The Texas Medical Center Library

DigitalCommons@TMC

2-2023

Low molecular weight cyclin E deregulates DNA replication and damage repair to promote genomic instability in breast cancer

Mi Li

Low molecular weight cyclin E deregulates DNA replication and damage repair to promote genomic instability in breast cancer

by

Mi Li, B.S.

APPROVED:

_____
Khandan Keyomarsi, Ph.D.
Advisory Professor

_____
Junjie Chen, Ph.D.

_____
Sendurai Mani, Ph.D.

_____
Jian Kuang, Ph.D.

_____
Chandra Bartholomeusz, M.D., Ph.D.

_____
Pawel Mazur, Ph.D.

APPROVED:

_____
Dean, The University of Texas
MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences

Low molecular weight cyclin E deregulates DNA replication and damage repair to

promote genomic instability in breast cancer

A

Dissertation

Presented to the Faculty of

The University of Texas

MD Anderson Cancer Center UTHealth

Graduate School of Biomedical Sciences

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

by

Mi Li, B.S.

December, 2022

DEDICATION

I would like to dedicate my work to my wife Qichen Zhu for her constant love and support.

ACKNOWLEDGEMENTS

Low molecular weight cyclin E deregulates DNA replication and damage repair to promote genomic instability in breast cancer

Mi Li, B.S.

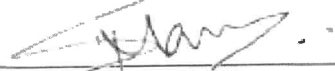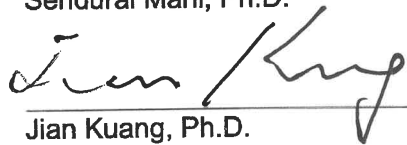Advisory Professor: Khandan Keyomarsi, Ph.D.

Low molecular weight cyclin E (LMW-E) are oncogenic forms of cyclin E that are post translationally generated by neutrophil elastase (NE) mediated cleavage of the 50 KDa full-length cyclin E1 (FL-cycE, encoded by *CCNE1* gene). The resultant N-terminus deleted (40 amino acids) form of LMW-E is detected in breast cancer cells and tumor tissues, but not in normal mammary epithelial cells or adjacent normal tissues. Unlike FL-cycE, LMW-E drives mammary epithelial cell transformation in human cells and spontaneous mammary tumor formation in transgenic mouse models, but the oncogenic mechanisms of LMW-E and its unique function(s) independent of FL-cycE are not fully understood. It is currently assumed that LMW-E drives the tumorigenic process by promoting G1/S cell cycle transition and accelerating mitotic exit. Biochemical features such as longer protein half-life, higher affinity to its kinase partner CDK2, and resistance to endogenous CDK inhibitors such as p21 and p27 all promote the tumorigenic ability of LMW-E. Clinical studies in breast cancer reveal that overexpression of LMW-E predicts recurrence and poor survival in breast cancer patients independent of molecular subtype, Ki67 status, nodal status, or tumor grade, suggesting LMW-E may be driving breast cancer development independent of its role in cell proliferation.

In the current study, we tested the hypothesis that LMW-E promotes genomic instability by deregulating DNA replication and damage repair. We generated immortalized pre-cancerous human mammary epithelial cells (hMECs) engineered to express doxycycline

inducible LMW-E or FL-cycE in *CCNE1* knock-out background. We found that, unlike LMW-E, FL-cycE overexpression led to DNA replication stress and DNA damage accumulation, resulting in reduced cell viability. LMW-E overexpression, on the other hand, promoted cell survival under replication stress, resulting in persistent genomic instability. RNA-sequencing results showed LMW-E but not FL-cycE overexpression enhanced DNA replication and damage repair pathways. Molecularly, LMW-E but not FL-cycE strongly interacted with CDC6, bound to chromatin, and facilitated replication stress tolerance by upregulating pre-replication complex assembly. LMW-E also mediated DNA repair by upregulating the levels of RAD51 and C17orf53, showing a dominant repairing effect over DNA damage induced by FL-cycE. Moreover, targeting the replication stress response pathway ATR-CHK1-RAD51 with small molecule inhibitors significantly decreased viability of LMW-E overexpressing hMECs and breast cancer cells. Lastly, we showed that positive LMW-E status was associated with genomic instability in tumors from a cohort of 725 patients diagnosed with early-stage breast cancer, further supporting our hypothesis that LMW-E promotes genomic instability to fuel breast cancer development.

Collectively, our findings delineated a novel role for LMW-E in breast tumorigenesis mediated by replication stress tolerance and genomic instability, providing novel therapeutic strategies for LMW-E overexpressing breast cancers.

Table of Contents

List of Figures

List of Tables

Chapter One: Introduction

1.1 Breast cancer overview: detection, treatment and drug resistance

Malignant neoplasm, also named cancer, is a genetic disease caused by changes to genes regulating the growth and behavior of cells that lead to un-controlled cell proliferation and invasion to nearby or distant tissues[1,2]. Recent data suggest cancer is the 2nd ranking cause of death, accounting for more than 20% of all deaths occurring in the United States, and the 1st leading cause of death in the age group between 45 and 64[3]. Breast cancer is one of the most commonly diagnosed cancers worldwide, accounting for 30% of all cancers diagnosed in women. There is an estimated 287,850 new patients expected to be diagnosed with the disease in 2022[4] . Although the survival of breast cancer patients has been greatly improved since the middle of 1970s, breast cancer is still the second most common cause of cancer deaths in women, accounting for an estimated 43,250 deaths in 2022[4] .

Screening tests such as mammography, breast magnetic resonance imaging (MRI) or clinical breast exams significantly improve the capacity to detect early stage breast cancers and enhancing the probability to remove premalignant lesions[5] . These measures effectively decrease breast cancer death rates in the past decades, while the improved breast cancer detection also leads to increased incidences of this disease. For instance, the incidence of stage 0 breast cancer, or ductal carcinoma in situ (DCIS), has increased 500% between 1983 to 2003 for women 50 years of age and older[6]. Additionally, improved breast imaging with follow-up studies reveal the existence of large patient populations that may never require treatment[7]. Studies in DCIS also show that in patients treated with breast-conserving surgery with a 13–20 years follow-up time, the recurrence rates range from 9–23% for patients treated with radiation therapy, and 26–36% for those without radiation treatment[8]. It is currently a clinical challenge to precisely predict tumor recurrence in early stage breast cancers and to separate the cases that are likely to recur from those which are indolent.

Since the late 1960s, much progress has been made to reduce the mortality of breast cancer through translation of basic scientific findings to clinical therapeutic intervention. Cytotoxic chemotherapeutic agents have been used for the treatment for breast cancers, based on their capacity of inducing DNA damage during DNA replication, leading to either cell death or inhibition of cell proliferation. However, due to the non-specific nature of these chemotherapies, patients suffer from adverse effects that greatly limit the treatment frequency and tolerable dosage. In the early 1970s, hormone receptors (ER and PR) in breast cancer cells have been identified as both critical prognostic markers and the first targetable molecules for endocrine therapies (further discussed in 1.2)[7,9] . Anti-hormonal therapies become the mainstay treatment method for hormone receptor positive breast cancers until today. It also shows powerful impact on development of other targeted therapies in oncology[7]. Until the late 1980s, HER2 positive status was regarded as a poor prognostic marker in early-stage breast cancers. Development of HER2-directed therapies have shown undeniable beneficial impact on the overall survival and disease-free survival of HER2-positive breast cancer patients since the 1990s[10]. Similarly, based on the discovery of synthetic lethality between the deficiencies of Poly(ADP-Ribose) Polymerase (PARP) and Breast cancer type 1 (and type 2) susceptibility protein (BRCA1 and 2), PARP class of targeted inhibitors have been approved for the treatment of high-risk germline pathogenic BRCA1 or BRCA2-mutated (or likely pathogenic variant) HER2-negative breast cancer[11]. These translational achievement in basic science research have greatly improved breast cancer management, and provide the paradigm of biomarker driven cancer therapies[11-14].

Despite the progress in the breast cancer detection and treatment, many challenges still remain. For example, drug resistance is a major obstacle for all systematic and targeted treatments. Mechanisms underlying drug resistance are multifactorial for different classes of therapeutic agents and may vary among cases using the same agent. The non-targeted

chemotherapy may cause neosis, a self-renewal process of the damaged tumor tissues, facilitating acquired resistance to chemotherapy. This process is characterized by the production of small mono-nuclear "Raju" cells derived from damaged parental cells. These cells have extended mitotic life span with inherit genomic instability, which further potentiate tumorigenesis and metastasis[15]. Chemotherapy resistance may also result from a compensatory mechanism by which damaged tumor cells promote cell division at adjacent tumor tissues. In particular, chemotherapy induces apoptosis of the tumor cells, and the activated caspase- 3 and 7 promotes the cleavage of calcium-independent phospholipase A2. This process enhances the synthesis of arachidonic acid and the production of prostaglandin E2, a pro-expanding hormone to stimulate cell proliferation in nearby cancer tissues[16]. Besides the aforementioned adaptive mechanisms functioning in response to on-going treatment, cancer cells may intrinsically rely on its population of micro-clonality/micro-genetic heterogeneity. For example, cancer stem cells (CSCs) exhibit relatively slow metabolism and are resistance to DNA damaging agents[17]. Additionally, the majority of the aggressive ER-positive breast cancers derive from dormant micro-metastases, which are not affected by initial anti-estrogen treatments[7]. Other mechanisms such as altered drug influx/efflux transporters, deregulated cell cycle checkpoints, and defective apoptosis pathways, may also contribute to drug resistance due to the growth advantage effect on tumor cells under selective pressure[18].

As a result, despite early detections through screening tests and advances in treatment, breast cancer incidences continue to rise and advanced breast cancers with metastases to distant organs remain lethal[5]. Identification of effective biomarkers to predict and/or to track disease specific alterations during breast cancer development is an area of unmet clinical need. Understanding the biological mechanisms underlying these alterations

may provide novel therapeutic strategies for breast cancers which fail to respond to currently available therapeutics (Figure 1).
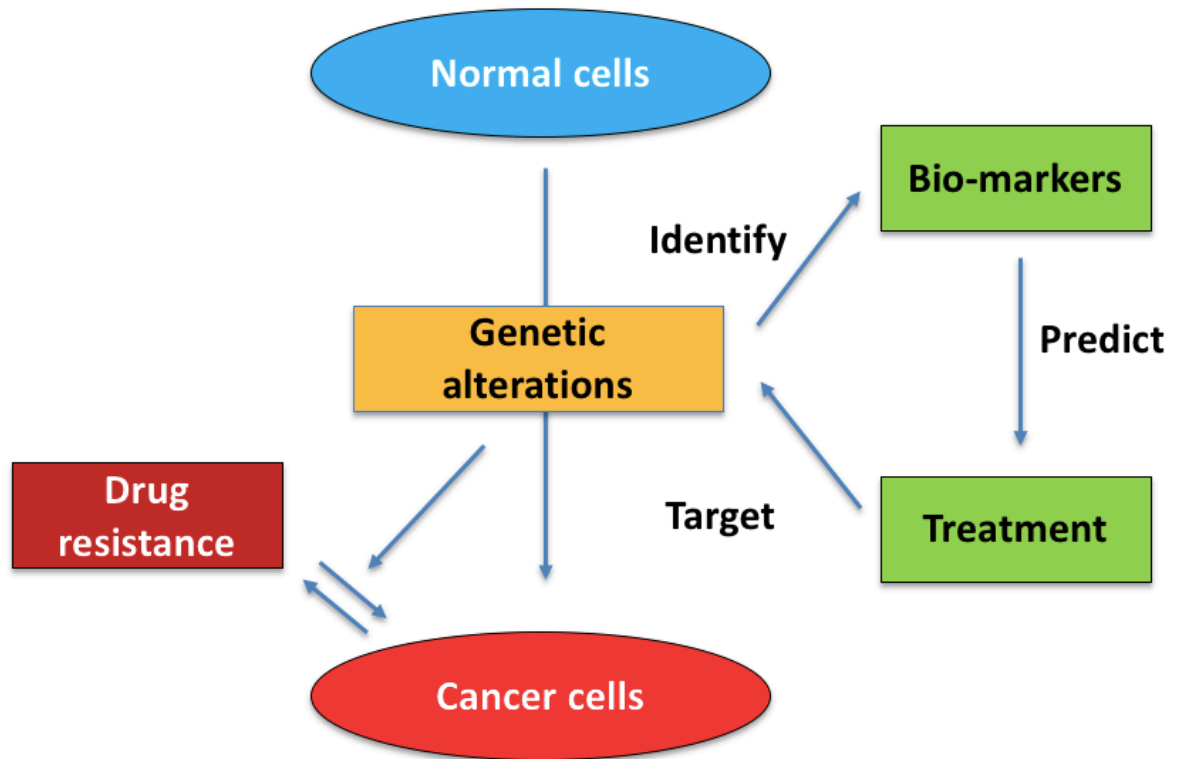
Figure 1. Genetic alterations fuel cancer transformation and drug resistance. Identification of disease specific alterations may provide bio-markers to tract cancer development, predict treatment responses and provide novel treatment strategies.

1.2 Breast cancer subtypes

The history of breast cancer management has been transformed through changes in understanding of the clinical features followed by characterizing underlying biological mechanisms. Breast cancer is intrinsically caused by genetic alterations in mammary epithelial cells, and one of the well-characterized features is the overexpression of specific cell surface receptors[19]. Clinically, the most common receptors expressed in breast cancers include hormone receptors such as estrogen receptor (ER) and progesterone receptor (PR),

and the epidermal growth factor receptor (EGFR) or the Receptor tyrosine-protein kinase erbB (ERBB) family like ERBB1/EGFR and ERBB2/HER2. As the expression of ER, PR, HER2 and EGFR are found in approximately 75%, 60%, 25% and 15% of breast cancers, the presence of these receptors are also considered as major drivers for breast cancer development [20]. Clinically, the status of these receptors is determined by immune-histochemistry (IHC) to test protein expression, as well as fluorescence in situ hybridization (FISH) or chromogenic in situ hybridization (CISH) to detect gene amplification in breast cancer cells[20,21]. One of the well-established systems to differentiate breast cancer subtypes is based on the status of these receptors, and thereby categorizes breast cancers into the following subtypes: i) hormone receptor positive (ER+/PR+/HER2-), ii) hormone receptor and HER-2 positive (ER+/PR+/HER2+), iii) HER-2 amplified (ER-/PR-/HER2+) and iv) triple negative (ER-/PR-/HER2-; Figure 2)[20].

Expression of the hormone receptors ER and PR are found in 75% of breast cancers patients[14]. Signaling cascade initiated by the ER and PR, increases cancer cell proliferation by promoting cell cycle progression (further discussed in 1.3)[22]. Endocrine therapies are administered to these hormone receptor positive breast cancer patients based on their menopausal status and the source of estrogen[23,24]. In premenopausal women estrogen is mainly produced in ovaries. Anti-estrogen agents, such as tamoxifen which competes with estrogen in the binding with ER as well as inhibit estrogen effects by forming a complex in the nucleus to inhibit DNA replication, are used to treat premenopausal, ER positive early or advanced breast cancer patients[24]. In post-menopausal women, because the aromatase enzyme mediates the production of estrogen from androgens produced by adipose tissue, the aromatase inhibitors such as letrozole and anastrozole are administered to ER positive breast cancer patients to inhibit the production of estrogen[23]. For advanced ER positive breast cancers who do not respond to aromatase inhibitors, fulvestrant, a compound that binds to

ER monomer to inhibit ER dimerization, translocation, and to promote ER degradation, has been approved as a standalone second line treatment for ER positive HER2-negative metastatic breast cancer in post-menopausal women[25,26].
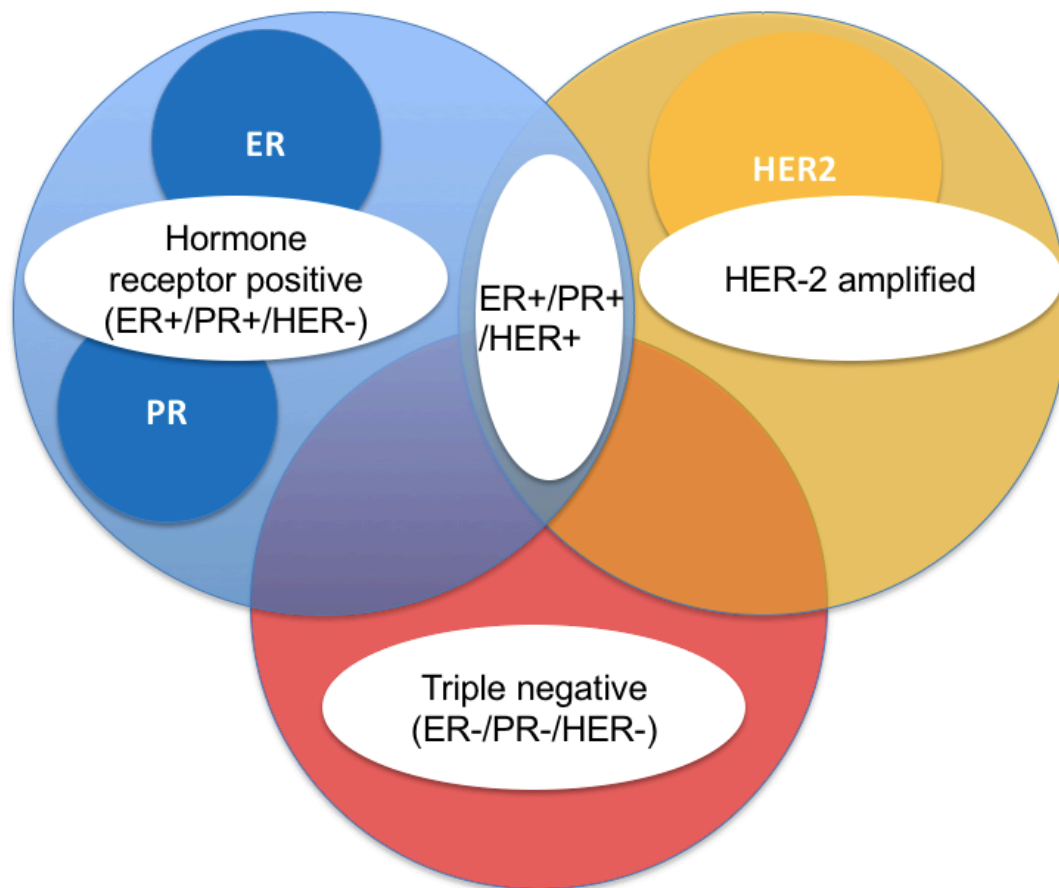
Figure 2. Breast cancer subtypes categorized by cell surface receptors. Breast cancers are categorized into four subtypes based on the status of hormone receptor (ER and PR) and HER-2 receptor. hormone receptor positive (ER+/PR+/HER2-), hormone receptor and HER-2 positive (ER+/PR+/HER2+), HER-2 amplified (ER-/PR-/HER2+) and triple negative (ER-/PR-/HER2-) breast cancer.

HER2 positive breast cancers are characterized by amplification of *ERBB2* gene and/or overexpression of ERBB2/HER2 protein, which accounts for 25 to 30% of all breast cancers[27,28]. HER2 positive breast cancers have up to 25-50 copies of the HER2 gene and up to 40 -100-fold in the expression of HER2 protein[29,30] . Notably, in early stage breast cancers, particularly *in situ* ductal carcinomas without invasion to nearby tissues, the amplification of HER2 gene is found in nearly half of the patients examined[31]. During breast cancer progression into invasive ductal carcinomas, and metastatic breast cancer into lymph nodes and/or distant tissues, the overexpression of HER2 remain persistant[31]. Compared to HR positive/HER2 negative breast cancers, HER2 positive breast cancers exhibit increased progression capacity, worse prognosis, and higher propensity to metastasize to the brain[32,33]. Mechanistically, as a membrane receptor tyrosine kinase, HER2 overexpression leads to the dimerization with other receptor tyrosine kinases such as EGFR and HER3[34]. The resulting complex EGFR-HER2 heterodimers are resistant to endocytic regulation so that the degradation of EGFR is inhibited. As a result, the signaling activated by these receptor tyrosine kinases, such as PI3K/AKT pathway are hyper activated with a prolonged growth signal duration, and subsequently promote cell proliferation by promoting the expression of cyclin D and activating down-stream cell cycle pathways (will be discussed in 1.3)[34]. Targeted therapies against HER2 are administered for the treatment of HER2 positive advanced breast cancer patients. These agents include tyrosine-kinase inhibitor lapatinib and receptor tyrosine-kinase binding/blocking antibodies such as trastuzumab and pertuzumab. The cancer killing mechanisms of these agents are known to be twofold: one is to prevent the activation of the intracellular tyrosine kinase and on the other is through immune responses against tumor cells mediated by antibody-dependent cellular cytotoxicity (ADCC)[35].

Breast cancers that are negative for ER, PR or HER2 are termed as "triple negative" breast cancer (TNBC). Historically, TNBC were treated with single agent chemotherapy, but

the response is relatively poor[36]. HR and HER2 targeted therapies are not applicable to TNBC for the lack of HR or HER2 expression. As a result, TNBC are often associated with poor prognosis[36]. High level of genomic instability and high frequency in mutations of tumor suppressors such as *RB, TP53* and *BRCA1* are found to be associated with TNBC[37]. Studies have also shown that an important regulator of the cell cycle, cyclin E is overexpressed in TNBC[38]. In particular, the low molecular weight isoforms of cyclin E, termed LMW-E, is present in 80% of all TNBC tumors samples[39] (further discussed in 1.4). This observation suggest that LMW-E expression may drive the development of TNBC by deregulating cell cycle progression and/or inducing genomic instability. Understanding the effect of LMW-E on breast cancer cells may provide novel therapeutic strategies for the treatment of TNBC, for which targeted therapy options are rare.

1.3 Cell cycle overview

Cell signaling cascades initiated by ER, PR, and HER2, increases cancer cell proliferation by promoting cell cycle progression[40,41]. These events include exiting of cells from the resting phase (G0) and entering into the interphase; comprised of the first gap phase (G1), DNA synthesis phase (S), and the second gap phase (G2)[42,43]. In G1, the cell prepares for the doubling of the genome, and the DNA replication is completed in S phase. The cell prepares in G2 to leave interphase and enter the mitotic (M) phase for further separation of chromatids and the formation of daughter cells. After the M phase, the first cell cycle is completed and each daughter cell may or may not enter the next cell cycle for further proliferation[44].

In normal dividing cells, the cell cycle progression is tightly regulated by the expression, modification, and destruction of a family of cell cycle proteins called cyclin(s). These cyclins bind with cyclin dependent kinases (CDKs). The levels of most CDKs are relatively stable throughout the cell cycle, but their activities are regulated due to the fluctuation of their cyclin binding partner. Cyclin-CDK complexes are also regulated by CDK

inhibitors. The CDK inhibitors directly bind to and block the activity of CDKs at multiple points of the cell cycle. Cyclin D-CDK4/6 are specifically inhibited by INK4 proteins (p16, p15, p18 and p19). Cip/Kip (p21, p27) proteins are specialized to inhibit other cyclin-CDKs including cyclin E-CDK2, cyclin A-CDK2, cyclin A-CDK1 and cyclin B-CDK1[45] (Figure 3).

The cyclin - CDK protein kinase complexes phosphorylate specific protein substrates to drive the cell cycle progression at each phase of the cell cycle[44]. The exiting out of interphase and entering into the mitotic phase is mainly regulated by the principal mitotic cyclin B-CDK1 complex[46]. Currently, the understanding regarding the protein substrates of cyclin B-CDK1 in G2/M transition is still very limited. Nuclear lamin is one of the substrates identified to be phosphorylated by cyclin B-CDK1 during mitosis. Reports suggest that cyclin B-CDK1 catalyze the phosphorylation of nuclear lamin at Ser-636 in lamin A, as well as Thr-19, Ser-22 and Ser-392 in both lamins A and C[47]. The phosphorylation of lamin lead to the breakdown of nuclear envelope, which is essential to the separation of chromosomes and the progression of mitosis[47]. Prior to the peak of cyclin B-CDK1 kinase activity, cyclin A binds to CDK1, although whether and how cyclin A-CDK1 kinase specificity differs from cyclin B-CDK1 remains unclear[48]. Cyclin A can also bind with CDK2 to form a protein kinase complex during S phase, and phosphorylates a large variety of proteins which are involved multiple cell cycle functions, such as DNA replication (ORC1, ORC2, MCM2, LIG3, CDT1, ATRIP, TOP2B, XRCC1), histone modification (DOT1L, JARID2, KAT6A, LSD1, MSL1, MSL3, SETDB1, and NSD2), chromatin remodeling (BCOR, DMAP1, INO80E, and SMARCA5), splicing and RNA metabolism (PRP3, PRP16/DHX38, and SRRM2)[49].

Cyclin D binds to CDK4 and CDK6 to form the cyclin D- CDK4/6 protein kinase complex that function in G1 phase. The most well-characterized substrate of cyclin D- CDK4/6 to promote G1/S transition is the retinoblastoma protein (pRB). Phosphorylation of pRB by
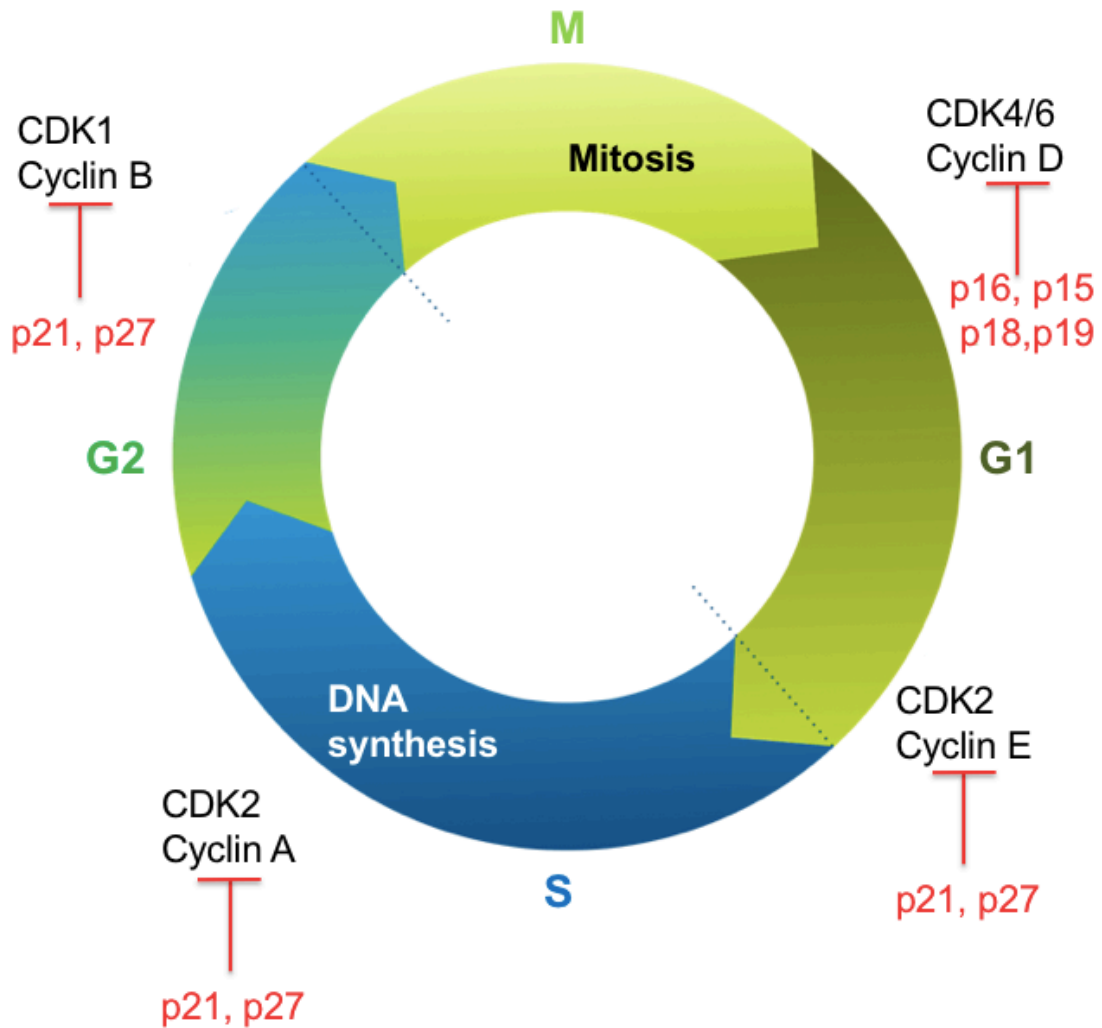
Figure 3. Cell cycle regulation by cyclin-cyclin dependent kinase (CDK) complexes.
Cyclin(s) drive the cell cycle progression by binding with CDKs to form active protein kinase complexes. Cyclin D- CDK4/6 protein kinase complex mainly functions in G1 phase. Cyclin E and cyclin A are important for entry and the progression of S phase. The exiting out of G2 and entering into the M phase is mainly regulated by the mitotic cyclin B-CDK1 complex. The activities of cyclin-CDKs are negatively regulated by cell endogenous inhibitor proteins p16, p15, p18, p19 that mainly inhibit cyclin D-CDK4/6, as well as p21 and p27 that inhibit other classes of cyclin-CDKs.

cyclin D- CDK4/6 releases the transcription factor E2F, which in turn promotes the transcription of cyclin E and cyclin A to further drive cell cycle progression. Cyclin E and cyclin A are important for entry and the progression of S phase[46], and both of them activate DNA replication[50]. Phosphorylated pRb also binds to histone deacetylase (HDAC), which in turn leads to the remodeling of chromatin required for DNA replication[51].

1.4 Deregulation of cyclin E in breast cancer

Core regulatory proteins involved in G1/S transitions, such as cyclin D and cyclin E have been shown as biomarkers to predict breast cancer prognostics[52,53]. Ten to twenty percent of breast cancers harbor the amplification of CCND1 gene[52,54]. One of the main mechanisms by which cyclin D protein overexpression promotes cell cycle progression, is by phosphorylating pRB. Besides gene amplification, the accumulation of cyclin D protein occurs from enhanced transcription of CCND1 gene. Binding of estrogen and ER activate the mitogenic signaling pathway to promote the transcription of CCND1 gene and the overexpression of cyclin D protein[55,56]. Another mechanism is via post-translational modification (e.g. phosphorylation) of cyclin D protein and its upstream regulators. The signaling cascade is mainly activated through PI3K-AKT-mTOR- pathway, which can be activated by HER2 overexpression and lead to the degradation of glycogen synthase kinase3β (GSK3β). GSK3β promotes cyclin D phosphorylation at T286, which drives the translocation of cyclin D from cell nuclear to cytoplasm. This is followed by SCF E3 ubiquitin ligase mediated cyclin D ubiquitination and proteasome mediated degradation[57,58]. As a result, via inhibition of GSK3β mediated cyclin D phosphorylation at T286, cyclin D protein is stabilized by the activation of PI3K-AKT-mTOR pathway[57]. Although CCND1 gene is required for breast tumor development[59], cyclinD1 overexpression predicts worse clinical outcome mainly in ER positive breast cancers but not in other types of breast cancers[52].

Overexpression of cyclin E protein independently predicts disease-specific survival in breast cancer patients regardless of breast cancer subtypes[53,60]. Cyclin E-CDK2 drives G1/S transition and E2F activation by phosphorylating pRB. In addition to pRb, cyclin E-CDK2 complex phosphorylate other substrates, each with a different function. These substrates include: NPAT involved in histone biosynthesis, FOXO1 in apoptotic response, NPM in centrosome duplication, as well as CDC6 and MCMs in DNA replication[61]. Cyclin E overexpression is mediated by multiple mechanisms. Cyclin D binds to CDK4 or CDK6 and catalyzes the phosphorylation of pRB, which then release the transcription factor E2F. The binding of E2F to the promotor of *CCNE* gene induce the transcription and overexpression of cyclin E[51]. Other mechanisms such as *CCNE1* gene amplification, disruption of cyclin E degradation, also lead to upregulation of cyclin E protein[62]. Additionally, the function of cyclin E is affected by post-translational modification of cyclin E protein such as the generation of low molecular weight cyclin E (LMW-E) from full-length cyclin E (FL-cycE; Figure 4)[39,63].

Enzymatic cleavage of the 50 kDa FL-cycE by neutrophil elastase (NE) recapitulates the LMW-E patterns (ranging from 45 to 33 kDa) observed in cancer cells[63,64]. Additionally, alternative transcription starting at Methionine 46 (M46) from N-terminus of cyclin E is also linked to the generation of a specific 45 kDa LMW-E in tumor cells (Figure 5)[64,65]. Although alternative splicing of *CCNE1* mRNA are also found in the cells, none of these spliced mRNA products is associated with the production of the LMW-E proteins[63].

Overexpression of LMW-E and FL-cycE are not mutually exclusive. FL-cycE is expressed in both cancer cells and normal cells, but LMW-E is cancer specific and not detectable in normal cells or tissues[39,53]. LMW-E has a higher binding affinity towards CDK2 compared with FL-CycE, and the protein kinase complex formed between LMW-E and CDK2 is less sensitive to the inhibition by p21 or p27[66]. Moreover, LMW-E is less susceptible than FL-CycE to FBW7 mediated ubiquitination and degradation[67]. As a result, LMW-E hyper- activates CDK2 and

may facilitate CDK2 activity throughout the cell cycle[39]. Importantly, the overexpression of LMW-E has been detected in 50% of ER positive, 75% of HER2 positive and 80% of TNBC patients[39]. Compared to cyclin D and FL-cycE, overexpression of LMW-E is the most prominent and consistent prognostic marker for breast cancer recurrence and breast cancer patient survival (Figure 4)[39].
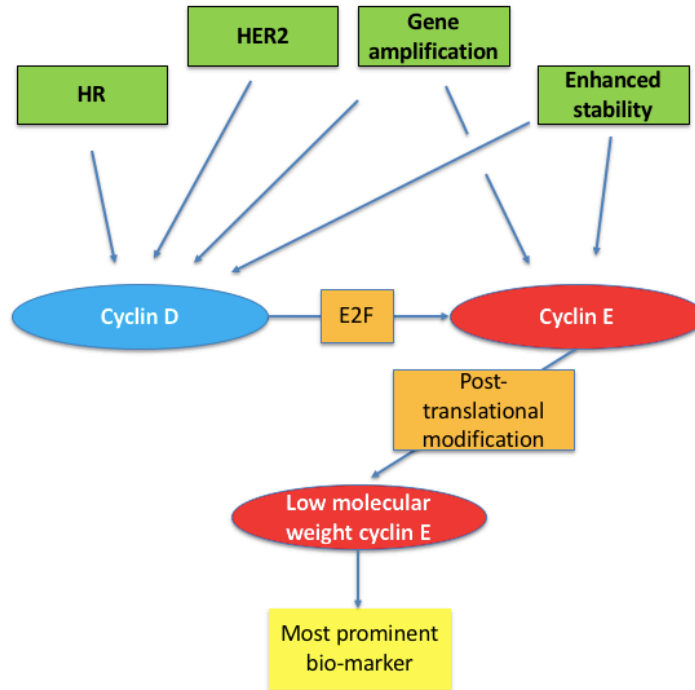
Figure 4. Deregulation of cyclin E in breast cancer. Driven by activated hormone receptors and HER2 receptor tyrosine kinase, cyclin D is overexpressed in breast cancer cells, which in turn promotes the E2F mediated transcription of cyclin E. Overexpression of cyclin E and cyclin D also result from *CCND1* and *CCNE1* gene amplification and/or enhanced protein stability. Moreover, cyclin E protein might be further post-translationally modified to generate the low molecular weight cyclin E (LMW-E). Overexpression of cyclin D, cyclin E, and LMW-E are associated with poor clinical outcome of breast cancer, while LMW-E is the most prominent biomarker that independently predicts tumor recurrence regardless of breast cancer subtypes.
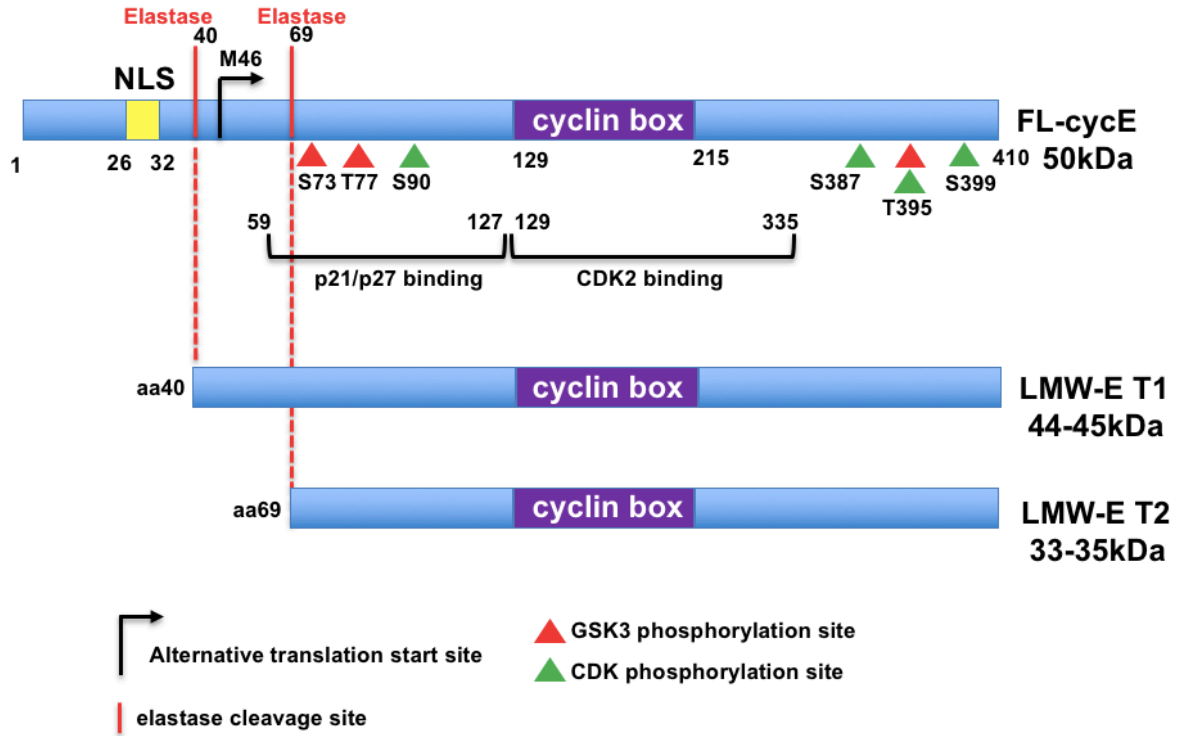
Figure 5. Schematic representation of FL-cycE and LMW-E proteins. LMW-E isoforms (ranging from 45 to 33 kDa) are generated by enzymatic cleavage at amino acid(aa) 40 and 69 of the FL-cycE (50 kDa) protein by neutrophil elastase (NE) [63,64]. Alternative transcription starting at methionine 46 (M46) from N-terminus of cyclin E is also linked to the generation of a 45 kDa LMW-E in tumor cells[64,65]. Both the LMW-E truncation 1 (T1, 44-45 kDa) and truncation 2 (T2, 33-35 kDa) lack the nuclear localization signal (NLS) which is located between aa26 – 32 at the N-terminus of the FL-cycE protein. LMW-E isoforms still contain other important domains such as the cyclin box (aa129-215), p21 and p21 binding domain (aa15-127), CDK2 binding domain (aa129-335), as well as multiple GSK3 and CDK phosphorylation sites indicated in the figure.

1.5 Role of neutrophil elastase in tumor development

As discussed in 1.4, protein cleavage of FL-cycE mediated by neutrophil elastase (NE) results in the generation of LMW-E. These post-translational modifications of cyclin E at two N-terminus cleavage sites aa40 and aa69 respectively lead to the LMW-E T1 and T2 isoforms (Figure 5).

Tumor infiltrating neutrophils are associated with several steps in tumor development. Cytokines and chemokines produced and secreted tumor cells, such as IL-1, IL-8 and TNF-a may induce the recruitment of inflammatory leukocytes, neutrophils and myeloid-derived suppressor cells (MDSCs)[68-70]. Additionally, in response to cytokines secreted into the circulation system, bone marrow upregulates the production of neutrophils and MDSCs, and increased number of circulating neutrophils and their granulocytic product are associated with tumor progression and worsened patient survival[71]. These granulocytic products include NE, proteinase 3 (PR3) and cathepsin G (CG), which are among the chymotrypsin superfamily of proteases collectively termed as neutrophil serine proteases[72]. The generation of these homolog proteins are resulted from duplication of ancestral protease gene belong to the neutrophil lineage. ELA2 gene located on chromosome 19p13.3 encodes NE, while PRTN3 and CTSG genes encode PR3 and CG respectively[72]. Physiologically, NE, PR3 and CG are produced in the maturation process of neutrophils differentiation and the sequential synthesis and secretion of primary azurophilic granules, secondary specific granules, tertiary gelatinase granules, and secretory granules[72](Figure 6). The level of NE is often associated with pathological conditions such as neutrophil influx into the sites of inflammation, where NE functions as a positive regulator in immune responses, microbial elimination and would repair[73,74]. The half-life of circulating neutrophils normally lasts a few hours[74], but the detection of NE in non-neutrophilic cell types were also reported. For example, tumor cells expression of NE and PR3 in tumor cells are also demonstrated by Desmedt et al., showing that breast

cancer cells may produce both proteases NE and PR3[75]. The level of NE in breast cancers negatively correlated with ER status[76], and associated with rapid relapse and worsened survival in a cohort of 1,143 primary breast cancer patients[77]. High levels of NE are also found to be associated with in bladder, lung, prostate, colon and pancreatic cancer progression and serves as a prognostic indicator in these cancer patients[70,78,79].

When produced in the differentiation process of neutrophils, NE is in its inactive "pre-pro" form prior to being packaged into the azurophilic granules, featured by the pro-dipeptide Ser14-Glu15 sequence. This pre-pro NE is processed by proteases including dipeptidyl peptidase I (DPPI) or cathepsin C to cleave the N-terminal pro-dipeptide of NE. The removal of the pro-dipeptide leads to the opening of the active site (free amino group of Ile16 and Asp194), featuring the activated form of NE. The substrate specificity of NE overlaps with PR3, for which both enzymes show a strong preference for Val and Ala in the P1 position of the cleavage site. NE but not PR3 also cleave substrates with Ile in the P1 position, and prefer multiple aliphatic aa both upstream and downstream of the P1 cleavage site[80].

The major physiological function of NE is the intracellular destruction of pathogens[73,74]. The role of NE in tumor development has been demonstrated in the contributions of extracellular NE to chronic inflammatory tumor microenvironment as well as altering intracellular signaling in favor of tumor growth[78,79]. NE has been shown to cleave the tumor extracellular matrix (ECM) elements, such as matrix protein elastin[81], and activates tissue degrading proteases such as MMP-2, MMP-3, and MMP-9, leading to the cleavage of tissue adhesion molecules such as E-cadherin, VCAM-1, JAM-C and G-CSFR[70,81]. Over-activated NE in tumor promotes cell invasion and metastasis through ECM degradation and the cleavage of adhesion proteins. Additionally, neutrophils in the tumor microenvironment also release NE in response to signals from cytokines and chemokines such as TNF-α and IL-8[68-70]. As a result, toll-like receptor 4 (TLR4), a down-stream factor activated by NE, promotes

NF-κB activation via MyD88-IRAK-TRAF6 signaling pathway, and induces the expression of IL-8[82]. Additionally, NE can cleave MMP and meprin-α to activate the TGFα signaling[83]. These pathways collectively form a positive feedback loop of NE driven inflammatory cytokines, facilitate a tumor promoting chronic inflammatory microenvironment, leading to the activation of IL-1 signaling, G-CSF and vascular endothelial growth factor (VEGF)[84-86]. As a result, over-activated NE may favor tumor progression by promoting angiogenesis and tumor metastasis (Figure 6).

In addition to the function of NE in ECM remodeling, further studies indicate NE may directly promote tumor cell proliferation. NE may enter tumor cells by endosomes, and intracellular NE promotes PI3K activity by degradation of Insulin Receptor Substrate-1 (IRS-1), leading to oncogenic AKT activation in human lung cancer cells[87]. NE promotes acute promyelocytic leukemia (APL) and acute myeloid leukemia (AML) respectively by cleaving the fusion protein consisting of the promyelocytic leukemia gene and the retinoic acid receptor alpha (PML-RARα) and transcription factor CUX-1[88,89]. In breast cancer, the intracellular NE mediated LMW-E expression is associated with hyper-activate CDK2 and worsened prognosis in breast cancer patient and NE may also lead to the mammary specific overexpression of CUX-1 which drives tumor formation and lung metastasis[63,75]. Recent studies using NE knock-out murine models suggest NE may promote the chemotactic migratory capacity of tumor cells, which is critical to the escape and active entry of primary tumor cells into intra-tumor vasculature[70].

Following physiological inflammation, NE activity is negatively regulated by serine protease inhibitors such as elafin and alpha1-antitrypsin (AAT), which promote inflammation resolving and restore tissue homeostasis[90]. However, in disease states such as cancer, where the tumor microenvironment is characterized by chronic or excessive inflammation and neutrophil accumulation, the imbalance between NE and its inhibitors may lead to

symptomatic ECM destruction, perpetuates inflammation, and tumor progression[78]. AAT can bind with NE, PR3 or CG, forming a protein complex that leads to the cleavage of the reactive center loop and degradation of both AAT and the binding neutrophil[91,92]. AAT preferably inhibits NE over PR3, as is shown by the faster rate of AAT-NE complex formation compared with AAT-PR3[91,92]. The association between AAT and NE are affected by the oxidization status of AAT, and the oxidized AAT is one of the potential markers for the activation of NE[91,92]. Additionally, the interaction between AAT and NE is also prevented by heparin and other glycosaminoglycans found at the site of inflammation[91,92].

Elafin is the highly specific and potent inhibitor of NE, PR3 and porcine pancreatic elastase (PPE)[93], encoded by the gene PI3 located on chromosome 20q13[94,95]. Mechanistically, elafin non-covalently binds to the catalytic cleft of NE to block the docking of substrate proteins[94,96]. Elafin is expressed at a relatively high concentration by epithelial cells, providing local anti-protease effect, and its levels are also induced by negative feed-back signals driven by TNF-α, p38 MAPK, c-JUN, and NF-κB pathways in response to inflammatory cytokines[94]. As a result, imbalance between NE and elafin may lead to excessive or chronic inflammation both in non-tumor diseases such as chronic obstructive pulmonary disease (COPD), cystic fibrosis, acute respiratory distress syndrome (ARDS), pulmonary fibrosis, asthma, inflammatory bowel disease, psoriasis, impetigo herpetiformis, as well as tumors[96]. For instances, elafin is a prognostic indicator in breast cancer, loss of elafin is significantly associated with shorter time of tumor relapse[97]. In ovarian tumors, elafin-positive status is correlated with tumor recurrence and reduced survival[98]. Low levels of elafin is also associated with poorly differentiated squamous cell carcinomas of the skin, head/neck, and esophagus compared to well differentiated tumors[98]. Elafin expression is downregulated in 24% of DCIS and 83% of invasive breast tumors when compared to the level of elafin in the normal mammary epithelium[98]. Elafin is also downregulated in 33% of ovarian cystadenomas, 43%

of borderline ovarian tumors, and 86% of invasive ovarian carcinomas when compared to the expression of elafin in the normal fallopian tubes[95,98,99]. In non-tumorigenic mammary cell lines such as 76N, 76NE6, 76NF2V and MCF10A, elafin is highly expressed while the expression of elastase is low. In breast cancer cell lines such as ER-positive MCF-7, ZR75T, T47D, BT20T, and ER negative MDA-MB-157, MDA-MB-231, MDA-MB-436 and MDA-MB-468, the expression status of elafin and elastase is reversed, showing the down-regulated elafin and up-regulation of elastase[97]. Mechanistically, C/EBPβ the transcription factor which functions in normal mammary gland development and differentiation, positively regulates elafin by transactivating gene transcription. C/EBPβ can induce the transcription of elafin by 3 to 10 fold in non-tumorigenic mammary cell lines compared with breast cancer cell lines[93] . Consistently, C/EBPβ is down-regulated in breast tumor tissues compared with adjacent normal tissues in 152 patients with stage I, II, or III breast cancer[93]. DNA binding analyses of the elafin promoter region suggest that within the 440 bp region upstream of the elafin promoter which harbors seven potential C/EBPβ binding sites, C/EBPβ-2, C/EBPβ-4 and C/EBPβ-5 sites are responsible for the interaction with C/EBPβ protein[93]. In breast cancer cell lines, C/EBPβ overexpression leads to enhanced elafin mRNA expression, while in non-tumorigenic mammary cells, knock-down of C/EBPβ by siRNA results in decreased elafin expression[93]. Overexpression of elafin in breast cancer cells leads to reduced cell proliferation and induction of apoptosis, consistent to the effect of elastase depletion by shRNA[97]. Knock down of elastase by shRNA also reduces the growth of breast tumor xenografts and inhibits breast cancer cells invasion in the ECM[97]. Consistently, elafin treatment of mice with breast cancer xenografts leads to reduced tumor volume and increased survival[97].

Collectively, studies in NE, its inhibitor elafin, and the transcription factor C/EBPβ highlight the role of NE in driving tumor progression (Figure 6), in which NE promotes the

post-translational cleavage of FL-cycE protein to generate the LMW-E isoforms in breast cancer (Figure 5).
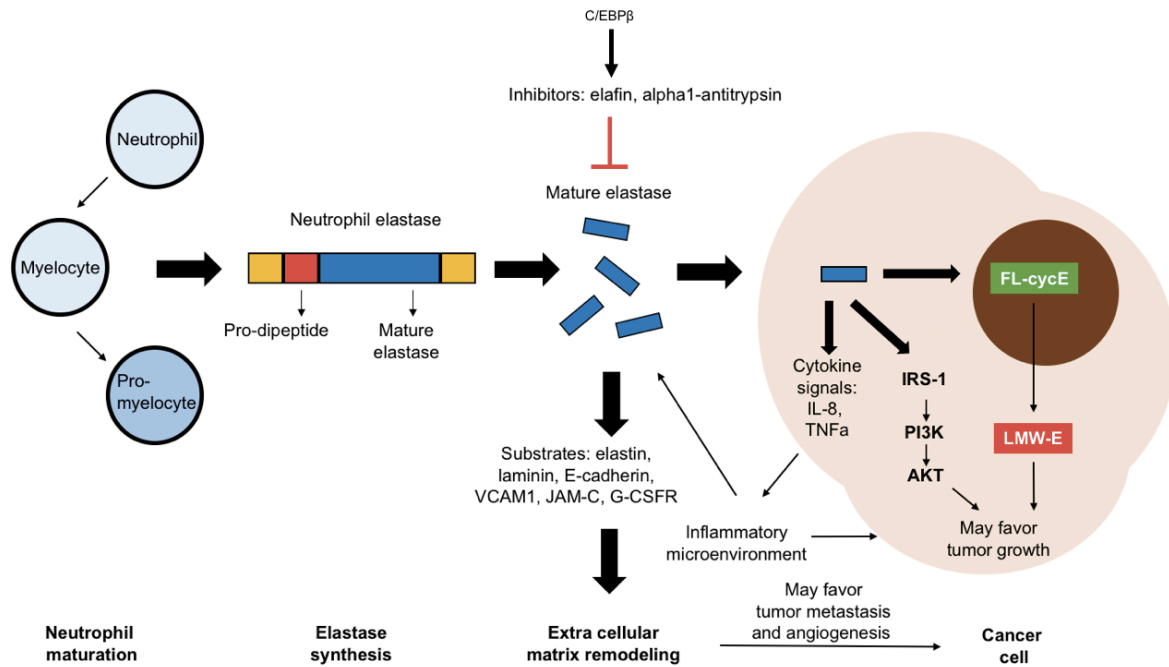
Figure 6. Neutrophil elastase and its role in tumor progression. Neutrophil elastase is synthesized during neutrophil maturation, activated by the cleavage of pro-dipeptide and balanced by elastase inhibitors such as elafin and alpha1-antitrypsin. Over-activated elastase may directly promote tumor cell growth by the generation of LMW-E from FL-cycE, activation PI3K-AKT pathway, and indirectly promote tumor development by inducing inflammatory microenvironment and extra cellular matrix remodeling, which favor angiogenesis and tumor metastasis.

1.6 Cyclin E deregulation and genomic instability

Genomic instability describes the rate and state of genomic alteration, which is generated from random mutations and chromosomal miss-rearrangements[68]. During cancer development, aberrant cell proliferation induces DNA damage, which in turn fuels genomic instability. Those cancer cells with genetic alterations are selected for the escape from apoptosis or resistance to drug treatment[69]. Certain mutations and genetic alterations confer

selective advantage on sub-clones of cells, further enabling their outgrowth and eventual dominance in a local tissue environment[69]. Studies in genetic changes in the cancer genome have revealed a group of cancer driver genes, which can be further sub-grouped into oncogenes and tumor suppressor genes[70].The former can transform cells when activated or overexpressed, while the latter transform cells by their inactivation (by point mutations or deletions) in human cancers[70]. Cyclin E is likely to be an oncogene due to its function to promote cell cycle progression and cancer cell proliferation. Cyclin E overexpression can also mediate drug resistance (both with targeted agents as well as chemotherapeutic agents), suggesting an oncogenic role for this protein that is independent of its role in cell proliferation[39,71].

Several tissue-specific transgenic and knock-in mouse models have provided significant information on the role of cyclin E deregulation in genomic instability. A knock-in mouse expressing a nondegradable form of cyclin E in mouse embryonic fibroblasts (MEFs) shows increased chromosome breaks, translocations, and aneuploidy in a $p21^{-/-}$ background[72]. In this model, cyclin E overexpression cooperates with $p53$ deficiency and Ras GTPase (RAS) activation to promote tumor transformation, with whole chromosome gains and losses accelerating lung carcinogenesis[72]. Additionally, transgenic mice expressing either wild-type or degradation-resistant cyclin E in the lungs incurred multiple pulmonary adenocarcinomas with specific gains of chromosomes 4 and 6[73]. In mammary gland transgenic mouse models, LMW-E overexpression induces p53 loss of function and drastically increase tumor formation in a $p53^{+/-}$ background[65]. Moreover, a knock-in mouse with expression of nondegradable cyclin E in the hematopoietic stem cell compartment exhibited abnormal hematopoiesis, chromosome instability evidenced by chromosome gains and losses, and decreased latency of T-cell malignancies in a $p53^{-/-}$ background[74,75]. These results strongly suggest deregulation of cyclin E is associated with enhanced genomic

instability that fuels cancer development, although the detailed mechanisms are not fully understood[74,76].

Hyper-activated cyclin E-CDK2 induces chromosome instability in cancers by mediating centrosome amplification during mitosis. In mammalian cells, CDK2 activity is required for centrosome duplication[77] . During mitosis, cyclinE-CDK2 is translocated to centrosomes[78], at which CDK2 phosphorylates and dissociates nucleophosmin (NPM) protein, which is required for the separation of paired centrioles and duplication of centrosomes[79]. In cancer cells with LMW-E over-expression, polo-like kinase 1 (PLK1) protein fails to interact with the central spindle during late mitosis. This leads to improper attachment of sister kinetochores to microtubules, and eventually causes chromosome segregation errors[80]. It is therefore hypothesized that the nuclear localization signal (NLS)-deficient LMW-E would be free to interfere with mitotic events through the cell cycle. Supporting this hypothesis, when breast cancer cells are engineered to over-express LMW-E, dramatic increases in the rate of chromosomal aberrations are observed[66,80]. LMW-E overexpressing cells show higher proportion of polyploid and tetraploid cells[66]. Additionally, expression of FL-cycE accelerates entry into mitosis and a delayed exit out of mitosis, which is characterized by a pronounced pro-metaphase delay and an increased rate of spindle defects. Cells expressing LMW-cyclin E also show an early entry into mitosis, but their exit from mitosis is also accelerated, associating with severe mitotic defects such as anaphase bridges and failed cytokinesis[81]. It is also proposed that impairment of mitotic progression could be resulted from CDK2 mediated phosphorylation CDH1[82,83] and CDC25[75,84].

Additionally, cyclin E overexpression is associated with deletion at DNA regions, which are annotated as under-replicated fragile sites[85,86]. This observation is in line with the concept of oncogene induced replication stress (further discussed in 1.7). Overexpression of oncogenes including cyclin E, as well as H-Ras and c-Myc, induce aberrant activation of DNA

replication[69]. During the initial stage of oncogene activated "proliferation signal", cells experience DNA hyper-replication stress featured by slowing and stalling of replication forks, followed by a "transition phase" in which DNA damage response is activated. In normal or pre-cancerous cells, this over-activated DNA replication signal is counteracted by ATR-CHK1 pathway to establish a cell cycle arrest and global stoppage of DNA replication (further discussed in section 1.8)[87,88]. If the stress is not removed or DNA damage resulting from fork collapse is beyond repair, the cells would reach the "senescent phase" or eventually die of mitotic catastrophe[89,90]. In cancer cells, the slow-down or stalling of replication fork potentially results in under-replicated genomic regions and sources for DNA damage[79], and thereby generates structural and numerical chromosome aberrations during mitosis[91]. It is currently presumed that defective cell cycle checkpoint and DNA damage responses contribute to the escape from replication stress induced tumor barrier, evidenced by loss of heterozygosity of tumor suppressors such as *p53*[65], but the detailed mechanisms of this process remain unclear.

1.7 DNA replication and replication stress

DNA replication is the most vulnerable cellular process during cell cycle progression, and its initiation, progression and termination are tightly controlled at different cell cycle phases[92]. In mammalian cells, during G1 phase, the origin recognition complex (ORC) acts together with the CDC6 ATPase and the CDT1 protein to load an inactive double hexamer of the MCM2-7 complex, the catalytic core of the replicative DNA helicase[93]. As cells enter S-phase, the CDK and the Dbf4-dependent kinase (DDK) promote helicase activation and replisome assembly by stimulating formation of CMG (Cdc45-MCM-GINS) complex, which unwinds duplex DNA in front of the moving replication fork[94,95]. Re-replication is prevented by inhibiting re-loading at origins of the MCM2-7 helicase core. This is achieved by CDK or DDK

mediated phosphorylation of MCM2-7 and its loading factors, together with the direct inhibition

of CDT1 protein by the Geminin protein (Figure 7)[95].
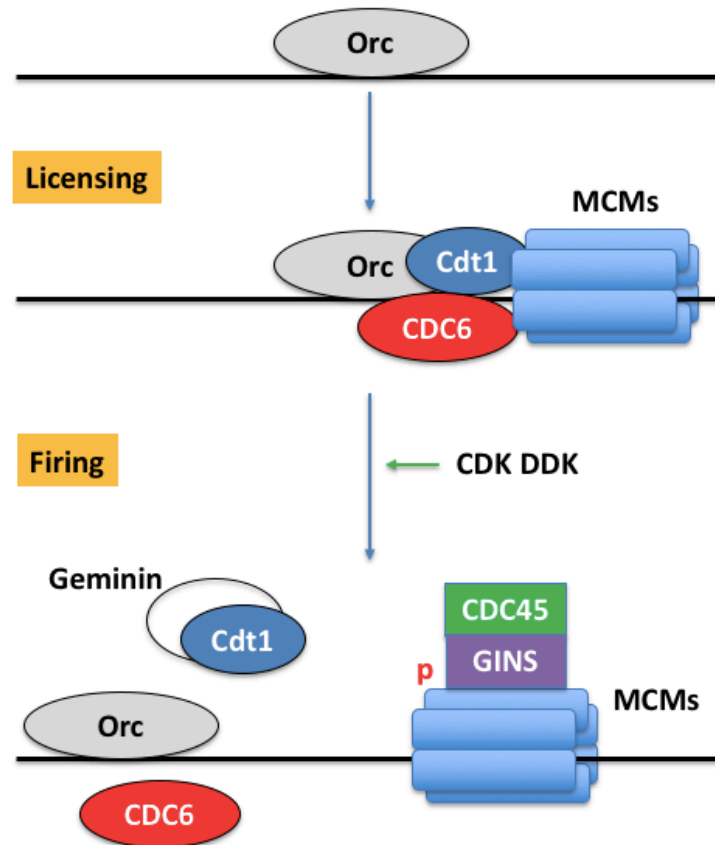
Figure 7. Schematic of DNA replication licensing and firing. DNA replication contains two steps, replication licensing and replication firing. During the licensing step, the origin recognition complex (ORC) acts together with CDC6 and CDT1 proteins to load an inactive MCMs complex. During the firing step, CDC45 and GINS complex are recruited to the MCMs complex, thereby facilitating the CDK and DDK mediated phosphorylation of MCMs. After replication firing, re-replication is inhibited by preventing the assembly of pre-replication complex on the replication origin.

Once replication has initiated, forks may undergo either transient pausing or a longer delay referred as fork stalling[88]. The replication complex usually remains associated with

stalled forks so that replication could restart once the obstacle that stops the replication fork has been removed[88]. However, if fork arrest remains persistent, it might lead to double strand break (DSB) in the DNA. It is proposed that in these cases, the replication complex might be disassembled and can restart in a CDK dependent manner[96]. Results from *in vitro* DNA replication experiments suggest cyclin A but not cyclin E facilitates DNA replication if replication complexes are already assembled on the DNA, while cyclin A activity also inhibits the assembly of new complexes (re-replication)[97]. During the replication licensing process, chromosomes have a distribution of licensed origins that exceeds the number of active replication origins in S-phase[88,98]. These licensed origins remain dormant during the S phase but constitute a backup group that can be activated to compensate for collapsed forks under replication stress [88,98]. Reduced efficiency of origin licensing is a major source of genome instability. For example, yeast cells that lack the CDK inhibitor *Sic1* initiate replication from fewer origins, increasing the distance between replication forks. These mutants have an extended S-phase, accumulate ssDNA, and show a strong increase in chromosome loss[99].

1.8 Cyclin E induces replication stress

Cyclin E overexpression are known to induce replication stress, but the detailed mechanisms and whether it is different between the FL-cycE and LMW-E remain unclear. Cyclin E-mediated replication stress is linked to elevated CDK2 kinase activity, as a CDK2 hyperactive knock-in allele was sufficient to delay replication fork progression and further induce DNA damage[84]. Interference of the nucleotide biosynthesis is one of the proposed mechanisms underlying cyclin E-CDK2 induced replication stress and genomic instability[100]. Cyclin E overexpression, through disruption of the RB-E2F pathway, enforces G1/S transition and cell proliferation with insufficient nucleotide pools. These nucleotides are crucial structural components of nucleic acids for DNA replication and transcription[100,101]. DNA replication without sufficient nucleotide supply results in fork stalling and under-replicated DNA regions.

Collapse of stalled forks lead to double-strand DNA breaks, for which the repair process is also blocked by the nucleotide deficiency induced by cyclin E overexpression. Reports have shown either exogenous supplementation of nucleosides or upregulation of nucleotide metabolism genes can attenuate cyclin E-mediated replication stress and rescue the DNA damage[100].

It is also proposed that cyclin E overexpression initiate unscheduled DNA firing, and the conflict between DNA replication and transcription causes replication obstacles. The role of transcription on oncogene-induced DNA replication stress was directly examined in a study using U2OS cells overexpressing cyclin E[102]. Short term treatment of these cells with cordycepin, an RNA-specific chain terminator that inhibits transcription elongation, can rescue the slow fork progression and reduce the induced DNA double strand breaks. This is consistent with the observation that the majority of under-replicated regions in cyclin E overexpressing cells are annotated as large late-replicating domains[85].

Additionally, most of the under-replicated sites are located in chromosomal regions that are characterized as extremely low origin density[85]. It is therefore hypothesized that cyclin E may induce replication stress by de-regulating DNA licensing. This is supported by the observation that overexpression of cyclin E in a human nasopharyngeal epidermoid carcinoma cell line leads to reduced association of the MCM helicase subunits Mcm4 and Mcm7 with chromatin during G1 and reduced DNA synthesis in S phase[103]. However, other reports show cyclin E deficient cells are incapable of loading MCMs to the chromatin when exiting the G0, suggesting cyclin E plays a positive role to promote MCM subunits loading[104] (Figure 8). It is currently unknown whether and how different cyclin E isoforms regulate each step of DNA replication, and the effect of FL-cycE and LMW-E on replication stress is the main subject of this thesis.
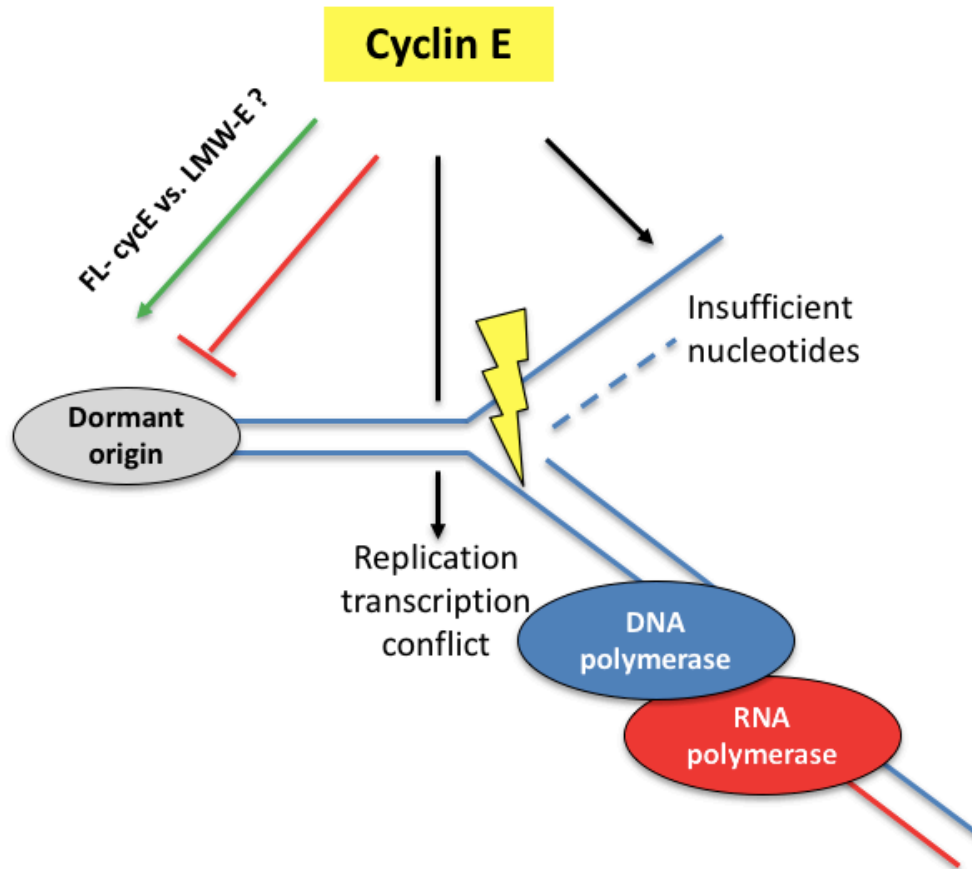
Figure 8. Proposed mechanisms underlying cyclin E induced replication stress. In cells overexpressed with cyclin E, replication fork slowing and stalling may result from insufficient nucleotides supply, unscheduled replication firing and conflicts between DNA synthesis with RNA transcription. Additionally, cyclin E is shown to both positively and negatively regulate replication licensing, a topic to be experimentally examined in this study to compare the functions of FL-cycE and LMW-E.

1.9 Replication stress response and therapeutic targets

As discussed in section 1.5, the overexpression of oncogenes such as cyclin E induce aberrant activation of DNA replication that are counteracted by cell cycle checkpoints[87,88]. In response to under-replicated DNA, replication protein A(RPA) is recruited to the exposed

ssDNA, and thereby activates cell cycle checkpointing. This may lead to cell cycle arrest in different cell cycle phases, allowing time for the DNA damage repair process[105,106].

When DSB occurs, the serine/threonine-protein kinases ATR and ATM protein kinases phosphorylate and activate CHK1[107]. Protein kinase CHK2 is mainly activated by ATM[107]. The activated CHK1 and CHK2 further phosphorylate p53 at multiple sites, which in turn increase the stability of p53 and promote its activation at p53-target promotes[106]. One of the major targets that is transcriptionally induced by *p53* is *p21*, the endogenous inhibitor for cyclin E-CDK2 and cyclin A-CDK2. Induction of *p21* inhibits CDK2 activity, which not only restores the repressive function of pRB toward E2F transcription factors, thereby leading to G1 cell cycle arrest, but also attenuates CDK dependent replication fork firing so that the global DNA replication in S phase is blocked[74,108,109].

Under replication stress, fork stalling results in the exposed ssDNA and the coating of RPAs. The ssDNA binding RPA proteins recruit ATR and CHK1 proteins to the stalled forks, where CHK1 is activated by ATR[107,110]. The activated ATR then perform a dual function: to stop global DNA replication by cell cycle arrest and to promote local DNA synthesis at the site of DNA replication stress by facilitating fork restart [50,84]. The detailed mechanisms of this process are not yet clear, but it is generally believed that activated CHK1 mediates temporary S phase arrest by inhibiting CDK2 activity. This is fulfilled by indirectly regulating T15 phosphorylation at CDK2, which is phosphorylated by WEE1 kinase and dephosphorylated by the CDC25 phosphatase family[84]. CHK1 directly phosphorylates WEE1 kinase and promotes its activity in phosphorylating CDK2 at T15[111]. Additionally, CHK1 also phosphorylate CDC25A,B, and C, reducing their activities and ultimately preventing CDC25 mediated dephosphorylation of T15 on CDK2[111,112]. Thus, recruitment and activation of CHK1 inhibit global DNA replication in response to replication stress[105,113]. On the site of replication stress lesions CHK1 phosphorylates RPA proteins and recruits RAD51 protein, which

replaces RPA to bind with ssDNA[90]. RAD51 functions as the central recombinase to mediate fork restart by at least three mechanisms[90,114-116]. (1) RAD51 promotes replication fork uncoupling and reversal, and facilitates DNA branch migration in the direction opposite to replication to form a Holliday junction[90]. (2) RAD51 protects newly synthesized DNA from degradation by forming a stable RAD51 filament, which also protects the regressed forks from extensive degradation[117]. (3) RAD51 is also required for the restart of the stalled replication fork without triggering homologous recombination (HR)[115], and repair of collapsed fork by HR mediated double strand break repair[116]. Therefore, in response to replication stress, activation of ATR-CHK1-RAD51 facilitates the restarting of stalled replication forks (Figure 9). It is currently unknown how the global and local DNA replication are influenced by LMW-E compared to FL-cycE, and what are the roles of ATR, CHK1, RAD51 molecules in response to high FL-cycE or LMW-E induced replication stress.

For cancers with a very unstable genome, such as TNBC[118], high-grade ovarian cancers[119], and those with oncogene-induced replication stress[119], targeting of the replication stress responsive mechanisms have shown promising therapeutic effects. It is worth noting that while replication stress is a source of genomic instability and potentially promotes cancer development, tumor cells are required to tolerate replication stress and maintain minimum genomic integrity for cell viability[119]. As a consequence, agents targeting replication stress responses are likely to further increase replication stress in cancer cells, which eventually lead to the loss of minimum genomic integrity and cell death.

There have been four ATR inhibitors that have progressed to clinical trials in different cancers: berzosertib, ceralasertib, M4344, and BAY1895344 (Table 1)[120]. In pre-clinical studies, all of these inhibitors have shown radio-sensitizing and chemo-sensitizing effect[120,121]. Reports show M4344 and BAY1895344 suppress the proliferation of DU145 prostate cancer at lower concentrations than berzosertib and ceralasertib[121]. Screening of protein expressions suggest

DNA damage response proteins CHK1, MSH2, and RAD51 positively correlate with response to ATR inhibitors [121]. ATR inhibitors are synergistic with DNA damage inducers such as TOP1 inhibitor and PARP inhibitor. These combination treatments result in suppressed DNA replication, increased DNA damage and chromosomal fragmentation and cell death[121,122].
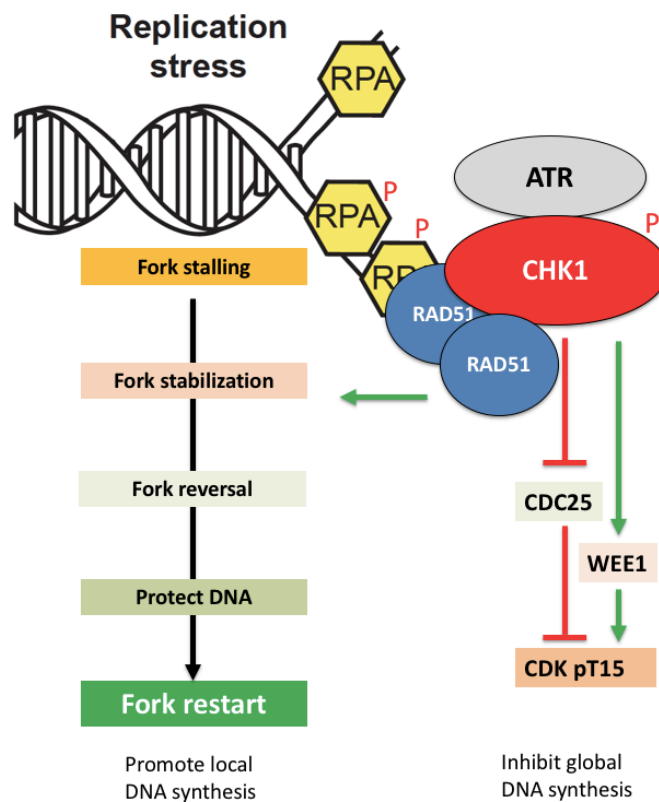
Figure 9. Replication stress activates ATR-CHK1-RAD51 pathway. Under replication stress, the exposed ssDNA is bound with RPAs, which in turn recruit ATR and CHK1 protein kinases to the stalled replication fork. Signaling cascade by ATR mediated CHK1 phosphorylation and CHK1 mediated RPA phosphorylation further induce the binding of RAD51 to promote fork restart and DNA synthesis at replication stress lesion, while global DNA synthesis and cell cycle progression are inhibited by CHK1 to prevent cell division with under-replicated DNA.

As the down-stream effector of ATR, when CHK1 is inhibited by small molecule inhibitors, such as MK-8776 and rabusertib, strong anti-tumor effect is observed[123,124]. Treatment of cancer cells with MK-8776 cause accumulation of DNA DSB, leading to apoptotic cell death *in vitro*[124]. Studies also show CHK1 inhibitor synergizes with gemcitabine, hydroxyurea and cytarabine in causing apoptosis of acute myeloid leukemia (AML) and breast cancer cells *in vitro*, as well as in ovarian and pancreatic cancer xenografts[122]. On the basis of these studies, the first clinical phase I trial with MK-8776 in combination with gemcitabine was initiated for patients with advanced solid tumors (Table 1). The trial suggested MK-8776 is well tolerated, and 43% of the patients in the trial showed stable disease, even in patients previously developed resistance to gemcitabine[125]. Another phase I clinical trial investigated the sequential administration of cytarabine and MK-8776 in patients with relapsed or refractory AML. Report suggest this combination achieves complete response in 33% of patients[126]. Rabusertib (LY2603618) is a recently developed inhibitor with higher selectivity for CHK1 than CHK2 (Table 1). *In vitro* studies suggest rabusertib activates CDC25A in cancer cells, leading to increased CDK2 activity[127]. The activation of the CDC25A-CDK2 axis promotes S phase progression with an increased number of replication forks, resulting in replication catastrophe caused by DNA DSB at stalled replication forks, followed by chromosome fragmentation and eventually mitotic cell death[127]. Rabusertib also reduces tumor growth in a xenograft model of lung cancer[128]. RAD51 is recruited to the replication stress lesion by activated CHK1, participating in replication fork remodeling and fork restart[90]. Currently, only one RAD51 inhibitor CYT-0851 has developed into clinical trial (Table 1). Preliminary result showed partial responses in 3 of 21 evaluable patients, and 10 more patients exhibited stable disease[129]. The specific clinical trials for these agents are listed in Table 1, and the status of the each of these trials are also indicated.

Agent	Target	Condition or disease	NCT number	Description	Phase
Berzosertib	ATR	Advanced solid tumor	NCT02157792	Combine with cytotoxic chemotherapy	I
		Ovarian, peritoneal, or fallopian tube cancer	NCT02627443	Combine with gemcitabine and carboplatin	II
		Metastatic urothelial cancer	NCT02567409	Combine with gemcitabine and cisplatin	II
		Small cell and extrapulmonary small cell cancer	NCT02487095	Combine with topotecan	I/II
		Ovarian, peritoneal, or fallopian tube cancer	NCT02595892	Combine with gemcitabine	II
		lung cancer brain metastases or neuroendocrine tumor	NCT02589522	Combine with radiotherapy	I
		Head and neck cancer	NCT02567422	Combine with chemotherapy and radiation	I
		Advanced solid tumor	NCT02595931	Combine with Irinotecan	I
		Refractory solid tumor	NCT02723864	Combine with veliparib and cisplatin	I
Ceralasertib	ATR	Advanced solid tumor	NCT02264678	Combine with carboplatin, olaparib, or durvalumab	I/Ib
		Refractory solid tumor	NCT02223923	Single/combine with radiotherapy	I
		Leukemia	NCT01955668	Single agent	I
		Refractory solid tumor	NCT02630199	Combine with paclitaxel	I
M4344	ATR	Breast cancer	NCT04655183	Combine with niraparib	I/II
		Advanced solid tumor	NCT02278250	Single/combine with carboplatin	I
BAY189534	ATR	Advanced solid tumor and lymphoma	NCT03188965	Single agent	I
MK-8776	CHK1	Relapsed acute myeloid lymphomas	NCT01870596	Combine with chemotherapy	II
		Solid tumor or lymphoma	NCT00779584	Single/combine with chemotherapy	I
Rabusertib	CHK1	Cancer	NCT00415636	Single agent	I
		Pancreatic cancer	NCT00839332	Single/combine with gemcitabine	I/II
		Non-small cell lung cancer	NCT00988858	Combine with pemetrexed	II
		Non-small cell lung cancer	NCT01139775	Combine with gemcitabine and pemetrexed	I/II
		Non-small cell lung cancer	NCT01358968	Combine with cisplatin and pemetrexed	I/II
		Solid tumor	NCT01341457	Combine with gemcitabine	I
CYT-0851	Rad51	Refractory B-cell malignancy and advanced solid tumor	NCT03997968	Single/combine with chemotherapy	I/II

Table 1. Summary of clinical trials for small molecule inhibitors of ATR, CHK1 or RAD51.

Considering the hyper-activated CDK2 activity and stronger tumorigenesis capacity observed in LMW-E overexpression models, we hypothesize that replication stress is induced but tolerated in LMW-E overexpression setting. Targeting of the replication stress responsive mechanisms may have promising therapeutic effects in LMW-E overexpressing breast cancers.

1.10 Gaps of knowledge

Current understanding of replication stress has established its role in promoting genomic instability, which fuels cancer initiation and/or development by providing the sources of genomic alterations. Replication stress can be induced by overexpression of oncogenes such as cyclin E (i.e. the LMW-E). In pre-cancerous lesions, replication stress is a tumor barrier by inducing DNA damage and triggering cell cycle arrest. Bypassing this barrier via replication stress tolerance may benefit cancer development with enhanced genomic instability. However, differences in the effect of FL-cycE and LMW-E, particularly on replication stress and subsequent DNA damage throughout the cell cycle is currently unknown. Considering the stronger tumorigenesis capacity observed in LMW-E overexpression compared to FL-cycE, in both *in vitro* and *in vivo* models, whether and how replication stress is tolerated in LMW-E overexpression setting is to be investigated in this study (Figure 10).

1.11 Hypothesis and specific aims

I hypothesize that LMW-E, as a cyclin E isoform resistant to p21/p27 inhibition and complete degradation by the end of S phase, may function as an oncogene in disrupting the temporal regulation of cell cycle transition and RB-E2F mediated gene transcription. These deregulations may further affect a number of different but related pathways such as DNA replication, replication stress responses, DNA damage repair, and cell cycle checkpoints. As a consequence, LMW-E may fuel oncogenic transformation of mammary epithelial cells by enhanced replication stress and facilitating replication stress tolerance (Aim 1), error-prone

DNA damage repair (Aim 2) thereby leading to enhanced genomic instability in pre-cancerous mammary cells, and predict genomic instability in breast tumor tissues (Aim 3; Figure 11).
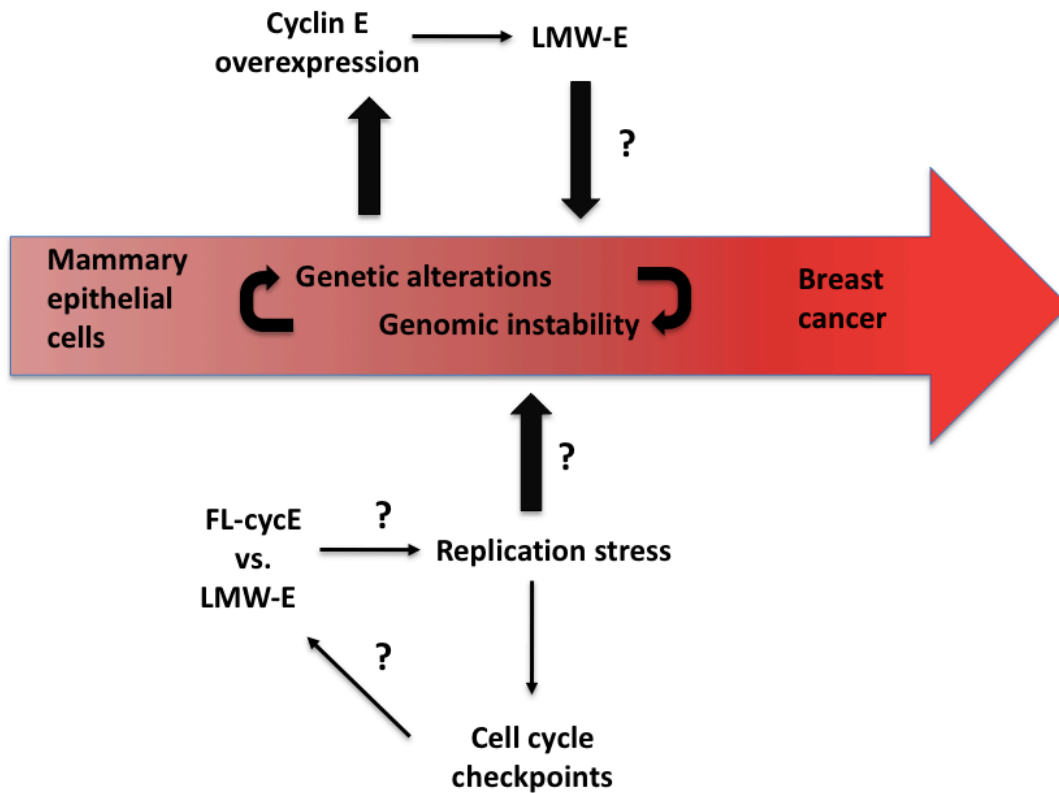
Figure 10. Gaps of knowledge. Current model suggests genetic alterations fueled by genomic instability drives cancer transformation from mammary epithelial cells. LMW-E generated from overexpressed cyclin E is associated with poor prognosis of breast cancer patients. However, the connection between LMW-E and genomic instability and the differences in the roles of FL-cycE and LMW-E, particularly on replication stress and the responses to the cell cycle check points are currently unknown.
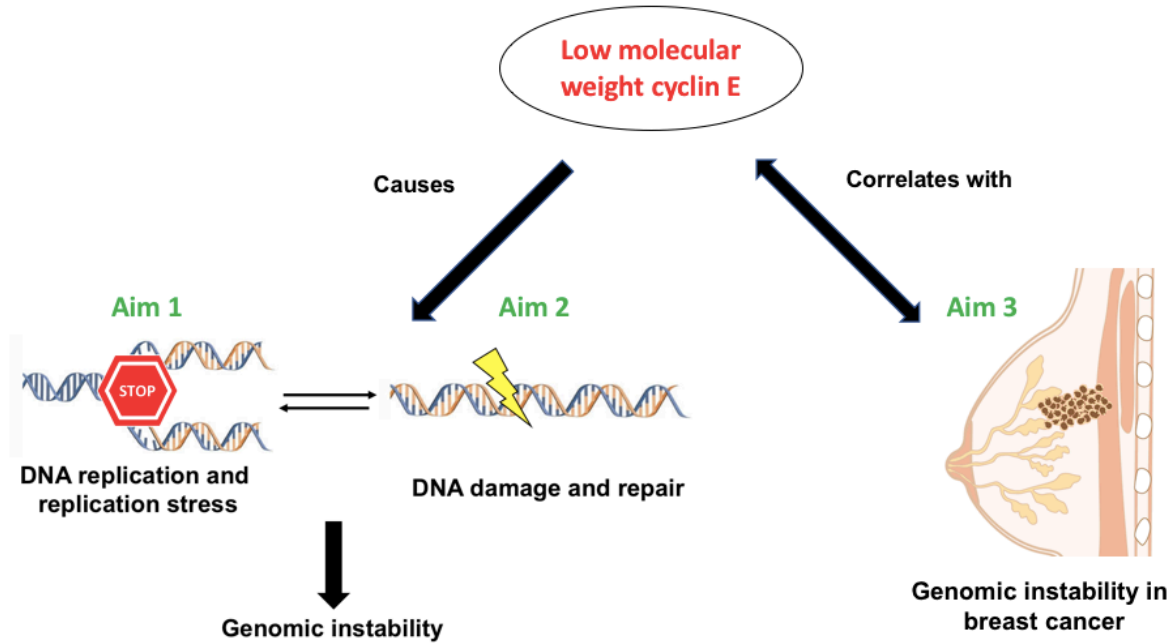
Figure 11. Specific aims. In this study, we aim to investigate the roles of low molecular weight cyclin E (LMW-E) in DNA replication and replication stress (aim1), examine the effect of LMW-E on DNA damage and damage repair (aim2), and test if LMW-E overexpression drives genomic instability in mammary epithelial cells (associated with aim1 and aim2) and if LMW-E positive status predicts genomic instability in human breast cancers (aim 3).

Chapter Two: Low molecular weight cyclin E promotes genomic instability

2.1 Introduction

In normal dividing cells, cyclin E and its associated kinase CDK2 function as the gate keeper for G1/S transition and the initiation of DNA replication[109]. Normal cells suppress the effects of excess/stabilized cyclin E via the G1/S checkpoint, which blocks the global DNA replication and G1/S transition by the ATR-CHK1 pathway, safeguarding the genomic integrity[109,112]. In cancer cells, deregulated cyclin E-CDK2 activity may result from the overexpression of cyclin E and/or loss of CDK2 inhibition (via *p53* mutation or the loss of *p21*), promoting cell proliferation and genomic instability[71]. Additionally, oncogenic LMW-E isoforms, ranging in molecular size from 33 to 44 kDa (Figure 5), have been identified in tumors overexpressing cyclin E[63]. These LMW-E are generated through post-translational cleavage of the 50kDa FL-cycE by the elastase family of serine proteases in tumor cells[63]. Distinct from its full-length counterpart, LMW-E is predominantly detected in tumor tissues but not their adjacent normal tissues[53]. Ectopic expression of LMW-E isoforms promotes cell proliferation and early entry into mitosis with de-regulated centrosome amplification[80]. LMW-E also hyper-activates CDK2 and the LMW-E-CDK2 kinase complex is resistant to cellular endogenous CDK inhibitors such as p21 and p27[66].

In transgenic mouse models, LMW-E drives early oncogenic events in the pre-neoplastic mammary glands, leading to hyperplastic lesions and spontaneous mammary tumors with high metastatic capacity[39,65]. Targeted sequencing of the top 70 breast cancer genes in LMW-E overexpressing mammary glands in a time course dependent manner (0, 3 and 6 months after LMW-E over-expression) revealed mutations of several genes in the non-tumorigenic mammary glands (data not shown). LMW-E included mutations in the following

7 genes at 3 months post LMW-E expression: *Alk*, *Atm*, *Rad50*, *Ctnna1*, *Akap9*, *Smg1*, and *Tecta*, and an additional 6 genes at 6-9 months post LMW-E induction: *Brca1*, *Brca2*, *Rb1*, *Sbno*, *Nrk*, and *Huwe1*. Mice generated tumors at 9-12 months post LMW-E induction (manuscript in preparation). These observations indicate LMW-E overexpression leads to increasing genetic alterations in mammary tissues, suggesting that LMW-E may drive tumor progression by promoting genomic instability. An unfilled gap of knowledge is that our current understandings of the role of LMW-E in mediating genomic instability is based on model systems with endogenous FL-cycE expression. Whether or not LMW-E independently drives genomic instability and/or performs other unique oncogenic functions apart from FL-cycE remain unclear.

In this chapter, we describe the generation of an *in vitro* model system where the endogenous *CCNE1* is knocked out and the FL-cycE and LMW-E are then inducibly expressed. To this end, we have generated the *CCNE1* knock-out human mammary epithelial cell (hEMC) models using 76NE6 (p53 deficient) and 76NF2V (p53 proficient) cell lines, and transfected the cells with a doxycycline inducible system (tet-on[130]) to express FL-cycE (aa1-410) or LMW-E (aa40-410). Our aim in this chapter is to discover distinct features of LMW-E and FL-cycE in regulating the growth of hMECs and genomic instability. We compare the effects of these different cyclin E isoforms on cell proliferation, cell viability, cell cycle distribution, DNA damage, chromosomal and nuclear abnormalities (Figure 12).

2.2 Schematics of model system
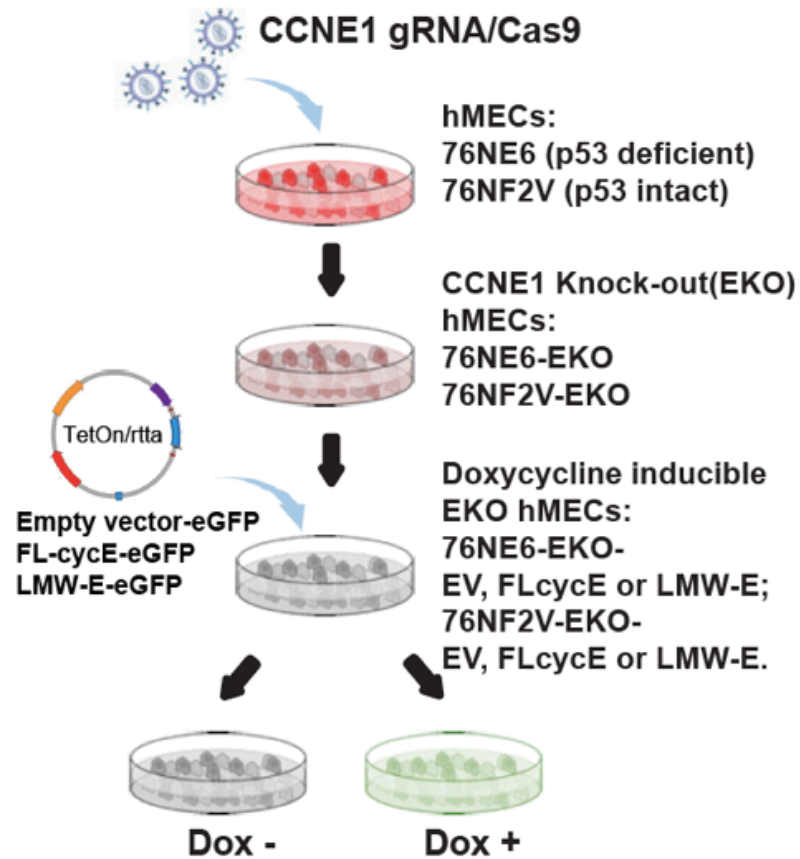
Figure 12. Schematics of the cellular models containing the doxycycline inducible FL-cycE or LMW-E (tet-on) in *CCNE1* knock-out background. In human mammary epithelial cell line 76NE6 and 76NF2V, endogenous *CCNE1* gene is depleted through gRNA/CRISPR mediated gene knock-out, followed by the generation of stable cell lines containing the doxycycline inducible (tet-on) expression of empty-vector, FL-cycE or LMW-E.

2.3 Materials and methods

Cell lines and Culture Conditions

We used the immortalized human mammary epithelial cell lines (hMECs) 76NE6 and 76NF2V, which were originally generated and provided by Dr. V. Band[131]. In brief, HPV-16 E6 was introduced into mammary epithelial cell line 76N to generate the immortalized strain 76NE6, and an E6 mutant F2V with compromised p53 association capacity (2% of wild-type E6) was used to generate 76NF2V line. As a result, the 76NE6 cell line lacks p53, and the 76NF2V expresses p53[131]. 76NE6 and 76NF2V and the subsequently generated EKO and/or inducible cell lines (see below for details on the generations of these models) were maintained in DCFI-1 media[64]. This media contains a-MEM/Ham's nutrient mixture F-12 (1:1,vol/vol), and is supplemented with 12.5ng/mL epidermal growth factor (EGF),10nM triiodothyronine, 10mM Hepes, 50 μM freshly made ascorbic acid, 2nM estradiol, 1ug/mL insulin, 2.8 μM hydrocortisone, 0.1mM ethanolamine, 0.1mM phosphor-ethanolamine, 10μg/mL transferrin, 2mM L-glutamine, 1% fetal bovine serum (Hyclone); and 35 μg/mL bovine pituitary extract (Hammond Cell/Tech). All cell lines were cultured in a humidified tissue culture incubator at 37°C and 5% CO2 and routinely tested for mycoplasma and authenticated. There are enough frozen vials for each cell line to ensure that all cell-based experiments are performed on cells that have been tested and in culture for 8 week or less.

CCNE1 knock-out experiments

In order to knock-out *CCNE1* gene in 76NE6 and 76NF2V cell lines, we first generated CRISPR/sgRNA constructs harboring Cas9 gene and sgRNA targeting human *CCNE1* gene. The sgRNAs targeting *hCCNE1* was provided by the Toronto Knockout Library (http://tko.ccbr.utoronto.ca/crispr_targets.pl) and cloned into pX330 vector. After transfection into 76NE6 or 76NF2V cells, single cell clones were isolated and expanded. Successful knock-out of *CCNE1* (EKO) in 76NE6-EKO or 76NF2V-EKO lines are confirmed by Sanger

sequencing for sgRNA targeting site (Figure 13A) and western blotting using anti-cyclin E antibody (Figure 13B).

Inducible cyclin E expression model system

We generated the doxycycline inducible (tet-on) cells from 76NE6-*CCNE1*knock-out (EKO) cell line and 76NF2V-EKO cell line. Empty-vector, FL-cycE and LMW-E doxycycline inducible (tet-on) cells were established using the pLVX-TRE3G-C-EGFP vector[38], derived from pLVX-TRE3G (Clontech Laboratories, Inc.). The constructs were co-transfected with pCMV-deltaR8.9 and pMD2.G-VSVG plasmids into HEK-293T cells for packaging lentivirus. The packaged lentivirus was then introduced into 76NE6-EKO or 76NF2V-EKO cell lines, followed by selection of cells surviving puromycin (1ug/mL) containing media. After brief induction (12 hours) with doxycycline (1μg/mL), EGFP high clones (top 10%) were sorted by fluorescence-activated single cell sorting (FACS) into 96 well plates and expanded in tetracycline free DCFI-1 media. Doxycycline induced expression of FL-cycE and LMW-E were confirmed by FACS detecting LMW-E or FL-cycE fused EGFP (Figure 15), western blot signals of LMW-E and FL-cycE in total cell lysates (Figure 16) and cyclin E associated kinase activity using pRB as substrate (Figure 17).

Western blot analysis

For western blot analysis of protein expression, total cell lysates were prepared in ice-cold lysis radio-immuno-precipitation assay buffer (RIPA) buffer [50mM Tris-HCl pH 7.6, 150mM NaCl, 1% Triton X-100, 0.5% sodium deoxycholate, 0.1% Sodium dodecyl sulfate (SDS) and mixed with fresh protease inhibitor cocktail] and placed on ice. After 15 minutes of lysis, brief sonication was performed followed by centrifugation at 13000rpm at 4° C for 15 minutes. Protein concentration were measured by bicinchoninic acid (BCA) protein assay (Bio-Rad protein assay kit #5000001). Samples were diluted to 1μg/μl in 1x loading buffer (50 mM Tris-HCl pH 6.8, 2% SDS, 10% glycerol, 1% β-mercaptoethanol, 12.5 mM EDTA, 0.02 %

40

bromophenol blue). For SDS-polyacrylamide gel electrophoresis (SDS-PAGE), each well/sample is loaded of 40μg protein and run by electrophoresis. Samples on the SDS-PAGE were then transferred to polyvinylidene fluoride (PVDF) membranes by wet-transfer (Bio-Rad). The membranes were blocked in 5% bovine serum albumin (BSA) in tris-buffered saline (20 mM Tris and 150 mM NaCl, pH 7.6) with 0.1% Tween 20 detergent (TBST) for 30 minutes, followed by overnight incubation of primary antibody at 4º C. In this chapter, we used the following primary antibodies: cyclin E (HE-12; Santa Cruz Biotechnology, SC-247, diluted 1:1000 in TBST containing 1% BSA) and Vinculin (Sigma, V9131, diluted 1:5000 in TBST containing 1% BSA). Membranes were washed four times in TBST for 10 minutes each. Horseradish peroxidase (HRP)-conjugated secondary antibodies (diluted 1:5000 in TBST) incubation was performed for 1 hour at room temperature. The membranes were then washed again four times in TBST for 10 minutes each, and subjected to brief treatment of enhanced chemiluminescence (ECL) reagents (Pierce). The signals on the membranes were then exposed to X-ray film. Densitometry of the signals was performed by using ImageJ software, the value for each signal band was normalized to the first visible band on the same gel.

Protein kinase assays

For FL-cycE and LMW-E kinase assays, 300μg of cell extracts was used per immunoprecipitation with polyclonal antibody to cyclin E (generated in lab[132]). The immuno-precipitates were then incubated with kinase assay buffer containing 60 μM cold ATP, 5 μCi of [32P]ATP and 1μg of GST-RB(Santa Cruz Biochemicals) in a final volume of 30 μl and incubated at 37°C for 30 min. The products of the reaction were analyzed on 10% SDS-PAGE gels that are exposed to X-ray film. Densitometry of the signals was performed by using ImageJ software, the value for each signal band was normalized to the first visible band on the same gel.

Growth Curves and Cell Viability Assay

To measure cell proliferation, inducible 76NE6-EKO cell lines or 76NF2V-EKO cell lines engineered to express empty-vector, FL-cycE and LMW-E upon doxycycline treatment were plated in 96 well plate (1500 cells per well). After cell confluency reached 15%, which usually take 24 hours incubation in DCFI-1 media supplemented with tetracycline-free FBS. The cells were then treated with doxycycline in a dose titration manner (0 to 100ng/mL) for a four-day experiment. Fresh media containing doxycycline or vehicle (DMSO) was replaced every 48 hours. Every 24 hours, cell confluency was determined using the Incucyte S3 Live-Cell Analysis System.

Cell viability was measured by the 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl-2H-tetrazolium bromide (MTT) assay as previously reported[64]. Similar to cell proliferation assay, for each of the inducible cell lines (76NE6-EKO-empty vector, FL-cycE or LMW-E, and 76NF2V-EKO-empty vector, FL-cycE or LMW-E), 1500 cells per well were plated on 96 well plate. At the end of the 4-day experiment, MTT assay was performed by adding 2.5 mg/mL MTT in serum-free media and culture for 4 hours at 37°C in dark. Then, the supernatant was removed, and the MTT formazan precipitates were dissolved in by adding 100µl dimethyl sulfoxide (DMSO) to each of the wells. After shaking the plates for 1 hour at room temperature in dark on a horizontal shaker, the absorbance for each well was quantified using a spectrophotometer (Victor3, Perkin-Elmer) for the absorbance at wavelength of 590 nm.

Cell cycle analysis and BrdU incorporation assay

Inducible 76NE6-EKO cell lines or 76NF2V-EKO cell lines, engineered to express FL-cycE or LMW-E upon doxycycline treatment were plated in p100 tissue culture plates (10,000 - 15,000 cells per mL). After the cell confluency reached 15 to 30%, which usually take 24 to 48 hours culturing in DCFI-1 media supplemented with tetracycline-free FBS. The cells were then treated with doxycycline (100ng/mL) for 12, 24, 36 or 48 hours before harvesting for cell

cycle analysis. 48 hours of DMSO treatment was used as uninduced (mock) control. Prior to cell harvesting, the cells were pulse labeled with 50 µM bromodeoxyuridine (BrdU; Sigma-Aldrich) for 30 minutes and then prepared as single-cell suspensions of 100,000 cells per mL in PBS. The cells were fixed with ice-cold 95% ethanol overnight at 4°C, and then incubated with Anti-BrdU antibody (1:1000; BD Biosciences 347580) overnight at 4°C, followed by incubation in Alexa Fluor 660 (1:1000; EMD Millipore) for 2 hours at room temperature. DNA staining is performed by staining with propidium iodide (PI; Sigma-Aldrich) and RNase treatment overnight. Cell cycle profiles were examined by Attune NxT Flow Cytometer and the data analyzed was performed by FlowJo software (v. 10.5.3) with gating strategies indicated in Figure 21 - 24.

Immunofluorescence staining

Immunofluorescence (IF) staining was performed based on general immunofluorescence protocol provided by Abcam (https://www.abcam.com/protocols/ immunocytochemistry-immunofluorescence-protocol). In brief, cells were seeded into 6-well plates containing glass cover slides in each well (15,000-25,000 cells per well). After the time course induction of cyclin E isoforms described in main text (page 59), the cells were fixed by incubating in 100% methanol (chilled at -20°C) for 15 min, and then permeabilized by incubating with PBS containing either 0.2% Triton X-100 for 15 min. The slides with fixed cells were blocked by incubating with 2.5% BSA, 22.52 mg/mL glycine in PBST (PBS+ 0.1% Tween 20) for 2 hours at room temperature. The cells were then incubated overnight at 4°C with the primary antibody against γ-H2AX (Millipore, 05-636) and antibody against 53BP1 (Novus, NB100-304) at 1:1000 dilution in BSA-PBST (PBS containing 0.1% Tween 20 and 2.5% BSA). After being washed three times in PBST, the cells were treated with secondary goat anti-mouse or goat anti-rabbit antibody (Alexa Fluor 594 or 660 EMD Millipore) diluted at 1:1000 and incubated at room temperature for 2 hours. Following washes with PBST for two times

and PBS for two additional times, the slides were mounted with fluorescent mounting medium with DAPI (Dako). The images were taken by Zeiss LSM880 Confocal (63X objective magnification) and examined by Zeiss Zen software. Cells with five or more foci were considered positive.

Comet assay

The alkaline comet assay (single cell gel electrophoresis) were performed by using Comet Assay Kit (Enzo Life Science, ADI-900-166) according to manufacturer's protocol. 1×10^4 cells were resuspended in melted agarose at 37℃ and spread on the slide. After 10 minutes incubation at 4℃ in dark, the slides were treated with prechilled Lysis Solution at 4℃ for 30 minutes, and then in Alkaline Solution (NaOH, pH>13) at 4℃ for 60 minutes. Alkaline electrophoresis was performed at 300 mA for 30 minutes and Cygreen Dye (Enzo Life Science) were used to stain the DNA. CometScore software were applied for data processing and tail moment quantitation. (http://rexhoover.com/index.php?id=cometscore). Three biological repeating experiments were performed and more than 100 cells were counted.

Metaphase spread analysis

Each inducible 76NE6-EKO cell lines (empty vector, FL-cycE or LMW-E) were plated on p100 mm plates (10,000 - 15,000 cells per mL). Doxycycline was added to a final concentration at 100ng/mL in the culture media to induce the expression of LMW-E, FL-cycE or empty vector control. DMSO was used for un-induced (mock) controls. Culture media were refreshed every other day. At 48 hours post induced expression of LMW-E, FL-cycE or empty vector, the cells were treated with Colcemid (0.04 µg/mL) for 3 hours at 37 °C, followed by treatment to 0.075M KCl for 15 minutes at room temperature. They were then fixed in mixture of methanol and acetic acid (3:1 v/v) and washed three times in fixative. The slides were air-dried, stained in 4% Giemsa and analyzed for chromosomal aberrations using a Nikon 80i

microscope. The assessment of chromosomal aberrations was performed by Cytogenetics and Cell Authentication Core Facility at MD Anderson.

2.4 Results

2.4.1 Generation of *CCNE1* knock-out human mammary epithelial cell (hMEC) lines

76NE6 and 76NF2V cell lines were transfected with pX330-Cas9-*CCNE1*sgRNA expression construct targeting *CCNE1* gene exon5. 48 hours post transfection, the cells were sorted into single cells and plated in each well of 96 well plates (one cell per well). Individual colonies were expanded for screening. Genomic DNA was isolated and amplified by high fidelity PCR, followed by Sanger sequencing of exon5 of *CCNE1* locus at the targeted region. Sequencing results showed that one of the clones harbor deletions at the targeted site, which resulted in a frameshift mutation for the cyclin E protein (Figure 13A). Western blot using anti-cyclin E antibody also confirmed the lack of endogenous cyclin E expression in the total cell lysates (Figure 13B).

2.4.2 Generation of LMW-E and FL-cycE inducible hMECs

We next generated the inducible human mammary epithelial cell lines capable of expressing the ectopic LMW-E or FL-cycE in a p53 deficient (76NE6) and proficient (76NF2V) background. To this end, we used the tetracycline regulatory system (tet-on) system (Figure 14) developed based on a transcription regulatory system (Tn10 tet operon) in gram-negative bacteria for tetracycline resistance, and the key elements TetR protein that function as a dimer to bind to the Tn10 tet operon[130]. The binding between TetR protein dimer and Tn10 tet operon inhibits the down-stream promoter activity and prevents gene transcription. Binding of tetracycline or doxycycline leads to a conformation change in the TetR protein dimers, subsequently releasing its binding with tet operon to drive the transcription of downstream genes. The tet-on system developed for mammalian cells uses the fusion protein containing a modified TetR (termed as reverse-TetR, or rTTA) and the herpes simplex virus VP16

activation domain. Binding of tetracycline/doxycycline facilitates conformational switch of rTTA, and promote gene transcription under the control of an engineered promoter (tetracycline response element, or TRE) containing 7 tetO sequences fused to a minimal TATA-box containing a eukaryotic promoter[130].
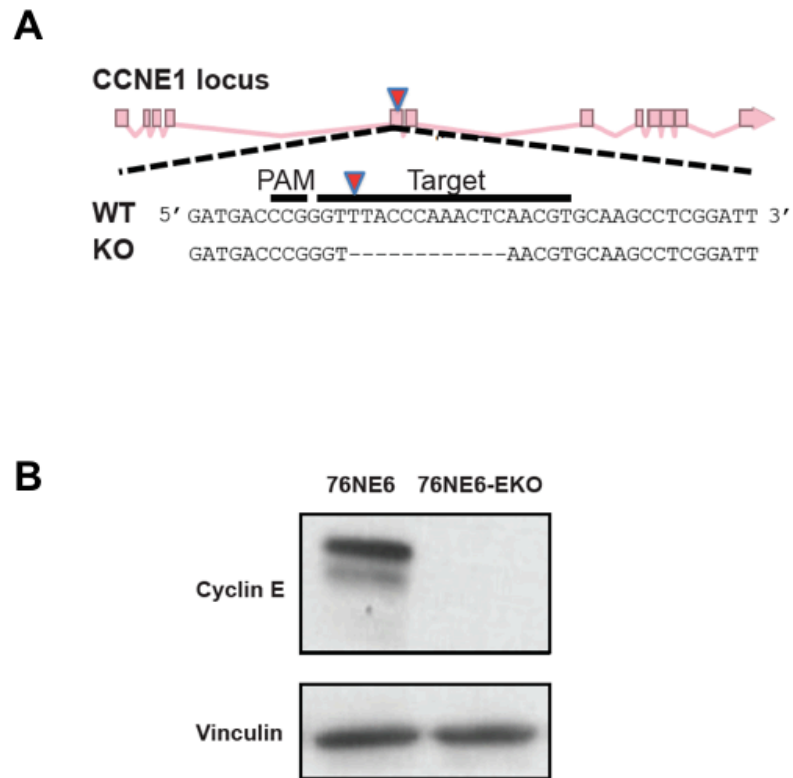
Figure 13. Generation of *CCNE1* knock-out human mammary epithelial cell (hMEC). 76NE6 hMECs were transfected with pX330-Cas9-*CCNE1*sgRNA targeting *CCNE1* gene locus exon 5 followed by clonal selection and expansion. The successful knock out of *CCNE1* gene was confirmed by Sanger sequencing showing the deletion at the sgRNA targeted site (A) and western blotting showing depletion of cyclin E expression in total cell lysates (B).

We generated the expression constructs for LMW-E, FL-cycE and empty-vector controls by using the pLVX-TRE3G vector. The cells were maintained in tetracycline free culture media to prevent leaky expression of LMW-E or FL-cycE. In the presence of doxycycline, the rtTA specifically binds to TRE and subsequently facilitate the expression of FL-cycE or LMW-E (Figure 14).
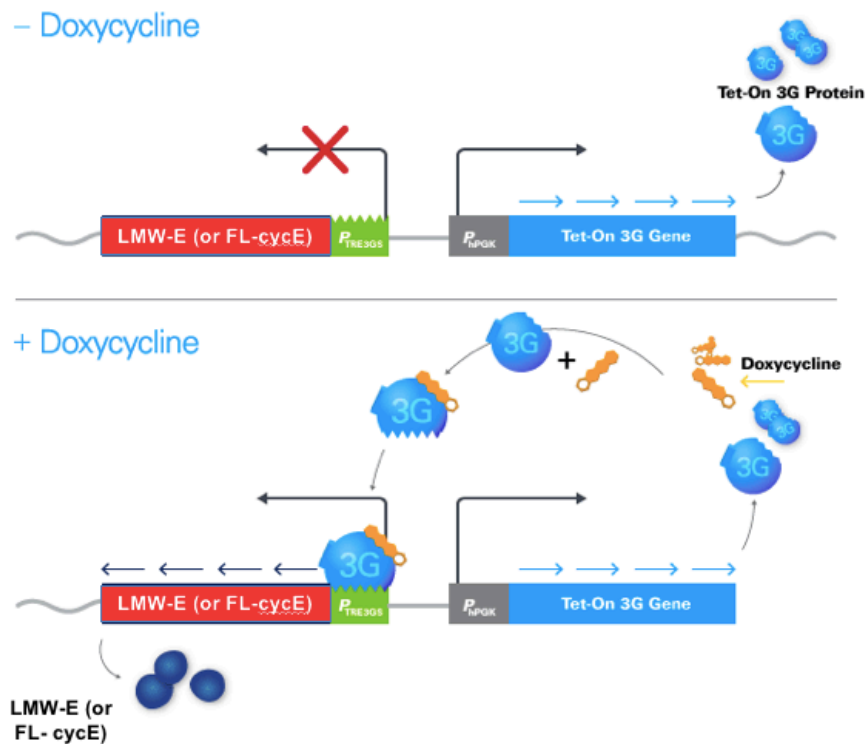
Figure 14. Schematics of the tet-on system to express LMW-E or FL-cycE under the control of doxycycline induction. Binding of doxycycline facilitates conformational switch of Tet-on 3G protein (containing reverse-TetR, or rTTA), and promotes gene transcription under the control of TRE-3G promoter (containing tetracycline response element or TRE). This figure is edited from on-line schematics of the Tet-On-3G system: https://www.takarabio.com/products/gene-function/tet-inducible-expression-systems.

We used lentivirus to generate the stable cell lines for inducible FL-cycE and LMW-E in 76NE6-EKO and 76NF2V-EKO background. The TRE constructs containing FL-cycE, or LMW-E (or empty vector) were co-transfected with pCMV-deltaR8.9 and pMD2.G-VSVG packaging vectors into HEK293T cells for lentivirus packaging. Lentivirus capable of expressing TRE driven LMW-E or FL-cycE were harvested 96 hours post transfection. 76NE6-EKO and 76NF2V-EKO cells were then transduced with the lentiviruses sequentially for two passages to enhance efficiency. This was followed by selection of cells surviving puromycin (1µg/mL) containing media. After one week of cell expanding, cell sorting was performed after a brief doxycycline induction (12 hours, 1µg/mL) to select single clones that harbor the transfected constructs. This was based on the doxycycline inducible fluorescent signal of EGFP, which is fused to the C-terminus in frame with the inducible target proteins (FL-cycE or LMW-E). EGFP high clones (top 10%) were sorted by FACS into 96 well plates and expanded in tetracycline free DCFI-1 media. For each of the inducible cell lines, namely 76NE6-EKO-FL-cycE, 76NE6-EKO-LMW-E, 76NF2V-EKO-FL-cycE, and 76NF2V-EKO-LMW-E, we have cell strains derived from single clone capable of expressing EGFP in more than 80% of the cell population when induced with doxycycline (Figure 15).

To examine the inducible expression of FL-cycE and LMW-E proteins, we treated the inducible hMEC lines with 0 to 100 ng/mL doxycycline for 24 hours, followed by western blot analysis for the expression of FL-cycE or LMW-E. When treated with vehicle alone (DMSO, dox= 0 ng/mL), the expression of FL-cycE and LMW-E were not detectable in both inducible 76NE6-EKO lines and 76NF2V-EKO lines. When cultured with increasing concentration of doxycycline, dose dependent expression of FL-cycE and LMW-E were detected. In inducible 76NE6-EKO background, 25 ng/mL doxycycline lead to 1 X (normalization reference of cyclin E western blotting densitometry) and 5 X level in FL-cycE and LMW-E expression. When treated with 50 ng/ml doxycycline, 5.5 X and 15.3 X level in FL-cycE and LMW-E were

observed. Additionally, 100 ng/mL doxycycline treatment lead to 23X and 30X level in FL-cycE and LMW-E respectively (Figure 16A). Similarly, in inducible 76F2V-EKO background, 25 ng/mL doxycycline lead to 1.0X (normalization reference) and 5 X level in FL-cycE and LMW-E, 50 ng/mL doxycycline lead to 3 X and 8 X level in FL-cycE and LMW-E, while 100 ng/mL doxycycline lead to 7X and 9X level in FL-cycE and LMW-E, respectively (Figure 16B). These results suggest the successful generation of doxycycline inducible FL-cycE and LMW-E in 76NE6-EKO and 76NF2V-EKO cell lines (Figure 16).

A
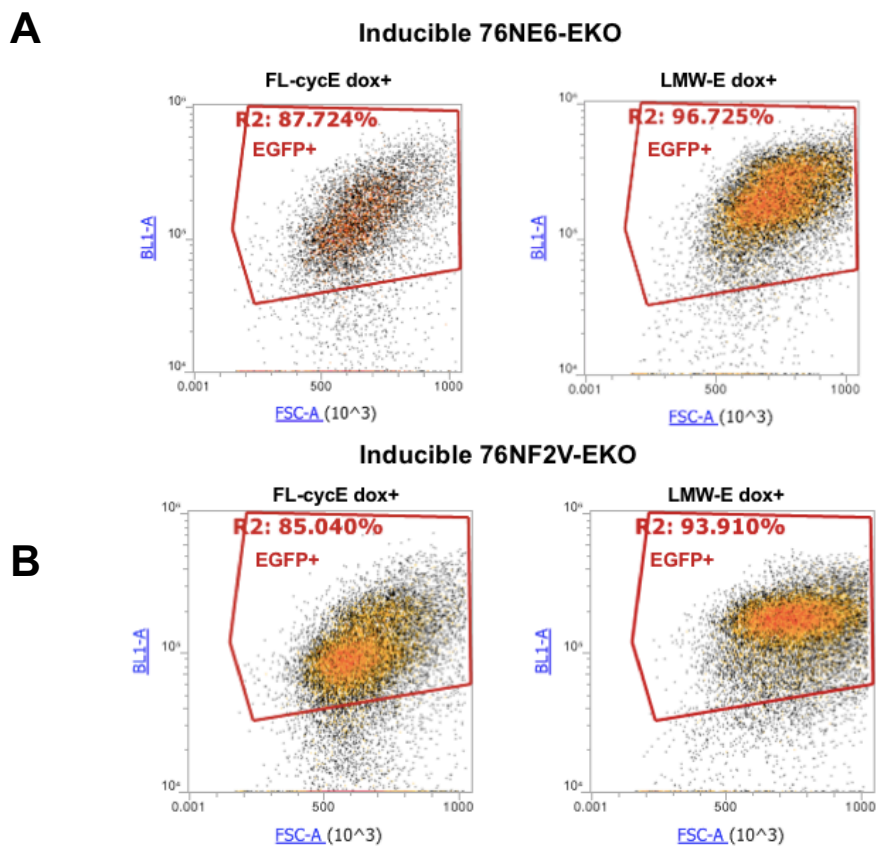
Figure 15. Expression of inducible FL-cycE and LMW-E in 76NE6-EKO and 76NF2V-EKO cell lines analyzed by FACS. Induction of positive fluorescent signals of EGFP fused to the C-terminus of FL-cycE or LMW-E proteins were detected by FACS in 76NE6-EKO (A) and 76NF2V-EKO (B) cell lines, harboring doxycycline inducible (tet-On) FL-cycE or LMW-E. The cells were treated with 100ng/mL doxycycline for 12 hours.

A

Inducible 76NE6-EKO

	FL-cycE				LMW-E			
Dox(ng/ml)	0	25	50	100	0	25	50	100

Cyclin E — FL-cycE / LMW-E

		1.0	5.5	23.2		5.7	15.3	30

Vinculin

1.0	0.9	0.8	1.0	0.9	1.0	0.9	1.0

B

Inducible 76NF2V-EKO

	FL-cycE				LMW-E			
Dox(ng/ml)	0	25	50	100	0	25	50	100

Cyclin E — FL-cycE / LMW-E

		1.0	3.1	7.3		5.1	8.5	9.0

Vinculin

1.0	1.2	1.2	1.4	1.2	1.4	1.3	1.4
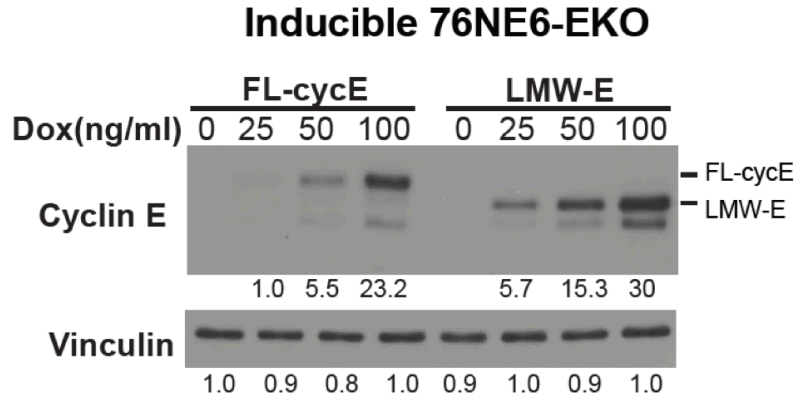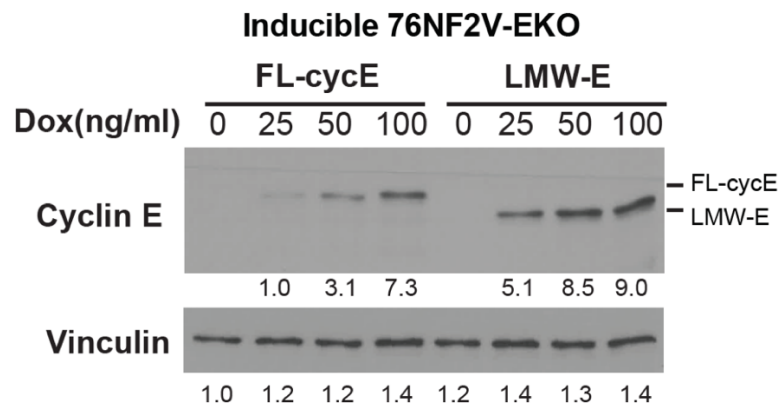
Figure 16. Expression of inducible FL-cycE and LMW-E in 76NE6-EKO and 76NF2V-EKO cell lines analyzed by western blot. Inducible 76NE6-EKO-FL-cycE or LMW-E cell lines (panel A), as well as 76NF2V-FL-cycE or LMW-E cell lines (panel B) were treated with 0, 25, 50 or 100ng/ml doxycycline for 24 hours. Total cell lysates were subjected to western blot analysis using cyclin E antibody to detect the expression of FL-cycE and LMW-E. Vinculin was used as loading control.

To examine the inducible kinase activity of induced FL-cycE and LMW-E, we performed *in vitro* kinase activity assay using GST-Rb as a substrate. Inducible 76NE6-EKO FL-cycE and LMW-E cells were treated with 0, 1, 10, or 100ng/mL doxycycline. After 24 hour of doxycycline (or vehicle DMSO) treatment, the cells were lysed for protein extraction. The

co-immunoprecipitation(IP) was then performed by using a poly clonal cyclin E antibody as bait to pull down FL-cycE or LMW-E[132]. The bound proteins in the anti-cyclin E IP were incubated with [^{32}P] ATP and pRB (GST-RB), followed by *in vitro* kinase assay to examine the phosphorylation of RB protein. The results show that without doxycycline treatment, cyclin E kinase activity was not detectable in inducible LMW-E or FL-cycE cells (Figure 17). Treatment of doxycycline at 1,10, and 100ng/mL led to the increasing kinase activity in both FL-cycE or LMW-E overexpressing cells at similar levels (Figure 17). Collectively, these results show the successful establishment of inducible LMW-E and FL-cycE expression hMECs in endogenous cyclin E knock out background as our model system for this study.

Inducible 76NE6-EKO

	FL-cycE					LMW-E			
Dox(ng/ml)	0	1	10	100		0	1	10	100

^{32}P-GST-RB

1.0 3.7 11.0   0.9 3.3 7.9

GST-RB

1.0 0.8 0.8 0.8   0.8 0.7 0.6 0.6
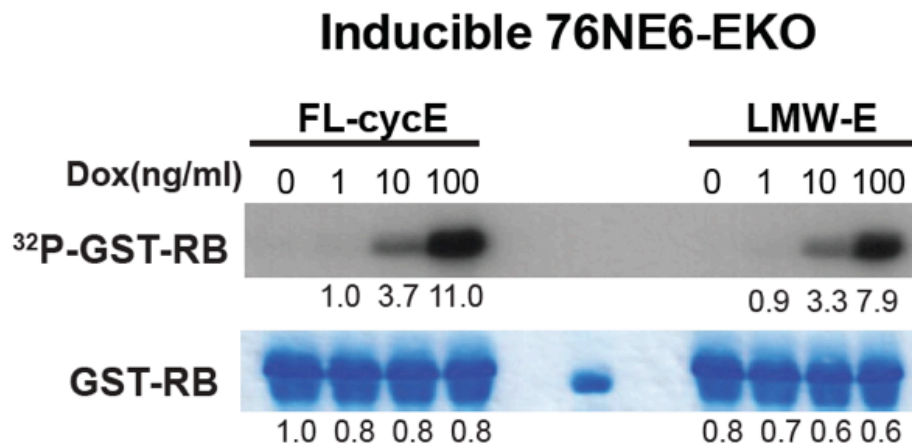
Figure 17. *In vitro* cyclin E associated kinase assay following induction of FL-cycE or LMW-E in inducible 76NE6-EKO cell lines. Without doxycycline treatment, cyclin E kinase was not detected in either inducible 76NE6-EKO-FL-cycE or 76NE6-EKO-LMW-E cell lines. Increasing kinase activities of FL-cycE or LMW-E were observed when the cells were treated with 1, 10, or 100ng/mL doxycycline for 24 hours.

2.4.3 Overexpression o FL-cycE but not LMW-E inhibited hMECs proliferation

We applied live cell imaging system to examine the effect of LMW-E and FL-cycE on the proliferation of the inducible hMECs. In the uninduced setting (i.e. no doxycycline treatment), we observed a similar growth pattern between 76NE6-EKO FL-cycE and 76NE6-EKO LMW-E cells, as well as between 76NF2V-EKO FL-cycE and 76NF2V-EKO LMW-E cells. Seeding at 1500 cells per each well of 96 well plates, the uninduced or LMW-E induced cells reached 80% confluency after 4 days in culture. However, induced FL-cycE overexpression showed strong inhibitory effect on cell proliferation compared with the settings of LMW-E or empty vector over-expression. Consistent results were obtained in both the 76NE6-EKO (Figure 18A) or 76NF2V-EKO (Figure 18B) background. To quantify and compare the effect of FL-cycE and LMW-E on cell proliferation, we calculated the cell doubling times based on the growth curves obtained in Figure 18. The results suggest in 76NE6-EKO cell lines, induced expression of FL-cycE with 25, 50 and 100 ng/mL doxycycline, respectively leads to 20%, 40% and 70% increase of the doubling time (Figure 19A). Similarly, in the 76NF2V-EKO cell lines, induced expression of FL-cycE with 25, 50 and 100 ng/mL doxycycline, respectively lead to 16%, 54% and 82% increase of the doubling time (Figure 19B). On the other hand, overexpression of empty vector or LMW-E in both 76NE6-EKO and 76NF2V-EKO background resulted in comparable doubling time as un-induced cells (Figure 20). These results suggest FL-cycE but not LMW-E inhibits the hMEC proliferation in both 76NE6-EKO and 76NF2V-EKO background.

2.4.4 Overexpression o FL-cycE but not LMW-E inhibited hMECs viability

Next, we assessed cell viability by MTT assay using the same experimental strategies as cell proliferation experiments. Following 4 days of culturing 76NE6-EKO and 76NF2V-EKO cell lines with or without 25, 50 or 100ng/mL doxycycline, we subjected all cells to MTT assay (see methods). The results for inducible FL-cycE or LMW-E cells were normalized to empty

vector cells with the same treatment conditions and cellular background to calculate the relative cell viability.
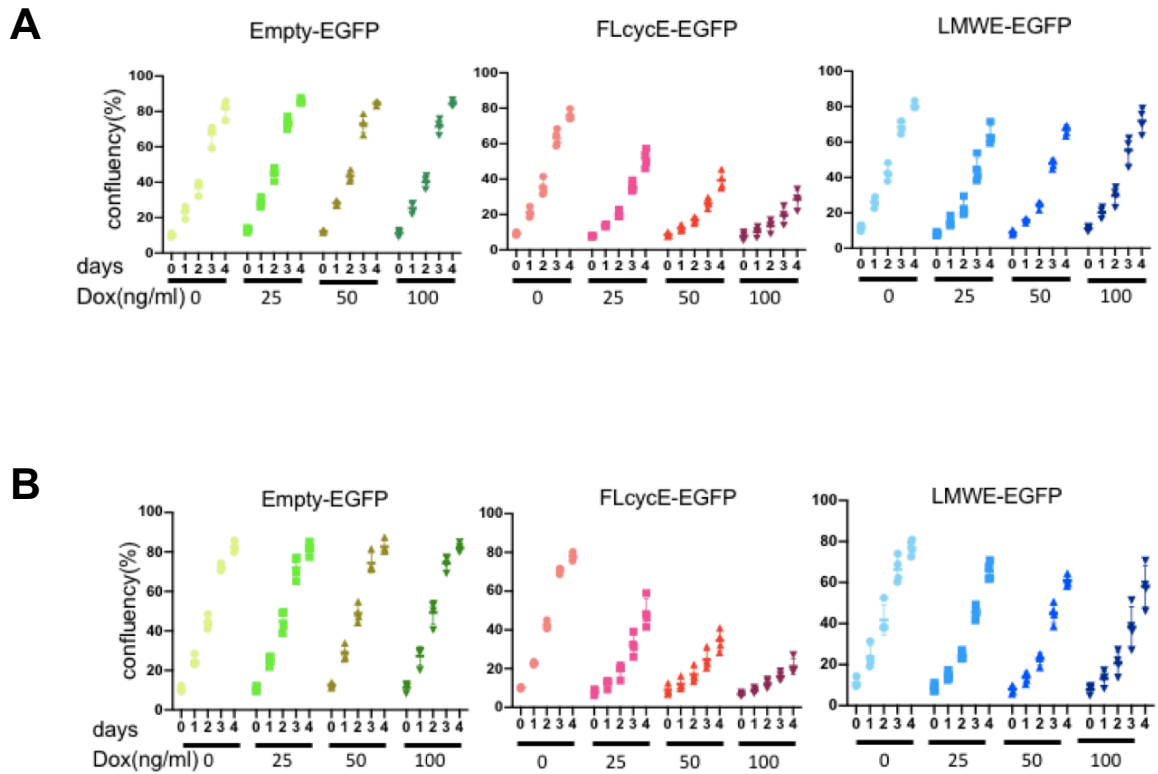
Figure 18. FL-cycE but not LMW-E inhibited cell proliferation in inducible 76NE6-EKO and 76NF2V-EKO cells. Inducible 76NE6-EKO (panel A) and 76NF2V-EKO (panel B) cells were treated with doxycycline at indicated concentration and the cell confluency was monitored by live cell imaging system (incucyte) for a time span of four days. The growth curves show overexpression of FL-cycE, but not LMW-E (or empty vector), down-regulated the cell proliferation of inducible 76NE6-EKO (panel A) and 76NF2V-EKO (panel B) cell lines.

By comparing the viability between induced and uninduced cells, we found that FL-cycE but not LMW-E induction significantly inhibited cell viability in a doxycycline dose dependent manner in both 76NE6-EO and 76NF2V-EKO cells. Overexpression of FL-cycE, treated with 25, 50, and 100 ng/mL doxycycline, respectively led to 30%, 60% and 80% decrease of cell viability in 76NE6-EKO background and 20%, 60% and 80% decrease of cell viability in 76NF2V-EKO background (Figure 20).No significant inhibitory effect of LMW-E on cell viability in either 76NE6-EKO or 76NF2V-EKO cell lines was observed. Additionally, treatment with 50ng/mL doxycycline increased the viability of LMW-E overexpressing cells by 1.35-fold and 1.25-fold of the uninduced 76NE6-EKO and 76NF2V-EKO cells, respectively (Figure 20). These results reveal the distinct effect between FL-cycE and LMW-E when overexpressed in mammary epithelial cells, that FL-cycE but not LMW-E inhibited cell proliferation and cell viability (Figures 17-19).

2.4.5 Overexpression of FL-cycE but not LMW-E led to S phase cell cycle arrest

Based on the observation that FL-cycE but not LMW-E inhibited the proliferation and viability of 76NE6-EKO and 76NF2V-EKO cell lines, we next hypothesized that FL-cycE and LMW-E may differentially regulate the cell cycle distribution of the cells. Initially, we examined the cell cycle distribution based on the DNA content of the cells. 76NE6-EKO and 76NF2V-EKO cell lines were induced to express FL-cycE or LMW-E (100ng/mL doxycycline treatment) for 0-48 hours. To compare the effect of FL-cycE and LMW-E on cell cycle distribution in a time course manner, doxycycline was added into the media 12, 24, 36 and 48 hours prior to cell harvesting. For each time point, the cells were collected, fixed and stained with propidium iodide (PI). The cell cycle phases were determined based on DNA content: cells that have 2n DNA content are gated as G0/G1, 4n DNA content are gated as G2/M and between 2n and 4n are gated as S phase.
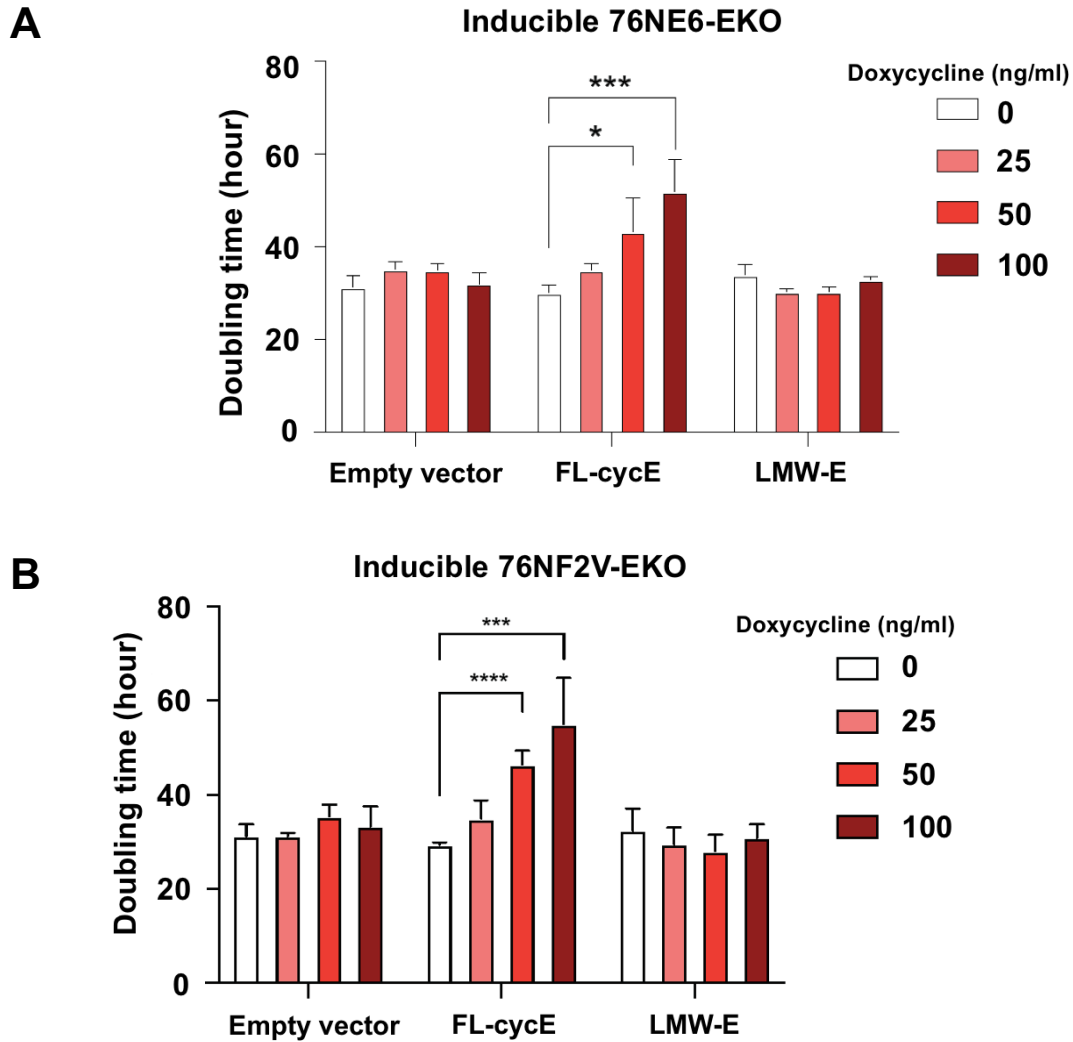
Figure 19. FL-cycE but not LMW-E increased the doubling time of inducible 76NE6-EKO and 76NF2V-EKO cells. The doubling times of inducible 76NE6-EKO (panel A) and 76NF2V-EKO(panel B) cells were calculated based on cell confluency mask from live cell imaging (Incucyte). The cells were treated with 25, 50, or 100ng/mL doxycycline to induce the overexpression of LMW-E, FL-cycE, or empty vector. DMSO (doxycycline 0ng/mL) was used as the control. The results suggest FL-cycE increased doubling time of inducible 76NE6-EKO (panel A) and 76NF2V-EKO(panel B) cells in a dose dependent manner. (n = 4, mean with standard deviation, *p < 0.05, ***p < 0.001, and ****p < 0.0001, Student *t* test)
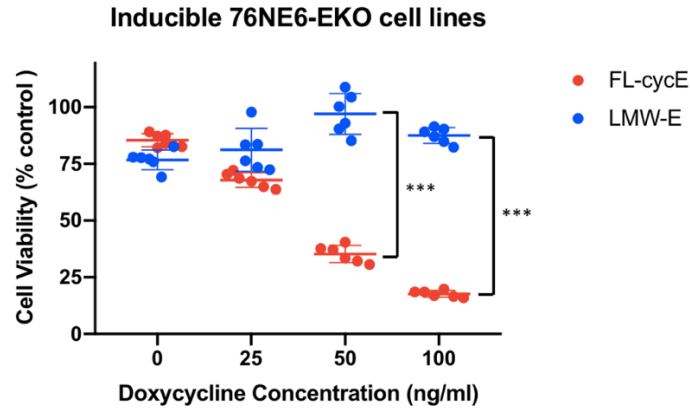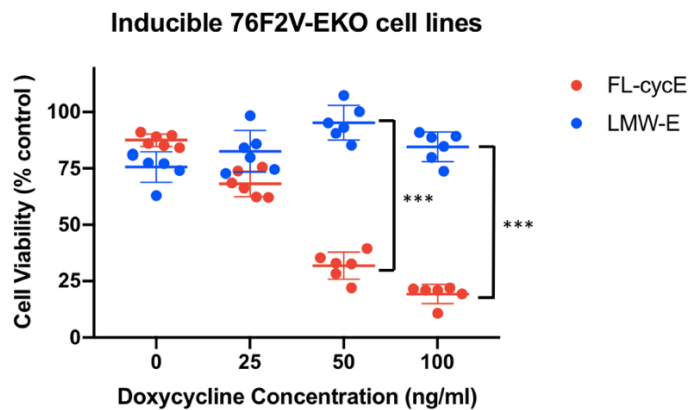
Figure 20. FL-cycE but not LMW-E reduced viability of inducible 76NE6-EKO and 76NF2V-EKO cells. Cell viability was examined by MTT assay after 4 days culture of inducible 76NE6-EKO (panel A) and 76NF2V-EKO (panel B) cell lines with or without 25, 50 or 100ng/mL doxycycline. The reading absorbance at 590 nm wavelength in inducible FL-cycE and LMW-E groups were normalized to empty vector groups of the same treatment conditions to calculate the % of cell viability. The results showed significantly lower cell viability of inducible 76NE6-EKO(panel A) or 76NF2V-EKO(panel B) cells when FL-cycE was induced by 50 or 100ng/ml doxycycline treatment, compared to the condition of LMW-E induction by 50 or 100ng/ml doxycycline treatment. ***$p < 0.001$, student t test.

The result showed that after 12 hours induced expression of FL-cycE, the ratio of cells in G0/G1 dropped from 59.9% to 36.4%, and the ratio cells in S increase from 17.7% to 41.1% (Figure 21 bottom panels). LMW-E overexpression for 12 hours similarly decreased the ratio

of G0/G1 phase cells from 64.7% to 40.1, while increased the ratio of S phase cells from 17.8% to 41.4% (Figure 22 bottom panels). Similar results were observed in the FL-cycE and LMW-E inducible 76NF2V-EKO cells, were an increased ratio of S/G1 phase was observed following 12 or 24 hours of induction (Figure 23 and 24 bottom panels). These results suggest FL-cycE and LMW-E, when overexpressed for 12 or 24 hours, similarly drive the cell cycle progression from G0/G1 to S phase. In later time points, particularly 48 hours post induction for both FL-cycE or LMW-E, we find accumulation of cells at G2/M phase, suggesting FL-cycE and LMW-E overexpression may activate G2/M checkpoint of the cell cycle (Figure 21 - 24).

To more specifically interrogate the effect of FL-cycE and LMW-E on S phase changes, we used BrdU incorporation to measure DNA synthesis. We combined the activity of DNA replication (BrdU signal) with DNA content (PI staining signal) to stratify cell cycle distributions. To this end, we used the following gating parameters: "DNA content=2n, BrdU incorporation low" to gate the G0/G1 phase, "DNA content=2n; BrdU incorporation high" to gate early S phase, "2n<DNA content<4n; BrdU incorporation high" to gate middle S phase, "DNA content=4n; BrdU incorporation high" to gate late S phase and "DNA content=4n; BrdU incorporation low" to gate G2/M phase (Figure 21 - 24). The results show dramatic differences in cell cycle distribution between FL-cycE and LMW-E overexpressing cells, starting from 24 hours post treatment of doxycycline (Figure 25). Although both FL-cycE and LMW-E can increase the ratio of cells undergoing DNA replication (DNA content between 2n and 4n, measured by PI staining), the DNA replicating activity (measured by BrdU incorporation) was inhibited by FL-cycE but not LMW-E. We gated the "2n<DNA content<4n, BrdU incorporation low" cells as rested S phase (RS) and compared its ratio between each time points (Figure 25).
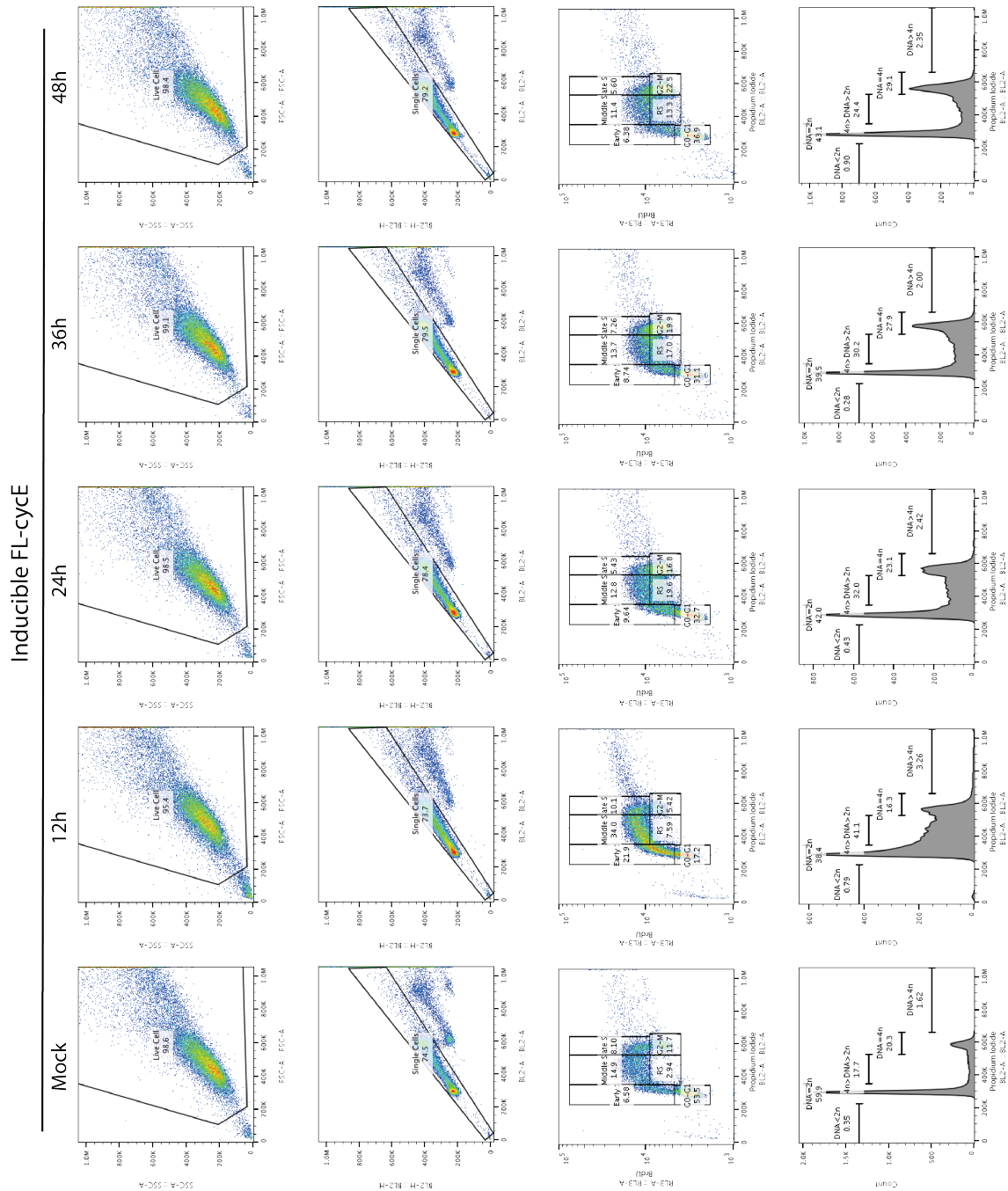
Figure 21. Cell cycle analysis for inducible 76NE6-EKO cells overexpressed with FL-cycE in a time course manner. Inducible 76NE6-EKO-FL-cycE cells were treated with 100ng/mL doxycycline for 12 to 48 hours followed by cell cycle analysis by PI staining and BrdU incorporation. Vehicle (DMSO) treated cells were used as mock control (uninduced).
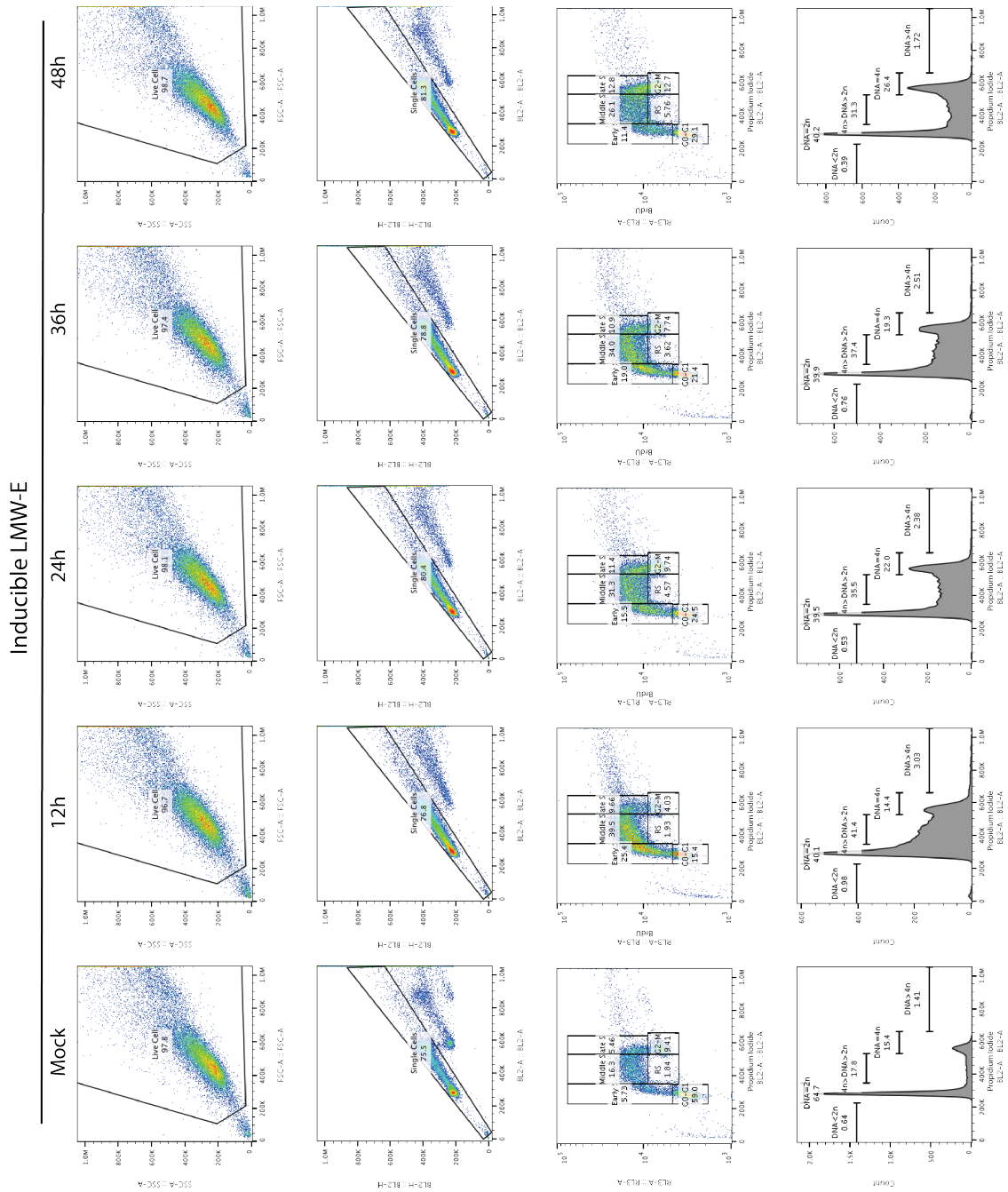
Figure 22. Cell cycle analysis for inducible 76NE6-EKO cells overexpressed with LMW-E in a time course manner. Inducible 76NE6-EKO-LMW-E cells were treated with 100ng/mL doxycycline for 12 to 48 hours followed by cell cycle analysis by PI staining and BrdU incorporation. Vehicle (DMSO) treated cells were used as mock control (uninduced).
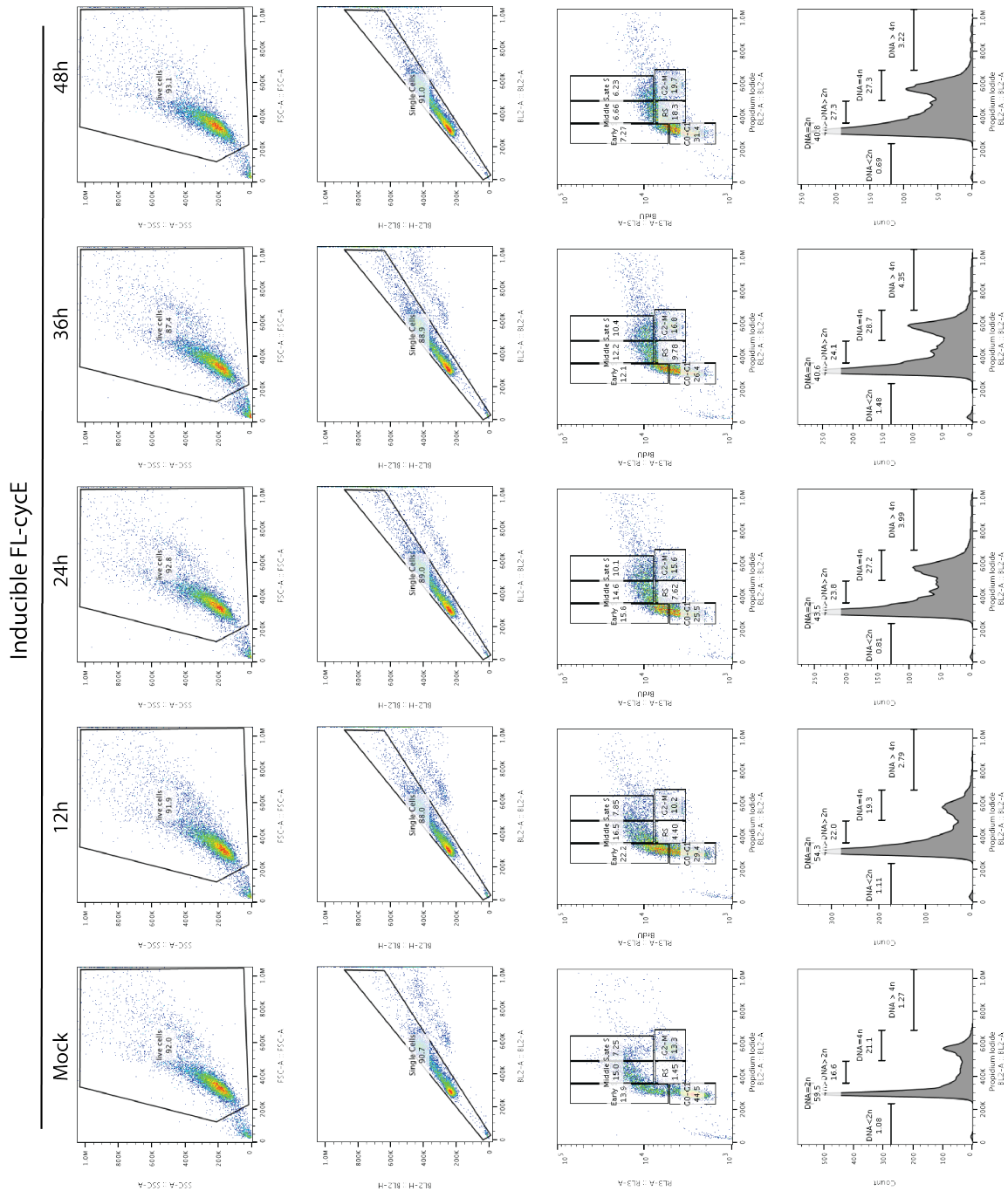
Figure 23. Cell cycle analysis for inducible 76NF2V-EKO cells overexpressed with FL-cycE in a time course manner. Inducible 76NF2V-EKO-FL-cycE cells were treated with 100ng/mL doxycycline for 12 to 48 hours followed by cell cycle analysis by PI staining and BrdU incorporation. DMSO treated cells were used as mock control (uninduced).
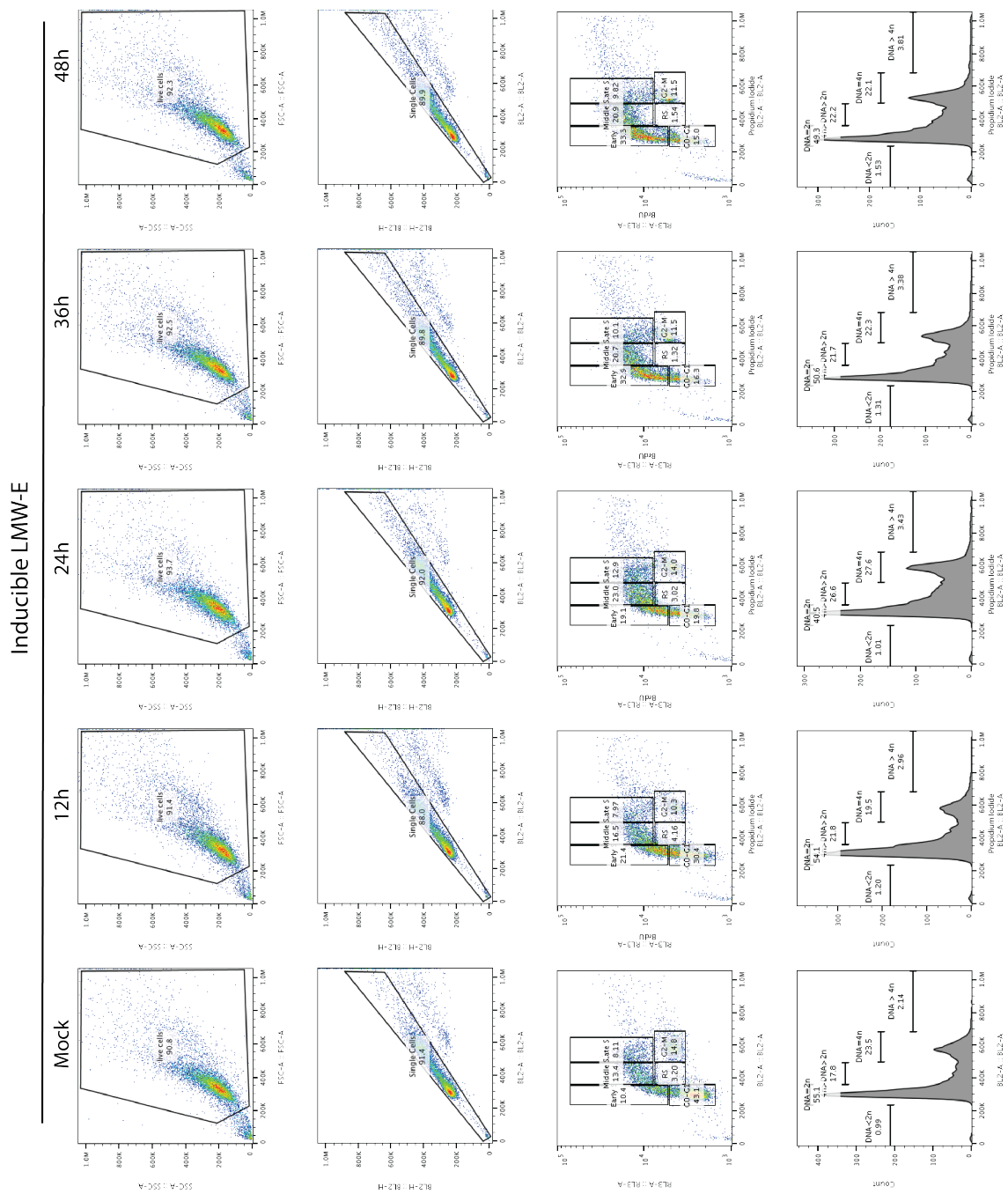
Figure 24. Cell cycle analysis for inducible 76NF2V-EKO cells overexpressed with LMW-E in a time course manner. Inducible 76NF2V-EKO-LMW-E cells were treated with 100ng/mL doxycycline for 12 to 48 hours followed by cell cycle analysis by PI staining and BrdU incorporation. DMSO treated cells were used as mock control (uninduced).
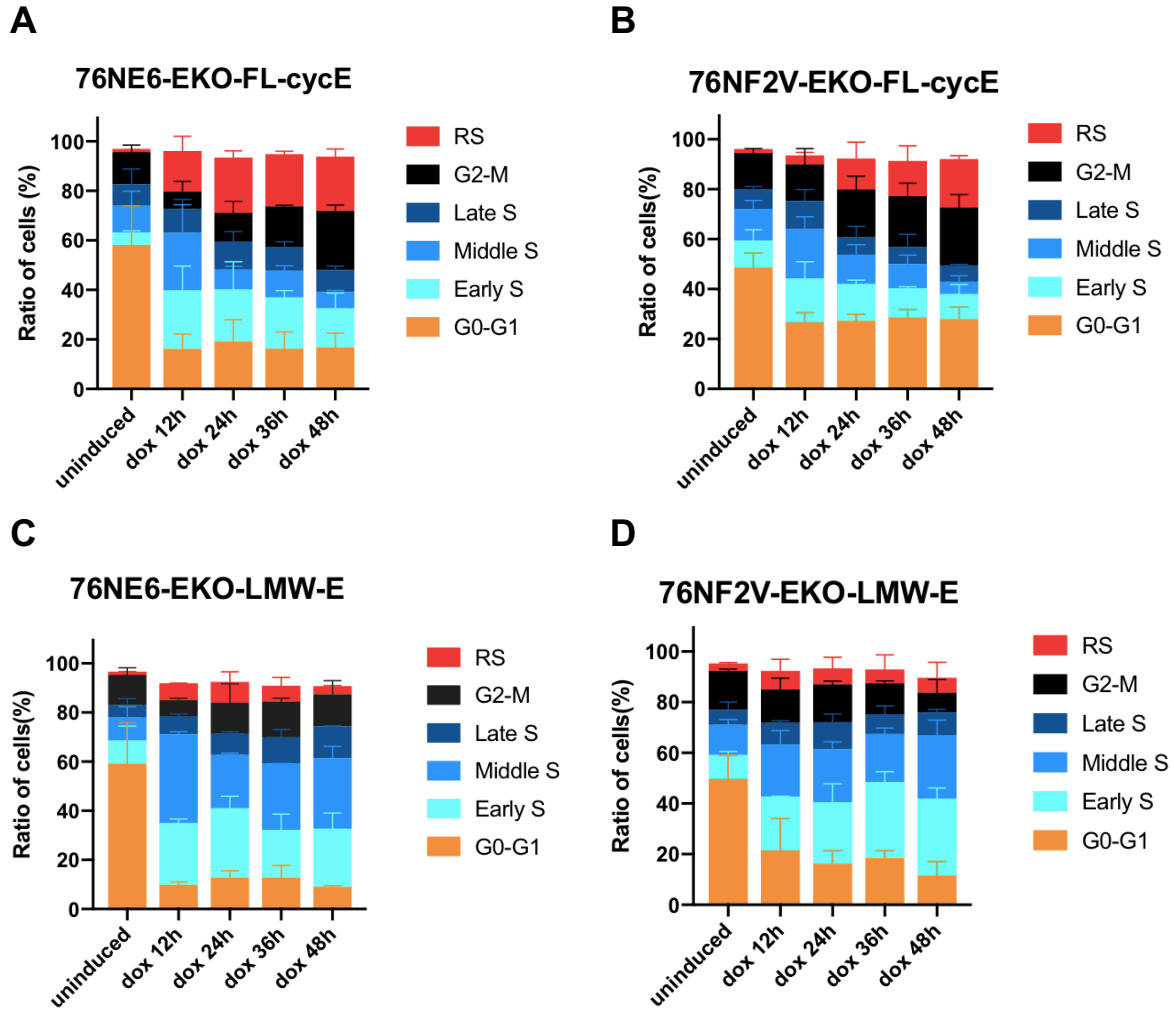
Figure 25. Quantitation of ratio of cells in different cell cycle phases in the inducible 76NE6-EKO and 76NF2V-EKO cell lines. We used "DNA content=2n, BrdU incorporation low" for the gating of G0/G1 phase, "DNA content=2n; BrdU incorporation high" for early S phase, "2n<DNA content<4n; BrdU incorporation high" for middle S phase, "DNA content=4n; BrdU incorporation high" for late S phase, "DNA content=4n; BrdU incorporation low" for G2/M phase, and 2n<DNA content<4n; BrdU incorporation low" for rested S phase (RS). The results showed FL-cycE led to 20% cells arrested in S phase (RS gating, red colored bars) after 24 hours overexpression in 76NE6-EKO cells (panel A) and 48 hours overexpression in 76NF2V-EKO cells (panel B), while LMW-E overexpression led to no more than 5% cells arrested in S phase in either 76NE6-EKO (panel C) or 76NF2V-EKO (panel D) background.

These data revealed that without induced overexpression of FL-cycE or LMW-E, the cells within RS gating remain low (less than 4 %), in both 76NE6-EKO and 76NF2V-EKO. Induction of FL-cycE led to an increase of cells in the RS group to 20% in 76NE6-EKO (starting from 24 hours) and 76NF2V-EKO (starting from 48 hours) cells. On the other hand, the ratio of RS cells only slightly increased (to around 5%) in LMW-E overexpressing 76NE6-EKO and 76NF2V-EKO cells at any of the time points examined.

Collectively, our results suggest that while FL-cycE and LMW-E both drive cell cycle progression from G0/G1 to S phase, 20% of the cells in S phase were arrested by the overexpression of FL-cycE but not LMW-E.

2.4.6 Overexpression of FL-cycE but not LMW-E led to DNA damage accumulation

Previous studies suggest that induction of DNA damage may cause the cells to arrest at S phase[133]. To interrogate if FL-cycE and LMW-E have different abilities to induce DNA damage, we quantitated DNA lesions in cells by immunofluorescent (IF) assay using the DNA damage markers γ-H2AX and 53BP1. For these experiments, cells were induced to express FL-cycE or LMW-E by 100ng/mL doxycycline for 0-48 hours, followed by IF assays. We compared the ratio of γ-H2AX positive cells (nuclear foci>5) and 53BP1 positive cells (nuclear foci>5) between each time point in the inducible 76NE6-EKO FL-cycE or LMW-E cells (Figure 26). The results suggest that compared to un-induced cells (48 hours vehicle treatment), 12 hours of induction of FL-cycE or LMW-E both led to 3-fold increase of the ratio of γ-H2AX positive cells. For 53BP1 positive cells, the fold change was 4.3-fold for FL-cycE and 2.5-fold for LMW-E (Figure 26B). Comparing the 24 hour time point with 12 hour time point, we found that FL-cycE overexpressing led to 4.2-fold increase of γ-H2AX positive cells and 3.8 fold increase of 53BP1 positive cells. LMW-E overexpressing increased 2.5 fold of γ-H2AX positive cells and 3.1 fold of 53BP1 positive cells between 24 hours induction versus 12 hours induction (Figure 26B). These results suggest both FL-cycE and LMW-E promote DNA

damage in an inducible manner, while higher DNA damage intensity was observed in FL-cycE overexpressing cells compared with LMW-E overexpressing cells.

The accumulation of DNA damage by FL-cycE but not LMW-E were more clearly shown in later time points. At 36 hours post induction, FL-cycE overexpression caused 66% of total cells to be γ-H2AX positive, and 65% to be 53BP1 positive. At 48 hours post induction of FL-cycE, 76% of total cells were positive for γ-H2AX positive, and 77% for 53BP1 positive. LMW-E overexpression for 48 hours remained at similar levels to the 36-hour time point, showing 38% or 36% of total cells were γ-H2AX or 53BP1 positive respectively (Figure 26B).

Consistent results were observed in the 76NF2V background, where 36 hours of FL-cycE overexpressing increased the ratio of γ-H2AX positive cells to 49% and 53BP1 positive cells to 48%, and 48 hours of FL-cycE overexpressing increased the ratio of γ-H2AX positive cells to 57% and 53BP1 positive cells to 72%. In LMW-E overexpressing cells the ratio of γ-H2AX positive cells and 53BP1 positive cells were between 25% to 35% in the 36 hours and 48 hours induction groups, which is significantly less than FL-cycE overexpressing cells (Figure 27A and 27B). Consistent results were also obtained by measuring DNA breaks in individual cells using the comet assay. We compared the 76NE6-EKO cells with or without 48 hours overexpression of FL-cycE or LMW-E induced by 100ng/ml doxycycline. Our data suggested that DNA damage indexed by the migration and the fraction of total DNA in the comet tail (termed tail moment) in 76NE6-EKO cells with FL-cycE overexpression was 7.1-fold higher than in un-induced cells, and was 3.0-fold higher than in LMW-E–overexpressing cells (Figure 28).

These results suggest that both FL-cycE and LMW-E can induce DNA damage in 76NE6-EKO or 76NF2V-EKO background. However, the DNA damage accumulation was more significantly when FL-cycE was induced, compared with LMW-E overexpression settings.
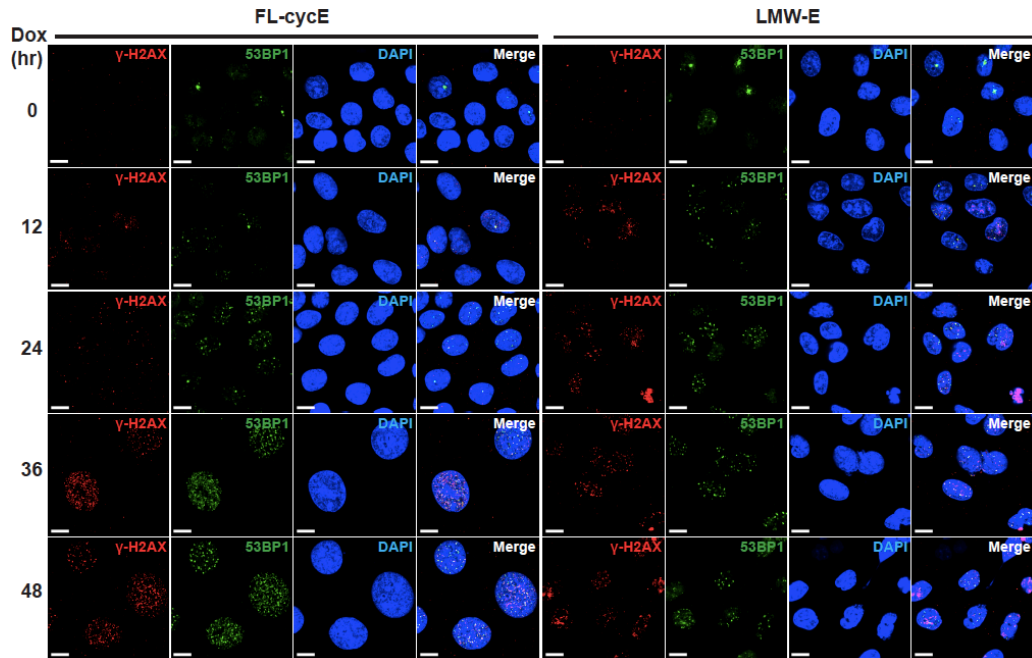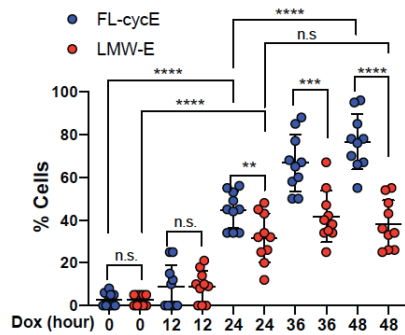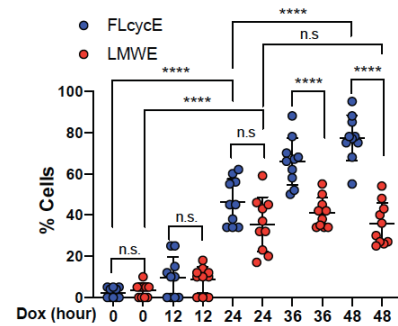
Figure 26. Time course analysis of DNA damage markers γ-H2AX and 53BP1 foci in the inducible 76NE6-EKO cell lines. Doxycycline (100 ng/mL) was used to induce FL-cycE or LMW-E, and un-induced control (Dox 0 hours) was treated with DMSO for 48 hours. A. Representative images of immunofluorescent γ-H2AX and 53BP1 foci in inducible 76NE6-EKO cells overexpressed with FL-cycE or LMW-E in a time course manner (scale bar = 10 μm). B and C. quantification of γ-H2AX-positive (panel B, nuclear foci > 5) and 53BP1-positive cells (panel C, nuclear foci > 5; n = 3, cell number > 600, mean with standard deviation, **p < 0.01, ***p < 0.001, and ****p < 0.0001, n.s. indicates not significant, student t test.)

Figure 27. Time course analysis of DNA damage markers γ-H2AX and 53BP1 foci in the inducible 76NF2V-EKO cell lines. Doxycycline (100 ng/mL) was used to induce overexpression of FL-cycE or LMW-E, and un-induced control (Dox 0 hours) was treated with DMSO for 48 hours. A. Representative images of immunofluorescent analysis of γ-H2AX and 53BP1 foci in inducible 76NF2V-EKO cells overexpressed with FL-cycE or LMW-E in a time course manner (scale bar = 10 μm). B and C. quantification of γ-H2AX-positive cells (panel B, nuclear foci > 5) and 53BP1-positive cells (panel C, nuclear foci > 5) at the indicated time points (n = 3, cell number > 600, mean with standard deviation, student *t* test.)

A

Inducible 76NE6-EKO

FL-cycE

Dox -   Dox +

LMW-E

Dox -   Dox +

B

Inducible 76NE6-EKO
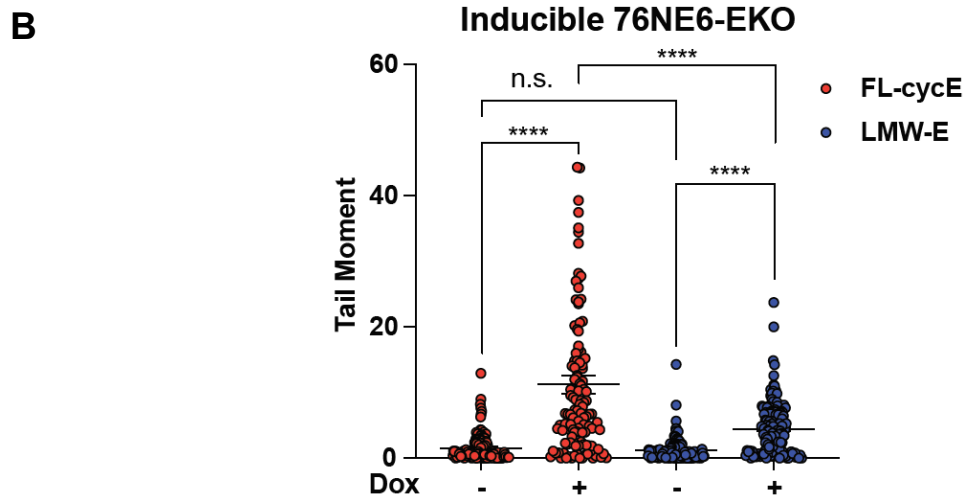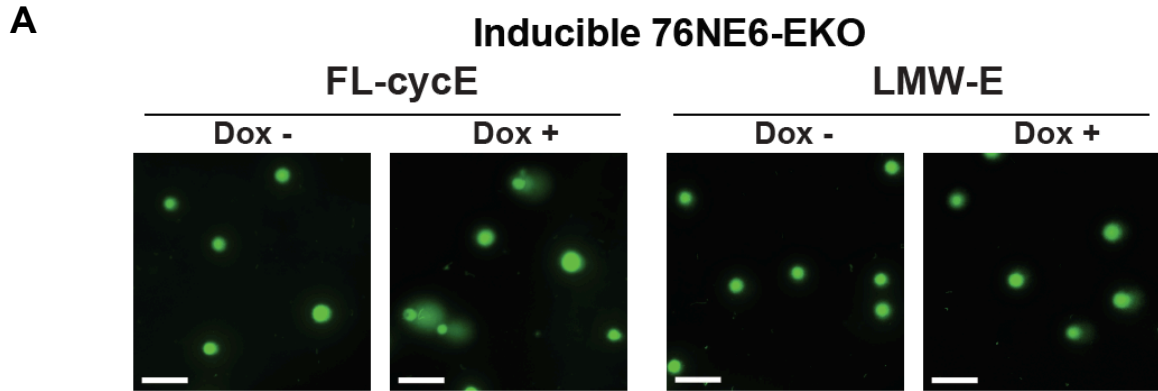
Figure 28. Confirmation of FL-cycE induced higher DNA damage than LMW-E by comet assay. Comet assay was performed to measure DNA breaks in 76NE6-EKO cells after 48 hours of doxycycline treatment to induce overexpression of LMW-E or FL-cycE (panel A, scale bar = 100 μm), and the intensity of DNA damage was quantified by tail moment (panel B, cell number > 100, n = 3, mean with standard deviation, ****p < 0.0001; n.s. indicates not significant; Student t test.)

2.4.7 Overexpression of LMW-E induced genomic instability in hMECs

Next, we tested the hypothesis that LMW-E overexpression drives genomic instability in hMECs. To this end we examined chromosomal damage mediated by either FL-cycE or LMW-E by evaluating the changes in the metaphase chromosomes. The metaphase spread assays were performed at the Molecular Cytogenetics Facility at M.D. Anderson Cancer Center. The results showed doxycycline-induced expression of LMW-E and FL-cycE can both induce significant chromosomal structural aberrations in 76NE6-EKO cells (Figure 29A, B). However, more than 2-fold of chromosomal breaks were observed in FL-cycE overexpressing cells (1.94 breaks per metaphase), compared with LMW-E overexpressing cells (0.9 breaks per metaphase; Figure 29C). Chromosome fusions in LMW-E overexpressing cells (0.26 per metaphase) were 1.5-fold of FL-cycE overexpressing cells (0.17 per metaphase; Figure 29D).

In addition, we found LMW-E overexpressing caused abnormal nuclear phenotypes in the inducible hMECs. We induced expression of LMW-E or FL-cycE for 36 hours and analyzed the ratio of cells showing abnormal nuclear phenotypes, such as micro-nuclear and multi-nuclear. For each of the conditions, we examined more than 750 cells and concluded that LMW-E overexpression lead to 3.5-fold increase of cells with micro-nuclear and 8.4-fold increase of multi-nuclear cells (Figure 30).

These results collectively suggest that overexpression of FL-cycE and LMW-E both induced DNA damage and chromosomal abnormalities in hMECs. However, FL-cycE overexpression cells lead to S phase cell cycle arrest with inhibited DNA replication, while LMW-E overexpressing cells continue proliferating with damaged DNA, leading to accumulated genomic instability.
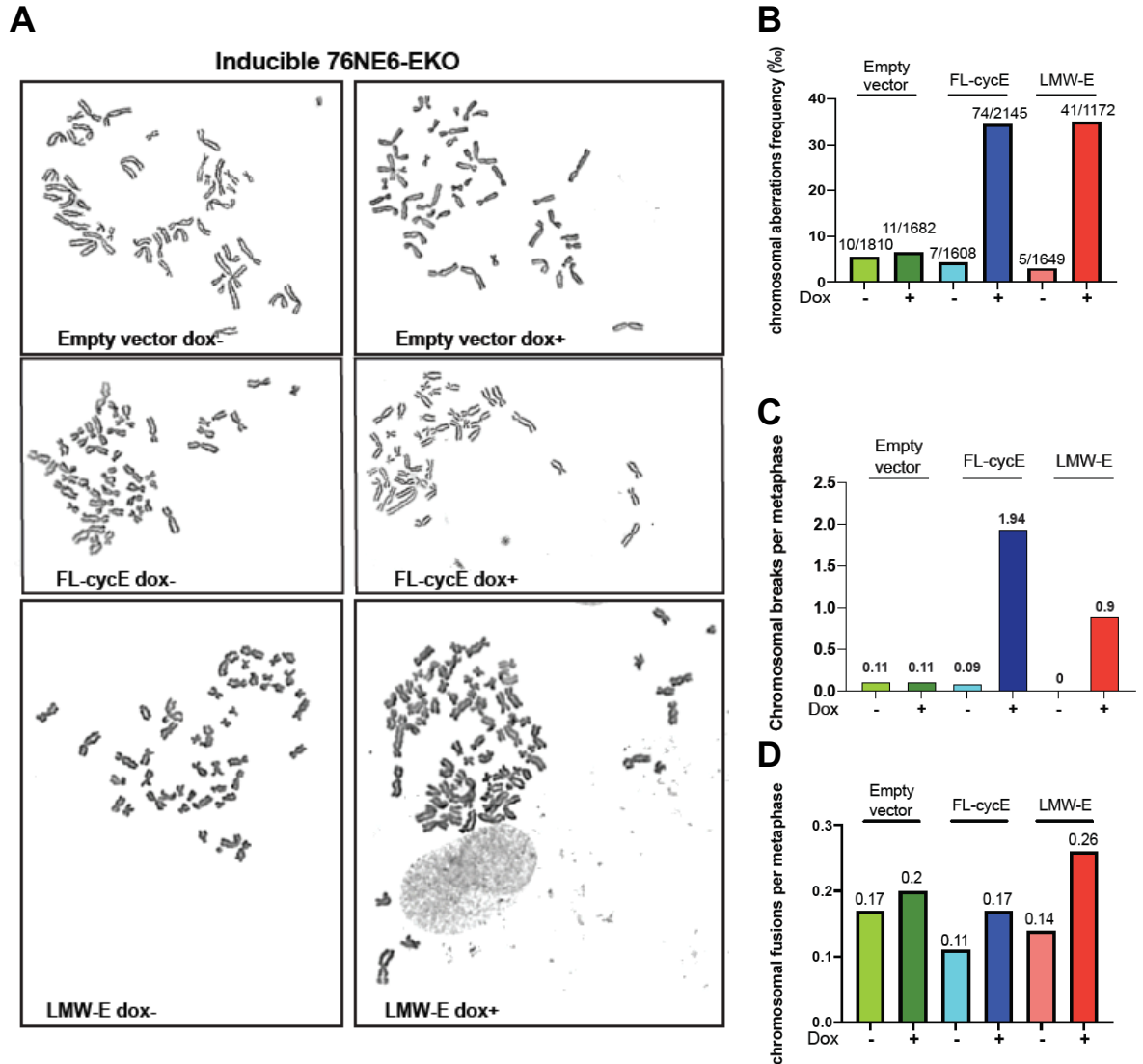
Figure 29. Chromosome structural aberrations were induced after overexpression of FL-cycE or LMW-E in hMECs. A. Representative images of chromosomal structural aberrations found in 76NE6-EKO cells after 48 hours of treatment with doxycycline to induce overexpression of empty-vector, FL-cycE or LMW-E. B-D. Quantification of chromosomal aberration frequency(B), chromosomal break frequency(C), chromosomal fusion frequency(D) in inducible 76NE6-EKO cells under the indicated expression conditions (n = 2, metaphases examined per condition = 35).
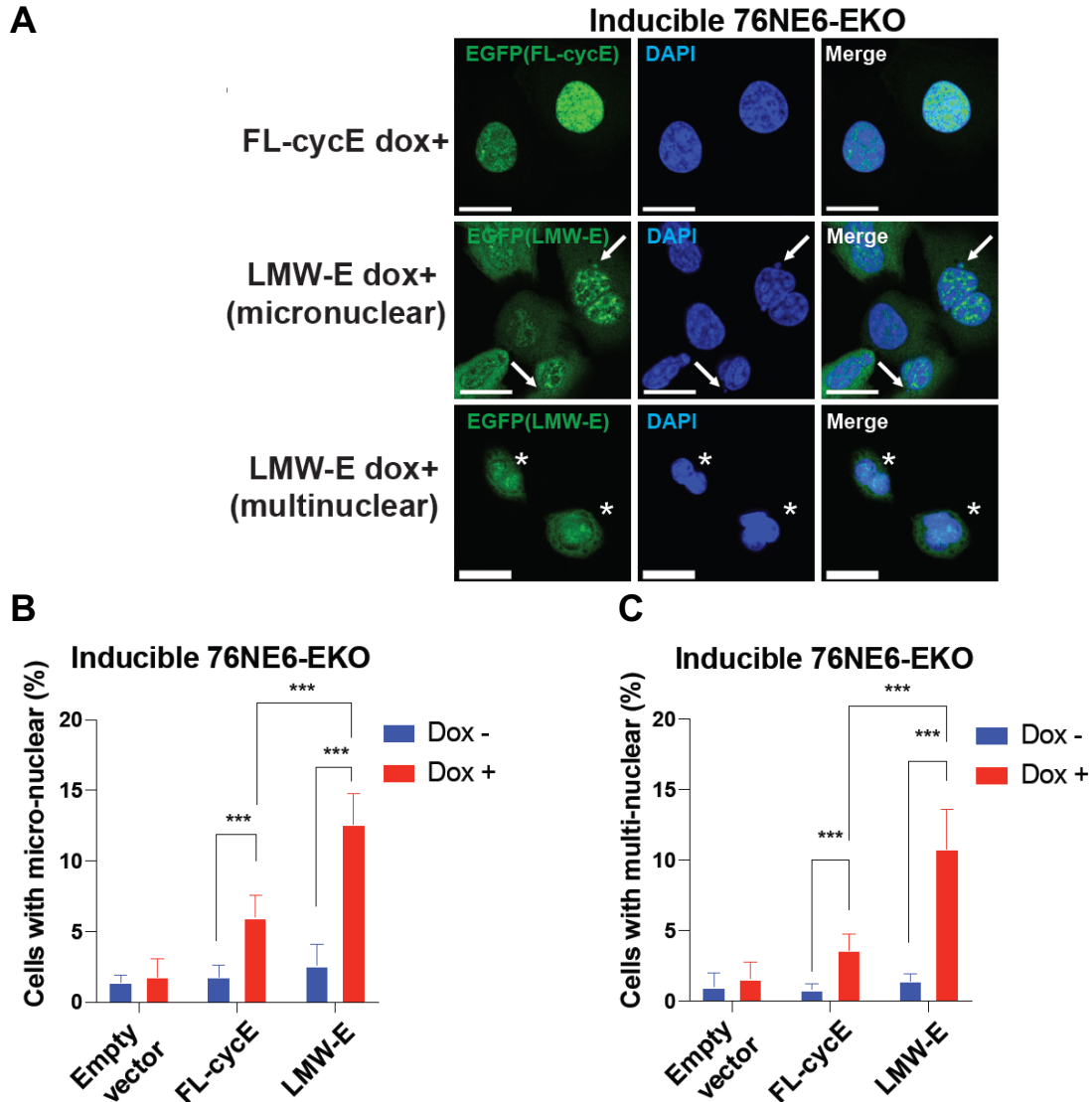
Figure 30. LMW-E overexpression induced abnormal nuclear phenotypes in hMECs.
Abnormal nuclear phenotypes (including micro-nuclear and multinuclear) were observed after 36 hours of induced expression of LMW-E in inducible 76NE6-EKO cells. A. Representative images of nuclear abnormalities found in 76NE6-EKO cells after 36 hours of treatment with doxycycline to induce FL-cycE or LMW-E overexpression and B. quantification of the ratio of cells containing micronuclear (arrow-head) and/or multinuclear (star) abnormalities in inducible 76NE6-EKO cells under the indicated expression conditions (n = 3, cell number > 750, mean with standard deviation, ***$p < 0.001$, Student t test.)

2.5 Conclusion

In this chapter, we show the generation of the *in vitro* model system containing doxycycline inducible system to express FL-cycE (aa1-410) or LMW-E (aa40-410) (schematics shown in Figure 5) in cyclin E knock-out 76NE6 (p53 deficient) and 76NF2V (p53 proficient) background. By comparing the effect of LMW-E and FL-cycE on cell proliferation, cell viability, cell cycle distribution, we find overexpression of FL-cycE in both 76NE6-EKO and 76NF2V-EKO background inhibit cell growth, reduce DNA replication in S phase, induce S phase cell cycle arrest and reduce cell viability. Both FL-cycE and LMW-E can also induce DNA damage in the 76NE6-EKO or 76NF2V-EKO at early time points. However, the DNA damage accumulate over time in cells overexpressed with FL-cycE but not LMW-E. LMW-E overexpressing cells, which can proliferate with damaged DNA, harbor elevated aberrant chromosomal structures and enhanced nuclear abnormalities including micro-nuclear and multi-nuclear phenotypes. These results support our hypothesis that LMW-E overexpression promotes genomic instability in human mammary epithelial cells independent of endogenous FL-cycE. LMW-E and FL-cycE overexpression result in different effects on cell growth and viability, featured by distinct S phase DNA replication activity and DNA damage.

Chapter Three: Common and specific transcriptional signatures induced by LMW-E and FL-cycE in hMECs

3.1 Introduction

In Chapter two, we present data on the generation of the cellular model in cyclin E knock-out background that capable to express empty vector, FL-cycE, or LMW-E when treated with doxycycline. By comparing the effect of LMW-E and FL-cycE, we find the phenotypes such as cell proliferation, cell viability, and cell cycle distributions are differentially regulated by the overexpression of FL-cycE or LMW-E.

Previous studies from our laboratory have revealed different biochemical features of LMW-E and FL-cycE. Compared to FL-cycE, LMW-E hyper-activates CDK2 and the LMW-E-CDK2 protein kinase is resistant to p21 and p27 inhibition[66]. Additionally, LMW-E is less susceptible than FL-CycE to FBW7 mediated ubiquitination and degradation, and may activate CDK2 throughout the cell cycle[67]. These features might lead to a unique regulatory pattern downstream of LMW-E overexpression, which can be interrogated by comparing the transcriptional profiles between the cells with or without LMW-E induction (Figure 31).

Additionally, comparing the pathways altered by LMW-E and FL-cycE, we seek to know why the S phase DNA replication and DNA damage are differentially regulated by LMW-E and FL-cycE, as we reported in Chapter two. This is important to determine essential down-stream factors that facilitate LMW-E overexpressing cells to proliferate with damaged DNA, and determine how the cells survive with aberrant chromosomal structures and enhanced nuclear abnormalities. Moreover, other pathways differentially regulated by LMW-E and FL-cycE, and/or those potentially fueling cancer development, may also provide druggable targets for the treatment of LMW-E overexpressing breast cancers.

3.2 Schematics of model system
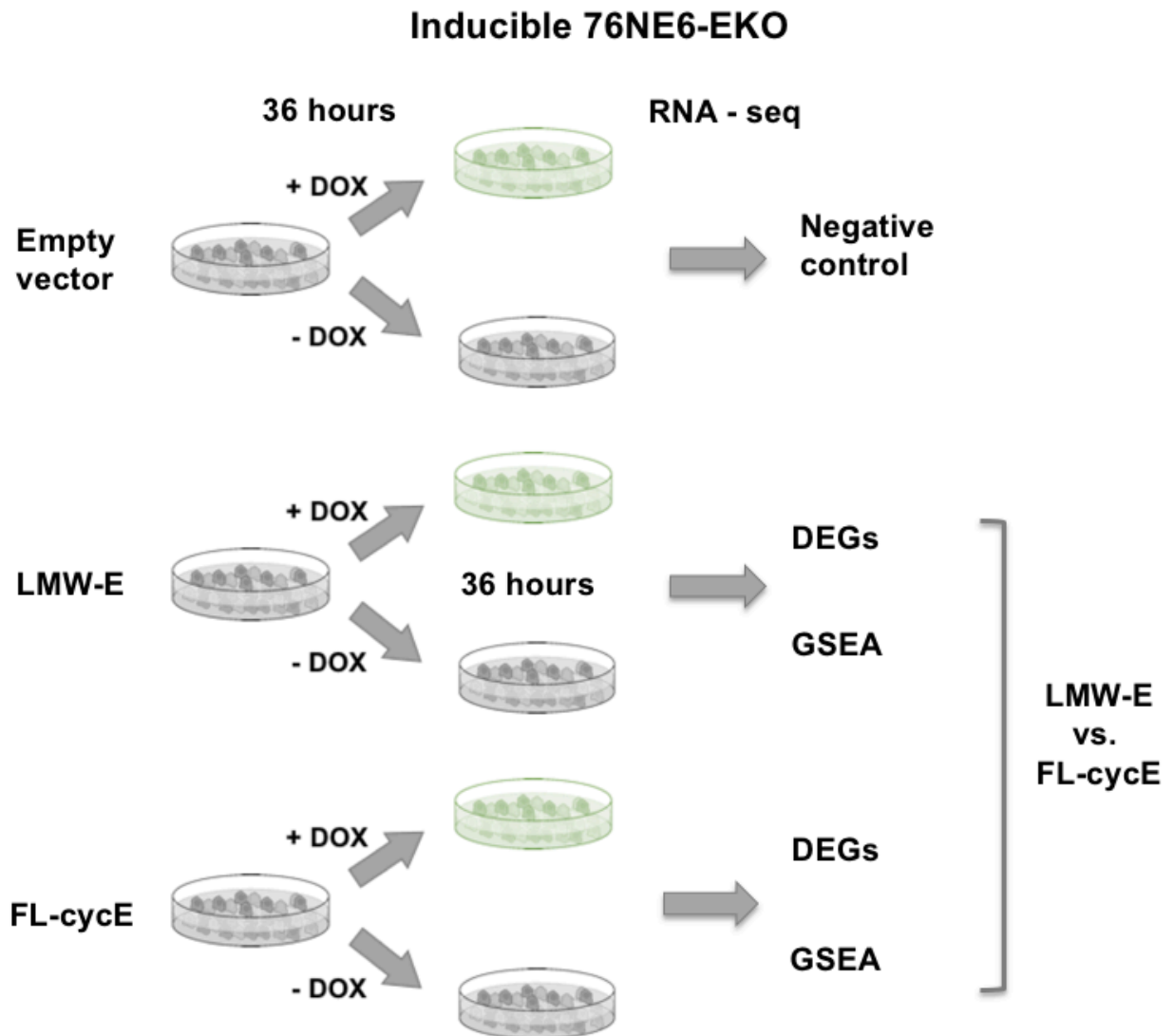
Figure 31. Schematics of experimental models to compare the effect of LMW-E versus FL-cycE on transcriptional profiles of hMECs. Inducible 76NE6-EKO cells treated for 36 hours in the presence or absence of doxycycline to induce the expression of empty-vector, FL-cycE or LMW-E. RNA-sequencing is performed to determine the differentially expressed genes (DEGs) between induced and un-induced cells, followed by gene-set enrichment analysis (GSEA), based on which the effect of LMW-E versus FL-cycE were also compared.

3.3 Materials and methods

Total RNA for RNA-sequencing were extracted from inducible 76NE6-EKO treated with 36 hours with 100ng/mL doxycycline to induce the expression of FL-cycE or LMW-E, empty vector cells, not harboring any transgenes, were used as negative controls. Cells cultured in the absence of doxycycline (DMSO, 36hours) were used as reference control. RNA extraction was performed by using RNeasy Kit (Qiagen, Cat. No.: 74004) according to manufacturer's protocol and quality control was performed by agarose gel electrophoresis both in lab and at Novogene before library preparation. Library construction was performed by Novogene (https://en.novogene.com/landing-page/amea-rnaseq/). Briefly, mRNA from the total RNA was purified using poly-T oligo-attached magnetic beads. The mRNA was first fragmented randomly by addition of fragmentation buffer. Then, the first strand cDNA was synthesized using random hexamer primer and M-MuLV Reverse Transcriptase (RNase H-). Second strand cDNA synthesis was subsequently performed using DNA Polymerase I and RNase H. Double-stranded cDNA was purified using AMPure XP beads. Remaining overhangs of the purified double-stranded cDNA were converted into blunt ends via exonuclease/polymerase activities. After adenylation of 3' ends of DNA fragments, NEBNext Adaptor with hairpin loop structure was ligated to prepare for hybridization. In order to select cDNA fragments of preferentially 150~200 bp in length, the library fragments were purified with AMPure XP system (Beckman Coulter). Finally, the PCR amplification and purification of PCR products by AMPure XP beads was performed to generate the libraries, which were then fed into illumina machines for RNA sequencing.

Raw data in FASTQ format were obtained from Novogene and then analyzed by The University of Texas MD Anderson Cancer Center Biostatistics Department. In brief, the dataset was initially filtered using counts-per-million (CPM) > 0.5 to remove the lowest expressed genes. We took log2 to generate the RNAseq count data from N=15960 highly

expressed genes. Trimmed Mean of the M-values (TMM) normalization was then performed to eliminate composition biases among the sample groups[134]. Following the voom transformation on normalized logCPM, we performed Limma liner model to compare the gene expression changes between LMW-E, and FL-cycE samples that were cultured in the presence versus absence of doxycycline[135]. We used the cut-off p-values at 0.05 for statistical significance and the p-value were adjusted by Benjamini-Hochberg Procedure[136]. Based on the t-statistic and adjusted p-value, the gene set enrichment analysis (GSEA) were then performed by using the following R-package:

https://bioconductor.org/packages/devel/bioc/manuals/fgsea/man/fgsea.pdf

3.4 Results

3.4.1 Differentially expressed genes induced by FL-cycE or LMW-E in hMECs

The gene expression in inducible 76NE6-EKO hMECs with or without doxycycline (100ng/mL, 36 hours) treatment were estimated by the abundance of transcripts (count of sequencing) that mapped to the human genome. The results revealed that within the 15960 genes successfully mapped to human genome, 207 genes were significantly up-regulated, and 132 genes were down regulated by FL-cycE. In inducible LMW-E cells, 1248 genes were up-regulated, and 1039 genes were down regulated by the overexpression of LMW-E. These results suggest LMW-E overexpression, compared with FL-cycE, resulted in more abundant transcriptional changes in inducible 76NE6-EKO cells.

Next, we aim to compare the differentially expressed genes induced by LMW-E overexpression or FL-cycE overexpression (Table 2 and Figure 32), which may explain why overexpression of LMW-E and FL-cycE lead to different phenotypes in inducible hMECs. From the list of most variable genes we have observed that *CCNE1* (the gene that encode Cyclin E) was the most differentially altered (and up-regulated) gene in both doxycycline treated

inducible 76NE6-EKO FL-cycE and 76NE6-EKO LMW-E cell lines, suggesting the successful induced expression of FL-cycE or LMW-E. Interestingly, for the rest of top ranked variable genes found in LMW-E overexpression cells compared to un-induced cells, we did not observe noticeable changes between FL-cycE dox+ and FL-cycE dox- cells (adjusted p < 0.05 and log2 fold change >1or <-1). The names for these genes are listed in Table 2, and the LogFC, p-value and adjusted p-values from the comparison between LMW-E dox + versus LMW-E dox-, and FL-cycE dox+ versus FL-cycE dox- are also specified (Table 2).

To investigate if the products (proteins) of these genes may contribute to the different phenotypes observed between LMW-E and FL-cycE overexpressing cells. Such phenotypes were characterized in Chapter 2, showing that FL-cycE but not LMW-E inhibits S-phase DNA replication, enhances DNA damage accumulation, and reduces cell viability. We performed a preliminary search for the functions and biological features of their encoding proteins in database (https://www.uniprot.org, Table 3). The results show the top two genes (except *CCNE1*) upregulated by LMW-E, *C17orf53* and *RAD51*, are encoding DNA binding-proteins involved in DNA damage repair. Further literature studies reveal both of them are required for the cell survival under replication stress[90,115,137,138], suggesting they may contribute to and are required for the cell survival under cyclin E induced replication stress. This hypothesis will be experimentally tested in Chapter 5. Additionally, we found the majority of the genes in the list are involved in DNA replication, such as pre-replication complex components *CDC6*, *CDT1*, *MCM5* and *MCM7*, DNA replication initiator *GINS1*, chromatin assembly regulator *ASF1B*, *CHAF1A*, *CENPW*, and *NCAPH*. We also found transcription factor *E2F1* and *E2F2* which drive the expression of DNA replication genes, and *TIMELESS* which is also involved in DNA replication, DNA damage repair and circadian clock. In chapter 4 of this study, we will examine the role of pre-replication complex, with a focus on *CDC6*, in LMW-E and FL-cycE overexpressing cells. Spindle assembly and G2/M checkpoints may also be regulated by

LMW-E, suggested by the function of *SHCBP1*, *CENPW* and *PKMYT1* shown in the list (Table 3).

Gene Name	LMW-E dox + versus LMW-E dox -			FL-cycE dox + versus FL-cycE dox -		
	logFC	P.Value	adj.P.Val	logFC	P.Value	adj.P.Val
CCNE1	3.12	3.27E-20	5.22E-16	2.42	2.43E-19	3.88E-15
C17orf53	2.33	4.63E-12	4.10E-09	0.46	0.0003353	0.02504753
RAD51	2.19	1.15E-12	1.53E-09	0.33	0.00216223	0.07607565
CDC6	1.92	1.95E-14	1.55E-10	-0.16	0.1038252	0.4709405
PKMYT1	1.90	5.55E-14	2.21E-10	0.19	0.03827059	0.3127141
SHCBP1	1.85	5.49E-13	8.76E-10	0.49	2.54E-05	0.00405604
MCM5	1.68	3.95E-12	3.71E-09	-0.36	0.00083335	0.04318069
ASF1B	1.68	4.28E-13	7.58E-10	0.21	0.0156161	0.2078187
CDT1	1.65	4.86E-14	2.21E-10	0.00	0.9700753	0.9920012
E2F1	1.65	2.54E-12	2.70E-09	-0.06	0.4427765	0.8003956
CHAF1A	1.64	2.21E-12	2.52E-09	0.02	0.8489793	0.9649489
CENPW	1.64	5.27E-12	4.43E-09	0.62	2.39E-07	0.00024064
GINS1	1.57	3.89E-13	7.58E-10	0.12	0.1069682	0.4758116
NCAPH	1.53	1.04E-12	1.51E-09	0.12	0.2041425	0.6030194
E2F2	1.53	3.25E-12	3.24E-09	0.56	6.57E-06	0.00160121
MCM7	1.46	3.61E-13	7.58E-10	0.02	0.8076163	0.9521019
CKS1B	1.42	7.18E-14	2.29E-10	0.31	0.00047106	0.03068639
TIMELESS	1.22	2.21E-12	2.52E-09	0.21	0.00716547	0.14318211
CDC20	1.10	4.04E-13	7.58E-10	0.31	0.00011077	0.01219204

Table 2. Summary of the top ranked variable genes between LMW-E dox + versus dox – in inducible 76NE6-EKO cells. For each of the genes, the logFC of expression, p-value, and adjusted p-value (adj.P val) by comparing LMW-E dox + versus LMW-E dox - as well as FL-cycE dox + versus FL-cycE dox - are listed.

Gene Name	Protein Name	Function suggested by Uniprot database	Potential therapeutic agents
C17orf53	Homologous recombination OB-fold protein	DNA-binding protein involved in homologous recombination that acts by recruiting the MCM8-MCM9 helicase complex to sites of DNA damage to promote DNA repair synthesis.	
RAD51	DNA repair protein RAD51 homolog	DNA-binding protein involved in homology search and homologous strand exchange. Exhibits DNA-dependent ATPase activity.	RAD51 inhibitor CYT-0851, B02
CDC6	Cell division control protein 6 homolog	DNA-binding protein involved in the initiation of DNA replication. Participates in checkpoint controls.	
PKMYT1	Membrane-associated tyrosine- and threonine-specific cdc2-inhibitory kinase	Negative regulator of entry into mitosis (G2 to M transition). Mediates phosphorylation of CDK1 predominantly on 'Thr-14'. Also involved in Golgi fragmentation.	PKMYT1 inhibitor RP-6306
SHCBP1	Testicular spindle-associated protein SHCBP1L	In association with HSPA2, participates in the maintenance of spindle integrity during meiosis in male germ cells.	
MCM5	DNA replication licensing factor MCM5	Component of the MCM2-7 complex (MCMs), the putative replicative helicase essential DNA replication initiation and elongation in eukaryotic cells.	
ASF1B	Histone chaperone ASF1B	Cooperates with chromatin assembly factor 1 (CAF-1) to promote replication-dependent chromatin assembly.	
CDT1	DNA replication factor Cdt1	DNA replication licensing factor, cooperates with CDC6 and the origin recognition complex (ORC) during G1 phase of the cell cycle to promote the loading of the mini-chromosome maintenance (MCM) complex onto DNA to generate pre-replication complexes (pre-RC).	
E2F1	Transcription factor E2F1	Transcription activator that binds the promoter region of a number of genes whose products are involved in cell cycle regulation or in DNA replication.	
CHAF1A	Chromatin assembly	Core component of the CAF-1 complex, a complex that is thought to mediate	

	factor 1 subunit A	chromatin assembly in DNA replication and DNA repair.	
CENPW	Centromere protein W	Component of the CENPA-NAC (nucleosome-associated) complex, a complex that plays a central role in assembly of kinetochore proteins, mitotic progression and chromosome segregation	
GINS1	DNA replication complex GINS protein PSF1	Component of the GINS complex, plays an essential role in the initiation of DNA replication, and progression of DNA replication forks.	
NCAPH	Condensin complex subunit 2	Regulatory subunit of the condensin complex, a complex required for conversion of interphase chromatin into mitotic-like condense chromosomes.	
E2F2	Transcription factor E2F2	Transcription activator that binds the promoter region of a number of genes whose products are involved in cell cycle regulation or in DNA replication.	
MCM7	DNA replication licensing factor MCM7	Component of the MCM2-7 complex (MCMs) , the putative replicative helicase essential DNA replication initiation and elongation in eukaryotic cells.	
CKS1B	Cyclin-dependent kinases regulatory subunit 1	Binds to the catalytic subunit of the cyclin dependent kinases and is essential for their biological function.	
TIMELESS	Protein timeless homolog	Plays an important role in the control of DNA replication, maintenance of replication fork stability, maintenance of genome stability throughout normal DNA replication, DNA repair and in the regulation of the circadian clock.	
CDC20	Cell division cycle protein 20 homolog	Required for full ubiquitin ligase activity of the anaphase promoting complex/cyclosome (APC/C) and may confer substrate specificity upon the complex. Is regulated by MAD2L1: in metaphase the MAD2L1-CDC20-APC/C ternary complex is inactive and in anaphase the CDC20-APC/C binary complex is active.	CDC20 inhibitor Apcin

Table 3. Functional information of proteins encoded by genes in Table 2. The

information was summarized from online uniprot database (https://www.uniprot.org).

Potential therapeutic agents targeting the genes in the list were also listed.
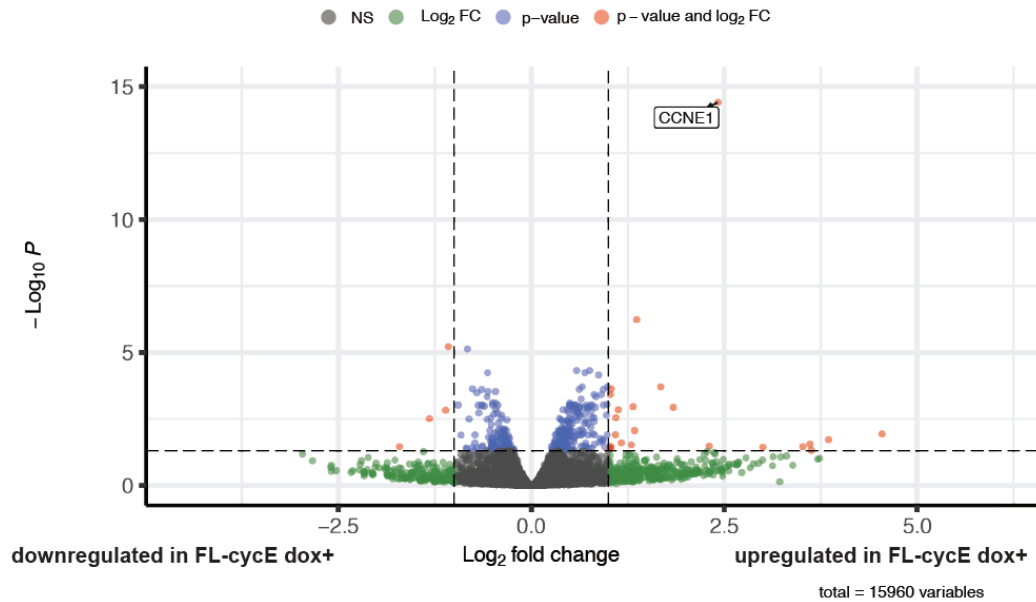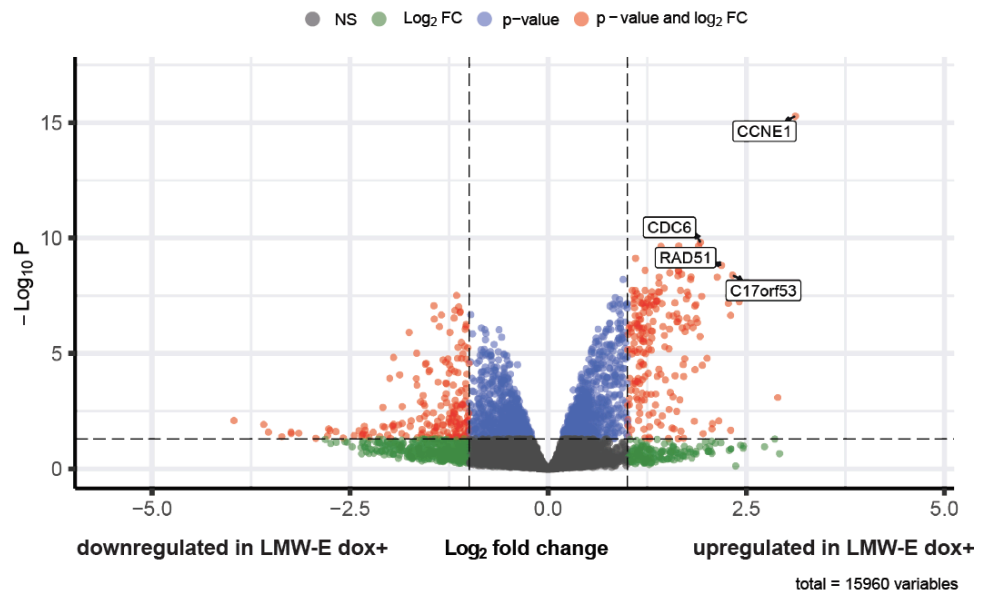
Figure 32. Volcano diagrams showing the overall distribution of differentially expressed genes induced by FL-cycE or LMW-E. A. FL-cycE dox+ versus dox- cells , and B. LMW-E dox+ versus dox- cells (panel B). *CCNE1* (gene encoding Cyclin E) is the most upregulated gene in both FL-cycE dox+ and LMW-E dox+ cells. *CDC6*, *RAD51*, and *C17orf53* were also identified to be upregulated by LMW-E. The roles of *CDC6*, *RAD51*, and *C17orf53* in LMW-E overexpression cells will be discussed in chapter 4 and 5. The threshold of differential expression genes is adjusted p value < 0.05 and log2 fold change (FC) >1 or <-1.

3.4.2 KEGG and HALLMARK gene-sets enriched by FL-cycE or LMW-E in hMECs

To further understand the impact of LMW-E and FL-cycE on signaling pathways in the inducible 76NE6-EKO cells, gene set enrichment analysis (GSEA) for KEGG pathways were performed using the differentially expressed gene induced by FL-cycE or LMW-E (Figure 33).

Five KEGG gene-sets were positively enriched in FL-cycE dox+ cells, compared with uninduced dox- cells. These gene-sets represent homologous recombination, p53 signaling pathway, cell cycle, oocyte meiosis and spliceosome pathways (Figure 33A). Because these five gene-sets were also up-regulated in LMW-E dox+ cells (Figure 33B), they may feature the common functions of FL-cycE and LMW-E. From the gene-sets specifically upregulated in LMW-E dox+ cells, we find DNA replication, pyrimidine metabolism, and purine metabolism, supporting our hypothesis that LMW-E but not FL-cycE facilitates DNA replication. We also observed that several DNA repair pathways, such as mismatch repair, nucleotide excision repair, and base excision repair were only positively enriched in LMW-E dox+ cells (Figure 33B), but not FL-cycE dox+ cells, suggesting LMW-E may upregulate DNA damage repair.

Additionally, because cell cycle pathway is the 1st ranking pathway enriched in LMW-E dox+ cells, but only ranked 3rd in FL-cycE dox+ cells, we further examined how FL-cycE or LMW-E overexpression changed the cell cycle genes expression in inducible 76NE6-EKO cells (Figure 34). The results show at 36 hours post overexpressing of FL-cycE or LMW-E (the time point when we harvested total RNA for RNA-sequencing), the majority of G2/M genes were upregulated by both FL-cycE and LMW-E. In detail, FL-cycE and LMW-E strongly upregulated *CDK1* and cyclin B. While *PLK1*, *MPS2*, *MAD2*, *BUBR1*, *BUB1*, *BUB2*, and *CDC20* were weakly upregulated by FL-cycE, they were strongly induced by LMW-E. In addition, LMW-E but not FL-cycE promoted the expression of *PKMYT1* and *WEE1*, the negative regulator of G2-M transition, suggesting an activated G2/M checkpoint in LMW-E

overexpressing cells and potential therapeutic targets for LMW-E overexpressing cells (Table 3, and previously reported WEE1 as druggable target[38]).

For the S phase genes, we found LMW-E but not FL-cycE strongly induced *CDK2* and cyclin A expression, as well as *CDC7* and *DBF4*, which are involved in replication activation. Additionally, *CHK1*, *CHK2*, *CDC25A* and *CDC25C* expression were all induced by LMW-E (Figure 34). In FL-cycE overexpressing cells, the expressions of S phase genes were not much changed (Figure 34). These results are in line with FL-cycE induced "rested S phase", observed in chapter two. Most strikingly, for the G1 phase genes, the MCM complex components were strongly upregulated by LMW-E but down-regulated by FL-cycE, suggesting FL-cycE and LMW-E may differentially regulate DNA licensing. Gene expression status in DNA replication pathway suggest that LMW-E induced the overexpression of DNA polymerase complex α, δ and ε, and DNA ligase LIG1. In addition to the DNA pre-replication complex members CDC6, CDC45, MCMs (MCM2, 3, 4, 5, 6, and 7) and its protein kinase CDC7-DBF4, we also observed the LMW-E up-regulated clamp protein PCNA and its loader RFC complex (Figure 35). These results strongly suggest LMW-E but not FL-cycE promote DNA replication, and this will be experimentally examined in Chapter 4.

Next, we performed GSEA for HALLMARK pathways using the differentially expressed genes induced by FL-cycE or LMW-E (Figure 36). We observed that gene-sets associated with the E2F-targets, G2M checkpoints, spermatogenesis, UV response up, mitotic spindle, and IL2-STAT5 signaling were positively enriched by both FL-cycE and LMW-E induction, while DNA repair gene-set was only upregulated by LMW-E induction. We selected top 10 DNA repair gene-set members upregulated by LMW-E and summarized their LogFC, p-value and adjusted p-values in Table 4 and listed their functional information in Table 5.

A

B
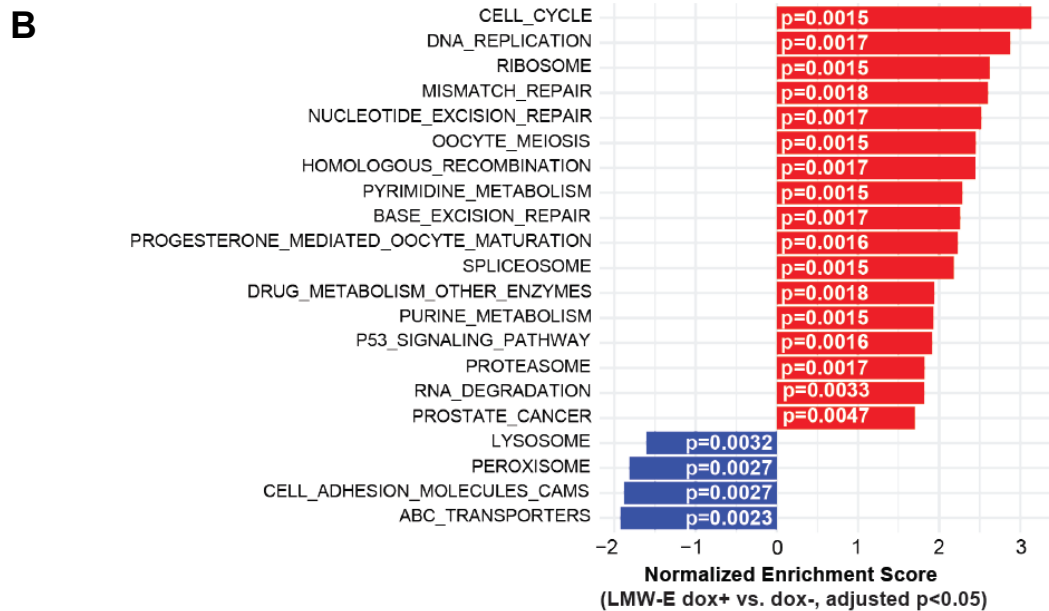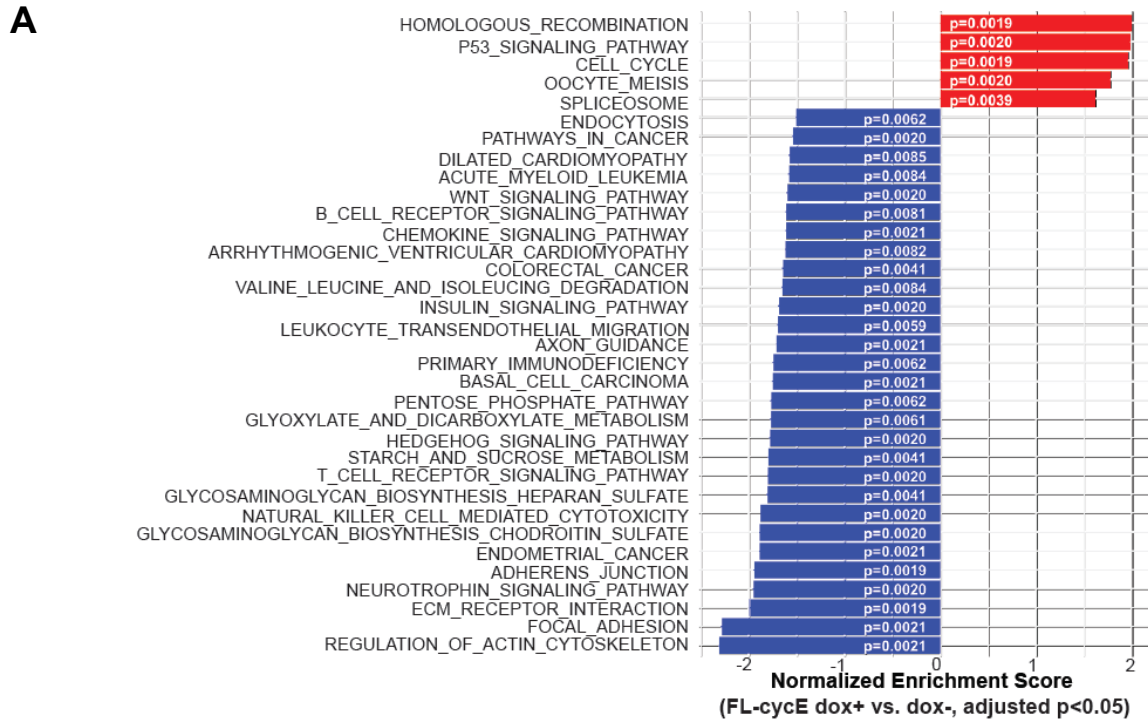
Figure 33. GSEA for KEGG pathways using data of the differentially expressed genes induced by FL-cycE or LMW-E. Gene-sets in red bars are up-regulated in FL-cycE dox+ cells (panel A) or LMW-E dox+ cells (panel B), compared with their own dox- controls, and blue bars indicate down-regulated gene-sets. Only the significantly enriched pathways (adjusted p value <0.05) are shown in the bar-graph.
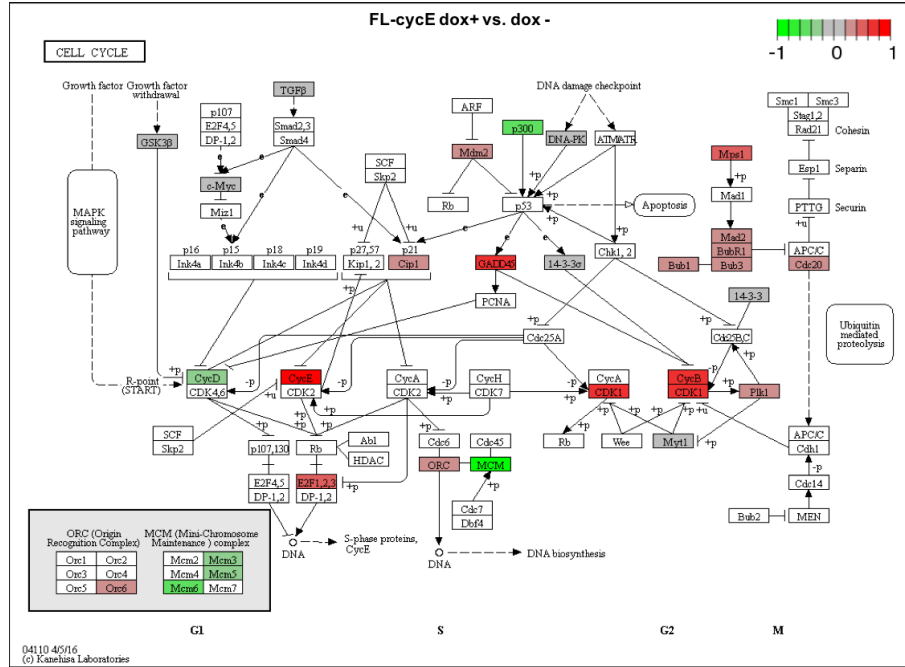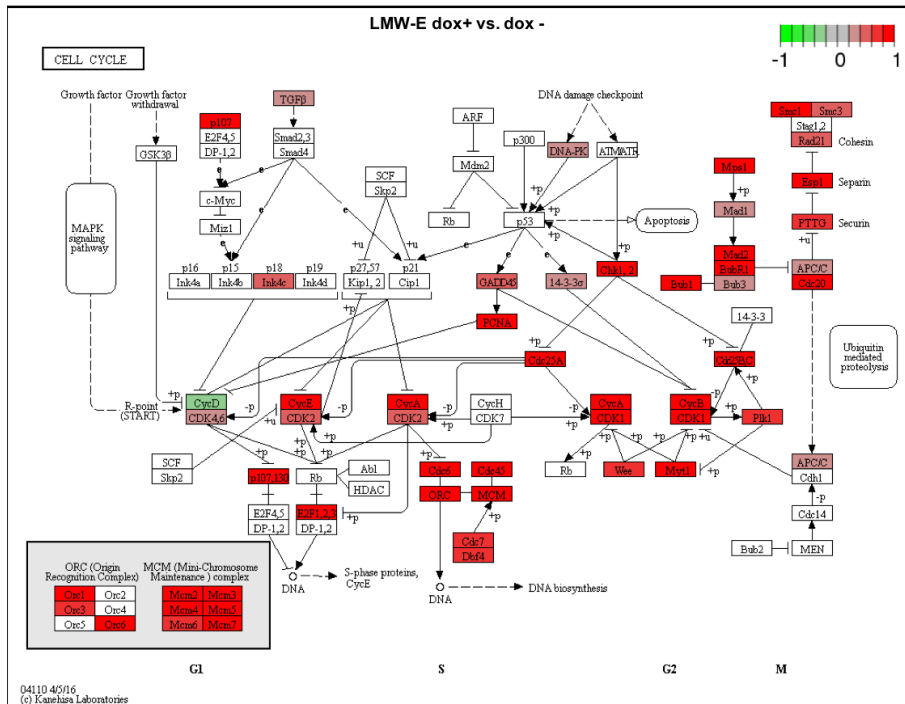
Figure 34. Expression status for KEGG DNA replication pathway members using data of the differentially expressed gene induced by FL-cycE or LMW-E. Genes labeled in red are up-regulated in FL-cycE dox+ cells (panel A) or LMW-E dox+ cells (panel B), compared with their own dox- controls. Genes labeled in green indicate down-regulation induced by FL-cycE (panel A) or LMW-E(panel B).

84

DNA replication related genes
LMW-E dox+ vs. dox-
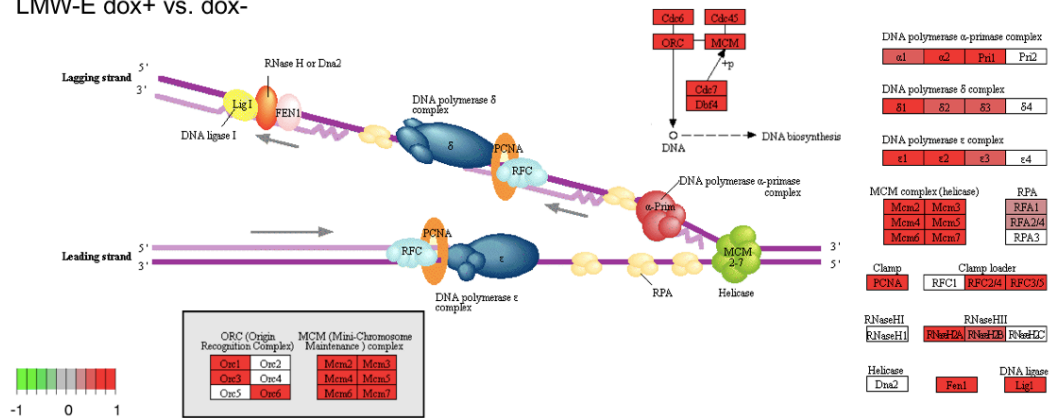
Figure 35. LMW-E overexpression upregulated the expression of genes involved in DNA replication. Induced LMW-E overexpressing (dox+) significantly upregulated genes (labeled in red) involved in DNA replication, such as DNA polymerase complex α, δ and ε, DNA ligase Lig1, DNA pre-replication complex members CDC6, CDC45, MCMs and protein kinase CDC7-DBF4, as well as DNA clamp protein PCNA and its loader RFC complex.

Gene Name	LMW-E dox + versus LMW-E dox -			FL-cycE dox + versus FL-cycE dox -		
	logFC	P.Value	adj.P.Val	logFC	P.Value	adj.P.Val
RAD51	2.19	1.15E-12	1.53E-09	0.33	0.00216223	0.07607565
ZWINT	1.38	1.63E-10	5.31E-08	0.22	0.01805183	0.2237846
FEN1	1.29	8.88E-11	3.49E-08	0.14	0.07231864	0.4037011
LIG1	1.28	3.84E-10	8.63E-08	0.07	0.4132822	0.780306
RFC3	1.23	1.49E-07	1.03E-05	0.31	0.00751017	0.14406531
PRIM1	1.10	4.00E-06	0.00015891	-0.14	0.284263	0.6856336
POLA2	1.08	4.25E-08	3.63E-06	0.03	0.7712624	0.9387721
RFC2	1.06	1.55E-07	1.06E-05	0.00	0.9903804	0.9971901
RFC4	1.05	6.76E-09	8.11E-07	-0.05	0.5982077	0.8789164
RFC5	1.03	4.27E-09	5.82E-07	0.07	0.3945042	0.7667077

Table 4. Summary of the top 10 upregulated DNA repair gene-set members induced by LMW-E. For each of the genes, the logFC, p-value, and adjusted p-value by comparing LMW-E dox + versus LMW-E dox - as well as FL-cycE dox + versus FL-cycE dox - are listed.
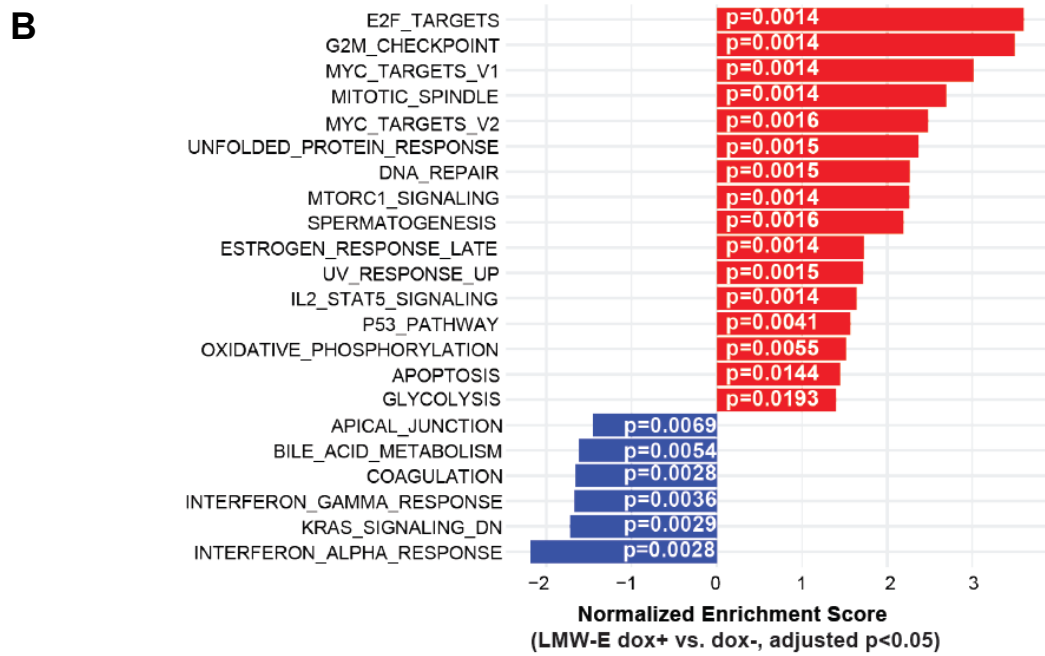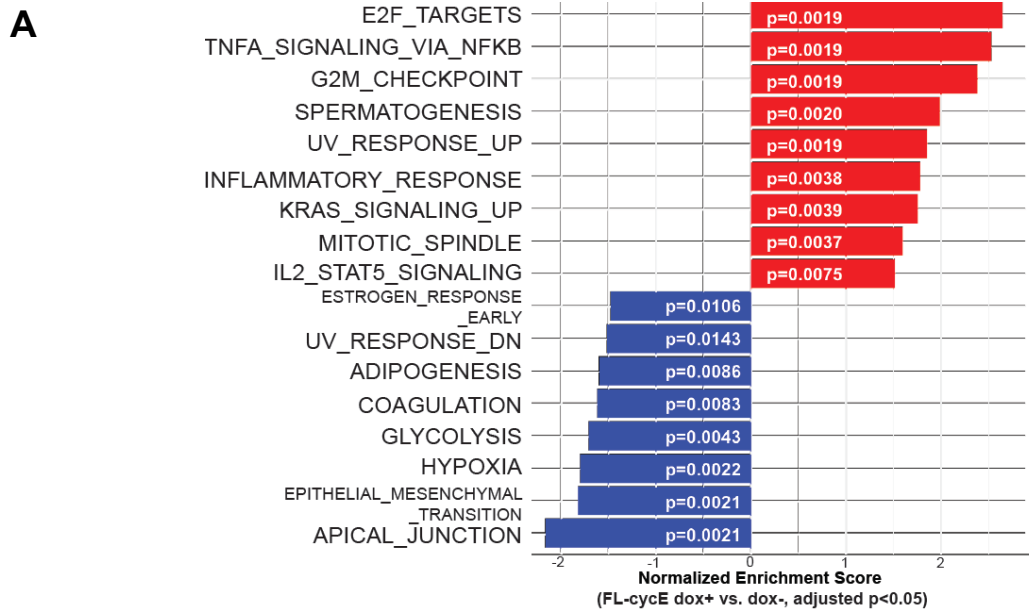
Figure 36. GSEA for HALLMARK pathways using data of the differentially expressed genes induced by FL-cycE or LMW-E. Gene-sets in red bars were up-regulated in FL-cycE dox+ cells (panel A) or LMW-E dox+ cells (panel B), compared with their own dox- controls, and blue bars indicate down-regulated gene-sets. Only the significantly enriched pathways (adjusted p value <0.05) are shown in the bar-graph.

Gene Name	Protein Name	Function suggested by Uniprot database
RAD51	DNA repair protein RAD51 homolog	DNA-binding protein involved in homology search and homologous strand exchange. Exhibits DNA-dependent ATPase activity.
ZWINT	ZW10 interactor	Part of the MIS12 complex, required for kinetochore formation and spindle checkpoint activity, targeting ZW10 to the kinetochore at prometaphase.
FEN1	Flap endonuclease 1	Structure-specific nuclease with 5'-flap endonuclease and 5'-3' exonuclease activities involved in DNA replication and repair.
LIG1	DNA ligase 1	DNA ligase that seals nicks in double-stranded DNA during DNA replication, DNA recombination and DNA repair.
RFC3	Replication factor C subunit 3	Subunit 3 of Replication factor C (RFC), a complex that opens the sliding clamp and loads it onto the DNA chain, critical for DNA synthesis.
PRIM1	DNA primase small subunit	Catalytic subunit of the DNA primase complex and component of the DNA polymerase alpha complex
POLA2	DNA polymerase alpha subunit B	Subunit of the DNA polymerase alpha complex, plays an essential role in the initiation of DNA synthesis
RFC2	Replication factor C subunit 2	Subunit 2 of Replication factor C (RFC), a complex that opens the sliding clamp and loads it onto the DNA chain, critical for DNA synthesis.
RFC4	Replication factor C subunit 4	Subunit 4 of Replication factor C (RFC), a complex that opens the sliding clamp and loads it onto the DNA chain, critical for DNA synthesis.
RFC5	Replication factor C subunit 5	Subunit 5 of Replication factor C (RFC), a complex that opens the sliding clamp and loads it onto the DNA chain, critical for DNA synthesis.

Table 5. Functional information of proteins encoded by genes in Table 4. The

information was summarized from online uniprot database (https://www.uniprot.org).

3.5 Conclusion

In this chapter, we performed RNA sequencing to compare the transcriptional alterations induced by FL-cycE versus LMW-E. The results show the genes associated with KEGG cell cycle gene-set and HALLMARK E2F targets gene-set were enriched in both FL-cycE and LMW-E overexpressing cells (Figure 37), consistent to the observation that FL-cycE and LMW-E promotes G1/S transition. However, there were specific differences between the genes induced by LMW-E or FL-cycE. For example, the enrichment plots for KEGG DNA replication and HALLMARK DNA repair gene-sets showed significant enrichment for LMW-E expression ($p = 0.0017$ and $p = 0.0015$, respectively) but not for FL-cycE expression ($p = 0.41$ and $p = 0.99$, respectively (Figure 38).

By comparing the specific gene expression changes mediated by LMW-E with those altered by FL-cycE, we identified CDC6, RAD51, and C17orf53 genes, which are essential to DNA replication and damage repair in response to replication stress, were strongly upregulated by LMW-E but not FL-cycE.

These results led to the hypothesis that LMW-E–mediated upregulation of CDC6, RAD51, and C17orf53 may be required for replication stress tolerance and cell viability in the cells overexpressed with LMW-E.

A **KEGG CELL CYCLE**
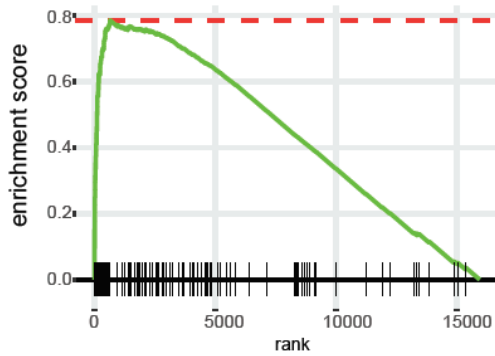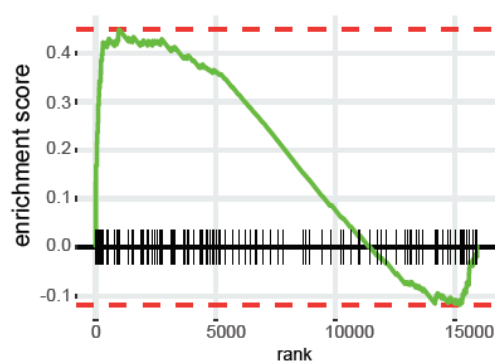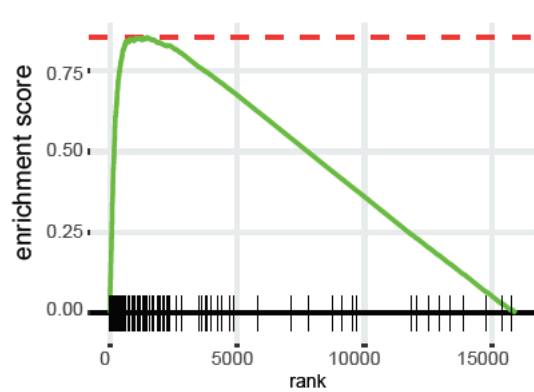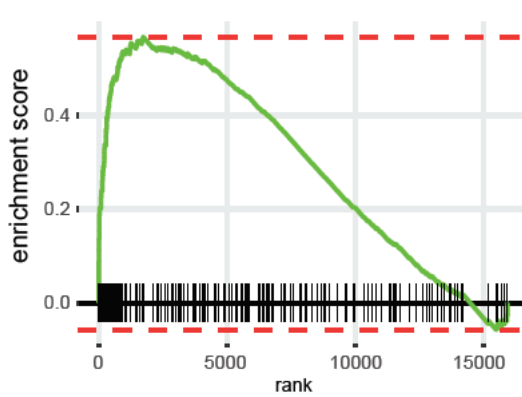
Figure 37. KEGG cell cycle gene-set and HALLMARK E2F targets gene-set were enriched in both FL-cycE and LMW-E overexpressing cells. A. Enrichment plot for KEGG cell cycle gene set in LMW-E dox+ group compared to LMW-E dox- group (upper panel) and FL-cycE dox+ group compared to FL-cycE dox- group **B**. Enrichment plot of HALLMARK E2F targets gene set in LMW-E dox+ group compared to LMW-E dox- group and FL-cycE dox+ group compared to FL-cycE dox- group.

A   **KEGG DNA REPLICATION**

B   **HALLMART DNA REPAIR**
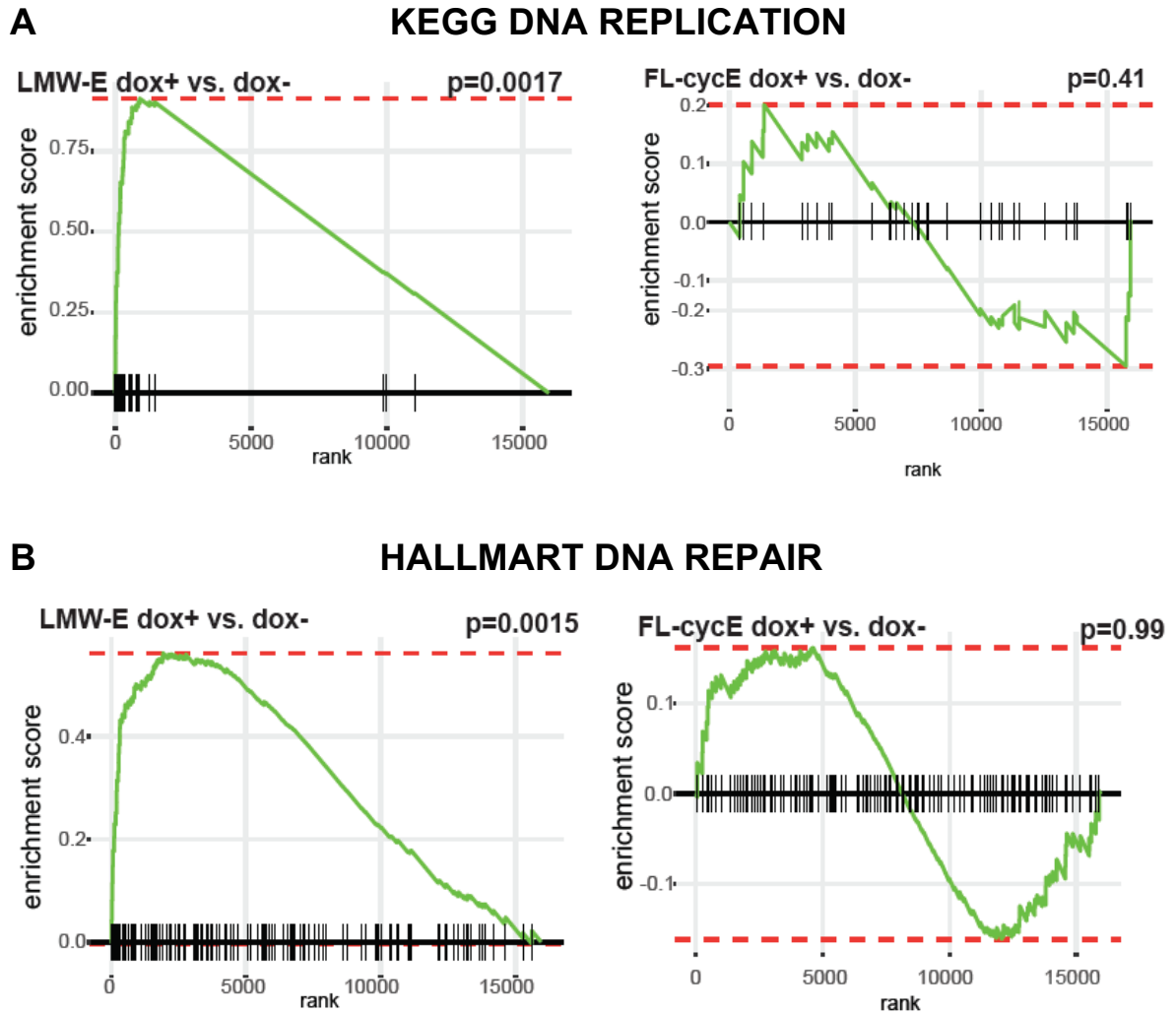
Figure 38. KEGG DNA replication and HALLMARK DNA repair gene-sets were specifically enriched in LMW-E overexpressing cells. A. Enrichment plot for KEGG DNA replication gene-set in LMW-E dox+ group compared to LMW-E dox- group and FL-cycE dox+ group compared to FL-cycE dox- group. **B**. Enrichment plot of HALLMARK DNA repair gene-set in LMW-E dox+ group compared to LMW-E dox- group and FL-cycE dox+ group compared to FL-cycE dox- group.

Chapter Four: LMW-E but not FL-cycE facilitated pre-replication complex assembly

4.1 Introduction

In normal dividing cells, DNA replication is tightly coordinated with the cell cycle to ensure that the genome is faithfully duplicated[92]. Cyclin E, in complex with cyclin-dependent kinase 2 (CDK2), promotes G1/S transition and the initiation of DNA replication. The first step in replication initiation is the loading of hexametric mini-chromosome maintenance 2-7 (MCM2-7) complex, recruited by CDT1 and CDC6 to the origin recognition complex (ORC) at replication origins. This process is termed as pre-replication complex (pre-RC) assembly or replication licensing[94,95]. The precise role of cyclin E in replication licensing is a subject of debate. Both promoting and inhibiting effect have been reported in different experimental systems[94,95]. Replication licensing peaks at G1, and is inhibited at S phase to prevent re-replication[95].

Activation of oncogenes, such as cyclin E, HRAS, and MYC, cause replication fork stalling, often referred as replication stress[88]. Depletion of nucleotide pools and collision between replication machinery with transcription complexes are proposed models for cyclin E induced replication stress, although the detailed mechanisms remain unknown[100,101]. It is worth noting that cyclin E is shown to be stabilized at replication origin under replication stress, and chromatin bound cyclin E exhibits low molecular weight bands in western blotting [87]. However, whether or not LMW-E is a product from FL-cycE under replication stress, or if LMW-E directly regulates proteins at replication origin (such as MCMs) is currently unknown.

In pre-cancerous cells, replication stress serves as a tumor barrier. Fork stalling may lead to fork collapse and under-replicated DNA region, activating DNA damage responses and S phase checkpoint, which in turn cause senescence or cell death[105,106]. In fact,

overexpression of cyclin-E in MCF10A (an immortalized non-tumorigenic mammary epithelial cell line) stalls replication fork, inhibits cell growth and causes cell death or senescence[105,106].

Previous reports also suggest excessive replication licensing in G1 provides dormant replication origins in S phase. Under normal conditions, these dormant origins are not activated during DNA synthesis, but they serve as back-up origins important for completing DNA replication in response to the stalling of an activated replication fork nearby, thereby promoting cell survival under replication stress[85]. Our analyses from RNA sequencing using inducible 76NE6-EKO cell lines revealed that LMW-E may trigger the expression of DNA replication genes including pre-replication complex factors (see Chapter 3). We thus hypothesize that the replication licensing might be up-regulated by LMW-E, while down-regulated by FL-cycE, which lead to the different cell viability observed in FL-cycE and LMW-E overexpressing cells (see Chapter 2).   In this chapter, we set out to examine the roles of LMW-E versus FL-cycE in regulating DNA replication and pre-replication complex loading to chromosomes (Figure 39).

4.2 Schematics of model system
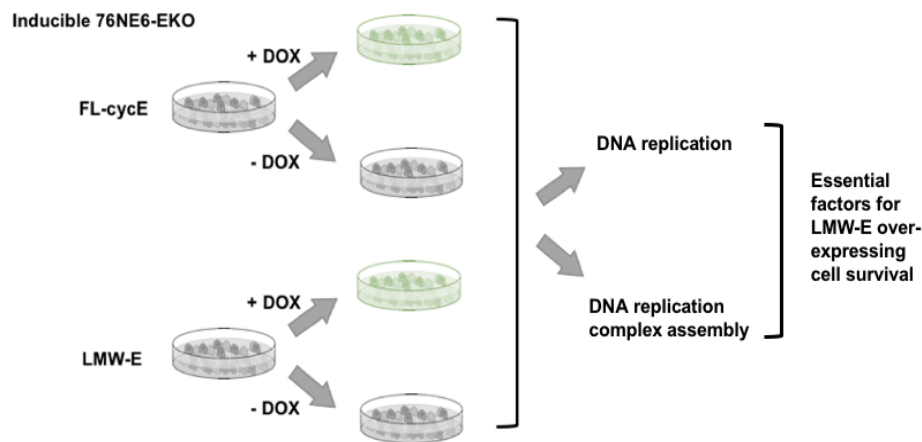
Figure 39. Schematics of experimental models to examine the roles of LMW-E versus FL-cycE in regulating DNA replication and pre-replication complex assembly.

4.3 Materials and methods

Western blot

Cell lysates were prepared and subjected to western blot analysis as previously described (Chapter 2). We used the following primary antibodies: cyclin E (HE-12; Santa Cruz Biotechnology, SC-247); CDC6 (180.2; Santa Cruz Biotechnology, SC-9964); CDK2, (BD Biosciences-Transduction Laboratories, 010146); CHK1 (G-4; Santa Cruz Biotechnology, SC-8408); phospho-CHK1 Ser345 (Cell Signaling, #2348); RPA32 (Cell Signaling, #22085); phospho-RPA32 Ser4/8 (Bethyl, A300-245a); phospho-RPA32 Ser33 (Bethyl, A300-246a); MCM2 (Cell Signaling, #4007); MCM4 (Cell Signaling, #12973); MCM7 (Cell Signaling, #3735); Rad51 (H-92; Santa Cruz Biotechnology, SC-8349); C17orf53 (Sigma, HPA023393); and Vinculin (Sigma, V9131). Densitometry was performed by using ImageJ software, the value for each band was normalized to the first visible band in the same gel.

Isolation of chromatin-bound proteins from subcellular fractions

Cell fractionation assays were performed using the protocol provided by Abcam (https://www.abcam.com/protocols/subcellular-fractionation-protocol). Briefly, the freshly harvested cell pellets were resuspended using fractionation buffer (20mM HEPES ph7.4, 10mM KCl, 2mM MgCl2, 1mM EDTA and 1mM EGTA, supplemented with 1 x protease inhibitor cocktail), followed by passing the cell suspension through #27gauge needle 10 times until the cells are fully lysed. After washing three times with fractionation buffer followed by centrifugation (1,000 × g at 4 °C for 5 min), the nuclei fraction (pellet) was harvested. The isolation of chromatin-bound proteins from nuclear fraction were performed by washing the nuclear pellets with solubilization buffer (10 mM Tris-HCl pH8.0, 200 mM NaCl, 1 mM EDTA, 1 mM EGTA, supplemented with 1 x protease inhibitor cocktail). After centrifugation (16,000 × g at 4 °C for 10 min), the soluble proteins in the supernatant were collected. The insoluble chromatin fractions were homogenized in RIPA buffer (50mM Tris-HCl pH 7.6, 150mM NaCl,

1% Triton X-100, 0.5% sodium deoxycholate, 0.1% SDS and mixed with fresh protease inhibitor cocktail), and sonicated on ice in 30 seconds on, 30 seconds off intervals for 5 minutes. The resultant homogenates were then subjected to western blot analysis.

Immunoprecipitation and immunoblotting

The total cell lysates were prepared in ice-cold lysis buffer (5mM EDTA; 150mM NaCl; 50mM Tris, pH 8.0; 1% Triton X-100; 0.1% SDS and protease inhibitor cocktail) and placed on ice. After 15 minutes of lysis, brief sonication is performed followed by centrifugation at 13000rpm at 4º C for 15 minutes. We used 1000μg of cell extract for each immunoprecipitation with polyclonal antibodies to cyclin E generated in lab[132], CDC6 (180.2; Santa Cruz Biotechnology, SC-9964), or normal mouse IgG (Santa Cruz Biotechnology, SC-2025). After antibody incubation at 4º C overnight, the protein A- or protein G-coated beads (CalBiochem) were added and incubated at 4º C for 1 hour. The beads were subjected to centrifugation at 8000rpm at 4º C for 30 seconds, and washed in ice-cold lysis buffer for four times (3 minutes each). The resultant immune-precipitates were subjected to western blotting with the indicated antibodies for each experiment.

Specific siRNAs and smart pool siRNAs

All siRNAs (specific siRNAs and smart pool siRNAs) used in this study are products from Dharmacon (Horizon Discovery). The transfection was performed by using Lipofectamine 3000 (Thermofisher) according to manufacturer's protocol. The knock-down efficiency of specific siRNAs and smart pool siRNAs were examined by western blot analysis for target proteins using total cell lysates harvested 48 - 72 hours post transfection of siRNAs. The target sequence for each siRNAs used in this chapter is listed as follows:

CDC6 siRNA#1: GAGAUCAGGUUCUGGACAA, #2: GGAAACGUCUGGGCGAUGA;

CDC6 siRNA smart pool: GAGAUCAGGUUCUGGACAA, GCUACUGGAUUGCCUUAAA, GGAAACGUCUGGGCGAUGA, UCAAUUCUGUGCCCGCAAA

Immunofluorescence staining

Immunofluorescence (IF) staining were performed using the primary antibodies: anti-γ-H2AX (Millipore, 05-636) and anti-53BP1 (Novus, NB100-304), based on protocol described previously (Chapter 2). The images were taken using Zeiss LSM880 Confocal at 63X objective magnification and examined by Zeiss Zen software. More than 600 cells were analyzed for each group, and cells with five or more foci were considered positive.

DNA fiber assay

Cells were plated into a p100 plate (1.5 x 10^5) on day 0, such that by the end of the experiment (48 hours from seeding) that the confluency is a 50%. One hour before harvesting, cells were pulse labeled with 50 µM iodo-deoxyuridine (IdU) for 30 minutes and with 100 µM chloro-deoxyuridine (CIdU) for 30 minutes after the removal of IdU (the chase period). 400,000 labeled cells were collected and diluted 1:10 with 1mL unlabeled cells. 3µl of cell mixture were mixed with 7µl of lysis buffer (200mM Tris-HCl, pH7.5, 50mM EDTA, 0.5% SDS) on glass slides (Fisherbrand Superfrost Plus) and incubated for 5 minutes at room temperature. After the DNA were spread on slides tilted at 30 degrees, the samples were air dried and fixed in methanol/acetic acid (3:1) for 15 minutes, denatured with 2.5M HCl for 1 hour and then neutralized by washing in PBS. Following the washes, the slides were blocked in 5% BSA for 1 hour and incubated in primary antibodies: rat anti-bromodeoxyuridine (1:500; detects CIdU; Abcam ab6326) and mouse Anti-BrdU (1:1000; detects IdU; BD Biosciences 347580) for overnight at 4°C, followed by secondary antibody incubation (Alexa Fluor 594 or 660 EMD Millipore diluted at 1:1000) for 2 hours. The mounted slides were examined by Zeiss LSM880

Confocal at 63X objective magnification and DNA fiber length was measured by Zeiss Zen software (>100 fibers per group).

4.4 Results

4.4.1 LMW-E facilitated replication stress tolerance

To directly test if FL-cycE or LMW-E causes replication stress, we examined the effects of FL-cycE or LMW-E on the fork speed by DNA fiber assay. Inducible 76NE6-EKO cell lines were treated with doxycycline for 0, 24, and 48 hours, followed by pulse-labeling with the thymidine analog iododeoxyuridine (IdU) for 30 minutes and chase-labeling with the thymidine analog iododeoxyuridine (CldU) for 30 minutes. The DNA fibers were then allowed to spread out, followed by CldU and IdU detection by immunostaining using specific antibodies. We compared the length of the DNA fibers and calculated the replication fork speed by the total labeling time for each fiber. The results suggested that without doxycycline induction, the replication fork speeds were very similar between inducible FL-cycE and LMW-E cells, for which the mean values were 54μm/hour and 58μm/hour respectively (Figure 40 A and B). Following 24 hours of induction, the mean fork speed in the FL-cycE cells dropped to 31 μm /hour (p value<0.0001). Similarly, LMW-E overexpression for 24 hours led the fork speed drop to 34 μm /hour (p value<0.0001). These finding suggest the changes between the fork speeds at 0-hour time point and 24-hour time point were comparable between LMW-E and FL-cycE expressing cells. Both FL-cycE or LMW-E overexpression resulted in 40% decrease of the fork speed, indicating similar amount of replication stress were induced by FL-cycE and LMW-E at 24 hour-time point.

However, the changes between the fork speeds at 24-hour time point and 48-hour time point were significantly different in FL-cycE and LMW-E cells. The mean fork speed continued to drop over time to 8 μm /hour in FL-cycE overexpressing cells, whereas in LMW-E overexpressing cells the fork speed remained relatively constant at 28 μm /hour. FL-cycE

96

expressing cells show a 74% reduction of the fork speed compared to 24-hour time point.

LMW-E expressing cells showed only an 18% reduction (Figure 40). These data suggest that

both FL-cycE and LMW-E expression can slow-down the speed of replication fork traverse,

one of the key features of replications stress. However, while replication stress continued to

accumulate in FL-cycE cells, LMW-E cells exhibited relative mild replication stress in later

time point.
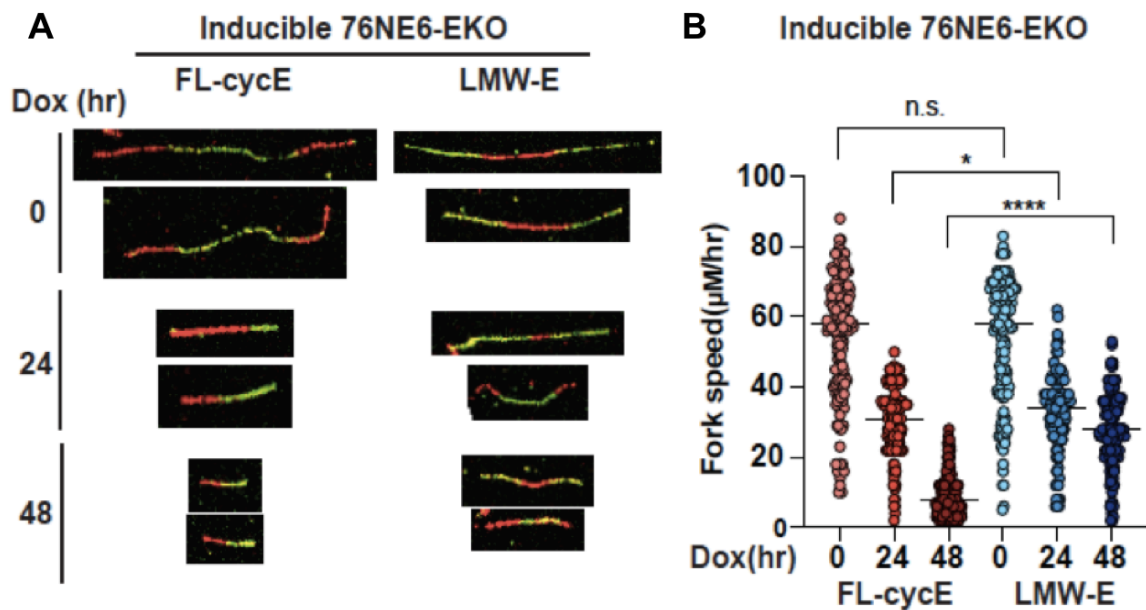
Figure 40. LMW-E prevented replication stress accumulation in inducible 76NE6-EKO

cells. DNA fiber assay using inducible 76NE6-EKO cells under the indicated expression

conditions (panel A) and the calculated replication fork speed (panel B). Both FL-cycE and

LMW-E reduced replication fork speed to a similar level at 24 hours post induction. At 48 hours

post induction, the fork progression was further attenuated in FL-cycE overexpressing cells.

LMW-E cells exhibited relative mild replication stress at 48 hours post induction (n = 3, fiber

number > 150, mean with standard deviation, *p < 0.05, **p < 0.01, ***p < 0.001, and ****p <

0.0001; n.s. indicates not significant; Student *t* test)

Under replication stress, the stalled replication fork may be followed by uncoupling of DNA polymerases from the MCMs complex. This leads to the coating of RPAs on exposed ssDNA, and further activates ATR-CHK1 pathway, resulting in phosphorylation of RPAs[50]. We thus examined the effect of LMW-E and FL-cycE on the proteins associated with replication stress responses, following time of FL-cycE or LMW-E induction (0, 12, 24, 36 and 48 hours). Western blot analysis was performed to detect the phosphorylation of CHK1 and RPA. The results revealed that ATR-dependent phosphorylation of CHK1 (Ser 345) were upregulated by both FL-cycE and LMW-E, suggesting both FL-cycE and LMW-E induced replication stress. However, phosphorylation of RPA (Ser 33 and Ser4/8) was strongly induced by FL-cycE but not LMW-E, indicating un-resolved replication stress in FL-cycE overexpressing cells (Figure 41A). Particularly, at 48 hours post induction of FL-cycE or LMW-E. the level of RPA pS4/8 and pS33 were at least 10-fold higher in FL-cycE overexpressing cells compared with LMW-E cells. Consistent results were also observed in 76NF2V-EKO cell lines (Figure 41B), in which phosphor-RPA signals continued to increase in FL-cycE cells, while initially increased but then decreased in LMW-E overexpressing cells. These data collectively suggested that replication stress accumulated in FL-cycE cells but resolved in LMW-E overexpressing cells.

To test whether LMW-E actively mediates tolerance from replication stress, we treated cells with hydroxyurea (HU) as an external agent to induce replication stress in 76NE6-EKO cells with or without prior induction of LMW-E. These cells were initially grouped by doxycycline treatment to induce LMW-E (dox=100ng/mL, 24 hours) and DMSO for un-induced controls. Following 24 hours of induction, HU were added to the culture media at a final concentration of 5mM and each of the LMW-E induced or uninduced groups were subsequently cultured for 3 hours in the presence or absence of HU. These cells were then subjected to DNA fiber assays as described earlier. The results showed that 24 hours of LMW-

E induction prior to 3 hour of HU treatment can partially rescue the replication stress induced by HU (Figure 42). Collectively, these results suggest that both FL-cycE and LMW-E can induce replication stress. However, LMW-E but not FL-cycE facilitated replication stress tolerance. In FL-cycE overexpressing cells replication stress continue to accumulate, which may lead to subsequent DNA damage beyond repair.
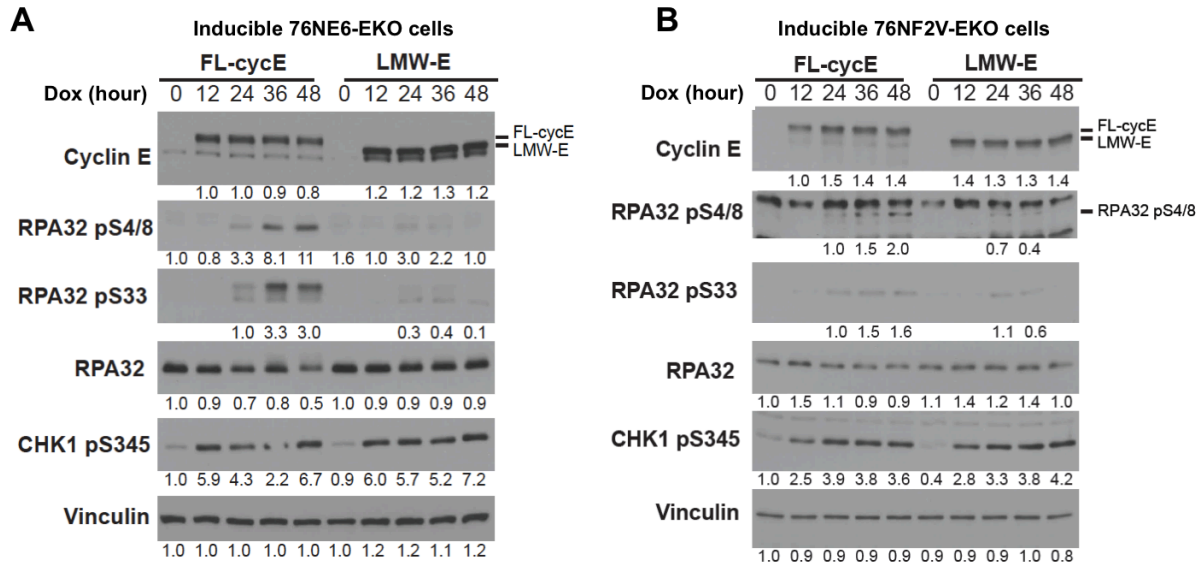
Figure 41. Replication stress markers accumulated in FL-cycE overexpression cells but resolved in LMW-E overexpressing cells. Western blot analysis of replication stress markers phosphorylation of CHK1 (pS345) and of RPA (Ser 33 and Ser4/8) in inducible 76NE6-EKO cells(A) and 76NF2V-EKO cells(B) with or without overexpression of LMW-E or FL-cycE. Cells were treated with 100ng/ml doxycycline in a time course manner to induce the expression of LMW-E or FL-cycE. Uninduced control (Dox 0) were treated with DMSO for 48 hours. The results suggested phosphor-RPA signals continued to increase in FL-cycE cells, while initially increased but then decreased in LMW-E overexpressing cells. Densitometry were performed by using ImageJ software, the value for each band was normalized to the first visible band in the same gel.
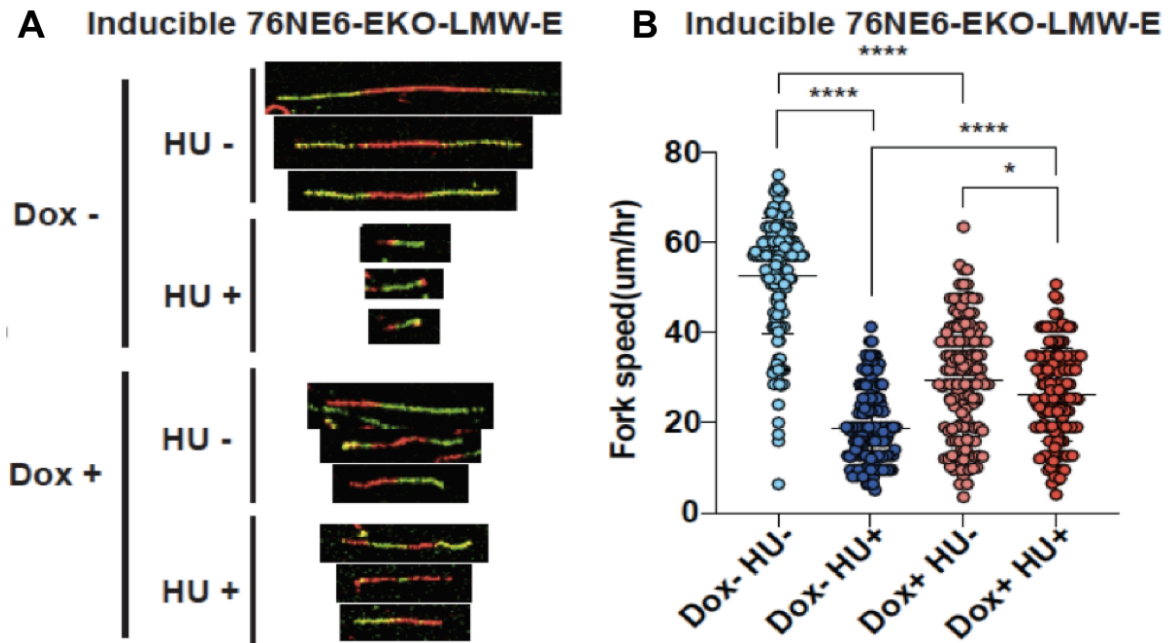
Figure 42. LMW-E overexpression rescued the replication stress induced by hydroxyurea. DNA fiber assay were performed to measure the replication fork speed in inducible 76NE6-EKO LMW-E cells. The cells were initially grouped into those overexpressing LMW-E (100 ng/mL doxycycline for 24 hours; Dox+) or not (DMSO for 24 hours; Dox-). This was followed by adding hydroxyurea (HU, 5mM, 3 hours) or vehicle (water) into the culture media. Representative DNA fibers (A) and calculated fork speed (B) are shown. (n = 3, fiber number > 100, mean with standard deviation). For all statistical analyses, *p < 0.05, and ****p < 0.0001; n.s. indicates not significant; Student *t* test.

4.4.2 LMW-E facilitated pre-replication complex assembly

In the process of DNA replication, CDC6 promotes replication licensing by recruiting the loading of MCMs complexes to the chromatin. Excess MCMs loading are required for cells to survive replicative stress[94,95]. Western blot analysis confirmed that induction LMW-E upregulated CDC6 protein in 76NE6-EKO cells by 1.6-fold, while induction of FL-cycE resulted in a 30% decrease in CDC6 expression (Figure 43). These results suggest differentially regulated replication licensing might be responsible for the distinct replication stress intensity

between LMW-E versus FL-cycE overexpressing cells. We also observed LMW-E overexpression upregulated RAD51 and C17orf53. Compared with FL-cycE overexpressing cells, the protein level of RAD51 and C17orf53 are 100% higher in LMW-E overexpressing cells. These results also confirmed our findings by RNA-sequencing (Table 2).
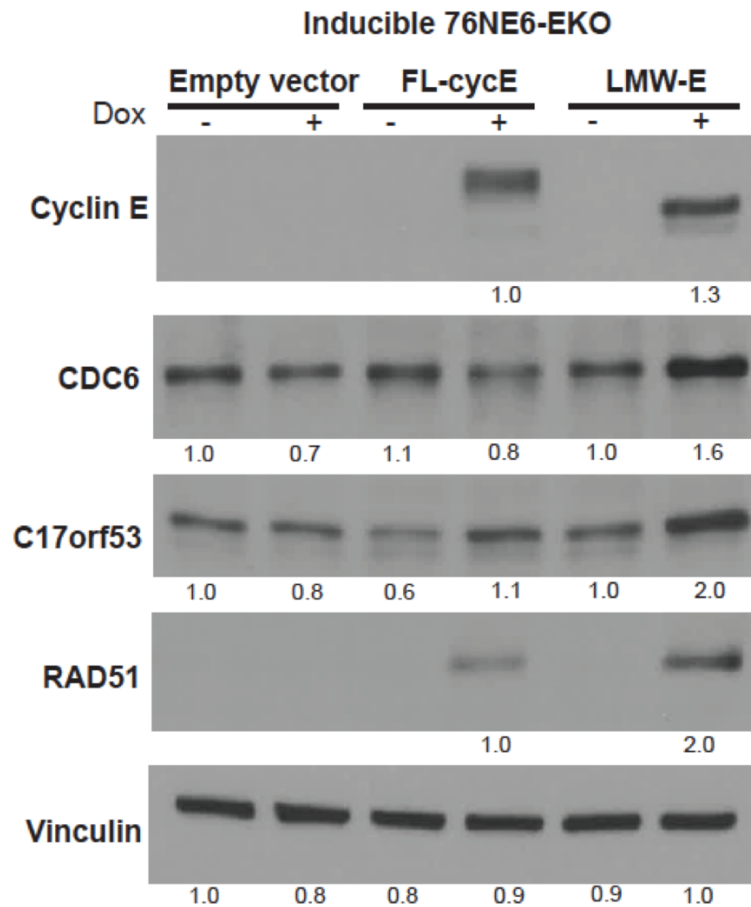
Figure 43 LMW-E upregulated CDC6, RAD51 and C17orf53. Western blotting analysis of cyclin E (FLcycE or LMW-E) and DNA pre-replication complex factor CDC6 in inducible 76NE6-EKO with or without doxycycline (100ng/mL, 24 hours). LMW-E upregulated CDC6 protein in 76NE6-EKO cells by 1.6-fold, while induction of FL-cycE resulted in 30% decrease in CDC6 expression. LMW-E overexpression also increased the level of RAD51 and C17orf53, 100% higher than those in FL-cycE overexpressing cells.

To examine the effect of FL-cycE and LMW-E on MCMs loading to DNA, we induced FL-cycE or LMW-E by treating the 76NE6-EKO and 76NF2V-EKO cells with 100ng/ml doxycycline for 24 hours, followed by cell fractionation to isolate the chromatin bound fraction of the cells. Western blotting results revealed that induced expression of LMW-E promoted chromatin loading of MCM2, 4 and 7, approximately 3-fold of un-induced cells (Figure 44A). We also observed that chromatin bound MCM4 and 7 were down-regulated by FL-cycE overexpressing by 50% (Figure 44A). These results suggest LMW-E but not FL-cycE promote the chromatin loading of MCMs complex. Consistent results were also observed in 76NF2V-EKO cell lines (Figure 44B). Induced expression of LMW-E promoted chromatin loading of MCM2, 4 and 7, respectively 2-fold, 4-fold and 1.3-fold of un-induced cells. Chromatin bound MCM2, 4 and 7 were down-regulated by FL-cycE overexpressing by 50 percent. These results suggest the LMW-E mediate MCMs loading to DNA, while FL-cycE inhibit their loading.

Interestingly, western blotting results also show higher level of LMW-E than FL-cycE (> 5-fold), in the chromatin bound fraction of inducible hMECs, suggesting LMW-E but not FL-cycE strongly binds to chromatin (Figure 44). To further validate this, we performed cell fractionation assay in MDA-MB-157 cells which endogenously express both FL-cycE and LMW-E. In MDA-MB-157 cells, we find 20% higher levels of FL-cycE than LMW-E in total cell lysates, while the LMW-E level in chromatin-bound fraction is 3-fold higher than the chromatin binding FL-cycE (Figure 45). These data suggest LMW-E but not FL-cycE is the major form of cyclin E that binds with chromatin.
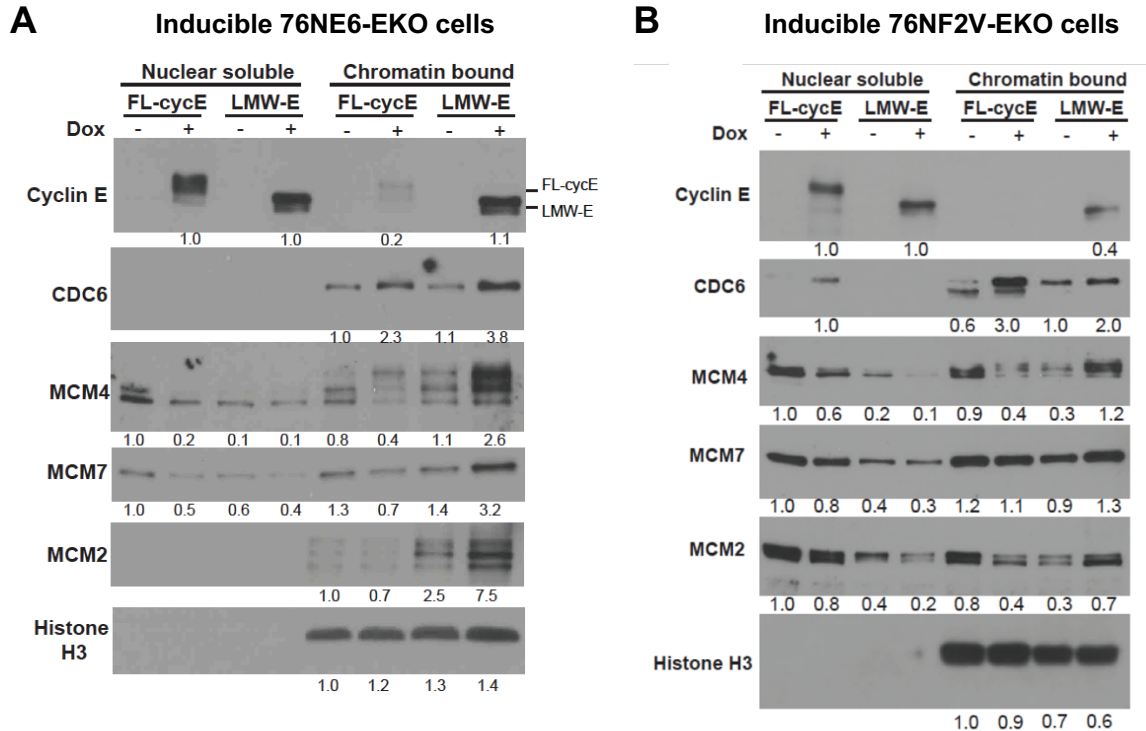
Figure 44. LMW-E but not FL-cycE was recruited to chromatin and promoted pre-replication complex loading. Western blotting analysis of cyclin E (FLcycE or LMW-E) and DNA pre-replication complex proteins using protein samples from cell nuclear soluble fraction and non-soluble (chromatin bound) fraction. The data using inducible 76NE6-EKO cells (panel A) and inducible 76NF2V-EKO cells (pane B) treated with or without doxycycline (100ng/ml, 24 hours) were shown respectively. Increased chromatin loading of MCM2, 4, and 7 were induced by LMW-E over-expression, and decreased chromatin loading of MCM2, 4, and 7 were observed in FL-cycE settings. In addition, higher level of LMW-E than FL-cycE were observed in the chromatin bound fraction from both cell lines.
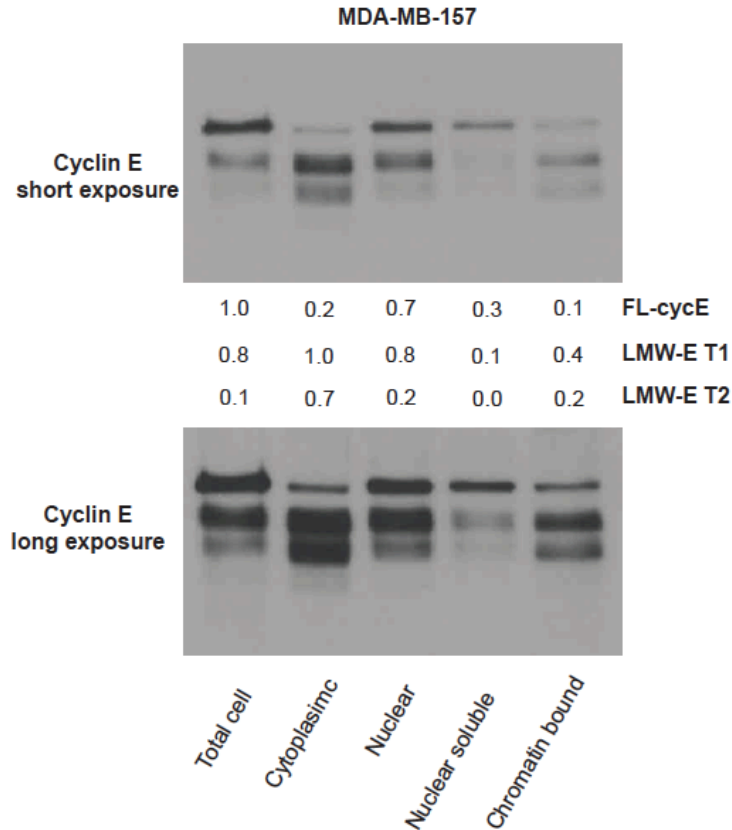
Figure 45. Validation of higher chromatin bound LMW-E level than FL-cycE using fractionated MDA-MB-157 cells. In Western blot analysis of the ratio of cyclin E (FL-cycE or LMW-E) in MDA-MB-157 total cell lysates of and fractionated lysates, we found 20% higher of FL-cycE than LMW-E(T1, see figure 5 for the schematics of cyclin E isoforms) in total cell lysates, while the LMW-E(T1) level in chromatin-bound fraction was 3-fold higher than the chromatin binding FL-cycE.

4.4.3 LMW-E but not FL-cycE strongly bind to CDC6

To examine the CDC6 binding capacity of LMW-E and FL-cycE, we co-transfected HEK293T cells with expression plasmids encoding CDC6 and FL-cycE (or the empty vectors for each of the plasmid as negative controls) and examined the binding between CDC6 and FL-cycE by co-immunoprecipitation (IP)/WB. In parallel, plasmids encoding CDC6 and LMW-

E (or empty vector controls) were co-transfected to examine the binding between CDC6 and LMW-E. These co-IP/WB were performed using anti-CDC6 antibody as bait, and the protein level of FL-cycE or LMW-E (by subsequent western blots with cyclin E) were compared in the anti-CDC6 pull-down precipitates in each setting. The result showed that while the level of CDC6 in the pull-down precipitates were similar between FL-cycE and LMW-E groups, 2.6-fold more LMW-E protein was bound to CDC6 compared to FL-cycE (Figure 46). These results suggest that LMW-E binds more efficaciously to CDC6 than FL-cycE.

To further confirm the different binding capacity between LMW-E versus FL-cycE to endogenous CDK6, we induced the LMW-E (or FL-cycE) in 76NE6-EKO cells and repeated the IP/western analysis. Total protein lysates were harvested 24 hours post induction of FL-cycE or LMW-E and normalized according to the level of CDC6. We performed co-IP with anti-cyclin E followed by western blotting with CDC6 or cyclin E. The data revealed that while the total FL-cycE is 40% higher than total LMW-E (in cyclin E co-IP/cyclin E WB), the level of CDC6 pulled down by FL-cycE is only one third of the CDC6 pulled down by LMW-E (in cyclin E co-IP/CDC6 WB; Figure 47). These results further confirmed LMW-E but not FL-cycE strongly interacts with CDC6.

4.4.4 CDC6 is required for LMW-E mediated replication stress tolerance

To further asses the role of CDC6 in LMW-E mediated replication stress tolerance, siRNAs targeting CDC6 (or non-target control siRNA) were used to deplete CDC6 in the inducible 76NE6-EKO-LMWE cells. We performed western blotting to evaluate the knock-down efficiency of CDC6. The results suggest CDC6 smart pool siRNA, CDC6 single specific siRNA#1, and single specific siRNA#2, respectively reduce the level of CDC6 to 40%, 50% and 60%, compared to the non-targeted siRNA control (Figure 48 and 49).
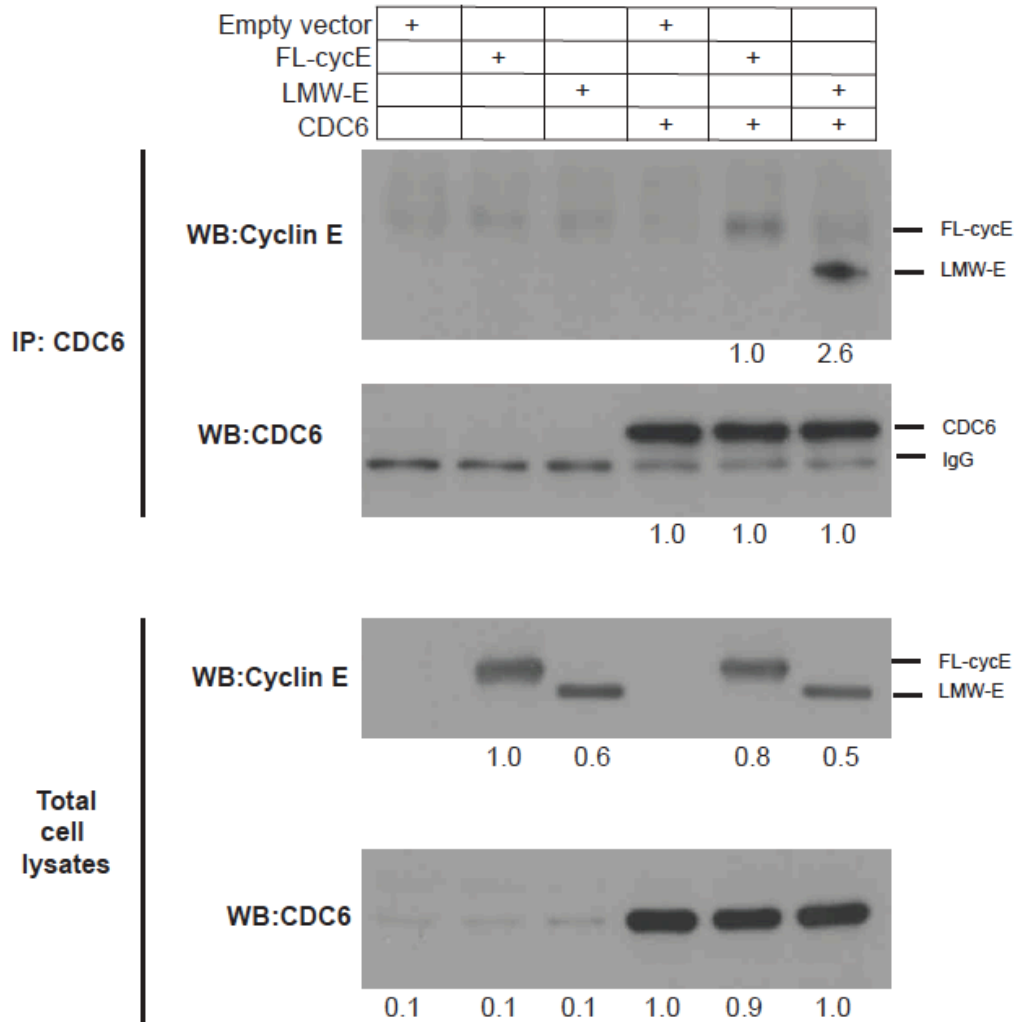
Figure 46. Analysis of the binding between cyclin E (FL-cycE or LMW-E) and CDC6 in plasmids transfected HEK293T cells. These cells were transfected with indicated plasmids to express CDC6, FL-cycE or LMW-E. After 48 hours in culture, the cell lysates were collected followed by co-IP using anti-CDC6 as bait. The level of Cyclin E and CDC6 in immuno-precipitates and total cell lysates were analyzed by western blot using indicated anti-bodies. Densitometry results show the level of LMW-E is 40% less than FL-cycE in total cell lysates, while LMW-E is 160% higher than FL-cycE in anti-CDC6 co-IP samples.
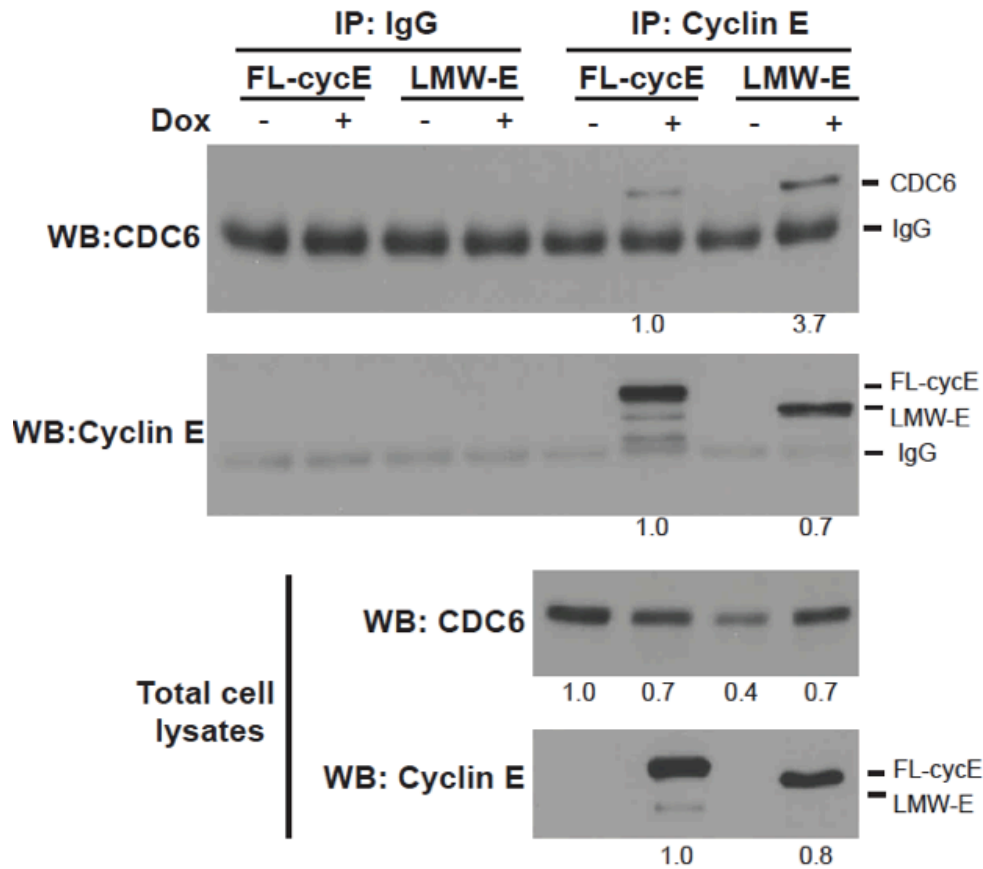
Figure 47. Analysis of the binding between induced cyclin E (FL-cycE or LMW-E) and endogenous CDC6 in inducible hMECs. Inducible 76NE6-EKO cells were treated with 100ng/ml doxycycline to induce the expression of FL-cycE or LMW-E for 24 hours, followed by co-IP using anti-cyclin E as bait and WB using indicated anti-bodies. Non-induced cells (DMSO treated) and co-IP using IgG were served as negative controls. Densitometry results showed similar level of CDC6 in total cell lysates of induced FL-cycE or LMW-E groups, while the level of CDC6 in the LMW-E co-IP sample was 270% higher than in FL-cycE co-IP sample.

Inducible 76NE6-EKO-LMWE

	NT siRNA pool		CDC6 siRNA smartpool	
Dox	−	+	−	+
CDC6	1.0	1.6	0.3	0.6
Cyclin E		1.0		0.8
Vinculin	1.0	1.2	1.2	1.2
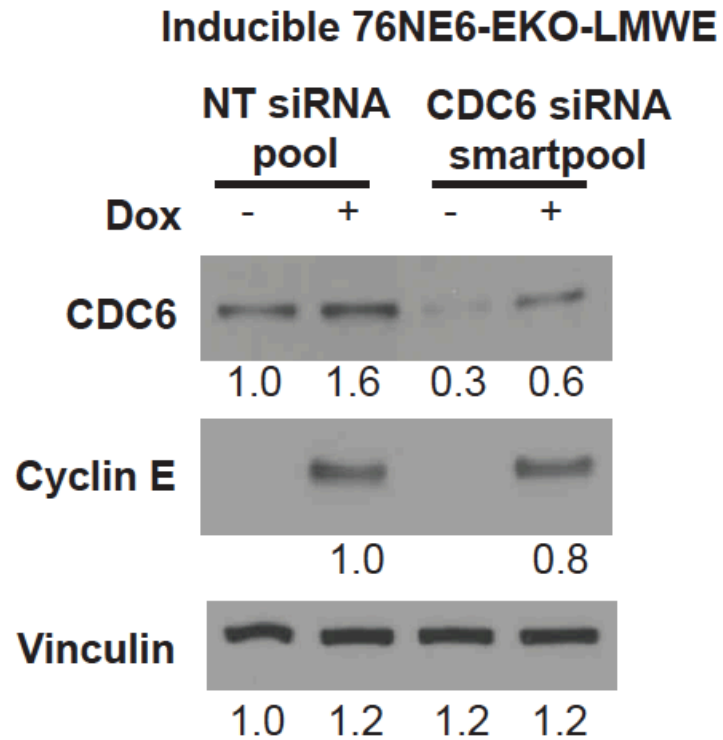
Figure 48. Effect of CDC6 knock-down by CDC6 siRNA smart pool. Western blot analysis of the level of CDC6 in inducible 76NE6-EKO LMW-E cells transfected with siRNA smart pool targeting CDC6 or non-target (NT) siRNA control. These cells were transfected with indicated siRNAs, followed by 24 hours treatment of 100ng/mL doxycycline to induce LMW-E overexpression. Densitometry results show the level of CDC6 in CDC6 siRNA transfected cells is 70% less than that in NT siRNA transfected cells.

Upon confirmation of the knock down efficiency, we then examined if CDC6 is required for i) LMW-E mediated MCMs loading to DNA (Figure 50), ii) LMW-E loading to DNA (Figure 50), iii) DNA damage intensity in inducible 76NE6-EKO cells with or without LMW-E overexpression (Figure 51), and iv) cell viability of inducible 76NE6-EKO cells with or without LMW-E overexpression (Figure 52).

We first tested the effect of CDC6 depletion on the MCMs loading to chromatin in inducible 76NE6-EKO-LMW-E cells with or without LMW-E induction. These cells were transfected with CDC6 siRNAs or non-target siRNA, followed by 24 hours treatment of 100ng/ml doxycycline to induce LMW-E or DMSO for uninduced controls. Western blot analysis using total cell lysates showed level of MCM2, MCM4 and MCM7 in total cell lysates remained relatively constant with or without knock-down of CDC6 (Figure 49). Next, we performed western blot using chromatin bound fraction to compare the MCMs loading to DNA. In the absence of LMW-E, CDC6 knock down slightly reduced the MCMs loading to DNA. Particularly, the transfection of CDC6 siRNA resulted in an average reduction of 30%, 20% and 50% respectively for the level of MCM4, MCM7 and MCM2 loading to DNA (Figure 50). When LMW-E was induced, in the non-target siRNA transfected cells, the level of chromatin loaded MCM4, MCM7 and MCM2 increased to 2.1-fold, 1.4-fold and 1.2-fold of uninduced control, consistent with previous observations that LMW-E-upregulate MCMs loading to DNA (Figure 50 and Figure 44). In CDC6 knock-down groups, the chromatin loading of MCM4, MCM7 and MCM2 were significantly reduced, which was at 70%, 80%, and 55% percent of control respectively (Figure 50). These results suggest that LMW-E mediated MCMs loading require the presence of CDC6. In the same experiment, we also observed that CDC6 knock-down diminished the chromatin bound LMW-E, suggesting CDC6 is also needed for the recruitment of LMW-E to the chromatin (Figure 50).
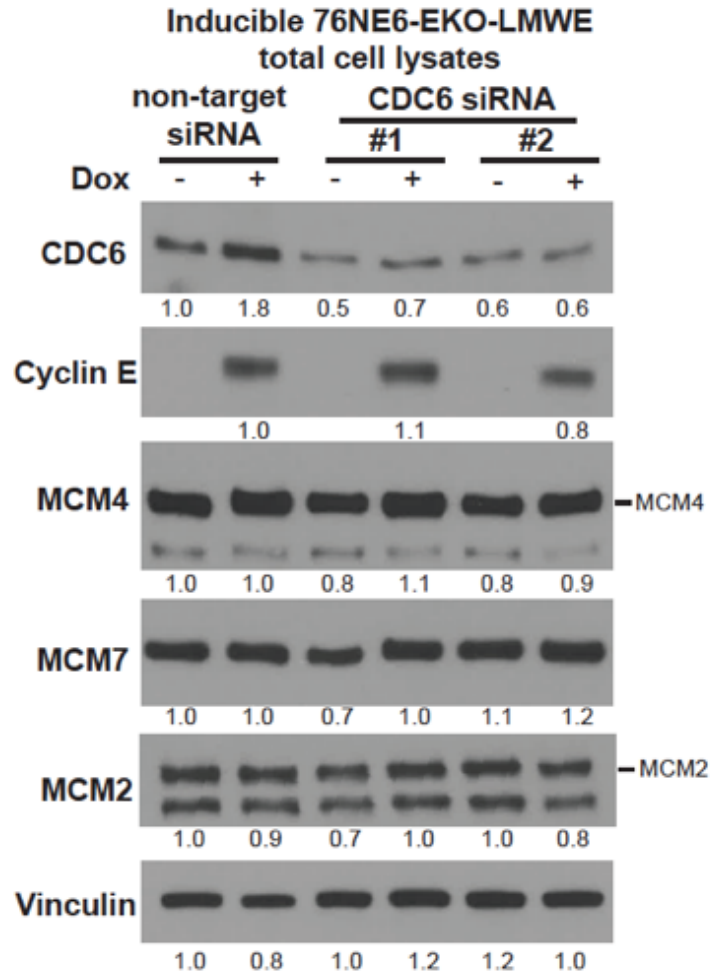
Figure 49. Effect of CDC6 knock-down by CDC6 specific siRNAs. Western blot analysis of CDC6 and MCMs complex components in LMW-E induced 76NE6-EKO cells transfected with siRNAs targeting CDC6 or non-target(NT) siRNA control. These cells were transfected with indicated siRNAs, followed by 24 hours treatment of 100ng/mL doxycycline to induce LMW-E overexpression. Densitometry results show the level of CDC6 in CDC6 siRNAs transfected cells is ~50% less than that in NT siRNA transfected cells. The level of Cyclin E (LMW-E), MCM2, 4 and 7 were comparable between the cells transfected with CDC6 siRNAs or NT siRNA.
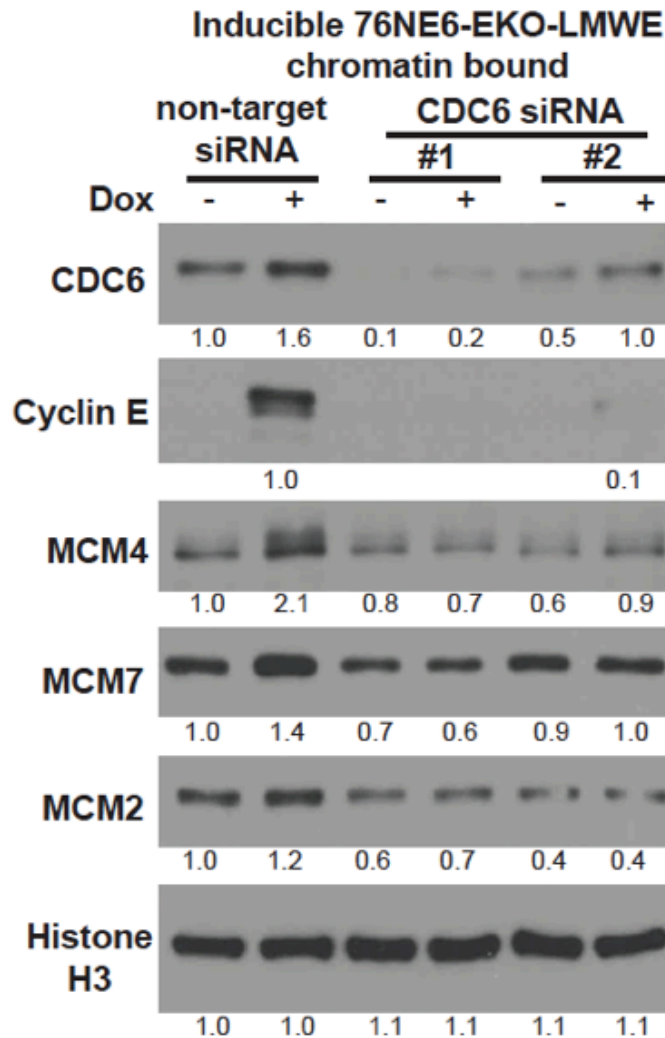
Figure 50. CDC6 is required for LMW-E loading to chromatin and LMW-E mediated MCMs loading. Western blot analysis of the level of CDC6 and MCM complex components in the chromatin bound fraction of LMW-E induced 76NE6-EKO cells transfected with siRNAs targeting CDC6 or non-target siRNA control. These cells were transfected with indicated siRNAs, followed by 24 hours treatment of 100ng/ml doxycycline to induce LMW-E overexpression. Densitometry results show in the chromatin bound fraction of LMW-E overexpressing cells, transfection of CDC6 siRNA lead to an average reduction of MCM4, MCM7 and MCM2 at 70%, 80%, and 55% percent of control, respectively.

To assess if CDC6 is required for LMW-E mediated replication stress tolerance, we examined the effect of CDC6 on DNA damage intensity following LMW-E induction in 76NE6-EKO. To this end, we knocked down CDC6 with smart-pool siRNA and quantitated DNA damage using γ-H2AX and 53BP1 IF foci. Results revealed that γ-H2AX positive cells increased from 2% to 12% and 53BP1 positive cells increased from 4% to 16% upon CDC6 knock-down in cells without LMW-E over-expression. These results suggest CDC6 is important for DNA damage control in normal dividing cells. In the non-target siRNA control groups, after 24 hours of induced expression of LMW-E, the ratio of γ-H2AX positive cells and 53BP1 positive cells both reached about 20-25%. In the CDC6 knocked-down groups, the ratio of γ-H2AX positive cells and 53BP1 positive cells increased to 40% of total cell population, after LMW-E was overexpressed for 24 hours, which are both ~60% higher compared to the LMW-E overexpressing cells without knock-down of CDC6 (Figure 51). These results suggest CDC6 plays an essential role in managing DNA damage in LMW-E overexpressing cells.

We also found that following siRNA knock-down of CDC6, cell viability was decreased by 30% in cells without LMW-E induction, and further decreased by 50% in cells with LMW-E induction (Figure 52), suggesting CDC6 is required for the cell survival upon LMW-E over-expression. Given that LMW-E positively regulates the level of CDC6 (Figure 43), our data collectively support the hypothesis that CDC6 serves as an essential LMW-E down-stream effector in tolerating replication stress.
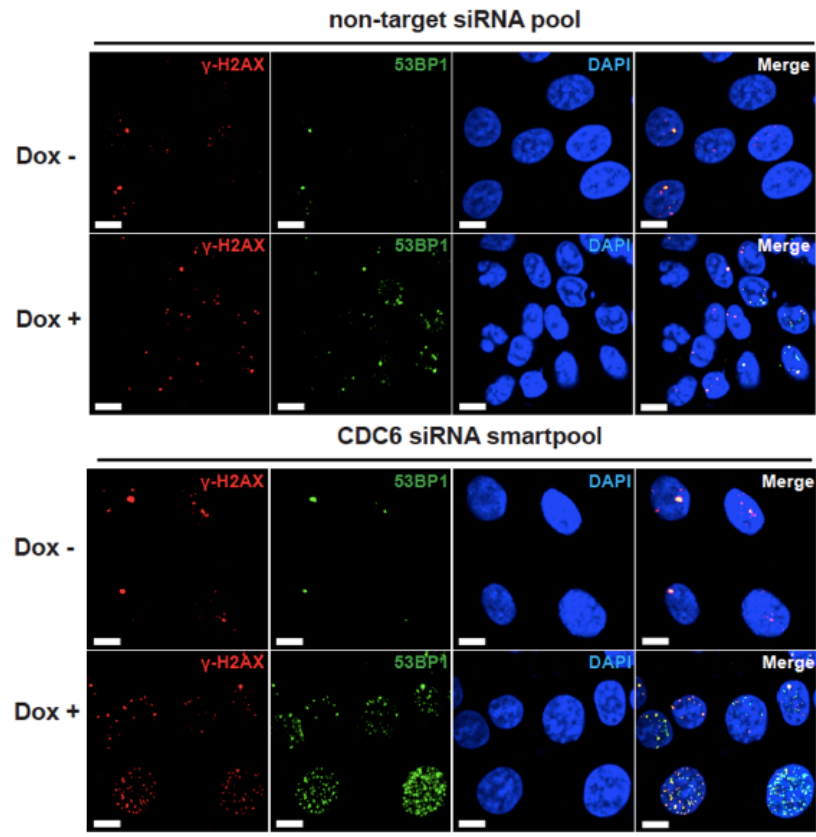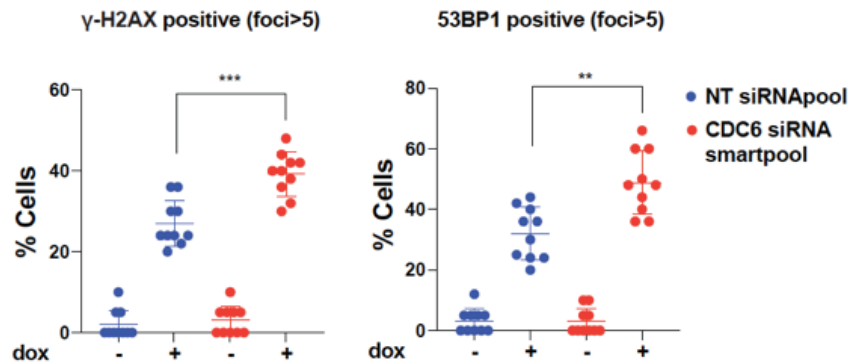
Figure 51. Depletion of CDC6 increased DNA damage in LMW-E overexpressing cells.

DNA damage assay by γ-H2AX and 53BP1 foci in inducible 76NE6-EKO cells, with or without CDC6 knockdown and/or induction of LMW-E overexpression. A. Representative images of immunofluorescent γ-H2AX and 53BP1 foci (scale bar = 10 μm). B. The ratio of cells positive for γ-H2AX and 53BP1 foci (foci > 5) was then calculated (cell number > 400, mean with standard deviation, **$p < 0.01$, ***$p < 0.001$, Student t test.)
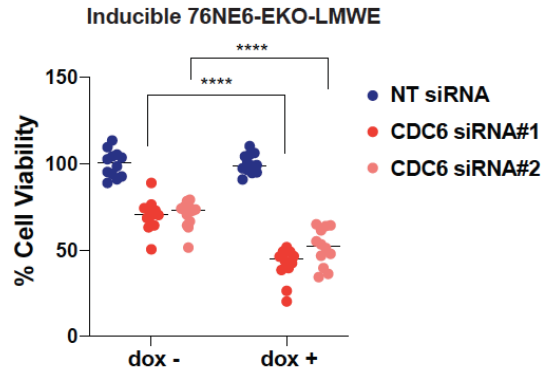
Figure 52. Depletion of CDC6 reduced viability of LMW-E overexpressing cells. Analysis of cell viability by MTT assay in inducible 76NE6-EKO cells transfected with siRNAs targeting CDC6 or non-targeting (NT) controls. These cells were treated with or without 100 ng/mL doxycycline for 48 hours to induce LMW-E overexpression. ****p < 0.0001, Student *t* test.

4.5 Conclusion

In this chapter, we show that LMW-E facilitates replication stress tolerance, so that replication stress only accumulates in FL-cycE overexpressing cells but not LMW-E overexpressing cells. We determined that replication licensing is one of the mechanisms that is differentially regulated by LMW-E and FL-cycE. Such processes depend on the pre-replication complex component CDC6, which is upregulated by LMW-E. We also find that LMW-E, but not FL-cycE, strongly binds to CDC6. Subsequently, LMW-E is recruited to the chromatin and promotes MCMs loading to DNA. Knock-down of CDC6 blocks the chromatin loading of LMW-E and LMW-E mediated MCMs loading, increases γ-H2AX and 53BP1 positive cells in LMW-E overexpressing cells. And decreases cell viability.

Our data collectively support the hypothesis that LMW-E but not FL-cycE plays an active role in facilitating replication stress tolerance. We next asked if one of the mechanisms that LMW-E facilitates replication stress tolerance, is through promoting DNA repair in Chapter 5.

Chapter Five: LMW-E but not FL-cycE facilitated DNA damage repair

5.1 Introduction

Using the *CCNE1* knock-out, FL-cycE or LMW-E inducible hMECs, we have demonstrated that DNA replication and DNA damage intensity are phenotypically different between cells expressing FL-cycE versus those overexpressed LMW-E. Additionally, FL-cycE but not LMW-E expression result in reduced DNA replication in S phase and increased DNA damage intensity. Consistently, RNA sequencing results suggest LMW-E but not FL-cycE induction can upregulate the DNA replication and repair pathways.

In this chapter, we aim to test the hypothesis that DNA damage repair is differentially processed by LMW-E and FL-cycE. To this end we compared the intensity of DNA damage markers, and the expression of replication stress proteins induced by FL-cycE or LMW-E expressed individually or co-expressed in U2OS cells. Cells expressing empty vector were used as negative controls (schematics shown in Figure 53). The rationale to use U2OS cells is twofold. First, U2OS cells are well characterized as a cellular model to study DNA damage, due to their lack of onco-virus proteins such as SV40-large T antigen that disrupts DNA damage responses[131,139]. Second, previous reports suggest overexpression of cyclin E can induce replication stress and DNA damage in U2OS cells, but which isoform of cyclin E is not specified in these studies[140-142]. To directly compare the function of FL-cycE and LMW-E in DNA repair efficiency, U2OS stable cell line harboring DNA double strand break repair reporter (EJ5-GFP) was also used[143].

From the list of genes upregulated by LMW-E (Table 2), we first selected RAD51 based on its well-established role in DNA damage repair, particularly in response to replication stress (see section 1.8 of Chapter 1). The stalling of replication forks may result in the

exposure of single strand DNA (ssDNA). RPA proteins then bind to the exposed ssDNA and activate ATR-CHK1 signaling pathway. This step further promotes the phosphorylation of RPA and the recruitment of RAD51 proteins, which replace RPA on ssDNA to facilitate lesion repair by stabilizing or remodeling of the stalled fork[90,115] .

Additionally, our RNA sequencing results showed that LMW-E upregulates C17orf53 (Table 2), a less studied gene functioning in DNA damage repair[137]. Previous studies have shown that C17orf53 is required for survival when cells are treated with Mitomycin C, an external replication stress agent. C17orf53 functions as a ssDNA- and RPA-binding protein that interacts with the MCM complex to promote homologous recombination (HR) and/or inter-strand crosslink (ICL) repair[137,138].

Understanding the (different) consequences of LMW-E versus FL-cycE expression on DNA damage repair may provide the rationale for targeting DNA damage repair pathways in LMW-E overexpressing breast tumors. Characterizing the dependence of RAD51 and/or C17orf53 on the DNA damage repair and cell viability of LMW-E overexpressing cells may also shed light on potential therapeutic targets.

5.2 Schematics of model system

A U2OS

B Inducible 76NE6-EKO-LMW-E
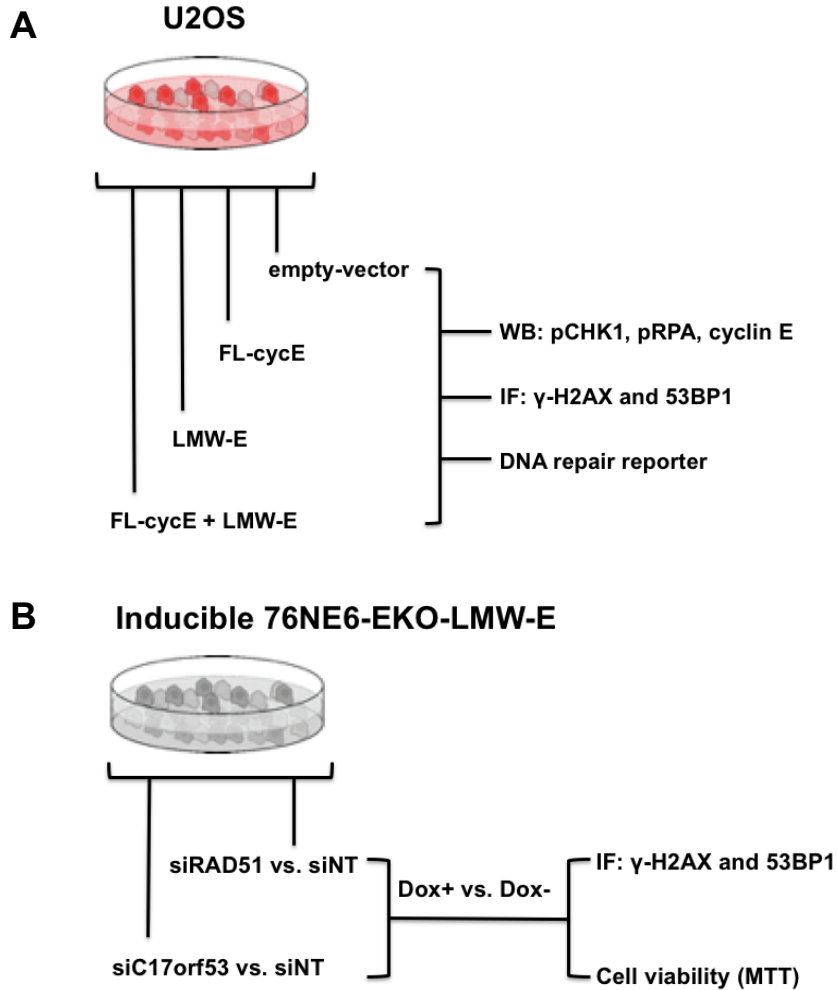
Figure 53. Schematics of experimental models to examine the roles of LMW-E versus FL-cycE in regulating DNA damage repair. To compare the effect of LMW-E and FL-cycE on DNA damage repair, U2OS cells were transfected to express empty-vector, FL-cycE, LMW-E, or to co-express of FL-cycE and LMW-E. After 48 hours, intensity of DNA damage markers, and the level of replication stress proteins were compared. Stable U2OS cell line harboring DNA double strand break repair reporter was also used to measure the DNA repair efficiency(A). We also performed siRNA knock-down experiments to examine the role of RAD51 and C17orf53 in DNA damage and cell viability of inducible 76NE6-EKO-LMW-E cells (B).

117

5.3 Materials and Methods

Specific siRNAs and smart pool siRNAs

All siRNAs (specific siRNAs and smart pool siRNAs) used in this study were provided by Dharmacon (Horizon Discovery). The transfection was performed by using Lipofectamine 3000 (Thermofisher) according to manufacturer's protocol. The knock-down efficiency of specific siRNAs and smart pool siRNAs were confirmed by western blot analysis for target proteins using total cell lysates harvested 72 hours post transfection of the following siRNAs. RAD51 siRNA#1: UAUCAUCGCCCAUGCAUCA, #2: CUAAUCAGGUGGUAGCUCA; RAD51 siRNA smart pool: UAUCAUCGCCCAUGCAUCA, CUAAUCAGGUGGUAGCUCA, GCAGUGAUGUCCUGGAUAA, CCAACGAUGUGAAGAAAUU. C17orf53 siRNA#1: UGGAUUAUUUCCUCGGAUA, #2: CUGGGAAGUCUGUCCGCAA. C17orf53 siRNA smart pool: UGGAUUAUUUCCUCGGAUA, CUGGGAAGUCUGUCCGCAA, GCAGUGAGGCCAUACCAAU, CCAUCCACAAAGCGGGUAU.

Cell viability assay

Cell viability is measured by the MTT assay in a 96-well plate as previously described (Chapter 2). Briefly, 12 hours following siRNA transfection, 1500 inducible 76NE6-EKO cells were plated into each well of 96-well plates. After the cell confluency reaches 20%, the cells were treated with 100ng/mL doxycycline to induce LMW-E overexpression. DMSO was used as uninduced control. 48 hours later, the cells were incubated in 100 µl of 2.5 mg/mL MTT in serum-free media for additional 4 hours at 37°C. The MTT precipitants were solubilized in 100µL DMSO at room temperature on a horizontal shaker for 30 min. The reading was quantified using a spectrophotometer (Victor3, Perkin-Elmer) for the absorbance at wavelength of 590 nm.

Western blot analysis

Cell lysates were prepared and subjected to western blot analysis as previously described in chapter 2. We used primary antibodies against cyclin E (HE-12; Santa Cruz Biotechnology, SC-247); CDC6 (180.2; Santa Cruz Biotechnology, SC-9964); CHK1 (G-4; Santa Cruz Biotechnology, SC-8408); phospho-CHK1 Ser345 (Cell Signaling, #2348); RPA32 (Cell Signaling, #22085); phospho-RPA32 Ser4/8 (Bethyl, A300-245a); phospho-RPA32 Ser33 (Bethyl, A300-246a); Rad51 (H-92; Santa Cruz Biotechnology, SC-8349); C17orf53 (Sigma, HPA023393); and Vinculin (Sigma, V9131). All bands on on western blots corresponding to individual proteins were quantitated by densitometry and presented as a ratio of the first band. Vinculin was used for loading control. Densitometry is performed by using ImageJ software.

Immunofluorescence staining

Immunofluorescence (IF) staining were performed using primary antibodies: γ-H2AX (Millipore, 05-636) and 53BP1 (Novus, NB100-304), based on the protocol described in Chapter 2. The images were taken by Zeiss LSM880 Confocal at 63X magnification and examined by Zeiss Zen software. More than 600 cells were analyzed per group. Cells with five or more foci were considered positive.

5.4 Results

5.4.1 LMW-E but not FL-cycE promoted DNA damage repair

U2OS cells were transfected with plasmids expressing empty vector, LMW-E, FL-cycE, or co-transfected with LMW-E and FL-cycE. 48 hours post transfection, total cell lysates were collected, and subjected to western blot analysis to detect replication stress markers such as phospho-CHK1 and phospho-RPA32, as well as to determine the expression status of cyclin E, CDC6, C17orf53, and RAD51. Vinculin was used as the loading control. The results (Figure 54) showed that both the overexpression of FL-cycE or LMW-E upregulated

the level of CHK1 pS345, to approximately 5-fold of empty vector control. This result suggested both FL-cycE and LMW-E induced replication stress in U2OS cells. We then examined the replication stress marker RPA32pS4/8 and RPA32pS33 in cells transfected with LMW-E or FL-cycE, with empty vector controls. RPA32pS4/8 and RPA32pS33 were both strongly induced by FL-cycE over-expression, respectively showing 6.4-fold and 4.5-fold increase, compared with empty vector control. LMW-E overexpression only moderately enhanced the level of RPA32 pS4/8 to 2.6-fold, and RPA32 pS33 to 1.1-fold of empty vector control. Hence, the replication stress marker was 3 to 5 times stronger in FL-cycE overexpressing U2OS cells than LMW-E overexpressing cells. Of note, in cells that were co-overexpressed with FL-cycE and LMW-E (lane 4, Figure 54), we observed the rescue effects of LMW-E on the RPA phosphorylation mediated FL-cycE. Specifically, the level of RPA32 pS4/8 in cells with FL-cycE and LMW-E co-expression was 30% less than the cells with FL-cycE overexpression alone, and that of RPA32 pS33 was 90% less (Figure 54). These results suggest that, in U2OS cells, the replication stress induced by FL-cycE is stronger than LMW-E. However, LMW-E co-expression can partially rescue the replication stress induced by FL-cycE.

In the same experimental system, we also observed a 50% decrease of CDC6 by FL-cycE, while LMW-E alone increased the CDC6 level to 2.3-fold (Figure 54). This is similar to the previous observation in inducible 76NE6-EKO cells (Figure 43). In U2OS cells, co-expression of LMW-E and FL-cycE upregulated the levels of CDC6 to 2.1-fold, compared to the empty vector control, which was 4-fold of the level of CDC6 in cells with FL-cycE expressing alone. This suggest LMW-E played a dominant role over FL-cycE in regulating the level of CDC6. The western blot results also confirmed that LMW-E upregulates the RAD51 and C17orf53 levels in U2OS cells (Figure 54). These results consistently show that LMW-E but not FL-cycE promote the expression of CDC6 and replication stress tolerance.

Additionally, the rescue effect of LMW-E on FL-cycE expressing cells suggest that while LMW-E and FL-cycE may differentially regulate replication stress, LMW-E plays a dominant role in stabilizing the DNA damage.
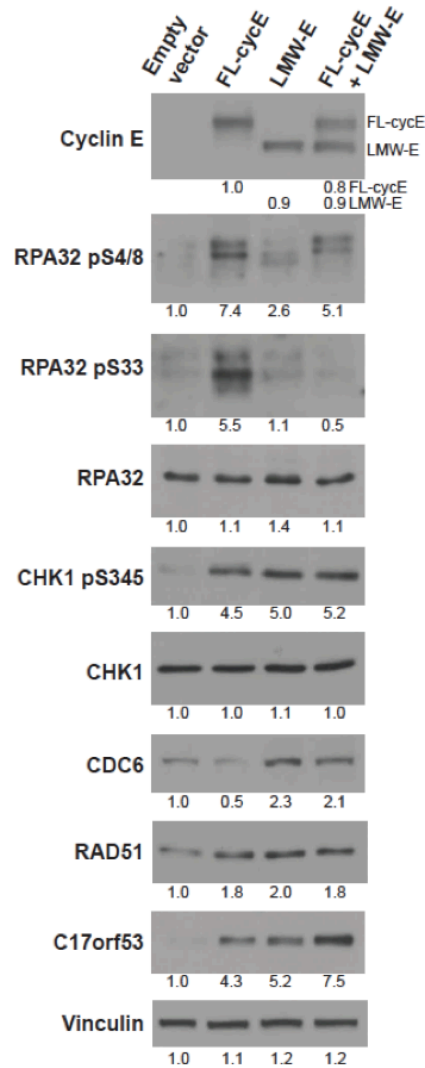
Figure 54. LMW-E promoted the expression of CDC6, RAD51 and C17orf53, facilitating replication stress tolerance in U2OS cells. Western blot analysis of the effect of LMW-E and FL-cycE on replication stress markers in U2OS cells. U2OS cells were transfected with the indicated plasmids, followed by western blot to identify the level of CHK1 pS345, RPA32 pS4/8 and RPA32 pS33. The level of CDC6, Rad51 and C17orf53 were also examined by western blot to validate the RNA-sequencing results.

121

We then analyzed the DNA damage intensity by γ-H2AX and 53BP1 foci in U2OS cells transfected with empty vector, FL-cycE, and LMW-E. The results suggested that after 48 hours expression of FL-cycE, 22% of cells were DNA damage positive, while for LMW-E overexpression cells the value was around 10% (Figure 55). Of note, U2OS cells co-expressing LMW-E and FL-cycE exhibited 40% lower DNA damage positive cells compared to FL-cycE overexpressing cells (Figure 55). These results also revealed that DNA damage induced by FL-cycE is 100% more than LMW-E, and that LMW-E co-expression can partially rescue the DNA damage induced by FL-cycE (Figure 55B).

To further confirm if LMW-E promoted DNA damage repair, we used the U2OS cell line stably expressing a NHEJ-I-SceI-based chromosomal break reporter (EJ5-GFP) [143]. In this system, double strand breaks can be induced by I-SceI expression, and end joining between two distal tandem I-SceI recognition sites may restore an GFP expression cassette.

The DNA damage repair efficiency is then estimated by the ratio of GFP positive cells. Our result suggested that in the empty vector or FL-cycE overexpressing cells, there were low percentage (1.5%) of GFP positive cells, suggesting very little double strand break repair efficiency. However, upon overexpression of LMW-E, there was a significant increase in GFP positive cells (7%). When both FL-cycE and LMW-E were co-overexpressed, the GFP positive cells are also significantly increased as compared to FL-cyc E alone (to 4%) (Figure 56). These results suggest that LMW-E but not FL-cycE promotes the efficiency of double strand break repair. The enhanced DNA repair by LMW-E + FL-cycE co-overexpression (~4.3-fold increase), indicated the role of LMW-E on DNA repair is dominant over FL-cycE.

Collectively, our result support the hypothesis that LMW-E but not FL-cycE plays a dual role in regulating DNA damage repair. LMW-E overexpression 1) leads to DNA damage by inducing replication stress, and 2) facilitates DNA repair, and in doing so contributes to replication stress tolerance by reducing DNA double strand breaks.
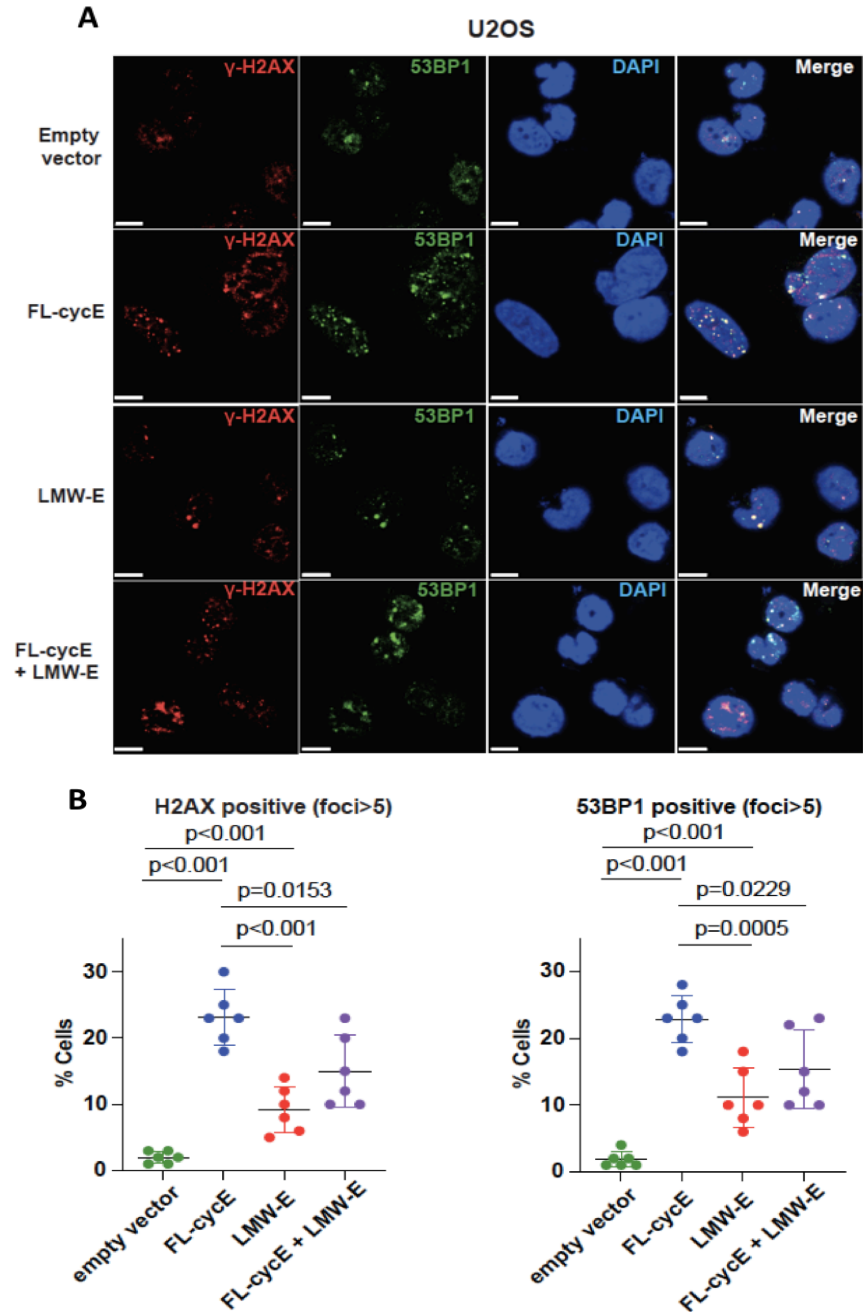
Figure 55. LMW-E overexpression induced less DNA damage than FL-cycE, and partially rescued FL-cycE induced DNA damage in U2OS cells A. Representative images of γ-H2AX and 53BP1 foci in U2OS cells 48 hours post transfected with empty-vector, FL-cycE, LMW-E or FL-cycE +LMW-E plasmids. (scale bar = 10 μm). B. The ratio of cells positive for γ-H2AX and 53BP1 foci (foci > 5) was then calculated (cell number > 300, mean with standard deviation, student *t* test.)
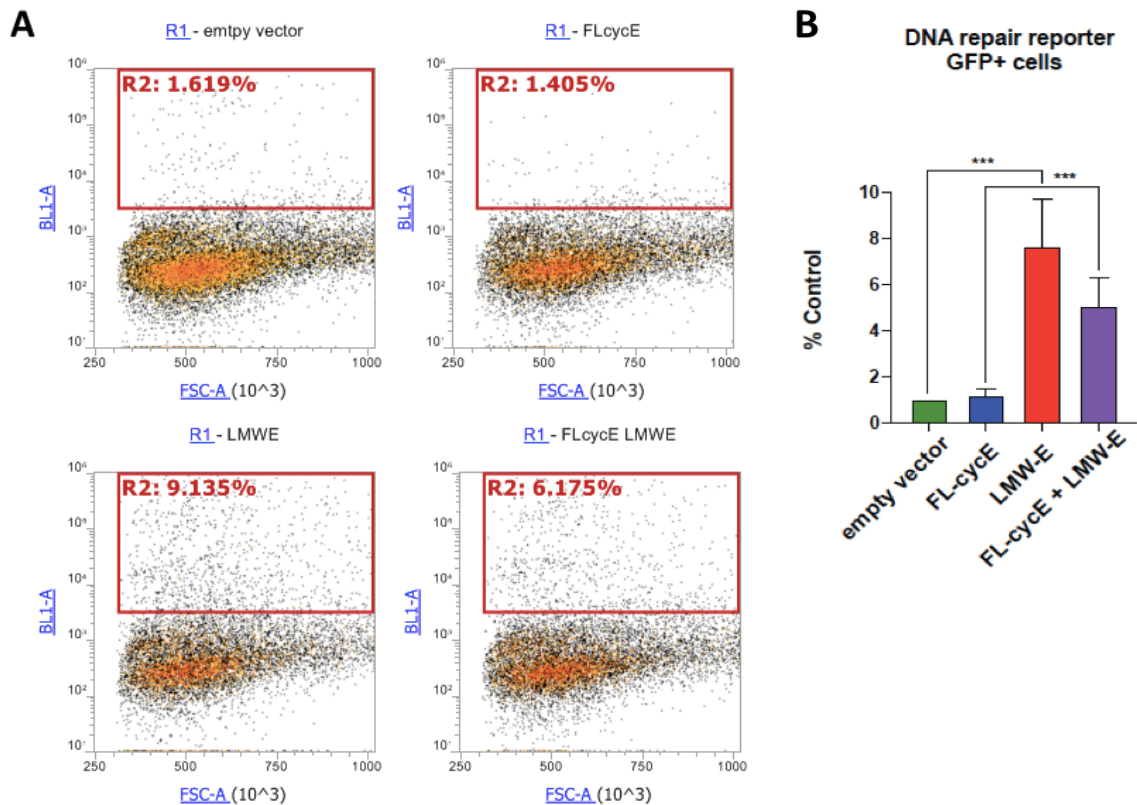
Figure 56. **LMW-E promoted DNA repair in U2OS cells, and played a dominant role when co-expressed with FL-cycE.** Overexpression of LMW-E promoted DNA damage repair. U2OS-EJ5-GFP were transfected with empty-vector, FL-cycE, LMW-E or FL-cycE +LMW-E plasmids, followed by transfection of I-SceI plasmid. The cells were harvested 48 hours post transfection and subjected to FACS analysis of GFP expression (R2 gating). Representative data are shown in (A), quantification of the experiments is shown in (B). Mean with standard deviation, n=3, ***p<0.001, student *t*-test.

5.4.2 RAD51 and C17orf53 are required for DNA damage repair and cell viability in LMW-E overexpressing hMECs.

To examine if RAD51 and/or C17orf53 are essential LMW-E downstream factors in facilitating DNA double strand break repair, we knocked down RAD51 or C17orf53 using siRNA transfection in the inducible 76NE6-EKO cells. Following LMW-E induction and siRNA

transfections, we subjected the cells to western blot analysis to evaluate the knock-down efficiency. These results showed that RAD51 smart pool siRNA, RAD51 single specific siRNA#1, and single specific siRNA#2, all resulted in >80% knock down efficiency of RAD51. C17orf53 smart pool siRNA, C17orf53 single specific siRNA#1, and single specific siRNA#2, respectively lead to >90%, 60% and 40% knock-down of C17orf53 expression (Figure 57).

We next examined the consequences of RAD51 or C17orf53 depletion on DNA damage markers by γ-H2AX and 53BP1 foci enumeration in the inducible 76NE6-EKO cells with or without LMW-E induction. After knock-down of RAD51, we observed an increase of γ-H2AX positive cells from 2% to 12% and 53BP1 positive cells from 4% to 16% in 76NE6-EKO cells without LMW-E induction, respectively. These results suggest that RAD51 is important for DNA damage control in normal dividing cells. Following 24 hours of induced expression of LMW-E, the ratio of γ-H2AX positive cells and 53BP1 positive cells both reached 20-25% in the non-targeted siRNA control group, which is consistent to the positive regulatory role of LMW-E on DNA damage by inducing replication stress. However, when RAD51 was knocked down, the ratio of γ-H2AX positive cells and 53BP1 positive cells following LMW-E induction further increased to 40% of total cell population (Figure 58). Similar results were observed in the when we knocked down C17orf53, suggesting depletion of C17orf53 increased DNA damage by 2-3 fold in LMW-E induced cells compared to uninduced cells (Figure 59).

Following siRNA knock-down of RAD51 or C17orf53, cell viability is also decreased in RAD51 or C17orf53 deficient LMW-E overexpressing cells by 63% or 50% respectively, compared to non-target siRNA treated cells (Figure 60). Given that LMW-E positively regulates the level of RAD51 and C17orf53, our data altogether support the hypothesis that RAD51 and C17orf53 serve as essential LMW-E down-stream effectors in the DNA damage repair process.
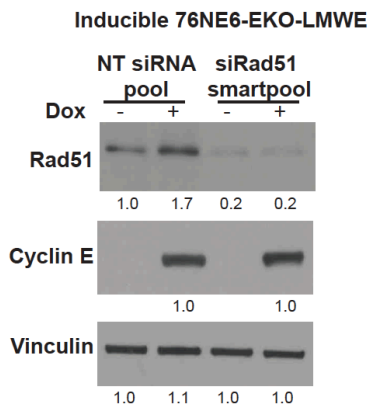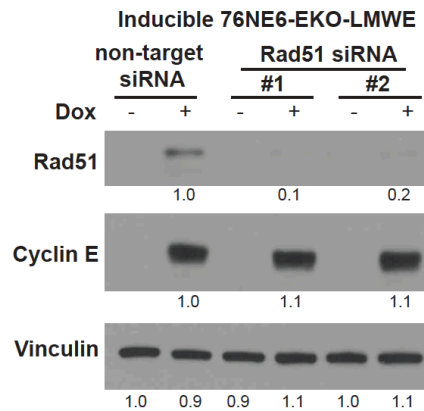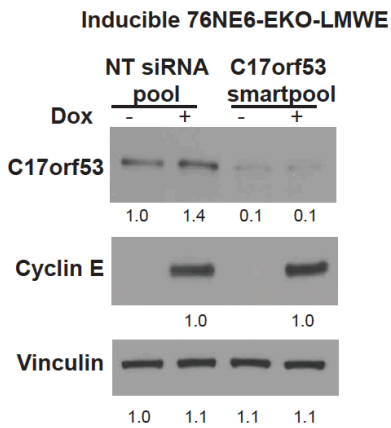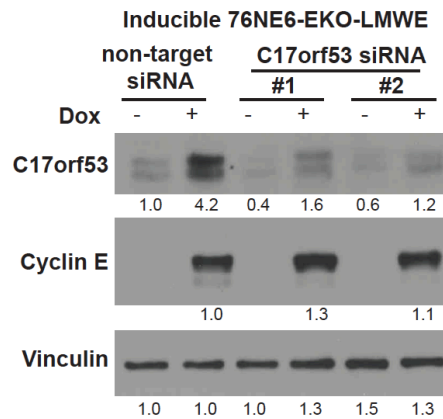
Figure 57. Effect of RAD51 and C17orf53 knock-down by siRNAs.

A and B. These cells were transfected with smart pool siRNA (A) or single specific siRNAs(B) targeting RAD51, NT siRNAs were used as control, followed by 24 hours treatment of 100ng/mL doxycycline to induce LMW-E overexpression. The Densitometry results show the RAD51 protein level in RAD51 siRNA transfected cells are 80 – 90% less than that in NT siRNA transfected cells. C and D. These cells were transfected with smart pool siRNA (C) or single specific siRNAs(D) targeting C17orf53, NT siRNAs were used as control, followed by 24 hours treatment of 100ng/ml doxycycline to induce LMW-E overexpression. The Densitometry results showed the C17orf53 protein level in C17orf53 smart pool siRNA transfected cells were 90% less than that in NT siRNA transfected cells. Single specific siRNAs against C17orf53 also reached 40 – 60% knock-down efficiency.
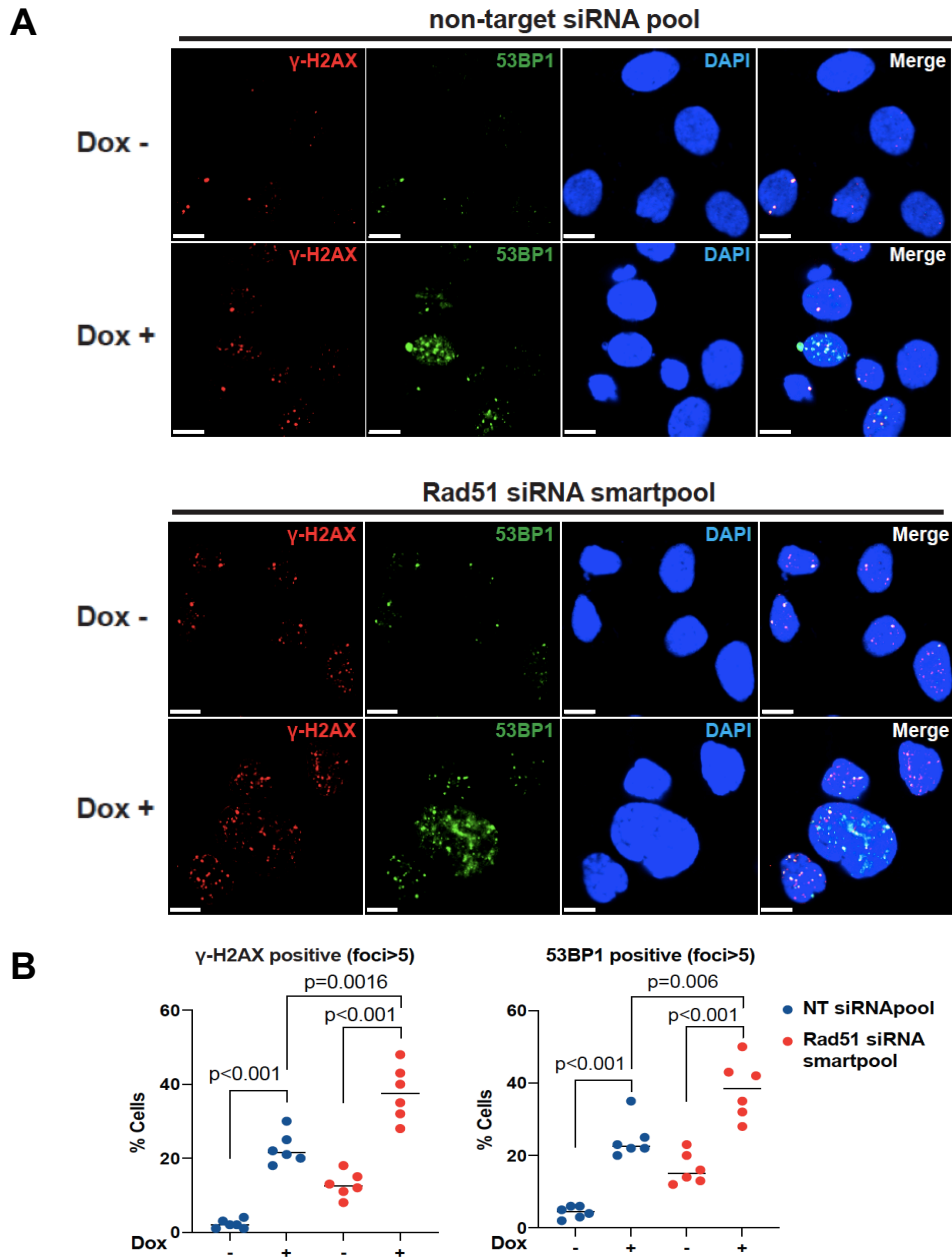
Figure 58. Depletion of RAD51 increased DNA damage in LMW-E overexpressing cells.

DNA damage assay by using immunofluorescent γ-H2AX and 53BP1 foci in inducible 76NE6-EKO cells, with or without RAD51 knockdown and/or induction of LMW-E overexpression. A. Representative images γ-H2AX and 53BP1 foci in inducible 76NE6-EKO cells, with or without RAD51 knockdown and/or induction of LMW-E overexpression (scale bar = 10 µm). B. The ratio of cells positive for γ-H2AX and 53BP1 foci (foci > 5) was then calculated (cell number > 400, mean with standard deviation, student *t* test).
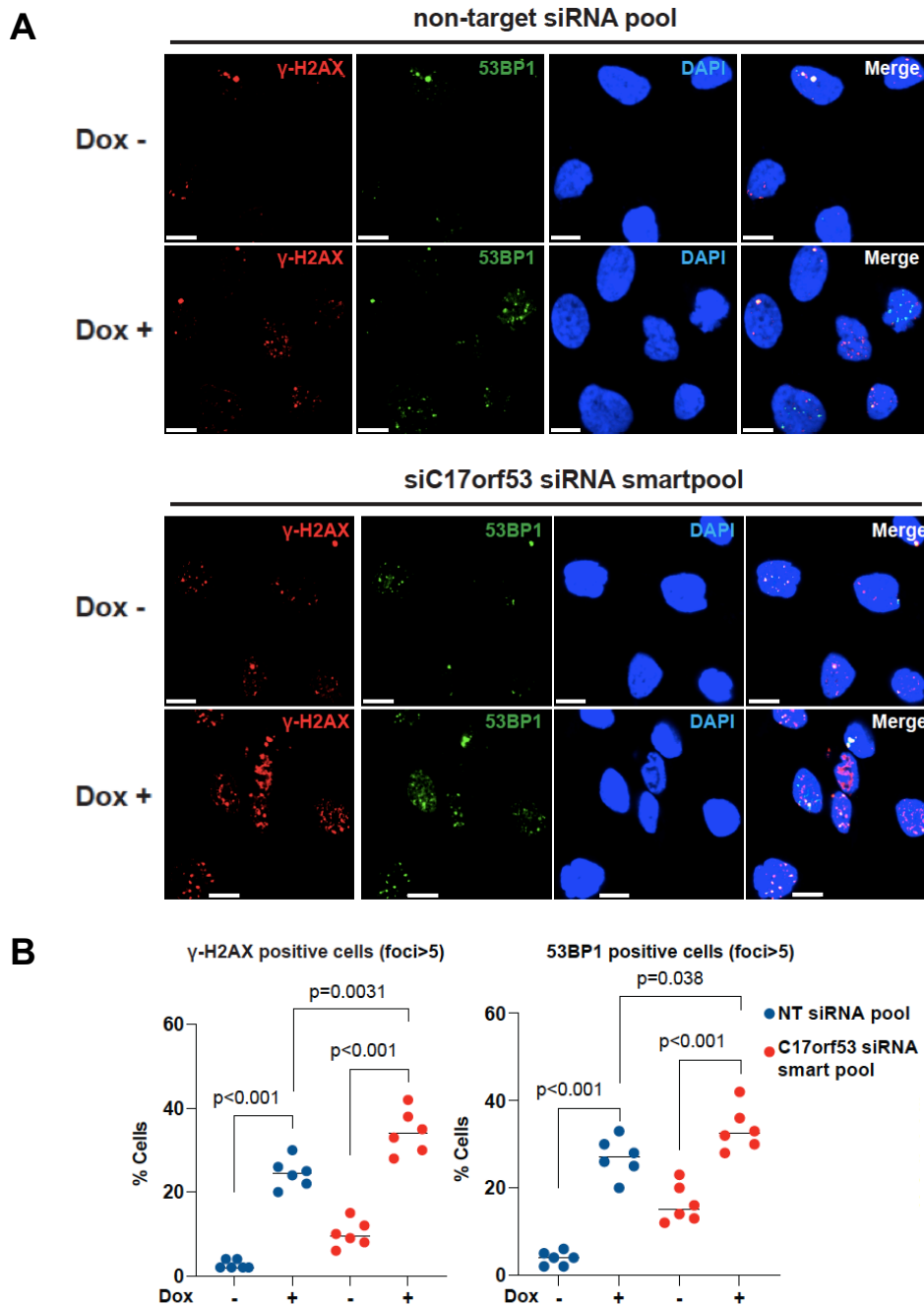
Figure 59. Depletion of C17orf53 increased DNA damage in LMW-E overexpressing cells. DNA damage assay by using immunofluorescent γ-H2AX and 53BP1 foci in inducible 76NE6-EKO cells, with or without C17orf53 knockdown and/or induction of LMW-E overexpression. A. Representative images γ-H2AX and 53BP1 foci in inducible 76NE6-EKO cells (scale bar = 10 μm). B. The ratio of cells positive for γ-H2AX and 53BP1 foci (foci > 5) was then calculated (cell number > 400, mean with standard deviation, student *t* test).

A
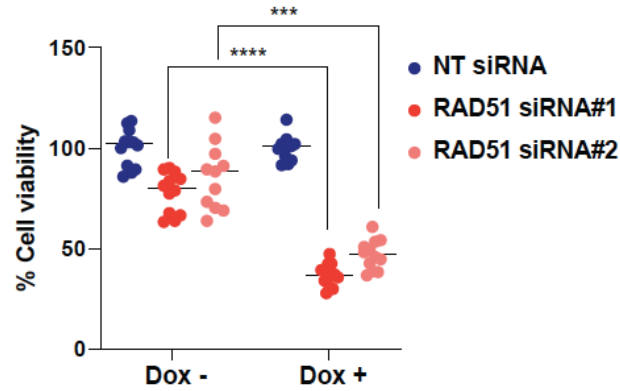
B
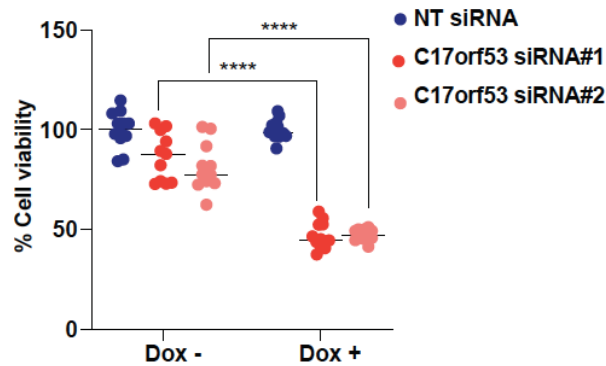
Figure 60. Depletion of RAD51 or C17orf53 reduced viability of LMW-E overexpressing cells. Analysis of cell viability by MTT assay in inducible 76NE6-EKO-LMWE cells after transfection of specific siRNAs targeting RAD51(panel A) or C17orf53(panel B), followed by 100ng/ml doxycycline induced LMW-E overexpression for 48 hours. Non-target siRNA and DMSO (dox-) were used as controls. The results suggest, compared with non-target siRNA transfected cells, cell viabilities in RAD51 or C17orf53 deficient LMW-E overexpressing cells decreased by 63% or 50% respectively. (****p < 0.0001, Student *t* test.)

5.5 Conclusion

In this chapter, we show that LMW-E but not FL-cycE facilitated DNA damage repair. Specifically, overexpression of LMW-E resulted in less DNA damage marker intensity and enhanced DNA repair efficiency in U2OS cells, compared with FL-cycE overexpression setting. When LMW-E was co-expressed with FL-cycE, it reduced DNA damage marker and

enhanced DNA repair, partially rescuing the DNA damage induced by FL-cycE. These results also help explain why DNA damage accumulates in inducible hMECs with FL-cycE but not LMW-E overexpressing condition. Additionally, data in U2OS cells confirmed that LMW-E overexpression upregulated RAD51 and C17orf53 as is observed in inducible 76NE6-EKO cells. To investigate the role of RAD51 and C17orf53 in LMW-E overexpressing hMECs, we examined the DNA damage markers and cell viability, with or without siRNA mediated knock-down of these genes. We observed a dramatic increase of γ-H2AX positive cells and 53BP1 positive cells after depletion of RAD51 or C17orf53, following the induction of LMW-E in 76NE6-EKO cells. Data from MTT assay also suggested that RAD51 or C17orf53 were required for the cell viability of when LMW-E was induced in 76NE6-EKO cells.

In summary, our data collectively support the hypothesis that LMW-E, but not FL-cycE, facilitates DNA damage repair, and that LMW-E plays a dominant role over FL-cycE. LMW-E overexpression upregulates RAD51 and C17orf53. These genes serve as essential down-stream effectors in the DNA damage repair and survival of LMW-E overexpressing cells.

Chapter Six: LMW-E overexpressing cells are sensitive to inhibitors targeting the ATR-CHK1-RAD51 pathway

6.1 Introduction

By comparing the effect of LMW-E with FL-cycE in inducible cyclin E knock-out hMECs, we observed FL-cycE but not LMW-E overexpressing resulted in DNA damage accumulation and loss of cell viability. Using U2OS cells and DNA damage repair reporter system, we have determined LMW-E facilitates DNA damage repair and can partially rescue the DNA damage induced by FL-cycE over-expression. These results suggest targeting DNA damage repair pathways may lead to reduced cell viability in LMW-E overexpressing cells.

We have determined that LMW-E overexpression led to replication stress and activated ATR-CHK1 pathway, evidenced by reduced fork speed (Figure 40) and enhanced phosphorylated CHK1(Figure 41). RNA sequencing and subsequent western blot analysis show that LMW-E overexpression upregulate RAD51(Figure 33, 42 and 53), which functions down-stream of ATR-CHK1 pathway and promote fork restart and damage repair[90,116,117]. We have also confirmed that depletion of RAD51 can increase DNA damage and reduce cell viability in 76NE6-EKO cells induced for LMW-E expression (Figure 58 and 60).

In this chapter, we aim to test the effect of inhibiting ATR-CHK1-RAD51 pathway in hMECs and breast cancer cells with or without LMW-E over-expression. We will use RAD51 inhibitor B02, CHK1 inhibitor Rabusertib and ATR inhibitor Ceralasertib (described in chapter 1), and evaluate the DNA damage and cell viability. Results from these experiments may establish ATR-CHK1-RAD51 pathway as therapeutic targets in LMW-E [high] breast cancer (Figure 61).

6.2 Schematics of model system.

Inducible 76NE6-EKO-LMW-E

MDA-MB-231
(stable cell lines)
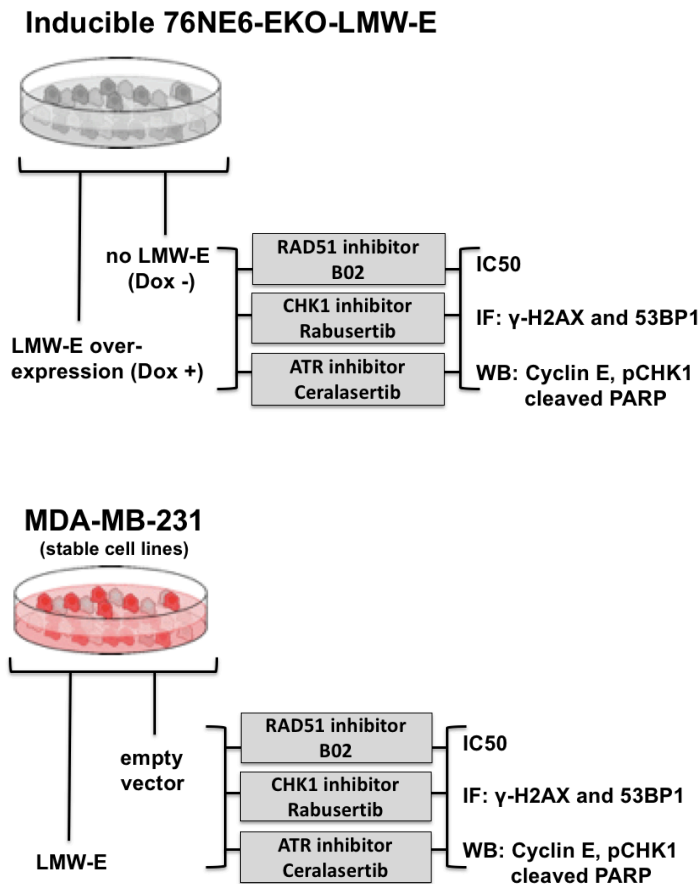
Figure 61. Model systems to test the effect of inhibiting ATR-CHK1-RAD51 pathway in cells with or without LMW-E overexpression. To evaluate the effect of inhibitors for RAD51, CHK1, or ATR in cells with or without LMW-E over-expression, we used inducible 76NE6-EKO-LMW-E cells with or without doxycycline treatment to mimic the condition of pre-cancerous mammary cells with or without LMW-E over-expression. We also used MDA-MB-231 breast cancer cells, engineered to express empty vector or LMW-E. For each inhibitor (RAD51 inhibitor B02, CHK1 inhibitor Rabusertib and ATR inhibitor Ceralasertib), we evaluate the half maximal inhibitory concentration (IC50) on cell viability by MTT assay and compare the value between LMW-E dox+ and dox – conditions. The level of DNA damage markers (γ-H2AX foci and 53BP1 foci), replication stress and DNA damage signal (phosphor-CHK1) and apoptosis marker (cleaved PARP) were also assessed.

6.3 Materials and methods

Inhibitors

We used the RAD51 inhibitor B02 (Selleckchem, Catalog No.S8434), CHK1 inhibitor Rabusertib (LY2603618, Selleckchem, Catalog No.S2626), and ATR inhibitor Ceralasertib (AZD6738, Selleckchem, Catalog No.S7693).

Western blot analysis

Cell lysates were prepared and subjected to western blot analysis as previously described in chapter 2. We used primary antibodies against cyclin E (HE-12; Santa Cruz Biotechnology, SC-247); phospho-CHK1 Ser345 (Cell Signaling, #2348); RPA32 (Cell Signaling, #22085); PARP (Cell Signaling, #9542); and Vinculin (Sigma, V9131). All bands on on western blots corresponding to individual proteins were quantitated by densitometry and presented as a ratio of the first band. Vinculin was used for loading control. Densitometry is performed by using ImageJ software.

Immunofluorescence staining

Immunofluorescence (IF) staining were performed using primary antibodies: γ-H2AX (Millipore, 05-636) and 53BP1 (Novus, NB100-304), based on the protocol described in Chapter 2. The images were taken by Zeiss LSM880 Confocal at 63X magnification and examined by Zeiss Zen software. More than 600 cells were analyzed per group. Cells with five or more foci were considered positive.

6.4 Results

6.4.1 LMW-E overexpressing increased the inhibitory effect of RAD51 inhibitor B02, CHK1 inhibitor rabusertib, and ATR inhibitor ceralasertib on cell viability

In this chapter we set out to test the hypothesis that small molecule inhibitors targeting Rad51 and ATR-CHK1 pathway would specifically reduce the cell viability in LMW-E overexpressing cells. To this end we used two different model systems: 1) the 76NE6-EKO and 2) MDA-MB-231 breast cancer cells overexpressing LMW-E. The 76NE6-EKO cells induced to express LMW-E were examined for cell viability following treatment with the RAD51 inhibitor B02, CHK1 inhibitor rabusertib and ATR inhibitor ceralasertib. For these experiments, 76NE6-EKO cells were cultured in media containing DMSO (Dox-) or 100ng/mL doxycycline (Dox+, to induce LMW-E expression) for 24 hours, followed by 96 hours treatment with Rad51 inhibitor B02, CHK1 inhibitor rabusertib, or ATR inhibitor ceralasertib(Figure 62). These inhibitors were added to the media containing doxycycline (to induce LMW-E over-expression) or DMSO (control). Following 96 hours of treatment, cells were processed for MTT assay and dose response curves generated and IC50 values determined (Figure 62). Results revealed that upon induction of LMW-E all agents examined could more specifically inhibit cell proliferation when LMW-E was induced as evident with lower IC50 values (Figure 64). The IC50 of CHK1 inhibitor rabusertib is 5-fold higher (p value=0.0071) in the un-induced control group compared with LMW-E overexpressing cells. For ATR inhibitor ceralasertib and Rad51 inhibitor B02, the IC50s are 3-fold higher (p value=0.0085) and 60% higher (p value=0.0117), respectively, in cells without LMW-E overexpression than cells with LMW-E overexpression. These data suggest that LMW-E expression significantly increased the sensitivity of these inhibitors.

Next, we used the TNBC line MDA-MB-231, stably expressing empty-vector or LMW-E to examine the sensitivity to CHK1 inhibitor rabusertib, ATR inhibitor ceralasertib, and

Rad51 inhibitor B02 in a breast cancer cell line (Figure 63). Consistent results to those of 76NE6-EKO cells were observed. Specifically, following 96 hours of inhibitor exposure, results from MTT assay show that the IC50s of all drugs examined is higher in empty vector overexpressing MDA-MB-231 cells than LMW-E overexpressing cells. Similar to the findings in inducible 76NE6-EKO cell lines, the most dramatic change is observed in the cells treated with CHK1 inhibitor (Figure 63A). LMW-E overexpression resulted in   63% (p value=0.0016) decrease of IC50 of CHK1 inhibitor rabusertib. We observed 50% (p value=0.031) and 34% (p value=0.037) decrease respectively for the IC50 ATR inhibitor ceralasertib, and Rad51 inhibitor B02 when we compared the IC50 of the MDA-MB-231 empty vector cells to the LMW-E overexpressing cells (Figure 63 B and C). These results suggest that LMW-E overexpressing also significantly increased the sensitivity to CHK1 inhibitor rabusertib, ATR inhibitor ceralasertib, and Rad51 inhibitor B02 in TNBC cells.

6.4.2 RAD51 inhibitor B02, CHK1 inhibitor rabusertib, and ATR inhibitor ceralasertib induced DNA damage in cells overexpressed with LMW-E

We next set out to interrogate if the treatment of cells with the three DNA damage repair inhibitors differentially induce DNA damage based on the LMW-E status of the cells. To this end, we compared the ratio of cells positive for γ-H2AX and 53BP1 foci as the markers for DNA damage in each of the treated cells as compared to untreated cells, as a function of LMW-E expression. For these experiments, 76NE6-EKO cells were cultured in media containing DMSO (Dox-) or 100ng/ml doxycycline (Dox+, to induce LMW-E expression) for 24 hours, and continued for 24 hours treatment with additional Rad51 inhibitor B02 (3μM), CHK1 inhibitor rabusertib (70nM), or ATR inhibitor ceralasertib (125nM). The results show that in LMW-E overexpressing cells: the ratio of γ-H2AX and 53BP1 positive cells is 120% higher when treated with Rad51 inhibitor B02, and 60% higher in cells treated with CHK1 inhibitor rabusertib, compared to cells without drug exposure.

135
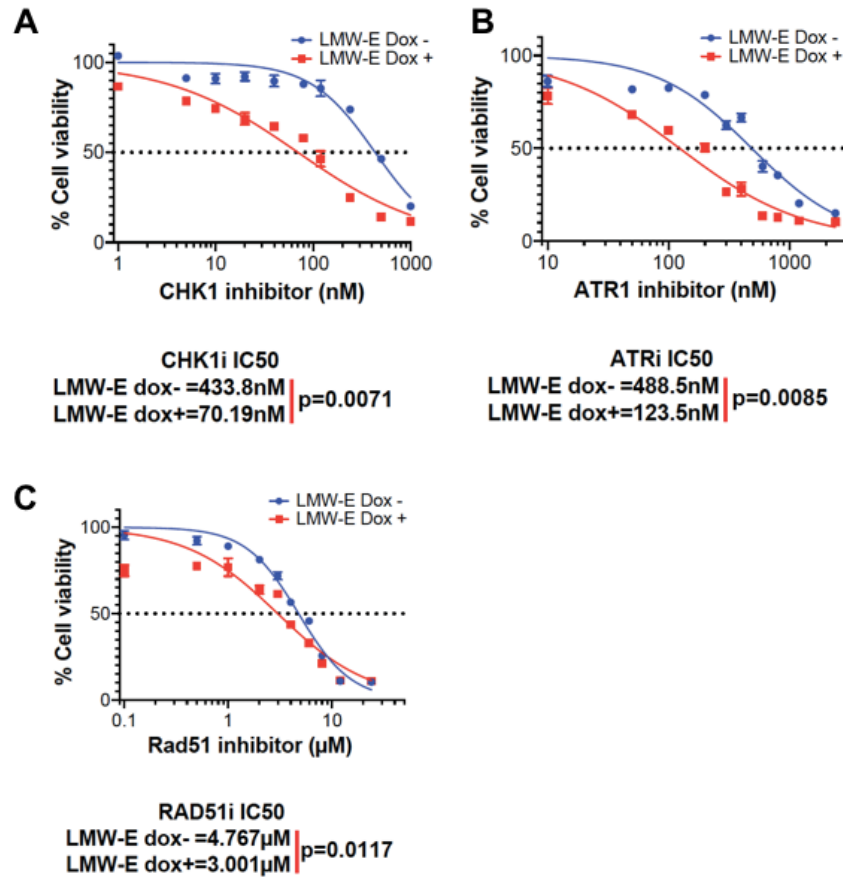
Figure 62. Dose response curves of CHK1 inhibitor rabusertib (panel A), ATR inhibitor ceralasertib (panel B), and Rad51 inhibitor B02 (panel C) in inducible 76NE6-EKO with or without LMW-E over-expression. Inducible 76NE6-EKO cells were cultured in media containing DMSO (Dox-) or 100 ng/mL doxycycline (Dox+, to induce LMW-E overexpression) for 24 hours, followed by the treatment of RAD51 inhibitor (B02), CHK1 inhibitor (rabusertib), or ATR inhibitor (ceralasertib). After 96 hours of inhibitor exposure, cells were subjected for cell viability assay (MTT) to calculate the half-maximal inhibitory concentration (IC50) of CHK1 inhibitor rabusertib (panel A), ATR inhibitor ceralasertib (panel B), and RAD51 inhibitor B02 (panel C). IC50s showing the mean value from 3 biological repeats and 4 technical repeats, p value from student *t* test.
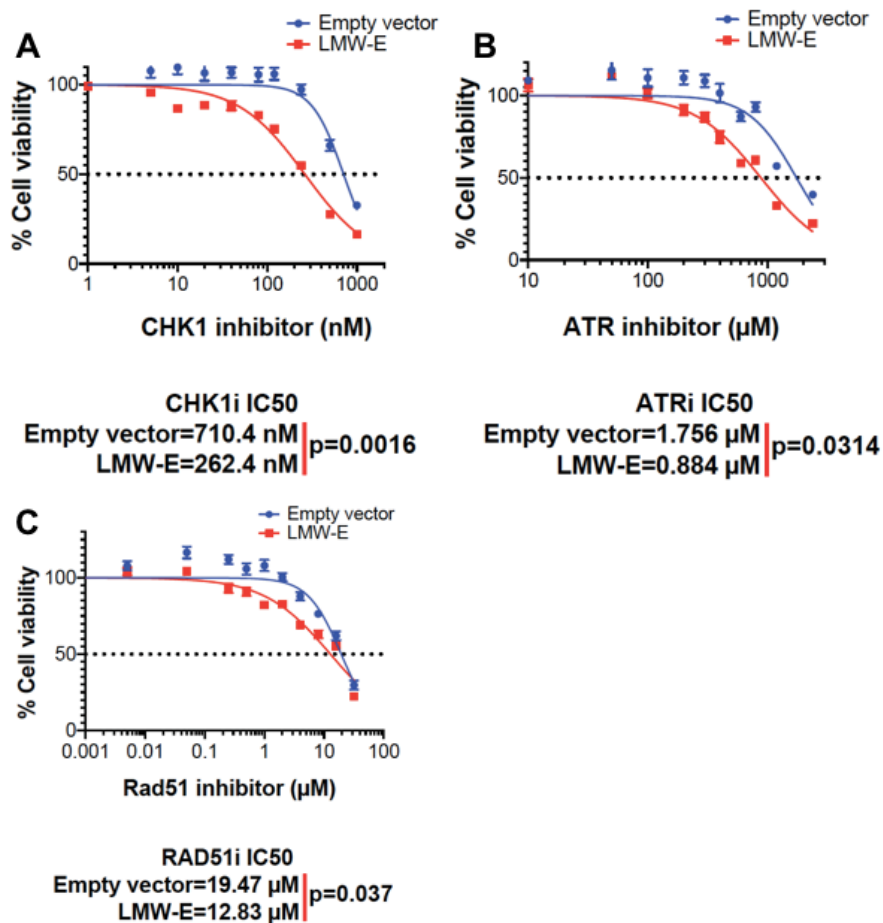
Figure 63. Dose response curves of CHK1 inhibitor rabusertib, ATR inhibitor ceralasertib, and RAD51 inhibitor B02 in MDA-MB-231 cells stably express empty vector or LMW-E. MDA-MB-231-emtpy vector or MDA-MB-231-LMW-E cells were subjected to cell viability assay (MTT) to calculate the half-maximal inhibitory concentration (IC50) for 96 hours treatment of CHK1 inhibitor rabusertib (panel A), ATR inhibitor ceralasertib (panel B), and Rad51 inhibitor B02 (panel C). IC50s showing the mean value from 3 biological repeats and 4 technical repeats, p value from student *t* test.

In cells treated with B02, LMW-E overexpression resulted in 56% increase of γ-H2AX and 53BP1 foci positive cells, compared with the cells without LMW-E induction. Similarly, in cells treated with rabusertib or ceralasertib, LMW-E induction resulted in 70% and 25% increase of γ-H2AX and 53BP1 foci positive cells, respectively (Figure 64).

We then performed DNA damage assay in MDA-MB-231 cell lines engineered to stably express empty-vector or LMW-E. We compared the ratio of γ-H2AX and 53BP1 positive cells in each cell line with or without 48 hours treatment of Rad51 inhibitor B02 (13µM), CHK1 inhibitor rabusertib (260nM), or ATR inhibitor ceralasertib (1µM). The results showed in LMW-E overexpression cells, the ratio of γ-H2AX and 53BP1 positive cells is 1.2-fold higher in cells treated with B02, 4-fold higher in cells treated with rabusertib, and 6-fold higher in cells treated with ceralasertib, compared cells without drug exposure. Additionally, In the CHK1 inhibitor rabusertib treated LMW-E overexpressing cells, the ratio of γ-H2AX positive cells and 53BP1 positive cells are 1.56-fold and 2.35-fold of empty vector cells, respectively (Figure 65). These results suggest blocking DNA damage repair pathways by small molecule inhibitors increase the DNA damage in LMW-E overexpressing hMECs and breast cancer cell line. Overexpression of LMW-E also increase the sensitivity of MDA-MB-231 cells to CHK1 inhibitor rabusertib.

To examine the effect of ATR-CHK1-RAD51 inhibitors on DNA damage responses and cell apoptosis, we next performed western blot to analyses of phospho-CHK1, cleaved PARP in the inducible 76NE6-EKO-LMWE cell lines. As expected, the phosphorylation of CHK1 at S345 were blocked by ATR inhibitor in both inducible 76NE6-EKO-LMW-E cells as well as MDA-MB-231 cells. CHK1 inhibitor promoted CHK1 pS345 MDA-MB-231 cell lines but not inducible 76NE6-EKO-LMWE cell lines. The results also show that when LMW-E was induced, treatment of cells with RAD51 inhibitor led to 1.5 times increase of cleaved PARP, a marker for apoptosis[144] (Figure 66A). Similarly, B02 treatment of MDA-MB-231 LMW-E overexpressing cell line showed a 5-fold increase of cleaved PARP, when compared to the LMW-E overexpression cells without drug exposure (Figure 66B). These data suggest RAD51 inhibitor B02 may induce apoptosis in LMW-E overexpressing hMECs and breast cancer cells.
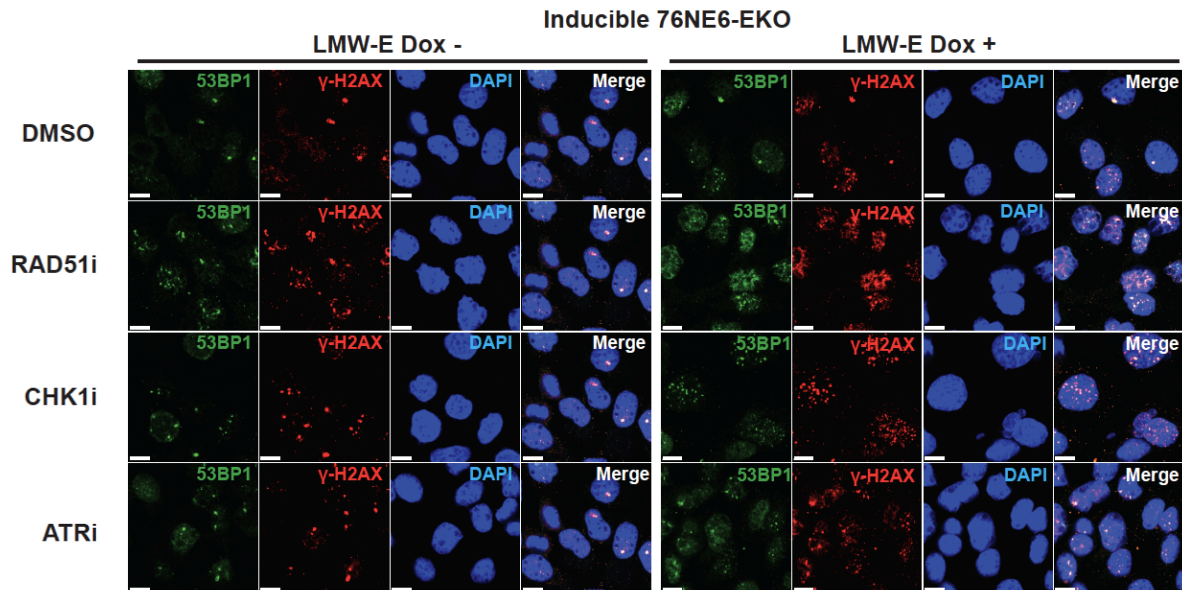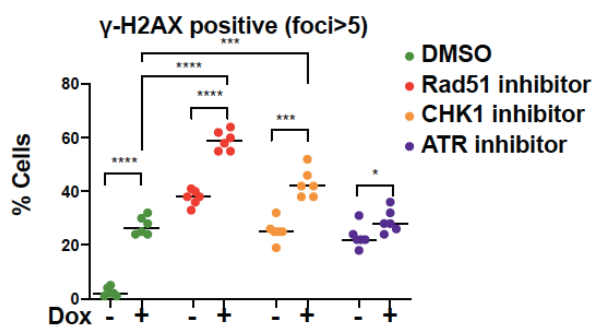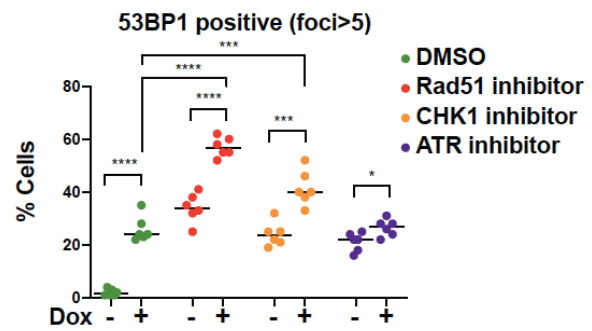
Figure 64. Treatment of Rad51 inhibitor B02, CHK1 inhibitor rabusertib, or ATR inhibitor ceralasertib enhanced DNA damage in inducible 76NE6-EKO-LMW-E cells. Inducible 76NE6-EKO-LMW-E cells were cultured in media containing DMSO (Dox-) or 100ng/ml doxycycline (Dox+, to induce LMW-E expression) for 24 hours, and continued for 24 hours treatment with additional Rad51 inhibitor B02 (3μM), CHK1 inhibitor rabusertib (70nM), or ATR inhibitor ceralasertib (125nM). A. representative images for DNA damage assay examined by immunofluorescence for γ-H2AX foci and 53BP1 foci (scale bar=10μm). B and C. quantitation of the ratio of γ-H2AX- positive cells (B, foci > 5) and 53BP1-positive cells (C, foci > 5). Cell number > 400, *p < 0.05, **p < 0.01, ***p < 0.001, and ****p < 0.0001, Student *t* test.
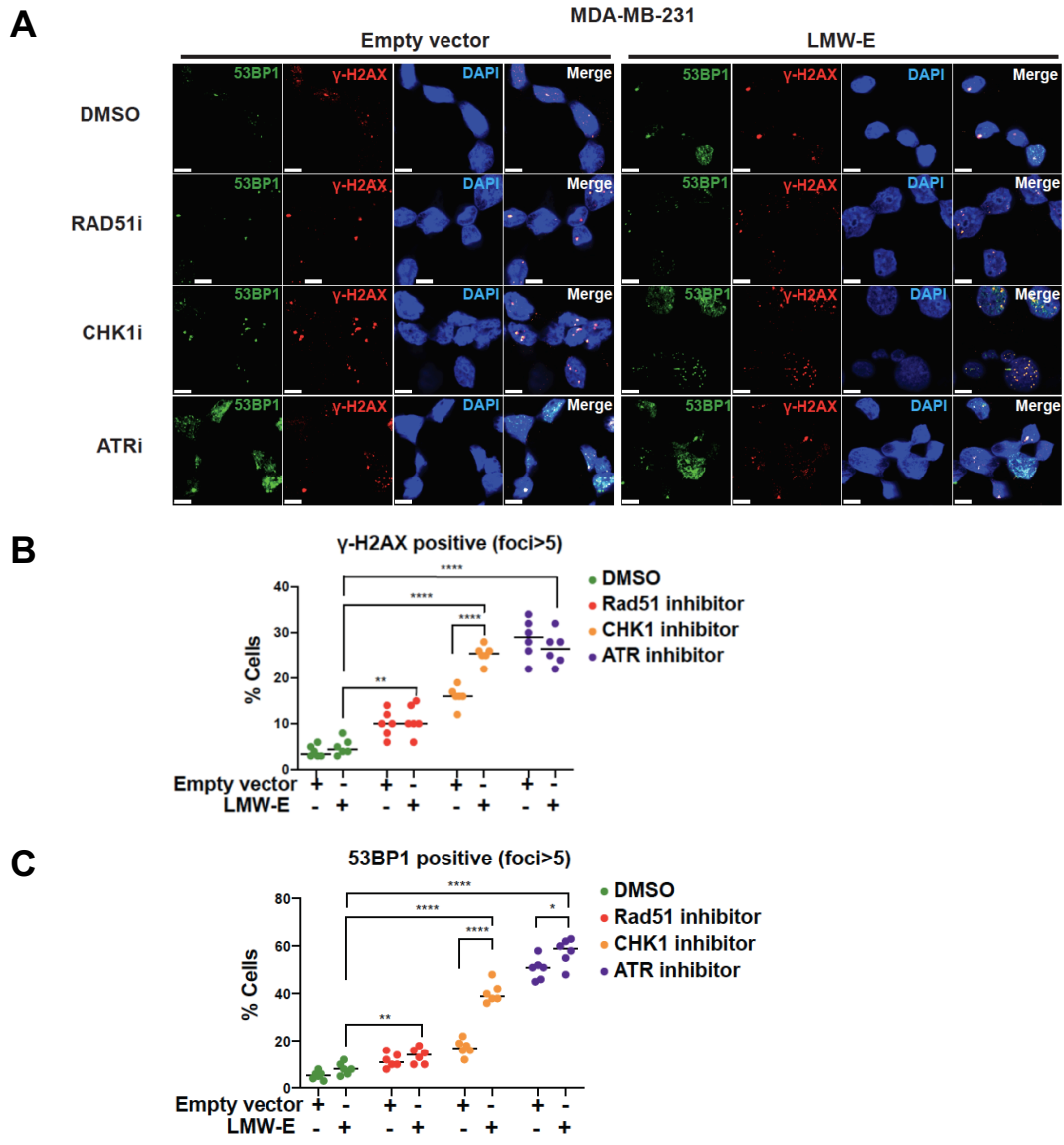
Figure 65. Treatment of Rad51 inhibitor B02, CHK1 inhibitor rabusertib, and ATR inhibitor reralasertib induced DNA damage in MDA-MB-231 stable cell lines. MDA-MB-231 empty vector or LMW-E stable cell lines were treated with Rad51 inhibitor B02 (13µM), CHK1 inhibitor rabusertib (260nM) or ATR inhibitor ceralasertib (1µM) for 48 hours, followed by immunofluorescence assay for γ-H2AX foci and 53BP1 detect the DNA damage. A. Representative images of γ-H2AX foci and 53BP1 foci (scale bar=10µm). B and C. quantitation of the ratio of γ-H2AX- positive cells (B, foci > 5) and 53BP1-positive cells (C, foci > 5). Cell number > 400, *p < 0.05, **p < 0.01, ***p < 0.001, and ****p < 0.0001, Student *t* test.
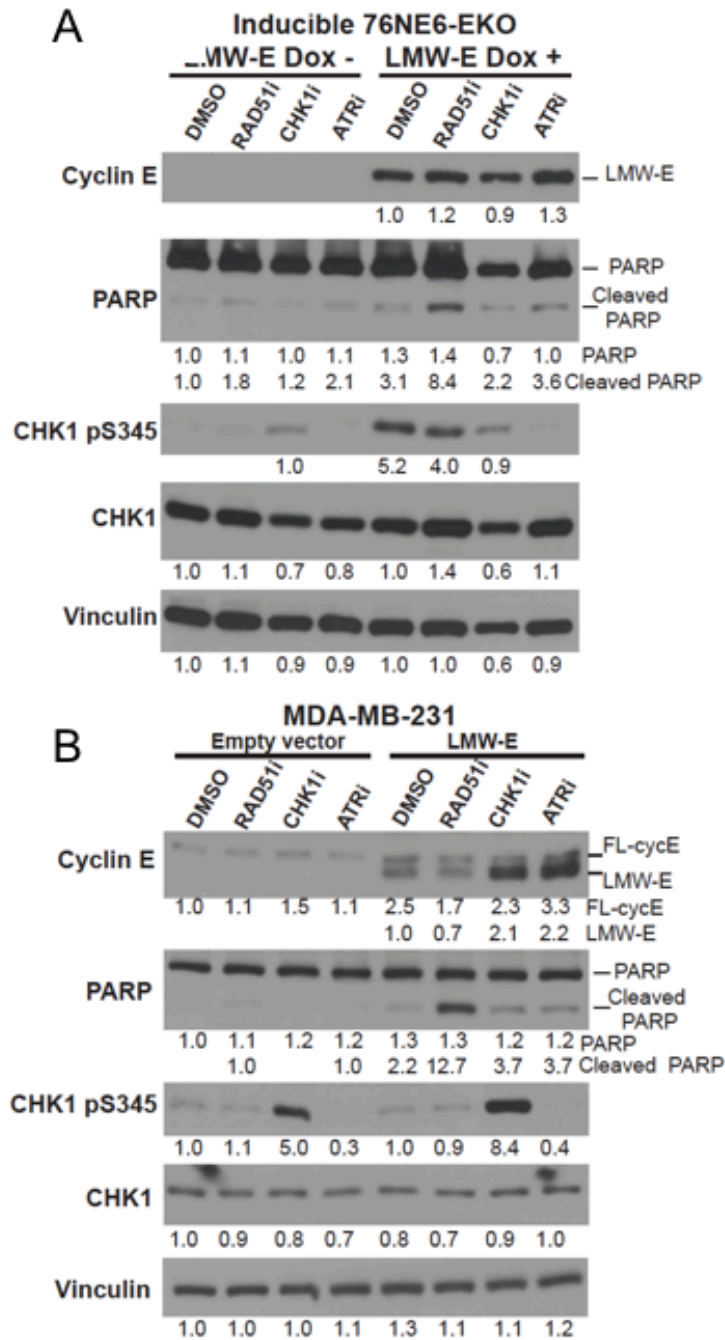
Figure 66. Western blot analysis for DNA damage and apoptosis marker in inducible 76NE6-EKO-LMW-E cells and MDA-MB-231 stable cell lines. The result showed RAD51 inhibitor B02 treatment induced cleaved PARP, the apoptotic marker, in inducible 76NE6-EKO-LMW-E cells and MDA-MB-231cells when LMW-E were overexpressed.

6.5 Conclusion

In this chapter, we used small molecule inhibitors targeting ATR-CHK1-RAD51 pathway to investigate the potential therapeutic targets when LMW-E is expressed. We applied inducible 76NE6-EKO cells with or without doxycycline to induce LMW-E, as well as TNBC lines MDA-MB-231 that were engineered to stably express empty vector or LMW-E as our model systems. Based on cell viability measured by MTT assay, we have determined that LMW-E overexpression significantly decreased the IC50s of RAD51 inhibitor B02, CHK1 inhibitor Rabusertib, and ATR inhibitor Ceralasertib.

Treatment of these drugs also led to significantly enhanced DNA damage in LMW-E overexpressing cells, as shown by increased ratio of cells positive for γ-H2AX- and 53BP1 foci. RAD51 inhibitor B02 treatment led to increased level of apoptosis marker- cleaved PARP in LMW-E overexpressing cells. These findings are consistent in both hMEC model 76NE6-EKO-LMW-E cells and TNBC model MDA-MB-231-LMW-E stable cell lines. Based on these findings, ATR-CHK1-RAD51 pathway may serve as druggable target in LMW-E [high] breast cancers. Among the drugs tested in this chapter, CHK1 inhibitor Rabusertib is the most effective (lowest IC50). RAD51 inhibitor B02 specifically induced apoptosis in LMW-E overexpressing cells, and may benefit from apoptosis modulators to maximize the killing effect.

Chapter Seven: LMW-E predicts genomic instability in early-stage breast cancers

7.1 Introduction

Thus far, we have established that LMW-E overexpression in human mammary epithelial cell lines promote genomic instability featured by aberrant chromosomal structures and abnormal nuclear phenotypes. Mechanistically, overexpression of LMW-E, similar to but independent of its precursor FL-cycE, promotes replication stress and DNA damage repair. In this chapter, we aim to translate our findings to breast cancer patients by interrogating if there is an association between LMW-E status and genomic instability in their tumors.

To this end, we used the tumor samples from a retrospective cohort of 725 stage I/II breast cancer patients, who were all treated at MD Anderson Cancer Center and for whom we have comprehensive clinical and pathological annotation and follow up. These tumors were previously analyzed by immunohistochemistry (IHC) for the status of LMW-E and by high-density molecular inversion probe (MIP) array[145] based targeted sequencing to calculate copy number (CN) gains and losses.

By sub-grouping the samples according to their LMW-E-positive or negative status, we performed bio-informatics analysis to determine the association between LMW-E and genomic instability as a function of CN gains and losses for each chromosomal region. Together with other clinical and pathological features, we set out to test the hypothesis that LMW-E status independently predicts genomic instability and clinical outcome in early stage breast cancer patients (Figure 67).

7.2 Schematics of model system
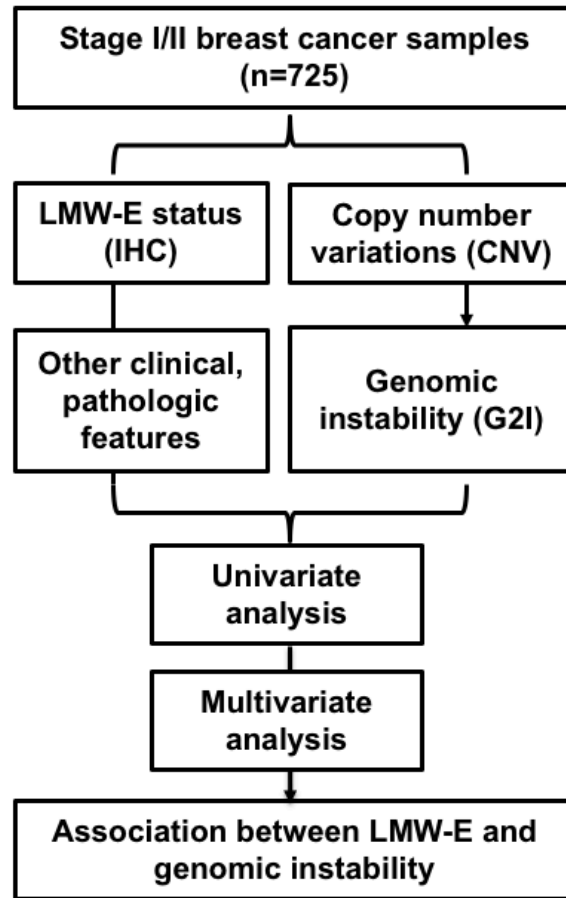
Figure 67. Schematics for association analysis between LMW-E status and genomic instability in breast cancer patient(n=725). We use a retrospective cohort of 725 patients with stage I-II breast cancer to examine if LMW-E levels predicts genomic instability. For each sample, we have the status of LMW-E (positive or negative) and MIP copy number data. We further applied an index for genomic instability (G2I) to stratify the patient cohort into genomic instability high and low groups and determined the association between LMW-E and genomic instability status by univariate logistic regression and multi-variate logistic regression models.

7.3 Materials and Methods

Patients and tissue samples

The study was approved by the Institutional Review Board of The University of Texas MD Anderson Cancer Center. The patients for this study include a retrospective cohort of 725 patients with stage I-II breast cancer treated at MD Anderson (Houston, TX) between 1985 and 1999[60,145]. Clinical information, including patient age, T category, nodal status, tumor grade, subtype, and low molecular weight cyclin E (LMW-E) status, was abstracted from a previous report[60]. In brief, to determine the status of LMW-E, the formalin-fixed, paraffin-embedded (FFPE) breast tumor samples were deparaffinized, rehydrated, and treated with 3% hydrogen peroxide and methanol to block the peroxidase activity and nonspecific protein-protein interactions. Antigen retrieval was performed by applying 0.01mM citric acid-based buffer at pH 6.0 for 15 min before immunostaining, followed by 1-hour blocking. The staining cyclin E was performed by using rabbit polyclonal antibody to cyclin E (Santa Cruz, C-19, sc-198) and the signals were detected by VECTASTAIN Elite ABC kit (PK6101 and PK6102; Vector Laboratories). For each tumor sample, LMW-E status was assigned as follows: LMW-E negative (no staining or just nuclear staining), LMW-E positive (nuclear + cytoplasmic or just cytoplasmic staining).

Molecular inversion probe-based arrays for copy number measurement

Raw data of molecular inversion probe (MIP)-based arrays from tumor DNA isolated from patient tissues in formalin-fixed, paraffin-embedded blocks (n = 725) and matched non–tumor-bearing lymph node formalin-fixed, paraffin-embedded blocks (n = 129 cases) were abstracted from a previous report[145]. In brief, the MIP assay was performed by the Affymetrix™ MIP laboratory, and the laboratory was blinded to all sample and subject information. Data from the MIP assay data, we generated the raw copy number data consisted of total copy number and B-allele frequencies for 201,032 molecular inversion probe arrays.

Among the 725 samples, 683 were diploid, 25 were triploid, and 17 were tetraploid tumors. By using the allele-specific copy number analysis of tumors (ASCAT) algorithm provided by Van Loo et al.,[146] we obtained 48,623 intervals to generate the copy number gains and losses for each of the sample (within each interval, the copy number did not change for any of the samples).

Genomic instability index (G2I)

The G2I algorithm was performed on copy number data using previously reported R scripts by Bonnet et al [147]. G2I used copy number data for a two-parameter index representing the overall level of genomic alteration and the number of altered genomic regions. The overall level of genomic alteration (A) is computed as the mean value of the altered probes divided by total probes. The number of altered regions (N) is computed by altered genomic regions along the genome. For a given sample i, if $A_i < a1$ and $N_i < n1$, then the genomic instability index = 1; if $A_i > a3$ and $N_i > n3$, then the genomic instability index = 3; otherwise genomic instability index = 2. The algorithm g2i.learn was provided by Bonnet et al[147].

Statistical analysis

Univariable and multivariable logistic regression were performed to determine whether LMW-E status and other clinical information, including patient age, T category, lymph node status, nuclear grade, and tumor subtype (based on estrogen receptor and progesterone receptor status), were associated with genomic instability as determined by the G2I scores. If the G2I score is 1 or 2, the genome is considered stable. If the G2I score is 3, then the genome is considered unstable.

Univariable and multivariable regression analyses using the Cox proportional hazards model were performed to determine whether age, T category, nodal status, tumor grade, subtype, chemotherapy, radiotherapy, endocrine therapy, G2I score, and LMW-E status were

146

associated with breast cancer freedom from recurrence. Freedom from recurrence is modified from guidelines reported by Hudis et al[148] for recurrence-free survival calculating only recurrences (locoregional or distant) and not including deaths as events, regardless of cause of death. The Kaplan–Meier method was used to calculate 15-year freedom from recurrence for each factor. Differences in survival between LMW-E–positive and LMW-E–negative samples were evaluated using the log-rank test.

7.4 Results

7.4.1 LMW-E status and patient characteristics

To study the association of LMW-E status and genomic instability in breast cancer, we analyzed a retrospective cohort of 725 patients with stage I-II breast cancer. This cohort contains 158 (22%) had HER2+ breast cancer, 420 (58%) had ER+/PR+ breast cancer, and 140 (20%) had TNBC (Figure 68). Samples from each patient were represented by spots on a tissue microarray slide and all slides were stained for cyclin E using C-terminal antibody to determine the status of LMW-E. In particular, LMW-E negative samples show no staining or just nuclear staining, and LMW-E positive samples show nuclear plus cytoplasmic staining or just cytoplasmic staining[60]. Three pathologists independently scored each sample for its nuclear and cytoplasmic scoring of cyclin E[39,60] . Comparing the LMW-E status with breast cancer subtypes, the result show LMW-E–positive (n = 427) contained a higher proportion of HER2+ and TNBC subtypes than LMW-E–negative subgroup (n = 298) (Figure 68), and breast cancer subtypes (HER2+, ER+/PR+, or TNBC), nuclear grade (I, II, or III), and nodal status (negative or positive) represented statistically distinct LMW-E status (Table 6). We used freedom from recurrence (FFR) as a measure of outcome for the patients in this cohort[148]. The Kaplan-Meier FFR curves suggest LMW-E-positive subgroup exhibit significantly increased recurrence compared with the LMW-E–negative subgroup (Figure 69). These results suggest although the patient in our cohort were distinct in pathological features such

as breast cancer subtypes, nuclear grade levels, nodal status, and tumor metastasis, LMW-E status may predict the clinical outcome of the early stage breast cancer patients independently of these other factors.
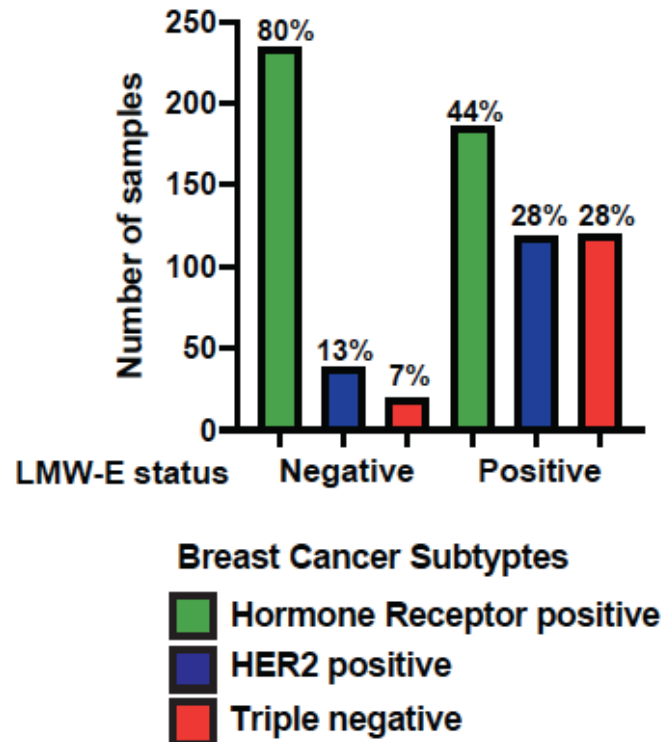
Figure 68. The distribution of breast cancer subtypes in LMW-E-negative subgroup (n=298) and LMW-E-positive subgroup (n=427) in our cohort of 725 patients with stage I and II breast cancer. Higher proportion of HER2+ and TNBC subtypes were observed in LMW-E-positive subgroup compared with the LMW-E–negative subgroup.

	LMW-E -, n (%)	LMW-E +, n (%)	Overall p-value	Odds ratio (95% confidence interval)	p-value	Odds ratio (95% confidence interval)	p-value
Subtype			6.61E-23				
ER+/PR+	234 (55.7)	186 (44.3)		Reference		0.26 (0.17-0.4)	1.78E-11
Her2+	39 (24.7)	119 (75.3)		3.83 (2.5-5.9)	1.78E-11	Reference	
TNBC	20 (14.3)	120 (85.7)		7.52 (4.5-13.3)	7.36E-19	1.96 (1.05-3.77)	0.0289
NA	5	2					
Nuclear grade			1.15E-13				
I	38 (59.4)	26 (40.6)		Reference			
II	174 (49.9)	175 (50.1)		1.47 (0.83-2.6)	0.175		
III	60 (22.6)	205 (77.4)		4.96 (2.7-9.3)	3.60E-08		
NA	26	21					
Nodal status			0.0032				
Negative	194 (45.9)	229 (54.1)		Reference			
Positive	100 (34.6)	189 (65.4)		1.6 (1.16-2.21)	0.0032		
NA	4	9					

Table 6. Association of low molecular cyclin E (LMW-E) status with tumor characteristics in early stage breast cancer samples (n = 725). Clinical information, including tumor subtypes, tumor grade, nodal status, and low molecular weight cyclin E (LMW-E) status, was abstracted from a previous report[60]. Fisher's exact tests were performed to compare the ratio of each clinical feature in LMW-E-positive subgroup and LMW-E-negative subgroup.
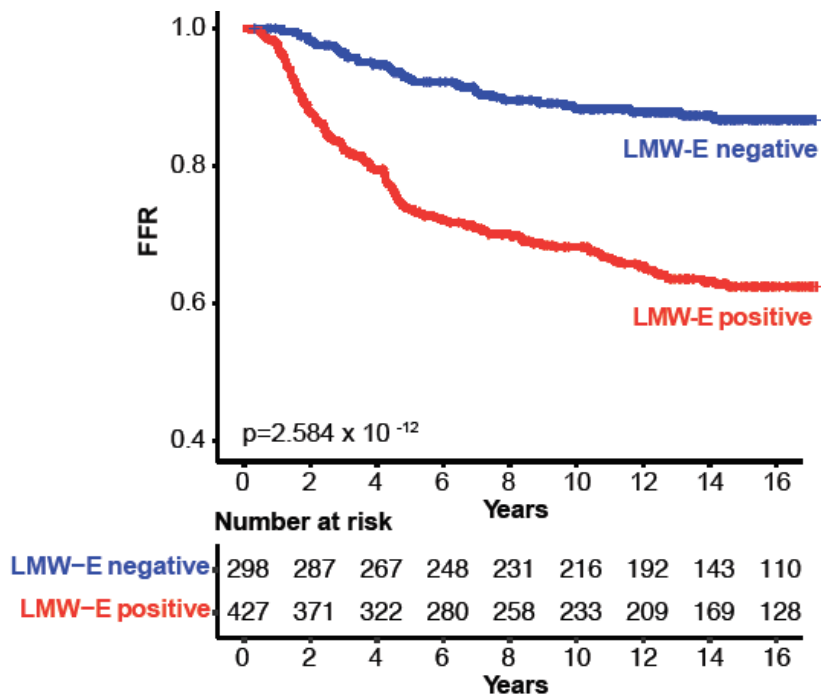
Figure 69. Kaplan-Meier survival curves showing the association between LMW-E status and freedom from recurrence (FFR) in a cohort of 725 patients with stage I and II breast cancer. The subgroup of our cohort that was LMW-E–positive (n = 427) showed significantly increased tumor recurrence and worse clinical outcomes compared with the LMW-E–negative subgroup(n=298; p=2.584 X 10 [-12]).

7.4.2 LMW-E status is associated with copy number variations (CNVs) in breast cancer patients

The CNVs in the 725 breast cancer samples were determined by Molecular inversion probe (MIP) based arrays[145]. The raw copy number data consist of total copy number and B-allele frequencies for 201,032 MIPs. We applied the ASCAT algorithm to simultaneously estimate the normal contamination percentage and tumor ploidy. Among the 725 samples there are 683 diploid, 25 triploid, and 17 tetraploid tumors. Based on the MIP copy number data, we estimated integer A and B allele counts across the genome. We assigned copy

number gains (or losses) when the total copy number (A+B allele counts) is greater (or less) than the ploidy. For each sample we obtained a list of segments and an associated gain (or loss) frequency. The results revealed that the CNV frequency in LMW-E–positive subgroup is significantly higher than LMW-E–negative subgroup (Figure 70).
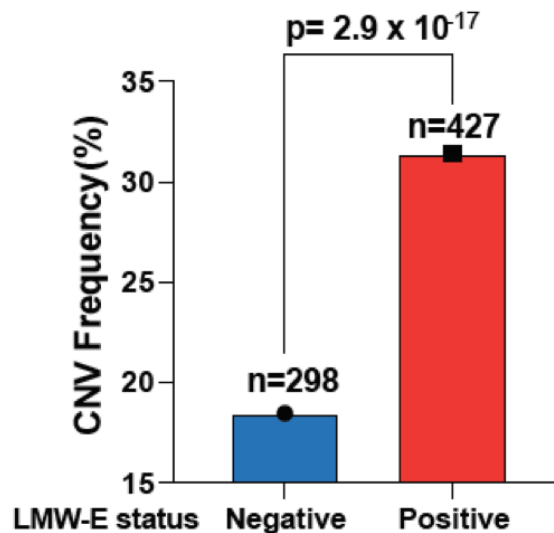
Figure 70. Comparison of copy number variation (CNV) frequency in LMW-E–negative and LMW-E–positive subgroups in our patient cohort. LMW-E positive samples (n=427) show significantly higher CNV frequency compared with LMW-E negative samples (n=298).

To identify the gains and losses associated with LMW-E status, we used Bonferroni correction (p-value cutoff=0.05, number of tests=48623) to determine the significantly different regions between LMW-E–positive and LMW-E–negative subgroups (Figure 71). Both LMW-E-positive and negative subgroups show recurrent gains of the 1q (arm), 8p11-q24, 14q11.2, and 20q13 and losses at 8p23-p12, consistent with features in breast cancer CNVs regardless of LMW-E status[145]. Specifically, regions with CN gains containing oncogenes such as *JUN* (1p32), *FOXC1* (6p25), *CCNE2* (8q22), *WNT5B* (12p13), *CCND2* (12p13), *FOXM1* (12p13), and *ERBB2* (17q12), as well as regions with CN loss containing *TP53* (17p13), were strongly associated with LMW-E–positive tumors. CN gains in *CCNE1* (cyclin E1) were also associated

with LMW-E–positive tumors (p = 0.0023) (Figure 71). Considering upregulated DNA replication and DNA repair pathways are enriched in LMW-E overexpressed mammary epithelial cells, we next interrogated if such alterations at genetic level are also associated with LMW-E positive breast tumors. Our analysis revealed that LMW-E–positive tumors harbored significantly higher CN gains in genes involved in the DNA pre-replication complex (such as *CDC6*, *MCM*s, *POLE3*, *POLE4*) (Figures 71 and 72) and DNA damage repair (such as *RAD51B*, *RAD51AP1*, *RAD52*, and *RAD54B*) (Figure 71) and non-homologous end joining pathway genes (such as *DCLRE1C*, *DNTT*, *NHEJ1*, *XRCC5*, and *LIG4*; Figure 73).
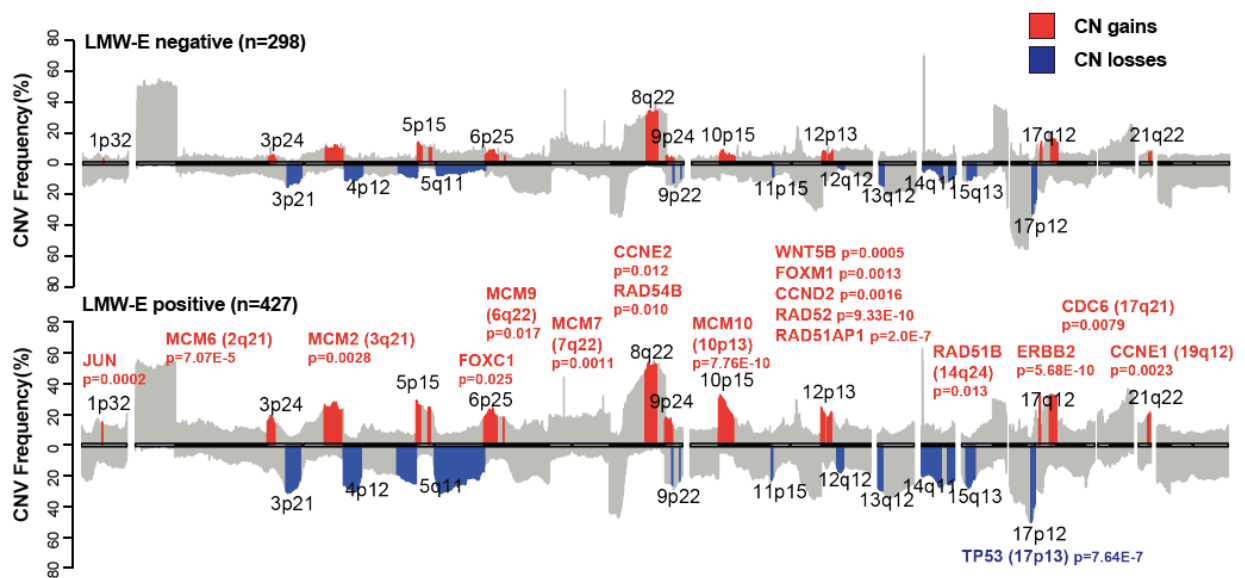
Figure 71. Association plot demonstrating the frequency of copy number variations (CNVs) in LMW-E–negative (n = 298) and LMW-E–positive (n = 427) tumors compared with normal tissue control. The colored patches highlight significantly different CN gains (red) and losses (blue) in LMW-E–positive tumors compared with LMW-E–negative tumors.

7.4.3 LMW-E positive status predicts genomic instability

Previous studies established an index of genetic instability based on two parameters linked to the overall level of genomic alteration and the number of altered regions, which can be calculated by CNVs across the genome[147]. To examine the association between LMW-E status and genomic instability, we used the genomic instability index (G2I), to stratify our patient cohort (n = 725) into subgroups of stable genomes (G2I = 1; n = 137), intermediately stable genomes (G2I = 2; n = 425), and unstable genomes (G2I = 3; n = 163; Figure 74).
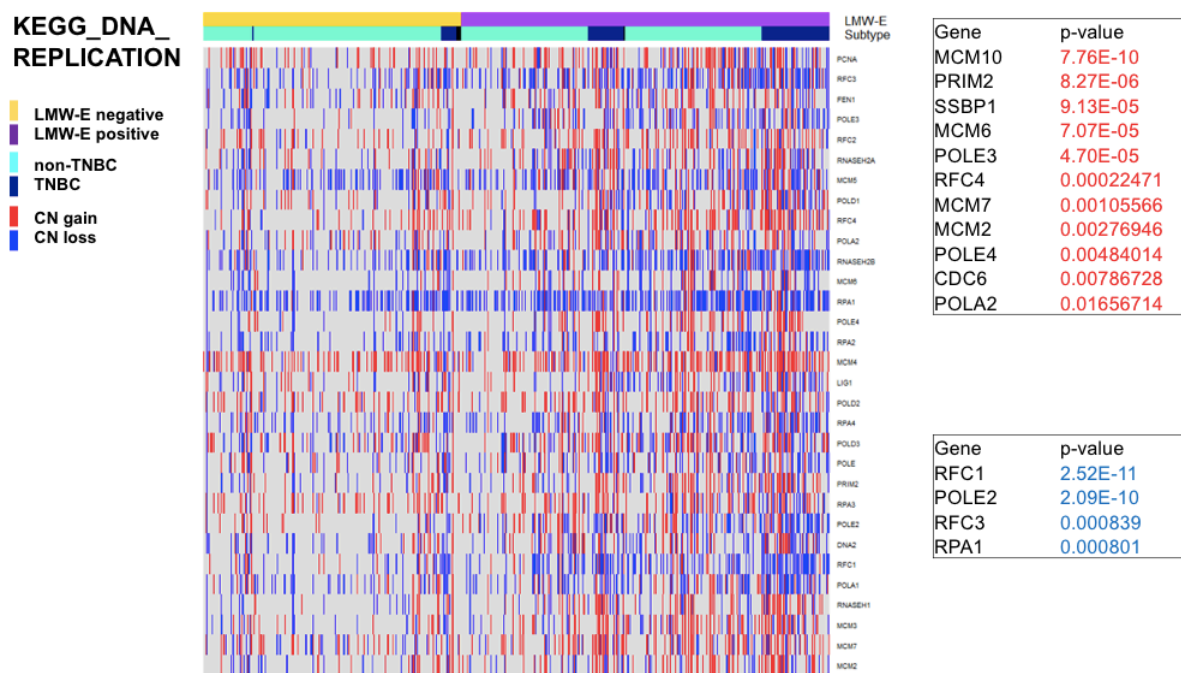
Figure 72. Association heatmap demonstrating the gains(red) and losses(blue) of genes in KEGG DNA replication pathway in LMW-E–positive (n = 427) subgroup compared with LMW-E–negative (n = 298) subgroup. The copy number gains and losses of genes in KEGG DNA replication pathway were compared between LMW-E positive samples versus LMW-E negative samples. Genes that showed significantly gains in LMW-E-positive sub-group were labeled red, and genes showed significant losses in LMW-E-positive sub-group were labeled blue, both with p-values listed.
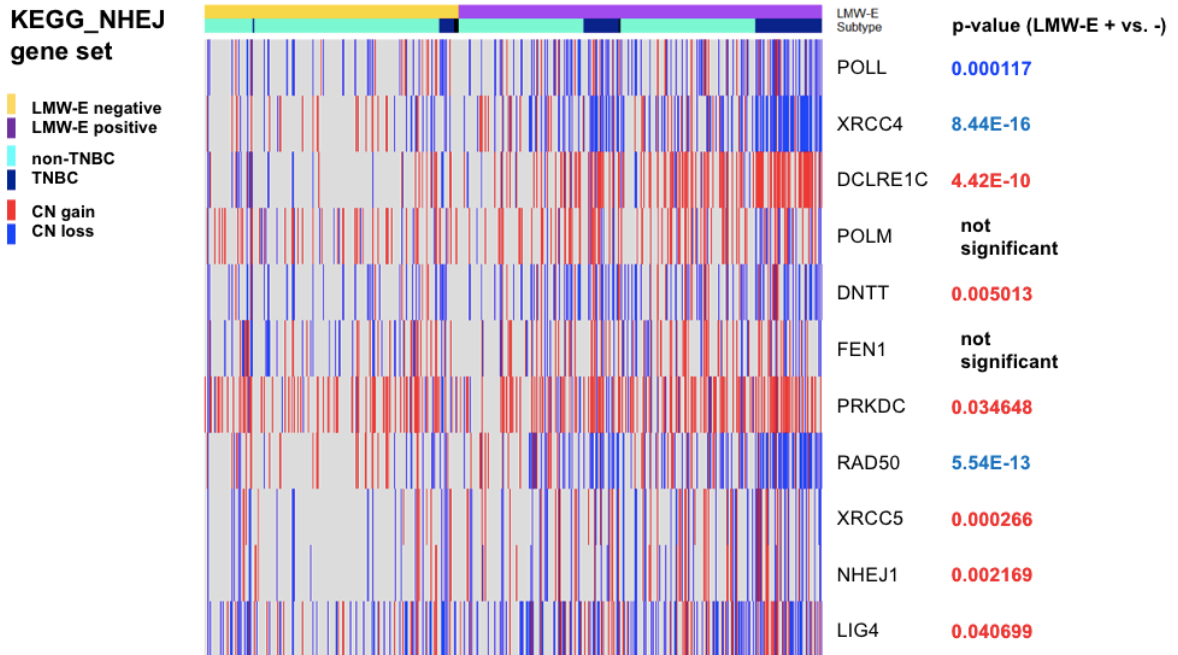
Figure 73. Association heatmap demonstrating the gains(red) and losses(blue) of genes in KEGG non-homologous end joining (NHEJ) pathway in LMW-E–positive (n = 427) subgroup compared with LMW-E–negative (n = 298) subgroup. The copy number gains and losses of genes in KEGG non-homologous end joining pathway were compared between LMW-E positive samples versus LMW-E negative samples. Genes that showed significantly gains in LMW-E-positive sub-group were labeled red, and genes showed significant losses in LMW-E-positive sub-group were labeled blue, both with p-values listed.

We observed significantly different distribution of samples grouped by LMW-E status and genomic stability status (Fisher test, p = 4.9 ×10^{-14}). Most tumors (>84%) with unstable genomes (G2I = 3) were LMW-E–positive (Figure 75). These results suggest LMW-E positive status is also associated with higher genomic instability in patient samples.
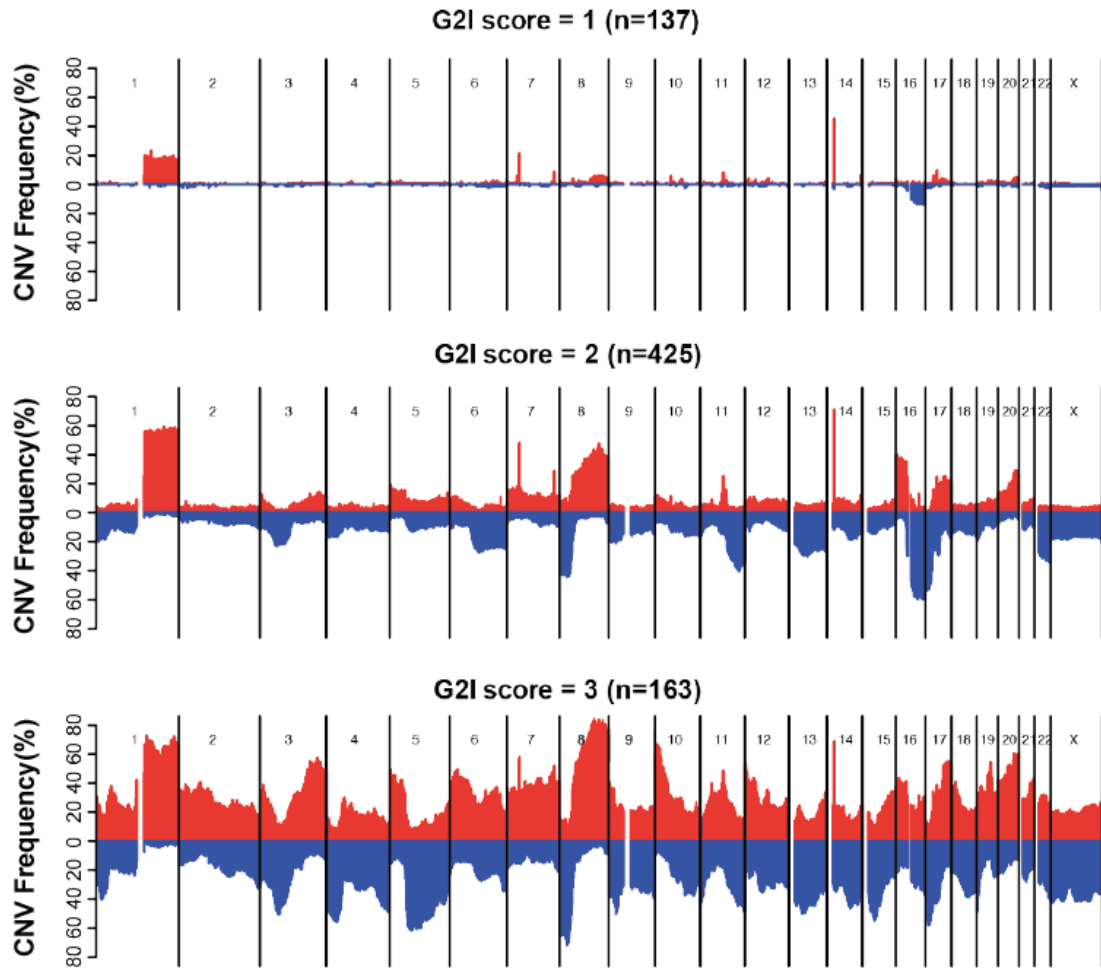
Figure 74. Association plot showing copy number variation (CNV) frequency stratified by genomic instability index (G2I): stable genomes (G2I = 1, n = 137), intermediately stable genomes (G2I = 2, n = 425), and unstable genomes (G2I = 3, n = 163).
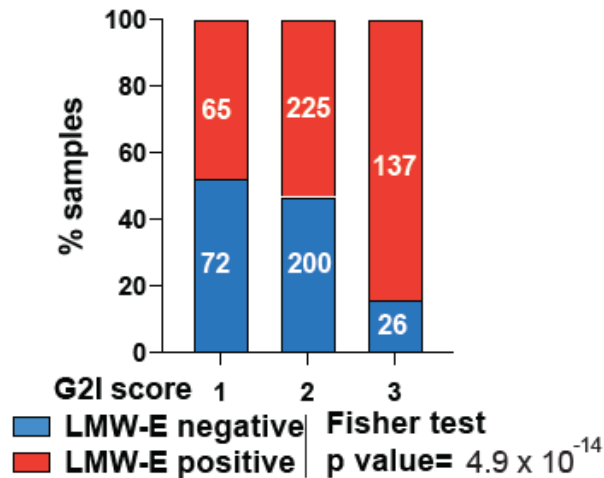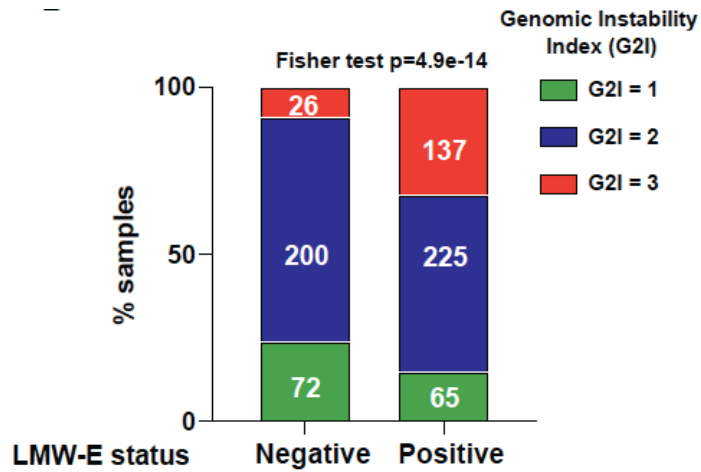
Figure 75. Distribution of LMW-E–negative tumors and LMW-E–positive tumors stratified by genomic instability index (G2I). The proportions of genome stability status were compared between LMW-E–negative and LMW-E–positive subgroups using the Fisher test. The results suggest LMW-E positive subgroup contain more genomic unstable samples (G2I=3, 31%) than LMW-E negative subgroup (9%). Within the unstable samples, 84% belong to LMW-E positive subgroup. (p = 4.9 ×10^{-14}).

To directly examine if LMW-E status predicts genomic instability, we further assigned the high G2I score (G2I=3) as the positive parameter for genomic instability, and low and medium G2I scores (1 and 2) as the negative reference. We initially used univariate logistic regression model to calculate the odds ratio, confidence interval and p-value for the association from LMW-E status and multiple clinical and pathological features (age, T stage, nodal status, tumor grade, tumor subtypes, chemotherapy, radio therapy, and endocrine therapy) to the genomic instability status. The results suggest that variables such as positive LMW-E status, T stage=2, and TNBC subtype, are strongly associated with genomic instability of breast cancers (p<0.01, Table 7).

Next, we subjected significant factors from the univariate analysis, to multivariate logistic regression model to determine which of the factors are independent variables. To this end, we calculated the odds ratio, confidence interval and p-value for the association from the selected features, LMW-E status, T stage, and tumor subtype, to genomic instability (Table 8). The results show that LMW-E status is an independent factor that strongly correlates with (p=7.38E-06) genomic instability in breast tumors (stage I/II, n=725). Collectively, these results have established the positive correlation between LMW-E expression and genomic instability in early stage breast tumors, and that LMW-E could be used as an independent predictor for genomic instability in breast cancer patients.

Additionally, positive LMW-E status, unstable genome, together with age, tumor T stage, nodal status, and TNBC status were shown to significantly correlated with poor FFR and clinical outcome in univariate cox regression model (Table 9). Multivariate cox regression model further suggest that LMW-E could be used as an independent predictor for genomic instability and poor clinical outcome in breast cancer patients (Table 10).

Variable	Odds ratio	Confidence interval (95%)	p-value
Age			
<50	Reference		
50-75	0.8	0.56-1.15	0.231574
>=75	0.68	0.34-1.39	0.293686
T stage			
1	Reference		
2	2.16	1.51-3.08	2.19E-05
Nodal status			
Negative	Reference		
Positive	0.73	0.51-1.06	0.096754
Tumor grade			
1	Reference		
2	0.57	0.31-1.05	0.070566
3	0.91	0.49-1.67	0.759247
Tumor subtype			
non-TNBC	Reference		
TNBC	7.4	4.93-11.1	4.26E-22
LMW-E status			
Negative	Reference		
Positive	4.94	3.15-7.76	3.71E-12

Table 7. Univariate logistic regression model results for the status of genomic instability in early stage breast cancer samples (n=725). The results suggest that variables such as positive LMW-E status, T stage=2, and TNBC subtype, are strongly associated with genomic instability of breast cancers.

Variable	Odds ratio	Confidence interval (95%)	p-value
T stage			
1	Reference		
2	1.6	1.06-2.43	0.026626
Tumor subtype			
non-TNBC	Reference		
TNBC	4.32	2.7-6.91	1.06E-09
LMW-E status			
Negative	Reference		
Positive	3.08	1.88-5.04	7.38E-06

Table 8. Multivariate logistic regression model using variables (selected from univariate logistic regression model) for the status of genomic instability in early stage breast cancer samples (n=725). The results show that LMW-E status is an independent factor that strongly predicts (p=7.38E-06) genomic instability in breast tumors.

Variable	Hazard ratio	Confidence interval (95%)	p-value
Age			
<50	Reference		
50-75	0.62	0.46-0.83	0.001385
>=75	0.46	0.22-0.94	0.033363
T stage			
1	Reference		
2	2.23	1.65-3	1.41E-07
Nodal status			
Negative	Reference		
Positive	1.97	1.46-2.64	7.85E-06

Tumor grade			
1	Reference		
2	0.86	0.5-1.47	0.580861
3	1.3	0.76-2.23	0.33044
Tumor subtype			
non-TNBC	Reference		
TNBC	1.42	1.01-2	0.044061
Chemotherapy			
no	Reference		
yes	1.5	1.12-2.01	0.006838
Radiotherapy			
no	Reference		
yes	0.98	0.73-1.32	0.901789
Endocrine therapy			
no	Reference		
yes	0.49	0.36-0.67	1.14E-05
LMW-E status			
Negative	Reference		
Positive	3.44	2.38-4.98	4.91E-11
Genomic instability			
Low	Reference		
High	1.39	1.01-1.93	0.045544

Table 9. Univariate cox regression model results for Freedom From Recurrence (FFR).

The result suggests positive LMW-E status, unstable genome, together with age, tumor T stage, nodal status, and TNBC status were significantly correlated with poor FFR.

Variable	Hazard ratio	Confidence interval (95%)	p-value
Age			
<50	Reference		
50-75	0.68	0.48-0.95	0.023361
>=75	0.49	0.23-1.06	0.069608
T stage			
1	Reference		
2	1.93	1.4-2.66	5.37E-05
Nodal status			
Negative	Reference		
Positive	1.94	1.4-2.69	7.22E-05
Tumor subtype			
non-TNBC	Reference		
TNBC	0.83	0.55-1.25	0.373222
Chemotherapy			
no	Reference		
yes	0.74	0.52-1.06	0.102078
Endocrine therapy			
no	Reference		
yes	0.53	0.37-0.77	0.000888
LMW-E status			
Negative	Reference		
Positive	2.99	1.99-4.5	1.36E-07
Genomic instability			
Low	Reference		
High	0.94	0.65-1.36	0.733382

Table 10. Multivariate cox regression model results using variables selected from univariate cox regression model for Freedom From Recurrence (FFR). The results show that LMW-E positive status is an independent factor that strongly predicts (p=1.36E-07) tumor recurrence and worse clinical outcome in early stage breast cancers.

7.5 Conclusion:

In this chapter, we have successfully established the positive correlation between LMW-E and enhanced genomic instability in early-stage breast cancer. By analyzing a cohort of 725 stage I/II breast tumor samples, our results show positive LMW-E status independently predicts genomic instability, regardless of these clinical and pathological features.

In addition, by comparing the genomic gains and losses between LMW-E-positive and LMW-E-negative subgroups, we find LMW-E positive tumors show higher copy number variation frequency. We have identified that LMW-E–positive tumors harbored significantly higher CN gains in genes involved in the DNA pre-replication complex (such as *CDC6*, *MCM*s, *POLE3*, *POLE4*) and DNA damage repair (such as *RAD51B*, *RAD51AP1*, *RAD52*, *RAD54B*, *DCLRE1C*, *DNTT*, *NHEJ1*, *XRCC5*, and *LIG4*). Additionally, gains in oncogenes such as *JUN*, *FOXC1, CCNE2*, *WNT5B*, *CCNE2, FOXM1*, and *ERBB2*, as well as loss of *TP53*, are strongly correlated with LMW-E–positive tumors.

Lastly, we find LMW-E positive status is significantly associated with tumor recurrence regardless of breast tumor subtypes and other clinical features. Our result indicate that LMW-E could be used as an independent predictor for genomic instability and poor clinical outcome in breast cancer patients.

Chapter Eight: Conclusions and future directions

8.1 Major findings

In this study, we have established a previously unknown function of the 44kDa low molecular weight cyclin E (LMW-E), a tumor driving protein initially identified by our lab in breast cancer patients and cell lines[53,39] and also found in other types of tumors in a variety of studies[149-154]. Although LMW-E is originated from the 50 kDa full-length cyclin E(FL-cycE) by post-translational cleavage mediated by elastase family of serine proteases, we have found LMW-E promote genomic instability by a mechanism distinct from its precursor FL-cycE. Importantly, we have determined the role of LMW-E is independent of FL-cycE (in the endogenous cyclin E knock-out cellular model), and is dominant over FL-cycE (in the LMW-E and FL-cycE co-expressing models). In this chapter, we will summarize the phenotypic similarities and differences induced by LMW-E or FL-cycE in human mammary epithelial cell lines; the over-lapping and distinct biological pathways regulated by LMW-E or FL-cycE; and the mechanisms by which LMW-E is likely to acquire its gain-of-function phenotype, diverse from FL-cycE, to facilitate cell viability under replication stress conditions. Next, we will illustrate the essential down-stream factors required for LMW-E mediated replication stress tolerance, and the specific small molecule inhibitors effectively targeting LMW-E overexpressing cells. Lastly, we will present the clinical significance of LMW-E as an independent bio-marker to predict genomic instability and tumor recurrence in early stage human breast cancer patients.

We generated the *CCNE1* knock-out human mammary epithelial cell (hMEC) models using 76NE6 (p53 deficient) and 76NF2V (p53 proficient) cell lines, and subsequently established doxycycline inducible empty-vector, FL-cycE (aa1-410) or LMW-E (aa40-410) lines on the EKO background. We find that overexpression of FL-cycE but not LMW-E inhibit cell growth and cell viability. Both FL-cycE and LMW-E can slow down DNA replication fork

speed and induce DNA damage at early time points (12-24 hours post induction). However, FL-cycE overexpression further stalls replication forks with accumulated DNA damage in later time point (36-48 hours post induction), leading to cell cycle arrest and chromosome breakages. LMW-E overexpressing cells are able to proceed DNA replication and proliferate with relatively mild DNA damage, resulting in less chromosome breakages, more fusions and increased nuclear abnormalities including micro-nuclear and multi-nuclear phenotypes in the 36-48 hour later time points. These results support the hypothesis that LMW-E overexpression promotes genomic instability in human mammary epithelial cells without the need of endogenous FL-cycE.

Through our RNA sequencing analysis, we compared the transcriptional alterations induced by FL-cycE versus LMW-E in the first 36 hours following induction of each transgene. Results show the genes associated with cell cycle and E2F targets pathway are enriched in both FL-cycE and LMW-E overexpressing cells, consistent to the observation that FL-cycE and LMW-E similarly promotes G1/S transition[80]. However, LMW-E overexpressing cells showed increased enrichment of gene sets associated with multiple types of DNA damage repair pathways, such as mismatch repair, nucleotide excision repair, base excision repair. Furthermore, LMW-E but not FL-cycE result in upregulation of genes involved in DNA replication, particularly the DNA polymerases and components of pre-replication complex and its regulators. These results suggest that the LMW-E deregulation of DNA replication and DNA damage repair are potential mechanisms for the phenotypical differences in LMW-E and FL-cycE overexpressing hMECs.

By cell fractionation assay, we showed that LMW-E facilitates (and FL-cycE inhibits) pre-replication complex (pre-RC) assembly, also termed replication licensing (Figure 76). By examining the effect on DNA damage marker and damage repair reporters, we conclude that LMW-E but not FL-cycE promotes DNA damage repair. From the list of genes upregulated by

LMW-E revealed by RNA-sequencing, we selected CDC6, RAD51 and C17orf53 and confirmed that LMW-E promotes their expression by western blot. We also knocked down each of these factors individually to examine if their depletion can turn LMW-E phenotype to that of Fl-cycE. The results show depletion of CDC6, RAD51 or C17orf53 in LMW-E overexpressing hMECs, lead to increased DNA damage and reduced cell viability. We also find that LMW-E, but not FL-cycE, strongly binds to CDC6. LMW-E is then recruited to the chromatin, and promotes MCMs loading to DNA. Knock-down of CDC6 blocks the chromatin loading of LMW-E and LMW-E mediated MCMs loading. These data collectively support the hypothesis that LMW-E, but not FL-cycE, facilitates replication licensing and DNA damage repair. CDC6, RAD51, and C17orf53 serve as essential down-stream effectors in these processes and are required for the viability of LMW-E overexpressing cells, such that when any of them is depleted, LMW-E cells respond to replicative stress by cell death, similar to those observed with FL-cycE cells.

The essential role of RAD51 as well as the increased CHK1 phosphorylation provide the rationale to target ATR-CHK1-RAD51 replication stress pathway in LMW-E overexpressing cells. We used RAD51 inhibitor B02, CHK1 inhibitor rabusertib and ATR inhibitor ceralasertib, and evaluated their effect on the DNA damage and cell viability in hMECs (inducible LMW-E) and breast cancer cell line (LMW-E stable overexpression). Our results show that LMW-E overexpression significantly decreased the half maximal inhibitory concentrations (IC50s) of these inhibitors. Treatment of these drugs at IC50 concentration led to significantly enhanced DNA damage in LMW-E overexpressing cells. Among the drugs tested, CHK1 inhibitor rabusertib was the most effective. RAD51 inhibitor B02 specifically induced apoptosis in LMW-E overexpressing cells.

Lastly, we used a retrospective cohort of 725 stage I/II breast cancer patients with previously analyzed LMW-E status (positive versus negative) and high-density molecular

165

inversion probe (MIP) data to calculate copy number (CN) gains and losses as a function of LMW-E expression. We find LMW-E-positive subgroup show significantly higher CN gains and losses compared with LMW-E-negative subgroup. We applied the genomic instability index measured by the overall level of CN variations and the number of altered genomic regions, and found the positive correlation between LMW-E and enhanced genomic instability in early stage breast cancer (Figure 76). We used univariate logistic regression model to determine LMW-E status as well as other clinical and pathological features associated with genomic instability, and further determined that LMW-E independently predicts genomic instability by multivariate analysis.

In addition, by comparing the genomic gains and losses between LMW-E-positive and LMW-E-negative subgroups, we find LMW-E–positive tumors harbored significantly higher CN gains in oncogenes (such as *JUN*, *FOXC1, CCNE2*, *WNT5B*, *CCNE2, FOXM1*, and *ERBB2*) and genes involved in the DNA pre-replication complex (such as *CDC6*, *MCM*s, *POLE3*, *POLE4*) and DNA damage repair (such as *RAD51B*, *RAD51AP1*, *RAD52*, *RAD54B*, *DCLRE1C, DNTT, NHEJ1, XRCC5*, and *LIG4*). Loss of tumor suppressor *TP53* is strongly correlated with LMW-E–positive tumors. This is in line with the data showing LMW-E positive status significantly associates with tumor recurrence regardless of breast tumor subtypes and other clinical features. This result also indicate that LMW-E could be used as an independent predictor for genomic instability and poor clinical outcome in breast cancer patients.
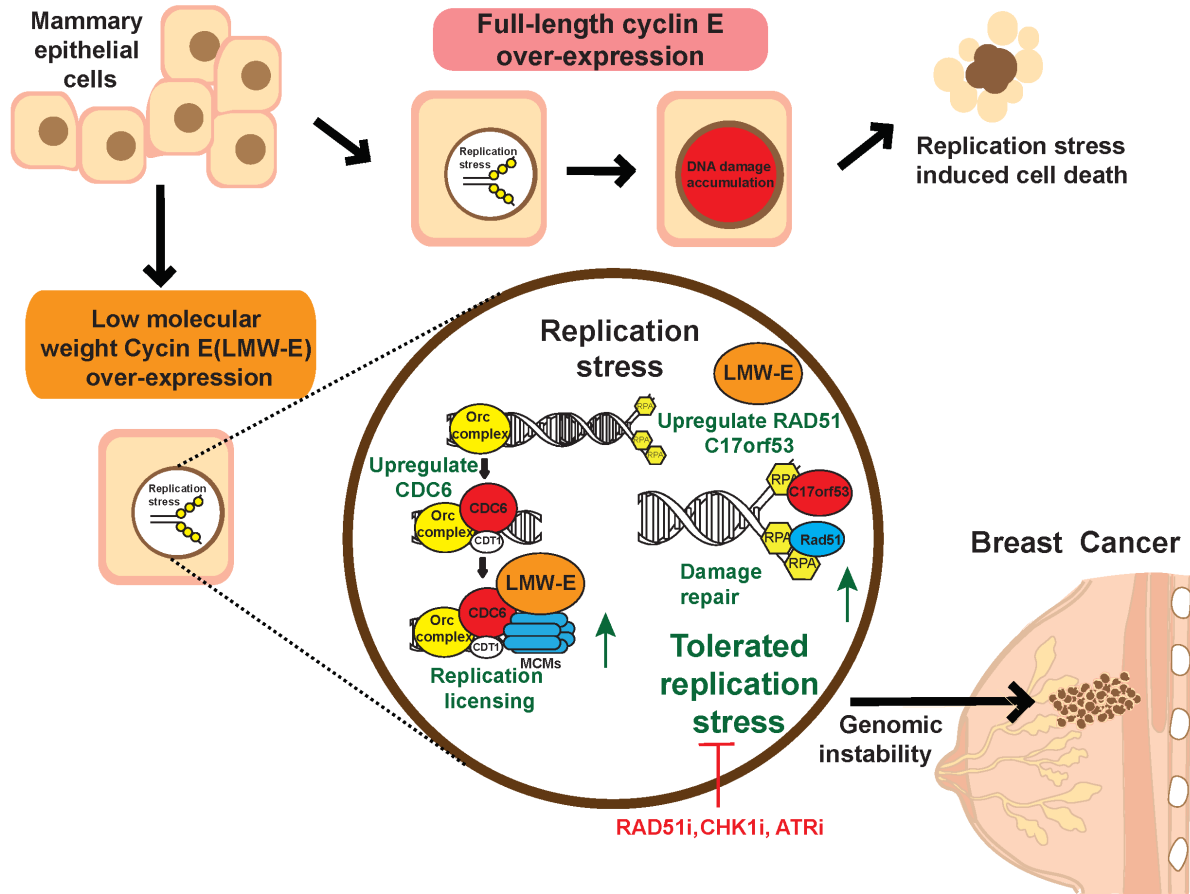
Figure 76. Summary of major findings. LMW-E overexpression promotes genomic instability in human mammary epithelial cells and associates with genomic instability in early stage breast tumors. FL-cycE overexpression leads to accumulation of replication stress, DNA damage, and ultimately cell death. Mechanistically, LMW-E upregulates the expression of CDC6, RAD51, and C17orf53. LMW-E, but not FL-cycE, facilitates replication stress tolerance by promoting pre-replication complex assembly in a CDC6 dependent manner and DNA damage repair in a RAD51- and C17orf53-dependent manner. Targeting the RAD51 or ATR-CHK1 pathway with small molecule inhibitors significantly decreased viability of LMW-E–overexpressing cells. In breast tumor samples, positive LMW-E status independently predicts genomic instability and tumor recurrence.

8.2 Future directions

8.2.1 Comprehensive characterization of genetic alterations in LMW-E driven breast tumor development

Overexpression of LMW-E drives tumor transformation of mammary epithelial cells, and causes spontaneous mammary tumor in mice[65,155]. Current study has associated the oncogenic role of LMW-E with replication stress and genomic instability. An emerging hypothesis from these results suggest certain genetic alterations might be associated with LMW-E over-expression. Indeed, our MIP based copy number variation analysis have revealed that LMW-E–positive tumors harbored significantly higher CN gains in oncogenes such as *JUN*, *FOXC1, CCNE2, WNT5B, CCNE2, FOXM1*, and *ERBB2* and loss of tumor suppressor gene *TP53*. However, our current approach only covers the copy number variations and does not identify mutations in the genome. Moreover, not all the alterations in the cancer genome have equal impact for the cancer initiation and progression. Therefore, comprehensive characterization of genetic alterations is needed to provide deeper understanding for the patterns of oncologic genetic alterations in LMW-E–positive tumors.

To investigate LMW-E-specific genetic alterations associated to breast cancer development, we will apply two different model systems, both can be sub-grouped into LMW-E-positive and LMW-E-negative for bio-informatic comparisons which will be described later in this chapter. The first system is a cohort containing 152 cases of Stage 0 (DCIS) breast cancer samples with matched Stage I/II breast cancer samples. We will continue to collect metastatic samples when available. We have the IHC images of these matched samples to stratify them into LMW-E-positive and LMW-E-negative subgroups, as well as clinical and pathological information to examine other associated features such as age, breast tumor subtypes and treatment.

The second model system is the doxycycline-inducible LMW-E transgenic mouse model. In particular, this mouse model harbors the DNA sequence encoding human LMW-E driven by the tetracycline response element (TRE), and the reverse tetracycline-controlled trans-activator (rtTA) controlled by mouse mammary tumor virus (MMTV) promoter. This allows the mammary tissue specific and doxycycline inducible LMW-E expression, that is shown to drive abnormal mammary gland development, extensive ductal hyperplasia, mammary tumorigenesis and Lung metastases. We also generated and maintained the inducible LMW-E transgenic mice in *Ccne1*-WT versus *Ccne1*-knock-out background, and p53[+/+] versus p53[+/-] background to determine the role of endogenous Cyclin E and P53 tumor suppressor in LMW-E driven breast tumors.

Samples from these models will be grouped by LMW-E status (LMW-E-positive versus LMW-E-negative), and by longitudinal cancer progression status (stage 0, and stage I/II for human breast cancer samples; and mammary gland, ductal hyperplasia, mammary tumor, and statistic lesions for LMW-E transgenic mice models). We will perform whole exome sequencing (WES) and determine the somatic mutations in each of the samples, and we expect to identify the specific altered genes for each LMW-E subgroups and frequently altered genes for each tumor progression status. Mutation frequency, distributions of copy number variation classes (single copy gain, high-level amplification, hemizygous deletion, homozygous deletion) and patterns of co-occurrence and mutual exclusion of genomic aberrations will also be analyzed and compared across the sample sub-groups. The statistical methods and bioinformatics algorithms are provided by Bailey et al. in their investigation for cancer driver genes and mutations across different cancer types[156] and Cha et al. in their study for genetic alterations in metastatic breast cancer across various metastatic sites.

Results from these analyses may reveal the (secondary) genetic alterations that are fueled by LMW-E mediated genomic instability, passively selected during the course of tumor

development. Such genetic alterations may serve as early bio-markers to track disease progression. If they are functionally important for tumor development, these genetic features may also be applied as therapeutic targets for LMW-E positive breast cancers. Additionally, if we find enhanced microsatellite instability (MSI) associated with certain subgroup in our cohort, this may imply potential response to immune checkpoint therapy[157,158].

8.2.2 Functional comparison of LMW-E and FL-cycE for DNA replication complex assembly and activation in a cell free system

By comparing LMW-E and FL-cycE the inducible cellular model deficient in endogenous cyclin E, we generated results providing a better understanding of the distinct roles that FL-cycE and LMW-E isoforms. Our finding that LMW-E, but not FL-cycE, is recruited to chromatin in a CDC6-dependent manner to promote MCMs loading, suggest that LMW-E may directly regulate DNA pre-replication complex (pre-RC) assembly. However, our current experimental system cannot exclude the indirect effect of LMW-E versus FL-cycE for their expression to induce different transcription landscape in the cells. Additionally, the role of LMW-E versus FL-cycE in DNA firing remains unknown. We are also curious to know whether the CDK activity (and which CDK) is required for the process of DNA licensing and/or DNA firing regulated by LMW-E versus FL-cycE.

To test these hypotheses and gain insight into the biochemical mechanism by which LMW-E and FL-cycE regulate DNA replication, we will apply the cell-free system developed by Coverley et al. to study the DNA replication based on G1 nuclei, G1 cytosol and recombinant proteins[97]. We will use the cyclin E knock-out 76NE6-EKO cells as the donor of the nuclei, and prepare nuclei and extracts from the G1 cells, which are collected by cell sorting with FUCCI cell cycle sensor[159].

Initially, we will examine whether LMW-E (and FL-cycE) directly affect pre-RC assembly by incubating recombinant LMW-E (or FL-cycE) with G1 nuclei extract. Second, we

will test the capacity of recombinant LMW-E and FL-cycE to activate DNA replication (DNA firing) in G1 nuclei. Third, we will apply CDK2 depletion versus put-back to investigate whether CDK2 (or other type of CDKs) are required for the pre-RC assembly and/or DNA replication activation. We will use Cyclin A-CDK2, the well-established kinase complex in DNA replication as positive control[97]. Results from these experiments may reveal the detailed mechanisms by which LMW-E (or FL-cycE) regulate DNA replication.

In the following experiments, we will also test the effect of LMW-E on DNA replication when incubated with nuclei collected from other phases of cell cycle. Results from these experiments might shed a light on whether LMW-E overexpression causes un-scheduled licensing in S phase[160], unscheduled DNA synthesis and re-replication in mitosis[160,161], and other types of replication deregulations[162] that are associated with cancer development.

8.2.3 Targeting the G2/M cell cycle check points in LMW-E[high] breast cancer

Our cell cycle analysis suggests both LMW-E and FL-cycE overexpression induce a G2/M accumulation after 48 hours induction. Alterations of G2/M checkpoints were also observed in our RNA-sequencing results from LMW-E overexpression cells, particularly for the overexpression of PKMYT1 and WEE1. These results suggest an intact G2/M checkpoint is required for LMW-E overexpression cells to maintain minimum genomic integrity for cell survival, targeting G2/M checkpoint may specifically target LMW-E overexpressing breast cancer cells by inducing mitotic catastrophe.

We will first confirm if LMW-E induces G2/M cell cycle arrest by synchronizing and releasing the cells from G0/G1, with or without LMW-E overexpression, and track the DNA content and DNA synthesis in a time course manner. During this time course, we will collect cell lysates to analyze the level of cyclin A, cyclin B, p21, p27 and the level of phosphorylated and total CDK2 and CDK1. Results from these experiments will provide cell cycle progression

(or arrest) with correspondent information on level and activity of S phase and G2/M cyclin-CDKs.

Next, we will use the inhibitors for WEE1 and PKMYT1, two protein kinase important for G2/M cell cycle arrest by phosphorylation CDK1 essential for repair of DNA from replicative stress mediated by cyclin E[163,164] . We will perform the experiments in the following groups: inducible hMECs and breast cancer cell lines with or without LMW-E induction; a panel of breast cancer cell lines grouped by LMW-E-negative or LMW-E-positive status, LMW-E[high] breast cancer cell lines before and after *CCNE1* knock-down. We will first evaluate the therapeutic potentials of these drugs by DNA damage assay and cell viability assay in the cell lines, and use in vivo xenograft and transgenic mouse model systems to design most effective treatment strategies targeting LMW-E[high] breast tumors.

Collectively, successful completion of these experiments will be a step toward to develop a biomarker (LMW-E) driven, personalized therapeutic strategy for LMW-E[high] breast tumors patients.

References

1.   The Genetics of Cancer. (National Cancer Institute https://www.cancer.gov/about-cancer/causes-prevention/genetics, 2022).

2.   Haber, D.A. & Fearon, E.R. The promise of cancer genetics. *Lancet* **351 Suppl 2**, SII1-8 (1998).

3.   Heron, M. Deaths: Leading Causes for 2017. *Natl Vital Stat Rep* **68**, 1-77 (2019).

4.   Siegel, R.L., Miller, K.D., Fuchs, H.E. & Jemal, A. Cancer statistics, 2022. *CA Cancer J Clin* **72**, 7-33 (2022).

5.   Croswell, J.M., Ransohoff, D.F. & Kramer, B.S. Principles of cancer screening: lessons from history and study design issues. *Semin Oncol* **37**, 202-215 (2010).

6.   Kerlikowske, K. Epidemiology of ductal carcinoma in situ. *J Natl Cancer Inst Monogr* **2010**, 139-141 (2010).

7.   Sledge, G.W., Mamounas, E.P., Hortobagyi, G.N., Burstein, H.J., Goodwin, P.J. & Wolff, A.C. Past, present, and future challenges in breast cancer treatment. *J Clin Oncol* **32**, 1979-1986 (2014).

8.   Subhedar, P., Olcese, C., Patil, S., Morrow, M. & Van Zee, K.J. Decreasing Recurrence Rates for Ductal Carcinoma In Situ: Analysis of 2996 Women Treated with Breast-Conserving Surgery Over 30 Years. *Ann Surg Oncol* **22**, 3273-3281 (2015).

9.   El Sayed, R., El Jamal, L., El Iskandarani, S., Kort, J., Abdel Salam, M. & Assi, H. Endocrine and Targeted Therapy for Hormone-Receptor-Positive, HER2-Negative Advanced Breast Cancer: Insights to Sequencing Treatment and Overcoming Resistance Based on Clinical Trials. *Front Oncol* **9**, 510 (2019).

10.   Mendes, D., Alves, C., Afonso, N., Cardoso, F., Passos-Coelho, J.L., Costa, L., Andrade, S. & Batel-Marques, F. The benefit of HER2-targeted therapies on overall survival of patients with metastatic HER2-positive breast cancer--a systematic review. *Breast Cancer Res* **17**, 140 (2015).

11. Cortesi, L., Rugo, H.S. & Jackisch, C. An Overview of PARP Inhibitors for the Treatment of Breast Cancer. *Target Oncol* **16**, 255-282 (2021).

12. Easton, D.F., Ford, D. & Bishop, D.T. Breast and ovarian cancer incidence in BRCA1-mutation carriers. Breast Cancer Linkage Consortium. *Am J Hum Genet* **56**, 265-271 (1995).

13. Tutt, A.N.J., Garber, J.E., Kaufman, B., Viale, G., Fumagalli, D., Rastogi, P., Gelber, R.D., de Azambuja, E., Fielding, A., Balmana, J., Domchek, S.M., Gelmon, K.A., Hollingsworth, S.J., Korde, L.A., Linderholm, B., Bandos, H., Senkus, E., Suga, J.M., Shao, Z., Pippas, A.W., Nowecki, Z., Huzarski, T., Ganz, P.A., Lucas, P.C., Baker, N., Loibl, S., McConnell, R., Piccart, M., Schmutzler, R., Steger, G.G., Costantino, J.P., Arahmani, A., Wolmark, N., McFadden, E., Karantza, V., Lakhani, S.R., Yothers, G., Campbell, C., Geyer, C.E., Jr., Olympi, A.C.T.S.C. & Investigators. Adjuvant Olaparib for Patients with BRCA1- or BRCA2-Mutated Breast Cancer. *N Engl J Med* **384**, 2394-2405 (2021).

14. Henry, N.L. & Hayes, D.F. Cancer biomarkers. *Mol Oncol* **6**, 140-146 (2012).

15. Sundaram, M., Guernsey, D.L., Rajaraman, M.M. & Rajaraman, R. Neosis: a novel type of cell division in cancer. *Cancer Biol Ther* **3**, 207-218 (2004).

16. Esmatabadi, M.J., Bakhshinejad, B., Motlagh, F.M., Babashah, S. & Sadeghizadeh, M. Therapeutic resistance and cancer recurrence mechanisms: Unfolding the story of tumour coming back. *J Biosci* **41**, 497-506 (2016).

17. Phi, L.T.H., Sari, I.N., Yang, Y.G., Lee, S.H., Jun, N., Kim, K.S., Lee, Y.K. & Kwon, H.Y. Cancer Stem Cells (CSCs) in Drug Resistance and their Therapeutic Implications in Cancer Treatment. *Stem Cells Int* **2018**, 5416923 (2018).

18. Rueff, J. & Rodrigues, A.S. Cancer Drug Resistance: A Brief Overview from a Genetic Viewpoint. *Methods Mol Biol* **1395**, 1-18 (2016).

19. Hanahan, D. & Weinberg, R.A. Hallmarks of cancer: the next generation. *Cell* **144**, 646-674 (2011).

20. Onitilo, A.A., Engel, J.M., Greenlee, R.T. & Mukesh, B.N. Breast cancer subtypes based on ER/PR and Her2 expression: comparison of clinicopathologic features and survival. *Clin Med Res* **7**, 4-13 (2009).

21. Fakruddin, M., Mannan, K.S., Chowdhury, A., Mazumdar, R.M., Hossain, M.N., Islam, S. & Chowdhury, M.A. Nucleic acid amplification: Alternative methods of polymerase chain reaction. *J Pharm Bioallied Sci* **5**, 245-252 (2013).

22. Thu, K.L., Soria-Bretones, I., Mak, T.W. & Cescon, D.W. Targeting the cell cycle in breast cancer: towards the next phase. *Cell Cycle* **17**, 1871-1885 (2018).

23. Baum, M., Budzar, A.U., Cuzick, J., Forbes, J., Houghton, J.H., Klijn, J.G., Sahmoud, T. & Group, A.T. Anastrozole alone or in combination with tamoxifen versus tamoxifen alone for adjuvant treatment of postmenopausal women with early breast cancer: first results of the ATAC randomised trial. *Lancet* **359**, 2131-2139 (2002).

24. Sunderland, M.C. & Osborne, C.K. Tamoxifen in premenopausal patients with metastatic breast cancer: a review. *J Clin Oncol* **9**, 1283-1297 (1991).

25. Bross, P.F., Cohen, M.H., Williams, G.A. & Pazdur, R. FDA drug approval summaries: fulvestrant. *Oncologist* **7**, 477-480 (2002).

26. Faslodex receives US FDA approval as monotherapy for expanded use in breast cancer. (https://www.astrazeneca.com/media-centre/press-releases/2017/faslodex-receives-us-fda-approval-as-monotherapy-for-expanded-use-in-breast-cancer.html, 2017).

27. Moasser, M.M. The oncogene HER2: its signaling and transforming functions and its role in human cancer pathogenesis. *Oncogene* **26**, 6469-6487 (2007).

28. Goddard, K.A., Weinmann, S., Richert-Boe, K., Chen, C., Bulkley, J. & Wax, C. HER2 evaluation and its impact on breast cancer treatment decisions. *Public Health Genomics* **15**, 1-10 (2012).

29. Kauraniemi, P., Hautaniemi, S., Autio, R., Astola, J., Monni, O., Elkahloun, A. & Kallioniemi, A. Effects of Herceptin treatment on global gene expression patterns in HER2-amplified and nonamplified breast cancer cell lines. *Oncogene* **23**, 1010-1013 (2004).

30. Lohrisch, C. & Piccart, M. An overview of HER2. *Semin Oncol* **28**, 3-11 (2001).

31. Park, K., Han, S., Kim, H.J., Kim, J. & Shin, E. HER2 status in pure ductal carcinoma in situ and in the intraductal and invasive components of invasive ductal carcinoma determined by fluorescence in situ hybridization and immunohistochemistry. *Histopathology* **48**, 702-707 (2006).

32. Ross, J.S., Fletcher, J.A., Linette, G.P., Stec, J., Clark, E., Ayers, M., Symmans, W.F., Pusztai, L. & Bloom, K.J. The Her-2/neu gene and protein in breast cancer 2003: biomarker and target of therapy. *Oncologist* **8**, 307-325 (2003).

33. Gabos, Z., Sinha, R., Hanson, J., Chauhan, N., Hugh, J., Mackey, J.R. & Abdulkarim, B. Prognostic significance of human epidermal growth factor receptor positivity for the development of brain metastasis after newly diagnosed breast cancer. *J Clin Oncol* **24**, 5658-5663 (2006).

34. Hendriks, B.S., Opresko, L.K., Wiley, H.S. & Lauffenburger, D. Coregulation of epidermal growth factor receptor/human epidermal growth factor receptor 2 (HER2) levels and locations: quantitative analysis of HER2 overexpression effects. *Cancer Res* **63**, 1130-1137 (2003).

35. Petricevic, B., Laengle, J., Singer, J., Sachet, M., Fazekas, J., Steger, G., Bartsch, R., Jensen-Jarolim, E. & Bergmann, M. Trastuzumab mediates antibody-dependent cell-

mediated cytotoxicity and phagocytosis to the same extent in both adjuvant and metastatic HER2/neu breast cancer patients. *J Transl Med* **11**, 307 (2013).

36.   Podo, F., Buydens, L.M., Degani, H., Hilhorst, R., Klipp, E., Gribbestad, I.S., Van Huffel, S., van Laarhoven, H.W., Luts, J., Monleon, D., Postma, G.J., Schneiderhan-Marra, N., Santoro, F., Wouters, H., Russnes, H.G., Sorlie, T., Tagliabue, E., Borresen-Dale, A.L. & Consortium, F. Triple-negative breast cancer: present challenges and new perspectives. *Mol Oncol* **4**, 209-229 (2010).

37.   Kumar, P., Mukherjee, M., Johnson, J.P., Patel, M., Huey, B., Albertson, D.G. & Simin, K. Cooperativity of Rb, Brca1, and p53 in malignant breast cancer evolution. *PLoS Genet* **8**, e1003027 (2012).

38.   Chen, X., Low, K.H., Alexander, A., Jiang, Y., Karakas, C., Hess, K.R., Carey, J.P.W., Bui, T.N., Vijayaraghavan, S., Evans, K.W., Yi, M., Ellis, D.C., Cheung, K.L., Ellis, I.O., Fu, S., Meric-Bernstam, F., Hunt, K.K. & Keyomarsi, K. Cyclin E Overexpression Sensitizes Triple-Negative Breast Cancer to Wee1 Kinase Inhibition. *Clin Cancer Res* **24**, 6594-6610 (2018).

39.   Caruso, J.A., Duong, M.T., Carey, J.P.W., Hunt, K.K. & Keyomarsi, K. Low-Molecular-Weight Cyclin E in Human Cancer: Cellular Consequences and Opportunities for Targeted Therapies. *Cancer Res* **78**, 5481-5491 (2018).

40.   Prall, O.W., Rogan, E.M. & Sutherland, R.L. Estrogen regulation of cell cycle progression in breast cancer cells. *J Steroid Biochem Mol Biol* **65**, 169-174 (1998).

41.   Nikolai, B.C., Lanz, R.B., York, B., Dasgupta, S., Mitsiades, N., Creighton, C.J., Tsimelzon, A., Hilsenbeck, S.G., Lonard, D.M., Smith, C.L. & O'Malley, B.W. HER2 Signaling Drives DNA Anabolism and Proliferation through SRC-3 Phosphorylation and E2F1-Regulated Genes. *Cancer Res* **76**, 1463-1475 (2016).

42.   JavanMoghadam, S., Weihua, Z., Hunt, K.K. & Keyomarsi, K. Estrogen receptor alpha is cell cycle-regulated and regulates the cell cycle in a ligand-dependent fashion. *Cell Cycle* **15**, 1579-1590 (2016).

43.   Le, X.F., Lammayot, A., Gold, D., Lu, Y., Mao, W., Chang, T., Patel, A., Mills, G.B. & Bast, R.C., Jr. Genes affecting the cell cycle, growth, maintenance, and drug sensitivity are preferentially regulated by anti-HER2 antibody through phosphatidylinositol 3-kinase-AKT signaling. *J Biol Chem* **280**, 2092-2104 (2005).

44.   Malumbres, M. & Barbacid, M. Cell cycle, CDKs and cancer: a changing paradigm. *Nat Rev Cancer* **9**, 153-166 (2009).

45.   Hwang, H.C. & Clurman, B.E. Cyclin E in normal and neoplastic cell cycles. *Oncogene* **24**, 2776-2786 (2005).

46.   Barnum, K.J. & O'Connell, M.J. Cell cycle regulation by checkpoints. *Methods Mol Biol* **1170**, 29-40 (2014).

47.   Chen, J.T., Ho, C.W., Chi, L.M., Chien, K.Y., Hsieh, Y.J., Lin, S.J. & Yu, J.S. Identification of the lamin A/C phosphoepitope recognized by the antibody P-STM in mitotic HeLa S3 cells. *BMC Biochem* **14**, 18 (2013).

48.   Matthews, H.K., Bertoli, C. & de Bruin, R.A.M. Cell cycle control in cancer. *Nat Rev Mol Cell Biol* **23**, 74-88 (2022).

49.   Chi, Y., Carter, J.H., Swanger, J., Mazin, A.V., Moritz, R.L. & Clurman, B.E. A novel landscape of nuclear human CDK2 substrates revealed by in situ phosphorylation. *Sci Adv* **6**, eaaz9899 (2020).

50.   Zeman, M.K. & Cimprich, K.A. Causes and consequences of replication stress. *Nat Cell Biol* **16**, 2-9 (2014).

51.   Poon, R.Y.C. Cell Cycle Control: A System of Interlinking Oscillators. *Methods Mol Biol* **2329**, 1-18 (2021).

52. Reis-Filho, J.S., Savage, K., Lambros, M.B., James, M., Steele, D., Jones, R.L. & Dowsett, M. Cyclin D1 protein overexpression and CCND1 amplification in breast carcinomas: an immunohistochemical and chromogenic in situ hybridisation analysis. *Mod Pathol* **19**, 999-1009 (2006).

53. Keyomarsi, K., Tucker, S.L., Buchholz, T.A., Callister, M., Ding, Y., Hortobagyi, G.N., Bedrosian, I., Knickerbocker, C., Toyofuku, W., Lowe, M., Herliczek, T.W. & Bacus, S.S. Cyclin E and survival in patients with breast cancer. *N Engl J Med* **347**, 1566-1575 (2002).

54. Al-Kuraya, K., Schraml, P., Torhorst, J., Tapia, C., Zaharieva, B., Novotny, H., Spichtin, H., Maurer, R., Mirlacher, M., Kochli, O., Zuber, M., Dieterich, H., Mross, F., Wilber, K., Simon, R. & Sauter, G. Prognostic relevance of gene amplifications and coamplifications in breast cancer. *Cancer Res* **64**, 8534-8540 (2004).

55. Fantl, V., Stamp, G., Andrews, A., Rosewell, I. & Dickson, C. Mice lacking cyclin D1 are small and show defects in eye and mammary gland development. *Genes Dev* **9**, 2364-2372 (1995).

56. Sicinski, P., Donaher, J.L., Parker, S.B., Li, T., Fazeli, A., Gardner, H., Haslam, S.Z., Bronson, R.T., Elledge, S.J. & Weinberg, R.A. Cyclin D1 provides a link between development and oncogenesis in the retina and breast. *Cell* **82**, 621-630 (1995).

57. Diehl, J.A., Cheng, M., Roussel, M.F. & Sherr, C.J. Glycogen synthase kinase-3beta regulates cyclin D1 proteolysis and subcellular localization. *Genes Dev* **12**, 3499-3511 (1998).

58. Lin, D.I., Barbash, O., Kumar, K.G., Weber, J.D., Harper, J.W., Klein-Szanto, A.J., Rustgi, A., Fuchs, S.Y. & Diehl, J.A. Phosphorylation-dependent ubiquitination of cyclin D1 by the SCF(FBX4-alphaB crystallin) complex. *Mol Cell* **24**, 355-366 (2006).

59. Yu, Q., Geng, Y. & Sicinski, P. Specific protection against breast cancers by cyclin D1 ablation. *Nature* **411**, 1017-1021 (2001).

60.   Hunt, K.K., Karakas, C., Ha, M.J., Biernacka, A., Yi, M., Sahin, A.A., Adjapong, O., Hortobagyi, G.N., Bondy, M., Thompson, P., Cheung, K.L., Ellis, I.O., Bacus, S., Symmans, W.F., Do, K.A. & Keyomarsi, K. Cytoplasmic Cyclin E Predicts Recurrence in Patients with Breast Cancer. *Clin Cancer Res* **23**, 2991-3002 (2017).

61.   Siu, K.T., Rosner, M.R. & Minella, A.C. An integrated view of cyclin E function and regulation. *Cell Cycle* **11**, 57-64 (2012).

62.   Narasimha, A.M., Kaulich, M., Shapiro, G.S., Choi, Y.J., Sicinski, P. & Dowdy, S.F. Cyclin D activates the Rb tumor suppressor by mono-phosphorylation. *Elife* **3**(2014).

63.   Porter, D.C., Zhang, N., Danes, C., McGahren, M.J., Harwell, R.M., Faruki, S. & Keyomarsi, K. Tumor-specific proteolytic processing of cyclin E generates hyperactive lower-molecular-weight forms. *Mol Cell Biol* **21**, 6254-6269 (2001).

64.   Caruso, J.A., Hunt, K.K. & Keyomarsi, K. The neutrophil elastase inhibitor elafin triggers rb-mediated growth arrest and caspase-dependent apoptosis in breast cancer. *Cancer Res* **70**, 7125-7136 (2010).

65.   Akli, S., Van Pelt, C.S., Bui, T., Multani, A.S., Chang, S., Johnson, D., Tucker, S. & Keyomarsi, K. Overexpression of the low molecular weight cyclin E in transgenic mice induces metastatic mammary carcinomas through the disruption of the ARF-p53 pathway. *Cancer Res* **67**, 7212-7222 (2007).

66.   Akli, S., Zheng, P.J., Multani, A.S., Wingate, H.F., Pathak, S., Zhang, N., Tucker, S.L., Chang, S. & Keyomarsi, K. Tumor-specific low molecular weight forms of cyclin E induce genomic instability and resistance to p21, p27, and antiestrogens in breast cancer. *Cancer Res* **64**, 3198-3208 (2004).

67.   Delk, N.A., Hunt, K.K. & Keyomarsi, K. Altered subcellular localization of tumor-specific cyclin E isoforms affects cyclin-dependent kinase 2 complex formation and proteasomal regulation. *Cancer Res* **69**, 2817-2825 (2009).

68. Coffelt, S.B., Wellenstein, M.D. & de Visser, K.E. Neutrophils in cancer: neutral no more. *Nat Rev Cancer* **16**, 431-446 (2016).

69. Gonzalez, H., Hagerling, C. & Werb, Z. Roles of the immune system in cancer: from tumor initiation to metastatic progression. *Genes Dev* **32**, 1267-1284 (2018).

70. Deryugina, E., Carre, A., Ardi, V., Muramatsu, T., Schmidt, J., Pham, C. & Quigley, J.P. Neutrophil Elastase Facilitates Tumor Cell Intravasation and Early Metastatic Events. *iScience* **23**, 101799 (2020).

71. Crusz, S.M. & Balkwill, F.R. Inflammation and cancer: advances and new agents. *Nat Rev Clin Oncol* **12**, 584-596 (2015).

72. Pham, C.T. Neutrophil serine proteases: specific regulators of inflammation. *Nat Rev Immunol* **6**, 541-550 (2006).

73. Doring, G. The role of neutrophil elastase in chronic inflammation. *Am J Respir Crit Care Med* **150**, S114-117 (1994).

74. Pham, C.T. Neutrophil serine proteases fine-tune the inflammatory response. *Int J Biochem Cell Biol* **40**, 1317-1333 (2008).

75. Desmedt, C., Ouriaghli, F.E., Durbecq, V., Soree, A., Colozza, M.A., Azambuja, E., Paesmans, M., Larsimont, D., Buyse, M., Harris, A., Piccart, M., Martiat, P. & Sotiriou, C. Impact of cyclins E, neutrophil elastase and proteinase 3 expression levels on clinical outcome in primary breast cancer patients. *Int J Cancer* **119**, 2539-2545 (2006).

76. Foekens, J.A., Ries, C., Look, M.P., Gippner-Steppert, C., Klijn, J.G. & Jochum, M. The prognostic value of polymorphonuclear leukocyte elastase in patients with primary breast cancer. *Cancer Res* **63**, 337-341 (2003).

77. Yamashita, J., Ogawa, M. & Shirakusa, T. Free-form neutrophil elastase is an independent marker predicting recurrence in primary breast cancer. *J Leukoc Biol* **57**, 375-378 (1995).

78. Sato, T., Takahashi, S., Mizumoto, T., Harao, M., Akizuki, M., Takasugi, M., Fukutomi, T. & Yamashita, J. Neutrophil elastase and cancer. *Surg Oncol* **15**, 217-222 (2006).

79. Lerman, I. & Hammes, S.R. Neutrophil elastase in the tumor microenvironment. *Steroids* **133**, 96-101 (2018).

80. Fu, Z., Thorpe, M., Akula, S., Chahal, G. & Hellman, L.T. Extended Cleavage Specificity of Human Neutrophil Elastase, Human Proteinase 3, and Their Distant Ortholog Clawed Frog PR3-Three Elastases With Similar Primary but Different Extended Specificities and Stability. *Front Immunol* **9**, 2387 (2018).

81. Zhu, Y., Huang, Y., Ji, Q., Fu, S., Gu, J., Tai, N. & Wang, X. Interplay between Extracellular Matrix and Neutrophils in Diseases. *J Immunol Res* **2021**, 8243378 (2021).

82. Walsh, D.E., Greene, C.M., Carroll, T.P., Taggart, C.C., Gallagher, P.M., O'Neill, S.J. & McElvaney, N.G. Interleukin-8 up-regulation by neutrophil elastase is mediated by MyD88/IRAK/TRAF-6 in human bronchial epithelium. *J Biol Chem* **276**, 35494-35499 (2001).

83. Benabid, R., Wartelle, J., Malleret, L., Guyot, N., Gangloff, S., Lebargy, F. & Belaaouaj, A. Neutrophil elastase modulates cytokine expression: contribution to host defense against Pseudomonas aeruginosa-induced pneumonia. *J Biol Chem* **287**, 34883-34894 (2012).

84. Henry, C.M., Sullivan, G.P., Clancy, D.M., Afonina, I.S., Kulms, D. & Martin, S.J. Neutrophil-Derived Proteases Escalate Inflammation through Activation of IL-36 Family Cytokines. *Cell Rep* **14**, 708-722 (2016).

85. Hunter, M.G., Druhan, L.J., Massullo, P.R. & Avalos, B.R. Proteolytic cleavage of granulocyte colony-stimulating factor and its receptor by neutrophil elastase induces growth inhibition and decreased cell surface expression of the granulocyte colony-stimulating factor receptor. *Am J Hematol* **74**, 149-155 (2003).

86.   Kurtagic, E., Jedrychowski, M.P. & Nugent, M.A. Neutrophil elastase cleaves VEGF to generate a VEGF fragment with altered activity. *Am J Physiol Lung Cell Mol Physiol* **296**, L534-546 (2009).

87.   Houghton, A.M., Rzymkiewicz, D.M., Ji, H., Gregory, A.D., Egea, E.E., Metz, H.E., Stolz, D.B., Land, S.R., Marconcini, L.A., Kliment, C.R., Jenkins, K.M., Beaulieu, K.A., Mouded, M., Frank, S.J., Wong, K.K. & Shapiro, S.D. Neutrophil elastase-mediated degradation of IRS-1 accelerates lung tumor growth. *Nat Med* **16**, 219-223 (2010).

88.   Lane, A.A. & Ley, T.J. Neutrophil elastase is important for PML-retinoic acid receptor alpha activities in early myeloid cells. *Mol Cell Biol* **25**, 23-33 (2005).

89.   Yu, L., Zhong, L., Xiong, L., Dan, W., Li, J., Ye, J., Wan, P., Luo, X., Chu, X., Liu, C., He, C., Mu, F. & Liu, B. Neutrophil elastase-mediated proteolysis of the tumor suppressor p200 CUX1 promotes cell proliferation and inhibits cell differentiation in APL. *Life Sci* **242**, 117229 (2020).

90.   Sun, Z. & Yang, P. Role of imbalance between neutrophil elastase and alpha 1-antitrypsin in cancer development and progression. *Lancet Oncol* **5**, 182-190 (2004).

91.   Rosenberg, S., Barr, P.J., Najarian, R.C. & Hallewell, R.A. Synthesis in yeast of a functional oxidation-resistant mutant of human alpha-antitrypsin. *Nature* **312**, 77-80 (1984).

92.   Janciauskiene, S., Wrenger, S., Immenschuh, S., Olejnicka, B., Greulich, T., Welte, T. & Chorostowska-Wynimko, J. The Multifaceted Effects of Alpha1-Antitrypsin on Neutrophil Functions. *Front Pharmacol* **9**, 341 (2018).

93.   Yokota, T., Bui, T., Liu, Y., Yi, M., Hunt, K.K. & Keyomarsi, K. Differential regulation of elafin in normal and tumor-derived mammary epithelial cells is mediated by CCAAT/enhancer binding protein beta. *Cancer Res* **67**, 11272-11283 (2007).

94.   Shaw, L. & Wiedow, O. Therapeutic potential of human elafin. *Biochem Soc Trans* **39**, 1450-1454 (2011).

95.   Labidi-Galy, S.I., Clauss, A., Ng, V., Duraisamy, S., Elias, K.M., Piao, H.Y., Bilal, E., Davidowitz, R.A., Lu, Y., Badalian-Very, G., Gyorffy, B., Kang, U.B., Ficarro, S., Ganesan, S., Mills, G.B., Marto, J.A. & Drapkin, R. Elafin drives poor outcome in high-grade serous ovarian cancers and basal-like breast tumors. *Oncogene* **34**, 373-383 (2015).

96.   Moreau, T., Baranger, K., Dade, S., Dallet-Choisy, S., Guyot, N. & Zani, M.L. Multifaceted roles of human elafin and secretory leukocyte proteinase inhibitor (SLPI), two serine protease inhibitors of the chelonianin family. *Biochimie* **90**, 284-295 (2008).

97.   Hunt, K.K., Wingate, H., Yokota, T., Liu, Y., Mills, G.B., Zhang, F., Fang, B., Su, C.H., Zhang, M., Yi, M. & Keyomarsi, K. Elafin, an inhibitor of elastase, is a prognostic indicator in breast cancer. *Breast Cancer Res* **15**, R3 (2013).

98.   Caruso, J.A., Karakas, C., Zhang, J., Yi, M., Albarracin, C., Sahin, A., Bondy, M., Liu, J., Hunt, K.K. & Keyomarsi, K. Elafin is downregulated during breast and ovarian tumorigenesis but its residual expression predicts recurrence. *Breast Cancer Res* **16**, 3417 (2014).

99.   Caruso, J.A., Akli, S., Pageon, L., Hunt, K.K. & Keyomarsi, K. The serine protease inhibitor elafin maintains normal growth control by opposing the mitogenic effects of neutrophil elastase. *Oncogene* **34**, 3556-3567 (2015).

100.   Kwei, K.A., Kung, Y., Salari, K., Holcomb, I.N. & Pollack, J.R. Genomic instability in breast cancer: pathogenesis and clinical implications. *Mol Oncol* **4**, 255-266 (2010).

101.   Wilhelm, T., Said, M. & Naim, V. DNA Replication Stress and Chromosomal Instability: Dangerous Liaisons. *Genes (Basel)* **11**(2020).

102.   Lee, E.Y. & Muller, W.J. Oncogenes and tumor suppressor genes. *Cold Spring Harb Perspect Biol* **2**, a003236 (2010).

103.   Chu, C., Geng, Y., Zhou, Y. & Sicinski, P. Cyclin E in normal physiology and disease states. *Trends Cell Biol* **31**, 732-746 (2021).

104. Loeb, K.R., Kostner, H., Firpo, E., Norwood, T., K, D.T., Clurman, B.E. & Roberts, J.M. A mouse model for cyclin E-dependent genetic instability and tumorigenesis. *Cancer Cell* **8**, 35-47 (2005).

105. Ma, Y., Fiering, S., Black, C., Liu, X., Yuan, Z., Memoli, V.A., Robbins, D.J., Bentley, H.A., Tsongalis, G.J., Demidenko, E., Freemantle, S.J. & Dmitrovsky, E. Transgenic cyclin E triggers dysplasia and multiple pulmonary adenocarcinomas. *Proc Natl Acad Sci U S A* **104**, 4089-4094 (2007).

106. Minella, A.C., Swanger, J., Bryant, E., Welcker, M., Hwang, H. & Clurman, B.E. p53 and p21 form an inducible barrier that protects cells against cyclin E-cdk2 deregulation. *Curr Biol* **12**, 1817-1827 (2002).

107. Siu, K.T., Xu, Y., Swartz, K.L., Bhattacharyya, M., Gurbuxani, S., Hua, Y. & Minella, A.C. Chromosome instability underlies hematopoietic stem cell dysfunction and lymphoid neoplasia associated with impaired Fbw7-mediated cyclin E regulation. *Mol Cell Biol* **34**, 3244-3258 (2014).

108. Minella, A.C., Grim, J.E., Welcker, M. & Clurman, B.E. p53 and SCFFbw7 cooperatively restrain cyclin E-associated genome instability. *Oncogene* **26**, 6948-6953 (2007).

109. Matsumoto, Y., Hayashi, K. & Nishida, E. Cyclin-dependent kinase 2 (Cdk2) is required for centrosome duplication in mammalian cells. *Curr Biol* **9**, 429-432 (1999).

110. Matsumoto, Y. & Maller, J.L. A centrosomal localization signal in cyclin E required for Cdk2-independent S phase entry. *Science* **306**, 885-888 (2004).

111. Tokuyama, Y., Horn, H.F., Kawamura, K., Tarapore, P. & Fukasawa, K. Specific phosphorylation of nucleophosmin on Thr(199) by cyclin-dependent kinase 2-cyclin E and its role in centrosome duplication. *J Biol Chem* **276**, 21529-21537 (2001).

112. Bagheri-Yarmand, R., Biernacka, A., Hunt, K.K. & Keyomarsi, K. Low molecular weight cyclin E overexpression shortens mitosis, leading to chromosome missegregation and centrosome amplification. *Cancer Res* **70**, 5074-5084 (2010).

113. Bagheri-Yarmand, R., Nanos-Webb, A., Biernacka, A., Bui, T. & Keyomarsi, K. Cyclin E deregulation impairs mitotic progression through premature activation of Cdc25C. *Cancer Res* **70**, 5085-5095 (2010).

114. Oakes, V., Wang, W., Harrington, B., Lee, W.J., Beamish, H., Chia, K.M., Pinder, A., Goto, H., Inagaki, M., Pavey, S. & Gabrielli, B. Cyclin A/Cdk2 regulates Cdh1 and claspin during late S/G2 phase of the cell cycle. *Cell Cycle* **13**, 3302-3311 (2014).

115. Bailey, M.L., Singh, T., Mero, P., Moffat, J. & Hieter, P. Dependence of Human Colorectal Cells Lacking the FBW7 Tumor Suppressor on the Spindle Assembly Checkpoint. *Genetics* **201**, 885-895 (2015).

116. Hughes, B.T., Sidorova, J., Swanger, J., Monnat, R.J., Jr. & Clurman, B.E. Essential role for Cdk2 inhibitory phosphorylation during replication stress revealed by a human Cdk2 knockin mutation. *Proc Natl Acad Sci U S A* **110**, 8954-8959 (2013).

117. Teixeira, L.K., Wang, X., Li, Y., Ekholm-Reed, S., Wu, X., Wang, P. & Reed, S.I. Cyclin E deregulation promotes loss of specific genomic regions. *Curr Biol* **25**, 1327-1333 (2015).

118. Durkin, S.G. & Glover, T.W. Chromosome fragile sites. *Annu Rev Genet* **41**, 169-192 (2007).

119. Bartkova, J., Rezaei, N., Liontos, M., Karakaidos, P., Kletsas, D., Issaeva, N., Vassiliou, L.V., Kolettas, E., Niforou, K., Zoumpourlis, V.C., Takaoka, M., Nakagawa, H., Tort, F., Fugger, K., Johansson, F., Sehested, M., Andersen, C.L., Dyrskjot, L., Orntoft, T., Lukas, J., Kittas, C., Helleday, T., Halazonetis, T.D., Bartek, J. & Gorgoulis, V.G. Oncogene-induced senescence is part of the tumorigenesis barrier imposed by DNA damage checkpoints. *Nature* **444**, 633-637 (2006).

120. Gaillard, H., Garcia-Muse, T. & Aguilera, A. Replication stress and cancer. *Nat Rev Cancer* **15**, 276-289 (2015).

121. Di Micco, R., Fumagalli, M., Cicalese, A., Piccinin, S., Gasparini, P., Luise, C., Schurra, C., Garre, M., Nuciforo, P.G., Bensimon, A., Maestro, R., Pelicci, P.G. & d'Adda di Fagagna, F. Oncogene-induced senescence is a DNA damage response triggered by DNA hyper-replication. *Nature* **444**, 638-642 (2006).

122. Zellweger, R., Dalcher, D., Mutreja, K., Berti, M., Schmid, J.A., Herrador, R., Vindigni, A. & Lopes, M. Rad51-mediated replication fork reversal is a global response to genotoxic treatments in human cells. *J Cell Biol* **208**, 563-579 (2015).

123. Burrell, R.A., McClelland, S.E., Endesfelder, D., Groth, P., Weller, M.C., Shaikh, N., Domingo, E., Kanu, N., Dewhurst, S.M., Gronroos, E., Chew, S.K., Rowan, A.J., Schenk, A., Sheffer, M., Howell, M., Kschischo, M., Behrens, A., Helleday, T., Bartek, J., Tomlinson, I.P. & Swanton, C. Replication stress links structural and numerical cancer chromosomal instability. *Nature* **494**, 492-496 (2013).

124. Negrini, S., Gorgoulis, V.G. & Halazonetis, T.D. Genomic instability--an evolving hallmark of cancer. *Nat Rev Mol Cell Biol* **11**, 220-228 (2010).

125. Fragkos, M., Ganier, O., Coulombe, P. & Mechali, M. DNA replication origin activation in space and time. *Nat Rev Mol Cell Biol* **16**, 360-374 (2015).

126. Kunkel, T.A. & Bebenek, K. DNA replication fidelity. *Annu Rev Biochem* **69**, 497-529 (2000).

127. Blow, J.J. & Dutta, A. Preventing re-replication of chromosomal DNA. *Nat Rev Mol Cell Biol* **6**, 476-486 (2005).

128. Deng, L., Wu, R.A., Sonneville, R., Kochenova, O.V., Labib, K., Pellman, D. & Walter, J.C. Mitotic CDK Promotes Replisome Disassembly, Fork Breakage, and Complex DNA Rearrangements. *Mol Cell* **73**, 915-929 e916 (2019).

129. Coverley, D., Laman, H. & Laskey, R.A. Distinct roles for cyclins E and A during DNA replication complex assembly and activation. *Nat Cell Biol* **4**, 523-528 (2002).

130. Ge, X.Q., Jackson, D.A. & Blow, J.J. Dormant origins licensed by excess Mcm2-7 are required for human cells to survive replicative stress. *Genes Dev* **21**, 3331-3341 (2007).

131. Lengronne, A. & Schwob, E. The yeast CDK inhibitor Sic1 prevents genomic instability by promoting replication origin licensing in late G(1). *Mol Cell* **9**, 1067-1078 (2002).

132. Bester, A.C., Roniger, M., Oren, Y.S., Im, M.M., Sarni, D., Chaoat, M., Bensimon, A., Zamir, G., Shewach, D.S. & Kerem, B. Nucleotide deficiency promotes genomic instability in early stages of cancer development. *Cell* **145**, 435-446 (2011).

133. Ohtsubo, M., Theodoras, A.M., Schumacher, J., Roberts, J.M. & Pagano, M. Human cyclin E, a nuclear protein essential for the G1-to-S phase transition. *Mol Cell Biol* **15**, 2612-2624 (1995).

134. Jones, R.M., Mortusewicz, O., Afzal, I., Lorvellec, M., Garcia, P., Helleday, T. & Petermann, E. Increased replication initiation and conflicts with transcription underlie Cyclin E-induced replication stress. *Oncogene* **32**, 3744-3753 (2013).

135. Kotsantis, P., Petermann, E. & Boulton, S.J. Mechanisms of Oncogene-Induced Replication Stress: Jigsaw Falling into Place. *Cancer Discov* **8**, 537-555 (2018).

136. Geng, Y., Lee, Y.M., Welcker, M., Swanger, J., Zagozdzon, A., Winer, J.D., Roberts, J.M., Kaldis, P., Clurman, B.E. & Sicinski, P. Kinase-independent function of cyclin E. *Mol Cell* **25**, 127-139 (2007).

137. Kastan, M.B. & Bartek, J. Cell-cycle checkpoints and cancer. *Nature* **432**, 316-323 (2004).

138. Ou, Y.H., Chung, P.H., Sun, T.P. & Shieh, S.Y. p53 C-terminal phosphorylation by CHK1 and CHK2 participates in the regulation of DNA-damage-induced C-terminal acetylation. *Mol Biol Cell* **16**, 1684-1695 (2005).

139. Bartek, J. & Lukas, J. Chk1 and Chk2 kinases in checkpoint control and cancer. *Cancer Cell* **3**, 421-429 (2003).

140. Hughes, J., Brown, P. & Shankland, S.J. Cyclin kinase inhibitor p21CIP1/WAF1 limits interstitial cell proliferation following ureteric obstruction. *Am J Physiol* **277**, F948-956 (1999).

141. Jackson, P.K., Chevalier, S., Philippe, M. & Kirschner, M.W. Early events in DNA replication require cyclin E and are blocked by p21CIP1. *J Cell Biol* **130**, 755-769 (1995).

142. Smith, J., Tho, L.M., Xu, N. & Gillespie, D.A. The ATM-Chk2 and ATR-Chk1 pathways in DNA damage signaling and cancer. *Adv Cancer Res* **108**, 73-112 (2010).

143. O'Connell, M.J., Raleigh, J.M., Verkade, H.M. & Nurse, P. Chk1 is a wee1 kinase in the G2 DNA damage checkpoint inhibiting cdc2 by Y15 phosphorylation. *EMBO J* **16**, 545-554 (1997).

144. Sorensen, C.S. & Syljuasen, R.G. Safeguarding genome integrity: the checkpoint kinases ATR, CHK1 and WEE1 restrain CDK activity during normal DNA replication. *Nucleic Acids Res* **40**, 477-486 (2012).

145. Nakanishi, M., Shimada, M. & Niida, H. Genetic instability in cancer cells by impaired cell cycle checkpoints. *Cancer Sci* **97**, 984-989 (2006).

146. Mason, J.M., Chan, Y.L., Weichselbaum, R.W. & Bishop, D.K. Non-enzymatic roles of human RAD51 at stalled replication forks. *Nat Commun* **10**, 4410 (2019).

147. Petermann, E., Orta, M.L., Issaeva, N., Schultz, N. & Helleday, T. Hydroxyurea-stalled replication forks become progressively inactivated and require two different RAD51-mediated pathways for restart and repair. *Mol Cell* **37**, 492-502 (2010).

148. Baumann, P. & West, S.C. Role of the human RAD51 protein in homologous recombination and double-stranded-break repair. *Trends Biochem Sci* **23**, 247-251 (1998).

149. Hashimoto, Y., Ray Chaudhuri, A., Lopes, M. & Costanzo, V. Rad51 protects nascent DNA from Mre11-dependent degradation and promotes continuous DNA synthesis. *Nat Struct Mol Biol* **17**, 1305-1311 (2010).

150. Chen, X., Yang, D., Carey, J.P.W., Karakas, C., Albarracin, C., Sahin, A.A., Arun, B.K., Guray Durak, M., Li, M., Kohansal, M., Bui, T.N., Ha, M.J., Hunt, K.K. & Keyomarsi, K. Targeting Replicative Stress and DNA Repair by Combining PARP and Wee1 Kinase Inhibitors Is Synergistic in Triple Negative Breast Cancers with Cyclin E or BRCA1 Alteration. *Cancers (Basel)* **13**(2021).

151. Murai, J. Targeting DNA repair and replication stress in the treatment of ovarian cancer. *Int J Clin Oncol* **22**, 619-628 (2017).

152. Barnieh, F.M., Loadman, P.M. & Falconer, R.A. Progress towards a clinically-successful ATR inhibitor for cancer therapy. *Curr Res Pharmacol Drug Discov* **2**, 100017 (2021).

153. Jo, U., Senatorov, I.S., Zimmermann, A., Saha, L.K., Murai, Y., Kim, S.H., Rajapakse, V.N., Elloumi, F., Takahashi, N., Schultz, C.W., Thomas, A., Zenke, F.T. & Pommier, Y. Novel and Highly Potent ATR Inhibitor M4344 Kills Cancer Cells With Replication Stress, and Enhances the Chemotherapeutic Activity of Widely Used DNA Damaging Agents. *Mol Cancer Ther* **20**, 1431-1441 (2021).

154. Qiu, Z., Oleinick, N.L. & Zhang, J. ATR/CHK1 inhibitors and cancer therapy. *Radiother Oncol* **126**, 450-464 (2018).

155. Otto, T. & Sicinski, P. Cell cycle proteins as promising targets in cancer therapy. *Nat Rev Cancer* **17**, 93-115 (2017).

156. Guzi, T.J., Paruch, K., Dwyer, M.P., Labroli, M., Shanahan, F., Davis, N., Taricani, L., Wiswell, D., Seghezzi, W., Penaflor, E., Bhagwat, B., Wang, W., Gu, D., Hsieh, Y., Lee, S., Liu, M. & Parry, D. Targeting the replication checkpoint using SCH 900776, a

potent and functionally selective CHK1 inhibitor identified via high content screening. *Mol Cancer Ther* **10**, 591-602 (2011).

157. Daud, A.I., Ashworth, M.T., Strosberg, J., Goldman, J.W., Mendelson, D., Springett, G., Venook, A.P., Loechner, S., Rosen, L.S., Shanahan, F., Parry, D., Shumway, S., Grabowsky, J.A., Freshwater, T., Sorge, C., Kang, S.P., Isaacs, R. & Munster, P.N. Phase I dose-escalation trial of checkpoint kinase 1 inhibitor MK-8776 as monotherapy and in combination with gemcitabine in patients with advanced solid tumors. *J Clin Oncol* **33**, 1060-1066 (2015).

158. Karp, J.E., Thomas, B.M., Greer, J.M., Sorge, C., Gore, S.D., Pratz, K.W., Smith, B.D., Flatten, K.S., Peterson, K., Schneider, P., Mackey, K., Freshwater, T., Levis, M.J., McDevitt, M.A., Carraway, H.E., Gladstone, D.E., Showel, M.M., Loechner, S., Parry, D.A., Horowitz, J.A., Isaacs, R. & Kaufmann, S.H. Phase I and pharmacologic trial of cytosine arabinoside with the selective checkpoint 1 inhibitor Sch 900776 in refractory acute leukemias. *Clin Cancer Res* **18**, 6723-6731 (2012).

159. King, C., Diaz, H.B., McNeely, S., Barnard, D., Dempsey, J., Blosser, W., Beckmann, R., Barda, D. & Marshall, M.S. LY2606368 Causes Replication Catastrophe and Antitumor Effects through CHK1-Dependent Mechanisms. *Mol Cancer Ther* **14**, 2004-2013 (2015).

160. Barnard, D., Diaz, H.B., Burke, T., Donoho, G., Beckmann, R., Jones, B., Barda, D., King, C. & Marshall, M. LY2603618, a selective CHK1 inhibitor, enhances the anti-tumor effect of gemcitabine in xenograft tumor models. *Invest New Drugs* **34**, 49-60 (2016).

161. Preliminary Activity Seen with RAD51 Inhibitor. *Cancer Discov* **11**, OF1 (2021).

162. Das, A.T., Tenenbaum, L. & Berkhout, B. Tet-On Systems For Doxycycline-inducible Gene Expression. *Curr Gene Ther* **16**, 156-167 (2016).

163. Band, V. & Sager, R. Distinctive traits of normal and tumor-derived human mammary epithelial cells expressed in a medium that supports long-term growth of both cell types. *Proc Natl Acad Sci U S A* **86**, 1249-1253 (1989).

164. Wingate, H., Zhang, N., McGarhen, M.J., Bedrosian, I., Harper, J.W. & Keyomarsi, K. The tumor-specific hyperactive forms of cyclin E are resistant to inhibition by p21 and p27. *J Biol Chem* **280**, 15148-15157 (2005).

165. Willis, N. & Rhind, N. Regulation of DNA replication by the S-phase DNA damage checkpoint. *Cell Div* **4**, 13 (2009).

166. Robinson, M.D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* **11**, R25 (2010).

167. Law, C.W., Chen, Y., Shi, W. & Smyth, G.K. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol* **15**, R29 (2014).

168. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *J R Stat Soc B* **57**, 289-300 (1995).

169. Wang, C., Chen, Z., Su, D., Tang, M., Nie, L., Zhang, H., Feng, X., Wang, R., Shen, X., Srivastava, M., McLaughlin, M.E., Hart, T., Li, L. & Chen, J. C17orf53 is identified as a novel gene involved in inter-strand crosslink repair. *DNA Repair (Amst)* **95**, 102946 (2020).

170. Hustedt, N., Saito, Y., Zimmermann, M., Alvarez-Quilon, A., Setiaputra, D., Adam, S., McEwan, A., Yuan, J.Y., Olivieri, M., Zhao, Y., Kanemaki, M.T., Jurisicova, A. & Durocher, D. Control of homologous recombination by the HROB-MCM8-MCM9 pathway. *Genes Dev* **33**, 1397-1415 (2019).

171. Lin, J.Y. & Simmons, D.T. The ability of large T antigen to complex with p53 is necessary for the increased life span and partial transformation of human cells by simian virus 40. *J Virol* **65**, 6447-6453 (1991).

172. Costantino, L., Sotiriou, S.K., Rantala, J.K., Magin, S., Mladenov, E., Helleday, T., Haber, J.E., Iliakis, G., Kallioniemi, O.P. & Halazonetis, T.D. Break-induced replication repair of damaged forks induces genomic duplications in human cells. *Science* **343**, 88-91 (2014).

173. Bartkova, J., Horejsi, Z., Koed, K., Kramer, A., Tort, F., Zieger, K., Guldberg, P., Sehested, M., Nesland, J.M., Lukas, C., Orntoft, T., Lukas, J. & Bartek, J. DNA damage response as a candidate anti-cancer barrier in early human tumorigenesis. *Nature* **434**, 864-870 (2005).

174. Sotiriou, S.K., Kamileri, I., Lugli, N., Evangelou, K., Da-Re, C., Huber, F., Padayachy, L., Tardy, S., Nicati, N.L., Barriot, S., Ochs, F., Lukas, C., Lukas, J., Gorgoulis, V.G., Scapozza, L. & Halazonetis, T.D. Mammalian RAD52 Functions in Break-Induced Replication Repair of Collapsed DNA Replication Forks. *Mol Cell* **64**, 1127-1134 (2016).

175. Mao, Z., Jiang, Y., Liu, X., Seluanov, A. & Gorbunova, V. DNA repair by homologous recombination, but not by nonhomologous end joining, is elevated in breast cancer cells. *Neoplasia* **11**, 683-691 (2009).

176. Boulares, A.H., Yakovlev, A.G., Ivanova, V., Stoica, B.A., Wang, G., Iyer, S. & Smulson, M. Role of poly(ADP-ribose) polymerase (PARP) cleavage in apoptosis. Caspase 3-resistant PARP mutant increases rates of apoptosis in transfected cells. *J Biol Chem* **274**, 22932-22940 (1999).

177. Thompson, P.A., Brewster, A.M., Kim-Anh, D., Baladandayuthapani, V., Broom, B.M., Edgerton, M.E., Hahn, K.M., Murray, J.L., Sahin, A., Tsavachidis, S., Wang, Y., Zhang, L., Hortobagyi, G.N., Mills, G.B. & Bondy, M.L. Selective genomic copy number imbalances and probability of recurrence in early-stage breast cancer. *PLoS One* **6**, e23543 (2011).

178.   Van Loo, P., Nordgard, S.H., Lingjaerde, O.C., Russnes, H.G., Rye, I.H., Sun, W., Weigman, V.J., Marynen, P., Zetterberg, A., Naume, B., Perou, C.M., Borresen-Dale, A.L. & Kristensen, V.N. Allele-specific copy number analysis of tumors. *Proc Natl Acad Sci U S A* **107**, 16910-16915 (2010).

179.   Bonnet, F., Guedj, M., Jones, N., Sfar, S., Brouste, V., Elarouci, N., Banneau, G., Orsetti, B., Primois, C., de Lara, C.T., Debled, M., de Mascarel, I., Theillet, C., Sevenet, N., de Reynies, A., MacGrogan, G. & Longy, M. An array CGH based genomic instability index (G2I) is predictive of clinical outcome in breast cancer and reveals a subset of tumors without lymph node involvement but with poor prognosis. *BMC Med Genomics* **5**, 54 (2012).

180.   Hudis, C.A., Barlow, W.E., Costantino, J.P., Gray, R.J., Pritchard, K.I., Chapman, J.A., Sparano, J.A., Hunsberger, S., Enos, R.A., Gelber, R.D. & Zujewski, J.A. Proposal for standardized definitions for efficacy end points in adjuvant breast cancer trials: the STEEP system. *J Clin Oncol* **25**, 2127-2132 (2007).

181.   Bales, E., Mills, L., Milam, N., McGahren-Murray, M., Bandyopadhyay, D., Chen, D., Reed, J.A., Timchenko, N., van den Oord, J.J., Bar-Eli, M., Keyomarsi, K. & Medrano, E.E. The low molecular weight cyclin E isoforms augment angiogenesis and metastasis of human melanoma cells in vivo. *Cancer Res* **65**, 692-697 (2005).

182.   Bedrosian, I., Lu, K.H., Verschraegen, C. & Keyomarsi, K. Cyclin E deregulation alters the biologic properties of ovarian cancer cells. *Oncogene* **23**, 2648-2657 (2004).

183.   Corin, I., Di Giacomo, M.C., Lastella, P., Bagnulo, R., Guanti, G. & Simone, C. Tumor-specific hyperactive low-molecular-weight cyclin E isoforms detection and characterization in non-metastatic colorectal tumors. *Cancer Biol Ther* **5**, 198-203 (2006).

184.   Davidson, B., Skrede, M., Silins, I., Shih Ie, M., Trope, C.G. & Florenes, V.A. Low-molecular weight forms of cyclin E differentiate ovarian carcinoma from cells of

mesothelial origin and are associated with poor survival in ovarian carcinoma. *Cancer* **110**, 1264-1271 (2007).

185.   Milne, A.N., Carvalho, R., Jansen, M., Kranenbarg, E.K., van de Velde, C.J., Morsink, F.M., Musler, A.R., Weterman, M.A. & Offerhaus, G.J. Cyclin E low molecular weight isoforms occur commonly in early-onset gastric cancer and independently predict survival. *J Clin Pathol* **61**, 311-316 (2008).

186.   Zhou, Y.J., Xie, Y.T., Gu, J., Yan, L., Guan, G.X. & Liu, X. Overexpression of cyclin E isoforms correlates with poor prognosis in rectal cancer. *Eur J Surg Oncol* **37**, 1078-1084 (2011).

187.   Duong, M.T., Akli, S., Wei, C., Wingate, H.F., Liu, W., Lu, Y., Yi, M., Mills, G.B., Hunt, K.K. & Keyomarsi, K. LMW-E/CDK2 deregulates acinar morphogenesis, induces tumorigenesis, and associates with the activated b-Raf-ERK1/2-mTOR pathway in breast cancer patients. *PLoS Genet* **8**, e1002538 (2012).

188.   Bailey, M.H., Tokheim, C., Porta-Pardo, E., Sengupta, S., Bertrand, D., Weerasinghe, A., Colaprico, A., Wendl, M.C., Kim, J., Reardon, B., Ng, P.K., Jeong, K.J., Cao, S., Wang, Z., Gao, J., Gao, Q., Wang, F., Liu, E.M., Mularoni, L., Rubio-Perez, C., Nagarajan, N., Cortes-Ciriano, I., Zhou, D.C., Liang, W.W., Hess, J.M., Yellapantula, V.D., Tamborero, D., Gonzalez-Perez, A., Suphavilai, C., Ko, J.Y., Khurana, E., Park, P.J., Van Allen, E.M., Liang, H., Group, M.C.W., Cancer Genome Atlas Research, N., Lawrence, M.S., Godzik, A., Lopez-Bigas, N., Stuart, J., Wheeler, D., Getz, G., Chen, K., Lazar, A.J., Mills, G.B., Karchin, R. & Ding, L. Comprehensive Characterization of Cancer Driver Genes and Mutations. *Cell* **173**, 371-385 e318 (2018).

189.   Cai, L., Sun, Y., Wang, K., Guan, W., Yue, J., Li, J., Wang, R. & Wang, L. The Better Survival of MSI Subtype Is Associated With the Oxidative Stress Related Pathways in Gastric Cancer. *Front Oncol* **10**, 1269 (2020).

190. Le, D.T., Uram, J.N., Wang, H., Bartlett, B.R., Kemberling, H., Eyring, A.D., Skora, A.D., Luber, B.S., Azad, N.S., Laheru, D., Biedrzycki, B., Donehower, R.C., Zaheer, A., Fisher, G.A., Crocenzi, T.S., Lee, J.J., Duffy, S.M., Goldberg, R.M., de la Chapelle, A., Koshiji, M., Bhaijee, F., Huebner, T., Hruban, R.H., Wood, L.D., Cuka, N., Pardoll, D.M., Papadopoulos, N., Kinzler, K.W., Zhou, S., Cornish, T.C., Taube, J.M., Anders, R.A., Eshleman, J.R., Vogelstein, B. & Diaz, L.A., Jr. PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *N Engl J Med* **372**, 2509-2520 (2015).

191. Zielke, N. & Edgar, B.A. FUCCI sensors: powerful new tools for analysis of cell proliferation. *Wiley Interdiscip Rev Dev Biol* **4**, 469-487 (2015).

192. Neelsen, K.J., Zanini, I.M., Mijic, S., Herrador, R., Zellweger, R., Ray Chaudhuri, A., Creavin, K.D., Blow, J.J. & Lopes, M. Deregulated origin licensing leads to chromosomal breaks by rereplication of a gapped DNA template. *Genes Dev* **27**, 2537-2542 (2013).

193. Fu, H., Redon, C.E., Thakur, B.L., Utani, K., Sebastian, R., Jang, S.M., Gross, J.M., Mosavarpour, S., Marks, A.B., Zhuang, S.Z., Lazar, S.B., Rao, M., Mencer, S.T., Baris, A.M., Pongor, L.S. & Aladjem, M.I. Dynamics of replication origin over-activation. *Nat Commun* **12**, 3448 (2021).

194. Champeris Tsaniras, S., Kanellakis, N., Symeonidou, I.E., Nikolopoulou, P., Lygerou, Z. & Taraviras, S. Licensing of DNA replication, cancer, pluripotency and differentiation: an interlinked world? *Semin Cell Dev Biol* **30**, 174-180 (2014).

195. Lobrich, M. & Jeggo, P.A. The impact of a negligent G2/M checkpoint on genomic instability and cancer induction. *Nat Rev Cancer* **7**, 861-869 (2007).

196. Schmidt, M., Rohe, A., Platzer, C., Najjar, A., Erdmann, F. & Sippl, W. Regulation of G2/M Transition by Inhibition of WEE1 and PKMYT1 Kinases. *Molecules* **22**(2017).

VITA

Mi Li was born in 1985 to Jun Li and Junying Ma. He attended Sichuan University in China, and graduated in 2007 with a Bachelor of Science degree in Bio-technology. He then worked as a technician at the Genetically Engineered Mouse Center of Sichuan University West-China Hospital. He received chemical-biology and molecular-biology training at Shanghai Institute of Biochemistry and Cell Biology, Chinese Academy of Sciences in 2014 and was further trained as a Research Intern at The University of Texas MD Anderson Cancer Center in 2015. He then worked as a Research Assistant at MD Anderson Cancer Center in 2016 and received admission into MD Anderson UThealth graduate school in 2017. He conducted his Ph.D. dissertation research under the supervision of Dr. Khandan Keyomarsi.