qPCR VALIDATION OF IN VIVO DIAGNOSTIC IMPORTANCE AND REGULATION BY ESTROGEN FOR CHRNA5 ISOFORM EXPRESSION IN BREAST CANCER

A THESIS SUBMITTED TO
THE DEPARTMENT OF MOLECULAR BIOLOGY AND GENETICS
AND GRADUATE SCHOOL OF ENGINEERING AND SCIENCE
OF BILKENT UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE

By Emine Sıla ÖZDEMİR

September 2014

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Assist. Prof. Özlen Konu (Advisor)

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Assist. Prof. Öznur Taştan Okan

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Assist. Prof. Özgür Şahin

Approved for the Graduate School of Engineering and Science

Prof. Dr. Levent Onural

Director of the Graduate School of

Engineering and Science

To my mother Sezer, and to my father Ekrem...

Thank you for being with me all the time ..

ABSTRACT

qPCR VALIDATION OF IN VIVO DIAGNOSTIC IMPORTANCE AND REGULATION BY ESTROGEN FOR CHRNA5 ISOFORM EXPRESSION IN BREAST CANCER

Emine Sıla Özdemir

MSc. in Molecular Biology and Genetics

Supervisor: Asst. Prof. Dr. Özlen Konu

September 2014

Breast cancer has multiple molecular subtypes; normal-like, basal-like, luminal A, luminal B and HER2 positive depending on receptor status of tumor cells. Cancer therapy is tailored according to the type of cancer; hence finding new diagnostic markers is important to decide on the best treatment approach. Cholinergic nicotinic receptor alpha 5 (CHRNA5) is one of the subunits of nicotinic acetylcholine receptors with significant roles in addiction and cancer. In the present study, CHRNA5 has been validated as an estrogen and/or Estrogen receptor (ER) modulated nicotinic acetylcholine receptor by qPCR in *in vitro* and *in vivo* in breast cancer samples. CHRNA5 isoform expression was measured using *in vitro* cell culture studies in which ER- and ER+ cell lines treated with different doses of estradiol (E2); MCF7 cell line was exposed to long-term E2 depletion, in another experiment it was treated with tamoxifen (4-OHT), an ER antagonist, and with or without E2. We found that all CHRNA5 isoforms exhibited increased expression in response to E2 dose-dependently in the ER+ MCF7 cell line while in the ER- MDA-MB-231 cell line CHRNA5 isoform expression response was variable in direction and magnitude. CHRNA5 isoform expression in general steadily decreased in ER+ cell line MCF7 after 4-OHT treatment. After six months of E2 depletion, ER+ MCF7 cell line had increased CHRNA5_v3 isoform and ESR1 (ER gene) mRNA

expression. *In vivo*, a human breast cancer cDNA panel was scanned with specially designed primers with qPCR using a custom-written GUI in MATLAB. It was found that CHRNA5, showing a statistically significant difference between normal and tumor cDNA, was a good candidate gene in diagnosis of breast cancer. CHRNA_v3 was able to distinguish between ER+ vs ER- breast tumor samples. We also addressed whether CHRNA5 isoforms exhibited differences in distinguishing tumor stage, and HER2 status. Our findings showed that expression of CHRNA5 isoforms were correlated with each other and regulated by E2 in breast cancer depending on ER receptor status.

Keywords: Breast cancer, CHRNA5, estrogen, estrogen depletion, tamoxifen, molecular subtype.

ÖZET

MEME KANSERİNDE CHRNA5 İSOFORM İFADESİNİN IN VIVO TANISAL ÖNEMİ VE ÖSTROJENLE DÜZENLENMESİNİN qPCR İLE DOĞRULANMASI

Emine Sıla Özdemir

Moleküler Biyoloji ve Genetik Yüksek Lisansı

Tez Danışmanı: Yrd. Doç. Dr. Özlen Konu

Eylül 2014

Meme kanserinin birden çok alt tipi bulunmaktadır; bunlar tümör hücresinin reseptör durumuna bağlı olan normal benzeri, bazal benzeri, luminal A, luminal B ve HER2 pozitiftir. Kanser tedavisi, kanser tipine göre belirlenir; bu da teşhis amaçlı yeni bir markör bulmanın en iyi tedavi yöntemini bulmaktaki önemini göstermektedir. Kolinerjik reseptör nikotinik alfa 5 (CHRNA5) nikotinik asetilkolin reseptörlerinin bir alt birimi olup bağımlılık ve kanserde önemli rol oynamaktadır. Bu çalışmada, CHRNA5'in östrojenle ve/veya östrojen reseptörü ile modüle edilen bir reseptör olduğu meme kanseri örneklerinde *in vivo* ve in *vitro* qPCR doğrulama çalışmalarıyla gösterilmiştir. CHRNA5 izoform ifadesi farklı dozda estradiol E2 uygulamasına maruz bırakılan ER- ve ER+ hücre hatlarında hücre kültürü çalışmalarıyla ölçülmüş; MCF7 hücre hattı uzun süreli E2 açlığına maruz bırakılmış ve diğer bir deneyde ise E2 varlığında ve yokluğunda bir ER antagonisti olan tamoxifen (4-OHT) uygulanmasına maruz bırakılmıştır. E2 tedavisinden sonra ER+ MCF7 hücre hattında doza bağlı olarak CHRNA5 izoform ifadesinin arttığı, ER- MDA-MB 231 hücre hattında ise CHRNA5 izoform ifadesinin yön ve miktar açısından değişkenlik gösterdiği bulgulanmıştır. CHRNA5 izoform ifadesi genel olarak ER+ MCF7 hücre hattında 4-OHT tedavisi sonucunda azalmıştır. Altı aylık E2 yokluğundan sonra ER+ hücre hatlarında CHRNA5_v3 ve ESR1 (ER geni) ifadesinin arttığı gözlenmiştir. *In*

vivo, bir insan meme kanseri cDNA paneli, özel olarak tasarlanmış primerlerle qPCR yöntemi ve MATLAB'da özel yazılmış GUI uygulaması kullanılarak taranmıştır. Normal ve tümör cDNA'leri arasında istatistiksel olarak önemli bir ifade farklılığı gösteren CHRNA5 geninin, meme kanseri teşhisinde önemli bir biyobelirteç olduğu bulgulanmıştır. Ayrıca, CHRNA5_v3 izoformu ER+ ve ER- meme tümörü örneklerini ayırt edebilmektedir. CHRNA5 izoformlarının miktarının stage ve HER2 statusüne göre nasıl değiştiği de test edilmiştir. Bulgularımız CHRNA5 izoformlarının birbirleriyle korale olduğunu ve meme kanserinde reseptör statüsüne bağlı olarak E2 tarafından regüle edildiğini göstermiştir.

Anahtar sözcükler: Meme kanseri, CHRNA5, östrojen, östrojen açlığı, tamoxifen, moleküler tip.

ACKNOWLEDGEMENTS

First of all, I would like to thank my thesis advisor Asst. Prof. Özlen Konu, I really appreciate her guidance and devotion on my thesis project and graduate studies.

I would also like to thank Prof. Ihsan Gürsel for helping me whenever I needed and primarily for changing my perspective through science. He caused me to gain new scientific interests and widen my vision. Assoc. Prof. Işık Yuluğ, Assoc .Prof. Uygar Tazebay and Asst. Prof. Ali Güre guided me in both academic and personal life, I really appreciate their help.

I would especially like to mention my dear friend Yusuf İsmail Ertuna, Azer Aylin Açıkgöz, Derya Soner and Seda Koyuncu who have always been by my side during these two years both in social and academic life.

I also want to mention and thank my other friends Kemal Çağlar Göğebakan, Melek Cansu Petek, Mehmet Kayşim Kişioğlu and Mustafa Erol who always support me.

I would like to thank Ermira Jahja for cloning, sequencing and providing primer sequences and Gülçin Çakan, Ender Avcı, Huma Shehwana and Şahika Çıngır for assistance in experiments.

Once again I would like to thank and express my gratitude and love toward my parents Ekrem and Sezer, also my aunt Ümran who is like a mother during my undergraduate and graduate studies. You are the main reason behind my success and improvement. Thank you for your support and everything.

TUBITAK was the foundation financially supporting this project (TBAG-111T316 to OK) and me (2210-E) during my master studies.

TABLE OF CONTENTS

SIGNATURE PAGE.....II

DEDICATION.....III

ABSTRACT.....IV

ÖZET.....VI

ACKNOWLEDGEMENTS.....VIII

TABLE OF CONTENTS.....IX

LIST OF FIGURES.....XIV

LIST OF TABLES.....XVI

1.INTRODUCTION..... 1

1.1 Breast Cancer 1

1.1.1 Classification and Subtypes in Breast Cancer 3

1.1.1.1 Receptor Status Defining Diagnosis and Survival 4

1.1.2 Risk Factors 5

1.2 Estrogen (E2) Signaling in Breast Cancer..... 6

1.2.1 Genomic and Nongenomic Signaling..... 7

1.2.2 4-OHT and Breast Cancer..... 8

1.2.3 Effects of Long-term E2 Depletion on Estrogen Receptors 9

1.3 Cholinergic Receptor Nicotinic Alpha (CHRNA or nAChR)..... 10

1.3.1 CHRNA5 and Cancer 11

1.3.2 Regulation of CHRNA5 Expression..... 12

1.3.3 Effects of E2 on Cholinergic Signaling 13

1.4 Alternative Splicing and Isoform Formation 13

1.4.1 Importance of isoform expression patterns of genes in cancer 14

1.4.2 Alternative Splicing and Breast Cancer 15

1.5 Isoform Expression Measurement Technologies 16

1.5.1 Microarrays and Analysis Methods 16

1.5.2 qPCR Studies and Analysis Methods 16

1.5.3 Northern Blots 17

1.6 Diagnosis and Classifying Cancer with *in vivo* cDNA Expression Patterns 17

1.6.1 cDNA Panel Studies of Different Types of Cancer 17

1.6.2 Quantitative and Statistical Methods to Analyze cDNA Panels 18

1.7 Tools for Expression Analysis 19

1.7.1 Creating Tools for cDNA Panel Analysis 19

1.7.2 Outlier Analysis 20

2. AIM 21

3. MATERIALS AND METHODS 22

3.1 Cell Culture Protocols 22

3.1.1 Cell Lines 22

3.1.2 Cell Growth Conditions 22

3.1.3 Thawing Cells 23

3.1.4 Subculturing of Cells 23

3.1.5 Cryopreservation of Cells 23

3.1.6 Counting Cells with Hemocytometer 24

3.1.7 β-Estradiol (E2) Treatment 24
3.1.8 Tamoxifen and E2 Treatment 25
3.1.9 Long-term E2 Depletion of Cells 26
3.2 Gene Expression Analyses 27
3.2.1 Total RNA Isolation 27
3.2.2 cDNA Synthesis 27
3.2.3 Oligonucleotides 28
3.2.4 Gradient Polymerase Chain Reaction 29
3.2.5 Agarose Gel Electrophoresis 30
3.2.6 Quantitative Real Time-Polymerase Chain Reaction (qPCR) 30
3.2.7 Detection of qPCR Outliers Based on Tm Values from Melt Curves 32
3.2.8 Primer Efficiency 32
3.2.9 ΔCt and ΔΔCt Analysis of qPCR Data 33
3.2.10 ANOVA analysis 34
3.3 Analysis of E2 Treatment 34
3.3.1 Clustering 34
3.3.2 Treeview analysis 34
3.4 Analysis of 4-OHT and E2 Treatment 35
3.5 Analysis of Long-term Estrogen Depletion Experiment 35
3.5 Analysis of cDNA Panel 35
3.5.1 Statistical Analysis of qPCR Data 36
3.5.1.1 Correlation Analysis 36

3.5.1.2 MATLAB GUI Application 36

3.6 General Solutions 36

4.RESULTS 38

4.1 Effects of Different Doses of E2 on CHRNA5 Isoform Expression 38

4.1.1 Preliminary PCR for MCF7 and MDA-MB-231 Treated with Different Doses of Estradiol (E2) 38

4.1.2 Statistical Analysis of E2 Treatment with MATLAB 38

4.1.3 Heat Map of CHRNA5 Isoform and E2 Markers Expression 46

4.2 Effects of 4-OHT and E2 Treatment 48

4.2.1 Statistical Analysis of 4-OHT and E2 Treatment with MATLAB 48

4.2.2 Comparison of CHRNA5 Isoform and E2 and 4-OHT Markers Expression 51

4.3 Effects of Long-term E2 Depletion (LTED) on CHRNA5 Isoform Expression 53

4.3.1 Determining the Housekeeping Gene for LTED Experiment 53

4.3.2 Expression of CHRNA5 Isoform and E2 Exposure Markers across LTED 54

4.4 Expression of CHRNA5 Isoforms in cDNA Panel 58

4.4.1 Data Quality Control Studies 58

4.4.1.1 Tm Consistency Check in cDNA Panels 59

4.4.2 Statistical Analysis of CHRNA5 Isoforms Expression 60

4.4.3 Comparison of CHRNA5 Isoform Expression 62

4.4.4 GUI Application for cDNA Panel Data Analysis 68

5. CONCLUSION AND DISCUSSION 73

6. FUTURE PERSPECTIVES 80

REFERENCES 82

APPENDICES……..…….89

LIST OF FIGURES

Figure 1.1: A model for breast carcinogenesis..... 2

Figure 1.2: Representation of genomic and nongenomic ER signaling pathway 8

Figure 1.3: GEO2R analysis of GSE20361 dataset consisting of initial and three months long estrogen depleted ER+ cell line for ESR1 (ER) gene 9

Figure 1.4: Schematic representation of nAChRs hetero (upper left) and homo-oligomer (botton left) complexes. 10

Figure 1.5: Representation of alternative splicing of CHRNA5 in lung. 12

Figure 1.6: Patterns of alternative splicing. 14

Figure 3.1: Scheme of a hemocytometer..... 24

Figure 3.2: Scheme of binding regions of CHRNA5 isoform primers 29

Figure 4.1: Representative expression of CHRNA_iso2 products across ER-and ER+ breast cancer cell lines treated with 1nm, 10nm and 100nm E2. 38

Figure 4.2: Effects of E2 doses on CHRNA5 isoforms and PS2 in MCF7 cell line (ΔΔCt)……41

Figure 4.3: Effects of E2 doses on CHRNA5 isoforms and PS2 in MDA MB 231 cell line (ΔΔCT).. 44

Figure 4.4: Heatmap representations of expression level of CHRNA5 isoforms and PS2 in MCF7 and MDA MB 231 cell lines. 47

Figure 4.5: Effects of E2 and 4-OHT on CHRNA5 isoforms, E2 and 4-OHT indicators in MCF7cell line……49

Figure 4.6: Overall representation of expression level of CHRNA5 isoforms and E2, 4-OHT indicators…….. 52

Figure 4.7: Changes in gene expression of treatment with respect to control of TPT1 expression per time point during six months of E2 depletion with ΔΔCt analysis. . . 54

Figure 4.8: Changes in gene expression of treatment with respect to control of CHRNA5 isoforms expression per time point during six months E2 depletion with ΔΔCt analysis……..55

Figure 4.9: Changes in gene expression of treatment with respect to control of E2 indicators expression per time point during six months of E2 depletion with ΔΔCt analysis ……..……..56

Figure 4.10: Gel image of qPCR products from different experiment with varying Tm values. 59

Figure 4.11: Representative melting curve. Tm curves having different Tm than the median differ from each other…….....60

Figure 4.12: Representation of Ct value distribution before and after Tm outlier removal..... 60

Figure 4.13: Correlation plots of three CHRNA5 isoforms 67

Figure 4.14: Interfaces of input module of cDNA panel analysis application..... 69

Figure 4.15: Interface to enter input file names 69

Figure 4.16: Interfaces of quality control module of cDNA panel analysis application 70

Figure 4.17: Interfaces of method option module of cDNA panel analysis application 71

Figure 4.18: Interfaces of analysis module of cDNA panel analysis application 72

Figure 8.1: GEO2R analysis of GSE26459 dataset for FOSL2 expression with E2 and 4-OHT in 4-OHT sensitive cell lines. 93

Figure 8.2: GEO2R analysis of GSE26459 dataset for CHRNA5 expression with E2 and 4-OHT in 4-OHT sensitive cell lines 93

Figure 8.3: GEO2R analysis of GSE26459 dataset for PS2 expression with E2 and 4-OHT in 4-OHT sensitive cell lines..... 94

Figure 8.4: GEO2R analysis of GSE20361 dataset for PS2 expression after and before LTED. 94

Figure 8.5: GEO2R analysis of GSE2251 dataset for PS2 expression with and without E2 treatment in ER- MDA MB 231 cell line. 95

Figure 8.6: GEO2R analysis of GSE26459 dataset for FOSL2 expression with E2 and 4-OHT in 4-OHT resistant cell lines. 95

LIST OF TABLES

Table 1.1: Classification of molecular subtypes 4

Table 1.2: Breast cancer risk factor and relativity of the risk 6

Table 3.1: Breast cancer cell lines origin 22

Table 3.2: The receptor status and molecular subtypes of cell lines 22

Table 3.3: Primer sequences, their product sizes and efficiencies 28

Table 3.4: PCR reagents and their volumes 29

Table 3.5: PCR conditions 30

Table 3.6: qPCR reagents and their volumes 31

Table 3.7: qPCR conditions 31

Table 4.1: p-values of each gene comparing MCF7 and MDA MB 231 cell lines and different doses of E2 based on two-way and one-way ANOVA 39

Table 4.2: p-values of each gene comparing E2 doses in MCF7 cell line from ANOVA test 42

Table 4.3: List of group having statistically significant expression difference for each gene 423

Table 4.4: p-values of each gene comparing E2 doses in MDA MB 231 cell line from ANOVA test 45

Table 4.5: List of group having statistically significant expression difference for each gene 46

Table 4.6: p-values of each gene comparing control, E2, 4-OHT, and E2 and 4-OHT together in MCF7 cell line from ANOVA test 50

Table 4.7: List of group having statistically significant expression difference for each gene 51

Table 4.8: p-values of each gene comparing expression after six months of estrogen depletion from sign test 57

Table 4.9: p-values of each CHRNA5 isoform normal-tumor, ER+ - ER-, HER2+ - HER2- and stages of tumor in cDNA panel 61

Table 4.10: List of group having statistically significant expression difference for each gene 61

Table 4.11: Correlation coefficiencies and p-values between primer pairs for all samples 68

Table 8.1: Covdesc file of cDNA panel 90

Table 8.2: Time points for LTED experiment 96

1.INTRODUCTION

1.1 Breast Cancer

Breast cancer is a cancer type initiated from mammary gland lobules that function in providing milk, and ducts that connect lobules to nipples [1]. It is the worldwide second most common cancer type, and the reason of ~15% of cancer-related deaths in women [2].

Breast cancer can be hereditary or sporadic. In hereditary breast cancer, the germline mutations that patients inherited increase susceptibility of developing breast cancer. BRCA1 and BRCA2 are the main genes involved in breast cancer, however, PTEN and TP53 mutations also are effective in breast cancer development [3]. On the other hand, sporadic breast cancer occurs as a result of the accumulation of mutations in somatic tissue. More than 80% of women who face with breast cancer do not have a previous complaint on the disease occurrence in the family; risk factors include age and diet [3].

Tumorigenesis, the transformation from a normal cell to a malignant cell, basically occurs in three steps [4]: First, normal cell changes into a lesion and then lesion becomes malignant. The last step occurs when the malignant lesion undergoes metastasis.

The genes and steps involving mammary gland tumorigenesis are explained in Figure 1.1. A breast stem cell divides asymmetrically, and then there could be mutation in a gene involving breast cancer susceptibility. It is possible that the two events could be simultaneous. These genes might involve cell cycle regulation and the mutation of in one or more of them encourages proliferation [5].

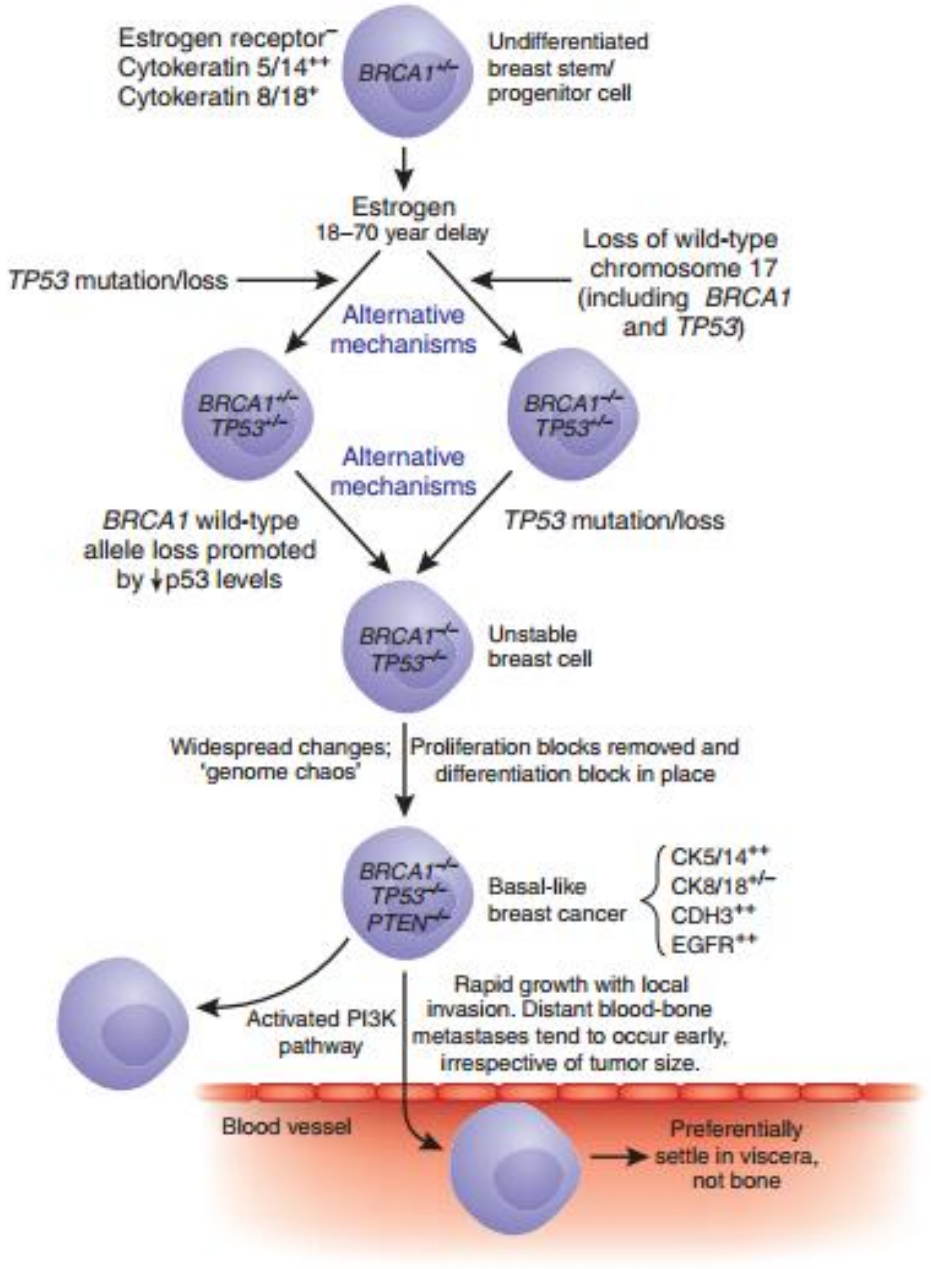


Figure 1.1: A model for breast carcinogenesis.
(Reprinted from Foulkes, 2008, with permission from Nature Publishing Group)

1.1.1 Classification and Subtypes in Breast Cancer

Breast cancer is a heterogeneous disease hence it requires to be categorized effectively for more accurate diagnosis of subtypes and identification of personalized therapeutic strategies. Breast cancer classification can be done considering different aspects of tumor biology, e.g., histological, molecular. Histological subtypes of breast cancer indicate the location of tumor initiation site and state as well as the status of invasiveness of cancer. The histological subtypes can be distinguished via *in situ* studies as invasive ductal, ductal and invasive lobular carcinomas [6].

Grade of a breast cancer tumor defines how differentiated the tumor tissue is from normal breast tissue. If the tumor is well differentiated it is called as low-grade and they are more similar to normal tissue than high-grade tumors that are not well differentiated [7].

Molecular classification of breast cancer on the other hand considers receptor status of tumor tissue as the primary factor. There are three types of receptors on breast cancer tumor cells which could be associated with the character of tumor. Estrogen Receptor 1 (ER), Human Epidermal growth factor Receptor 2 (HER2) and Progesteron Receptor (PR) are these three receptors that may take part in the proliferation and maintenance of tumor cells.

Recent studies performed with microarrays also have led to identification of expression profiles that categorize molecular subtypes; these have been basically classified as normal-like, luminal A, luminal B, HER2+ and basal-like type [8]. Basal-like contains triple negative cancers that do not include any of the indicated three receptors. Luminal types are ER+, however they differ by the level of gene expression of HER2, luminal A subtypes do not express HER2, while luminal B subtypes are HER2+ [9]. HER2+ type breast cancer as the name implies are positive for HER2 [8]; ERBB2/HER2 positive tumors grow and spread more aggressively similar to basal-like subtypes [10]. In Table 1.1 molecular subtypes are summarized.

Table 1.1: Classification of molecular subtypes

Subtype	ER	PR	HER2
Basal	-	-	-
Luminal A	+	+ or -	-
Luminal B	+	+ or -	+
HER2+	-	-	+

1.1.1.1 Receptor Status Defining Diagnosis and Survival

The estrogen receptor (ER) has an important role in breast cancer development and progression. Current work on gene expression profiles seem to be altered according to hormone receptor status of the breast cancer [11]. Microarray studies show existence of gene sets correlating with receptor status thus allowing for interpreting subtypes upon examining the expression of these genes. ER status, in addition, has an impact on the DNA methylation state of some genes such as FAM124B, NAV1, and PER1 in breast cancer [12].

Genetic and epigenetic alterations in ER + tumors increase the sensitivity to endocrine therapy, yet ER- tumors are known as hormone independent [12]. ER- tumors do not have a functional protein to which the hormone estrogen will bind. There is no need for estrogen for the growth of ER- breast cancer, and usually they do not halt growing upon treatment with hormones that prevent estrogen from binding. Therefore, ER- breast tumors are considered as more aggressive [13]. Studies show that ER- breast cancers do not have ER mRNA. The lack of ER is not a result of deletions or other structural changes in the ER gene. CpG island methylation is thought to be one of the alternatives in blockage of the transcription of the ER gene in ER – breast cancers [14].

Progesterone receptor (PR) is the mediating factor for progesterone's effects in the development of the mammary gland and the progression of the breast cancer. PR is

expressed in more than half of the ER + breast cancer, and estrogen signaling is necessary to induce PR expression [15]. It was found that some oncogenic pathways including PI3 K/Akt/mTOR, were more active in both ER-/PR- and ER+/PR- compared to ER+/PR+ tumors. PR- tumors can be considered more aggressive and invasive [15].

The ERBB2/HER2 gene is overexpressed in 20% - 30% of invasive breast carcinomas; a positive HER2 status is associated with improvement of metastatic potential and decreased overall survival. In addition, patients with HER2+ tumors are less responsive to tamoxifen treatment than patients with HER2- tumors [16].

1.1.2 Risk Factors

Risk factors causing breast cancer can be hereditary as well as environmental. Moreover, main risk factors such as gender and age are not modifiable. Family history is also an important risk factor, with increasing susceptibility to breast cancer in women with one or more relatives with breast cancer.

Conditions triggering other risk factors including postmenopausal obesity, smoking and alcohol consumption, use of estrogen and progesterone menopausal hormones can be intervened. For example, use of hormones as therapy increases the cell proliferation and promotes susceptibility to DNA damage and induces breast cancer cell growth [17]. Recent epidemiological studies have focused on how much a factor hormonal therapy is in breast cancer risk and they concluded that recent and long-term use of menopausal hormone therapy has a relative risk value between 1.1-2.0 [2]. Avoding such modifiable risk factors may help reduce the risk of breast cancer.

Table 1.2 summarizes general breast cancer risk factors with the relative likelihood of the risk [2].

Table 1.2: Breast cancer risk factor and relativity of the risk [2]

Relative Risk	Factor
>4.0	Age, Certain inherited genetic mutations (BRCA1-2), Personal history of early onset
2.1-4.0	Exposure to high-dose radiation, One first degree relative with breast cancer
1.1-2.0	Alcohol consumption, Height, Obesity, Long-term and recent use of hormone therapy

1.2 Estrogen (E2) Signaling in Breast Cancer

Estrogens are steroid hormones found primarily in females. Serum estrogen level in premenopausal women is high, while its level decreases in postmenopausal women. Estrogen can be provided both endogenously or exogenously. Estrogen is primarily produced by ovaries during menstrual life endogenously. Exogenous sources include hormone-replacement therapies and oral contraceptive usage. The small amount of estrogen present after menopause comes from extragonadal tissues like muscles, fat and liver [18].

It is known that estrogens are responsible for the initiation and growth of breast cancer. The exact mechanism is still complex but estrogens lead to the development of mammary cancer with proliferative effects [19]. Metabolism of estrogen and its metabolites are affected by genetic and environmental factors and breast cancer initiation and growth can relate to estrogen and its metabolites.

There are other ligands besides E2 which can trigger ER pathway; however they may have antagonistic effect on cell proliferation. Phytoestrogens and genistein are plant derived compounds that are structurally similar to E2. Although phytoestrogens induce estrogen responsive genes, their anti-proliferative and apoptotic effects are mediated through the estrogen receptor [20,21].

1.2.1 Genomic and Nongenomic Signaling

Estrogen stimulation triggers two pathways in general: genomic and nongenomic signaling. Estrogen molecule can pass through membrane because of its steroid structure. In ER+ cells, estrogen and ER located in nucleus interact and form a ligand-receptor complex. After binding of estrogen, ER dimerizes using two types of estrogen receptors, ER-α and ER-β, with ability to form homo- or heterodimers after binding of ligand (estrogen). This complex then binds to specific response elements known as estrogen response elements (EREs) located in the promoters of target genes [22,23]. However, there are cases in which ER does not directly bind to DNA to regulate gene expression. This occurs through secondary messenger triggering a set of transcription factors in nucleus [22].

E2 modulated classical genomic signaling pathway includes direct binding of ER to DNA via ERE. Transferred E2-ER complex from the plasma to the nucleus recruits transcriptional machinery and other cofactors to specific DNA target sequences of estrogen-responsive gene promoters [24]. However, it is known that approximately one third of the genes in humans that are regulated by ERs do not contain ERE-like sequences [25]. Molecular mechanism behind this phenomenon still is not fully understood.

Nongenomic ER signaling pathway alters gene expression with the help of secondary messengers, E2-ER complex does not directly enter nucleus [24]. Cytoplasmic ER and E2 interaction triggers mobilization of intracellular calcium and cAMP production. E2 can activate MAPK signaling pathway with the nongenomic actions in several cell types including breast cancer [26]. E2 also triggers the phosphoinositol (PI) 3-kinase signaling pathway in breast cancer cell [27]. Additionally, E2 activated membrane associated ERs come in close contact with metabotropic glutamate

receptors (mGluRs), turning on mGluR signaling without glutamate. Protein kinase cascades including PKA,PKC, and tyrosine kinases pathways can mediate the non-genomic ER pathway [24].

Figure 1.2 shows a representation of genomic and nongenomic ER signaling pathways.

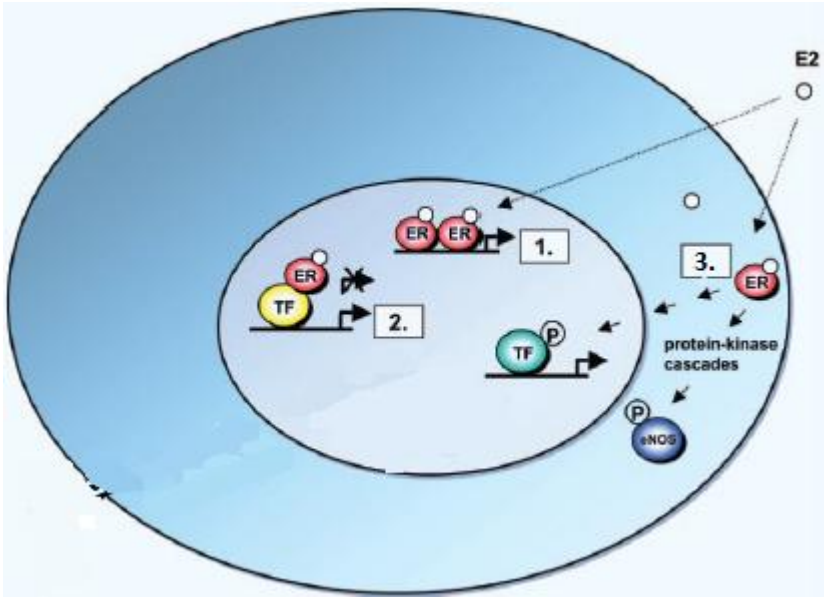


Figure 1.2: Representation of genomic and nongenomic ER signaling pathway
1. Genomic signaling pathway, 2. ERE-independent genomic signaling pathway, 3. Nongenomic signaling pathway ***(Reproduced with permission from Björnström et al., 2005 Copyright Massachusetts Medical Society)***

1.2.2 4-OHT and Breast Cancer

Tamoxifen (4-OHT) is an agent commonly used as an anti-estrogen treatment for hormone-dependent breast cancer patients. As mentioned earlier E2 induces cell proliferation in ER+ breast cancer cells and 4-OHT, an ER antagonist, blocks the binding of E2 to ER [28,29]. 4-OHT binds to the ER and affect the position of helix 12, preventing ER transcriptional activation [30].

It is also thought that E2 can modulate a pathway that triggers the invasiveness of tumor cells, so inhibition of E2 pathway is crucial for the therapy of both early and advanced breast cancer [31].

Although 4-OHT/tamoxifen therapy is very efficient and beneficial for patients, eventually tamoxifen resistance develops in tumor cells. It is also known that tamoxifen resistance can promote EMT-like behavior [32,33].

1.2.3 Effects of Long-term E2 Depletion on Estrogen Receptors

Endocrine therapies including long term estrogen depletion (LTED) are the most effective systemic treatment of ER+ breast cancer. However, tumor cells can develop resistance to endocrine therapies. Indeed, acquired resistance is not related to the conversion of ER+ cells to ER- cells; hence changes in ER expression are thought to be main reason of resistance [34]. Previous research supports that resistance is acquired by molecular events trigger constitutive activation of ER and growth factor signaling pathway [35]. Long term estrogen depletion may cause an increase in growth factor receptor levels, such that ER could also be significantly overexpressed [36].

As a pilot study for hypothesis generation, we have analyzed a part of a microarray study that focused on effects of long-term estrogen depletion on transcriptome by using GEO2R [36]. Figure 1.3 shows ESR1 expression in 3, 15 days and three months of long-term estrogen depletion (LTED) time course in an ER+ cell line.

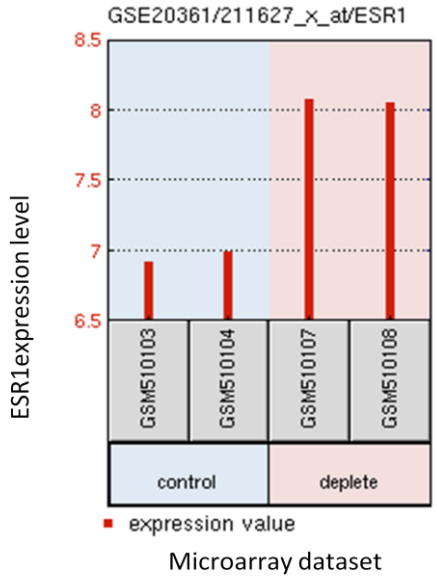


Figure 1.3: GEO2R analysis of GSE20361 dataset consisting of initial and three months long estrogen depleted ER+ cell line for ESR1 (ER) gene.

It is apparent that LTED leads to an upregulation in expression of ER mRNA levels; this increase could then be translated into protein and has potential to modify and compensate the lack of E2. Overexpression of ER could result in improved cell survival ability and the mediation of estrogen-independent growth, and end up with resistance to endocrine therapies in breast cancer patients [37].

1.3 Cholinergic Receptor Nicotinic Alpha (CHRNA or nAChR)

Nicotinic acetylcholine receptors or cholinergic nicotinic receptors are transmembrane proteins functioning as ligand-gated ion channels which can bind to acetylcholine (ACh) and nicotine. These ligands mainly found at the neuromuscular junction and in some neurons, also are present in non-neuronal cells [38].

nAChRs have 2 different types of subunits; 9α (2-10), and 3β (2-4) subunits. Subunits form pentamers with a barrel like structure in the membrane around a central ion channel [38] (Figure 1.4).

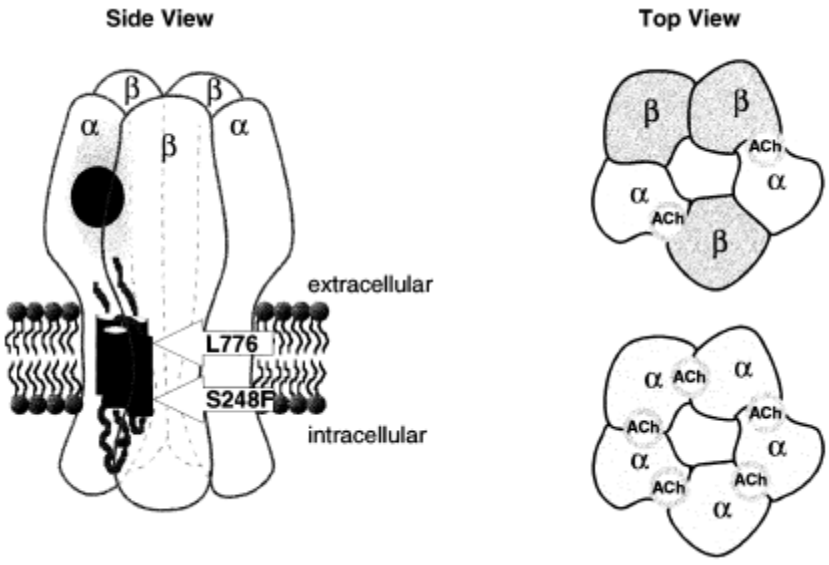


Figure 1.4: Schematic representation of nAChRs hetero (upper left) and homo-oligomer (botton left) complexes. ***(Reprinted from Weiland, 2000, with permission from Elsevier)***

Neuronal nAChRs have alternatively spliced forms which increase the diversity and number of receptor subtypes [39]. Different subtypes may explain the variable

tolerance and addiction to nicotine or different release of neurotransmitters in different synapses [40].

Besides nicotine addiction, nAChRs are effective in other disease, disorder and behaviors, such as Alzheimer's Disease, Parkinson's Disease, schizophrenia, epilepsy, and cancer [38–40].

1.3.1 CHRNA5 and Cancer

Neuronal nAChRs are also expressed in other tissues than lung; they can also be found in muscle, skin and pancreas [41].

Chromosome 15q25.1 contains CHRNA3, CHRNA5 and CHRNB4 genes encoding α3, α5 and β4 subunits of nAChR [42]. Genome-wide studies showed that these genes could be related to cancer development and progress. Studies mainly focused on lung cancer; a single nucleotide polymorphism (SNP) was found strongly associated with lung cancer susceptibility in CHRNA5 [42]. Recent studies have shown that nAChRs can accelerate cell proliferation, tumor invasion, and angiogenesis and can provide resistance against apoptosis [43].

It was found that CHRNA5 was strongly upregulated in lung adenocarcinoma, compared to normal lung tissue. Five different transcripts isoforms of CHRNA5 were identified in normal lung tissue. These isoforms form due to alternative splicing of the 5th exon. It was demonstrated that all mRNA isoforms were upregulated in lung adenocarcinoma with respect to normal lung tissue [44]. Figure 1.5 shows alternative splicing of five CHRNA5 isoforms.

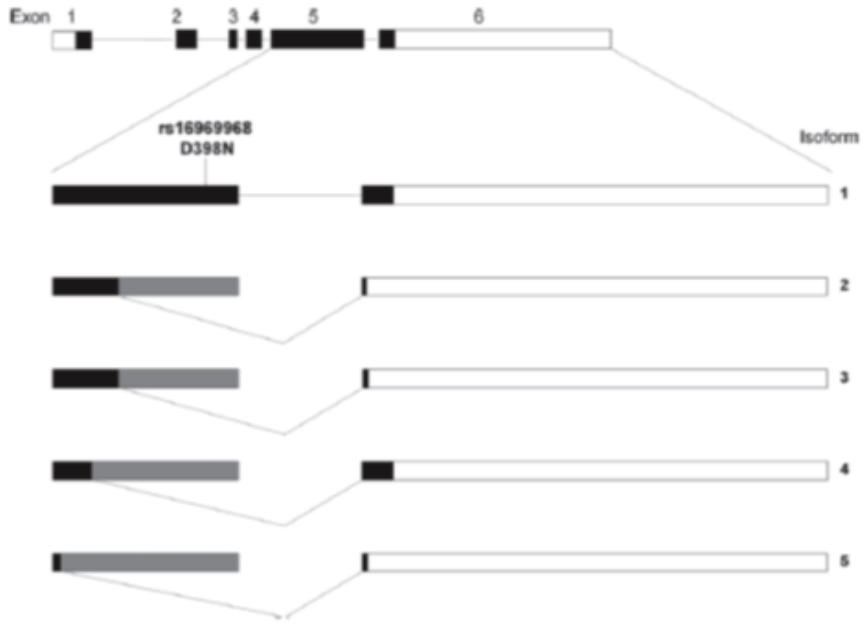


Figure 1.5: Representation of alternative splicing of CHRNA5 in lung. ***(Reprinted from Falvella et al., 2013, with permission from Oxford University Press)***

Also, it was demonstrated that CHRNA5 plays a role in modulating adhesion and motility in bronchial cells and in regulating p63 oncogene [42].

Recent findings suggest that the expression of nAChR subunits including α5 encoded by CHRNA5 might be altered in human breast cancer tissues [45]. Furthermore, some cohort and case–control studies have shown a positive association between cancer and smoking suggesting a connection between nAChRs and breast cancer [43]. However, there is no comprehensive *in vitro* or *in vivo* study that examines the estrogen dependency of isoform specific expression in breast cancer.

1.3.2 Regulation of CHRNA5 Expression

CHRNA5 gene expression can be regulated by the changes in the density of regulatory elements or by indirect effects of changes in expression of other genes [46,47].

IREB2, PSMA4, CHRNA5, CHRNA3 and CHRNB4 on the chromosome 15 are thought to be good candidates for lung cancer and they are co-regulated; their expression increases all together when expression of one of them is induced [47].

SNPs in the cis-regulatory elements may create or diminish cis-acting transcription/enhancer/suppressor factor binding sites, thus accumulation of SNP causing gain of function or alternative splicing induction are very common in genome. rs16969968 SNP on the sequence of CHRNA5 in exon 5, for example, triggers alternative splicing by creating new binding sites for transcription factors resulting the risk of nicotine dependence and cancer [46].

1.3.3 Effects of E2 on Cholinergic Signaling

E2 is thought as a factor affecting nAChRs expression; this also might give a clue to nAChRs-breast cancer relationship. Recent studies have showed that the ER signaling mediates upregulation of nAChR gene expression; nicotine and E2 both induced a9-nAChR expression in human breast cancer cells [43,48].

The increase in nAChR expression with the effect of E2 can also be seen in protein level, however, it is found that E2-induced a9-nAChR protein upregulation can be detected in the ER+ cells compared with the ER- cells [48]. Previous studies in our lab suggested that E2 treatment might increase expression of CHRNA5 based on analysis of microarray studies[49] and by qPCR in a time-dependent manner [50] however no study yet confirmed dose-dependence of increase in CHRNA5 expression based on qPCR analysis of *in vitro* cell lines. Correlation between ER status and CHRNA5 expression in vivo was not addressed yet in the literature either.

1.4 Alternative Splicing and Isoform Formation

Alternative splicing is a post-transcriptional regulation in mRNA which might occur almost in all human genes controlling expression and function [51]. The splicing process provides variation to gene expression. Alterations in splicing can affect the mRNA and protein products of a gene. Alternative splicing has an enormous contribution in protein diversity in metazoan organisms [52]. Different patterns of alternative splicing exist (Figure 1.6): a cassette exon can be either included in the mRNA or excluded. Mutually exclusive exons may occur in the presence of two or more cassette exons. Alternative 5'_ and 3'_ splice sites permit the alterations in the

inclusion of a specific exon. CHRNA's 5th exon represents an example in which alternative splice sites allow the lengthening or shortening of the exon [44].

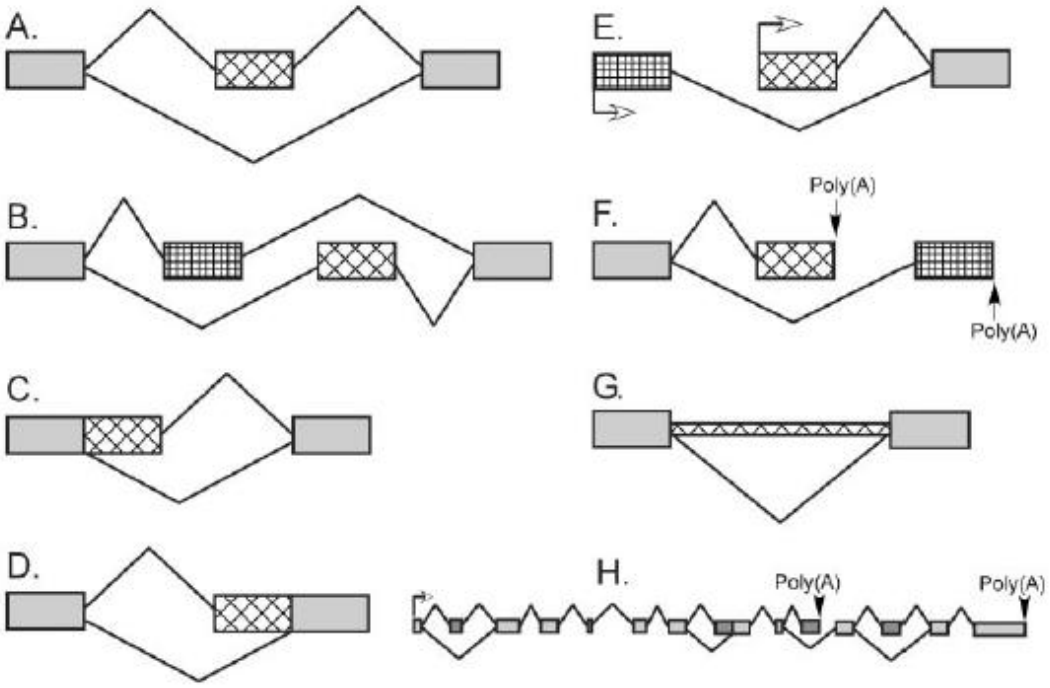


Figure 1.6: Patterns of alternative splicing. Grey boxes show sequences present final spliced mRNA. Stripped boxes are for representation of alternative RNA segments that may or may not be included in the mRNA.
(Reprinted from Black, 2003, with permission from Annual Reviews)

1.4.1 Importance of isoform expression patterns of genes in cancer

Genes related to cancer susceptibility do not always have only one form, due to alternative splicing process same gene can be expressed in different ways. It is known that expression of these alternative spliced forms - isoforms- of genes can play roles in differentiation of cancer cells. Alternative splicing in the protein coding gene of mRNA gives rise to different proteins, therefore they differ in their activities [51]. It is found that different isoforms of cancer related genes can help diagnosis of cancer [44]. For example, different isoforms of CHRNA5 gene show significantly different expression level between normal lung and tumor tissue [44]. Recent studies have reported that transcription factor XBP1 mRNA splicing activates autophagy in endothelial cells through upregulating Beclin1 transcription, a gene found as playing

a role in T-cell natural T-cell lymphoma [51]. Alternative splicing has effects on different types of cancer, such as lymphoblastic leukemia cells, esophageal carcinoma lung cancer and breast cancer [44,51,53].

1.4.2 Alternative Splicing and Breast Cancer

Besides other types of cancers, alternatively spliced forms of genes may play a role in breast cancer initiation and growth. Recent studies revealed that certain isoforms are involved in breast tumor formation from breast epithelial cells [54].

CEACAM1 is a gene that plays a role in some cellular processes including cell-cell adhesion, differentiation, apoptosis and immune response; its pre-mRNA undergoes alternative splicing generating isoforms consisting of an N-terminal and a variable number of multiple extracellular Ig-like domains, a transmembrane domain, and either short (CEACAM1-S) or long (CEACAM1-L) cytoplasmic domains. In normal breast, CEACAM1-S is the predominant isoform, while in breast cancer the S/L ratio highly decreases [54].

Another gene TMEM16A forming Ca2+-activated Cl- channel has been found as alternatively spliced. Multiple protein isoforms of TMEM16A affect the voltage and Ca2+-dependence of the channels. Alternative splicing of exons 6b and 15 of TMEM16A are coordinated in several normal tissues and that this coordination increases in breast tumors [55].

Chemokine CXCL12 signaling is strongly linked to tumor growth and metastasis in breast cancer and some other malignancies. Recent studies have showed that low levels of expression of CXCL12 correlates with worse prognosis in breast cancer with isoform-specific differences among α, β, γ, and δ isoforms [56].

MNK2 gene encoding protein that is a substrate of the MAPK pathway and phosphorylates the translation initiation factor eIF4E is alternatively spliced forming two splicing isoforms Mnk2a, which contains a MAPK-binding domain, and Mnk2b, which lacks it. It is found that the Mnk2a isoform is downregulated in breast tumors [57].

Also it is known that expression of CHRNA5 gene isoforms show changes in cancer such as lung carcinoma, as represented in Figure 1.5. CHRNA5 might play a significant role in breast cancer initiation and growth [43,44]. Therefore, it is possible that expression of CHRNA5 isoforms can change in breast cancer and its subtypes and this warrants further research.

1.5 Isoform Expression Measurement Technologies

It was mentioned that alternative spliced form of genes can play important role in some disease including cancer. This implies the importance of isoform expression measurement in diagnosis. Different isoforms involving tumor initiation or growth can also be used as a biomarker to obtain a reliable diagnosis without complex interventions such as biopsies or surgery. New technologies provide new techniques to determine and measure isoforms of gene of interest.

1.5.1 Microarrays and Analysis Methods

Microarray is one of the most common technology used detection and analysis on gene isoforms. Recent advances in microarray technology rendered the search for the expression of alternatively spliced form possible.

In one study, arrays were designed using 2.5 million oligonucleotide selected probes to differentiate exon alternatives for ~17,250 human genes. ALEXA was designed for the analysis of AS microarrays. It used ~231,000 exon-intron probes, together with ~818,000 oligonucleotide exon-exon junctions probes and ~155,000 exon-internal probes to identify the differentially expressed exon alternatives[58].

1.5.2 qPCR Studies and Analysis Methods

qPCR can be used to identification of alternative spliced isoforms by designing primers or probes targeting the unique exon junctions of spliced isoforms. This method can be used to specifically amplify the long splicing isoform by using one primer in the alternative exon and an opposing primer in a constitutive exon. To detect the short splicing isoform, a boundary-spanning primer (BSP) for the sequence

covering the exon–exon junction with the reverse primer in a fundamental exon can be used [59].

1.5.3 Northern Blots

Northern blot technique is also used to identify alternative spliced isoforms. RNA is run on agarose gel for separation, and then RNA in the gel is transferred onto membranes.

Membrane is hybridized with probes to label RNA of interest, since the RNA is first separated by size, if one probe type can label different size of RNA on membrane it means RNA same gene has more than one product. This way the size of the products can be assessed and the presence of alternative splice products of the same gene or repetitive sequence motifs can be inferred [60].

1.6 Diagnosis and Classifying Cancer with *in vivo* cDNA Expression Patterns

Cancer diagnosis and subtype classification is very crucial to start therapy early. Studies on cancer diagnosis and subtyping should include a very large sample size to be sure of the results and interpretation of the data. Recently, tissue samples taken from cancer patients with different backgrounds are widely used in diagnosis studies. These tissues can be directly used by antibody labeling or staining for a certain set of proteins or gene products of cDNA isolated from these tissues can be used for further analysis. Expression studies, isoform analysis, clustering and subtyping can be performed using cDNA arrays with proper sample sizes.

1.6.1 cDNA Panel Studies of Different Types of Cancer

It is possible to find and utilize various cDNA panels commercially to use for *in vivo* validation of biomarker studies. These panels are produced for more than 20 different types of cancer including pancreas, colorectal, lung, brain, bladder and breast cancer (e.g., ORIGENE). Arrays for only one type of cancer can be found for subtyping and diagnosis analysis; others are for a comparative cancer survey in which cDNA panels including different cancer types can be found.

In a hematopoietic cell and hematological malignancy study, cDNA array consisting of blood cancer samples, expression level of two related genes was examined and changes in their level between normal and cancer tissues were detected, also correlation between these two genes was spotted [61].

A colon cancer array was used to detect isoform expression of gene of interest and most abundant isoform of the gene found in colon cancer was detected [62].

1.6.2 Quantitative and Statistical Methods to Analyze cDNA Panels

Primary method to analysis cDNA panels is quantitative polymerase chain reaction (qPCR). qPCR is a method mainly used for gene expression analysis studies, because it gives the user exact relative quantity of expression. However, in regular PCR it is hard to determine expression change exactly with agarose gel images.

qPCR provides a cycle number (Ct or Cp) for each sample indicating in which cycle the amount of expression of the gene of interest exceed a certain threshold. Therefore, each Ct (threshold cycle) indirectly shows the level of expression when standardized by giving the opportunity to compare samples using these Ct values.

cDNA panels including normal and tumor samples, different types of cancer or different stages or subtypes of same cancer can be analyzed with qPCR [63].

However, qPCR only gives Ct values for each sample; in order to compare and interpret them, additional analyses are required. After normalizing Ct values can be normalized using ΔCt or ΔΔCt methods. In ΔΔCt method, a relative quantification based on relative expression of a target gene versus reference gene can be obtained to investigate the physiological changes in gene expression [64,65]. Upon normalization, statistical tests are used to interpret the significance of a difference in the expression level between groups in cDNA panel. ANOVA, t-test, ANCOVA, sign test are such statistical tests used to analyze raw data. Therefore, analyzing cDNA panels with large sample size will benefit from automation of data analysis [66].

In ΔΔCt method, a relative quantification based on relative expression of a target gene versus reference gene can be gathered to interrogate the physiological fluctuations in gene expression [64,65].

1.7 Tools for Expression Analysis

Importance of expression studies and some different analysis methods to evaluate and interpret the data from gene or isoform expression experiments were mentioned above. Expression studies generally require large sample size to obtain reliable results, however, this makes the analysis of data difficult to handle. Software tools become important for an easy and reliable data analysis. While analyzing an expression data, raw data should be processed and the quality should be assessed.

One of the most common expression data analysis tools that can be reached online is NCBI GEO. It is used not exactly for cDNA panel but for microarray analysis. The database has limited information about a microarray experiment (MIAME)-compliant infrastructure that grasps completely annotated raw and processed data [67].

GEO is currently the largest gene expression resource and it has grown exponentially each year. GEO analysis provides users a gene-centered representation that gives quantitative gene expression data for one gene across a DataSet. [67].

Moreover, some other software can be used to cluster data according the gene expression. Hierarchical clustering can be performed to pin-point to decide the identification of specific tissues from their expression profile through the defined elements, finally in order to unravel the tissue-specific gene expression in determinance of genes or gene clusters, clustering method can be made use of [68].

1.7.1 Creating Tools for cDNA Panel Analysis

MATLAB® (**mat**rix **lab**oratory) is a high-level language and interactive environment for numerical computation and programming. It is possible to analyze data, develop algorithms, and create models and applications using MATLAB as a programming medium [69].

MATLAB provides functions for linear algebra, statistics, filtering, plotting and visualization. These features make MATLAB a proper programming medium for gene expression analysis.

With basic mathematical functions raw data of qPCR of a cDNA panel together with its covariate information can be processed. For example, means for replicate measurements can be taken and ΔCt or ΔΔCt values can be calculated by applying simple mathematics. Where MATLAB is separated from other programs or tools is wide options of statistical analysis it provides; statistical tests include t-test, sign test, ANOVA, ANCOVA, Kruskal-Wallis, chi-square, linear and nonlinear regression etc. It allows choosing the proper test for data analysis [69].

1.7.2 Outlier Analysis

In order to check the quality of data, further analyses should be performed. Performing the experiments with replicates allows distinguishing the outlier data caused by technical or personal mistakes. In real engineering and scientific applications, it is widely known that outliers are most probably to be focused primarily in every operating step [70]. One important aspect of qPCR analysis is that the reaction should produce a single amplicon with a predetermined and stable Tm value obtained upon melt curve analysis. Thus, it is important to detect outliers based on Tm values automatically in a cDNA panel for exclusion since these outliers might represent different amplification products (e.g., dimers) than the one intended.

2. AIM

Literature search indicated that determining molecular subtypes of breast cancer is important in diagnosis. Early diagnosis and therapy of breast cancer is very important to increase survival rates. ER pathway plays role in the development and growth of breast cancer. CHRNA5 activity has effects on pathways controlling cell motility, proliferation, differentiation and apoptosis [42]. Our previous studies[49,50] demonstrated that CHRNA5 might be regulated by estrogen [49] based on microarray data expression analysis and was modulated by estrogen in a time dependent manner. *In vivo*, ER- cells were likely to express more CHRNA5 based on GEO dataset analyses [49]. However, *in vivo* cDNA panel validation of CHRNA5 expression in association with ER status has not been performed yet.

In this study, our aim is a) to test whether isoform specific CHRNA5 expression responds similarly to estrogen exposure in ER positive and ER negative cell lines in a dose dependent manner; b) to assess effects of the E2 antagonist tamoxifen and long term E2 depletion in MCF7 cells; and c) to investigate isoform specific expression profile *in vivo* for diagnosis of breast cancer.

Our experimental approaches include testing CHRNA5 isoform expression by qPCR in response to different doses of E2 in MCF7 (ER+) and MDA MB 231 (ER-) cell lines, and as well as CHRNA5 response to tamoxifen, an inhibitor of estrogen response, when given alone or together with estrogen in MCF7 cells. We also developed a long term estrogen depletion (LTED) model in MCF7 cells to test whether LTED affects ESR1 as well as CHRNA5 isoform expression. We then tested whether CHRNA5 expression differs among a set of tumor samples with different molecular subtype characteristics using a commercial cDNA panel. We developed MATLAB cDNA panel analysis GUI for quality control and statistical analysis of qPCR data obtained from OriGene cDNA panels.

3. MATERIALS AND METHODS

3.1 Cell Culture Protocols

3.1.1 Cell Lines

For this study, two different cell lines (MCF7, MDA MB 231) were used (ATCC; Manassas, USA). They have been cultured in Bilkent University MBG laboratories. In order to study the effect of E2, ER- and ER+ cell lines were studied. The origin of the cell lines can be found in Table 3.1 [71].

Table 3.1: Breast cancer cell lines origin

Name	Source	Tumor	Age	Ethicity
MCF7	PE	IDC	69	White
MDA MB 231	PE	AC	51	White

AC, adenocarcinoma; IDC, invasive ductal carcinoma; PE, pleural effusion.

The receptor status and molecular subtypes of cell lines are listed in Table 3.2.

Table 3.2: The receptor status and molecular subtypes of cell lines

Name	ER	PR	HER2	Molecular subtype
MCF7	+	+	-	Luminal A
MDA MB 231	-	-	-	Basal

3.1.2 Cell Growth Conditions

DMEM (SH30021.01), FBS (SW30160.03), Penicillin/Streptomycin Solution (SV30010) and Trypsin/EDTA solution (SH30042.01) were from HyClone (Logan,

USA). Non-essential amino acid solution (K0293) (Berlin, Germany) and PBS (17-516F) (Lonza ,Switzerland) were used to prepare medium.

The complete growth medium of MCF7 and MDA-MB-231 cells included DMEM with 10% Fetal Bovine Serum, 1% Penicillin/Streptomycin solution and 1% Non-essential amino acids.

3.1.3 Thawing Cells

The frozen cell vial obtained from liquid nitrogen was placed in the water bath at 37°C. After melting, the cells were taken to a 15 ml, sterile tube and total volume was adjusted to 5 ml with complete medium. The cells were then centrifuged at 1500 rpm for 5 minutes; the supernatant was aspirated; the cells were re-suspended in 1 ml of complete growth medium and seeded in T25 flasks (Greiner Bio One, Frickenhausen, Germany) with 7-8 ml of complete medium.

3.1.4 Subculturing of Cells

All cell types were grown in T75 flasks (Grenier Bio One, Frickenhausen, Germany) with 12-13 ml of complete medium in an incubator with 5% CO2 at 37°C. When the cells reached 80-90% confluency, on average, after 3 days, they were passaged to new flasks. For passaging, the old growth medium was aspirated; flasks were washed with 5 ml PBS. 1 ml Trypsin/EDTA solution was then added to cover the bottom of them for 5-8 minutes in 37°C incubator for detachment of cells, which were then collected from flasks using 7-8 ml complete growth medium and centrifuged at 1500 rpm for 5 minutes. The supernatant was aspirated; the cells were re-suspended in complete medium and were seeded into new flasks with the desired dilution depending on the amount of pellet and the type of experiment.

3.1.5 Cryopreservation of Cells

For cryopreservation, detached cells which were aspirated the old medium washed with PBS and trypsinized were centrifuged at 1500 rpm from 5 minutes. Medium were aspirated; and pellet was re-suspended in freezing medium (10% DMSO in FBS). 1.5-2 ml of this was put into cryotube and stored at -20°C for 1-2 hours, then

transferred to - 80°C. For long-term storage the cryotubes should be stored in liquid nitrogen.

3.1.6 Counting Cells with Hemocytometer

Cells were counted before seeding them for an experiment using hemocytometer.

Cell pellet obtained after 5 minutes of centrifugation at 1500rpm was completely dissolved in 10ml medium; medium was chosen depending on cell type. Then cells were diluted 1:10 and 10 µl form this diluted solution was put on hemocytometer. Cells from four corners (Figure 3.1.B) of hemocytometer were counted and the total number was divided to 4. The division gave the number of how many millions cells 10ml medium had.

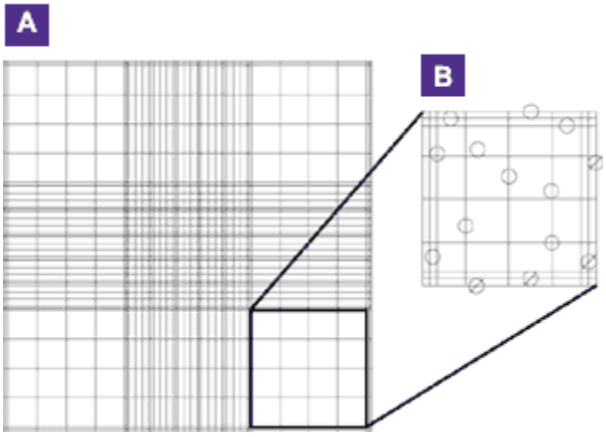


Figure 3.1: A. Scheme of a whole hemocytometer. B. Right bottom corner of hemocytometer

3.1.7 *β*-Estradiol (E2) Treatment

For *β*-Estradiol (E2) (E2758; Sigma Aldrich, Missouri,USA) treatment, MCF7 and MDA-MB231 cell lines were used. E2 stock solution was prepared by dissolving 27.238 mg E2 in 50 ml 100% sterile EtOH (32221) (Sigma Aldrich, Missouri, USA) so final concentration was 2mM. The stock solution was filtered; aliquoted; and stored in -20°C.

Final E2 treatment concentration were prepared as 1nM, 10nM, and 100nM dilutions were made in phenol red-free Dulbecco's MEM (F0475; Biochrom AG, Berlin, Germany) with 5% charcoal/dextran-treated FBS (SH30068.02) (Hyclone, Logan,

USA), 1% penicillin/streptomycin solution and 1% non-essential amino acids. First, 2mM stock solution was serially diluted as 1:10, 1:100 to obtain 0.2 mM and 0.02mM E2 solutions, then from each solutions (0.02 mM, 0.2 mM, 2 mM) 1 µl was put into 20ml phenol red-free DMEM with 5% charcoal/dextran-treated FBS to obtain 1nm, 10nm and 100nm E2 solutions, respectively.

The medium for E2 control group was prepared by adding 1 µl of 100% EtOH to 20 ml phenol red-free DMEM with 5% charcoal/dextran-treated FBS, because for all doses of E2, same amount of (0.005v/v % EtOH) 100% EtOH was used.

MCF7 and MDA MB231 were counted and seeded in 8 T75 flasks for each cell line at a concentration of 1 x 10^6 cells/ flask. They were incubated with complete DMEM (SW30160.0, HyClone (Logan, USA)) for 24 hours. After 24 hours, the media of the cells were changed to phenol red-free DMEM with 5% charcoal dextran-treated FBS and incubated for 72 hours. Then, the media then was changed to phenol red-free DMEM with 0.1% charcoal/dextran-treated FBS to synchronize cells for 24 hours. 0.1% FBS medium was aspirated and replaced with 12 ml of dose-specific media (Control, 1nM, 10nM,100nM) as explained above was added to the flasks. Each dose was performed in duplicates. After 24 hours all treatments were collected.

3.1.8 Tamoxifen and E2 Treatment

For tamoxifen (4-OHT) (S1972, Selleck Chem) and E2 treatment MCF7 cell line was used.

2mM E2 stock solution prepared as explained above was diluted 1:10 and 0.5 µl of this was added to 10ml phenol red-free DMEM with 5% charcoal/dextran-treated FBS to obtain 10nm E2 concentration in medium.

To obtain 1µM 4-OHT concentration in medium, 1µl from 10mM 4-OHT was added in 10ml phenol red-free DMEM with 5% charcoal/dextran-treated FBS.

The medium for E2 control group was prepared by adding 1 µl of 100% EtOH to 20 ml phenol red-free DMEM with 5% charcoal/dextran-treated FBS (0.005v/v % EtOH), and the medium for 4-OHT control group was prepared by adding 0.5 µl of

DMSO to 5 ml (0.01v/v % DMSO) phenol red-free DMEM with 5% charcoal/dextran-treated FBS.

MCF7 breast cancer cells/well were counted and seeded in 8 wells of 2 6-well plate as $2x10^5$ cells/well. They were incubated in complete DMEM for 24h. Then same procedure with β-Estradiol (E2) treatment has been performed until 6th day of seeding. After 24h synchronization, cells were treated with 1uM 4-OHT+10nm E2, 10nm E2, 1uM 4-OHT, EtOH and DMSO as duplicated for 24h. After 24 hours all treatments were collected.

For expression analysis deltadeltaCT (TPT1-Target gene (deltaCt) normalized with control) method was used. 4-OHT treated samples were normalized with DMSO, E2 treated samples were normalized with EtOH and 4-OHT-E2 treated samples were normalized with average of DMSO-EtOH.

3.1.9 Long-term E2 Depletion of Cells

In order to establish a long-term E2 depletion model, MCF7 cell line was used.

To prepare depletion and control medium, phenol red-free DMEM with 10% charcoal/dextran-treated FBS was used. For control, 2mM E2 stock solution was diluted 1/10 to obtain 0.2mM E2 and 2µl from this was added to 40ml phenol red-free DMEM with 10% charcoal/dextran-treated FBS. For depletion, only 2µl EtOH was added in 40ml phenol red-free DMEM (0.005v/v % EtOH) with 10% charcoal/dextran-treated FBS.

For 3 months, MCF7, divided into two flasks after growing in complete DMEM for 2 days, was cultured in depletion and control medium. Pellets were taken in every three days. After 3 months, cells were collected for cryopreservation for 2 months. Then, cells $4x10^6$ were seeded, grown for 3 days, and pellets were collected for 4 weeks.

There are totally 15 samples from 15 different time points, first three samples belong to first month, 4.-6. Samples belong to second month and samples from 7 to 9 are from third month, while last five samples belong to sixth month (Table 8.2).

3.2 Gene Expression Analyses

3.2.1 Total RNA Isolation

Cell pellets taken from flasks should be preserved at -80°C. Before total RNA isolation pellets were taken from -80°C and thawed on ice. Then, they were completely dissolved in 1ml TRIzol Reagent (15596-026) (Life Technologies, California, USA). Dissolved lysates were transferred into 1.5ml tubes and 200µl chloroform was added to each tubes. Tubes were shaken vigorously for 15 seconds and set for 3 minutes at room temperature. Then tubes were centrifuged at 13.200 rpm for 17 minutes at 4°C. In this critical step, lysates were separated into three phases; the top and clear phase included RNA; this was carefully taken and transferred into a new tube. 500µl isopropanol was added to new tubes and tubes were gently inverted for 4-5 times. After 10 minutes at RT, they were centrifuged at 4°C for 12 minutes at 13.200 rpm. The supernatant was carefully removed to preserve pellet at the bottom of the tubes, the pellet as washed with 75% EtOH and centrifuged at 8.000 rpm for 8 minutes at 4°C. EtOH was discarded; the pellet was washed with 100% EtOH and centrifuged at 8.000 rpm for 8 minutes at 4°C. The pellet was dried after discarding EtOH; and dissolved in 30-50µl Hypure Molecular Biology Grade water (SH30538.03) (Hyclone, Logan, USA). RNA concentration and quality were measured with NanoDrop® ND-1000 spectrophotometer (NanoDrop Technologies, USA). RNA was stored at -80°C.

3.2.2 cDNA Synthesis

Total RNA was used to synthesize cDNA. RevertAid First Strand cDNA Synthesis Kit (K1621) (Thermo Fisher Scientific, Massachusetts, USA) was preferred to synthesize cDNA. In order to synthesize 1ng cDNA, required volume of RNA was calculated and H_2O to complete RNA volume to 11µl was added to tubes. 1µl oligo(dT) primer was used. Other components were added in this order; 4 µl 5x Reaction Buffer, 1 µl RiboLock RNase Inhibitor, 2 µl 10 mM dNTP Mix, 1 µl RevertAid M-MuLV Reverse Transcriptase. The tubes were incubated at 42°C for 60 min, then 5 minutes at 70°C. They were stored at - 20°C.

3.2.3 Oligonucleotides

Table 3.3 presents the primers used in this study.

Table 3.3: Primer sequences, their product sizes and efficiencies

Name	Primer sequence (5'-3')	Size of amplicon(bp)	Amplification efficiency
CHRNA5_v1	F: AGATGGAACCCTGATGACTATGGT R:AAACGTCCATCTGCATTATCAAAC	104	1.87
CHRNA5_v2	F:GGAAACTGAGAGTGGTAGTGGA R:CTTCAACAACCTCACGGACA	122	1.95
CHRNA5_v3*	F:CATCAGGTGTTGAAGATTGGAAAT R:AAAAAGCCCAAGAGATCCAACAAT	101	1.92
CHRNA5_iso2**	F:TGGAGAATGGGAGATTGTGAGTGCA R:CCAATCTTCAACAACCAGCAACAGC	78	1.97
CHRNA5_iso3**	F:TGGAGAATGGGAGATTGTGAGTGCA R:CCAATCTTCAACAACGGATACCAGC	78	1.96
TPT1	F:GATCGCGGACGGGTTGT R:TTCAGCGGAGGCATTTCC	100	1.95
PS2 (TFF1)***	F:CCATGGAGAACAAGGTGATCTGC R:TTAGGATAGAAGCACCAGGGGAC	208	1.85
ESR1	F:AGACATGAGAGCTGCCAACC R: GCCAGGCACATTCTAGAAGG	299	1.78
FOSL2	F: GGCCCAGTGTGCAAGATTAGCC R: TTTCACCACTACAGCGCCCACC	105	1.99

*(Warzecha et al. 2009)

** (Falvella et al. 2013)

***Alotaibi et al. 2006

CHRNA5 primers used in this study bind the regions showed in Figure 3.2 on exons 4- 6 of CHRNA5.

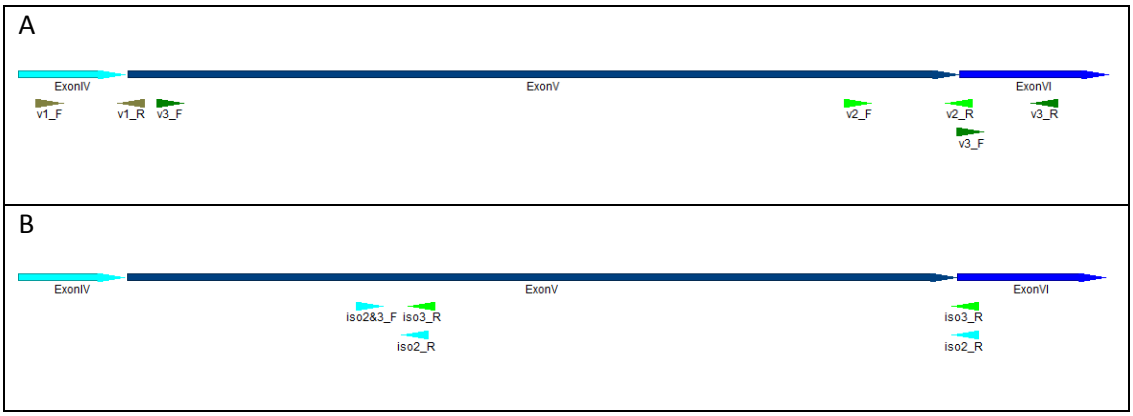


Figure 3.2: Scheme of binding regions of CHRNA5 isoform primers A. Newly designed CHRNA5 isoforms used in the experiments on exon 4 – 6 of CHRNA5; CHRNA5_v1_F&R detects multiple isoforms while other ones correspond to a transcript that contain a certain part of the gene. B. Previously known (Falvella et al. 2013) CHRNA5 isoforms used in the experiments on exon 4-6 of CHRNA5 (Only ~150bp of ExonVI was showed in this figure 1446bp-1596bp).

3.2.4 Gradient Polymerase Chain Reaction

Gradient Polymerase chain reaction (gPCR) was performed using OneTaq Quick-Load Master Mix (M0486L) (New England Biolabs, Massachusetts, USA). New primers were tested with this technique using 5 different annealing temperatures (54-62°C) to see in which Tm primers work best. Gradient PCR reaction was prepared as explained in Table 3.4.

Table 3.4: PCR reagents and their volumes

Reagent	Volume (µl)
2x SYBR Green I Master Mix	5
Forward Primer (10pmol)	1
Reverse Primer(10pmol)	1
DNase/RNase-free water	1
cDNA	2
Total	**10**

PCR was performed under conditions listed in Table 3.5.

Table 3.5: PCR conditions

Step	Temperature & Duration
Initial Denaturation	95°C, 5’
Denaturation Annealing Extension	95°C, 30” 58°C Δ10, 1.5’ 68°C 1’ } 40 cycles
Final Extension	68°C 10’

3.2.5 Agarose Gel Electrophoresis

1.5% agarose gel was prepared with 200ml 1xTAE buffer and 3gr agarose. 1mg/ml ethidium bromide solution was added and allowed to polymerase. The wells were loaded with 10-12 µl PCR product (for 22 well/gel) or 8-10µl PCR product (for 40 well/gel). 3 µl GeneRuler 100bp DNA ladder (SM0241) (Thermo Scientific, Waltham, Massachusetts, USA) was used as marker.

3.2.6 Quantitative Real Time-Polymerase Chain Reaction (qPCR)

Quantitative real time-polymerase chain reactions (qPCR) were performed using LightCycler® 480 SYBR Green I Master (04887352001) (Roche Applied Science, Germany). Experiments were performed using every sample as at least duplicates or more, including negative controls.

Plates were LightCycler® 480 Multiwell Plate 96 (04729692001) and (Roche Applied Sciences, Germany) PCR machine was LightCycler® 480 PCR Instrument (Roche Applied Science, Germany). qPCR reaction was prepared as explained in Table 3.6

Table 3.6: qPCR reagents and their volumes

Reagent	Volume (µl)
2x OneTaq QuickLoad Master Mix	12.5
Forward primer (10pmol)	1
Reverse primer (10pmol)	1
cDNA	1
DNase/RNase-free water	9.5
Total	25

qPCR was performed under conditions listed in Table 3.7.

Table 3.7: qPCR conditions

Step	Temperature & Duration
Initial Denaturation	95°C, 5’
Denaturation Annealing Extension	95°C, 10” 58°C, 20” 72°C, 20” (50 cycles)
Acquisition	95°C, 5” 55°C, 1’ 95°C, Continuous acquisition
Final Extension	40°C, 10”

3.2.7 Detection of qPCR Outliers Based on Tm Values from Melt Curves

In order to detect outliers in qPCR data, samples having relatively different melting temperature (Tm) values rather than the expected Tm for the primer pairs were loaded onto 1.5% agarose gel. 2µl 6x loading dye was added into 10 µl qPCR samples and 8µl of samples was loaded into well.

It was observed that samples having approximately Tm 1°C smaller or l°C larger than expected Tm (detected by taking median of all Tm values in the cDNA panel) generally provided bands with different sizes than the expected PCR product sizes. They were characterized as primer dimers, DNA contaminations or different dimerization products. To increase the reliability of data, samples having ±1 different Tm than regular Tm were removed and was not included in future analyses.

3.2.8 Primer Efficiency

Primers should be tested before using them in order to be sure of their amplification efficiencies. It is important to obtain reliable results in every experiment for which the primers are used. Optimizing the conditions which primers work best under and calculating efficiency of primers are needed to be sure of the quality of primers and trustworthiness results.

Primer efficiencies for each pair of primers were calculated to 1/5 fold dilution series and for each gene it was repeated at least once before taking the average efficiency value. Perfect PCR efficiency should demonstrate a change of 2.3 cycles between 5 fold dilutions of the template.

Primer amplification efficiencies listed in Table 3.2 were obtained by tested primers with serially diluted cDNA. Two cDNA pools obtained from MDA MB 231, MCF7, HCC1937, BT20, and MCF7; and MDA MB 231, MCF7, MDA MB 361, BT20 and MDA MB 468 cell lines serially diluted as 1:5, 1:25, 1:125 and 1:625. Listed primers were tested using these four dilutions in a regular qPCR reaction described in Table 3.4.

Threshold cycle (Ct) values and logarithm of dilutions were plotted; the efficiency was calculated with this formula;

E= 10^(-1/m)

Where, E is primer efficiency and m is the slope of Ct-log(dilutions) graph.

3.2.9 ΔCt and ΔΔCt Analysis of qPCR Data

In order to determine the relative quantification of a target gene in comparison to a reference gene ΔCt and ΔΔCt values depending on primer efficiencies were calculated.

For E2, 4-OHT and LTED experiments, after taking mean of Ct value for duplicates, ΔΔCt values were calculated using the formula;

$$\Delta\Delta Ct = (E_{target}^{\Delta Ct(target)} / E_{ref}^{\Delta Ct(ref)})$$

$$\log_2(\Delta\Delta Ct)$$

E is primer efficiency,

$\Delta Ct_{(target)}$ is control $_{target}$-treatment $_{target}$

$\Delta Ct_{(ref)}$ is control $_{ref}$-treatment $_{ref}$

In this analysis method, a relative quantification based on relative expression of a target gene versus reference gene can be obtained to investigate the physiological changes in gene expression[64,65].

For cDNA panel experiments, efficiency corrected ΔCt values, as defined here correspond to -Δ**Ct** considering the direction of Ct values, were calculated using this formula;

$$\Delta Ct = \log2(E_{ref}^{Ct(ref)} / E_{target}^{Ct(target)})$$

3.2.10 ANOVA analysis

ANOVA test was used in analysis of each qPCR experiment in order to test the significance of the data. After taking data and checking the Tm as described, an excel file was created including data and the groups wanted to be analyzed. Each treatment or each cell line was labeled to introduce them into ANOVA test.

General function construct used in the analysis of this study was

[P,ANOVATAB,STATS] = ANOVA1(X, GROUP)

Where ANOVA1performs a one-way ANOVA for comparing the means of two or more groups of data, it returns the p-value (P) for the null hypothesis that the means of the group means are equal. Multiple test corrections were performed using *multcompare* function in MATLAB. Two different multcompare methods were used, the 'hsd' option stands for 'honestly significant difference'. The 'lsd'option stands for 'least significant difference' and uses plain t-test, it provides no protection against multiple comparison problems.

3.3 Analysis of E2 Treatment

In order to clearly present E2 treatment data, efficiency corrected $\log_2(\Delta\Delta Ct)$ values were used in clustering and drawing heatmap. Similarly, ΔCt values were used also in the clustering to show the differences (before normalization with control group) in PS2, an E2 responsive gene, between ER+ and ER- cell lines.

3.3.1 Clustering

Cluster 3.0 [72] program was used to cluster $\Delta\Delta Ct$ values of E2 treatment experiment. Euclidean distance was used while clustering average linkage of data.

3.3.2 Treeview analysis

TreeView program was used to draw heatmap using clustering data obtained from Cluster 3.0. High-expression was represented with red color while green was chosen for low-expression (down-regulation).

3.4 Analysis of 4-OHT and E2 Treatment

In order to clearly present E2 treatment data, efficiency corrected $\log_2(\Delta\Delta Ct)$ values were graphed using GraphPad Prism 6 program. Five genes and data belonging to their expressions under E2+4-OHT, 4-OHT, E2 treatment and control were plotted using GraphPad. Grouped graphic type was chosen to be able show data belonging to same gene in the same cluster. Gene names were placed on x-axes, mean of set1 and set2 for each treatment group was taken by GraphPad and error bars showing standart deviation were added to y-axes.

3.5 Analysis of Long-term Estrogen Depletion Experiment

In order to clearly present LTED data, efficiency corrected ΔΔCt values were graphed using GraphPad Prism 5 program and sign test was applied to see whether there is a consistent increasing or decreasing positive or negative trend after LTED .

ΔΔCt values (treatment (T) value was paired with its control (C) at a specific time point) were calculated and they were plotted for each time point. This graph type was chosen to see how the difference between C and T groups changed throughout six months. Sign test was applied the C-T values for each gene using p=signtest(x) command in Matlab; x is an array containing C-T values.

3.5 Analysis of cDNA Panel

Breast Cancer cDNA Array II (BCRT101) (OriGene, Rockville,USA) consists of 48 different breast cDNA, 7 of them are normal and remaining 41 are cDNAs obtained from breast tumor tissues [Breast cancer cDNA Array II SKU# BCRT501]. The datasheet of panel provides information about gender, age, ethnicity, ER, PR and HER2 status, grade, stage, pathology of samples (Table 8.1). While the detailed information provides an opportunity to evaluate data from different aspects, analyzing data becomes more complex and requires automation; however, there is no free software or GUI that enables that for ORIGINE cDNA panels. Therefore, functions to extract, normalize, perform Tm check, and statistically analyze and plot

the cDNA qPCR data for target genes were written in MATLAB in a user friendly GUI.

3.5.1 Statistical Analysis of qPCR Data

In order to test the significance of results obtained from cDNA panels mainly ANOVA test was performed using efficiency corrected $\log_2(\Delta Ct)$. Efficiency corrected ΔCt= Eff ^(reference Ct)/Eff ^(target Ct).

Also correlation between isoform of CHRNA5 for cancer samples was plotted to see if isoforms behave in same way under same conditions. Finally in order to make the codes more sophisticated and user friendly a GUI application was designed allowing to user choose data, correct, and publish them (please see Results).

3.5.1.1 Correlation Analysis

Each gene tested in cDNA panel plotted against each other to see whether they were correlated or not, plots were categorized according their receptor status and stages. Each group in a category was plotted using different color scale. i.e in ER status plot ER- and ER+ samples were labeled using different colors. This allows seeing how genes were clustering according to receptor status and stages.

3.5.1.2 MATLAB GUI Application

The MATLAB codes for this application are available and the process has been shown in the Results section.

3.6 General Solutions

• **50x Tris-acetic-EDTA (TAE):** 242 g Tris base and 18.6 g EDTA were dissolved in ddH_2O. 57.1 ml glacial acetic acid was added and the volume was completed to 1 lt. pH was set to 8.0.

• **6X DNA Loading Dye:** 10mM Tris-HCl (pH 7.6), 0.03% bromophenol blue, 0.03% xylene cyanol, 60% glycerol, 60mM EDTA (0.5M pH8.0)

• **Ethidium Bromide:** 10 mg/ml dissolved in ddH_2O (stock) 1:10.000 working dilution (1µl per 10ml of gel)

4.RESULTS

4.1 Effects of Different Doses of E2 on CHRNA5 Isoform Expression

4.1.1 Preliminary PCR for MCF7 and MDA-MB-231 Treated with Different Doses of Estradiol (E2)

After the E2 treatment treated cDNAs obtained from RNA of treated cells were tested with CHRNA5_iso2 primer by regular PCR. The aim was to see whether expression difference between E2 doses can be recognized with regular PCR with 35 cycles.

Figure 4.1 shows the results of PCR such that, a slight increase can be seen with increasing doses. However, the agarose gel image is not enough to judge the effects of E2 on expression of the gene showing importance of qPCR for detection of quantitative changes in gene expression.

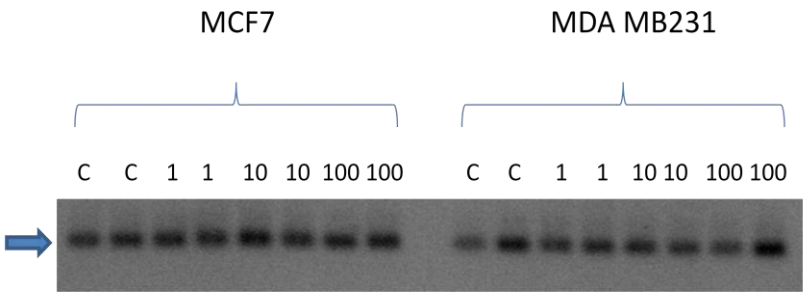


Figure 4.1: Representative expression of CHRNA_iso2 products across ER-and ER+ breast cancer cell lines treated with 1nm, 10nm and 100nm E2. C indicates the control groups. Each dose has a biological duplicate. The arrow indicates the expected amplicon size which is 93bp.

4.1.2 Statistical Analysis of E2 Treatment with MATLAB

In order to decide the statistical significance of expression changes in CHRNA5 isoforms after E2 treatment between different doses of treatment, ANOVA was used and calculations were performed in MATLAB.

First the mean expression difference between different doses and cell lines were tested with two-way ANOVA analysis. ER+ MCF7 and ER- MDA MB 231 under

treatment of E2 were tested to see how different ER+ and ER- cell lines behave upon exposure to E2 treatment (Table 4.1, top).

Table 4.1: p-values of each gene comparing MCF7 and MDA MB 231 cell lines and different doses of E2 based on two-way and one-way ANOVA

	ΔΔCt		ΔCt
Genes	**p-values (Cell line)**	**p-values (Doses)**	**p-values (Cell line)**
CHRNA5_v2	0.0006	0.085	0.8484
CHRNA5_v3	0.0075	0.5849	0.122
CHRNA5_iso2	0.0027	0	0.3525
CHRNA5_iso3	0.0052	0.009	0.4642
PS2	0.0063	0.002	0

	v2	**v3**	**iso2**	**iso3**	**PS2**
p-value	0.001	0.0478	0.0002	0.0098	0.0054
Mult compare	M C-M 10nm M C-M 100nm M 1nm- MD 1nm M 10nm- MD C M 10nm- MD 1nm M 10nm-MD 10nm M 10nm- MD 100nm M 100nm- MD C M 100nm- MD 1nm M 100nm- MD 10nm M 100nm-MD 100nm		M C-M 1nm M C-M 10nm M C-M 100nm M C-MD 10nm M C-MD 100nm M 1nm- MD C M 10nm- MD C M 10nm- MD 1nm M 100nm- MD C M 100nm-MD 1nm MD C- MD 10nm MD C-MD 100	M C-M 1nm M C-M 100nm M 10nm- MD C M 100nm- MD C	M C-M 10nm M C-M 100nm M 10nm- MD C M 100nm- MD C

The results demonstrated that each CHRNA5 iso2 and iso3 isoforms and PS2 (TFF1), an E2 responsive gene, showed significantly different expression under different doses of E2 treatment; moreover effect of E2 treatment differed significantly from cell line to cell line for all genes tested at $p<0.05$ or less. While ΔΔCt shows the fold change of expression, ΔCt implies absolute values of

expression; that is why only PS2 expression shows highly significant expression difference between ER- and ER+ cell lines based on deltaCt Two-Way-ANOVA (Table 4.1).

Analyzing MCF7 and MDA MB 231 data together by One-Way-ANOVA gave us the opportunity to increase sample size and multiple test comparisons showed differences between cell lines and between doses (Table 4.1, bottom). After showing the expression level difference between ER+ and ER- cell lines as in Table 4.1, each of the cell lines were analyzed one by one to see the effect of different E2 doses on CHRNA5 isoform expression level using One-Way ANOVA and *multcompare* in MATLAB (Tables 4.2-4.5).

Accordingly, MCF7 cell line which is ER+, was generally more affected by E2 treatments than ER- cell line MDA-MB-231 (Figure 4.2).

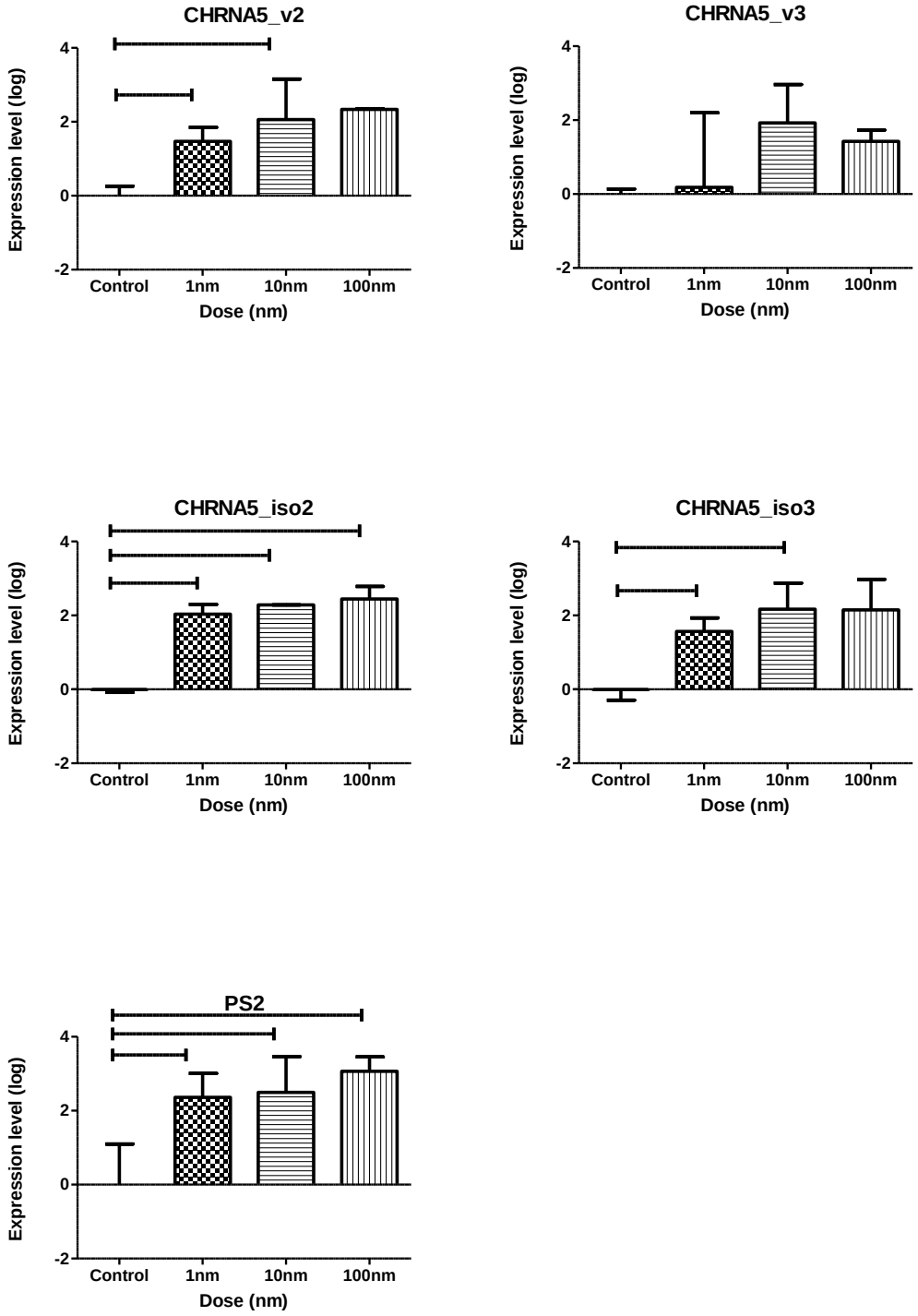


Figure 4.2: Effects of E2 doses on CHRNA5 isoforms and PS2 in MCF7 cell line (ΔΔCt). Lines above the graphs indicate the group having statistically different expression using lsd (t-test equivalent) method.

CHRNA5_v2 and CHRNA5_iso2 was directly affected by increases in E2 doses, however CHRNA5_v3 did not follow that kind of trend by showing a higher expression under 10nm E2 treatment than 100nm E2 treatment. CHRNA5_iso3 reached saturation after 10nm dose and did not show any difference in expression level even after 100nm E2 treatment. PS2 is known a responsive gene to E2 treatment, so it was used as control for E2 and despite closeness of PS2 expression levels in 1nm and 10nm E2 treatments, an increase with rising E2 doses can be observed in PS2 expression. p-values for each group can be seen in Table 4.2.

Table 4.2: p-values of each gene comparing E2 doses in MCF7 cell line from ANOVA test

Genes	**p-values**
CHRNA5_v2	0.055
CHRNA5_v3	0.3794
CHRNA5_iso2	0.0011
CHRNA5_iso3	0.0592
PS2	0.0675
Average of CHRNA5 isoforms	0.0277

Multiple comparisons between each group were performed with and without correction against the multiple comparison problem and groups giving significant results were listed in Table 4.3.

Table 4.3: List of group having statistically significant expression difference for each gene

Multcompare type	Genes	Groups with significant expression difference
With correction (hsd)	CHRNA5_iso2	Control-1nm,Control-10nm,Control-100nm
	Ave. of CHRNA5 isoforms	Control-10nm, Control-100nm
Without correction (lsd)	CHRNA5_v2	Control-10nm, Control-100nm
	CHRNA5_iso2	Control-1nm, Control-10nm, Control-100nm
	CHRNA5_iso3	Control-10nm, control-100nm
	PS2	Control-1nm, Control-10nm, Control-100nm
	Ave. of CHRNA5 isoforms	Control-1nm, Control-10nm, Control-100nm

CHRNA5_iso2 showed highly significant expression difference depending on E2 doses while the others had test statistics approaching significance (<0.06) except CHRNA5_v3. However, ANOVA may cause the loss of significance because of low sample size; therefore a multiple comparison without correction (raw t-test values) was applied, too. It was seen that almost all doses significantly affect CHRNA5 gene expression except in v3 (Table 4.3). Also average of CHRNA5 isoforms were calculated and ANOVA was applied to obtain a generic idea about CHRNA5 expression-E2 treatment relationship and it was seen CHRNA5 expression was affected significantly (Table 4.2, Table 4.3).

Effects of different doses of E2 on MDA MB 231 cell line was also tested with ANOVA test and plotted (Figure 4.3).

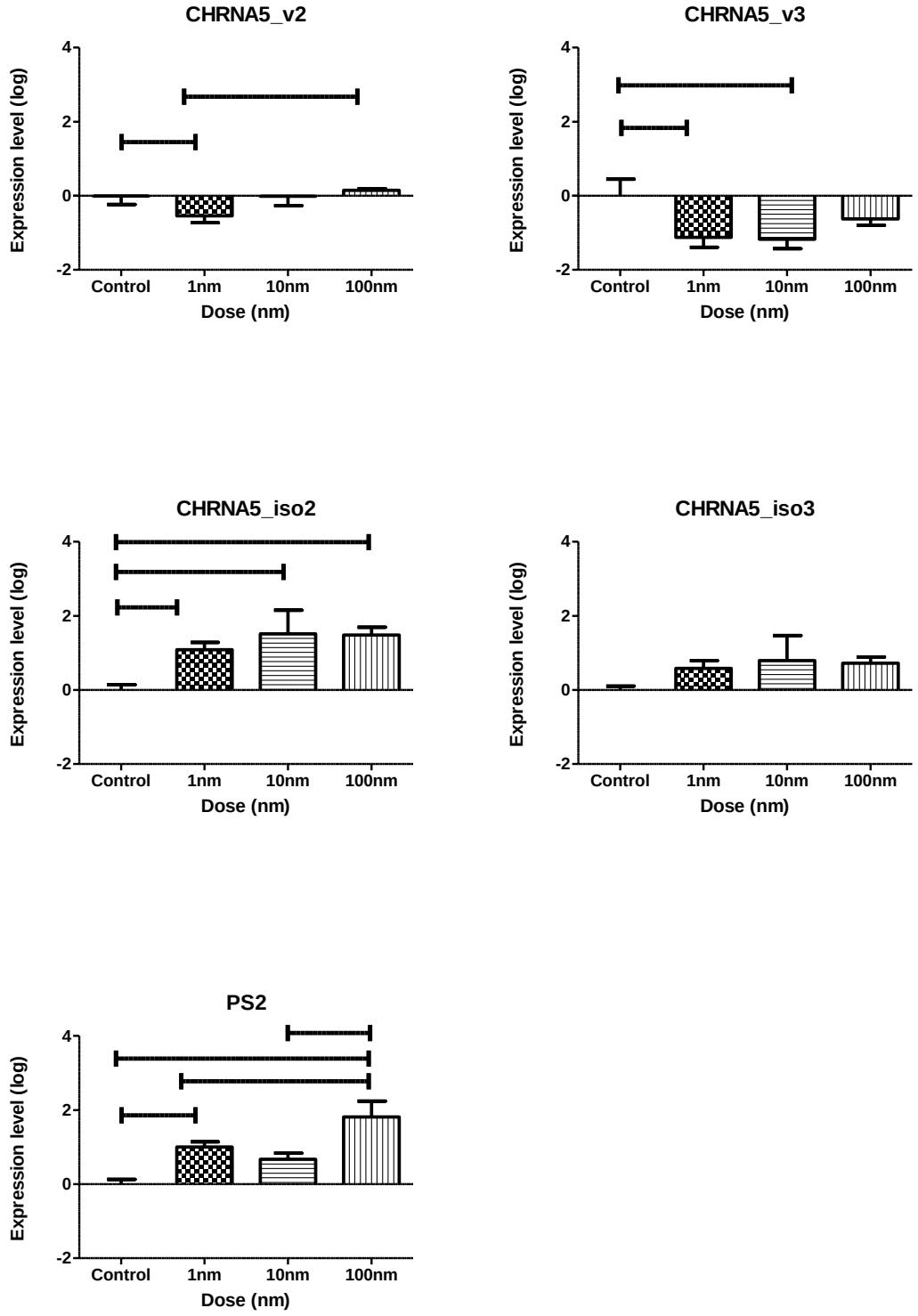


Figure 4.3: Effects of E2 doses on CHRNA5 isoforms and PS2 in MDA MB 231 cell line (ΔΔCT). Lines above the graphs indicate the group having statistically different expression using lsd (t-test equivalent) method.

Any of the tested genes in MDA MB 231 cell line did not show a regular expression trend in a dose dependent manner although some isoforms exhibited increasing or decreasing expression across treatments (Figure 4.3). According to ANOVA comparing control and each dose of E2 in MDA MB 231 cell line, CHRNA5_v3, CHRNA5_iso2 and PS2 expression have significantly changed (Table 4.4) but CHRNA5_v3 was downregulated while CHRNA5_iso2 and PS2 were upregulated (Figure 4.3). The observation that CHRNA5_v2 and CHRNA5_v3 isoforms exhibited dose dependent decreases in response to E2 in MDA-MB-231 and an increase in MCF7 showed that it was important to analyze isoform specificity in response to E2. On the other hand iso2 was positively affected by E2 in both cell lines based on ANOVA analyses followed by pairwise comparisons (Tables 4,4 and 4.5).

Table 4.4: p-values of each gene comparing E2 doses in MDA MB 231 cell line from ANOVA test

Genes	p-values
CHRNA5_v2	0.0783
CHRNA5_v3	0.0509
CHRNA5_iso2	0.0378
CHRNA5_iso3	0.2702
PS2	0.0082
Average of CHRNA5 isoforms	0.3691

When the average of CHRNA5 isoforms was taken ANOVA was not significant for ER- MDA MB 231 (Table 4.4). This indicates that there could be isoform specific differences in response of CHRNA5 to E2 in ER- cell line MDA-MB-231 but not in ER+ MCF7 cells.

Multiple comparisons between each group were performed and groups giving significant results were listed in Table 4.5. PS2 shows statistically significant expression change between high doses of E2 treatment.

Table 4.5: List of group having statistically significant expression difference for each gene

Multcompare type	Genes	Groups with significant expression difference
With correction (hsd)	CHRNA5_iso2	Control-10nm,Control-100nm
	PS2	Control-100nm, 10nm-100nm
Without correction (lsd)	CHRNA5_v2	Control-1nm, 1nm-100nm
	CHRNA5_v3	Control-1nm, Control-10nm
	CHRNA5_iso2	Control-1nm, Control-10nm, Control-100nm
	PS2	Control-1nm, Control-100nm, 1nm-100nm, 10nm-100nm

4.1.3 Heat Map of CHRNA5 Isoform and E2 Markers Expression

Heatmap graphs provide opportunity to visualize a coordinated expression data set from multiple genes at once. Both expression of each cell line and the difference under treatment of each dose can be observed in a single graph.

Also reliability of experiment data can be clearly interpreted because of clustering of data. Heatmap showed that treatment group clustered together in a cell line-specific manner (Figure 4.4).

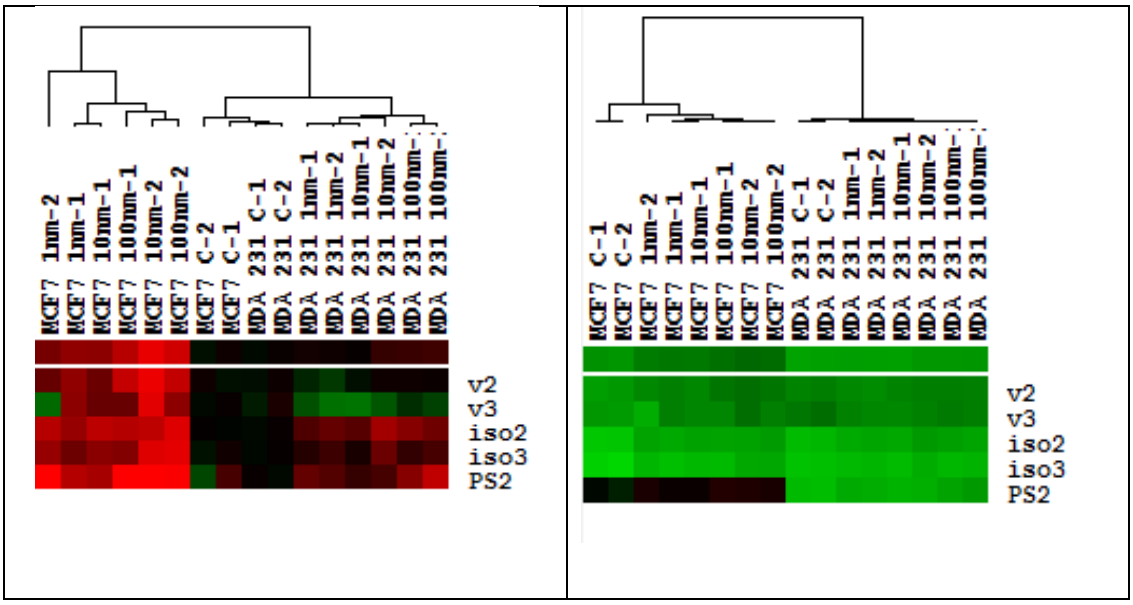


Figure 4.4: Heatmap representations of expression level of CHRNA5 isoforms and PS2 in MCF7 and MDA MB 231 cell lines. Red indicates an increase in expression and green indicates a decline. Euclidian distance was used as distance method and average linkage was taken for clustering. Map on the right side was obtained from ΔΔCt analysis and the one at the left side was from ΔCt analysis.

Heatmap graphs clearly showed that expression responses of MCF7 and MDA MB 231 to E2 treatment were different as they clustered in separate groups. Also in each cluster E2 doses were almost ordered, showing that genes were affected by changes in E2 doses. Moreover, genes in MCF7 cell line were highly upregulated with E2 treatment when compared with controls, while MDA MB 231 cells both have up- and downregulated genes but the magnitude of change was relatively less when compared with MCF7 (Figure 4.4, left side). On the other hand ΔCt analysis derived heat map (Figure 4.4, right side) showed that PS2 could strictly distinguish ER- and ER+ cell lines.

For further analysis of E2 treatment data, regression and/or ANCOVA (testing slope difference for both cell lines at the same time) can be used to predict how CHRNA5 expression changes depending on different doses of E2. Moveover, other statistical tests appropriate for analysis of gene expression datasets with small sample sizes can be applied for testing differences between groups within and between cell lines, e.g., limma [73].

4.2 Effects of 4-OHT and E2 Treatment

4-OHT (generally known as tamoxifen) is an ER antagonist which prevents E2 from binding to ER. Considering E2 and ER pathway trigger proliferation, 4-OHT should decrease the effect of E2 by blocking ER pathway, in particular for tamoxifen sensitive cells. The effects of E2 and 4-OHT treatment (both individually and together) were examined on CHRNA5 isoforms expression as well as selected E2 and 4-OHT marker genes for making comparisons.

4.2.1 Statistical Analysis of 4-OHT and E2 Treatment with MATLAB

ANOVA was used to test significance between the expression differences of genes (CHRNA5 isoforms as well as PS/TFF1 and FOSL2) under E2, 4-OHT, E2+4-OHT treatment. PS2 gene expression has been a well known marker of E2 response and thus was used to assess the effectiveness of E2 [74]. Our preliminary analysis of the GEO dataset (GSE26459) also has led to identification of FOSL2 as a potential responder to 4-OHT and E2 application (Appendix; Figure 8.1 and 8.6). In tamoxifen sensitive cell lines it is expected that FOSL2 expression level might increase after E2+4-OHT treatment up to two-fold [75] (Figure 8.1). In case of resistance to 4-OHT, however, both E2 or E2+4-OHT treatment can increase FOSL2 expression up to three fold while no change from control is expected for 4-OHT alone (Figure 8.6).

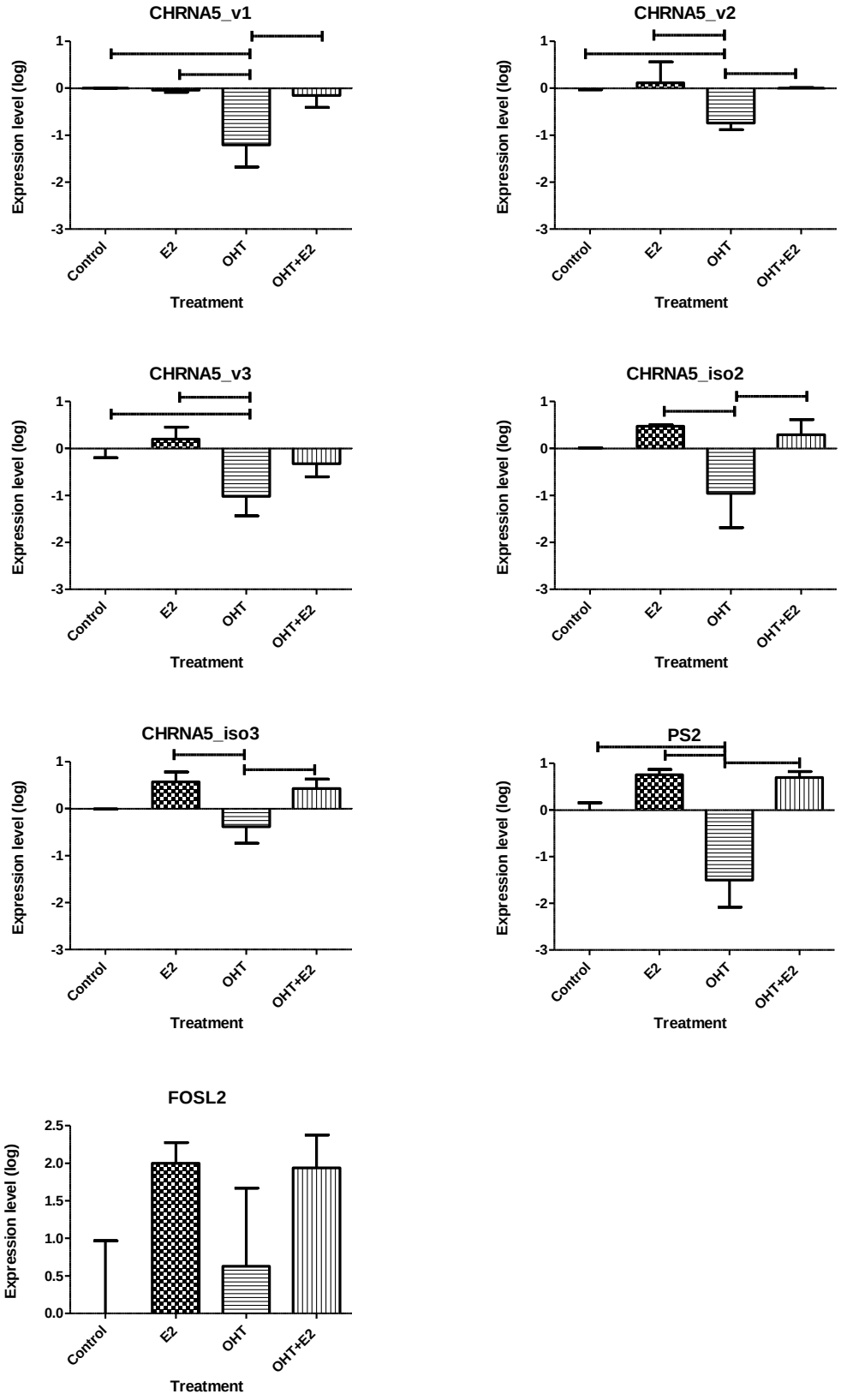


Figure 4.5: Effects of E2 and 4-OHT on CHRNA5 isoforms, E2 and 4-OHT indicators in MCF7 cell line

The results demonstrated that 4-OHT had significant effects on CHRNA5 isoform expression as well as PS2 (Figure 4.5; Table 4.6). Multiple comparisons indicated that the difference between E2 and tamoxifen was significant for CHRNA5_v1, v3, iso3 and PS2 genes (Table 4.6). For PS2 all groups were significantly different from each other while for v1, control and tamoxifen showed difference together with E2 and tamoxifen. For the other two genes only E2 and tamoxifen showed difference. If E2 has increased CHRNA5 expression and tamoxifen decreased it we would expect to see that when tamoxifen given together with E2 also should decrease (when compared with E2) the expression of CHRNA5, or any other E2 responsive gene. If we look at PS2, a known E2 responsive gene, the effect of E2 (compared with control) caused upregulation with and without 4-OHT. FOSL2 shows a similar expression pattern with the previous research however the changes were not significant [75]. This could be because the dose of E2 (10nM) overrides the effect of 1uM tamoxifen, i.e., the neutralization of E2 effect by tamoxifen could be dose-dependent. Another possibility could be that sensitivity of MCF7 to tamoxifen might vary from passage to passage or batch to batch in MCF7 cells. Interestingly, CHRNA5_v1 and CHRNA5_v2 expression drastically decreased with 4-OHT treatment however E2 did not seem to be effective for these genes in this experiment. However, CHRNA5_iso2 and CHRNA5_iso3 expressions changed as expected. Although CHRNA5_v3 shows an increase in expression after E2 treatment, this was not enough to revert its expression after co-treatment with 4-OHT (Table 4.6).

These results suggest that for most of the genes we studied at least one group's mean expression was different from one or more. To see which groups were significantly different from each other we performed multiple test comparisons using *multcompare* (Table 4.7).

Table 4.6: p-values of each gene comparing control, E2, 4-OHT, and E2 and 4-OHT together in MCF7 cell line from ANOVA test

Genes	p-values
PS2	0.0058
CHRNA5_v1	0.0291
CHRNA5_v2	0.0633

CHRNA5_v3	0.0497
CHRNA5_iso3	0.0438
CHRNA5_iso2	0.075
FOSL2	0.1321

Table 4.7: List of group having statistically significant expression difference for each gene

Genes	Groups with significant expression difference with correction	Groups with significant expression difference without correction
PS2	Cont-4-OHT,E2-4-OHT, 4-OHT-E2+4-OHT	Cont-4-OHT,E2-4-OHT, 4-OHT-E2+4-OHT
CHRNA5_v1	Cont-4-OHT,E2-4-OHT	Cont-4-OHT,E2-4-OHT, 4-OHT-E2+4-OHT
CHRNA5_v2	-	Cont-4-OHT,E2-4-OHT, 4-OHT-E2+4-OHT
CHRNA5_v3	E2-4-OHT	Cont-4-OHT,E2-4-OHT
CHRNA5_iso2	-	E2-4-OHT, 4-OHT-E2+4-OHT
CHRNA5_iso3	E2-4-OHT	E2-4-OHT, 4-OHT-E2+4-OHT

4.2.2 Comparison of CHRNA5 Isoform and E2 and 4-OHT Markers Expression

A comparative histogram showing all isoforms also was drawn below (Figure 4.6). It was seen that CHRNA5_v3 was relatively sensitive to tamoxifen and when given together with E2 it reduced the expression level.

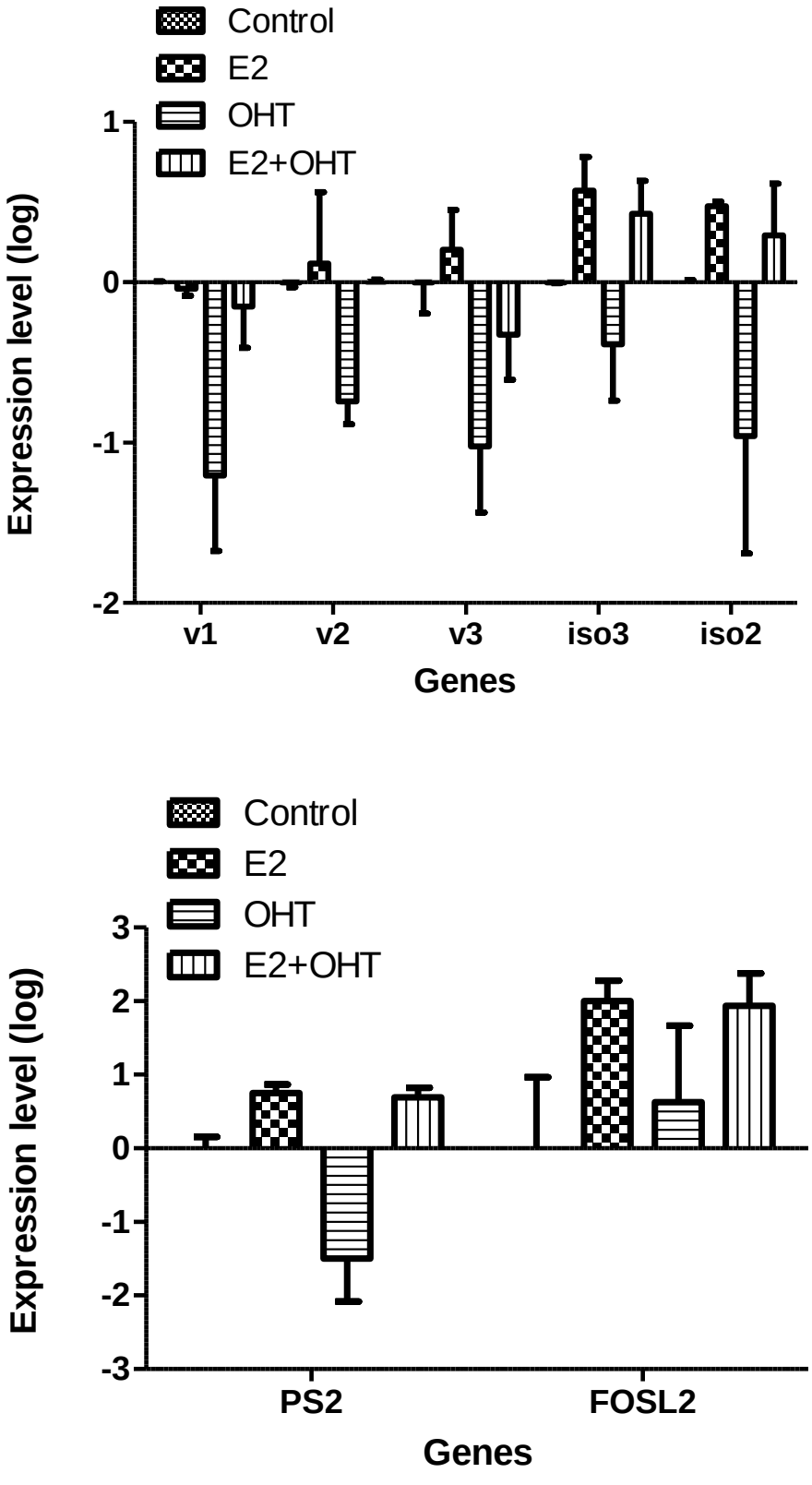


Figure 4.6: Overall representation of expression level of CHRNA5 isoforms and E2, 4-OHT indicators.

4.3 Effects of Long-term E2 Depletion (LTED) on CHRNA5 Isoform Expression

To address the mechanism behind the relationship between E2 depletion - CHRNA5 in breast cancer we generated an isogenic cell line of MCF7 that has gone through long-term E2 depletion. It is thought that the response of breast cancer line may differ after a short period of E2 depletion than long term of E2 depletion in gene expression level [36].

During a period of six months, E2 depleted and normal (where E2 was manually added into growth medium) MCF7 cell lines were cultured. The time points belong to each samples can be seen in table 8.7. Expression of CHRNA5 isoforms and estrogen signaling pathway markers (PS2 and ESR1) were studied by qPCR. Results were analyzed using paired tests across this time series qPCR data in MATLAB.

4.3.1 Determining the Housekeeping Gene for LTED Experiment

It is important to find a reliable housekeeping gene not affected by the lack of E2 in the cell line experiments. In this study, TPT1 gene was chosen as housekeeping gene and its expression was measured over six months.

To determine whether TPT1 gene is a consistent and reliable housekeeping gene difference between control and treatment (depletion) group per time was plotted (Figure 4.7). Also sign test result showed that there was no significant positive or negative collection of values in TPT1 expression throughout six months (Table 4.8).

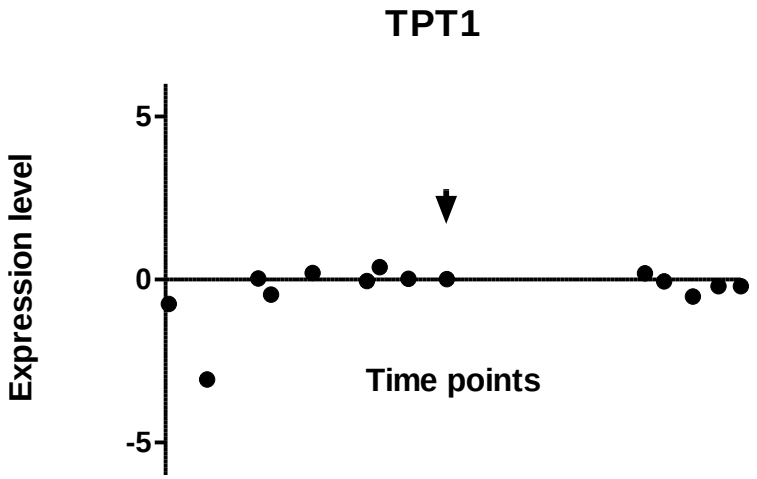


Figure 4.7: Changes in gene expression of treatment with respect to control of TPT1 expression per time point during six months of E2 depletion with ΔΔCt analysis. Arrow indicates the end of third month.

Except one time point the expression of TPT1 gene did not differ between the control (with E2) and treatment (without E2) group during six months. This plot indicated that TPT1 could be a reliable housekeeping gene that can be used in further analysis.

4.3.2 Expression of CHRNA5 Isoform and E2 Exposure Markers across LTED

During the first three months of experiment, cells were cultured and split without counting and without considering time (hr) of the pellet collection. This might cause a fluctuation in gene expression especially if the data is analyzed in an unpaired fashion. Thus we have analyzed the expression difference between treatment (T) and control (C) groups per time point when normalized with respective TPT1 values (ΔΔCt) (Figure 4.8-4.9).

During 6^{th} month of exposure, to increase consistency and reliability of data across time points all cells were seeded in a predetermined number after counting, and also they were cultured at the same time of the day in order not to interfere with the cell cycle. It is important to note that at later time points the difference between control and treatment groups stabilized compared to earlier timer points also in response to the effects of LTED.

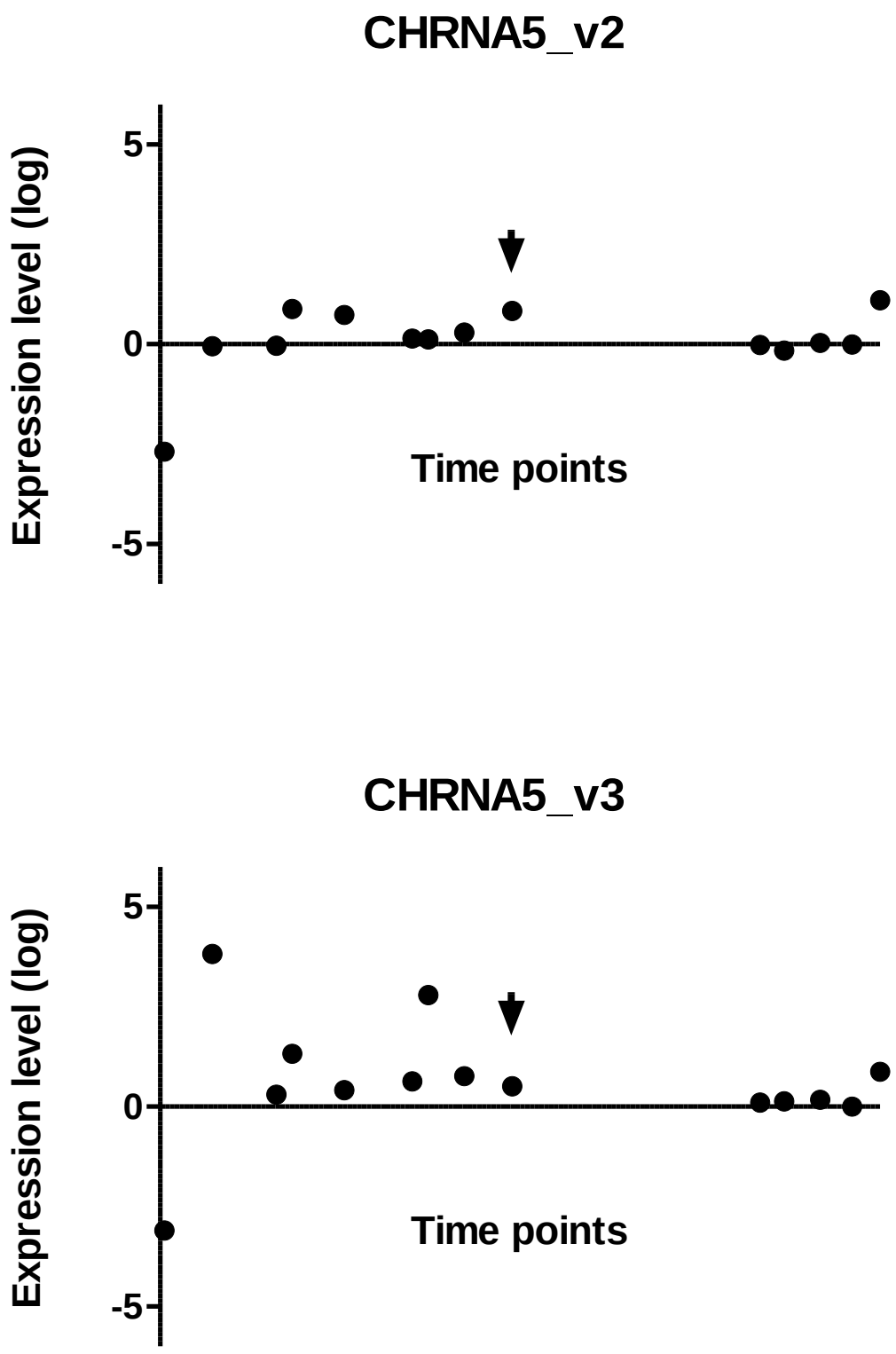


Figure 4.8: Changes in gene expression of treatment with respect to control of CHRNA5 isoforms expression per time point during six months of E2 depletion with ΔΔCt analysis. Arrow indicates the end of third month.

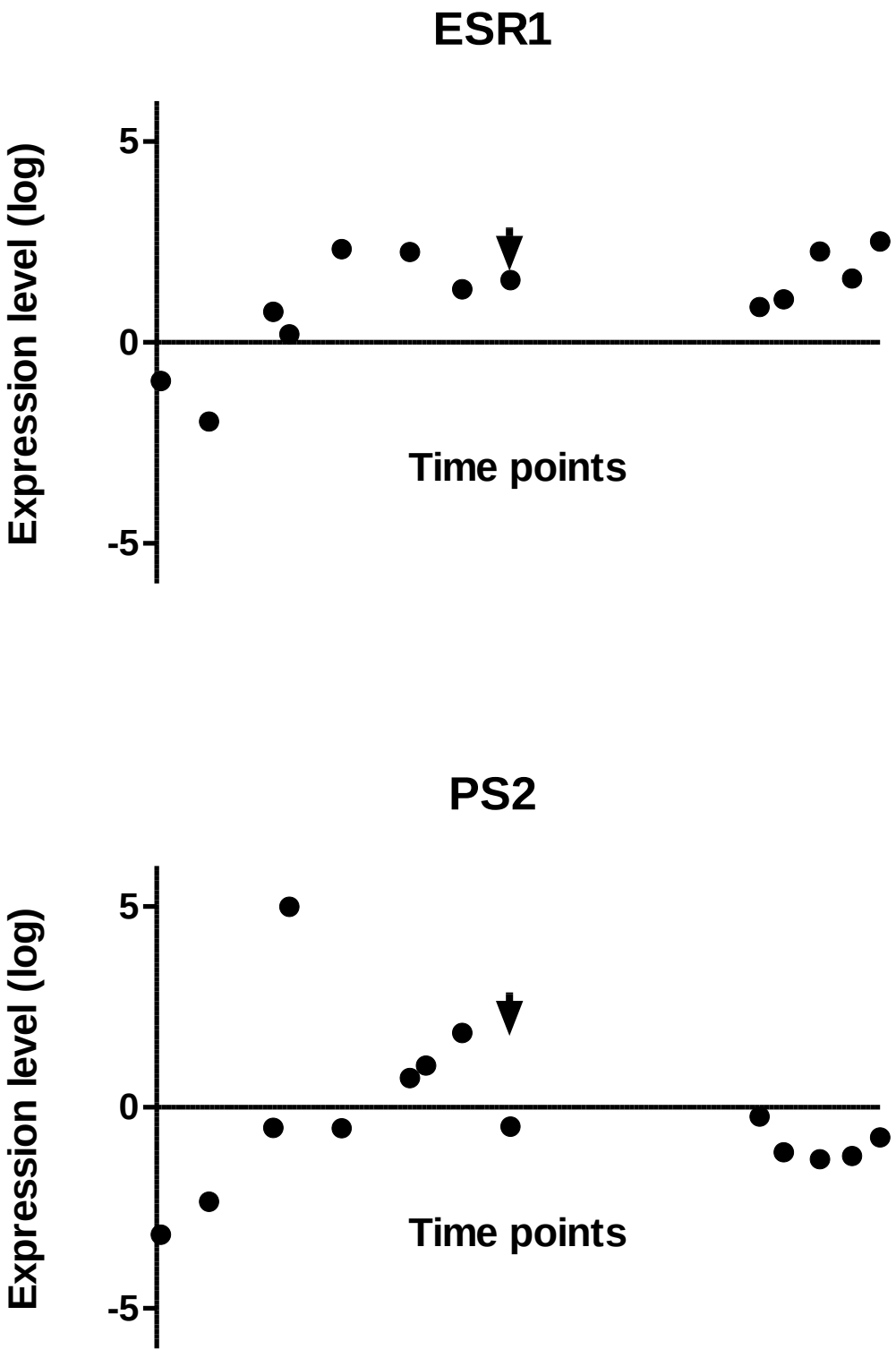


Figure 4.9: Changes in gene expression of treatment with respect to control of E2 indicators expression per time point during six months of E2 depletion with ΔΔCt analysis. Arrow indicates the end of third month.

The most striking result we obtained from LTED experiment was that ESR1 expression was significantly different after fifth time point onwards corresponding to the middle of second month between C (control) and T (E2 depleted) groups. ESR1 expression in second month (4-6. points) comparing to first month (1-3. points) highly increased by LTED, and despite the fluctuation, CHRNA5_v2 and CHRNA5_v3 also showed upregulation throughout these three months. PS2 did not follow any trend; its expression without E2 first was smaller than control then became larger than the control. After time point 10 (beginning of sixth month) the PS2 expression without E2 became reduced in amount when compared with the control (Figure 4.9).

Last five samples covering sixth months had less variation within groups than first three months. Despite PS2 showed a decreasing trend towards the end of the sixth month, the expression level of PS2 could still be considered high in Treatment versus Control when comparing with the samples belonging to initiation of the study (first weeks). CHRNA5_v2 and in particular CHRNA_v3 did not show much fluctuation yet they had an increasing trend. In GEODataSet results for PS2, the expression decreases first month then it increases at the third month and remains at the same level at sixth month (Figure 8.4).

Table 4.8 shows the result of sign test applied for each gene to see whether there is a positive or negative trend between the differences in expression of control and treatment group in LTED.

Table 4.8: p-values of each gene comparing expression after six months of estrogen depletion from sign test

	CHRNA5_v2	**CHRNA5_v3**	**ESR1**	**PS2**	**TPT1**
p-value	0.7905	0.0018	0.0129	0.1796	0.7905

In this analysis, p-value showed how the samples were different from their median or whether they increased or decreased consistently. Table 4.10 indicates that PS2 and CHRNA5_v2 do not regularly decrease or increase; however ESR1 and CHRNA5_v3 showed a significant difference between LTED and its control across

time points. We generated a promising model to test the effects of LTED on isoform specific gene expression including CHRNA5 isoforms.

4.4 Expression of CHRNA5 Isoforms in cDNA Panel

In order to study how different isoforms of CHRNA5 gene are expressed in different subtypes of breast cancer a cDNA panel from ORIGENE was examined. CHRNA5 expression difference between normal and tumor breast tissues as well as differences among tumor samples with different receptor status were compared in terms of CHRNA5 isoform expression upon normalization with TPT1 as the reference gene.

ANOVA test was mainly used because of large sample size. To study the correlation between isoforms and their ability to separate based on receptor status and stage Pearson correlation coefficients and scatter plots were used. Efficiency corrected ΔCt method was used for analyses. Multiple cDNA panels, each with two technical replicates for each well were studied with 1/20 dilution with a negative control; and only non-contaminated runs for each primer were used. Replicates were averaged for each sample before statistical analyses.

4.4.1 Data Quality Control Studies

The data coming as a result of qPCR procedure may be problematic in several ways especially for genes expressed at low levels. First of all, a well can be read as empty while its duplicate gives a reliable Ct value; second, there may be relatively large Ct differences between duplicates; and third, Tm values may vary among duplicates and/or samples tests. Pipetting errors, technical problems during sample preparation as well as performing the reaction in a qPCR machine can also contribute to variability obtained between duplicate measurements or gathering of the qPCR data from studies that use large plates (96 well or 384 well). Therefore it is important to assess the quality of the qPCR data in terms of different quality checks for duplicate variability as well as Tm and amplicon size differences.

4.4.1.1 Tm Consistency Check in cDNA Panels

To be sure of the accuracy of Ct values we checked the consistency of Tm values for a primer pair in a run. A reliable primer should work with similar and relatively close Tm's under similar conditions when the experiment is repeated. However, we don't know how much variation in the Tm values should be tolerated If Tm values of some wells in qPCR plate are unexpectedly deviant than others, the Ct values of these wells may not belong to tested primer, or double Tm measurements may refer to dimer formation or other artifactual hybridization between strands.

In order to label a Tm value as deviant or marginal a threshold should be determined. The threshold was determined by examining qPCR products with different Tm values. Median of all Tm values in the cDNA panel was taken, then samples less or more than one degree Centigrade different from this median were loaded into agarose gel together with positive controls. In Figure 4.10, a representation of this study can be seen.

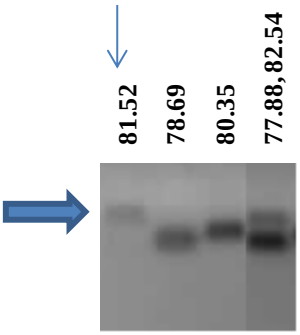


Figure 4.10: Gel image of qPCR products from different experiment with varying Tm values. Thin arrow indicates qPCR product having closer Tm with the median, it was loaded as control. Normal Tm expectancy for this primer was 81.11 °C. Bold arrow indicates expected amplicon size, which is 93bp.

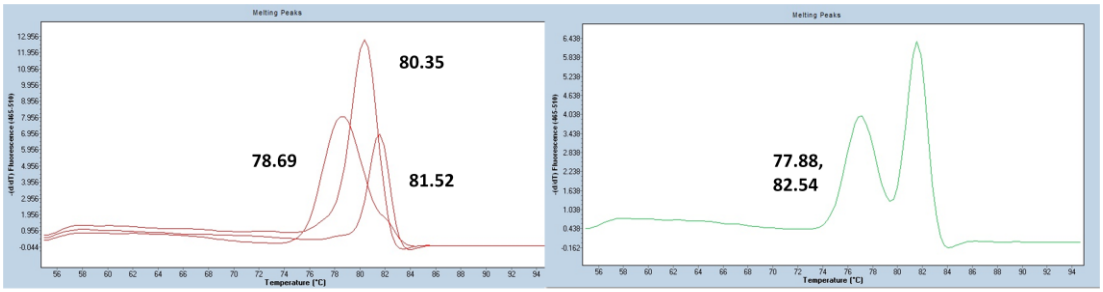


Figure 4.11: Representative melting curve. Tm curves having different Tm than the median differ from each other.

After examining gel images, we estimated that qPCR products with Tm values ± 1degree were likely to be different than that of positive control and did not give bands at the same size with the control. These were generally dimers, sometimes heteroduplexes, artifacts, or smear.

As a result of this strategy, it was decided to remove samples with Tm values ± different than that of control and samples with double Tm values before any further analysis.

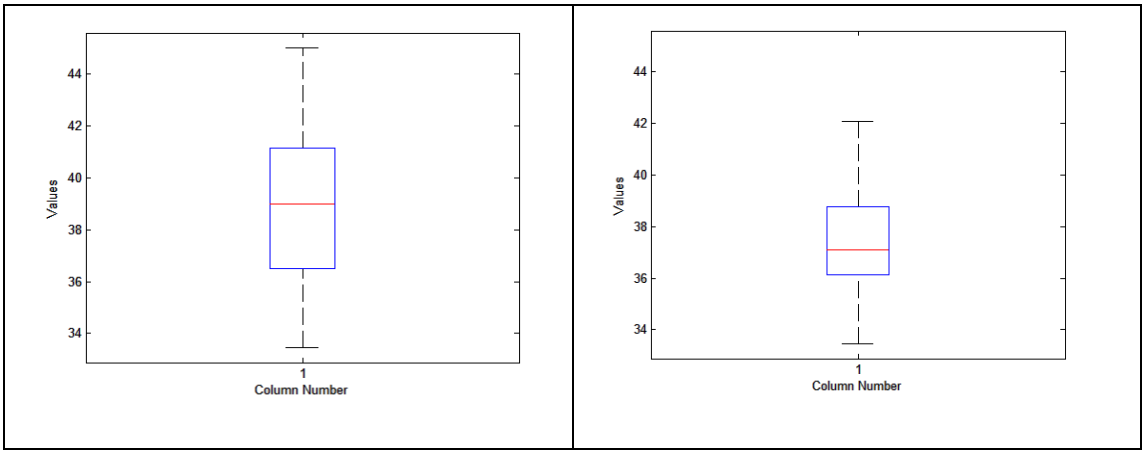


Figure 4.12: Representation of Ct value distribution before and after Tm outlier removal. First figure belongs to the data before Tm checks and exclusion and second one is boxplot of the same data after exclusion of wells with inconsistent Tms. The variation considerably decreased after Tm outlier removal.

4.4.2 Statistical Analysis of CHRNA5 Isoforms Expression

The level of changes in the CHRNA5 isoform level in breast tumor samples compared to normal breast tissue as well as to the receptor status and CHRNA5 isoforms expression (v1-v3) were studied.

One-way ANOVA was performed for each gene after normalizing their Ct value with housekeeping TPT1 gene with efficiency correction. p-values were obtained to compare normal-tumor, ER+ - ER-, HER2+ - HER2- and stages of tumor. Table 4.9 shows p-values derived from ANOVA test.

Table 4.9: p-values of each CHRNA5 isoform normal-tumor, ER+ - ER-, HER2+ - HER2- and stages of tumor in cDNA panel

p-values	Normal-Tumor	Normal-ER+, ER-	Normal-HER2+, HER2-	Normal-Stages
CHRNA5_v1	0.0076	0.0072	0.0174	0.014
CHRNA5_v2	0.0031	0.0097	0.0162	0.0013
CHRNA5_v3	0.0249	0.0003	0.083	0.0109
Ave. of three isoforms	0.0055	0.0022	0.0188	0.0035

Table 4.10 represents the result of multiple comparison test performed to see p-values are significant between exactly which groups.

Table 4.10: List of group having statistically significant expression difference for each gene

Multcompare	Normal-ER+, ER-	Normal-HER2+, HER2-	Normal-Stages
CHRNA5_v1	N-ER-, N-ER+	N-HER2+	II-N
CHRNA5_v2	N-ER-	N-HER2-, N-HER2+	II-N, III-N
CHRNA5_v3	ER+ - ER-, N-ER-	-	II-N
Ave. of three isoforms	N-ER-, N-ER+	N-HER2-, N-HER2+	II-N

According to results of ANOVA, expression of each isoform of CHRNA5 significantly differed between normal and breast tumor tissue, a finding supporting previous studies on CHRNA5 isoforms in lung cancer [44]. CHRNA5_v2 showed the most drastic change with a p-value equal to 0.0031 (Table 4.9).

After interpreting p-values and multiple comparison results, it seems that even CHRNA5_v1 expression significantly differs between tumor and normal breast tissue although this isoform was not enough to distinguish subtypes of breast cancer; i.e.., ER+ and ER- samples did not show any remarkable difference according to

multiple comparison results. However, stage II seems to be distinguished from Normals when analyzing CHRNA5_v1 expression (Table 4.10).

CHRNA5_v2 followed a similar trend with CHRNA5_v1 except its expression more remarkably differs according the p-values. Also, CHRNA5_v2 provides opportunity to differentiate stage III as well as stage II from normal tissue.

CHRNA5_v3 was not effective to distinguish HER2 status however more importantly its expression can be used to separate ER+ and ER- breast tumor tissues. Moreover, just like other isoforms its expression significantly differs between stage II and normal tissue.

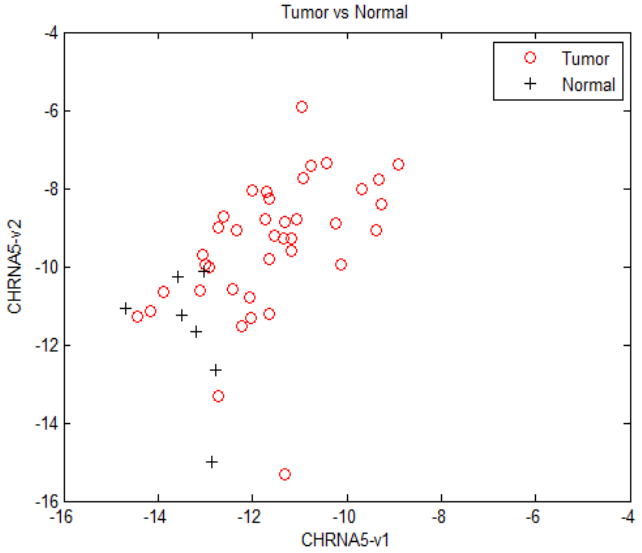
Overall, ANOVA applied to the average of all isoform gave a generic representation of CHRNA5 expression in cDNA panel, while stage II from Normals can be distinguished by this test (Table 4.9, Table 4.10).

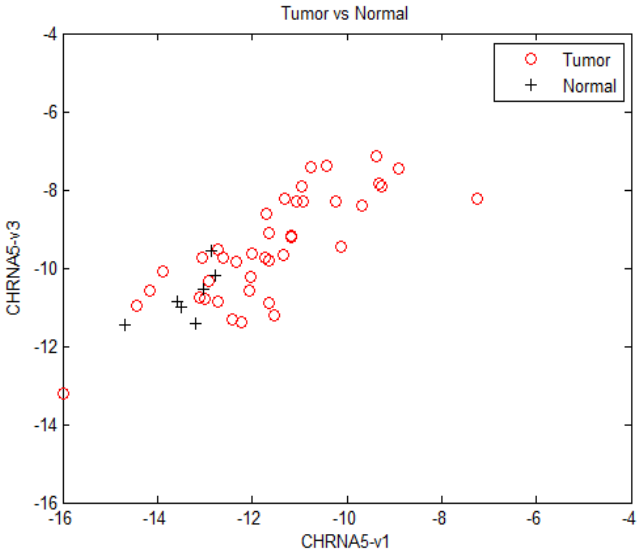
4.4.3 Comparison of CHRNA5 Isoform Expression

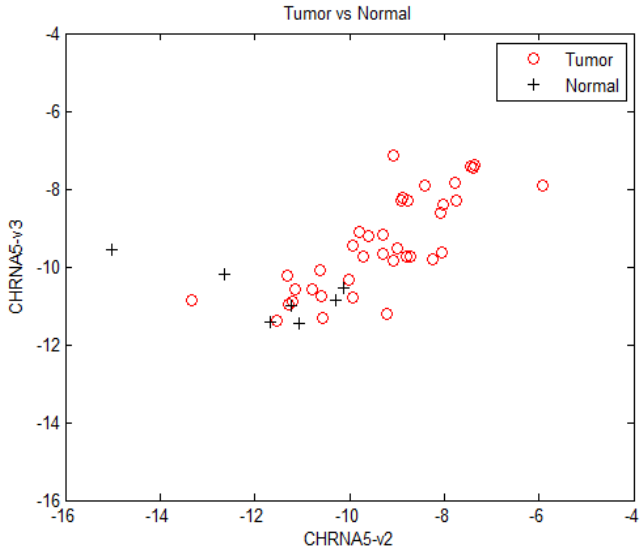
After one-by-one analysis of CHRNA5 isoforms it was seen that subtyping only considering one isoform of the gene may not be as powerful as using multiple isoforms. Therefore, correlation between isoforms and how well they can distinguish receptor status should be analyzed.

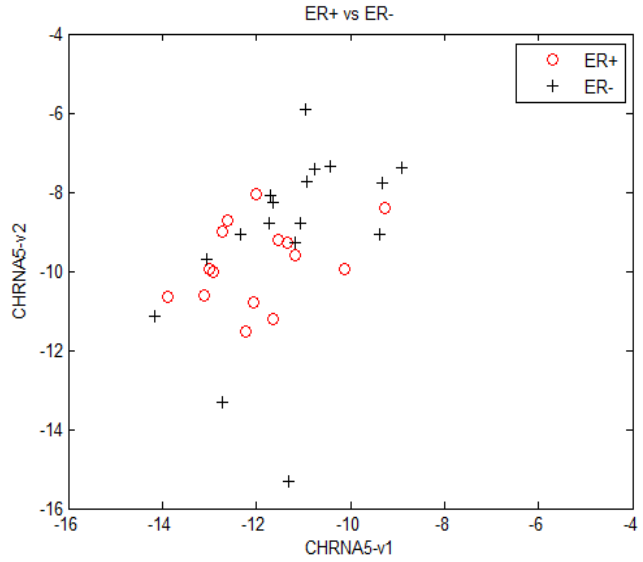
Expression values of each isoform (CHRNA5_v1, v2, v3) for cancer samples were plotted against each other after normalizing their Ct values with TPT 1 (ΔCt); each group in each plot was differently labeled to see whether clustering between groups can be obtained or not (Figure 4.13).

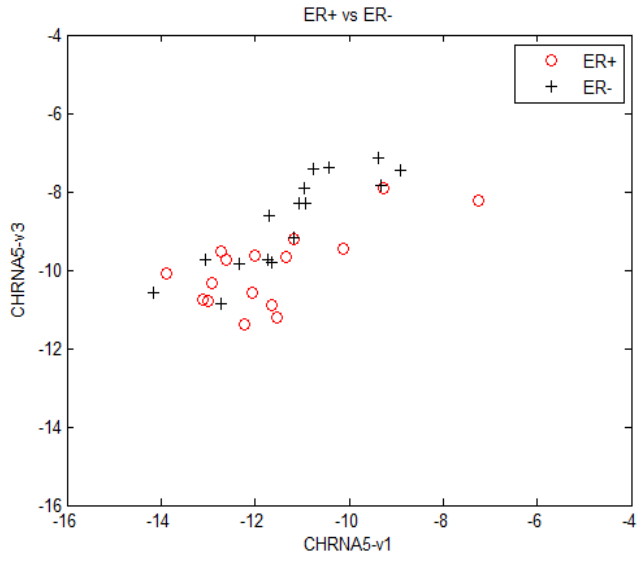
Figure 4.13 represents correlation plots of all three isoform considering normal vs tumor, ER+ vs ER-, HER2+ vs HER2- and three stages of breast cancer.

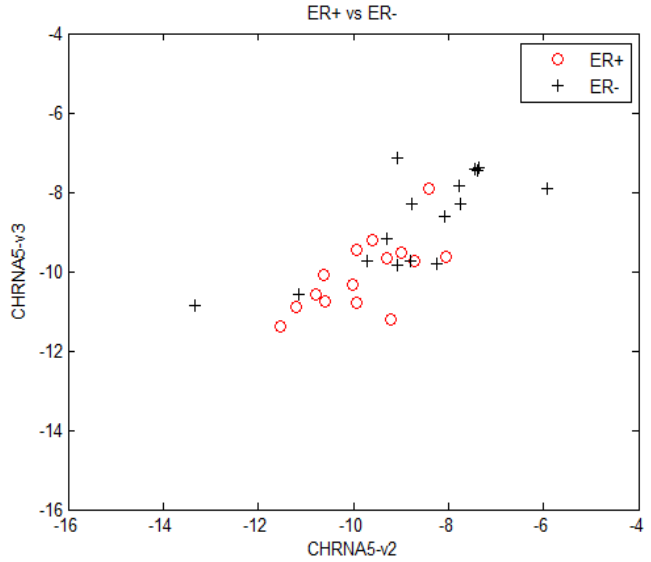

Tumor vs Normal
Tumor
Normal
CHRNA5-v2
CHRNA5-v1
-4
-6
-8
-10
-12
-14
-16
-16
-14
-12
-10
-8
-6
-4

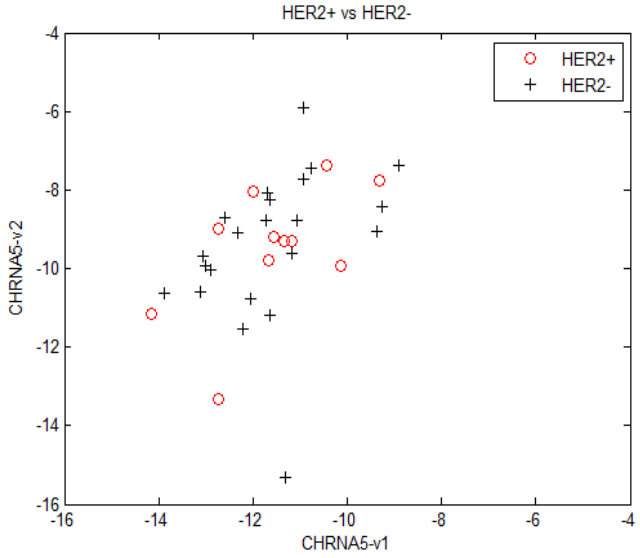

Tumor vs Normal
Tumor
Normal
CHRNA5-v3
CHRNA5-v1
-4
-6
-8
-10
-12
-14
-16
-16
-14
-12
-10
-8
-6
-4

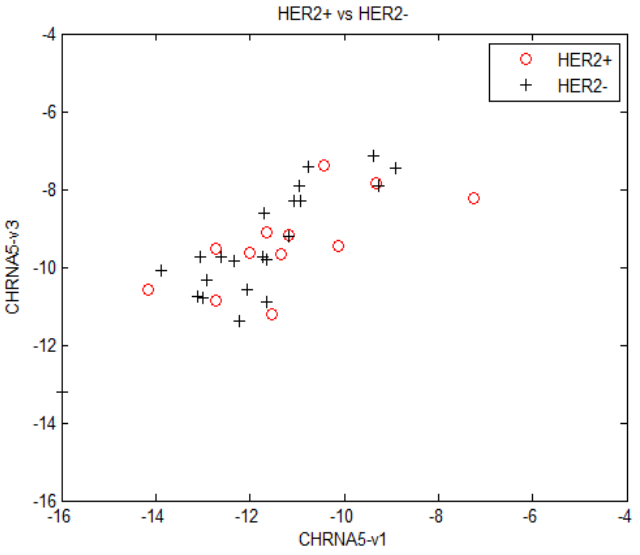

Tumor vs Normal
Tumor
Normal
CHRNA5-v3
CHRNA5-v2
-4
-6
-8
-10
-12
-14
-16
-16
-14
-12
-10
-8
-6
-4

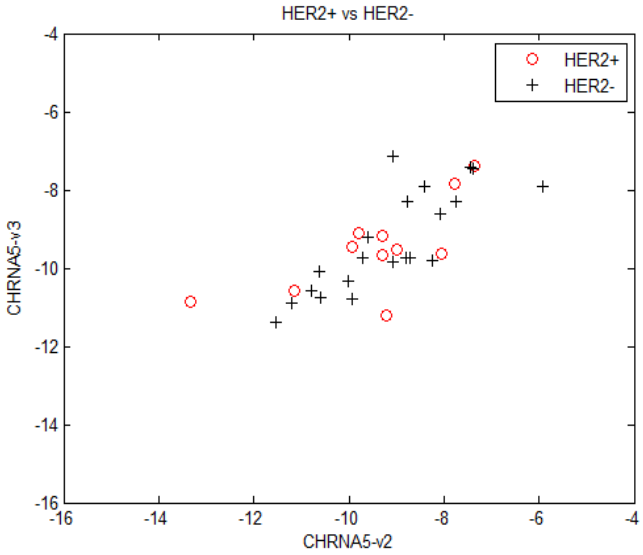

ER+ vs ER-
ER+
ER-
CHRNA5-v2
CHRNA5-v1
-4
-6
-8
-10
-12
-14
-16

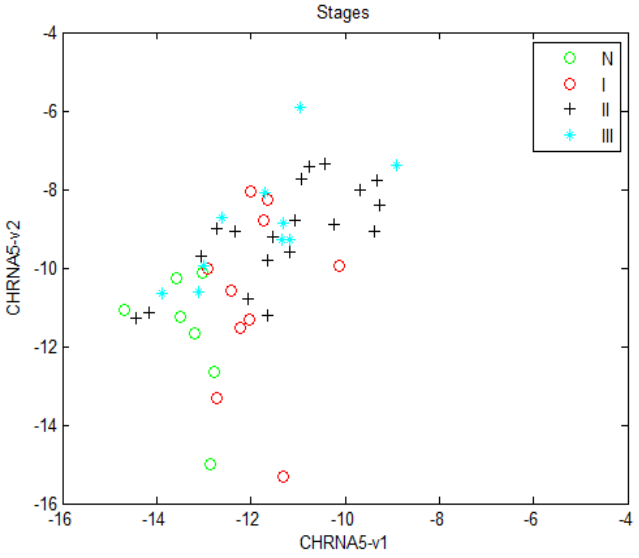

ER+ vs ER-
ER+
ER-
CHRNA5-v3
CHRNA5-v1
-4
-6
-8
-10
-12
-14
-16

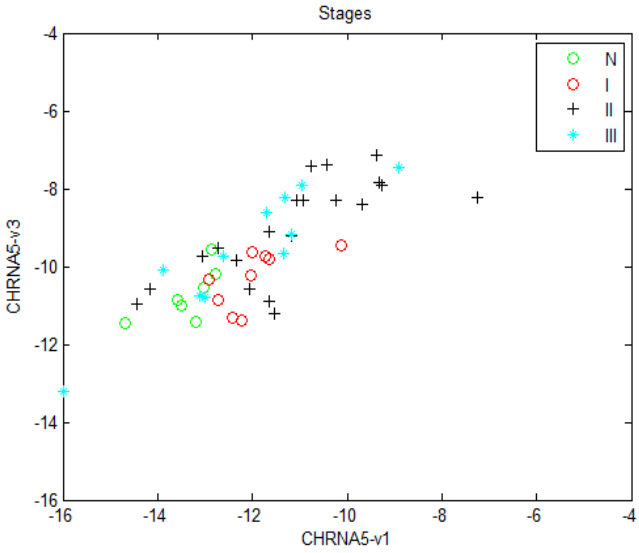

ER+ vs ER-
ER+
ER-
CHRNA5-v3
CHRNA5-v2
-4
-6
-8
-10
-12
-14
-16


HER2+ vs HER2-
HER2+
HER2-
CHRNA5-v2
CHRNA5-v1
-4
-6
-8
-10
-12
-14
-16


HER2+ vs HER2-
HER2+
HER2-
CHRNA5-v3
CHRNA5-v1
-4
-6
-8
-10
-12
-14
-16


HER2+ vs HER2-
HER2+
HER2-
CHRNA5-v3
CHRNA5-v2
-4
-6
-8
-10
-12
-14
-16


Stages
N
I
II
III
CHRNA5-v2
CHRNA5-v1
-4
-6
-8
-10
-12
-14
-16


Stages
N
I
II
III
CHRNA5-v3
CHRNA5-v1
-4
-6
-8
-10
-12
-14
-16

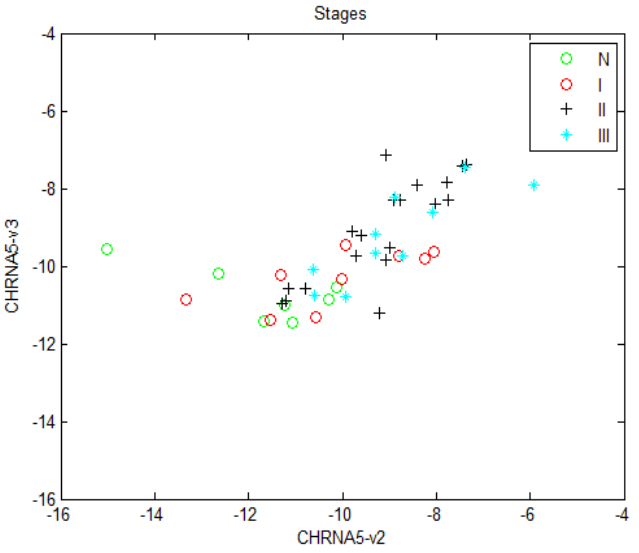

Stages
N
I
II
III
CHRNA5-v3
CHRNA5-v2
-4
-6
-8
-10
-12
-14
-16

Figure 4.13: Correlation plots of three CHRNA5 isoforms

For each correlation graph belonging to three isoforms, tumor and normal tissue, represented by red circle and black plus respectively, can easily be distinguished. Tumor samples have higher expression for each isoform than normal samples and they are almost always in different clusters. These graphs and findings also supports ANOVA test results represented in Table 4.9 and Table 4.10.

ER status could be differentiated by CHRNA5 isoform v3 (Table 4.10) and plotting two isoforms across each other provided a better clustering of ER+ and ER- samples (Figure 4.13).

HER2 status cannot be clearly differentiated by ANOVA or by correlation studies, CHRNA5 isoform expression does not seem differ significantly based on HER2 status (Figure 4.13, Table 4.9). Considering average expression of all isoforms might be useful (Table 4.10).

Breast cancer stages are important in diagnosis as well as for selecting treatment modality. For all genes, normal samples were different from one or more of the advancing stages; no two stages could be statistically significantly distinguished by CHRNA5 isoforms (Figure 4.13; Table 4.10). However, combining stage II with stage III as a single group can help distinguish Normal, Stage I, and Stage II+III.

Expression of isoforms was highly correlated to each other; correlation coefficients can be seen in Table 4.11. Overall, CHRNA5_v1 and CHRNA5_v3 were more correlated with each other than they were with CHRNA5_v2. At least one isoform of CHRNA5 was able to distinguish between ER- tumors from ER+ and ER- tumors had higher expressions in general. CHRNA5 has emerged an important tumor vs. normal marker in breast cancer (as it was shown in lung cancer) for all isoforms (v2 and v3; and v1 for multiple isoforms) tested here. Analysis of clinical parameters in combination might reveal further details about the isoform specificity and diagnostic potential for CHRNA5 (e.g., ER status together with stage or HER2 status).

Table 4.11: Correlation coefficiencies and p-values between primer pairs for all samples

	Gene pairs		
	v1-v2	v1-v3	v2-v3
Corr. coef	0.5317	0.8149	0.6926
Corr. p-value	0.0001	0.0000	0.0000

4.4.4 GUI Application for cDNA Panel Data Analysis

cDNA panels have large number of samples that are tested in replicates. Furthermore, multiple panels for the same set of samples can be studied for one or more genes. It is important to automatize the labeling, normalization, statistical analysis and graphical visualization steps in a cDNA panel qPCR analysis. However, there is not a useful application for this purpose in the literature thus we aimed to develop a GUI that will allow for input, quality control, selection of methods, and analysis and visualization of results in a modular manner. The application herein consists of four main parts; INPUT, QUALITY_CONTROL, METHOD, ANALYSIS.

A pre-module presents options to normalize each target gene with their own reference from the same cDNA panel dilution (paired) or analyze them simply by using the average of expression values for a gene from different runs (dilutions) (unpaired). In the present thesis, the analyses were performed using the unpaired option. INPUT is where the user gives the input data in a certain format (e.g., 131024_TPT1_cDNAPanel3.xlsx) and then the application allows user to enter datasets as individual files (Figure 4.14, Figure 4.15).

After taking input, quality control module is presented and the current application has an option for checking the consistency of Tm, if user wants to apply this option (Figure 4.16).

Methods module can be chosen as the third part of application. Efficiency of analyzed primer should be entered and either ΔΔCt or ΔCt methods can be chosen to analyze the expression data with primer amplification efficiency correction (Figure 4.17).

Last phase of the application is to choose analysis type. ANOVA, Kruskal Wallis with or without multiple comparison tests can be chosen. The source of variation can be decided based on the column values of the given covariate description file that corresponds to the labels of each well in the analysis (e.g., receptor status or stages). Moreover, correlation of genes entered as target genes can be plotted. All data are analyzed after clicking ANALYZE button and published (Figure 4.18).

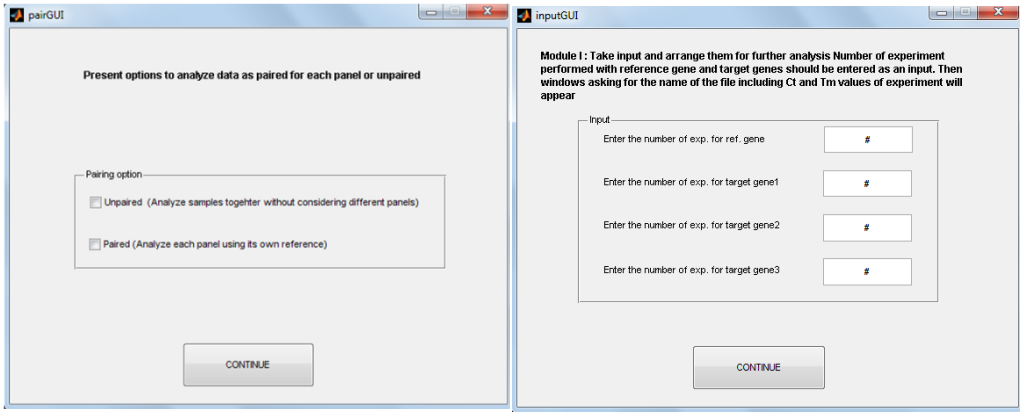


Figure 4.14: Interfaces of input module of cDNA panel analysis application

Input module allows user to enter input for both reference gene and target genes desired to be analyzed. After entering the repetition number of experiments for both reference gene and target genes windows corresponding to each experiment will appear; first one would be for reference gene (Figure 4.15); the window has empty areas to enter file names. After clicking "OK" other windows will appear to enter input file names of target genes or both target genes and their references.

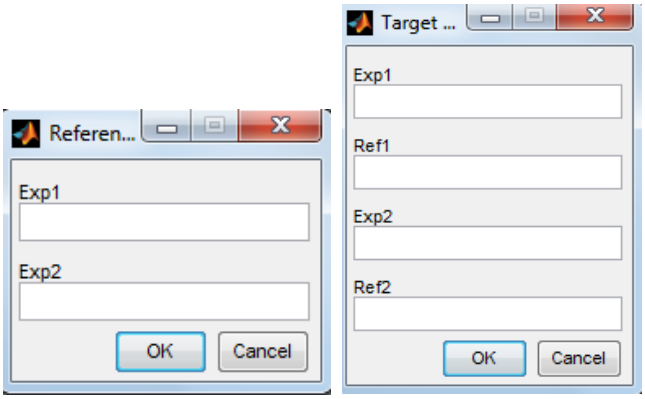


Figure 4.15: Interface to enter input file names

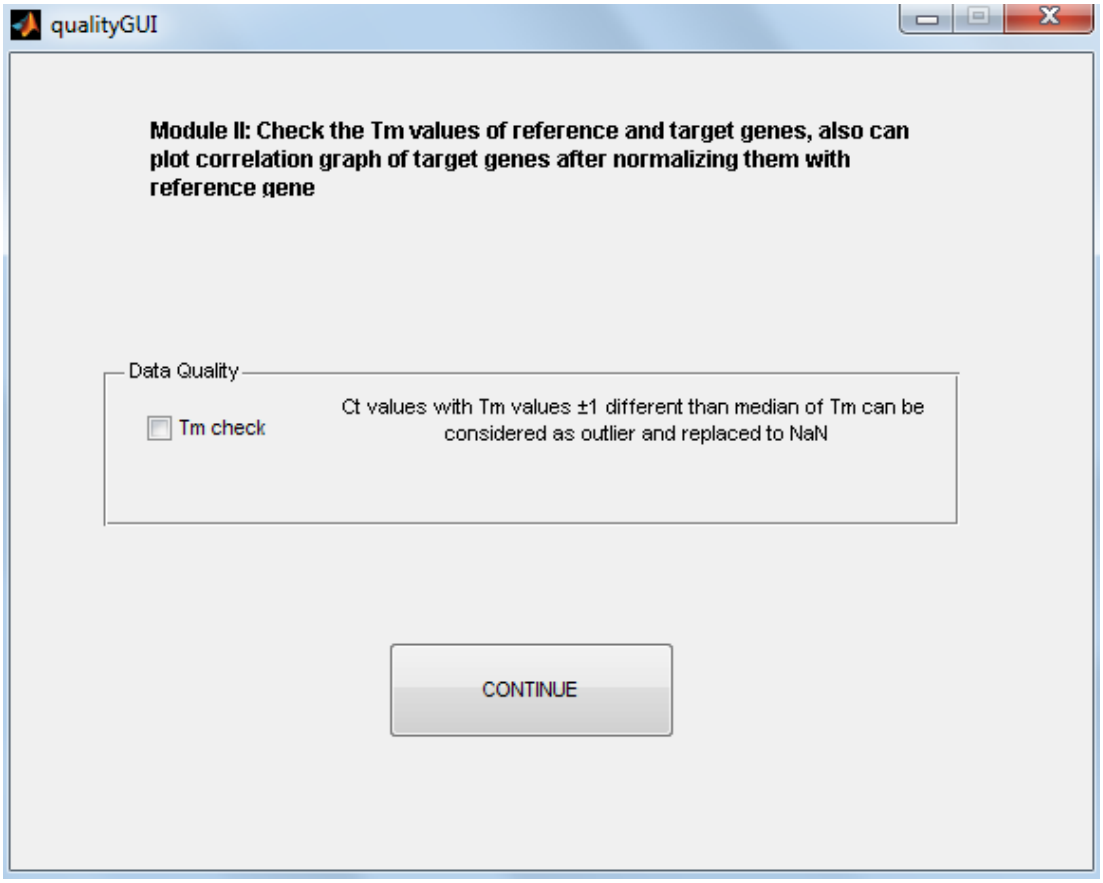


Figure 4.16: Interfaces of quality control module of cDNA panel analysis application

This second module allows user to control quality of the data. Tm values can be checked with the method defined above. After checking Tm of data and excluding Ct values with outlier Tm values, the Ct data showing excluded and included Ct values can be exported into an excel file created with the same file name with program automatically.

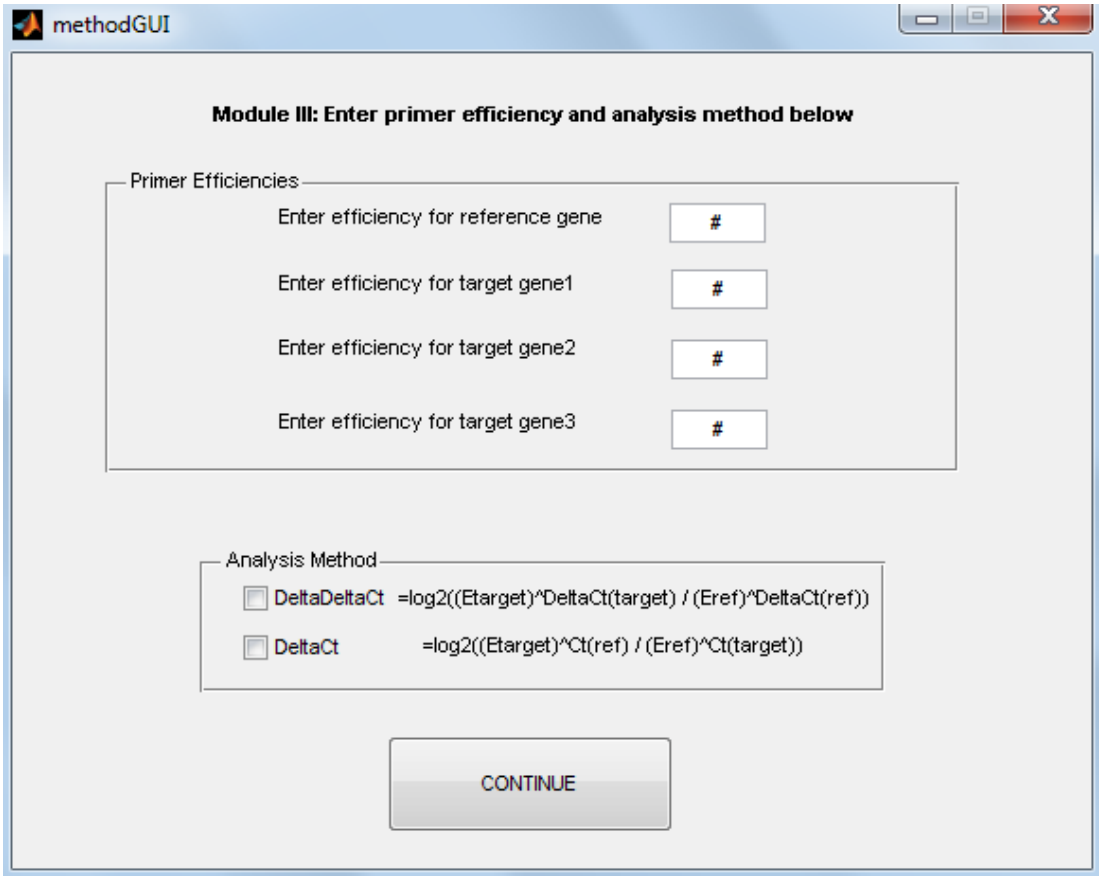


Figure 4.17: Interfaces of method option module of cDNA panel analysis application

Method module is to choose analysis method for raw data. After entering efficiency value for primer of target gene, module can process raw data using ΔΔCt or ΔCt methods.

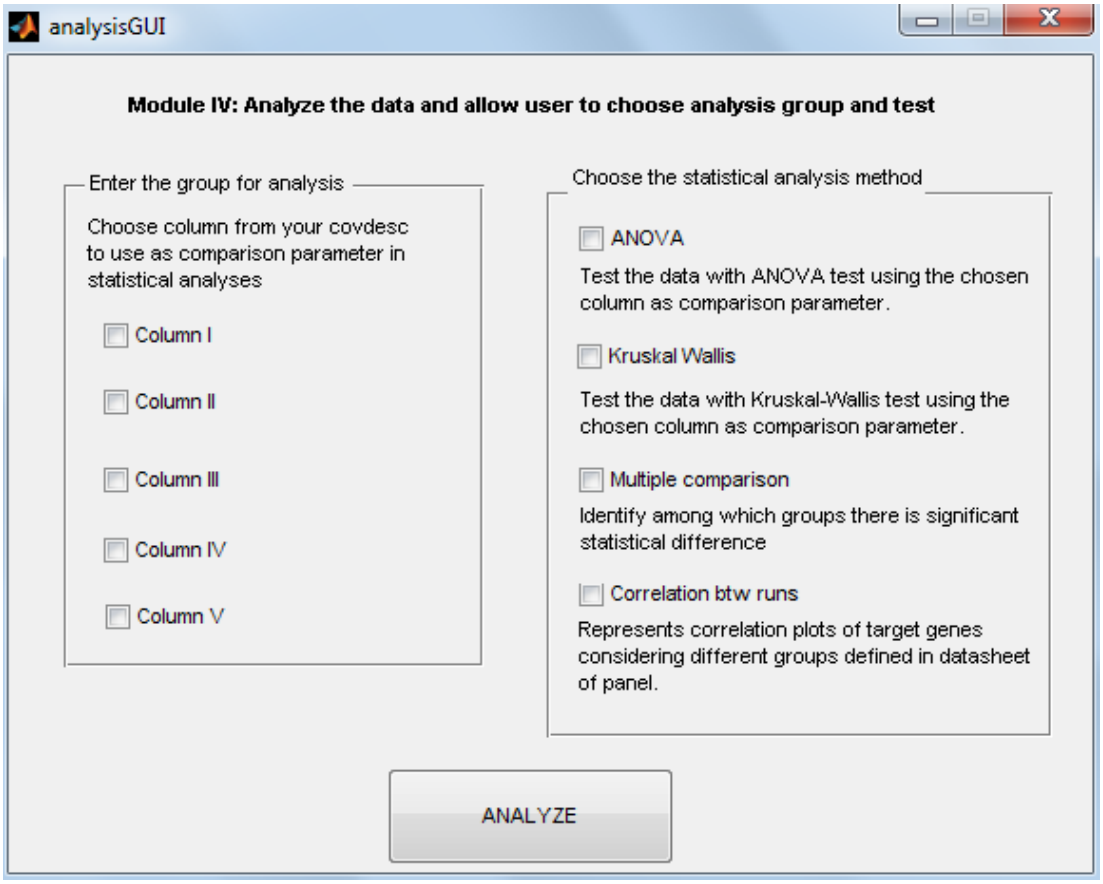


Figure 4.18: Interfaces of analysis module of cDNA panel analysis application

Last module finalizes the analysis of cDNA panel; it has options to choose groups for analysis based on the covdesc (covariate description file) to include groups defined by user. After determining which groups to compare and statistical analysis method, "ANALYZE" button gives boxplots and p-values for introduced data. This module will be have added functionality in the future for use of other ORIGENE panels in the GUI.

5. CONCLUSION AND DISCUSSION

Breast cancer is one of the leading reasons of cancer causing death and there are still uncovered mechanisms related with breast cancer initiation, growth and invasion. In this study, a target gene, CHRNA5, thought to be playing a role in breast cancer was tested as a potential biomarker for normal vs cancer and subtype diagnosis in breast cancers *in vivo*. Furthermore, its association with ER receptor status and Estrogen responsiveness were studied using *in vitro* breast cancer cell lines.

Alternative splicing was previously associated with a variety of diseases, including cancer. Splicing products, isoforms, of pre-mRNAs can lead to malignancy by the production of variants that are not normally present in a tissue type, or by changing the normal ratio between natural variants in a specific tissue [51,54,76].

CHRNA5 is a subtype of nAChR that has previously been associated with neurological diseases such as Alzheimer, Parkinson, epilepsy as well as smoking behavior and cancer [42]. CHRNA5 was also found to be alternatively spliced in cell lines, normal lung tissue and lung adenocarcinoma and the level of expression of CHRNA5 isoforms has been found as increased in lung cancer [44].

All of these findings have led our research to focus on alternative splicing patterns of CHRNA5 in defining breast cancer subtypes and in response to β-estradiol (E2), tamoxifen (4-OHT) treatment and long term estrogen depletion. Studies were done with two different breast cancer cell lines, MCF7 (Luminal A) and MDA MB 231(Basal-like) and an OriGene breast cancer cDNA panel [10].

Our first approach to observe the modulatory effects of E2 on CHRNA5 isoform expression and breast cancer was through applying different doses of E2 to MCF7 (ER+) and MDA MB 231 (ER-) cell lines. E2 is an agent playing a very important role in carcinogenesis of ER+ tumors; and it upregulates expression of certain genes, such as PS2/TFF1 [23,77]. However, dose-depending effects of E2 on CHRNA5 isoform expression in ER+ and ER- breast cancer cell lines have not been studied using qPCR in the literature yet. In this study, 1nm, 10nm, 100nm and 1000nm E2 were applied to the cancer cells that differ in their ER status for 24h after growing cells in E2 depleted medium and upon syncronization. Our results showed that the response

of cell lines to E2 treatment was completely different for ER+ and ER- cell lines while the response to different doses varies from gene to gene (Table 4.1). Then cell lines were analyzed separately for each CHRNA5 isoform and PS2 (TFF1). Expression of genes of interest in ER+ MCF7 cell line showed an upregulated trend with increasing E2 dose. Only expression of CHRNA_v3 did not regularly increase with increasing E2 dose (Figure 4.2).

In ER- cell line MDA-231 PS2 also increased to a lesser degree also its expression pattern was not directly correlated with increasing E2 doses degree. Previous studies also show that PS2 can increase in ER- cell lines (Figure 8.5) [78]. Moreover, PS2 (a gene in the downstream of ER pathway) expression increases with dosage confirming reliability of experiment in MCF7. Our ΔCt analysis of CHRNA5 and PS2 using clustering it is apparent that ER- cell line MDA-MD-231 expresses PS2 at much lower levels than MCF7. Expression of some of the genes decreased or showed a fluctuating expression pattern with changing dose of E2 (Figure 4.3). Heatmap of CHRNA5 isoforms and PS2 expression, also, presents the increase in gene expression in MCF7 and varying pattern in MDA MB 231 cell line (Figure 4.4).

In ER+ cell line, E2 can bind as ligand to the ER; ER pathway either genomic or non-genomicaly causes an increase in expression of a group of genes that targets of ER pathway. Supported by our results, it can be interpreted that CHRNA5 expression could be one of the genes whose expression is regulated by ER pathway activation. It is a possibility that E2 in combination with ER binds to the CHRNA5 promoter directly or indirectly. Previous studies suggested that effect of E2 on CHRNA5 is secondary since cycloheximide, a translational inhibitor, prevented the increase in expression of CHRNA5 upon E2 treatment [49]; future studies should include treatment with cycloheximide to validate this aspect of interaction between ER pathway and CHRNA5 expression.

We, then, decided to study the effect of E2 in the presence of tamoxifen (4-OHT), ER antagonist which prevents E2 to bind to ER [32]. In the presence of 4-OHT, ER+ cell lines was expected to behave like the ER- cell line in terms of gene expression changes. The question here was whether E2 when given with tamoxifen has a

reduced effect on the target genes' expression, e.g., PS2 and FOSL2, together with CHRNA5 isoforms.

Our results indicated that 4-OHT was significantly effective in decreasing all CHRNA5 isoforms' expression we analyzed except CHRNA5_v3. Previous experiments showed that some of CHRNA5 isoforms were upregulated with E2 treatment; with this study for the first time we showed that 4-OHT reduced the CHRNA5 expression in the opposite direction of E2. PS2 expression graph also confirmed effectivity of E2 although to a lesser degree than the previous experiment (Figure 4.2, Figure 4.5). Bioinformatics studies show that, PS2 expression increases with E2 and decreases with 4-OHT+E2 (Figure 8.3). However in our results, PS2 showed almost same expression level with E2 treatment after E2+4-OHT treatment (Figure 4.5). This may indicate tamoxifen resistance or incompetency between dose of E2 and tamoxifen. We therefore examined expression of FOSL2, expected to increase in 4-OHT+E2 in sensitive and resistant cell lines (although to differing degrees; Figure 8.1 and Figure 8.6). Bioinformatics studies showed that FOSL2 expression with 4-OHT treatment was half of its expression with E2 or E2+4-OHT treatment [75] (Figure 8.1) in 4-OHT sensitive cell lines. Our result for FOSL2 expression by qPCR did not yield significance (Figure 4.5). However, in order to be certain about tamoxifen sensitivity of our cell line other genes from GDS1094 and GDS4051 dataset should be chosen and their expression levels should be analyzed. Furthermore, different dose combinations of 4-OHT and E2 as well as different batches of MCF7 cell lines while increasing sample size can be used for a better understanding of the 4-OHT response of CHRNA5.

Some CHRNA5 isoforms did not show a remarkable increase with E2 treatment, while others were affected by E2 in positive way. In particular, expression of CHRNA5_v2 did not show the same increase rate with "Effects of Different Doses of E2 on CHRNA5 Isoform Expression" experiment. It may indicate that 10nm E2 was not that that efficient for this set of experiments. However, in every case, E2+4-OHT treatment is able to revert the effect of 4-OHT; the gene expression levels were very close to those of E2 treated samples after E2+ 4-OHT treatment (Figure 4.5).

PS2, CHRNA5_v1 and CHRNA5_iso3 gene expressions changed significantly between E2 and 4-OHT treatment also 4-OHT treatment resulted in a significant expression change compared to control group in PS2 and CHRNA5_v1 genes (Table 4.6, Table 4.7). This also suggests that identification of CHRNA5 isoform expression might be useful in *in vivo* diagnosis studies that include samples with different ER status and also for patients with or without tamoxifen treatment.

"Effects of 4-OHT and E2 Treatment" experiment showed that 4-OHT treatment can significantly reduce the expression of a set of genes that are found to be related with breast cancer; however in the presence of E2 this decline in the expression of these genes can be rescued. This, again, supported the hypothesis that CHRNA5 expression was affected by ER pathway activation; presence of E2 induces an increase in transcription of CHRNA5 gene even in co-treatment with 4-OHT, which may also point out to resistance developed in vitro.

To further examine the effects of E2 depletion on breast cancer, ER+ breast cancer cell line (MCF7) has been exposed to long-term E2 depletion (LTED). It may be simply thought that without E2 breast cancer related gene set expression may decrease, however previous studies show that this is not the case. After a long term depletion of E2, breast cancer cells find a way to compensate the lack of E2 [34,36].

In our study, we applied LTED for 6 months along with a control group including E2 1nm/ml (10nm) added manually. The main result from this experiment was that ESR1 expression significantly increased over time and was consistently higher than the control after time (Table 4.9).

During about first fifteen days of experiment, a difference can be observed in expression levels between control and LTED group, however after approximately first month of experiment the difference between control and LTED group has been closed; LTED group became control group like, the lack of E2 started not to affect expression of genes of interest (Figure 4.8).

In sixth months, there was almost no difference between control and LTED groups except ESR1 and PS2 genes. However, the increasing trend in the level of gene expression with the LTED can be observed. LTED experiment revealed an

interesting phenomenon related to E2- breast cancer relationship. In the absence of E2, first, cells decrease the expression of genes, whose regulations and expressions affected by ER pathway. However, after LTED, these genes increase their expressions. This phenomenon may be explained with the increase in ESR1 expression, in the absence of E2 cells might want to compensate this depletion and start to compensate by increasing their estrogen receptor expression with a feedback mechanism. One of the most common hypotheses for the mechanism behind ESR1 upregulation is that, following E2 deprivation, growth factor signaling cross-talk modulates ESR1, turning it into a supersensitive receptor able to regulate transcription in the presence of traces of estrogens or even independently of them [36]. With increasing levels of the receptor, ER pathway may be activated with other estrogen derivatives and ligands. GEO2R analysis and our results show that ESR1 expression significantly increases in the absence of E2 (Figure 1.3). This upregulation at the level of estrogen receptor may cause the increase in the level of expression of genes on the downstream of ER pathway such as PS2 and CHRNA5 (Figure 8.4). In our results, the level of PS2 increased towards the third month but then decreased again, however it was still higher than first two months (Figure 4.9). At the end of 6-month LTED course we obtained an LTED isogenic MCF7 cell line for further studies.

Panel studies are also a part of expression analysis studies and research on cancer diagnosis. Cancer subtyping should include a very large sample size for reliable prediction of disease status and treatment modality as well as interpretation of the data statistically. In order to study how different isoforms of CHRNA5 gene were expressed in different subtypes of breast cancer (ER+/-; HER2+/-; and stage N-I-II-III) a breast cancer cDNA panel from OriGene was examined.

ANOVA results showed that CHRNA5 isoforms distinctly had different expression levels in normal and breast tumor samples. This was highly significant for all tested isoforms. To see whether CHRNA5 gene can be used for subtyping, ER, HER2 status and tumor stages were also examined. CHRNA5_v3 was seen as a good candidate to distinguish ER- and ER+ breast tumor. Also stages can be separated with the differences at the level of CHRNA5 isoforms but only in comparison with

the normals for most cases. HER2 status was not significantly distinguished with a single CHRNA5 isoform expression level based on ANOVA results.

After one-by-one analysis of CHRNA5 isoform expression, we decided to analyze the correlation between isoforms in a pairwise fashion in their ability to distinguish receptor status. ER status cannot be distinguished with every single CHRNA5 isoform; however, coplots of two isoforms at once give a more detailed idea to distinguish ER status with CHRNA5 isoform expression. This suggests that combination of CHRNA5 isoform expression may closely predict ER status in breast cancer. Future analyses might include Principal Component Analysis or MANOVA. As a result, ER+ samples were observed to have relatively lower level of CHRNA5 expression for three isoforms. Between three correlation graphs between CHRNA5_v1, CHRNA5_v2 and CHRNA5_v3, CHRNA5_v3- CHRNA5_v2 pairs, respectively, the latter gives the most separated clusters of ER+ and ER- (Figure 4.13). Indeed these two isoforms correspond to the long and short forms of the CHRNA5 mRNA and are spliced from the same transcript with exclusion/inclusion of Exon V while CHRNA5_v1 amplicon corresponds to multiple isoforms that include these two and possibly others; thus this indicates that use of specific primers corresponding to a specific form of the transcript might help increase the diagnostic potential of candidate genes in cancer. Moreover, ANOVA results indicated CHRNA5_v3 can distinguish ER- and ER+ tumor samples making this isoform suitable for future studies with larger cDNA panels (Table 4.11). HER2 status could not be clearly differentiated by univariate tests or correlation studies. Also, no clear separation can be obtained to distinguish tumor stage with CHRNA5 isoform expression level, but CHRNA5 expression trend for each isoform showed an increase while stage was increasing. Regression analyses on stage might help statistically test this trend observed in the graphs.

cDNA panel analysis experiment showed that CHRNA5 isoform expression increased in breast tumor compared to normal breast tissue, similar to lung cancer [44]. Also, different level of expression of CHRNA5 isoforms may help distinguish different subtypes of breast cancer. Considering more than one isoform of CHRNA5 at a time maybe more effective for distinguishing ER status in breast cancer.

As a part of cDNA panel analysis, data quality and reliability are important issues to consider. For this purpose a new approach has been developed by checking melting temperature (Tm) of every sample for each primer. Analysis with or without Tm check is compared and it was observed that outlier number and the range of Ct remarkably decreased with Tm check (Figure 4.12). When melting curves from data were examined by PCR, we have seen a shift from the expected amplicon size with deviation from the median Tm of the cDNA panel. Melting curves should be uniform for each cycle of PCR, because the reaction happening in each cycle for one specific primer should be the same unless it amplifies another sequence rather than the amplicon of interest [79].

A workflow was then established for OriGene Breast cancer cDNA panel quality control, normalization, statistical analysis, and data visualization using multiple cDNA panel dilutions by integrating the above mentioned analyses into a GUI using MATLAB. This application was presented as an automated expression analysis tool which makes data quality and analysis much easier to handle and report the results with given options to select from (e.g., ANOVA vs. Kruskal-Wallis; delta-CT or deltadeltaCt methods; selection of clinical parameters to perform tests). The GUI presented in this study thus allows for comparison of methods while providing a platform to analyze any OriGene cDNA panel since it is possible to customize the analyses based on the columns that include clinical parameters for cancer subtypes. This tool is important because it is the first freely available tool developed to process that detailed analysis for cDNA panels. Finally, it reduces the time spent for analysis and possibility to introduce mistakes by handling the data manually.

6. FUTURE PERSPECTIVES

For further studies to reveal effects and regulatory function of CHRNA5 gene expression on breast cancer, we studied ER pathway related factors effecting to CHRNA5 expression and regulation.

It is known that CHRNA5 expression can be altered by the changes in expression level of nearby genes, they are co-regulated [47]. Expression of these genes and their relationship with different subtypes of breast cancer can be studied and analyzed in the future. Such a gene set maybe a biomarker set for diagnosis of breast cancer subtypes. Also it is known that some of these genes, such as IREB2, have already played a role in breast cancer initiation and growth [42]. Therefore, to reveal nAChRs – breast cancer relationship along with CHRNA5, genes such as CHRNA4, CHRNB3, and PSMA4 (a chromosomal neighbor) should be studied.

Moreover, to obtain a more detailed profile covering CHRNA5 expression - subtype relationship different cDNA panels can be studied. This can both increase sample size and variation in the cohort resulting in a more realistic estimate of the association. Also, new isoforms of CHRNA5 can be revealed with Northern blot studies.

To further assess E2-CHRNA5 relationship, experiments with different doses and time-points of E2 in the absence and presence of cycloheximide can be designed. This could help validation of a previous hypothesis suggesting that effects of E2 on CHRNA5 are secondary [49]. Also, different doses of 4-OHT and other ER antagonist should be tested on ER+ cell lines to obtain the exact dose and the most effective molecule to suppress effect of E2 on CHRNA5 expression and breast cancer.

In this study only two cell lines were used, however to further examine response of ER+ and ER- breast cancer cells to E2 and 4-OHT and indirectly on CHRNA5 expression, more than two cell lines including both ER+ and ER- ones could be studied.

Effects of E2 and/or 4-OHT on cells treated with an siRNA or an overexpression vector for CHRNA5 could also be studied to determine the importance of CHRNA5 expression both at the mRNA and protein level on ER signaling.

In addition to *in vitro* studies, *in vivo* approaches should be used; for example, injecting an overexpression vector or siCHRNA5 to a model organism proper to study breast tumor formation or changes in the size and invasiveness of tumor can be studied. Rodents can be a good choice to study breast cancer because of their high number of breasts.

Tissue microarrays (TMAs) including tissues from breast tumors can be studied for changes in protein expression of CHRNA5. Detecting location and density of CHRNA5 by labeling them with specific CHRNA5 Ab in different types and stages of breast cancer tissue can give an insight into the importance of this gene in diagnosis.

Finally, the GUI application designed herein can be extended by adding more options and module interfaces and can be made more appealing using different programs. The processing of template covariate description file also can generalized so that user may use the GUI on different cDNA panels with different sample arrangement, type and clinical characteristics.

REFERENCES

1. Graham, L. J. *et al.* Current Approaches and Challenges in Monitoring Treatment Responses in Breast Cancer. *J. Cancer* **5,** 58–68 (2014).

2. Ross, R. K., Wan, P. C. & Pike, M. C. Effect of Hormone Replacement Therapy on Breast Cancer Risk : Estrogen Versus Estrogen Plus Progestin. **92,** (2000).

3. Wooster, R. & Weber, B. L. 030605 Breast and Ovarian Cancer. 2339–2347 (2003).

4. Yokota, J. & Sugimura, T. Multiple steps in carcinogenesis involving alterations of multiple tumor suppressor genes. *FASEB J.* **7,** 920–5 (1993).

5. Foulkes, W. D. BRCA1--sowing the seeds crooked in the furrow. *Nat. Genet.* **40,** 8–9 (2008).

6. Weigelt, B., Geyer, F. C. & Reis-Filho, J. S. Histological types of breast cancer: how special are they? *Mol. Oncol.* **4,** 192–208 (2010).

7. Bloom, H. J. & Richardson, W. W. Histological grading and prognosis in breast cancer; a study of 1409 cases of which 359 have been followed for 15 years. *Br. J. Cancer* **11,** 359–77 (1957).

8. Perou, C. M. *et al.* Molecular portraits of human breast tumours. *Nature* **406,** 747–52 (2000).

9. States, U. IN C ER IG E. (2013).

10. Blows, F. M. *et al.* Subtyping of breast cancer by immunohistochemistry to investigate a relationship between subtype and short and long term survival: a collaborative analysis of data for 10,159 cases from 12 studies. *PLoS Med.* **7,** e1000279 (2010).

11. Kim, T. W. *et al.* Nuclear-encoded mitochondrial MTO1 and MRPL41 are regulated in an opposite epigenetic mode based on estrogen receptor status in breast cancer. *BMC Cancer* **13,** 502 (2013).

12. Li, L. *et al.* Estrogen and progesterone receptor status affect genome-wide DNA methylation profile in breast cancer. *Hum. Mol. Genet.* **19,** 4273–7 (2010).

13. Ford, C. H. J., Al-Bader, M., Al-Ayadhi, B. & Francis, I. Reassessment of estrogen receptor expression in human breast cancer cell lines. *Anticancer Res.* **31,** 521–7 (2011).

14. Ottaviano, Y. L. *et al.* Methylation of the Estrogen Receptor Gene CpG Island Marks Loss of Estrogen Receptor Expression in Human Breast Cancer Cells Advances in Brief Methylation of the Estrogen Receptor Gene CpG Island Marks Loss of Estrogen Receptor Expression in Human Breast . 2552–2555 (1994).

15. Creighton, C. J. *et al.* Molecular profiles of progesterone receptor loss in human breast tumors. *Breast Cancer Res. Treat.* **114,** 287–99 (2009).

16. Fiegl, H. *et al.* Breast cancer DNA methylation profiles in cancer cells and tumor stroma: association with HER-2/neu status in primary breast cancer. *Cancer Res.* **66,** 29–33 (2006).

17. Tessari, A., Palmieri, D. & Di Cosimo, S. Overview of diagnostic/targeted treatment combinations in personalized medicine for breast cancer patients. *Pharmgenomics. Pers. Med.* **7,** 1–19 (2013).

18. Yager, J. D. & Davidson, N. E. Estrogen carcinogenesis in breast cancer. *N. Engl. J. Med.* **354,** 270–82 (2006).

19. Nandi, S., Guzman, R. C. & Yang, J. Review Hormones and mammary carcinogenesis in mice , rats , and humans : A unifying hypothesis. **92,** 3650–3657 (1995).

20. Obiorah, I. E., Fan, P. & Jordan, V. C. Breast cancer cell apoptosis with phytoestrogens is dependent on an estrogen deprived state. *Cancer Prev. Res. (Phila).* (2014). doi:10.1158/1940-6207.CAPR-14-0061

21. Choi, E. J., Jung, J. Y. & Kim, G.-H. Genistein inhibits the proliferation and differentiation of MCF-7 and 3T3-L1 cells via the regulation of ERα expression and induction of apoptosis. *Exp. Ther. Med.* **8,** 454–458 (2014).

22. Björnström, L. & Sjöberg, M. Mechanisms of estrogen receptor signaling: convergence of genomic and nongenomic actions on target genes. *Mol. Endocrinol.* **19,** 833–42 (2005).

23. Yager, J. D. & Davidson, N. E. Estrogen Carcinogenesis in Breast Cancer. 270–282 (2006).

24. Cui, J., Shen, Y. & Li, R. Estrogen synthesis and signaling pathways during aging: from periphery to brain. *Trends Mol. Med.* **19,** 197–209 (2013).

25. O'Lone, R., Frith, M. C., Karlsson, E. K. & Hansen, U. Genomic targets of nuclear estrogen receptors. *Mol. Endocrinol.* **18,** 1859–75 (2004).

26. Migliaccio, A. *et al.* Tyrosine kinase / p21ras / MAP-kinase pathway activation by estradiol-receptor complex in MCF-7 cells. **15,** 1292–1300 (1996).

27. Castoria, G. *et al.* PI3-kinase in concert with Src promotes the S-phase entry of oestradiol-stimulated MCF-7 cells. **20,** (2001).

28. Lymperatou, D., Giannopoulou, E., Koutras, A. K. & Kalofonos, H. P. The exposure of breast cancer cells to fulvestrant and tamoxifen modulates cell migration differently. *Biomed Res. Int.* **2013,** 147514 (2013).

29. Higgins, M. J. & Baselga, J. Review series Targeted therapies for breast cancer. **121,** (2011).

30. Pearce, S. T., Liu, H. & Jordan, V. C. Modulation of estrogen receptor alpha function and stability by tamoxifen and a critical amino acid (Asp-538) in helix 12. *J. Biol. Chem.* **278,** 7630–8 (2003).

31. Sanchez, A. & Villanueva, J. PI3K-based molecular signatures link high PI3K pathway activity with low ER levels in ER+ breast cancer. *Expert Rev. Proteomics* **7,** 819–21 (2010).

32. Hiscox, S. *et al.* Tamoxifen resistance in MCF7 cells promotes EMT-like behaviour and involves modulation of beta-catenin phosphorylation. *Int. J. Cancer* **118,** 290–301 (2006).

33. Mo, Z. *et al.* GPR30 as an initiator of tamoxifen resistance in hormone-dependent breast cancer. *Breast Cancer Res.* **15,** R114 (2013).

34. Johnston, S. R. *et al.* Changes in estrogen receptor, progesterone receptor, and pS2 expression in tamoxifen-resistant human breast cancer. *Cancer Res.* **55,** 3331–8 (1995).

35. Clarke, R. *et al.* Antiestrogen resistance in breast cancer and the role of estrogen receptor signaling. *Oncogene* **22,** 7316–39 (2003).

36. Aguilar, H. *et al.* Biological reprogramming in acquired resistance to endocrine therapy of breast cancer. *Oncogene* **29,** 6071–83 (2010).

37. Tolhurst, R. S. *et al.* Transient over-expression of estrogen receptor-α in breast cancer cells promotes cell survival and estrogen-independent growth. *Breast Cancer Res. Treat.* **128,** 357–68 (2011).

38. Paterson, D. & Nordberg, a. Neuronal nicotinic receptors in the human brain. *Prog. Neurobiol.* **61,** 75–111 (2000).

39. Itier, V. & Bertrand, D. Neuronal nicotinic receptors: from protein structure to function. *FEBS Lett.* **504,** 118–25 (2001).

40. Weiland, S., Bertrand, D. & Leonard, S. Neuronal nicotinic acetylcholine receptors: from the gene to the disease. *Behav. Brain Res.* **113,** 43–56 (2000).

41. Albuquerque, E. X., Pereira, E. F. R., Alkondon, M. & Rogers, S. W. Mammalian Nicotinic Acetylcholine Receptors : From Structure to Function. 73–120 (2009). doi:10.1152/physrev.00015.2008.

42. Krais, A. M. *et al.* CHRNA5 as negative regulator of nicotine signaling in normal and cancer bronchial cells: effects on motility, migration and p63 expression. *Carcinogenesis* **32,** 1388–95 (2011).

43. Lee, C.-H. *et al.* Crosstalk between nicotine and estrogen-induced estrogen receptor activation induces α9-nicotinic acetylcholine receptor expression in human breast cancer cells. *Breast Cancer Res. Treat.* **129,** 331–45 (2011).

44. Falvella, F. S. *et al.* Multiple isoforms and differential allelic expression of CHRNA5 in lung tissue and lung adenocarcinoma. *Carcinogenesis* **34,** 1281–5 (2013).

45. Lee, C.-H. *et al.* Overexpression and activation of the alpha9-nicotinic receptor during tumorigenesis in human breast epithelial cells. *J. Natl. Cancer Inst.* **102,** 1322–35 (2010).

46. Smith, R. M. *et al.* Nicotinic α5 receptor subunit mRNA expression is associated with distant 5' upstream polymorphisms. *Eur. J. Hum. Genet.* **19,** 76–83 (2011).

47. Falvella, F. S. *et al.* Transcription deregulation at the 15q25 locus in association with lung adenocarcinoma risk. *Clin. Cancer Res.* **15,** 1837–42 (2009).

48. Tu, S.-H. *et al.* Tea polyphenol (-)-epigallocatechin-3-gallate inhibits nicotine- and estrogen-induced α9-nicotinic acetylcholine receptor upregulation in human breast cancer cells. *Mol. Nutr. Food Res.* **55,** 455–66 (2011).

49. Biyik, R., Nicotine-modulated gene expression profiles in MCF7 breast cancer cell line and involvement of estrogen in CHRNA5 expression. Master Thesis, Bilkent University (2009)

50. Feil, R. *et al.* Ligand-activated site-specific recombination in mice. *Proc. Natl. Acad. Sci. U. S. A.* **93,** 10887–90 (1996).

51. Niu, Y.-N. *et al.* Alternative messenger RNA splicing of autophagic gene Beclin 1 in human B-cell acute lymphoblastic leukemia cells. *Asian Pac. J. Cancer Prev.* **15,** 2153–8 (2014).

52. Black, D. L. Mechanisms of alternative pre-messenger RNA splicing. *Annu. Rev. Biochem.* **72,** 291–336 (2003).

53. Chu, X. *et al.* Alternative splicing variants of human Fbx4 disturb cyclin D1 proteolysis in human cancer. *Biochem. Biophys. Res. Commun.* **447,** 158–64 (2014).

54. Dery, K. J. *et al.* IRF-1 regulates alternative mRNA splicing of carcinoembryonic antigen-related cell adhesion molecule 1 (CEACAM1) in breast epithelial cells generating an immunoreceptor tyrosine-based inhibition motif (ITIM) containing isoform. *Mol. Cancer* **13,** 64 (2014).

55. Ubby, I. *et al.* TMEM16A alternative splicing coordination in breast cancer. *Mol. Cancer* **12,** 75 (2013).

56. Zhao, S., Chang, S. L., Linderman, J. J., Feng, F. Y. & Luker, G. D. A Comprehensive Analysis of CXCL12 Isoforms in Breast Cancer(1,2.). *Transl. Oncol.* **7,** 429–438 (2014).

57. Maimon, A. *et al.* Article Mnk2 Alternative Splicing Modulates the p38-MAPK Pathway and Impacts Ras-Induced Transformation. *CellReports* **7,** 501–513 (2014).

58. Griffith, M. *et al.* A microarray design for the detection of alternate isoforms Application to a model of chemotherapy resistance in colon cancer. 25 (2003).

59. Wan, G. & Too, H. Discrimination of Alternative Spliced Isoforms by Real-Time PCR Using Locked Nucleic Acid (LNA) Substituted Primer. 3–6

60. Mekouar, M. *et al.* Detection and analysis of alternative splicing in Yarrowia lipolytica reveal structural constraints facilitating nonsense-mediated decay of intron-retaining transcripts. *Genome Biol.* **11,** R65 (2010).

61. Merindol, N. *et al.* The emerging role of Twist proteins in hematopoietic cells and hematological malignancies. *Blood Cancer J.* **4,** e206 (2014).

62. Wei, N., Chu, E., Wipf, P. & Schmitz, J. C. Protein kinase d as a potential chemotherapeutic target for colorectal cancer. *Mol. Cancer Ther.* **13,** 1130–41 (2014).

63. Cavnar, S. P. *et al.* Microfluidic source-sink model reveals effects of biophysically distinct CXCL12 isoforms in breast cancer chemotaxis. *Integr. Biol. (Camb).* **6,** 564–76 (2014).

64. Pfaffl, M. W. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* **29,** e45 (2001).

65. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25,** 402–8 (2001).

66. Minafra, L. *et al.* Gene expression profiling of epithelial-mesenchymal transition in primary breast cancer cell culture. *Anticancer Res.* **34,** 2173–83 (2014).

67. Barrett, T. *et al.* NCBI GEO: mining tens of millions of expression profiles--database and tools update. *Nucleic Acids Res.* **35,** D760–5 (2007).

68. Eickhoff, H. Tissue Gene Expression Analysis Using Arrayed Normalized cDNA Libraries. *Genome Res.* **10,** 1230–1240 (2000).

69. Matlab, U. The Language of Technical Computing.

70. Acikgoz, A., Analysis of CHRNA5 expression in breast cancer cell lines in response to serum starvation and estrogen treatment. Master Thesis, Bilkent University, (2013).

71. Holliday, D. L. & Speirs, V. Choosing the right cell line for breast cancer research. *Breast Cancer Res.* **13,** 215 (2011).

72. Blackwell, K. L., Haroon, Z. A., Shan, S., Saito, W. & Broadwater, G. Tamoxifen Inhibits Angiogenesis in Estrogen Receptor-negative Animal Models Tamoxifen Inhibits Angiogenesis in Estrogen Receptor-negative Animal Models 1. 4359–4364 (2000).

73. Smyth, G. K., Ritchie, M. & Thorne, N. Linear Models for Microarray Data User ' s Guide (Now Including RNA-Seq Data Analysis). (2014).

74. Gao, Q.-G., Chan, H.-Y., Man, C. W.-Y. & Wong, M.-S. Differential ERα-mediated rapid estrogenic actions of ginsenoside Rg1 and estren in human breast cancer MCF-7 cells. *J. Steroid Biochem. Mol. Biol.* **141,** 104–12 (2014).

75. Gonzalez-malerva, L. *et al.* High-throughput ectopic expression screen for tamoxifen resistance identi fi es an atypical kinase that blocks autophagy. (2010). doi:10.1073/pnas.1018157108/-/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1018157108

76. Dembowski, J. a *et al.* Alternative Splicing of a Novel Inducible Exon Diversifies the CASK Guanylate Kinase Domain. *J. Nucleic Acids* **2012,** 816237 (2012).

77. Akman, B. H., Can, T. & Erson-Bensan, a E. Estrogen-induced upregulation and 3'-UTR shortening of CDC6. *Nucleic Acids Res.* **40,** 10679–88 (2012).

78. Moggs, J. G. *et al.* Anti-proliferative effect of estrogen in breast cancer cells that re-express ERalpha is mediated by aberrant regulation of cell cycle genes. *J. Mol. Endocrinol.* **34,** 535–51 (2005).

79. Ririe, K. M., Rasmussen, R. P. & Wittwer, C. T. Product differentiation by analysis of DNA melting curves during the polymerase chain reaction. *Anal. Biochem.* **245,** 154–60 (1997).

APPENDICES

Appendix A

Covdesc File

Appendix B

GeoDataSet Graphs

Appendix C

Time Point Table for LTED Experiment

Appendix D

Permissions for the Copyrighted Materials

Appendix A

Covdesc File

Table 8.1: Covdesc file of cDNA panel

qPCRwell	PatientID	Replicate	Tumor	ER	PR	HER2	Stage
A1	C1	1	0	5	5	5	5
A2	C2	1	0	5	5	5	5
A3	C3	1	0	5	5	5	5
A4	C4	1	0	5	5	5	5
A5	C5	1	0	5	5	5	5
A6	C6	1	0	5	5	5	5
A7	C7	1	0	5	5	5	5
A8	C8	1	1	0	0	0	1
A9	C9	1	1	1	1	0	1
A10	C10	1	1	0	0	0	1
A11	C11	1	1	1	1	1	1
A12	C12	1	1	0	1	0	1
B1	C1	2	0	5	5	5	5
B2	C2	2	0	5	5	5	5
B3	C3	2	0	5	5	5	5
B4	C4	2	0	5	5	5	5
B5	C5	2	0	5	5	5	5
B6	C6	2	0	5	5	5	5
B7	C7	2	0	5	5	5	5
B8	C8	2	1	0	0	0	1
B9	C9	2	1	1	1	0	1
B10	C10	2	1	0	0	0	1
B11	C11	2	1	1	1	1	1
B12	C12	2	1	0	1	0	1
C1	D1	1	1	1	1	1	1
C2	D2	1	1	NaN	NaN	NaN	1
C3	D3	1	1	0	0	1	1
C4	D4	1	1	1	1	0	1
C5	D5	1	1	NaN	NaN	NaN	1
C6	D6	1	1	1	1	1	2
C7	D7	1	1	0	0	0	2
C8	D8	1	1	0	NaN	0	2
C9	D9	1	1	0	0	1	2
C10	D10	1	1	1	1	0	2
C11	D11	1	1	1	1	0	2
C12	D12	1	1	0	0	1	2
D1	D1	2	1	1	1	1	1

D2	D2	2	1	NaN	NaN	NaN	1
D3	D3	2	1	0	0	1	1
D4	D4	2	1	1	1	0	1
D5	D5	2	1	NaN	NaN	NaN	1
D6	D6	2	1	1	1	1	2
D7	D7	2	1	0	0	0	2
D8	D8	2	1	0	NaN	0	2
D9	D9	2	1	0	0	1	2
D10	D10	2	1	1	1	0	2
D11	D11	2	1	1	1	0	2
D12	D12	2	1	0	0	1	2
E1	E1	1	1	1	0	0	2
E2	E2	1	1	0	0	0	2
E3	E3	1	1	0	0	1	2
E4	E4	1	1	1	1	0	2
E5	E5	1	1	NaN	NaN	1	2
E6	E6	1	1	NaN	NaN	NaN	2
E7	E7	1	1	0	1	0	2
E8	E8	1	1	NaN	NaN	NaN	2
E9	E9	1	1	1	1	1	2
E10	E10	1	1	0	NaN	0	2
E11	E11	1	1	1	1	1	2
E12	E12	1	1	0	0	0	2
F1	E1	2	1	1	0	0	2
F2	E2	2	1	0	0	0	2
F3	E3	2	1	0	0	1	2
F4	E4	2	1	1	1	0	2
F5	E5	2	1	NaN	NaN	1	2
F6	E6	2	1	NaN	NaN	NaN	2
F7	E7	2	1	0	1	0	2
F8	E8	2	1	NaN	NaN	NaN	2
F9	E9	2	1	1	1	1	2
F10	E10	2	1	0	NaN	0	2
F11	E11	2	1	1	1	1	2
F12	E12	2	1	0	0	0	2
G1	F1	1	1	NaN	NaN	NaN	2
G2	F2	1	1	0	0	1	3
G3	F3	1	1	0	NaN	0	3
G4	F4	1	1	1	1	0	3
G5	F5	1	1	1	0	1	3
G6	F6	1	1	1	1	0	3
G7	F7	1	1	0	0	0	3
G8	F8	1	1	NaN	NaN	NaN	3

G9	F9	1	1	0	1	0	3
G10	F10	1	1	NaN	NaN	0	3
G11	F11	1	1	1	1	0	3
G12	F12	1	1	1	0	0	3
H1	F1	2	1	NaN	NaN	NaN	2
H2	F2	2	1	0	0	1	3
H3	F3	2	1	0	NaN	0	3
H4	F4	2	1	1	1	0	3
H5	F5	2	1	1	0	1	3
H6	F6	2	1	1	1	0	3
H7	F7	2	1	0	0	0	3
H8	F8	2	1	NaN	NaN	NaN	3
H9	F9	2	1	0	1	0	3
H10	F10	2	1	NaN	NaN	0	3
H11	F11	2	1	1	1	0	3
H12	F12	2	1	1	0	0	3

Appendix B

GeoDataSet Graphs

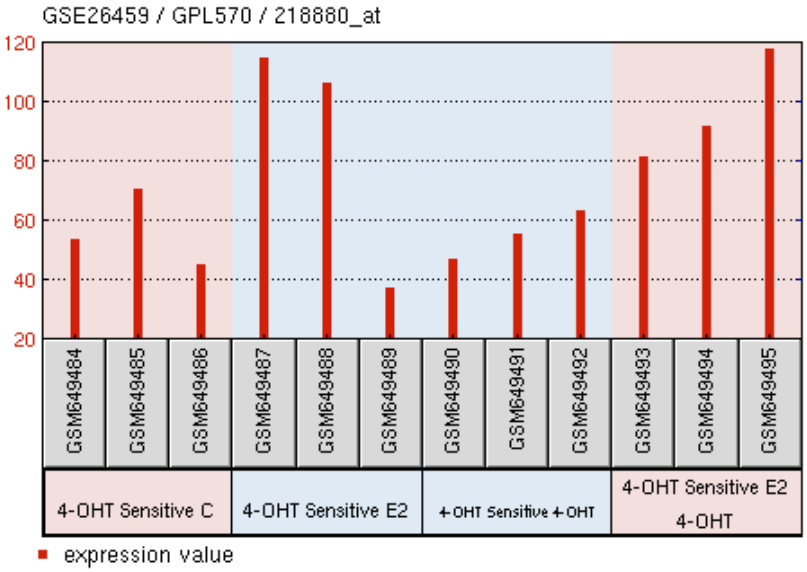


Figure 8.1: GEO2R analysis of GSE26459 dataset for FOSL2 expression with E2 and 4-OHT in 4-OHT sensitive cell lines.

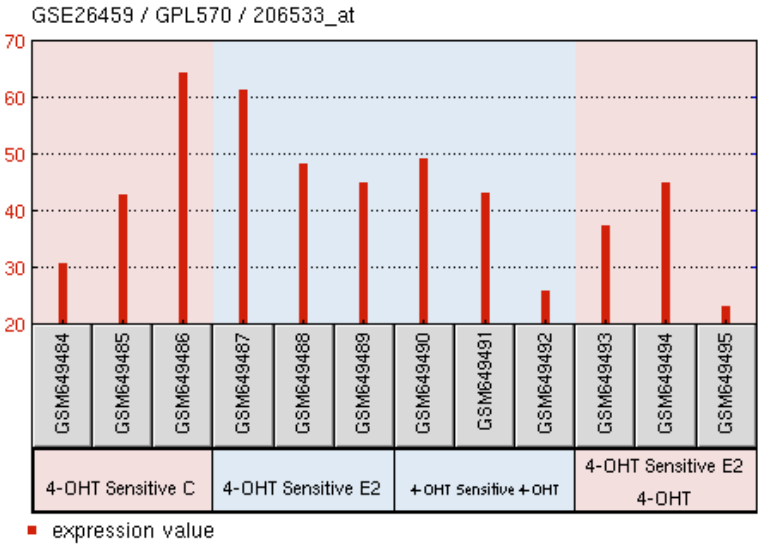


Figure 8.2: GEO2R analysis of GSE26459 dataset for CHRNA5 expression with E2 and 4-OHT in 4-OHT sensitive cell lines.

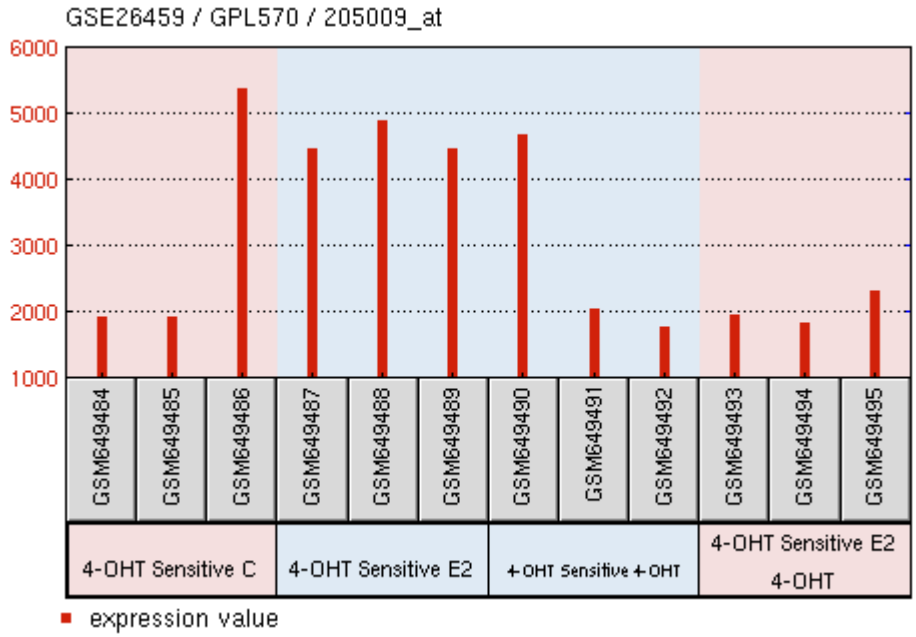


Figure 8.3: GEO2R analysis of GSE26459 dataset for PS2 expression with E2 and 4-OHT in 4-OHT sensitive cell lines.

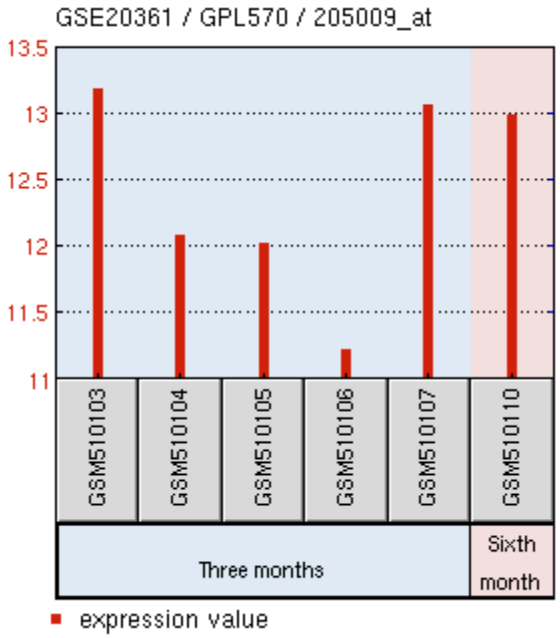


Figure 8.4: GEO2R analysis of GSE20361 dataset for PS2 expression after and before LTED.

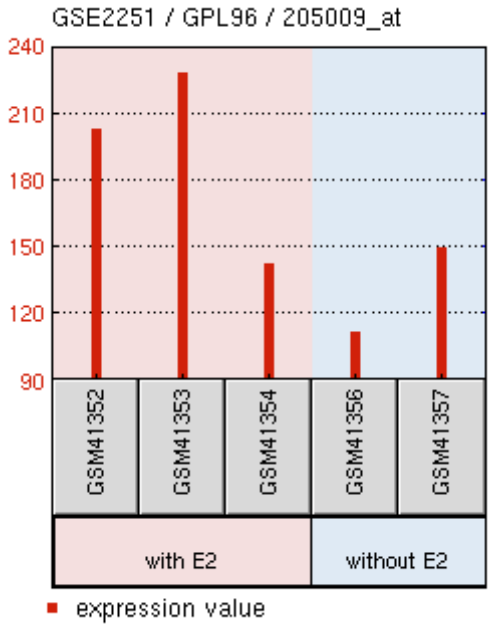


Figure 8.5: GEO2R analysis of GSE2251 dataset for PS2 expression with and without E2 treatment in ER- MDA MB 231 cell line.

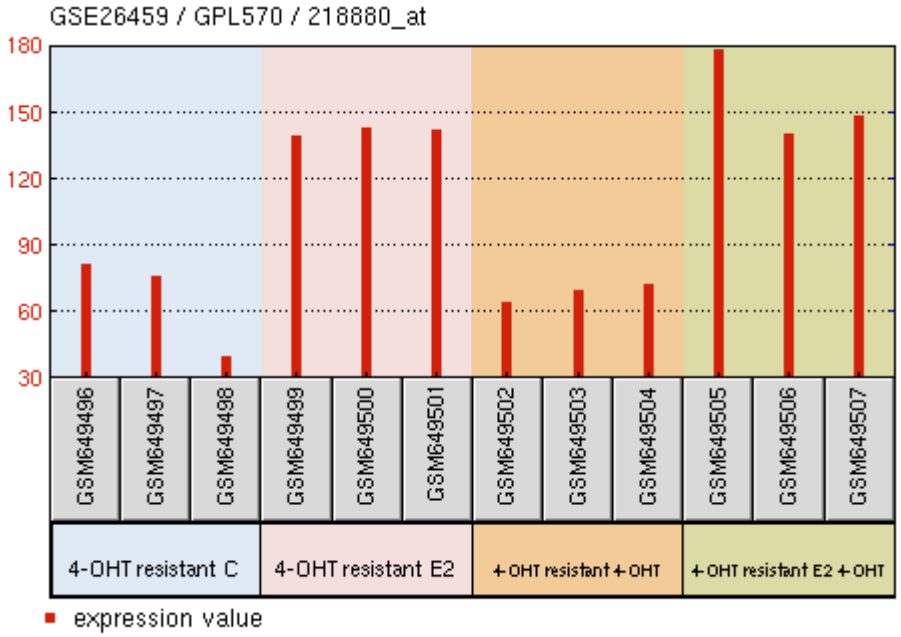


Figure 8.6: GEO2R analysis of GSE26459 dataset for FOSL2 expression with E2 and 4-OHT in 4-OHT resistant cell lines.

Appendix C

Time Point Table for LTED Experiment

Table 8.2: Time points for LTED experiment

Sample	Collection date	Experimental date
1	12.12.13	1
2	25.12.13	13
3	10.01.14	29
4	14.01.14	33
5	27.01.14	46
6	13.02.14	63
7	17.02.14	67
8	26.02.14	76
9	10.03.14	88
10	05.05.14	150
11	13.05.14	158
12	19.05.14	164
13	25.05.14	170
14	03.06.14	180

Appendix D

Permissions for the Copyrighted Materials

RightsLink®

Home | Account Info | Help

Title:	BRCA1[mdash]sowing the seeds crooked in the furrow
Author:	William D Foulkes
Publication:	Nature Genetics
Publisher:	Nature Publishing Group
Date:	Jan 1, 2008

Copyright © 2008, Rights Managed by Nature Publishing Group

Logged in as:
Sıla ÖZdemir
LOGOUT

Order Completed

Thank you very much for your order.

This is a License Agreement between Sıla ÖZdemir ("You") and Nature Publishing Group ("Nature Publishing Group"). The license consists of your order details, the terms and conditions provided by Nature Publishing Group, and the payment terms and conditions.

Get the printable license.

License Number	3423670193494
License date	Jul 07, 2014
Licensed content publisher	Nature Publishing Group
Licensed content publication	Nature Genetics
Licensed content title	BRCA1[mdash]sowing the seeds crooked in the furrow
Licensed content author	William D Foulkes
Licensed content date	Jan 1, 2008
Type of Use	reuse in a dissertation / thesis
Volume number	40
Issue number	1
Requestor type	academic/educational
Format	print and electronic
Portion	figures/tables/illustrations

Title: Neuronal nicotinic acetylcholine receptors: from the gene to the disease
Author: S. Weiland,D. Bertrand,S. Leonard
Publication: Behavioural Brain Research
Publisher: Elsevier
Date: August 2000

Order Completed

Thank you very much for your order.

This is a License Agreement between Sıla ÖZdemir ("You") and Elsevier ("Elsevier"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

Get the printable license.

License Number	3425941385837
License date	Jul 11, 2014
Licensed content publisher	Elsevier
Licensed content publication	Behavioural Brain Research
Licensed content title	Neuronal nicotinic acetylcholine receptors: from the gene to the disease
Licensed content author	S. Weiland,D. Bertrand,S. Leonard
Licensed content date	August 2000
Licensed content volume number	113
Licensed content issue number	1–2
Number of pages	14
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	both print and electronic