FUNCTIONAL IMPORTANCE OF CXXC5 IN E2-DRIVEN CELLULAR PROLIFERATION

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

BY

NEGIN RAZIZADEH

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

SEPTEMBER 2019

Approval of the thesis:

FUNCTIONAL IMPORTANCE OF CXXC5 IN E2-DRIVEN CELLULAR PROLIFERATION

submitted by **NEGIN RAZIZADEH** 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. Ayşe Gül Gözen
Head of Department, **Biology** _____

Prof. Dr. Mesut Muyan
Supervisor, **Biology, METU** _____

Examining Committee Members:

Assist. Prof. Dr. Erkan Kiriş
Biological Sciences, Middle East Technical University _____

Prof. Dr. Mesut Muyan
Biology, METU _____

Assist. Prof. Dr. Murat Alper Cevher
Molecular Biology and Genetics, Bilkent University _____

Date: 06.09.2019

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, Surname: Negin Razizadeh

Signature:

ABSTRACT

FUNCTIONAL IMPORTANCE OF CXXC5 IN E2-DRIVEN CELLULAR PROLIFERATION

Razizadeh, Negin
Master of Science, Biology
Supervisor: Prof. Dr. Mesut Muyan

September 2019, 82 pages

17β-estradiol (E2) as the main circulating estrogen hormone has an important role in the regulation of various tissues including mammary tissue. E2 effects target tissue functions by binding to the nuclear receptors, ERα and β. ERs regulate the expression of target genes. Previous studies conducted in our laboratory indicate that one of these estrogen responsive genes is CXXC5 which is regulated by ERα. CXXC5 has a highly conserved zinc-finger CXXC domain, which makes it a member of zinc-finger CXXC domain protein family. The family binds to non-methylated CpG dinucleotides, specifically in CpG island promoters and alters gene expressions through their enzymatic activities for DNA methylation or epigenetic modifications. However, structural and functional properties of CXXC5 remains largely unknown.

In an attempt to decipher the role of CXXC5 in E2-ERα mediated cellular events, we uncovered that CXXC5 do not have an intrinsic transcription activation or repression function but through binding to CpG dinucleotides regulates gene expressions distinctly and mutually modulated by E2 as well. This results in E2-driven cellular proliferation. We therefore suggest that CXXC5 as a CpG binder involves in the regulation of E2-mediated transcriptional activation or repression of genes culminating in the regulation of cellular proliferation.

Keywords: Estrogen, Estrogen Receptor, CXXC5

ÖZ

CXXC5'İN E2 TARAFINDAN DÜZENLENEN HÜCRE PROLİFERASYONUNDAKİ FONKSİYONEL ÖNEMİ

Razizadeh, Negin
Yüksek Lisans, Biyoloji
Tez Danışmanı: Prof. Dr. Mesut Muyan

Eylül 2019, 82 sayfa

17β-östradiyol (E2) dolaşımda bulunan ana östrojen hormonudu olarak, meme dokusu dahil çoğu dokunun düzenlemelerinde önemli bir rol oynuyor. E2 nükleer reseptörlere, ERα ve β, bağlanarak hedef dokuyu etkiler ve hedef genlerin ekspresyonunu düzenler. Laboratuvarımızda yapılan önceki deneylere göre, bu hedef genlerden biri ERα ile düzenlenen CXXC5'dır. CXXC5 proteini, yüksek derece korunmuş bir çinko-parmak domainine sahip olduğu için ZF-CXXC domain protein ailesinin bir üyesi olarak Kabul ediliyor. CXXC5'in metile-olmamış CpG dinükleotitlerine, özellikle transkripsiyonel olarak aktif DNA bölgelerinde, bağlanabilmesi gösterilmiştir. ZF-CXXC5 ailesi DNA metilasyonu ve/veya epigenetic modifikasyonları üzerine olan enzimatik aktivitesini kullanarak gen ekspresyonunu etkileyebiliyor. Bu güne kadar CXXC5'in yapısal ve foksiyonel özelliklerinin çoğu bilinmemektedir. CXXC5'in E2-ERα tarafından sağlanan hücresel olaylarındaki rolünü çözmek için ilk olarak CXXC5'in gerçekten metile-olmamış CpG dinükleotitlerine bağlanabilmesini göstermemiz gerekti. Bir sonraki adım olarak, CXXC5'in transkripsiyon aktivasyonu veya önlemesinde hakiki ve intrinsic bir etkisi olmadığını bulduk. Bunun yanı sıra, CXXC5'in E2 ile birlike ve ayrı olarak birçok genin düzenlemesinde önemli rol oynadığını gördük. Bu düzenlemeler E2 tarafından düzenlenen hücre çoğalmasına rol oynamaktadır. Bu sonuçlar, CXXC5'in CpG dinükleotitlerine bağlanan bir protein

olarak E2 tarafından düzenlenen hücre çoğalmasında etkili genlerin transkripsiyon aktivasyonu veya represyonunda rol oynadığını önermektedir.

Anahtar Kelimeler: Östrojen, Östrojen Reseptörü, CXXC5

To my family

ACKNOWLEDGEMENTS

First, I would like to thank my supervisor Prof. Dr. Mesut Muyan for his patience, guidance, and support throughout this research. I would not be able to take my first steps in the path of becoming a scientist without his help. I feel truly privileged to have been able to learn from him.

I wanted to express my sincere gratitude to Prof. Dr. Rengül Çetin Atalay and Dr. Deniz Kahraman for their valuable suggestions and for providing their laboratory facilities for our use.

I feel grateful to have known my lab mates, Dr. Gamze Ayaz, Pelin Yaşar, Çağla Ece Olgun, Gizem Kars, Gizem Turan, Kerim Yavuz, Burcu Karakaya, and Öykü Deniz Demiralay. I owe them many thanks for their advices, assistance, and friendship.

I would like to thank my second family here in Turkey; I cannot thank Özlem Yeter enough for her presence and friendship. I thank Furkan Doğramacı for bringing joy to my life. I could fill pages thanking Baturay Özel, Erdem Canbalaban, Nami Çağan Yüksel, Aysel Akgemci, Barış Özkuşlar, Cenk Tüysüz, Emek Barış Küçüktabak, Ece Büber, Sevil Çalışkan, Volkan Doğan, and Hakan Mert one by one. I am grateful for their friendship and consider it a privilege to have spent the past eight years surrounded with them.

I thank my uncles, Yadallah, Selim and Togan Nurioğlu and their families. I especially thank Ayda Göyçek Nurioğlu for being the sister I never had.

Last but certainly not least, I would like to thank my parents without whom I would not be the person I am today. Pages would not be enough to express my feelings of gratitude for all the unconditional love, support and sacrifice they provided me with throughout my life.

I would like to thank TÜBİTAK for supporting this study through 114Z243 project.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

CHAPTER 1

INTRODUCTION

1.1. Estrogen–Estrogen Receptor Signaling

Estrogens are primary female sex hormones and the main estrogen circulating estrogen in body is 17β-estradiol (E2) [1]. E2 has many effects on regulation of different tissues including the breast tissue. Breast cancer is one of the most seen cancer types affecting women and men in some cases. Breast cancer has many subtypes with distinct properties. E2 is considered the primary factor affecting breast cancer along with some contributing factors such as the effects of environment and genetic background [2], [3].

Estrogen receptors α and β are two transcription factors primarily responsible for mediation of E2 effects. ERα and β although encoded by two different genes, *ESR1* and *ESR2*, have a lot in common regarding their structure and functions [4]. The primary estrogen receptor in breast tissue is ERα.

ERα has six different domains, each responsible for a distinct feature [2]. The representation of ERα domains from N-terminal to C-terminal is as shown in Figure 1.1.

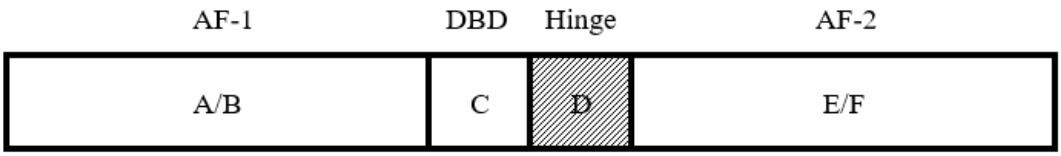


Figure 1.1. Six functional domains of ERα

These six domains are called A/B, C, D and E/F. The A/B domain is located in the amino terminus of ERα and is responsible for transactivation function independently from any ligand (AF-1). The C domain is the DNA binding domain (DBD) which is responsible for binding of ERα to estrogen responsive elements (ERE). The D domain contains a nuclear localization signal (NLS) and works as a hinge domain and causes flexibility between amino- and carboxyl-termini. The E/F domain is a ligand-binding domain (LBD) and is multifunctional that includes dimerization functions, ligand-dependent transactivation (AF-2), and ligand binding [2]. ERα mostly translocates to the nucleus whether there is E2 in the environment or not [5]. When E2 binds to ER, the receptor undergoes major conformational changes which generate surfaces for the recruitment and binding of co-regulatory proteins and for the shifting of the protein from an inactive form to an active one [6], [7].

Estrogen responsive elements (EREs) are palindromic DNA sequences, separated by three non-specific nucleotides, 5’ - GGTCAnnnTGACC - 3’. E2-ER mediated genomic events can occur through either an ERE-dependent or an ERE-independent pathway [6]. The ERE-dependent pathway involves the direct binding of ERα to an ERE. In the ERE-independent pathway, on the other hand, ERs use other transcription factors which are bound to their cognate response elements on DNA [5]. Both pathways play critical roles in cellular events.

1.2. CXXC-Type Zinc Finger Protein 5 (CXXC5) and Its Protein Product

Previous studies of our laboratory suggest that CXXC5 is an E2-responsive gene and its expression is regulated by ERα using an ERE-dependent signaling pathways [5], [8]. In CXXC5 gene regulation, ERα binds to a non-consensus ERE sequence, GGTCAggaTGAC**A**, -242 bp upstream of the first ATG of CXXC5 [5], [8].

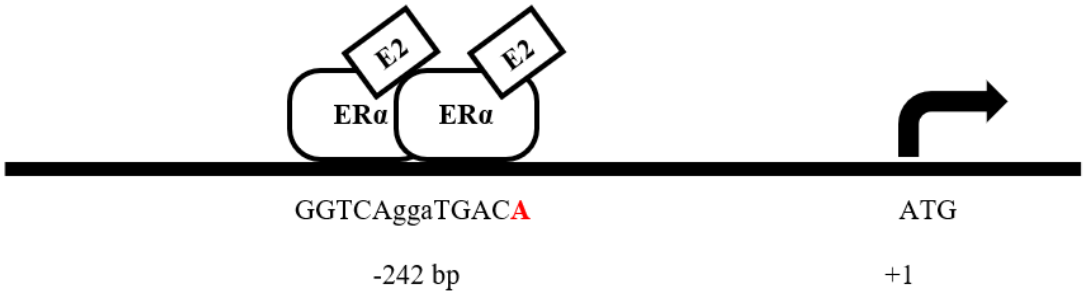


Figure 1.2. ERE dependent regulation of CXXC5

The CXXC5 gene is located on 5q31.2 forward strand, encloses a ~35.5 kb region of genomic DNA and contains 11 exons resulting in 16 transcript variants, 14 of which are protein coding. Its main transcript encodes a 322 amino acid long protein with a molecular mass of ~33 kDa [8]. The CXXC5 protein, which is also referred to as CXXC Finger Protein 5 (CF5), Retinoid-Inducible Nuclear Factor (RINF) or WT1-Induced Inhibitor of Dishevelled (WID), belongs to CXXC-type zinc finger (ZF) protein family, and like all other members of this protein family, CXXC5 contains a ZF-CXXC domain. It has two CXXC type zinc finger motifs, each with four cysteine residues coordinated by a zinc ion, between amino acids 256 and 297.

The majority of CpGs in mammalian genomic DNA are methylated [9], except for those within CpG islands (CGIs), which are usually unmethylated. ZF-CXXC family proteins which include 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-)-Methyltransferase1 (DNMT1, CXXC9), and Tet Methylcytosine Dioxygenase 1, 2, 3 (TET1, TET2, TET3) proteins, are known to bind to these non-methylated CpGs, which are mostly seen in promoter regions [10].

CXXC5 protein contains a nuclear localization signal on its C terminus, which causes the protein to localize in nucleus in cell lines such as MCF7 [8].

1.3. Previous Studies on CXXC5

Although there are limited studies covering structural and functional properties of CXXC5, studies suggest that CXXC5 plays a role as a transcription factor, co-regulator and/or epigenetic factor in a wide variety of cellular functions such as modulation of signal transduction, DNA damage response, cellular energy metabolism, proliferation, differentiation, angiogenesis and cell death [11]–[19].

One of DNA damage response pathways important tumor suppressors is the Cellular Tumor Antigen p53. p53 induces gene expressions responsible for cell cycle arrest, senescence, apoptosis, and DNA repair and CXXC5 induces p53 transcriptional activity and can induce apoptosis using its interaction with Ataxia Telangiectasia Mutated Protein Kinase (ATM) [18]. CXXC5 is shown to be involved in DNA damage pathway using ATM-p53 signaling pathway, because interaction of CXXC5 with ATM prevents ATM phosphorylation and consequently preventing the recruitment of the protein to DNA break sites [18].

Moreover, *in vivo* experiments indicate that CXXC5 contributes to osteoblast differentiation, growth plate senescence, cutaneous wound healing, hair loss and antiviral responses and kidney and heart development [13], [20]–[26]. CXXC5 is suggested to be a transcriptional regulator in different tissues; for example, it is reported that it enhances the expressions of genes involved in the skeletal muscle differentiation, and its absence acts as an inhibitor of myocyte differentiation [15]. Moreover, CXXC5 activated by BMP4, can regulate the expression of receptor for vascular endothelial growth factor gene (Flk-1) during endothelial differentiation and vessel formation [14].

Acute promyelocytic leukemia (APL) is a subtype of acute myeloid leukemia (AML) and APL cells mature with the help of CXXC5 expression [11]. CXXC5 is also reported to be involved in TGF-β signaling and that deregulation of CXXC5 expression affects TGF-β signaling and consequently causes the development of cardiovascular disease (REF) [16], [25]. On the other hand, CXXC5 by regulating

TGF-β expression can contribute to tumor suppression in hepatocellular carcinoma (REF) [16], [27].

In accordance with the importance of CXXC5 in cellular events, de-regulated expressions of *CXXC5* appear to correlate with the development, and resistance to therapies, of various pathologies including cardiovascular disease, diminished ovarian reserve (DOR), Blepharophimosis Ptosis Epicantus inversus Syndrome (BPES), Acute Myeloid Leukemia (AML), prostate and breast cancer [11], [28]–[33].

1.4. Aim of This Study

Although we have shown previously that the E2-ERα signaling increases *CXXC5* expression and the synthesis of CXXC5 protein [5], [8], there is not much known about the functions of CXXC5 in cellular events mediated by E2-ERα. To study this, we needed to verify that full length CXXC5, as its CXXC domain, is indeed a non-methylated CpG binding protein. We used siRNAs to target CXXC5 mRNA and study the effects of low CXXC5 levels in MCF7 cells, derived from a breast adenocarcinoma. MCF7 cells, which are E2-responsive and synthesizes ERα, have been used as a useful cell model. We conducted cell proliferation, cell cycle, and apoptosis-detection experiments and also used Nanostring PanCancer Pathway to study the effects of CXXC5 in the absence (-E2) or presence of E2 (+E2) on expression levels of around 700 genes active in cancerous pathways. We here show that CXXC5 is a non-methylated CpG dinucleotide binding protein with effects on E2-mediated gene expressions. This results in the modulation of E2-driven cellular proliferations.

CHAPTER 2

MATERIALS AND METHODS

2.1. Cell lines and Maintenance

MCF7 cells used in these experiments were a kind gift of Prof. Dr. Rengül Çetin Atalay (Middle East Technical University, Ankara, Turkey). MCF7 cells were grown in high glucose (4.5g/L) phenol-red free Dulbecco's modified Eagle's medium (DMEM, Biological Industries, Canada, BI01-053-1A) with additional 10% final volume Fetal Bovine Serum (FBS, Biowest, France, S181G-500), 1.2% L-Glutamine (L-Glutamine, Biological Industries, Canada, BI03-020-1B), and 1% Penicillin-Streptomycin (Pen-Strep, Biological Industries, Canada, BI03-031-1B).

In the cases of E2 treatment, the experiments were carried out using phenol-red free DMEM supplemented with 10% charcoal dextran (Dextran Coated Charcoal, Sigma, Germany, C6241-20G) treated FBS (CDFBS), 1.2% L-Glutamine, and 1% Penicillin-Streptomycin. MCF7 cells were maintained in steroid hormone free medium for 48 hours prior to transfection and hormone treatment.

Cells were maintained for a maximum of 5 weeks in 95% humidified, 5% CO2, 37˚C cell culture incubator as monolayer and passaged or refreshed as needed every three days.

2.2. Hormone Treatments

To determine the effects of estrogen on CXXC5 knock down MCF7 cells, we used 17β-estradiol (E2) (Sigma-Aldrich, St. Louis, MO, USA). To make sure the observed effects are ER-dependent, a complete antagonist of ERα Imperial Chemical Industries 182,780 (ICI, Tocris Biosciences, MN, USA) was used.

2.3. siRNA Transfection

To perform siRNA knock-down experiments we used FlexiTube GeneSolution (#1027416) for CXXC5 (Qiagen). We used two different siRNAs for CXXC5 knock-down, siRNA 2 and 10, which both target the open reading frame of *CXXC5* transcript. The target sequences are given in Table 2.1.

Table 2.1. *The siRNA sequences and targeting regions of CXXC5*

siRNAs	siRNA sequence (5' to 3')	Target
Hs_CXXC5_2	CAGCAGTTGTAGGAATCGAAA	In the ORF of *CXXC5*
Hs_CXXC5_10	TCAGATTTGCAAATTCAGAAA	In the ORF of *CXXC5*

In addition to these siRNAs there is Allstars negative control siRNA (Qiagen, Germany) which is a validated non-targeting siRNA with no known silencing effect on any mammalian gene, thus can be used as a negative control during siRNA transfection experiments. All mentioned siRNAs were received in lyophilized form and dissolved in RNase free water to a final concentration of 10µM and stored at -20˚C until use.

$5x10^3$ MCF7 cells/well were seeded to 96-well plates for cell counting experiments and $160x10^3$ MCF7 cells were seeded to 6-well plates for western blot and Annexin V experiment. 48 hours after seeding the cells, the transfection mixture was prepared by the use of the HiPerfect transfection reagent (Qiagen, Germany). Transient transfection mixture protocol consists of 1µl siRNA (from the mentioned 10µM main stock), 5µl HiPerfect transfection reagent, and 94µl phenol red free DMEM (for the cells in a 6-well plate). The transfection mixture is then vortexed shortly and incubated for 30 minutes in room temperature; then the appropriate amount of media and hormones are added in the absence (ethanol, 0.01%) or the presence of E2 (10^{-8} M) and/or ICI (10^{-7} M) and the final mixture is introduced to the cells. 24 hours after transfection, the cells' media is refreshed (containing the hormones) and 48 hours after transfection, the cells are trypsinized and collected for subsequent processing.

The transfection ratios for 96 and 12-well cultures are given in Table 2.2. and transfectin was carried out as described above.

Table 2.2. *siRNA transfection mixture*

Culture Format	Volume of Transfection Mixture (µl)	Volume of Diluted siRNA (µl)	Volume of HiPerfect Transfection Reagent (µl)	Final siRNA Concentration (nM)
96-well plate	100	0.1	0.5	10
12-well plate	500	0.5	2.5	10
6-well plate	1000	1	5	10

In all experiments the siRNA concentration used is 10nM. This concentration was the most appropriate since it is the highest concentration in which the transfected cells show minimal phenotypic change up to 96 hours after transfection.

2.4. Cellular Proliferation

To observe the effects of CXXC5 knock-down on cellular growth, a cellular proliferation experiment was done on siRNA transfected cells. Cells were plated in 96-well culture plates in CD-FBS containing media for 48h prior to transfection, they were then transiently transfected with siRNAs as previously described. 48h after transfection, cells were collected and counted using a hemocytometer. The cell proliferation can also be measured by using MTT (3-(4,5-Dimethylthiazol-2-yl)-2,5 diphenyltetrazolium bromide) assay (Sigma-Aldrich, MO, USA, M5655) according to the protocol provided by the manufacturer.

2.5. Luciferase Assay

To address the question of whether CXXC5 has an intrinsic transcription activation or proliferation repression activity, we used mammalian two hybrid experiments. pM expression vector containing GAL4 DNA Binding Domain (Gal4DBD, amino acids 1–147), and a nuclear localization signal was used as the template to generate desired

proteins. Appropriate restriction enzyme sites were used to insert the sequences encoding the proteins of full length CXXC5, wild type MECP2 (amino acids 1-486) which was kindly provided by Dr. K. Miyake, University of Yamanashi, Japan), and the carboxyl-terminus of ERα-EF domain which contains ligand-dependent transaction activity (amino acids 301–595) to the 3' end of the GAL4DBD encoding sequence. Among our controls, expression vectors containing only the full length CXXC5, wild type MECP2 and ERαEF cDNA were included to make sure of the specificity of the results. $10x10^3$ MCF7 cells per well of a 48-well plate were seeded. 48 hours after seeding, cells were transiently transfected with *Renilla* luciferase containing vector (0.5 ng / well), pGal4-RE-Luc reporter plasmid (125 ng / well) and combinations of expression vectors bearing cDNAs (75 ng of each interactor / well). TurboFect transfection reagent (Thermo Scientific, USA, R0531) was used for this transfection; and when added to the mixture, using the proportions provided by the manufacturer, the mixture was incubated at room temperature for 30 minutes and introduced to cells. In wells transfected with either the expression vector bearing ERαEF or Gal4DBD-ERαEF cDNA, cells were treated without (ethanol, 0.01%) or with 10^{-8} M E2. 24 hours after transfection, cells were lysed, and cellular extracts were analyzed for the change in luciferase enzyme activities using a Dual Luciferase Assay kit (Promega Corp. Madison, WI, USA).

2.6. Cell Cycle Analysis

To study the effects of reduced CXXC5 protein levels on cell cycle phases, cell cycle analysis was performed. Cells, $16 x10^4$/well of a 6-well plate, were seeded and grown in phenol red-free medium containing 10% CD-FBS. After 48 hours, cells were transfected using the siRNAs, as mentioned before, in the absence (ethanol, 0.01%) or the presence of 10^{-8} M E2. 48 hours after transfection, cells were trypsinized, collected, and washed with PBS. Cell pellets were then re-suspended in 100 µl of PBS containing 2% CD-FBS, to prevent cell clumps, and fixed by adding 4 ml 70% ice-cold ethanol dropwise while gently vortexing the cells. After fixation, cells can be stored at 4˚C. When ready to continue the assay, fixed cells were centrifuged and

washed once with 1X PBS to remove ethanol. Then, they were incubated with 100 µl of PBS containing propidium iodide (0.02 mg/ml; Sigma-Aldrich), 200 µg/ml RNase A (Thermo-Fisher Scientific) and 0.1% (v/v) Triton X-100 (AppliChem, Germany) for 30 minutes in dark. After incubation, cells can be analyzed by flow cytometry (BD AccuriTM C6 Cytometer; BD Biosciences, San Jose, CA, USA). FL2 channel was used for PI readings on BD Accuri Flow Cytometer, and $10x10^3$ cells were analyzed for each sample.

2.7. Annexin V Assay

To study the effects of reduced CXXC5 protein levels on cell death, Annexin V assay was performed. MCF7 cells, 16 x104/well of a 6-well plate, were seeded and grown in phenol red-free medium containing 10% CD-FBS. After 48 hours, cells were transfected using the siRNAs, as mentioned before, in the absence (ethanol, 0.01%) or the presence of 10^{-8} M E2. 48 hours after transfection, cells were trypsinized, collected, and washed with PBS. Cells were then subjected to Annexin V staining according to the manufacturer's instructions (11988549001, Roche, Germany). Washed cell pellets were resuspended in 100 µl Annexin V incubation buffer containing 2 µl Annexin V Fluorescein and 2 µl Propidium Iodide. Then, cell suspension was gently mixed and incubated for 20 minutes at room temperature in the dark, cells were centrifuged to remove the buffer and dyes, then the pellet was resuspended in 200 µl of incubation buffer and analyzed by flow cytometry (BD Accuri^{TM} C6 cytometer, BD Biosciences).

2.8. Nuclear and Cytoplasmic Protein Isolation

To isolate nuclear and cytoplasmic proteins of the cells separately, the NE-PER system was used (Nuclear and Cytoplasmic Protein Extraction Kit; Thermo Scientific, CA, USA). CERI, CERII, and NER are the buffers used for protein extraction as indicated in the kit's protocol provided by the manufacturer. For cells collected from two wells of a 6-well plate, a ratio of 100:5.5:50 µl of CERI:CERII:NER buffers are used. CERI and NER buffers are prepared by adding protease inhibitor (Roche Applied Science,

Switzerland). After isolation, the proteins are stored at -80˚C until further use. Before use, the proteins are thawed on ice and their concentrations are calculated using the Quick Start Bradford Protein Assay (Bio-Rad, CA, USA).

2.9. Total Protein Isolation

To extract the total amount of proteins of cells collected from 2 wells of a 6-well plate, 100µl M-PER reagent (Mammalian Protein Extraction Reagent; Thermo Scientific, CA,USA) containing protease inhibitor (Roche Applied Science, Switzerland) is used according to the protocol provided by the manufacturer. After isolation, the proteins are stored at -80˚C until further use. Before use, the proteins are thawed on ice and their concentrations are calculated using the Quick Start Bradford Protein Assay (Bio-Rad, CA, USA).

2.10. Western Blotting

Proteins were extracted with NE-PER as described above, denatured with 6X Laemmli buffer and boiled at 95˚ C for 5 minutes. Laemmli buffer composition is indicated in Table 2.3.

Table 2.3. *Laemmli buffer components*

Tris-base	0.375 M Tris (pH 6.8)
Glycerol	60%
β-mercaptoethanol	30%
SDS	12%
Bromophenol Blue	0.012%

Denatured samples were then electrophoresed with 100-volt power for an hour using a 10% or an 10% SDS-PAGE according to the sizes of the proteins. When the proteins are separated, the samples are then transferred onto a PVDF (Polyvinylidene Fluoride, Advansta, USA, L-08008-001) membrane using wet-transfer method. After the transfer, the membrane is blocked in 5% skim milk in 0.1% TBS-T (Tris Buffered

Saline- Tween) for an hour at room temperature. The CXXC5 (CXXC5 antibody, Abcam, USA, ab106533) and PARP-1 (PARP-1 antibody, Cell Signaling, USA, 46D11) antibody dilutions used were 1:500 and 1:1000, respectively. The membrane was incubated with the CXXC5 primary antibody at room temperature for an hour and with the PARP-1 antibody at 4˚C overnight. Then after washing the membrane three times with 0.1% TBS-T for a total of 15 minutes, the secondary antibody incubation begins. The secondary goat anti-rabbit horseradish peroxidase, HRP, conjugated (Advansta, USA, R-05072-500) antibody was used with a dilution of 1:4000 for both CXXC5 and PARP-1 primary antibodies. The membrane was incubated with secondary antibody for an hour at room temperature. Then, it was again washed three times and was treated with Enhanced Chemiluminescence solution (ECL, Advansta, USA, K-12045-D50) for 2 minutes in dark for visualization which was carried out with ChemiDoc™ XRS+ System (Molecular Imager ® ChemiDoc™ XRS+ System, BioRad, USA, 170-8265). ImageLab software (ImageLab software, BioRad, USA, 170-9690) was used to obtain images of the membrane.

2.11. RNA Isolation

siRNA treated cells from a 12-well plate were collected by trypsinization. For RNA isolation, Quick-RNA ™ MiniPrep (Quick-RNA ™ MiniPrep, Zymo Research, USA, R1015) kit and the provided protocol were used. The RNA concentration and purity (A260/280 and A260/230) were measured using a micro-volume spectrophotometer (MaestroNano, Maestrogen, Taiwan, MN-913). Obtained RNA samples were then stored at -80°C until further use.

2.12. Genomic DNA Contamination Control

To make sure that our RNA samples did not contain any genomic DNA, 300 ng of each RNA sample was used as PCR template in a PCR targeting GAPDH (glyceraldehyde-3-phosphate dehydrogenase), which is a housekeeping gene. The primers used for targeting GAPDH are mentioned in Table 2.4.

Table 2.4. *GAPDH primer sequences*

GAPDH Forward Primer	5’-GGGAGCCAAAAGGGTCATCA-3’
GAPDH Reverse Primer	5’-TTTCTAGACGGCAGGTCAGGT-3’

This PCR was performed according to the conditions mentioned in Table 2.5., and in the case of presence of any genomic DNA contamination, a PCR product in the size of 409 bp would be detectable on agarose gel.

Table 2.5. *Genomic DNA contamination control PCR components*

	˚C	Duration	Number of Cycles
Initial Denaturation	95	3 minutes	1
Denaturation	95	30 seconds	40
Annealing	65	30 seconds	
Extension	72	1 minute	
Final Extension	72	10 minutes	1
Hold	4	∞	1

2.13. cDNA Synthesis

If the RNA samples are free of genomic DNA contamination, we proceeded to cDNA synthesis. The kit used for cDNA synthesis is RevertAid First Strand cDNA Synthesis Kit (Thermo Scientific, USA, K1622). cDNA was synthesized using 300 ng total RNA and as described in Table 2.6.

Table 2.6. *cDNA synthesis conditions*

Components	Amount
Total RNA	300 ng
Oligo (dT)18 Primer	1 µl
Nuclease-free Water	Up to 12 µl

After preparing the mixture above, it should be incubated at 65˚C for 5 minutes, then chilled on ice for 2 minutes. Meanwhile, the mixture below should be prepared to be added to the previous 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

Then, the prepared mixture should be incubated at 42˚C for 60 minutes and then a final incubation at 70˚C for 5 minutes. The product is now cDNA.

2.14. cDNA Conversion Control

cDNA converted from 300 ng total RNA should be used as PCR template for a second GAPDH PCR to make sure cDNA conversion was successful. After PCR was completed, a GAPDH product of 409 bp should be detectable on agarose gel.

2.15. RT-qPCR Reaction and Analysis

RT-qPCR reactions were done using SsoAdvanced ™ Universal SYBR ® Green Supermix (BioRad, USA, 172-5272) kit and thermal cycler CFX Connect™ Real-Time PCR Detection System. cDNA made from 300 ng total RNA was used as sample and a various number of primers were used for RT-qPCR reactions. For the normalization of results, the expression levels of 60S acidic ribosomal protein P0 (RPLP0) were used. The sequences of all primers used in RT-qPCR experiments are shown in Table 2.7.

Table 2.7. *Sequence of Primers Used in RT-qPCR Experiments*

	Forward Primer Sequence (5’ to 3’)	Reverse Primer Sequence (5’ to 3’)
CXXC5	CGGTGGACAAAAGCAACCCTAC	CGCTTCAGCATCTCTGTGGACT
RPLP0	GGAGAAACTGCTGCCTCATA	GGAAAAAGGAGGTCTTCTCG

AXIN2	CAAACTTTCGCCAACCGTGGTTG	GGTGCAAAGACATAGCCAGAACC
CASP9	GTTTGAGGACCTTCGACCAGCT	CAACGTACCAGGAGCCACTCTT
ID2	TTGTCAGCCTGCATCACCAGAG	AGCCACACAGTGCTTTGCTGTC
IL12A	TGCCTTCACCACTCCCAAAACC	CAATCTCTTCAGAAGTGCAAGGG
IL1R1	GTGCTTTGGTACAGGGATTCCTG	CACAGTCAGAGGTAGACCCTTC
JAK2	CCAGATGGAAACTGTTCGCTCAG	GAGGTTGGTACATCAGAAACACC
RNF43	GGTTACATCAGCATCGGACTTGC	ATGCTGGCGAATGAGGTGGAGT
WNT16	TCGGAAACACCACGGGCAAAGA	GCGGCAGTCTACTGACATCAAC

For the standard curve, 1:10, 1:20, 1:40, and 1:80 dilutions were used. A dilution of 1:20 was used for the samples. The RT-qPCR conditions used are stated in Table 2.8.

Table 2.8. *RT-qPCR Conditions*

	˚C	Duration	Number of Cycles
Initial Denaturation	94	10 minutes	1
Denaturation	94	30 seconds	40
Annealing	60	30 seconds	
Extension and Plate Reading	72	30 seconds	
Rapid Heating	95	10 minutes	1
Melt Curve Analysis	55-95 increment 1 ˚C	5 seconds	1
Hold	4	∞	1

The relative gene expression was calculated from RT-qPCR data using ΔΔCq calculation method. RT-qPCR data was recorded, extracted and displayed using CFX Manager™ Software Gene Expression Analysis (BioRad, USA); and MIQE guidelines were followed.

2.16. Gene Expression Analysis using nCounter PanCancer Pathway Panel and RT-qPCR

MCF7 cells were seeded in 12-well tissue culture plates (9 x 104 cells/well) in phenol red free medium containing 10% CD-FBS medium. After 48 hours, cells were transiently transfected with siRNAs using HiPerfect transfection reagent (Qiagen) as described before. After 48 hours, cells were treated without (0.01% EtOH as vehicle control) or with 10^{-8} M of E2 for 3 hours. Then cells were collected and subjected to RNA extraction, Quick-RNA™ MiniPrep (Zymo Research, Irvine, CA, USA), using the protocol provided by the manufacturer. The total RNA sample was measured using a micro-volume spectrophotometer (MaestroNano, Maestrogen, Taiwan, MN-913). After checking the RNA samples for genomic DNA contamination, as described before, and making sure the RNA samples are pure, 50 ng RNA per sample was used for processing the nCounter PanCancer Pathway Panel gene expression analysis (NanoString Technologies, Inc. Redwood City, CA, USA). This panel contains probes to target the mRNA of 770 different genes including 40 internal reference genes as controls. Each probe has a specific fluorescence barcode which makes digital counting of each gene possible. Sample preparation was carried out with nCounter PrepStation and microscopy scanning of the barcode signals were carried out with nCounter Digital Analyzer (Nanostring) as instructed by the manufacturer. Nanostring nSolver™ 3.0 Analysis software and its Advanced Analysis plug-in were used for quality control, data normalization and differential expression analyses.

For RT-qPCR, cDNAs were synthesized (The RevertAid First Strand cDNA Synthesis Kit, Thermo-Fisher) from total RNA samples used for Nanostring PanCancer Pathway Panel. To verify the results from the gene expression analysis of PanCancer Pathway

Panel, eight genes (the sequences of qPCR primers were listed in the appendix A) were selected with RT-qPCR. RT-qPCR products' amplification was detected with SYBR Green on BioRad Connect Real-Time PCR as described before. The expression of RPLP0 was used for normalization of RT-qPCR results.

2.17. Statistical Analysis

Experiments were conducted at least three independent times. The graphics were drawn using mean ± standard error (S.E.). The statistical analyses were performed using one-way ANOVA with Tukey test for post-hoc analysis or two-tailed unpaired t test with a confidence interval, minimum, of 95% using GraphPad Prism 7 Software (GraphPad Software Inc., USA).

CHAPTER 3

RESULTS AND DISCUSSION

3.1. CXXC5 is a non-methylated CpG Binding Protein

As previously described, CXXC5 protein as a member of the ZF-CXXC family, bears a CXXC domain at the carboxyl-terminus (250-322) which binds to non-methylated CpG dinucleotides [10], [34], [35]. Despite this information, previous analysis in our laboratory has indicated the lack of any known structural motif in the amino terminus of CXXC5 protein [10], [34], [35]. To further understand the functions of this protein, we wanted to assess whether or not the full length CXXC5 protein can bind to non-methylated CpG dinucleotides as well. To address this question, electrophoretic mobility shift assay (EMSA) was conducted by Gamze Ayaz, a previous PhD student in our laboratory. Gamze expressed and purified the recombinant full-length CXXC5 (FL-CXXC5) and the CXXC domain of CXXC5 (CXXC-D) proteins using a bacterial expression system.

Using these recombinant proteins and two DNA fragments, one bearing a non-methylated ***CG*** dinucleotide (5'-GTGATAC***CG***GATCAGT-3'), and another with identical sequence but with a methylated-***CG*** (***mCG***) dinucleotide (5'-GTGATAC***mCG***GATCAGT-3'), EMSA at 1:0.5, 1:1, 1:2, 1:4, and 1:8 molar ratio of DNA (50 μM) to FL-CXXC5 or 1:4 CXXC-D was performed (Fig. 3.1.).

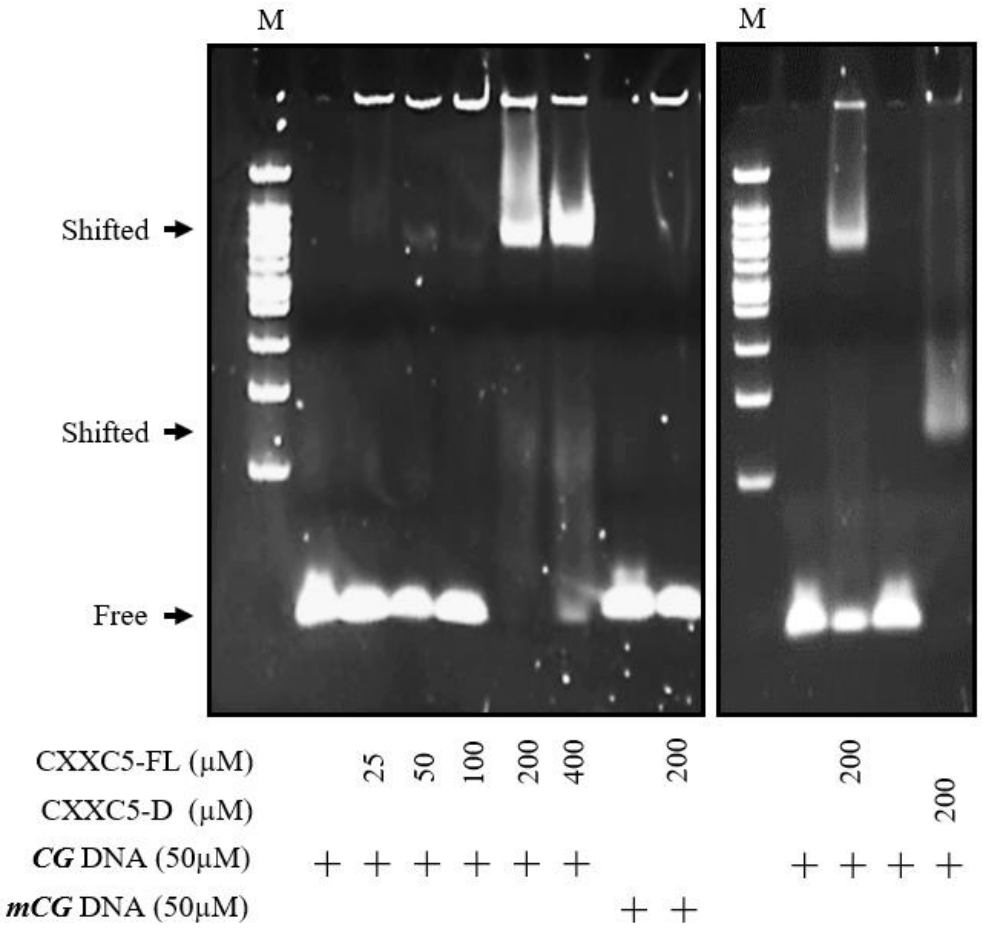


Figure 3.1. Electrophoretic mobility shift (EMSA).

50 μM DNA with the central unmethylated (***CG***) or methylated (***mCG***) CpG dinucleotides were subjected to EMSA with recombinant FL-CXXC5, at 1:0.5, 1:1, 1:2, 1:4 and 1:8 molar ratios, or its CXXC domain (CXXC-D), at 1:4 molarratio. "M" represents molecular marker. "Free" indicates free DNA and "Shifted" shows protein-bound DNA.

The results revealed that CXXC-D, as we already know [35], and FL-CXXC5 both can bind to a non-methylated ***CG*** dinucleotide bearing DNA but not to one bearing ***mCG***. This result indicates that FL-CXXC5 is a non-methylated CpG dinucleotide binder.

3.2. CXXC5 May not Have an Intrinsic Transcription Activation/Repression Function

A number of studies about the function of CXXC5 protein suggest that it plays critical roles in regulating numerous signaling pathways [8], [13], [14], [18], [36], as well as

regulating gene expression by acting as a transcription activator [14], [15], [37] or repressor [17]. CXXC5 protein can be performing these regulations either indirectly by recruiting other proteins, or directly using its possible intrinsic transcription regulatory effects. To assess this possible function of CXXC5 protein, we used a mammalian one-hybrid system, a derivative of the mammalian two-hybrid approach [38], which uses the two functionally distinct domains of eukaryotic transcription factors, DNA binding domain (DBD) and transcription activation. A protein can be genetically fused to the carboxyl-terminus of Gal4DBD using the appropriate enzyme sites. Gal4DBD can bind to Gal4 response elements preceding a TATA box which positions the fused protein to the promoter, driving the expression of the firefly *luciferase* reporter enzyme [39]. Using this approach, there will be no need for a transcription factor-specific responsive element to drive reporter gene expressions. A chimeric Gal4DBD-CXXC5 protein was generated by genetically fusing FL-CXXC5 to the carboxyl-terminus of Gal4DBD. The E/F domain of ERα, which is a ligand (E2)-dependent transcription activator domain [39], was also fused to Gal4DBD. This way Gal4DBD-ERαEF can act as a transcription activator in the presence of the appropriate ligand (E2). Gal4DBD-VP16 fusion protein was also used as a positive control since VP16 domain is a transcription activator. We also fused MeCP2 (methyl-CpG binding protein 2) to Gal4DBD since it was shown to be a transcriptional repressor [40]. These constructs were transiently transfected into MCF7 cells along with a Gal4RE-Luc reporter vector as previously described. MCF7 is an E2 responsive cell line so E2 treatments were done along with ERαEF bearing vectors' transfections to observe the effects of absence (-E2; ethanol as vehicle control, 0.01%) or the presence of E2 (E2; 10-8 M). 24 hours after transfection, cells were lysed and analyzed. (Fig. 3.2.).

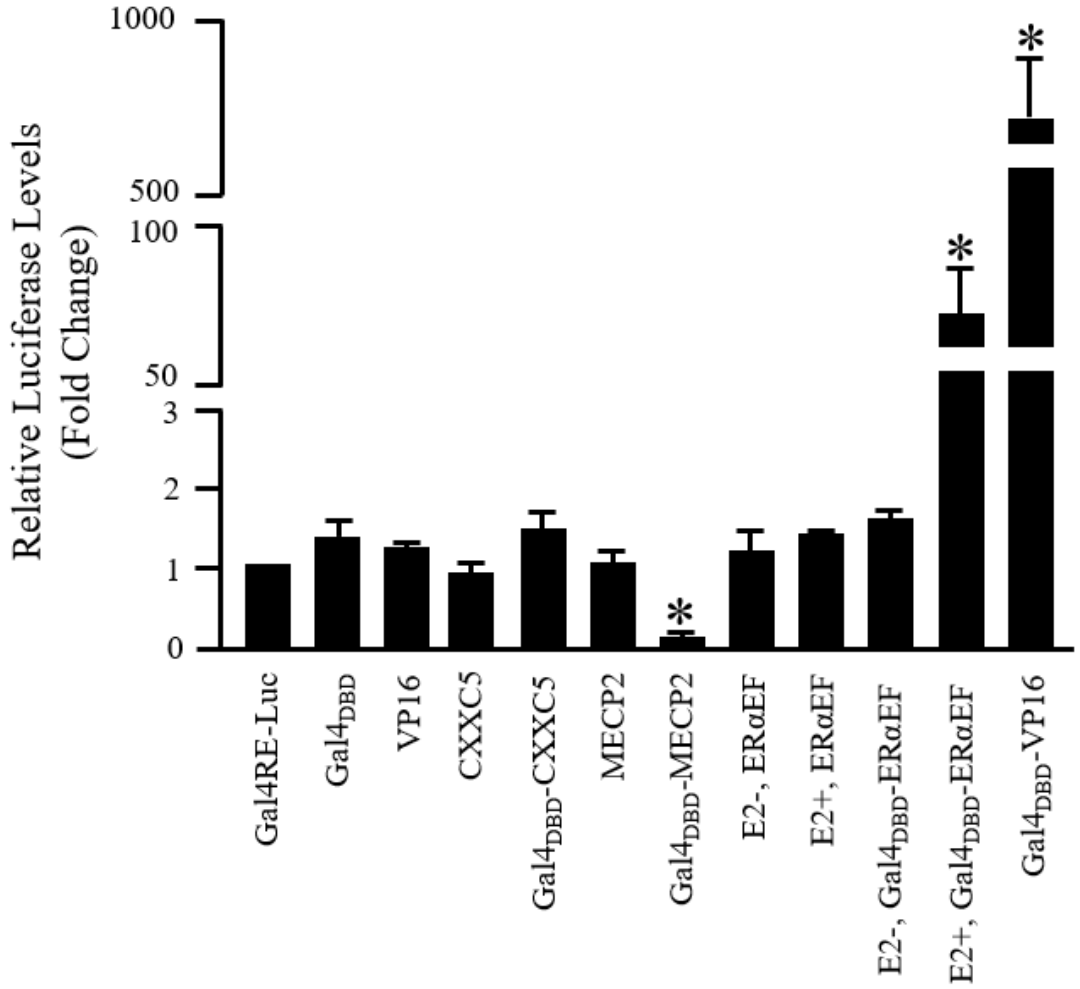


Figure 3.2. Luciferase Assay.

To assess the intrinsic function of full length CXXC5, as a DNA binding protein, on transcription regulation, pGal4-RE Luciferase Reporter vector (pGal4RE-Luc, 125 ng) which expresses the firefly Luciferase enzyme was transfected as the reporter along with an expression vector (75 ng) bearing Gal4DBD, VP16, FL-CXXC5, WT-MECP2 or ERα-EF domain cDNA or Gal4DBD-VP16, Gal4DBD-CXXC5, Gal4DBD-MECP2 or Gal4DBD-ERαEF cDNA, into MCF7 cells that were cultured in 10% CD-FBS containing medium for 48h. Cells transfected with expression vector bearing ERα-EF cDNA or Gal4DBD-ERαEF were treated without (%0.01 EtOH, E2-) or with 10^{-8} M E2 (E2+) for 24h. A *Renilla* Luciferase enzyme cDNA (0.5 ng) bearing plasmid was also co-transfected. Results indicate relative firefly/*Renilla* luciferase activity and are presented as the mean ± S.E. from three independent biological replicates, each performed in duplicate. The significance of each group's luciferase activity compared to Gal4-RE, which is set to 1, is shown by an asterisk.

The effects of each fused protein to Gal4DBD, on regulation of transcription can be estimated by measuring the levels of the reporter enzyme activity. The fused Gal4DBD-VP16 gave the most enzymatic activity as expected because of the constant

transcription activation provided by VP16. Gal4DBD-ERαE/F in the presence of 10^{-8} M E2, has also dramatically increased levels of the reporter enzyme activity compared to Gal4DBD-ERαE/F in the absence of E2 and to that induced by Gal4DBD alone. As shown in previous studies [40], Gal4DBD-MECP2 is indeed repressing the reporter enzyme levels. Our other controls, VP16, MECP2, and ERαE/F (in the absence or presence of E2), had no effects on the enzyme levels indicating that without its DNA binding ability, a transcription regulator loses its potential. FL-CXXC5 protein, alone or fused to Gal4DBD, had no effects on the levels of the reporter enzyme activity. Since fused FL-CXXC5 could be affecting Gal4DBD and preventing it from binding to DNA, we also designed a promoter competition assay (Fig. 3.3.). We know and tested that Gal4DBD-VP16 dramatically enhances the reporter enzyme activity; an increase in the amount of Gal4DBD-CXXC5 fused protein in an environment harboring a certain amount of Gal4DBD-VP16, would start a competition between the two fused proteins for binding to Gal4REs.

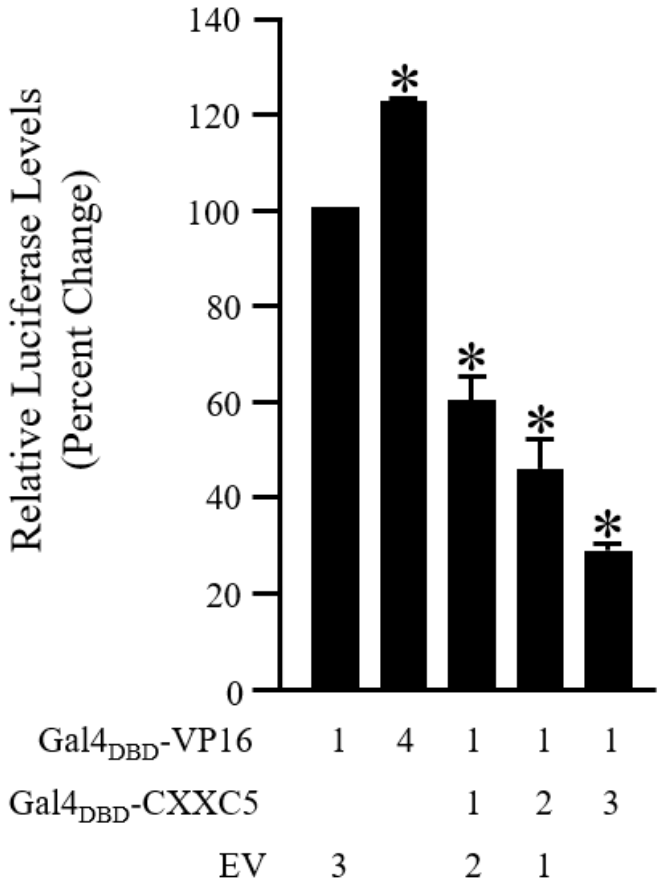


Figure 3.3. Promoter competition assay.

MCF7 cells were transfected with pGal4RE-Luc and an expression vector bearing none (empty vector, EV), Gal4DBD-VP16 and/or Gal4DBD-CXXC5 cDNA. In all groups' transfections, a total of 300 ng expression vector (1 indicates 75 ng) was used to equalize the total plasmid DNA amount; the DNA vacancy in each group is filled with appropriate amounts of the parent expression vector (EV). 24h after transfection, cells were lysed and assayed using a dual luciferase assay kit to detect the luciferase enzyme activities. Results indicate relative percent change of firefly/*Renilla* luciferase levels and are the mean ± S.E. of three independent biological replicates, each performed in duplicate. The significance of each group's luciferase activity compared to Gal4DBD-VP16, which is set to 100, is shown by an asterisk

As seen in the results of this experiment (Fig. 3.3.), the higher the amount of Gal4DBD-CXXC5 fused protein, the lower the enzymatic activity induced by Gal4DBD-VP16, which suggests that Gal4DBD-CXXC5 has the ability to bind to DNA. These results suggest that CXXC5 has no intrinsic transcription regulatory function.

3.3. CXXC5 Knock-Down Affects Cell Proliferation in the Presence of E2

We know from our previous studies that *CXXC5* is an E2-ERα responsive gene [5], [8]. A good approach to further understand the functions of CXXC5 is to observe the changes in cells when CXXC5 levels are increased or decreased. Stable transfections of CXXC5 cDNA or shRNA bearing expression vectors, designed according to our previous study [8], were the first step of addressing the question of CXXC5 function. However, we observed that chronic over-expression and knock-down of CXXC5 led to cell death regardless of cell line used.

Since stable transfections were not possible, we decided to transiently transfect siRNAs to MCF7 cells in the absence (-E2; ethanol as vehicle control, 0.01%) or the presence of E2 (E2; 10^{-8} M). Previously we had used four different siRNAs to target CXXC5 mRNAs (FlexitubeGene Solution, Qiagen siRNA #2, 7, 9 and 10) [8]. We also used a non-targeting siRNA as negative control (CtS). First, we made sure that CtS had no effects on CXXC5 mRNAs or protein levels when comparing them to un-transfected (UT) cells. Then, by transfecting MCF7 cells with the mentioned four siRNAs designed to target CXXC5 transcripts, we observed that among all these siRNAs, siRNA#2 and siRNA#10 transfected cells showed an efficient decrease in CXXC5 mRNA levels and protein synthesis compared to CtS6. The sequences of siRNA#2, siRNA#10, and CtS are given in Table 3.1.

Table 3.1. *siRNA Sequences, 5' to 3'*

siRNA#2	CAGCAGTTGTAGGAATCGAAA
siRNA#10	TCAGATTTGCAAATTCAGAAA
CtS	TTCTCCGAACGTGTCACGT

We decided to continue using SiRNAs #2 and 10 for our further experiments. We initially wanted to assess the effects of CXXC5 knock-down on cellular growth. To do so, MCF7 cells were grown in 10% CD-FBS containing medium for 48 hours to eliminate most of the endogenous E2. After 48 hours, cells were transiently transfected

with siRNA#2, siRNA#10, and CtS in the absence (-E2; ethanol as vehicle control, 0.01%) or the presence of E2 (E2; 10^{-8} M) up to 72 hours (Fig. 3.4.).

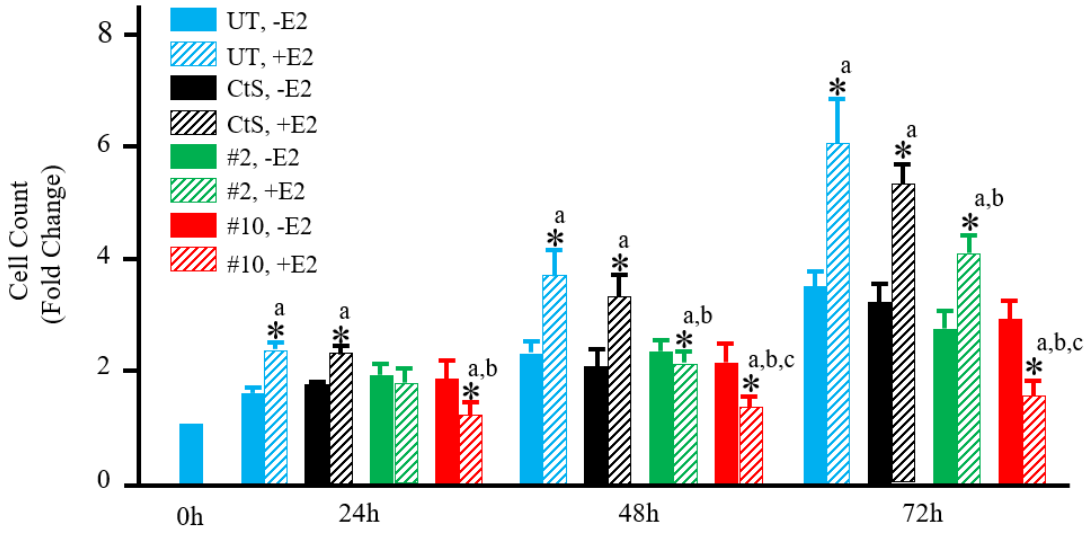


Figure 3.4. Time dependent effects of CXXC5-targeting siRNAs on cellular growth.

MCF7 cells were cultured in 10% CD-FBS containing medium for 48h; then they were transiently transfected without (UT) or with CtS, siRNA#2 or siRNA#10 in the absence (0.01% EtOH, -E2) or the presence of 10^{-8} M E2(+E2) for 24h, 48h or 72h. Cells of each group were then collected and counted using hemocytometer. Results, which are the means ± S.E. of three independent biological replicates, indicate fold changes in cell numbers compared with those observed with cells at 0h time point in the absence of E2, which is set to 1. Asterisk indicates significant change, superscript a indicates a significant difference of the group compared to the corresponding ethanol treated group (-E2); superscript b indicated a significant difference of the group compared to the Cts transfected cells in the presence of E2 (CtS, +E2); and superscript c indicates a significant difference of the group compared to the siRNA#2 transfected cells in the presence of E2 (#2, +E2).

We counted the cells of each group in all the time points (24, 48, and 72 hours). In all time points, E2 treatment, as expected, increased cell proliferation. CtS transfection in absence or presence of E2 did not have any effect on cell proliferation when compared to UT cells. siRNA#2 decreased cellular proliferation at 48 and 72 hours in the presence of E2. siRNA#10 transfected cells showed even more repression in cellular proliferation compared to siRNA#2. The numbers of these cells were effectively decreased specifically in the presence of E2 and in every time point tested. The fact

that siRNA#10 transfected only in the presence of E2 decreased cellular proliferation, was both assessed by cell counting (Fig. 3.5.A) and MTT assay (Fig. 3.5.B).

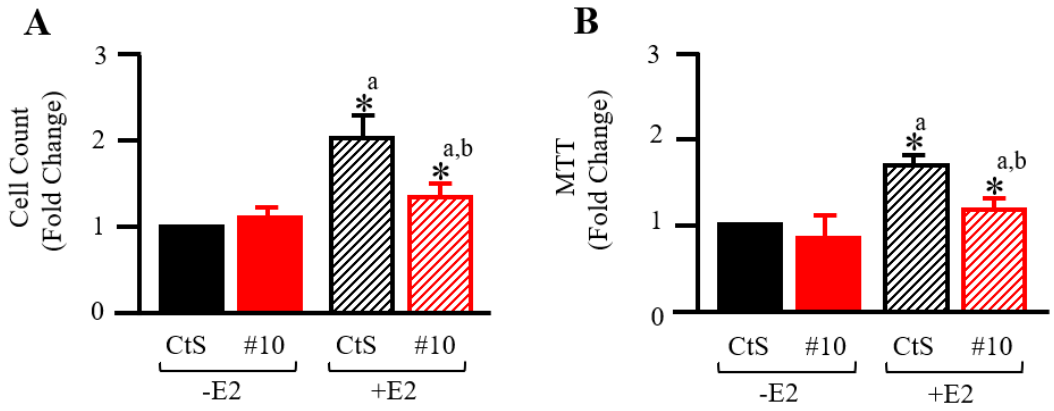


Figure 3.5. Effects of CXXC5-targeting siRNAs on cellular growth.

MCF7 cells were cultured and transfected as described before. After 48h, cells were subjected to cell counting using (A) hemocytometer or (B) MTT assay. Results indicate fold change in cell numbers compared to the group corresponding CtS in the absence of E2, which is set to 1, and are represented as the means ± S.E. of three independent biological replicates. Asterisk indicates significant change, superscript "a" indicates a significant difference of the group compared to the corresponding ethanol treated group (-E2); and superscript "b" indicates a significant difference of the group compared to the Cts transfected cells in the presence of E2 (CtS, +E2).

Based on these time point experiments, siRNA#10 proved to be the most efficient siRNA for CXXC5 knock-down and was selected as the siRNA of choice for the further experiments. Also, as a time point, we decided to use 48 hours because in both 48 and 72 hours after transfection, effective repression in cell proliferation can be observed. Again, MCF7 cells were grown in 10% CD-FBS for 48 hours; then transiently transfected with CtS or siRNA#10 in the absence (0.01% EtOH) or presence of 10^{-8} M E2. CXXC5 expression and protein levels were enhanced in the presence of E2, as we showed previously [8]. The same effects were observed in E2 treated CtS samples. siRNA#10 transfected cells, on the other hand, showed reduced CXXC5 transcript (Fig. 3.6.A) and protein (Fig. 3.6.B) levels both in the absence and presence of E2.

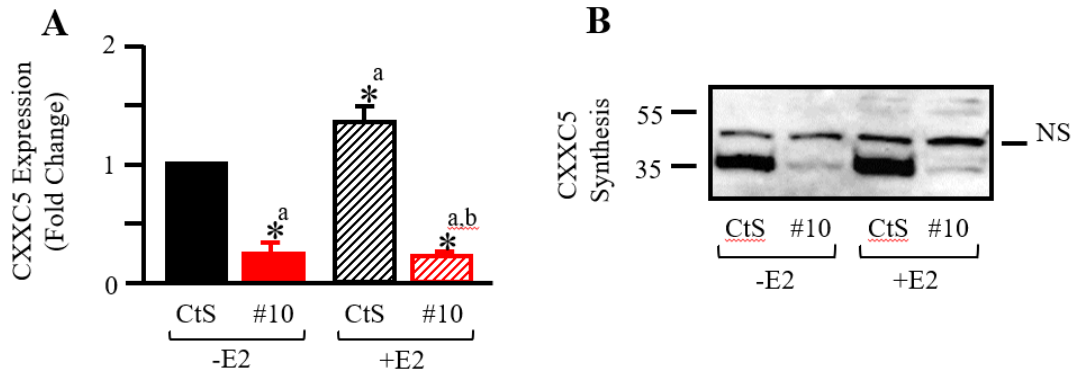


Figure 3.6. Effects of CXXC5-targeting siRNAs on CXXC5 expression and protein synthesis.

MCF7 cells were cultured and transfected as described before. (A) Total RNA and cDNA were generated from cells after 48h; cDNA was subjected to qPCR using a CXXC5-specific primer set. Results indicate fold change in mRNA levels compared to the amount observed in cells transiently transfected with CtS in the absence of E2 (CtS, -E2), which is set to 1; and are represented as the means ± S.E. of three independent biological replicates. Asterisk indicates significant change, superscript a indicates a significant difference of the group compared to the corresponding ethanol treated group (-E2); and superscript b indicated a significant difference of the group compared to the Cts transfected cells in the presence of E2 (CtS, +E2). (B) 48h after transfection, nuclear and cytoplasmic extracts of cells were isolated; Nuclear extracts were subjected to WB using CXXC5 antibody. One representing result from three independent biological replicates is shown.

To make certain that the decrease in cellular proliferation in E2 treated siRNA#10 samples as we observed in our initial experiments (Fig. 3.4.) is actually ERα mediated, we used an antiestrogen, ICI (Imperial Chemical Industries 182,780) and observed that the E2 effects on both CtS and siRNA#10 transfected cells were neutralized (Fig. 3.7.).

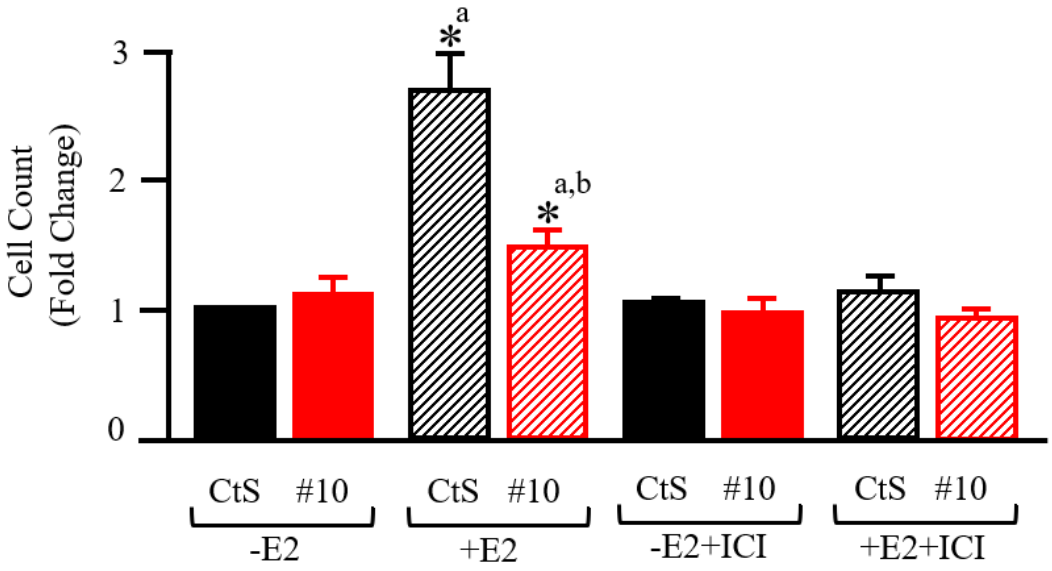


Figure 3.7. The effects of ligands on cellular growth of cells transfected with CXXC5-specific siRNA (siRNA#10) or the control scrambled siRNA (CtS).

MCF7 cells were cultures in 10% CD-FBS containing medium for 48h prior transient transfection with CtS or siRNA#10 in the absence (0.01% EtOH, -E2), the presence of 10^{-8} M E2 (+E2) and/or 10^{-6} M E2 ICI for 48h. Cells were then collected and counted using hemocytometer. Results indicate fold change in cellular growth compared to the cell number of the cells transfected with CtS in the absence of E2 (CtS, -E2), which is set to 1; and are the means ± S.E. of three independent biological replicates. Asterisk indicates significant change, superscript "a" indicates a significant difference of the group compared to CtS transfected cells in the absence of E2 (CtS, -E2); and superscript "b" indicated a significant difference of the group compared to the Cts transfected cells in the presence of E2 (CtS, +E2).

These results indicate that CXXC5 plays an important role in the cellular proliferation mediated by the E2-ERα signaling.

3.4. CXXC5 Knock-Down Alters Cell Cycle Progression

Since we observed that CXXC5 knock-down represses cell proliferation in the presence of E2, we were curious to know whether this repression was caused by any alterations in cell cycle phases. To address this question, we performed cell cycle analysis. MCF7 cells were grown in 10% CD-FBS containing medium for 48 hours, then cells were transiently transfected with CtS or siRNA#10 in the absence (-E2, 0.01% EtOH) or the presence of E2 (E2; 10^{-8} M). After an additional 48 hours, cells were collected and gone cell cycle analysis (Fig. 3.8.).

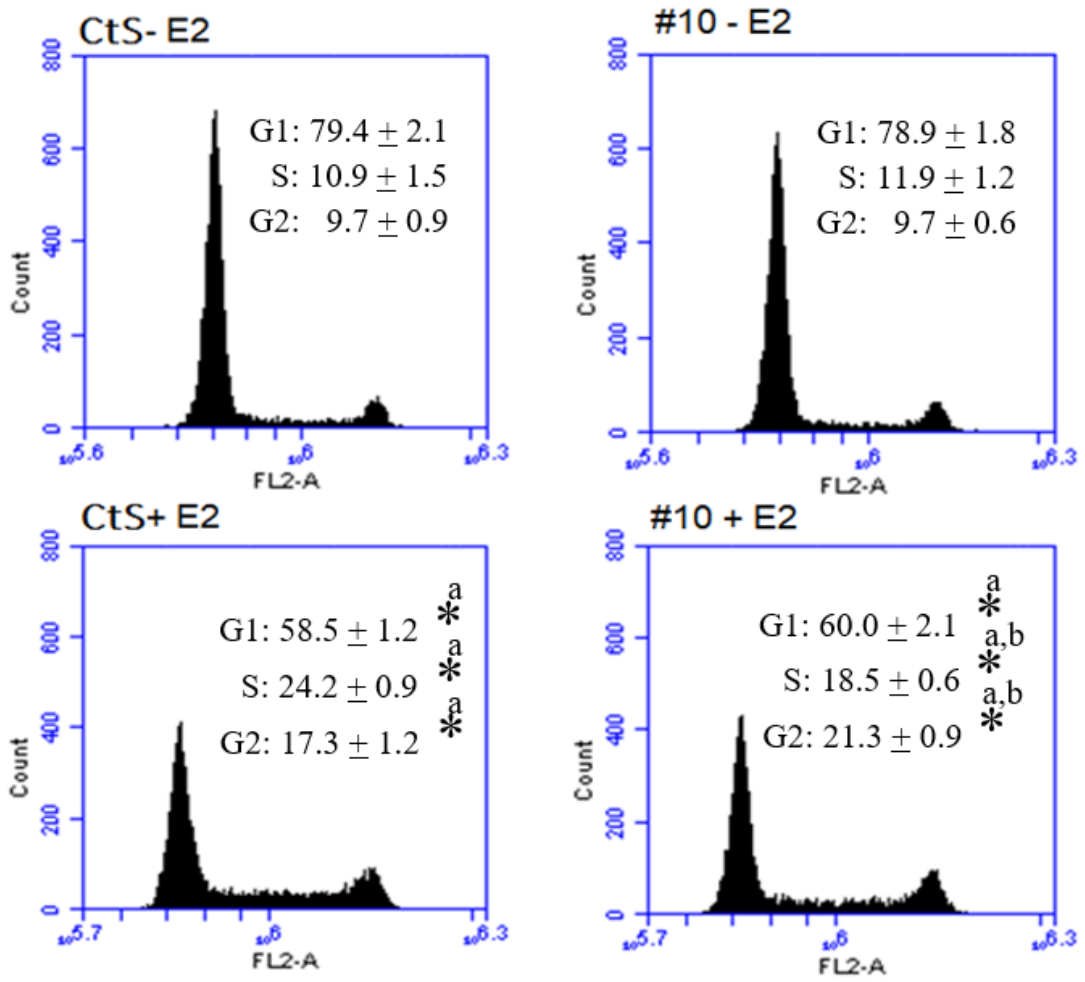


Figure 3.8. Effects of CXXC5-targeting siRNA on cell cycle phases.

MCF7 cells were cultured in 10% CD-FBS containing medium for 48h, and then transiently transfected with CtS or siRNA#10 in the absence (0.01% EtOH, -E2) or the presence of 10^{-8} M E2 (+E2) for 48h. Cells were then collected, washed, processed and subjected to cell cycle analysis by flow cytometry. Results indicate the percentage of cells in G1, G2 and S phases, and are the mean ± S.E of three independent biological replicates. Asterisk indicates significant change, superscript a indicates a significant difference of the group compared to the corresponding ethanol treated group (-E2); and superscript b indicated a significant difference of the group compared to the Cts transfected cells in the presence of E2 (CtS, +E2). In cells transfected with CtS, E2 treatment causes an increase in the population of the cells in the S and G2 phases which causes the percentage of G1 phase to decrease. On the other hand, siRNA#10 transfection reduced the E2-mediated increase in S phase cell population, caused an increment in the population in the G2 phase but had no effect on the G1 phase cell percentage. This shows that CXXC5 plays a role in events in the E2 mediated S phase and in the G2/G1 transition.

Then we asked whether CXXC5 also contributes to cell death and as a result decreases cell proliferation. To address this issue, again MCF7 cells were grown in 10% CD-

FBS containing medium for 48 hours prior to transient transfection by CtS or siRNA#10 in the absence or presence of E2 (E2; 10^{-8} M). After 48 hours, cells were collected and subjected to Annexin V assay and western blot for detection of the cleaved poly (ADP-ribose) polymerase-1 (PARP1) as a marker for apoptosis.

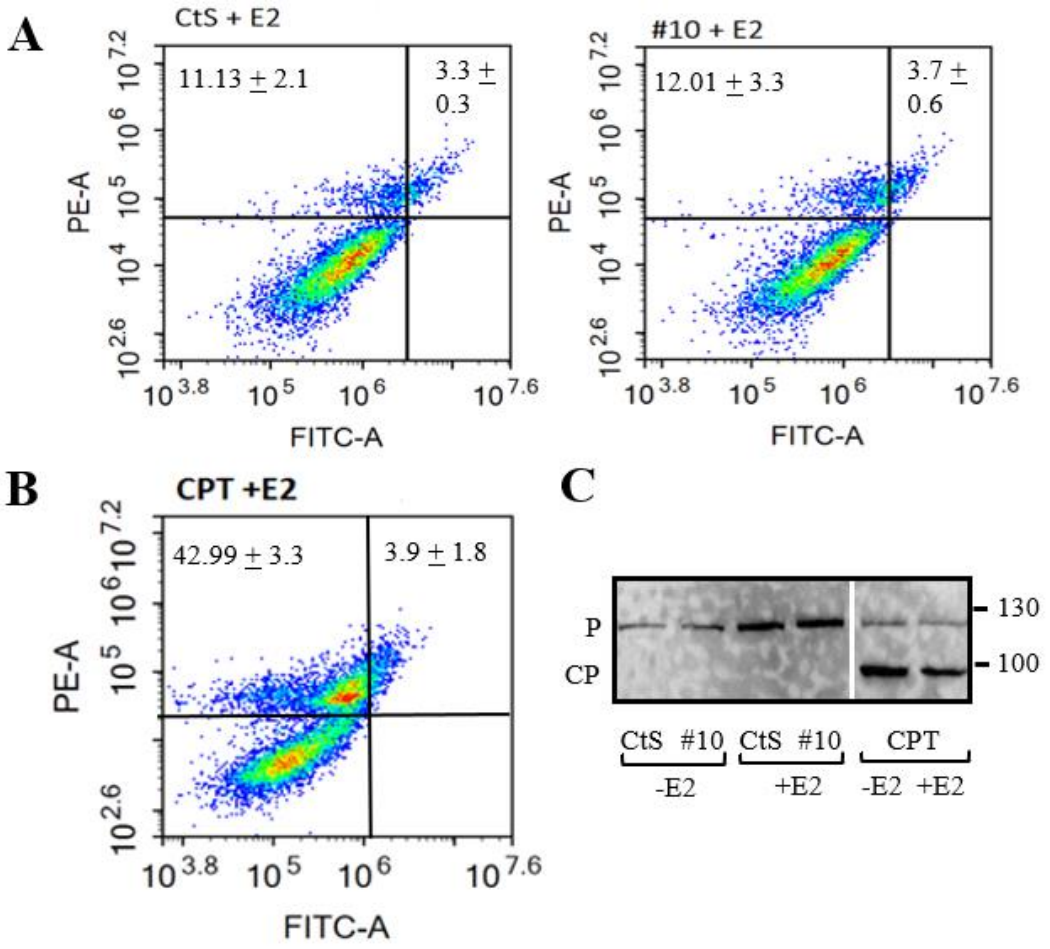


Figure 3.9. Effects of CXXC5-targeting siRNA on cell death.

(A & C) Cells grown in 10% CD-FBS containing medium for 48h were transfected with CtS or siRNA#10 in the presence of 10^{-8} M E2 (+E2) for 48h. Cells were then collected and subjected to Annexin V assay using flow cytometry (A) or WB (C) using an antibody detecting PARP1 (both full length and cleaved). (B & C) Untransfected MCF7 cells were cultured in 10% CD-FBS containing medium for 48h, then treated with the apoptosis inducer camptothecin (2 nM) for 24h in the absence (-E2) or presence of E2 (+E2). After 24 hours, these cells were then collected, processed, and subjected to Annexin V assay by flow cytometry (B) or WB (C) using PARP1-specific antibody. Results show

the percent change in apoptosis (upper right quadrant) and death (upper left quadrant) cells and are shown as the mean ± S.E. of three biological replicates.

There was no difference in the results regarding cell death between siRNA#10 and CtS transfected cells whether or not cells were treated with E2. This was supported with PARP1 western blot in which CXXC5 knock-down had no effects on the cleavage of PARP1, which was increased in the presence of E2 (Fig. 3.9.A and C). The apoptosis inducer camptothecin (CPT) treated cells, on the other hand, were apoptotic (shown in the presence of E2) (Fig. 3.9.B) and cleaved PARP1 (CP) was detectable in the absence or presence of E2 (Fig. 3.9.C). Annexin V and western blot results come hand in hand together and show that CXXC5 which contributes to E2-mediated cellular proliferation through events associated with DNA synthesis, does not seem to have any roles in cell death.

3.5. CXXC5 alone and together with E2-ERα is involved in the regulation of gene expressions

We knew that CXXC5 binds to non-methylated CpG dinucleotides and our experiments suggest that it contributes to E2-mediated cellular proliferation. It is possible that CXXC5 alone or together with E2-ERα alters gene expressions. To test this prediction, we used Nanostring assay which is a multiplex gene expression analysis. We used PanCancer Pathway Panel (nCounter, Nanostring) which contains cancer-associated 770 genes involved in different regulatory pathways such as cell cycle regulation, apoptosis, DNA damage control, transcriptional regulation and chromatin modification. Nanostring directly measures the mRNA transcript levels of genes [41], [42]. To prepare samples for this experiment, MCF7 cells were grown in 10% CD-FBS containing medium for 48 hours prior to transient transfection with CtS or siRNA#10 in the absence of E2 (in CD-FBS containing medium). 48 hours after transfection, transfected cells were treated without (-E2, 0.01% EtOH) or with E2 (10^{-8} M E2) for 3 hours. 3 hours is a critical time point because the primary effects of E2 treatment is detectable meaning that the expression of E2 target genes change without having enough time to alter protein synthesis [43], [44]. After these 3 hours, cells were

collected, and total RNA was isolated. We had four groups in this experiment, 1) Cts in the absence of E2 (-E2, 0.01% EtOH), 2) siRNA #10 in the absence of E2 (-E2, 0.01% EtOH), 3) CtS in the presence of E2 (+E2, 10^{-8} M E2), and 4) siRNA#10 in the presence of E2 (+E2, 10^{-8} M E2). Each group had three biological replicates. The isolated total RNA from each group was subjected to the Nanostring Panel profiler. RNA quality control and background correction was done and data from other three groups were normalized to the first group (CtS transfected cells in the absence of E2) because among all four groups, the first group, having no CXXC5 knock-down and no E2 treatment, served as our control.

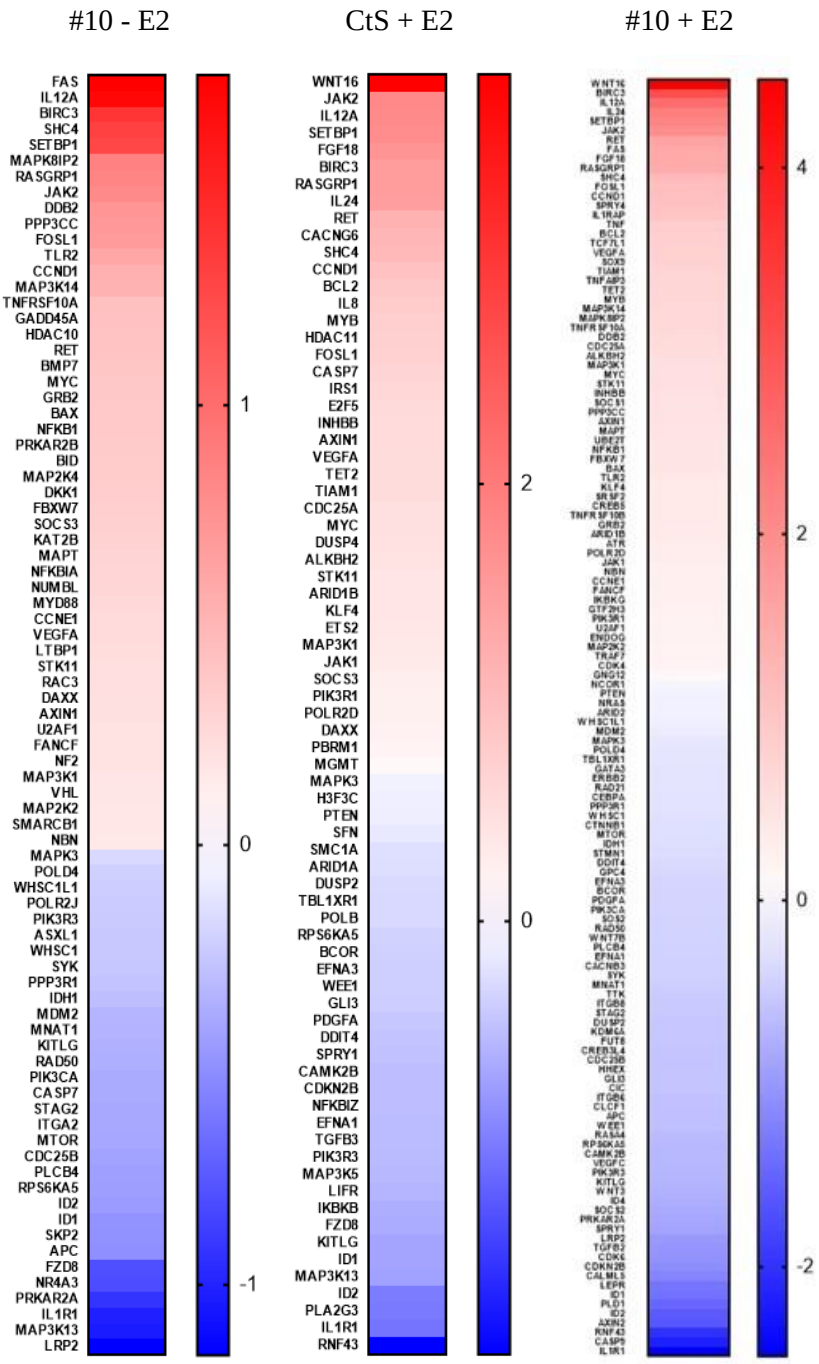
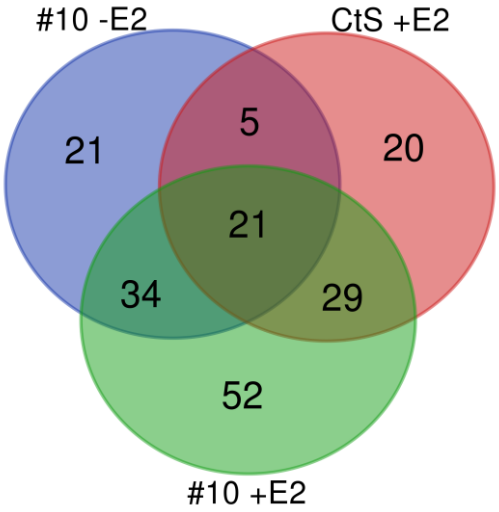


Figure 3.10. Analysis of Gene Expression with nCounter PanCancer Pathway Panel shown as heatmaps.

MCF7 cells cultured in 10% CD-FBS medium for 48h were transiently transfected by 10 nM of CtS or siRNA#10. 48h after transfection, cells were treated without (ethanol, 0.01%, -E2) or with 10^{-8} M

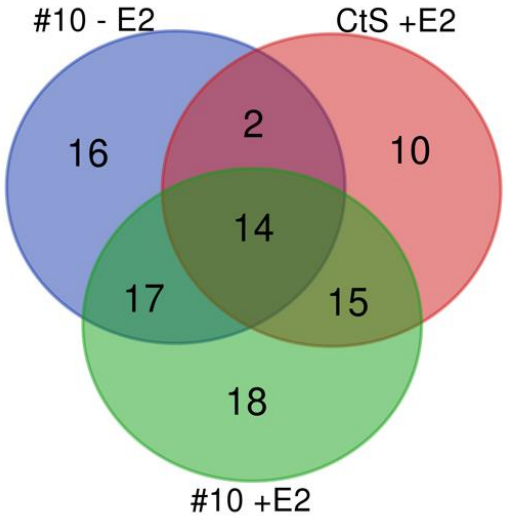
of E2 (E2) for 3 hours. Total RNA was isolated, and 50 ng of RNA was processed for and subjected to nCounter PanCancer Pathway Panel gene expression analysis with nCounter Digital Analyzer. Nanostring nSolver™ 3.0 Analysis software and its Advanced Analysis plug-in were used for quality control, data normalization and differential expression analyses. Results of each treatment group, which are the mean of three independent biological replicates' data in log2 scale and are shown as heat maps with increasing (red) and decreasing levels (blue), were normalized to the results of the first group (CtS, -E2).

A



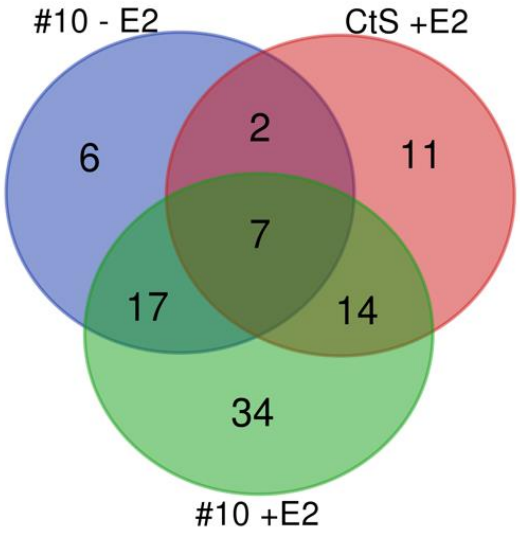
Names	total	elements
#10 +E2 #10 -E2 CtS +E2	21	ID1 AXIN1 STK11 VEGFA SETBP1 MYC MAPK3 RPS6KA5 MAP3K1 JAK2 IL12A IL1R1 FOSL1 RASGRP1 RET ID2 SHC4 CCND1 PIK3R3 BIRC3 KITLG
#10 -E2 CtS +E2	5	MAP3K13 CASP7 SOCS3 DAXX FZD8
#10 +E2 #10 -E2	34	SYK MTOR NBN TNFRSF10A MAP3K14 MNAT1 APC U2AF1 TLR2 MAP2K2 MAPK8IP2 FANCF LRP2 RAD50 CDC25B BAX PLCB4 STAG2 IDH1 MAPT PPP3R1 POLD4 WHSC1 MDM2 DDB2 NFKB1 PRKAR2A FBXW7 CCNE1 WHSC1L1 PIK3CA PPP3CC FAS GRB2
#10 +E2 CtS +E2	29	PDGFA BCL2 IL24 PTEN RNF43 TET2 DUSP2 CDKN2B CAMK2B TBL1XR1 GLI3 EFNA1 BCOR FGF18 DDIT4 JAK1 INHBB ARID1B PIK3R1 EFNA3 SPRY1 WNT16 ALKBH2 POLR2D WEE1 TIAM1 KLF4 MYB CDC25A
#10 -E2	21	NFKBIA SMARCB1 BID SKP2 PRKAR2B MAP2K4 ITGA2 POLR2J MYD88 NR4A3 NUMBL LTBP1 NF2 GADD45A ASXL1 BMP7 HDAC10 KAT2B VHL DKK1 RAC3
CtS +E2	20	SFN PLA2G3 E2F5 CACNG6 IRS1 ETS2 IL8 MGMT TGFB3 NFKBIZ MAP3K5 SMC1A ARID1A POLB H3F3C HDAC11 IKBKB LIFR DUSP4 PBRM1
#10 +E2	52	TRAF7 CDK4 SRSF2 LEPR SOCS2 IL1RAP RASA4 TGFB2 GPC4 SOX9 ITGB8 CDK6 TCF7L1 CALML5 TNF SOS2 IKBKG CASP9 GTF2H3 ATR VEGFC GATA3 NCOR1 ITGB6 TTK HHEX WNT7B KDM6A ENDOG WNT3 CLCF1 ARID2 GNG12 PLD1 TNFAIP3 TNFRSF10B FUT8 SOCS1 ID4 SPRY4 CREB3L4 RAD21 STMN1 NRAS UBE2T CTNNB1 CACNB3 AXIN2 CEBPA CIC CREB5 ERBB2

B **Increased**



Names	total	elements
#10 +E2 #10 - E2 CtS +E2	14	MAP3K1 JAK2 IL12A FOSL1 AXIN1 RASGRP1 RET STK11 VEGFA SETBP1 MYC SHC4 CCND1 BIRC3
#10 - E2 CtS +E2	2	SOCS3 DAXX
#10 +E2 #10 - E2	17	BAX NBN TNFRSF10A MAP3K14 MAPT U2AF1 TLR2 DDB2 NFKB1 FBXW7 CCNE1 MAP2K2 PPP3CC MAPK8IP2 FANCF FAS GRB2
#10 +E2 CtS +E2	15	PIK3R1 BCL2 IL24 TET2 WNT16 ALKBH2 POLR2D TIAM1 FGF18 KLF4 JAK1 INHBB ARID1B MYB CDC25A
#10 - E2	16	NFKBIA SMARCB1 BID GADD45A PRKAR2B MAP2K4 BMP7 HDAC10 KAT2B VHL MYD88 DKK1 NUMBL LTBP1 NF2 RAC3
CtS +E2	10	MGMT CASP7 E2F5 CACNG6 HDAC11 IRS1 DUSP4 PBRM1 ETS2 IL8
#10 +E2	18	TRAF7 CDK4 SRSF2 GNG12 TNFAIP3 TNFRSF10B IL1RAP SOX9 SOCS1 SPRY4 TCF7L1 TNF UBE2T IKBKG GTF2H3 ATR CREB5 ENDOG

C **Decreased**



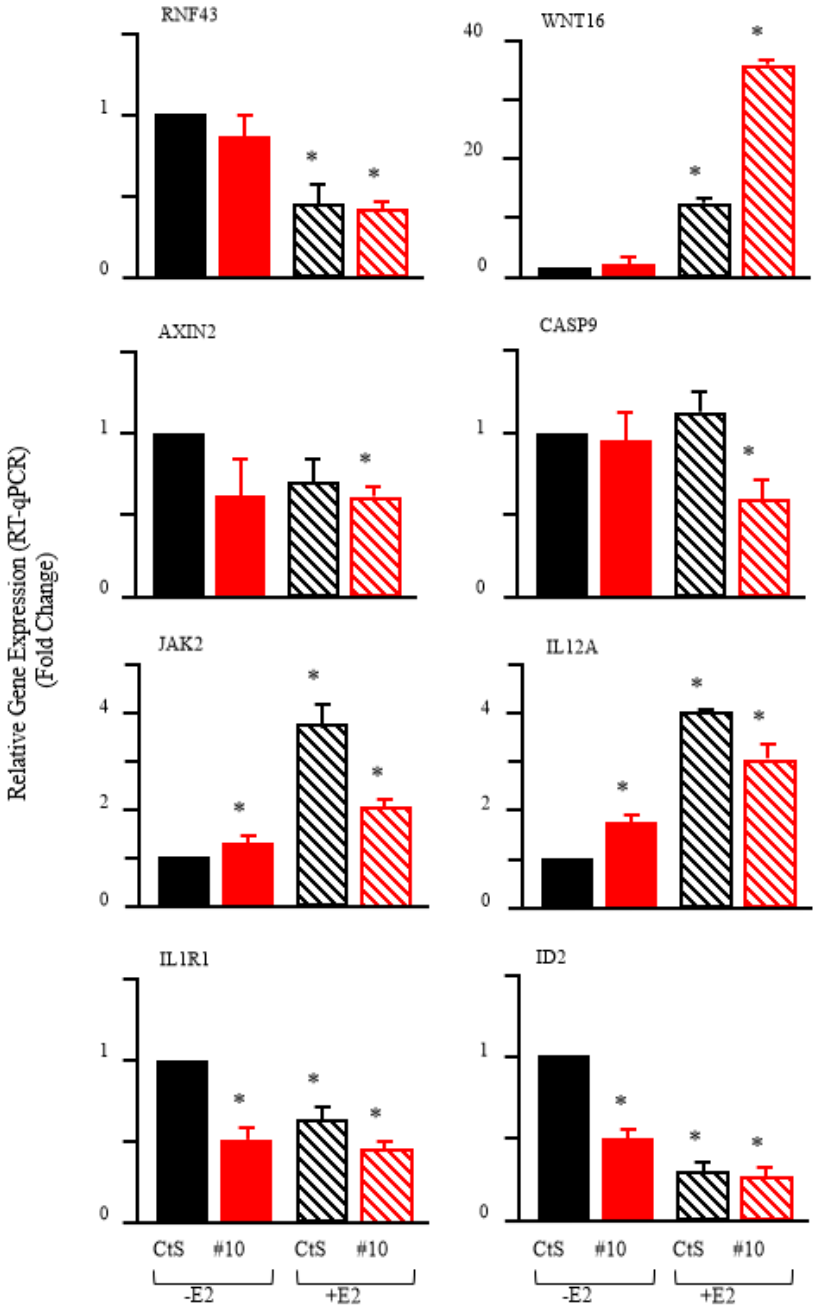
Names	total	elements
#10 +E2 #10 - E2 CtS +E2	7	IL1R1 ID1 ID2 MAPK3 PIK3R3 RPS6KA5 KITLG
#10 - E2 CtS +E2	2	MAP3K13 FZD8
#10 +E2 #10 - E2	17	LRP2 SYK CDC25B RAD50 MTOR PLCB4 STAG2 MNAT1 IDH1 PPP3R1 APC POLD4 WHSC1 MDM2 PRKAR2A WHSC1L1 PIK3CA
#10 +E2 CtS +E2	14	PDGFA PTEN RNF43 EFNA3 DUSP2 CDKN2B CAMK2B SPRY1 TBL1XR1 GLI3 WEE1 EFNA1 BCOR DDIT4
#10 - E2	6	SKP2 CASP7 ASXL1 POLR2J ITGA2 NR4A3
CtS +E2	11	TGFB3 NFKBIZ MAP3K5 SFN SMC1A PLA2G3 ARID1A POLB H3F3C IKBKB LIFR
#10 +E2	34	CLCF1 ARID2 LEPR PLD1 SOCS2 RASA4 TGFB2 FUT8 GPC4 ITGB8 ID4 CDK6 CALML5 CREB3L4 SOS2 RAD21 NRAS STMN1 CTNNB1 CACNB3 CASP9 AXIN2 CEBPA CIC VEGFC GATA3 NCOR1 ITGB6 TTK HHEX ERBB2 WNT7B KDM6A WNT3

Figure 3.11. Analysis of Gene Expression with nCounter PanCancer Pathway Panel shown as Venn Diagrams.

Results of nCounter PanCancer Pathway Panel are also presented as Venn diagrams generated by using a Venn Diagram web tool (http://bioinformatics.psb.ugent.be/webtools/Venn/) showing the cumulative (A), increasing (B) or decreasing (C) expressions of gene sets. Note that *CASP7* is a common gene regulated in siRNA#10, -E2, Group 2, CtS, +E2, Group 3, and its expression shows directional polarity in treatment groups meaning that its expression decreases in Group 2; while increases in Group 3, which is why *CASP7* is absent in increasing and decreasing Venn diagrams commonly regulated in Group 2 and Group 3.

The results, presented as heat maps and Venn Diagrams (Fig. 3.10. and 3.11.), show that each group's treatment has regulated many mutual and distinct genes compared to the other groups. We also verified this data by performing RT-qPCR on some of the genes (Fig. 3.12.).

A


Relative Gene Expression (RT-qPCR)
(Fold Change)
RNF43
WNT16
AXIN2
CASP9
JAK2
IL12A
IL1R1
ID2
CtS
#10
CtS
#10
-E2
+E2
CtS
#10
CtS
#10
-E2
+E2

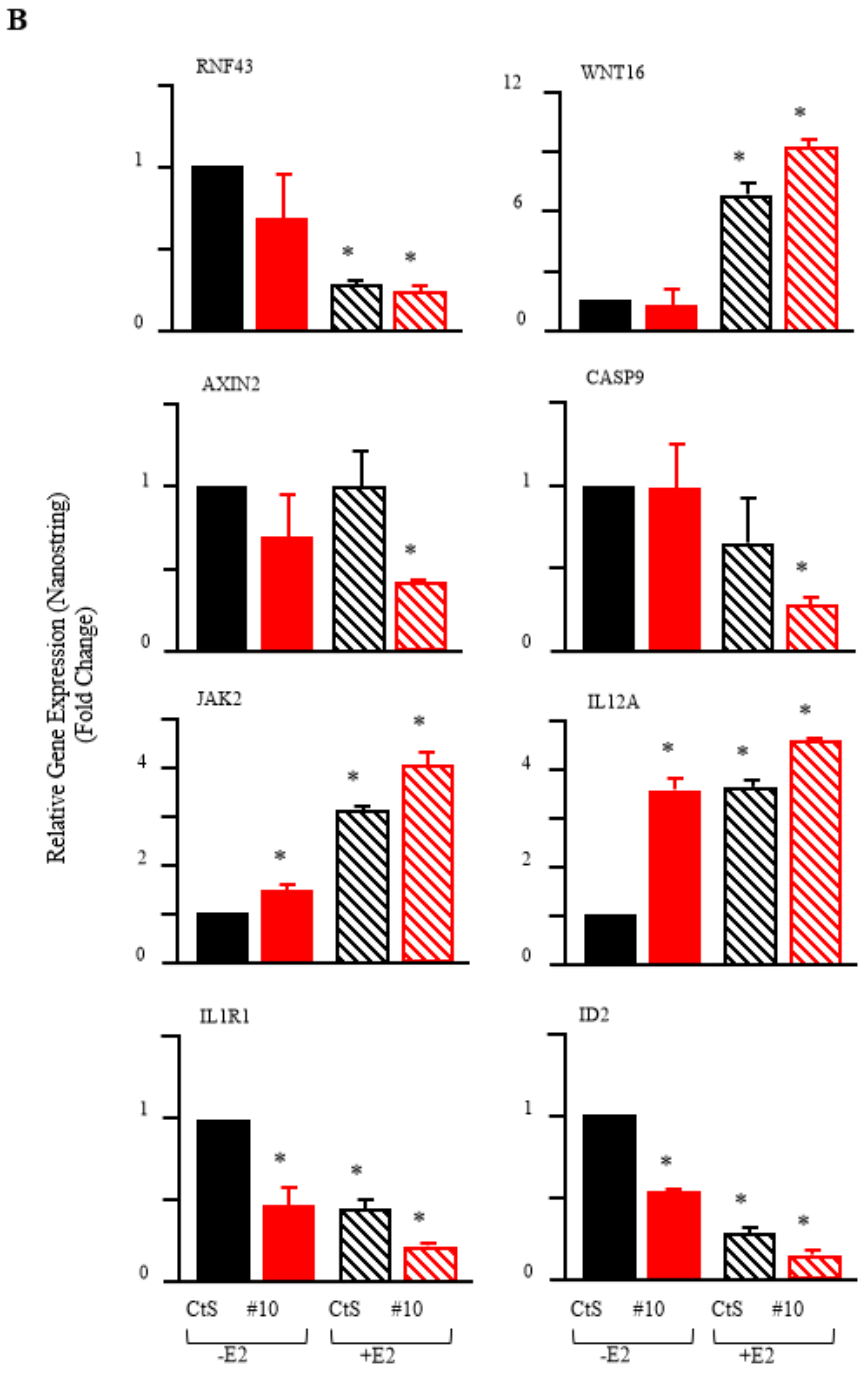


Figure 3.12. Analysis of gene expressions using RT-qPCR.

(A) Total RNA used for Nanostring nCounter PanCancer analysis were subjected to cDNA synthesis and qPCR using primer sets specific to selected genes. Results, which are the mean ± S.E. of three independent biological replicates, represent fold changes in mRNA levels compared to those observed in group 1 (CtS, -E2), which is set to 1. Asterisk indicates significant change. (B) Represented is also

the gene expression analyses of PanCancer Pathway Panel from the same genes selected for qPCR verification.

We saw that the Nanostring data and our RT-qPCR data support each other. As it can be seen in the heatmaps, Group 2, 3 and 4 each regulate 81, 75 and 136 genes, respectively (Fig. 3.10.). The second group, which is siRNA #10 in the absence of E2 (-E2, 0.01% EtOH), shows the genes that are regulated by CXXC5 alone (without E2). These genes' proteins are involved in many intracellular events ranging from signal transduction to chromatin modelling. The third group, CtS in the presence of E2 (+E2, 10^{-8} M E2), shows the genes affected by E2 treatment such as *DUSP4, ETS2, EFNA1, HDAC11, IKBKB, IL24, INHBB, IRS1, LIFR, MAP3K5, MGMT, NFKBIZ, PDGFA, PTEN* and *TGFB3*, which were shown to be regulated by E2 [5], [44]–[51]. A set of estrogen responsive genes such as *AXIN1, CCND1, FOSL1, ID1, ID2, MYC, PIK3R3, RET, STK11* and *VEGFA* [43], [44], [52]–[55] were also listed in the fourth group which shows the effects of siRNA#10 transfection along with E2 treatment. These genes were already over-expressed (*AXIN1, CCND1, FOSL1, MYC, RET, STK11* and *VEGFA*) or repressed *(ID1, ID2* and *PIK3R3*) as a result of E2 regulation, but their expression levels increased or decreased even further when combined with the effects of CXXC5 knock-down (Fig. 3.10.). Furthermore, there are 52 genes in the fourth group that were not present in the second and third groups, meaning that CXXC5 knock-down together with E2 treatment brings forth the regulation of a large number of genes that were not significantly regulated either by E2 treatment or CXXC5 knock-down alone. Among these 52 genes, there are growth factors (*ERBB2, TGFB2, VEGFC*), cytokines (*CLCF1, SOCS1, SOCS2, TNF, TNFRSF10B, TNFAIP3*), signal transducers (*WNT3, WNT7B*) and their signaling pathways' downstream components (*AXIN2, CTNNB1, NRAS*). As there are genes distinctively regulated by the fourth group, there are also 5 genes that are absent in this group but present in the second and third groups, *MAP3K13, CASP7, SOCS3, DAXX* and *FZD8*, meaning that the effects of E2 treatment and CXXC5 merge to determine the state and level of a target gene's transcription. These data suggest that CXXC5 plays a role in E2-mediated cellular proliferation by regulating the expression of E2 target genes.

CHAPTER 4

CONCLUSION AND FUTURE DIRECTIONS

To study the functions of CXXC5 *in vitro* and *in cellula* we have shown here that CXXC5 is indeed a non-methylated CpG dinucleotide binding protein that although does not have an intrinsic effect on gene expressions, it contributes to the modulation of gene expressions, and plays an important role in E2-modulated cellular proliferation.

DNA methylation is an epigenetic mark that helps regulate chromatin structurally and functionally. In vertebrates, one of these functional regulations can be gene silencing. Cytosine residues of CpG dinucleotides can be methylated by the help of DNA methyltransferase enzymes (DNMTs) and consequently recruit methyl-CpG-binding proteins (MBPs) that act as transcription repressors [56]. It is known that roughly 60–80% of CpGs in mammalian genomic DNA are methylated [9], but in a survey to identify the locations of all CpGs in the human genome, it has been shown that 72% of promoters are associated with unmethylated DNA sequences called CpG islands (CGIs) [57], [58], which are regions with relatively high density of CpG dinucleotides. These islands are often found in the promoter regions of genes that have a characteristic transcription-associated chromatin organization and are mostly hypomethylated. It has been recently indicated that the CXXC domains of the ZF-CXXC family proteins, play important roles in epigenetic regulation by recognizing, binding and targeting various activities to CpG islands [35]. Upon binding to non-methylated cytosine in CpG island promoters, the ZF-CXXC family proteins can act directly as chromatin modifying enzymes and/or indirectly by recruiting chromatin modifiers to establish a chromatin architecture critical for gene expressions [10], [34], [59]. As previously indicated for the CXXC domain of CXXC5 [35], we further show

here that the full length CXXC5, as its CXXC domain, is a non-methylated CpG dinucleotide binding protein.

We already know that CXXC5 is an E2- responsive gene; our results here indicate that CXXC5 protein plays an important role in E2-ERα signaling-mediated gene expression. E2 signaling causes the expression of a primary set of genes with important roles in cellular processes including metabolism of proteins, transcriptions, membrane signaling cascade and receptor proteins. These primary proteins then pave the way for the regulation of late gene expressions with products involved in the modulation of DNA replication, recombination and repair, cell cycle and division, consequently in the initiation of E2-mediated cellular proliferation [43], [60]–[62]. Here we performed a gene expression analysis using the PanCancer Pathway Panel. This analysis together with our cell proliferation experiments revealed that although CXXC5 knock-down in the absence of E2 (#10, -E2) results in the regulation of a number of genes important for cellular proliferation, such as *CCND1, CCNE1, CDC25B, MYC*, it does not have any effects on cell growth when compared with CtS -E2 group. As expected, we can see that E2 treatment (CtS, +E2) increases cellular proliferation by regulating genes such as *CDC25A, MYB* and *WEE1* which are specifically regulated by E2 as well as *CCND1* and *MYC* which are commonly regulated by both E2 and CXXC5. This increase in cellular proliferation caused by E2, was abrogated with CXXC5 knock-down. This group (#10, +E2) regulates a large number of genes, some of which are commonly regulated by E2 and CXXC5 separately (*CCND1, CCNE1, CDC25B, MYB, MYC* and *WEE1*), and some are regulated specifically with this group; meaning that there are a number of genes that are not significantly regulated by E2 or CXXC5, but show an effective increase or decrease in the expression when E2 and CXXC5 knock-down effects are both present. Some of these genes have important roles that can lie behind the repression of E2-mediated cell proliferation; for example, *CDK4* and *CDK6*, encode cyclin dependent kinases critical for initiation and phase transitions. This indicates that CXXC5 has an

important role in the regulation of both primary and secondary E2 responsive gene expressions critical for cellular proliferation.

We showed that CXXC5 is involved in gene expressions and is a non-methylated CpG binding protein, but its functions and used mechanisms are still unclear. Our mammalian one hybrid results indicate that CXXC5 does not have an intrinsic effect on gene expression, but since it is a non-methylated CpG dinucleotide binder, it may bind to DNA to recruit transcription and/or epigenetic factors. In keeping with this prediction, our ongoing studies suggest that CXXC5 indeed interacts with a various transcription factors as well as DNA and histone modifiers.

Since CXXC5 knock-down in the presence of E2 caused the regulation of a large number of genes and consequently the repression of cellular proliferation, the same experiments should be carried our when CXXC5 is over expressed, but since the previous attempts of CXXC5 over expression caused cell loss, inducible system could be used to both over-express and knock-down CXXC5 at will and efficiently in a short amount of time. Nanostring experiments could be conducted for CXXC5 over expressed cells using PanCancer and even other panels to be able to further understand the functions of CXXC5 in cancer, immunology, and other pathways.

REFERENCES

[1] Y. Huang, X. Li, and M. Muyan, "Estrogen receptors similarly mediate the effects of 17β-estradiol on cellular responses but differ in their potencies," *Endocrine*, vol. 39, no. 1, pp. 48–61, Feb. 2011.

[2] J. Huang, X. Li, C. A. Maguire, R. Hilf, R. A. Bambara, and M. Muyan, "Binding of Estrogen Receptor β to Estrogen Response Element in Situ Is Independent of Estradiol and Impaired by Its Amino Terminus," *Mol. Endocrinol.*, vol. 19, no. 11, pp. 2696–2712, Nov. 2005.

[3] J. S. Lewis-Wambi and V. C. Jordan, "Estrogen regulation of apoptosis: how can one hormone stimulate and inhibit?," *Breast Cancer Res.*, vol. 11, no. 3, p. 206, 2009.

[4] P. Ascenzi, A. Bocedi, and M. Marino, "Structure–function relationship of estrogen receptor α and β: Impact on human health," *Mol. Aspects Med.*, vol. 27, no. 4, pp. 299–402, Aug. 2006.

[5] S. L. Nott *et al.*, "Genomic Responses from the Estrogen-responsive Element-dependent Signaling Pathway Mediated by Estrogen Receptor α Are Required to Elicit Cellular Alterations," *J. Biol. Chem.*, vol. 284, no. 22, pp. 15277–15288, May 2009.

[6] C. M. Klinge, "Estrogen receptor interaction with estrogen response elements," *Nucleic Acids Res.*, vol. 29, no. 14, pp. 2905–2919, Jul. 2001.

[7] M. Muyan, L. M. Callahan, Y. Huang, and A. J. Lee, "The ligand-mediated nuclear mobility and interaction with estrogen-responsive elements of estrogen receptors are subtype specific," *J. Mol. Endocrinol.*, vol. 49, no. 3, pp. 249–266, Dec. 2012.

[8] 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. 1, p. 37808, Dec. 2016.

[9] Z. D. Smith and A. Meissner, "DNA methylation: roles in mammalian development," *Nat. Rev. Genet.*, vol. 14, no. 3, pp. 204–220, Mar. 2013.

[10] 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, Jun. 2013.

[11] F. Pendino *et al.*, "Functional involvement of RINF, retinoid-inducible nuclear factor (CXXC5), in normal and tumoral human myelopoiesis," *Blood*, vol. 113, no. 14, pp. 3172–3181, Apr. 2009.

[12] P. A. Marshall *et al.*, “Discovery of novel vitamin D receptor interacting proteins that modulate 1,25-dihydroxyvitamin D3 signaling,” *J. Steroid Biochem. Mol. Biol.*, vol. 132, no. 1–2, pp. 147–159, Oct. 2012.

[13] H.-Y. Kim *et al.*, “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, Jun. 2015.

[14] H.-Y. Kim *et al.*, “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, Feb. 2014.

[15] G. Li *et al.*, “CXXC5 regulates differentiation of C2C12 myoblasts into myocytes,” *J. Muscle Res. Cell Motil.*, vol. 35, no. 5–6, pp. 259–265, Dec. 2014.

[16] X. Yan, J. Wu, Q. Jiang, H. Cheng, J.-D. J. Han, and Y.-G. Chen, “CXXC5 suppresses hepatocellular carcinoma by promoting TGF-β-induced cell cycle arrest and apoptosis,” *J. Mol. Cell Biol.*, vol. 10, no. 1, pp. 48–59, Feb. 2018.

[17] S. Aras *et al.*, “Oxygen-dependent expression of cytochrome c oxidase subunit 4-2 gene expression is mediated by transcription factors RBPJ, CXXC5 and CHCHD2,” *Nucleic Acids Res.*, vol. 41, no. 4, pp. 2255–2266, Feb. 2013.

[18] M. Zhang *et al.*, “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, Jun. 2009.

[19] X. Wang *et al.*, “CXXC5 Associates with Smads to Mediate TNF-α Induced Apoptosis,” *Curr. Mol. Med.*, vol. 13, no. 8, pp. 1385–1396, Aug. 2013.

[20] S. Choi *et al.*, “CXXC5 mediates growth plate senescence and is a target for enhancement of longitudinal bone growth,” *Life Sci. Alliance*, vol. 2, no. 2, p. e201800254, Apr. 2019.

[21] H. Kim *et al.*, “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. 375–387, Apr. 2016.

[22] S.-H. Lee *et al.*, “The Dishevelled-binding protein CXXC5 negatively regulates cutaneous wound healing,” *J. Exp. Med.*, vol. 212, no. 7, pp. 1061–1080, Jun. 2015.

[23] S.-H. Lee, S. H. Seo, D.-H. Lee, L.-Q. Pi, W.-S. Lee, and K.-Y. Choi, “Targeting of CXXC5 by a Competing Peptide Stimulates Hair Regrowth and Wound-Induced Hair Neogenesis,” *J. Invest. Dermatol.*, vol. 137, no. 11, pp. 2260–2269, Nov. 2017.

[24] M. S. Kim *et al.*, “A Novel Wilms Tumor 1 (WT1) Target Gene Negatively

Regulates the WNT Signaling Pathway," *J. Biol. Chem.*, vol. 285, no. 19, pp. 14585–14593, May 2010.

[25] X. Peng *et al.*, "CXXC5 is required for cardiac looping relating to TGFβ signaling pathway in zebrafish," *Int. J. Cardiol.*, vol. 214, pp. 246–253, Jul. 2016.

[26] S. Ma *et al.*, "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, May 2017.

[27] T. A. Melhuish, C. M. Gallo, and D. Wotton, "TGIF2 Interacts with Histone Deacetylase 1 and Represses Transcription," *J. Biol. Chem.*, vol. 276, no. 34, pp. 32109–32114, Aug. 2001.

[28] S. Knappskog *et al.*, "RINF (CXXC5) is overexpressed in solid tumors and is an unfavorable prognostic factor in breast cancer†," *Ann. Oncol.*, vol. 22, no. 10, pp. 2208–2215, Oct. 2011.

[29] P. May-Panloup *et al.*, "Molecular characterization of corona radiata cells from patients with diminished ovarian reserve using microarray and microfluidic-based gene expression profiling," *Hum. Reprod.*, vol. 27, no. 3, pp. 829–843, Mar. 2012.

[30] D. L'Hôte *et al.*, "Discovery of novel protein partners of the transcription factor FOXL2 provides insights into its physiopathological roles," *Hum. Mol. Genet.*, vol. 21, no. 14, pp. 3264–3274, Jul. 2012.

[31] M. B. Treppendahl, L. Möllgård, E. Hellström-Lindberg, P. Cloos, and K. Grønbaek, "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, Mar. 2013.

[32] F. Centritto *et al.*, "Cellular and molecular determinants of all- trans retinoic acid sensitivity in breast cancer: Luminal phenotype and RAR α expression," *EMBO Mol. Med.*, vol. 7, no. 7, pp. 950–972, Jul. 2015.

[33] I. Benedetti, A. M. De Marzo, J. Geliebter, and N. Reyes, "CXXC5 expression in prostate cancer: implications for cancer progression," *Int. J. Exp. Pathol.*, vol. 98, no. 4, pp. 234–243, Aug. 2017.

[34] X. Xiong, S. Tu, J. Wang, S. Luo, and X. Yan, "CXXC5: A novel regulator and coordinator of TGF-β, BMP and Wnt signaling," *J. Cell. Mol. Med.*, vol. 23, no. 2, pp. 740–749, Feb. 2019.

[35] C. Xu *et al.*, "DNA Sequence Recognition of Human CXXC Domains and Their Structural Determinants," *Structure*, vol. 26, no. 1, pp. 85-95.e3, Jan. 2018.

[36] T. Andersson *et al.*, "CXXC5 Is a Novel BMP4-regulated Modulator of Wnt Signaling in Neural Stem Cells," *J. Biol. Chem.*, vol. 284, no. 6, pp. 3672–3681, Feb. 2009.

[37] M.-Y. Kim *et al.*, "CXXC5 plays a role as a transcription activator for myelin genes on oligodendrocyte differentiation," *Glia*, vol. 64, no. 3, pp. 350–362, Mar. 2016.

[38] Y. Luo, A. Batalao, H. Zhou, and L. Zhu, "Mammalian Two-Hybrid System: A Complementary Approach to the Yeast Two-Hybrid System," *Biotechniques*, vol. 22, no. 2, pp. 350–352, Feb. 1997.

[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. 1810–1827, Aug. 2002.

[40] X. Nan, F. J. Campoy, and A. Bird, "MeCP2 Is a Transcriptional Repressor with Abundant Binding Sites in Genomic Chromatin," *Cell*, vol. 88, no. 4, pp. 471–481, Feb. 1997.

[41] H. Wang *et al.*, "NanoStringDiff: a novel statistical method for differential expression analysis based on NanoString nCounter data," *Nucleic Acids Res.*, p. gkw677, Jul. 2016.

[42] D. C. Kahraman, T. Kahraman, and R. C. Atalay, "Differential alteration of IL-8 in liver cancer stem cell enrichment in response to PI3K/Akt/mTOR inhibitors and sorafenib," in *Experimental and Molecular Therapeutics*, 2018, pp. 3937–3937.

[43] V. Bourdeau, J. Deschenes, D. Laperriere, M. Aid, J. H. White, and S. Mader, "Mechanisms of primary and secondary estrogen target gene regulation in breast cancer cells," *Nucleic Acids Res.*, vol. 36, no. 1, pp. 76–93, Nov. 2007.

[44] J. Frasor, J. M. Danes, B. Komm, K. C. N. Chang, C. R. Lyttle, and B. S. Katzenellenbogen, "Profiling of Estrogen Up- and Down-Regulated Gene Expression in Human Breast Cancer Cells: Insights into Gene Networks and Pathways Underlying Estrogenic Control of Proliferation and Cell Phenotype," *Endocrinology*, vol. 144, no. 10, pp. 4562–4574, Oct. 2003.

[45] A. R. Patterson, X. Mo, A. Shapiro, K. E. Wernke, T. K. Archer, and C. J. Burd, "Sustained Reprogramming of the Estrogen Response After Chronic Exposure to Endocrine Disruptors," *Mol. Endocrinol.*, vol. 29, no. 3, pp. 384–395, Mar. 2015.

[46] A. S. Pechenino and K. M. Frick, "The effects of acute 17β-estradiol treatment on gene expression in the young female mouse hippocampus," *Neurobiol. Learn. Mem.*, vol. 91, no. 3, pp. 315–322, Mar. 2009.

[47] E. C. Chang, J. Frasor, B. Komm, and B. S. Katzenellenbogen, “Impact of Estrogen Receptor β on Gene Networks Regulated by Estrogen Receptor α in Breast Cancer Cells,” *Endocrinology*, vol. 147, no. 10, pp. 4831–4842, Oct. 2006.

[48] L. Mauro *et al.*, “Estradiol Increases IRS-1 Gene Expression and Insulin Signaling in Breast Cancer Cells,” *Biochem. Biophys. Res. Commun.*, vol. 288, no. 3, pp. 685–689, Nov. 2001.

[49] A. Paranjpe *et al.*, “New insights into estrogenic regulation of O6-methylguanine DNA-methyltransferase (MGMT) in human breast cancer cells: Co-degradation of ER-α and MGMT proteins by fulvestrant or O6-benzylguanine indicates fresh avenues for therapy,” *J. Biomed. Res.*, vol. 30, no. 5, pp. 393–410, Sep. 2016.

[50] G. Shanker, M. Sorci-Thomas, and M. R. Adams, “Estrogen modulates the inducible expression of platelet-derived growth factor mrna by monocyte/macrophages,” *Life Sci.*, vol. 56, no. 7, pp. 499–507, Jan. 1995.

[51] N. D. Andruska *et al.*, “Estrogen receptor α inhibitor activates the unfolded protein response, blocks protein synthesis, and induces tumor regression,” *Proc. Natl. Acad. Sci.*, vol. 112, no. 15, pp. 4737–4742, Apr. 2015.

[52] N.-O. Chimge *et al.*, “RUNX1 prevents oestrogen-mediated AXIN1 suppression and β-catenin activation in ER-positive breast cancer,” *Nat. Commun.*, vol. 7, no. 1, p. 10751, Apr. 2016.

[53] A. J. M. O’Donnell, K. G. Macleod, D. J. Burns, J. F. Smyth, and S. P. Langdon, “Estrogen receptor-α mediates gene expression changes and growth response in ovarian cancer cells exposed to estrogen,” *Endocr. Relat. Cancer*, vol. 12, no. 4, pp. 851–866, Dec. 2005.

[54] K. Sengupta, S. Banerjee, N. Saxena, and S. K. Banerjee, “Estradiol-induced vascular endothelial growth factor-A expression in breast tumor cells is biphasic and regulated by estrogen receptor-alpha dependent pathway.,” *Int. J. Oncol.*, vol. 22, no. 3, pp. 609–14, Mar. 2003.

[55] K. J. McInnes, K. A. Brown, N. I. Hunger, and E. R. Simpson, “Regulation of LKB1 expression by sex hormones in adipocytes,” *Int. J. Obes.*, vol. 36, no. 7, pp. 982–985, Jul. 2012.

[56] P.-A. Defossez and I. Stancheva, “Biological Functions of Methyl-CpG-Binding Proteins,” 2011, pp. 377–398.

[57] S. Saxonov, P. Berg, and D. L. Brutlag, “A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters,” *Proc. Natl. Acad. Sci.*, vol. 103, no. 5, pp. 1412–1417, Jan. 2006.

[58] A. M. Deaton and A. Bird, “CpG islands and the regulation of transcription,”

Genes Dev., vol. 25, no. 10, pp. 1010–1022, May 2011.

[59] N. P. Blackledge and R. Klose, "CpG island chromatin," *Epigenetics*, vol. 6, no. 2, pp. 147–152, Feb. 2011.

[60] N. Hah *et al.*, "A Rapid, Extensive, and Transient Transcriptional Response to Estrogen Signaling in Breast Cancer Cells," *Cell*, vol. 145, no. 4, pp. 622–634, May 2011.

[61] V. Jagannathan and M. Robinson-Rechavi, "Meta-analysis of estrogen response in MCF-7 distinguishes early target genes involved in signaling and cell proliferation from later target genes involved in cell cycle and DNA repair," *BMC Syst. Biol.*, vol. 5, no. 1, p. 138, 2011.

[62] L. Cicatiello *et al.*, "Estrogen Receptor α Controls a Gene Network in Luminal-Like Breast Cancer Cells Comprising Multiple Transcription Factors and MicroRNAs," *Am. J. Pathol.*, vol. 176, no. 5, pp. 2113–2130, May 2010.

APPENDICES

A. The List of RT-qPCR Primers

Table A.1. *RT-qPCR Primer Sequences*

	Forward Primer Sequence (5' to 3')	Reverse Primer Sequence (5' to 3')
CXXC5	CGGTGGACAAAAGCAACCCTAC	CGCTTCAGCATCTCTGTGGACT
RPLP0	GGAGAAACTGCTGCCTCATA	GGAAAAAGGAGGTCTTCTCG
AXIN2	CAAACTTTCGCCAACCGTGGTTG	GGTGCAAAGACATAGCCAGAACC
CASP9	GTTTGAGGACCTTCGACCAGCT	CAACGTACCAGGAGCCACTCTT
ID2	TTGTCAGCCTGCATCACCAGAG	AGCCACACAGTGCTTTGCTGTC
IL12A	TGCCTTCACCACTCCCAAAACC	CAATCTCTTCAGAAGTGCAAGGG
IL1R1	GTGCTTTGGTACAGGGATTCCTG	CACAGTCAGAGGTAGACCCTTC
JAK2	CCAGATGGAAACTGTTCGCTCAG	GAGGTTGGTACATCAGAAACACC
RNF43	GGTTACATCAGCATCGGACTTGC	ATGCTGGCGAATGAGGTGGAGT
WNT16	TCGGAAACACCACGGGCAAAGA	GCGGCAGTCTACTGACATCAAC

B. siRNA Sequences

Table 0.1. *siRNA Sequences*

siRNAs	siRNA sequence (5’ to 3’)	Target
Hs_CXXC5_2	CAGCAGTTGTAGGAATCGAAA	In the ORF of *CXXC5*
Hs_CXXC5_10	TCAGATTTGCAAATTCAGAAA	In the ORF of *CXXC5*

C. RT-qPCR Results Samples

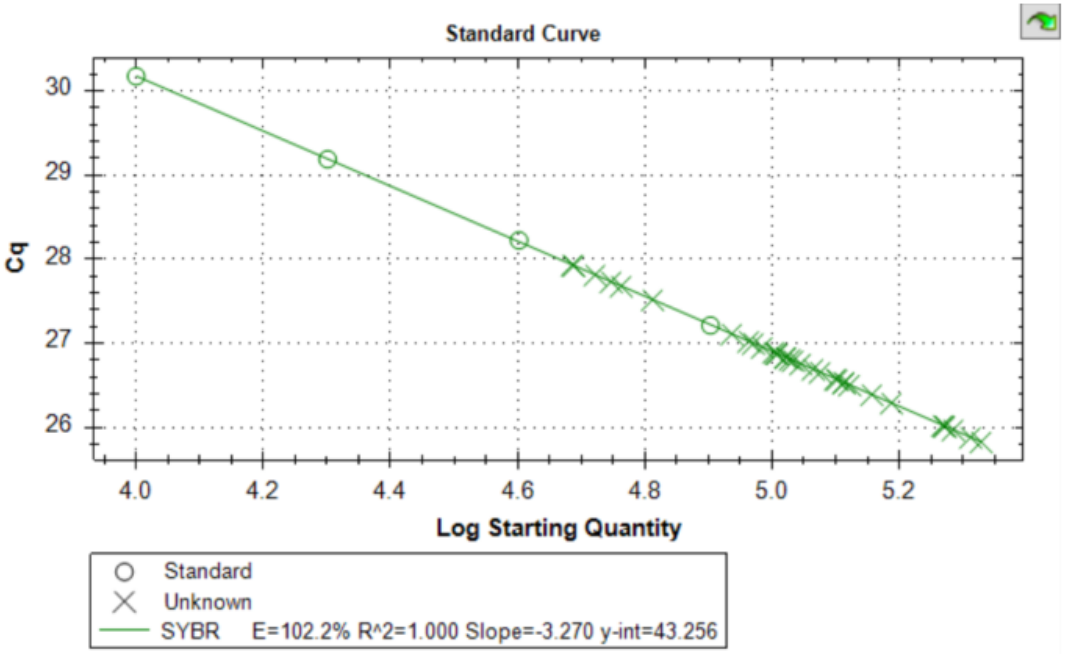


Figure 0.1. Sample Standard Curve of a RT-qPCR

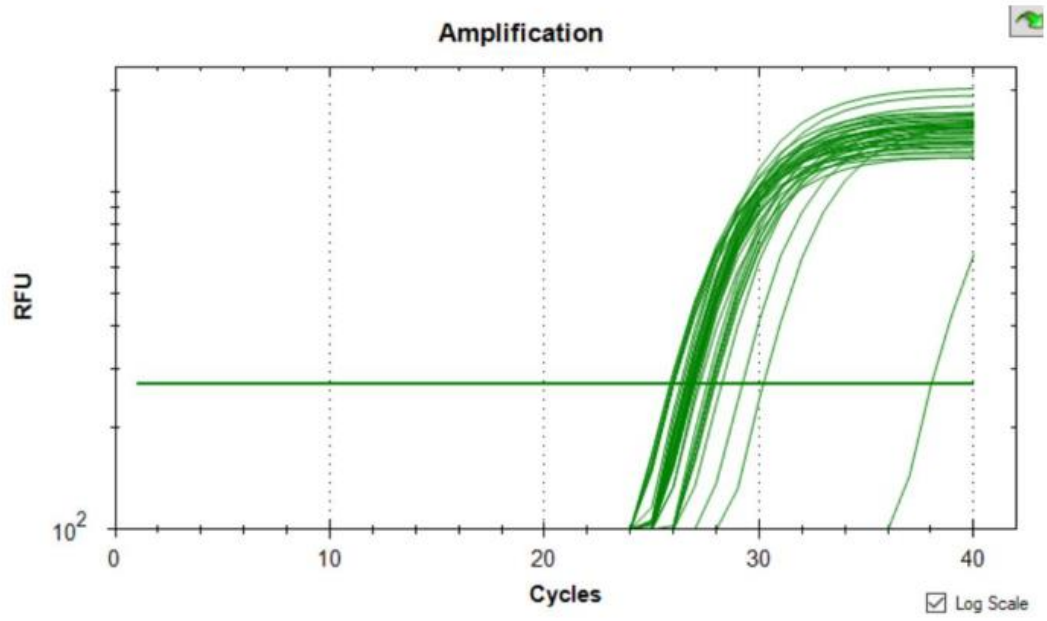


Figure 0.2. Sample Amplification Curve

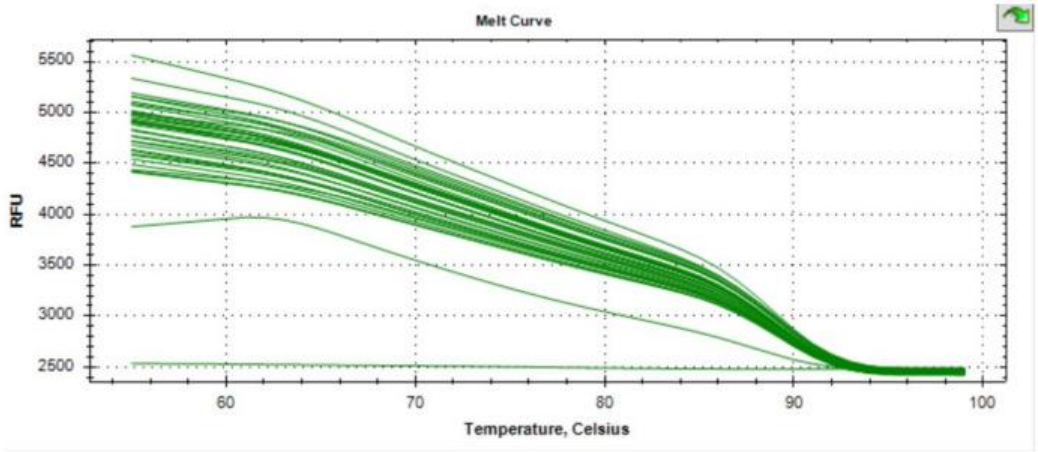


Figure 0.3. Sample Melt Curve

D. Buffers and Solutions

6X SDS Loading Buffer

0.35 M Tris-HCl pH:6.8

10.210% (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

Dissolve by heating

Sodium Dodecyl Sulfate 20 g

dH 2 O 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 2 O

Dissolve with the help of heat

Filter sterilization w/22µm filter

Acrylamide 29 g

Bisacrylamide 1 g

Protein Running Buffer

250 mM Tris

2500mM Glycine

1 % SDS

Complete volume to 1 l w/ dH_2O

10x Transfer Buffer

for 1x @4 o 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

E. Genomic DNA Contamination Control of RNA Isolation with PCR Using GAPDH Primers

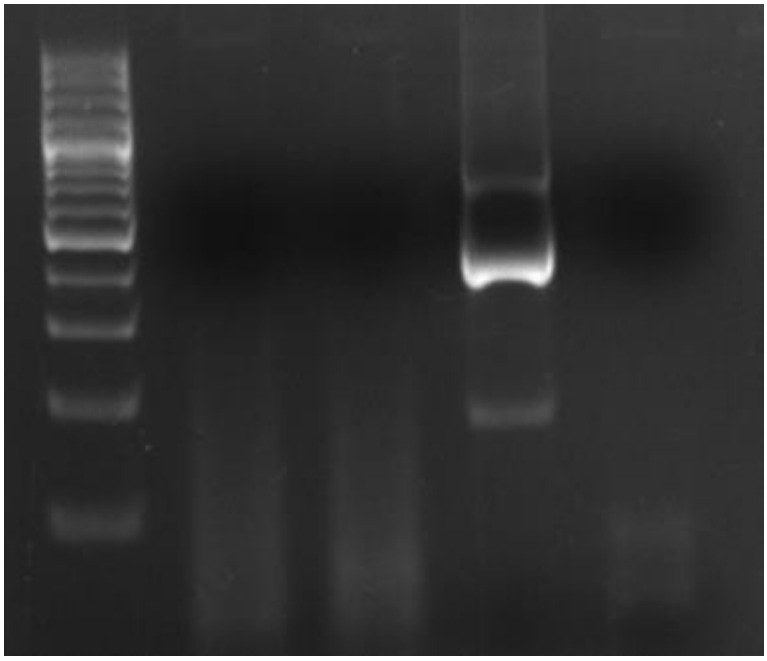


Figure 0.1. Agarose gel image of genomic DNA contamination control PCR with gDNA positive control

F. cDNA Conversion Control

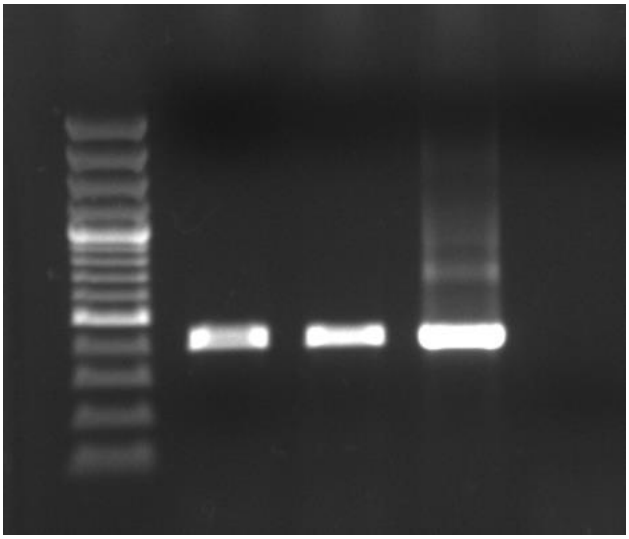


Figure 0.1. Agarose gel image of cDNA conversion control PCR with GAPDH cDNA positive controls

G. MIQE Checklist

Table 0.1. *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 G.1. 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 G.1. 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 G.1. 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

H. Charcoal-Coated Dextran Treatment of Fetal Bovine Serum

500 ml of fetal bovine serum and 10 g of charcoal coated dextran (Charcoal, dextran coated, Sigma, USA, C6241) were mixed together and stirred overnight at 4°C. Then the mixture was centrifuged at 10800 g for 30 min at 4°C to pellet the charcoal. The same steps were carried out again using 10 g of Charcoal. This time the mixture was stirred at 4°C for 4-6 hours and the same centrifugation step was followed. Supernatant was taken into biological safety cabinet and vacuum filtered with 0.45 μM sterile filter unit (Corning, polystyrene, cellulose acetate membrane, low protein binding, Sigma, USA, CLS430770)

I. nCounter Pan-Cancer Pathways Panel Gene List

Table 0.1. *List of genes present in nCounter Pan-Cancer Panel*

Official Symbol	Accession	Official Symbol	Accession
ABL1	NM_005157.3	ITGA3	NM_005501.2
ACVR1B	NM_004302.3	ITGA6	NM_000210.1
ACVR1C	NM_145259.2	ITGA7	NM_002206.1
ACVR2A	NM_001616.3	ITGA8	NM_003638.1
AKT1	NM_005163.2	ITGA9	NM_002207.2
AKT2	NM_001626.2	ITGB3	NM_000212.2
AKT3	NM_181690.1	ITGB4	NM_001005731.1
ALK	NM_004304.3	ITGB6	NM_001282353.1
ALKBH2	NM_001001655.2	ITGB7	NM_000889.1
ALKBH3	NM_139178.3	ITGB8	NM_002214.2
AMER1	NM_152424.3	JAG1	NM_000214.2
AMH	NM_000479.3	JAG2	NM_145159.1
ANGPT1	NM_001146.3	JAK1	NM_002227.1
APC	NM_000038.3	JAK2	NM_004972.2
APH1B	NM_001145646.1	JAK3	NM_000215.2
AR	NM_001011645.1	JUN	NM_002228.3
ARID1A	NM_006015.4	KAT2B	NM_003884.3
ARID1B	NM_020732.3	KDM5C	NM_004187.2
ARID2	NM_152641.2	KDM6A	NM_021140.2
ARNT2	NM_014862.3	KIT	NM_000222.1
ASXL1	NM_001164603.1	KITLG	NM_003994.4
ATM	NM_138292.3	KLF4	NM_004235.4
ATR	NM_001184.2	KMT2C	NM_170606.2
ATRX	NM_000489.3	KMT2D	NM_003482.3
AXIN1	NM_181050.1	KRAS	NM_004985.3
AXIN2	NM_004655.3	LAMA1	NM_005559.2
B2M	NM_004048.2	LAMA3	NM_000227.3
BAD	NM_004322.3	LAMA5	NM_005560.3
BAIAP3	NM_003933.4	LAMB3	NM_000228.2
BAMBI	NM_012342.2	LAMB4	NM_007356.2
BAP1	NM_004656.2	LAMC2	NM_005562.2
BAX	NM_138761.3	LAMC3	NM_006059.3

Table I.1. continued

BCL2	NM_000657.2	LAT	NM_001014987.1
BCL2A1	NM_004049.2	LEF1	NM_016269.3
BCL2L1	NM_138578.1	LEFTY1	NM_020997.2
BCOR	NM_001123383.1	LEFTY2	NM_003240.2
BDNF	NM_170732.4	LEP	NM_000230.2
BID	NM_197966.1	LEPR	NM_001003679.1
BIRC3	NM_182962.1	LFNG	NM_001040168.1
BIRC7	NM_022161.2	LIF	NM_002309.3
BMP2	NM_001200.2	LIFR	NM_002310.3
BMP4	NM_001202.2	LIG4	NM_002312.3
BMP5	NM_021073.2	LRP2	NM_004525.2
BMP6	NM_001718.2	LTBP1	NM_000627.3
BMP7	NM_001719.1	MAD2L2	NM_001127325.1
BMP8A	NM_181809.3	MAML2	NM_032427.1
BMPR1B	NM_001203.1	MAP2K1	NM_002755.2
BNIP3	NM_004052.2	MAP2K2	NM_030662.2
BRAF	NM_004333.3	MAP2K4	NM_003010.2
BRCA1	NM_007305.2	MAP2K6	NM_002758.3
BRCA2	NM_000059.3	MAP3K1	NM_005921.1
BRIP1	NM_032043.1	MAP3K12	NM_006301.2
C19orf40	NM_152266.3	MAP3K13	NM_004721.3
CACNA1C	NM_199460.2	MAP3K14	NM_003954.1
CACNA1D	NM_000720.2	MAP3K5	NM_005923.3
CACNA1E	NM_000721.2	MAP3K8	NM_005204.2
CACNA1G	NM_198397.1	MAPK1	NM_138957.2
CACNA1H	NM_021098.2	MAPK10	NM_002753.2
CACNA2D1	NM_000722.2	MAPK12	NM_002969.3
CACNA2D2	NM_001005505.1	MAPK3	NM_001040056.1
CACNA2D3	NM_018398.2	MAPK8	NM_002750.2
CACNA2D4	NM_001005737.1	MAPK8IP1	NM_005456.2
CACNB2	NM_000724.3	MAPK8IP2	NM_012324.2
CACNB3	NM_000725.2	MAPK9	NM_139068.2
CACNB4	NM_001005747.2	MAPT	NM_016834.3
CACNG1	NM_000727.2	MCM2	NM_004526.2
CACNG4	NM_014405.2	MCM4	NM_182746.1
CACNG6	NM_145814.1	MCM5	NM_006739.3
CALML3	NM_005185.2	MCM7	NM_182776.1

Table I.1. continued

CALML5	NM_017422.4	MDC1	NM_014641.2
CALML6	NM_138705.2	MDM2	NM_006878.2
CAMK2B	NM_001220.3	MECOM	NM_005241.2
CAPN2	NM_001748.4	MED12	NM_005120.2
CARD11	NM_032415.2	MEN1	NM_130802.2
CASP10	NM_032977.3	MET	NM_000245.2
CASP12	NM_001191016.1	MFNG	NM_002405.2
CASP3	NM_032991.2	MGMT	NM_002412.3
CASP7	NM_001227.3	MLF1	NM_022443.3
CASP8	NM_001228.4	MLH1	NM_000249.2
CASP9	NM_001229.2	MLLT3	NM_004529.2
CBL	NM_005188.2	MLLT4	NM_005936.2
CBLC	NM_012116.3	MMP3	NM_002422.3
CCNA1	NM_003914.3	MMP7	NM_002423.3
CCNA2	NM_001237.2	MMP9	NM_004994.2
CCNB1	NM_031966.2	MNAT1	NM_002431.2
CCNB3	NM_033671.1	MPL	NM_005373.2
CCND1	NM_053056.2	MPO	NM_000250.1
CCND2	NM_001759.2	MSH2	NM_000251.1
CCND3	NM_001760.2	MSH6	NM_000179.1
CCNE1	NM_001238.1	MTOR	NM_004958.2
CCNE2	NM_057735.1	MUTYH	NM_012222.2
CCNO	NM_021147.3	MYB	NM_005375.2
CCR7	NM_001838.2	MYC	NM_002467.3
CD14	NM_000591.2	MYCN	NM_005378.4
CD19	NM_001770.4	MYD88	NM_002468.3
CD40	NM_001250.4	NASP	NM_172164.1
CDC14A	NM_033313.2	NBN	NM_001024688.1
CDC14B	NM_003671.3	NCOR1	NM_006311.3
CDC25A	NM_001789.2	NF1	NM_000267.2
CDC25B	NM_021873.2	NF2	NM_181828.2
CDC25C	NM_001790.2	NFATC1	NM_172389.1
CDC6	NM_001254.3	NFE2L2	NM_006164.3
CDC7	NM_003503.2	NFKB1	NM_003998.2
CDH1	NM_004360.2	NFKBIA	NM_020529.1

Table I.1. continued

CDK2	NM_001798.2	NFKBIZ	NM_001005474.1
CDK4	NM_000075.2	NGF	NM_002506.2
CDK6	NM_001259.5	NGFR	NM_002507.1
CDKN1A	NM_000389.2	NKD1	NM_033119.3
CDKN1B	NM_004064.2	NODAL	NM_018055.3
CDKN1C	NM_000076.2	NOG	NM_005450.4
CDKN2A	NM_000077.3	NOS3	NM_000603.4
CDKN2B	NM_004936.3	NOTCH1	NM_017617.3
CDKN2C	NM_001262.2	NOTCH2	NM_024408.3
CDKN2D	NM_001800.3	NOTCH3	NM_000435.2
CEBPA	NM_004364.2	NPM1	NM_002520.5
CEBPE	NM_001805.2	NPM2	NM_182795.1
CHAD	NM_001267.2	NR4A1	NM_173157.1
CHEK1	NM_001114121.1	NR4A3	NM_173198.1
CHEK2	NM_007194.3	NRAS	NM_002524.3
CHUK	NM_001278.3	NSD1	NM_022455.4
CIC	NM_015125.3	NTF3	NM_002527.4
CLCF1	NM_013246.2	NTHL1	NM_002528.5
CNTFR	NM_147164.1	NTRK1	NM_001012331.1
COL11A1	NM_001854.3	NTRK2	NM_001007097.1
COL11A2	NM_001163771.1	NUMBL	NM_004756.3
COL1A1	NM_000088.3	NUPR1	NM_001042483.1
COL1A2	NM_000089.3	OSM	NM_020530.3
COL24A1	NM_152890.5	PAK3	NM_002578.2
COL27A1	NM_032888.2	PAK7	NM_177990.1
COL2A1	NM_001844.4	PAX3	NM_013942.3
COL3A1	NM_000090.3	PAX5	NM_016734.1
COL4A3	NM_000091.3	PAX8	NM_013953.3
COL4A4	NM_000092.4	PBRM1	NM_181042.3
COL4A5	NM_033381.1	PBX1	NM_002585.2
COL4A6	NM_001847.2	PBX3	NM_006195.5
COL5A1	NM_000093.3	PCK1	NM_002591.2
COL5A2	NM_000393.3	PCNA	NM_002592.2
COL6A6	NM_001102608.1	PDGFA	NM_002607.5
COMP	NM_000095.2	PDGFB	NM_033016.2

Table I.1. continued

CREB3L1	NM_052854.1	PDGFC	NM_016205.1
CREB3L3	NM_001271995.1	PDGFD	NM_025208.4
CREB3L4	NM_130898.2	PDGFRA	NM_006206.3
CREB5	NM_182898.2	PDGFRB	NM_002609.3
CREBBP	NM_004380.2	PGF	NM_002632.5
CRLF2	NM_001012288.1	PHF6	NM_032335.3
CSF1R	NM_005211.2	PIK3CA	NM_006218.2
CSF2	NM_000758.2	PIK3CB	NM_006219.1
CSF3	NM_000759.3	PIK3CD	NM_005026.3
CSF3R	NM_156038.2	PIK3CG	NM_002649.2
CTNNB1	NM_001904.3	PIK3R1	NM_181504.2
CUL1	NM_003592.2	PIK3R2	NM_005027.2
CXXC4	NM_025212.1	PIK3R3	NM_003629.3
CYLD	NM_015247.1	PIK3R5	NM_001142633.1
DAXX	NM_001350.3	PIM1	NM_002648.2
DDB2	NM_000107.1	PITX2	NM_000325.5
DDIT3	NM_004083.4	PKMYT1	NM_004203.3
DDIT4	NM_019058.2	PLA1A	NM_015900.2
DKK1	NM_012242.2	PLA2G10	NM_003561.1
DKK2	NM_014421.2	PLA2G2A	NM_000300.2
DKK4	NM_014420.2	PLA2G3	NM_015715.3
DLL1	NM_005618.3	PLA2G4A	NM_024420.2
DLL3	NM_203486.2	PLA2G4C	NM_003706.2
DLL4	NM_019074.2	PLA2G4E	NM_001206670.1
DNMT1	NM_001379.2	PLA2G4F	NM_213600.2
DNMT3A	NM_022552.3	PLA2G5	NM_000929.2
DTX1	NM_004416.2	PLAT	NM_000931.2
DTX3	NM_178502.2	PLAU	NM_002658.2
DTX4	NM_015177.1	PLCB1	NM_182734.1
DUSP10	NM_144728.2	PLCB4	NM_000933.3
DUSP2	NM_004418.3	PLCE1	NM_001165979.1
DUSP4	NM_057158.2	PLCG2	NM_002661.2
DUSP5	NM_004419.3	PLD1	NM_002662.3
DUSP6	NM_001946.2	PML	NM_002675.3
DUSP8	NM_004420.2	POLB	NM_002690.1

Table I.1. continued

E2F1	NM_005225.1	POLD1	NM_002691.2
E2F5	NM_001951.3	POLD4	NM_021173.2
EFNA1	NM_004428.2	POLE2	NM_002692.2
EFNA2	NM_001405.3	POLR2D	NM_004805.3
EFNA3	NM_004952.4	POLR2H	NM_001278698.1
EFNA5	NM_001962.2	POLR2J	NM_006234.4
EGF	NM_001963.3	PPARG	NM_015869.3
EGFR	NM_201282.1	PPARGC1A	NM_013261.3
EIF4EBP1	NM_004095.3	PPP2CB	NM_001009552.1
ENDOG	NM_004435.2	PPP2R1A	NM_014225.3
EP300	NM_001429.2	PPP2R2B	NM_181676.2
EPHA2	NM_004431.2	PPP2R2C	NM_181876.2
EPO	NM_000799.2	PPP3CA	NM_000944.4
EPOR	NM_000121.2	PPP3CB	NM_001142354.1
ERBB2	NM_004448.2	PPP3CC	NM_005605.3
ERCC2	NM_000400.2	PPP3R1	NM_000945.3
ERCC6	NM_000124.2	PPP3R2	NM_147180.2
ETS2	NM_005239.4	PRDM1	NM_182907.1
ETV1	NM_004956.4	PRKAA2	NM_006252.2
ETV4	NM_001079675.1	PRKACA	NM_002730.3
ETV7	NM_016135.2	PRKACB	NM_182948.2
EYA1	NM_172059.2	PRKACG	NM_002732.2
EZH2	NM_004456.3	PRKAR1B	NM_001164759.1
FANCA	NM_000135.2	PRKAR2A	NM_004157.2
FANCB	NM_152633.2	PRKAR2B	NM_002736.2
FANCC	NM_000136.2	PRKCA	NM_002737.2
FANCE	NM_021922.2	PRKCB	NM_212535.1
FANCF	NM_022725.2	PRKCG	NM_002739.3
FANCG	NM_004629.1	PRKDC	NM_006904.6
FANCL	NM_001114636.1	PRKX	NM_005044.1
FAS	NM_152876.1	PRL	NM_000948.3
FASLG	NM_000639.1	PRLR	NM_001204318.1
FBXW7	NM_018315.4	PRMT8	NM_019854.3
FEN1	NM_004111.4	PROM1	NM_006017.1
FGF1	NM_033137.1	PTCH1	NM_000264.3

Table I.1. continued

FGF10	NM_004465.1	PTCRA	NM_138296.2
FGF11	NM_004112.2	PTEN	NM_000314.3
FGF12	NM_004113.4	PTPN11	NM_002834.3
FGF13	NM_033642.1	PTPN5	NM_001039970.1
FGF14	NM_004115.3	PTPRR	NM_001207015.1
FGF16	NM_003868.1	PTTG2	NM_006607.2
FGF17	NM_003867.2	RAC1	NM_198829.1
FGF18	NM_003862.1	RAC2	NM_002872.3
FGF19	NM_005117.2	RAC3	NM_005052.2
FGF2	NM_002006.4	RAD21	NM_006265.2
FGF20	NM_019851.1	RAD50	NM_005732.2
FGF21	NM_019113.2	RAD51	NM_133487.2
FGF22	NM_020637.1	RAD52	NM_134424.2
FGF23	NM_020638.2	RAF1	NM_002880.2
FGF3	NM_005247.2	RASA4	NM_001079877.2
FGF4	NM_002007.2	RASAL1	NM_004658.1
FGF5	NM_004464.3	RASGRF1	NM_153815.2
FGF6	NM_020996.1	RASGRF2	NM_006909.1
FGF7	NM_002009.3	RASGRP1	NM_005739.3
FGF8	NM_033163.3	RASGRP2	NM_001098670.1
FGF9	NM_002010.2	RB1	NM_000321.1
FGFR1	NM_015850.2	RBX1	NM_014248.2
FGFR2	NM_000141.4	RELA	NM_021975.2
FGFR3	NM_022965.2	RELN	NM_005045.2
FGFR4	NM_002011.3	RET	NM_020630.4
FIGF	NM_004469.2	RFC3	NM_002915.3
FLNA	NM_001456.3	RFC4	NM_181573.2
FLNC	NM_001127487.1	RHOA	NM_001664.2
FLT1	NM_002019.4	RIN1	NM_004292.2
FLT3	NM_004119.1	RNF43	NM_017763.4
FN1	NM_212482.1	RPA3	NM_002947.3
FOS	NM_005252.2	RPS27A	NM_002954.5
FOSL1	NM_005438.2	RPS6KA5	NM_004755.2
FOXL2	NM_023067.2	RPS6KA6	NM_014496.1
FOXO4	NM_005938.2	RRAS2	NM_001102669.2

Table I.1. continued

FST	NM_006350.2	RUNX1	NM_001754.4
FUBP1	NM_003902.3	RUNX1T1	NM_004349.2
FUT8	NM_004480.4	RXRG	NM_006917.3
FZD10	NM_007197.2	SETBP1	NM_015559.2
FZD2	NM_001466.2	SETD2	NM_014159.6
FZD3	NM_017412.2	SF3B1	NM_001005526.1
FZD7	NM_003507.1	SFN	NM_006142.3
FZD8	NM_031866.1	SFRP1	NM_003012.3
FZD9	NM_003508.2	SFRP2	NM_003013.2
GADD45A	NM_001924.2	SFRP4	NM_003014.2
GADD45B	NM_015675.2	SGK2	NM_170693.1
GADD45G	NM_006705.3	SHC1	NM_183001.4
GAS1	NM_002048.2	SHC2	NM_012435.2
GATA1	NM_002049.2	SHC3	NM_016848.5
GATA2	NM_032638.3	SHC4	NM_203349.2
GATA3	NM_001002295.1	SIN3A	NM_015477.1
GDF6	NM_001001557.2	SIRT4	NM_012240.1
GHR	NM_000163.2	SIX1	NM_005982.3
GLI1	NM_005269.1	SKP1	NM_170679.2
GLI3	NM_000168.5	SKP2	NM_005983.2
GNA11	NM_002067.1	SMAD2	NM_001003652.1
GNAQ	NM_002072.2	SMAD3	NM_005902.3
GNAS	NM_080425.1	SMAD4	NM_005359.3
GNG12	NM_018841.3	SMAD9	NM_005905.2
GNG4	NM_004485.2	SMARCA4	NM_003072.3
GNG7	NM_052847.1	SMARCB1	NM_003073.3
GNGT1	NM_021955.3	SMC1A	NM_006306.2
GPC4	NM_001448.2	SMC1B	NM_148674.3
GRB2	NM_002086.4	SMC3	NM_005445.3
GRIA3	NM_000828.4	SMO	NM_005631.3
GRIN1	NM_000832.5	SOCS1	NM_003745.1
GRIN2A	NM_000833.3	SOCS2	NM_003877.3
GRIN2B	NM_000834.3	SOCS3	NM_003955.3
GSK3B	NM_002093.2	SOS1	NM_005633.2
GTF2H3	NM_001516.3	SOS2	NM_006939.2

Table I.1. continued

GZMB	NM_004131.3	SOST	NM_025237.2
H2AFX	NM_002105.2	SOX17	NM_022454.3
H3F3A	NM_002107.3	SOX9	NM_000346.2
H3F3C	NM_001013699.2	SP1	NM_003109.1
HDAC1	NM_004964.2	SPOP	NM_001007226.1
HDAC10	NM_032019.5	SPP1	NM_000582.2
HDAC11	NM_024827.3	SPRY1	NM_005841.1
HDAC2	NM_001527.1	SPRY2	NM_005842.2
HDAC4	NM_006037.3	SPRY4	NM_030964.3
HDAC5	NM_005474.4	SRSF2	NM_003016.3
HDAC6	NM_006044.2	SSX1	NM_005635.2
HELLS	NM_018063.3	STAG2	NM_001042749.1
HES1	NM_005524.2	STAT1	NM_007315.2
HES5	NM_001010926.3	STAT3	NM_139276.2
HGF	NM_000601.4	STAT4	NM_003151.2
HHEX	NM_002729.4	STK11	NM_000455.4
HHIP	NM_022475.1	STMN1	NM_203401.1
HIST1H3B	NM_003537.3	SUV39H2	NM_024670.3
HIST1H3G	NM_003534.2	SYK	NM_003177.3
HIST1H3H	NM_003536.2	TBL1XR1	NM_024665.4
HMGA1	NM_145904.1	TCF3	NM_003200.2
HMGA2	NM_003484.1	TCF7L1	NM_031283.1
HNF1A	NM_000545.4	TCL1B	NM_004918.2
HOXA10	NM_018951.3	TET2	NM_001127208.2
HOXA11	NM_005523.5	TFDP1	NM_007111.4
HOXA9	NM_152739.3	TGFB1	NM_000660.3
HPGD	NM_001145816.2	TGFB2	NM_003238.2
HRAS	NM_005343.2	TGFB3	NM_003239.2
HSP90B1	NM_003299.1	TGFBR2	NM_001024847.1
HSPA1A	NM_005345.5	THBS1	NM_003246.2
HSPA2	NM_021979.3	THBS4	NM_003248.3
HSPA6	NM_002155.3	THEM4	NM_053055.4
HSPB1	NM_001540.3	TIAM1	NM_003253.2
IBSP	NM_004967.3	TLR2	NM_003264.3
ID1	NM_002165.2	TLR4	NM_138554.2

Table I.1. continued

ID2	NM_002166.4	TLX1	NM_005521.3
ID4	NM_001546.2	TMPRSS2	NM_005656.2
IDH1	NM_005896.2	TNC	NM_002160.3
IDH2	NM_002168.2	TNF	NM_000594.2
IFNA17	NM_021268.2	TNFAIP3	NM_006290.2
IFNA2	NM_000605.3	TNFRSF10A	NM_003844.2
IFNA7	NM_021057.2	TNFRSF10B	NM_003842.3
IFNG	NM_000619.2	TNFRSF10C	NM_003841.2
IGF1	NM_000618.3	TNFRSF10D	NM_003840.3
IGF1R	NM_000875.2	TNFSF10	NM_003810.2
IGFBP3	NM_000598.4	TNN	NM_022093.1
IKBKB	NM_001556.1	TNR	NM_003285.2
IKBKG	NM_003639.2	TP53	NM_000546.2
IL10	NM_000572.2	TPO	NM_175722.1
IL11	NM_000641.2	TRAF7	NM_032271.2
IL11RA	NM_147162.1	TSC1	NM_000368.3
IL12A	NM_000882.2	TSHR	NM_001018036.2
IL12B	NM_002187.2	TSLP	NM_033035.4
IL12RB2	NM_001559.2	TSPAN7	NM_004615.3
IL13	NM_002188.2	TTK	NM_003318.3
IL13RA2	NM_000640.2	U2AF1	NM_001025203.1
IL15	NM_172174.1	UBB	NM_018955.2
IL19	NM_013371.3	UBE2T	NM_014176.3
IL1A	NM_000575.3	UTY	NM_007125.3
IL1B	NM_000576.2	VEGFA	NM_001025366.1
IL1R1	NM_000877.2	VEGFC	NM_005429.2
IL1R2	NM_173343.1	VHL	NM_000551.2
IL1RAP	NM_002182.2	WEE1	NM_003390.3
IL20RA	NM_014432.2	WHSC1	NM_007331.1
IL20RB	NM_144717.2	WHSC1L1	NM_017778.2
IL22RA1	NM_021258.2	WIF1	NM_007191.2
IL22RA2	NM_181309.1	WNT10A	NM_025216.2
IL23A	NM_016584.2	WNT10B	NM_003394.2
IL23R	NM_144701.2	WNT11	NM_004626.2
IL24	NM_181339.1	WNT16	NM_057168.1

Table I.1. continued

IL2RA	NM_000417.1	WNT2	NM_003391.2
IL2RB	NM_000878.2	WNT2B	NM_024494.1
IL3	NM_000588.3	WNT3	NM_030753.3
IL3RA	NM_002183.2	WNT4	NM_030761.3
IL5RA	NM_000564.3	WNT5A	NM_003392.3
IL6	NM_000600.1	WNT5B	NM_032642.2
IL6R	NM_000565.2	WNT6	NM_006522.3
IL7	NM_000880.2	WNT7A	NM_004625.3
IL7R	NM_002185.2	WNT7B	NM_058238.1
IL8	NM_000584.2	WT1	NM_000378.3
INHBA	NM_002192.2	XPA	NM_000380.3
INHBB	NM_002193.2	XRCC4	NM_003401.3
IRAK2	NM_001570.3	ZAK	NM_016653.2
IRAK3	NM_007199.1	ZBTB16	NM_006006.4
IRS1	NM_005544.2	ZBTB32	NM_014383.1
ITGA2	NM_002203.2	ZIC2	NM_007129.2

Table I.1. continued

Internal Reference Genes			
Official Symbol	**Accession**	**Official Symbol**	**Accession**
ACAD9	NM_014049.4	NOL7	NM_016167.3
AGK	NM_018238.3	NUBP1	NM_001278506.1
AMMECR1L	NM_001199140.1	PIAS1	NM_016166.1
C10orf76	NM_024541.2	PIK3R4	NM_014602.1
CC2D1B	NM_032449.2	PRPF38A	NM_032864.3
CNOT10	NM_001256741.1	RBM45	NM_152945.2
CNOT4	NM_001190848.1	SAP130	NM_024545.3
COG7	NM_153603.3	SF3A3	NM_006802.2
DDX50	NM_024045.1	SLC4A1AP	NM_018158.2
DHX16	NM_001164239.1	TLK2	NM_006852.2
DNAJC14	NM_032364.5	TMUB2	NM_024107.2
EDC3	NM_001142443.1	TRIM39	NM_021253.3
EIF2B4	NM_172195.3	TTC31	NR_027749.1
ERCC3	NM_000122.1	USP39	NM_001256725.1
FCF1	NM_015962.4	VPS33B	NM_018668.3
FTSJ2	NM_013393.1	ZC3H14	NM_001160103.1
GPATCH3	NM_022078.2	ZKSCAN5	NM_014569.3
HDAC3	NM_003883.2	ZNF143	NM_003442.5
MRPS5	NM_031902.3	ZNF346	NM_012279.2
MTMR14	NM_022485.3	ZNF384	NM_133476.3