

Single copy loss of CTCF leads to increased accessibility of chromatin and  
p53-mediated DNA damage response

Cheng Kit Wong

Division of Experimental Medicine

McGill University, Montreal

December 2021

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

Master of Science in Experimental Medicine

© Cheng Kit Wong 2021

# Table of contents

<b>1. Abstract</b> .....	<b>i</b>
<b>2. Le Résumé</b> .....	<b>ii</b>
<b>3. Acknowledgement</b> .....	<b>iii</b>
<b>4. Author contributions</b> .....	<b>iv</b>
<b>5. Introduction</b> .....	<b>1</b>
5.1 CTCF .....	1
5.1.1 Transcription factor .....	2
5.1.2 Insulator .....	2
5.1.3 Chromatin boundaries.....	3
5.1.4 Genome organization.....	3
5.1.5 DNA DSB repair .....	6
5.2 p53 .....	8
5.2.1 Structure and regulation.....	9
5.2.2 Mice model and clinical representation.....	11
5.2.3 Functions of p53 .....	11
5.2.3.1 Cell cycle arrest.....	11
5.2.3.2 DNA repair.....	12
5.2.3.3 Senescence .....	12
5.2.3.4 Apoptosis .....	13
5.2.4 Post-translational modifications of p53 .....	15
5.3 Breast cancer .....	16
<b>6. Aim</b> .....	<b>19</b>
<b>7. Materials and methods</b> .....	<b>20</b>
7.1 Cell culture .....	20
7.2 RNA isolation.....	20
7.3 Polymerase Chain Reaction (PCR) .....	20
7.4 Reverse Transcription Quantitative PCR (RT-qPCR).....	21
7.5 Western blotting .....	22
7.6 Chromatin ImmunoPrecipitation (ChIP) .....	23
7.7 MNase assay.....	26
<b>8. Results</b> .....	<b>27</b>

8.1 CTCF may negatively regulate p53-mediated DNA damage response.....	27
8.2 MCF10A CTCF <sup>+/-</sup> cells show a more robust p53-mediated DNA damage response.....	28
8.3 Increased expression of p53 target genes in CTCF <sup>+/-</sup> cells after DNA damage is p53-dependent.....	29
8.4 Stabilization of p53 after DNA damage in CTCF <sup>+/-</sup> cells is consistent with WT .....	29
8.5 p53 binding is increased at p53 target genes in CTCF <sup>+/-</sup> cells after DNA damage .....	30
8.6 Overall chromatin region in CTCF <sup>+/-</sup> cells is more accessible .....	30
8.7 Increased H3K27ac at <i>BBC3</i> and <i>BAX</i> in CTCF <sup>+/-</sup> cells after DNA damage.....	31
8.8 No changes in binding of CTCF surrounding <i>BBC3</i> and <i>BAX</i> after DNA damage .....	31
<b>9. Figures.....</b>	<b>33</b>
<b>10. Discussion.....</b>	<b>43</b>
<b>11. Concluding remarks .....</b>	<b>48</b>
<b>12. References.....</b>	<b>49</b>

## 1. Abstract

CTCF is a multifunctional epigenetic regulator that has a single copy loss in more than 50% of breast cancer patients. From TCGA breast tumor sequencing data, we observed a negative correlation between the gene expression levels of CTCF and p53 target genes in breast cancer patients, that is, breast cancer patients with low CTCF expression correlates with higher p53 target gene expression. Higher expression of p53 target genes may increase the sensitivity of these tumors to anti-cancer treatments that induces p53-mediated DNA damage response, leading to a better outcome of treatment. While p53 is found mutated in approximately 35% of breast cancers, CTCF may potentially be a biomarker for breast cancer patients with functional p53 to predict the efficacy of chemo- and radiotherapy treatments for patients with functional p53. In this thesis, we show that upon DNA damage, p53 target genes in CTCF<sup>+/-</sup> cells have a higher gene expression compared to WT cells, but p53 and p-p53 (serine 15) stabilization at protein levels are consistent between WT and CTCF<sup>+/-</sup> cells. We observed increased binding of p53 at *BBC3*, *BAX* and *CDKN1A* after DNA damage in CTCF<sup>+/-</sup> cells compared to WT cells, which is likely driving the increased p53 target gene expression in CTCF<sup>+/-</sup> cells. We also show that overall chromatin regions are more accessible after single copy loss of CTCF, possibly increasing accessibility to p53 binding upon stabilization of p53 during DNA damage. We propose that the single copy loss of CTCF results in deregulation of chromatin loops, which may lead to aberrant spreading of activating histone marks that increases the accessibility of the chromatin region near p53 target genes, resulting in increased expression of p53 target genes after DNA damage.

## 2. Le Résumé

CTCF est un régulateur épigénétique multifonctionnel pour lequel une perte mono-allélique est observée chez plus de 50 % des patientes atteintes d'un cancer du sein. À partir des données de séquençage de tumeurs du sein de la banque de données TCGA, nous avons observé une corrélation négative entre les niveaux d'expression de CTCF et des gènes régulés par p53 chez les patientes atteintes d'un cancer du sein. C'est-à-dire, que les patientes atteintes d'un cancer du sein avec une faible expression de CTCF présentent une expression plus élevée des gènes ciblés par p53. Cette plus forte expression peut augmenter la sensibilité de ces tumeurs aux traitements anticancéreux qui créent des dommages à l'ADN pour lesquels la réponse cellulaire est médiée par p53, résultant en une meilleure réponse aux traitements. Alors que p53 est muté dans environ 35% des cancers du sein, CTCF peut potentiellement être un biomarqueur pour les patientes atteintes d'un cancer du sein avec un p53 fonctionnel pour prédire l'efficacité des traitements de chimiothérapie et de radiothérapie. Dans cette thèse, nous démontrons qu'en réaction aux dommages à l'ADN, les gènes ciblés par p53 dans les cellules CTCF<sup>+/-</sup> ont une expression plus élevée que dans les cellules WT, mais la stabilisation de la protéine p53 et p-p53 (sérine 15) est similaire entre les cellules WT et CTCF<sup>+/-</sup>. Nous avons aussi observé une présence accrue de p53 sur les gènes *BBC3*, *BAX* et *CDKN1A* en réponse aux dommages à l'ADN dans les cellules CTCF<sup>+/-</sup> par rapport aux cellules WT. Ceci est probablement à l'origine de l'augmentation de l'expression de ces gènes dans les cellules CTCF<sup>+/-</sup>. Nous montrons également que la chromatine est globalement plus accessible après la perte d'un allèle de CTCF, augmentant éventuellement l'accessibilité à p53 pour se lier à l'ADN lors de sa stabilisation en réponse aux dommages à l'ADN. Nous proposons que la perte d'un allèle de CTCF entraîne une dérégulation des boucles tri-dimensionnelles de la chromatine, conduisant à une propagation aberrante des marques d'histone activatrices qui augmente l'accessibilité des régions de la chromatine proximales aux gènes cibles p53, entraînant une expression accrue de ces gènes à la suite de dommages à l'ADN.

### **3. Acknowledgement**

I would like to express my gratitude to my supervisor, Dr. Michael Witcher, for this opportunity to do graduate studies in his lab. His continuous guidance and support have tremendously helped me to make quick progress in this project. I am grateful for his willingness for prompt discussions and to share any opportunities such as workshops, application for scholarships and participation in meetings with experts of DNA repair field. All of these have enhanced my graduate studies experience.

I want to extend a sincere thanks to all members of the Witcher lab for providing a conducive, productive and fun environment in these 2 years. Special thanks to Tiejun Zhao for his help in the procurement of all materials necessary for this project, and to Benjamin Lebeau for his help in experimental ideas and discussion throughout this entire duration. Thank you everyone else in the lab for the company and moral support. I am thankful for being a part of the lab with great work and social culture.

I would also like to thank my committee members, Dr. Lorraine Chalifour, Dr. Alexandre Orthwein and Dr. Frederick Mallette, for their guidance and discussion in the course of my project.

Finally, I would like to thank all my friends and family who have provided invaluable support and encouragement. Thank you everyone.

## 4. Author contributions

This thesis is entirely written by Cheng Kit Wong (the author) and revised by Dr. Michael Witcher. All experiments described in this thesis were carried out by the author, under the supervision of Dr. Michael Witcher.

Gene expression dataset of 1217 breast tumors from the TCGA-BRCA database were extracted and analyzed with the help of Benjamin Lebeau (Fig. 10A and C). For RNA-seq experiments, pre-processing of data, 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. The author utilized the pre-processed data to generate heatmaps and carry out pathway enrichment analysis in the results section (Fig. 12).

MCF10A and 67NR CTCF<sup>+/-</sup> cell lines were previously generated in the lab by former members of the lab.

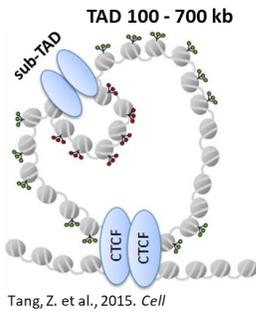
All other results were generated, analyzed and plotted by the author.

# 5. Introduction

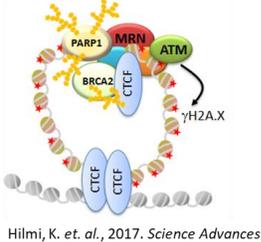
## 5.1 CTCF

CCCTC-binding factor (CTCF) is a multifunctional epigenetic regulatory protein consisting of an N-terminal region, a Zinc-finger domain comprising of 11 DNA-binding Zinc-fingers to bind DNA, and a C-terminal domain (Fig. 1). CTCF was first identified as a transcription factor having repressive activity (Filippova et al., 1996; Klenova et al., 1993; Lobanenko et al., 1990), but was subsequently demonstrated to act as a transcriptional activator as well (Liu et al., 2011; Peña-Hernández et al., 2015). Apart from its role as a transcription factor, CTCF has also been described to act as chromatin boundaries, insulator element, play a role in three-dimensional genome organization and to facilitate the repair of DNA double-strand break via homologous recombination. Genome-wide, approximately 19,000 to 50,000 CTCF binding sites have been identified in various human cell lines (Cuddapah et al., 2009; Dixon et al., 2012; Nakahashi et al., 2013). As described below, many of its disparate functions may be related to its central role in organizing chromatin into large loops.

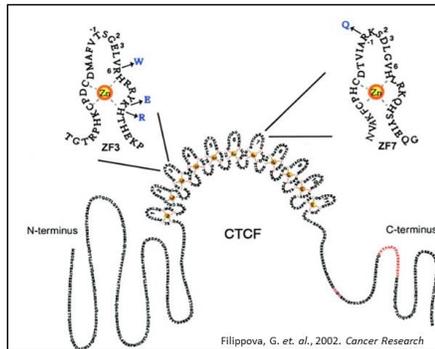
### Genome organization



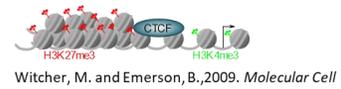
### DNA DSB repair (HR repair)



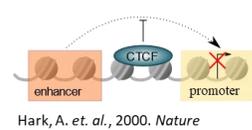
### CCCTC binding Factor (CTCF)



### Chromatin boundaries



### Enhancer blocker/Insulator



### Transcription factor

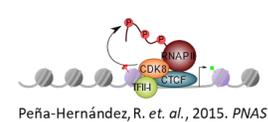


Fig. 1. **CTCF is a multifunctional epigenetic regulatory protein.** CTCF consists of N-terminus, 11 Zinc-finger domains and C-terminus.

### 5.1.1 Transcription factor

CTCF was first identified as a transcription factor that binds to CCCTC motifs upstream of chicken *c-Myc* gene promoter and it was shown to be essential for transcriptional regulation of chicken *c-Myc* (Klenova et al., 1993; Lobanenkov et al., 1990). Subsequently, conserved CTCF binding sites were found in human *c-Myc* gene and binding of CTCF was demonstrated to repress transcription of *c-Myc* gene in human (Filippova et al., 1996). CTCF has also been shown to interact directly with the largest subunit of RNA polymerase II via its C-terminal domain, likely to recruit RNA pol II for transcription (Chernukhin et al., 2007). There is evidence showing that CTCF binds to TATA-binding associated factor 3 (TAF3) to regulate transcriptional activity of cell-type specific proximal promoters that are critical for lineage specification during stem cell differentiation (Liu et al., 2011). In addition, previous data from our lab has shown that CTCF is recruited to promoter regions of metabolic-related genes by general transcription factor II-I (TFII-I) to coordinately regulate the expression of these genes (Peña-Hernández et al., 2015).

### 5.1.2 Insulator

Insulator elements are elements on the DNA that, when present in between enhancers and promoters, prevents interaction between the enhancer and promoter affected. Such elements were described in *Drosophila* where the insertion of DNA element *gypsy* in the *yellow* gene region results in mutant phenotype due to binding of *su(Hw)* protein to *gypsy*, thereby blocking the interaction between enhancer and promoter of *yellow* gene (Geyer and Corces, 1992). Insertion of *gypsy* upstream of the enhancer of *yellow* gene resulted in *Drosophila* with normal phenotype. CTCF was proposed to act as an insulator protein when an insulator sequence was found upstream of chicken  $\beta$ -globin locus, which was readily bound by CTCF via its C-terminal region (Bell et al., 1999). CTCF's role as an insulator protein became evident when CTCF is found to regulate the monoallelic expression of *Igf2* and *H19* genes in mice through gene imprinting (Szabó et al., 2004). CTCF binding sites were found in the imprinting control region (ICR) between the *Igf2* and *H19* genes where CTCF binds to the ICR on the maternal allele, blocking interaction of *Igf2* gene with downstream enhancers and silencing gene expression (Fig. 2). On the other hand, the paternal allele has hypermethylated ICR which prevents binding of CTCF, thus allowing *Igf2* gene expression.

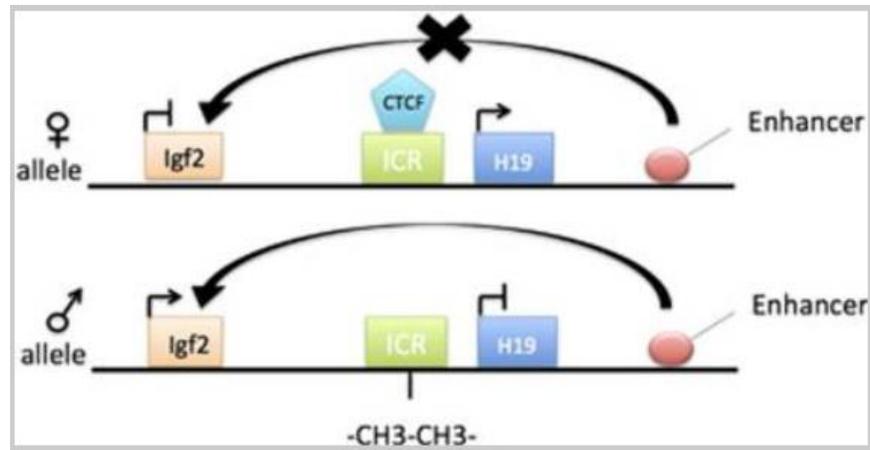


Fig. 2. **Schematic of CTCF acting as an insulator in gene imprinting.** Binding of CTCF to ICR region prevents interaction of *Igf2* to downstream enhancer, silencing the gene expression in the maternal allele. In the paternal allele, ICR region is methylated, preventing binding of CTCF and allowing expression of *Igf2* (Kim et al., 2015).

### 5.1.3 Chromatin boundaries

Chromatin boundaries function to separate differentially programmed regions of chromatin, most commonly, opened chromatin regions (euchromatin) from closed chromatin regions (heterochromatin) (Fig. 1). Since euchromatin regions are typically regions of active transcription, the establishment of chromatin boundaries is important to maintain transcriptional output. The loss of CTCF binding at tumor suppressor genes *p16<sup>INK4a</sup>*, *RASSF1A* and *CDH1* leads to repression of these genes and thereby aberrantly silencing the expression of these genes in breast cancer cells (Witcher and Emerson, 2009). In the case of *p16<sup>INK4a</sup>*, loss of CTCF as chromatin boundary resulted in deregulation and spreading of repressive histone marks into *p16<sup>INK4a</sup>* gene locus. In contrast, mutation of a highly conserved CTCF binding site in the *HoxA* gene cluster led to the disruption of a chromatin boundary, resulting in the spreading of activating histone marks that produced a 25-fold increase in expression of genes near the disrupted CTCF boundary (Narendra et al., 2015). Thus, chromatin boundaries separate differentially modified chromatin, but it is difficult to predict the precise impact of compromised boundaries.

### 5.1.4 Genome organization

The three-dimensional architecture of the genome dictates the differential frequency of interaction between regions of the chromatin and this is essential for the regulation of gene expression. Studies on human genome organization has revealed the crucial role of CTCF in the formation of

chromatin loops, and now it is thought that the roles of CTCF as transcription factor, chromatin boundary and insulator may be dependent on the ability of CTCF to form chromatin loops. Genome-wide chromatin interactions were first visualized at 1Mb resolution when Hi-C was first developed and used on normal human lymphoblastoid cell line in 2009 (Lieberman-Aiden et al., 2009). Briefly, Hi-C involves crosslinking interacting chromatin followed by an enzymatic digestion to expose the ends of the interacting chromatin. These exposed ends are biotinylated and re-ligated to generate an artificial chromatin consisting of sequences from both the original interacting chromatin regions. This re-ligated chromatin is then purified by enrichment against biotin, sequenced and mapped to the original chromatin regions.

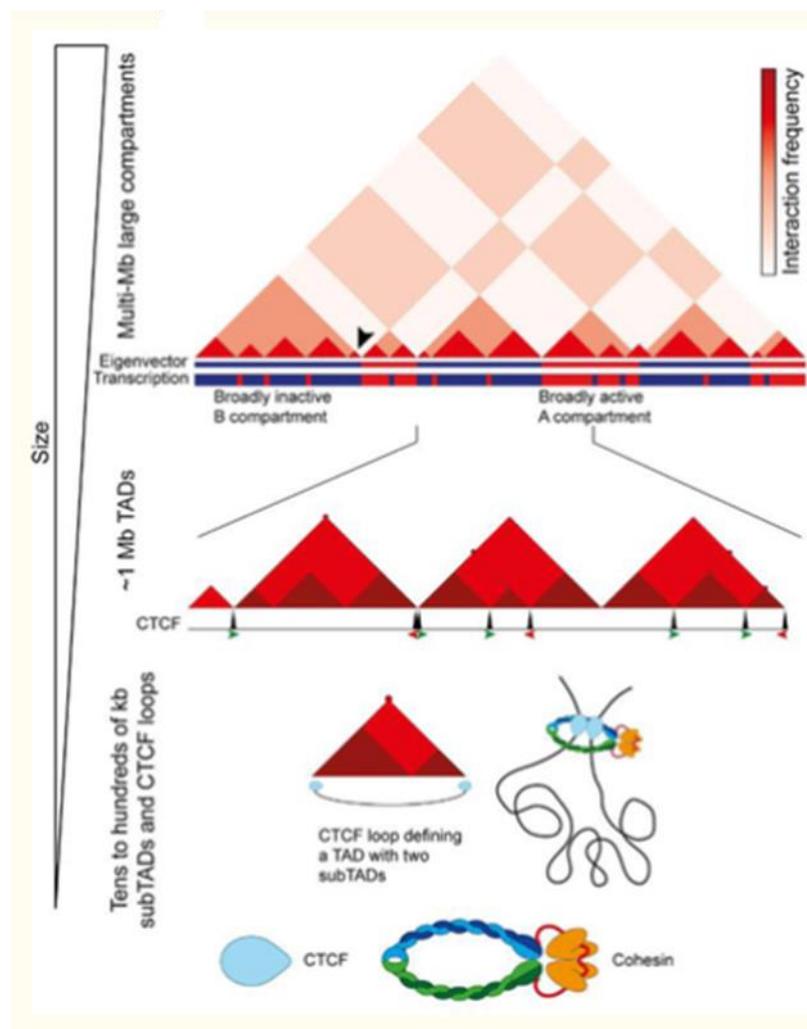


Fig. 3. **Chromatin organization visualized by Hi-C data.** CTCF is found at boundaries of frequently interacting chromatin regions known as TADs and subTADs, where CTCF and cohesin mediate the formation of chromatin loops by acting as boundaries of these loops (Rowley and Corces, 2018).

Based on Lieberman-Aiden et. al., the chromatin interaction map of the genome appears with a plaid pattern that can generally be distinguished into 2 groups called compartment A and B. Further analysis revealed that chromatin interactions in compartment A are associated with opened chromatin and active transcription while compartment B is more densely packed and less actively transcribed (Fig. 3). Hi-C experiments carried out on mouse embryonic stem cells, (mESCs) and human fibroblasts, showed that within compartments, there are chromatin regions of about 1Mb that interact *in cis* with a much higher frequency than with chromatin outside this contiguous region. These are known as topologically associated domains (TADs). The boundaries of these interacting chromatin regions are enriched with CTCF binding sites and these sites are conserved across species (Dixon et al., 2012). Despite the enrichment of CTCF binding sites at chromatin boundaries, this only represents 15% of the total CTCF binding sites in the genome, reflecting the multi-functionality of CTCF (Dixon et al., 2012). Improvements in Hi-C technology allowed visualization of chromatin interaction at resolutions down to 1kb. This revealed that TADs are defined as chromatin loops over 700kb in length, that are established by binding of CTCF to convergent CTCF motifs (Rao et al., 2014). Additionally, cohesin binding is frequently found to coincide with CTCF binding sites and it has been shown that CTCF interacts directly with cohesin to likely mediate the formation of chromatin loops (Li et al., 2020). Within TADs, there are chromatin regions that interact with each other at a higher frequency to form chromatin loops within TADs, called sub-TADs. While TADs have been shown to be rather conserved across cell type, sub-TADs have been shown to be more dynamic during differentiation, as it contributes to cell-type specific gene expression (Dixon et al., 2012; Narendra et al., 2016). Genetic disruption of CTCF binding sites at sub-TADs in *Hox* gene clusters result in deregulation of gene expression and leads to developmental defects in mice, due to differentiation into incorrect cell-type during development (Narendra et al., 2016). Moreover, comparison between fetal and adult hematopoietic stem cells showed more than 60% cell-type specific enhancer-promoter interactions arising from changes in sub-TAD interactions while compartment and TADs show limited changes between the two cell types (Chen et al., 2019). This suggests that the organization of sub-TADs are likely cell-type dependent.

A strongly supported model of chromatin loop formation by the “loop extrusion model” in the recent years has further established the role of CTCF in genome organization (Davidson and Peters, 2021). In the loop extrusion model, cohesin is thought to continuously reel in chromatin

and stops when it encounters CTCF boundaries to form a chromatin loop (Fig. 4). Indeed, CTCF has been shown to physically interact with cohesin via its N-terminal region and stabilizes cohesin on chromatin (Li et al., 2020; Rubio et al., 2008). Depletion of cohesin led to the loss of chromatin loops while depletion of CTCF led to the disruption of the boundaries of chromatin loops such that the boundaries became less defined (Wutz et al., 2017).

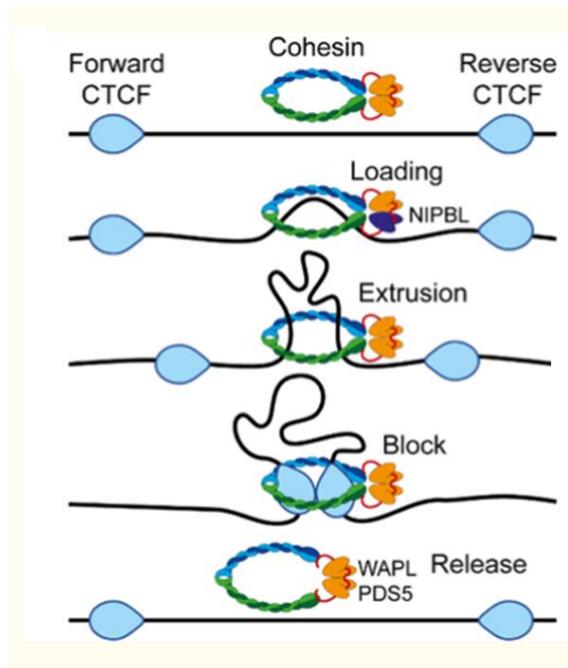


Fig. 4. **Model for chromatin loop extrusion.** In this model, cohesin continuously reel both sides of the chromatin in to form a loop, and extrusion process is blocked by CTCF boundaries, which results in the formation of a chromatin loop (Rowley and Corces, 2018).

### 5.1.5 DNA DSB repair

The role of CTCF in facilitating multiple steps of the DNA double strand break (DSB) repair pathway has been recently described. DNA double strand breaks in cells are repaired by two major pathways, error-prone non-homologous end joining (NHEJ) and high-fidelity homologous recombination (HR) repair (Fig. 5). Our lab has previously demonstrated that CTCF is involved in the regulation of HR repair but not NHEJ (Hilmi et al., 2017). In the HR repair pathway, DNA DSBs are detected by the MRE11-Rad50-NBS1 (MRN) complex that initiates a signalling cascade which involves recruitment of ATM kinase that phosphorylates histone variant H2A.X to form  $\gamma$ H2A.X DNA damage foci. One of the key steps in HR repair is the end resection mediated by the endo- and exonuclease activity of MRE11 in complex with CtIP. End resection nicks away ends

of the strands near the DSB, to produce single-stranded DNA (ssDNA) overhangs that are critical for strand invasion to search for sequence homology during repair (Scully et al., 2019). These ssDNA overhangs are protected from spontaneous degradation by binding of RPA. BRCA2 then targets Rad51 to ssDNA to promote and stabilizes Rad51 binding to ssDNA, replacing RPA and resulting in Rad51 nucleoprotein filaments that are involved in homology search on sister chromatid and strand invasion.

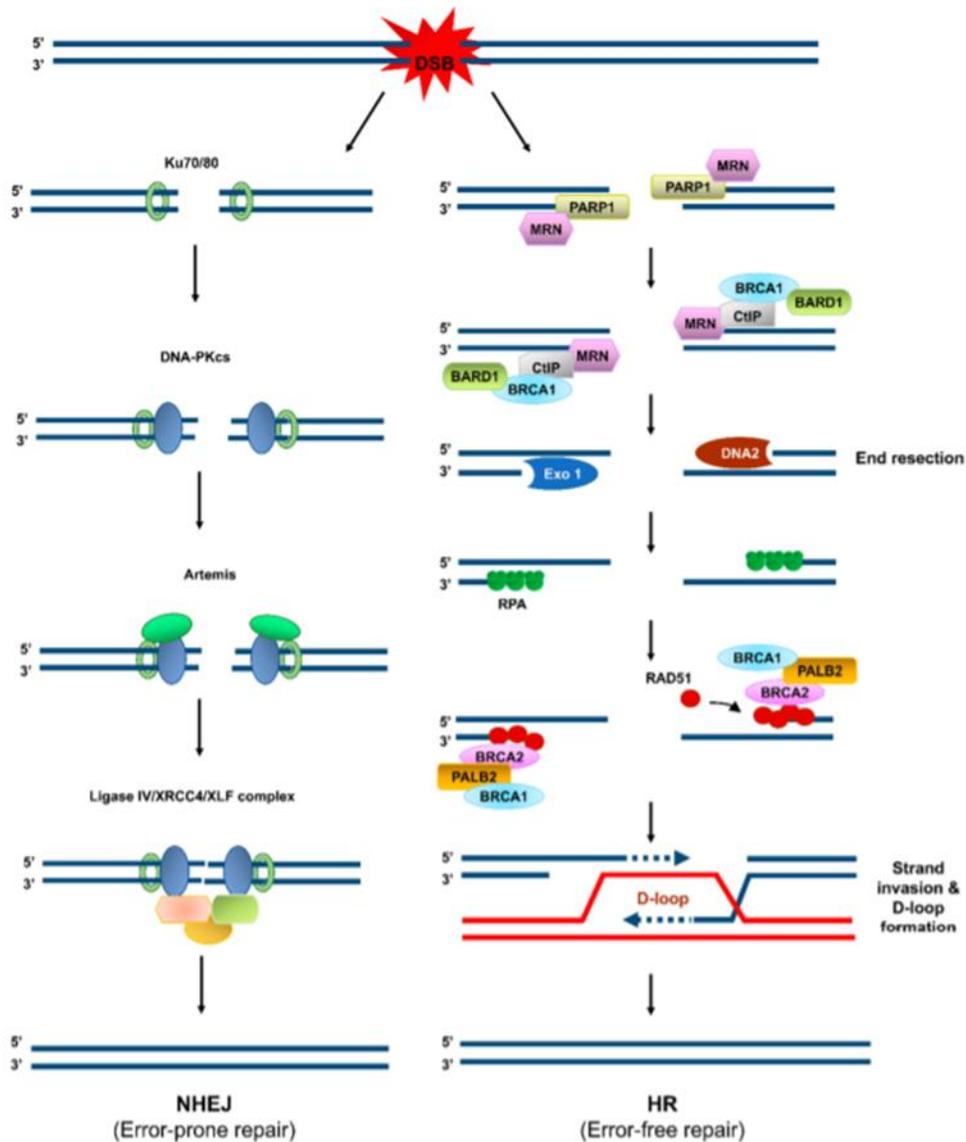


Fig. 5. Major repair pathways for DNA double strand breaks in human cells. DNA DSBs are repaired by NHEJ or HR repair pathways in human cells (Kang and Lee, 2021).

Our lab and others have previously shown that CTCF is rapidly recruited to DNA DSB sites (Han et al., 2017; Hilmi et al., 2017). The binding of CTCF to DSB sites seems to be independent of the known 19,000 to 50,000 CTCF binding sites in the genome as CTCF is efficiently recruited to sites damaged by laser micro-irradiation. The rapid recruitment of CTCF to DSB was shown to be mediated by poly-(ADP)-ribosylation (PARylation) of the DSB site as inhibition of PARylation by PARP inhibitor, Olaparib, abrogates recruitment of CTCF to DSB sites (Han et al., 2017). Another group has shown that recruitment of CTCF to DSB sites is dependent on MRE11 as knockdown of MRE11 diminishes CTCF binding to DSB sites (Hwang et al., 2019). As both PARP1 and MRN complexes are sensors of DNA DSBs, both elements are likely required for the efficient early recruitment of CTCF to DNA DSB sites. Localization of CTCF to DSB sites is essential for the recruitment of CtIP, which mediates end resection in HR repair (Hwang et al., 2019). Loss of CTCF abrogates recruitment of CtIP to DSB sites. We have also shown that localization of CTCF at DSB sites promote the recruitment of BRCA2 as knockdown of CTCF compromises the recruitment of BRCA2 to DSB sites (Hilmi et al., 2017). Depletion of CTCF resulted in defective repair kinetics, consistent with the importance of CTCF in facilitating HR repair.

Aside from directly participating in modulating HR repair, CTCF has also been shown to modulate the formation of  $\gamma$ H2A.X DNA damage foci (Natale et al., 2017). CTCF act as boundaries to delimit the spread of  $\gamma$ H2A.X foci where depletion of CTCF leads to impaired formation of  $\gamma$ H2A.X foci in terms of size and number of foci. Upon ionizing radiation-induced DNA damage, CTCF binding strength has been shown to increase at CTCF binding sites, likely for the formation of  $\gamma$ H2A.X foci (Sanders et al., 2020).

## **5.2 p53**

p53 was first reported in 1979 by 6 independent research groups that were studying Simian Virus 40 (SV40)-transformed cancer cells in which they detected the presence of a host protein migrating at 53kDa. This protein was either immunoprecipitated by anti-serum against purified SV40 (Kress et al., 1979; Lane and Crawford, 1979; Linzer and Levine, 1979; Melero et al., 1979; Smith et al., 1979) or found reactive in transformed murine cell lines when tested with anti-serum against chemically-induced murine sarcomas (DeLeo et al., 1979). Tryptic peptide analysis of p53 showed that it is structurally not related to the large and small antigens of SV40. A comparison of amino

acid composition and peptide mapping of p53 extracted from mouse and human cell lines found that p53 is highly conserved, suggesting the importance of its role in cells (Jörnvall et al., 1982). Further work on p53 led researchers to believe that it functions as an oncogene as “wild-type p53” isolated from transformed murine cell lines were found to promote cellular transformation when co-transfected into cells with *Ras* oncogene (Eliyahu et al., 1984; Jenkins et al., 1984). Shortly after, the isolated murine p53 was found to be a mutant p53 instead of wild-type, due to the seminal discovery of mutations in the conserved regions of p53. This is corroborated with the discovery of mutation and deletion of p53 in tumors from colorectal carcinoma that promote *Ras* oncogene-induced transformation (Baker et al., 1989; Hinds et al., 1989). As such, p53 was re-classified as a tumor suppressor gene and since then, it became one of the most researched proteins over several decades. Presently, we know that p53 is highly conserved across species and it is one of the most frequently mutated gene in all cancers (Lawrence et al., 2014).

### 5.2.1 Structure and regulation

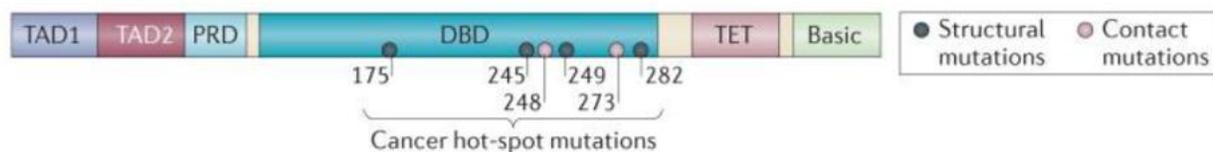


Fig. 6. **p53 protein structure.** p53 protein consists of transactivation domain 1 (TAD1); transactivation domain 2 (TAD2); proline-rich domain (PRD); DNA-binding domain (DBD); tetramerization/oligomerization domain (TET); C-terminal basic domain (Basic). The DBD frequently has point mutations in all cancers which are termed hot-spot mutations that result in structural and contact mutants of p53 (Bieging et al., 2014).

p53 is a transcription factor that binds to and activate the transcription of its downstream target genes through its transactivation domains as a homo tetramer (Friedman et al., 1993). Somatic mutations of p53 in cancers were found to be frequently occurring in the conserved regions of the DNA-binding domain of p53 (Levine et al., 1991; Walker et al., 1999). Most mutations are missense mutations, in which a single base pair mutation results in a change of one amino acid of the protein sequence. Some of these mutations are termed hotspot mutations due to the high frequency of mutation found in cancer (Fig. 6). These mutations generally give rise to two types of p53 mutants, a structural and a contact mutant. Structural mutants arise due to point mutations in the DBD that disrupts the three-dimensional folding of DBD, while contact mutants arise due

to point mutations that disrupt binding of p53 to DNA. The prevalence of hotspot mutations at the conserved DNA-binding domain are suggestive of the importance of these residues in the function of p53 that may provide selective advantage for cellular transformation upon inactivation. Under normal, and unstressed, physiological conditions, p53 is stably expressed at the mRNA level but highly regulated at the protein level by MDM2 and MDM4. p53 and MDM2 are involved in a negative regulatory feedback loop, such that the activation of p53 upregulates the transcription of MDM2, which in turn suppresses p53 protein expression (Wu et al., 1993). MDM2 suppresses p53 protein expression in several ways. MDM2 binds to both transactivation domains of p53, masking its ability to mediate transcriptional activation of downstream target genes (Oliner et al., 1993). The binding of MDM2 to p53 actively exports p53 out of the nucleus due to the nuclear export signal found in MDM2 protein (Roth et al., 1998). As p53 is a transcription factor, shuttling of p53 by MDM2 out of the nucleus prevents p53 from activating target genes, providing another level of regulation of p53 signalling. MDM2 also has an E3 ubiquitin ligase activity, which actively directs polyubiquitination of p53 thereby targeting p53 for proteasomal degradation (Honda et al., 1997; Kubbutat et al., 1997). The importance of MDM2 as one of the major regulators of p53 expression is supported by observations of embryonic lethality in *Mdm2*<sup>-/-</sup> mice model, which, upon deletion of p53, becomes viable and develops normally (Jones et al., 1995; Montes de Oca Luna et al., 1995).

MDM4 is a protein that is structurally similar to MDM2 but lacking an E3 ubiquitin ligase activity. MDM4 binds to p53 to inhibit p53-mediated transactivation to a lesser extent than MDM2, but it is not a direct p53 target gene (Shvarts et al., 1996). While the relationship between MDM2 and MDM4 in regulating p53 is complex, MDM4 has been shown to stabilize MDM2 through heterodimerization between their RING domains but, MDM4 also protects MDM2-bound p53 from degradation (Jackson and Berberich, 2000; Sharp et al., 1999). Despite this, MDM4 plays a role in regulating p53 expression as *Mdm4*<sup>-/-</sup> mice are embryonically lethal, albeit with differing phenotypes, and lethality is rescued by deletion of p53 (Parant et al., 2001). Studies performed in conditional knockouts of neuronal progenitors from embryos showed that *Mdm2*<sup>-/-</sup> mice die due to induction of apoptosis by p53 (E10.5), while *Mdm4*<sup>-/-</sup> mice show a delayed phenotype and die due to absence of cell proliferation induced by p53 (E17.5) (Xiong et al., 2006).

## 5.2.2 Mice model and clinical representation

The importance of p53 as a tumor suppressor gene is demonstrated in mice models with p53 deletion. p53<sup>-/-</sup> mice develop spontaneous tumors as early as 6 months of age, while p53<sup>+/+</sup> mice develop tumors on average of more than 2 years of age. p53<sup>+/-</sup> mice show delayed development of spontaneous tumors as compared to p53<sup>-/-</sup> mice but are still more prone to tumor development compared to p53<sup>+/+</sup> mice (Harvey et al., 1993). The spectrum of tumors arising in p53<sup>+/-</sup> and p53<sup>-/-</sup> mice are strikingly different, with p53<sup>+/-</sup> mice tend to develop soft tissue sarcomas and osteosarcomas while p53<sup>-/-</sup> mice tend to develop malignant lymphomas (Harvey et al., 1993). This is likely due to the role of p53 in the early stages of lymphoma development that, the loss of p53 function leads to increased tumorigenesis.

Another piece of evidence of p53 as an important tumor suppressor is from the discovery of families carrying an inherited mutant allele of p53, called the Li-Fraumeni Syndrome, which predisposes these families to developing early onset cancers and multiple cancers over their lifetime (Li et al., 1988).

## 5.2.3 Functions of p53

### 5.2.3.1 Cell cycle arrest

p53 plays a critical role in the regulation of processes that pertain to the maintenance of genome integrity. Activation of p53 has been shown to induce G1/S cell cycle arrest by directly upregulating the transcription of CDKN1A/p21, which then binds to cdk2-cyclin complexes to inhibit their kinase activity. (el-Deiry et al., 1993; Harper et al., 1993). p53 is also capable of inducing G2/M cell cycle arrest upon DNA damage through upregulation of p21, which binds to cdk1-cyclin complexes to inhibit their kinase activity, and direct upregulation of GADD45a which binds to cdk1 to dissociate cdk1-cyclin complexes (Harper et al., 1993; Zhan et al., 1999). Induction of cell cycle arrest during detection of DNA damage is thought to provide sufficient time for DNA repair to occur and be completed, before proceeding with cell cycle. This ensures that no genetic lesions are propagated into daughter cells upon cell division. In mice, p21<sup>-/-</sup> mice develops normally but they are more susceptible to spontaneous tumor development as compared to wild type mice. MEFs derived from p21<sup>-/-</sup> mice show defects in G1 arrest upon exposure to DNA damage (Deng et al., 1995; Martín-Caballero et al., 2001).

### 5.2.3.2 DNA repair

Apart from cell cycle arrest, p53 also plays a role in modulating and facilitating DNA repair. Our cells employ different types of DNA repair processes targeted at different types of lesions. DNA lesions that cause distortion of the DNA helix structure are typically repaired by nucleotide excision repair (NER), specifically the global genomic NER (GG-NER) while another form of NER, transcription-coupled NER (TC-NER) is initiated when the RNA synthesis is being blocked due to a bulky lesion. p53 homozygous mutant fibroblasts has been shown to have defects in repairing UV-induced lesions by GG-NER pathway, but limited effects on the TC-NER pathway (Ford and Hanawalt, 1997). While p53 does not directly localize to DNA lesions during NER, p53 upregulates the transcription of DDB2 and XPC, which are proteins involved in the early recognition of UV-induced lesions repaired by GG-NER (Fitch et al., 2003). In the case of DNA double strand breaks, p53 has been shown to regulate the HR repair pathway. p53 was shown to preferentially bind unphosphorylated RPA, likely to sequester RPA in undamaged cells. Upon detection of DNA DSB, phosphorylation of RPA leads to the dissociation of p53-RPA complex, allowing phosphorylated RPA to bind to ssDNA produced by end resection to protect from degradation and allow for recruitment of proteins required for strand invasion (Serrano et al., 2013).

### 5.2.3.3 Senescence

Senescence is a protective mechanism in cells to undergo irreversible cell cycle arrest, preventing genomic instability and tumor progression. Senescence can be triggered by multiple stimuli, including shortening of telomeres which triggers replicative senescence, abnormal oncogene activation which leads to oncogene-induced senescence and finally stress-induced senescence. During cellular stress such as DNA damage, p53 upregulates the transcription of CDKN1A/p21 to induce transient cell cycle arrest by inhibiting the activity of cdk-cyclin complexes. Inability to repair the extensive damage leads to a prolonged upregulation of p21, which results in cellular senescence (Di Leonardo et al., 1994). Upregulation of p21 is also followed by a delayed induction of p16<sup>INK4a</sup>, which acts similarly to p21 by inhibiting the activity of cdk-cyclin complexes, reducing phosphorylation of Rb and preventing cell cycle progression (Robles and Adami, 1998). Apart from p21 and p16<sup>INK4a</sup>, the p53 target gene plasminogen activator inhibitor 1 (PAI-1) has been shown to be critical in promoting cellular senescence as loss of PAI-1 in fibroblasts led to a

bypass in senescence (Kortlever et al., 2006). In *Ku80*<sup>-/-</sup> mice that undergo premature aging due to accumulation of DNA damage, MEFs derived from these mice express high levels of p21. Deletion of p21 from these *Ku80*<sup>-/-</sup> MEFs delayed the onset of senescence in these cells (Zhao et al., 2009).

### 5.2.3.4 Apoptosis

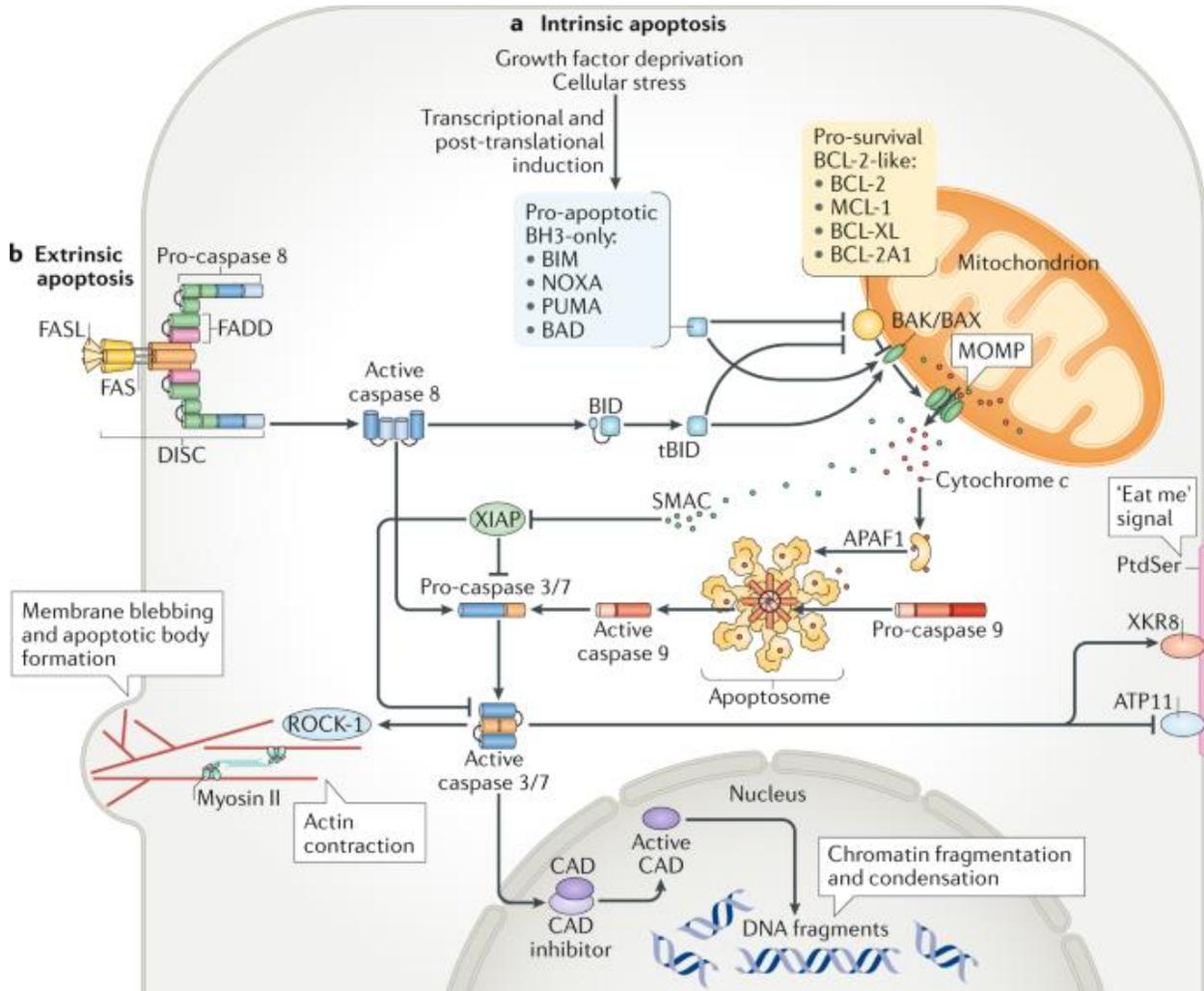


Fig. 7. **Primary mechanisms of activation of apoptotic pathway in human cells.** Apoptotic pathway can be activated by a) intrinsic apoptotic pathway or b) extrinsic apoptotic in the cells. Intrinsic apoptotic pathway is activated through permeabilization of outer mitochondrial membrane by pro-apoptotic BH-3 proteins to activate effector caspases. The extrinsic apoptotic pathway is activated by binding of death receptor ligands to death receptors, which activate downstream effector caspases for proteolytic cleavage (Bedoui et al., 2020).

The role of p53 in mediating apoptosis is crucial in anti-cancer treatment as chemo- and radiotherapies for treatment of cancer aim to induce DNA damage to the cells and cause the cancer cells to die by apoptosis. Apoptosis is induced via two different pathways in the cell, intrinsic pathway involving the permeabilization of mitochondria outer membrane and extrinsic signalling pathway involving the activation of death receptors (Fig. 7). Both pathways converge at activation of caspases that result in the cleavage of proteins. p53 has been shown to directly upregulate the expression of PUMA, NOXA and BAX, which are pro-apoptotic factors that contribute to the permeabilization of the mitochondrial outer membrane in the intrinsic apoptotic pathway (Chipuk et al., 2004; Nakano and Vousden, 2001). Muller and colleagues have shown that p53 is able to indirectly activate the extrinsic apoptotic signalling pathway by inducing expression of FAS and FASL in a panel of cancer cell lines from different organ origins, using a panel of different DNA damaging agents (Müller et al., 1998). Amino acid mutation of R172P of p53 (*Trp53<sup>515C</sup>*) in mice leads to the expression of a full length mutant p53 protein in mice. MEFs derived from *Trp53<sup>515C/515C</sup>* mice show significant defects in apoptosis induced by DNA damage, similar to *Trp53<sup>-/-</sup>* mice (Liu et al., 2004). Thymocytes from *p21<sup>-/-</sup>puma<sup>-/-</sup>noxa<sup>-/-</sup>* triple knockout mice are shown to be defective in p53-mediated apoptosis and cell cycle arrest when induced with DNA damage (Valente et al., 2013). Interestingly, these mice do not show increased susceptibility to tumor development such as that seen in *p53<sup>-/-</sup>* mice.

## 5.2.4 Post-translational modifications of p53

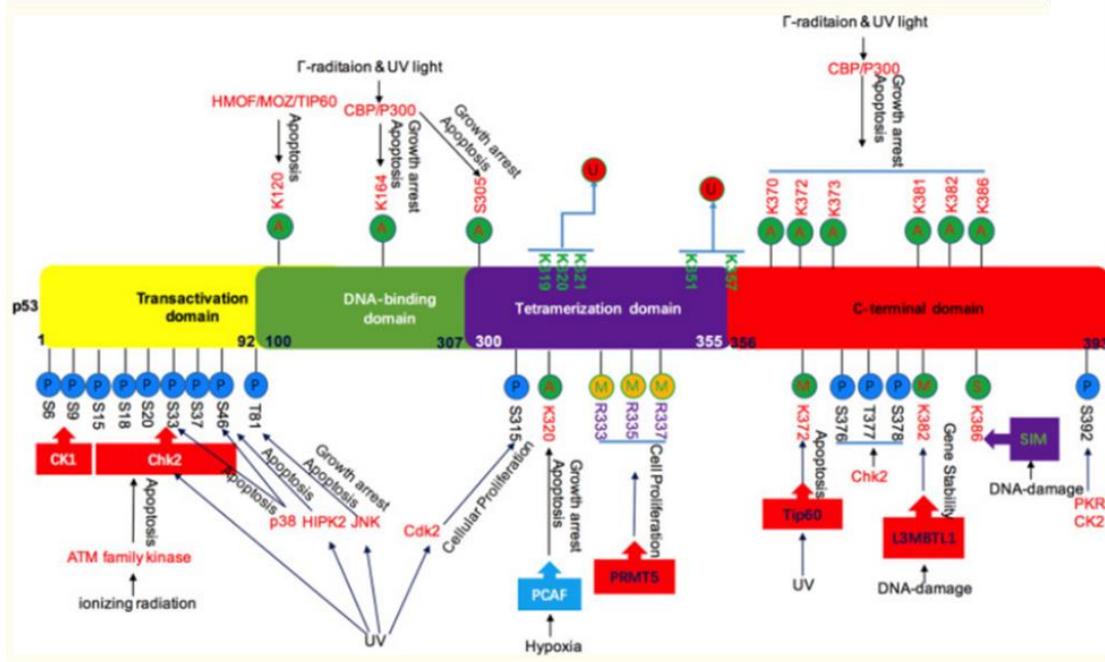


Fig. 8. **Post-translational modifications of p53.** p53 can be post-translationally modified at multiple sites to dictate activity and downstream signalling of p53. P – phosphorylation, A – acetylation, U – ubiquitination, S – sumoylation, M – methylation (Chen et al., 2020).

p53 is regulated by MDM2 at basal conditions through ubiquitination at C-terminal region of p53. Polyubiquitination of p53 directs p53 for proteasomal degradation (Honda et al., 1997; Kubbutat et al., 1997).

Following a variety of cellular stress signals such as DNA damage, hypoxia, nutrient deprivation, and oncogene activation, the normally short-lived p53 protein is stabilized and accumulates in the cell. In the case of DNA damage, p53 is stabilized by a variety of post-translational modifications. Phosphorylation of p53 at serine-15 (ser15) and serine-20 (ser20) by kinases such as ATM during a DNA double-strand break are one of the earliest modifications on p53 that contributes to p53 activity and stabilization. Mutation of S15A on p53 prevents the phosphorylation of serine-15 and abolishes p53's ability to transactivate its downstream target genes. However, binding of MDM2 to p53-S15A mutant was not affected (Dumaz and Meek, 1999). Phosphorylation of p53 at ser20 has been shown to promote p53 stabilization, as mutation of S20A on p53 prevents phosphorylation of ser20 and p53-S20A are rapidly degraded in the presence of MDM2 (Dumaz et al., 2001).

Apart from phosphorylation, p53 is also acetylated. Lysine 120 (K120) on p53 has been shown to be rapidly acetylated upon exposure to DNA damage and K120R mutation which prevents acetylation at lysine 120 leads to defects in p53-mediated apoptosis (Sykes et al., 2006). Simultaneous mutation of all major lysine residues including 6 C-terminal lysines, K120 and K164 (p53<sup>8KR</sup>) results in defective apoptotic and cell cycle arrest. Cells with p53<sup>8KR</sup> still retain the ability to phosphorylate p53 at ser15 and show diminished interaction with MDM2 (Tang et al., 2008). All these are suggestive that phosphorylation and acetylation may be synergistic in activation of p53.

Stabilization of p53 is due to a variety of post-translational modifications that dictates specific functions of p53 to either promote DNA repair, induce cell cycle arrest or to determine cell fate via senescence or apoptosis (Fig. 8). Regulation of these essential processes allow p53 to maintain the integrity of the genome and p53 has been dubbed as “guardian of the genome” due to its critical role in facilitating DNA repair and tumor suppression (Lane, 1992).

### 5.3 Breast cancer

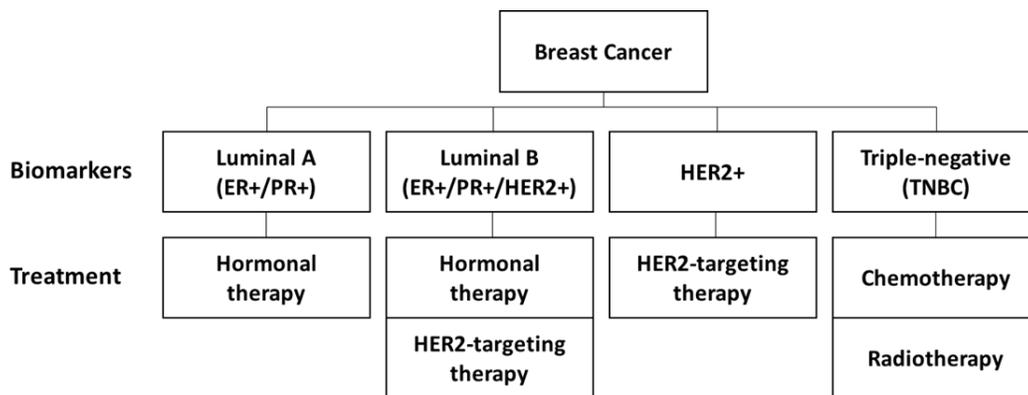


Fig. 9. **Breast cancer classification.** Breast cancer can be classified into Luminal A and B, HER2+ or TNBC depending on the biomarkers present on the tumors.

Breast cancer can be classified into subtypes based on the molecular biomarkers present on the surface of the tumors and the subtypes dictate the type of treatment administered (Fig. 9). Luminal A subtype is the most common subtype (60-70%) of breast cancer and has the best prognosis among all subtypes. This subtype expresses high levels of estrogen receptor (ER) and/or progesterone receptor (PR), low levels of Ki67 proliferative marker and negative for human epidermal growth factor receptor 2 (HER2) expression by immunohistochemistry staining (Yersal

and Barutca, 2014). Targeted hormonal therapy, with adjuvant palbociclib, a CDK4/6 inhibitor, is the current stand of care to treat Luminal A tumor subtypes due to their heavy reliance on hormonal signalling pathways (Martin et al., 2021). Tamoxifen is a competitive inhibitor of estradiol to bind ER, which blocks the downstream proliferative signalling effects of bound-ER. Luminal B subtype represents about 15-20% of breast cancers and shows a slightly poorer prognosis than Luminal A subtypes (Creighton, 2012). Luminal B subtype is characterized by high expression of ER/PR and Ki-67, positive or negative for HER2 overexpression. Standard treatment for Luminal B subtype includes targeted endocrine therapy, with or without cytotoxic therapy (chemo- or radiotherapy) due to a population of Luminal B subtypes being irresponsive to endocrine therapy. HER2 receptor blockers such as Herceptin/Trastuzumab may be included depending on the presence of HER overexpression (Creighton, 2012).

The third subtype of breast cancer is HER2-enriched and this subtype shows a slightly worse prognosis than Luminal subtypes (Seung et al., 2020). HER2-enriched subtype is characterized by a highly proliferative tumor with overexpression of the HER2 receptor due to amplification of the gene (Yersal and Barutca, 2014). The current standard of care against HER2-enriched tumors is a combination therapy of trastuzumab and pertuzumab. Both trastuzumab and pertuzumab are monoclonal antibodies that bind to HER2 receptor, blocking signal transduction (Cesca et al., 2020). Trastuzumab-emtansine (TMD1) is a second line of treatment that is conjugated with a cytotoxic drug molecule that promotes stabilization of microtubules, leading to induction of cell cycle arrest and apoptosis (Verma et al., 2012).

Breast cancers that lack ER/PR and HER2 expression are classified as triple-negative breast cancers (TNBC), which is the subtype with the worst prognosis (Seung et al., 2020). Due to the lack of ER, PR and HER2 expression, endocrine and targeted therapies are not effective in treatment of TNBC. Therefore, the standard of treatment for TNBC patients is chemotherapy and radiotherapy. Chemotherapy against TNBC consist of a combinatorial regimen of different classes of chemotherapeutic drugs including taxanes, anthracyclines, cyclophosphamides, platinum agents and fluorouracil (Yin et al., 2020).

Loss of heterozygosity (LOH) in the q-arm of chromosome 16 has been described as a frequent event in breast cancer and it is implicated in breast cancer progression (Cleton-Jansen et al., 2001; Tsuda et al., 1994). It was thought that the LOH events in breast cancer led to the inactivation of

tumor suppressor genes that subsequently promote breast cancer progression. 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 in the q-arm of chromosome 16. One of smallest region of overlap in chromosome 16q LOH events in breast cancers encompasses the locus 16q22.1, the locus which CTCF resides (Filippova et al., 1998). While *CTCF*<sup>-/-</sup> mice are embryonically lethal, a *CTCF*<sup>+/-</sup> mouse model has demonstrated the role of CTCF as a haploinsufficient tumor suppressor gene, where *CTCF*<sup>+/-</sup> mice show an increased rate of spontaneous lymphoma development as compared to WT mice (Kemp et al., 2014). Consistent with this, single cell DNA sequencing of breast tumors revealed that invasive regions of breast tumor are highly populated by subclones that carry a copy number loss at locus 16q22.1 amongst other copy number alterations (Casasent et al., 2018).

Based on whole genome sequencing of breast tumors, somatic mutation of p53 occurs in about 35% of breast cancers (Kandoth et al., 2013). The majority of the p53 somatic mutations observed are due to missense and truncation mutations, both of which lead to a defective p53 protein. The distribution of p53 somatic mutations is also subtype-specific, where breast cancer subtypes with poorer prognosis (HER2+ and TNBC) tend to have a higher incidence of p53 somatic mutations (Silwal-Pandit et al., 2014). However, p53 is still not used in clinical setting as a prognostic marker as the prognostic value of p53 is not comparable to other existing clinicopathological factors such as tumor size, node status and ER/PR/HER2 status (Olivier et al., 2006; Overgaard et al., 2000; Végran et al., 2013). While p53 mutation is enriched in both HER2+ and TNBC subtypes, p53 mutation is associated with poorer outcome in HER2+ patients but for TNBC subtype, no prognostic effect was observed in terms of breast cancer specific survival (Silwal-Pandit et al., 2014).

## **6. Aim**

Based on available breast tumor sequencing data, preliminary analysis suggested a negative correlation between gene expression of CTCF and several p53 target genes (discussed in Results section). We hypothesize that CTCF may play a role in the negative regulation of p53 target genes. As single copy loss of CTCF is found in about 50% of breast cancer patients, it is likely that the loss of CTCF results in deregulation of p53 target genes in these patients, possibly potentiating p53-mediated response. This may lead to an increased sensitivity towards chemo- and radiotherapies that induce p53-mediated DNA damage response in patients with functional p53, resulting in a more effective eradication of tumor cells. In this thesis, we will describe the effects of CTCF single copy loss on p53-mediated DNA damage response and the possible mechanism involved.

## **7. Materials and methods**

### **7.1 Cell culture**

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). U2OS and 293T cells were cultured in DMEM (Wisent, cat# 319-005-CL) supplemented with 10% fetal bovine serum (Gibco, cat# 12483-020). MCF10A, U2OS and 293T cell cultures were maintained at 1:6 and discarded once they reach passage 8. NMuMG cells were cultured in DMEM supplemented with 10% fetal bovine serum and 0.01mg/mL insulin. 67NR cells were cultured in DMEM supplemented with 10% fetal bovine serum. NMuMG and 67NR cell cultures were maintained at 1:10 and discarded once they reach passage 8. All cell lines were cultured at 37°C incubator with 5% CO<sub>2</sub>. For CTCF single copy loss cell lines previously generated by CRISPR/Cas9-mediated gene knockout, single allele knockout was confirmed by sanger sequencing and lower protein expression of CTCF was confirmed by western blot. For treatment of cells with chemotherapeutic agents, cells were treated with 6µM of cisplatin (Jewish General Hospital) and 500nM doxorubicin (Jewish General Hospital) for durations indicated.

### **7.2 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.

### **7.3 Polymerase Chain Reaction (PCR)**

PCR was performed following 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.

## 7.4 Reverse Transcription Quantitative PCR (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 4 $\mu$ L of 5X All-In-One RT Mastermix and topped up to 20 $\mu$ L 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. 5 $\mu$ L of master mix was added to 1 $\mu$ L of 0.5 $\mu$ M forward and reverse primer mix, and then 2 $\mu$ L of cDNA was added and topped up with 2 $\mu$ L of nuclease-free water to 10 $\mu$ L 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:

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

qPCR Primers for 67NR:

Target	Forward sequence	Reverse sequence
B-actin	AGGCACCAGGGCGTGAT	GCCCACATAGGAATCCTTCTGAC
GAPDH	GGTGCTGAGTATGTCGTGGA	CGGAGATGATGACCCTTTTG

PRDX1	AATGCAAAAATTGGGTATCCTGC	CGTGGGACACACAAAAGTAAAGT
18S	GCTTAATTTGACTCAACACGGGA	AGCTATCAATCTGTCAATCCTGTC
BBC3	ATGGCCCGCGCACGCCAGG	CCGCCGCTCGTACTGCGCGTTG
BAX	TGGAGATGAACTGGACAGCA	GAAGTTGCCATCAGCAAACA
TIGAR	GCTTCGCCTTGACCGTTAT	GAAACCCAGTCTCCGAAAGG

## 7.5 Western blotting

Cells were cultured in 10cm dish until 80% confluent. Treated and untreated cells were collected by first aspirating the media, following by addition of ice-cold PBS. Cell scraper was used to scrape the cells off the dish and the cell suspension was transferred into a 1.5mL tube. The cells were centrifuged at 2,500 rpm for 5mins 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 MgCl<sub>2</sub>, 1mM EDTA, 1mM dithiothreitol, 2mM phenylmethylsulfonyl fluoride (PMSF), 1mM P8340 Cocktail inhibitor (Roche), 1mM bis-glycerol phosphate, and 1mM NaF] on ice for 15mins, with agitation at intervals. Cell suspension was then centrifuged at 13,000 rpm for 15mins 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 ran 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 1hr 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 10mins each wash. Secondary antibody in 5% skimmed milk/BSA was added to the blot for 1hr 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
$\beta$ -actin	Sigma Aldrich #A2228	1:5000
GAPDH (14C10)	Cell Signalling #2118	1:5000
CTCF	BD Biosciences #612149	1:1000
Goat anti-mouse HRP-conjugated	Seracare, #5450-0011	1:10000
Goat anti-rabbit HRP-conjugated	Seracare, #5220-0458	1:10000

## 7.6 Chromatin Immunoprecipitation (ChIP)

Cells were cultured in 15cm dish until 80% confluent. Treated and untreated cells were fixed by aspirating culture media and adding ice-cold 1% Formaldehyde in PBS for 10mins. After 10mins, 2mL of ice-cold 1.25M glycine was added for 5mins to quench the reaction. The solution was removed for an ice-cold PBS wash, and cells were collected by scraping using a cell scraper. Cell suspension was transferred to a 1.5mL tube and centrifuged at 2,500rpm for 5mins at 4°C. The resulting supernatant was removed, and cell pellet was flash frozen with liquid nitrogen and stored at -80°C until required. Cell pellet was resuspended with 1mL of ice-cold IP buffer [0.25% NP-40, 0.25% Triton X-100, 0.25% Sodium Deoxycholate, 0.05% SDS, 50mM Tris (pH8), 0.1M NaCl, 5mM EDTA, 1mM PMSF, 2mM NaF, 1X P8340 Cocktail Inhibitor (Roche)] and sonicated using a probe sonicator (Fisher Scientific, Sonic Dismembrator Model 500) with 5 cycles at 20%, 5 cycles at 25% and 5 cycles at 30% amplitude, each cycle at 10secs. The cell suspension was centrifuged at 13,000 rpm for 30mins at 4°C. The resulting supernatant was transferred to a new 1.5mL tube. Protein concentration was determined using Bradford reagent. For ChIP, 2mg of protein was resuspended in 1mL of IP buffer per antibody per sample. From this protein-DNA suspension, 20 $\mu$ L of sample were obtained as input and kept at -20°C. 50 $\mu$ L/mL of Protein G Plus-

Agarose Suspension Beads (Calbiochem, cat# IP04-1.5ML) were added to the protein-DNA suspension and incubated at 4°C with constant agitation for 3hr of pre-clearing. The protein-DNA suspension was centrifuged at 2000 rpm for 2mins, and the resulting supernatant was transferred into a new 1.5mL tube. The protein-DNA suspension was then incubated with fresh agarose beads and 2-5µL of antibody overnight at 4°C with constant agitation. The next day, protein-DNA suspension was centrifuged at 4000 rpm for 2mins at 4°C to pellet the agarose beads. The agarose beads containing antibody bound to crosslinked protein-DNA were washed with Wash 1, Wash 2, Wash 3 [0.1% SDS, 1% Triton X-100, 2mM EDTA, 20mM Tris (pH 8), 150/200/500mM NaCl for Wash 1,2,3 respectively] solutions containing low salt to high salt concentration. Each wash for 5mins at 4°C with constant agitation. The agarose beads were then washed with LiCl wash solution [0.25M LiCl, 1% NP-40, 1% Sodium Deoxycholate, 1mM EDTA, 10mM Tris (pH8)], and then twice with Tris-EDTA solution [10mM Tris (pH8), 1mM EDTA]. The protein-DNA was eluted with elution buffer [1% SDS, 0.1M NaHCO<sub>3</sub>] from the agarose beads at 65°C for 15mins and then de-crosslinked at 65°C overnight. The input stored at -20°C was de-crosslinked at 65°C overnight together with the samples. On the next day, 20µg of Proteinase K (Sigma, cat# 39450-01-6) was added to samples and incubated for 1h at 42°C. Next, the samples were purified by column extraction. 5 volumes of PB buffer [5M Guanidine, 30% isopropanol] was added to DNA and then transferred into column. Column was centrifuged at 13,000 rpm for 1min at room temperature. Column was then washed twice with PE buffer [0.01M Tris (pH7.5), 80% ethanol] and DNA was eluted with nuclease-free water. For ChIP-sequencing (ChIP-seq), the starting amount of protein, amount of agarose beads and antibodies used were doubled, but eluted in the same final amount as ChIP.

#### Antibodies for ChIP:

Target	Source	Amount used
p53 (DO-7)	Cell Signalling #48818	5uL
CTCF	Millipore #07-729	5uL
H3K27ac	Millipore #07-360	5uL
Normal Rabbit IgG	Millipore #12-370	2uL

#### ChIP-qPCR Primers for p53:

Target	Forward sequence	Reverse sequence
BBC3 -1.4kb	TGTCTCAATTAACAACAAC AAAACC	AACTTCTACCTGCAATTTTA CTGACCT
BBC3 +1.3kb peak	TCAGTGTGTGTGTCCGACTGT C	GGCAGGGCCTAGCCCAAGG
BAX +0.4kb peak	TAGCGTCCCCTAGCCTCTT	CCAGACAACCTGAGTCCCTGA
BAX +2.5kb	ATGGGTGTGCACCATTATCC	GGAGTTCAAGACCAGCCTGA
CDKN1A -3kb	CCGGCCAGTATATATTTTAA TTGAGA	AGTGGTTAGTAATTTTCAGT TTGTCAT
CDKN1A -2.3kb peak	AGCAGGCTGTGGCTCTGATT	CAAATAGCCACCAGCCTCT TCT

ChIP-qPCR Primers for H3K27ac:

Target	Forward sequence	Reverse sequence
BBC3 -1.4kb	TGTCTCAATTAACAACAAC AAAACC	AACTTCTACCTGCAATTTTA CTGACCT
BBC3 +1.3kb peak	TCAGTGTGTGTGTCCGACTGT C	GGCAGGGCCTAGCCCAAGG
BBC3 +10kb	AGAGGACAAACACGGAATGC	CTTGGGGGATGTCTTTCTCA
BAX +0.4kb peak	TAGCGTCCCCTAGCCTCTT	CCAGACAACCTGAGTCCCTGA
BAX +2.5kb	ATGGGTGTGCACCATTATCC	GGAGTTCAAGACCAGCCTGA
BAX +3.5kb	AAATGCTCCTGGCTGTTGTT	CTACCACCAGGGCTTGTCAT

ChIP-qPCR Primers for CTCF:

Target	Forward sequence	Reverse sequence
BBC3 -1kb	TCAAAACGCCAACAACAAA	GTGTCGAACTCCGGACCTTA
BBC3 +2kb peak	TTCCTCTGGATCGACACCAC	TGTGGATCTGCAGGTGTCTC
BBC3 +7kb peak	AATCATGGAGCATTCTTAGCT TAGC	TATTGAGGGAAAAGGAATT CTGG

BAX -110kb peak	AGTGGTCCTCACCTCACAC	GATGGCAGTAGCACACAGG A
BAX +3.5kb	AAATGCTCCTGGCTGTTGTT	CTACCACCAGGGCTTGTCAT
BAX +16kb peak	ATTCCCATAACCGTGCACTC	TTGCGATTAAGACGGTAGGC

## 7.7 MNase assay

Cells were cultured in a 10cm dish until approximately 80% confluent before harvesting by scraping. The cells were washed with PBS and centrifuged to pellet at 300g for 10mins at 4°C. Cell pellet was then resuspended in 1mL of lysis buffer [0.5% NP-40, 0.01M Tris (pH7.5), 0.01M NaCl, 3mM MgCl<sub>2</sub>, 0.5mM PMSF, 1mM DTT] for 5min at 4°C. Samples were next centrifuged at 120g for 10min at 4°C to pellet the nuclei and the resulting supernatant was aspirated. The nuclei pellet was then resuspended in 500µL of CaCl<sub>2</sub> buffer [0.02M HEPES (pH7.5), 0.05M KCl, 10% glycerol, 1mM CaCl<sub>2</sub>, 0.2mM EDTA, 0.5mM PMSF, 1mM DTT] and two 20µL samples were taken from the resuspended nuclei as input. One input was added with 0.1M EDTA while the other input was left at room temperature to control for degradation. 100µL of nuclei suspension was transferred into a new tube and 10U of MNase diluted in CaCl<sub>2</sub> buffer was added. 20µL of MNase-treated nuclei suspension were transferred into a new tube containing 2µL of 0.1M EDTA for each specified time point. Next, 178µL of EDTA buffer [0.02M HEPES (pH7.5), 0.05M KCl, 10% glycerol, 0.01M EDTA, 0.5mM PMSF, 1mM DTT] was added to each sample to bring the total volume up to 200µL. 180µL of WSN buffer [Per sample, 30µL water, 20µL 10% SDS, 40µL 5M NaCl] was added to each sample and then 100µg of RNase A was added to the samples to be incubated at 37°C for 20mins. 100µg of Proteinase K was then added to the samples and incubated at 65°C for 1h. Next, the samples were purified by DNA column extraction. 5 volumes of PB buffer [5M Guanidine, 30% isopropanol] was added to samples and then transferred into column. Column was centrifuged at 13,000 rpm for 1min at room temperature. Column was then washed twice with PE buffer [0.01M Tris (pH7.5), 80% ethanol] and DNA was eluted in nuclease-free water. The concentration of eluted DNA was measured with nanodrop and diluted to 15ng/µL. The DNA was run on 1.2% agarose gel and 15µL of the eluted DNA was loaded for each sample.

## 8. Results

### 8.1 CTCF may negatively regulate p53-mediated DNA damage response

We obtained global gene expression data of 1217 breast tumors from the TCGA-BRCA database within the UCSC Xena browser (<https://xenabrowser.net/heatmap/>) and compared the gene expression levels of CTCF to a panel of p53 target genes extracted from a list of “top” p53 target genes (Goldman et al., 2020). This list of p53 target genes was obtained from a meta-analysis of multiple high throughput datasets to determine p53 target genes that commonly appear as top hit upon p53 activation (Fischer, 2017), where p53 target genes that appeared as a hit in at least 6 high throughput datasets were included as top p53 target genes. We noticed that patients with breast tumors expressing low levels of CTCF had a tendency to express higher levels of p53 target genes (Fig. 10A). From the list of top p53 target genes containing 116 p53 target genes in total, we found that 65.5% of the p53 target genes’ expression significantly correlated with CTCF gene expression levels while 37% of the p53 target genes significantly correlated with p53 gene expression levels in the tumors (Fig. 10B). Next, we correlated the gene expression level of each p53 target gene to CTCF and p53 respectively. We observed a significant ( $p = 0.0014$ ,  $R = -0.3$ ) negative correlation between the correlation of expression level of p53 target genes to CTCF and p53 respectively (Fig. 10C). For example, *BBC3* (encodes PUMA), a pro-apoptotic gene regulated by p53, showed a negative correlation ( $R = -0.37$ ) to CTCF expression levels but a positive correlation ( $R = 0.03$ ) to p53 expression levels in the breast tumors. Similar trend is observed for p53 target genes commonly known to elicit p53-mediated DNA damage response such as *CDKN1A* and *GADD45a* in mediating cell cycle arrest, and proapoptotic gene *BAX*. This negative correlation suggests that CTCF may be negatively regulating the expression of p53 target genes. Relapse-free survival for breast cancer patients with all statuses of p53 showed a poorer prognosis for patients with low CTCF expression in their tumors (Fig. 10D), but when these patients are stratified by wild-type p53 status in their tumors, patients with low CTCF expression showed better prognosis (logrank  $P = 0.0057$ ) (Fig. 10E). As single copy loss of CTCF is prevalent in more than 50% of breast cancers (Goldman et al., 2020), we sought to investigate the effects of CTCF single copy loss in terms of p53-mediated DNA damage response.

## 8.2 MCF10A CTCF<sup>+/-</sup> cells show a more robust p53-mediated DNA damage response

Our laboratory has previously generated CTCF hemizygous knockout (CTCF<sup>+/-</sup>) in MCF10A cell line using CRISPR/Cas9-mediated gene editing (Fig. 11A). MCF10A cell line is frequently used as a model for breast epithelial cells as it is non-tumorigenic but immortalized (Qu et al., 2015), thus allowing us to explore the effects of CTCF single copy loss on p53-mediated DNA damage response under basal conditions without confounding factors from the chaotic genome of cancer cells. To characterize the kinetics of p53-mediated response in CTCF<sup>+/-</sup> cells, we treated these cells with the clinically-relevant DNA damaging agents, 6 $\mu$ M cisplatin and 500nM doxorubicin for 2h, 4h and 8h. We assessed the mRNA expression levels of a panel of known p53 target genes (*BBC3*, *BAX* and *CDKN1A*) as a readout of p53-mediated DNA damage response. For both cisplatin and doxorubicin-treated cells, p53 target genes were generally upregulated by 4h in WT and CTCF<sup>+/-</sup> cells (Fig. 11B and C). More importantly, our data indicate that CTCF<sup>+/-</sup> cells show a more robust upregulation of p53 target genes compared to WT cells. This prompted us to investigate the difference in global transcriptome levels in CTCF<sup>+/-</sup> cells by RNA-seq after DNA damage and we chose to look at 8h post-treatment timepoint as it showed a more robust upregulation of p53 target genes.

RNA-seq data showed 2 distinct profiles of global transcriptomic changes possibly contributed by the dosage and propensity of different types of DNA damage induced by cisplatin and doxorubicin. Cisplatin-treated cells show a mild change of gene expression in MCF10A WT cells 8h after treatment while in stark contrast, CTCF<sup>+/-</sup> cells show a broad and significant change of gene expression after treatment (fold change > 1.5, False Discovery Rate (FDR) < 0.05) (Fig 12A). Pathway enrichment analysis of cisplatin-treated cells showed genes significantly upregulated in CTCF<sup>+/-</sup> cells are enriched with pathways regulated by p53 (Fig. 12B), while pathways promoting proliferation are enriched in genes significantly downregulated (Fig. 12C). This result is expected as treatment of cells with cisplatin induces p53-mediated DNA damage response, leading to upregulation of p53 target genes while downregulating proliferative genes to result in cell cycle arrest for DNA repair to occur. We observed more robust changes in gene expression in WT cells after 8h doxorubicin treatment than with cisplatin. Consistent with the effects we observed with cisplatin, the changes in gene expression in CTCF<sup>+/-</sup> cells after doxorubicin treatment is enhanced

(fold change > 1.5, FDR < 0.05) (Fig. 12D). Pathway enrichment analysis showed an enrichment in genes regulated by p53 response pathways for upregulated genes (Fig. 12E), while an enrichment in downregulation of cell cycle progression pathways (Fig. 12F), indicating cell cycle arrest to facilitate DNA DSB repair. Finally, we looked at differentially expressed p53 target genes in our RNA-seq data, and the results are in corroboration with our RT-qPCR data, in which we observed a more robust upregulation of p53 target genes in CTCF<sup>+/-</sup> cells compared to WT for both cisplatin and doxorubicin-treated cells (Fig. 12G and H).

### **8.3 Increased expression of p53 target genes in CTCF<sup>+/-</sup> cells after DNA damage is p53-dependent**

Next, we sought to determine whether the increased expression of p53 target genes in CTCF<sup>+/-</sup> cells after DNA damage is dependent on p53. We utilized a CTCF<sup>+/-</sup> mouse mammary carcinoma cell line (67NR) that was previously generated in our laboratory. The 67NR cell line has been reported to be devoid of p53 protein expression (Johnstone et al., 2015) and we observed the same result by western blot after 8h doxorubicin treatment, with a non-transformed mouse mammary gland epithelial cell line (NMuMG) as positive control of p53 expression (Fig. 13A). Exposure of the 67NR CTCF<sup>+/-</sup> cells to 6μM cisplatin and 500nM doxorubicin for 8h has revealed no significant upregulation of known p53 target genes tested (*BBC3*, *BAX*, *TIGAR*) (Fig. 13B), suggesting that the increased expression of p53 target genes observed in MCF10A CTCF<sup>+/-</sup> cells after DNA damage is p53-dependent.

### **8.4 Stabilization of p53 after DNA damage in CTCF<sup>+/-</sup> cells is consistent with WT**

Next, we assessed if p53 is differentially stabilized and activated in CTCF<sup>+/-</sup> cells compared to WT cells which may contribute to the increased p53 target gene expression in CTCF<sup>+/-</sup> cells. As the changes in gene expression of our CTCF<sup>+/-</sup> #1 and #2 appears similar in both our RT-qPCR and RNA-seq data, we focused on CTCF<sup>+/-</sup> #2 for downstream experiments. We looked at p53 and p-p53 (ser15) protein levels by western blotting after 4h and 8h of treatment, in which p-p53 (ser15) is an essential post-translational modification to initiate p53-mediated DNA damage response (Dumaz and Meek, 1999). Both cisplatin and doxorubicin-treated cells showed consistent amount of p53 and p-p53(ser15) protein stabilization at 4h and 8h post-treatment between MCF10A WT and CTCF<sup>+/-</sup> cells (Fig. 14A and B). Additionally, MDM2 protein, which is a direct target of p53

transcriptional activation showed higher protein expression level in CTCF<sup>+/-</sup> cells compared to WT cells after DNA damage, likely a result of increased gene expression levels in CTCF<sup>+/-</sup> cells. This suggests that the more robust upregulation of p53 target genes in CTCF<sup>+/-</sup> cells may be regulated at the transcriptional level.

### **8.5 p53 binding is increased at p53 target genes in CTCF<sup>+/-</sup> cells after DNA damage**

We then postulated that the more robust p53 response in CTCF<sup>+/-</sup> cells may be driven by changes in p53 binding at p53 target genes, without significant changes at the protein level. It has been reported that, at least at the *CDKN1A* gene locus, transcriptional activation by p53 may occur without an increase in p53 binding to the gene promoter (Espinosa and Emerson, 2001; Younger and Rinn, 2017). Moreover, a recent study has shown that MCF10A cells transformed with oncogenic Ras led to a global change in transcriptome expression that may be due to transcriptional reprogramming by the re-distribution of p53 binding across the genome, while having negligible changes in p53 mRNA and protein levels (Schwartz et al., 2020). Therefore, it is possible that in our case, a differential binding of p53 in CTCF<sup>+/-</sup> cells may be driving the more robust p53 response despite similar levels of p53 stabilization at the protein level after DNA damage. To explore this possibility, we designed primers targeting known p53 binding sites surrounding *BBC3*, *BAX* and *CDKN1A* based on published p53 ChIP-seq data (Andrysik et al., 2017). We performed ChIP against p53 in WT and CTCF<sup>+/-</sup> cells treated with 6μM cisplatin for 1h, 2h, 4h and 8h. We observed increased binding of p53 at 2h and 8h after 6μM cisplatin treatment in CTCF<sup>+/-</sup> cells compared to WT (Fig. 15). Thus, the increased expression of p53 target genes in CTCF<sup>+/-</sup> cells after DNA damage may be driven by the increased binding to p53 to these genes.

### **8.6 Overall chromatin region in CTCF<sup>+/-</sup> cells is more accessible**

CTCF plays an important role in establishing chromatin loops and boundaries and it has been shown that loss of CTCF results in aberrant silencing of tumor suppressor gene p16<sup>INKa</sup> in breast cancer due to deregulation of a chromatin boundary and spreading of repressive histone marks (Witcher and Emerson, 2009). Likewise, single copy loss of CTCF may result in deregulation of chromatin boundaries that lead to spread of activating histone marks that potentially result in a more accessible chromatin region (Narendra et al., 2015). To test this, we performed a micrococcal nuclease (MNase) assay which utilizes the MNase to digest chromatin regions that are not

protected by histones randomly. Increased accessibility of chromatin regions is characterized by increased degree of MNase digestion for a certain duration of treatment (Fig. 16A). MCF10A CTCF<sup>+/-</sup> cells show a higher degree of digestion compared to WT cells at 10mins, 15mins and 30mins after MNase treatment (Fig. 16B and C). Furthermore, we utilized histone mark H3K27ac as a surrogate to represent accessible chromatin regions as H3K27ac is highly associated with active enhancer regions (Creyghton et al., 2010). ChIP-seq against H3K27ac in MCF10A WT and CTCF<sup>+/-</sup> cells revealed that CTCF<sup>+/-</sup> cells have increased number of H3K27ac sites gained, as compared to WT cells (Fig. 16D and E). Taken together, these evidence show that, overall, chromatin regions in CTCF<sup>+/-</sup> cells are more accessible.

### **8.7 Increased H3K27ac at *BBC3* and *BAX* in CTCF<sup>+/-</sup> cells after DNA damage**

We investigated the H3K27ac histone mark at *BBC3* and *BAX* by ChIP-qPCR. While there is no significant differences in H3K27ac at basal level (0h) between WT and CTCF<sup>+/-</sup> cells, we observed a gain in H3K27ac binding after 8h of 6μM cisplatin treatment in CTCF<sup>+/-</sup> cells but not WT cells (Fig. 17). This is likely a result of the increased p53 binding at *BBC3* and *BAX* after DNA damage, which leads to recruitment of transcription factors that promote deposition of H3K27ac to the surrounding chromatin region.

### **8.8 No changes in binding of CTCF surrounding *BBC3* and *BAX* after DNA damage**

CTCF has been reported to act as a *BBC3*-specific repressor in human colon carcinoma (HCT116) cells (Gomes and Espinosa, 2010). These cells were treated with 5-fluorouracil (5-FU), a DNA damaging agent, for 8h and CTCF binding was decreased in the 5' region of *BBC3* which led to the upregulation of *BBC3* gene expression. They also demonstrated that shRNA-mediated CTCF knockdown in HCT116 cells led to increased *BBC3* gene expression but not other p53 target genes such as *GADD45α*, *MDM2* and *CDKN1A*, therefore showing that binding of CTCF to *BBC3* specifically repress *BBC3* transcription. To ascertain the status of CTCF binding in our MCF10A CTCF<sup>+/-</sup> cells after DNA damage, we first designed primers targeting known CTCF binding sites surrounding *BBC3* and *BAX* genes in MCF10A cell line based on a MCF10A ChIP-seq experiment previously done in our laboratory. ChIP was performed against CTCF for cells treated with 6μM cisplatin for 8h. We observed no significant changes in CTCF binding surrounding *BBC3* and *BAX* for both WT and CTCF<sup>+/-</sup> cells before and after treatment (Fig. 18). This suggests that the increased

expression of p53 target genes in MCF10A CTCF<sup>+/-</sup> cells is not contributed by loss of CTCF binding as a gene-specific repressor.

## 9. Figures

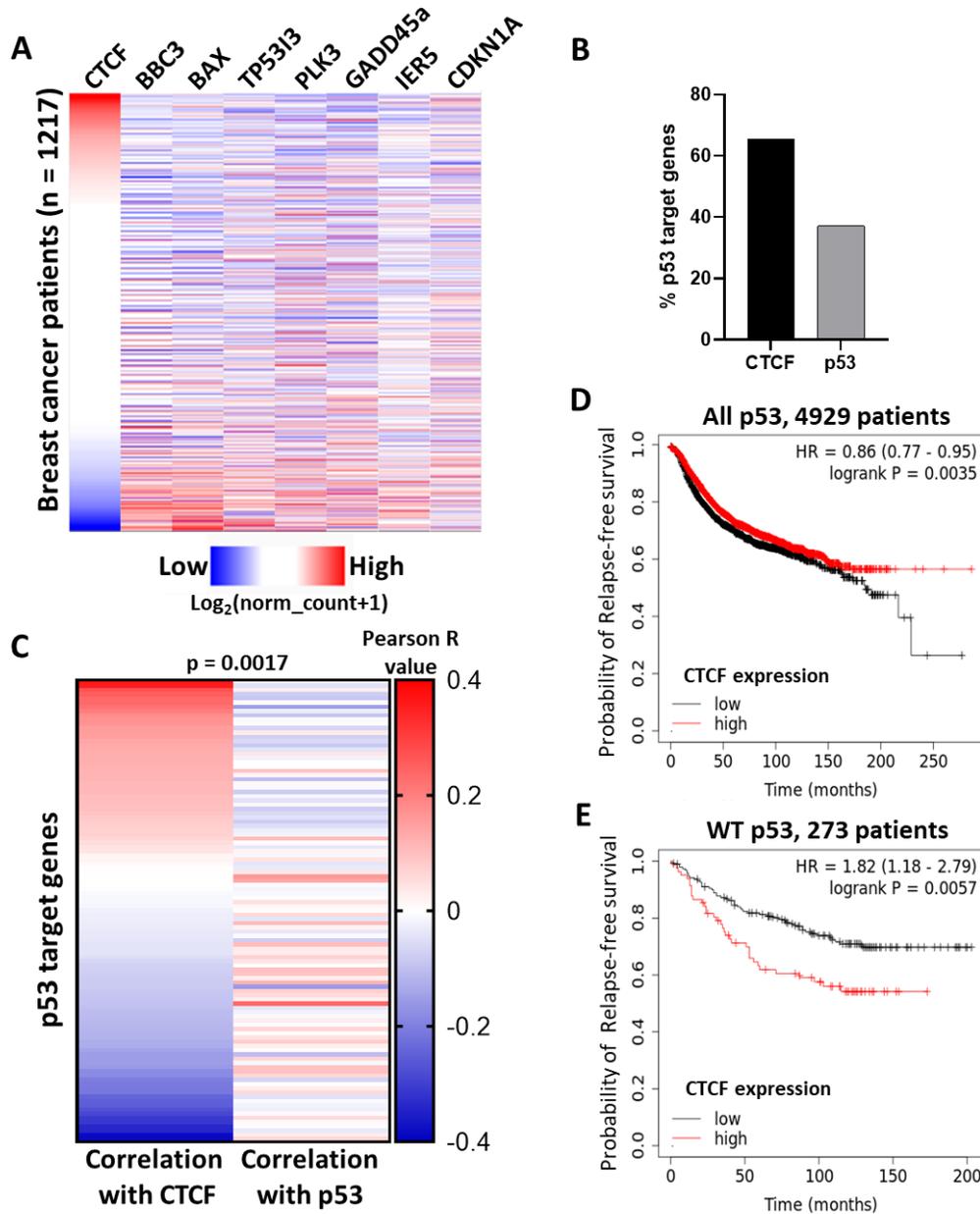


Fig. 10. **CTCF may negatively regulate gene expression of p53 target genes.** A) Gene expression data of 1217 breast tumors obtained from TCGA-BRCA database in UCSC Xena browser (<https://xenabrowser.net/heatmap/>) and comparing CTCF gene expression level to a panel of p53 target genes. B) Percentage of p53 target genes significantly ( $p < 0.05$ ) correlated with CTCF and p53 from a list of 116 p53 target genes by Pearson correlation. C) Correlation of the expression level of each p53 target gene to CTCF and p53 expression levels in 1217 breast tumors. D) Probability of relapse-free survival for 4929 breast tumors with all status of p53 stratified by low CTCF expression (black) and high CTCF expression (red), plotted using Kaplan-Meier plotter. E) Probability of relapse-free survival for 273 breast tumors with wild-type p53 stratified by low CTCF expression (black) and high CTCF expression (red), plotted using Kaplan-Meier plotter.

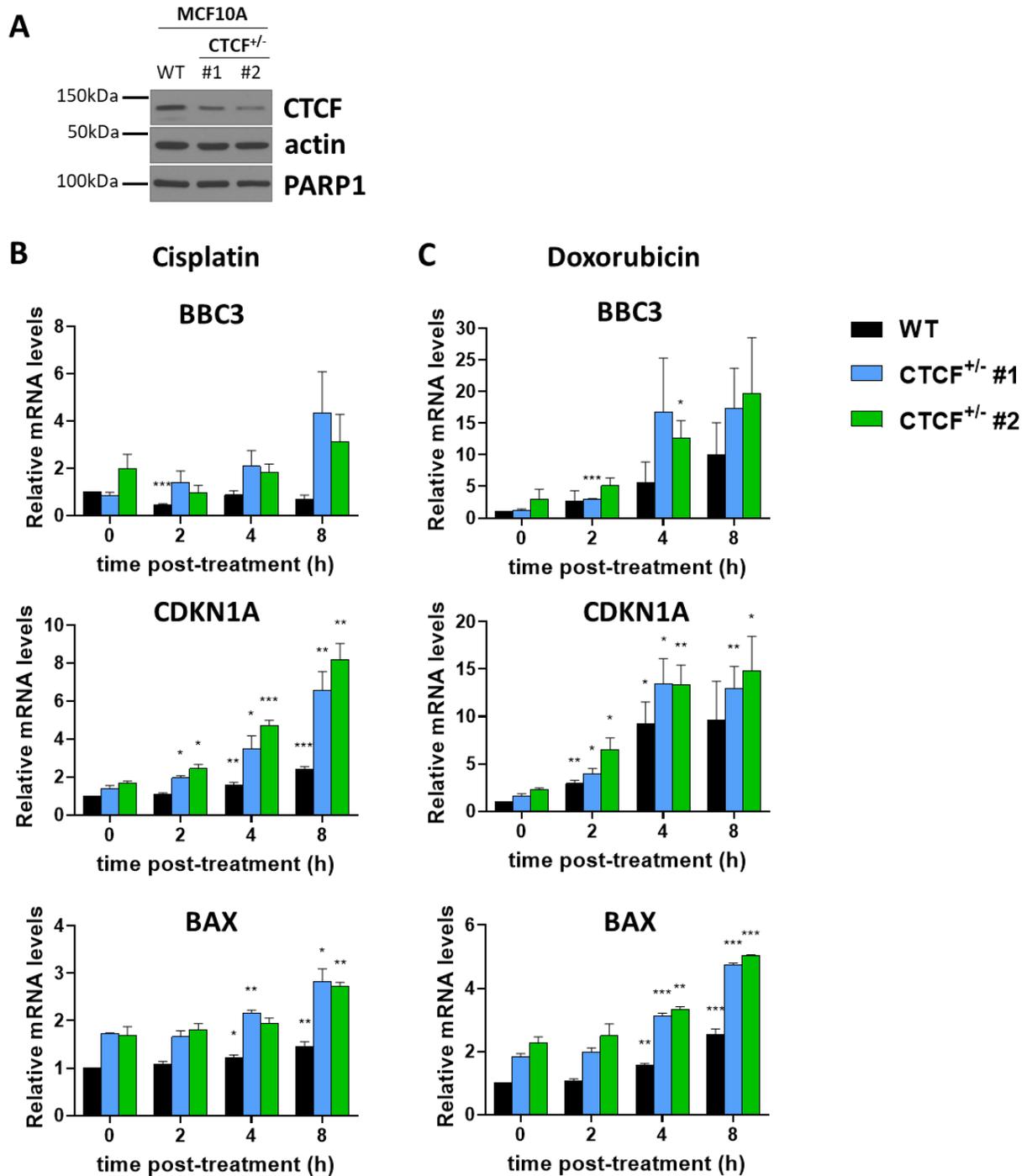
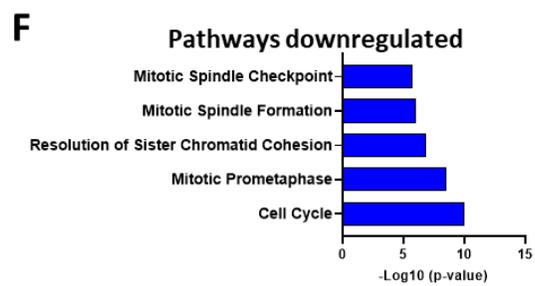
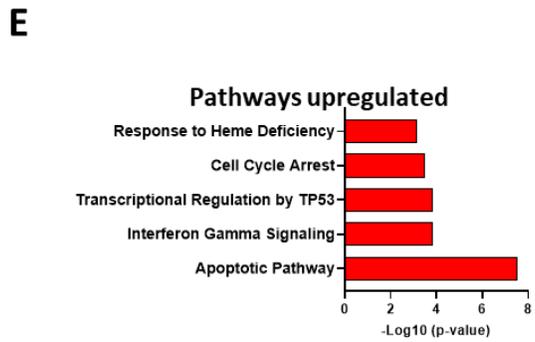
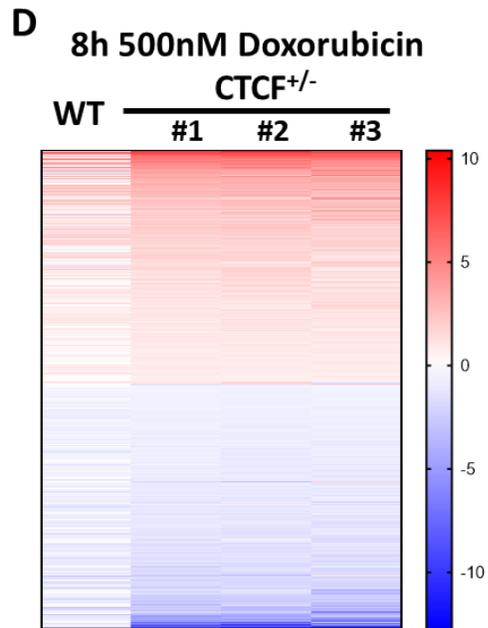
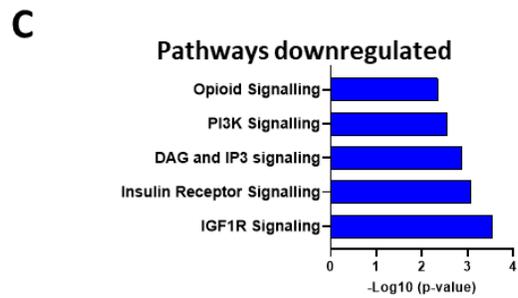
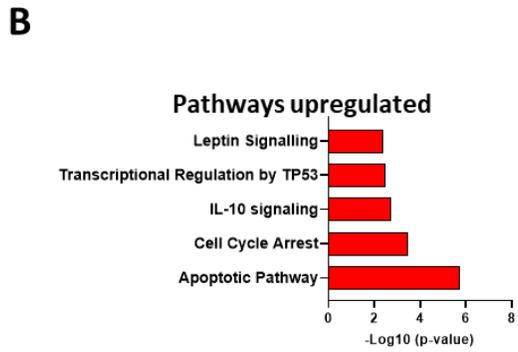
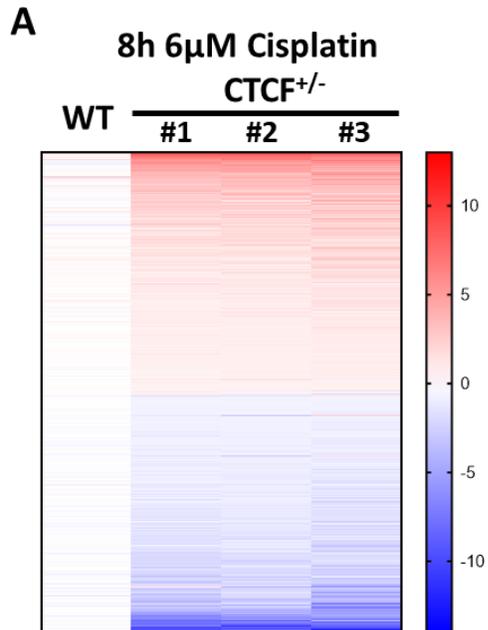
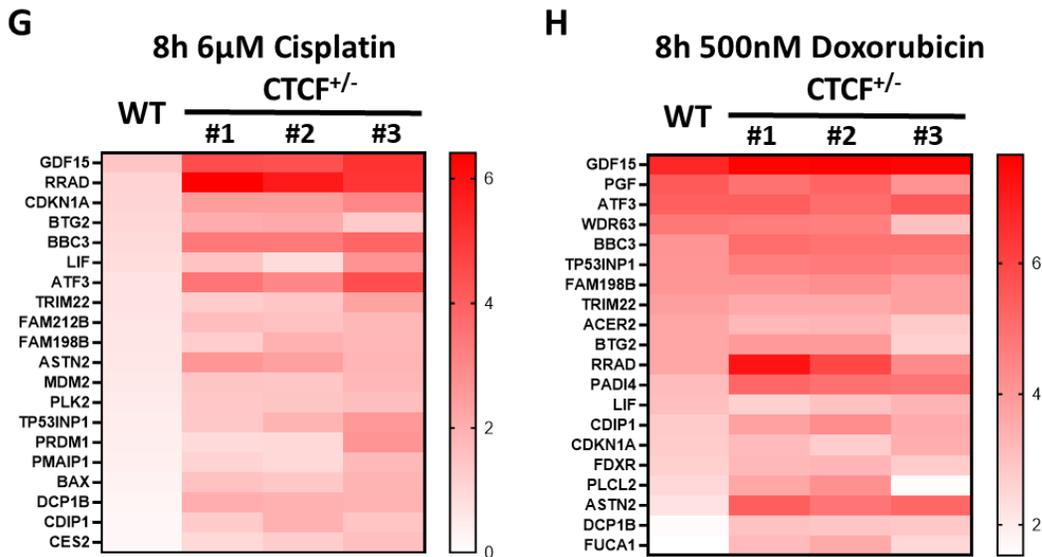
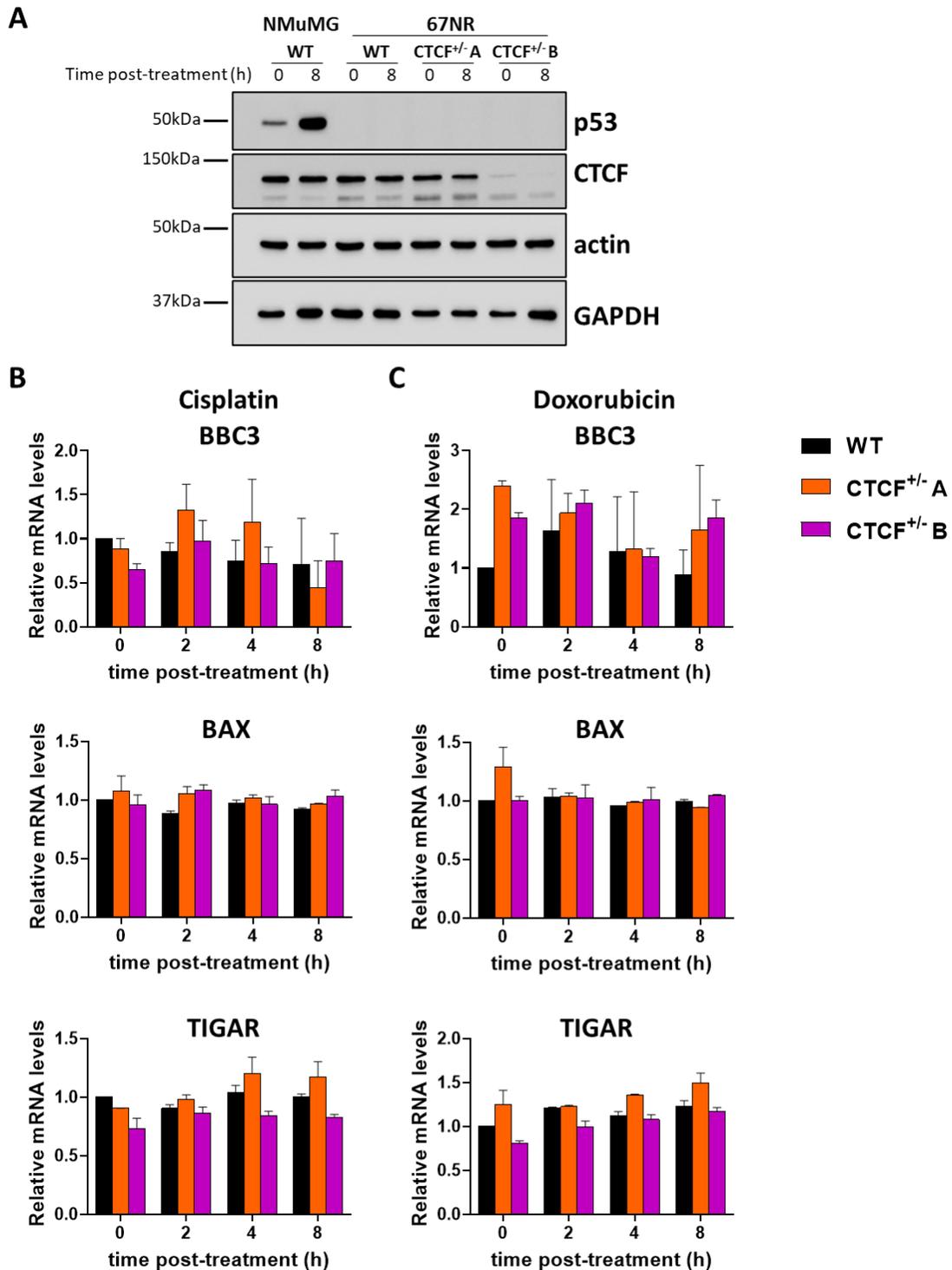


Fig. 11. MCF10A CTCF<sup>+/-</sup> cells show a more robust p53-mediated DNA damage response. A) Western blot against CTCF showing reduced CTCF protein expression in MCF10A CTCF<sup>+/-</sup> #1 and CTCF<sup>+/-</sup> #2 cells. B) Gene expression levels of *BBC3*, *CDKN1A* and *BAX* by RT-qPCR in MCF10A CTCF<sup>+/-</sup> #1 and CTCF<sup>+/-</sup> #2 cells after 6 $\mu$ M cisplatin treatment (n = 3). C) Gene expression levels of *BBC3*, *CDKN1A* and *BAX* by RT-qPCR in MCF10A CTCF<sup>+/-</sup> #1 and CTCF<sup>+/-</sup> #2 cells after 500nM doxorubicin treatment (n = 3). Statistical analysis performed using Student's t-Test against WT 0h. \* p < 0.05, \*\* p < 0.01, \*\*\* p < 0.001.





**Fig. 12. RNA-seq analysis of MCF10A WT and CTCF<sup>+/-</sup> cells after DNA damage.** A) Differential gene expression of WT and CTCF<sup>+/-</sup> cells after 8h treatment with 6µM cisplatin, with scale on the right representing Log<sub>2</sub> fold change. B) Reactome pathway enrichment analysis for genes significantly upregulated after 6µM cisplatin treatment (fold change > 1.5, adjusted p-value < 0.05). C) Reactome pathway enrichment analysis for genes significantly downregulated after 6µM cisplatin treatment (fold change > 1.5, adjusted p-value < 0.05). D) Differential gene expression of WT and CTCF<sup>+/-</sup> cells after 8h treatment with 500nM doxorubicin, with scale on the right representing Log<sub>2</sub> fold change. E) Reactome pathway enrichment analysis for genes significantly upregulated after 500nM doxorubicin treatment (fold change > 1.5, adjusted p-value < 0.05). F) Reactome pathway enrichment analysis for genes significantly downregulated after 500nM doxorubicin treatment (fold change > 1.5, adjusted p-value < 0.05). G) Top 20 p53 target genes significantly upregulated in CTCF<sup>+/-</sup> cells after 8h 6µM cisplatin treatment, with scale on the right representing Log<sub>2</sub> fold change. H) Top 20 p53 target genes significantly upregulated in CTCF<sup>+/-</sup> cells after 8h 500nM doxorubicin treatment, with scale on the right representing Log<sub>2</sub> fold change.



**Fig. 13. Increased expression of p53 target genes in MCF10A CTCF<sup>+/-</sup> cells after DNA damage is p53-dependent.** A) Western blot against p53 showing no p53 protein expression in 67NR WT and CTCF<sup>+/-</sup> cells. B) Gene expression levels of *BBC3*, *BAX* and *TIGAR* by RT-qPCR in 67NR CTCF<sup>+/-</sup> A and CTCF<sup>+/-</sup> B cells after 6 $\mu$ M cisplatin treatment (n = 3). C) Gene expression levels of *BBC3*, *BAX* and *TIGAR* by RT-qPCR in 67NR CTCF<sup>+/-</sup> A and CTCF<sup>+/-</sup> B cells after 500nM doxorubicin treatment (n = 3). Statistical analysis performed using Student's t-Test against WT 0h. \* p < 0.05, \*\* p < 0.01, \*\*\* p < 0.001.

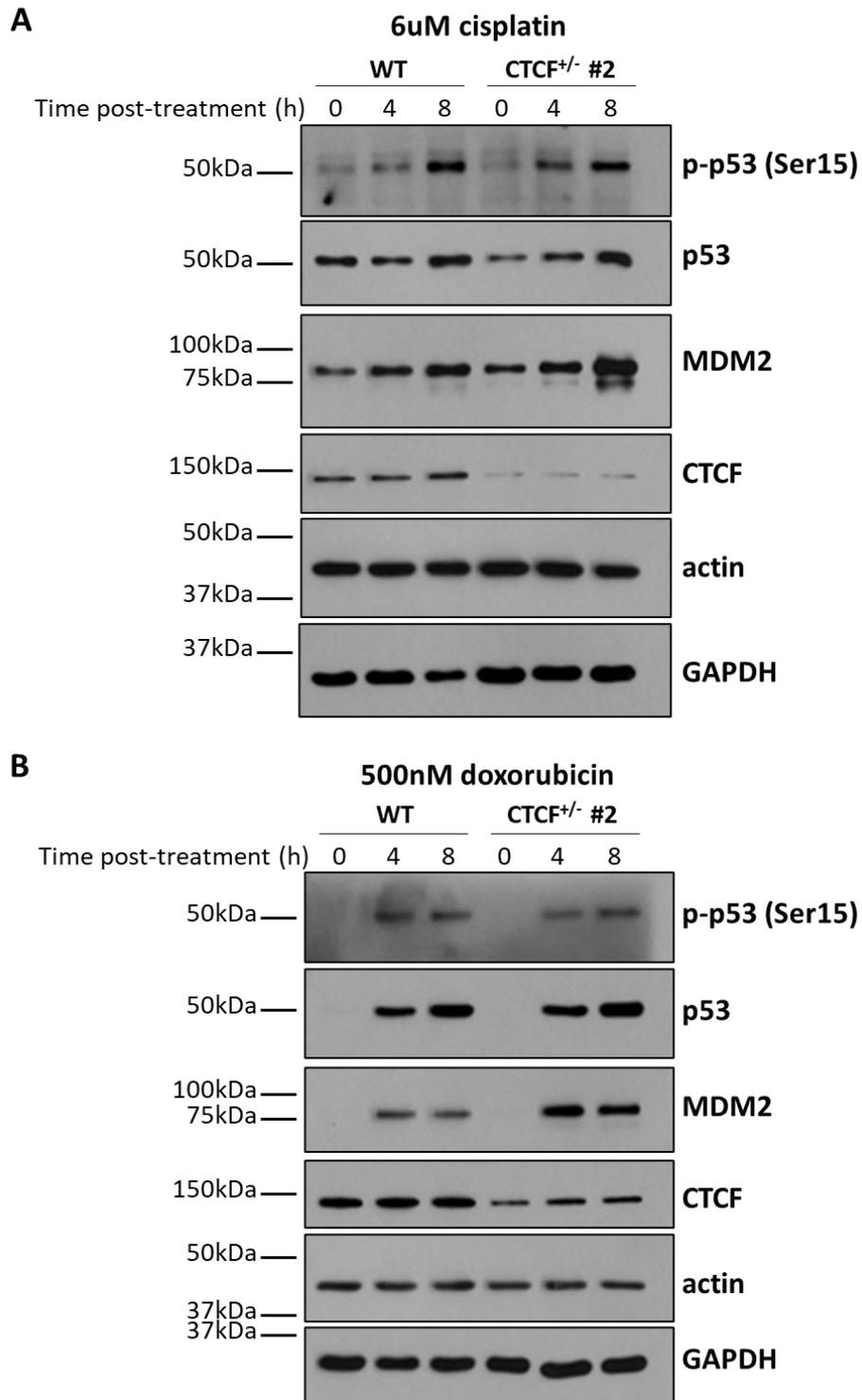


Fig. 14. **Stabilization of p53 and p-p53 after DNA damage is consistent between WT and CTCF<sup>+/-</sup> #2 cells.** Western blot against p53, p-p53 (Ser15) and MDM2 for A) 6 $\mu$ M cisplatin-treated cells and B) 500nM doxorubicin-treated cells.

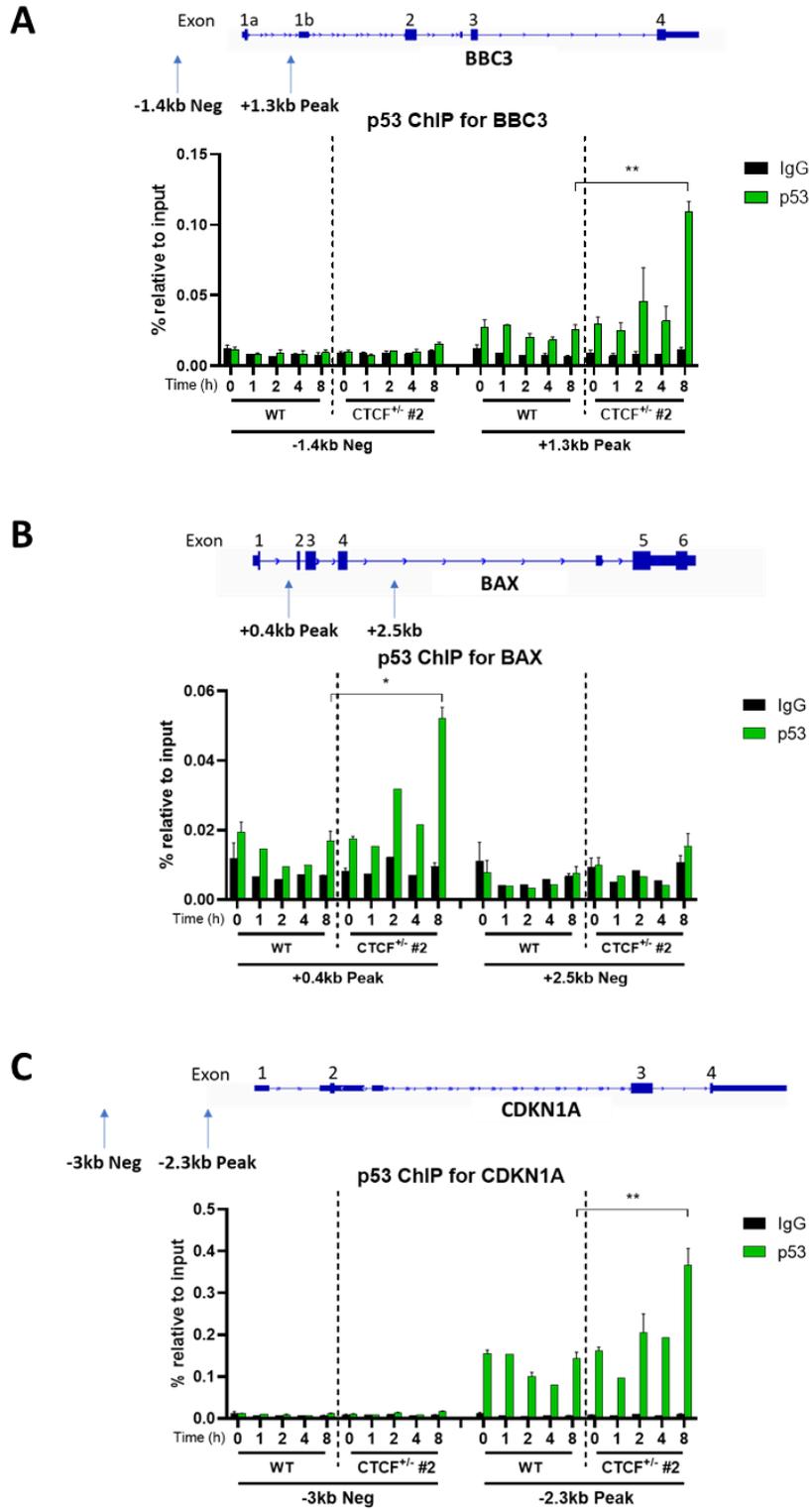


Fig. 15. **Increased p53 binding in MCF10A CTCF<sup>+/-</sup> cells after DNA damage.** ChIP against p53 after 6 $\mu$ M cisplatin treatment at A) *BBC3*, B) *BAX* and C) *CDKN1A*. Statistical analysis performed using Student's t-Test. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

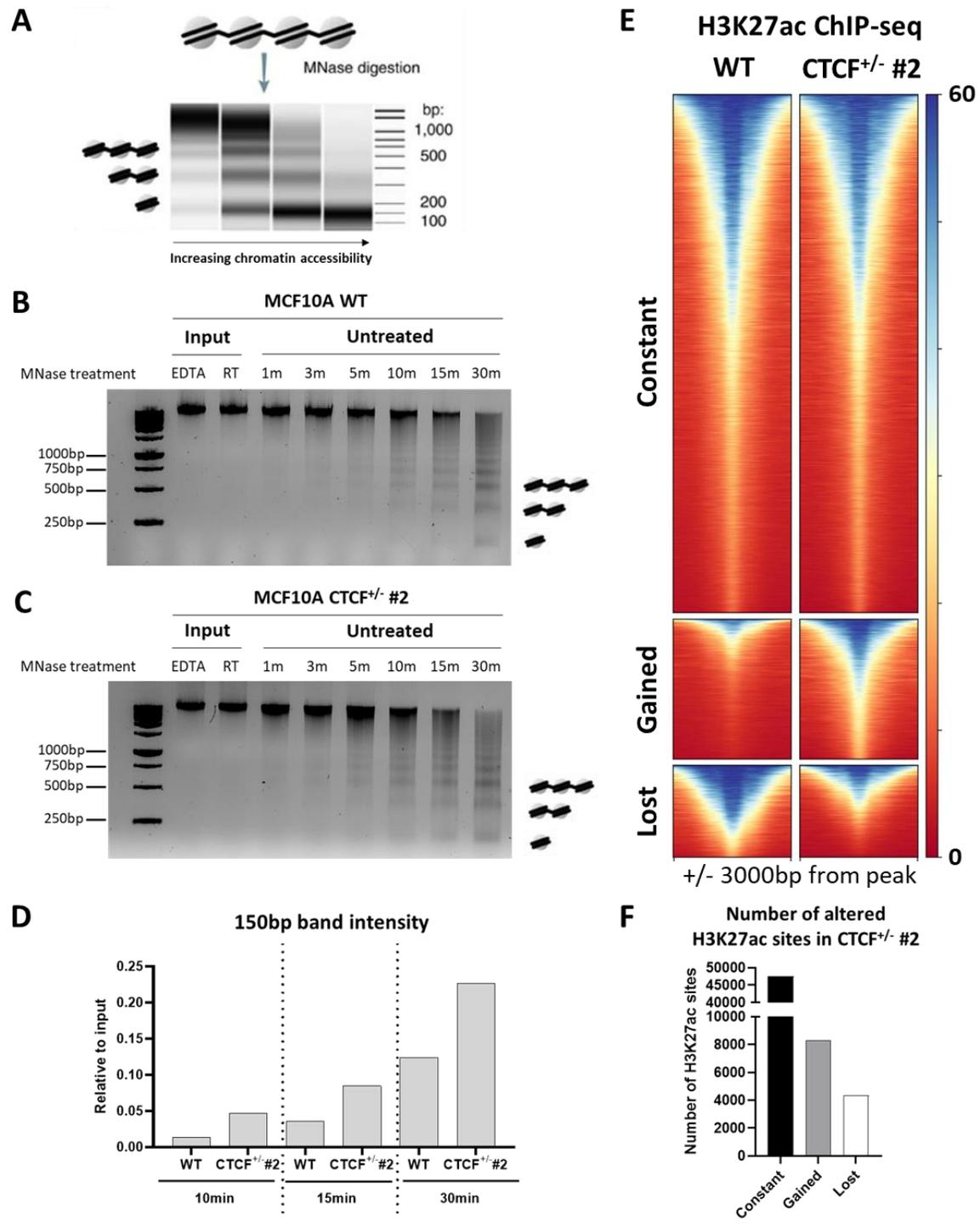


Fig. 16. **Overall chromatin region in CTCF<sup>+/-</sup> cells is more accessible.** A) Schematic of MNase assay. B) Gel electrophoresis of purified genomic DNA from MCF10A WT cells after treatment with 10U MNase. C) Gel electrophoresis of purified genomic DNA from MCF10A CTCF<sup>+/-</sup> #2 cells after treatment with 10U MNase. D) Band intensity quantification of band at 150bp (lowest band) from gel images of A and B at 10min, 15min and 30min using ImageJ. E) ChIP-seq analysis of H3K27ac in MCF10A WT and CTCF<sup>+/-</sup> #2 cells for signal of H3K27ac within +/- 3000bp from identified H3K27ac peaks. F) Quantification of number of constant, gained and lost sites from heatmap in E, for MCF10A CTCF<sup>+/-</sup> #2 cells compared to MCF10A WT.

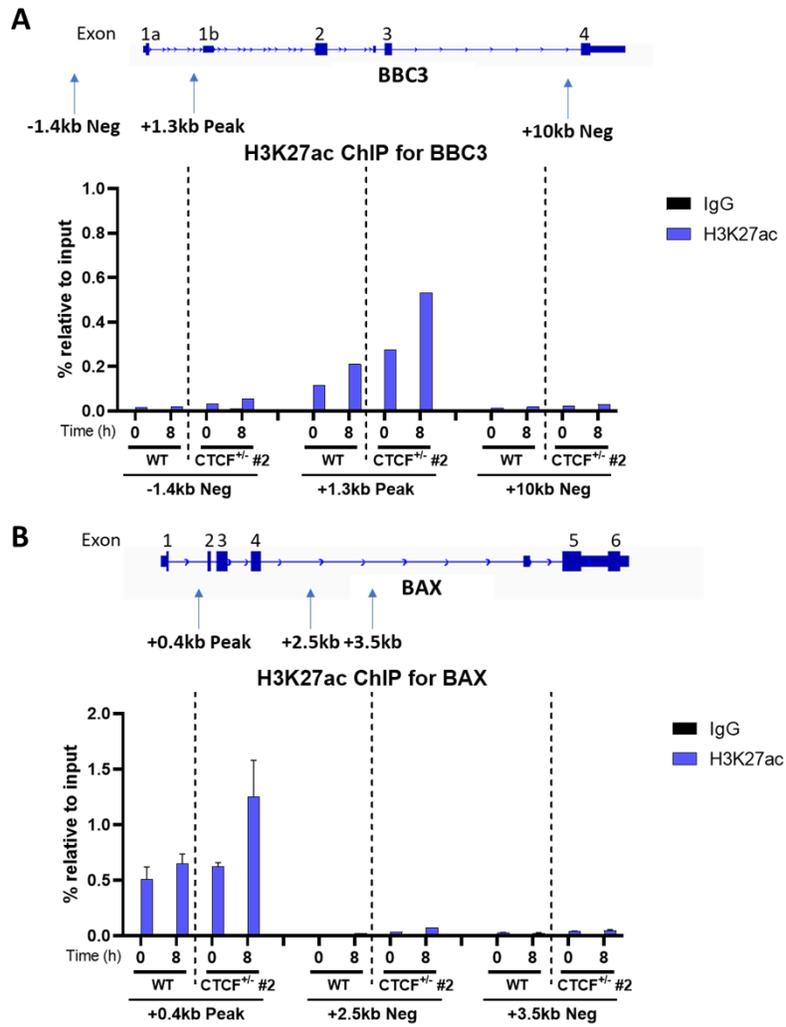


Fig. 17. **Increased H3K27ac at BBC3 and BAX in CTCF<sup>+/-</sup> cells after DNA damage.** ChIP against H3K27ac after 6 $\mu$ M cisplatin treatment in MCF10A WT and CTCF<sup>+/-</sup> #2 cells for A) *BBC3* and B) *BAX*.

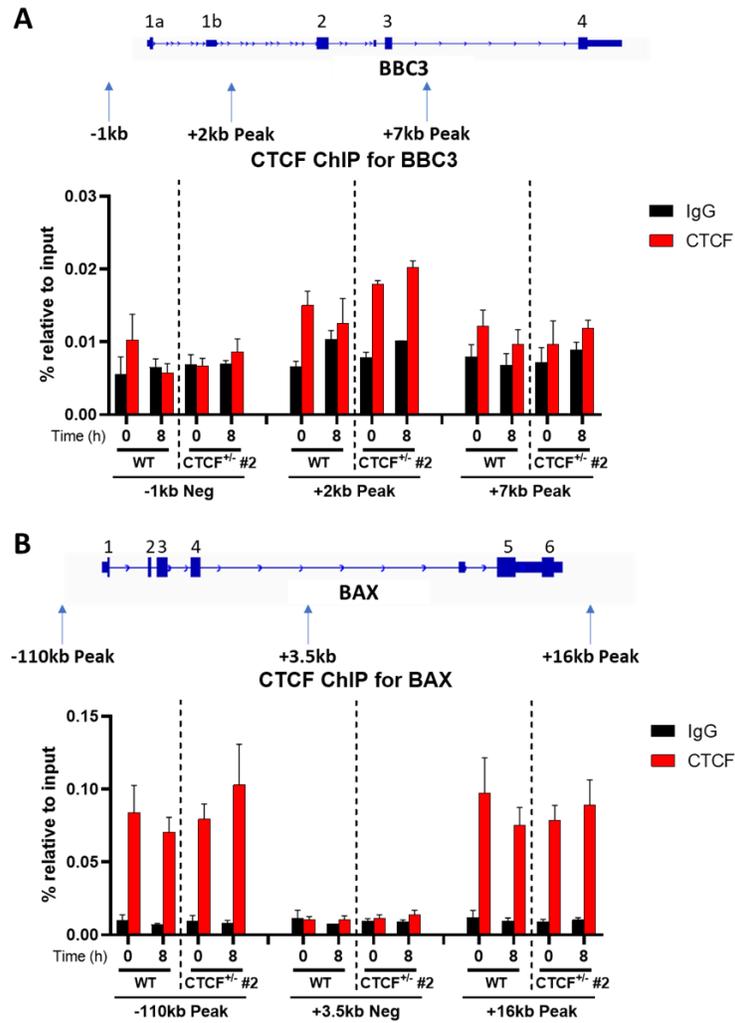


Fig. 18. **No changes in CTCF binding at BBC3 and BAX after DNA damage.** ChIP against CTCF after 6 $\mu$ M cisplatin treatment in MCF10A WT and CTCF<sup>+/-</sup> #2 cells for A) *BBC3* and B) *BAX*.

## 10. Discussion

In this thesis, we show that single copy loss of CTCF led to increased expression of p53 target genes upon DNA damage, compared to WT cells. At least at 3 known p53 target genes (*BBC3*, *BAX* and *CDK1NA*), the increased p53 target gene expression is due to increased binding of p53 at these genes after DNA damage in CTCF<sup>+/-</sup> cells while amount of p53 stabilized at the protein level remain consistent between WT and CTCF<sup>+/-</sup> cells. We also show that overall chromatin regions in CTCF<sup>+/-</sup> cells are more accessible, possibly increasing accessibility of p53 to p53 binding sites, resulting in increased binding of p53 upon DNA damage. We propose that the single copy loss of CTCF results in deregulation of chromatin loops, which may lead to aberrant spreading of activating histone marks that increases the accessibility of the chromatin region near p53 target genes, resulting in increased expression of p53 target genes after DNA damage (Fig. 19). Indeed, it has been shown that siRNA-mediated depletion of CTCF in prostate cancer cells led to merging of chromatin loops and formation of new chromatin boundaries due to loss of CTCF binding (Khoury et al., 2020).

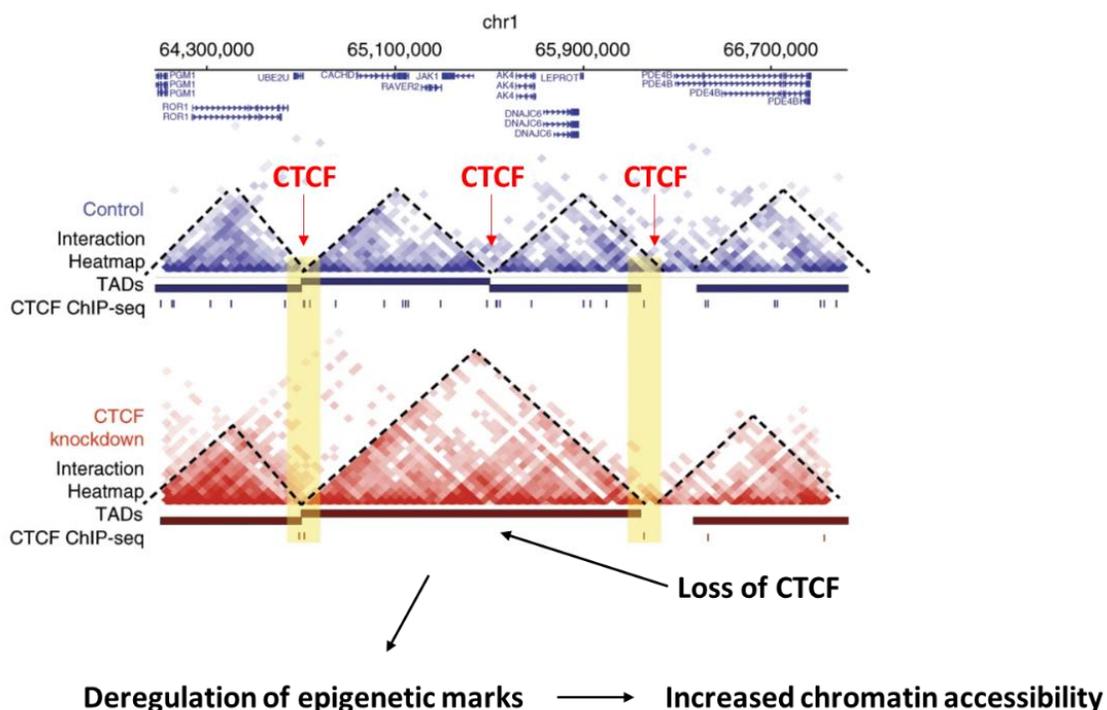


Fig. 19. Schematic of working model. Loss of CTCF leads to deregulation of chromatin boundaries and merging of chromatin loops which potentially increases chromatin accessibility due to spreading of activating epigenetic marks within new chromatin loop (Khoury et al., 2020).

Our RNA-seq data showed two distinct profiles of global differential gene expression for MCF10A WT and CTCF<sup>+/-</sup> cells after 8h of treatment with 6μM cisplatin or 500nM doxorubicin. MCF10A WT cells treated with 6μM cisplatin show minimal changes in gene expression while treatment with 500nM doxorubicin resulted in some observable extent of changes in gene expression. This can likely be attributed to the difference in mechanism of action between cisplatin and doxorubicin in inducing DNA damage. Cisplatin treatment results in DNA adducts that consist of 98% intrastrand crosslinks involving purine nucleotides and 2% interstrand crosslinks which involves crosslinking of different strands of DNA (Kartalou and Essigmann, 2001). Intrastrand crosslinks are repaired by nucleotide excision repair (NER) pathway that is active throughout the cell cycle, while interstrand crosslinks are potentially lethal as they stall replication forks and lead to DNA DSB induction (Enoiu et al., 2012). Since our experiment timepoints are only up to 8h after cisplatin treatment, it is unlikely that significant amount of DNA DSBs are generated, as interstrand crosslinks are generated in low frequencies and MCF10A cell line has a doubling time of approximately 20h (Bessette et al., 2015). Therefore, it is unlikely that majority of the cells have undergone replication fork stalling and DNA DSB generation during our 8h cisplatin treatment. This explains why we see minimal changes in gene expression in our MCF10A WT cell line after 8h cisplatin treatment. This observation is also supported by our western blot results, as the magnitude p53 and p-p53 protein stabilization for cisplatin-treated cells were not as comparable to doxorubicin-treated cells and required longer exposure times to obtain visible bands.

Doxorubicin promotes DNA DSB formation by intercalation into DNA to prevent DNA and RNA synthesis (Momparler et al., 1976), and inhibition of DNA topoisomerase II-mediated DNA repair (Tewey et al., 1984) after the initial DNA cleavage during alleviation of DNA supercoiling. It has also been reported that doxorubicin generates reactive oxygen species (ROS) in the presence of NADPH and oxygen which results in DNA damage (Gutteridge and Quinlan, 1985; Sinha et al., 1989). As DNA DSBs are lethal insults to the cell, it must be repaired rapidly and thus, a greater magnitude of p53-mediated DNA damage response is not surprising. This is supported by the larger extent of changes in gene expression in MCF10A WT cells after 8h of doxorubicin treatment.

We have showed that the higher p53 target gene expression in CTCF<sup>+/-</sup> cells is dependent on p53, in 67NR cells. While 67NR cells are of mice origin which has high similarity to human DNA

sequence, the results may be affected by species-specific differences due to differences in transcriptional regulation of genes (Lin et al., 2014). It is therefore ideal to verify the dependence of the higher p53 target gene expression in CTCF<sup>+/-</sup> cells to p53 in human cell lines. A shRNA-mediated knockdown of p53 can be carried out in the MCF10A CTCF<sup>+/-</sup> cells to deplete p53 levels and followed by exposure to cisplatin or doxorubicin to evaluate the expression of p53 target genes.

In our ChIP against p53 data, we did not observe increase in p53 binding at *BBC3*, *BAX* and *CDKN1A* in our WT cells after cisplatin treatment. However, based on our RT-qPCR and western blot data, we know that MCF10A WT cells do undergo p53-mediated DNA damage response upon treatment of cisplatin at the dosage and duration used in our experiments. We observed an increase in gene expression of p53 target genes by RT-qPCR and increased stabilization of p53 and p-p53 at the protein level, suggesting robust p53-mediated DNA damage response. p53 has been shown to bind the *CDKN1A* locus at basal level (Espinosa and Emerson, 2001; Younger and Rinn, 2017). Upon DNA damage, the transcriptional upregulation of *CDKN1A* was observed. While no changes in p53 binding was observed, they showed that basally bound p53 was phosphorylated in-situ at the *CDKN1A* locus, as observed from ChIP against p-p53. Thus, phosphorylation of p53 can occur directly on basally bound p53 at *CDKN1A*, leading to transcriptional activation and gene expression, without change in p53 binding levels. It is therefore likely that due to the low magnitude of p53-mediated DNA damage response induced by cisplatin treatment in our cells, activation of p53 by direct phosphorylation of p53 at basally bound p53 is sufficient to induce a required p53-mediated DNA damage response.

We demonstrated an increased p53 binding at three known p53 target genes (*BBC3*, *BAX* and *CDKN1A*) in our CTCF<sup>+/-</sup> cells after cisplatin treatment compared to WT. It will be interesting to investigate if p53 binding is increased at other p53 target genes in our CTCF<sup>+/-</sup> cells. A ChIP-seq against p53 should be performed to look at p53 binding genome-wide after cisplatin treatment in CTCF<sup>+/-</sup> cells. p53 ChIP-seq data allows us to determine p53 binding levels at all p53 target genes after cisplatin and determine if the increase in p53 binding correlate with p53 target genes that are significantly upregulated in our RNA-seq data (Fig. 12G and H). A positive correlation will suggest that the increased p53 target gene expression is driven by increased p53 binding at these genes in our CTCF<sup>+/-</sup> cells. Besides that, ChIP-seq against RNA pol II may provide further support to the effect of increased p53 binding at p53 target genes on gene expression. The transactivation

domains of p53 has been shown to interact with subunits of RNA pol II and binding of p53 to RNA pol II leads to a conformational change in RNA pol II, which may contribute to assembly of transcriptional machinery (Liou et al., 2021). Therefore, if ChIP-seq data against RNA pol II correlates with ChIP-seq data against p53, i.e. regions with increased p53 binding also contain increased RNA pol II binding, it is highly likely that the increased p53 binding is driving increased p53 target gene expression.

We think that the increased binding of p53 at p53 target genes after DNA damage in our CTCF<sup>+/-</sup> cells may be due to increased accessibility of chromatin regions. Our data from MNase treatment and H3K27ac ChIP-seq experiments highly suggest that the overall chromatin regions in the CTCF<sup>+/-</sup> cells are more accessible. However, we lack a direct link to p53 target genes. Therefore, we can potentially map our MNase results directly to p53 target genes by looking for enrichment of p53 binding sites at p53 target genes in the 150bp band representing single nucleosome unit by RT-qPCR. By detecting level of p53 binding sites in the 150bp band in WT and CTCF<sup>+/-</sup> cells, we expect to observe enrichment of p53 binding sites if the chromatin regions surrounding p53 target genes are more accessible, which allows more rapid digestion by MNase.

Our ChIP against CTCF at *BBC3* showed no changes in CTCF before and after cisplatin treatment. As CTCF has been shown to be a *BBC3*-specific repressor in colon carcinoma cells, our ChIP data showed that this is not the case in our MCF10A cells (Gomes and Espinosa, 2010). In fact, our observations are rather different. They showed that knockdown of CTCF resulted in upregulation of *BBC3* expression, but not other p53 target genes, in contrast to our observation. This can be explained with the role of CTCF in formation of chromatin loops at the TAD and subTAD levels of chromatin organization. Boundaries of TADs are rather conserved between different cell types, while boundaries between subTADs are more dynamic between different cell types, as seen in the transition of fetal to adult hematopoietic stem cells (Chen et al., 2019). On top of that, we have evidence in our lab showing that in our CTCF<sup>+/-</sup> cells, loss of CTCF levels led to a more prominent loss of CTCF binding at the subTADs, compared to TADs (Lebeau et al., submitted for review). Therefore, loss of CTCF in colon carcinoma (HCT116) cells and breast epithelial (MCF10A) cells may result in deregulation of different chromatin boundaries, leading to a cell-type specific phenotype. Therefore, the phenotype observed in our MCF10A cells may be specific to cells from breast tissue. Further experiments should be done to verify this, possibly by generating CTCF<sup>+/-</sup>

by CRISPR editing in different cell types to evaluate if the increased p53 target genes expression compared to WT cells after DNA damage can be recapitulated. Furthermore, as CTCF-mediated chromatin boundaries at subTADs are dynamic between different cell types and subTADs can span up to several hundred thousand basepairs, ChIP-seq against CTCF may provide insights to the potential subTADs that are deregulated due to loss of CTCF in our MCF10A CTCF<sup>+/-</sup> cells. Additionally, we may correlate changes in gene expression from our RNA-seq data to CTCF ChIP-seq data, to see if changes in gene expression correlate with changes in boundaries of subTADs.

## 11. Concluding remarks

We showed that single copy loss of CTCF in MCF10A cells led to increased p53 target gene expression upon treatment with DNA damage inducing agents, possibly driven by increase p53 binding at p53 target genes. Single copy loss of CTCF led to an overall chromatin region that is more accessible, potentially allowing p53 to readily access and bind p53 target genes to enhance transcriptional activity. As single copy loss of CTCF is a frequent event in breast cancer, tumors with functional p53-mediated DNA damage response pathway may show hypersensitivity to DNA damage. Therefore, CTCF may potentially be a suitable biomarker to predict efficacy of treatment with chemotherapeutic agents that induce DNA damage as their mechanism of action. Further studies are required *in vivo* to test this hypothesis and additional mechanistic studies, as well, to determine the role of CTCF in regulating p53-mediated DNA damage response.

## 12. References

- Andrysiak, Z., Galbraith, M.D., Guarnieri, A.L., Zaccara, S., Sullivan, K.D., Pandey, A., MacBeth, M., Inga, A., and Espinosa, J.M. (2017). Identification of a core TP53 transcriptional program with highly distributed tumor suppressive activity. *Genome Res* 27, 1645-1657.
- Baker, S.J., Fearon, E.R., Nigro, J.M., Hamilton, S.R., Preisinger, A.C., Jessup, J.M., vanTuinen, P., Ledbetter, D.H., Barker, D.F., Nakamura, Y., *et al.* (1989). Chromosome 17 deletions and p53 gene mutations in colorectal carcinomas. *Science* 244, 217-221.
- Bedoui, S., Herold, M.J., and Strasser, A. (2020). Emerging connectivity of programmed cell death pathways and its physiological implications. *Nature Reviews Molecular Cell Biology* 21, 678-695.
- Bell, A.C., West, A.G., and Felsenfeld, G. (1999). The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* 98, 387-396.
- Bessette, D.C., Tilch, E., Seidens, T., Quinn, M.C., Wiegmanns, A.P., Shi, W., Cocciardi, S., McCart-Reed, A., Saunus, J.M., Simpson, P.T., *et al.* (2015). Using the MCF10A/MCF10CA1a Breast Cancer Progression Cell Line Model to Investigate the Effect of Active, Mutant Forms of EGFR in Breast Cancer Development and Treatment Using Gefitinib. *PLoS One* 10, e0125232.
- Bieging, K.T., Mello, S.S., and Attardi, L.D. (2014). Unravelling mechanisms of p53-mediated tumour suppression. *Nat Rev Cancer* 14, 359-370.
- Casasent, A.K., Schalck, A., Gao, R., Sei, E., Long, A., Pangburn, W., Casasent, T., Meric-Bernstam, F., Edgerton, M.E., and Navin, N.E. (2018). Multiclonal Invasion in Breast Tumors Identified by Topographic Single Cell Sequencing. *Cell* 172, 205-217.e212.
- Cesca, M.G., Vian, L., Cristóvão-Ferreira, S., Pondé, N., and de Azambuja, E. (2020). HER2-positive advanced breast cancer treatment in 2020. *Cancer Treat Rev* 88, 102033.
- Chen, C., Yu, W., Tober, J., Gao, P., He, B., Lee, K., Trieu, T., Blobel, G.A., Speck, N.A., and Tan, K. (2019). Spatial Genome Re-organization between Fetal and Adult Hematopoietic Stem Cells. *Cell Rep* 29, 4200-4211.e4207.
- Chen, L., Liu, S., and Tao, Y. (2020). Regulating tumor suppressor genes: post-translational modifications. *Signal Transduct Target Ther* 5, 90.
- Chernukhin, I., Shamsuddin, S., Kang, S.Y., Bergström, R., Kwon, Y.W., Yu, W., Whitehead, J., Mukhopadhyay, R., Docquier, F., Farrar, D., *et al.* (2007). CTCF interacts with and recruits the largest subunit of RNA polymerase II to CTCF target sites genome-wide. *Mol Cell Biol* 27, 1631-1648.
- Chipuk, J.E., Kuwana, T., Bouchier-Hayes, L., Droin, N.M., Newmeyer, D.D., Schuler, M., and Green, D.R. (2004). Direct activation of Bax by p53 mediates mitochondrial membrane permeabilization and apoptosis. *Science* 303, 1010-1014.
- Cleton-Jansen, A.M., Callen, D.F., Seshadri, R., Goldup, S., McCallum, B., Crawford, J., Powell, J.A., Settasatian, C., van Beerendonk, H., Moerland, E.W., *et al.* (2001). Loss of heterozygosity mapping at chromosome arm 16q in 712 breast tumors reveals factors that influence delineation of candidate regions. *Cancer Res* 61, 1171-1177.
- Creighton, C.J. (2012). The molecular profile of luminal B breast cancer. *Biologics* 6, 289-297.
- Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., *et al.* (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proceedings of the National Academy of Sciences* 107, 21931-21936.

Cuddapah, S., Jothi, R., Schones, D.E., Roh, T.Y., Cui, K., and 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.

Davidson, I.F., and Peters, J.-M. (2021). Genome folding through loop extrusion by SMC complexes. *Nature Reviews Molecular Cell Biology* 22, 445-464.

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

Deng, C., Zhang, P., Harper, J.W., Elledge, S.J., and Leder, P. (1995). Mice lacking p21CIP1/WAF1 undergo normal development, but are defective in G1 checkpoint control. *Cell* 82, 675-684.

Di Leonardo, A., Linke, S.P., Clarkin, K., and Wahl, G.M. (1994). DNA damage triggers a prolonged p53-dependent G1 arrest and long-term induction of Cip1 in normal human fibroblasts. *Genes Dev* 8, 2540-2551.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485, 376-380.

Dumaz, N., and Meek, D.W. (1999). Serine15 phosphorylation stimulates p53 transactivation but does not directly influence interaction with HDM2. *Embo j* 18, 7002-7010.

Dumaz, N., Milne, D.M., Jardine, L.J., and Meek, D.W. (2001). Critical roles for the serine 20, but not the serine 15, phosphorylation site and for the polyproline domain in regulating p53 turnover. *Biochem J* 359, 459-464.

el-Deiry, W.S., Tokino, T., Velculescu, V.E., Levy, D.B., Parsons, R., Trent, J.M., Lin, D., Mercer, W.E., Kinzler, K.W., and Vogelstein, B. (1993). WAF1, a potential mediator of p53 tumor suppression. *Cell* 75, 817-825.

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

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

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

Filippova, G.N., Fagerlie, S., Klenova, E.M., Myers, C., Dehner, Y., Goodwin, G., Neiman, P.E., Collins, S.J., and Lobanekov, V.V. (1996). An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-myc oncogenes. *Mol Cell Biol* 16, 2802-2813.

Filippova, G.N., Lindblom, A., Meincke, L.J., Klenova, E.M., Neiman, P.E., Collins, S.J., Doggett, N.A., and Lobanekov, V.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, 26-36.

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

Fitch, M.E., Cross, I.V., and Ford, J.M. (2003). p53 responsive nucleotide excision repair gene products p48 and XPC, but not p53, localize to sites of UV-irradiation-induced DNA damage, in vivo. *Carcinogenesis* 24, 843-850.

Ford, J.M., and Hanawalt, P.C. (1997). Expression of wild-type p53 is required for efficient global genomic nucleotide excision repair in UV-irradiated human fibroblasts. *J Biol Chem* 272, 28073-28080.

Friedman, P.N., Chen, X., Bargonetti, J., and 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.

Geyer, P.K., and Corces, V.G. (1992). DNA position-specific repression of transcription by a *Drosophila* zinc finger protein. *Genes Dev* 6, 1865-1873.

Goldman, M.J., Craft, B., Hastie, M., Repečka, K., McDade, F., Kamath, A., Banerjee, A., Luo, Y., Rogers, D., Brooks, A.N., *et al.* (2020). Visualizing and interpreting cancer genomics data via the Xena platform. *Nature Biotechnology* 38, 675-678.

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

Gutteridge, J.M., and Quinlan, G.J. (1985). Free radical damage to deoxyribose by anthracycline, aureolic acid and aminoquinone antitumour antibiotics. An essential requirement for iron, semiquinones and hydrogen peroxide. *Biochem Pharmacol* 34, 4099-4103.

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

Harper, J.W., Adami, G.R., Wei, N., Keyomarsi, K., and Elledge, S.J. (1993). The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. *Cell* 75, 805-816.

Harvey, M., McArthur, M.J., Montgomery, C.A., Jr., Butel, J.S., Bradley, A., and Donehower, L.A. (1993). Spontaneous and carcinogen-induced tumorigenesis in p53-deficient mice. *Nat Genet* 5, 225-229.

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

Hinds, P., Finlay, C., and Levine, A.J. (1989). Mutation is required to activate the p53 gene for cooperation with the ras oncogene and transformation. *J Virol* 63, 739-746.

Honda, R., Tanaka, H., and Yasuda, H. (1997). Oncoprotein MDM2 is a ubiquitin ligase E3 for tumor suppressor p53. *FEBS Lett* 420, 25-27.

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

Jackson, M.W., and Berberich, S.J. (2000). MdmX protects p53 from Mdm2-mediated degradation. *Mol Cell Biol* 20, 1001-1007.

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

Johnstone, C.N., Smith, Y.E., Cao, Y., Burrows, A.D., Cross, R.S., Ling, X., Redvers, R.P., Doherty, J.P., Eckhardt, B.L., Natoli, A.L., *et al.* (2015). Functional and molecular characterisation of EO771.LMB tumours, a new C57BL/6-mouse-derived model of spontaneously metastatic mammary cancer. *Dis Model Mech* 8, 237-251.

Jones, S.N., Roe, A.E., Donehower, L.A., and Bradley, A. (1995). Rescue of embryonic lethality in Mdm2-deficient mice by absence of p53. *Nature* 378, 206-208.

Jörnvall, H., Luka, J., Klein, G., and 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.

Kandoth, C., McLellan, M.D., Vandin, F., Ye, K., Niu, B., Lu, C., Xie, M., Zhang, Q., McMichael, J.F., Wyczalkowski, M.A., *et al.* (2013). Mutational landscape and significance across 12 major cancer types. *Nature* 502, 333-339.

Kang, M.A., and Lee, J.S. (2021). A Newly Assigned Role of CTCF in Cellular Response to Broken DNAs. *Biomolecules* 11.

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

Kemp, C.J., Moore, J.M., Moser, R., Bernard, B., Teater, M., Smith, L.E., Rabaia, N.A., Gurley, K.E., Guinney, J., Busch, S.E., *et al.* (2014). CTCF haploinsufficiency destabilizes DNA methylation and predisposes to cancer. *Cell Rep* 7, 1020-1029.

Khoury, A., Achinger-Kawecka, J., Bert, S.A., Smith, G.C., French, H.J., Luu, P.-L., Peters, T.J., 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, S., Yu, N.K., and Kaang, B.K. (2015). CTCF as a multifunctional protein in genome regulation and gene expression. *Exp Mol Med* 47, e166.

Klenova, E.M., Nicolas, R.H., Paterson, H.F., Carne, A.F., Heath, C.M., Goodwin, G.H., Neiman, P.E., and Lobanenkova, V.V. (1993). CTCF, a conserved nuclear factor required for optimal transcriptional activity of the chicken c-myc gene, is an 11-Zn-finger protein differentially expressed in multiple forms. *Mol Cell Biol* 13, 7612-7624.

Kortlever, R.M., Higgins, P.J., and Bernards, R. (2006). Plasminogen activator inhibitor-1 is a critical downstream target of p53 in the induction of replicative senescence. *Nat Cell Biol* 8, 877-884.

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

Kubbutat, M.H., Jones, S.N., and Vousden, K.H. (1997). Regulation of p53 stability by Mdm2. *Nature* 387, 299-303.

Lane, D.P. (1992). Cancer. p53, guardian of the genome. *Nature* 358, 15-16.

Lane, D.P., and Crawford, L.V. (1979). T antigen is bound to a host protein in SV40-transformed cells. *Nature* 278, 261-263.

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

Levine, A.J., Momand, J., and Finlay, C.A. (1991). The p53 tumour suppressor gene. *Nature* 351, 453-456.

Li, F.P., Fraumeni, J.F., Jr., Mulvihill, J.J., Blattner, W.A., Dreyfus, M.G., Tucker, M.A., and Miller, R.W. (1988). A cancer family syndrome in twenty-four kindreds. *Cancer Res* 48, 5358-5362.

Li, Y., Haarhuis, J.H.I., Sedeño Cacciatore, Á., Oldenkamp, R., van Ruiten, M.S., Willems, L., Teunissen, H., Muir, K.W., de Wit, E., Rowland, B.D., *et al.* (2020). The structural basis for cohesin–CTCF-anchored loops. *Nature* 578, 472-476.

Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., *et al.* (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326, 289-293.

Lin, S., Lin, Y., Nery, J.R., Urich, M.A., Breschi, A., Davis, C.A., Dobin, A., Zaleski, C., Beer, M.A., Chapman, W.C., *et al.* (2014). Comparison of the transcriptional landscapes between human and mouse tissues. *Proceedings of the National Academy of Sciences* *111*, 17224-17229.

Linzer, D.I., and Levine, A.J. (1979). Characterization of a 54K dalton cellular SV40 tumor antigen present in SV40-transformed cells and uninfected embryonal carcinoma cells. *Cell* *17*, 43-52.

Liou, S.-H., Singh, S.K., Singer, R.H., Coleman, R.A., and Liu, W.-L. (2021). Structure of the p53/RNA polymerase II assembly. *Communications Biology* *4*, 397.

Liu, G., Parant, J.M., Lang, G., Chau, P., Chavez-Reyes, A., El-Naggar, A.K., Multani, A., Chang, S., and Lozano, G. (2004). Chromosome stability, in the absence of apoptosis, is critical for suppression of tumorigenesis in Trp53 mutant mice. *Nat Genet* *36*, 63-68.

Liu, Z., Scannell, D.R., Eisen, M.B., and Tjian, R. (2011). Control of embryonic stem cell lineage commitment by core promoter factor, TAF3. *Cell* *146*, 720-731.

Lobanenkov, V.V., Nicolas, R.H., Adler, V.V., Paterson, H., Klenova, E.M., Polotskaja, A.V., and Goodwin, G.H. (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.

Martín-Caballero, J., Flores, J.M., García-Palencia, P., and Serrano, M. (2001). Tumor susceptibility of p21(Waf1/Cip1)-deficient mice. *Cancer Res* *61*, 6234-6238.

Martin, M., Zielinski, C., Ruiz-Borrego, M., Carrasco, E., Turner, N., Ciruelos, E.M., Muñoz, M., Bermejo, B., Margeli, M., Anton, A., *et al.* (2021). Palbociclib in combination with endocrine therapy versus capecitabine in hormonal receptor-positive, human epidermal growth factor 2-negative, aromatase inhibitor-resistant metastatic breast cancer: a phase III randomised controlled trial-PEARL. *Ann Oncol* *32*, 488-499.

Melero, J.A., Stitt, D.T., Mangel, W.F., and Carroll, R.B. (1979). Identification of new polypeptide species (48-55K) immunoprecipitable by antiserum to purified large T antigen and present in SV40-infected and -transformed cells. *Virology* *93*, 466-480.

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

Montes de Oca Luna, R., Wagner, D.S., and Lozano, G. (1995). Rescue of early embryonic lethality in mdm2-deficient mice by deletion of p53. *Nature* *378*, 203-206.

Müller, M., Wilder, S., Bannasch, D., Israeli, D., Lehlbach, K., Li-Weber, M., Friedman, S.L., Galle, P.R., Stremmel, W., Oren, M., *et al.* (1998). p53 activates the CD95 (APO-1/Fas) gene in response to DNA damage by anticancer drugs. *J Exp Med* *188*, 2033-2045.

Nakahashi, H., Kieffer Kwon, K.R., Resch, W., Vian, L., Dose, M., Stavreva, D., Hakim, O., Pruett, N., Nelson, S., Yamane, A., *et al.* (2013). A genome-wide map of CTCF multivalency redefines the CTCF code. *Cell Rep* *3*, 1678-1689.

Nakano, K., and Vousden, K.H. (2001). PUMA, a novel proapoptotic gene, is induced by p53. *Mol Cell* *7*, 683-694.

Narendra, V., Bulajić, M., Dekker, J., Mazzoni, E.O., and Reinberg, D. (2016). CTCF-mediated topological boundaries during development foster appropriate gene regulation. *Genes Dev* *30*, 2657-2662.

Narendra, V., Rocha, P.P., An, D., Raviram, R., Skok, J.A., Mazzoni, E.O., and Reinberg, D. (2015). CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. *Science* *347*, 1017-1021.

Natale, F., Rapp, A., Yu, W., Maiser, A., Harz, H., Scholl, A., Grulich, S., Anton, T., Hörl, D., Chen, W., *et al.* (2017). Identification of the elementary structural units of the DNA damage response. *Nat Commun* 8, 15760.

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

Olivier, M., Langerød, A., Carrieri, P., Bergh, J., Klaar, S., Eyfjord, J., Theillet, C., Rodriguez, C., Lidereau, R., Bièche, I., *et al.* (2006). The clinical value of somatic TP53 gene mutations in 1,794 patients with breast cancer. *Clin Cancer Res* 12, 1157-1167.

Overgaard, J., Yilmaz, M., Guldborg, P., Hansen, L.L., and Alsner, J. (2000). TP53 mutation is an independent prognostic marker for poor outcome in both node-negative and node-positive breast cancer. *Acta Oncol* 39, 327-333.

Parant, J., Chavez-Reyes, A., Little, N.A., Yan, W., Reinke, V., Jochemsen, A.G., and Lozano, G. (2001). Rescue of embryonic lethality in Mdm4-null mice by loss of Trp53 suggests a nonoverlapping pathway with MDM2 to regulate p53. *Nat Genet* 29, 92-95.

Peña-Hernández, R., Marques, M., Hilmi, K., Zhao, T., Saad, A., Alaoui-Jamali, M.A., del Rincon, S.V., Ashworth, T., Roy, A.L., Emerson, B.M., *et al.* (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, E677-686.

Qu, Y., Han, B., Yu, Y., Yao, W., Bose, S., Karlan, B.Y., Giuliano, A.E., and Cui, X. (2015). Evaluation of MCF10A as a Reliable Model for Normal Human Mammary Epithelial Cells. *PLoS One* 10, e0131285.

Rao, S.S., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., *et al.* (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 159, 1665-1680.

Robles, S.J., and Adami, G.R. (1998). Agents that cause DNA double strand breaks lead to p16INK4a enrichment and the premature senescence of normal fibroblasts. *Oncogene* 16, 1113-1123.

Roth, J., Dobbstein, M., Freedman, D.A., Shenk, T., and Levine, A.J. (1998). Nucleocytoplasmic shuttling of the hdm2 oncoprotein regulates the levels of the p53 protein via a pathway used by the human immunodeficiency virus rev protein. *The EMBO Journal* 17, 554-564.

Rowley, M.J., and Corces, V.G. (2018). Organizational principles of 3D genome architecture. *Nat Rev Genet* 19, 789-800.

Rubio, E.D., Reiss, D.J., Welsh, P.L., Distche, C.M., Filippova, G.N., Baliga, N.S., Aebersold, R., Ranish, J.A., and Krumm, A. (2008). CTCF physically links cohesin to chromatin. *Proc Natl Acad Sci U S A* 105, 8309-8314.

Sanders, J.T., Freeman, T.F., Xu, Y., Gollosi, R., Stallard, M.A., Hill, A.M., San Martin, R., Balajee, A.S., and McCord, R.P. (2020). Radiation-induced DNA damage and repair effects on 3D genome organization. *Nature Communications* 11, 6178.

Schwartz, M., Portugez, A.S., Attia, B.Z., Tannenbaum, M., Cohen, L., Loza, O., Chase, E., Turman, Y., Kaplan, T., Salah, Z., *et al.* (2020). Genomic retargeting of p53 and CTCF is associated with transcriptional changes during oncogenic HRas-induced transformation. *Communications Biology* 3, 696.

Scully, R., Panday, A., Elango, R., and Willis, N.A. (2019). DNA double-strand break repair-pathway choice in somatic mammalian cells. *Nature Reviews Molecular Cell Biology* 20, 698-714.

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

Seung, S.J., Traore, A.N., Pourmirza, B., Fathers, K.E., Coombes, M., and Jerzak, K.J. (2020). A Population-Based Analysis of Breast Cancer Incidence and Survival by Subtype in Ontario Women. *Current Oncology* 27, 191-198.

Sharp, D.A., Kratowicz, S.A., Sank, M.J., and George, D.L. (1999). Stabilization of the MDM2 oncoprotein by interaction with the structurally related MDMX protein. *J Biol Chem* 274, 38189-38196.

Shvarts, A., Steegenga, W.T., Riteco, N., van Laar, T., Dekker, P., Bazuine, M., van Ham, R.C., van der Houven van Oordt, W., Hateboer, G., van der Eb, A.J., *et al.* (1996). MDMX: a novel p53-binding protein with some functional properties of MDM2. *Embo j* 15, 5349-5357.

Silwal-Pandit, L., Vollan, H.K., Chin, S.F., Rueda, O.M., McKinney, S., Osako, T., Quigley, D.A., Kristensen, V.N., Aparicio, S., Børresen-Dale, A.L., *et al.* (2014). TP53 mutation spectrum in breast cancer is subtype specific and has distinct prognostic relevance. *Clin Cancer Res* 20, 3569-3580.

Sinha, B.K., Mimnaugh, E.G., Rajagopalan, S., and Myers, C.E. (1989). Adriamycin activation and oxygen free radical formation in human breast tumor cells: protective role of glutathione peroxidase in adriamycin resistance. *Cancer Res* 49, 3844-3848.

Smith, A.E., Smith, R., and Paucha, E. (1979). Characterization of different tumor antigens present in cells transformed by simian virus 40. *Cell* 18, 335-346.

Sykes, S.M., Mellert, H.S., Holbert, M.A., Li, K., Marmorstein, R., Lane, W.S., and McMahon, S.B. (2006). Acetylation of the p53 DNA-binding domain regulates apoptosis induction. *Mol Cell* 24, 841-851.

Szabó, P.E., Tang, S.H., Silva, F.J., Tsark, W.M., and Mann, J.R. (2004). Role of CTCF binding sites in the Igf2/H19 imprinting control region. *Mol Cell Biol* 24, 4791-4800.

Tang, Y., Zhao, W., Chen, Y., Zhao, Y., and Gu, W. (2008). Acetylation is indispensable for p53 activation. *Cell* 133, 612-626.

Tewey, K.M., Rowe, T.C., Yang, L., Halligan, B.D., and Liu, L.F. (1984). Adriamycin-induced DNA damage mediated by mammalian DNA topoisomerase II. *Science* 226, 466-468.

Tsuda, H., Callen, D.F., Fukutomi, T., Nakamura, Y., and Hirohashi, S. (1994). Allele loss on chromosome 16q24.2-qter occurs frequently in breast cancers irrespectively of differences in phenotype and extent of spread. *Cancer Res* 54, 513-517.

Valente, L.J., Gray, D.H., Michalak, E.M., Pinon-Hofbauer, J., Egle, A., Scott, C.L., Janic, A., and Strasser, A. (2013). p53 efficiently suppresses tumor development in the complete absence of its cell-cycle inhibitory and proapoptotic effectors p21, Puma, and Noxa. *Cell Rep* 3, 1339-1345.

Végran, F., Rebucci, M., Chevrier, S., Cadouot, M., Boidot, R., and Lizard-Nacol, S. (2013). Only missense mutations affecting the DNA binding domain of p53 influence outcomes in patients with breast carcinoma. *PLoS One* 8, e55103.

Verma, S., Miles, D., Gianni, L., Krop, I.E., Welslau, M., Baselga, J., Pegram, M., Oh, D.-Y., Diéras, V., Guardino, E., *et al.* (2012). Trastuzumab Emtansine for HER2-Positive Advanced Breast Cancer. *New England Journal of Medicine* 367, 1783-1791.

Walker, D.R., Bond, J.P., Tarone, R.E., Harris, C.C., Makalowski, W., Boguski, M.S., and Greenblatt, M.S. (1999). Evolutionary conservation and somatic mutation hotspot maps of p53: correlation with p53 protein structural and functional features. *Oncogene* 18, 211-218.

Witcher, M., and Emerson, B.M. (2009). Epigenetic silencing of the p16(INK4a) tumor suppressor is associated with loss of CTCF binding and a chromatin boundary. *Mol Cell* 34, 271-284.

Wu, X., Bayle, J.H., Olson, D., and Levine, A.J. (1993). The p53-mdm-2 autoregulatory feedback loop. *Genes Dev* 7, 1126-1132.

Wutz, G., Várnai, C., Nagasaka, K., Cisneros, D.A., Stocsits, R.R., Tang, W., Schoenfelder, S., Jessberger, G., Muhar, M., Hossain, M.J., *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.

Xiong, S., Van Pelt, C.S., Elizondo-Fraire, A.C., Liu, G., and 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, 3226-3231.

Yersal, O., and Barutca, S. (2014). Biological subtypes of breast cancer: Prognostic and therapeutic implications. *World J Clin Oncol* 5, 412-424.

Yin, L., Duan, J.J., Bian, X.W., and Yu, S.C. (2020). Triple-negative breast cancer molecular subtyping and treatment progress. *Breast Cancer Res* 22, 61.

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

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

Zhao, B., Benson, E.K., Qiao, R., Wang, X., Kim, S., Manfredi, J.J., Lee, S.W., and Aaronson, S.A. (2009). Cellular senescence and organismal ageing in the absence of p21(CIP1/WAF1) in ku80(-/-) mice. *EMBO Rep* 10, 71-78.