelial and microenvironmental heterogeneity on treatment responses in HER2+ breast cancer. *JCI Insight. 6(11).*

Jenkins, J., Rudge, K., & Currie, G. (1984). Cellular immortalization by a cDNA clone encoding the transformation-associated phosphoprotein p53. *Nature. 312*, 651-654.

Jiang, M., Wei, Q., Wang, J., Du, Q., Yu, J., Zhang, L., & Dong, Z. (2006). Regulation of puma-α by p53 in cisplatin-induced renal cell apoptosis. *Oncogene*. 25(29), 4056–4066.

Joerger, A., & Fersht, A. (2016). The p53 pathway: origins, inactivation in cancer, and emerging therapeutic approaches. *Annual Review of Biochemistry*. *85(1),* 375–404.

Jörnvall, H., Luka, J., Klein, G., & Appella, E. (1982). A 53-kilodalton protein common to chemically and virally transformed cells shows extensive sequence similarities between species. *Proc Natl Acad Sci U S A. 79*, 287-291.

Kaaks, R., Rinaldi, S., Key, T., Berrino, F., Peeters, P., Biessy, C., Dossus, L., Lukanova, A., Bingham, S., Khaw, K., Allen, NE. Bueno-de-Mesquita, H., van Gils, C., et al. (2005) Postmenopausal serum androgens, oestrogens and breast cancer risk: the European prospective investigation into cancer and nutrition. *Endocr Relat Cancer*. 12(4), 1071-1082.

Kartalou, M., Essigmann, J. (2001). Recognition of cisplatin adducts by cellular proteins. *Mutat Res*. *478(1-2),* 1-21.

Kemp, C., et al. (2014). CTCF Haploinsufficiency Destabilizes DNA Methylation and Predisposes to Cancer. *Cell Reports. 7,* 1020-1029.

Khaled, Y., Elkord, E., & Ammori, B. (2012). Macrophage inhibitory cytokine-1: a review of its pleiotropic actions in cancer. *Cancer Biomark. 11*, 183–190.

Khoury, A., Achinger-Kawecka, J., Bert, S., Smith, G., French, H., Luu, P., Peters, T., Du, Q., Parry, A.J., & Valdes-Mora, F.*, et al.* (2020). Constitutively bound CTCF sites maintain 3D chromatin architecture and long-range epigenetically regulated domains. *Nature Communications. 11*, 54.

Kim, T., Abdullaev, Z., Smith, A., Ching, K., Loukinov, D., Green, R., Zhang, M., Lobanenkov, V., & Ren, B. (2007). Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. *Cell*. *128*, 1231-45.

Klemm, S., Zohar, S., & Greenleaf, W. (2019). Chromatin accessibility and the regulatory epigenome. *Nature Reviews.Genetics. 20*(4), 207-220.

Knudson, A. (1971). Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci USA. 68*, 820–823.

Kress, M., May, E., Cassingena, R., & May, P. (1979). Simian virus 40- transformed cells express new species of proteins precipitable by anti-simian virus 40 tumour serum. *J Virol. 31*, 472-483.

Kruse, J., & Gu, W. (2008). SnapShot: p53 posttranslational modifications. *Cell. 133*, 930.

Kritsilis, M., Rizou, S., Koutsoudaki, P., Evangelou, K., Gorgoulis, V., Papadopoulos, D. (2018). Ageing, Cellular Senescence and Neurodegenerative Disease. *Int. J. Mol. Sci. 19*, 2937.

Kubbutat, M., Jones, S., & Vousden, K. (1997). Regulation of p53 stability by Mdm2. *Nature. 387*, 299–303.

Lane, D., & Crawford, V. (1979). T antigen is bound to a host protein in SY40-transformed cells. *Nature. 278*, 261-263.

Lang, F., Li, X., Zheng, W., Li, Z., Lu, D., Chen, G., Gong, D., Yang, L., Fu, J., Shi, P., & Zhou J. (2017). CTCF prevents genomic instability by promoting homologous recombination-directed DNA double-strand break repair. *Proc Natl Acad Sci USA*. *114(41),* 10912-10917.

Lawrence, M., Stojanov, P., Mermel, C., Robinson, J., Garraway, L., Golub, T., Meyerson, M., Gabriel, S., Lander, E., & Getz, G. (2014). Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature*. *505*, 495-501.

Lee, R., Kang, M., Kim, Y., Yang, B., Shim, H., Kim, S., Kim, K., Yang, C., Min, B., Jung, W., Lee, E., Joo, J., Park, G., Cho, W., & Kim, H. (2022). CTCF-mediated chromatin looping provides a topological framework for the formation of phase-separated transcriptional condensates. *Nucleic Acids Res. 50(1),* 207-226.

Lee, J., Mustafa, M., Kim, C., & Kim, M. (2017). Depletion of CTCF in breast cancer cells selectively induces cancer cell death via p53. *J. Cancer. 8,* 2124–2131.

Ling, J., Li, T., Hu, J., Vu, T., Chen, H., Qiu, X, Cherry, A., & Hoffman, A. (2006). CTCF mediates interchromosomal colocalization between Igf2/H19 and Wsb1/Nf1. *Science. 312*, 269–272.

Liu, J., Doty, T., Gibson, B., & Heyer, W. (2010). Human BRCA2 protein promotes RAD51 filament formation on RPA-covered single-stranded DNA. *Nat. Struct. Mol. Biol. 17*, 1260–1262.

Liu, S., Bishop, W., & Liu, M. (2003). Differential effects of cell cycle regulatory protein p21(WAF1/Cip1) on apoptosis and sensitivity to cancer chemotherapy. *Drug Resist. 6*, 183-195.

Li, Y., Haarhuis, J., Cacciatore, A., Oldenkamp, R., van Ruiten, M., Willems, L., et al. (2020). The structural basis for cohesin-CTCF-anchored loops. *Nature.* 2020. *578(7795),* 472–476.

Lobanenkov, V., Nicolas, R., Adler, V., Paterson, H., Klenova, E., Polotskaja, A., & Goodwin, G. (1990). A novel sequence-specific DNA binding protein which interacts with three regularly spaced direct repeats of the CCCTC-motif in the 5′-flanking sequence of the chicken c-myc gene. *Oncogene. 5*, 1743–1753.

Luo, H., Wang, F, Zha, J., Li, H., Yan, B., Du, Q., et al. (2018). CTCF boundary remodels chromatin domain and drives aberrant HOX gene transcription in acute myeloid leukemia. *Blood*. *132(8),* 837–848.

Mahmoudi, S., Henriksson, S., Corcoran, M., Méndez-Vidal, C., Wiman, KG., & Farnebo, M. (2009). Wrap53, a natural p53 antisense transcript required for p53 induction upon DNA damage. *Mol Cell. 33(4),* 462-471.

Mason, J., Goodfellow, P., Grundy, P., Skinner, M. (2000). 16q loss of heterozygosity and microsatellite instability in Wilms' tumor. *J Pediatr Surg*. *35(6),* 891-897.

Massenkeil, G., Caduff, R., Oberhuber, H., Schwartewaldhoff, I., Diener, P., Bannwart, F., Walt, H., & Schafer, R. (1995). Loss of heterozygosity on chromosome 16q, expression of her2/neu and p53 mutations in endometrial cancer. *Int J Oncol. 6(1),* 157-62.

Matheu, A., Maraver, A., & Serrano, M. (2008). The Arf/p53 Pathway in Cancer and Aging. *Cancer Res. 28(15*), 6031-6034.

McLure, K., & Lee, P. (1998). How p53 binds DNA as a tetramer. *EMBO J.* 17(12), 3342-50.

Momparler, R., Karon, M., Siegel, S., & Avila, F. (1976). Effect of adriamycin on DNA, RNA, and protein synthesis in cell-free systems and intact cells. Cancer Res *36*, 2891-2895.

Moore, J., Rabaia, N., Smith, L., Fagerlie, S., Gurley, K., Loukinov, D., Disteche, C., Collins, S., Kemp, C., Lobanenkov, V., & Filippova, G. (2012). Loss of maternal CTCF is associated with peri-implantation lethality of Ctcf null embryos. *PLoS One. 7(4*), e34915.

Mustafa, M., Lee, J., & Kim, M. (2015). CTCF negatively regulates HOXA10 expression in breast cancer cells. *Biochem.* Biophys. *Res. Commun. 467,* 828–834.

Nagaich, A., Zhurkin, V., Durell, S., Jernigan, R., Appella, E., & Harrington, R. (1999). p53-induced DNA bending and twisting: p53 tetramer binds on the outer side of a DNA loop and increases DNA twisting. *Proc Natl Acad Sci U S A. 96(5),* 1875-80.

Narendra, V., Rocha, P., An, D., Raviram, R., Skok, J., Mazzoni, E., et al. (2015). CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. *Science. 347(6225),* 1017–1021.

Natale, F., Rapp, A., Yu, W., Maiser, A., Harz, H., Scholl, A., Grulich, S., Anton, T., Horl, D., Chen, W., Durante, M., Taucher-Scholz, G., Leonhardt, H., & Cardoso, M. (2017). Identification of the elementary structural units of the DNA damage response, *Nat. Commun*. *8*, 15760.

National Cancer Institute (NIH). [Internet]. National Cancer Institute (NIH); c2019-2022. Cancer Stat Facts: Female Breast Cancer Subtypes. https://seer.cancer.gov/statfacts/html/breast-subtypes.html

Oliner, J., Pietenpol, J., Thiagalingam, S., Gyuris, J., Kinzler, K., & Vogelstein, B. (1993). Oncoprotein MDM2 conceals the activation domain of tumour suppressor p53. *Nature. 362*, 857–860.

O'Rourke, R., Miller, C., Kato, G., Simon, K., Chen, D., Dang, C., & Koeffler, H. (1990). A potential transcriptional activation element in the p53 protein. *Oncogene*. *5*, 1829–1832.

Parada, L., Land, H., Weinberg, R., et al. (1984). Cooperation between gene encoding p53 tumour antigen and ras in cellular transformation. *Nature. 113*, 649–651.

Peña-Hernández, R., Marques, M., Hilmi, K., Zhao, T., Saad, A., Alaoui-Jamali, M., del Rincon, S., Ashworth, T., Roy, A., Emerson, B., & Witcher, M. (2015). Genome-wide targeting of the epigenetic regulatory protein CTCF to gene promoters by the transcription factor TFII-I. *Proc Natl Acad Sci U S A. 112(7),* E677-686.

Quon, K., & Berns, A. (2001). Haplo-insufficiency? Let me count the ways. *Genes Dev. 15*, 2917–2921.

Rakha, E., Green, A., Powe, D., Roylance, R., & Ellis, I. (2006). Chromosome 16 tumor-suppressor genes in breast cancer. *Genes Chromosomes Cancer. 45*, 527–535.

Rasko, J., Klenova, E., Leon, J., Filippova, G., Loukinov, D., Vatolin, S., Robinson, A., Hu, Y., Ulmer, J., Ward, M., Pugacheva, E., Neiman, P., Morse, H., Collins, S., & Lobanenkov,

V. (2001). Cell growth inhibition by the multifunctional multivalent zinc-finger factor CTCF. *Cancer Res. 61(16),* 6002-6007.

Raycroft, L., Wu, H., & Lozano, G. (1990). Transcriptional activation by wild type but not transforming mutants of the p53 anti-oncogene. *Science. 249*, 1049–1051.

Rayess, H., Wang, M., & Srivatsan, E. (2012). Cellular senescence and tumor suppressor gene p16. *Int. J. Cancer. 130*, 1715–1725.

Riley, T., Sontag, E., Chen, P., & Levine, A. (2008). Transcriptional control of human p53-regulated genes. *Nat Rev. 9*, 402–412.

Rufini, A., Tucci, P., Celardo, I., & Melino, G. (2013). Senescence and aging: the critical roles of p53. *Oncogene. 32*, 5129–5143.

Sabapathy, K., & Lane, D. (2019). Understanding p53 functions through p53 antibodies. *J Mol Cell Biol. 11(4),* 317-329.

Saldaña-Meyer, R., González-Buendía, E., Guerrero, G., Narendra, V., Bonasio, R., Recillas-Targa, F., & Reinberg, D. (2014). CTCF regulates the human p53 gene through direct interaction with its natural antisense transcript, Wrap53. *Genes Dev. 28*, 723–734.

Sanborn, A., et al. (2105). Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc. Natl Acad. Sci. USA. 112*, E6456–E6465.

Scheffner, M., Huibregste, J., Vierstra, R. & Howley, P. (1993) The HPV E6 and E6-AP complex functions as a ubiquitin-protein: ligase in the ubiquitination of p53. *Cell. 75*, 495–505.

Seow, W., Agarwal, P., & Wang, K. (2022). CLOuD9: CRISPR-Cas9-Mediated Technique for Reversible Manipulation of Chromatin Architecture. *Methods Mol Biol. 2532*, 293-309.

Serrano, M., Li, Z., Dangeti, M., Musich, P., Patrick, S., Roginskaya, M., Cartwright, B., & Zou, Y. (2013). DNA-PK, ATM and ATR collaboratively regulate p53-RPA interaction to facilitate homologous recombination DNA repair. *Oncogene*. *32*, 2452-2462.

Sherr, C. (1998). Tumor surveillance via the ARF-p53 pathway. *Genes Dev. 12(19),* 2984-2991.

Soto-Reyes, E., & Recillas-Targa, F. (2010). Epigenetic regulation of the human p53 gene promoter by the CTCF transcription factor in transformed cell lines. *Oncogene. 29(15),* 2217-2227.

Speidel, D. (2010). Transcription-independent p53 apoptosis: an alternative route to death. *Trends Cell Biol*. *20*, 14–24.

Takahashi, A., Okada, R., Nagao, K., Kawamata, Y., Hanyu, A., Yoshimoto, S., Takasugi, M., Watanabe, S., Kanemaki, M.T., Obuse, C., et al. (2017). Exosomes maintain cellular homeostasis by excreting harmful DNA from cells. *Nat. Commun. 8*, 15287.

Tang, Y., Luo, J., Zhang, W., & Gu, W. (2006). Tip60-dependent acetylation of p53 modulates the decision between cell-cycle arrest and apoptosis. *Mol Cell. 24*, 827–839.

Tang, Z., Luo, O., Li, X., Zheng, M., Zhu, J., Szalaj, P., Trzaskoma, P., Magalska, A., Wlodarczyk, J., Ruszczycki, B., Michalski, P., Piecuch, E., Wang, P., et al. (2015). CTCF-Mediated Human 3D Genome Architecture Reveals Chromatin Topology for Transcription. *Cell. 163(7),* 1611-1627.

Tanimura, S., Ohtsuka, S., Mitsui, K., Shirouzu, K., Yoshimura, A., & Ohtsubo, M. (1999). MDM2 interacts with MDMX through their RING finger domains. *FEBS Lett. 447*, 5–9.

Tanwar, V., Jose, C., & Cuddapah, S. (2019). Role of ctcf in dna damage response. *Mutation Research-Reviews in Mutation Research. 780*, 61–68.

Tarasov, V., Jung, P., Verdoodt, B., Lodygin, D., Epanchintsev, A., Menssen, A., Meister, G., & Hermeking, H. (2007). Differential regulation of microRNAs by p53 revealed by massively parallel sequencing: miR-34a is a p53 target that induces apoptosis and G1-arrest. *Cell Cycle. 6*, 1586–1593.

Thurman, R. et al. (2012). The accessible chromatin landscape of the human genome. *Nature*. *489*, 75–82.

Udvardy, A., Maine, E., & Schedl, P. (1985). The 87A7 chromomere. Identification of novel chromatin structures flanking the heat shock locus that may define the boundaries of higher order domains. *J. Mol. Biol. 185*, 341-358

Valentine, J., Kumar, S., & Moumen, A. (2011). A p53-independent role for the MDM2 antagonist Nutlin-3 in DNA damage response initiation. *BMC Cancer. 11*, 79.

van Deursen, J. (2014). The role of senescent cells in ageing. *Nature. 509*, 439–446.

Wade, M., Wang, Y., & Wahl, G. (2010). The p53 orchestra: Mdm2 and Mdmx set the tone. *Trends Cell Biol. 20*, 299–309.

Witcher, M., & Emerson, B. (2009). Epigenetic silencing of the p16(INK4a) tumor suppressor is associated with loss of CTCF binding and a chromatin boundary. *Mol Cell. 34(3),* 271-84.

Wong, C. (2021). Single copy loss of CTCF leads to increased accessibility of chromatin and p53-mediated DNA damage response. MSc Thesis. McGill University, Montreal.

Wu, Y., Sarkissyan, M., & Vadgama, J. (2015). Epigenetics in breast and prostate cancer. *Methods Mol Biol. 1238*, 425-466.

Wutz, G., et al. (2017). Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins. *EMBO J. 36*, 3573–3599.

Xi, W., & Beer, M. (2021). Loop competition and extrusion model predicts CTCF interaction specificity. *Nat Commun. 12(1),* 1046.

Xiao, T., Wallace, J., & Felsenfeld, G. (2011). Specific sites in the C terminus of CTCF interact with the SA2 subunit of the cohesin complex and are required for cohesin-dependent insulation activity. *Mol Cell Biol. 31*, 2174–2183.

Xiong, S., Van Pelt, C., Elizondo-Fraire, A., Liu, G., & Lozano, G. (2006). Synergistic roles of Mdm2 and Mdm4 for p53 inhibition in central nervous system development. *Proc Natl Acad Sci U S A. 103(9),* 3226-3231.

Xu, B., Wang, H., Wright, S., Hyle, J., Zhang, Y., Shao, Y., Niu, M., Fan, Y., Rosikiewicz, W., Djekidel, M., Peng, J., Lu, R., & Li, C. (2021). Acute depletion of CTCF rewires genome-wide chromatin accessibility. *Genome Biol. 22(1),* 244.

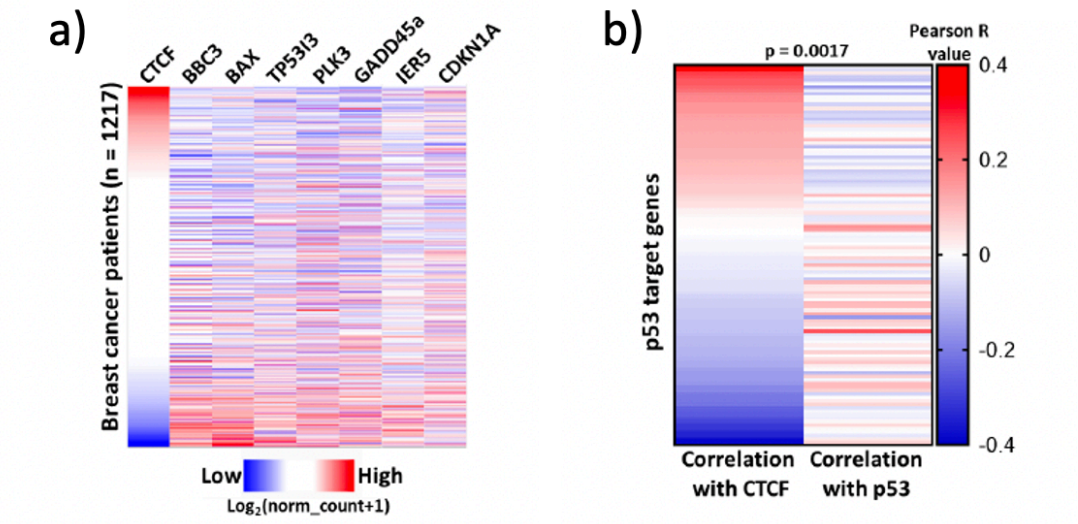
Younger, S., & Rinn, J. (2017). p53 regulates enhancer accessibility and activity in response to DNA damage. *Nucleic Acids Res. 45*, 9889-9900.

Yu, J., Zhang, L., Hwang, P., Kinzler, K., & Vogelstein, B. (2001). PUMA induces the rapid apoptosis of colorectal cancer cells. *Mol Cell. 3*, 673–682.

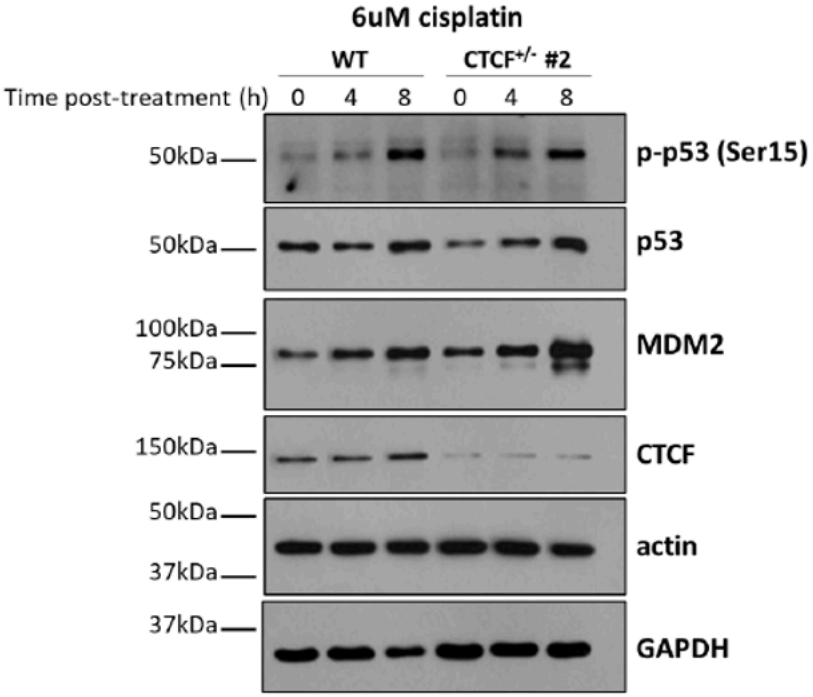
Zhan, Q., Antinore, M., Wang, X., Carrier, F., Smith, M., Harris, C., & Fornace, A. (1999). Association with Cdc2 and inhibition of Cdc2/Cyclin B1 kinase activity by the p53-regulated protein Gadd45. *Oncogene. 18*, 2892-2900.

Zhang, Y., Gao, Y., Zhang, G., Huang, S., Dong, Z., Kong, C., Su, D., Du, J., Zhu, S., Liang, Q., Zhang, J., Lu, J., & Huang, B. (2011). DNMT3a plays a role in switches between doxorubicin-induced senescence and apoptosis of colorectal cancer cells. *Int J Cancer. 128(3),* 551-561.

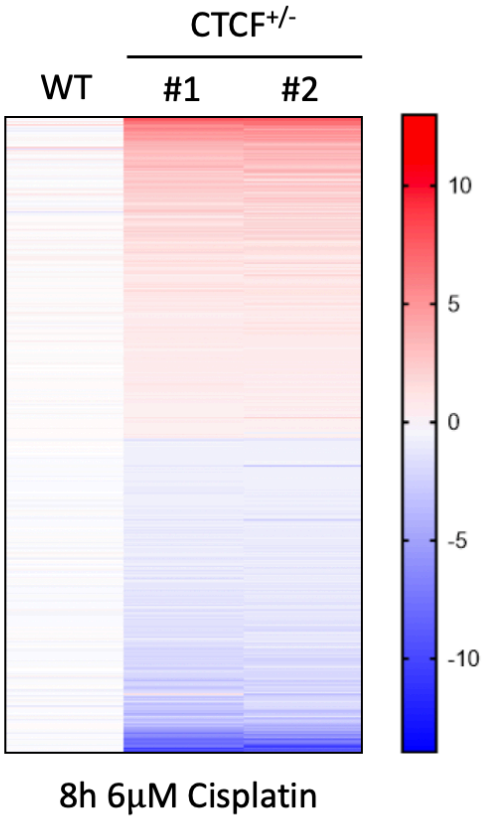
14. Supplemental Figures



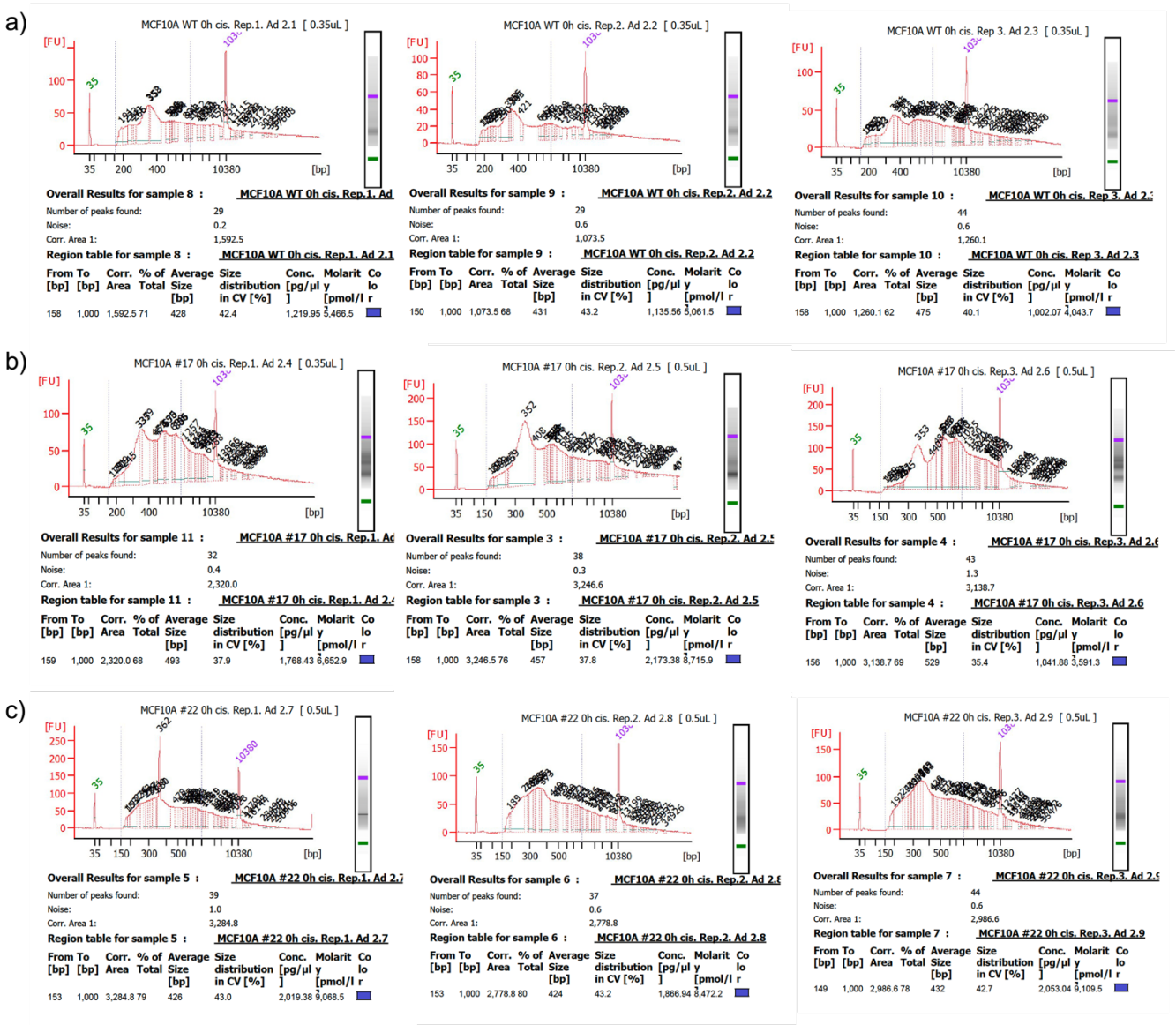
Supplemental Figure 1. CTCF may negatively regulate the expression of *TP53* target genes. a) Gene expression data of 1217 breast tumors obtained from TCGA-BRCA database in UCSC Xena browser (https://xenabrowser.net/heatmap/) and comparing CTCF gene expression level to a panel of *TP53* target genes. c) Correlation of the expression level of each *TP53* target gene to CTCF and *TP53* expression levels in 1217 breast tumors. (Wong. 2021).



Supplemental Figure 2. Western blot validating loss of CTCF in CTCF$^{+/-}$ cells. Western blot against p53, p-p53 (Ser15) and MDM2 for 6μM cisplatin-treated cells (Wong. 2021).

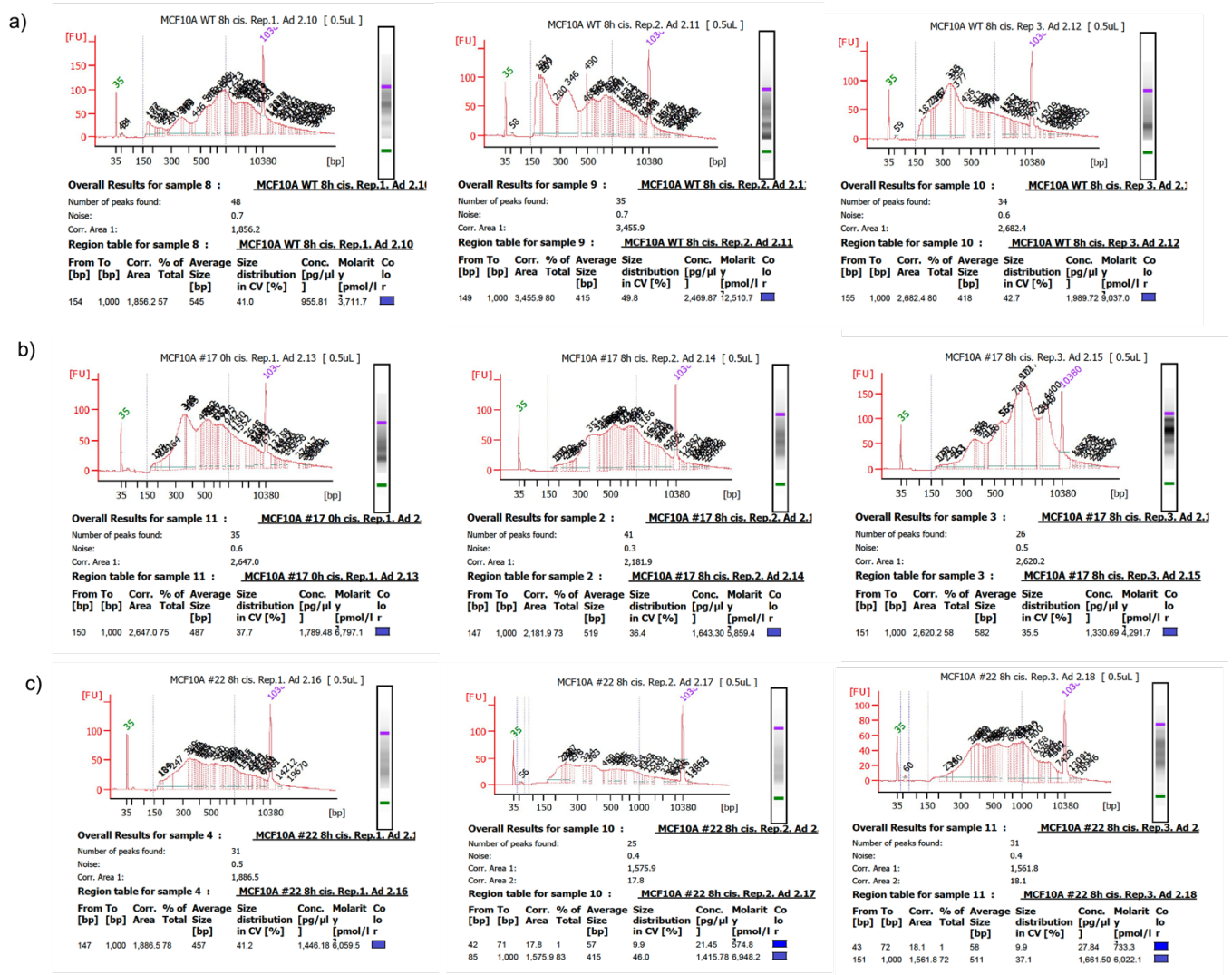


Supplemental Figure 3. RNA-seq analysis of differential MCF10A WT and CTCF$^{+/-}$ (#1 and #2) following the induction of DNA damage. 8h treatment with 6µM cisplatin, with scale on the right representing Log_2 fold change (Wong, 2021).



Supplemental Figure 4. Bioanalysis quality control for untreated MCF10A samples.

Triplicates (Rep.1-2-3; left to right) of untreated MCF10A samples were put through bioanalysis prior to sequencing for a) control cells (WT), b) CTCF+/- #1 (KO #17), and c) CTCF+/- #2 (KO #22). Ad "2.x" (2.1-2.9) indicates the unique corresponding Illumina barcode adapter sequence used for each sample.



Supplemental Figure 5. Bioanalysis quality control for treated MCF10A samples.

Triplicates (Rep. 1-2-3; left to right) of 6µM Cisplatin treated (8h) MCF10A samples were put through bioanalysis prior to sequencing for a) control cells (WT), b) CTCF$^{+/-}$ #1 (KO #17), and c) CTCF$^{+/-}$ #2 (KO #22). Ad "2.x" (2.10-2.18) indicates the unique corresponding Illumina barcode adapter sequence used for each sample.