MICROARRAY APPLICATIONS FOR DETERMINATION OF THE EFFECTS OF EMODIN ON BREAST CANCER CELL LINES

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

BY

EMILIA QOMI EKENEL

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

FEBRUARY 2012

Approval of the thesis:

MICROARRAY APPLICATIONS FOR DETERMINATION OF EFFECTS BY RHEUM ON MCF7 AND MDA231 CELL LINES

submitted by **EMILIA QOMI EKENEL** in partial fulfillment of the requirements for the degree of **Master of Science in Biotechnology Department, Middle East Technical University** by,

Prof. Dr. Canan Özgen _____
Dean, Graduate School of **Natural and Applied Sciences**

Prof. Dr. Nesrin Hasırci _____
Head of Department, **Biotechnology**

Prof. Dr. Mesude İşcan _____
Supervisor, **Biological Sciences Dept., METU**

Assoc. Prof. Dr. Nursen Çoruh _____
Co-Supervisor, **Chemistry Dept., METU**

Examining Committee Members:

Prof. Dr. Meral Yücel _____
Biological Sciences Dept., METU

Prof. Dr. Mesude İşcan _____
Biological Sciences Dept., METU

Assoc. Prof. Dr. Nursen Çoruh _____
Chemistry Dept., METU

Assist. Prof. Dr. Yeşim Aydın Son _____
Med. Informatics Dept., METU

Prof. Dr. Leyla Açık _____
Biological Sciences Dept., GAZI U

Date: 09/02/2012

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

Name, Last name : Emilia Qomi Ekenel

Signature :

ABSTRACT

MICROARRAY APPLICATIONS FOR DETERMINATION OF THE EFFECTS OF EMODIN ON BREAST CANCER CELL LINES

Ekenel Qomi, Emilia
M.S., Department of Biotechnology
Supervisor: Prof. Dr. Mesude İşcan
Co-Supervisor: Assoc. Prof. Dr. Nursen Çoruh

February 2012, 191 pages

Cancer is a genetic disease that is characterized by uncontrolled cells growth. Breast cancer is a type of cancer originating from breast tissue. Some breast cancers are sensitive to hormones such as estrogen which makes it possible to treat them by blocking the effects of these hormones in the target tissues. These require less aggressive treatment than hormone negative cancers. Breast cancers without hormone receptors, are higher-risk, and are treated more aggressively.

The aim of our study is to investigate the effect of emodin on MCF-7 which is ER (estrogen receptor) positive, and MDA-MB-231 (ER negative) cancerous cell lines. Emodin which is a phytoestrogen component, extracted from rheum (genus) plant, has been reported to suppress the growth of tumor in some clinical situation, and it's found that emodin induced apoptosis through the decrease of Bcl-2/Bax ratio and the increase of cytoplasm cytochrome c concentration in human breast cancer Bcap-37 cells. Comparing the effect of emodin between ER positive and ER negative cells at the molecular level was investigated by Microarray analysis of gene expressions using Affymetrix Human Genome U133 plus 2.0 Array. The microarray data analysis was performed by using BRB-Array Tools, v.4.2.0.

GST and its classes; Alpha, Mu, Pi, Theta, Sigma, Omega, Zeta and Kappa is our interested genes because of its role in regulating susceptibility to cancer, by their ability to metabolize reactive electrophilic intermediates to usually less reactive and more water soluble glutathione conjugates. And also its have a role in detoxifying the damage caused by oxidative stress which is a result of the radiotherapy.

The differentially expressed genes from emodin treated and untreated control breast cancer cell lines were compared after normalization and filtering and annotated, it was shown that the top 10 highly (significantly) varied genes belong to the biological processes such as (namely) cell cycle, cell division, cell proliferation, mitosis and meiosis, this insure the relation of emodin to the cell growth processes in the cancerous cells. The analysis of the change on the cell growth confirmed the anti-tumor effect of emodin.

About the effect of emodin treatment on MCF-7 and MDA-MB-231 cancerous cell lines separately; Both cells its significant genes was belong to cell growth biological processes, in MCF-7 cells in-addition other biological processes was shown, for example; stimulus to estradoil response, and the metabolism of xenobiotic by cytochrome p450, so CYP1A1 gene code for a protein which is used in emodin metabolism. The varied gene number was nearly 4400 gene from the scatter plot result in MCF-7 cells while in MDA-MB-231 cells it was nearly 3400 gene, these result insured the effect of emodin as a phytoestrogenic component as MCF-7 cells are ER positive cells, so emodin bind to the ER in MCF-7 cells and affected more gene number than MDA-MB-231.

More number of GST enzyme classes changed in MCF-7 cells than MDA-MB-231, and the effect of emodin as anti-cancer showed different change of GST genes between MCF-7 and MDA-MB-231.

The results confirmed by network analysis done, to find the most related genes to our top 10 regulated gene list, and these genes were analyzed; most of them where in our gene list, and their regulation after emodin treatment analyzed and the result was supported to emodin as anti-tumor and phytoestrogenic component.

Keywords: Breast cancer, Microarray technique, GST (glutathione S-transferase), ER (Estrogen receptor), MCF-7, MDA-MB-231, CYP1A1, Network Analysis.

öz

EMODİN ETKİLERİNİN MEME KANSERİ HUCRE HATLARİNDA MİKRODİZİN ANALİZİ İLE İNCELENMESİ

Ekenel Qomi, Emilia
Yüksek lisans, Biyoteknoloji Bölümü
Tez Yöneticisi: Prof. Dr. Mesude İşcan
Ortak Tez Yöneticisi :Doç. Dr. Nursen Çoruh

Şubat 2012, 191 sayfa

Kanser, kontrolsüz hücre büyümesi ile karakterize edilen genetik bir hastalıktır. Meme kanseri, meme dokusundan kaynaklanan bir kanser türüdür. Östrojen gibi hormonlara hassas bazı meme kanserleri, bu hormonların hedef dokulardaki etkilerini bloke ederek tedavi edilebilir. Bu tip kanserlerin tedavi edilmesi hormon reseptörleri olmayan kanserlere göre daha kolay olmaktadır. Diğer yandan hormon reseptörleri olmayan meme kanserlerinin tedavisi daha zor ve daha risklidir.

Çalışmamızın amacı ER (östrojen reseptörü) pozitif MCF-7 ve ER negatif MDA-MB-231 kanserli hücre hatları üzerinde emodinin etkisini araştırmaktır. Emodin, bazı klinik durumlarda tümör büyümesini baskıladığı tespit edilmiş Rheum (cinsi) bitkiden elde edilen bir fitoöstrojen bileşenidir. Bu çalışmada, Emodinin kanser hücreleri üzerinde moleküler düzeyde etkilerini belirlemek için, Mikroarray tekniği kullanılmıştır.

Emodine maruz bırakılan MCF-7 ve MDA-MB-231 kanser hücre hatlarında GST ve onun sınıflarının (alfa, mu, pi, theta, sigma, mega, zeta ve kapa) gen düzeyinde ki değişiklikleri görmek için analiz edildi. GST enzimleri daha reaktif elektofilik ara ürünleri daha az reaktif ve daha suda çözünür glutatyona metabolize ettikleri için kansere yatkınlığın düzenlenmesinde önemlidirler. Ayrıca GST enzimleri radyoterapi sonucu oksidatif stresden kaynaklanan hasarı detoksifiye etmede bir role sahiptir. Emodine maruz bırakılan kanser hücrelerinin ve kontrol kanser hücrelerinin gen anlatımları karşılaştırılıldığında, emodine maruz bırakılan kanserli hücreler için hücre

bölünmesi, hücre döngüsü, hücre çoğalması, mitoz ve mayoz gibi biyolojik süreçlerde önemli ilk 10 genin anlatımda anlamlı bir azalma bulundu. Ayrıca MDA-MB-231 ve MCF-7 hücre hatları arasında da genlerin düzenlenmesinde anlamlı farklılıklar bulundu. MCF-7 hücre hattında, MDA-MB-231 hücre hattına göre daha fazla ve farklı genler düzenlenlendiği tespit edildi. Mesela MCF hücre hattında zenobiyotiklerinin metabolizmasından sorumlu CYP1A1 gen anlatımı daha fazla artmıştır bunda MCF-7 hücreleri ER pozitif oldukları için emodine fitoöstrojen olarak yanıt vermesinin de etkisi vardır.

GST enzim sınıfları, MCF-7 hücrelerinde MDA-MB-231 hücrelerine gore daha fazla değişti ve emodinin anti kanser etkisi MCF-7 ve MDA-MB-231 hücrelerinde farklı GST genlerinde değişme gösterdi.

Sonuçlar Ağ analiz yapılarak doğrulandı.Bizim ilk 10 regüle gen listesine en ilişkili genleri bulmak için,bu genler analiz edildi. Bizim gen listesindeki genlerin çoğu normalizasyon, filtreleme yapıldıktan sonra emodin maruziyeti sonrası bunların düzenlenmesi analiz edildi ve sonuçlar emodinin anti tümör etkisi ve fitoöstrojenik bileşeni olduğunu destekledi.

Anahtar Kelimeler: Meme kanseri, Mikroarray tekniği, GST (glutatyon S-transferase), ER (östrojen reseptörü), MCF-7, MDA-MB-231, CYP1A1, Ağ Analizi.

To my dear Parents, beloved Husband and Son.

ACKNOWLEDGEMENTS

First of all, I would like to express my deepest gratitude to my supervisor Prof. Dr. Mesude İşcan for her great advises, support and guidance throughout this study.

I feel great appreciation to Assist. Prof. Dr. Yeşim Aydın Son for her guidance and support especially in the Bioinformatics Part Of this Study.

I would like to thank the Molecular Biology and Biotechnology Research Center/METU for the Microarray Work and REFGEN Biotechnology Company for DNA Sequencing work.

I Also want to Acknowledge the BRB-Array Tool which developed by Dr. Richard Simon and BRB-Array Tools Development Team; our Microarray analysis done using this tool.

I acknowledge examining committee members, my Co-supervisor; Assoc. Prof. Dr. Nursen Çoruh, Prof. Dr. Meral Yücel, and prof. Dr. Leyla Açık for evaluating my thesis and their contributions.

I have special thanks for Elif Sakalli for her help and Friendship through this study especially on the tissue culture Part; also i am deeply thankful for Dr. Pembegül Uyar for her help and friendship especially in the Microarray Part and RNA isolation. Also i am thankful for Metin Konuş and Elif Aşik for their friendship and help in the RT-PCR work.

I would like to thank my Lab mates; Selis Yılmaz, Deniz Hisarlı, Deniz İrtem Kartal, Bade Kaya, and Can Yılmaz for their friendship and support through this study.

I am deeply gratitude to my mother Sabiha Akpinar Qomi and to my father Mohd Hassan Qomi for their endless love, support in every step of my life and encouragement to complete my study in the Field that i love.

I would like to express my great appreciation to my husband Hüseyin Ekenel for his love, care, support and encouragement through my study period, and to my Dear son Ali Ekenel for his presence in my Life and for being a good boy which helped me during this period.

Lastly, I would like to express my appreciation to my sisters; Alice, Ruba, Rawan and muna Qomi and to my brother Sameeh Qomi for their love, support and encouragement during this study and in every step of my life.

TABLE OF CONTENTS

ABSTRACT.....iv

ÖZ.....vi

ACKNOWLEDGEMENTS..... ix

TABLE OF CONTENTS..... xi

LIST OF TABLES..... xv

LIST OF FIGURES xvii

LIST OF SYMBOLS AND ABBREVIATIONS..... xx

CHAPTERS

1. INTRODUCTION..... 1

1.1 Breast cancer..... 1

1.1.1 Estrogen in breast cancer..... 2

1.1.2 Glutathione S-transferase (GST)..... 5

1.1.3 Emodin..... 5

1.1.4 MCF7 and MDA Cell Lines..... 6

1.1.5 Cancer studies based on microarray analysis..... 7

1.2 microarray..... 7

1.2.1 History of Microarray Evolution..... 9

1.2.2 Gene Expression microarray..... 10

1.2.2.1 Affymetrix gene expression microarrays..... 10

1.2.2.2 Affymetrix Probes..... 11

1.2.3 Experimental Design..... 12

1.2.4 DNA Microarray limitation..... 12

1.2.5 Controlling non-biological variance..... 13

1.2.5.1 RNA Quality assesment..... 14
1.2.6 Data analysis..... 16
1.2.6.1 Image analysis..... 16
1.2.6.2 Preprocessing..... 17
1.2.6.2.1 Log2 transformation..... 17
1.2.6.2.2 Filtering signal..... 17
1.2.6.2.3 Background correction..... 18
1.2.6.2.4 Normalization..... 18
1.2.6.2.5 Summarization..... 19
1.2.6.2.6 Different preprocessing algorithms available..... 19
1.2.6.3 Assessment of array quality..... 20
1.2.6.3.1 B2 Oligo performance..... 20
1.2.6.3.2 Poly-A Controls: lys, phe, thr, dap..... 21
1.2.6.3.3 Hybridization Controls: bioB, bioC, bioD, and cre..... 22
1.2.6.3.4 Internal Control Genes..... 22
1.2.6.4 Differentially expressed genes..... 22
1.2.6.5 Clustering and Classification..... 23
1.2.6.6 Extraction of Biological Information..... 24
1.2.7 Network analysis using GeneMANIA Tool..... 24
1.2.8 Software for Data Analysis..... 24
1.2.8.1 BRB (Biometric Research Branch) Array Tools..... 24
1.2.8.2 Gene Spring GX software..... 25
1.3 The Aim of the study..... 26
2. MATERIAL AND METHODS..... 27
2.1 Experiment description..... 27
2.1. Cell culture..... 27
2.1.2 Biological Replicate..... 27

2.2 RNA Isolation..... 28
2.3 Microarray experiment.....28
2.3.1 Target Preparation..... 28
2.3.2 Target Hybridization, washing and staining..... 29
2.4 Data Analysis..... 32
2.4.1 Image analysis..... 32
2.4.2 Assessment of the quality control..... 32
2.4.3 Preprocessing..... 32
2.4.4 Differentially expressed genes..... 33
2.4.5 Top 10 pathways from DAVID Annotation Tool.....35
2.4.6 Clustering..... 35
2.4.7 Our gene of interest list..... 35
2.5 Primer design for Real time PCR..... 35
2.5.1 Checking the Primers..... 36
2.5.1.1 RNA Isolation..... 37
2.5.1.2 cDNA Preperation..... 37
2.5.1.3 PCR Work.....37
2.5.1.4 DNA Sequencing.....38
3. RESULTS AND DISCUSSIONS..... 39
3.1 RNA Quality assessment (bioanalyzer Results)..... 39
3.2 Quality control and Preprocessing evaluation of affymetrix CEL..... 42
3.2.1 RNA degradation plot..... 42
3.2.2 2D images for spatial bias diagnostic..... 44
3.2.3 Density histogram..... 46
3.2.4 Boxplot..... 47
3.2.5 The PCA (Principal Component Analysis)..... 49

3.3 Scatter plot phenotype average between Emodin and control classes after preprocessing..... 51

3.4 Differentially expressed genes and their analysis..... 53

3.4.1 Analysis of the top 10 regulated genes in the 8 arrays for both cells type..... 54

3.4.2 Analysis of the top 10 regulated genes in the 4 arrays for MCF-7 cells..... 59

3.4.3 Analysis of the top 10 regulated genes in the 4 arrays for MDA-MB-231 cells..... 66

3.5 Top 10 pathways from DAVID Annotation Tool..... 72

3.6 Clustering between all samples and genes for pattern discovery..... 73

3.7 Our gene of interest list.....75

3.8 Checking the GST primers designed for real time PCR..... 82

3.8.1 RNA Isolation result..... 82

3.8.2 PCR Result..... 83

3.8.3 DNA Sequencing.....85

4. CONCLUSION..... 87

REFERENCES..... 89

APPENDICES..... 96

A: The Annotation Tool..... 96

B: The DNA Sequencing result for PCR product done to confirm the primers designed..... 190

LIST OF TABLES

TABLES

Table 1.1: Show the main properties of MCF7 and MDA-MB-231 cell lines..... 6

Table 1.2: Sources of variation in a cDNA microarray experiment..... 13

Table 1.3: Show the differences between the most popular algorithm for preprocessing step in microarray analysis..... 20

Table 2.1: GST isozymes primers Designed using Primer-BLAST tool..... 36

Table 3.1: Samples legend.....39

Table 3.2: RNA properties, result from Nano drop and Bio analyzer..... 42

Table 3.3: Show the top 10 regulated genes between Emodin and Control classes, for the three type of analysis done..... 54

Table 3.4: Summary of the Annotation of the top 10 Genes which are differentially expressed among Emodin and Control Classes in analysis of 8 arrays; MDA and MCF cells.....55

Table 3.5: The Analysis of the predicted genes from network analysis result in figure 3.13 59

Table 3.6: Summary of the Annotation of the top 10 Genes which are differentially expressed among Emodin and Control Classes in analysis of 4 arrays; MCF cells..... 61

Table 3.7: The Analysis of the predicted genes from network analysis result in figure 3.15 65

Table 3.8: Summary of the Annotation of the top 10 Genes which are differentially expressed among Emodin and Control Classes in analysis of 4 arrays; MDA cells..... 68

Table 3.9: The Analysis of the predicted genes from network analysis result in figure 3.18 71

Table 3.10: The Top 10 pathways From DAVID Annotation Tool..... 73

Table 3.11: GST Isozymes regulation in MCF7 and MDA cells after Emodin treatment..... 76

Table 3.12: GST Isozymes regulation in MCF7 cells after Emodin treatment..... 76

Table 3.13: GST Isozymes regulation in MDA MB 231 cells after Emodin treatment..... 76

Table 3.14: The analysis of the predicted genes from network analysis result in figure 3.21 ..... .78

Table 3.15: The BLAST Alignment Search Result for the DNA Sequencing done for the PCR Products..... 86

Table A.1: Annotation of the top 10 Genes which are differentially expressed among Emodin and Control Classes in analysis of 8 arrays; MDA and MCF cell..... 96

Table A.2: The Analysis of the predicted genes from network analysis result in figure 3.13 106

Table A.3: Annotation of the top 10 Genes which are differentially expressed among Emodin and Control Classes in analysis of 4 arrays; MCF cells..... 117

Table A.4: The Annotation of the predicted genes from network analysis result in figure 3.15 127

Table A.5: Annotation of the top 10 Genes which are differentially expressed among Emodin and Control Classes in analysis of 4 arrays; MDA cells..... 144

Table A.6: The Annotation of the predicted genes from network analysis result in figure 3.18 157

Table A.7: Annotation for the GST isozymes genes..... 178

Table A.8: The Annotation of the predicted genes from network analysis result in figure 3.21 183

LIST OF FIGURES

FIGURES

Figure 1.1: Cancer development from normal cells to cancer..... 1

Figure 1.2: Oxidative metabolism of estrogen through the catechol pathway..... 2

Figure 1.3: Pathways for estrogen carcinogenesis..... 3

Figure 1.4: Estrogen receptor signal transduction pathway.....4

Figure 1.5: Comparing the single (a) and dual channel microarray (b)..... 8

Figure 1.6: A typical affymetrix genechip.....10

Figure 1.7: The synthesis of these oligonucleotides on GeneChip microarrays are based on the concept of photolithography..... 11

Figure 1.8: Good RNA sample quality measured with the Bioanalyzer.....15

Figure 1.9: Degraded RNA sample quality measured with bioanalyzer..... 16

Figure 1.10: An example of B2 illuminating the corner and edges of the array..... 21

Figure 1.11: The array name, located in the upper left or upper middle of the array..... 21

Figure 2.1: Genechip Eukaryotic labeling assays for expression analysis..... 31

Figure 3.1: Electrophoresis Run Summary..... 40

Figure 3.2: Detection of ribosomal RNA band by Agilent 2100 Bioanalyzer Machine..... 41

Figure 3.3: RNA degradation plot from arrayanalysis.org..... 43

Figure 3.4: 2D images for spatial bias diagnostic.....44

Figure 3.5: Density histogram of raw intensities and after RMA algorithm applied in the preprocessing step..... 46

Figure 3.6: Boxplot of raw intensity data..... 47

Figure 3.7: Boxplot after RMA..... 48

Figure 3.8: The PCA graph presents 3 plots: the array data are projected respectively on PC1 versus PC2, PC1 versus PC3 and PC2 versus PC3..... 50

Figure 3.9: Scatter plot, phenotype average between two classes Emodin and Control from the analysis of 8 arrays from both MCF7 and MDA MB 231.....52

Figure 3.10: Scatter plot, phenotype average between two classes Emodin and Control from the analysis of 4 arrays from MCF7 cells..... .52

Figure 3.11: Scatter plot, phenotype average between two classes Emodin and Control from the analysis of 4 arrays from MDA MB 231cells..... 53

Figure 3.12: Clustered heatmap of significantly expressed genes from the analysis of the 8 arrays of the both cells type; arrays grouped by class.....56

Figure 3.13: Network analysis done for the top 10 regulated genes result from the analysis of the 8 arrays of the both cells type..... 58

Figure 3.14: Clustered heatmap of significantly expressed genes from the analysis of the 4 arrays of MCF7 cells type; arrays grouped by class.....62

Figure 3.15: Network analysis done for the top 10 regulated genes result from the analysis of the 4 arrays of the MCF7 cells 64

Figure 3.16: The Effect of Emodin on MCF-7 cell lines, pathway drawn using Wiki Pathways, and need to be confirmed in the future studies.....66

Figure 3.17: Clustered heatmap of significantly expressed genes from the analysis of the 4 arrays of MDA-MB-231 cells type; arrays grouped by class..... 69

Figure 3.18: Network analysis done for the top 10 regulated genes result from the analysis of the 4 arrays of the MDA-MB-231 cells..... 70

Figure 3.19: The Effect of Emodin on MDA-MB-231 cell lines, pathway drawn using Wiki Pathways, and need to be confirmed in the future studies..... 72

Figure 3.20: Hierarchical cluster/ average linked between all sample and genes of the analysis of 8 arrays for both MCF7 and MDA-MB-231..... 74

Figure 3.21: Network analysis done for the our genes of interest GST enzyme family.....77

Figure 3.22: Toll-like receptor signaling Pathway..... 79

Figure 3.23: Signal transduction pathways.....80

Figure 3.24: Example of GST-mediated kinase regulation for GSTM1.....81

Figure 3.25: Other Example of GST-mediated kinase regulation for GSTP1..... 82

Figure 3.26: RNA Gel Electrophoresis..... 83

Figure 3.27: PCR Result on Gel Electrophoresis for MCF-7 cells..... 84

Figure 3.28: PCR Result on Gel Electrophoresis for MDA-MB-231 cells..... 85

LIST OF SYMBOLS AND ABBREVIATIONS

BRB	Biometric Research Branch array tool
cRNA	Complementary ribonucleic acid
DAVID	Database for Annotation, Visualization and Integrated Discovery
dNTP	Deoxy ribonucleotide triphosphate
ER	Estrogen Receptor
FDR	False discovery rate
GO	Gene ontology
GST	Glutathione-S- transferase
IQR	Inter-quartile range
IVT	In-vitro transcription
MAS 5.0	Microarray Suite 5.0
MCF-7	Michigan Cancer Foundation-7
MDA-MB-231	Monroe Dunaway Anderson- Metastatic Breast
mm	Mismatch
mRNA	Messenger RNA
pm	Perfect match
RMA	Robust multi-array analysis
RT PCR	Reverse transcriptase Polymerase chain reaction

CHAPTER 1

INTRODUCTION

1.1 BREAST CANCER

Cancer is a genetic disease that characterized by un controlled cells growth, Tumors have the ability to destroy adjacent tissue by invasion, and spread to other locations in the body by metastasis. Despite remarkable advances in early detection of cancer and new therapeutic strategies, over sixty percent of people that are diagnosed with cancer still die from it worldwide. (Song, 2008).

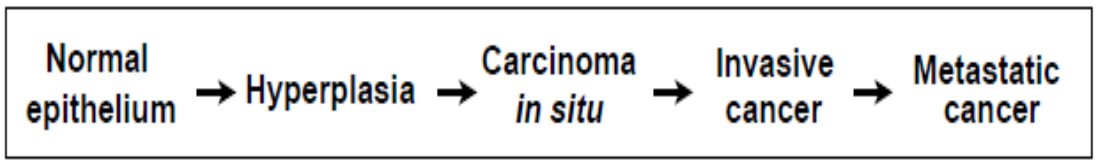


Figure 1.1: Cancer development from normal cells to cancer. (Song, 2008).

The type of cancer that originating from breast tissue called Breast cancer; worldwide, breast cancer comprises 22.9% of all cancers (excluding non-melanoma skin cancers) in women. In 2008, breast cancer caused 458,503 deaths worldwide (13.7% of cancer deaths in women). (http://en.wikipedia.org/wiki/Breast_cancer).

Some breast cancers are sensitive to hormones such as estrogen which makes it possible to treat them by blocking the effects of this hormone in the target tissues. These require less aggressive treatment than hormone negative cancers. Breast cancers without hormone receptors, are higher-risk, and are treated more aggressively.

1.1.1 ESTROGEN IN BREAST CANCER

Estrogen considered One of the major risk factors for breast cancer, There are three forms of estrogen circulating in our bloodstream; estrone, estradiol and estriol. Estrogen considered as carcinogenic; the carcinogenicity of estrogen has been linked to the role of catechol estrogens as carcinogenic metabolites. Quinones; the further oxidized metabolites, are the ultimate reactive electrophiles capable of DNA binding if not inactivated by glutathione conjugation (Mitrunen, et al., 2001). See Figure (1.2)

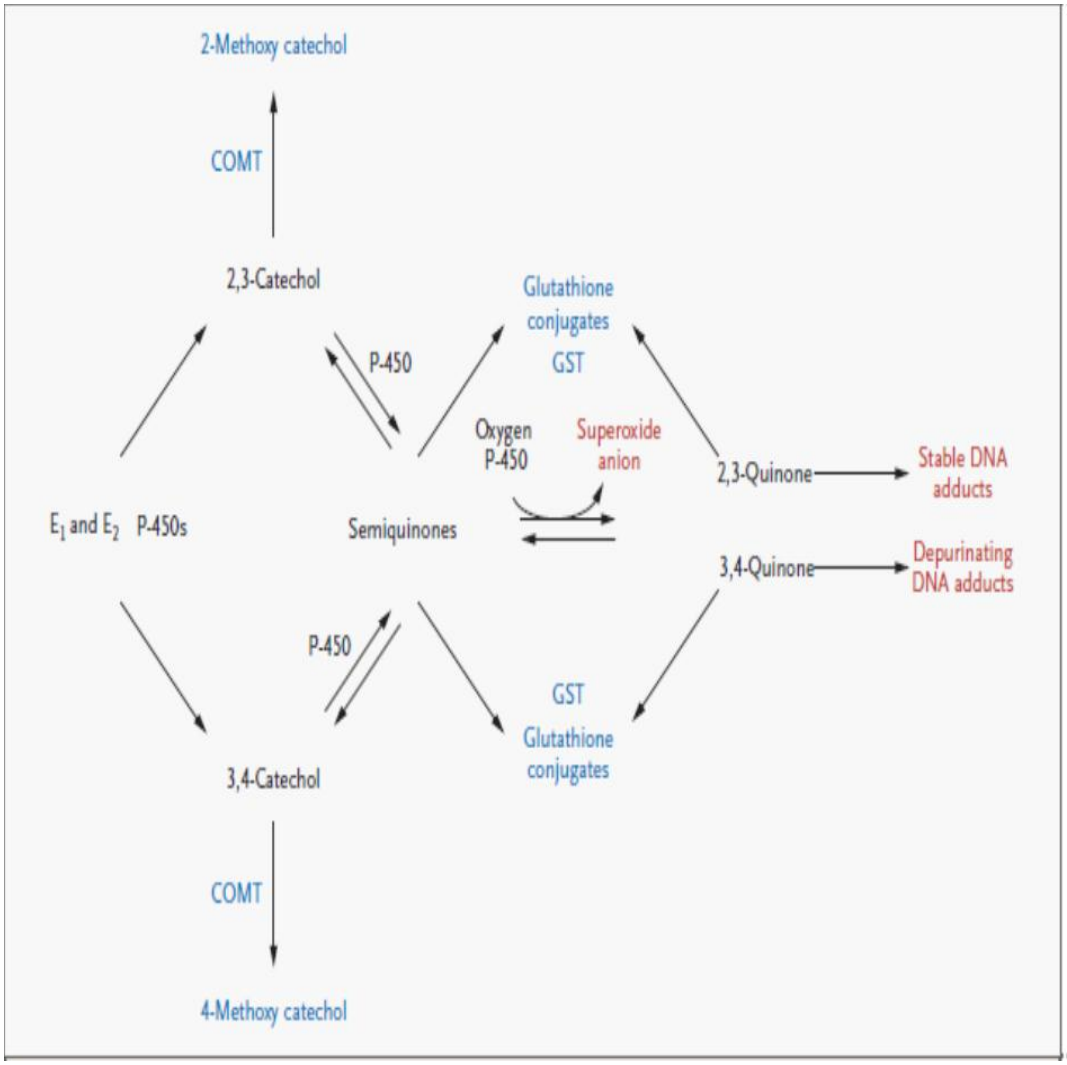


Figure 1.2: Oxidative metabolism of estrogen through the catechol pathway. COMT refer to catechol O-methyl transferase, P-450: cytochrome P-450, E1 refers to estrone and E2: estradiol. (Yagar and Davidson, 2006).

Figure 1.3, shows two different but complementary pathways play a role in estrogen carcinogenetic; one estrogen metabolism through the catechol pathway where estrogen 3, 4 Quinone can form unstable adducts with adenine and guanine in DNA. And the second through binding the estrogen to ER which lead to an alteration in gene expression, then alteration in apoptosis

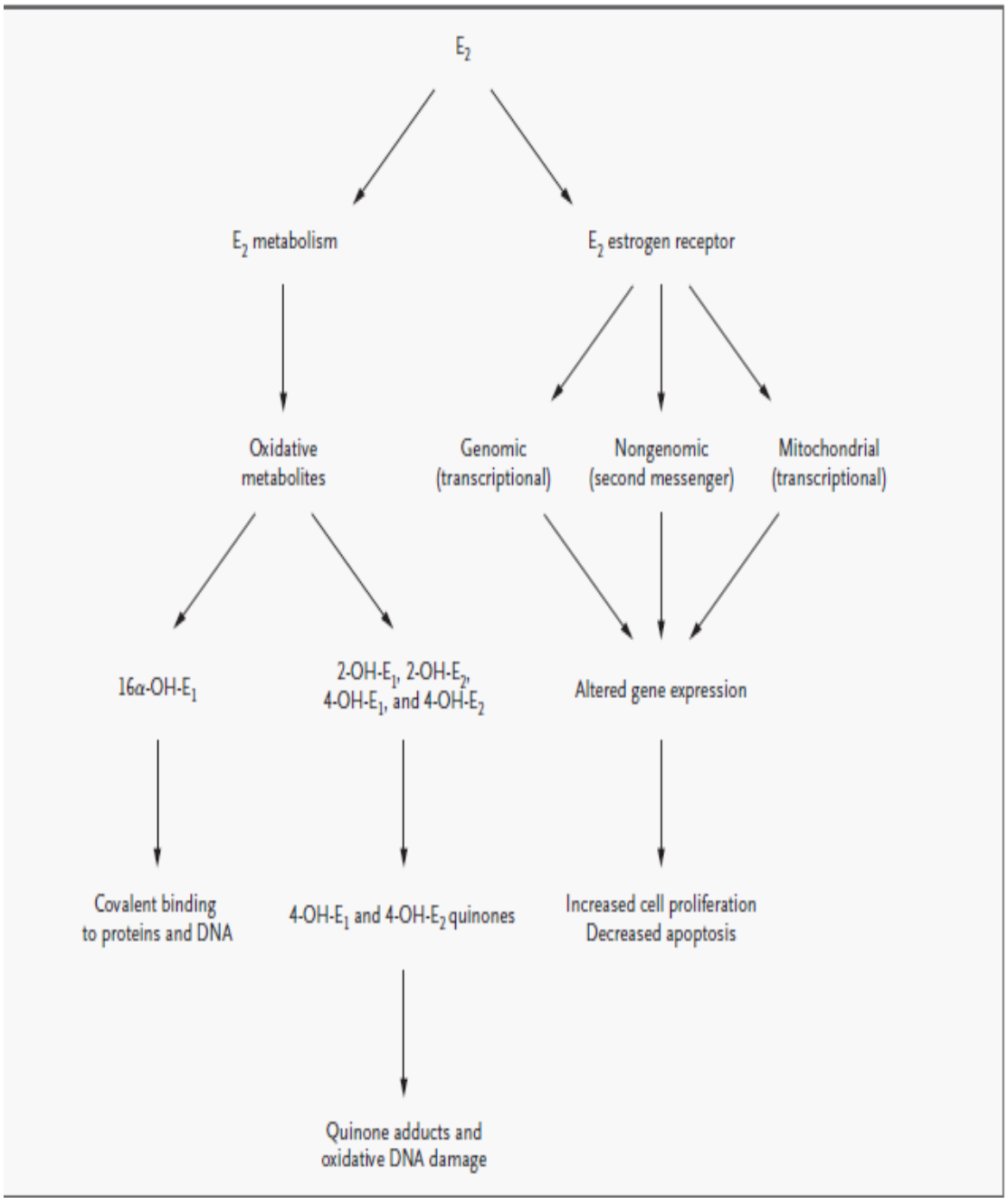


Figure 1.3: Pathways for estrogen carcinogenesis. E1 refer to estrone and E2: estradiol. (Yagar and Davidson, 2006).

Figure 1.4, summarizes the multiple estrogen receptor signal transduction pathways emphasizing effects associated with increased proliferation and inhibition of apoptosis. ER (estrogen receptors) are founded in many sites within the cell; nucleus, cytoplasm and membrane. When estrogen bind to ER in nucleus or cytoplasm through the mitochondria or through the membrane localized ER which bound to the growth factor receptor like tyrosine kinases all play a role in apoptosis.

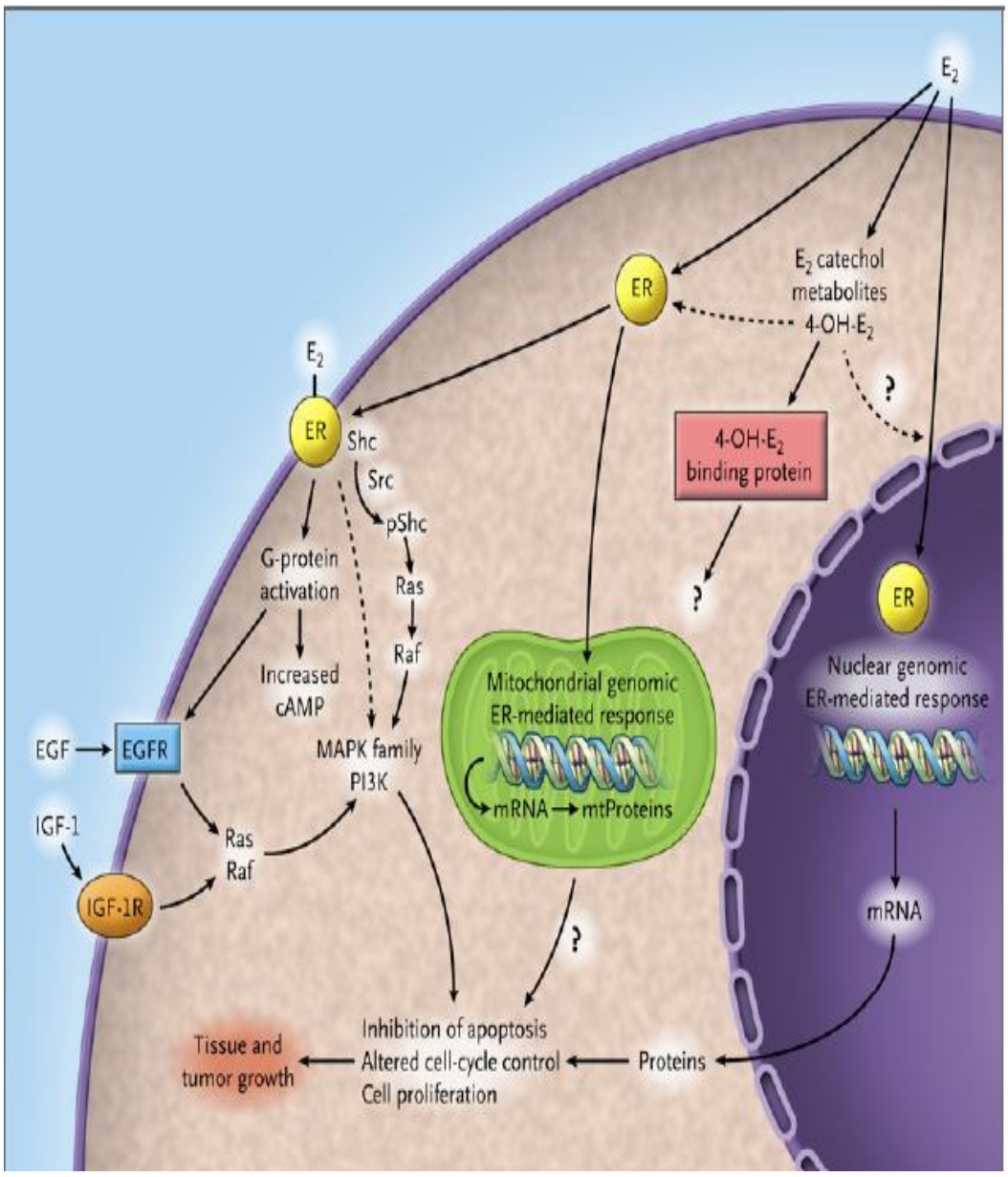


Figure 1.4: Estrogen receptor signal transduction pathway. (Yagar and Davidson, 2006).

1.1.2 GLUTATHIONE S-TRANSFERASE (GST)

Another important risk factor for causing breast cancer is exposing to environmental agents, the first line of defense is provided by the ability to metabolize and detoxify exogenous toxins .

Glutathione S-transferase (GST) genes are a superfamily of enzymes that are potentially important in regulating susceptibility to cancer because of their ability to metabolize reactive electrophilic intermediates to usually less reactive and more water soluble glutathione conjugates (Mitrunen, et al., 2001). See figure (2) the Oxidative metabolism of estrogen through the catechol pathway, GST play a role in inactivation of Quinones.

GST enzymes in mammals are subdivided into several classes. This is: Alpha, Mu, Pi, Theta, Sigma, Omega, Zeta and Kappa.

Glutathione-S-transferases (GSTs) are activated as a result of breast radiotherapy treatment; which is usually done after breast cancer surgery to kill the remain breast cancer cells, GSTs role is to detoxify the damage caused by oxidative stress which is a result of the radiotherapy.

1.1.3 EMODIN

Emodin or rheum (3- methyl-1, 6, 8-trihydroxy anthraquinone), an herb widely used as a laxative in traditional Chinese medicine. (Huang, et al., 2008A). Emodin is a phytoestrogen component extracted from rheum (genus) plant; has been reported to supress the growth of tumor in some clinical situation. And effectly inhibit tumor metastasis in vitro and in vivo (Huang, et al., 2008A). Some studies investigated the effect of emodin on cell death and on the pathways that lead to apoptosis, they found that emodin ınduced apoptosis through the decrease of Bcl-2/Bax ratio and the increase of cytoplasm cytochrome c concentration in human breast cancer Bcap-37 cells (Huang, et al., 2008B). Some other studies investigated the effect of emodin as a Tyrosine Kinase İnhibitor suppresses the growth of HER-2/neu overexpressing in breast cancer cells by western blot analysis and Immunohistochemical assays methods (Lisha,et al.,1999).

The new studies started to use the Microarrays to study the effect of emodin in cancerous cells, an apoptosis associated cDNA microarray comprised of 458 human apoptosis related genes to determine the impact of emodin in breast cancer Bcap-37 cells; they found that the gene expression profiling was altered when exposed to emodin, but has no effect on caspases. It's that P53 pathway may cooperate with IGF-2 pathway resulting in emodin induced apoptosis through the disruption of the mitochondrial signaling pathway (Hang, et al., 2008B). In this project we want to use microarray technique to study the effect of emodin on gene expression profiling of MCF7 and MDA-MB-231231 cancerous cell line.

1.1.4 MCF-7 AND MDA-MB-231 CELL LINES

MCF-7 and MDA-MB-231 is breast cancer cell lines, most of breast cancer researches use these cell lines as in vitro work. MCF-7 cell line was first isolated in 1970 from the breast tissue of a 69-year old Caucasian woman, the ability of MCF-7 cells to process estrogen, in the form of estradiol, via estrogen receptors in the cell cytoplasm. This makes the MCF-7 cell line an estrogen receptor (ER) positive control cell line. (http://mcf7.com/).

The MDA-MB-231 breast cancer cell line was obtained from a patient in 1973 at M. D. Anderson Cancer Center.
(http://www.cellbiolabs.com/sites/default/files/2FD8C527-3048-812A-2EB950C744EB9D73.pdf) .

Table 1.1 shows the main properties of MCF7 and MDA-MB-231 cell lines.

Cell line	**primary tumor**	**Origin of cells**	**Estrogen receptors**	**Tumorigenic in mice**
MCF-7	Invasive ductal carcinoma	Metastasis (plueural effusion)	Yes	Yes (with estrogen supplementation)
MDA_MB_23	Invasive ductal carcinoma	Metastasis (plueural effusion)	No	No

1.1.5 CANCER STUDIES BASED ON MICROARRAY ANALYSIS

The main goal of cancer research is to identify significant genomic alterations responsible for the initiation and progression of the disease. Gene expression profiling using DNA microarray data enables researchers to monitor the genome at the transcriptional level in cancer cells (Song, 2008), it provides a great opportunity to understand the disease at the molecular level, many microarray based study compared normal and diseased tissues, some drug discovery studies search the effects of drug, chemicals or *etc*. on the cancer cells. Microarray helps the researcher in cancer field to understand the disease mechanism and to find the correct treatment.

1.2 MICROARRAY

Microarray is A high throughput technology that allows detection of thousands of genes simultaneously; it's An arrangement of DNA sequences on a solid support where matching of known and unknown DNA samples is done based on base pairing rules.

Thousands of spotted samples which called probes (with known identity) are immobilized on a solid support (a microscope glass slides or silicon chips or nylon membrane). Probes could be DNA, cDNA (PCR product) or oligonucleotide, the difference between them and the difference in manufacturing them lead to the different microarray technologies exist today.

Robotic spotting and in-situ synthesis is the main technologies for making a microarray. İn the robotic spotting DNA probes are synthesized and then is spotted on to the microarray, Spotting is easy to automate but may generate poor quality spots (irregular spots of different shapes and sizes). While in-situ synthesis the DNA oligonucleotides are synthesized directly on the microarray by using Photolithography technology.

Another difference between the microarray platforms is the different in detection types (or labeling methods) for the microarray; single or dual channel microarrays see Figure (1.5); the Two-color detection microarray Compare two samples by labeling with two different flurophores and analyzing on the same array Cy5 (red)

and Cy3 (green), analyzing the Relative intensities of each fluorophore is the method to identify up-regulated and down-regulated genes. The One-color detection; Determine gene expression level and need one array per sample, the analysis method is to compare the samples to find the regulated genes.

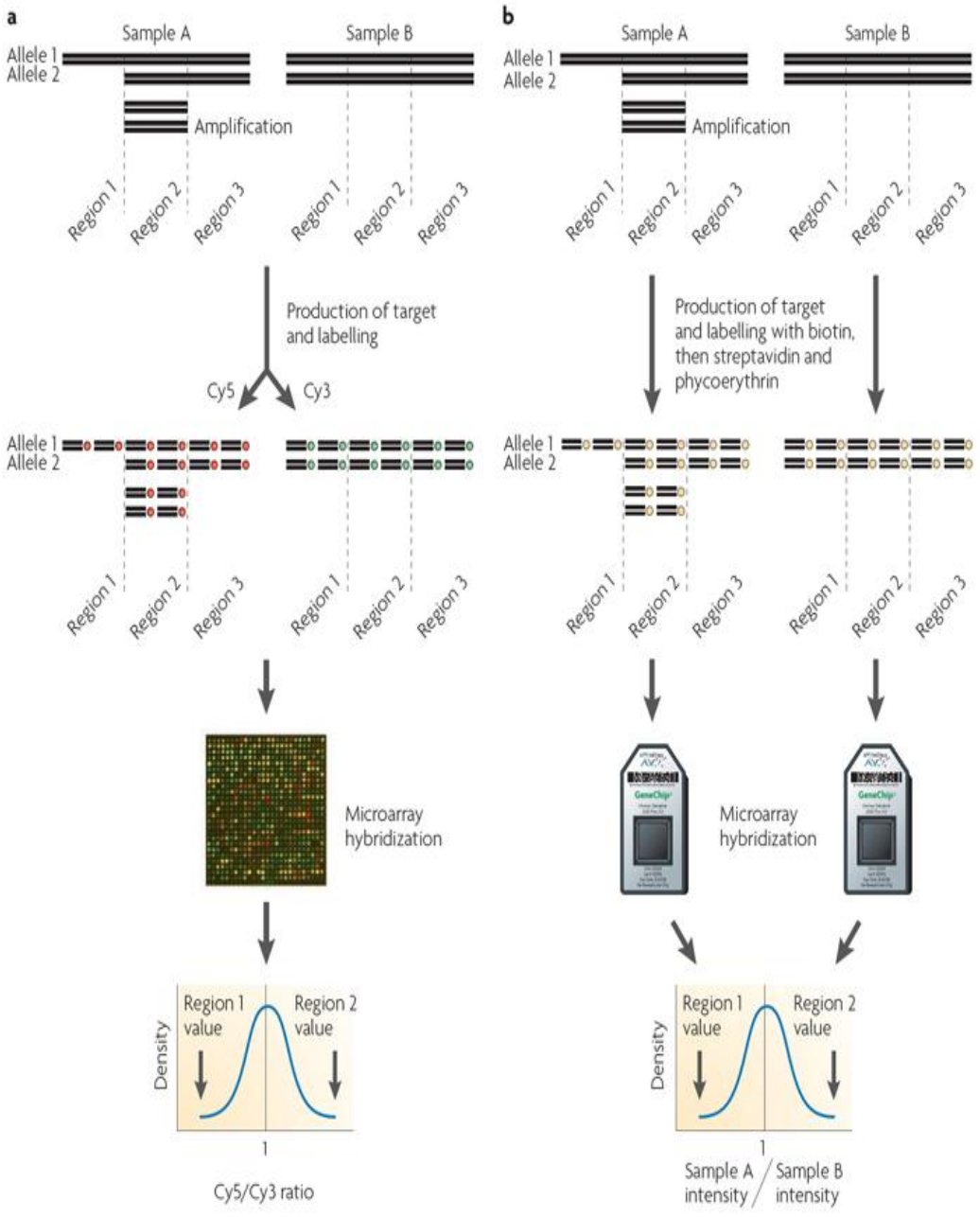


Figure 1.5: Comparing the single (b) and dual channel microarray (a). (http://www.nature.com/scitable/content/outline-of-a-typical-microarray-experiment-20729).

The huge volume of the data generated from microarray experiment cause the need for bioinformatics, which help in the interpretation of the results using statistical and mathematics methods.

Microarray have many application in gene expression studies, disease diagnosis, pharmacogenomics (drug discovery) and toxicogenomics .

1.2.1 HISTORY OF MICROARRAY EVALUIATON

Microarray technology evolved from the immunoassay technologies; immunoassays are a biochemical test that measures the presence or concentration of a substance in solutions that frequently contain a complex mixture of substances. Then the combining between labeling techniques and immunoassays like in enzyme- linked immunosorbent assay (ELISA); where the fluorescent labeling is either the antibody or the antigen, which commonly used to detect antibodies in the blood. These bring the idea for microarray. For example; The idea of the attachments of the antibodies to a solid support and that it depend on the specificity of target molecules binding to the antibody lead to the thought of using it in DNA analysis, as the DNA consist of two complementary strands of nucleotides.

The southern blot was the first array of genetic material, in this technique; fragmented DNA is bound to a substrate (often a nitrocellulose or nylon membrane), and then probed with a known gene or fragment.

Then the First use of microarray to profile gene expression was published in 1995, after that in 1997 a complete eukaryotic genome (*Saccharomyces cerevisiae*) spotted on a microarray was published. In the Early 2000s they spotted thousands of spots on a chip.

Now High density arrays can print up to 6 million spots on one chip.

1.2.2 GENE EXPRESSION MICROARRAY

In gene expression microarrays, either oligonucleotides or cDNA fragments have been used as probes. Because of this it's called DNA microarrays , commonly known as gene chip, DNA chip, or biochip.

DNA microarrays are assays for quantifying the types and amounts of mRNA transcripts present in a collection of cells. The number of mRNA molecules derived from transcription of a given gene is an approximate estimate of the level of expression of that gene (Simon *et al.*, 2004).

Experiment the expression levels of thousands of genes are simultaneously monitored to study the effects of certain treatments, diseases, and developmental stages on gene expression.

1.2.2.1 AFFYMETRIX GENE EXPRESSION MICROARRAYS

Affymetrix GeneChip arrays have oligonucleotide probes lithographically synthesized directly on the array. The array in this case is not a glass slide, but a silicon chip.

Figure 1.6: A typical affymetrix genechip.

The oligonucleotides at all locations on the chip are synthesized in parallel. At the first step, the chip is bathed in a solution containing a precursor to one of the four

nucleotides, then the oligo nucleotide constructed by the concept of photolithography (Figure 7); a mask is employed to ensure that light reaches only those addresses where the next nucleotide in the desired sequence is that represented by the current bath. The in situ synthesis continues in this manner with multiple baths and washes. (Simon *et al*, 2004).

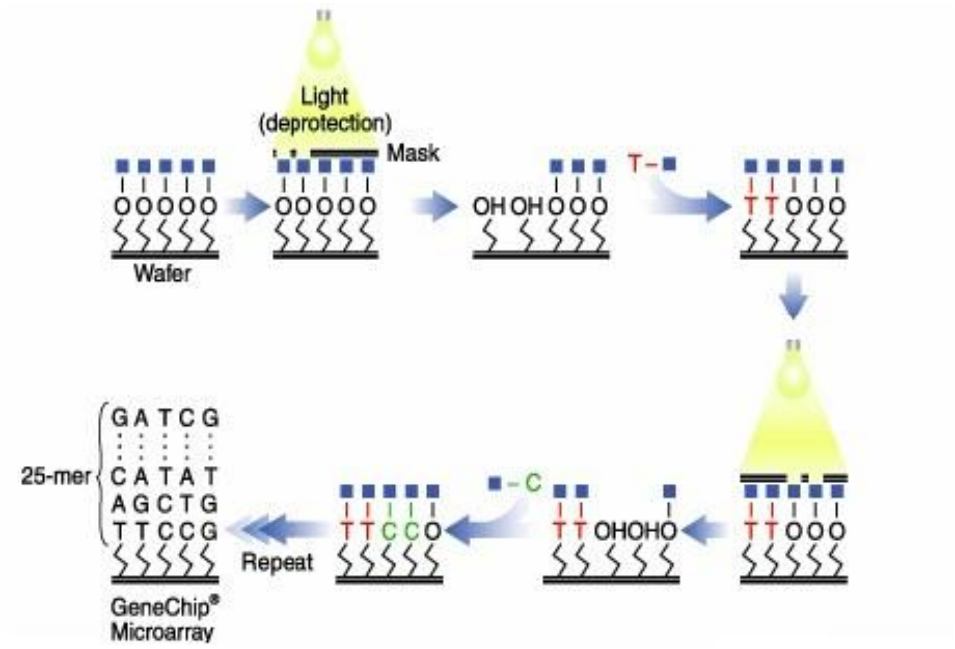


Figure 1.7: The synthesis of these oligonucleotides on GeneChip microarrays are based on the concept of photolithography. (German Cancer Research Center, http://www.dkfz.de/gpcf/24.html)

1.2.2.2 AFFYMETRIX PROBES

Affymetrix gene chip consist of 25 mer oligonucleotides probes; for each transcript a set of probes called (probe set) is designed, typically consist of 11 different probes, some of the probes are unique portion from the gene completely complementary to the mRNA transcript, these probes called perfect match(PM), other oligonucleotides the same as the (PM) but have a different base in the 13th position, this kind of probes called mismatch (MM), its serve as controls for specific hybridization and help in subtraction of the background and cross hybridization signals. Together the PM probes with its own MM probe are called probe pair. All probes of one probe set are distributed along the chip.

1.2.3 EXPERIMENTAL DESIGN

Microarray experiment Design is one of the important issues because if it's designed well, it will give more significant Result with a minimum cost. To Design a microarray experiment we must have well defined goals, expect the technical source that could result in variation of the results.

First issue in the experiment design is to identify the experiment aim; what questions it will answer and at the first it must answer one direct question this is called pilot experiment which usually focus on single variable with a control, and after the result of the pilot experiment we can decide the next variables must be studied and how many replicate needed for the next microarray experiment. This is will minimize the array number and help in the data analysis.

Experiment design also include determination of the sample type, and the biological material needed, how many replicate needed and the microarray platform will be used; Replicate number and type (biological or technical or both) is depend on the expected change in the gene expression level, the variables number, the quality of the sample and the technical noise expected, and choosing the microarray platform depend in the budget and the aim of the experiment.

Other important issue in the designing is sample pooling; some time its needed especially if there's no enough yield of mRNA, sample pooling can decrease noise if individual variation is not of interest

1.2.4 DNA MICROARRAY LIMITATION

Microarray measure the gene expression levels by measuring the signal intensities of the fluorescently labeled RNA of the sample, so this indirect way of measuring could lead to a variation of the result. The variation of the result could be occurred because of an error or problem in the stage before the microarray experiment; for example in the tissue culture stage or the RNA isolation, or during the microarray fabrication, and also in the microarray experiments during labeling or hybridization or scanning, and in the analysis of the result during the preprocessing level and depend on the powerful of the statistical test chose. (Table 1.2) show the source of variation in a cDNA microarray experiment.

Table 1.2: Sources of variation in a cDNA microarray experiment.(wit and McClure, 2004).

Sources of variation in the mRNA: Differences in conditions. Differences between experimental subjects within the same covariate level. Differences between samples from the same subject. Variation in mRNA extraction methods from original sample. Variations in reverse transcription. Differences in PCR amplifications. Different labeling efficiencies.
Sources of variation in the microarray production: Print-pin anomalies. Variation in printed probe quantities even with the same pin. Chip batch variation (due to many sources of unknown variations). Differences in sequence length of the immobilized DNA Variations in chemical probe attachment levels to the slide.
Sources of variation in the hybridization process: Different dye sensitivities. İnequalities in the application of mRNA to the slide. Variations in the washing efficiencies of non-hybridized mRNA off the slide. Other differences in hybridization parameters, such as: • Temperature. • Experimenter. • Time of the day.
Sources of variation in the scanning: Different scanners. Different photo-multipliers or gain. Different spot-finding software. Different grid alignments.

1.2.5 CONTROLING NON-BIOLOGICAL VARIANCE

The aim of microarray experiment is to identify the biological variation in gene expression, so to standardize the experiment it's important to control the non-biological variance and trying to decrease them. standardization of the system hybridization, washing, staining, scaling and also the quality controls which built during the manufacturing processes will result to neglect the system noise; for example to run all RNAs on the same day, using reagent from the same lots, preparing reagent master mix and one scientist to do the bench work. Other important things to be done in the microarray laboratory is that all equipment used

in the experiment should be calibrated regularly to ensure accuracy. The next thing to be considered to control sample preparation variability is RNA isolation.

1.2.5.1 RNA QUALITY ASSESMENT

All RNA samples should meet assay quality standards to ensure the highest quality RNA is hybridized to the gene expression arrays. Researchers should run the initial total RNA on an agarose gel and examine the ribosomal RNA bands. Non-distinct ribosomal RNA bands indicate degradation which can lead to poor dsDNA synthesis and cRNA yield. A 260/280 absorbance reading should be obtained for both total RNA and biotinylated cRNA. Acceptable A260/280 ratios fall in the range of 1.8 to 2.1, Ratios below 1.8 indicate possible protein contamination. Ratios above 2.1 indicate presence of degraded RNA, truncated cRNA transcripts, and/or excess free nucleotides. (genechip expression analysis, Data analysis fundamentals by affymetrix; 2002-2004).

The quality of total RNA can also be measured by the Bioanalyzer. A good quality sample should have 18S and 28S peaks that look like the image in Figure (1.8). The graph should have a low baseline and sharp ribosomal peaks. A good quality sample will typically have a ratio of 28S:18S ribosomal peaks of 2:1, however, this can be sample dependent. (genechip expression analysis, Data analysis fundamentals by affymetrix; 2002-2004).

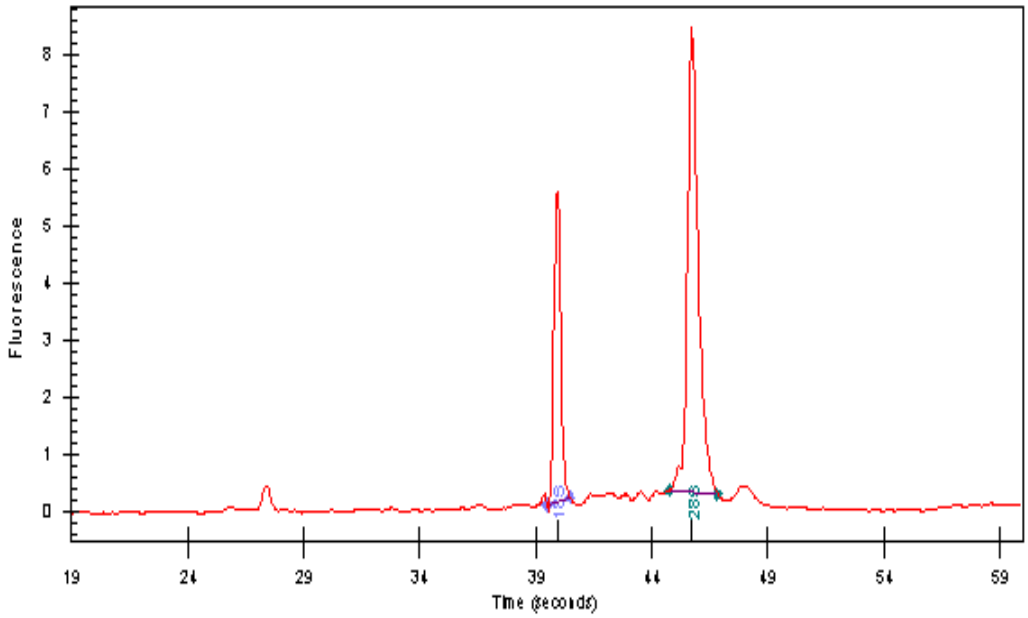


Figure 1.8: Good RNA sample quality measured with the Bioanalyzer. (genechip expression analysis, Data analysis fundamentals by affymetrix; 2002-2004).

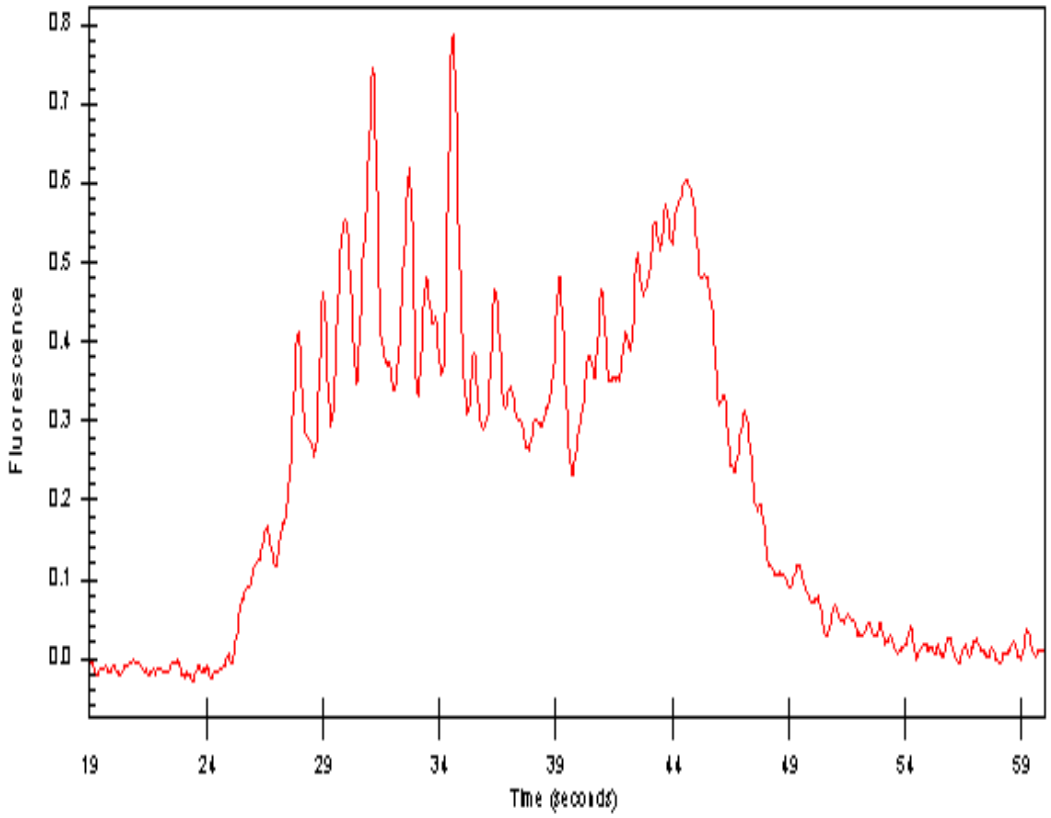


Figure 1.9: Degraded RNA sample quality measured with bioanalyzer. (genechip expression analysis, Data analysis fundamentals by affymetrix; 2002-2004).

1.2.6 DATA ANALYSIS

Data analysis start when the scanner gives the microarray image then using computer algorithms converting the image to numerical data that quantifies gene expression.

1.2.6.1 IMAGE ANALYSIS

When using affymetrix platform image analysis is simple procedure because it's all standardized and the chip and the scanner all automated to give high quality result.

After image acquisition, affymetrix software for example; GCOS will identify the position of every probe location by placing a grid on top of the scanned image. B2 Oligo serves as positive hybridization control which found on the top corner of each array, and used by the software to place a grid over the image. After the grid is adjusted and the location of each probe is identified, a square set of pixels is identified for each probe, from this set of pixels an overall of signal value is calculated; the outer ring is discarded as its considered unreliable cause it may be affected from the neighbor probes, then of the remaining pixels, the pixel intensity of the 75th percentile is calculated, which is the probe cell intensity that reported in CEL. File (Göhlmann and talloen; 2009). There's different methods for Calculating signal intensity: mean (pixel intensities), median (pixel intensities) or Pixel variation IQR of log (pixel intensities).

1.2.6.2 PREPROCESSING

Preprocessing step is important to get out with reliable microarray result; it's done to remove the non-biological variation to end up with the true gene expression variation. Another reason is to transform the data into a format that is ready to be analyzed.

1.2.6.2.1 LOG2 TRANSFORMATION

Log2 transformation is the first step in preprocessing, where microarray data transforming the values to log scale usually its Logs base 2 because it's easiest for people. Its benefit in the analysis that its Make the variation of intensities and ratios of intensities more independent of absolute magnitude and the distribution is approximately normal so a SD can be calculated, Gives a more realistic sense of variation so make the result easy to analyze.

1.2.6.2.2 FILTERING SIGNAL

Filter signaling applied by doing set of filters, spots with very weak signals, high background or uneven signals are removed to not be included in the analysis or just flagged so after the analysis to make sure of the result the spot intensity quality checked.

1.2.6.2.3 BACKGROUND CORRECTION

Background correction done to remove the effect of the non-biological variance in the measured signal, Background noise come from for example; unspecific binding of transcript, back ground signal from incomplete washing of the microarray and optical noise from the scanner.

MM probes are designed to give a measure for non-specific binding of their corresponding PM probe. So the MM values should be subtracted from their corresponding PM values as a first step in the analysis process, but many of the known preprocessing methods solve this problem by simply ignoring the MM probes altogether and PM values are corrected for non-specific binding using other methods like:

- No correction

- Constant Background: done by subtract a constant background for all spots, it ignores the variability of background estimates among individual spots.

- Local Background correction: estimates the background with pixels in a fixed area around the spot , or take the median of the surrounding areas.

- Morphological opening: first the pixel in the center of the window is replaced by the minimum value of all pixels in the window. The next time each pixel is replaced by the maximum value. The window is chosen to be larger than the feature diameter so that spots will disappear in the transformed image. (Laurell; 2006).

1.2.6.2.4 NORMALIZATION

The main goal of normalization to remove the systematic bias in the data as possible, while keeps the variation in gene expression that occurs because of biologically changes in transcription.

A basic assumption of most normalization procedures is that the average gene expression level does not change in an experiment; the change in some biological factors will lead to a small change in the gene expression which in general will not affect the average value. Simply, normalization ensures that when comparing expression levels of different arrays, that we are comparing two thing like each other.

There are many normalization methods but the most two methods used in single channel microarrays is:

- Global scaling: is done on the probeset level
- Quantile normalization: consider the quantiles of each chip are equal, so it gives same distribution to each chip.

1.2.6.2.5 SUMMARIZATION

Calculating gene expression values; by reduce the 11-20 probe intensities for each probeset on to a gene expression value.

1.2.6.2.6 DIFFRENT PREPROCESSING ALGORITHMS AVAILABLE

Different algorithms will use different methods of background correction, normalization and summarization the most popular algorithms is:

MicroArray Suite 5.0: algorithm developed by Affymetrix.

Robust Multi-Array Analysis (RMA) : academic alternative to Affymetrix's algorithms for converting probe level data to gene expression measures. And table 1.3 shows the difference between them.

Table 1.3: Shows the Differences between the most popular algorithm for preprocessing step in microarray analysis.

Algorithm	Citation	BACKGROUND SUBTRACTION	NORMALIZATION	PROBE SUMMARIZATION
RMA	Irizarry et al. 2003	PM based	Quantile	Log (PM)
CGRMA	Wu and Irizarry. 2004	PM-MM based	Quantile	Log (PM)
MAS 5.0	MAS 5 Affymetrix 2002.	PM-MM based	Scaling	One-step Tukey Biweight

1.2.6.3 ASSESSMENT OF ARRAY QUALITY

Assessment of array quality done to judge quality of chip data. By controlling the control parameters in the chip; poly-A labeling controls, hybridization controls and internal control genes (housekeeping genes) on the arrays, and produced graphical outputs of normalized signal histograms, box plots and probe cell intensities from the data.

In Affymetrix chip there are several controls which allow to monitor data quality; B2 Oligo performance, Poly-A Controls, Hybridization Controls, Internal Control Genes.

1.2.6.3.1 B2 OLIGO PERFORMANCE

B2 Oligo serves as a positive hybridization control and is used by the software to place a grid over the image. Figure (10) and figure (11) shows the image must see in controlling B2 oligo.

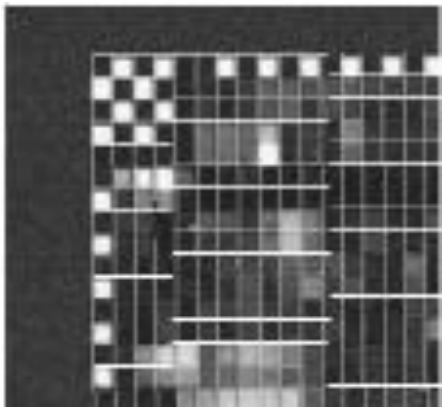


Figure 1.10: An example of B2 illuminating the corner and edges of the array. (genechip expression analysis, Data analysis fundamentals by affymetrix; 2002-2004).

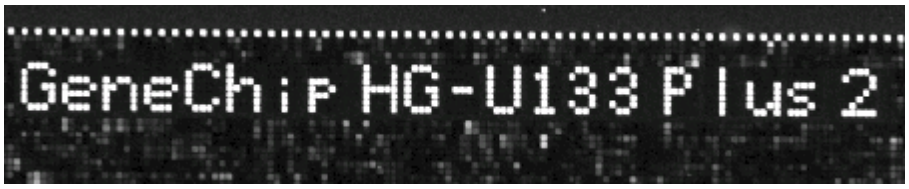


Figure 1.11: The array name, located in the upper left or upper middle of the array. (genechip expression analysis, Data analysis fundamentals by affymetrix; 2002-2004).

1.2.6.3.2 POLY-A CONTROLS: lys, phe, thr, dap

Poly-A RNA controls can be used to monitor the entire target labeling process. Dap, lys, phe, thr, and trp are B. subtilis genes that have been modified by the addition of poly-A tails, and then cloned into Bluescript vectors, which contain T3 promoter sequences. Amplifying these poly-A controls with T3 RNA polymerase will yield sense RNAs, which can be spiked into a complex RNA sample, carried through the sample preparation process, and evaluated like internal control genes. The GeneChip Poly-A RNA Control Kit (P/N 900433) contains a pre-synthesized mixture of lys, phe, thr, and dap. The final concentrations of the controls, relative to the total RNA population, are: 1:100,000; 1:50:000; 1:25,000; 1:7,500, respectively. All of the Poly-A controls should be called "Present" with increasing Signal values in the order of lys, phe, thr, dap. (genechip expression analysis, Data analysis fundamentals by affymetrix; 2002-2004).

1.2.6.3.3 HYBRIDIZATION CONTROLS: bioB, bioC, bioD, and cre

BioB, bioC and bioD represent genes in the biotin synthesis pathway of E. coli. Cre is the recombinase gene from P1 bacteriophage. The GeneChip® Eukaryotic Hybridization Control Kit (P/N 900299 and 900362) contains 20x Eukaryotic Hybridization Controls that are composed of a mixture of biotin-labeled cRNA transcripts of bioB, bioC, bioD, and cre, prepared in staggered concentrations (1.5 pM, 5 pM, 25 pM, and 100 pM final concentrations for bioB, bioC, bioD, and cre, respectively). The 20x Eukaryotic Hybridization Controls are spiked into the hybridization cocktail, independent of RNA sample preparation, and are thus used to evaluate sample hybridization efficiency on eukaryotic gene expression arrays. BioB is at the level of assay sensitivity (1:100,000 complexity ratio) and should be called "Present" at least 50% of the time. BioC, bioD, and cre should always be called "Present" with increasing Signal values, reflecting their relative concentrations. (genechip expression analysis, Data analysis fundamentals by affymetrix; 2002-2004).

1.2.6.3.4 INTERNAL CONTROL GENES

For the majority of GeneChip expression arrays, β-actin and GAPDH are used to assess RNA sample and assay quality. Specifically, the Signal values of the 3' probe sets for actin and GAPDH are compared to the Signal values of the corresponding 5' probe sets. The ratio of the 3' probe set to the 5' probe set is generally no more than 3 for the 1-cycle assay.

1.2.6.4 DIFFRENTIALLY EXPRESSED GENES

A microarray experiment aims to identify the gene expression differences which occurred between the biological conditions examined. To choose the differentially expressed list from the result which is for example, in affymetrix gene chip nearly 54000 probe, after gene filtering in the preprocessing step the gene number will decrease according to the filter options chose, then the fold changes calculated, p-values are calculated for each gene present on the microarray by using the t-test or some other analytical strategies such as the ANOVA, which helps to estimate the contribution of experimental factors to the distribution of the measured gene

expression . Next, a cut-off is found to separate the differentially expressed genes from the genes whose expression is not changed. This cut-off is usually based on a multiple testing criterion such as the Bonferroni or the false discovery rate. (Benjamini and Hochberg 1995). (Greco;2009). According the result the differentially expressed gene list been chose with genes have significant p-value and fold change.

1.2.6.5 CLUSTERING AND CLASSIFICATION

Clustering is performed to divide the massive amounts of gene expression data into groups based on similarity. This can be accomplished with two different strategies used for two different purposes; unsupervised clustering for exploratory analysis, and supervised clustering, which can be used to create a diagnostic device based on gene expression signatures. (Laurell;2006).

Unsupervised clustering: Unsupervised clustering is a way of obtaining a more comprehensible representation of the data set. With reduction techniques, for example; principal component analysis (PCA), singular value decomposition (SVD), it allows to visualize the data in two or three dimensional space.

Clustering can be performed either on genes or samples, or on both, showing relationships between genes and samples for pattern discovery. the most common methods for exploratory grouping, are hierarchical clustering

Hierarchical clustering is an agglomerative method; where building up the branches of a tree, beginning with the two most closely related objects which produce a dendogram with a bottom-up structure. Distance between clusters can be calculated by using the minimum, maximum or the average distance between samples in two different clusters.

Supervised classification: Supervised classification on the other hand is performed to create a tool which can be used for discrimination of new data. Supervised clustering techniques thus needs a data set with information of the sample labels, for example; if a tumor is malignant or not. The aim is to create a classifier which can classify new unlabeled samples and genes based on the microarray data. These can be divided into machine learning algorithms like support vector machines

(SVM), ANN and k-nearest neighbors (KNN), and statistical linear discriminate analysis. (Laurell;2006).

1.2.6.6 EXTRACTION OF BIOLOGICAL INFORMATION

To understand the biological reasons why genes appear as differentially expressed, Annotation must be done. Some examples of biological annotations tools is: gene ontology (GO) annotation tool, Kegg pathway (Kyoto Encyclopedia of Genes and Genomes), David (The Database for Annotation, Visualization and Integrated Discovery.

1.2.7 NETWORK ANALYSIS USING GENEMANIA TOOL

The GeneMANIA Cytoscape plugin brings fast gene function prediction capabilities to the desktop. GeneMANIA identifies the most related genes to a query gene set using a guilt-by-association approach. The plugin uses over 800 networks from six organisms and each related gene is traceable to the source network used to make the prediction. Users may add their own interaction networks and expression profile data to complement or override the default data. (Montojo, 2010).

1.2.8 SOFTWARES FOR DATA ANALYSIS

1.2.8.1 BRB (BIOMETRIC RESEARCH BRANCH)-ARRAY TOOL

Developed by: Richard Simon & BRB-ArrayTools Development Team, BRB-ArrayTools is an integrated package for the visualization and statistical analysis of DNA microarray gene expression data. It was developed by professional statisticians experienced in the analysis of microarray data and involved in the development of improved methods for the design and analysis of microarray based experiments. The array tools package utilizes an Excel front end. Scientists are familiar with Excel and utilizing Excel as the front end makes the system portable and not tied to any database. The input data is assumed to be in the form of Excel spreadsheets describing the expression values and a spreadsheet providing user-specified phenotypes for the samples arrayed. The analytic and visualization tools are integrated into Excel as an add-in. The analytic and visualization tools themselves are developed in the powerful R statistical system, in C and Fortran programs and in

Java applications. Visual Basic for Applications is the glue that integrates the components and hides the complexity of the analytic methods from the user. The system incorporates a variety of powerful analytic and visualization tools developed specifically for microarray data analysis. (http://linus.nci.nih.gov/BRB-ArrayTools.html).

1.2.8.2 GENESPRING GX SOFTWARE

Agilent's GeneSpring GX software provides powerful, accessible statistical tools for fast visualization and analysis of transcriptomics, genomics, proteomics and metabolomics data. Designed specifically for the needs of biologists, GeneSpring GX offers an interactive desktop computing environment that promotes investigation and enables understanding of microarray data within a biological context. (http://www.genomics.agilent.com/CollectionSubpage.aspx?PageType=Product&SubPageType=ProductDetail&PageID=1675).

1.3 THE AIM OF THE STUDY

The aim of our study is to investigate the effect of emodin on MCF7 (ER +) and MDA-MB-231 (ER -) breast cancer cell lines by microarray. Emodin which is a phytoestrogen component extracted from rheum (genus) plant; has been reported to suppress the growth of tumor, and effectively inhibit tumor metastasis in vitro and in vivo (Huang , et al.,2008(A)).

- We have examined the changes in gene expressions by using BRB-Array Tools (v.4.2.0) in order to decipher the mechanism of action of emodin on the cell lines, 10 genes exhibiting highest variation in expression were examined.

- Since glutathione S-transferase enzymes are potentially important in regulating susceptibility to cancer because of their ability to metabolize reactive electrophilic intermediates to usually less reactive and more water soluble glutathione conjugates (Mitrunen, et al., 2001) involved in xenobiotic metabolism, we specifically concentrated on the comparative changes in the gene expression of GST isozymes. The primers for each isozyme have been designed using primer-BLAST Tool. These primers were used to show the presence of specific isozymes in both cell lines separately.

CHAPTER 2

MATERIALS AND METHODS

2.1 EXPERIMENT DESCRIPTION

2.1.1 CELL CULTURE

Detailed information about cell culture in: (Sakally, E. Comparative Effects of Emodin on Biological Activities of MCF-7 and MDA-231 Cell Lines. M.S. Thesis; METU, 2010).

MCF and MDA-MB-231 From ATC American type culture collection, 500,000 cell/well Cultured in 6 well plate after 24 hr medium changed and half of them treated with 10 µg/ml Emodin the other treated as control with 1% DMSO. Incubated for 48 hr. At 37 ^{8}C and 5 % Co2 Incubator.

2.1.2 BIOLOGICAL REPLICATES

We had two biological replicates; cultured and treated at the same time;
1st biological replicate:

MCF DMSO control
MCF Emodin
MDA DMSO control
MDA Emodin

2nd biological replicate:

MCF DMSO control

MCF Emodin
MDA DMSO control
MDA Emodin

2.2 RNA ISOLATION

RNA Isolated using QIAGEN RNAeasy spin column mini kit. (500.000 cells/well) of MCF7 and MDA-MB-231l were placed in 6 well plate. For control cell's isolation, two spin columns were used 12 wells (from the 6 well plate wells) , and for treated cells two columns were used for 18 wells (from the 6 well plate wells) for each Replicate.

RNA concentration measured using nano drop and detection of ribosomal RNA band checked by Agilent 2100 Bioanalyzer Machine.

2.3 MICROARRAY EXPERIMENT

The microarray experiment steps are summarized in figure 2.1, the major steps are: Target Preparation, Target Hybridization, Fluidics Station Setup, Probe Array Washing and Staining, and Probe Array Scan.

All the procedures are described in detail in the Affymetrix GeneChip Expression Analysis Technical Manual, 2005-2009. A brief description of them is mentioned down:

2.3.1 TARGET PREPERATION

The microarray experiment steps are summarized in figure 2.1. Procedures are described in detail in the Affymetrix GeneChip Expression Analysis Technical Manual, 2005-2009.

We have 10 µg RNA so one cycle cDNA synthesis is done as the first step in the microarray experiment, first addition of a poly-A tail to the RNA as a positive control

for the labeling process, Affymetrix Eukaryotic poly-A RNA Control Kit is used for this step; and a serial Dilution of Poly-A RNA control stock is prepared. The cDNA synthesis done using affymetrix one-cycle cDNA synthesis kit to turn the RNA into a cDNA by reverse transcription. T7-oligo(dT) primer; 50 µM were used in the kit, Applied Biosystem thermocycler used for the synthesis and the program for the first-strand cDNA synthesis is 70 ºC for 10 minutes, 4 ºC hold, 42 ºC for 2 minutes, 42ºC for 1 hour and hold at 4 ºC. The second-strand cDNA synthesis program 16 ºC for 2 hours, 4 ºC hold, 16 ºC for 5 minutes, and hold at 4 ºC.

A cleanup of the double-strand cDNA done then the cDNA is allowed to go through in vitro transcription back to RNA (now known as cRNA), but this RNA is labeled with Biotin. This is done by having all the uracil bases tagged with the Biotin. So, anytime a Uracil is added to the RNA chain during the transcription, a biotin molecule is also added. Affymetrix genechip IVT labeling Kit used for this step. Then a cleanup and quantification of biotin labeled cRNA is done.

This labeled cRNA is then randomly fragmented in to pieces anywhere from 30 to 400 base pairs in length (there is enough Biotin to make sure each RNA fragment has some biotin found on it. (http://cswww.essex.ac.uk/staff/W.Langdon/genechip/)

2.3.2 TARGET HYBRIDIZATION, WASHING AND STAINING

The fragmented, Biotin-labeled cRNA is then added to the array, by hybridization step; affymetrix Hybridizatioın control kit for the hybridization positive control which are four biotin labeled cRNA (bioB, bioC, bioD, cre). And affymetrix genechip hybridization, wash, and stain kit used. our hybridization machine is genechip hybridization oven 640. The arrays used is genechip Human Genome U133 plus 2.0 arrays which cover over 47.000 transcripts and variants; which represent nearly 39.000 of the best characterized human genes. The duration of the hybridization on the oven is 16 hours.

Anywhere on the array where a RNA fragment and a probe are complimentary, the RNA hybridized to the probes in the array (there are millions of identical probes in each array).

The array is then washed to remove any RNA that is not hybridized to an array and then stained with the fluorescent molecule streptavidin phycoerythrin (SAPE); that hybridize to Biotin; done by automated washing and staining protocols, our machine is genechip fluidics station 450.

the array is scanned with a gene array scanner; pixel value is 3 μm and the scanning wavelength is 570 nm, to give us a dat. image for the array, kept in the computer for quantitative analysis.

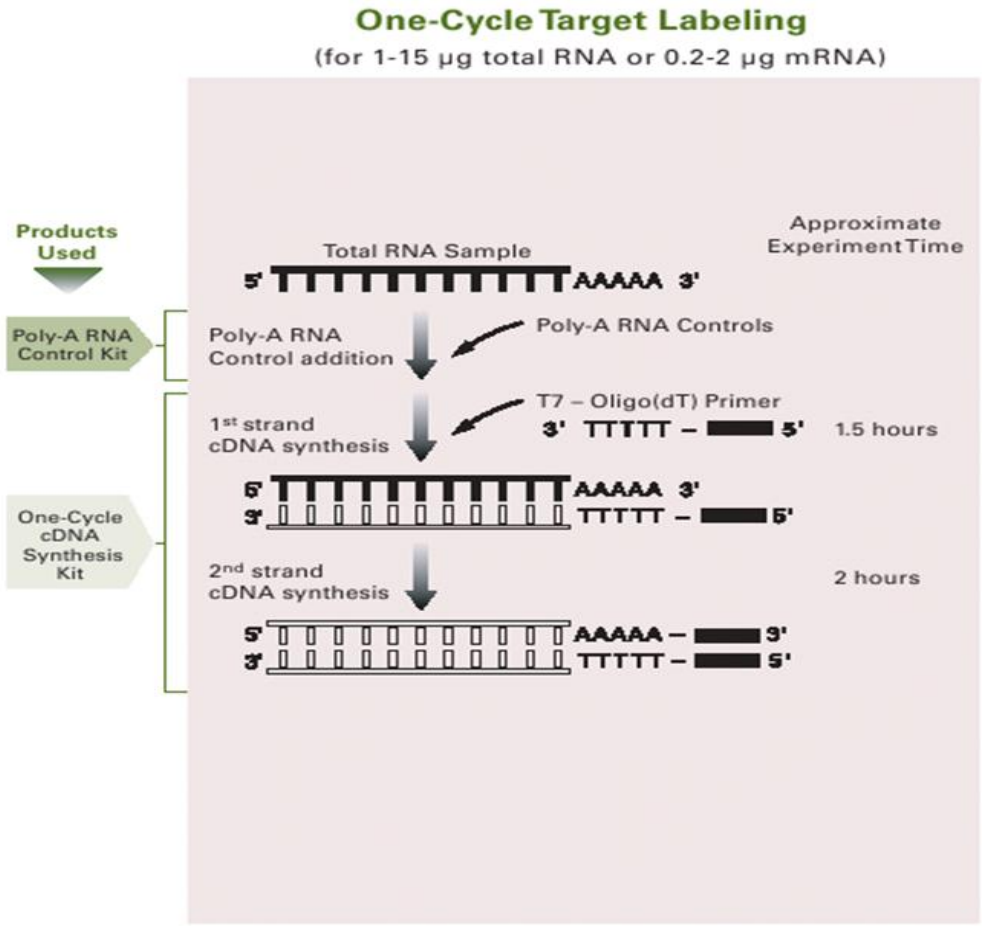


Figure 2.1: continue in the next page.

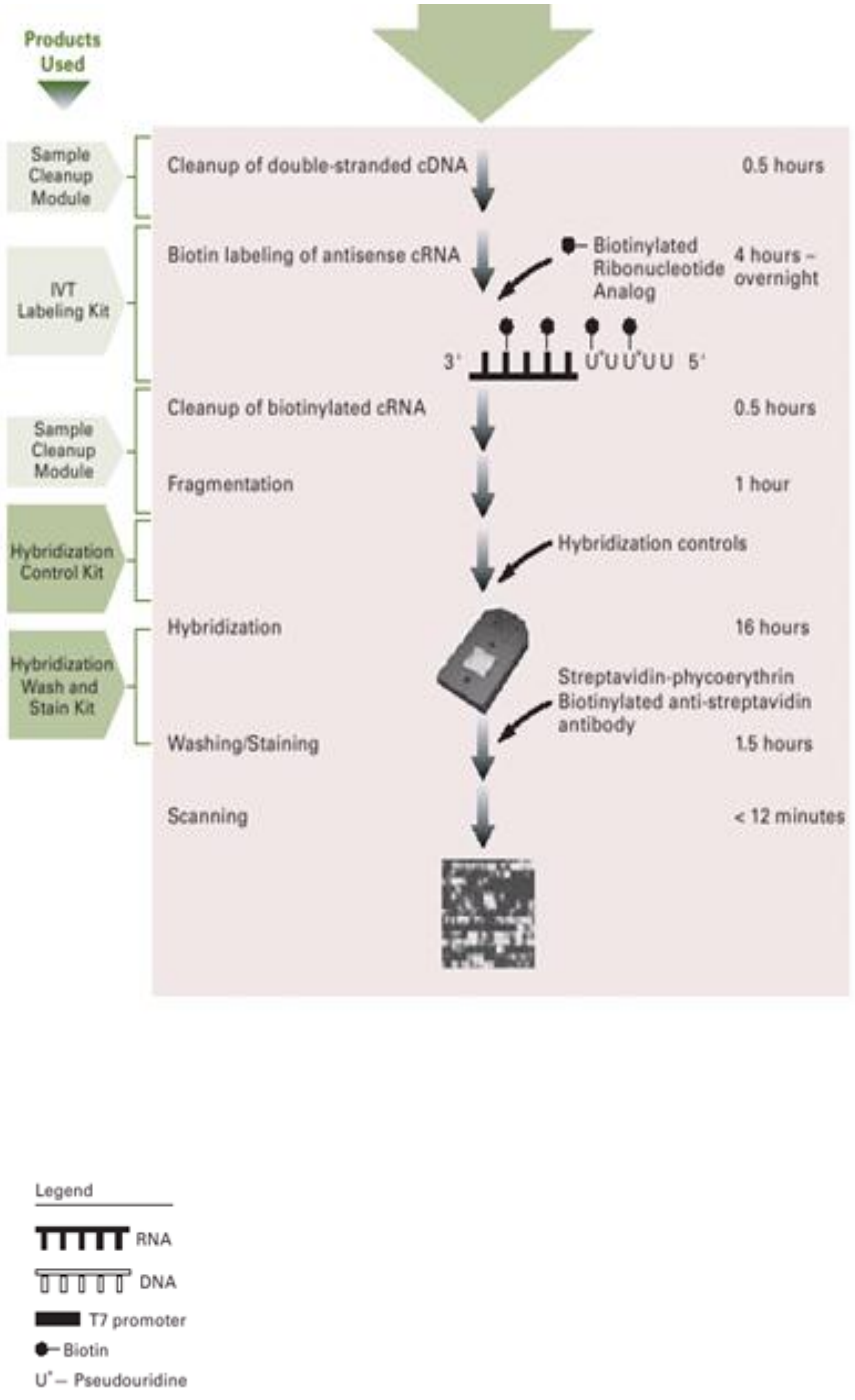


Figure 2.1: Genechip Eukaryotic labeling assays for expression analysis. (http://cswww.essex.ac.uk/staff/W.Langdon/genechip/).

2.4 DATA ANALYSIS

2.4.1 IMAGE ANALYSIS

When using affymetrix platform image analysis is simple and automated procedure because it's all standardized and the chip and the scanner all automated to give high quality result.

From the dat. image, affymetrix software GCOS will identify the position of every probe location by placing a grid on top of the scanned image. B2 Oligo serves as positive hybridization control which found on the top corner of each array, and used by the software to place a grid over the image. After the grid is adjusted and the location of each probe is identified, and signal value for each probe is calculated and reported in a CEL file.

2.4.2 ASSESSMENT OF THE QUALITY CONTROLS

Affymetrix genechip controls are chicked; B2 oligo performance which serves as a hybridization positive control, poly-A Control lys, phe, thr, dap to monitor the target labeling process, hybridization controls bioB, bioC, bioD, and cre and internal controls gene to assess RNA sample and assay quality all these controls chicked according to the information on the introduction (refer to....)

And also quality control and preprocessing evaluation of affymetrix CEL.files done using arrayanalysis.org from BiGcat bioinformatics online tool.

2.4.3 PREPROCESSING

Data analysis done by BRB (Biometric Research Branch)-Array Tools, data algorithm for preprocessing RMA methods chose because (RMA) normalization can provide a better estimation of expression levels than MAS 5.0, especially for the lower expression values. RMA normalization suggests that subtracting mismatch (MM) values from perfect match (PM) values as a way of correcting non-specific binding as used in MAS-5.0 is not always appropriate.(Song; 2008).

Filtering parameters:

- R version 2.13.2 (2011-09-30)
- BRB-ArrayTools Version: 4.2.0 - Stable Release (October 2011)
- Project annotated by Bioconductor (www.bioconductor.org) annotation package hgu133plus2.db (Version:2.5.0).
- Spot Filters: OFF when choosing JUST RMA algorithm
- Average the replicate spots within an array: ON
- Normalization: there's no option to change in BRB tool when using Just RMA method; Median arrays used as a reference array.
- Exclude a gene under any of the following conditions:
- Less than 20 % of expression data have at least a 1.25 -fold change in either direction from gene's median value
- p-value of the log-ratio variation in greater than 0.01
- Percent of data missing or filtered out exceeds 50 %
- 20th Percentile of intensities is less than 100
- Gene Subsets: OFF

The analysis done for 3 times: for the 4 array of MCF cells (2 control and 2 treated with Emodin), the other 4 arrays of MDA cells(2 control and 2 treated with Emodin), and once for all the eight arrays.

Scatter plot for all the genes passing the filter done, plotted values for each gene will be log intensities averaged over each phenotype class.

2.4.4 DIFFRENTIALLY EXPRESSED GENES

To choose our regulated gene list Class comparison between groups of arrays application done, the parameter used in the analysis of the 8 arrays :

- Number of classes: 2
- Number of genes used for random variance estimation: 54675
- Number of genes that passed filtering criteria: 54675
- Type of univariate test used: Two-sample T-test (with random variance model)
- Class variable : Control vs Emodin

- Random variance model parameters: a= 0.71293 , b= 53.32553 , Kolmogorov-Smirnov statistic= 0.03424
- Nominal significance level of each univariate test: 1e-04

And the parameters for the other analysis 4 arrays MCF and 4 arrays MDA done using this parameter:

- Number of classes: 2
- Number of genes used for random variance estimation: 54675
- Number of genes that passed filtering criteria: 7732
- Type of univariate test used: Two-sample T-test (with random variance model)
- Class variable : Control vs Emodin
- Random variance model parameters: a= 2.49582 , b= 29.25139 , Kolmogorov-Smirnov statistic= 0.00989
- Nominal significance level of each univariate test: 1e-06

Then network analysis (using GeneMANIA online tool) done for the three analysis top 10 regulated genes lists separately, and the search parameters was:

- Organism: H.Sapiens (Human).
- Networks:
  - Co-expression
  - Genetic interaction
  - Pathway
  - Predicted
- Network weighting: Biological process based
- Number of gene results: 20.

Next a pathway Drawn using Wiki Pathways: Create, for the effect of Emodin on MCF-7 and MDA-MB-231 cells line; using the information of the annotation tables, and the network analysis result for the relations between the genes.

2.4.5 TOP 10 PATHWAYS FROM DAVID ANNOTATION TOOL

Gene annotation done for all the up and down regulated genes at 1.25 fold change, from the analysis of the 8 arrays of both cell lines, using DAVID functional annotation tool. And the top 10 Pathway were analyzed.

2.4.6 CLUSTERING

Clustering between all samples and gene done for the 8 arrays to see the overall pattern in the experiment, Hierarchical clustering. Metric: one minus correlation, Average linkage.

2.4.7 OUR GENES OF INTEREST

Our genes of interest; GST isozymes classes: Alpha, Mu, Pi, Theta, Omega, Zeta and C-terminal domain containing. Checked their result which one is up and which is down regulated after treated with emodin and their Annotations, biological function, pathways is listed.

2.5 PRIMER DESIGN FOR REAL TIME PCR

Primer design for our genes of interest list (GST isozymes) done using Primer-BLAST online tool, which developed by NCBI , It uses Primer3 to design PCR primers and then submits them to BLAST search against user-selected database. The blast results are then automatically analyzed to avoid primer pairs that can cause amplification of targets other than the input template.(http://molbiol-tools.ca/PCR.htm)

The sequences are taken from The National Center for Biotechnology Information (NCBI) as FASTA format, then copied at Primer-BLAST tool.

Primer parameter that changed from the suggested parameter is :

- Product size: Min 70, Max 200
- Exon junction span: primer must span an exon-exon junction.

Table 2.1: GST isozymes primers Designed using Primer-BLAST tool, ***All the Primers sequence (5'-->3').***

Gene name	Forward primer Tm	Reverse Primer Tm	Length (bp)
GSTA1	AGCTTCCCTCTGCTGAAGGCCC 60.50°C	AGTTCTTGGCCTCCATGACTGCG 59.20°C	167
GSTA4	CCGCTGACCTGGCGCTTTGT 60.25°C	CTTCATCAAACTCGACTCCGGCGG 59.89°C	164
GSTM3	AGGGGTCAGCGCTCTTGCTT 58.34°C	GGGAAATGCCACAGTATCGCAGC 58.61°C	158
GSTO1	TCCCACAGTCTCAGCCCTGCT 59.22°C	TCCTGCCCCCTTCAGAGCCC 59.89°C	111
GSTT1	TTCCGGTCAGGTCGGTCGGT 59.55°C	CACCTGGGCAAAGGCATCGCT 59.98°C	171
GSTZ1	GCATCGACTACAAGACGGTGCCC 59.88°C	GAAGTCGCGGAGTGGGACGC 60.11°C	183

2.5.1 CHECKING THE PRIMERS

To chick if our primers work we tested them on (RT- PCR) Reverse transcriptase PCR, using control cells of MCF-7 and MDA-MB-231 cell line.

2.5.1.1 RNA ISOLATION

RNA Isolated using 5 Prime RNA Cultured cell kit. MCF-7 and MDA-MB-231 cells (500.000 cells/well) were placed in 6 well plate. One column were used for 3 wells from each cell type.

RNA concentration measured using nano drop and detection of ribosomal RNA band checked by 1% Agarose Gel Electrophoresis (2.5 µl Dye+2.5 µl RNA).

2.5.1.2 cDNA PREPERATION

cDNA prepared using 5 Prime Master Script kit/ two-step RT-PCR used. Oligo dT primers used and 2 µg RNA were added to the mixture.

2.5.1.3 PCR WORK

The primers Prepared using DEPC autoclaved water, 20 µM prepared for the PCR (2.5 Forward+2.5 Reverse), the total volume for the reaction was 50 µl, contain;

5 µl: 10X PCR buffer with (NH4)2SO4 without Mgcl2, 1.8 µl: dNTPs (10mM), 8 µl: Mgcl2 (25mM), 0.5 µl:Taq DNA Polymerase (5u/µl), 2µl from our cDNA and nuclease free water, the Thermocycler program was:

95˚C for 2 min. (1X)

[95˚C for 30 second, 59˚C for 30 second, 72˚C for 30 second] (30X)

72˚C for 10 min. (1X)

Then 1% gel electrophoresis run done to see the result (2µl dye+10µl product sample).

2.5.1.4 DNA SEQUENCING

To Confirm our primers; DNA Sequencing done from one side using Forward Primer of each gene, for the PCR Product of this Samples:

MDA GSTA1
MDA GSTA4
MCF GSTZ1
MCF GSTO1
MCF GSTT1
MCF GSTM3

And the result sequencing were searched on (NCBI) BLAST Alignment Tool, and the Options used:

- Database Name: GP/9606.9558/RefSeq_RNA.
- Description: Homo sapiens RefSeq RNA.
- Program: BLASTN 2.2.26+.

CHAPTER 3

RESULTS AND DISCUSSIONS

3.1 RNA QUALITY ASSESSMENT (BIOANALYZER RESULTS)

Table 3.1: samples legend.

Sample name	Array names	Treatment
Sample 1	MCF DMSO Control 1	Control (treated with DMSO)
Sample 2	MCF Emodin	Treated with Emodin
Sample 3	MDA DMSO Control 3	Control (treated with DMSO)
Sample 4	MDA Emodin 4	Treated with Emodin
Sample 5	MCF DMSO Control 5	Control (treated with DMSO)
Sample 6	MCF Emodin 6	Treated with Emodin
Sample 7	MDA DMSO Control 7	Control (treated with DMSO)
Sample 8	MDA Emodin 8	Treated with Emodin

The detection of ribosomal RNA band by electrophoresis and bioanalyzer results which seen in figure 3.1 and 3.2, the rRNA 18s band detected in all samples at ≈43 second, while 28s band in average also for all samples at 50 second. And in figure 3.2 all the graphs have a baseline and sharp ribosomal Peaks with the ratio of (28s/ 18s) is near the acceptable value 2:1, 280/260 results also between the range (1.8-2.1) see table 2.3, So all the RNA sample considered as good quality, purity and acceptable to work microarray.

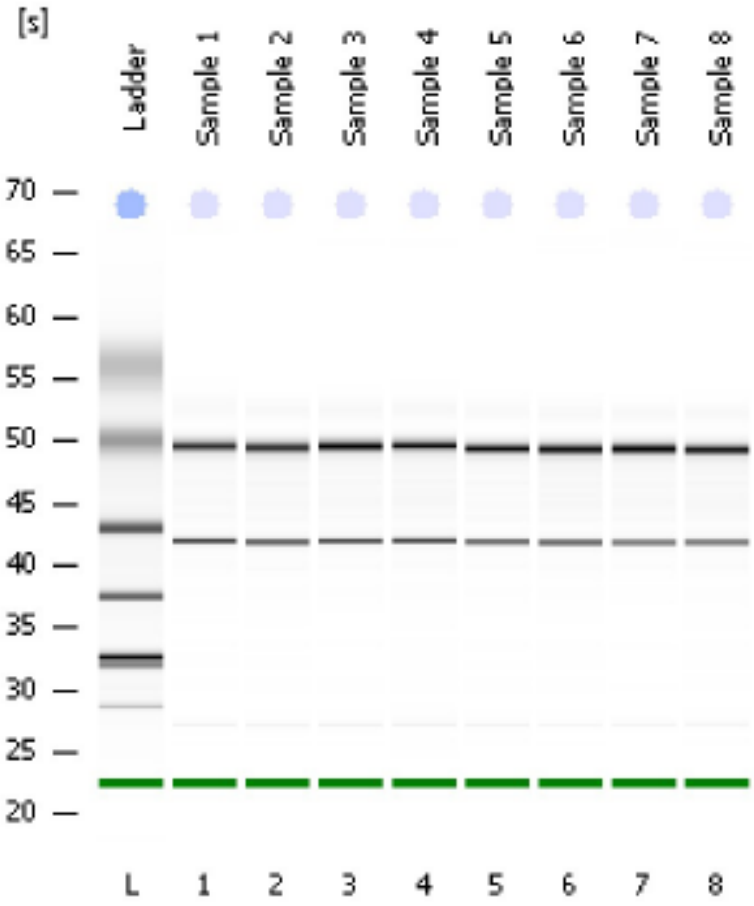


Figure 3.1: Electrophoresis Run for total RNA samples.

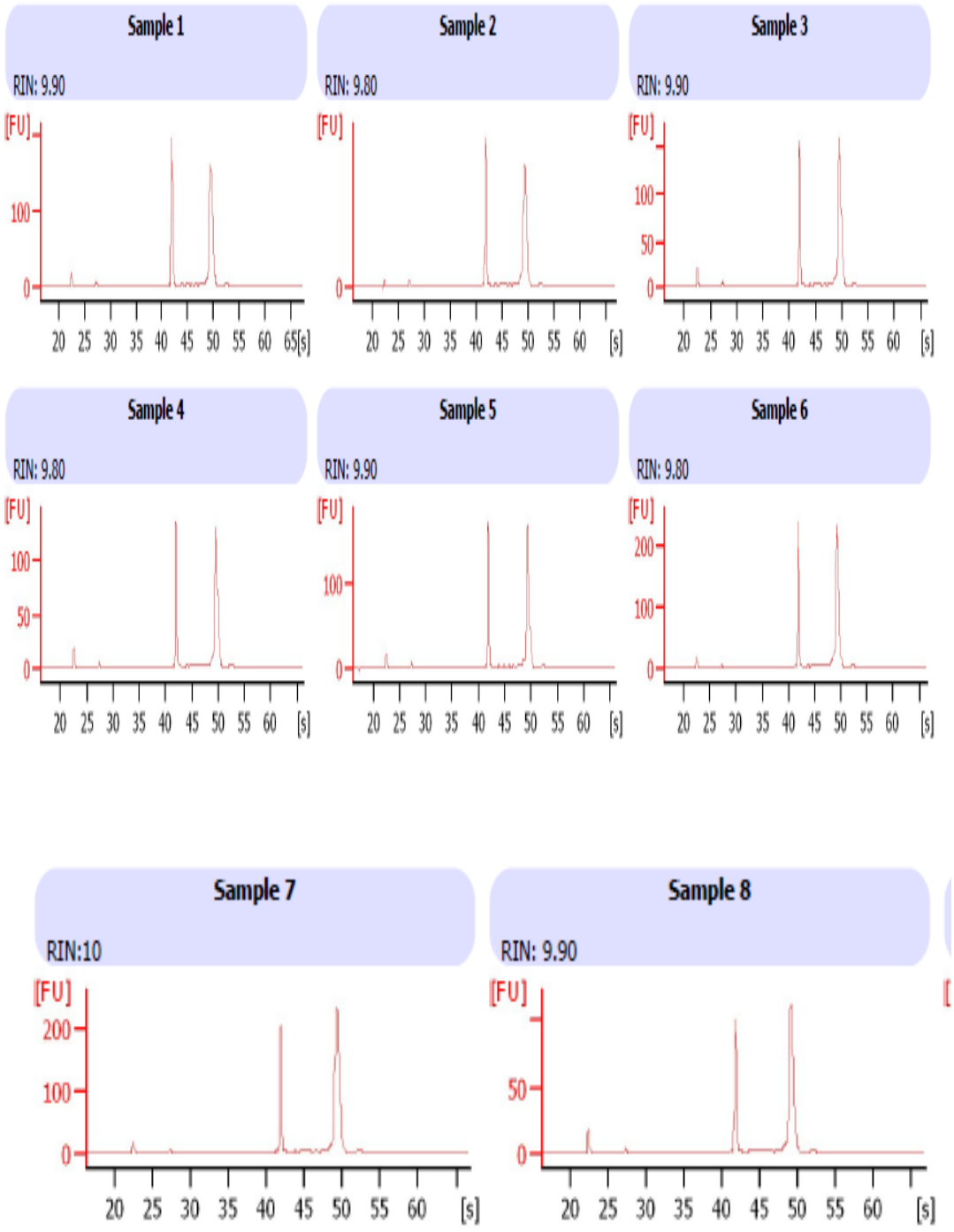


Figure 3.2: Detection of ribosomal RNA band by Agilent 2100 Bioanalyzer Machine.

Table 3.2: RNA properties, result from nano drop and Bioanalyzer

Sample name	RNA concentration (ng/µl)	rRNA Ratio [28s/ 18s]	280/260
Sample 1	177	1.7	2.06
Sample 2	345	1.7	2.02
Sample 3	157	2.1	2.03
Sample 4	132	1.9	2
Sample 5	153	1.7	2.02
Sample 6	237	1.9	2.06
Sample 7	217	2.3	2.08
Sample 8	103	1.9	2.04

3.2 QUALITY CONTROL AND PREPROCESSING EVALUATION OF AFFYMETRIX CEL.

3.2.1 RNA DEGRADATION PLOT

mRNA degradation occurs when the molecule begins to break down and is therefore ineffective in determining gene expression. Because this kind of degradation starts at the 5′ end of the molecule and progresses to the 3′ end it can be easily measured using oligonucleotide arrays, where each PM probe is numbered sequentially from the 5′ end of the targeted mRNA transcript. When RNA degradation is advanced, PM probe intensity at the 3′ end of a probeset should be elevated when compared with the 5′ end. When dealing with high quality RNA a slope of between 0 .5 and 1.7 is typical, depending on the type of array; slopes that exceed these values by a factor of 2 or higher could indicate excessive degradation. The down figure is for our Data show that in general the slope look to be between the Range (0.5-1.7) so our RNA is acceptable.

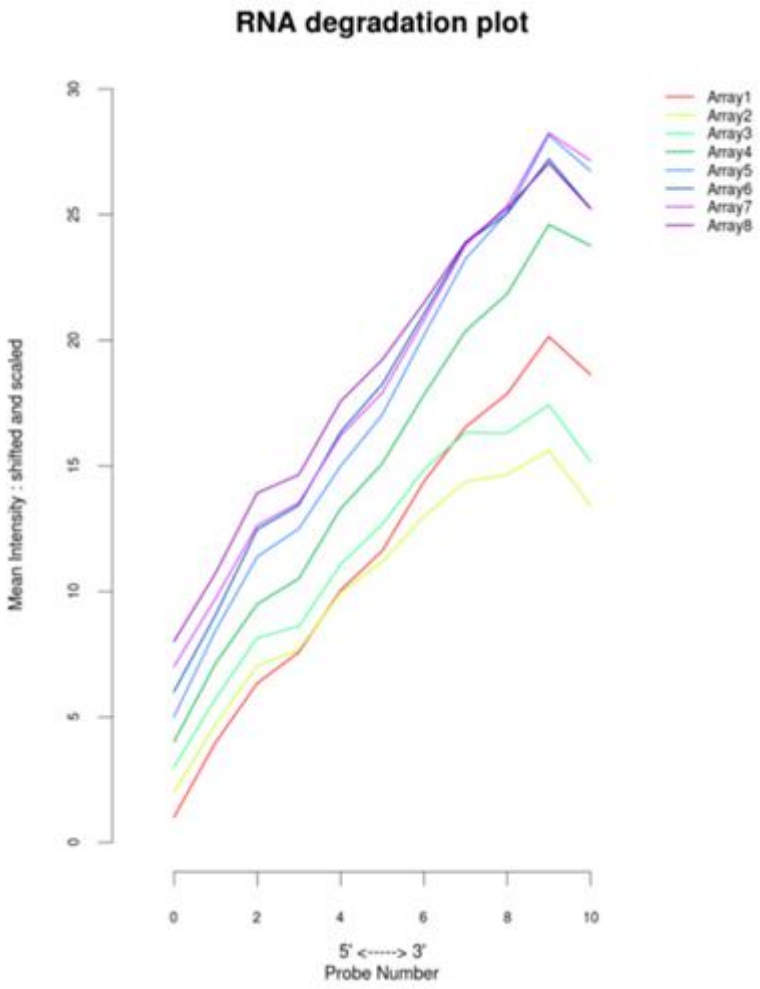


Figure 3.3: RNA degradation plot from arrayanalysis.org (quality control and preprocessing evaluation of affymetrix CEL.) Report.

3.2.2 2D IMAGES FOR SPATIAL BIAS DIAGNOSTIC

The expression estimate's characteristics plotted on the array positions allow to see spatial trends or biases on the array that are not possible to distinguish on the raw data. Expression measure estimated by a Probe Level Model (PLM) using a M-estimator robust regression.

From the 2D image for the 8 arrays we can see No spatial bias; the color-coded values are homogeneous, except array 8 it has a small artifact its clear as a ring but it's not considered as problem, because the probe intensity image (the black image), didn't show this problem due to strong probe effect. in general we can consider the arrays are good.

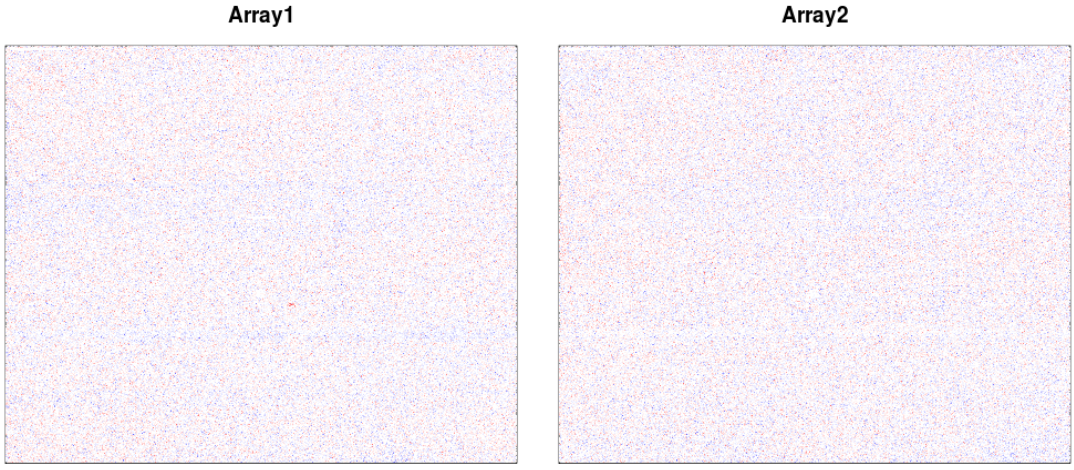


Figure 3.4: continue in the next page.

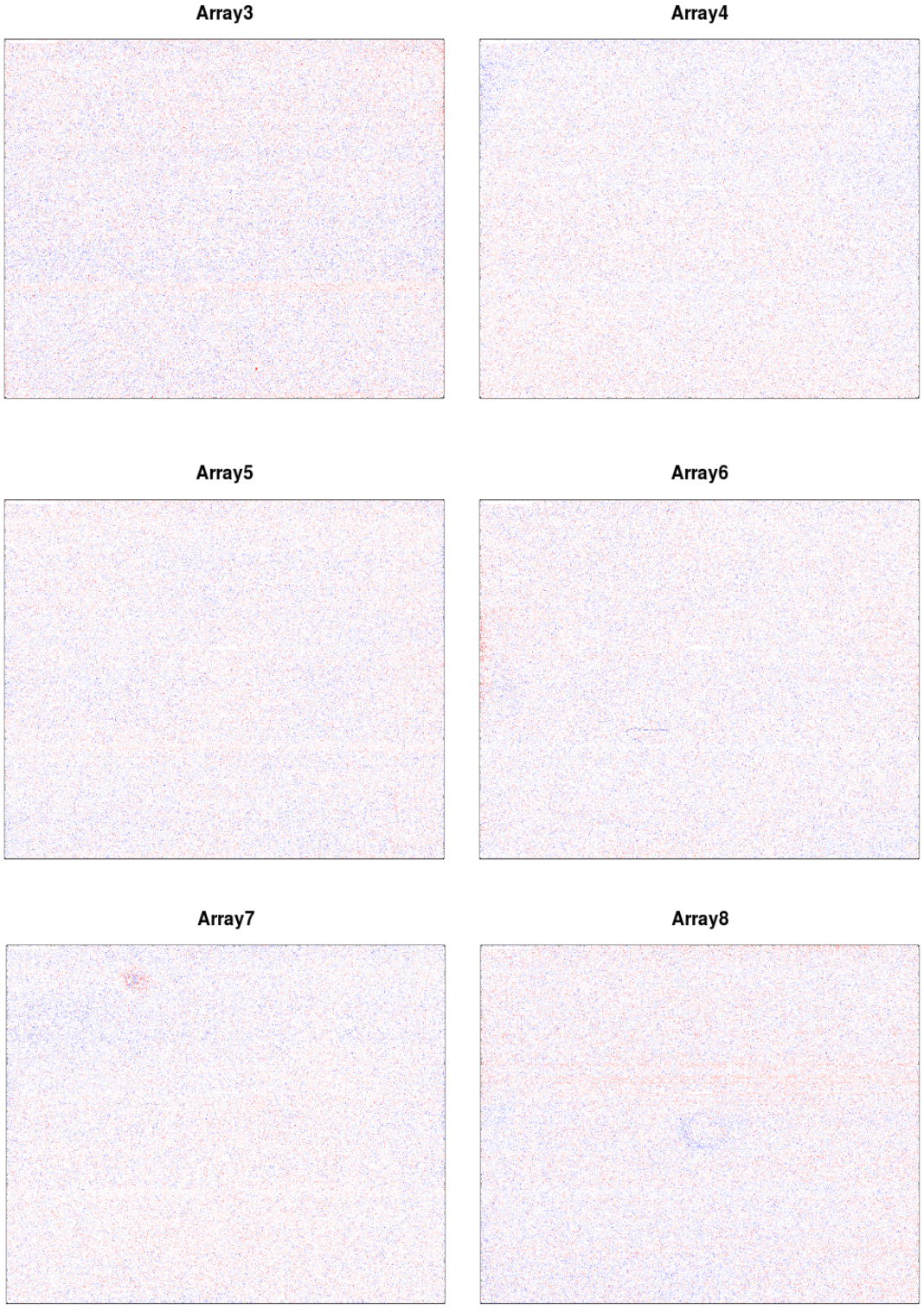


Figure 3.4: 2D images for spatial bias diagnostic.

3.2.3 DENSITY HISTOGRAM

Density histogram of raw intensities and after RMA algorithm applied in the preprocessing step used to visualize the spread of data and compare and contrast probe intensity between the arrays of the dataset. The x-axis indicates probe intensity and the y-axis represents probe density level. From comparing the two figures, we can clearly see that the Raw intensities histogram shifted to the right this indicate that the arrays have a noise like high background values then after RMA the arrays show to be near to the normal Distribution and the arrays Fit each other more than the raw Data, this give us indication of a good preprocessing step Using RMA algorithm.

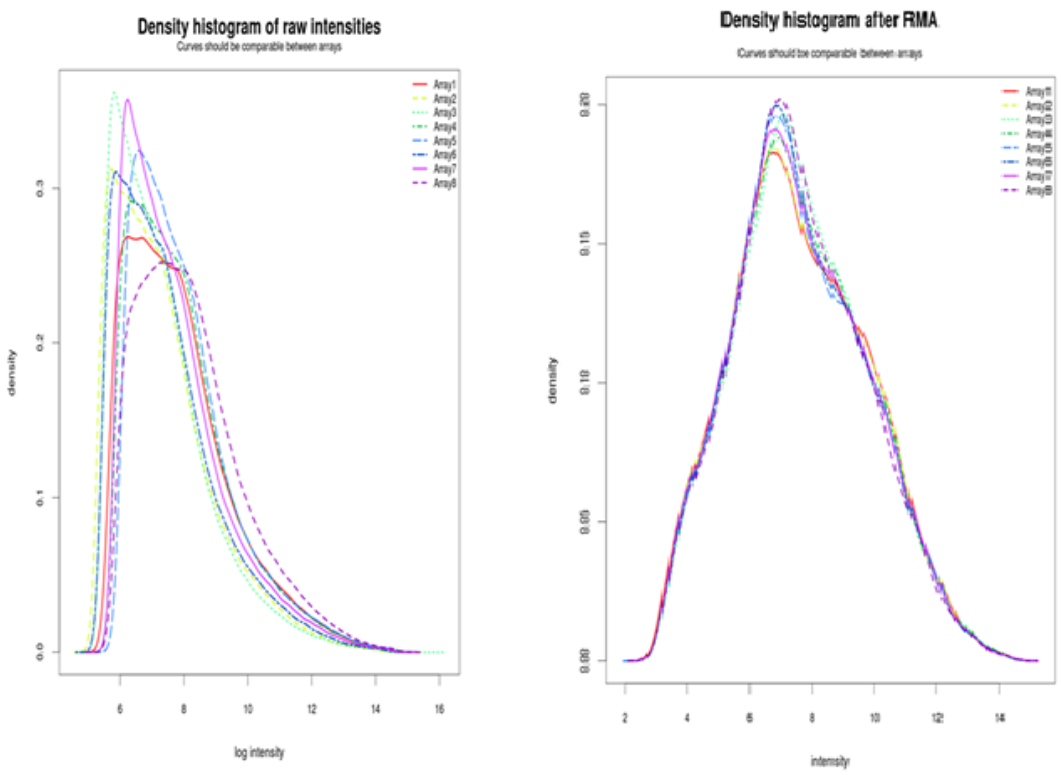
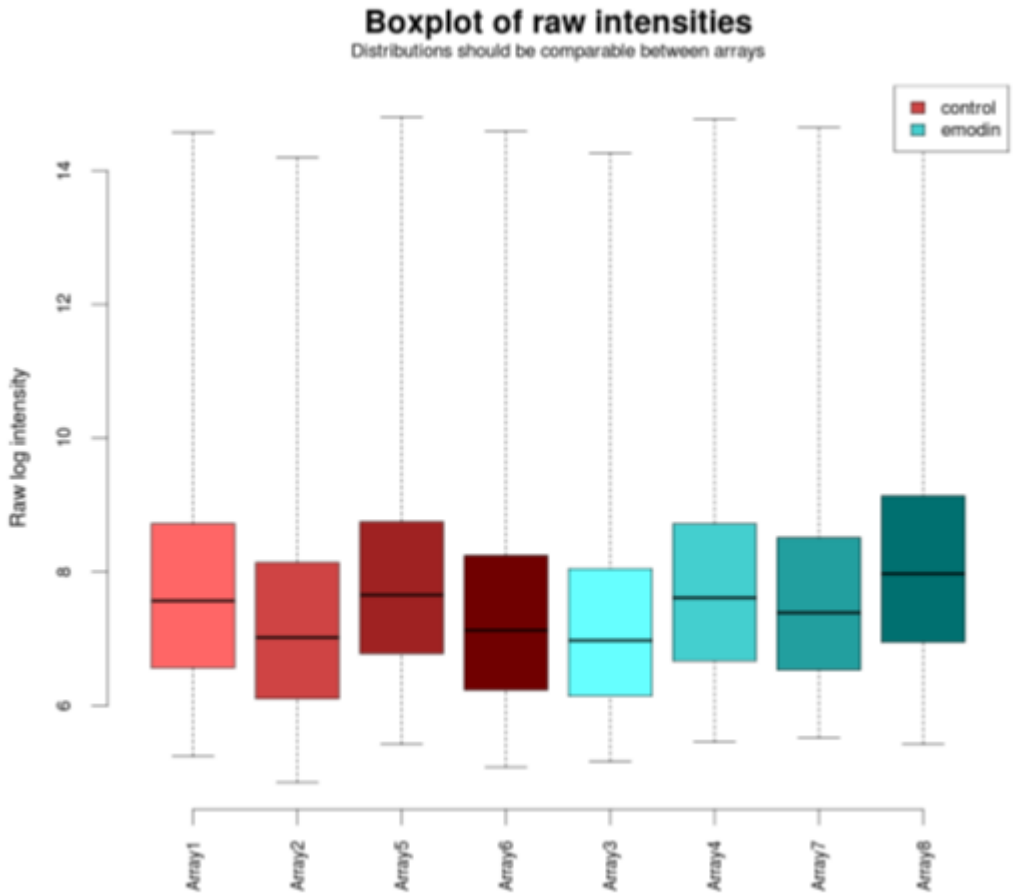


Figure 3.5: Density histogram of raw intensities and after RMA algorithm applied in the preprocessing step.

3.2.4 BOXPLOT

From Boxplot Figures 3.6 and 3.7 it's clear that the difference between the arrays removed and the average nearly one after the Normalization Step.



Figur 3.6: Boxplot of raw intensity data

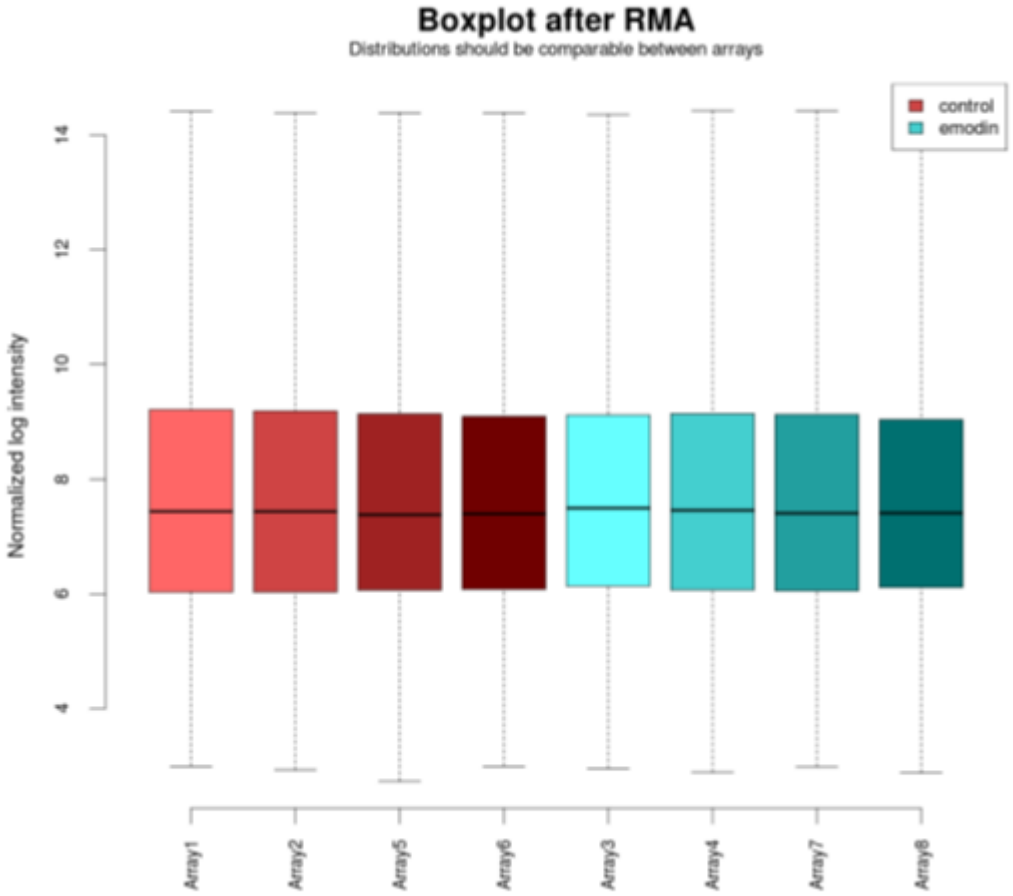


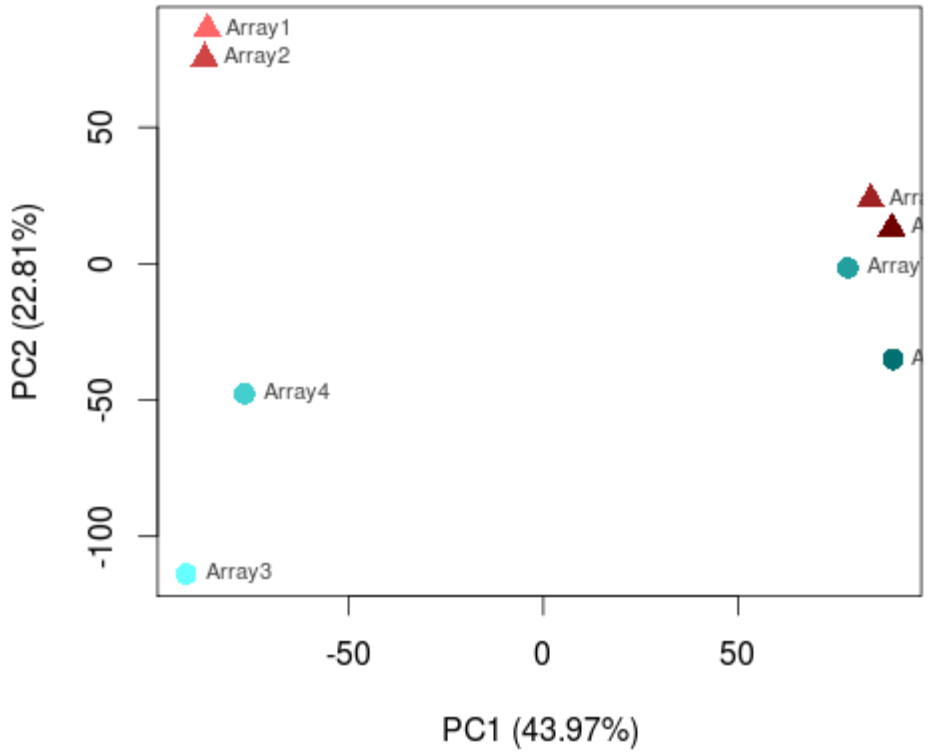
Figure 3.7: Boxplot after RMA

From the Density histograms and the Boxplot its considered that the preprocessing steps is successfully done by RMA algorithm, and the non-biological source of variance between the arrays is removed the Data is ready to analyzed to see the gene expression regulation.

3.2.5 THE PCA (PRINCIPAL COMPONENT ANALYSIS)

The PCA (Principal Component Analysis) gives another view of the correlations of expression between arrays: the data are projected on several axes (or components), ordered by decreasing significantly; the first principal component (PC1) explains most of the variations of expression. (Almost 47% of the variance).

The PCA graph presents 3 plots: the array data are projected respectively on PC1 versus PC2, PC1 versus PC3 and PC2 versus PC3. by decreasing order of significance: PC1, PC2, PC3. We clearly see that it the arrays scattered in the space in the three graphs, some are tight together which represent that the expression level near to each other; in general there's two group first according MCF and MDA variance and the other variance grouped is between emodin and controls.



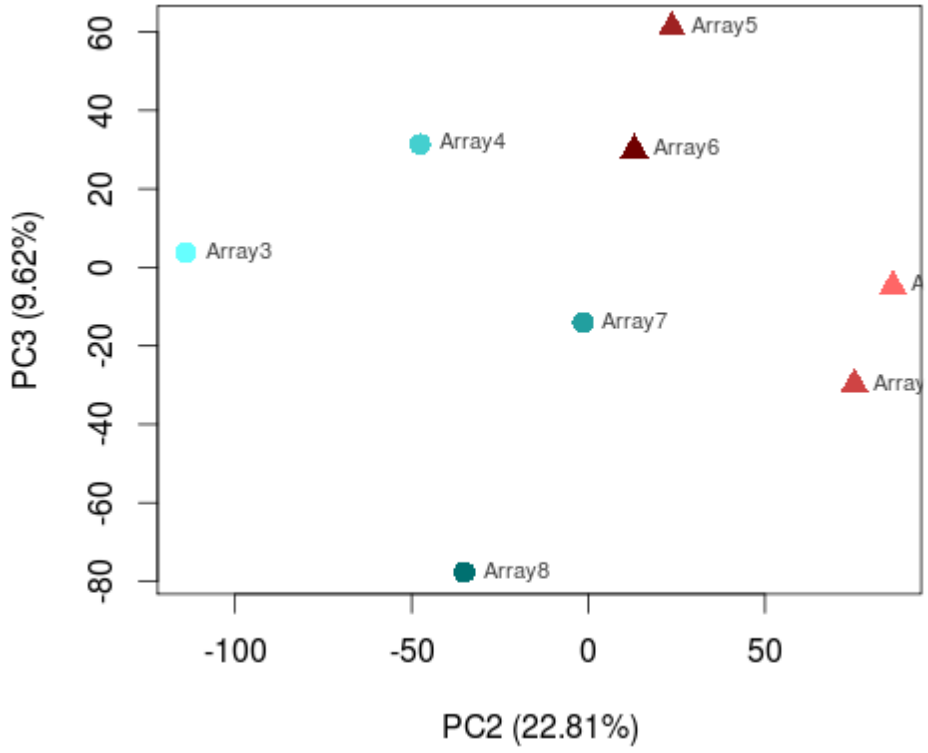


Figure 3.8: The PCA Graph; continue in the next page.

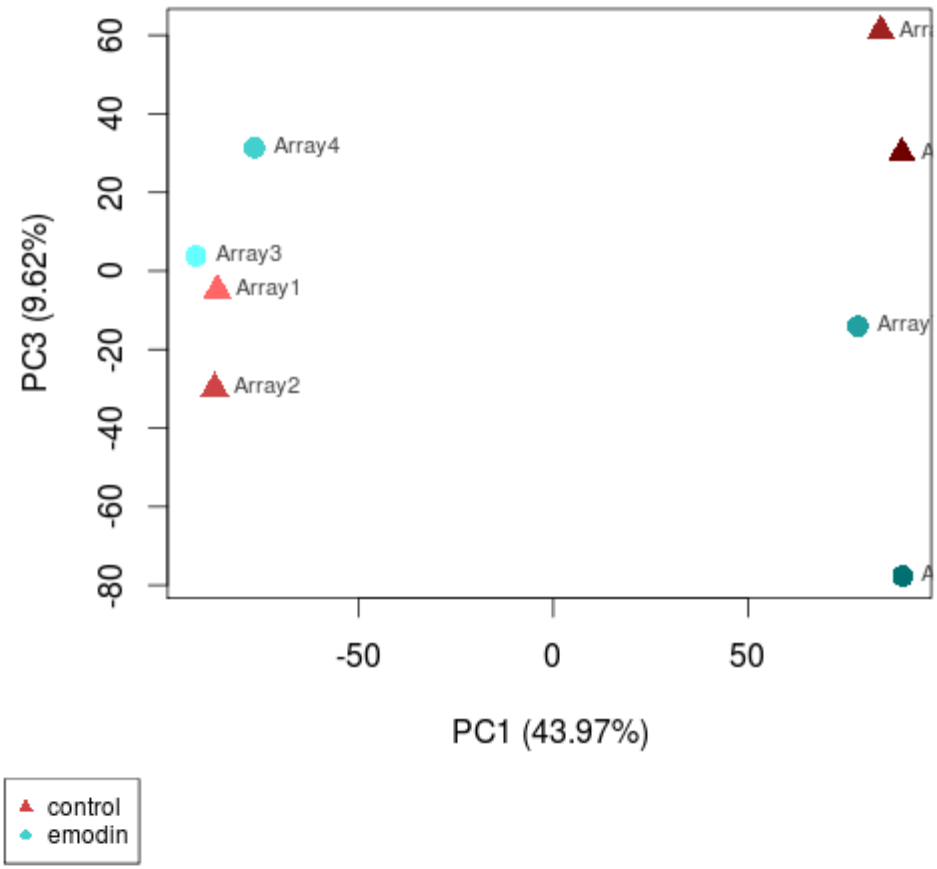


Figure 3.8: The PCA graph presents 3 plots: the array data are projected respectively on PC1 versus PC2, PC1 versus PC3 and PC2 versus PC3.

3.3 SCATTER PLOT PHENOTYPE AVERAGE BETWEEN EMODIN AND CONTROL CLASSES AFTER PREPROCESSING

From the scatter plot in Figures 3.9, 3.10, and 3.11, we can see the up and down regulated genes at 2 fold change; and it's clear that in MCF cells the change of gene expression is more than MDA-MB-231 after emodin treatment, which support that emodin is a phytoestrogenic component. Also in the graphs we can see some of the genes name of the highly differentially expressed chose randomly.

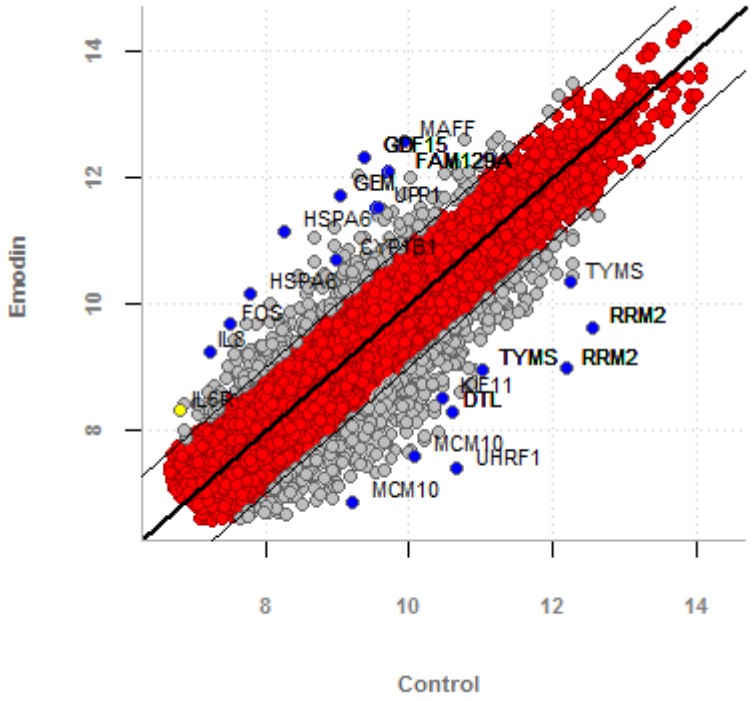


Figure 3.9: Scatter plot, phenotype average between two classes Emodin and Control from the analysis of 8 arrays from both MCF7 and MDA MB 231.

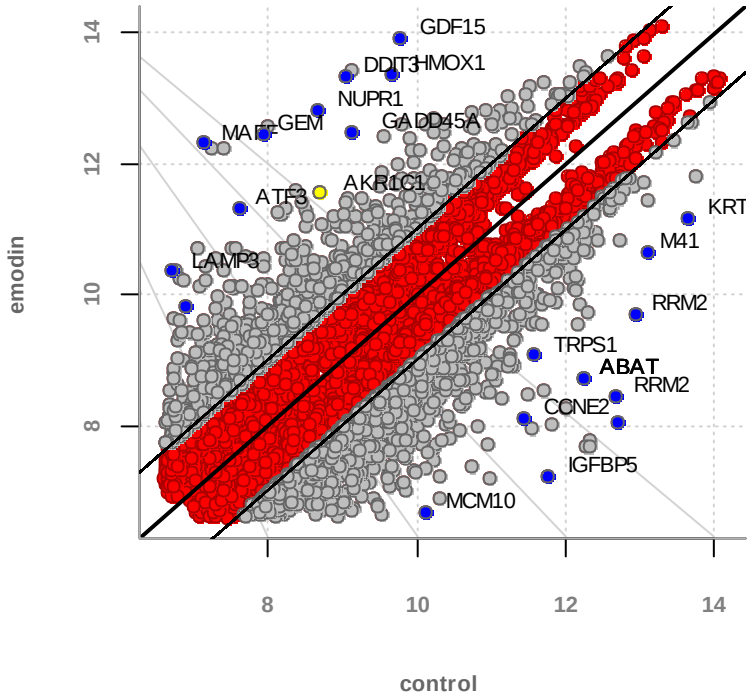


Figure 3.10: Scatter plot, phenotype average between two classes emodin and Control from the analysis of 4 arrays from MCF7 cells.

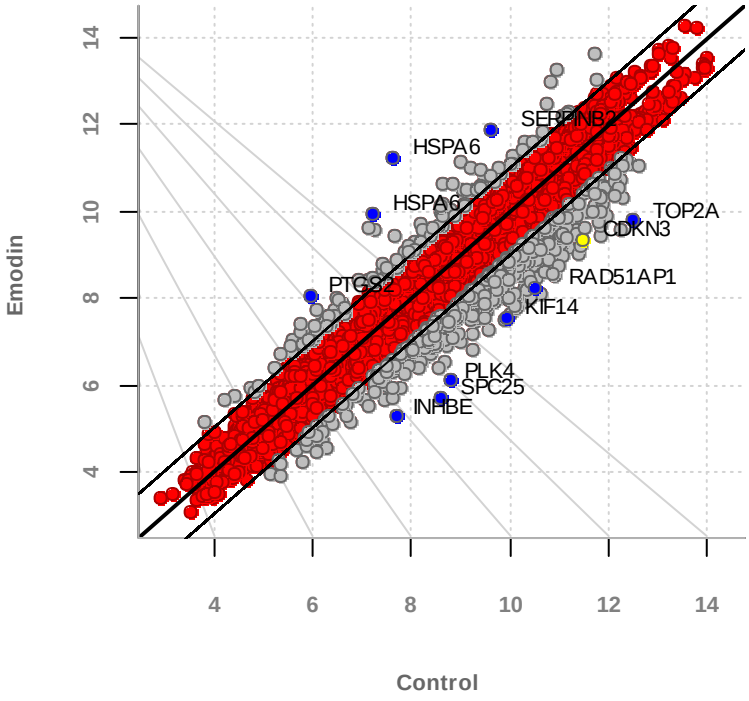


Figure 3.11: Scatter plot, phenotype average between two classes Emodin and Control from the analysis of 4 arrays from MDA MB 231cells.

3.4 DIFFRENTIALLY EXPRESSED GENES AND THEIR ANALYSIS

Differentially expressed genes selected from class comparison between Emodin and Controls classes, at 1e-04 level of the univariate test, we have three differentially expressed genes lists; the first from the analysis of the 8 arrays for both MCF7 and MDA-MB-231 cells, the second from the analysis of the 4 arrays of MCF7 cells alone, the third from and the analysis of the 4 arrays of MDA-MB-231 cells alone. see Table 3.3; which show the top 10 regulated genes between Emodin and Control classes for the three type of analysis done, and the summary of the Annotation of the three analysis results in Tables 3.4, 3.6 and 3.8. Then the significant genes from every analysis done are clustered, and the heat map drawn, see Figures 3.12, 3.14 and 3.17, and also network analysis done for the three lists, see Figures 3.13, 3.15 and 3.18 , the predicted genes analyzed, most of them were in our gene list after, and Tables 3.5, 3.7 and 3.9 show the predicted genes annotations and their regulation after emodin treatment

Table 3.3: Show the top 10 regulated genes between emodin and control classes, for the three type of analysis done.

Top regulated genes after emodin treatment of both cells type	Up/down regulation	Top regulated genes after emodin treatment of MCF7 cells	Up/down regulation	Top regulated after emodin treatment of MDA cells	UP/down regulation
TRIP13	Down	FAM129A	Up	HSPA6	Up
BIRC5	Down	ATF3	Up	PTGS2	Up
SKA3	Down	NUPR1	Up	IL24	Up
NCAPH	Down	FAM129A	Up	HSPA6	Up
TCF19	Down	DDIT3	Up	DEPDC1B	Down
ZNF304	Up	GDF15	Up	KIF14	Down
KIAA0101	Down	CYP1A1	Up	KIF23	Down
CCDC34	Down	GADD45A	Up	HSPA1A	Up
C1orf112	Down	PHLDA1	Up	TOP2A	Down
COQ2	Down	GEM	Up	TCF19	Down

3.4.1 ANALYSIS OF THE TOP 10 REGULATED GENES IN THE 8 ARRAYS FOR BOTH CELLS TYPE

From the analysis of the significant genes and their annotations, it's found that in general for the effect of emodin on the cells whether MCF7 or MDA- MB-231; the top 10 significant genes contain biological processes include cell cycle, cell division, cell proliferation, mitosis and meiosis, most of the top 10 regulated gene are down regulated which insured the relation of emodin to suppress tumor growth. The network analysis results also support this, the predicted genes was contain the same biological processes profile like cell cycle and mitotic cell cycle, and all of this gene was down regulated also in our project after emodin treatment.

Table 3.4: Summary of the Annotation of the top 10 Genes which are differentially expressed among emodin and Control Classes in analysis of 8 arrays; MDA and MCF cells.

Top regulated genes after emodin treatment of both cells type	Up/down regulation	The most related To tumor growth Biological processes From Gene Ontology	Kegg pathways	Biocarta pathways
TRIP13	Down	Male and Female meiosis ı		
BIRC5	Down	Apoptosis, Anti- apoptosis, and positive regulation of mitotic cell cycle		
SKA3	Down	Cell cycle and cell division		
NCAPH	Down	Cell cycle and cell division		
TCF19	Down	Cell proliferation		
ZNF304	Up	Regulation of transcription, DNA-dependent		
KIAA0101	Down			
CCDC34	Down			
C1orf112	Down			
COQ2	Down	Ubiquinone biosynthetic process	Ubiquinone biosynthesis	

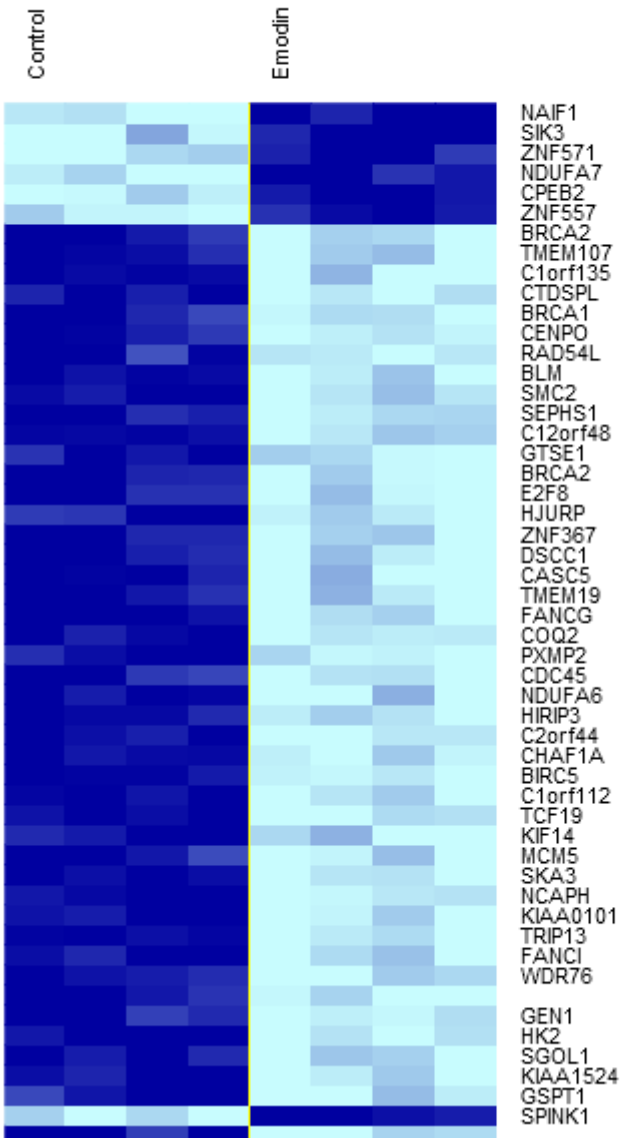


Figure 3.12; continue in the next page.

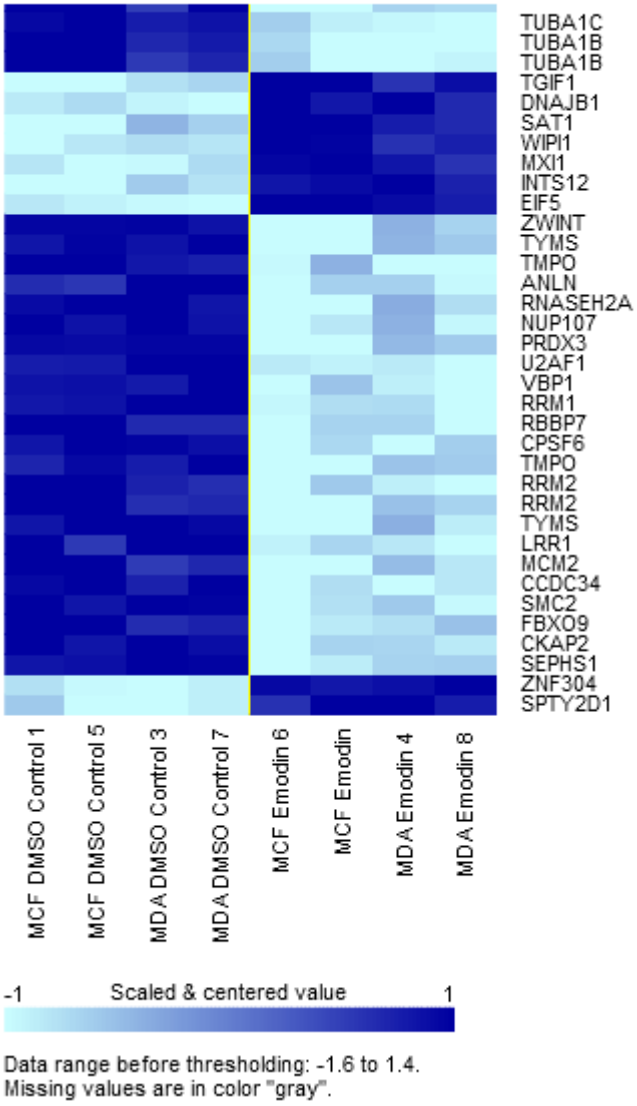


Figure 3.12: Clustered heatmap of significantly expressed genes from the analysis of the 8 arrays of the both cells type, arrays grouped by class.

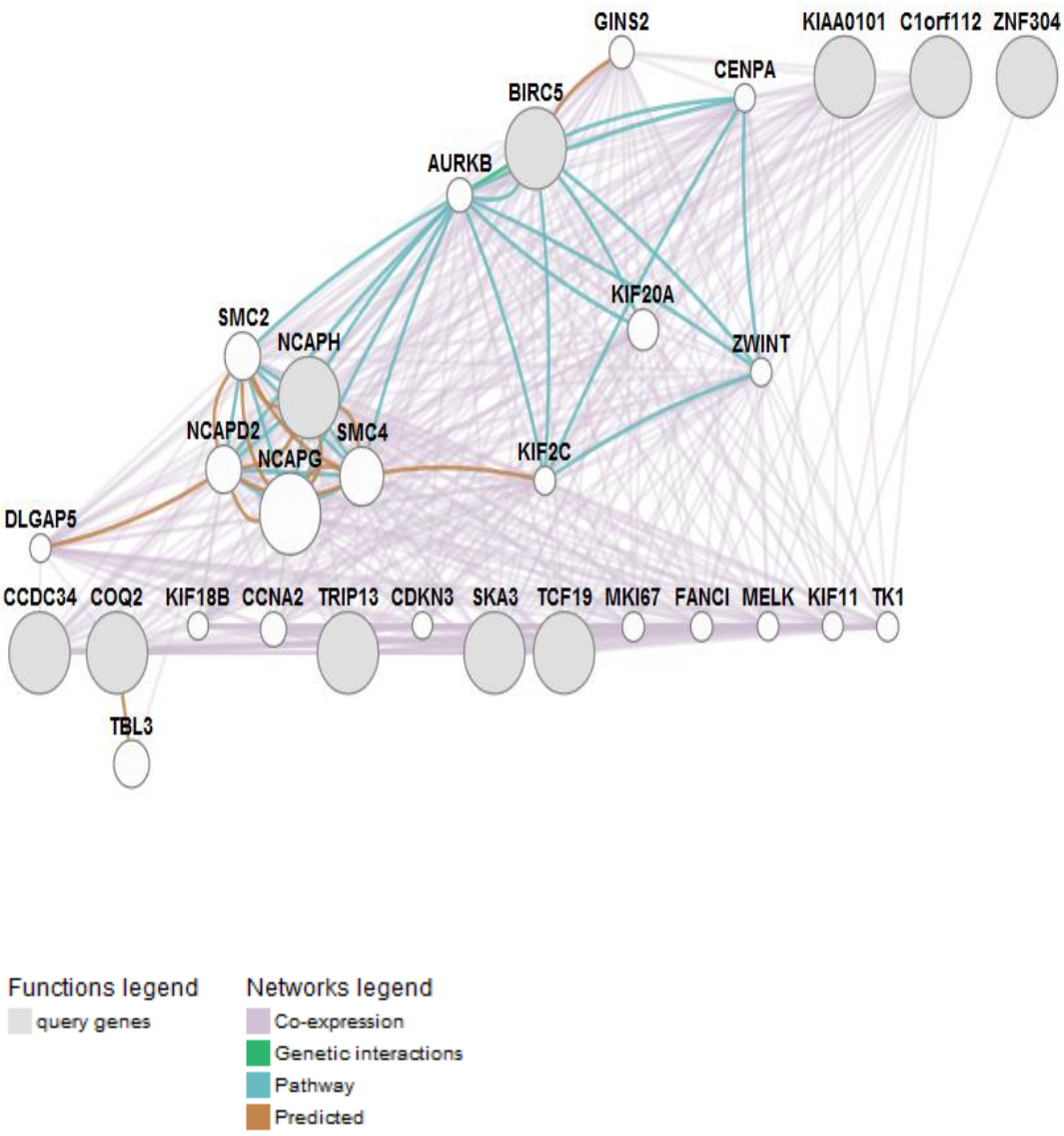


Figure 3.13: Network analysis done for the top 10 regulated genes result from the analysis of the 8 arrays of the both cells type.

Table 3.5: The analysis of the predicted genes from network analysis result in figure 3.13.

Predicted gene	Up/down regulated after Emodin treatment in the analysis of 8 arrays for both cells type	The most related To tumor growth Biological processes From Gene Ontology
NCAPG	DOWN	cell cycle, and cell division
NCAPD2	DOWN	cell cycle, and cell divisiom
SMC4	DOWN	cell cycle, and cell division
SMC2	DOWN	cell cycle, and cell division
DLGAP5	DOWN	cell cycle, and positive regulation of mitotic metaphase/anaphase transition.
KIF2C	DOWN	mitotic cell cycle, cell proliferation, and cell division.
AURKB	DOWN	mitotic cell cycle, and cell proliferation
GINS2	DOWN	DNA replication
CENPA	DOWN	mitotic cell cycle
ZWINT	DOWN	cell cycle and cell division
KIF20A	DOWN	mitotic cell cycle

3.4.2 ANALYSIS OF THE TOP 10 REGULATED GENES IN THE 4 ARRAYS FOR MCF-7 CELLS

In the analysis of the effect of emodin on MCF cells alone, from the top 10 significant varied genes annotation table we can see cell growth related biological process like the up regulation of NUPR1 which has induction of apoptosis biological process, and up regulation of GADD45A gene which encode the protein responsible for Growth arrest and inducing DNA-damage. And also other processes like phosphorylation, oxidative stress, stimulus to estradoil response, and toxin metabolic process. Also we can see some of the pathways included in the table, for example P38 MAPK signaling pathway which relates to cancer, specifically by affecting apoptosis (see Figure 3.22), and the metabolism of xenobiotic by cytochrome p450 pathway which affected by the Up regulation of CYP1A1 gene, so CYP1A1 gene codes for a protein which is used in emodin metabolism. These results, and the high number of the varied genes in MCF cells comparing to MDA-MB-231 it was nearly 4400 gene in MCF-7 from the scatter plot result while in MDA-MB-231 it was nearly 3400 gene, insured the effect of emodin as anti-tumor, and as a phytoestrogenic component, as MCF7 cells are ER positive cells. The network analysis result also supported these results, when the predicted genes analyzed, it

found that also they supported the result of emodin as anti-cancer, for example; JUN and FOS genes was up regulated in our gene list after emodin treatment, this genes contain Biological process like negative regulation of cell proliferation, stress activated MAPK cascade which relate to apoptosis, and Toll signaling pathway, and MyD88-independent toll-like receptor signaling pathway, which play a role in the innate immune system, see Figure 3.22, and GADD45B gene also was up regulated, which contain Apoptosis and activation of MAPKK activity. The network result also supported emodin as a phytoestrogenic component, ASNS gene was up regulated, contain biological process like; cellular response to hormone stimulus, and CYP1B1 gene also was up regulated which include in estrogen metabolic Process, and xenobiotic metabolic process ,DUSP1 gene was up regulated, it contain response to esradiol stimulus and inactivation of MAPK activity, Figure 1.4 in the introduction show the estrogen receptor signal transduction pathway and show the relation between ER and MAPK Activity. NRUA1 also up regulated gene also has molecular function: steroid hormone receptor activity, and in its biological processes; induction of apoptosis, which supported our result.

A proposed mechanism of the effect of emodin on MCF-7 cell lines pathway drawn using Wiki Pathways, and need to be confirmed in the future studies.

Table 3.6: Summary of the Annotation of the top 10 Genes which are differentially expressed among emodin and Control Classes in analysis of 4 arrays of MCF cells.

Top regulated genes of MCF7 cells	Up/down regulation	The most related To tumor growth Biological processes From Gene Ontology	Kegg pathways	Biocarta pathways
FAM129A	Up	Response to stress, and positive regulation of translation		
ATF3	Up	Positive regulation of cell proliferation, and regulation of transcription DNA dependent		
NUPR1	Up	Induction of apoptosis		
FAM129A	Up	Response to stress, positive regulation of translation		
DDIT3	Up	Transcription- DNA dependent, response to DNA Damage Stimulus, and positive regulation of apoptosis.	MAPK Signaling pathway	P38 signaling pathway
GDF15	Up	Transforming growth factor beta receptor signaling pathway, and cell-cell signaling		
CYP1A1	Up	Toxin metabolic process.	Tryptophan metabolism And Metabolism of xenbiotic by cytochrome p450	
GADD45A	Up	DNA repair, apoptosis, cell cycle arrest, and signal transduction in response to DNA damage	MAPK signaling Pathway And Cell cycle	Cell cycle G2/M checkpoint, ATM signaling pathway, P53 signaling pathway Hypoxia and P53 in the cardiovascular system
PHLDA1	Up	Induction of Apoptosis		

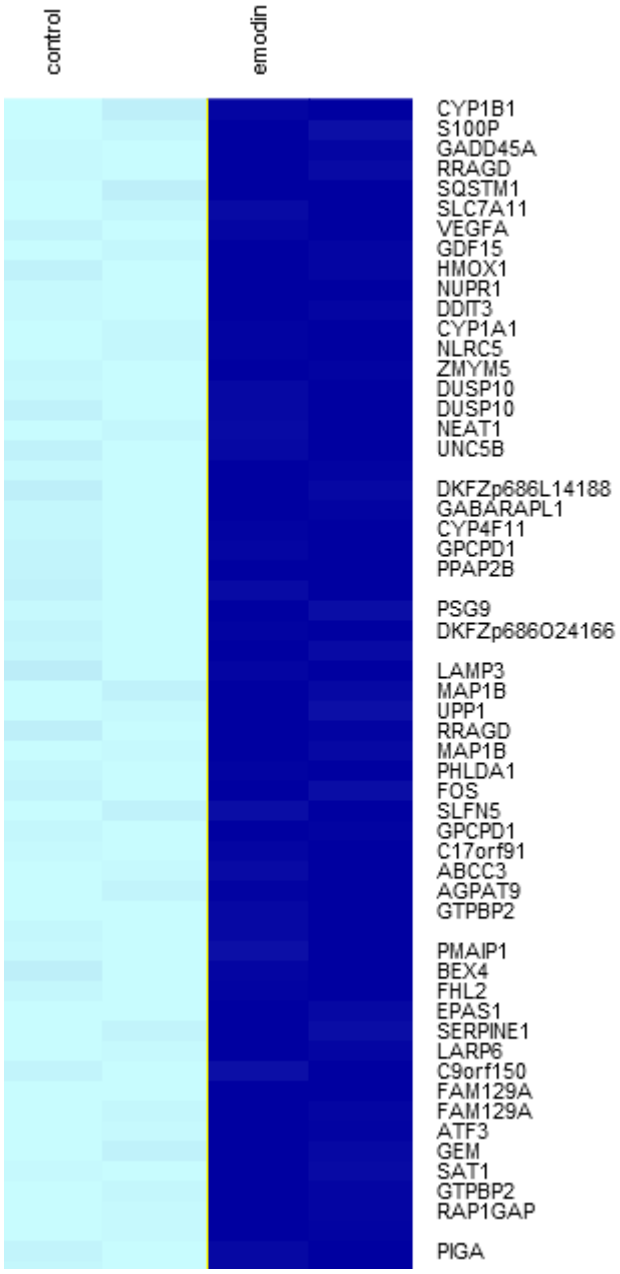


Figure 3.14; continue in the next page.

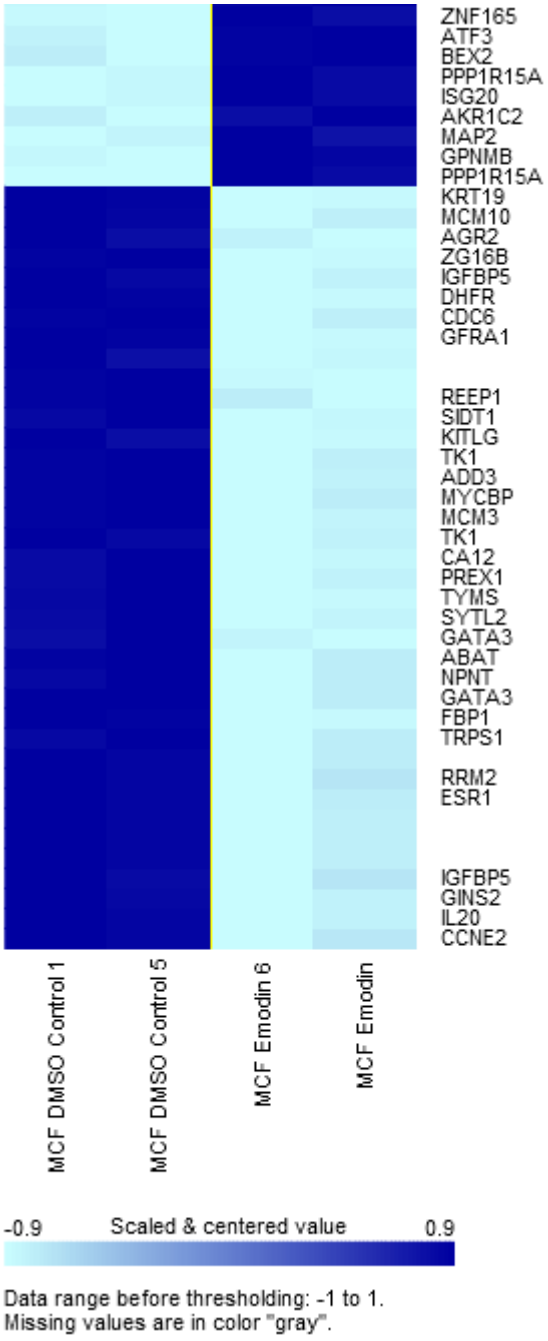


Figure 3.14: Clustered heatmap of significantly expressed genes from the analysis of the 4 arrays of MCF7 cells type, arrays grouped by class.

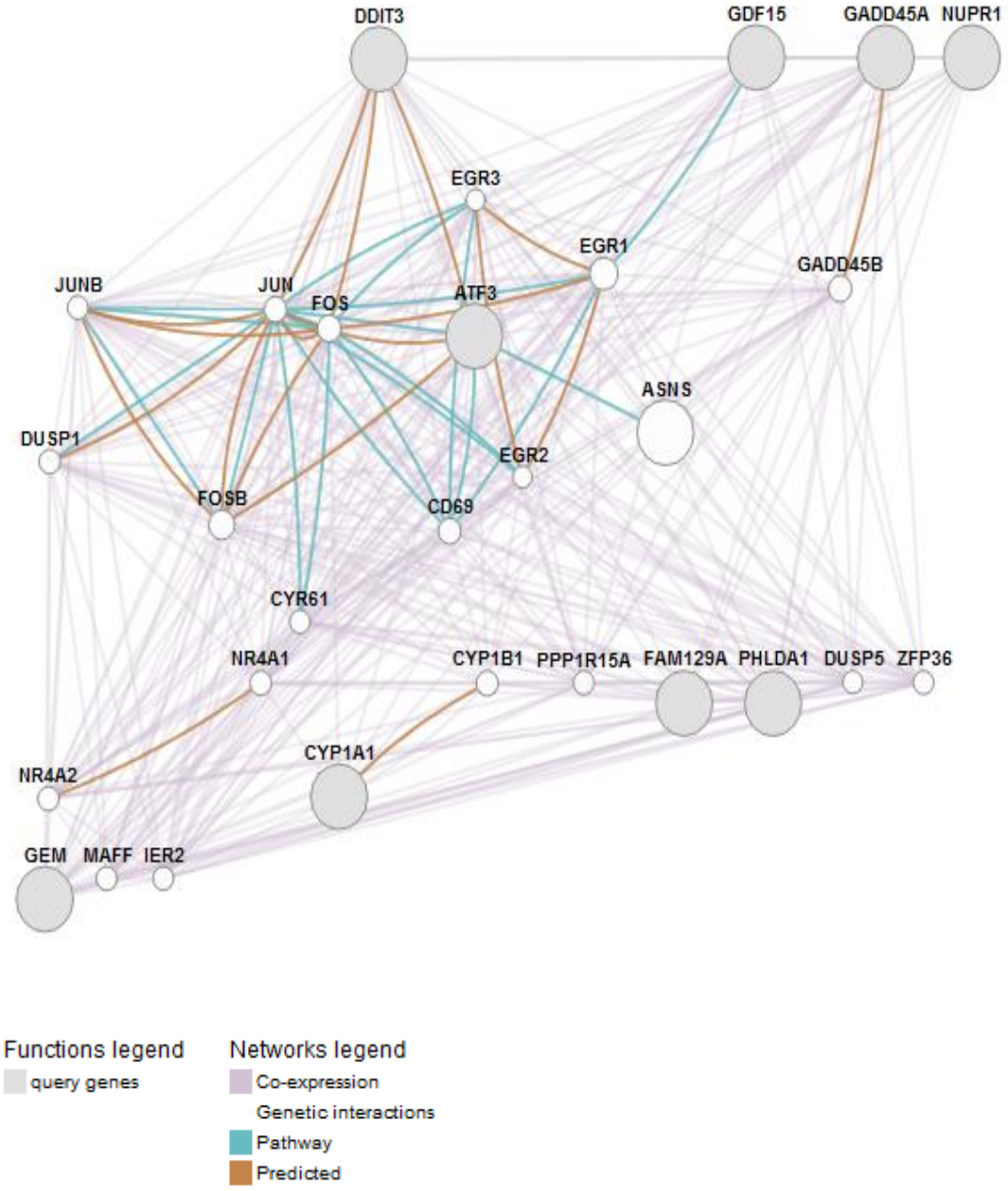


Figure 3.15: network analysis done for the top 10 regulated genes result from the analysis of the 4 arrays of the MCF7 cells .

Table 3.7: The Analysis of the predicted genes from network analysis result in figure 3.15.

Predicted gene	Up/down regulated after emodin treatment in the analysis of 4 arrays for MCF7 cells	The most related To tumor growth Biological processes From Gene Ontology
JUN	UP	release of cytochrome c from mitochondria toll-like receptor signaling pathway Toll signaling pathway negative regulation of cell proliferation stress-activated MAPK cascade
FOS	UP	toll-like receptor signaling pathway inflammatory response Toll signaling pathway cellular response to hormone stimulus toll-like receptor 3 signaling pathway stress-activated MAPK cascade
EGR3	UP	transcription, DNA-dependent regulation of transcription
EGR1	UP	transcription, DNA-dependent positive regulation of transcription
ASNS	UP	response to toxin cellular response to hormone stimulus negative regulation of apoptosis positive regulation of mitotic cell cycle
GADD45B	UP	activation of MAPKKK activity activation of MAPKK activity Apoptosis response to stress
CYP1B1	UP	xenobiotic metabolic process estrogen metabolic process toxin metabolic process
DUSP1	UP	inactivation of MAPK activity response to estradiol stimulus cellular response to hormone stimulus positive regulation of apoptosis positive regulation of anti-apoptosis
CYR61	UP	regulation of cell growth insulin-like growth factor binding cell proliferation
NR4A1	UP	regulation of transcription, DNA-dependent induction of apoptosis

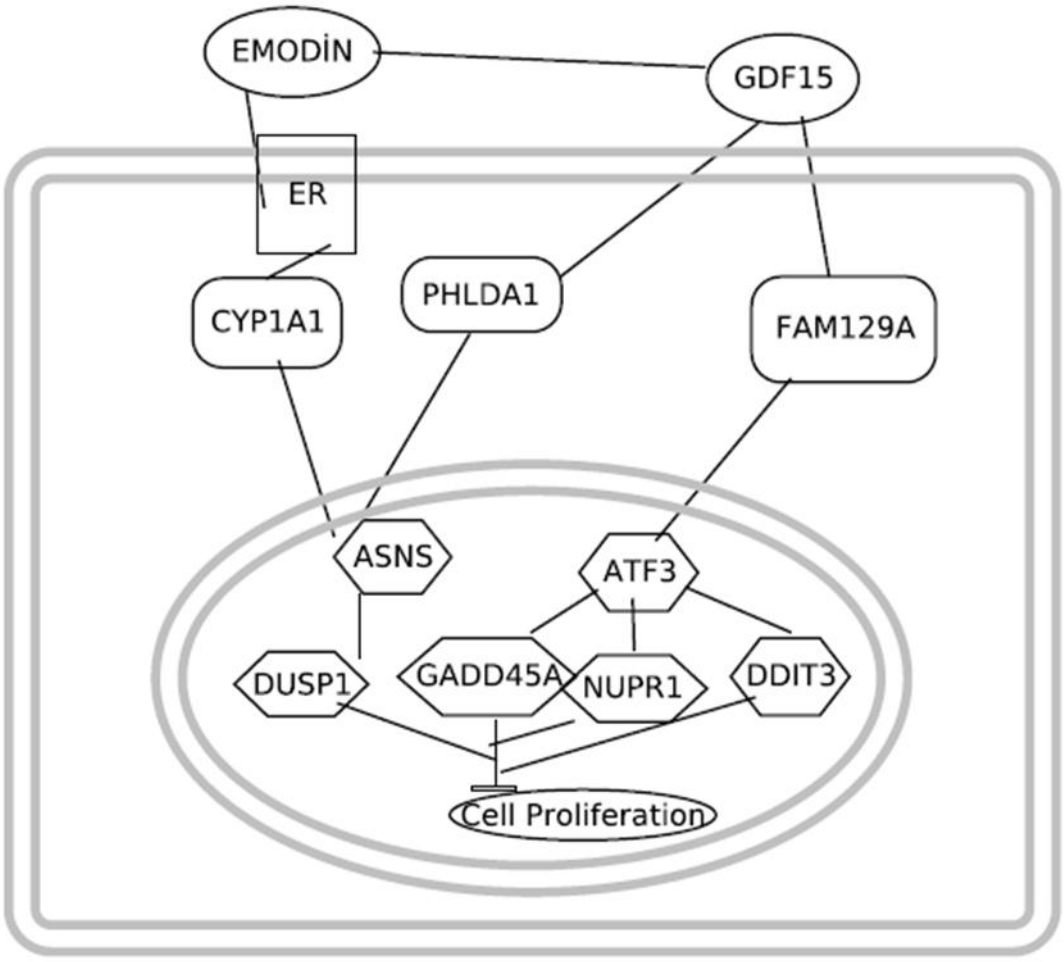


Figure 3.16: The proposed mechanism of the effect of emodin on MCF-7 cell lines, pathway drawn using Wiki Pathways.

3.4.3 ANALYSIS OF THE TOP 10 REGULATED GENES IN THE 4 ARRAYS FOR MDA-MB-231 CELLS

In MDA-MB-231 cells, the effect of emodin seen in the top 10 regulated genes that their biological process contain cell proliferation and there's two important pathways for cancer, Jak-Stat signaling pathway which affected by the up regulation of IL24 gene, several studies have shown that cell death occurs in cancer cells or cell lines following exposure to IL-24 (http://en.wikipedia.org/wiki/IL24). And apoptotic DNA fragmentation and tissue homeostasis pathway which refer to TOP2A gene in the list, this gene show down regulation. This results insure the anticancer effect of emodin on the cancerous cells. The network analysis result also support this, as all

the predicted genes when analyzed; showed biological processes like cell division and mitotic cell cycle and Wnt receptor signaling pathway, see Figure 3.23, and all of the genes was down regulated.

From the heat maps which clustered between the two classes, emodin and control, we can see that the significant regulated genes number in MCF7 more than on MDA-MB-231 cells, which as mentioned on the MCF cells analysis, it confirms that emodin is a phytoestrogenic component of emodin as MCF-7 cells are ER positive cells, so the number of genes affected is more than the genes changed on MDA-MB-231 cells. See Figures 3.14 and 3.17.

A proposed mechanism of the effect of emodin on MDA-MB-231 cell lines pathway drawn using Wiki Pathways, and need to be confirmed in the future studies.

Table 3.8: Summary of the Annotation of the top 10 Genes which are differentially expressed among emodin and Control Classes in analysis of 4 arrays of MDA-MB-231 cells.

Top regulated after emodin treatment of MDA cells	UP/down regulation	The most related To tumor growth Biological processes From Gene Ontology	Kegg pathways	Biocarta pathways
HSPA6	Up	Response to unfolded protein		
PTGS2	Up	Prostaglandin biosynthetic process, xenbiotic metabolic process, response to oxidative stress, negative regulation og cell proliferation, response to estradiol stimulus	Arachidonic acid metabolism	Mechanism of acetaminophen activity and toxicity, Mechanism of gene regulation by peroxisome proliferation via PPARa(alpha), And Eiocosanid metabolism
IL24	Up	Apoptosis	Cytokine-cytokine receptor interaction, And Jak-STAT signaling pathway	
HSPA6	Up	Response to unfolded protein		
DEPDC1B	Down	Regulation of small GTPase mediated signal transduction		
KIF14	Down	Microtuble based movement		
KIF23	Down	Mitotic cell cycle		
HSPA1A	Up		Antigen processing and presentation, MAPK signaling pathway	Hypoxia and p53 in the cardiovascular system, Mechanism of gene Regulation by Peroxisome Proliferators via PPARa(alpha), And Chaperones modulate interferon signaling Pathway
TOP2A	Down	DNA Repair, DNA replication, positive regulation of apoptosis, and mitotic cell cycle G2/M transition decatenation checkpoint		Apoptotic DNA fragmentation and tissue homeostasis.

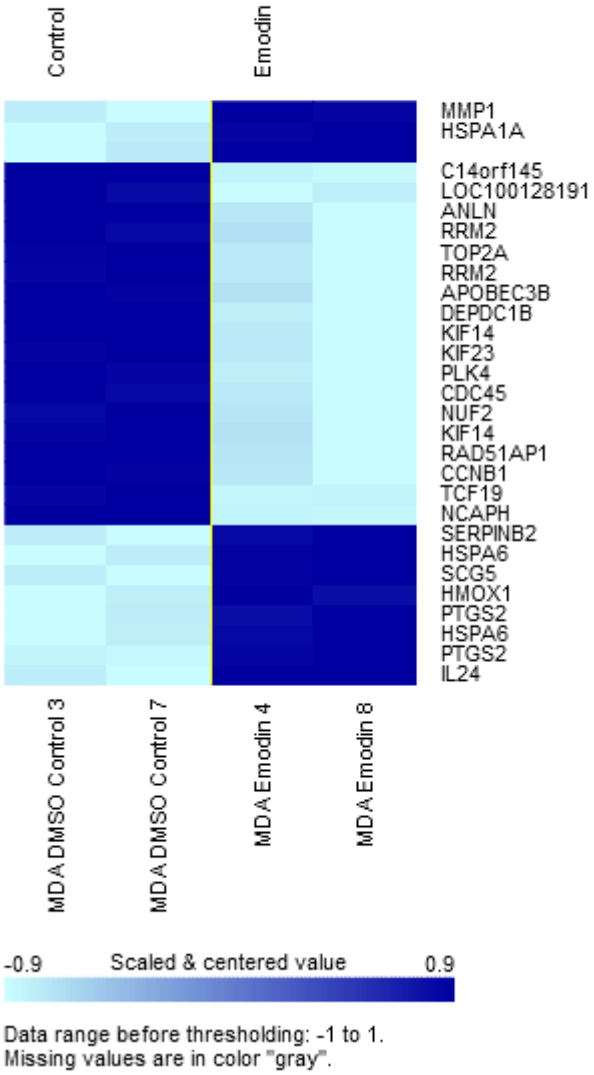


Figure 3.17: Clustered heatmap of significantly expressed genes from the analysis of the 4 arrays of MDA-MB-231 cells type, arrays grouped by class.

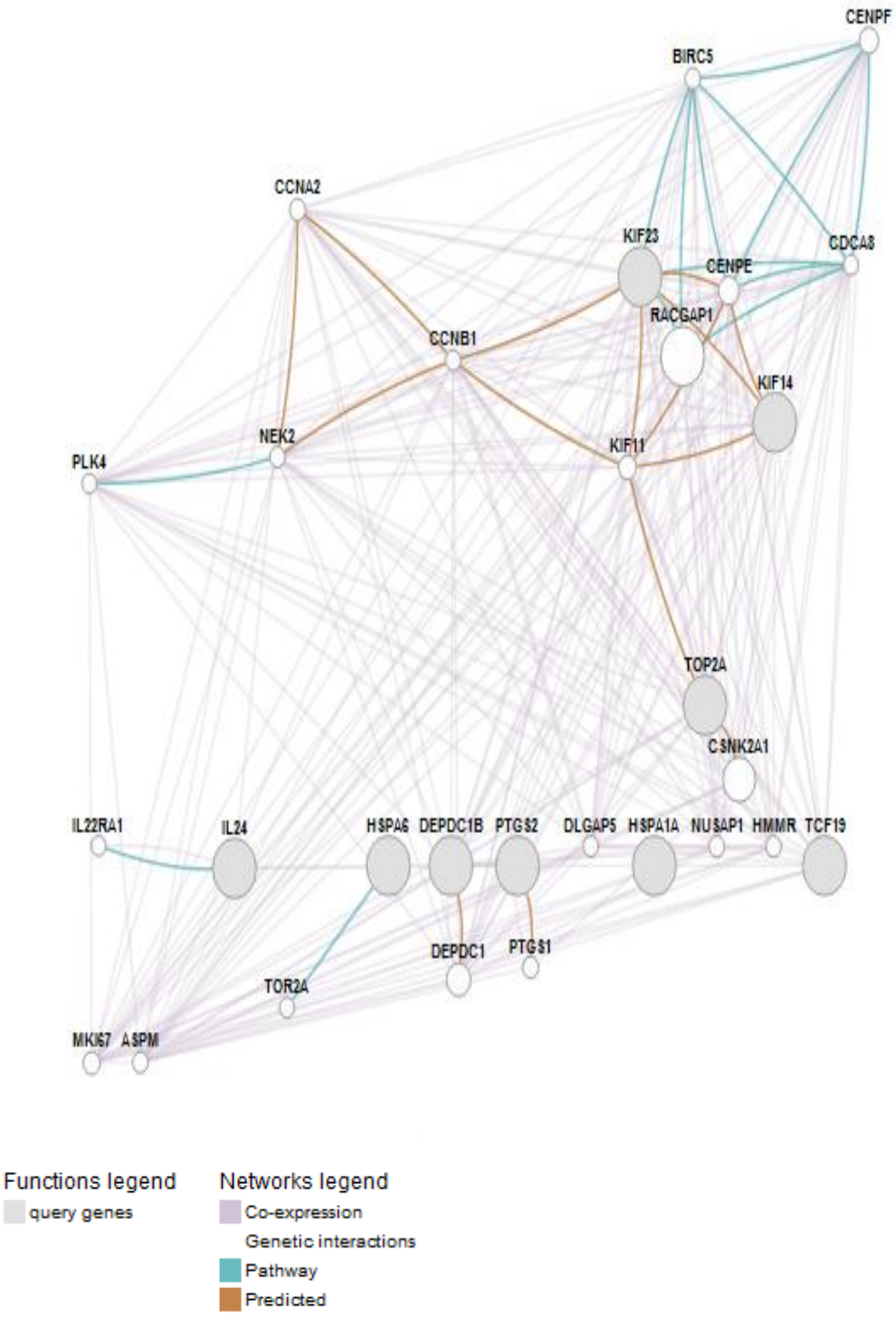


Figure 3.18: Network analysis done for the top 10 regulated genes result from the analysis of the 4 arrays of the MDA-MB-231 cells.

Table 3.9: The Analysis of the predicted genes from network analysis result in figure 3.18.

Predicted gene	Up/down regulated after emodin treatment in the analysis of 4 arrays for MDA-MB-231 cells	The most related To tumor growth Biological processes From Gene Ontology
CSNK2A1	DOWN	Wnt receptor signaling pathway
CENPF	DOWN	mitotic cell cycle mitotic cell cycle spindle assembly checkpoint cell proliferation cell division
CDCA8	DOWN	mitotic cell cycle cell division
CENPE	DOWN	mitotic cell cycle spindle assembly checkpoint cell division
BIRC5	DOWN	Apoptosis anti-apoptosis positive regulation of mitotic cell cycle cell division
RACGAP1	DOWN	cell cycle
KIF11	DOWN	cell cycle, and cell division
PLK4	DOWN	mitotic cell cycle
CCNB1	DOWN	negative regulation of gene expression cell division regulation of cell cycle
NEK2	DOWN	regulation of mitosis, Meiosis, and cell division

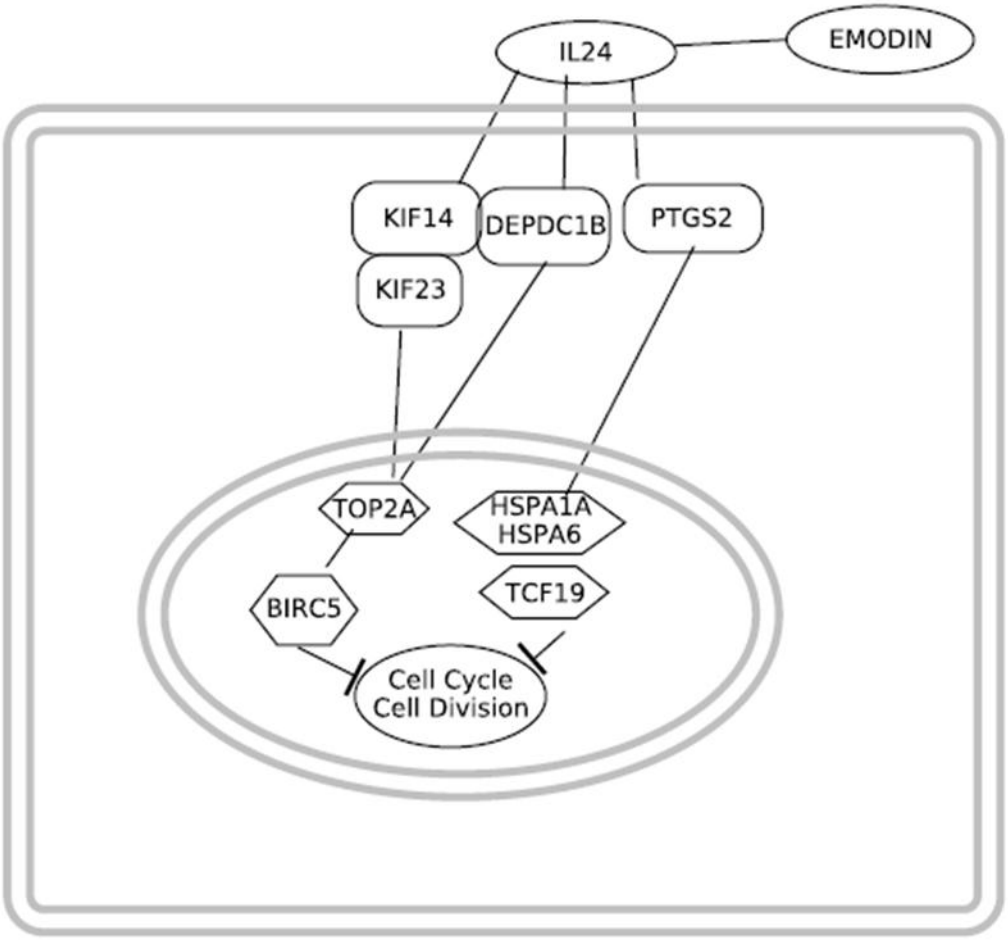


Figure 3.19: The proposed mechanism of the effect of emodin on MDA-MB-231 cell lines, pathway drawn using Wiki Pathways.

3.5 TOP 10 PATHWAYS FROM DAVID ANNOTATION TOOL

The top 10 pathways Result shows pathways related to Tumor Growth and Cancer see Table 3.10. The top Pathway which is cell cycle, was involve 77 gene from our regulated gene list, other pathways like DNA replication, Oocyte meiosis and p53 signaling pathway. These results confirmed the relation of emodin to cell growth processes on Breast Cancer Cell lines.

Table 3.10: The Top 10 pathways From DAVID Annotaion Tool.

Category	Pathway name	Gene Number	P-Value
KEGG Pathway	Cell cycle	77	2.0E-20
KEGG Pathway	DNA replication	31	4.6E-14
KEGG Pathway	Oocyte meiosis	57	6.6E-11
KEGG Pathway	P53 signaling Pathway	34	5.3E-6
KEGG Pathway	Spliceosome	52	8.1E-6
KEGG Pathway	Progesterone-mediated oocytematuration	38	2.4E-5
KEGG Pathway	Pyrimidine metabolism	41	3.0E-5
KEGG Pathway	Valine, Leucine and isoleucine degradation	24	3.2E-5
KEGG Pathway	Prostate cancer	39	4.3E-5
KEGG Pathway	Mismatch repair	15	6.9E-5

3.6 Clustering between all samples and genes for pattern discovery.

Cluster between all samples and genes done of the analysis of the 8 arrays from both MCF7 and MDA-MB-231 cells and in general it's clear that there's two major patterns between emodin and controls which shown clearly in the heatmap Figure 3.20, for example; from genes (1 - 920), and the second pattern between MDA and MCF cells from genes for example from genes (1839 - 4596).

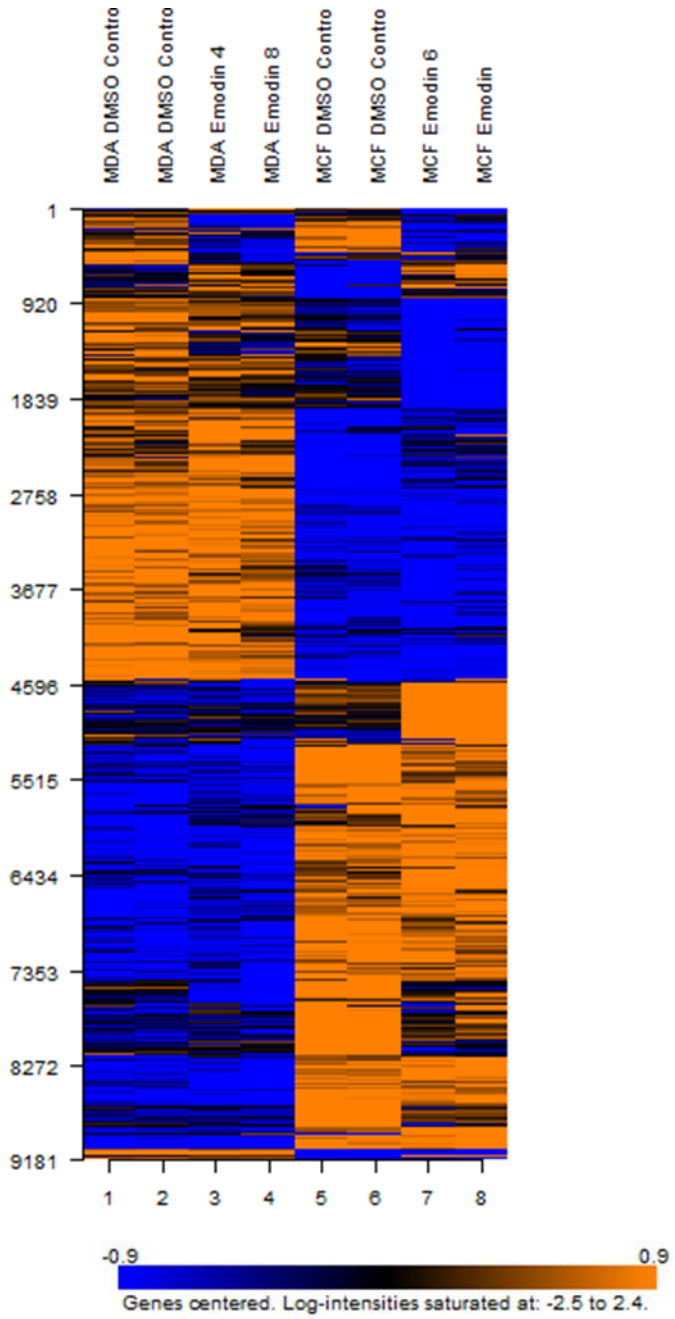


Figure 3.20: hierarchical cluster/ average linked between all sample and genes of the analysis of 8 arrays for both MCF7 and MDA-MB-231.

3.7 Our gene of interest list

From the result down in the three tables 3.11, 3.12, and 3.13, which show the change of GST enzymes genes after emodin treatment in the three analysis done. From analysis of these tables we can see that MCF7 shows more number of GST gene classes changed, and they show down regulation change like (GSTZ1, GSTT1, GSTM3, GSTM4 and GSTO1), except GSTP1 was up regulated, GSTP1 shows anti-apoptosis properties in the biological process, see Figure 3.25.

the intersect of GST classes between MCF7 and MDA-MB-231 is GSTO1 and GSTM3 but it showed different regulation in the two cancerous cells, except GSTCD was down regulated in the both cells type, trying to understand this change by analyzing the annotation of the genes, the biological process of GSTCD is rRNA processing, this is down regulated in both MDA-MB-231 and MCF7, GSTM3 biological processes is response to estrogen stimulus so this may be the reason for the different change, and from estrogen pathway, we know that estrogen related to breast cancer and it considered a carcinogenic. And for the GSTO1 its Molecular Function is monodehydro ascorbate reductase (NADH) activity, which have a role in triggering apoptosis, because of this we consider the down regulation in MCF for GSTO1 and GSTM3 a good result toward anti-cancer effect.

The up regulation of GST in MDA-MB-231 cells like GSTA1, GSTA4, GSTM3 and GSTO1, may interpretate by this text which taken from (McIlwain, et al.,2006); Chemotherapeutic-resistant tumor cell lines have been shown to overexpress GST isozymes. This overexpression leads to an accelerated detoxification of drug substrates and thus an acquired resistance. However, drug resistance is exhibited in cells expressing certain isoforms of GSTs even when that specific selecting drug is not an enzyme substrate. This anomaly may be explained by the ability of GSTs to act as ligand-binding proteins in the regulation of cell cycle components such as mitogen-activated protein kinases (MAPK) and extracellular-regulated kinases (ERK), see Figure 3.24 which show example of GST-mediated kinase regulation for GSTM1.

From the network analysis we can see that one of the genes related to xenobiotic metabolic process and oxidation-reduction process show up regulation in MCF-7 and Down regulation in MDA-MB-231 this is explained by the different regulation of some GSTs processes after emodin treatment between the two cells, as GSTs are included in these processes. And FIS1 gene which predicted from co-expression network, and its include apoptosis in its biological processes; was down regulated after emodin treatment.

Table 3.11: GST Isozymes regulation in MCF7 and MDA cells after emodin treatment.

Probeset	Gene symbol	Up/down regulated	1.25 Fold change
235387_at	GSTCD	Down	5
1554518_at	GSTCD	Down	4
204149_s_at	GSTM4	Down	4
209531_at	GSTZ1	Down	3
200824_at	GSTP1	Up	6
202967_at	GSTA4	Up	4

Table 3.12: GST Isozymes regulation in MCF7 cells after emodin treatment

Probeset	Gene symbol	Up/down regulated	1.25 Fold change
200824_at	GSTP1	Up	2
1554518_at	GSTCD	Down	2
1557915_s_at	GSTO1	Down	1
201470_at	GSTO1	Down	1
202554_s_at	GSTM3	Down	1
204149_s_at	GSTM4	Down	4
209531_at	GSTZ1	Down	1
232193_at	GSTT1	Down	2
235387_at	GSTCD	Down	2

Table 3.13: GST Isozymes regulation in MDA MB 231 cells after emodin treatment

Probeset	Gene symbol	Up/down regulated	1.25 Fold change
1557915_s_at	GSTO1	Up	2
202967_at	GSTA4	UP	4
215766_at	GSTA1	Up	2
235867_at	GSTM3	Up	2
235387_at	GSTCD	Down	2

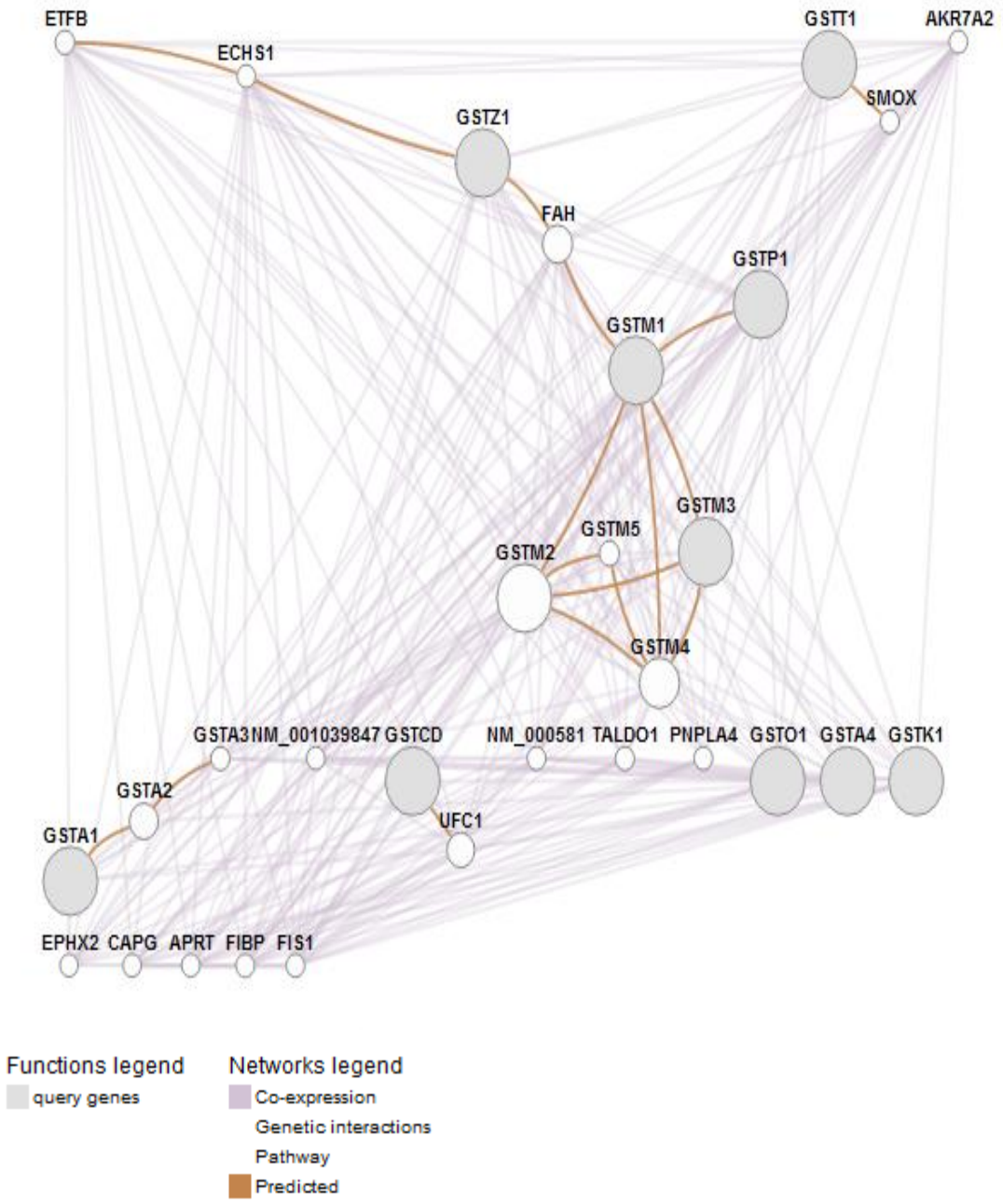


Figure 3.21: network analysis done for the our genes of interest GST enzyme family.

Table 3.14: Summary of the Annotation of the predicted genes from network analysis result in figure 3.21.

Predicted gene	Up/down regulated after Emodin treatment in the analysis of 8 arrays for both cells type	The most related To tumor growth Biological processes From Gene Ontology
PNPLA4	Up regulated in MCF7 and Down regulated in MDA-MB-231.	metabolic process lipid catabolic process
ECHS1(found in the gene list of the analysis of 4 arrays of MCF7)	Down just in MCF7	fatty acid metabolic process cellular lipid metabolic process
SMOX	Up regulated in MCF and Down regulated in MDA	xenobiotic metabolic process oxidation-reduction process
AKR7A2	Down	oxidation-reduction process
FIS1	Down	mitochondrial fission Apoptosis peroxisome fission Cellular Component: integral to mitochondrial outer membrane
FIBP	Down	fibroblast growth factor receptor signaling pathway
APRT	Down	 purine-containing compound salvage nucleobase, nucleoside and nucleotide metabolic process
CAPG	UP	protein complex assembly Biological Process: cell projection assembly barbed-end actin filament capping
EPHX2(found in the gene list of the analysis of 4 arrays of MCF7)	Down just in MCF7	xenobiotic metabolic process inflammatory response response to toxin

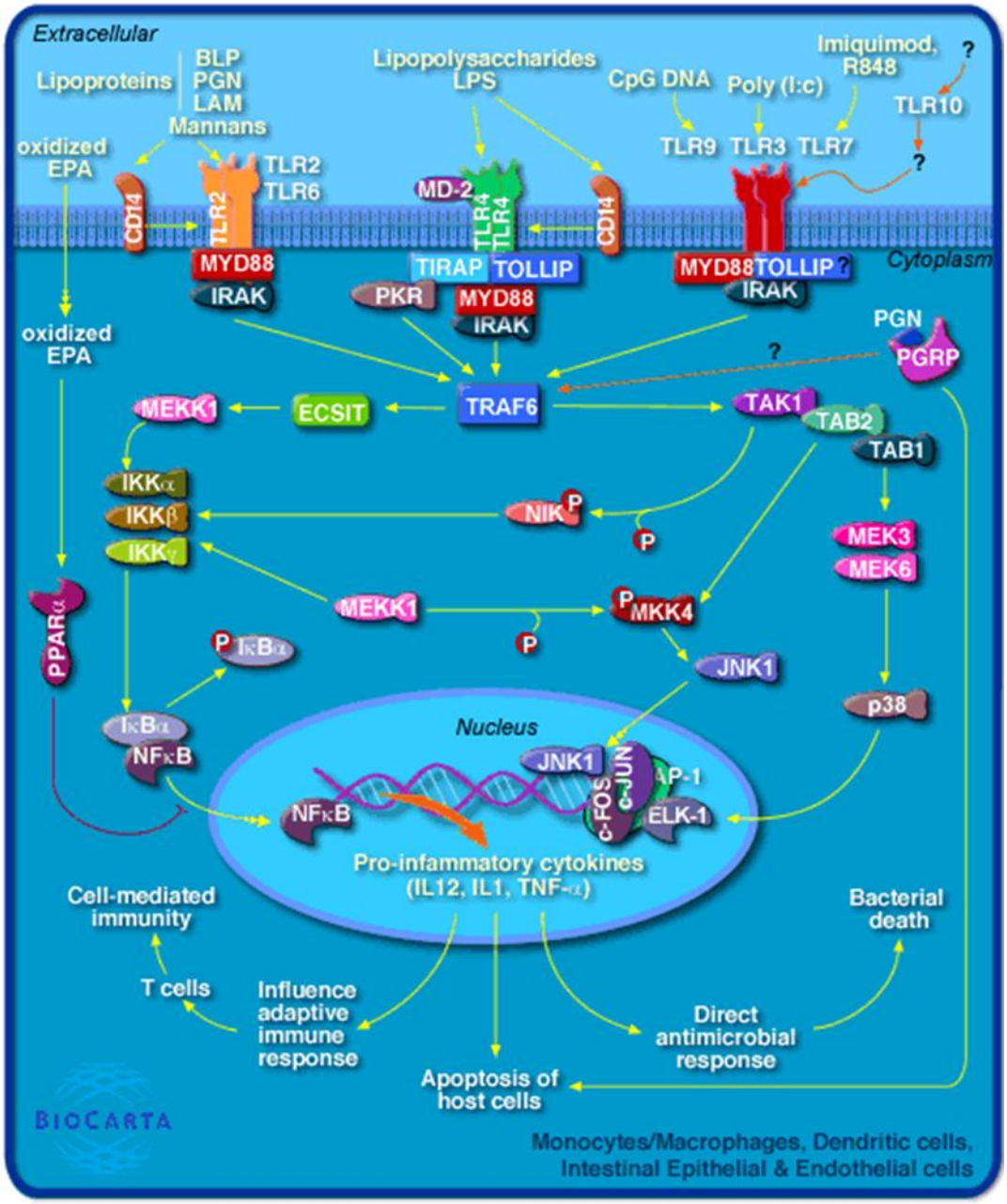


Figure 3.22: Toll–like receptor signaling Pathway. From BioCarta.

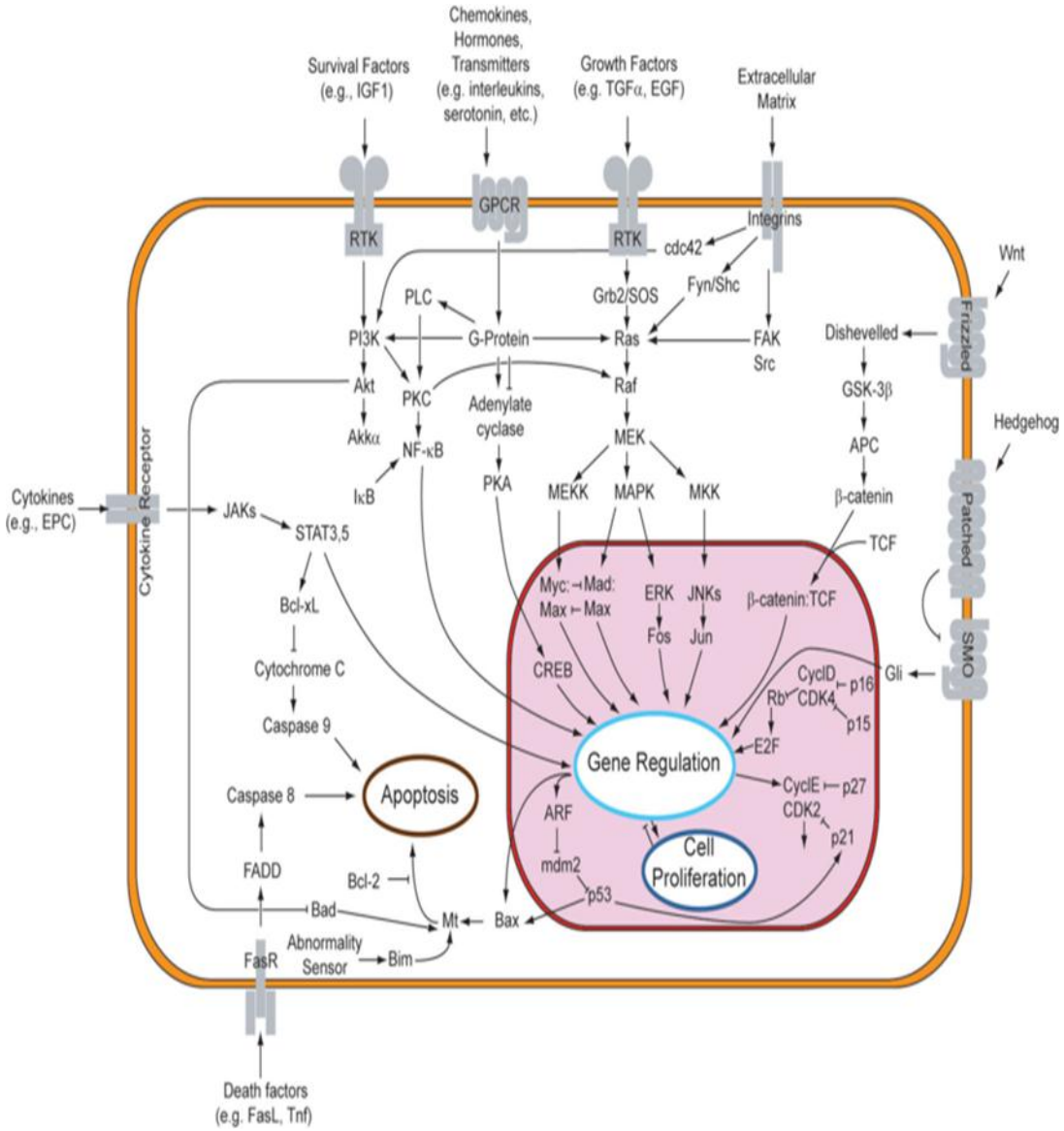


Figure 3.23: Signal transduction pathways.
(http://en.wikipedia.org/wiki/Signal_transduction_pathways)

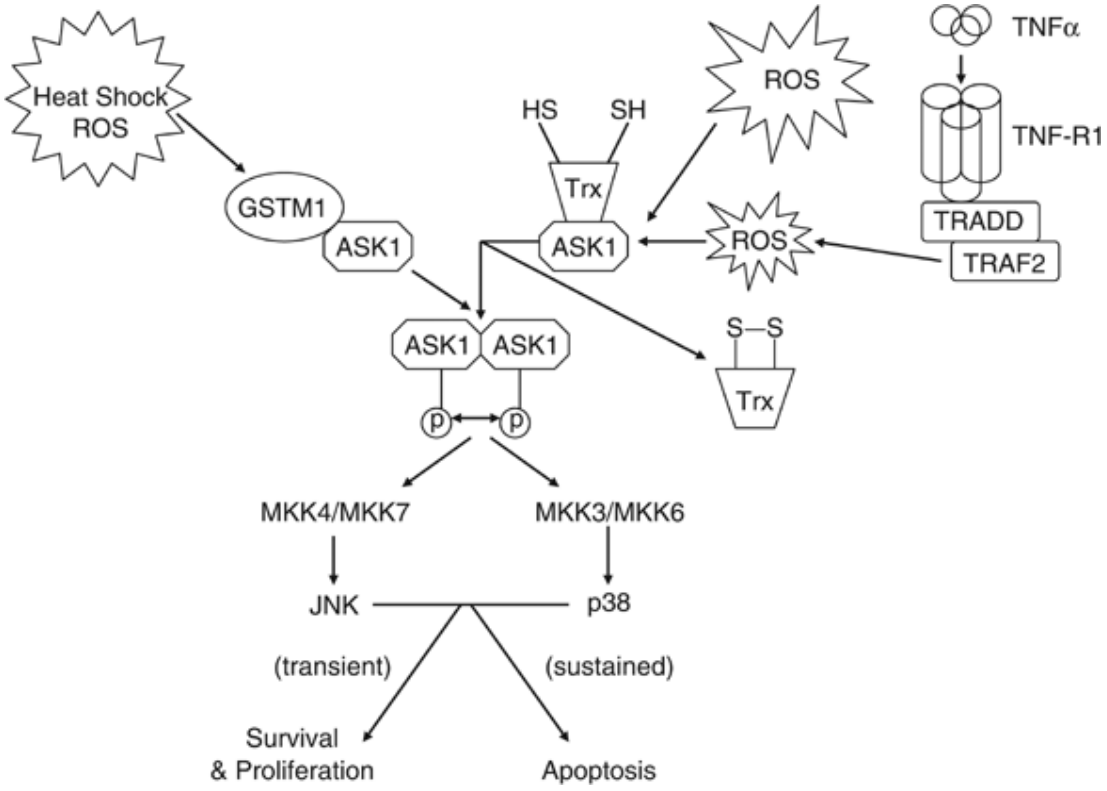


Figure 3.24: Example of GST-mediated kinase regulation for GSTM1. GSTm and thioredoxin (Trx) can act as inhibitors of ASK1. Stresses such as heat shock or reactive oxygen species can resultin the release of ASK1 from the GSTm:ASK1 or TRX:ASK1 complex (respectively). ASK1 oligomerizes and is activated through autophosphorylation, which in turn activates downstream kinases such as MKK4/MKK7, MKK3/MKK6, JNK and p38. The fate of the cell (either proliferation or apoptosis) is dependent upon the time/concentration exposure to the stress. (McIlwain, et al.,2006).

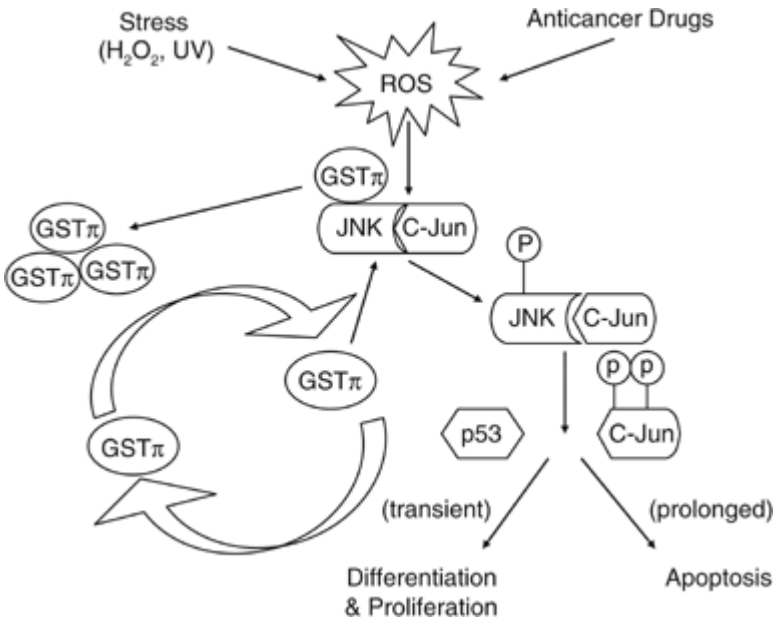


Figure 3.25: Other Example of GST-mediated kinase regulation for GSTP1. (McIlwain, et al.,2006).

3.8 CHECKING THE GST PRIMERS DESIGNED FOR REAL TIME PCR

3.8.1 RNA ISOLATION RESULT

The MCF7 concentration was: 995.55 ng/ul, A260/280=2.09, the MDA-MB-231 concentration was: 1155.4 ng/ul, A260/280=2.10. The detection of ribosomal RNA band by electrophoresis results For the Both cells are shown in figure 3.26, the nano drop results; 280/260 results also between the ranges (1.8-2.1).

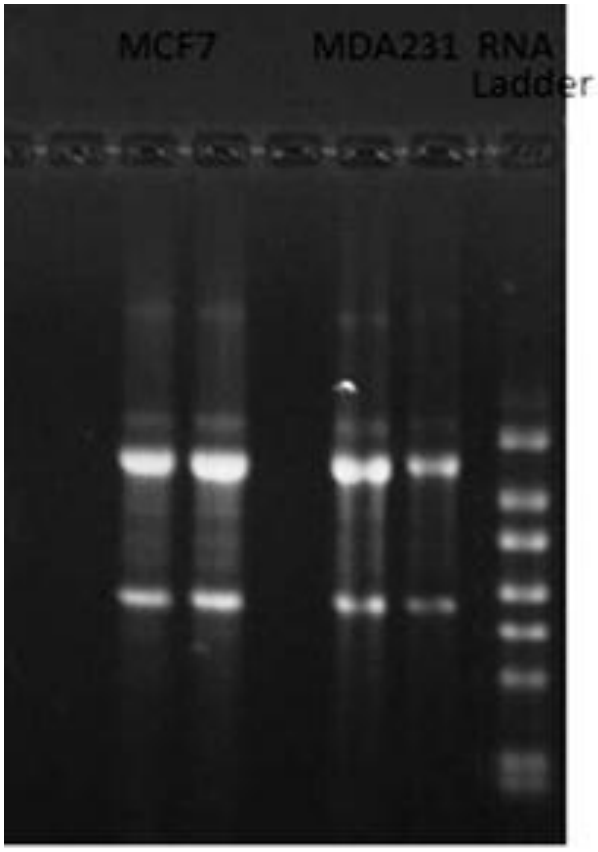


Figure 3.26: RNA Gel Electrophoresis.

3.8.2 PCR RESULT

The PCR result on the Gel electrophoresis in Figures 3.27, and 2.28; in MCF-7 cells (figure 3.27), we can see that GSTO1, GSTZ1, GSTM3, GSTT1 give bands on their product length, when the GST gene checked on the Microarray analysis of MCF-7 cells; in the control MCF-7 cells; GSTA1 and GSTA4 was not in our gene list after Normalization and filtering, this may explain that there were no Result for them. The intensity of the genes also checked and the log filtered intensity value was:

For GSTM3≈ 12.6
For GSTZ1≈ 10.75
For GSTT1≈ 7.6
For GSTO1≈11.4

GSTT1 intensity was less than the others so because of this its band is weak in the Gel photo.

For MDA-MB-231 cells in Figure 3.28 also all the Product give bands on their product length just GSTA1 its product was less than it should be, like in MCF-7 result. And the GST gene checked on the Microarray analysis of MDA-MB-231 cells; in the control MDA-MB-231cells; GSTA1 and GSTT1 was not in our gene list after Normalization and filtering, this may explain that in the gel photo it give a fade band. The intensity of the genes also checked and the log filtered intensity value was:

For GSTA1≈ 7.3
For GSTA4≈ 9
For GSTM3≈ 12.6
For GSTZ1≈ 10.75
For GSTO1≈11.4

GSTA4 intensity was less than the others so because of this its band is weak comparing to the others in the Gel photo.

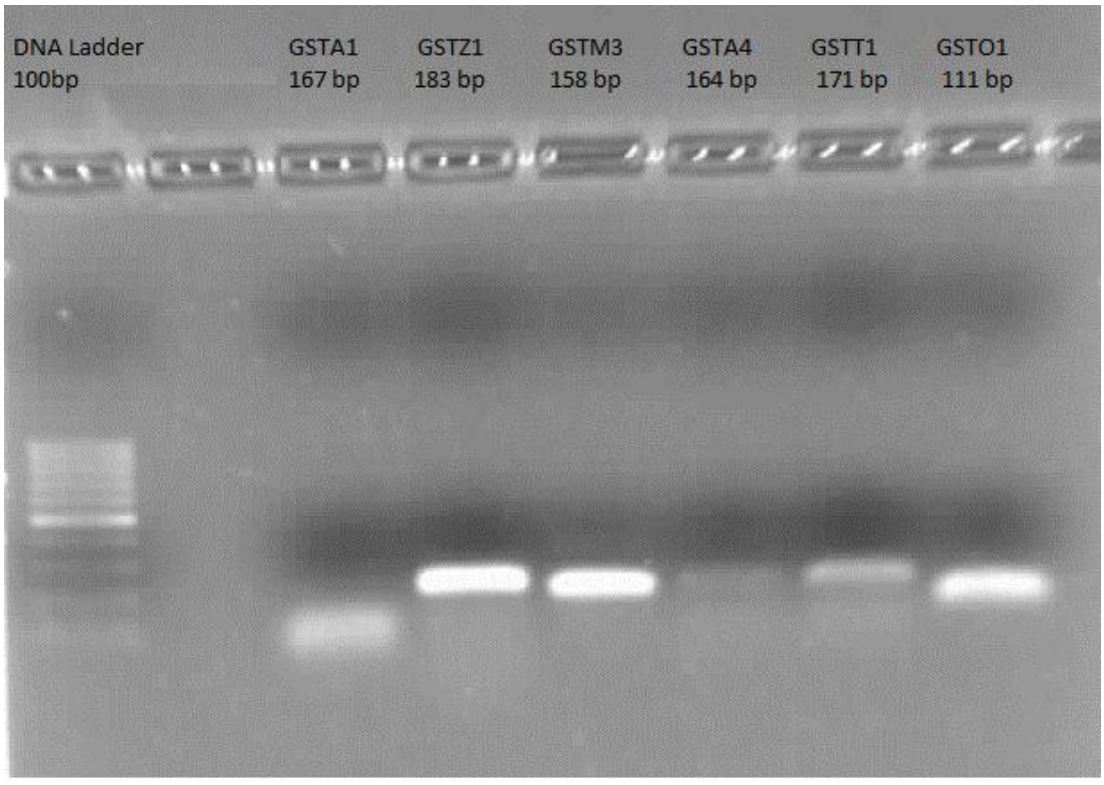


Figure 3.27: PCR Result on Gel Electrophoresis for MCF7 cells.

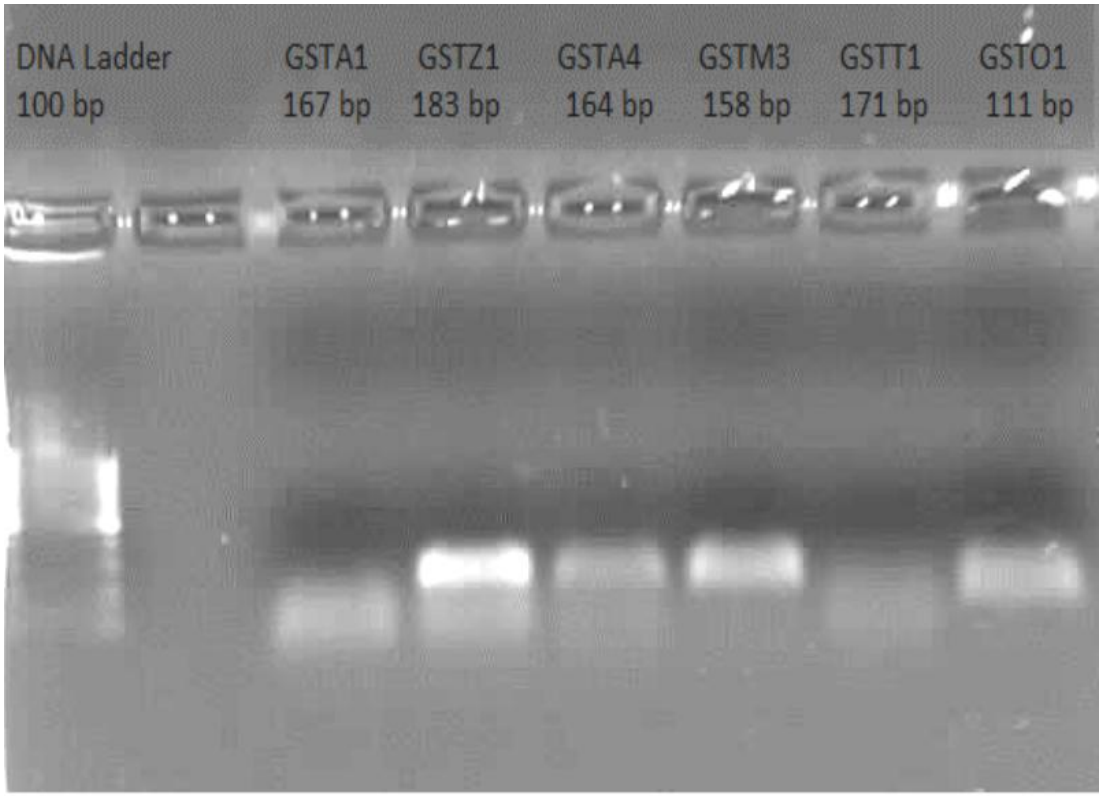


Figure 3.28: PCR Result on Gel Electrophoresis for MDA-MB-231 cells.

3.8.3 DNA Sequencing

The DNA sequencing result were taken and using BLAST Alignment tool were searched and the result on Table 3.15 showed that the sequencing of every PCR product of primer used the result was its gene, so we confirmed our Primers.

Table 3.15: The BLAST Alignment Search Result for the DNA Sequencind done for the PCR Products.

PCR Product Sample	Description Result	Query coverage	E-value	Max İdentity
MDA GSTA4	Homo sapiens glutathione S-transferase alpha 4 (GSTA4)	91%	2e-53	98%
MDA GSTA1	Homo sapiens glutathione S-transferase alpha 1 (GSTA1)	77%	4e-19	80%
MCF GSTM3	Homo sapiens glutathione S-transferase mu 3 (Brain) (GSTM3)	43%	2e-49	96%
MCF GSTO1	Homo sapiens glutathione S-transferase omega 1 (GSTO1)	45%	8e-14	91%
MCF GSTT1	Homo sapiens glutathione S-transferase theta 1 (GSTT1)	57%	3e-62	95%
MCF GSTZ1	Homo sapiens glutathione S-transferase zeta 1 (GSTZ1)	27%	2e-51	97%

CHAPTER 4

CONCLUSION

This study focus into microarray analysis to see the effect of emodin treatment to the breast cancer cell lines, MCF-7 and MDA-MB-231, at molecular level. In breast cancer cell lines treated with emodin, the genes whose expressions are highly varied as compared to untreated control cells are annotated (Table A.1). It has been shown that top 10 genes mostly belong to the biological functions such as cell division, cell proliferation, and cell cycle. Nine of these genes are down regulated, indicating the suppression of related biological functions by emodin. The network analysis performed using those 10 significantly regulated genes, and the predicted genes which are involved in the similar biological functions were analyzed and their variation after emodin treatment was toward the anti-tumor effect (Table A.2).

The comparison of emodin treated MCF7 cells to their untreated controls has shown the induction of gene expressions playing a role in apoptosis, positive regulation of translation and transcription, and in cell cycle arrest. Similar comparison of gene expressions in MDA-MB-231 cells has exhibited up and down regulation of completely different set of genes (Table A.5), although the genes are also involved in similar biological activities. An important pathway, namely Jak-Stat, is worth to examine in detail in future studies.

The GST isozyme compositions and the changes in their gene expressions are found different upon emodin treatment of MCF-7 and MDA-MB-231 cell lines. GSTP and GSTM classes are known anti-apoptotic. Emodin treatment results in slight up regulation of GSTP1 in MCF-7 cell line indicating insignificant role of GSTP in regulating apoptosis in these cells. However, GSTM4 is significantly down regulated stimulating apoptosis in MCF-7 cells.

GeneMANIA network analysis Tool has confirmed that highest number of the genes significantly changing after emodin treatment in breast cancer cell lines are involved in cell cycle regulation.

REFERENCES

Affymetrix Inc. GeneChip Expression analysis / Techical Manual . P/N 702232 Rev. 3:14-82; 2009.

(Agilent Technologies), http://www.genomics.agilent.com/CollectionSubpage.aspx?PageType=Product&SubPageType=ProductDetail&PageID=1675, last Visited on 20/ January/ 2012.

Bainbridge, M.N. Gene Expression Microarray Analysis Using Pattern Discovery. Thesis for the Degree of Master of Science in the Department of Computer Science University of Saskatchewan, Canada, 2004.

(BiGCaT Bioinformatics), www.arrayanalysis.org, Last Visited on 20/ January/ 2012.

(BLAST Alignment Tool), http://blast.ncbi.nlm.nih.gov, Last Visited on 28/ February/ 2012.

Crowe, M.L., et al. CATMA: a complete Arabidopsis GST database. Nucleic Acids Research/ Oxford University Press. 31: 156–158; 2003.

David Warde-Farley, D. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. Oxford journals, W214–W220 Nucleic Acids Research. 38; 2010.

(Design PCR Primers Site), http://molbiol-tools.ca/PCR.htm, last Visited on 20/ January/ 2012.

Dettling, M. Supervised learning in very high Dimensional problems with applications to microarray data. A thesis for the Degree of Doctor of mathematics, Swiss federal institute of technology, Zurich, 2004.

Dialyna, I.A., et al. Genetic Polymorphisms and transcriptional pattern analysis of CYP1A1, AhR, GSTM1, GSTP1 and GSTT1 Genes in Breast Cancer. International Journal of Molecular Medicine. 8: 79-87; 2001.

Duggan, D.J., et al. Review: Expression profiling using cDNA microarrays. Nature America Inc. 21; 1999.

El-Shemy, H.A., et al. Antitumor Properties and Modulation of Antioxidant Enzymes Activity by Aloe vera Leaf Active Principles Isolated via Supercritical Carbon Dioxide Extraction. Current Medicinal Chemistry. 17: 129-138; 2010.

Genechip Expression Analysis/ Data analysis fundamentals Manual. Affymetrix; 2002-2004.

(GeneMANIA Network Analysis), http://www.genemania.org/, Last Visited on 20/ January/ 2012.

(German Cancer Research Center), http://www.dkfz.de/gpcf/24.html, last Visited on 20/ January/ 2012.

Greco, D. Gene expression: from microarray to functional genomics. Thesis for public examination of the Viikki Graduate School in Biosciences, University of Helsinki, 2009.

Göhlmann, H., and Talloen, W. Gene Expression Studies using Affymetrix Microarrays. CRC Press, 2009.

Helzlsouer, K.J., et al. Reports: Association Between Glutathione S-Transferase M1, P1, and T1 Genetic Polymorphisms and Development of Breast Cancer. Oxford Journals, 2011.

Huang, Z., Chen, G., and Shi, P. Effects of Emodin on the gene expression profiling of human breast carcinoma cells. Cancer Detection and Prevention. 464-470; 2008 B.

Huang, Z., Chen, G., and Shi, P. Emodin induced Apoptosis in Human Breast cancer Bcap-37 cells through the mitochondrial signaling pathway. Arch Pharm Res. 31:742-748; 2008 A.

Igney, F. H. and Krammer, P.H. Death and anti-death tumor resistance to Apoptosis. Nature reviews cancer. 227-288; April 2002.

Kasper, G. et al. stromelysin-3 over expression enhances tumorigenesis in MCF-7 and MDA-MB-231 breast cancer cell lines: involvement of the IGF-I signaling pathway. BMC Cancer. 7:12; 2007.

Kate Peeters, J.K. Microarray Bioinformatics and Applications in Oncology. thesis for the Department of Bioinformatics, Erasmus University Medical Center, Rotterdam, The Netherlands, 2008.

Khan, J. Bittner, M. L. et al. DNA Microarray technology: the anticipated impact on the study of human disease. BBA. 1423: M17-M28; 1999.

Laborde, E. Review: Glutathione transferases as mediators of signaling pathways involved in cell proliferation and cell death. Cell Death and Differentiation. 17: 1373–1380; 2010.

Laurell, C. Microarray-based Gene Expression Analysis in Cancer Research. Doctoral Thesis Royal Institute of Technology School of Biotechnology, Stockholm, 2006.

Leader, D. Methods for Incorporating Biological Information into the Statistical Analysis of Gene Expression Microarray Data. A thesis for the degree of Doctor of Philosophy in Statistics, The University of Auckland New Zealand, December 2009.

Levin, E.R. Integration of the extracellular action of Estrogen. Mol endocrinal. 19(8):1951-1959; 2005.

Lisha, Z. Tyrosine Kinase Inhibitor Emodin Suppresses Growth of HER-2/neu-overexpressing Breast Cancer Cells in Athymic Mice and Sensitizes These Cells to the Inhibitory Effect of Paclitaxel. Clinical Cancer Research. 5: 343–353; February 1999.

Lu, G.D., et al. Critical role of oxidative stress and sustained JNK activation in aloe-emodin-mediated apoptotic cell death in human hepatoma cells. Carcinogenesis. 28: 1937–1945; 2007.

(MCF-7 cells/ cell culture protocol), http://mcf7.com/, last Visited on 20/ January/ 2012.

McIlwain, CC., et al. Review: Glutathione S-transferase polymorphisms: cancer incidence and therapy. Oncogene. 25: 1639–1648; 2006.

McLachlan, G.J. et al. Analyzing Microarray Gene Expression Data. From Wily Series in Probability and statistics, Wiley interscience, 2004.

Mitrunen, K., et al. Glutathione S-Transferase M1, M3, P1, and T1 Genetic Polymorphisms and Susceptibility to Breast Cancer. Cancer Epidemiol Biomarkers Prev. 10:229-236; 2001.

Montojo, J., et al. GeneMANIA Cytoscape plugin: fast gene function predictions on the desktop. Bioinformatics. 26: 2927–2928; 2010.

Muthukaruppan, A. Gene Expression Analysis in Breast Cancer. A thesis for the degree of Doctor of Philosophy in Obstetrics and Gynaecology, University of Auckland, 2011.

(National Cancer Institute), http://linus.nci.nih.gov/BRB-ArrayTools.html, last Visited on 20/ January/ 2012.

Piaggi, S., et al. Glutathione transferase omega 1-1 (GSTO1-1) plays an anti-apoptotic role in cell resistance to cisplatin toxicity. Carcinogenesis. 31: 804–811; 2010.

PEROU, C.M., et al. Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. Proc. Natl. Acad. Sci. USA. 96: 9212–9217; 1999.

(Primer BLAST), http://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi?LINK_LOC=BlastHome, last Visited on 20/ January/ 2012.

(Product Data Sheet / Cell- Biolabs Inc.), http://www.cellbiolabs.com/sites/default/files/2FD8C527-3048-812A-2EB950C744EB9D73.pdf, last Visited on 20/ January/ 2012.

Rampal, J. B. Microarrays/ Applications and Data Analysis. Human Press, 2nd Edition; 2007.

Rebbeck, T.R. Molecular epidemiology of the human glutathione S-transferase genotypes GSTM1 and GSTT1 in cancer susceptibility. Cancer Epidemiol Biomarkers Prev. 6:733-743; 1997.

Sakally, E. Comparative Effects Of Emodin On Biological Activities Of Mcf-7 And Mda-231 Cell Lines. M.S. Thesis; METU, 2010.

(Scitable by Nature Education), http://www.nature.com/scitable/content/outline-of-a-typical-microarray-experiment-20729, last Visited on 20/ January/ 2012.

Simon,R.M., et al. Design and analysis of DNA Microarray Investigations. New york: springer-Verlag, 2004.

Slonim, D.K. Review: From patterns to pathways: gene expression data analysis comes of age. Nature Publishing Group. 32; 2002.

Song, Q.S. Microarray-based gene set analysis in cancer studies. A thesis submitted for the degree of Doctor of Philosophy in Statistics, The University of Auckland, 2008.

Stekel, D. Microarray Bioinformatics. Cambrigde University Press, 2003.

(The Free Encyclopedia Wikipedia), http://en.wikipedia.org/wiki/Breast_cancer, last Visited on 20/ January/ 2012.

(The Free Encyclopedia Wikipedia), http://en.wikipedia.org/wiki/DNA_Microarray, last Visited on 20/ January/ 2012.

(The Free encyclopedia Wikipedia), http://en.wikipedia.org/wiki/IL24, last Visited on 20/ January/ 2012.

(The Free encyclopedia Wikipedia), http://en.wikipedia.org/wiki/Signal_transduction_pathways, last Visited on 20/ January/ 2012.

(The National Center for Biotechnology Information /NCBI), http://www.ncbi.nlm.nih.gov/, last Visited on 20/ January/ 2012.

(University of Essex), http://cswww.essex.ac.uk/staff/W.Langdon/genechip/, last Visited on 20/ January/ 2012.

(Wiki Pathways), http://wikipathways.org/index.php/WikiPathways, Last Visited on 28/ February/ 2012.

Wit, E. And McClure, J. Statistics for Microarrays/ Design, Analysis and Inference. Wiley interscience, 2004.

Yager, J.D., Davidson, N.E. Review (estrogen carcinogenesis in breast cancer). The New England J of Medicine. 554:70-82; 2006.

Yang, X., et al. DNA methylation in breast cancer. Endocrine-Related Cancer. 8: 115–127; 2001.

APPENDIX A

THE ANNOTATION TABLES

Table A.1: Annotation of the top 10 Genes which are differentially expressed among Emodin and Control Classes in analysis of 8 arrays; MDA and MCF cell. (Continued from page 96 - page105).

Gene Info	Gene Ontology	Biocarta Pathways	Kegg Pathways
ProbeSet: 204033_at **Name:** thyroid hormone receptor interactor 13 **Accession:** NM_004237 **UniGene:** Hs.728869 **Symbol:** TRIP13 **EntrezID:** 9319 **Chromosome:** 5 **Cytoband:** 5p15.33 SOURCE GENECARD **DrugBank:** Query Gene Symbol:TRIP13	**Gene Ontology:** Molecular Function: nucleotide binding Biological Process: pachytene Biological Process: oocyte maturation Cellular Component: male germ cell nucleus Molecular Function: transcription cofactor activity Molecular Function: protein binding Molecular Function: ATP binding Cellular Component:		

Continue of Table A.1:

Gene Info	Gene Ontology	Biocarta Pathways	Kegg Pathways
	Cellular Component: nucleus Biological Process: double-strand break repair Biological Process: transcription from RNA polymerase II promoter Biological Process: reciprocal meiotic recombination Biological Process: male meiosis I Biological Process: female meiosis I Biological Process: spermatid development Molecular Function: nucleoside-triphosphatase activity Molecular Function: identical protein binding		
ProbeSet: 202094_at **Name:** baculoviral IAP repeat containing 5 **Accession:** AA648913 **UniGene:** Hs.514527 **Symbol:** BIRC5 **EntrezID:** 332 **Chromosome:** 17 **Cytoband:** 17q25 SOURCE GENECARD **DrugBank:** Query Gene Symbol:BIRC5	**Gene Ontology:** Biological Process: G2/M transition of mitotic cell cycle Biological Process: M phase of mitotic cell cycle Cellular Component: nuclear chromosome Biological Process: mitotic prometaphase		

Continue of Table A.1:

Gene Info	Gene Ontology	Biocarta Pathways	Kegg Pathways
	Biological Process: mitotic cell cycle Cellular Component: chromosome, centromeric region Biological Process: cytokinesis Molecular Function: cysteine-type endopeptidase inhibitor activity Molecular Function: protein binding Cellular Component: intracellular Cellular Component: nucleus Cellular Component: cytoplasm Cellular Component: centriole Cellular Component: spindle Cellular Component: cytosol Cellular Component: cytosol Cellular Component: cytoskeleton Cellular Component: spindle microtubule		

Continue of Table A.1:

Gene Info	Gene Ontology	Biocarta Pathways	Kegg Pathways
	Cellular Component: cytoplasmic microtubule Biological Process: apoptosis Biological Process: anti-apoptosis Biological Process: chromosome segregation Biological Process: mitosis Molecular Function: microtubule binding Molecular Function: zinc ion binding Molecular Function: zinc ion binding Molecular Function: Ran GTPase binding Molecular Function: tubulin binding Molecular Function: enzyme binding Molecular Function: peptidase inhibitor activity Cellular Component: midbody Cellular Component: interphase microtubule organizing center Biological Process: protein complex localization		

Continue of Table A.1:

Gene Info	Gene Ontology	Biocarta Pathways	Kegg Pathways
	Biological Process: positive regulation of exit from mitosis Biological Process: spindle checkpoint Cellular Component: chromosome passenger complex Molecular Function: identical protein binding Molecular Function: protein homodimerization activity Molecular Function: protein homodimerization activity Molecular Function: caspase inhibitor activity Biological Process: negative regulation of caspase activity Biological Process: negative regulation of caspase activity Biological Process: positive regulation of mitotic cell cycle Molecular Function: metal ion binding Molecular Function: protein heterodimerization activity Molecular Function: cofactor binding		

Continue of Table A.1:

Gene Info	Gene Ontology	Biocarta Pathways	Kegg Pathways
	Molecular Function: cobalt ion binding Molecular Function: chaperone binding Biological Process: cell division Biological Process: cell division Biological Process: establishment of chromosome localization		
ProbeSet: 227165_at **Name:** spindle and kinetochore associated complex subunit 3 **Accession:** AI829603 **UniGene:** Hs.88523 **Symbol:** SKA3 **EntrezID:** 221150 **Chromosome:** 13 **Cytoband:** 13q12.11 SOURCE GENECARD **DrugBank:** Query Gene Symbol:SKA3	**Gene Ontology:** Cellular Component: condensed chromosome outer kinetochore Molecular Function: protein binding Cellular Component: cytoplasm cytoskeleton Cellular Component: spindle microtubule Biological Process: cell cycle Biological Process: chromosome segregation Biological Process: mitosis Biological Process: regulation of microtubule polymerization or depolymerization Biological Process: cell division		

Continue of Table A.1:

Gene Info	Gene Ontology	Biocarta Pathways	Kegg Pathways
ProbeSet: 212949_at **Name:** non-SMC condensin I complex, subunit H **Accession:** D38553 **UniGene:** Hs.308045 **Symbol:** NCAPH **EntrezID:** 23397 **Chromosome:** 2 **Cytoband:** 2q11.2 SOURCE GENECARD **DrugBank:** Query Gene Symbol:NCAPH	**Gene Ontology:** Cellular Component: condensin complex Cellular Component: nucleus Cellular Component: chromosome Cellular Component: cytoplasm Biological Process: cell cycle Biological Process: mitosis Biological Process: mitotic chromosome condensation Cellular Component: microtubule cytoskeleton Biological Process: cell division		
ProbeSet: 223274_at **Name:** transcription factor 19 **Accession:** BC002493 **UniGene:** Hs.584807 **Symbol:** TCF19 **EntrezID:** 6941 **Chromosome:** 6 **Cytoband:** 6p21.3 SOURCE GENECARD **DrugBank:** Query Gene Symbol:TCF19	**Gene Ontology:** Molecular Function: sequence-specific DNA binding transcription factor activity Cellular Component: nucleus Biological Process: transcription, DNA-dependent Biological Process: regulation of transcription from RNA polymerase II promoter		

Continue of Table A.1:

Gene Info	Gene Ontology	Biocarta Pathways	Kegg Pathways
	Molecular Function: zinc ion binding Biological Process: cell proliferation Biological Process: regulation of transcription Molecular Function: metal ion binding		
ProbeSet: 207753_at **Name:** zinc finger protein 304 **Accession:** NM_020657 **UniGene:** Hs.287374 **Symbol:** ZNF304 **EntrezID:** 57343 **Chromosome:** 19 **Cytoband:** 19q13.4 SOURCE GENECARD **DrugBank:** Query Gene Symbol:ZNF304	**Gene Ontology:** Molecular Function: DNA binding Cellular Component: intracellular Cellular Component: nucleus Biological Process: regulation of transcription, DNA-dependent Molecular Function: zinc ion binding Molecular Function: metal ion binding		
ProbeSet: 211713_x_at **Name:** KIAA0101 **Accession:** BC005832 **UniGene:** Hs.81892 **Symbol:** KIAA0101 **EntrezID:** 9768 **Chromosome:** 15 **Cytoband:** 15q22.31 SOURCE GENECARD	**Gene Ontology:** Cellular Component: nucleus Cellular Component: mitochondrion		

Continue of Table A.1:

Gene Info	Gene Ontology	Biocarta Pathways	Kegg Pathways
DrugBank: Query Gene Symbol:KIAA0101			
ProbeSet: 226287_at **Name:** coiled-coil domain containing 34 **Accession:** AI458313 **UniGene:** Hs.143733 **Symbol:** CCDC34 **EntrezID:** 91057 **Chromosome:** 11 **Cytoband:** 11p14.1 SOURCE GENECARD **DrugBank:** Query Gene Symbol:CCDC34			
ProbeSet: 220840_s_at **Name:** chromosome 1 open reading frame 112 **Accession:** NM_018186 **UniGene:** Hs.443551 **Symbol:** C1orf112 **EntrezID:** 55732 **Chromosome:** 1 **Cytoband:** 1q24.2 SOURCE GENECARD **DrugBank:** Query Gene Symbol:C1orf112			
ProbeSet: 213379_at **Name:** coenzyme Q2 homolog, prenyltransferase (yeast) **Accession:** AF091086	**Gene Ontology:** Molecular Function: 4-hydroxybenzoate decaprenyltransferase activity Molecular Function: prenyltransferase activity		**KEGG Pathways:** 1: Ubiquinone biosynthesis

Continue of Table A.1:

Gene Info	Gene Ontology	Biocarta Pathways	Kegg Pathways
UniGene: Hs.144304 **Symbol:** COQ2 **EntrezID:** 27235 **Chromosome:** 4 **Cytoband:** 4q21.23 SOURCE GENECARD **DrugBank:** Query Gene Symbol:COQ2	Cellular Component: mitochondrion Biological Process: glycerol metabolic process Biological Process: ubiquinone biosynthetic process Biological Process: isoprenoid biosynthetic process Biological Process: biosynthetic process Cellular Component: membrane Cellular Component: integral to membrane Molecular Function: transferase activity Cellular Component: mitochondrial membrane Molecular Function: 4-hydroxybenzoate nonaprenyltransferase activity		

Table A.2: The Analysis of the predicted genes from network analysis result in figure 3.13. (Continued from page 106 - page116).

Predicted gene	**Gene Ontology**
NCAPG	Cellular Component: condensin complex Molecular Function: protein binding Cellular Component: Nucleus Cellular Component: Chromosome Cellular Component: Cytoplasm Biological Process: cell cycle Biological Process: Mitosis Biological Process: mitotic chromosome condensation Biological Process: cell division
NCAPD2	Cellular Component: nuclear chromosome Cellular Component: condensed chromosome Cellular Component: condensin complex Cellular Component: condensin core heterodimer Molecular Function: protein binding Cellular Component: Nucleus Cellular Component: Cytoplasm

Continue of Table A.2:

Predicted gene	**Gene Ontology**
	Biological Process: intracellular protein transport Biological Process: cell cycle Biological Process: Mitosis Biological Process: mitotic chromosome condensation Biological Process: vesicle-mediated transport Cellular Component: membrane coat Molecular Function: histone binding Cellular Component: Pronucleus Biological Process: cell division.
SMC4	Biological Process: mitotic sister chromatid segregation Molecular Function: nucleotide binding Cellular Component: condensin complex Cellular Component: condensin complex Molecular Function: ATP binding Cellular Component: Nucleus Cellular Component: Chromosome

Continue of Table A.2:

Predicted gene	**Gene Ontology**
	Cellular Component: Cytoplasm. Biological Process: cell cycle Biological Process: mitotic chromosome condensation Molecular Function: protein heterodimerization activity Biological Process: cell division
SMC2	Molecular Function: nucleotide binding Cellular Component: nuclear chromosome Cellular Component: condensed chromosome Cellular Component: condensin complex Molecular Function: protein binding Molecular Function: ATP binding Cellular Component: Nucleus Cellular Component: Cytoplasm Biological Process: cell cycle Biological Process: Mitosis Biological Process: mitotic chromosome condensation

Continue of Table A.2:

Predicted gene	**Gene Ontology**
	Biological Process: symbiosis, encompassing mutualism through parasitism Molecular Function: protein heterodimerization activity Biological Process: cell division
DLGAP5	Biological Process: M phase of mitotic cell cycle Molecular Function: phosphoprotein phosphatase activity Molecular Function: protein binding Cellular Component: Nucleus Cellular Component: Cytoplasm Cellular Component: spindle Cellular Component: Cytoskeleton Biological Process: cell cycle Biological Process: mitotic chromosome movement towards spindle pole Biological Process: cell-cell signaling Biological Process: cell proliferation Cellular Component: spindle pole centrosome

Continue of Table A.2:

Predicted gene	**Gene Ontology**
	Biological Process: positive regulation of mitotic metaphase/anaphase transition
KIF2C	Biological Process: M phase of mitotic cell cycle Molecular Function: nucleotide binding Biological Process: mitotic prometaphase Biological Process: mitotic cell cycle Cellular Component: chromosome, centromeric region Cellular Component: Kinetochore Cellular Component: condensed chromosome kinetochore Molecular Function: microtubule motor activity Molecular Function: protein binding Molecular Function: ATP binding Cellular Component: Nucleus Cellular Component: Cytoplasm Cellular Component: Cytosol Cellular Component: kinesin complex Cellular Component: cytoplasmic microtubule

Continue of Table A.2:

Predicted gene	**Gene Ontology**
	Biological Process: microtubule-based movement Biological Process: microtubule depolymerization Biological Process: Mitosis Biological Process: blood coagulation Biological Process: cell proliferation Cellular Component: microtubule cytoskeleton Molecular Function: centromeric DNA binding Biological Process: establishment or maintenance of microtubule cytoskeleton polarity Molecular Function: microtubule plus-end binding Biological Process: cell division Biological Process: regulation of chromosome segregation
AURKB	Biological Process: M phase of mitotic cell cycle Molecular Function: nucleotide binding Biological Process: mitotic prometaphase Biological Process: mitotic cell cycle

Continue of Table A.2:

Predicted gene	**Gene Ontology**
	Cellular Component: condensed chromosome, centromeric region region Biological Process: Cytokinesis Molecular Function: protein serine/threonine kinase activity Molecular Function: protein binding Molecular Function ATP binding Cellular Component: Nucleus Cellular Component: Chromosome Cellular Component: Cytoplasm Cellular Component: Spindle Cellular Component: Cytosol Cellular Component: Cytoskeleton Biological Process: protein phosphorylation Biological Process: Aging Biological Process: cell proliferation Cellular Component: Midbody

Continue of Table A.2:

Predicted gene	**Gene Ontology**
	Biological Process: anaphase-promoting complex-dependent proteasomal ubiquitin-dependent protein catabolic process Cellular Component: chromosome passenger complex Biological Process: protein localization to kinetochore Molecular Function: metal ion binding
GINS2	Biological Process: S phase of mitotic cell cycle Biological Process: mitotic cell cycle Molecular Function: protein binding Cellular Component Nucleus Cellular Component: Nucleoplasm Biological Process: DNA replication Biological Process: DNA strand elongation involved in DNA replication
CENPA	Biological Process: M phase of mitotic cell cycle Biological Process: establishment of mitotic spindle orientation Biological Process: mitotic prometaphase Biological Process: mitotic cell cycle Cellular Component: chromosome, centromeric region

Continue of Table A.2:

Predicted gene	**Gene Ontology**
	Cellular Component: condensed nuclear chromosome kinetochore Cellular Component: condensed nuclear chromosome, centromeric region Cellular Component: Nucleosome Cellular Component: condensed chromosome inner kinetochore Molecular Function: DNA binding Molecular Function: chromatin binding Molecular Function: protein binding Cellular Component: Nucleus Cellular Component: Nucleoplasm Cellular Component: Cytosol Biological Process: nucleosome assembly Biological Process: CenH3-containing nucleosome assembly at centromere Biological Process: interspecies interaction between organisms Biological Process: kinetochore assembly Biological Process: protein localization to chromosome, centromeric region

Continue of Table A.2:

Predicted gene	**Gene Ontology**
ZWINT	Biological Process: mitotic sister chromatid segregation Biological Process: M phase of mitotic cell cycle Biological Process: mitotic prometaphase Biological Process: mitotic cell cycle: Cellular Component: Kinetochore Cellular Component: condensed chromosome kinetochore Molecular Function: protein binding Cellular Component: Nucleus Cellular Component: Cytoplasm Cellular Component: Cytosol Biological Process: cell cycle Biological Process: spindle organization Biological Process: mitotic cell cycle checkpoint Molecular Function: protein N-terminus binding Biological Process: phosphatidylinositol-mediated signaling Biological Process: cell division

Continue of Table A.2:

Predicted gene	**Gene Ontology**
	Biological Process: establishment of localization in cell
KIF20A	Biological Process: M phase of mitotic cell cycle Molecular Function: nucleotide binding Biological Process: mitotic cell cycle Biological Process: Cytokinesis Molecular Function: microtubule motor activity Molecular Function: transporter activity Molecular Function: protein binding Molecular Function: ATP binding Cellular Component: Nucleoplasm Cellular Component: Cytoplasm Cytoskeleton Cellular Component Golgi apparatus Cellular Component: Microtubule Biological Process: microtubule-based movement Biological Process: protein transport Biological Process: vesicle-mediated transport

Table A.3: Annotation of the top 10 Genes which are differentially expressed among Emodin and Control Classes in analysis of 4 arrays; MCF cells. (Continued From page 117 – 126).

Gene Info	Gene Ontology	Biocarta Pathways	Kegg Pathways
ProbeSet: 217967_s_at **Name:** family with sequence similarity 129, member A **Accession:** AF288391 **UniGene:** Hs.518662 **Symbol:** FAM129A **EntrezID:** 116496 **Chromosome:** 1 **Cytoband:** 1q25 SOURCE GENECARD **DrugBank:** Query Gene Symbol:FAM129A	**Gene Ontology:** Biological Process: negative regulation of protein phosphorylation Biological Process: positive regulation of protein phosphorylation Molecular Function: molecular_function Cellular Component: nucleus Cellular Component: cytoplasm Cellular Component: plasma membrane Biological Process: response to stress Biological Process: response to endoplasmic reticulum stress Biological Process: positive regulation of translation		
ProbeSet: 202672_s_at **Name:** activating transcription factor 3 **Accession:** NM_001674 **UniGene:** Hs.460 **Symbol:** ATF3 **EntrezID:** 467 **Chromosome:** 1 **Cytoband:** 1q32.3	**Gene Ontology:** Molecular Function: sequence-specific DNA binding transcription factor activity Molecular Function: transcription corepressor activity		

Continue of Table A.3:

Gene Info	Gene Ontology	Biocarta Pathways	Kegg Pathways
SOURCE GENECARD **DrugBank:** Query Gene Symbol:ATF3	Molecular Function: protein binding Cellular Component: nucleus Cellular Component: nucleolus Biological Process: gluconeogenesis Biological Process: transcription, DNA-dependent Biological Process: regulation of transcription, DNA-dependent Biological Process: positive regulation of cell proliferation Biological Process: negative regulation of transcription Molecular Function: transcription repressor activity Molecular Function: identical protein binding Molecular Function: sequence-specific DNA binding Molecular Function: protein dimerization activity		
ProbeSet: 209230_s_at	**Gene Ontology:** Molecular Function: molecular_function		

Continue of Table A.3:

Gene Info	Gene Ontology	Biocarta Pathways	Kegg Pathways
Name: nuclear protein, transcriptional regulator, 1 **Accession:** AF135266 **UniGene:** Hs.513463 **Symbol:** NUPR1 **EntrezID:** 26471 **Chromosome:** 16 **Cytoband:** 16p11.2 SOURCE GENECARD **DrugBank:** Query Gene Symbol:NUPR1	Cellular Component: nucleus Biological Process: induction of apoptosis Biological Process: cell growth		
ProbeSet: 217966_s_at **Name:** family with sequence similarity 129, member A **Accession:** NM_022083 **UniGene:** Hs.518662 **Symbol:** FAM129A **EntrezID:** 116496 **Chromosome:** 1 **Cytoband:** 1q25 SOURCE GENECARD **DrugBank:** Query Gene Symbol:FAM129A	**Gene Ontology:** Biological Process: negative regulation of protein phosphorylation Biological Process: positive regulation of protein phosphorylation Molecular Function: molecular_function Cellular Component: nucleus Cellular Component: cytoplasm Cellular Component: plasma membrane Biological Process: response to stress Biological Process: response to endoplasmic reticulum stress Biological Process: positive regulation of trtranslation		

Continue of Table A.3:

Gene Info	Gene Ontology	Biocarta Pathways	Kegg Pathways
ProbeSet: 209383_at **Name:** DNA-damage-inducible transcript 3 **Accession:** BC003637 **UniGene:** Hs.505777 **Symbol:** DDIT3 **EntrezID:** 1649 **Chromosome:** 12 **Cytoband:** 12q13.1-q13.2 SOURCE GENECARD **DrugBank:** Query Gene Symbol:DDIT3	**Gene Ontology:** Biological Process: response to amphetamine Molecular Function: DNA binding Molecular Function: sequence-specific DNA binding transcription factor activity Molecular Function: transcription corepressor activity Cellular Component: nucleus Cellular Component: cytoplasm Biological Process: transcription, DNA-dependent Biological Process: regulation of transcription, DNA-dependent Biological Process: response to DNA damage stimulus Biological Process: response to oxidative stress Biological Process: ER overload response Biological Process: cell cycle		

Continue of Table A.3:

Gene Info	Gene Ontology	Biocarta Pathways	Kegg Pathways
	Biological Process: cell cycle arrest Biological Process: aging Biological Process: response to nutrient Molecular Function: transcription factor binding Biological Process: negative regulation of transcription Biological Process: endoplasmic reticulum unfolded protein response Biological Process: negative regulation of CREB transcription factor activity Biological Process: response to endoplasmic reticulum stress Biological Process: response to drug Biological Process: response to hydrogen peroxide Biological Process: mRNA transcription from RNA polymerase II promoter Biological Process: positive regulation of apoptosis		

Continue of Table A.3:

Gene Info	Gene Ontology	Biocarta Pathways	Kegg Pathways
	Biological Process: negative regulation of transcription factor activity Molecular Function: sequence-specific DNA binding Biological Process: regulation of transcription in response to stress Biological Process: cell redox homeostasis Biological Process: positive regulation of transcription Molecular Function: protein dimerization activity Biological Process: embryonic organ development Biological Process: negative regulation of canonical Wnt receptor signaling pathway Biological Process: negative regulation of determination of dorsal identity		
ProbeSet: 221577_x_at **Name:** growth differentiation factor 15 **Accession:** AF003934	**Gene Ontology:** Molecular Function: cytokine activity Cellular Component: extracellular region		

Continue of Table A.3:

Gene Info	Gene Ontology	Biocarta Pathways	Kegg Pathways
UniGene: Hs.616962 **Symbol:** GDF15 **EntrezID:** 9518 **Chromosome:** 19 **Cytoband:** 19p13.11 SOURCE GENECARD **DrugBank:** Query Gene Symbol:GDF15	Cellular Component: extracellular space Biological Process: signal transduction Biological Process: transforming growth factor beta receptor signaling pathway Biological Process: cell-cell signaling Molecular Function: growth factor activity		
ProbeSet: 205749_at **Name:** cytochrome P450, family 1, subfamily A, polypeptide 1 **Accession:** NM_000499 **UniGene:** Hs.72912 **Symbol:** CYP1A1 **EntrezID:** 1543 **Chromosome:** 15 **Cytoband:** 15q24.1 SOURCE GENECARD **DrugBank:** Query Gene Symbol:CYP1A1	**Gene Ontology:** Cellular Component: endoplasmic reticulum Cellular Component: endoplasmic reticulum membrane Cellular Component: microsome Biological Process: xenobiotic metabolic process Molecular Function: electron carrier activity Biological Process: amine metabolic process Biological Process: toxin metabolic process Cellular Component: membrane Molecular Function: oxidoreductase activity		**KEGG Pathways:** 1: Tryptophan metabolism 2: Metabolism of xenobiotics by cytochrome P450

Continue of Table A.3:

Gene Info	Gene Ontology	Biocarta Pathways	Kegg Pathways
	Biological Process: drug metabolic process Molecular Function: oxygen binding Molecular Function: heme binding Biological Process: vitamin D metabolic process Biological Process: cellular lipid metabolic process Biological Process: heterocycle metabolic process Molecular Function: metal ion binding Biological Process: hydrogen peroxide biosynthetic process Biological Process: oxidation-reduction process Molecular Function: aromatase activity Molecular Function: vitamin D 24-hydroxylase activity		
ProbeSet: 203725_at **Name:** growth arrest and DNA-damage-inducible, alpha **Accession:** NM_001924 **UniGene:** Hs.80409	**Gene Ontology:** Biological Process: regulation of cyclin-dependent protein kinase activity	**BioCarta Pathways:** 1: Cell Cycle: G2/M Checkpoint	**KEGG Pathways:** 1: MAPK signaling pathway

Continue of Table A.3:

Gene Info	Gene Ontology	Biocarta Pathways	Kegg Pathways
Symbol: GADD45A **EntrezID:** 1647 **Chromosome:** 1 **Cytoband:** 1p31.2 SOURCE GENECARD **DrugBank:** Query Gene Symbol:GADD45A	Biological Process: G2/M transition of mitotic cell cycle Molecular Function: protein binding Cellular Component: nucleus Biological Process: DNA repair Biological Process: negative regulation of protein kinase activity Biological Process: apoptosis Biological Process: cell cycle arrest Biological Process: centrosome cycle Biological Process: signal transduction in response to DNA damage Biological Process: cellular response to ionizing radiation	2: ATM Signaling Pathway 3: p53 Signaling Pathway 4: Hypoxia and p53 in the Cardiovascular system	2: Cell cycle
ProbeSet: 217996_at **Name:** pleckstrin homology-like domain, family A, member 1 **Accession:** AA576961 **UniGene:** Hs.602085 **Symbol:** PHLDA1 **EntrezID:** 22822 **Chromosome:** 12	**Gene Ontology:** Molecular Function: protein binding Cellular Component: nucleus Cellular Component: cytoplasm Cellular Component: plasma membrane		

Continue of Table A.3:

Gene Info	Gene Ontology	Biocarta Pathways	Kegg Pathways
Cytoband: 12q15 SOURCE GENECARD **DrugBank:** Query Gene Symbol:PHLDA1	Biological Process: apoptosis Biological Process: induction of apoptosis Cellular Component: cytoplasmic vesicle membrane Cellular Component: cytoplasmic vesicle Biological Process: FasL biosynthetic process		
ProbeSet: 204472_at **Name:** GTP binding protein overexpressed in skeletal muscle **Accession:** NM_005261 **UniGene:** Hs.654463 **Symbol:** GEM **EntrezID:** 2669 **Chromosome:** 8 **Cytoband:** 8q13-q21 SOURCE GENECARD **DrugBank:** Query Gene Symbol:GEM	**Gene Ontology:** Molecular Function: nucleotide binding Molecular Function: magnesium ion binding Molecular Function: GTPase activity Molecular Function: protein binding Molecular Function: calmodulin binding Molecular Function: GTP binding Cellular Component: plasma membrane Biological Process: GTP catabolic process Biological Process: immune response Biological Process: signal transduction		

Continue of Table A.3:

Gene Info	Gene Ontology	Biocarta Pathways	Kegg Pathways
	Biological Process: cell surface receptor linked signaling pathway Biological Process: small GTPase mediated signal transduction Cellular Component: internal side of plasma membrane Molecular Function: GDP binding		

Table A.4: The Annotation of the predicted genes from network analysis result in figure 3.15. (Continued from page 127 – page 143).

Predicted gene	GO annotation
JUN	Cellular Component: nuclear chromosome Cellular Component: nuclear chromatin Biological Process: Angiogenesis Biological Process: release of cytochrome c from mitochondria Biological Process: toll-like receptor signaling pathway Biological Process: MyD88-dependent toll-like receptor signaling pathway Biological Process: MyD88-independent toll-like receptor signaling pathway Molecular Function: DNA binding

Continue of Table A.4:

Predicted gene	**GO annotation**
	Molecular Function: double-stranded DNA binding Molecular Function: sequence-specific DNA binding transcription factor activity Molecular Function: RNA polymerase II transcription factor activity Molecular Function: sequence-specific enhancer binding RNA polymerase II transcription factor activity Molecular Function: transcription coactivator activity Molecular Function: Rho GTPase activator activity Molecular Function: protein binding Cellular Component: Nucleus Cellular Component: Nucleoplasm Cellular Component: Nucleoplasm Cellular Component: transcription factor complex Cellular Component: Cytosol Biological Process: transcription, DNA-dependent Biological Process: transforming growth factor beta receptor signaling pathway Biological Process: SMAD protein import into nucleus

Continue of Table A.4:

Predicted gene	**GO annotation**
	Biological Process: Aging Biological Process: circadian rhythm Biological Process: Toll signaling pathway Molecular Function: transcription factor binding Biological Process: negative regulation of cell proliferation Biological Process: response to mechanical stimulus Biological Process: cellular process: Molecular Function: promoter binding Biological Process: response to organic cyclic compound Molecular Function: transcription activator activity Molecular Function: transcription repressor activity Biological Process: negative regulation of protein autophosphorylation Biological Process: response to lipopolysaccharide Biological Process: response to cytokine stimulus Biological Process: toll-like receptor 1 signaling pathway

Continue of Table A.4:

Predicted gene	**GO annotation**
	Biological Process: toll-like receptor 2 signaling pathway Biological Process: toll-like receptor 3 signaling pathway Biological Process: toll-like receptor 4 signaling pathway Biological Process: leading edge cell differentiation Biological Process: response to drug Biological Process: response to hydrogen peroxide Molecular Function: protein homodimerization activity Biological Process: negative regulation of DNA binding Biological Process: positive regulation of neuron apoptosis Biological Process: negative regulation by host of viral transcription Biological Process: positive regulation by host of viral transcription Biological Process: innate immune response Biological Process: positive regulation of monocyte differentiation Biological Process: positive regulation of DNA replication Biological Process: positive regulation of transcription Biological Process: positive regulation of transcription from RNA

Continue of Table A.4:

Predicted gene	**GO annotation**
	polymerase II promoter Biological Process: positive regulation of smooth muscle cell proliferation Biological Process: regulation of transcription factor activity Biological Process: cellular response to potassium ion starvation Biological Process: stress-activated MAPK cascade Biological Process: response to cAMP Biological Process: regulation of cell cycle Biological Process: membrane depolarization Biological Process: SMAD protein signal transduction Molecular Function: R-SMAD binding.
FOS	Biological Process: conditioned taste aversion Biological Process: toll-like receptor signaling pathway Biological Process: MyD88-dependent toll-like receptor signaling pathway: Biological Process: MyD88-independent toll-like receptor signaling pathway Biological Process: MyD88-independent toll-like receptor signaling pathway

Coninue of Table A.4:

Predicted gene	**GO annotation**
	Molecular Function: double-stranded DNA binding Molecular Function: sequence-specific DNA binding transcription factor activity Molecular Function: specific RNA polymerase II transcription factor activity Molecular Function: protein binding Cellular Component: Nucleus Cellular Component: Nucleoplasm Cellular Component: transcription factor complex Biological Process: DNA methylation Biological Process: transcription, DNA-dependent Biological Process: regulation of transcription from RNA polymerase II promoter Biological Process: inflammatory response Biological Process: transforming growth factor beta receptor signaling pathway: Biological Process: nervous system development Biological Process: female pregnancy Biological Process: Aging

Continue of Table A.4:

Predicted gene	**GO annotation**
	Biological Process: Toll signaling pathway Biological Process: response to cold Biological Process: response to light stimulus Biological Process: response to mechanical stimulus Biological Process: response to gravity Biological Process: response to toxin Molecular Function: promoter binding Biological Process: response to organic cyclic compound Cellular Component: Synaptosome Biological Process: Sleep Biological Process: cellular response to extracellular stimulus Biological Process: response to lipopolysaccharide Biological Process: response to progesterone stimulus Biological Process: cellular response to hormone stimulus Biological Process: response to cytokine stimulus Biological Process: toll-like receptor 1 signaling pathway

Continue of Table A.4:

Predicted gene	**GO annotation**
	Biological Process: toll-like receptor 2 signaling pathway Biological Process: toll-like receptor 3 signaling pathway Biological Process: toll-like receptor 4 signaling pathway Biological Process: cellular response to reactive oxygen species Biological Process: response to drug Biological Process: innate immune response Biological Process: positive regulation of transcription Biological Process: positive regulation of transcription from RNA polymerase II promoter Molecular Function: protein dimerization activity Biological Process: regulation of transcription factor activity Biological Process: stress-activated MAPK cascade Biological Process: response to corticosterone stimulus Biological Process: response to cAMP Biological Process: response to protein stimulus Biological Process: SMAD protein signal transduction Molecular Function: R-SMAD binding

Continue of Table A.4:

Predicted gene	**GO annotation**
EGR3	Molecular Function: DNA binding Molecular Function: sequence-specific DNA binding transcription factor activity Cellular Component: Intracellular Cellular Component: Nucleus Biological Process: transcription, DNA-dependent Biological Process: muscle organ development Biological Process: circadian rhythm Molecular Function: zinc ion binding Biological Process: regulation of transcription Molecular Function: metal ion binding
EGR1	Molecular Function: DNA binding Molecular Function: sequence-specific DNA binding transcription factor activity Molecular Function: specific RNA polymerase II transcription factor activity Molecular Function: protein binding Cellular Component: Intracellular

Continue of Table A.1:

Predicted gene	**GO annotation**
	Cellular Component: Nucleus Cellular Component: Nucleus Cellular Component: Cytoplasm Biological Process: transcription, DNA-dependent Molecular Function: zinc ion binding Biological Process: positive regulation of gene-specific transcription from RNA polymerase II promoter Biological Process: cytokine-mediated signaling pathway Biological Process: regulation of protein sumoylation Molecular Function: sequence-specific DNA binding Biological Process: positive regulation of transcription Molecular Function: metal ion binding Biological Process: type I interferon-mediated signaling pathway Biological Process: cellular response to heparin Biological Process: cellular response to mycophenolic acid Biological Process: positive regulation of glomerular metanephric mesangial cell proliferation

Continue of Table A.4:

Predicted gene	**GO annotation**
ASNS	Molecular Function: nucleotide binding Biological Process: liver development Molecular Function: asparagine synthase (glutamine-hydrolyzing) Molecular Function: asparagine synthase (glutamine-hydrolyzing) activity Molecular Function: ATP binding Cellular Component: soluble fraction Cellular Component: Cytosol Biological Process: asparagine biosynthetic process Biological Process: glutamine metabolic process Biological Process: metabolic process Biological Process: cellular amino acid biosynthetic process Biological Process: response to light stimulus Biological Process: response to mechanical stimulus Biological Process: response to toxin Molecular Function: ligase activity Biological Process: response to methotrexate

Continue of Table A.4:

Predicted gene	**GO annotation**
	Biological Process: response to nutrient levels Biological Process: response to follicle-stimulating hormone stimulus Biological Process: cellular response to hormone stimulus Biological Process: cellular nitrogen compound metabolic process Biological Process: cellular response to glucose starvation Molecular Function: protein homodimerization activity Biological Process: negative regulation of apoptosis Biological Process: response to amino acid stimulus Biological Process: positive regulation of mitotic cell cycle Molecular Function: cofactor binding
GADD45B	Biological Process: activation of MAPKKK activity Biological Process: activation of MAPKK activity Biological Process: negative regulation of protein kinase activity Biological Process: Apoptosis Biological Process: response to stress Biological Process: multicellular organismal development

Continue of Table A.4:

Predicted gene	**GO annotation**
	Biological Process: cell differentiation Biological Process: regulation of cell cycle
CYP1B1	Molecular Function: monooxygenase activity Molecular Function: protein binding Cellular Component: endoplasmic reticulum Cellular Component: endoplasmic reticulum membrane Cellular Component: Microsome Biological Process: cellular aromatic compound metabolic process Biological Process: xenobiotic metabolic process Biological Process: visual perception Biological Process: estrogen metabolic process Molecular Function: electron carrier activity Biological Process: toxin metabolic process Biological Process: response to organic substance Cellular Component: Membrane Molecular Function: oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular

Continue of Table A.4:

Predicted gene	**GO annotation**
	oxygen, reduced flavin or flavoprotein as one donor, and incorporation of one atom of oxygen Molecular Function: oxygen binding Molecular Function: oxygen binding Molecular Function: heme binding Molecular Function: metal ion binding Biological Process: oxidation-reduction process Molecular Function: aromatase activity
DUSP1	Biological Process: inactivation of MAPK activity Biological Process: endoderm formation Molecular Function: non-membrane spanning protein tyrosine phosphatase activity Molecular Function: protein binding Cellular Component: soluble fraction Cellular Component: Nucleus Cellular Component: Nucleoplasm Biological Process: protein dephosphorylation Biological Process: response to oxidative stress

Continue of Table A.4:

Predicted gene	**GO annotation**
	Biological Process: cell cycle Molecular Function: protein tyrosinethreonine phosphatase activity Biological Process: response to light stimulus Molecular Function: hydrolase activity Molecular Function: MAP kinase tyrosine/serine/threonine phosphatase activity Biological Process: response to estradiol stimulus Biological Process: response to retinoic acid Biological Process: cellular response to hormone stimulus Biological Process: response to testosterone stimulus Biological Process: response to hydrogen peroxide Biological Process: regulation of apoptosis Biological Process: positive regulation of apoptosis Biological Process: positive regulation of anti-apoptosis Biological Process: response to glucocorticoid stimulus Biological Process: response to cAMP Biological Process: response to calcium ion

Continue of Table A.4:

Predicted gene	**GO annotation**
CYR61	Biological Process: regulation of cell growth Biological Process: intussusceptive angiogenesis Molecular Function: insulin-like growth factor binding Cellular Component: extracellular region Biological Process: Chemotaxis Biological Process: cell adhesion Molecular Function: heparin binding Biological Process: cell proliferation Biological Process: anatomical structure morphogenesis Biological Process: positive regulation of cell-substrate adhesion Biological Process: extracellular matrix organization Molecular Function: extracellular matrix binding Biological Process: response to protein stimulus Biological Process: chorio-allantoic fusion Biological Process: labyrinthine layer blood vessel development
NR4A1	Molecular Function: DNA binding

Continue of Table A.4:

Predicted gene	**GO annotation**
	Molecular Function: sequence-specific DNA binding transcription factor activity Molecular Function: steroid hormone receptor activity Molecular Function ligand-dependent nuclear receptor activity Molecular Function: protein binding Cellular Component: Nucleus Cellular Component: Nucleoplasm Cellular Component: Nucleoplasm Biological Process: regulation of transcription, DNA-dependent Biological Process: induction of apoptosis Biological Process: signal transduction Molecular Function: zinc ion binding Biological Process: gene expression Molecular Function: transcription activator activity Biological Process: regulation of transcription from RNA polymerase II promoter by nuclear hormone receptor Molecular Function: protein homodimerization activity

Continue of Table A.4:

Predicted gene	**GO annotation**
	Biological Process: negative regulation of caspase activity Molecular Function: sequence-specific DNA binding Biological Process: positive regulation of transcription from RNA polymerase II promoter Molecular Function: metal ion binding Molecular Function: protein heterodimerization activity Biological Process: nerve growth factor receptor signaling pathway Biological Process: phosphatidylinositol-mediated signaling.

Table A.5: Annotation of the top 10 Genes which are differentially expressed among Emodin and Control Classes in analysis of 4 arrays; MDA cells. (Continued from page 144 – page 156).

Gene Info	**Gene Ontology**	**Biocarta Pathways**	**Kegg Pathways**
ProbeSet: 213418_at **Name:** heat shock 70kDa protein 6 (HSP70B') **Accession:** NM_002155 **UniGene:** Hs.654614 **Symbol:** HSPA6 **EntrezID:** 3310 **Chromosome:** 1 **Cytoband:** 1q23 SOURCE GENECARD **DrugBank:** Query Gene Symbol:HSPA6	**Gene Ontology:** Molecular Function: nucleotide binding Molecular Function: ATP binding Biological Process: response to unfolded protein		
ProbeSet: 204748_at	**Gene Ontology:**	**BioCarta Pathways:**	**KEGG Pathways:**

Continue of Table A.5:

Gene Info	Gene Ontology	Biocarta Pathways	Kegg Pathways
Name: prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase) **Accession:** NM_000963 **UniGene:** Hs.196384 **Symbol:** PTGS2 **EntrezID:** 5743 **Chromosome:** 1 **Cytoband:** 1q25.2-q25.3 SOURCE GENECARD **DrugBank:** Query Gene Symbol:PTGS2	Biological Process: prostaglandin biosynthetic process Molecular Function: peroxidase activity Molecular Function: peroxidase activity Molecular Function: prostaglandin-endoperoxide synthase activity Cellular Component: nucleus Cellular Component: cytoplasm Cellular Component: cytoplasm Cellular Component: endoplasmic reticulum Cellular Component: endoplasmic reticulum lumen Cellular Component: endoplasmic reticulum lumen Cellular Component: endoplasmic reticulum membrane Cellular Component: microsome Cellular Component: caveola Biological Process:	1: Mechanism of Acetaminophen Activity and Toxicity 2: Mechanism of Gene Regulation by Peroxisome Proliferators via PPARa(alpha) 3: Eicosanoid Metabolism	1: Arachidonic acid metabolism

Continue of Table A.5:

Gene Info	Gene Ontology	Biocarta Pathways	Kegg Pathways
	fatty acid biosynthetic process Biological Process: prostanoid metabolic process Biological Process: prostaglandin metabolic process Biological Process: xenobiotic metabolic process Biological Process: cellular component movement Biological Process: response to oxidative stress Biological Process: embryo implantation Biological Process: memory Biological Process: regulation of blood pressure Biological Process: negative regulation of cell proliferation Molecular Function: lipid binding Biological Process: response to fructose stimulus Biological Process: response to manganese ion Biological Process: response to organic nitrogen Biological Process: positive regulation vascular endothelial growth factor production		

Continue of Table A.5:

Gene Info	Gene Ontology	Biocarta Pathways	Kegg Pathways
	Biological Process: response to organic cyclic compound Cellular Component: membrane Molecular Function: oxidoreductase activity Molecular Function: oxidoreductase activity, acting on single donors with incorporation of molecular oxygen, incorporation of two atoms of oxygen Biological Process: cyclooxygenase pathway Molecular Function: enzyme binding Molecular Function: heme binding Biological Process: bone mineralization Biological Process: ovulation Biological Process: positive regulation of prostaglandin biosynthetic process Biological Process: positive regulation of fever generation Biological Process: positive regulation of synaptic plasticity Biological Process:		

Continue of Table A.5:

Gene Info	Gene Ontology	Biocarta Pathways	Kegg Pathways
	negative regulation of synaptic transmission, dopaminergic Biological Process: response to estradiol stimulus Biological Process: response to lipopolysaccharide Biological Process: response to vitamin D Biological Process: response to cytokine stimulus Biological Process: hormone biosynthetic process Biological Process: response to drug Biological Process: anagen Cellular Component: neuron projection Biological Process: positive regulation of apoptosis Cellular Component: protein complex Biological Process: positive regulation of nitric oxide biosynthetic process Biological Process: positive regulation of vasoconstriction Biological Process: positive regulation of smooth muscle contraction Biological Process:		

Continue of Table A.5:

Gene Info	Gene Ontology	Biocarta Pathways	Kegg Pathways
	decidualization Molecular Function: metal ion binding Biological Process: positive regulation of smooth muscle cell proliferation Biological Process: regulation of inflammatory response Biological Process: response to glucocorticoid stimulus Biological Process: regulation of cell cycle Biological Process: negative regulation of calcium ion transport Biological Process: positive regulation of synaptic transmission, glutamatergic Biological Process: oxidation-reduction process Biological Process: positive regulation of transforming growth factor-beta production Biological Process: positive regulation of cell migration involved in sprouting angiogenesis Biological Process: positive regulation of fibroblast growth factor production Biological Process:		

Continue of Table A.5:

Gene Info	Gene Ontology	Biocarta Pathways	Kegg Pathways
	positive regulation of brown fat cell differentiation Biological Process: positive regulation of platelet-derived growth factor production		
ProbeSet: 206569_at **Name:** interleukin 24 **Accession:** NM_006850 **UniGene:** Hs.58831 **Symbol:** IL24 **EntrezID:** 11009 **Chromosome:** 1 **Cytoband:** 1q32 SOURCE GENECARD **DrugBank:** Query Gene Symbol:IL24	**Gene Ontology:** Molecular Function: cytokine activity Cellular Component: extracellular region Cellular Component: extracellular space Biological Process: apoptosis		**KEGG Pathways:** 1: Cytokine-cytokine receptor interaction 2: Jak-STAT signaling pathway
ProbeSet: 117_at **Name:** heat shock 70kDa protein 6 (HSP70B') **Accession:** X51757 **UniGene:** Hs.654614 **Symbol:** HSPA6 **EntrezID:** 3310 **Chromosome:** 1 **Cytoband:** 1q23 SOURCE GENECARD **DrugBank:** Query Gene Symbol:HSPA6	**Gene Ontology:** Molecular Function: nucleotide binding Molecular Function: ATP binding Biological Process: response to unfolded protein		
ProbeSet: 226980_at **Name:** DEP domain containing 1B **Accession:** AK001166 **UniGene:** Hs.482233 **Symbol:** DEPDC1B	**Gene Ontology:** Molecular Function: GTPase activator activity Cellular Component: intracellular		

Continue of Table A.5:

Gene Info	Gene Ontology	Biocarta Pathways	Kegg Pathways
EntrezID: 55789 **Chromosome:** 5 **Cytoband:** 5q12.1 SOURCE GENECARD **DrugBank:** Query Gene Symbol:DEPDC1B	Cellular Component: cytosol Biological Process: small GTPase mediated signal transduction Biological Process: regulation of small GTPase mediated signal transduction		
ProbeSet: 236641_at **Name:** kinesin family member 14 **Accession:** AW183154 **UniGene:** Hs.3104 **Symbol:** KIF14 **EntrezID:** 9928 **Chromosome:** 1 **Cytoband:** 1q32.1 SOURCE GENECARD **DrugBank:** Query Gene Symbol:KIF14	**Gene Ontology:** Molecular Function: nucleotide binding Molecular Function: microtubule motor activity Molecular Function: protein binding Molecular Function: ATP binding Cellular Component: nucleus Cellular Component: cytoplasm Cellular Component: spindle Cellular Component: cytoskeleton Cellular Component: microtubule Biological Process: microtubule-based movement		
ProbeSet: 204709_s_at **Name:** kinesin family member 23	**Gene Ontology:** Biological Process: mitotic spindle elongation		

Continue of Table A.5:

Gene Info	Gene Ontology	Biocarta Pathways	Kegg Pathways
Accession: NM_004856 **UniGene:** Hs.270845 **Symbol:** KIF23 **EntrezID:** 9493 **Chromosome:** 15 **Cytoband:** 15q23 SOURCE GENECARD **DrugBank:** Query Gene Symbol:KIF23	Biological Process: M phase of mitotic cell cycle Molecular Function: nucleotide binding Biological Process: mitotic cell cycle Biological Process: cytokinesis Molecular Function: microtubule motor activity Molecular Function: protein binding Molecular Function: ATP binding Cellular Component: nucleus Cellular Component: nucleoplasm Cellular Component: cytoplasm Cellular Component: spindle Cellular Component: cytosol Cellular Component: cytoskeleton Cellular Component: kinesin complex Cellular Component: microtubule		

Continue of Table A.5:

Gene Info	Gene Ontology	Biocarta Pathways	Kegg Pathways
	Biological Process: microtubule-based movement Biological Process: mitosis Biological Process: blood coagulation Cellular Component: midbody		
ProbeSet: 200799_at **Name:** heat shock 70kDa protein 1A **Accession:** NM_005345 **UniGene:** Hs.274402 **Symbol:** HSPA1A **EntrezID:** 3303 **Chromosome:** 6 **Cytoband:** 6p21.3 SOURCE GENECARD **DrugBank:** Query Gene Symbol:HSPA1A		**BioCarta Pathways:** 1: Hypoxia and p53 in the Cardiovascular system 2: Mechanism of Gene Regulation by Peroxisome Proliferators via PPARa(alpha) 3: Chaperones modulate interferon Signaling Pathway	**KEGG Pathways:** 1: Antigen processing and presentation 2: MAPK signaling pathway
ProbeSet: 201292_at **Name:** topoisomerase (DNA) II alpha 170kDa **Accession:** AL561834 **UniGene:** Hs.156346 **Symbol:** TOP2A **EntrezID:** 7153 **Chromosome:** 17 **Cytoband:** 17q21-q22	**Gene Ontology:** Molecular Function: nucleotide binding Cellular Component: nuclear chromosome Biological Process: resolution of meiotic recombination intermediates	**BioCarta Pathways:** 1: Apoptotic DNA fragmentation and tissue homeostasis	

Continue of Table A.5:

Gene Info	Gene Ontology	Biocarta Pathways	Kegg Pathways
SOURCE GENECARD **DrugBank:** Query Gene Symbol:TOP2A	Cellular Component: synaptonemal complex Biological Process: sister chromatid segregation Molecular Function: DNA binding Molecular Function: chromatin binding Molecular Function: sequence-specific DNA binding transcription factor activity Molecular Function: DNA topoisomerase (ATP-hydrolyzing) activity Molecular Function: protein kinase C binding Molecular Function: protein binding Molecular Function: ATP binding Cellular Component: nucleus Cellular Component: nucleoplasm Cellular Component: nucleolus Cellular Component: cytoplasm Cellular Component: centriole Biological Process:		

Continue of Table A.5:

Gene Info	Gene Ontology	Biocarta Pathways	Kegg Pathways
	DNA replication Biological Process: DNA-dependent DNA replication Biological Process: DNA topological change Biological Process: DNA ligation Biological Process: DNA repair Biological Process: mitotic recombination Biological Process: regulation of transcription, DNA-dependent Biological Process: response to DNA damage stimulus Biological Process: chromosome segregation Molecular Function: protein C-terminus binding Molecular Function: DNA-dependent ATPase activity Molecular Function: drug binding Cellular Component: DNA topoisomerase complex (ATP-hydrolyzing) Cellular Component: viral integration complex		

Continue of Table A.5:

Gene Info	Gene Ontology	Biocarta Pathways	Kegg Pathways
	Molecular Function: enzyme binding Biological Process: apoptotic chromosome condensation Molecular Function: protein homodimerization activity Molecular Function: histone deacetylase binding Biological Process: positive regulation of apoptosis Molecular Function: ubiquitin binding Cellular Component: protein complex Biological Process: positive regulation of viral genome replication Biological Process: positive regulation of retroviral genome replication Molecular Function: protein heterodimerization activity Biological Process: phosphatidylinositol-mediated signaling Biological Process: mitotic cell cycle G2/M transition decatenation checkpoint		
ProbeSet: 223274_at **Name:** transcription	**Gene Ontology:** Molecular Function:		

Continue of Table A.5:

Gene Info	Gene Ontology	Biocarta Pathways	Kegg Pathways
factor 19 **Accession:** BC002493 **UniGene:** Hs.584807 **Symbol:** TCF19 **EntrezID:** 6941 **Chromosome:** 6 **Cytoband:** 6p21.3 SOURCE GENECARD **DrugBank:** Query Gene Symbol:TCF19	sequence-specific DNA binding transcription factor activity Cellular Component: nucleus Biological Process: transcription, DNA-dependent Biological Process: regulation of transcription from RNA polymerase II promoter Molecular Function: zinc ion binding Biological Process: cell proliferation Biological Process: regulation of transcription Molecular Function: metal ion binding		

Table A.6: The Annotation of the predicted genes from network analysis result in figure 3.18. (Continued from page 157 – 177).

Predicted gene	GO annotation
CSNK2A1	Molecular Function: nucleotide binding Molecular Function: protein serine/threonine kinase activity Molecular Function: protein binding Molecular Function: ATP binding Cellular Component: Nucleus

Continue of Table A.6:

Predicted gene	**GO annotation**
	Cellular Component: Cytosol Cellular Component: plasma membrane Biological Process: protein phosphorylation Biological Process: signal transduction Biological Process: axon guidance Biological Process: Wnt receptor signaling pathway Cellular Component: Sin3 complex Cellular Component: NuRD complex Molecular Function: protein N-terminus binding
CENPF	Biological Process: G2 phase of mitotic cell cycle Biological Process: M phase of mitotic cell cycle Biological Process: M phase of mitotic cell cycle Biological Process: M phase of mitotic cell cycle Biological Process: mitotic prometaphase Biological Process: mitotic cell cycle Biological Process: mitotic cell cycle

Continue of Table A.6:

Predicted gene	**GO annotation**
	Cellular Component: chromosome, centromeric region Cellular Component: Kinetochore Cellular Component: Chromatin Cellular Component: spindle pole Cellular Component: condensed chromosome outer kinetochore Cellular Component: condensed chromosome outer kinetochore Molecular Function: chromatin binding Molecular Function: protein binding Cellular Component: Nucleus Cellular Component: nuclear envelope Cellular Component: Cytoplasm Cellular Component: Spindle Cellular Component: Cytosol Cellular Component: Cytoskeleton Biological Process: DNA replication Biological Process: chromosome segregation

Continue of Table A.6:

Predicted gene	**GO annotation**
	Biological Process: chromosome segregation Biological Process: mitotic cell cycle spindle assembly checkpoint Biological Process: multicellular organismal development Biological Process: muscle organ development Molecular Function: protein C-terminus binding Molecular Function: transcription factor binding Biological Process: cell proliferation Biological Process: regulation of G2/M transition of mitotic cell cycle Biological Process: protein transport Biological Process: protein transport Biological Process: regulation of striated muscle tissue development Cellular Component: nuclear matrix Biological Process: negative regulation of transcription Biological Process: cell differentiation Cellular Component: Midbody Biological Process: response to drug

Continue of Table A.6:

Predicted gene	**GO annotation**
	Molecular Function: protein homodimerization activity Molecular Function: dynein binding Cellular Component: perinuclear region of cytoplasm Biological Process: cell division Biological Process: metaphase plate congression Biological Process: kinetochore assembly
CDCA8	Biological Process: M phase of mitotic cell cycle Biological Process: mitotic metaphase Biological Process: mitotic prometaphase Biological Process: mitotic cell cycle Cellular Component: chromosome, centromeric region Molecular Function: protein binding Cellular Component: Nucleus Cellular Component: Nucleolu Cellular Component: Cytoplasm Cellular Component:

Continue of Table A.6:

Predicted gene	**GO annotation**
	Spindle Cellular Component: Cytosol Cellular Component: Cytoskeleton Cellular Component: chromosome passenger complex Cellular Component: protein complex Biological Process: chromosome organization Biological Process: cell division
CENPE	Biological Process: M phase of mitotic cell cycle Biological Process: mitotic metaphase Molecular Function: nucleotide binding Biological Process: mitotic prometaphase Biological Process: mitotic cell cycle Cellular Component: chromosome, centromeric region Cellular Component: Kinetochore Cellular Component: condensed chromosome, centromeric region Cellular Component: condensed chromosome outer kinetochore Molecular Function: microtubule motor activity

Continue of Table A.6:

Predicted gene	**GO annotation**
	Molecular Function: protein binding Molecular Function: ATP binding Cellular Component: Nucleus Cellular Component: Cytoplasm Cellular Component: Spindle Cellular Component: Cytosol Cellular Component: Cytoskeleton Cellular Component: Microtubule Biological Process: microtubule-based movement Biological Process: mitotic chromosome movement towards spindle pole Biological Process: mitotic metaphase plate congression Biological Process: regulation of mitosis Biological Process: mitotic cell cycle spindle assembly checkpoint Biological Process: multicellular organismal development Biological Process: blood coagulation Molecular Function: protein kinase binding

Continue of Table A.6:

Predicted gene	**GO annotation**
	Molecular Function: kinetochore binding Biological Process: establishment of protein localization Biological Process: positive regulation of mitotic metaphase/anaphase transition Biological Process: positive regulation of protein kinase activity| Biological Process: regulation of developmental process Biological Process: cell division Biological Process: kinetochore assembly Biological Process: positive regulation of attachment of spindle microtubules to kinetochore
BIRC5	Biological Process: G2/M transition of mitotic cell cycle Biological Process: M phase of mitotic cell cycle Cellular Component: nuclear chromosome Biological Process: mitotic prometaphase Biological Process: mitotic cell cycle Cellular Component: chromosome, centromeric region Biological Process: Cytokinesis Molecular Function:

Continue of Table A.6:

Predicted gene	**GO annotation**
	cysteine-type endopeptidase inhibitor activity Molecular Function: protein binding Cellular Component: Intracellular Cellular Component: Nucleus Cellular Component: Cytoplasm Cellular Component: Centriole Cellular Component: Spindle Cellular Component: Cytosol Cellular Component: Cytosol Cellular Component: Cytoskeleton Cellular Component: spindle microtubule Cellular Component: cytoplasmic microtubule Biological Process: Apoptosis Biological Process: anti-apoptosis Biological Process: chromosome segregation Biological Process: Mitosis

Continue of Table A.6:

Predicted gene	**GO annotation**
	Molecular Function: microtubule binding Molecular Function: zinc ion binding Molecular Function: zinc ion binding Molecular Function: Ran GTPase binding Molecular Function: tubulin binding Molecular Function: enzyme binding Molecular Function: peptidase inhibitor activity Cellular Component: Midbody Cellular Component: interphase microtubule organizing center Biological Process: protein complex localization Biological Process: positive regulation of exit from mitosis Biological Process: spindle checkpoint Cellular Component: chromosome passenger complex Molecular Function: identical protein binding Molecular Function: protein homodimerization activity Molecular Function: protein homodimerization activity

Continue of Table A.6:

Predicted gene	**GO annotation**
	Molecular Function:: caspase inhibitor activity Biological Process: negative regulation of caspase activity Biological Process: negative regulation of caspase activity Biological Process: positive regulation of mitotic cell cycle Molecular Function: metal ion binding Molecular Function: protein heterodimerization activity Molecular Function: cofactor binding Molecular Function: cobalt ion binding Molecular Function: chaperone binding Biological Process: cell division Biological Process: cell division Biological Process: establishment of chromosome localization
RACGAP1	Biological Process: Cytokinesis Biological Process: cytokinesis, actomyosin contractile ring assembly Cellular Component: acrosomal vesicle Molecular Function: GTPase activator activity

Continue of Table A.6:

Predicted gene	**GO annotation**
	Molecular Function: GTPase activator activity Molecular Function: protein binding Cellular Component: İntracellular Cellular Component: Nucleus Cellular Component: Cytoplasm Cellular Component: Spindle Cellular Component: Cytosol Cellular Component: Cytoskeleton Cellular Component: Microtubule Biological Process: ion transport Biological Process: microtubule-based movement Biological Process: cell cycle Biological Process: cytokinesis, initiation of separation Biological Process: small GTPase mediated signal transduction Biological Process: Spermatogenesis Biological Process: neuroblast proliferation

Continue of Table A.6:

Predicted gene	**GO annotation**
	Biological Process: blood coagulation Molecular Function: protein C-terminus binding Biological Process: sulfate transport Biological Process: embryo development Biological Process: cell differentiation Cellular Component: Midbody Cellular Component: cytoplasmic vesicle Molecular Function: alpha-tubulin binding Molecular Function: gamma-tubulin binding Molecular Function: metal ion binding Molecular Function: beta-tubulin binding Biological Process: regulation of small GTPase mediated signal transduction
KIF11	Molecular Function: nucleotide binding Cellular Component: spindle pole Molecular Function: microtubule motor activity Molecular Function: ATP binding

Continue of Table A.6:

Predicted gene	**GO annotation**
	Cellular Component: Cytoplasm Cellular Component: Spindle Cellular Component: Cytosol Cellular Component: Cytosol Cellular Component: Cytoskeleton Cellular Component: kinesin complex Cellular Component: spindle microtubule Biological Process: microtubule-based movement Biological Process: cell cycle Biological Process: spindle organization Biological Process: mitotic spindle organization Biological Process: Mitosis Biological Process: mitotic centrosome separation Biological Process: blood coagulation Cellular Component: chromatin remodeling complex Molecular Function: protein kinase binding

Continue of Table A.6:

Predicted gene	**GO annotation**
	Biological Process: cell division Biological Process: spindle assembly involved in mitosis
PLK4	Biological Process: G2/M transition of mitotic cell cycle Molecular Function: nucleotide binding Biological Process: mitotic cell cycle Molecular Function: protein serine/threonine kinase activity Molecular Function: protein tyrosine kinase activity Molecular Function: protein binding Molecular Function: ATP binding Cellular Component: Nucleus Cellular Component| Nucleolus Cellular Component: Cytoplasm Cellular Component: Centriole Cellular Component: Cytosol Cellular Component: Cytoskeleton Biological Process: protein phosphorylation Molecular Function:

Continue of Table A.6:

Predicted gene	**GO annotation**
	transferase activity Cellular Component: cleavage furrow Biological Process: positive regulation of centriole replication Biological Process: trophoblast giant cell differentiation
CCNB1	Biological Process: cell cycle checkpoint Biological Process: G1/S transition of mitotic cell cycle Biological Process: G2/M transition of mitotic cell cycle Biological Process: G2/M transition of mitotic cell cycle Biological Process: mitotic prometaphase Biological Process: mitotic cell cycle Cellular Component: spindle pole Cellular Component: condensed nuclear chromosome outer kinetochore Biological Process: oocyte maturation Biological Process: in utero embryonic development Biological Process: negative regulation of protein phosphorylation Molecular Function: protein binding

Continue of Table A.6:

Predicted gene	**GO annotation**
	Cellular Component: membrane fraction Cellular Component: soluble fraction Cellular Component: Nucleus Cellular Component: Nucleoplasm Cellular Component: Nucleoplasm Cellular Component: Cytoplasm Cellular Component: microtubule organizing center Cellular Component: Cytosol Cellular Component: Cytosol Cellular Component: Cytoskeleton Biological Process: protein complex assembly Biological Process: mitotic metaphase plate congression Biological Process: spermatogenesis Biological Process: response to mechanical stimulus Biological Process: response to toxin Biological Process: negative regulation of gene expression

Continue of Table A.6:

Predicted gene	**GO annotation**
	Molecular Function: kinase activity Molecular Function: protein kinase binding Biological Process: anaphase-promoting complex-dependent proteasomal ubiquitin-dependent protein catabolic process Biological Process: anaphase-promoting complex-dependent proteasomal ubiquitin-dependent protein catabolic process Biological Process: positive regulation of mRNA 3'-end processing Molecular Function: protein complex binding Biological Process: positive regulation of histone phosphorylation Molecular Function: histone kinase activity Biological Process: tissue regeneration Biological Process: response to drug Biological Process: mitotic spindle stabilization Biological Process: positive regulation of mitotic Biological Process: response to DDT Biological Process: digestive tract development Biological Process:

Continue of Table A.6:

Predicted gene	**GO annotation**
	cell division Biological Process: positive regulation of ubiquitin-protein ligase activity involved in mitotic cell cycle Biological Process: regulation of ubiquitin-protein ligase activity involved in mitotic cell cycle Biological Process: regulation of cell cycle Biological Process: positive regulation of attachment of spindle microtubules to kinetochore Biological Process: ventricular cardiac muscle cell development Biological Process: positive regulation of cardiac muscle cell proliferation Biological Process: regulation of chromosome condensation Biological Process: mitotic cell cycle spindle checkpoint Biological Process: cellular response to iron(III) ion Biological Process: cellular response to fatty acid Biological Process: cellular response to organic cyclic compound Biological Process: cellular response to protein stimulus Biological Process: cellular response to hypoxia

Continue of Table A.6:

Predicted gene	**GO annotation**
NEK2	Biological Process: mitotic sister chromatid segregation Biological Process: G2/M transition of mitotic cell cycle Molecular Function: nucleotide binding Biological Process: mitotic cell cycle Cellular Component: Kinetochore Cellular Component: condensed chromosome kinetochore Cellular Component: condensed nuclear chromosome Cellular Component: spindle pole Molecular Function: protein kinase activity Molecular Function: protein serine/threonine kinase activity Molecular Function: protein binding Molecular Function: ATP binding Cellular Component: Nucleus Cellular Component: Nucleolus Cellular Component: Cytoplasm Cellular Component: Centrosome

Continue of Table A.6:

Predicted gene	**GO annotation**
	Cellular Component: Cytosol Cellular Component: Cytosol Biological Process: chromosome segregation Biological Process: Mitosis Biological Process: regulation of mitosis Biological Process: Meiosis Molecular Function: transferase activity Molecular Function: protein phosphatase binding Cellular Component: Midbody Cellular Component: protein complex Biological Process: protein autophosphorylation Molecular Function: metal ion binding Biological Process: centrosome separation Biological Process: cell division

Table A.7: Annotation for the GST isozymes genes. (Continued from page 178 – 183).

Gene Info	**Go annotation**
ProbeSet: 1554518_at Accession: BC032942 Symbol: GSTCD Cytoband: 4q24	Cellular Component: Cytoplasm Biological Process: rRNAprocessing Molecular Function: rRNA methyltransferase activity
ProbeSet: 1557915_s_at Accession: U56250 Symbol: GSTO1 Cytoband: 10q25.1	Molecular Function: glutathione transferase activity Cellular Component Cytoplasm Cellular Component: Cytosol Biological Process: xenobiotic metabolic process Biological Process : metabolic process Molecular Function: monodehydroascorbate reductase (NADH) activity Molecular Function: transferase activity Biological Process: L-ascorbic acid biosynthetic process.
ProbeSet: 200824_at Accession: NM_000852 Symbol: GSTP1 Cytoband: 11q13	Molecular Function: glutathione transferase activity Molecular Function: protein binding Cellular Component: Nucleus Cellular Component: Cytoplasm Cellular Component: Cytosol

Continue of Table A.7:

Gene Info	Go annotation
	Cellular Component: plasma membrane Biological Process: glutathione metabolic process Biological Process: xenobiotic metabolic process Biological Process: xenobiotic metabolic process Biological Process: anti-apoptosis Biological Process: central nervous system development Molecular Function: drug binding Biological Process: metabolic process Biological Process: response to toxin Biological Process: oligodendrocyte development Molecular Function: transferase activity Biological Process: organ regeneration Biological Process: response to nutrient levels Biological Process: response to estradiol stimulus Biological Process: cellular response to insulin stimulus Biological Process: response to L-ascorbic acid

Continue of Table A.7:

Gene Info	**Go annotation**
	Biological Process: response to amino acid stimulus Molecular Function: glutathione binding Biological Process: response to ethanol Biological Process: cellular response to epidermal growth factor stimulus Biological Process: cellular response to glucocorticoid stimulus Biological Process: cellular response to cell-matrix adhesion
ProbeSet: 202554_s_at Accession: AL527430 Symbol: GSTM3 Cytoband: 1p13.3	Molecular Function: glutathione transferase activity Molecular Function: protein binding Cellular Component: soluble fraction Cellular Component: Cytoplasm Biological Process: glutathione metabolic process Biological Process: establishment of blood-nerve barrier Biological Process: metabolic process Molecular Function: transferase activity Molecular Function: identical protein binding Biological Process: response to estrogen stimulus

Continue of Table A.7:

Gene Info	**Go annotation**
ProbeSet: 202967_at Accession: NM_001512 Symbol: GSTA4 Cytoband: 6p12.1	Molecular Function: glutathione transferase activity Cellular Component: Cytoplasm Cellular Component: Cytosol Biological Process: glutathione metabolic process Biological Process: xenobiotic metabolic process Biological Process: xenobiotic metabolic process Biological Process: metabolic process Molecular Function: transferase activity Molecular Function: protein homodimerization activity.
ProbeSet: 204149_s_at Accession: NM_000850 Symbol: GSTM4 Cytoband: 1p13.3	Molecular Function: glutathione transferase activity Cellular Component: Cytoplasm Cellular Component: endoplasmic reticulum membrane Biological Process: xenobiotic metabolic process Biological Process: metabolic process Molecular Function: transferase activity Biological Process: nitrobenzene metabolic process

Continue of Table A.7:

Gene Info	**Go annotation**
	Biological Process: xenobiotic catabolic process
ProbeSet: 209531_at Accession: BC001453 Symbol: GSTZ1 Cytoband: 14q24.3	Molecular Function: glutathione transferase activity Molecular Function: glutathione peroxidase activity Molecular Function: protein binding Cellular Component: Cytoplasm Cellular Component: Mitochondrion Cellular Component: Cytosol Biological Process: L-phenylalanine catabolic process Biological Process: tyrosine catabolic process Biological Process: glutathione metabolic process Biological Process: aromatic amino acid family metabolic process Molecular Function: maleylacetoacetate isomerase activity Molecular Function: maleylacetoacetate isomerase activity Molecular Function: transferase activity Molecular Function: isomerase activity Biological Process: cellular nitrogen compound metabolic process

Continue of Table A.7:

Gene Info	Go annotation
	Molecular Function: protein homodimerization activity.
ProbeSet: 215766_at Accession: AL096729 Symbol: GSTA1 Cytoband: 6p12.1	Molecular Function: glutathione transferase activity Molecular Function: glutathione transferase activity Cellular Component: Cytoplasm Cellular Component: Cytosol Biological Process: glutathione metabolic process Biological Process: xenobiotic metabolic process Biological Process: metabolic process Molecular Function: transferase activity

Table A.8: The Annotation of the predicted genes from network analysis result in figure 3.21. (Continued from page 183 – page 189).

Predicted Gene	Go annotation
PNPLA4	Molecular Function: triglyceride lipase activity Biological Process: metabolic process Biological Process: lipid catabolic process Molecular Function: hydrolase activity
ECHS1	Molecular Function: enoyl-CoA hydratase activity Molecular Function: protein binding

Continue of Table A.8:

Predicted Gene	**Go annotation**
	Cellular Component: soluble fraction Cellular Component: Mitochondrion Cellular Component: Mitochondrion Cellular Component: mitochondrial matrix Biological Process: fatty acid metabolic process Biological Process: fatty acid beta-oxidation Molecular Function: lyase activity Biological Process: cellular lipid metabolic process
SMOX	Cellular Component: Nucleus Cellular Component: Cytoplasm Cellular Component: Cytosol Biological Process: polyamine metabolic process Biological Process: polyamine biosynthetic process Biological Process: xenobiotic metabolic process Molecular Function: oxidoreductase activity Biological Process: cellular nitrogen compound metabolic process

Continue of Table A.8:

Predicted Gene	**Go annotation**
	Biological Process: spermine catabolic process Molecular Function: polyamine oxidase activity Biological Process: oxidation-reduction proces
AKR7A2	Molecular Function: aldehyde reductase activity Cellular Component: Cytoplasm Cellular Component: Golgi apparatus Biological Process: carbohydrate metabolic process Biological Process: cellular aldehyde metabolic process Molecular Function: electron carrier activity Molecular Function: oxidoreductase activity Molecular Function: oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor Biological Process: oxidation-reduction process
FIS1	Biological Process: mitochondrial fission Molecular Function: protein binding Cellular Component: Mitochondrion Cellular Component: mitochondrial outer membrane

Continue of Table A.8:

Predicted Gene	**Go annotation**
	Cellular Component: Peroxisome Cellular Component: peroxisomal membrane Cellular Component: integral to peroxisomal membrane Biological Process: Apoptosis Cellular Component: Membrane Cellular Component: integral to membrane Biological Process: peroxisome fission Cellular Component: integral to mitochondrial outer membrane
FIBP	Molecular Function: protein binding Cellular Component: membrane fraction Cellular Component: Nucleus Cellular Component: Mitochondrion Cellular Component: Microsome Biological Process: fibroblast growth factor receptor signaling pathway Cellular Component: endomembrane system Cellular Component: Membrane

Continue of Table A.8:

Predicted Gene	**Go annotation**
	Molecular Function: fibroblast growth factor binding
APRT	Molecular function: adenine binding Molecular Function: adenine phosphoribosyltransferase activity Molecular Function: protein binding Cellular Component: Nucleus Cellular Component: Cytoplasm Cellular Component: Cytoplasm Cellular Component: Cytosol Biological Process: purine base metabolic process Biological Process: purine ribonucleoside salvage Biological Process: adenine salvage Biological Process: grooming behavior Biological Process: nucleoside metabolic process Molecular Function: AMP binding Molecular Function: transferase activity, transferring glycosyl groups Biological Process: purine-containing compound salvage

Continue of Table A.8:

Predicted Gene	**Go annotation**
	Biological Process: adenine metabolic process Biological Process: nucleobase, nucleoside and nucleotide metabolic process
CAPG	Molecular Function: actin binding Cellular Component: Nucleus Cellular Component: Nucleolus Cellular Component: Cytoplasm Biological Process: protein complex assembly Cellular Component: F-actin capping protein complex Biological Process: cell projection assembly Cellular Component: nuclear membrane Cellular Component: Melanosome Biological Process: barbed-end actin filament capping
EPHX2	Molecular Function: epoxide hydrolase activity Cellular Component: soluble fraction Cellular Component: Nucleolus Cellular Component:

Continue of Table A.8:

Predicted Gene	**Go annotation**
	Cytoplasm Cellular Component: Peroxisome Cellular Component: Golgi apparatus Cellular Component: Cytosol Cellular Component: focal adhesion Biological Process: xenobiotic metabolic process Biological Process: cellular calcium ion homeostasis Biological Process: inflammatory response Biological Process: regulation of blood pressure Biological Process: response to toxin Molecular Function: hydrolase activity Biological Process: drug metabolic process Biological Process: aromatic compound catabolic process Molecular Function: protein homodimerization activity Biological Process: positive regulation of vasodilation Molecular Function: metal ion binding

APPENDIX B

THE DNA SEQUENCING RESULT FOR PCR PRODUCT DONE TO CONFIRM THE PRIMERS DESIGN

Sequencing Result for MDA GSTA4:

CTCGAGTGGACTCAGAAGCCTGATAGCTATCATGGCAGCAAGGCCCAAGCTCCACTATCC
CAACGGAAGAGGCCGGATGGAGTCCGTGAGATGGGTTTTAGCTGCCGCCGGGTCGAGTT
TGATGAAGACAGCCGCCGCGTCTGCATTTCAGAAAGCCTGAAAAGCTATCATGGCAGCAA
GGCCCAAGCTCCACTATCCCAACGGAAGAGGCCGGATGGAGTCCGTGAGATGGGTTTTAG
CTGCCGCCGGAGTCGAGGTGGATGAAGA.

Sequencing result for MDA GSTA1:

AAGACATCACTGATACTGCAAGAGCTCACGAAACTATGAAGAAGTTTTCTAGCGCCCTGCT
CAAGCCCCATGGCAGCCTTCTCTCTCGTATGAGAAAGTGTTTAGAAGAGGCAAGGAAGAT
TTTCTGGCTCTTATTAACGCAGTCATGGAGGCCAAGAACTA.

Sequence result for MCF GSTM3:

GATGCCTTATGTTACTGGTGTCCAGCTGAGTTTCTCTTGGGTATAAAGGCTAAAAGGGAA
AAAGGATATGTGGAGAATCATCAAGATATGAATTGAATCGCTGCAGACTGTGGCATTTCCC
TGCTCCATTTGGCCTATCTATACTTCACAGACGAAATCACAGGAGAGGCATTATATGTCTG
GGCGGTCAGTACCAACGTGGAATGGGTGGCCACGTTCGCATAGCCTGCAGGAGCTTGAG
GGCCTTTGGATGGAGGCTGCGATACTGTGGCATTTCCC.

Sequencing Result MCF GSTO1:

CCAACGGGGAAGACTGGGCAGGGTTTTCCTAGAGCTCTACTTACAGAAACAGCCCTGAGG
CCTGTGATCTACTCGTCGCGCGCTCTGCAAGAGGGAGCAGAGAATGGTTTCTCTCCCAAA
AGAGAAA.

Sequencing result for MCF GSTT1:

GCACGAGTAGAGCTGAGAGATGTACTGGGACCTGCTGTCCCAGCCCTGCCGCGCTGTTTA
CATCTTTGCAAGAAGAACGACATTCCCTTCGAGCTGCGCATCGTGGATCTGATTAAAGGTC
AGCACTTAAGCGATGCCTTTGCCCAGGTGACACCCCCATCAACACAAAACTTTGCCCTCTG
TGCCTAGCTAAAGCTTGTCACTCAGCCTGAGCCTGCTGCTAGTCACGTGAAAACCCGCCTT
TCCATGGCGGGGGGGGGGAG.

Sequencing result for MCF GSTZ1:

CATATCTCTATAAGGGATGGGGGGGCCACAGCTTTTCTAAGGGACTTCCAGGGCACTGAA
TCCTATGAAGCAGGTGCCAACCCTGAAGATTGATGGAATCACCATTCACCAGTCACTGGCC
ATCATTGAGTATCTAGAGGAGACGCGTCCCCACTCCGCGACTTCACAGGCGACTTCATATC
TCGGCCACGCCTGGGCACCCTCAACGTGATTAGCAGGTTCCATACGCTCTAAAGTCACACC
CCGCGCACTTCAGCAAGTTATTTCGGAAGGGGAGGGAAGATTTCCCATGGAAACCAGGTT
CTGCATGAAGGGTCCAAAAAGGCTTATAGGCTTGACCTGACAAAAACAGGCCAAAAAGTG
TTATGATGGTGCTGAAAAGAAAAGAATTTTGCTGGTCTGCCCCCTCTGGCCCAGGGTCCG
GACCCTCTCCTGCCTCCTGATCCCCTCTGGACAGTGGCCATCTTGAATACTGCAAGGAAAG
GCGTCCCCCTC.