ESTABLISHMENT OF CELL LINES WITH INDUCIBLE EXPRESSION OF shRNA FOR AN ESTROGEN RESPONSIVE GENE

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY
BURCU KARAKAYA

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
BIOLOGY

JUNE 2018

Approval of the Thesis:

ESTABLISHMENT OF CELL LINES WITH INDUCIBLE EXPRESSION OF shRNA FOR AN ESTROGEN RESPONSIVE GENE

submitted by **BURCU KARAKAYA** in partial fulfillment of the requirements for the degree of **Master of Science in Biology Department, Middle East Technical University** by,

Prof. Dr. Halil Kalıpçılar _____
Dean, Graduate School of **Natural and Applied Sciences**

Prof. Dr. Orhan Adalı _____
Head of Department, **Biological Sciences**

Prof. Dr. Mesut Muyan _____
Supervisor, **Biological Sciences Dept., METU**

Examining Committee Members:

Assoc. Prof. Dr. A. Elif Erson-Bensan _____
Biological Sciences Dept., METU

Prof. Dr. Mesut Muyan _____
Biological Sciences Dept., METU

Prof. Dr. Sreeparna Banerjee _____
Biological Sciences Dept., METU

Assoc. Prof. Dr. Çağdaş Devrim Son _____
Biological Sciences Dept., METU

Assoc. Prof. Dr. Bala Gür Dedeoğlu _____
Ankara University Biotechnology Institute

Date: 01.06. 2018

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name: Burcu Karakaya

Signature :

ABSTRACT

ESTABLISHMENT OF CELL LINES WITH INDUCIBLE EXPRESSION OF shRNA FOR AN ESTROGEN RESPONSIVE GENE

Karakaya, Burcu

M.S., Department of Biology

Supervisor: Prof. Dr. Mesut Muyan

June 2018, 99 pages

Estrogen hormones, primarily 17β-estradiol (E2) as the primary circulating estrogen, are involved in the homeodynamic regulation of various tissues/organs including mammary gland within which estrogen receptor α (ERα) conveys E2 signaling. The binding of E2 to ERα activates the receptor to regulate estrogen responsive gene expressions.

Previous microarray and gene expression studies of our laboratory indicate that CXXC5 is an estrogen responsive gene regulated by ERα. Our ongoing studies also indicate that CXXC5 as a member of zinc-finger CXXC domain protein family, binds to non-methylated CpG dinucleotides. Methylation of CpG islands found in enhancer and promoter regions is thought to silence gene transcriptions. Binding of this family proteins to non-methylated CpGs is suggested to have a role in transcription by preventing DNA methylation. Although studies on CXXC5 are limited, CXXC5 appears to participate in cellular processes as a transcription factor, co-regulator and/or epigenetic factor.

To address the role of CXXC5 in E2-mediated cellular events, we aimed 1) to identify protein partners and 2) to define cellular function of CXXC5. Since, proteins perform their functions within the context of interacting proteins in a spacio-temporal manner, we initially decided to use a yeast-two-hybrid (Y2H) service for the identification of putative CXXC5 protein partners followed by a mammalian-two-hybrid system (M2H) for verification. However, of the reported interacting partners identified with Y2H, we were not able to verify any protein as the CXXC5 interacting partner with M2H. Based on these results, we decided to use the BioID approach that allows the identification of interacting proteins through the biotin tagging of interactors. Our ongoing studies suggest that BioID can indeed be used for the identification of CXXC5 interactors.

In studies we performed in parallel to Y2H to address the role of CXXC5 in E2-mediated cellular events, we wanted to generate cell lines in which CXXC5 was stably overexpressed or silenced. However, due to cellular death; we failed to generate stable cell lines. Our failure necessitated the use of an inducible expression system to generate cell lines that allow us to conduct functional studies on CXXC5. The system was expected to provide us maintainable cell lines containing on/off switch for CXXC5 expression. For this, we used pINDUCER system, which provides a tetracycline (Tet) inducible expression of target gene. For this purpose, we initially used an shRNA approach to specifically down-regulate CXXC5 synthesis. Although we have succeeded the generation of monoclones, we could not decrease CXXC5 synthesis. Studies for obtaining overexpressed or silenced CXXC5 has been continuing with different cell lines and expression systems.

Keywords: estrogen, estrogen receptor, CXXC5

ÖZ

ÖSTROJEN DUYARLI BİR GENİN İNDÜKLENEBİLİR shRNA İFADESİ OLAN HÜCRE HATLARININ OLUŞTURULMASI

Karakaya, Burcu
Yüksek Lisans, Biyoloji Bölümü
Tez Yöneticisi: Prof. Dr. Mesut Muyan
Haziran 2018, 99 sayfa

Östrojen hormonları, başta 17β-östradiol (E2), dolaşımdaki başlıca östrojen olarak, östrojen reseptör α (ERα)'nın E2 sinyal yolağıyla üreme organları ve meme bezleri dahil çeşitli doku/organların homeodinamik düzenlemesinde yer alır. E2'nin ERα'ya bağlanması, östrojen duyarlı gen ifadelerini düzenlemek için reseptörü aktive eder.

Önceden laboratuvarımızda yürütülmüş olan mikrodizi ve gen ekspresyonu çalışmaları, CXXC5'in ERα tarafından düzenlenen östrojen duyarlı bir gen olduğu gösterilmiştir. Devam eden çalışmalarımız ise çinko-parmak CXXC ailesinin bir üyesi olarak, CXXC5'in; metil olmayan CpG dinükleotidlere bağlandığını göstermektedir. İfade güçlendirici ve aktifleştirici bölgelerde bulunan CpG adalarının metilasyonunun, genlerin susturulmasına neden olduğu düşünülmektedir. Bu aile proteinlerinin metillenmemiş CpG'lere bağlanmasının, DNA metilasyonunu önleyerek transkripsiyonda rol oynadığı düşünülmektedir. CXXC5 ile ilgili çalışmaların sınırlı olmasına rağmen, CXXC5'in hücresel olaylara, transkripsiyon faktörü, ko-regülatör ve/veya epigenetik faktör olarak katıldığı düşünülmektedir.

E2 aracılığıyla gerçekleşen hücresel olaylarda CXXC5'in rolünü ele almak için; 1) protein partnerlerini tanımlamayı ve 2) CXXC5'in hücresel işlevini tanımlamayı amaçladık. Proteinler hücresel bağlamdaki görevlerini, proteinlerle uzay-zamansal bir düzlemde etkileşerek gerçekleştirdiğinden, ilk olarak olası CXXC5 protein partnerlerinin tanımlanması için maya-iki-melez (Y2H) hizmetini ve sonrasında doğrulama için memeli-iki-melez sistemi kullanmaya karar verdik. Ancak, Y2H ile tanımlanan partnerlerin hiçbirini M2H ile CXXC5 ile etkileşen bir protein partneri olarak doğrulayamadık. Bu sonuçlara dayanarak, etkileşen proteinlerin biyotin etiketlenmesiyle işaretlenerek tanımlanmasını sağlayan BioID yaklaşımını kullanmaya karar verdik. Devam eden çalışmalarımız, CXXC5 protein partnerlerinin tanımlanması için BioID'nin kullanılabileceğini göstermektedir.

E2 aracılığıyla gerçekleşen hücresel olaylarda CXXC5'in görevini ele almak için Y2H'ye paralel olarak yaptığımız çalışmalarda, CXXC5'in ifadesinin arttırıldığı veya susturulduğu hücre hatları üretmek istedik. Ancak hücresel ölüm nedeniyle; hücre hatlarının üretiminde başarısız olduk. Başarısızlığımız, CXXC5 üzerinde fonksiyonel çalışmalar yapmamızı sağlayabilecek hücre hatlarını oluşturmak için tetiklenebilir bir ifade sisteminin kullanılmasını gerektirdi. Sistemin, CXXC5 ifadesi için açma / kapatma anahtarı içeren sürdürülebilir hücre hatları sağlaması bekleniyordu. Bunun için, hedef genin tetrasiklin (Tet) uyarılabilir ifadesini sağlayan pINDUCER sistemini kullandık. Bu amaçla, başlangıçta CXXC5 sentezini susturabilmek için shRNA yaklaşımı kullandık. Her ne kadar monoklon üretimini başarmış olsak da, CXXC5 sentezini azaltamadık. İfadesi arttırılmış veya susturulmuş CXXC5 elde etmek için çalışmalara farklı hücre hatları ve ekspresyon sistemleri ile devam edilmektedir.

Anahtar kelimeler: östrojen, östrojen reseptörü, CXXC5

ACKNOWLEDGEMENTS

I would like to thank first and foremost my thesis advisor, Prof. Dr. Mesut Muyan, for his support and guidance throughout this research. I was a little duckling not knowing what to do and he gave not only his wisdom but also his helping hand in the laboratory. I am grateful for all the things he thought me to make me a better scientist.

I would like to thank members of thesis examining committee; Assoc. Prof. Dr. Ayşe Elif Erson Bensan, Prof. Dr. Sreeparna Banerjee, Assoc. Prof. Dr. Çağdaş Devrim Son, and Assoc. Prof. Dr. Bala Gür Dedeoğlu for their time, precious suggestions and comments.

I also thank Prof. Dr. Rengül Çetin Atalay for her valuable suggestions to our research and providing her laboratory facilities for our use. I thank Assoc. Prof. Dr. Çağdaş Devrim Son for his helps in fluorescent microscopy imaging; Assoc. Prof. Dr. Nathan Lack and his PhD student Fatma Özgün for teaching me pseudo viral particle production techniques. I also thank laboratories B56, B58 and B59 they all have considered me as one of their laboratory members.

It was a great pleasure for me to know my labmates; Çağla Ece Olgun; the one with lots of shared memories, Gamze Ayaz; the one with solutions for all problems, Gizem Kars; the one with anarchism, Gizem Turan; the one with the energy enough to provide electricity for whole department, Kerim Yavuz; the one with fuzzy hair and knowledge contest winner, Negin Razizadeh the one who was with me all the time and will always be, Pelin Yaşar the one for the rewarding advices, Sırma Damla User the one with the wit. I thank all for sharing their knowledge, assistance and friendship.

My lifelong friends Almıla Yerli and Rümeysa Yıldırım I cannot thank enough for always being there for me, it is a privilege to have you in my life. And my dearest

friends Beril Demirçelik, Betül Çalışkan Batı, Betil Yıldız, Elif Yıldız, Nagihan Öcal, Özlem Buluş, Sıla Akülke, Taner Bulut; I would like to thank them for their moral support. I am also grateful to be part of METU archery team and thank my all team members for their friendship and support. My sincere thanks also goes to Ayşegül Dede, Esin Gülce Seza, Onur Özer; whose constant support allowed me to carry out my daily routine and Özge Özgüç without you I would not be able to graduate with a bachelor degree. I also thank my academic life coach İlke Süder she always inspired me when I lost my faith in myself. I thank my dearest cats Sherlock and Watson for accompanying me during long nights of thesis writing. Last but not least Seçkin Baytekin, thank you for bringing joy to my life.

I cannot find any suitable words to express how thankful I am to have a sister like you Ezgi; you are the only person whom I will never regret loving. Having you by my side while accomplishing my dreams were and will be precious. I would like to mention how lucky I am to have my baby sister, Lara; who is not able to read this right now but I am sure that she will achieve best of her choice in life. I thank my grandmother, my aunt, and my cousins Atakan and Beyza. My dearest family, thanks for being there for me.

I would like to acknowledge TÜBİTAK, for the supporting this study and providing me with a scholarship with TÜBİTAK 114Z243 project.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURES

LIST OF TABLES

TABLES

CHAPTER 1

INTRODUCTION

Estrogen–Estrogen Receptor signaling

Estrogens are female sex hormones that are involved in many biological and physiological processes especially in reproductive organs and mammary glands. 17β-estradiol (E2) is the primary circulating estrogen hormone in the body. E2 is synthesized in and secreted from ovaries and acts as a pro-mitotic agent for cellular proliferation and growth of target tissues/organs [1]. Besides, E2 is also one of the main factor contributing to target tissue malignancies like breast cancer [2], which is the most commonly encountered cancer type among women and it can also be seen in men.

The effects of E2 on cells are mediated by the transcription factors, estrogen receptors ERα and ERβ that are encoded by *ESR1* and *ESR2* respectively [3]. ERs share high sequence identity that result in common functional features. ERα is the main estrogen receptor synthesized in breast tissue.

The gene encode for ERα is located at q24-q27 of chromosome encompassing 300 kb genomic segment. The gene contains eight exons encoding 15 transcripts, most of which encodes 595 amino-acids protein with a molecular mass (MM) of ~66 kDa protein [3], [4].

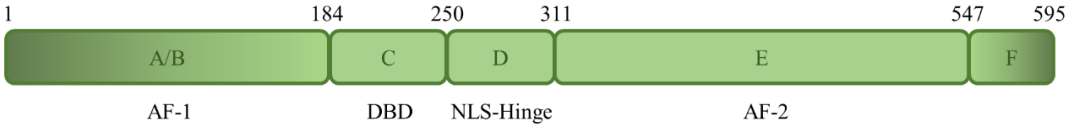

Figure 1 Schematic representation of functional domains of ER α

Nuclear hormone receptor ERα contains six functionally distinct domains which are A/B, C, D and E/F. The amino-terminal A/B domain of ERα is responsible for stimulating transcription from certain estrogen responsive promoters and this domain acts independently from any ligand. The C domain is the DNA binding domain (DBD) that contains two Nuclear Receptor C4 type Zinc Finger domain at the position of 185-205 and 221-245. The DBD is responsible for ERα to bind to derivatives of a DNA sequences, estrogen responsive element (ERE; GGTCAnnnTGACC, wherein the palindromic sequences are separated by three nonspecific nucleotides, n) (Figure 2) [4], [1], [5]. Nuclear Localization Signal (NLS) containing D domain acts as flexible hinge for the receptor. The ligand-binding domain (LBD) of ERα is a multifunctional domain. The E domain is responsible for hormone-binding, dimerization as well as ligand-dependent activation function (AF-2). The F domain is located at the extreme carboxyl-terminus of the ERα (Figure 1).

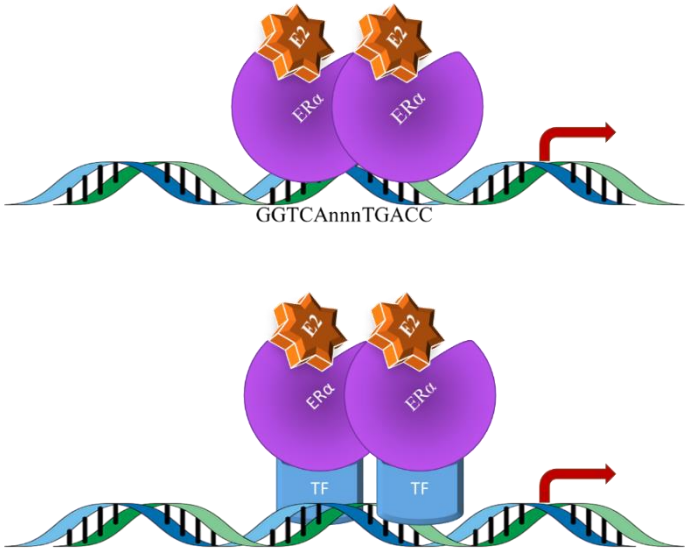

Figure 2 Schematic representation of ERE-dependent and ERE-independent signaling pathways

The binding of E2 is essential for ERs to be functionally active. E2 binding causes re-organization of the ligand-binding domain and activate ERα. This activation increases the stability as well as the interaction with Estrogen Response Element (ERE) sequences and coregulatory proteins of the ERα dimer [5].

The nuclear ERα convey the E2 signaling in the nucleus two distinct ways: The ERE-dependent signaling route which involves the direct binding of ERα to ERE. The ERE-independent route encompasses a direct/indirect interactions with other transcription factors bound to their cognate response elements on DNA. Both pathways are critical for the expression of genes important for cellular processes [2], [6].

While the majority (95%) of the newly synthesized ERα is translocated to the nucleus independent of the ligand, the fraction of the receptor shows peripheral membrane, cytoplasmic and mitochondrial locations [5], [7], [8], [9].

CXXC5 and its protein product

Previous studies conducted in our laboratory using microarray suggests that *CXXC5* is an estrogen responsive gene [10]. Moreover, our recent studies suggest that regulation of *CXXC5* gene expression is regulated by ERα through the ERE dependent signaling pathway [11] (Figure 3). This ERE-dependent pathway for *CXXC5* expression involves the binding of ERα to a non-consensus ERE sequence, located on -242 bp upstream of the first ATG (A is at position +1) of the *CXXC5* gene and composed of GGTCAggaTGAC**A** sequence that differs from the consensus sequence with one nucleotide (bold A) [10], [11].

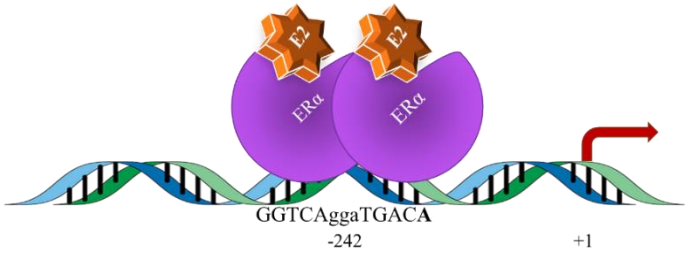

Figure 3 ERE dependent regulation of CXXC5

The *CXXC5* (ENSG00000171604) gene is on 5q31.2 forward strand according to GRCh38.p7 assembly. The *CXXC5* spans on ~35.5 kb genomic DNA region and is organized in 11 exons resulting 16 transcript variants, 14 of which is protein coding; with main transcript of 2658 nucleotide long mRNA and 969 base pair long coding sequence resulting 322 amino acids with a MM of ~33 kDa [4], [12].

The CXXC protein contains two CXXC type zinc finger motif between amino acids 256 and 297, each finger with four cysteine residues coordinated by a zinc ion. Moreover, CXXC5 bears a nuclear localization signal on its C terminus, leading to nuclear localization, at least, in MCF7 cells [13], [11]. Phosphorylation may occur at a threonine residue at positon of 53th [14], [13] (Figure 4).

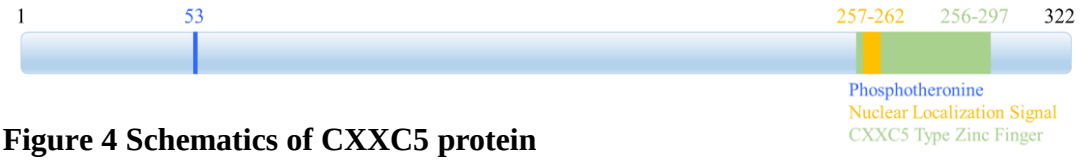

Figure 4 Schematics of CXXC5 protein

CXXC5 alias are CXXC Finger Protein 5 (CF5), Retinoid-Inducible Nuclear Factor (RINF) or WT1-Induced Inhibitor of Dishevelled (WID). CXXC5 is considered to belong CXXC-type zinc finger (ZF) protein family. ZF-CXXC family includes CXXC protein finger 1 (CFP1, CXXC1), Lysine (K)-Specific Demethylase 2B (KDM2B, CXXC2), Methyl-CpG Binding Domain Protein 1 (MBD1, CXXC3), Lysine (K)-Specific Demethylase 2A (KDM2A; CXXC8), DNA (Cytosine-5-)-Methyltransferase 1 (DNMT1, CXXC9), and Tet Methylcytosine Dioxygenase 1, 2, 3 (TET1, TET2, TET3) proteins [15]. CXXC-type zinc finger family proteins have an affinity for non-methylated CpG containing regions of DNA. Most CpG dinucleotides are methylated, with the exception of those within CpG islands (CGIs), which are usually unmethylated. Methylated CpG islands (CGIs) which are set and maintained by DNA methyltransferases (DNMTs) are associated with transcriptionally silenced genomic DNA regions. CGIs can be found on repetitive sequences, enhancers, promoters, and gene bodies, thereby influencing local chromatin structure to regulate gene activity. The binding of CXXC-type zinc finger protein family to non-methylated CpGs prevents DNA methylation, generates an environment permissive to transcription [16].

Studies on CXXC5

Studies on the *CXXC5* gene structure and regulation as well as the functional features of CXXC5 protein are limited. However, it appears that CXXC5 as an epigenetic factor and/or transcription regulator is involved in cell differentiation, cellular proliferation, differentiation, death. Studies suggest that CXXC5 participates in TGF-β signaling, mediates Wnt signaling activity in neural stem cells and, regulates the expression of bone morphogenetic protein 4 (BMP4) involved in myelopoiesis and human acute myeloid leukemia.

Wnt signaling pathway is important in embryonic development, stem cell pool regulation, cell migration and specialization, adult healing processes and homeostasis. Wilms tumor gene (WT1), which is a transcription factor having roles in cellular development and cell survival in bone homeostasis, negatively regulates WNT/β-catenin signaling pathway by mediating the expression of the *CXXC5* gene. Studies showed that CXXC5 can also act as a negative feedback regulator of the Wnt/β-catenin pathway by interacting with the Dishevelled (Dvl) protein [17], [18],[19]. Moreover, this Dvl–CXXC5 interaction is suggested to be a potential target for anabolic therapy of osteoporosis based on the reasoning that if the Dvl–CXXC5 interaction is disrupted, Wnt/β-catenin pathway could be activated again and boost primary osteoblast differentiation. To overcome negative feedback effect of CXXC5 on Wnt/β-catenin pathway and create anti-osteoporosis drugs small-molecule inhibitors are being designed against the Dvl–CXXC5 interaction [20], [21].

To accomplish normal brain functions to be maintained, myelination in corpus callosum is required. CXXC5 was reported to be involved in the Wnt/β-catenin signaling by directly binding to MBP promoter within neural stem cell differentiation (NSC) and decrease expression of myelin related genes. This regulation causes aberrancy in myelin structure and reduces electrical transmission at corpus callosum [22].

Transcriptional regulatory role of CXXC5 has been found in different tissues. CXXC5 is suggested to act an enhancer of genes involved in skeletal muscle differentiation, and absence of it inhibits myocyte differentiation [23]. Moreover, regulated by BMP4, the presence of CXXC5 induces the expression of *Flk-1* gene whose protein product is a receptor for vascular endothelial growth factor during endothelial differentiation and vessel formation [24].

Acute promyelocytic leukemia (APL) is usually triggered with retinoids through the modulation of myeloid regulatory gene transcriptions. *CXXC5* as a retinoid-responsive gene is highly expressed in human acute myeloid leukemia (AML) cells. Dysregulation of the *CXXC5* gene expression by promoter methylation was suggested to correlate with the poor prognosis of AML patients. Moreover, an altered expression of CXXC5 was reported to alter the expression as well as the phosphorylation statuses of several intracellular signaling mediators in PI3K-Akt-mTOR signaling pathway in induced differentiation of leukemic cells and in cytokine-induced myelopoiesis of normal $CD34^+$ progenitor. Based on these observations, it was suggested that deregulated *CXXC5* expression is critical factor in human leukemogenesis [16], [25], [26], [27], [28].

It was also shown that CXXC5 is involved in TGF-β signaling and that deregulated expression of *CXXC5* could lead to the development of cardiovascular disease by adversely affecting TGF-β signaling [29], [30]. In contrast, CXXC5 through the regulation of TGF-β expression was reported to act as a tumor suppressor in hepatocellular carcinoma [30], [31].

Studies also suggest that CXXC5 acts also as an epigenetic regulator [32], [33]. CXXC5 was reported to recruit DNA demethylase Tet2 to CpG islands. The role of CXXC5 in TLR7/9- and virus-induced interferon response important in terms of their methylation status. CXXC5 was observed to be synthesized at high levels in plasmacytoid dendritic cells (pDCs). The silencing of CXXC5 by siRNA approaches results in aberrant methylation of CpG island in the *Irf7* gene locus leading to decrease

IRF7 protein levels in pDCs. IRF7 plays a role in the transcriptional activation of virus-inducible cellular genes, including interferon beta chain genes. The decrease in IRF as a result of a diminished CXXC5 was shown to lead an increased susceptibility to viral infections [32], [33].

The Cellular Tumor Antigen p53 is known as critical tumor suppressor in DNA damage response pathway by inducing genes responsible for cell cycle arrest, senescence, apoptosis, and DNA repair. CXXC induces p53 transcriptional activity and apoptosis by interacting with Ataxia Telangiectasia Mutated Protein Kinase (ATM). The interaction of CXXC5 with ATM was reported to prevent the phosphorylation of ATM thereby preventing the recruitment of the protein to DNA break sites. Based on these observations CXXC5 is also suggested to involve in DNA damage pathway via ATM-p53 signaling [34].

Aim of This Study

Based on our findings that the expression of *CXXC5* gene is modulated by E2-ERα through ERE-dependent signaling pathway, we predict that CXXC5, synthesized as the primary response gene product, participates in the regulation of E2-mediated cellular proliferation and death.

Proteins fulfill their functions within the context of dynamically changing composition of proteins in a spatiotemporal manner. To explore the potential protein partners of CXXC5, we decided to use yeast-two-hybrid (Y2H) system. Based on these findings, we planned to verify these interactions in mammalian cells using mammalian-two-hybrid, co-immunoprecipitation and co-localization approaches.

Our laboratory had attempted to generate cell lines that stably express *CXXC5* gene or shRNA against CXXC5-mRNA. However, we failed at several attempts to generate cell lines due to severely disrupted cellular growth with or without altered phenotypical features or cell died before the completion of monoclonal selections. Our failure

therefore necessitated the establishment of an inducible expression system to generate cell lines that allow us to conduct functional analyses on CXXC5.

Inducible expression approaches depend on the presence or absence of biotic or abiotic factors. Gained inspiration from *Escherichia coli* tetracycline resistance operon containing transposon, regulation of gene expression via inducible promoters become very powerful tool in molecular biology. Expression of genes or shRNAs connected to inducible promoter provides opportunity of rapidly and reversibly turned on or off the gene expression at any time. This advantage of the system is expected to provide us maintainable cell lines containing an on/off switch to control CXXC5 synthesis. For this purpose, we decided to utilize powerful, versatile and widely used tetracycline (Tet) inducible expression system to generate cell lines that we could control CXXC5 synthesis by initially using an shRNA approach.

The aim of this study is to lay ground for understanding of CXXC5 functions in cellular processes by identifying its interacting protein partners and by controlling CXXC5 synthesis.

CHAPTER 2

MATERIALS AND METHODS

Cell lines and maintenance

MCF7 cell line was a kind gift of Assoc. Prof. A. Elif Erson Bensan (Middle East Technical University, Ankara, Turkey) and HEK 293T cell line was obtained from Assoc. Prof. Nathan Lack (Koç University, Istanbul, Turkey).

MCF7 and HEK 293T cells were grown in high glucose (4.5 g/L) containing Dulbecco's Modified Eagle's Medium without phenol red to prevent mimicking estradiol of phenol (DMEM, Lonza, Belgium, BE12-917F) with addition of 8% final concentration of Fetal Bovine Serum (FBS, Merck, Germany, S0115), 20mM L-Glutamine (L-Glutamine, Lonza, Belgium, BE17-605E), 100 units of Potassium Penicillin and 100 μg of Streptomycin Sulfate per mL of the media (Pen/Strep, Lonza, Belgium, BE17-602E).

When there was a need of estrogen treatment, cells were maintained in DMEM supplemented with 8% charcoal dextran (Dextran Coated Charcoal, Sigma, Germany, C6241-20G) treated FBS (CDFBS), 20mM L-Glutamine, 100 units of Potassium Penicillin and 100 μg of Streptomycin Sulfate per mL of the media.

All cells were grown as monolayer, passaged or refreshed every three days according to needs and maintained maximum of 6 weeks in 95% humidified, 5% CO_2, 37^{o}C cell culture incubator.

When cells need to be stocked; cells trypsinized (Trypsin EDTA, Lonza, Belgium, BE17-161E), collected and centrifuged. Cell pellet was resuspended with the complete medium of which FBS concentration increased to 20% and 10% DMSO (Dimethyl sulfoxide Cell culture grade, Applichem, USA, A3672,0250) added and 0.22µm filter sterilized before use. Cells were then aliquoted into 1 mL cell suspension per cryovial. Prepared stocks were gradually cooled with the help of isopropanol container at -80^{o}C for 24 hours. Frozen vials were then transferred in to liquid nitrogen for long-term storage. When cells were needed for experimentations, frozen stock cryovials were thawed in 37^{o}C water bath and resuspended in 5 ml pre-warmed complete medium quickly to decrease effect of DMSO on cells, and centrifuged. The spend medium was discarded and pellet was re-suspended with a 4 ml fresh complete medium and seeded in T25 flask. Maintained as described above.

All reagents and chemicals were cell culture grade and all media were sterilized using 0.22µm filter unit before use.

Hormone treatments

In experiments requiring estrogen treatment; cells were starved with charcoal coated dextran (Charcoal, Dextran Coated, Sigma, USA, C6241-20G) stripped (APPENDIX J) fatal bovine serum (CDFBS) containing medium with at least 48 hours. Treatment done with 10^{-9} M E2 (physiological level of E2) or as vehicle control % 0.01 ethanol containing CDFBS media.

Identification of putative interaction partners of CXXC5

Yeast two-hybrid screening was carried out by Creative BioLabs (USA) to determine possible interacting partners of CXXC5. The bait for Y2H screening was constructed by subcloning the MCF7 cDNA library into the vector pDBLeu (Invitrogen) and transformed into the strain MaV203 (MATα, leu2-3,112, trp1-901, his3Δ200, ade2-101, gal4Δ, gal80Δ, SPAL10::URA3, GAL1::lacZ, HIS3UAS GAL1::HIS3@LYS2, can1^{R}, cyh2^{R}) and absence of self-activation was verified. In the screening, MCF7

cDNA library were cloned into pPC86 (Invitrogen) plasmids as preys and were tested for His, Ade and LacZ activation. Positive transformants were retransformated into MaV203 to test for His, Ura and LacZ activation. The identity of the positive interactor was determined by sequencing.

Results revealed that of the 26 positive colonies (1 His/Ura and 25 His), five clone sequences have corresponding sequences in the human genome sequence. Based on the verifiable sequence information, Cold Inducible RNA Binding Protein (CIRBP), Poly(RC) Binding Protein 2 (PCBP2), Ribosomal Protein L15(RPL15), Ribosomal Protein S2 (RPS2) and Ribosomal Protein S4, X-Linked (RPS4X) were reported to us as possible CXXC5 interactors. To confirm these interactions, we initially used mammalian two hybrid system. We obtained cDNAs of these genes from Addgene (Addgene non-profit plasmid repository, Cambridge) and cloned with appropriate restriction enzyme sites to dual luciferase reporter assay component vectors; pGAL4-DBD that contains GAL4 DNA Binding Domain and pVP16-AD which contains transcriptional Activation Domain.

Cloning of the CXXC5 and its putative protein partners to dual luciferase reporter assay component vectors; pGAL4-DBD and pVP16-AD

CXXC5, CIRBP, PCBP2, RPL15, RPS2 and RPS4X cDNAs were amplified by PCR using designed primers containing required restriction enzyme cut sites as follows (Table 1);

Table 1 PCR components

Template (cDNA)	100 ng
FP	5 µl
REP	5 µl
dNTP (2mM)	5 µl
5x GC buffer	10 µl
Phusion DNA Polymerase (Thermo Scientific, USA, F530L)	0.5 µl

Molecular grade water	Up to 50 µl

Amplification PCR conditions run by thermal cycling machine (Table 2);

Table 2 PCR conditions

Initial denaturation	98ºC	3 minutes	
Denaturation	98ºC	30 seconds	X45
Annealing and extension	72ºC	30 seconds	
Final extension	72ºC	10 minutes	
Final hold	4ºC	∞	

Amplified products were then used for restriction enzyme digestion for mammalian two hybrid vectors. We used XhoI (XhoI, New England Biolabs, USA, R0146) and HindIII (HindIII, New England Biolabs, USA, R3104) restriction enzymes with 1x Cutsmart buffer (Cutsmart Buffer, New England Biolabs, USA, B7204) at 37ºC, and the vectors digested with the same enzymes.

Vectors treated with alkaline phosphatase (FastAP Thermosensitive Alkaline Phosphatase, Thermo Scientific, USA, EF0654) at 37ºC to prevent self-ligation. Required phosphate group for ligation reaction were provided by the insert only.

After digestion, vector was run on 1% agarose gel prepared with 0.5 X Tris Borate EDTA (Agarose, Sigma, Germany, A9539). DNA fragments were recovered from the gel using Zymoclean™ Gel DNA Recovery Kit (Zymoclean Gel DNA Recovery, Zymo Research, USA, D4007).

The recovered DNA concentration and purity was measured using a BioDrop (BioDrop µLITE, BioDrop, UK, 80-3006-51). 50µg linearized vector was ligated with insert at 1:6 ratio using T4 DNA Ligase (T4 DNA Ligase, Thermo Scientific, Lithuania, EL0014) and Rapid Ligation Buffer (Rapid DNA Ligation Kit, Thermo Scientific, Lithuania, K1422) at room temperature for 15 minutes.

Ligation reaction was introduced into chemically competent *E.coli XL-1 Blue* cells (XL1-Blue, Stratagene, USA, 200249), which was prepared with a RbCl approach (RbCl, Merck, Germany, 1.07615.0025). After the ligation mixture and bacteria were mixed together, the mixture was placed on ice for 1 hour. Heat shock was then applied at 42°C for 45 seconds followed by 2 minutes of incubation on ice. Then 450µl of LB (LB Broth (Miller), Sigma, Germany, L3522) was added on the cells and incubated at 37°C for 1 hour with shaking at 180 rpm. Cells were spreaded on LB-Agar plates containing 100µg/ml ampicillin (Roche Applied Science, Switzerland, 10835269001).

After colony selection, 5 ml of bacterial culture used for plasmid DNA isolation using QIAprep Spin Miniprep Kit (QIAprep Spin Miniprep Kit, Qiagen, Germany, 27106). Constructs were sent for sequencing to ensure sequence fidelity.

Transfection

To perform mammalian two hybrid experiments, MCF7 cells were transiently transfected with *Renilla* luciferase containing vector, firefly luciferase containing reporter vector and expression vector bearing cDNA for combinations of possible interactor protein.

For the experiment, $4x10^4$ MCF7 cells were seeded per well of 48-well plate with DMEM-FBS or DMEM-CDFBS complete medium.

48 hours after seeding cells were transfected with DNA mixture indicated below (Table 3) with the complex formation provided with TurboFect transfection reagent (TurboFect Transfection Reagent, Thermo Scientific, USA, R0531) at the ratio of 1:2 (DNA in µg: TurboFect in µl) in DMEM; incubated at room temperature for 30 minutes and introduced on cells.

With proper controls of all vectors to be used in the experiment; there is a need for measuring activity of all vectors one by one to be able to make deductions from results. Sample experimental design is given below (Table 3).

Table 3 Experimental design with proper controls for M2H assay

RENILLA (0.5ng)	REPORTER (125ng)	Interactor I (75ng)	Interactor II (75ng)
SV40-*Renilla*	pGL3-Gal4-RE	-	-
SV40-*Renilla*	pGL3-Gal4-RE	Gal4-DBD EV	-
SV40-*Renilla*	pGL3-Gal4-RE	VP16-AD EV	-
SV40-*Renilla*	pGL3-Gal4-RE	Gal4-DBD-X	-
SV40-*Renilla*	pGL3-Gal4-RE	Gal4-DBD-Y	-
SV40-*Renilla*	pGL3-Gal4-RE	VP16-AD-X	-
SV40-*Renilla*	pGL3-Gal4-RE	VP16-AD-Y	-
SV40-*Renilla*	pGL3-Gal4-RE	Gal4-DBD EV	VP16-AD-X
SV40-*Renilla*	pGL3-Gal4-RE	Gal4-DBD EV	VP16-AD-Y
SV40-*Renilla*	pGL3-Gal4-RE	VP16-AD EV	Gal4-DBD-X
SV40-*Renilla*	pGL3-Gal4-RE	VP16-AD EV	Gal4-DBD-Y
SV40-*Renilla*	pGL3-Gal4-RE	Gal4-DBD-X	VP16-AD-Y
SV40-*Renilla*	pGL3-Gal4-RE	Gal4-DBD-Y	VP16-AD-X
SV40-*Renilla*	pGL3-Gal4-RE	VP16-AD ERαAB	Gal4-DBD ERαEF
SV40-*Renilla*	pGL3-Gal4-RE	DBD_AD fusion	-

VP16-AD ERαAB and Gal4-DBD ERαEF transfected cells were starved with CDFBS for 48 hours for hormone treatments and then after transfection medium was changed with CDFBS containing 10^{-9} M E2.

24 hours after transfection cells were washed with 1x Phosphate Buffered Saline (PBS-10X w/o Ca^{++}, Mg^{++}, Lonza, Belgium, 17-517Q) two times.

Luciferase reading

Cells were then prepared for luciferase reading according to the manufacturer's manual (Dual-Luciferase® Reporter Assay System, Promega, USA, E1910). Cells were lysed with 50µl 1x Passive Lysis Buffer (PLB) for 30 minutes. PLB provides optimal environment for performance and stability of the firefly and *Renilla* luciferase enzymes to work and has minimal auto-fluorescence activity. 25 µl of lysates were transferred in to 96 well lumitrac microplate (Greiner LUMITRAC™ 200 microplate, Sigma, USA, GN655075) for analysis.

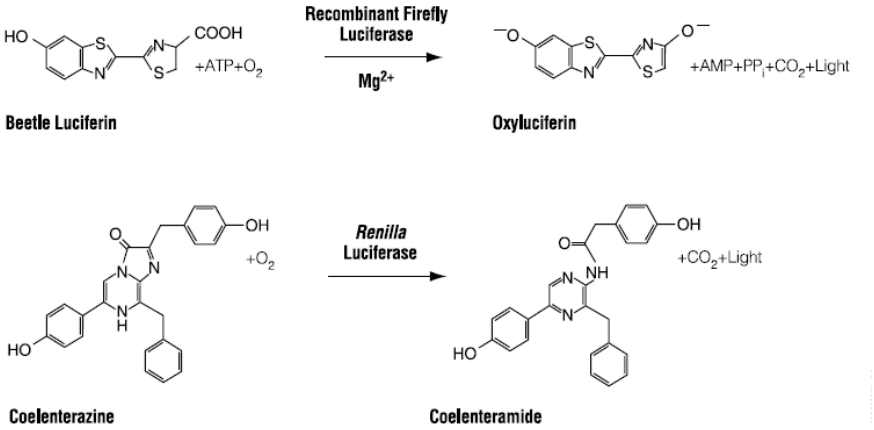

Figure 5 Chemical reaction of firefly luciferase and *Renilla* luciferase in the presence of their substrate

Using GloMax®-Multi Detection System (GloMax®-Multi Detection System, Promega, USA, E7041) each well was introduced with first 25 µl of Luciferase Assay Reagent II (LARII) containing beetle luciferin to be utilized by firefly luciferase, then 25 µl of Stop & Glo® containing coellentarazine to be utilized by *Renilla* luciferase for light formation.

Cloning of the CXXC5 shRNAs into pINDUCER10-mir-RUP-PheS vector

In inducible expression, we used an inducible promoter containing vector; pInducer10-mir-RUP-PheS [35].

shRNAs were designed with XhoI and MluI flanking sequences and a hairpin; XhoI and MluI cut sites were highlighted with purple and yellow respectively (APPENDIX A);

```
5′TCGAcccTCAGATTTGCAAATTCAGAAAttcaagagaTTTCTGAATTTGCAAATCTGAttttta    3′
3′    gggAGTCTAAACGTTTAAGTCTTTaagttctctAAAGACTTAAACGTTTAGACTaaaaaTGCGC5′
```

Forward and reverse strands of the oligos were dissolved as 1µg/µl with Tris-EDTA buffer (pH: 8.0). To anneal oligos, 10 µl forward and 10µl reverse oligos were mixed with 30µl TE buffer and this mixture was heated up to 95^{o}C for 5 minutes then let it cool-down overnight gradually. Annealed oligo inserts were kept at -20^{o}C until use.

pINDUCER10-mir-RUP-PheS vector was digested with MluI (MluI-HF, New England Biolabs, USA, R3198) and XhoI (XhoI, New England Biolabs, USA, R0146) in the presence of 1x Cutsmart buffer (Cutsmart Buffer, New England Biolabs, USA, B7204) at 37^{o}C. Vector was not treated with alkaline phosphatase since phosphorylated ends were required to ligate annealed oligos.

After digestion, the vector was run on and then recovered from 1% agarose gel. Annealed oligos already have flanking ends compatible with MluI and XhoI sticky ends ready to ligation. 50µg linearized vector and 1:25 ratio of insert was ligated with T4 DNA Ligase at room temperature for 15 minutes and followed by transformation into XL1-Blue cells.

After clone selection, positive constructs were sequenced for sequence fidelity.

Pseudo Viral Particle Production

Lentiviral system contains a minimum of three plasmids that includes 1) transfer plasmid containing cDNA or shRNA for gene of interest; 2) packaging plasmid (the second generation has one plasmid containing Gag, Pol, Rev, Tat while third generation has two plasmids encoding Gag, Pol and another encoding Rev of the packaging system); 3) envelope plasmid encoding VSV-G.

To produce virus; SV40 T antigen transfected human embryonic kidney (HEK293) cells, HEK 293T, were used since the T antigen enhances expression of viral proteins and increases pseudo viral particle production.

To seed cells according to experimental procedure, the spend medium was discarded, room temperature PBS was slowly added from edge of the dish and discarded after gentle washing of the cells. Adequate amount of trypsin (approximately 2.5 ml) was added on the cells and incubated in 37^{o}C incubator for ~5-7 min. When the cells were de-attached, they were collected with the complete medium. A small aliquot of cell suspension (~100µl) was taken into an Eppendorf and vortexed for counting. 10 µl of suspension was dropped on a hematocytometer. At least four regions hematocytometer were counted to obtain average cell number/ml. After calculation $2x10^5$ cells was seeded in 6 well plate with 2 ml medium.

Two days after seeding, cell confluency is expected to be around 80%.

Transfection for pseudo viral particle production

Transfection ratios for a 6-well plate were described in the table below (Table 4, Table 5):

Table 4 Transfection ratios of vectors for second generation to produce pseudo viral particles

2nd generation	
Transfer vector (pINDUCER)	2.000µg
VSVG (pMD2G)	0.500µg
Gal/Pol, Rev, Tat (psPAX2)	1.500µg
DMEM only	100µl

Table 5 Transfection ratios of vectors for third generation to produce pseudo viral particles

3rd generation	
Transfer vector (pINDUCER)	2.000µg
VSVG (pMD2G)	0.900µg
Rev (pRVS-REV)	0.600µg
Gal/Pol (pMDL/pRRE)	1.500µg
DMEM only	100µl

In an Eppendorf tube, DNA mixture indicated amounts were mixed and in another Eppendorf; 100 µl of DMEM and 2µl of TurboFect (TurboFect Transfection Reagent, Thermo, USA, R0531) per µg of DNA were mixed. Then DNA-DMEM and Transfection reagent-DMEM mixtures were combined, mixed well and incubated at room temperature for 30 minutes for the formation of the transfection complex. Then mixture was added on the cells dropwise. Cells were placed back into the incubator.

The day after transfection, medium containing transfection complex was discarded and pre-warmed fresh medium was added onto cells.

48-hour post-transfection, the medium containing the first pseudo viral particle was collected. The medium containing the first particles was gently collected into the appropriate size centrifuge tube. After collecting the medium, fresh pre-warmed 2mL

medium was put on the cells slowly and the plates were placed back into the incubator. Collected medium was capped and sealed with parafilm and store in +4°C for overnight.

72 hours after transfection, the medium collection was repeated and combined with 48h collection. Then the medium was cleared from debris and cells using a 45µm filter. The filtered pseudo viral particle containing medium was then aliquoted into 1ml in 1.5 Eppendorf tubes and kept frozen at -80°C. All disposable plastic ware was treated with bleach and discarded into a biological waste disposal.

Transduction with pseudo viral particles

For infection (transduction) $2x10^5$ MCF7 cells per well seeded in 6 well-tissue culture plates with 2 ml medium to allow cells to become about 50% confluent 24 hours later for infection.

Transduction mixture was prepared for each well to be infected with the mixture of 1 ml of viral supernatant and 1 ml of fresh complete DMEM-FBS medium. The mixture was then received proteamine sulfate (Proteamine sulfate salt from herring, Sigma, Germany, P4505) to a final concentration of 8µg/ml. The whole mixture was subsequently introduced onto cells directly. 24 hours later, the medium was changed with DMEM-Tet Free FBS (Fetal Bovine Serum Tetracycline Free, Biowest, USA, S181T-500) to prevent unwanted induction of the cDNA/shRNA expression.

Generation of stable cell lines

72 hours after infection, selection was started with 1µg/ml of puromycin in DMEM-Tet Free FBS. This was based on our preliminary studies in which 1µg/ml of puromycin was the minimal dose for the MCF7 cell line to maximally kill cells within one week. Monoclones within three weeks of antibiotic treatment were selected and placed into a new culture plates.

Monoclonal selection

Puromycin selection was carried for one week until there were no cells surviving in the uninfected control cells.

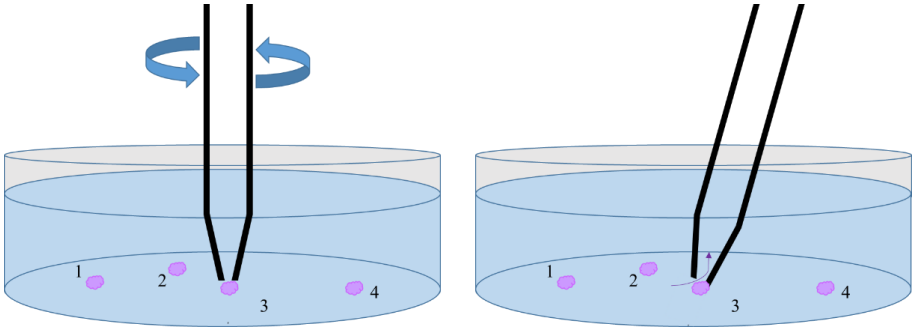

Figure 6 Monoclonal selection

When surviving cells started to grow as colonies. Colonies were marked with a marker underneath the tissue culture plate. Marked colonies were then lifted out from the plate with suction using 1 ml pipette tip attached to a pipette and transferred into a 48 well tissue culture plate.

Firstly, media of the cells growing in is refreshed then using a 1000µl volume capacity tip, containing ~100µl media and adjusted 250µl suction volume; tip was carefully lowered on the colony. Then with gentle twisting cells were separated from the possible surrounding cells. Pipette was then angled for about 45^{o} to allow the media with the cells be sucked into the pipette. Cells were then placed into 48 well tissue culture plates for further growth (Figure 6 Monoclonal selection).

Testing the stable cell clones

To make sure obtained clones contain all the regulatory cassette of the inducible system, we treated cells with 1µg/ml of doxycycline (DOX, Doxycycline hyclate, AppliChem, Germany, A2951) or its vehicle control DMSO (Dimethyl sulfoxide cell culture grade, AppliChem, Germany, A3672) as control. Since the induction of gene expression produces red fluorescence, we used FLoid™ Cell Imaging Station (FLoid™ Cell Imaging Station, Thermo, USA, 4471136), equipped with the emission

of 646/68 nm and the excitation of 586/15nm which are compatible with the fluorophore we are using, to visually assess the gene induction. The observation of the red fluorescent protein activity in cells are indicative of positive clones which are assumed to carry the regulatory cassette of inducible system for subsequent analysis of transgene synthesis by western blot analysis and RT-qPCR.

Protein isolation

CXXC5 is shown to be nuclear protein [11]. To be able to detect changes in protein amount in the presence of induction reagent; DOX, fractionated protein isolation was done with the NE-PER kit (Nuclear and Cytoplasmic Isolation Kit, Thermo Scientific, USA, 78833) that contains cytoplasmic and nuclear protein extraction buffers. For isolation, we used ~2 x 10^6 induced (DOX) or vehicle control (DMSO) cells; which roughly equals to three combined wells of a 6-well plate.

Collected cells were washed three times with PBS and the isolation procedure followed as indicated in manufacturer's manual. Adequate amount of protease inhibitor was added prior use to CERI and NER buffers of the kit. 100µl CERI was added on the cells, vortexed for 15 seconds and incubated on ice 10 minutes. Then 5.5 µl CERII was added to the tube, mixture was vortexed for 5 seconds, incubated on ice for 1 minute, and vortexed again for 5 seconds. At this step, cell wall was disrupted but the nuclear membrane retains its integrity, with centrifuging at 14000 g for 15 minutes; as the pellet is the nuclear fraction, while the supernatant contains the cytoplasmic fraction. For nuclear protein isolation; 50 µl of NER buffer was added on the pelleted nucleus, 5 repeat of vortex 15 seconds incubating on ice 10 minutes was done to break down nuclear membrane and release nuclear proteins following centrifugation at 14000 g for 15 minutes.

To measure protein concentration, Bradford assay (Quick Start Bradford Protein assay, Bio-Rad, USA, 500-0201) was used. For this, 2 µl of sample, 3 µl of sample buffer and 245 µl of Bradford dye mixed at room temperature and incubated for 10 minutes for color formation. In the meantime, standard curve was generated with the use of

bovine serum albumin (2mg/ml) provided with the kit by diluting it with sample buffer. Standard curve was generated and if R^2 value was close to 1, concentration calculation equation was depicted from the curve. Concentrations of the samples calculated using this equation and dilution factor. Extracted proteins were stored at -80^{o}C until use.

Western blotting

For western blot analysis, proteins were isolated with NE-PER as previously described. Samples were then electrophoresed using Sodium Dodecyl Sulfate Poly Acrylamide Gel (SDS-PAGE). In this study 10% SDS-PAGE was used, which was prepared as indicated in Table 6.

Table 6 SDS-PAGE gel components

SDS gel (separating)	10 %	SDS gel (stacking)	5 %
dH_2O	3.97 ml	dH_2O	5.44 ml
30% Acrylamide:Bis (29:1)	3.33 ml	30% Acrylamide:Bis (29:1)	1.36 ml
1.5 M Tris Buffer (pH 8.8)	2.5 ml	1 M Tris Buffer (pH 6.8)	1 ml
10% SDS	100 µl	10% SDS	80 µl
10% Ammonium persulfate	100 µl	10% Ammonium persulfate	80 µl
TEMED	4 µl	TEMED	8 µl

Before running samples, proteins were denatured with 6X Laemmli buffer with 30% β mercaptoethanol and boiled at 95^{o}C for 5 minutes. Laemmli buffer composition is indicated in Table 7

Table 7 Laemmli buffer components

6X Laemmli Buffer	
12 % SDS	Denaturation
30 % β - Mercaptoethanol	Disulfide bond disruption
60 % Glycerol	Ensure sink of the proteins in to the wells
0.012 % Bromophenol blue	Dye allowing tracking of the run
0.375 M Tris	Buffer environment

50 µg of nuclear protein samples was loaded into wells and the current was run with constant voltage of 100 watt for ~120 minutes. Then proteins transferred onto

electrostatic, hydrophobic PVDF membrane (Polyvinylidene Difluoride Western Blotting Membranes, Roche Applied Science, Switzerland, 03 010 040 001) with wet transfer technique using tris-glycine transfer buffer with 20% methanol. Transfer was accomplished at 4^{o}C with 100 volt voltage for an hour.

For western blotting, membranes were blocked one hour at room temperature with 5% skimmed milk (Blotting-Grade Blocker, BioRad, USA, 1706404) in 0.05% TBS-T (Tris Buffered Saline with 0.05% Tween). Then a CXXC5 antibody (Anti-CXXC5 antibody, Abcam, USA, ab106533) was prepared as 1:500 ratio in 5% skim milk in 0.05% TBS-T and membrane was incubated in it for an hour. After washing of primary antibody with 0.05% TBS-T for 5 minutes for three times, the membrane was incubated with a secondary antibody (Goat Anti Rabbit HRP conjugated, Santa Cruz Biotechnology, USA, sc-2004), at 1:4000 ratio in 5% skim milk in 0.05% TBS-T, for 1 hour at room temperature. The membrane was then washed and were treated with Enhanced Chemiluminescence solution 5 minutes (ECL, BioRad, USA, 1708280) for visualization.

Visualization was carried out with ChemiDoc™ XRS+ System (Molecular Imager® ChemiDoc™ XRS+ System, BioRad, USA, 170-8265). Obtained images were analyzed with ImageLab software (ImageLab software, BioRad, USA, 170-9690). When it was necessary, membranes were stripped and reblotted with different antibodies. Membranes were stripped using mild stripping buffer at 60°C for 15 min, and washed with 0.05% TBS-T three times for subsequent blotting.

RNA Isolation

For total RNA isolation, induced or uninduced ~1 x 10^6 cells were collected with trypsinization and washed for three times with PBS. RNA isolation was performed with Quick-RNA™ MiniPrep (Quick-RNA™ MiniPrep, Zymo Research, USA, R1015) according to manufacturer's instructions The RNA concentration and purity were measured using a BioDrop (BioDrop μLITE, BioDrop, UK, 80-3006-51) and samples were stored at -80°C until use.

Genomic DNA contamination control

500 ng of total RNA was used as PCR template for genomic DNA contamination using PCR targeting primers for glyceraldehyde-3-phosphate dehydrogenase (GAPDH), which gives a PCR product with an expected size of 409 bp,

PCR was conducted according to conditions indicated in Table 8 and

Table 9:

Table 8 Genomic DNA contamination control PCR components

Template (RNA to be checked)	500 ng
GAPDH FP	2 µl
GAPDH REP	2 µl
dNTP (2mM)	2 µl
Taq Buffer with KCl (10X)	2 µl
$MgCl_2$ (25mM)	1.6 µl
Taq DNA Polymerase (5 U/µL) (Thermo Scientific, USA, EP0402)	0.2 µl
Molecular grade water	Up to 20 µl

Table 9 PCR reaction conditions in thermal cycler

Initial denaturation	95ºC	3 minutes	
Denaturation	95ºC	30 seconds	
Annealing	65ºC	30 seconds	X40
Extension	72ºC	30 seconds	
Final extension	72ºC	10 minutes	
Final hold	4ºC	∞	

As positive control for PCR reaction, 50 ng of genomic DNA was also used. 5 µl of 20 µl samples were run on 1 % agarose gel.

cDNA synthesis

After ensuring that RNA was free of genomic DNA contamination, RNA samples were subjected to cDNA conversion using RevertAid First Strand cDNA Synthesis Kit (RevertAid First Strand cDNA Synthesis Kit, Thermo Scientific, USA, K1622) according to manufacturer's instruction, as Table 10;

Table 10 cDNA synthesis reaction

Total RNA as template	500 ng
Oligo $(dT)_{18}$ primer	1 µl
Nuclease free water	Up to 12 µl

Prepared mixture was incubated at 65°C for 5 minutes and then chilled on ice.

At the same time, following components were prepared to mix with oligo-template mixture;

5X Reaction Buffer	4 µl
RiboLock RNase Inhibitor (20 U/µl)	1 µl
10 mM dNTP Mix	2 µl
RevertAid M-MuLV RT (200 U/µl)	1 µl

After mixing oligo-template mixture with enzyme containing mixture, samples incubated at 42°C for 60 minutes and then for 5 minutes at 70°C.

cDNA conversion control

Total cDNA converted from 500 ng of total RNA was used as PCR template for cDNA conversion control using GAPDH primers. After PCR was completed, 5 µl of the reaction run on a 1% agarose gel.

RT-qPCR reaction and analysis

RT-qPCR reactions were done with SsoAdvanced™ Universal SYBR® Green Supermix (SsoAdvanced™ Universal SYBR® Green Supermix, BioRad, USA, 172-5272) kit using cDNA made from 500 ng total RNA and CXXC5 qPCR forward and

reverse primers. For the normalization of data, 60S acidic ribosomal protein P0 (RPLP0) was used since it has been used for qPCR normalization in breast carcinoma cell lines [36]. For standard curve, 1:10, 1:20, 1:40, 1:80 dilutions were used and the samples used in 1:20 dilution.

Reaction was done according to Table 11 and Table 12

Table 11 RT-qPCR reaction mixture

SsoAdvanced™ Universal SYBR® Green Supermix (2x)	10 µl
Forward primer	1.2 µl
Reverse primer	1.2 µl
Template	4 µl
Nuclease free water	3.6 µl

Table 12 PCR conditions using the thermal cycler CFX Connect™ Real-Time PCR Detection System

Polymerase activation and initial denaturation	95°C	10’	
Denaturation	94°C	30”	
Annealing	60°C	30”	X40
Extension and plate reading	72°C	30”	
Rapid heating	95°C	10’	
Melt curve analysis	55-95°C increment 1°C	5”	
Final hold	4°C	∞	

To calculate relative gene expression from qPCR data, ΔΔCq calculation method was used (CFX Manager™ Software Gene Expression Analysis, BioRad, USA) and MIQE guidelines were followed.

Statistical Analysis

All experiments were conducted with at least three biological replicates. Statistical analyses and graphs were done using GraphPad Prism 7 Software (GraphPad Software Inc., USA). One-way ANOVA and Dunnett’s test was employed to determine significance minmum at $p<0.05$.

CHAPTER 3

RESULTS AND DISCUSSION

Since proteins in cells exert their functions within the context of dynamically interacting protein networks in a spatiotemporal manner, the identification of protein partners of CXXC5 would be an important prelude for the understanding of CXXC5 functions in E2-mediated cellular events. To accomplish this, we initially decided to use a yeast-two-hybrid system (Y2H). Y2H, also known as yeast two-hybrid screening, is a technique applied to uncover protein-protein interactions. The principle behind the technique is the activation of downstream reporter genes by the binding of a transcription factor onto an upstream activating sequence (UAS). For the two-hybrid screening, due to modular nature of the transcription factor, distinct DNA binding domain (DBD) and transactivation domain (AD) modules can separately be cloned into expression vectors. In the Y2H system, domains of the yeast transcription activator protein Gal4 are generally utilized. The DBD of Gal4 protein in the Y2H is the domain responsible for binding to the UAS and the AD of Gal4 is the domain responsible for the activation of transcription. The Y2H is thus a protein-fragment complementation assay that requires further verifications to verify the resultant protein-protein interaction using variety of approaches.

We utilized Creative BioLabs Service's (New Jersey, USA), Y2H screening services for the identification of putative binding partners of CXXC5. The CXXC5 cDNA we provided to the company was inserted into the bait vector pDBLeu containing DNA binding domain. To test toxicity of CXXC5 or the ability of CXXC5 to self-activate

the reporter gene in yeast, the transformation was carried out in a yeast strain, AH109 which is a MATa yeast strain with the *HIS3, ADE2,* and *lacZ* reporter genes are under the control of GAL4-responsive UAS; if the bait can auto-activate the reporter then medium should contain suitable concentration of 3AT (which is a competitive inhibitor of His3 and prevent autoactivation). Based on the results that CXXC5 is not toxic or cannot self-activate the reporter gene in AH109, the cDNA library constructions using total RNA of MCF7 cell line were carried out. MCF7 library cDNAs were then inserted into the activation domain containing prey pPC86Trp vector in yeast. Transformants were then plated on the appropriate selective media and harvested. Clones that grow under selection were picked up from the primary screening plates, restreaked onto new selection plates and assayed for activity of the reporter gene *His, Ade* and *LacZ*. The restreaking combined with the lacZ assay guarantees that only interaction pairs that reliably activate reporter genes. Clones that survived the primary selection and/or lacZ tests were picked and the library plasmid is isolated and amplified in *E. coli*. To determine whether the potential positives were specific for the bait used, the positive clones tested against CXXC5 bait together with individual pray plasmid were transformed back into the yeast strain MaV203 (MATα yeast strain containing *HIS3*, *URA3,* and *lacZ* reporter genes are under UAS) for interaction readout on synthetic complex media lacking of Leu (indicates presence of bait plasmid), Trp (indicates presence of prey plasmid), His, Ura (reporters in the case of activation) with 10mM 3-AT(to prevent auto-activation). Once a positive candidate was obtained, then the target plasmid DNA was sequenced. A database search was subsequently performed to determine the origin of the target sequence [37].

Based on the sequence information from positive clones, Cold Inducible RNA Binding Protein (CIRBP), Poly(RC) Binding Protein 2 (PCBP2), Ribosomal Protein L15 (RPL15), Ribosomal Protein S2 (RPS2) and Ribosomal Protein S4, X-Linked (RPS4X) were reported to us as possible CXXC5 interactors.

In an attempt to identify the possible interactors for another estrogen responsive gene protein product YPEL2, we also used Creative BioLabs Service's Y2H screening

services. Results reported to us as possible YPEL2 interactors also included RNA binding and ribosomal proteins suggested that the putative CXXC5 interactors obtained from Y2H may not be CXXC5 specific. We nevertheless decided to carry out an initial screening study in mammalian cells using a mammalian-two-hybrid (M2H) approach. If we were to identify a CXXC5 interactor, we would continue with co-localization and co-immunoprecipitation studies to further verify these interactions in mammalian cells.

Detection of possible protein-protein interaction with mammalian two hybrid system

Derived from Y2H approach, M2H utilizes expression vectors that contain promoters specific to mammalian cells that derive the expression of reporter enzymes as experimental outputs. In this dual luciferase system, there are two reporter vectors: 1) Firefly Luciferase Reporter Vector (p3xGal4RE-Luc), which contains three juxtaposed GAL4 Response Element upstream of a TATA box that derives the expression of the firefly luciferase cDNA as the reporter enzyme; 2) *Renilla* luciferase Reporter Vector (pRLuc), which contains a moderately strong SV40 promoter that derives the expression of the *Renilla* luciferase cDNA for normalization of transfection efficiency. Firefly luciferase vector requires the binding of GAL4-DBD bearing protein to drive the expression of the Firefly luciferase cDNA (Figure 7 – 1); whereas, the SV40 promoter of the *Renilla* luciferase vector is constitutively active in mammalian cells due to the promoter sequence that bears many binding sites for ubiquitous mammalian transcription factors (Figure 7 – 2).

As in yeast, the basis of mammalian two-hybrid systems is the modularity of functional domains, i.e. isolated structural regions exhibit unique functional features, in most of transcription factors. In the M2H system we are using, GAL4 DNA-BD Cloning Vector (pMGal4DBD) contains sequences encoding the yeast GAL4 DNA-binding domain present at the upstream of a multiple cloning site (MCS). The pVP16-AD

Cloning Vector (pVP16AD) contains sequences encoding the strong herpes simplex virus VP16 activation domain present at the upstream of a MCS [38]. Sequences encoding the two potentially interactive proteins of interest are cloned into pGal4DBD and pVP16AD vectors to generate fusion proteins with the DNA-binding domain of GAL4 and the activation domain of VP16, respectively. The pGAL4DBD and pVP16AD fusion constructs are transfected along with p3xGal4RE-Luc and pRLuc vectors into mammalian cells. Twenty-four hours after transfection, transfected cells are lysed, and the amounts of firefly and *Renilla* luciferase levels are quantified using the dual luciferase reading. Interaction between the two test proteins, as GAL4 and VP16 fusion constructs, results in an increase in relative firefly/*Renilla* luciferase enzyme levels, which are then taken as an indication for transcription activation compared to either constructed transfected alone as the negative control.

To induce firefly luciferase gene transcription and translation, GAL4 DNA Binding domain should bind to the GAL4 region of the reporter plasmid and the AD should be in the close proximity to be able to activate transcription (Figure 7 – 3). The schematic representation describes that if X protein fused to Gal4DBD that bound to Gal4RE on the promoter of the firefly luciferase reporter vector interacts with Y protein fused to the AD, this interaction will bring the AD to the promoter to drive the expression of the firefly luciferase cDNA. Since the promoter of *Renilla* luciferase reporter vector is always on; the *Renilla* luciferase enzyme levels are taken as an indication for transfection efficiency. The ratio of firefly luciferase/*Renilla* luciferase levels would then be considered as possible interactions between X and Y proteins (Figure 7 – 4&5).

As the positive control, we used the expression vector bearing sequences that encode Gal4DBD-VP16AD fusion protein as well as the Gal4DBD and the amino-terminus (N terminus) of ERα fusion and the VP16 fused to the carboxyl-terminus (C-terminus) ERα (Figure 7 – 6). Since N and C termini of ERα interacts in the presence of E2 to induce transcription [39], this interaction also provides a positive control for us in which DBD and AD domains coming from separate vectors.

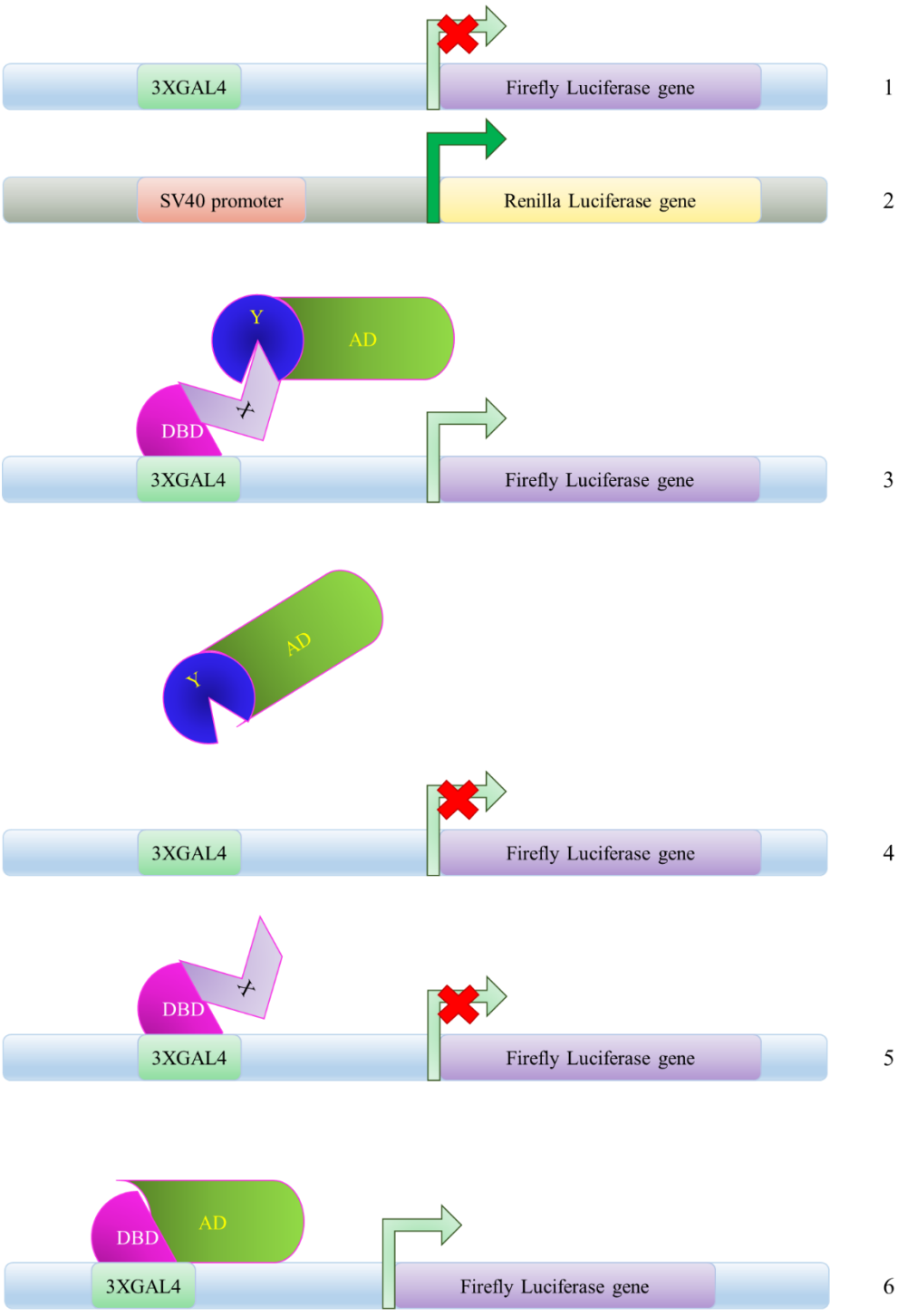

Figure 7 Reporter plasmids and working principle of Dual Luciferase System
The bait protein is fused to DBD while the prey protein is fused to AD. If bait and prey interact, the reporter gene is activated.

We cloned the ORF of *CXXC5* to the MCS of the pGal4DBD vector to generate Gal4DBD-CXXC5 fusion protein. We also cloned the ORF of a putative CXXC5 interactor into the MCS of the pVP16AD vector to generate VP16-putative interactor fusion protein and vice versa. The expression vectors were then transiently transfected into MCF7 cells together with the luciferase reporter vectors. 24 hours after transfections, cell lysates were assayed for luciferase enzyme activities.

Results as the ratio of firefly luciferase/*Renilla* luciferase outputs were analyzed with one-way ANOVA. Significance calculated with respect to only *Renilla* luciferase reporter vector containing sample and results having R square of 0.9632.

Table 13 One-way ANOVA statistical analysis of M2H assay

Dunnett's multiple comparisons test (p value:0.05)	Mean Diff.	Summary	Adjusted P Value
Gal4 DBD EV	-1.067	ns	>0.9999
VP16 AD EV	-1.53	ns	>0.9999
Gal4 DBD CXXC5	-1.2	ns	>0.9999
VP16 AD CXXC5	-1.377	ns	>0.9999
Gal4 DBD EV + VP16 AD CXXC5	-2.19	ns	>0.9999
VP16 AD EV + Gal4 DBD CXXC5	-4.06	ns	>0.9999
Gal4 DBD CIRBP	-0.3883	ns	>0.9999
VP16 AD CIRBP	0.4257	ns	>0.9999
Gal4 DBD EV + VP16 AD CIRBP	-0.2907	ns	>0.9999
VP16 AD EV + Gal4 DBD CIRBP	0.4956	ns	>0.9999
Gal4 DBD CXXC5 + VP16 AD CIRBP	0.009065	ns	>0.9999
Gal4 DBD CIRBP + VP16 AD CXXC5	0.1402	ns	>0.9999
Gal4 DBD PCBP2	0.413	ns	>0.9999
VP16 AD PCBP2	-1.687	ns	>0.9999
Gal4 DBD EV + VP16 AD PCBP2	0.4119	ns	>0.9999
VP16 AD EV + Gal4 DBD PCBP2	-0.8552	ns	>0.9999
Gal4 DBD CXXC5 + VP16 AD PCBP2	0.1938	ns	>0.9999
Gal4 DBD PCBP2 + VP16 AD CXXC5	-0.07916	ns	>0.9999
Gal4 DBD RPL15	-0.7905	ns	>0.9999
VP16 AD RPL15	-0.8006	ns	>0.9999
Gal4 DBD EV + VP16 AD RPL15	-1.76	ns	>0.9999
VP16 AD EV + Gal4 DBD RPL15	-5.276	ns	>0.9999
Gal4 DBD CXXC5 + VP16 AD RPL15	-2.181	ns	>0.9999
Gal4 DBD RPL15 + VP16 AD CXXC5	-4.641	ns	>0.9999
Gal4 DBD RPS2	0.2349	ns	>0.9999
VP16 AD RPS2	-0.9784	ns	>0.9999
Gal4 DBD EV + VP16 AD RPS2	-5.94	ns	>0.9999
VP16 AD EV + Gal4 DBD RPS2	-5.429	ns	>0.9999
Gal4 DBD CXXC5 + VP16 AD RPS2	-5.316	ns	>0.9999
Gal4 DBD RPS2 + VP16 AD CXXC5	-4.884	ns	>0.9999
Gal4 DBD RPS4X	-3.287	ns	>0.9999
VP16 AD RPS4X	-3.238	ns	>0.9999
Gal4 DBD EV + VP16 AD RPS4X	-6.039	ns	>0.9999
VP16 AD EV + Gal4 DBD RPS4X	-4.122	ns	>0.9999
Gal4 DBD CXXC5 + VP16 AD RPS4X	-1.811	ns	>0.9999
Gal4 DBD RPS4X + VP16 AD CXXC5	-1.798	ns	>0.9999
Split ERα	-1494	****	<0.0001
Gal4VP16 fusion	-2883	****	<0.0001

MCF7 cells were seeded in 48-well tissue culture plates with 4 x 10^4 cells per well. 48 hours later, cells were transfected with the SV40 promoter driven *Renilla* luciferase vector (0.5 ng), reporter vector Gal4 response element containing firefly luciferase vector (125 ng) as well as DNA binding domain and activation domain containing vectors (75 ng each). 48 hours after transfection, cells were washed with PBS three times and lysed with 50 µl of passive lysis buffer provided by Dual-Luciferase kit (Dual-Luciferase® Reporter Assay System, Promega, USA, E1910). 25 µl of cell lysates was taken into a well of a white 96 well lumitrac microplate (Greiner LUMITRAC™ 200 microplate, Sigma, USA, GN655075) and luciferase reads were done with 25 µl of Luciferase Assay Reagent II, then 25 µl of Stop & Glo® using GloMax®-Multi Detection System one by one (GloMax®-Multi Detection System, Promega, USA, E7041). Reads were obtained as luciferase intensity values; these values were then analyzed with GraphPad Prism 7 Software (GraphPad Software Inc., USA) using One-way ANOVA and Dunnett's test to determine p value of <0.05.

Results were also graphically represented below;

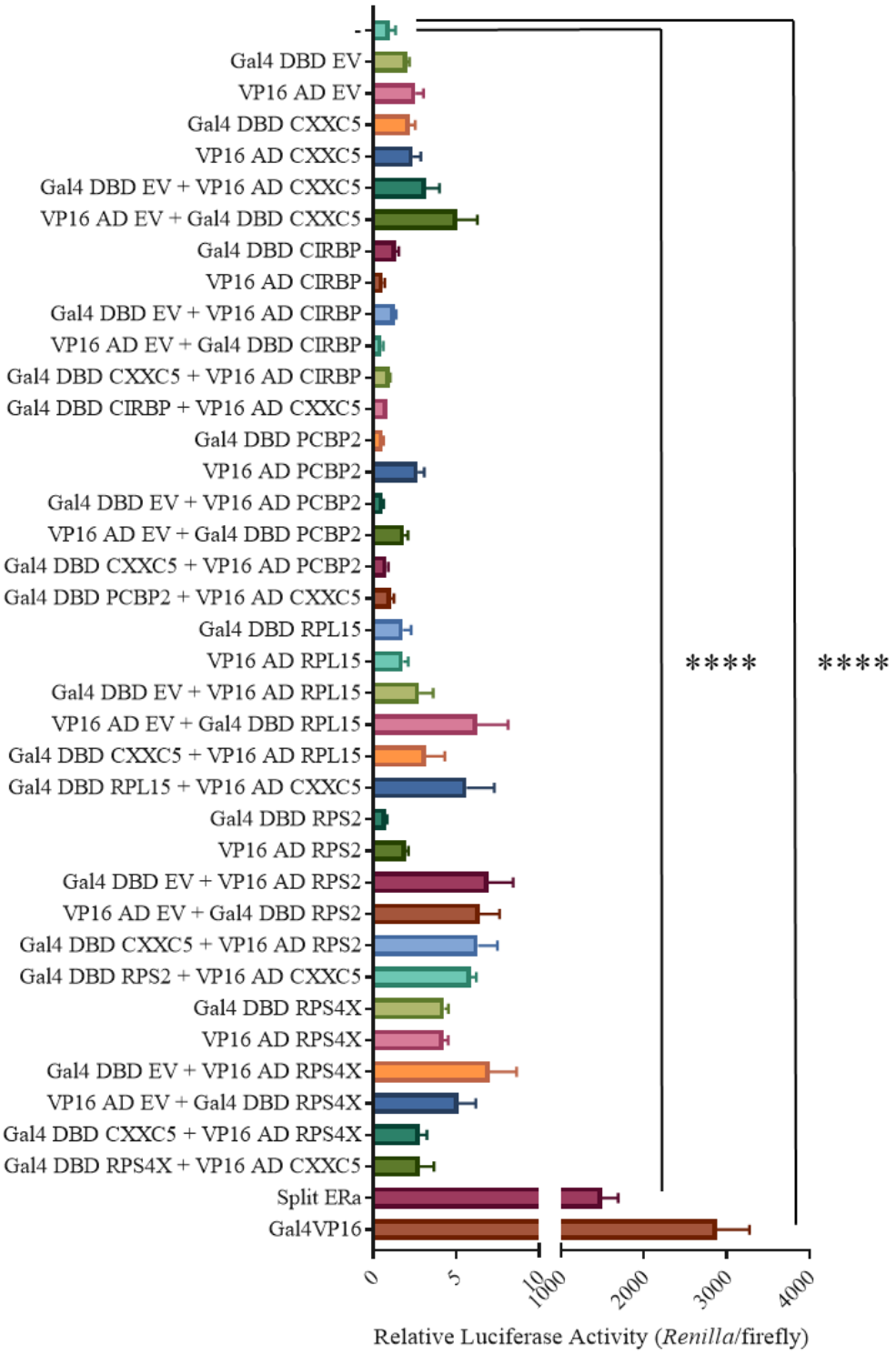

Figure 8 Graphical representation of Mammalian Two Hybrid results as Firefly/*Renilla* ratio.
One way ANOVA statistical analyses graphically represented with GraphPad.

We found that none of the possible interaction partners of CXXC5 identified with Y2H showed luciferase activity at levels that were significantly different from the basal level in M2H experiment. We, on the other hand and as expected, observed significant luciferase activity with split ERα-fusion proteins; since E2 treatment triggers interaction of the amino and carboxyl-terminus of ERα [39]. This was also the case for the Gal4DBD-VP16AD fusion protein which effectively enhanced luciferase activity.

Based on these results, we decided not to continue with the investigation of possible interactions of CXXC5 with putative protein partners identified with Y2H. But we carried out studies using the BioID approach, with which Gamze Ayaz, a doctoral candidate in our laboratory, had established in parallel with the Y2H approach. BioID has been used effectively to screen for physiologically relevant protein interactions in mammalian cells [40]. BioID utilizes a promiscuous biotin ligase to biotinylate proteins based on proximity. This is accomplished by genetically fusing the biotin ligase to a protein of interest and expressing in cells, where it biotinylates proximal endogenous proteins. High level of biotinylation of endogenous proteins by fusion proteins enables the identification of candidate interactors for the protein of interest by biotin-streptavidin capture approaches followed by mass spectrometry analyses. Based on the BioID analysis, Gamze Ayaz, to whom I also offered a hand in experiments, has identified a number of CXXC5 interactors in mammalian cells and verified some these interactors as *bona fide* protein partners of CXXC5 by using co-immunoprecipitation and co-localization studies.

It should be noted that none of the interactor partner reported with Y2H was observed with BioID system as putative CXXC5 interactor. These findings also iterate our conclusion that the reported interactors of CXXC5 by the use of Y2H are not specific for CXXC5.

Silencing of CXXC5 in MCF7 cells

Since CXXC5 is an E2-ERα responsive gene, one of my thesis project aims was to assess the functional importance of CXXC5 in mammalian cells that are responsive to E2. We therefore used E2-responsive and ERα-positive MCF7 cell line derived from breast adenocarcinoma. This was based on our previous studies with which MCF7 cells were used as the cell model aimed at the initial functional characterization of CXXC5. MCF7 cells were kindly provided by Assoc. Prof. A. Elif Erson-Bensan (Middle East Technical University, Ankara, Turkey).

To analyze the effects of CXXC5 on cellular processes that include proliferation and death, we initially decided to use an shRNA approach. Previous studies in our laboratory effectively used transient transfections of siRNAs specifically designed to decrease CXXC5 protein levels in MCF7 cells by interacting with the CXXC5 mRNA. Based on these findings, we decided to use CXXC5 siRNA#10, with which we had shown that it effectively decreases endogenous CXXC5 mRNA and protein levels in MCF7 cells [11]. The double stranded shRNA obtained from the PRZ Biotech (PRZ biotech, Turkey) was then inserted into an shRNA vector. We used pSUPER Retro-Neo-GFP for cloning. This vector utilizes polymerase-III H1 gene promoter, which transcribes small RNA, for the expression of CXXC5 shRNA#10 for stable silencing of CXXC5. However, in several attempts, we failed to generate stable clones, as none of clones survived after antibiotic selection. The reason behind these unsuccessful trials to generate MCF7 clones within which CXXC5 is stably silenced is not clear, but could be due to the critical importance of CXXC5 in cellular processes. That our previous studies to generate clones that stably synthesize Flag-CXXC5 had also failed are supportive of this prediction. Therefore, we decided to use an inducible shRNA expression utilizing the versatile a TET inducible lentiviral expression system: pINDUCER [35].

While overexpressing, or repressing, a gene of interest can provide an avenue to assess its role in cells, it may be necessary to control the expression of the gene due to its

intrinsic toxicity and/or importance in cell functions. This necessitates the timing and level of expression of a protein to be studies. There are many inducible expression systems, including the widely used LacSwitch system which utilizes the Lac repressor and isopropyl ß-D-thiogalactopyranoside (IPTG) as an inducible means of inhibiting the Lac repressor. Due to the ease in use, the recently developed tetracycline inducible systems provide a tightly controlled, rapid and reversible gene expressions to allow the study of protein functions in cells. But it should be noted that the tetracycline (Tet) system has also problems: The system often times gives rise to a "leaky expression" of the gene of interest even without the addition of tetracycline. Means of reducing this leakiness include the use of stronger tet-activators/repressors.

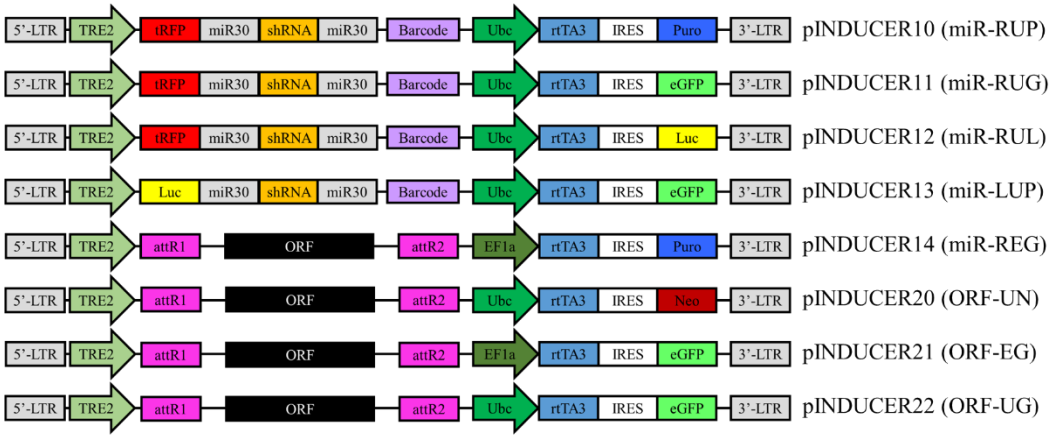

Figure 9 Diagrams of the pINDUCER vector series.
The schematics is adapted from Meerbrey et al.2011 [35].

Tet systems rely on two elements: 1) A transactivator protein, and 2) consensus binding sites for the transactivator protein that drive the expression of a reporter gene. The "Tet-On" system of pINDUCER is based on the conformational changes in the strong transactivator (reverse tetracycline-controlled transactivator 3, rtTA3) upon binding to tetracycline (or stable tet derivatives such as doxycycline, Dox). This conformational changes in rtTA3 exposes the DBD of the activator to bind to TRE sequences, which are composed of 7 repeats of a 19 nucleotide tetracycline operator (TetO) sequence, and turn on the expression of gene of interest.

The pINDUCER lentivirus-based system encodes rtTA3 and an interchangeable marker as a constitutive, bicistronic transcript. In this system, shRNA sequence is embedded in a precursor miRNA context. Endogenous miRNAs are transcribed predominantly by polymerase II (Pol II) promoters as long primary miRNA transcripts (pri-miRNAs). pri-miRNAs are subsequently processed into precursor miRNAs (pre-miRNAs) and eventually incorporated into the RNA-induced silencing complex to mediate silencing of the target gene. In the pINDUCER series of lentiviral vectors (depicted in Figure 9), miRNA motifs (miRNA30) are fused to the end of a sequence encoding the turbo red fluorescent protein (tRFP), giving rise to a transcript which encodes both the reporter protein and pre-miRNA. As this is a Pol II- driven system, it can be controlled by a tet-responsive promoter element (TRE2). The synthesis of turbo-red fluorescent protein (tRFP) is then taken as an indication of shRNA expression. In addition, these lentiviral vector systems contain the constitutively active Ubiquitin C (Ubc) or elongation factor-1 alpha (EF1a) promoter that drives the expression of another selection marker: an enhanced Green fluorescent protein (eGFP), puromycin (Puro) or Luciferase enzyme (Luc) cDNA [35] .

Pseudo Viral Particle Production

pINDUCER11 lentiviral vector is a single-vector system containing all regulatory cassette in one vector containing fluorescent reporters to detect infected cells with green fluorescent protein expression (GFP) and shRNA expressing cells with tRFP.

pINDUCERs have a cassette containing Long terminal repeats (LTRs) found both side of the lentiviral provirus cassette. A TRE is a 7 repeats of a 19 nucleotide tetracycline operator (tetO) sequence, and is recognized by the reverse tetracycline-controlled transactivator (rtTA3). The 3' LTR terminates the transcript by adding a poly A tract.

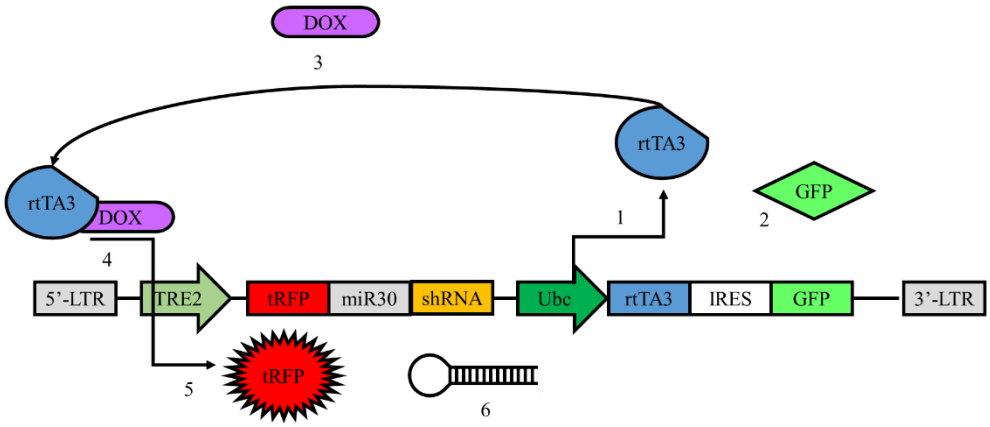

Figure 10 Working principle of pINDUCER11 vector system.
Ubiquitin C promoter (Ubc) provides the expression of rtTA3 (Figure 10-1) and the marker, for example GFP, through an internal ribosome entry site (IRES) (Figure 10-2). In addition to the ease of selecting clones with fluorescent microscope, the GFP expression could also be used for selection of infected cells with fluorescence-activated cell sorting (FACS). In the presence of Tet or Dox , (Figure 10-3), rtTA3 binds to and activates gene expression from TREs (Figure 10-4). When the TRE driven promoter is activated, it drives the expression of the tRFP (Figure 10-5) and also shRNA (Figure 10-6).

Initially, to test whether or not we could produce recombinant lentiviral particles and infect target cell with them, we used pINDUCER11. We cloned CXXC5 shRNA#10 into pINDUCER11 using XhoI and MluI restriction enzyme sites. The pINDUCER11 CXXC5 shRNA#10 vector together with packaging plasmids pMD2G, pRVS-REV, pMDL/pRRE was transiently transfected into HEK293T cells, which contain T antigen of SV40 driving the expression of viral packaging proteins expressed from the packaging plasmids to generate pseudo particles.

In the following image, HEK293T were shown 24 hours later transfection with pINDUCER11 and packaging plasmids to observe transfection efficiency;

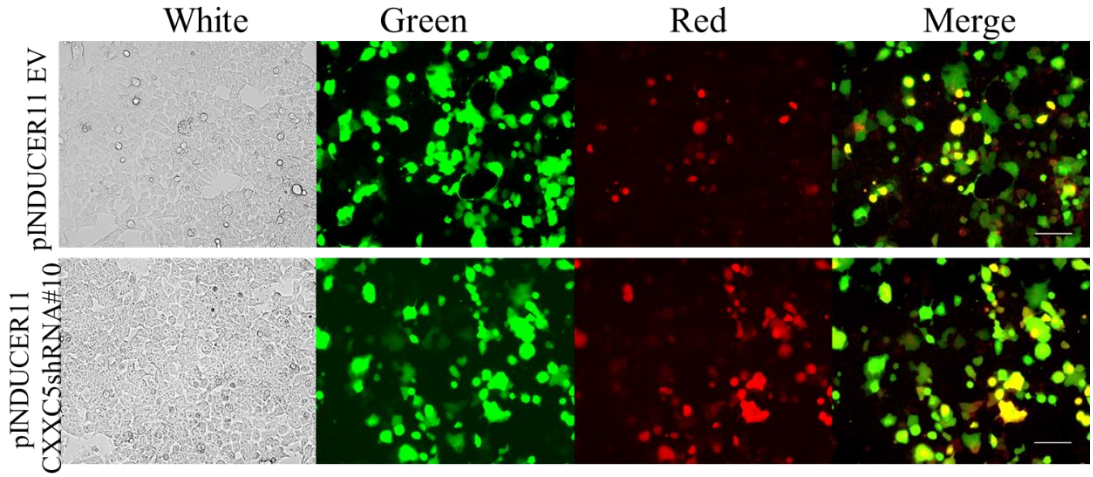

Figure 11 Transfection with pINDUCER11 and packaging elements
2 x 10^5 HEK293T cells were seeded per well in a 6-well tissue culture plate for pseudo viral production. 48 hours later, cells were transfected with 2000 ng of pINDUCER11, together with packaging plasmids pMD2G (900 ng), pRVS-REV (600 ng) and pMDL/pRRE (1500 ng). The microscopic image was taken 24 hours after transfection to observe transfection efficiency as well as shRNA expression with FLoid™ Cell Imaging Station (FLoid™ Cell Imaging Station, Thermo, USA, 4471136). Scale bar represents 100µm. In green channel eGFP is visualized indicating Ubc promoter is active and red channel represents synthesized tRFP indicating system can be induced in the presence of Dox and also TRE promoter can be activated.

Since medium was containing tetracycline, the induction of tRFP synthesis were observable; while the GFP synthesis revealed that most of HEK293T cells were successfully transfected with pINDUCER11. This result indicated that the transfected inducible lentiviral vector do synthesize the reporter markers as well. But at this stage, we do not know whether or not we have produced pseudo viral particles useful for infection of cell models. To assess this, we carried out initial infection studies in MCF7 cells.

Infection

Silencing of CXXC5 with inducible system was the primary step of our functional analysis for understanding the role of CXXC5 in cellular processes. To achieve this objective, produced pseudo viral particles in transiently transfected HEK293T cells were used for the infection of MCF7 cells. When pINDUCER11 containing pseudo viral particles were introduced to MCF7 cells in medium that contains tetracycline-free FBS to minimize the effect of endogenous tetracycline on the expression of reporter proteins.

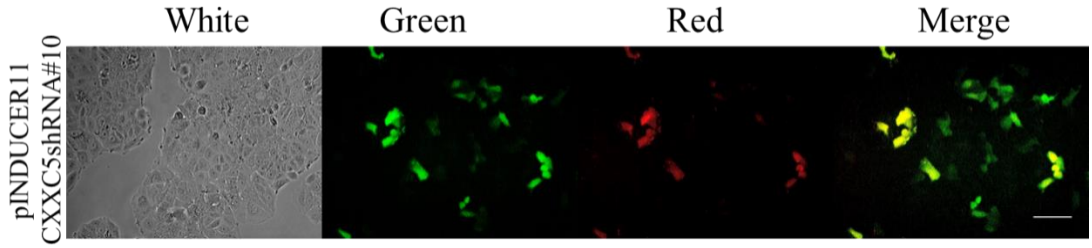

Figure 12 pINDUCER11 CXXC5 shRNA#10 infected MCF7
MCF7 cells were seeded in 48-well tissue culture plates as 2 x 10^4 cells per well. 1:1 ratio of pseudo viral particle containing medium and fresh DMEM containing TET-free FBS was mixed in the presence of proteamine sulfate at a final concentration of 8 µg/ml. The mixture was then introduced to MCF7 cells. 24 hours after, the infection medium was changed with fresh medium containing 1µg/ml dox. 48 hours later, cells were imaged using FLoid™ Cell Imaging Station (FLoid™ Cell Imaging Station, Thermo, USA, 4471136) to observe fluorescent activity. Scale bar represents 100µm. Green channel indicates MCF7 cells can be infected with produced virus and Ubc promoter can synthetize rtTA3 and GFP as separate proteins. Moreover in the presence of Dox; RFP fluorescence indicates TRE promoter is inducible and may give rise to shRNA.

We observed some cells synthesized GFP, indicating that we produced viral particles capable of infecting MCF7 cells. Furthermore, the addition of Dox to the medium also led to the observation of red staining indicating the synthesis tRFP in some of the infected cells. This suggested that infected viral particle has the potential to drive the expression of shRNA in response to Dox. Based on these findings, we carried out a set of studies to find out the optimal dox concentration, which was 1000ng/ml and used in subsequent studies (APPENDIX K).

Generation of stable cell lines

To generate stable cell lines, we used pINDUCER10, which differs only from pINDUCER11 by expressing puromycin N-acetyl-transferase (*PAC*) gene that provide selection against antibiotic puromycin instead of green fluorescent protein.

Induction control of CXXC5 shRNA#10 colonies

Infection of MCF7 cells was done using produced pseudo particles generated with the transfer vector of pINDUCER10.

For the generation of stable MCF7 cell line that expresses CXXC5 shRNA#10 for ablating/reducing the endogenous CXXC5 synthesis in response to Dox, MCF7 cells were seeded in 6-well tissue plates as 1 x 10^5 cells per well for 48 hours. Then, cells were introduced with 1:1 ratio of pseudo viral particles produced and fresh DMEM medium with FBS together with proteamine sulfate at the final concentration of 8 µg/ml. 72 hours later infection, the spend medium was changed with fresh DMEM medium containing TET-free FBS with 1µg/ml puromycin. The concentration of puromycin was based on preliminary studies in which we observed that 1µg/ml puromycin was the maximal concentration that led to an effective cell death within one week of the antibiotic introduction (APPENDIX L).

After one week in the presence of 1µg/ml puromycin, individual clones were observable. Within three weeks of colony formation, individual colonies were picked with 1 ml pipette tips and placed into a well of 48-well tissue culture plates. Cells were then grown for near confluency. Cells were trypsinized, re-suspended in DMEM medium supplemented with TET-free FBS containing 1µg/ml puromycin and divided into two wells in a 48-well tissue culture plate for maintenance and for Dox induction to select positive clones.

Clones expressing tRFP observed with fluorescent microscope were selected for the verification of shRNA expression.

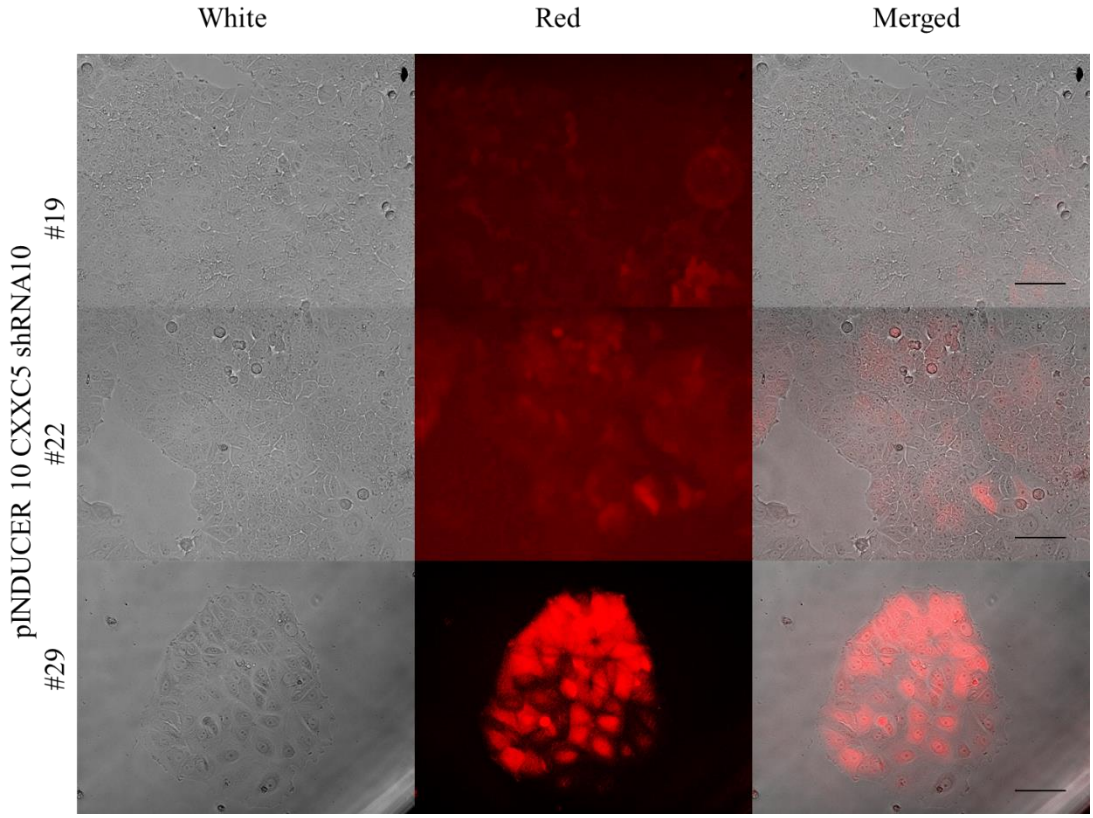

Figure 13 Induction control of pINDUCER10 CXXC5-shRNA#10 colony 19, 22, 29

After colony selection with puromycin, possible positive clones were seeded in 48-well tissue culture plates as 2 x 10^4 cells per well with DMEM containing TET-free FBS to induce cells and detect tRFP synthesis. 24 hours after seeding, medium was changed to fresh medium supplemented with TET-free medium containing 0.1% DMSO (as vehicle control) or 1µg/ml doxycycline. 48 hours after induction, cells were imaged using FLoid™ Cell Imaging Station (FLoid™ Cell Imaging Station, Thermo, USA, 4471136) to observe the tRFP activity indicating that cells contained all regulatory elements required for shRNA expression. Microscopy images of puromycin selected colonies 19, 22 and 29 of CXXC5-shRNA#10 are shown. Scale bar represents 100µm. Since colonies are surviving in the puromycin containing media the Ubc promoter is active and producing rtTA3 and also puromycin resistance. Presence of RFP in response to Dox indicates all regulatory cassette for production of shRNA is in the genome.

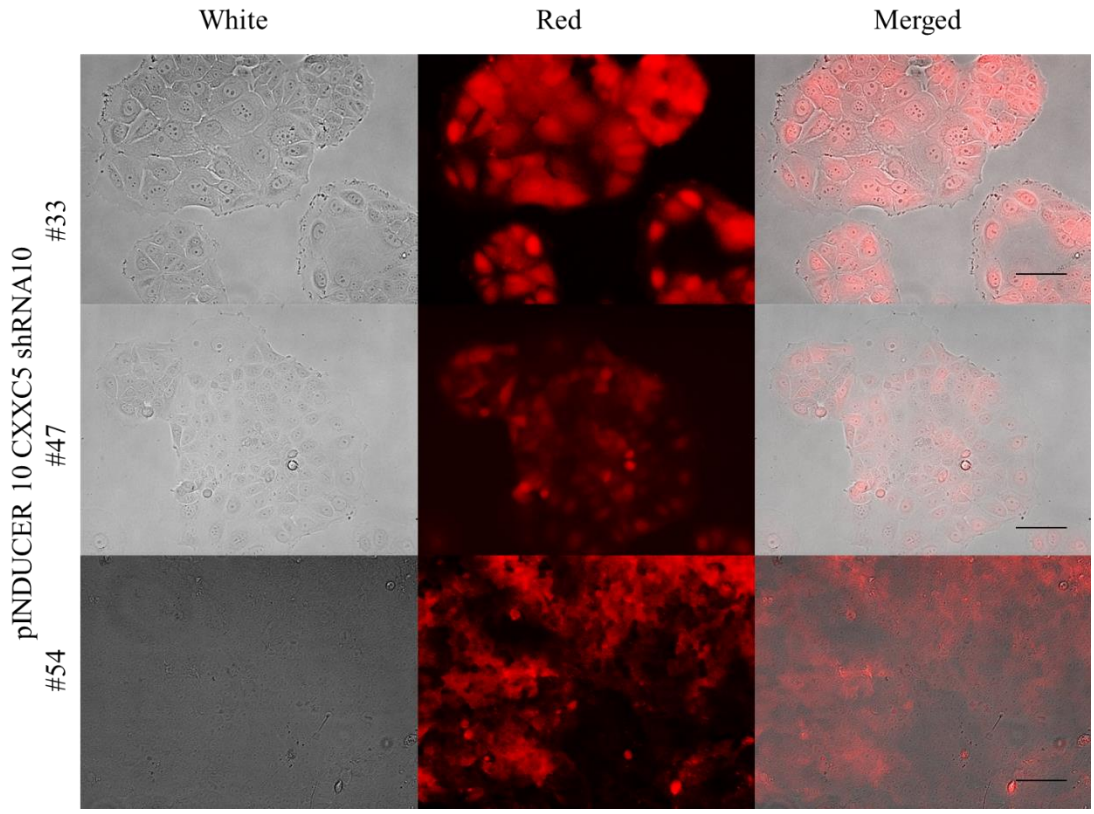

Figure 14 Induction control of pINDUCER10 CXXC5-shRNA#10 colony 33, 47, 54

Microscopy images of puromycin selected colonies 33, 47 and 54 of CXXC5shRNA #10. 48 hours after the induction of expression with 1µg/ml puromycin is shown with the scale bar of 100µm.

In a number of the puromycin resistant clones we did not observe tRFP synthesis, suggesting that some of the regulatory elements required for tRFP synthesis is not present. Clones that did not synthesize tRFP were then discarded.

Induction control of NTshRNA colonies

Non targeting shRNAs were required as the control group to ensure that physiological and phenotypical changes occurred in stable cell lines were the result of the shRNA specifically target CXXC5 and not the infection or the insertion of the viral regulatory into the genome.

For the generation of MCF7 cell lines with inducible non targeting shRNA expression, we used the same procedure that we detailed for the generation of cell line infected with pINDUCER CXXC5-shRNA#10.

Positive clones of pINDUCER10 NTshRNA that are expressing tRFP were selected for growth and maintenance.

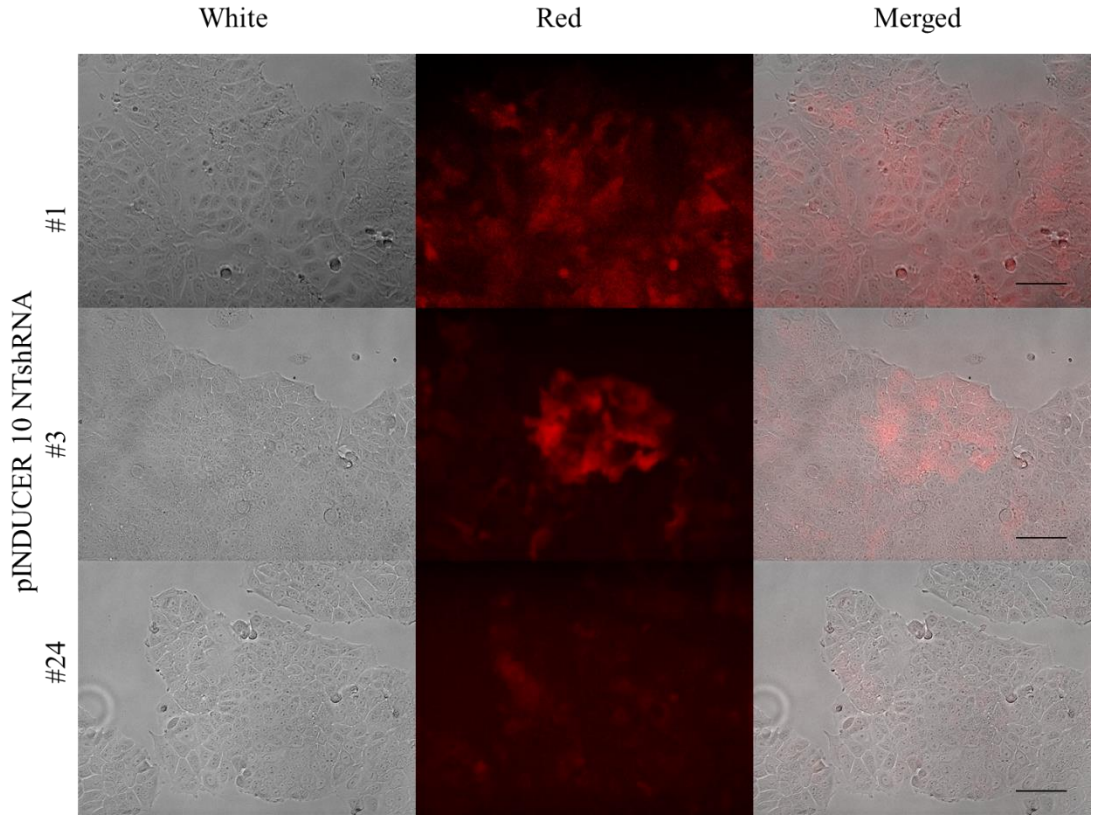

Figure 15 Induction control of pINDUCER10 NTshRNA colony 1, 3, 24
After puromycin selection, 2 x 10^4 cells were seeded in 48 well tissue culture plates with DMEM supplemented with TET-free FBS. 24 hours later, medium was changed to fresh DMEM with TET-free FBS containing 1µg/ml doxycycline or 0.1% DMSO (as vehicle control). 48 hours later, cells were imaged using FLoid™ Cell Imaging Station (FLoid™ Cell Imaging Station, Thermo, USA, 4471136) for tRFP synthesis. Microscopy images of puromycin selected colonies 1, 3 and 24 of non-targeting shRNA, 48 hours after the induction of expression with 1µg/ml puromycin are shown with the scale bar of 100µm. Since colonies used for induction control were puromycin resistant and producing RFP in the presence of Dox, they assumed to have all the components required for shRNA production.

Detecting CXXC5 as a nuclear protein

Consistent with our previous studies [11], we also observed that CXXC5 is a nuclear protein in MCF7 cells.

Cells grown for 48h were subjected to protein extractions using NE-PER protein isolation kit (Nuclear and Cytoplasmic Isolation Kit, Thermo Scientific, USA, 78833). As a positive control, we also used cellular extract transiently transfected with transfected Flag tagged-CXXC5. Cytoplasmic and nuclear extracts in the amount of 50 µg for the endogenous protein and 25 µg for the Flag-CXXC5 transfected cells were subjected to SDS-PAGE followed by western blot analysis using CXXC5-specific antibody. We used the detection of Histone Deacetylase 1 (HDAC1) with a HDAC1 antibody as the loading control for nuclear proteins [41]. β-Actin, which is a cytoskeletal protein, was used for cytoplasmic fraction verification [42].

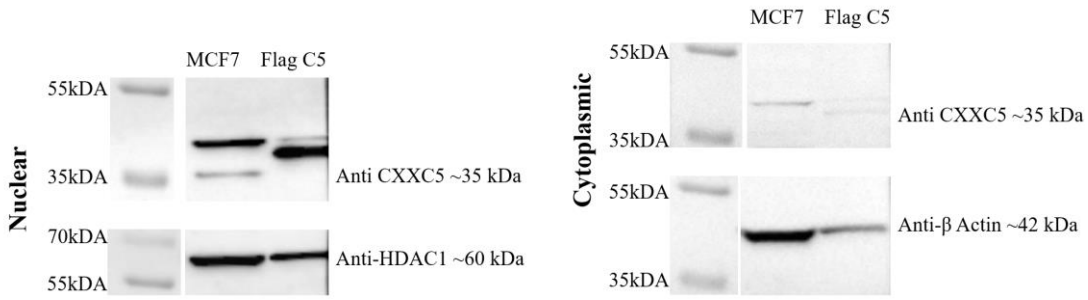

Figure 16 Nuclear vs cytoplasmic expression of CXXC5 in MCF7 cells
50 µg nuclear extracts of MCF7 and 25 µg of Flag tagged CXXC5 overexpressed in MCF7 were subjected to western blot analysis for detection of CXXC5. Primary antibodies: Anti-CXXC5 1:500 (Anti-CXXC5 antibody, Abcam, UK, ab106533), Anti-HDAC1 (Anti-HDAC1 antibody, Abcam, UK, ab19845), Anti-β Actin (Anti-beta Actin antibody, Abcam, UK, ab8227). Secondary antibody was anti rabbit HRP conjugated 1:4000 (Goat-anti-rabbit IgG (H+L), HRP conjugate antibody, Advansta, USA, R-05072-500).

Testing the stable cell lines with western blot analysis

To assess the effects of shRNA on CXXC5 synthesis, stable cell lines were generated with viral infection of MCF7; in the presence of puromycin as a selection marker. Surviving colonies then were subjected to induction with Dox or DMSO (vehicle control). In the presence of Dox, clones expressing red fluorescent protein were selected as candidates for possible CXXC5 shRNA expressing cell in an inducible manner.

Clones that were capable of synthesizing tRFP in response to Dox were seeded into 6-well tissue culture plates as 1 x 10^5 cells per well. 24 hours later, Dox induction was started with 1 µg/ml Dox or % 0.1 DMSO, which was used as vehicle control. Cells were maintained for 96h with the change of medium on the second day. Cells treated with Dox or DMSO were collected and subjected to nuclear protein isolation with NE-PER kit. Nuclear extracts were then subjected to SDS-PAGE followed by WB using the ab106533 antibody specific to CXXC5.

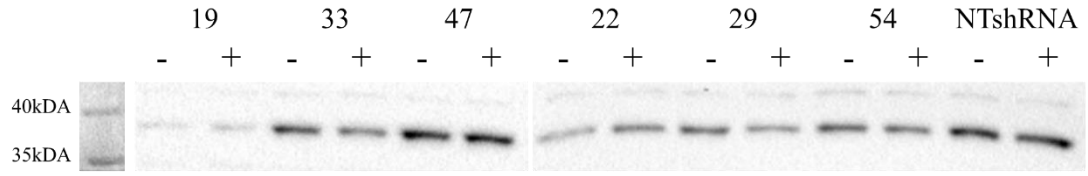

Figure 17 CXXC5 expression in nuclear extracts of stable cell lines generated with MCF7
50 µg of nuclear protein extracts of selected clones in the absence (-) or presence of Dox for 96h were subjected to western blot analysis. CXXC5 synthesis was assessed using a CXXC5 antibody (CXXC5 antibody, Abcam, UK, ab106533) at 1:500 dilution followed by goat anti-rabbit HRP conjugated antibody (Advansta, USA, R-05072-500 at 1:4000 dilution.

Analysis of WB results with ImageLab (Biorad) and statistical analysis with GraphPad (Tukey's multiple comparison test with p value of 0.05) revealed that Dox had no significant effect on CXXC5 levels compared to cells treated with vehicle control (APPENDIX M). This suggests that 1 µg/ml Dox for 96h does not affect CXXC5 levels, despite the fact that tRFP is synthesized.

Dose response experiments with stable cell lines

Our various attempts to silence CXXC5 synthesis using different tRFP synthesizing clones also failed. Although the amount of Dox concentration we used was dependent upon our preliminary studies with which we observed tRFP synthesis without an adverse effect on cell viability, it was possible that clones we selected required different concentration of Dox and different duration of Dox exposure rather than 1 µg/ml for 96h as we used.

Available information suggests that 2µg/ml is the maximal dose used for induction of human breast cancer cell lines. Therefore, we used increasing doses of doxycycline up to 2000 ng/ml.

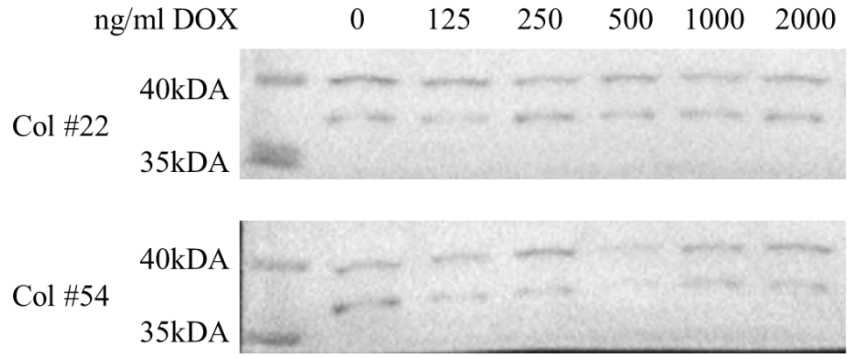

Figure 18 Dose response experiment for colonies; 22 and 54
Cells shown are the clone #22 and #54 were seeded in to 6 well tissue culture plates as 1 x 10^5 cell per well and induced with doxycycline for 96 hours with 125, 250, 500, 1000 and 2000 ng/ml. 50 µg of nuclear protein were loaded in the gel for each sample. Western blot analysis was done using the CXXC5 antibody. Bands that were close to 35 kDa corresponded to the CXXC5 protein.

Results based on WB image analysis with ImageLab software revealed that varying concentrations of Dox had no significant effects on CXXC5 levels.

Time course experiments with stable cell lines

To assess the effect of duration on the induction we conducted a time course experiment using 1µg/ml doxycycline.

For this experiment; clones, for example #22 and #54 were seeded in 6 well tissue culture plates and Dox induction was started 24 hours after seeding. Then at 24, 48 72 and 96h; cells were collected and nuclear proteins were extracted.

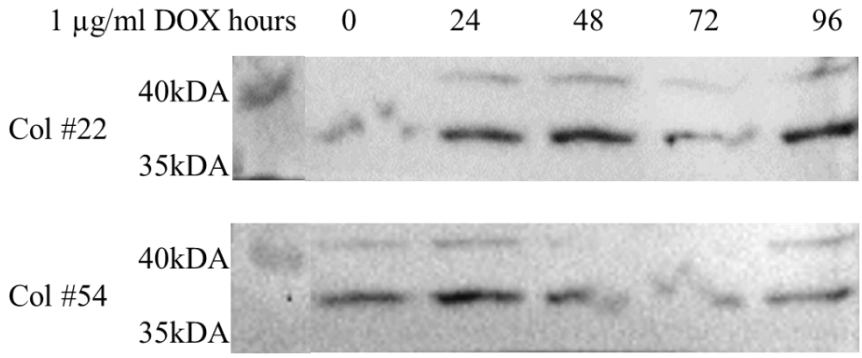

Figure 19 Time course experiment for clones; #22 and #54
Clone #22 and #54 were seeded in to 6 well tissue culture plates and induced with 1 µg/ml doxycycline for 24,48, 72 and 96 hours. Each sample were loaded to gel as 50 µg of nuclear protein. The CXXC5 antibody was used for detection of CXXC5 protein which corresponds close to 35 kDa.

Results revealed that there are no significant effects of the duration of Dox treatment on CXXC5 levels in cell clones.

Testing the stable cell lines with RT-qPCR analysis

Since western blot results did not answer the question as to why CXXC5 synthesis was not silenced in response to Dox although cells survived to puromycin selection and also synthesize tRFP, we decided to investigate mRNA expression of CXXC5 with RT-qPCR.

Cells, for example from Clone#29, were seeded in the same way with time course experiment and samples were collected at 24, 48, 72 and 96 hours after Dox induction. Collected cells were washed with PBS and Subjected to RNA isolation was done using Zymo research Quick-RNA™ MiniPrep kit. Isolated RNAs were tested for genomic DNA contamination and if they were free from gDNA; they were subjected to cDNA conversion. cDNA was then used as the template in RT-qPCR analysis to observe if there was any change in the CXXC5 RNA levels in response to Dox.

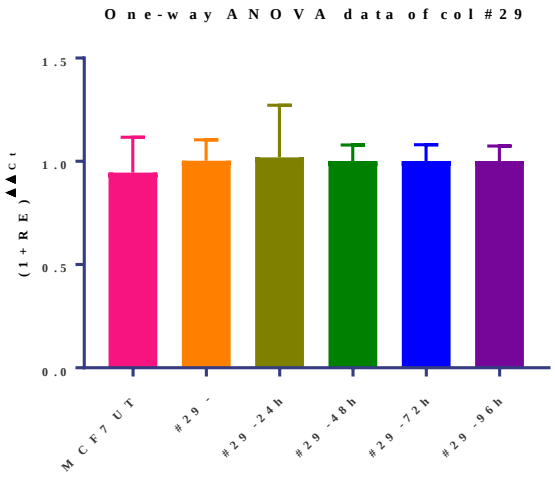

Figure 20 Graphical representation of RT-qPCR results of time course experiment of colony #29
cDNAs produced by 500 ng total RNA isolated from clones which were induced for indicated durations used for RT-qPCR. Change of expression in CXXC5 levels was normalized with RPLP0 expression which is known to be used for normalization of qPCR data in breast cancer cell lines. ΔΔCt calculations were done and the results used for statistical analysis.

Statistical analysis of the results was done with Dunnett's multiple comparison test with the p value of $p<0.05$. The results revealed that Dox induction had no significant effects on CXXC5 expression at any time point we tested.

Table 14 Graphical representation of RT-qPCR results and western blot results of time course experiment of colony #29

Dunnett's multiple comp. test	Mean Diff.	Summary	Adjusted P Value
MCF7 UT vs. #29 -	-0.05739	ns	0.9810
MCF7 UT vs. #29 -24h	-0.07324	ns	0.9497
MCF7 UT vs. #29 -48h	-0.05583	ns	0.9831
MCF7 UT vs. #29 -72h	-0.05599	ns	0.9828
MCF7 UT vs. #29 -96h	-0.05571	ns	0.9832

CHAPTER 4

CONCLUSION AND FUTURE DIRECTIONS

Cellular functions including architecture, metabolism and signaling are dependent upon transient and dynamic protein-protein interactions. The role of CXXC5 in cellular events remains largely unknown. Identification of interacting partners of a protein offers an important insight into the elucidation of the structure/function and the relative importance of the protein in cellular functions. Our initial attempts directed at understanding of the function of CXXC5 through identification of protein partners of CXXC5 with Y2H followed by an M2H approach yielded no positive results. This necessitated the use of an alternative approach. The recently introduced BioID approach has been used effectively to screen for physiologically relevant protein interactions in mammalian cells [43]. BioID utilizes a promiscuous biotin ligase to biotinylate proteins based on proximity. This is accomplished by genetically fusing the biotin ligase to a protein of interest and expressing in cells, where it biotinylates proximal endogenous proteins. High level of biotinylation of endogenous proteins by biotin ligase conjugated CXXC5 enabled the identification of candidate interactors by biotin-streptavidin capture approaches followed by mass spectrometry analyses. Results revealed that CXXC5 interacts with various transcription factors and epigenetic regulators. Although consistent with earlier predictions based on experimental studies that CXXC5 as a non-methylated CpG binding protein can act as a transcription modulator and/or epigenetic regulator in cells, these preliminary studies need to be extended to structural/functional domain analysis of CXXC5 for a better understanding of the protein function as well as to cellular context.

To begin to understand the functional features of CXXC5 in cellular context, we had carried out an extensive arrays of stable transfections for CXXC5 over-synthesis by CXXC5-cDNA expression or for the repression/prevention of CXXC5 synthesis by shRNA expression without a success. To circumvent this problem, we here wanted to control the synthesis of CXXC5 with an inducible expression system. To accomplish this, we used the pINDUCER lentivirus-based system. This system uses a tetracycline response element binding (TRE) of transcription activator (rtTA3) in response to tetracycline. The lentiviral vector also contains an interchangeable marker as a constitutive, bicistronic transcript. In this vector, shRNA sequence is embedded in a precursor miRNA context which are transcribed by Pol II promoters as long primary miRNA transcripts (pri-miRNAs). pri-miRNAs are then processed into precursor miRNAs (pre-miRNAs) and eventually incorporated into the RNA-induced silencing complex to mediate silencing of the target gene. In pINDUCER series of lentiviral vectors, miRNA motifs (miRNA30) are fused to the end of a sequence encoding the turbo red fluorescent protein (tRFP), giving rise to a transcript which encodes both the reporter gene protein and pre-miRNA. The synthesis of turbo-red fluorescent protein (tRFP) is taken as an indication of shRNA expression. Due to the dual promoter nature of the system, the constitutively active Ubiquitin C (Ubc) or elongation factor-1 alpha (EF1a) promoter drives the expression of the enhanced green fluorescent protein, puromycin or luciferase enzyme cDNA as the selection marker. As my studies indicate that I could generate lentivirus particles that are capable of infecting target cells. Moreover, I selected stable clones that are resistant to the antibiotic puromycin and also synthesize tRFP. These results indicate that the lentivirus vector DNA is integrated into the genome of MCF7 cells. Based on these results, I expected that the expression of shRNA specifically designed, and effectively used, for the silencing of the CXXC5 synthesis would allow us to address the role of CXXC5 as an estrogen responsive gene product in estrogen-mediated cellular events including cellular proliferation. However, several attempts to downregulate CXXC5 synthesis assessed at both mRNA and protein levels yielded no results. Although the underlying cause(s) is unclear, one likely reason could be that the amount of shRNA synthesized in

response to rtTA3 through TRE promoter may not be enough for silencing CXXC5. Indeed, our observations that the puromycin resistant *PAC* gene encoding sequence together with rtTA3 encoding sequence appear to have integrated into the host genome to allow resistance to puromycin and the detection of tRFP indicate that both TRE and Ubc promoters are active. Since, however, shRNA#10 sequences juxtaposed to miRNA motifs (miRNA30) are genetically fused to the end of the sequence encoding tRFP and since tRFP is synthesized, this suggest that shRNA#10 is also expressed but at levels that may be insufficient to alter CXXC5 synthesis. It is also possible that the shRNA#10 sequence conjugated to miRNA30 is not processed or processed inefficiently by the Dicer-RISC complex to generate siRNA for CXXC5. As we were editing my thesis work we realized that our shRNA design may not give rise to an siRNA that could effectively repress the CXXC5 synthesis [44]. All these possibilities are testable predictions that are needed to be considered for the establishment of an effective inducible system. Regardless of the underlying cause, a controlled level of synthesis could further our efforts to delineate CXXC5 functions in cellular events mediated by estrogens. This would, in turn, could provide novel approaches for the development of prognostics and/or therapeutic tools against estrogen tissue malignancies including breast cancer.

REFERENCES

[1] E. H. Wall, S. C. Hewitt, L. K. Case, C.-Y. Lin, K. S. Korach, and C. Teuscher, "The role of genetics in estrogen responses: a critical piece of an intricate puzzle." The FASEB Journal, 2014.

[2] J. Huang, X. Li, R. Hilf, R. A. Bambara, and M. Muyan, "Molecular Basis of Therapeutic Strategies for Breast Cancer," vol. 3, no. Lob 2, pp. 379–396, 2005.

[3] P. Ascenzi, A. Bocedi, and M. Marino, "Structure – function relationship of estrogen receptor a and b : Impact on human health," vol. 27, pp. 299–402, 2006.

[4] UniProt, "UniProtKB - P03372 (ESR1_HUMAN)." [Online]. Available: http://www.uniprot.org/uniprot/P03372. [Accessed: 29-Mar-2018].

[5] P. Yaşar, G. Ayaz, S. D. User, G. Güpür, and M. Muyan, "Molecular mechanism of estrogen–estrogen receptor signaling," *Reprod. Med. Biol.*, vol. 16, no. 1, pp. 4–20, 2017.

[6] J. M. Hall, J. F. Couse, and K. S. Korach, "The Multifaceted Mechanisms of Estradiol and Estrogen Receptor Signaling," *J. Biol. Chem.*, vol. 276, no. 40, pp. 36869–36872, 2001.

[7] L. Li, M. P. Haynes, and J. R. Bender, "Plasma membrane localization and function of the estrogen receptor alpha variant (ER46) in human endothelial cells.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 100, no. 8, pp. 4807–12, 2003.

[8] B. J. Deroo and A. V. Buensuceso, "Minireview: Estrogen Receptor-β: Mechanistic Insights from Recent Studies," *Mol. Endocrinol.*, vol. 24, no. 9, pp. 1703–1714, 2010.

[9] Ensembl, "Gene: ESR1 ENSG00000091831." [Online]. Available: https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00

000091831;r=6:151656691-152129619. [Accessed: 29-Mar-2018].

[10] S. L. Nott, Y. Huang, X. Li, B. R. Fluharty, X. Qiu, W. V. Welshons, S. Y. Yeh, and M. Muyan, "Genomic responses from the estrogen-responsive element-dependent signaling pathway mediated by estrogen receptor ?? are required to elicit cellular alterations," *J. Biol. Chem.*, vol. 284, no. 22, pp. 15277–15288, 2009.

[11] P. Yaşar, G. Ayaz, and M. Muyan, "Estradiol-Estrogen Receptor α Mediates the Expression of the CXXC5 Gene through the Estrogen Response Element-Dependent Signaling Pathway," *Sci. Rep.*, vol. 6, no. November, p. 37808, 2016.

[12] Ensembl, "Gene: CXXC5 ENSG00000171604." [Online]. Available: https://www.ensembl.org/Homo_sapiens/Gene/Summary?g=ENSG00000171604;r=5:139647299-139683882. [Accessed: 29-Mar-2018].

[13] UniProt, "UniProtKB - Q7LFL8 (CXXC5_HUMAN)," *Last modified: March 28, 2018*. [Online]. Available: http://www.uniprot.org/uniprot/Q7LFL8. [Accessed: 29-Mar-2018].

[14] H. Zhou, S. Di Palma, C. Preisinger, M. Peng, A. N. Polat, A. J. R. Heck, and S. Mohammed, "Toward a comprehensive characterization of a human cancer cell phosphoproteome," *J. Proteome Res.*, vol. 12, no. 1, pp. 260–271, 2013.

[15] H. K. Long, N. P. Blackledge, and R. J. Klose, "ZF-CxxC domain-containing proteins, CpG islands and the chromatin connection," *Biochem. Soc. Trans.*, vol. 41, no. 3, pp. 727–740, 2013.

[16] F. Pendino, E. Nguyen, I. Jonassen, B. Dysvik, A. Azouz, M. Lanotte, E. Ségal-Bendirdjian, and J. R. Lillehaug, "Functional involvement of RINF, retinoid-inducible nuclear factor (CXXC5), in normal and tumoral human myelopoiesis," *Blood*, vol. 113, no. 14, pp. 3172–3181, 2009.

[17] M. S. Kim, S. K. Yoon, F. Bollig, J. Kitagaki, W. Hur, N. J. Whye, Y. P. Wu, M. N. Rivera, J. Y. Park, H. S. Kim, K. Malik, D. W. Bell, C. Englert, A. O. Perantoni, and S. B. Lee, "A novel wilms tumor 1 (WT1) target gene negatively regulates the WNT signaling pathway," *J. Biol. Chem.*, vol. 285, no. 19, pp.

14585–14593, 2010.

[18] I. Lee, S. Choi, J. H. Yun, S. hwa Seo, S. Choi, K. Y. Choi, and W. Lee, “Crystal structure of the PDZ domain of mouse Dishevelled 1 and its interaction with CXXC5,” *Biochem. Biophys. Res. Commun.*, vol. 485, no. 3, pp. 584–590, 2017.

[19] S.-H. Lee, M.-Y. Kim, H.-Y. Kim, Y.-M. Lee, H. Kim, K. A. Nam, M. R. Roh, D. S. Min, K. Y. Chung, and K.-Y. Choi, “The Dishevelled-binding protein CXXC5 negatively regulates cutaneous wound healing,” *J. Exp. Med.*, vol. 212, no. 7, pp. 1061–1080, 2015.

[20] H.-Y. Kim, S. Choi, J.-H. Yoon, H. J. Lim, H. Lee, J. Choi, E. J. Ro, J.-N. Heo, W. Lee, K. T. No, and K.-Y. Choi, “Small molecule inhibitors of the Dishevelled-CXXC5 interaction are new drug candidates for bone anabolic osteoporosis therapy.,” *EMBO Mol. Med.*, vol. 8, no. 4, pp. 375–87, 2016.

[21] H. Y. Kim, J. Y. Yoon, J. H. Yun, K. W. Cho, S. H. Lee, Y. M. Rhee, H. S. Jung, H. J. Lim, H. Lee, J. Choi, J. N. Heo, W. Lee, K. T. No, D. Min, and K. Y. Choi, “CXXC5 is a negative-feedback regulator of the Wnt/β-catenin pathway involved in osteoblast differentiation,” *Cell Death Differ.*, vol. 22, no. 6, pp. 912–920, 2015.

[22] M. Y. Kim, H. Y. Kim, J. Hong, D. Kim, H. Lee, E. Cheong, Y. Lee, J. Roth, D. G. Kim, D. S. Min, and K. Y. Choi, “CXXC5 plays a role as a transcription activator for myelin genes on oligodendrocyte differentiation,” *Glia*, vol. 64, no. 3, pp. 350–362, 2016.

[23] G. Li, X. Ye, X. Peng, Y. Deng, W. Yuan, Y. Li, X. Mo, X. Wang, Y. Wan, X. Liu, T. Chen, Z. Jiang, X. Fan, X. Wu, and Y. Wang, “CXXC5 regulates differentiation of C2C12 myoblasts into myocytes,” *J. Muscle Res. Cell Motil.*, vol. 35, no. 5–6, pp. 259–265, 2014.

[24] H. Y. Kim, D. H. Yang, S. W. Shin, M. Y. Kim, J. H. Yoon, S. Kim, H. C. Park, D. W. Kang, D. Min, M. W. Hur, and K. Y. Choi, “CXXC5 is a transcriptional activator of Flk-1 and mediates bone morphogenic protein-induced endothelial cell differentiation and vessel formation,” *FASEB J.*, vol. 28, no. 2, pp. 615–

626, 2014.

[25] A. Astori, H. Fredly, T. Aloysius, L. Bullinger, V. M.-D. Mas, P. de la Grange, F. Delhommeau, K. M. Hagen, C. Récher, I. Dusanter-Fourt, S. Knappskog, J. R. Lillehaug, F. Pendino, and Ø. Bruserud, "CXXC5 (Retinoid-Inducible Nuclear Factor, RINF) is a Potential Therapeutic Target in High-Risk Human Acute Myeloid Leukemia.," *Oncotarget*, vol. 4, no. 9, pp. 1438–1448, 2013.

[26] Ø. Bruserud, H. Reikvam, H. Fredly, J. Skavland, K.-M. Hagen, T. Thy Van Hoang, A. K. Brenner, A. Kadi, A. Astori, B. T. Gjertsen, and F. Pendino, "Expression of the potential therapeutic target CXXC5 in primary acute myeloid leukemia cells -high expression is associated with adverse prognosis as well as altered intracellular signaling and transcriptional regulation," *Oncotarget*, vol. 6, no. 5, 2014.

[27] M. B. Treppendahl, L. Möllgård, E. Hellström-Lindberg, P. Cloos, and K. Grønbæk, "Downregulation but lack of promoter hypermethylation or somatic mutations of the potential tumor suppressor CXXC5 in MDS and AML with deletion 5q," *Eur. J. Haematol.*, vol. 90, no. 3, pp. 259–260, 2013.

[28] D. G. Andrea Kühnl, Peter J. M. Valk, Mathijs A. Sanders, Adam Ivey, Robert K. Hills, Ken I. Mills, Rosemary E. Gale, Martin F. Kaiser, Richard Dillon, Melanie Joannides, Amanda Gilkes, Torsten Haferlach, Susanne Schnittger, Estelle Duprez, David, "Downregulation of the Wnt inhibitor CXXC5 predicts a better prognosis in acute myeloid leukemia," *Blood*, 2014.

[29] X. Peng, G. Li, Y. Wang, J. Zhuang, R. Luo, J. Chen, F. Chen, Y. Shi, J. Li, Z. Zhou, X. Mo, X. Liu, W. Yuan, Q. Zeng, Y. Li, Z. Jiang, Y. Wan, X. Ye, W. Xu, X. Wang, X. Fan, P. Zhu, X. Wu, and Y. Deng, "CXXC5 is required for cardiac looping relating to TGFβ signaling pathway in zebrafish," *Int. J. Cardiol.*, vol. 214, pp. 246–253, 2016.

[30] X. Yan, J. Wu, Q. Jiang, H. Cheng, J. J. Han, and Y. Chen, "CXXC5 suppresses hepatocellular carcinoma by promoting TGF-β-induced cell cycle arrest and apoptosis.," *J. Mol. Cell Biol.*, vol. 1, no. October, pp. 1–12, 2017.

[31] T. A. Melhuish, C. M. Gallo, and D. Wotton, "TGIF2 Interacts with Histone

Deacetylase I and Represses Transcription," *J. Biol. Chem.*, vol. 276, no. 34, pp. 32109–32114, 2001.

[32] S. Ma, X. Wan, Z. Deng, L. Shi, C. Hao, Z. Zhou, C. Zhou, Y. Fang, J. Liu, J. Yang, X. Chen, T. Li, A. Zang, S. Yin, B. Li, J. Plumas, L. Chaperot, X. Zhang, G. Xu, L. Jiang, N. Shen, S. Xiong, X. Gao, Y. Zhang, and H. Xiao, "Epigenetic regulator CXXC5 recruits DNA demethylase Tet2 to regulate TLR7/9-elicited IFN response in pDCs," *J. Exp. Med.*, vol. 214, no. 5, pp. 1471–1491, 2017.

[33] C. Xu, K. Liu, M. Lei, A. Yang, Y. Li, T. R. Hughes, and J. Min, "DNA Sequence Recognition of Human CXXC Domains and Their Structural Determinants," *Structure*, vol. 26, no. 1, p. 85–95.e3, 2018.

[34] M. Zhang, R. P. Wang, Y. Y. Wang, F. C. Diao, F. Lu, D. Gao, D. Y. Chen, Z. H. Zhai, and H. B. Shu, "The CXXC finger 5 protein is required for DNA damage-induced p53 activation," *Sci. China, Ser. C Life Sci.*, vol. 52, no. 6, pp. 528–538, 2009.

[35] K. L. Meerbrey, G. Hu, J. D. Kessler, K. Roarty, M. Z. Li, J. E. Fang, J. I. Herschkowitz, A. E. Burrows, A. Ciccia, T. Sun, E. M. Schmitt, R. J. Bernardi, X. Fu, C. S. Bland, T. A. Cooper, R. Schiff, J. M. Rosen, T. F. Westbrook, and S. J. Elledge, "The pINDUCER lentiviral toolkit for inducible RNA interference in vitro and in vivo," *Proc. Natl. Acad. Sci.*, vol. 108, no. 9, pp. 3665–3670, 2011.

[36] M. B. Lyng, A. V. Lænkholm, N. Pallisgaard, and H. J. Ditzel, "Identification of genes for normalization of real-time RT-PCR data in breast carcinomas," *BMC Cancer*, vol. 8, pp. 1–11, 2008.

[37] C. T. Chien, P. L. Bartel, R. Sternglanz, and S. Fields, "The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest.," *Proc. Natl. Acad. Sci.*, vol. 88, no. 21, pp. 9578–9582, 1991.

[38] C. Laboratories, "Mammalian Assay Kit 2 User Manual," 2008.

[39] P. Yi, S. Bhagat, R. Hilf, R. a Bambara, and M. Muyan, "Differences in the abilities of estrogen receptors to integrate activation functions are critical for subtype-specific transcriptional responses.," *Mol. Endocrinol.*, vol. 16, no. 8,

pp. 1810–27, 2002.

[40] K. J. Roux, D. I. Kim, and B. Burke, “BioID: A screen for protein-protein interactions,” *Curr. Protoc. Protein Sci.*, no. SUPPL.74, pp. 1–14, 2013.

[41] R. Nadimpalli, N. Yalpani, S. J. Gurmukh, and C. R. Simmons, “HDAC1 interacts with checkpoint proteins.,” no. 908, 2000.

[42] L. Jønson, J. Vikesaa, A. Krogh, L. K. Nielsen, T. vO. Hansen, R. Borup, A. H. Johnsen, J. Christiansen, and F. C. Nielsen, “Molecular Composition of IMP1 Ribonucleoprotein Granules,” *Mol. Cell. Proteomics*, vol. 6, no. 5, pp. 798–811, 2007.

[43] P. Li, J. Li, L. Wang, and L. J. Di, “Proximity Labeling of Interacting Proteins: Application of BioID as a Discovery Tool,” *Proteomics*, vol. 17, no. 20, pp. 1–10, 2017.

[44] X. Bofill-De Ros and S. Gu, “Guidelines for the optimal design of miRNA-based shRNAs,” *Methods*, vol. 103, pp. 157–166, 2016.

APPENDIX A

OLIGOSNUCLEOTIDES USED IN THIS STUDY

ORF of CXXC5

ATGTCGAGCCTCGGCGGTGGCTCCCAGGATGCCGGCGGCAGTAGCAGCA
GCAGCACCAATGGCAGCGGTGGCAGTGGCAGCAGTGGCCCAAAGGCAGG
AGCAGCAGACAAGAGTGCAGTGGTGGCTGCCGCCGCACCAGCCTCAGTG
GCAGATGACACACCACCCCCCGAGCGTCGGAACAAGAGCGGTATCATCA
GTGAGCCCCTCAACAAGAGCCTGCGCCGCTCCCGCCCGCTCTCCCACTAC
TCTTCTTTTGGCAGCAGTGGTGGTAGTGGCGGTGGCAGCATGATGGGCGG
AGAGTCTGCTGACAAGGCCACTGCGGCTGCAGCCGCTGCCTCCCTGTTGG
CCAATGGGCATGACCTGGCGGCGGCCATGGCGGTGGACAAAAGCAACCC
TACCTCAAAGCACAAAAGTGGTGCTGTGGCCAGCCTGCTGAGCAAGGCA
GAGCGGGCCACGGAGCTGGCAGCCGAGGGACAGCTGACGCTGCAGCAGT
TTGCGCAGTCCACAGAGATGCTGAAGCGCGTGGTGCAGGAGCATCTCCC
GCTGATGAGCGAGGCGGGTGCTGGCCTGCCTGACATGGAGGCTGTGGCA
GGTGCCGAAGCCCTCAATGGCCAGTCCGACTTCCCCTACCTGGGCGCTTT
CCCCATCAACCCAGGCCTCTTCATTATGACCCCGGCAGGTGTGTTCCTGG
CCGAGAGCGCGCTGCACATGGCGGGCCTGGCTGAGTACCCCATGCAGGG
AGAGCTGGCCTCTGCCATCAGCTCCGGCAAGAAGAAGCGGAAACGCTGC
GGCATGTGCGCGCCCTGCCGGCGGCGCATCAACTGCGAGCAGTGCAGCA
GTTGTAGGAATCGAAAGACTGGCCATCAGATTTGCAAATTCAGAAAATG
TGAGGAACTCAAAAAGAAGCCTTCCGCTGCTCTGGAGAAGGTGATGCTT
CCGACGGGAGCCGCCTTCCGGTGGTTTCAGTGA

siRNA which was used in previous studies and the shRNA which was designed from it is highlighted.

Table 15 CXXC5 shRNA#10 oligos

Oligo Name	Sequence
pIND10 CXXC5 shRNA# 10 UP (XhoI modified)	5'TCGACCCTCAGATTTGCAAATTCAGAAATTCA AGAGATTTCTGAATTTGCAAATCTGATTTTTTA 3'
pIND10 CXXC5 shRNA# 10 DOWN (MluI)	5'CGCGTAAAAATCAGATTTGCAAATTCAGAAA TCTCTTGAATTTCTGAATTTGCAAATCTGAGGG 3'

Annealing;

```
5'TCGAcccTCAGATTTGCAAATTCAGAAAttcaagagaTTTCTGAATTTGCAAATCTGAttttta     3'
3'    gggAGTCTAAACGTTTAAGTCTTTaagttctctAAAGACTTAAACGTTTAGACTaaaaaTGCGC 5'
```

NTshRNA sequence was taken from premade vector, pTRIPZ NTshRNA, oligos designed accordingly, sequence is below;

Table 16 NTshRNA oligos

Oligo Name	Sequence
pIND10 NTshRNA UP (XhoI modified)	5'TCGAGcccTCTCGCTTGGGCGAGAGTAAGTAgtg aagccacagatgTACTTACTCTCGCCCAAGCGAGAttttt 3'
pIND10 NTshRNA DOWN (MluI)	5' GCGCAttttTCTCGCTTGGGCGAGAGTAAGTAgtga agccacagatgTACTTACTCTCGCCCAAGCGAGAccc'5

Annealing;

```
5'TCGAGcccTCTCGCTTGGGCGAGAGTAAGTAgtgaagccacagatgTACTTACTCTCGCCCAAGCGAGAttttt 3'
3'  gggTCTCGCTTGGGCGAGAGTAAGTAcatctgtggcttcacTACTTACTCTCGCCCAAGCGAGAaaaaaTGCGC 5'
```

APPENDIX B

SEQUENCING AND QPCR PRIMERS

Table 17 Sequencing primers and control primers

Oligo Name	Sequence
LNCX seq primer	5'AGCTCGTTTAGTGAACCGTCAGATC'3
pIND20 seq primer	5'TTGCGCGCTGATTTTTGCGGTATAAG'3
pIND 11 seq primer	5'TTCGTCTGACGTGGCAGCGCT'3
GAPDH_FP	5'GGGAGCCAAAAGGGTCATCA '3
GAPDH_REP	5'TTTCTAGACGGCAGGTCAGGT '3

Table 18 qPCR Primers

Oligo Name	Sequence
CXXC5_qPCR_FP	5'GATGCACCCGTCTTTAGAACC'3
CXXC5_qPCR_REP	5'ATGCTAGGGACGTGGAGATG '3
RPLP0_FP	5'GGAGAAACTGCTGCCTCATA '3
RPLP0_REP	5'GGAAAAAGGAGGTCTTCTCG '3

APPENDIX C

CLONING PRIMERS

Table 19 Primers for cloning of the genes to the Dual Luciferase vectors

Primer name	Sequence
CXXC5_FP	CGCATATACTCGAGATTACCATGGAAGCTAGCATGTCGAGCCTCGGCGGTGGCTC
CXXC5_REP	CGCATGGGATCCAAGCTTTTTATTAGAATTCCTGAAACCACCGGAAGGCGGC
CIRBP_FP	CGCATCTCGAGCCATGGAAGCTAGCGCATCAGATGAAGGCAAACT
CIRBP_REP	CGCATGGGATCCAAGCTTTTTATTAGAATTCCTCGTTGTGTGTAGCGTAACTG
PCBP2_FP	CGCATCTCGAGCCATGGAAGCTAGCGACACCGGTGTGATTGAAGG
PCBP2_REP	CGCATGGGATCCAAGCTTTTTATTAGAATTCGCTGCTCCCCATGCCACCCG
RPL15_FP	CGCATCTCGAGCCATGGAAGCTAGCGGTGCATACAAGTACATCCAGG
RPL15_REP	CGCATGGGATCCAAGCTTTTTATTAGAATTCGCGGTAACGGTGGAGCTGG
RPS2_FP	CGCATCTCGAGCCATGGAAGCTAGCGCGGATGACGCCGGTGCAGC
RPS2_REP	CGCATGGGATCCAAGCTTTTTATTAGAATTCTGTTGTAGCCACAGCTGGAGC
RPS4X_FP	CGCATCTCGAGACCATGGAAGCTAGCGCTCGTGGTCCCA
RPS4X_REP	CGCATGGGATCCAAGCTTTTTATTAGAATTCCCACTGCT

APPENDIX D

GENOMIC DNA CONTAMINATION CONTROL OF RNA ISOLATION WITH PCR USING GAPDH PRIMERS

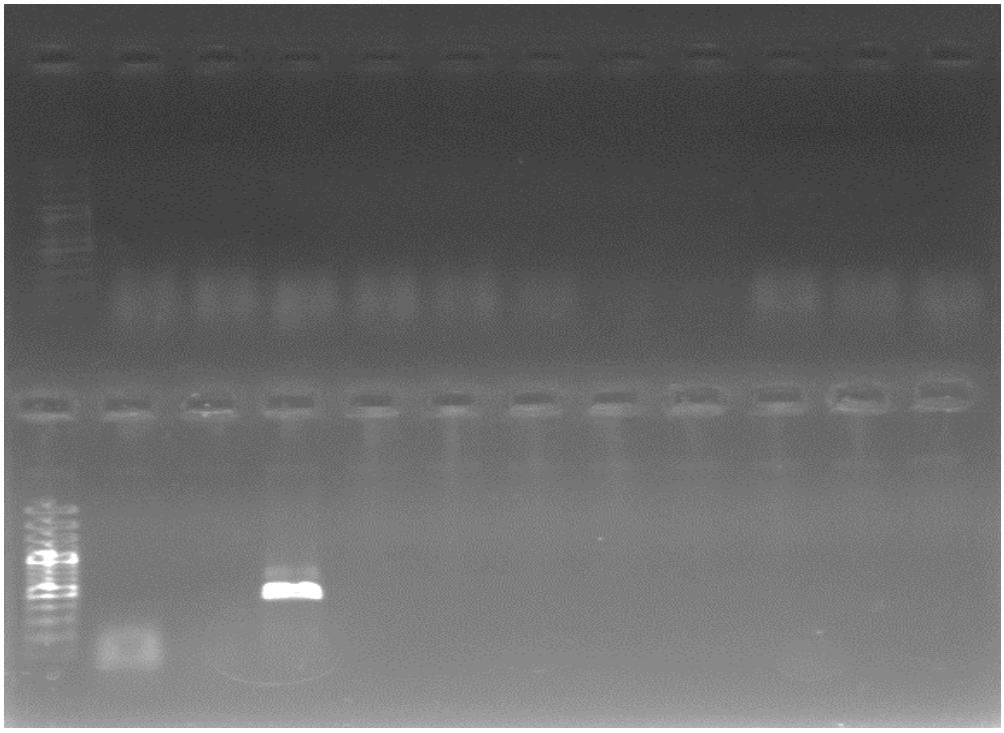

Figure 21 Genomic DNA contamination control PCR

with gDNA positive control

APPENDIX E

CDNA CONVERSION CONTROL

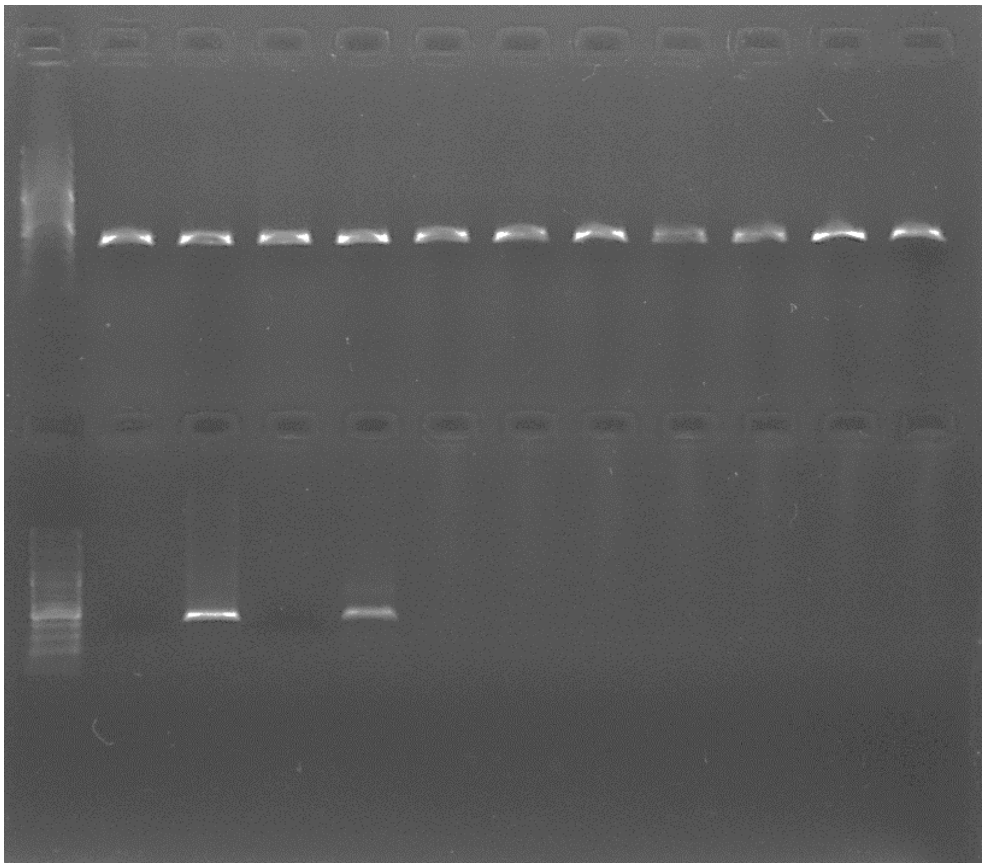

Figure 22 cDNA conversion control PCR

with GAPDH cDNA and gDNA positive controls

APPENDIX F

MIQE CHECK LIST

Table 20 MIQE checklist

ITEM TO CHECK	IMPORTANCE	CHECKLIST
EXPERIMENTAL DESIGN		
Definition of experimental and control groups	E	YES
Number within each group	E	YES
Assay carried out by core lab or investigator's lab?	D	YES
Acknowledgement of authors' contributions	D	N/A
SAMPLE		
Description	E	N/A
Volume/mass of sample processed	D	N/A
Microdissection or macrodissection	E	N/A
Processing procedure	E	N/A
If frozen - how and how quickly?	E	N/A
If fixed - with what, how quickly?	E	N/A
Sample storage conditions and duration (especially for FFPE samples)	E	N/A
NUCLEIC ACID EXTRACTION		
Procedure and/or instrumentation	E	YES
Name of kit and details of any modifications	E	YES
Source of additional reagents used	D	N/A
Details of DNase or RNAse treatment	E	YES
Contamination assessment (DNA or RNA)	E	YES
Nucleic acid quantification	E	YES
Instrument and method	E	YES
Purity (A260/A280)	D	YES
Yield	D	NO
RNA integrity method/instrument	E	YES
RIN/RQI or Cq of 3' and 5' transcripts	E	YES

Table 20 continued

Electrophoresis traces	D	YES
Inhibition testing (Cq dilutions, spike or other)	E	YES
REVERSE TRANSCRIPTION		
Complete reaction conditions	E	YES
Amount of RNA and reaction volume	E	YES
Priming oligonucleotide (if using GSP) and concentration	E	YES
Reverse transcriptase and concentration	E	YES
Temperature and time	E	YES
Manufacturer of reagents and catalogue numbers	D	YES
Cqs with and without RT	D*	N/A
Storage conditions of cDNA	D	YES
qPCR TARGET INFORMATION		
If multiplex, efficiency and LOD of each assay.	E	N/A
Sequence accession number	E	YES
Location of amplicon	D	YES
Amplicon length	E	NO
In silico specificity screen (BLAST, etc)	E	NO
Pseudogenes, retropseudogenes or other homologs?	D	YES
Sequence alignment	D	YES
Secondary structure analysis of amplicon	D	NO
Location of each primer by exon or intron (if applicable)	E	YES
What splice variants are targeted?	E	YES
qPCR OLIGONUCLEOTIDES		
Primer sequences	E	YES
RTPrimerDB Identification Number	D	N/A
Probe sequences	D**	N/A
Location and identity of any modifications	E	N/A
Manufacturer of oligonucleotides	D	NO
Purification method	D	NO
qPCR PROTOCOL		
Complete reaction conditions	E	YES

Table 20 continued

Reaction volume and amount of cDNA/DNA	E	YES
Primer, (probe), Mg++ and dNTP concentrations	E	N/A
Polymerase identity and concentration	E	N/A
Buffer/kit identity and manufacturer	E	YES
Exact chemical constitution of the buffer	D	N/A
Additives (SYBR Green I, DMSO, etc.)	E	YES
Manufacturer of plates/tubes and catalog number	D	NO
Complete thermocycling parameters	E	YES
Reaction setup (manual/robotic)	D	YES
Manufacturer of qPCR instrument	E	YES
qPCR VALIDATION		
Evidence of optimisation (from gradients)	D	NO
Specificity (gel, sequence, melt, or digest)	E	YES
For SYBR Green I, Cq of the NTC	E	YES
Standard curves with slope and y-intercept	E	YES
PCR efficiency calculated from slope	E	YES
Confidence interval for PCR efficiency or standard error	D	NO
r2 of standard curve	E	YES
Linear dynamic range	E	YES
Cq variation at lower limit	E	YES
Confidence intervals throughout range	D	N/A
Evidence for limit of detection	E	NO
If multiplex, efficiency and LOD of each assay.	E	N/A
DATA ANALYSIS		
qPCR analysis program (source, version)	E	YES
Cq method determination	E	YES
Outlier identification and disposition	E	N/A
Results of NTCs	E	YES
Justification of number and choice of reference genes	E	YES
Description of normalisation method	E	YES

Table 20 continued

Number and concordance of biological replicates	D	YES
Number and stage (RT or qPCR) of technical replicates	E	YES
Repeatability (intra-assay variation)	E	YES
Reproducibility (inter-assay variation, %CV)	D	YES
Power analysis	D	NO
Statistical methods for result significance	E	YES
Software (source, version)	E	YES
Cq or raw data submission using RDML	D	N/A

APPENDIX F

RT-QPCR EFFICIENCY RESULTS

Representative images of RT-qPCR of CXXC5 and RPLP0.

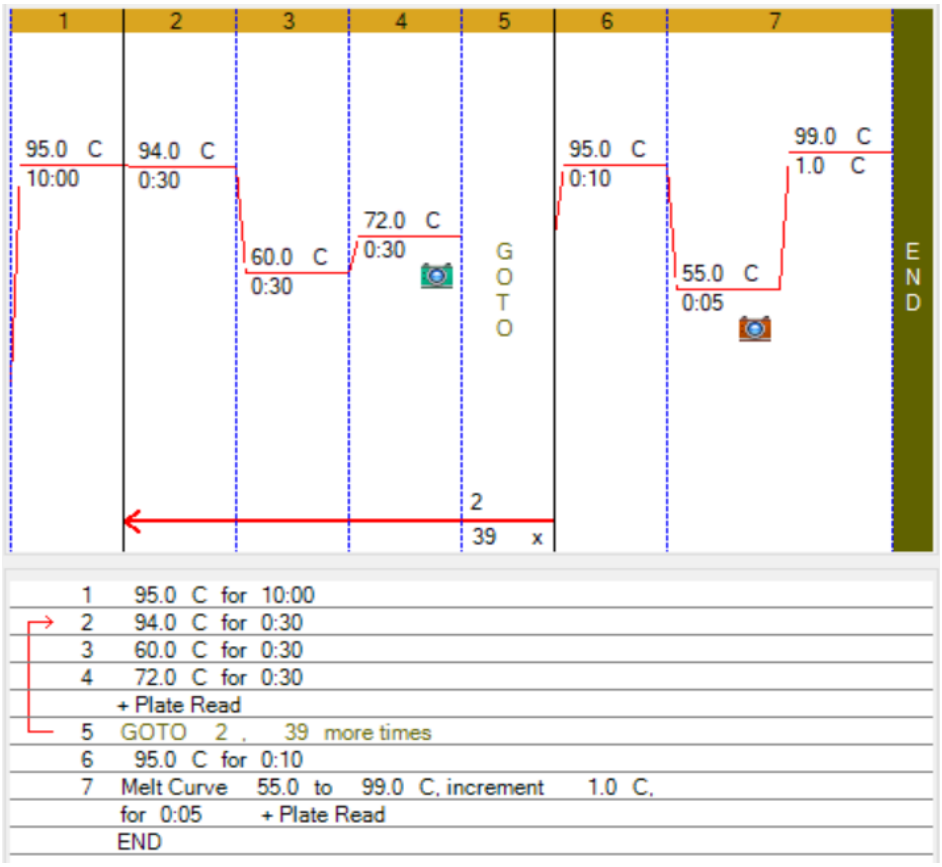

Figure 23 RT-qPCR cycling conditions

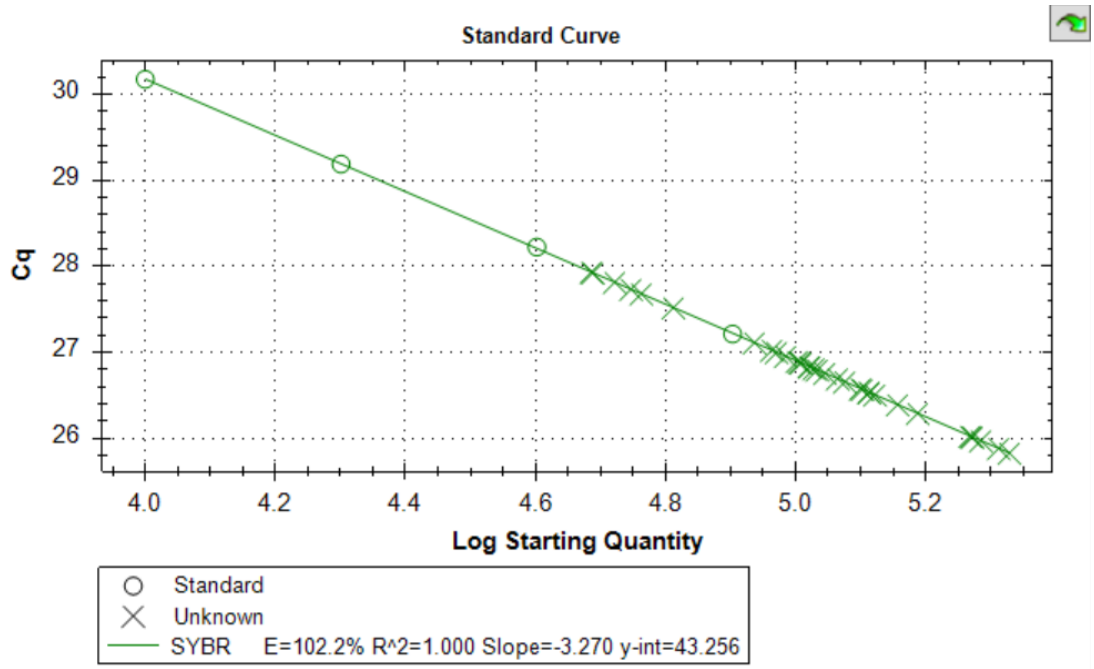

Figure 24 Sample standard curve of qPCR reaction

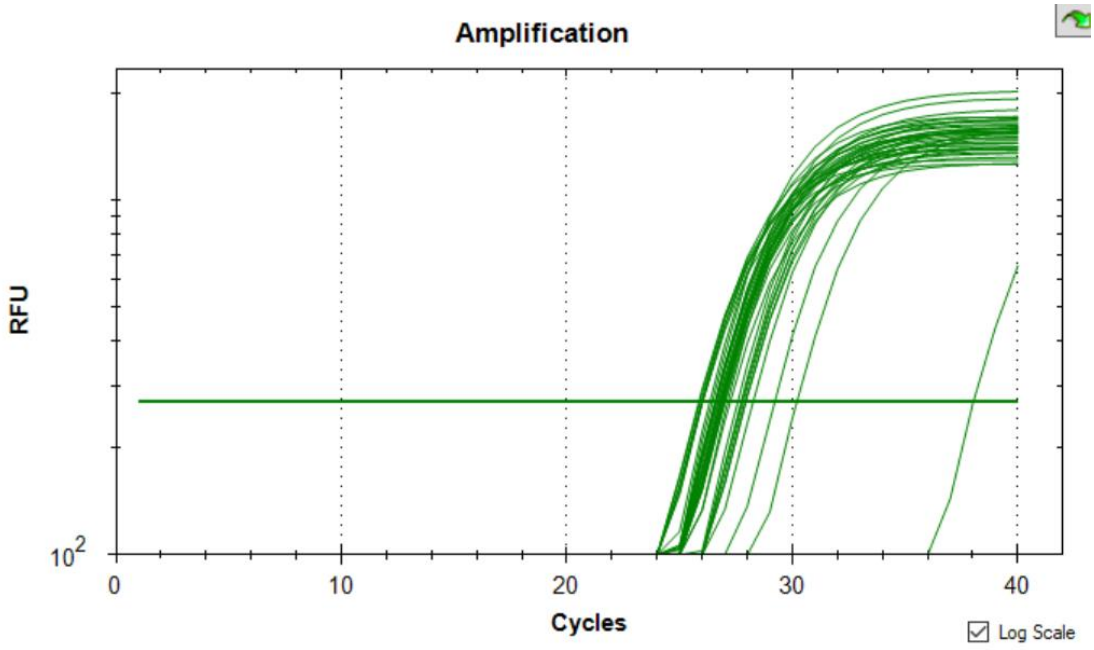

Figure 25 Sample amplification curve for qPCR

Description	Value	Use
Negative control with a Cq less than	38	☑
NTC with a Cq less than	38	☑
NRT with a Cq less than	38	☑
Positive control with a Cq greater than	30	☑
Unknown without a Cq	N/A	☑
Standard without a Cq	N/A	☑
Efficiency greater than	110.0	☑
Efficiency less than	90.0	☑
Std Curve R^2 less than	0.980	☑
Replicate group Cq Std Dev greater than	0.20	☑

Negative control with a Cq less than 38

No wells fail this QC Rule.

Figure 26 Quality check points of CFX Software

Fluor	Content	Sample	Starting Quantity (SQ)	Cq	Log Starting Quantity
SYBR	Std	1:10 CXXC5	8.000E+04	27.21	4.903
SYBR	Std	1:20 CXXC5	4.000E+04	28.22	4.602
SYBR	Std	1:40 CXXC5	2.000E+04	29.19	4.301
SYBR	Std	1:80 CXXC5	1.000E+04	30.17	4.000
SYBR	Unkn-01	MCF7	1.080E+05	26.79	5.033
SYBR	Unkn-01	MCF7	1.260E+05	26.58	5.100
SYBR	Unkn-01	MCF7	1.435E+05	26.39	5.157
SYBR	NTC		N/A	38.03	N/A
SYBR	NTC		N/A	N/A	N/A

Figure 27 Sample result, Cq values of qPCR

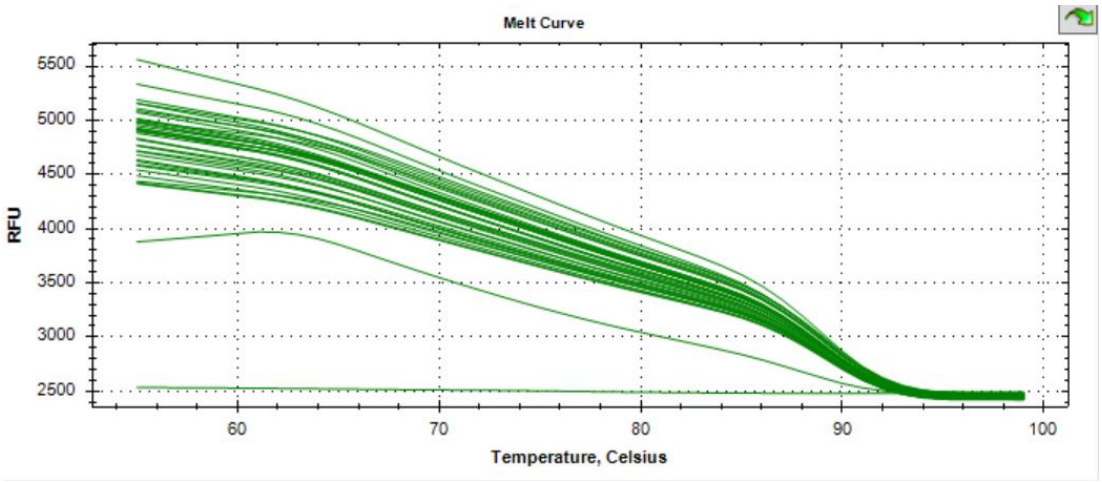

Figure 28 Sample melt curve

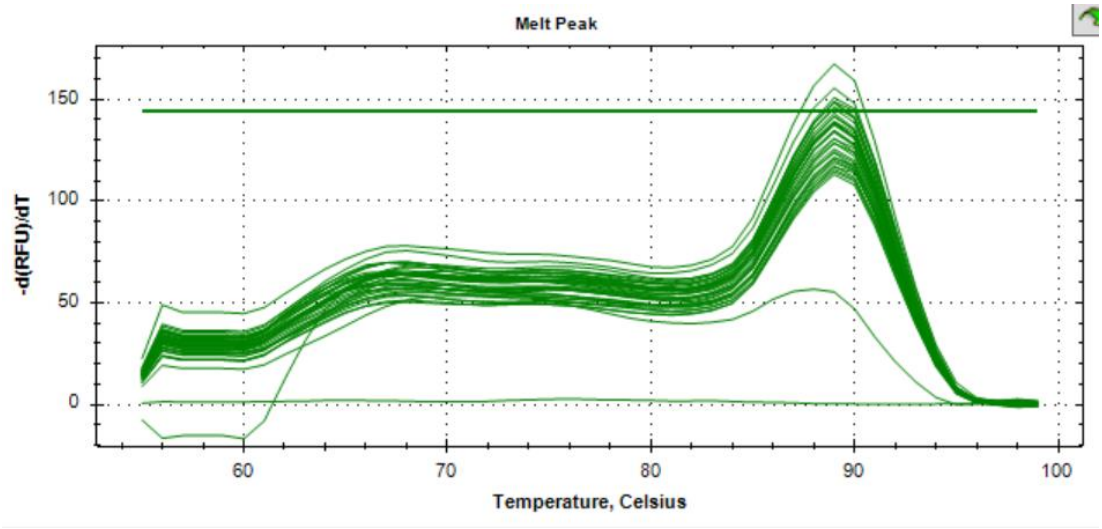

Figure 29 Sample melt peak

APPENDIX H

TIME COURSE QPCR ANALYSIS RESULTS OF COLONY #33

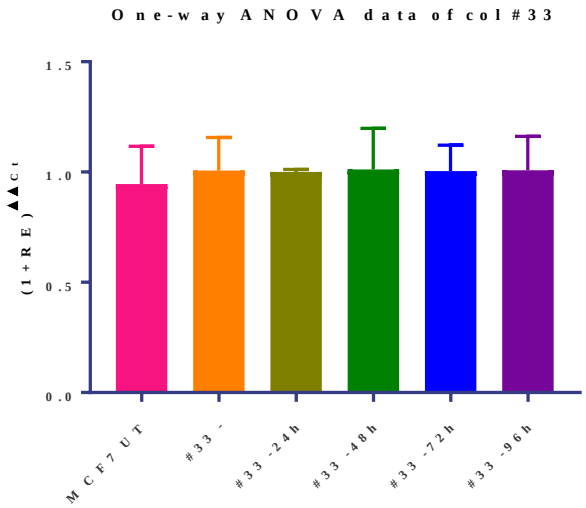

Figure 30 Graphical representation of RT-qPCR results of time course experiment of colony #33

Table 21 Graphical representation of RT-qPCR results of time course experiment of colony #33

Dunnett's multiple comparisons test	Mean Diff.	Summary	Adjusted P Value
MCF7 UT vs. #33 -	-0.06122	ns	0.9764
MCF7 UT vs. #33 -24h	-0.05394	ns	0.9861
MCF7 UT vs. #33 -48h	-0.06618	ns	0.9676
MCF7 UT vs. #33 -72h	-0.05854	ns	0.9803
MCF7 UT vs. #33 -96h	-0.06187	ns	0.9753

APPENDIX I

BUFFERS AND SOLUTIONS

6X SDS Loading Buffer

0.35 M Tris-HCl pH:6.8

10.28% (w/v) SDS

36% (v/v) Glycerol

5% β – Mercaptoethanol

0.0012% (w/v) Bromophenol blue

5X SDS Sample Loading Buffer

187.5 mM Tris-HCl (pH 6.8)

6% (w/v) SDS

30% glycerol

150 mM DTT

0.03% (w/v) bromophenol blue

2% β-mercaptoethanol

6X Laemmli Buffer

12 % SDS

30 % β - Mercaptoethanol

60 % Glycerol

0.012 % Bromophenol blue

0.375 M Tris

20% SDS

Sodium Dodecyl Sulfate	20 g	Dissolve by heating
dH_2O	100 ml	

Tris Buffer Saline (TBS) pH:7.6

20 mM Tris (MW:121.14g)

137 mM NaCl (MW:58.44g)

30 x Acrylamide Stock Solution	29:1	Bring to 100 ml w/dH_2O, dissolve
Acrylamide	29 g	with the help of heat, Filter
Bisacrylamide	1 g	sterilization w/22µm filter

Protein Running Buffer

250 mM Tris

2500mM Glycine

1 % SDS

Complete volume to 1 l w/ dH_2O

10x Transfer Buffer		for 1x @4°C
Tris	30.3 g	100 ml 10x stock
Glycine	144.1g	700 ml dH_2O
Complete volume to 1 l w/ dH_2O		200 ml Methanol

Blocking Buffer

TBS + 0.1 % Tween ➔ TBST

TBST + 5 % non-fat dry milk ➔ Blocking buffer

Protein Stripping Buffer pH: 2.2	600ml
β – Mercaptoethanol	4.8 ml
20 % SDS	60 ml
1 M Tris HCl	37.5 ml
Complete volume to 600 ml w / dH_2O	

SDS gel (separating)	8 %	10 %	12 %	15 %
dH_2O	4.6 ml	3.97 ml	3.3 ml	2.3 ml
30% Acrylamide:Bis Solution (29:1)	2.7 ml	3.33 ml	4 ml	5 ml
1.5 M Tris Buffer (pH 8.8)	2.5 ml	2.5 ml	2.5 ml	2.5 ml
10% SDS	100 µl	100 µl	100 µl	100 µl
10% Ammonium persulfate	100 µl	100 µl	100 µl	100 µl
TEMED	6 µl	4 µl	4 µl	4 µl

SDS gel (stacking)	5 %
dH_2O	5.44 ml
30% Acrylamide:Bis Solution (29:1)	1.36 ml
1 M Tris Buffer (pH 6.8)	1 ml
10% SDS	80 µl
10% Ammonium persulfate	80 µl
TEMED	8 µl

APPENDIX J

CHARCOAL COATED DEXTRAN TREATMENT OF FETAL BOVINE SERUM

To obtain charcoal coated dextran stripped fatal bovine serum 500 ml of fetal bovine serum was mixed with 10 g of charcoal coated dextran (Charcoal, dextran coated, Sigma, USA, C6241) and stirred overnight at 4°C. Then charcoal was pelleted by centrifuging at 10800 g for 30 min at 4°C. Procedure was repeated with 10 g of Charcoal again with stirring at 4°C for 4-6 hours and followed by same centrifugation step. Supernatant was taken in to biological safety cabinet and vacuum filtered with 0.45 μM sterile filter unit (Corning, polystyrene, cellulose acetate membrane, low protein binding, Sigma, USA, CLS430770)

APPENDIX K

pINDUCER11 CXXC5 shRNA#10 INFECTED MCF7 DOSE RESPONSE FOR DOXYCYCLINE

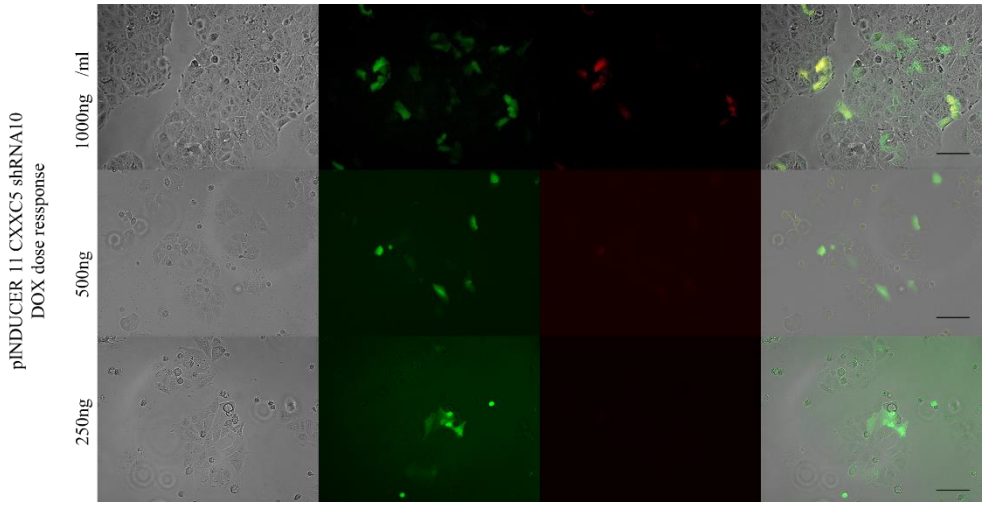

Figure 31 pINDUCER11 CXXC5 shRNA#10 infected MCF7 dose response for Doxycycline

2 x 10^4 MCF7 cells were seeded in 48 well tissue culture plates and infected with pINDUCER11 containing viral particles. 24 hours after infection, medium was changed with fresh medium containing 250, 500 or 1000ng/ml dox. 48 hours later, cells were imaged to observe fluorescent activity. Scale bar represents 100µm.

APPENDIX L

MCF7 KILL CURVE FOR PUROMYCIN

Table 22 MCF7 kill curve with puromycin

MCF7 7 day after selection start with variable puromycin concentrations	
Concentration of puro (ng/ml)	**Cell viability**
0	Over confluent
500	50 %
750	30%
1000	<1 %
1500	-
2000	-

APPENDIX M

WESTERN BLOT ANALYSIS; NORMALISATION AND STATISTICAL ANALYSIS FOR INDUCTION EFFECT ON CXXC5 EXPRESSION

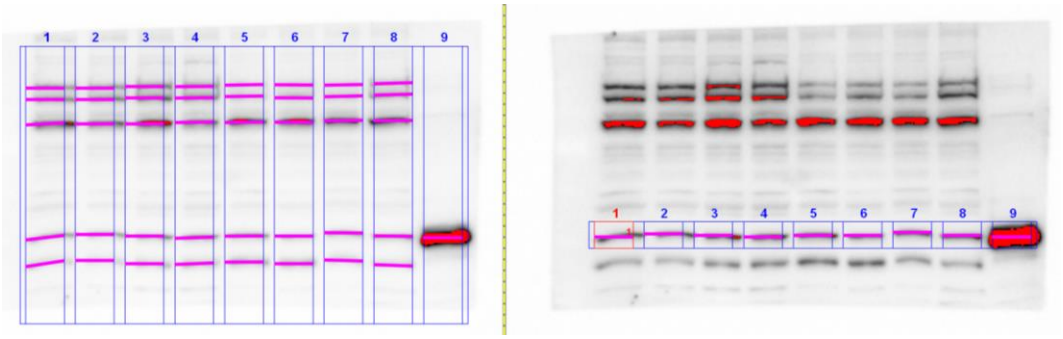

Figure 32 Normalization of the WB data using ImageLab

Table 23 Statistical analysis of the normalized WB data

Tukey's multiple comparisons test					
Compare column means (main column effect)					
Alpha	0.05				
	Mean Diff.	95.00% CI of diff.	Significant?	Summary	Adjusted P
19 vs. 19+	-0.1	-6.956 to 6.756	No	ns	>0.9999
33 vs. 33+	-2.75	-9.606 to 4.106	No	ns	0.9116
47 vs. 47+	0.95	-5.906 to 7.806	No	ns	>0.9999
22 vs. 22+	-0.65	-7.506 to 6.206	No	ns	>0.9999
29 vs. 29+	0.35	-6.506 to 7.206	No	ns	>0.9999
54 vs. 54+	1.7	-5.156 to 8.556	No	ns	0.9978
NT vs. NT+	-1.7	-8.556 to 5.156	No	ns	0.9978
Test details	Mean 1	Mean 2	Mean Diff.	SE of diff.	q
19 vs. 19+	8.45	8.55	-0.1	1.698	0.0833
33 vs. 33+	9.65	12.4	-2.75	1.698	2.291
47 vs. 47+	15.85	14.9	0.95	1.698	0.7913
22 vs. 22+	10.3	10.95	-0.65	1.698	0.5414
29 vs. 29+	11.7	11.35	0.35	1.698	0.2915
54 vs. 54+	12.6	10.9	1.7	1.698	1.416
NT vs. NT+	10.8	12.5	-1.7	1.698	1.416

APPENDIX N

Y2H

Table 24 yeast strains used in Y2H service

AH109	MATa, trp1-901, leu2-3, 112, ura3-52, his3-200, gal4Δ, gal80Δ, LYS2::GAL1$_{UAS}$-GAL1$_{TATA}$-HIS3, GAL2$_{UAS}$-GAL2$_{TATA}$ADE2, URA3::MEL1$_{UAS}$-MEL1$_{TATA}$-lacZ
MaV203	MATα, trp1-901, leu2-3,112, ura3-52, his3Δ200, gal4Δ, gal80Δ, ade2-101, cyh2R, can1R, GAL1::HIS3@LYS2, GAL1::lacZ, SPAL10::URA3@ura3