# CHARACTERIZATION OF CHEMOSENSITIVITY PROFILES OF BREAST CANCER CELL LINES, WITH AND WITHOUT STEM CELL LIKE FEATURES

**A THESIS**

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

**BY**

**MUHAMMAD WAQAS AKBAR**

**August, 2014**

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

Assist. Prof. Dr. Ali Osmay GÜRE

(Advisor)

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

Assoc. Prof. Dr. Burçak VURAL

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

Assist. Prof. Dr. Özgür ŞAHIN

Approved for the Graduate School of Engineering and Science

Prof. Dr. Levent ONURAL
Director of the Graduate School

## ABSTRACT

## CHARACTERIZATION OF CHEMOSENSITIVITY PROFILES OF BREAST CANCER CELL LINES, WITH AND WITHOUT STEM CELL LIKE FEATURES

**Muhammad Waqas Akbar**

**Supervisor: Assist. Prof. Dr. Ali Osmay GÜRE**

**August, 2014**

Breast cancer is the second most common cause of death worldwide from cancer due to complications with its diagnosis and resistance to therapy. Recent studies have shown that breast tumors when compared with other solid tumors also contain a subpopulation termed as cancer stem cells (CSCs). CSCs are hard to kill due to their therapy resistant capacities. These unharmed cells then result into relapse of tumor after treatment. Some established breast cancer cell lines also behave in similar fashion to CSCs in overall manner thus termed as CSC like cell lines. This study primarily focuses on characterizing CSC like cell lines from non CSC like cell lines based upon their gene expression and prediction of drugs which can target these groups separately. In this study two databases, Cancer Cell Line Encyclopedia (CCLE) and Cancer Genome Project (CGP), were used which contain gene expression data and drugs cytotoxicity data for most of the established cancer cell lines. Breast cancer cell lines gene expression data was used to predict two gene lists which can separate breast cancer cell lines into CSC like and non CSC like cell lines by *in silico* analysis. These gene lists were named as Patentable and Non Patentable. Additionally four drugs were predicted which can target CSC like group (Midostaurin and Elesclomol) and non CSC like group (Panobinostat and Lapatinib) separately. Later these findings were validated *in vitro*. Non Patentable gene list could not be validated due to low concordance with microarray data. On the other hand, Patentable gene list was validated and was found concordant with microarray data. Out of four selected drugs, Panobinostat and Lapatinib showed increased toxicity to non CSC like cell lines while only Midostaurin showed toxicity to CSC like cell lines. To investigate further that cell lines were grown in 3D cell culture conditions, to increase their stem cell like properties (stemness). But only one cell line MDA-MB-157 which was found as CSC like, showed expected behavior. Additionally this cell line increased resistance to Lapatinib and Panobinostat and became more sensitive to Midostaurin. Correlation analysis showed some genes as potential biomarkers for selected drugs. In conclusion, in this study various genes are proposed to differentiate CSC like

cell lines from non CSC like cell lines. And Midostaurin can be potential drug to treat CSC like cells while Lapatinib and Panobinostat showed increased activity against non CSC like cell lines.

Keywords: Breast cancer, Cancer stem cells, Midostaurin, Elesclomol, Lapatinib, Panobinostat, 3D cell culture, mammosphere, CCLE, CGP, qPCR.

## ÖZET

## KÖK-HÜCRE ÖZELLİĞİ OLAN VE OLMAYAN MEME KANSERI HÜCRE HATLARININ ILAÇ HASSASIYET PROFILLERININ TANIMLANMASI

**Muhammad Waqas Akbar**

**Danışman: Assist. Prof. Dr. Ali Osmay GÜRE**

**Ağustos, 2014**

Meme kanseri, teşhisindeki zorluk ve tedaviye direncinden dolayı dünyada en ölümcül ikinci kanserdir. Literatürdeki son çalışmalar göstermiştir ki meme tümörleri diğer katı tümörler gibi kanser kök hücresi (KKH) olarak adlandırılan bir sup-grup içermektedir. KKH`lerinin kanser tedavilerine dirençli yapılarından dolayı öldürülmesi güçtür. İlacın etki etmediği bu grup hücreler tedavi sonrası tümörün nüksetmesine neden olmaktadır. Bilinen bazı meme kanseri hücre hatları ise genel olarak KKH gibi davranmakta ve kanser kök hücresi-benzeri hücre hatları olarak adlandırılmaktadırlar. Bu çalışmada temel olarak KKH benzeri ve KKH benzeri olmayan hücre hatlarını, gen ifadesi ve bu iki gruba ayrı ayrı etki eden ilaçlara verdikleri cevaplara dayanarak karakterize etmeyi amaçladık. Bilinen hücre hatlarının çoğunun gen ifade ve ilaçlara hassasiyet bilgisini içeren iki veritabanı kullandık, Kanser Hücre Hatları Ansiklopedisi (Cancer Cell Line Encyclopedia, CCLE) and Kanser Genom Projesi (Cancer Genome Project, CGP). Meme kanseri hücre hatlarının gen ifade bilgisi, bu hücre hatlarını KKH benzeri ve KKH benzeri olmayan gruplara ayırabilen 2 farklı gen listesini in siliko analizler sonucunda elde etmek için kullanılmıştır. Bu gen listelerini Patentlenebilen ve Patentlenemeyen listeler olarak adlandırdık. Buna ek olarak KKH benzeri hücrelere ve KKH benzeri olmayan hücrelere ayrı ayrı etki edebilecek ikişer ilaç (sırasıyla, Midostaurin, Elesklomol, Panobinostat ve Lapatinib) belirledik. Daha sonra bu bulguları in vitro çalışmalarla doğruladık. Patentlenemeyen gen listesini mikroarray bilgisi ile uyumlu olmamasından dolayı doğrulayamadık. Öte yandan, Patentlenebilen gen ifadesi mikroarray verisiyle uyumlu sonuç verdi. Aynı zamanda iki grup arasında istatistiksel olarak anlamlı ifade farkı elde ettik. Seçilen 4 ilaçtan, Panobinostat ve Lapatinib KKH benzeri olmayan hücre hatlarına daha etkili iken, Midostaurin KKH benzeri hücre hatlarında yüksek toksisite gösterdi. Daha derin araştırmak için hücre hatlarını kök hücre özelliklerini arttırmak amacıyla 3 boyutlu kültürde büyüttük. Ancak KKH benzeri olarak belirlenen hucre hatlarından sadece MDA-MB-157 hücre hattı beklenen sonucu verdi. Buna ek olarak, bu hücre hattı Lapatinib ve Panobinostat ilaçlarına karşı direnç artışı gösterirken, Midostaurin

ilacına karşı daha hassas hale geldi. Korelasyon analizleri gösterdi ki bazı genler seçilen ilaçlar için potansiyel biyobelirteç görevi görebilir. Sonuç olarak, bu çalışmada KKH benzeri hücre hatlarını KKH benzeri olmayan hücre hatlarından ayırabilecek çeşitli genler önerilmiştir. Uzun vadede Midostaurin KKH benzeri hücreleri; Lapatinib ve Panobinostat KKH benzeri olmayan hücreleri hedefleyen potansiyel ilaç olabilir.

Anahtar kelimeler: Meme kanseri, Kanser kök hücreleri, Midostaurin, Elesklomol,Lapatinib, Panobinostat, 3B hücre kültürü, mammosphere, CCLE, CGP, qPCR.

# To MY Family

# ACKNOWLEDGMENTS

I have great pleasure to express deepest gratitude to my caring and helpful supervisor Assist. Prof. Dr. Ali Osmay Güre. I am very grateful to him for his valuable guidance, keen interest, persistent guidance and ever willing support throughout the course of this project. His ever optimistic and evocative views for completion of this project in time have made possible a living dream. He always encouraged me to look in broader perspective. I learned a lot under his supervision. It is a privilege for me to be a part of his lab.

Secondly I owe my profound regards to Assoc. Prof. Dr. Burçak Vural and Assist. Prof. Dr. Özgür Şahin for their valuable counsel and insights for the preparation of this thesis and project future plans.

I cannot thank enough my lab member; Murat İşbilen for his valuable contributions to this project and for his support and motivation which made my stay a lot easier in Turkey. He is my true kind, caring and loyal brother.

I would like to extend heartiest thanks to all members of AOG Lab, Mehdi Ghassemi, Kerem Mert Şenses, Sinem Yilmaz, Şükrü Atakan, Seçil Demirkol, Barış Küçükkaraduman and Alper Poyraz. They were real support and caring. Especially Kerem, Secil and Mehdi guided me a lot with their experiences and they were always there when I needed them the most. Presence of Alper and Barış in lab made every moment memorable.

I am sincerely thankful to Umar Raza who encouraged and supported me at every turn of events. He made my stay in Bilkent quite peaceful. I would also like to thank Huma Shehwana, Gurbet Karahan, Nilufer Sayar, Emre Yurdusev and Erol Eyüpoğlu for their help and support.

I also thank Asad Ali, Ali Haider, Naveed Mehmood and Naveed-ul-Mustafa from the core of my heart. I cannot forget the moments we spent together “while killing waves”.

Last but not least, I admit that I owe all my achievements to my sincere and most loving parents and my sisters, whose prayers and motivational support always remained a source of determination for me.

Muhammad Waqas Akbar

## TABLE OF CONTENTS


**1 Introduction**

1.1 Breast Cancer ..... 1

1.2 Classification of Breast Cancer ..... 1

1.3 Cancer Stem Cells ..... 1

1.4 Breast Cancer Stem Cells ..... 2

1.5 Markers for Breast Cancer Stem Cells ..... 2

1.6 Epithelial - Mesenchymal Transition and Cancer Stem Cells ..... 3

1.7 Factors Associated With Induction of Stemness Phenotype ..... 3

1.8 Resistance of Cancer Stem Cells to Therapy ..... 4

1.9 Cancer Stem Cells in Cell Lines ..... 5

1.10 Cancer Stem Cell Culturing ..... 6

1.11 Cancer Cell Line Encyclopedia and Cancer Genome Project ..... 6

1.12 Patentable Gene List and Non Patentable gene lists ..... 7

**2 Aims of the study ..... 8**

2.1 Scheme Of the study ..... 8

**3 Materials ..... 9**

3.1 General Laboratory Reagents and Equipment ..... 9

3.2 General Laboratory Solutions ..... 9

3.3 Cell Culture Materials and Reagents ..... 9

3.4 Cell Culture Media and Solutions ..... 10

**4 Methods ..... 13**

4.1 Cell Culture Techniques ..... 13

4.2 Culturing Preserved Cells ..... 13

4.3 Culturing Cells in 2D Culture ..... 13

4.4 Cryopreservation of Cells ..... 13

4.5 Culturing Cells in 3D Culture ..... 14

4.6 Collection of Cells for RNA Extraction ..... 14

4.7 RNA Extraction ..... 15
4.8 DNAse Treatment ..... 15
4.9 RNA Quantification by Qubit ..... 15
4.10 cDNA Synthesis ..... 16
4.11 End Point Polymerase Chain Reaction ..... 16
4.12 Agarose Gel Electrophoresis ..... 16
4.13 Quantitative Real time Polymerase Chain Reaction (qRT-PCR) ..... 17
4.14 Cell Counting ..... 19
4.15 Drug Cytotoxicity ..... 19
4.16 Calculation of Cytotoxicity Parameters ..... 20
4.17 Software Programs Used in this Study ..... 23
4.18 GraphPad Prism 6.0 ..... 23
4.19 Principle Component Analysis ..... 23
**5 Results ..... 24**
5.1 *In Silico* Results ..... 24
5.1.1 Determination of Gene lists differentiating CSC and Non CSC like breast cancer cell lines 24
5.1.2 Validation of Gene Lists ..... 29
5.1.3 Molecular Subtyping of Breast Cancer Cell Lines ..... 32
5.1.4 Prediction of Drug Response ..... 33
5.2 *In vitro* validation ..... 37
5.2.1 Validation in 2D cell culture conditions ..... 38
5.2.2 Validation in 3D cell culture conditions ..... 55
5.3 Correlation between Gene expression and cytotoxicity data ..... 63
5.4 Principle Component Analysis ..... 64
5.5 Concordance of Microarray Data with SybrGreen and TaqMan qPCR Data ..... 68
5.6 Conclusion ..... **Error! Bookmark not defined.**
**6 Discussion & Future Perspectives ..... 73**
6.1 Non-Patentable Gene List ..... 73

6.2 Patentable Gene List ..... 74
6.3 Drugs selected for CSC like and non CSC like cell lines ..... 75
6.4 Correlation Analysis ..... 77
6.5 Concordance of Microarray Data with qPCR Validation Data ..... 77
6.6 Future Perspectives ..... **Error! Bookmark not defined.**
**7 References ..... 78**
**8 Appendix ..... 83**
8.1 Validation of drug response in 2D cultured cells ..... 83
8.2 Comparison of gene expression between 2D and 3D culture conditions of Non Patentable Gene List 90
8.3 Comparison of gene expression between 2D and 3D culture conditions of Non Patentable Gene List 92
8.4 Drug cytotoxicity for 3D cell culture conditions ..... 98
8.5 Comparison of gene expression of MDA-MB-153 between 2D and 3D culture conditions ..... 106

# LIST OF FIGURES

Figure0.1: Cancer stem cells niche……..2

Figure0.2: Targeting cancer stem cells and non-cancer stem cells…… ……5

Figure 1.3: Parameters used for determination of drug's cell cytotoxicity. ……..7

Figure 4.1: Percent Viability curves for HCC1937 treated with Lapatinib……... 22

Figure 5.1: Common genes between genes selected from CCLE and CGP datasets…… 25

Figure 5.2: CCLE and CGP Datasets were clustered by using Non-Patentable Gene List. ……..26

Figure 5.3: CCLE and Datasets were clustered by using Patentable Gene List…… 28

Figure 5.4: Validation with Bhat-Nakshatri et al dataset with Non-Patentable Gene List…… 29

Figure 5.5: Validation with Bhat-Nakshatri et al dataset with Patentable Gene List…….. 30

Figure 5.6: Comparison of Elesclomol cytotoxicity against Breast cancer cell lines…….. 34

Figure 5.7: Comparison of Midostaurin cytotoxicity against Breast cancer cell lines……. 35

Figure 5.8: Comparison of Lapatinib cytotoxicity against Breast cancer cell lines……36

Figure 5.9: Comparison of Panobinostat cytotoxicity against Breast cancer cell lines……37

Figure 5.10: Comparison of expression of PVRL3, FBN1, EMP3 and GNG11 across breast cancer cell lines……38

Figure 5.11: Comparison of expression of RAB25, SPINT1, CLDN7 and GRHL2 across breast cancer cell lines……..39

Figure 5.12: Correlation of genes of Non Patentable genes with each other……40

Figure 5.13: Comparison of VIM, FN1, E-CAD and CLDN4 between breast cancer cell lines……..42

Figure 5.14: Comparison of BNC2, DDR2, PVRL3, DKK3, SLIT2 and TMEM158 in breast cancer cell lines……43

Figure 5.15: Comparison of ST14, BSPRY, IRF6, ZNF165 and AP1M2 in breast cancer cell lines……..45

Figure 5.16: Correlation between 2D and 3D gene expression for Patentable gene list……..46

Figure 5.17: Bar plots showing Comparison of Drug IC50 Midostaurin, Elesclomol, Lapatinib and Panobinostat against CSC like and non CSC like cell lines……50

Figure 5.18: Bar plots showing Comparison of Drug EC50 Midostaurin, Elesclomol, Lapatinib and Panobinostat against CSC like and non CSC like cell lines……51

Figure 5.19: Bar plots showing Comparison of Drug Activity area Midostaurin, Elesclomol, Lapatinib and Panobinostat against CSC like and non CSC like cell lines……52

Figure 5.20: Bar plots showing Comparison of Drug Amax Midostaurin, Elesclomol, Lapatinib and Panobinostat against CSC like and non CSC like cell lines……53

Figure 5.21: Cumulative drug response of drugs parameters against non CSC like vs. CSC like IC50, EC50, Activity area and Amax……54

Figure 5.22: Mammosphere formed by Breast cancer cell lines……56

Figure 5.23: Comparison of VIM, FN1, E-CAD and CLDN4 between breast cancer cell lines……...57-58

Figure 5.24: Comparison of selected genes VIM, FN1, E-CAD and PVRL3, ST14, BSPRY and IRF6 in MDA-MB-157 breast cancer cell line……..59

Figure 5.25: Bar plots showing Comparison of Drug parameters for Midostaurin, Lapatinib and Panobinostat against MDA-MB157 cells cultured in 2D and 3D culture conditions……62

Figure 5.26: Pearson correlation analysis results between gene expression and drug cytotoxicity parameters……..63

Figure 5.27: Correlation of PC1 of AP1M2 and SLIT2 with Elesclomol IC50……..65

Figure 5.28: Correlation of PC1 of ZNF165 with Midostaurin EC50……..66

Figure 5.29: Correlation of PC1 of E-CAD, SLIT2 and TMEM158 with Panobinostat Activity area…….67

Figure 5.30: Correlation of PC1 of PVRL3 and SLIT2 with Panobinostat Amax……68

Figure 5.31: Analysis for explaining concordance between microarray data and TaqMan qPCR data…..69

Figure 5.32: Analysis for explaining concordance between microarray data and SybrGreen qPCR data.....70

Figure 8.1: Percent Cell Viability curves for Panobinostat against breast cancer cell lines……..83

Figure 8.2: Percent Cell Viability curves for Lapatinib against breast cancer cell lines……85

Figure 8.3: Percent Cell Viability curves for Midostaurin against breast cancer cell lines…… 87

Figure 8.4: Percent Cell Viability curves for Elesclomol against breast cancer cell lines……..89

Figure 8.5: Comparison of PVRL3, FBN1, EMP3 and GNG11 in both 2D and 3D……91

Figure 8.6: Comparison of RAB25, SPINT1, CLDN7 and GRHL2 in both 2D and 3D cultured cells……..92

Figure 8.7: Comparison of BNC2, DDR2, PVRL3, DKK3, SLIT2 and TMEM158 in breast cancer cell lines in both 2D and 3D……93-95

Figure 8.8: Comparison of ST14, BSPRY, IRF6, ZNF165 and AP1M2 in breast cancer cell lines in both 2D and 3D……96-98

Figure 8.9: Percent Cell Viability curves for Panobinostat against breast cancer cell lines……..99

Figure 8.10: Percent Cell Viability curves for Lapatinib against breast cancer cell lines……101

Figure 8.11: Percent Cell Viability curves for Midostaurin against breast cancer cell lines……103

Figure 8.12: Percent Cell Viability curves for Elesclomol against breast cancer cell lines……..105

Figure 8.13: Comparison of not selected genes CLDN4, BNC2, DDR2, DKK3, SLIT2, TMEM158, IRF6 and AP1M2 in MDA-MB-157 breast cancer cell line……..106

# LIST OF TABLES

Table 0.1: Materials purchased for use in general cell culture procedures and the companies…… 9

Table 3.2: Cell culture medium and reagents……10

Table 3.3: Composition of media for 3D culture……11

Table4.1: Predesigned TaqMan probes……17

Table 4.2: Sequence and melting temperature of each gene used with SybrGreen assay……..18-19

Table 4.3: OD values for HCC1937 treated with Lapatinib……20

Table 4.4: % viability values for HCC1937 treated with Lapatinib……21

Table 4.5: Standard error of model, log (IC50), log (EC50), Activity area and Amax values for HCC1937 treated with Lapatinib……21

Table 5.1: Non-Patentable Gene List……...25

Table 5.2: Patentable Gene List. ……..27

Table 5.3: Details of the datasets selected for further validation……...31

Table 5.4: Comparison of selected genes behavior among different datasets (Downregulated in CSC like cells)……31

Table 5.5: Comparison of selected genes behavior among different datasets (Upregulated in CSC like cells)……32

Table 5.6: Comparison of Molecular subtypes of breast cancer cell lines with CSC like characterization for both CCLE and CGP datasets……32

Table 5.7: Selected drugs, their mode of action and molecular subtype of breast cancer cell lines targeted……33

Table 5.8: Capability of Non Patentable genes to separate CSC like from non CSC like……40

Table 5.9: Concordance between microarray and qPCR gene expression data for Non Patentable gene list……41

Table 5.10: Capability of Non Patentable genes to separate CSC like from non CSC like……..47

Table 5.11: Concordance between microarray and qPCR gene expression data for Patentable gene list47

Table 5.12: Genes showed clear difference in graph bars between CSC like and non CSC like……48

Table 5.13: IC50, EC50, Activity area and Amax values for Panobinostat, Lapatinib, Midostaurin and Elesclomol in 2D cell culture conditions……49

Table 5.14: Comparison of drug parameters for CSC like cells vs. non CSC like cells using 2 sided T.Test……55

Table 5.15: IC50, EC50, Activity area and Amax values for Panobinostat, Lapatinib, Midostaurin and Elesclomol in 3D cell culture conditions……61

Table 5.16: Concordance analysis of Microarray with TaqMan Probes result……..71

Table 5.17: Concordance analysis of Microarray with SybrGreen result……...71

# ABBREVIATIONS

| | |
|---|---|
| $^0C$ | Degree Centigrade |
| CSC | Cancer Stem Cell |
| PCR | Polymerase Chain Reaction |
| TNBC | Triple Negative Breast Cancer |
| ALDH | Aldehyde Dehydrogenase |
| BCSC | Breast Cancer Stem Cells |
| $CO_2$ | Carbon dioxide |
| Conc. | Concentration |
| $ddH_2O$ | Double Distilled Water |
| DMSO | DimethylSulphoxide |
| EDTA | Ethylene diamine tetra acetic acid |
| EGF | Epidermal Growth Factor |
| ER | Estrogen Receptor |
| PR | Progesterone Receptor |
| FACS | Fluorescent Activated Cell Sorting |
| FBS | Fetal Bovine Serum |
| HDAC | Histone Deacetylase Inhibitor |
| HER2(ERBB2) | Human Epidermal Growth Factor Receptor 2 |
| IC50 | The half maximal inhibitory concentration |
| EC50 | The half maximal effective concentration |
| NA | Not Available |

# 1 INTRODUCTION

## 1.1 Breast Cancer

Among all the causes of morbidity and mortality in women, role of breast cancer is highly intense. Its importance is huge as it is, after lung cancer, the most deadly cancer around the globe[1]. According to Cancer statistics, 2014, estimated onset of breast cancer cases will be 235,030 (2,360 cases in men and 232,670 cases in women) and estimated deaths are 40,430 (430 in men and 40,000 in women) in USA only[2].

## 1.2 Classification of Breast Cancer

Breast cancer is mostly classified based on clinical and pathological characteristics. Estrogen receptor (ER), progesterone receptor (PR) and ERBB2 receptor are the markers which are being used for early stratification of breast tumors. But this kind of classification cannot reflect true heterogeneous nature of breast tumors. This heterogeneity between tumors arises due to transcriptional variations in biological programs[3]. Molecular characterization of breast tumors has led us to better classification which relates with prognosis and drug response. Currently, based upon gene expression analyses of tumors, breast tumors are divided into different groups such as Luminal A, Luminal B, Basal, ERBB2 enriched, Claudin low and normal like[4, 5].

## 1.3 Cancer Stem Cells

Various studies have made it clear that breast tumors contain a subpopulation of cancer stem cells (CSCs). CSCs are defined as “a cell within a tumor that possess the capacity to self-renew and to cause the heterogeneous lineages of cancer cells that comprise the tumor”[6]. These cells have many features similar to their normal counterparts, normal adult stem cells. Asymmetrical division is one of these characteristics by which these cells produce more CSCs and also phenotypically diverse cells (non CSCs) which compose the major portion of tumor. These CSCs are also responsible for malignancy, metastasis, aggressive tumors and breast cancer relapse [7, 8].

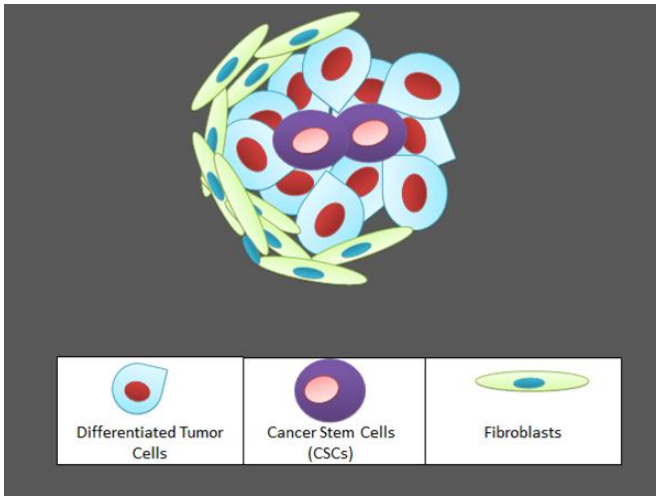



**Figure 1.1: Cancer stem cells niche**

### 1.4 Breast Cancer Stem Cells

Breast cancer stem cells (BCSCs) were first reported by *Al hajj et al* based on cell surface markers in 2003. They sorted $CD44^{+}/CD24^{-}$ presenting cells through FACS and found that these cells produce aggressive tumors when injected in mammary fat pad of NOD/SCID mice even when injected as low as 200 cells. But the cells showing CD44+/CD24+ failed to form tumors even at high numbers such as 20,000 and unsorted cells when injected at 10,000 number could form tumors only in 25% of animals upon injections[9]. There are different theories about the origin of CSCs. First one is cancer stem cell hypothesis that states that CSCs originate through transformation of normal stem cells or early progenitor cells. Second theory states that CSCs originate from epithelial-mesenchymal transition (EMT)[7]. Another theory against CSC hypothesis is clonal evolution model which states that tumors arise from deviant normal cells which under the influence of accumulated mutations divide insanely. The progeny of that deviant cell will acquire more mutations giving rise to heterogeneous tumor bulk[10, 11]. But this bias in concepts has been justified by combining CSC and clonal evolution model. This new concept reveals that frequency of CSCs in patients, differ dramatically. And this is also dependent upon dominant mutation, gene amplifications and deletions. So only the dominant CSCs survive and emerge as resistant ones[12]. Claudin low tumors have been reported as highly enriched for cancer stem cells in breast cancer[13].

### 1.5 Markers for Breast Cancer Stem Cells

CSCs are identified by various cell surface markers but there is not a single universal marker. This may be because of the heterogeneous nature of breast cancer. But the most accepted marker which is being used by many research groups is $CD44^{+}/CD24^{-}$ , first discovered by *Al Hajj et al*[9]. Other cell surface markers include are aldehyde dehydrogenase1+ (ALDH1), CD49f+, CD133+, GD+, CD271, PKH+ and

ANTXR1[14-20]. These markers identify CSC phenotypes differently on the basis of cell differentiation[21]. These markers are used in various combinations. Among the techniques which are used for characterizing cells with the help of above mentioned markers are Fluorescence-activated cell sorting (FACS) and immunohistochemistry. Microarray and RT-PCR can also be used for specifying CSCs from Non CSCs. Most of the studies conducting research over breast cancer stem cells (BCSCs), cell are sorted over expression of stem cell markers and then used for further experiments.

### 1.6 Epithelial - Mesenchymal Transition and Cancer Stem Cells

Epithelial - mesenchymal transition (EMT) is the process of transformation of non-mobile epithelial like cells into mobile mesenchymal like cells. This process was first recognized due to its role in embryogenesis. Opposite of EMT is mesenchymal - epithelial transformation (MET)[22]. Apart from natural role of EMT, this process also plays its part also in caner development. In tumors EMT and MET coexist. But the most alarming part is CSC cells share a lot of similar characteristics with the cells undergone EMT. Both kinds show characteristics like capability to generate whole tumors, metastasis, lack of differentiation, participation in formation of mammosphere and resistance e to anti-cancer therapies[23]. Additionally anoikis is any type of apoptosis which can be inhibited by cell to matrix interaction and is critical for tumor metastasis. Anoikis is suppressed both in EMT and CSCs and this is one of the reason that spheres can be generated from these CSC cells[24]. Previously in a study, human mammary epithelial cells (HMLE) cells were induced for ectopic expression of Twist or Snail (EMT inducers), most of the resultant mesenchymal cells also expressed $CD44^{high}/CD24^{low}$. Additionally, these cells were also capable of generating basal and luminal lineage cells[25]. Moreover stem like cells isolated from mouse and human mammary glands were found to express EMT markers. p53 inactivation, EMT and stem cell like properties are also related closely[26]. From these evidences, at least this can be concluded that EMT and CSCs are induced by same mechanism.

### 1.7 Factors Associated With Induction of Stemness Phenotype

Various factors which are known to be associated with induction of stemness properties have been reported in literature previously. For instance Transforming growth factor-beta (TGF-β) plays its role in increasing breast CSCs through EMT especially in Claudin low tumors[27]. BRCA1 mutant or deficient cells, are reported as well for increased accumulation of stem cells. This factor can be arisen due to loss of double stranded DNA repair in these cells[28]. Cytokines have also been previously reported as key regulators for BCSCs. Most important of these cytokines is IL6[29]. IL8 is also reported to be excreted at 10 fold higher amounts when MCF7 cells were cultured on collagen scaffold as compared to culturing in adherent conditions[30]. Additionally, overexpression of IL8, also plays an important role in EMT as well[31]. IL6 also has been reported to impart resistance for HER2+ breast cancers by increasing the number of BCSCs population[32]. SLUG/SNAI2 and Tumor necrosis factor (TNF) are also reported to generate

BCSCs. MCF10A cells when transfected with SLUG, showed $CD44^{+}/CD24^{-}$ high signature and same pattern was seen when these cells were treated with TNF[33].

### 1.8 Resistance of Cancer Stem Cells to Therapy

Conventional drugs used for breast cancer only target non CSCs in most of the cases. So CSCs causes relapse of tumor by differentiation. These CSCs are resistant to chemotherapy, radiation therapy, endocrine therapy and apoptosis. Wnt, Notch, Hedgehog and TGF beta signaling pathways interact with each other in CSCs aiding their mechanisms of evasion from cancer therapies[22, 23, 34]. So to target these cells, different strategies should be opted. CSC signaling pathways are important target for therapeutic agents. Among other targets to treat CSCs some are differentiation therapies, DNA repair mechanisms, apoptotic resistance and CSC microenvironment[12]. Additionally cellular surface markers from which CSCs are identified, can be used as targets to destroy these cells[35]. So by combining traditional therapy with CSC targeted therapy, CSC issue can be overcome.

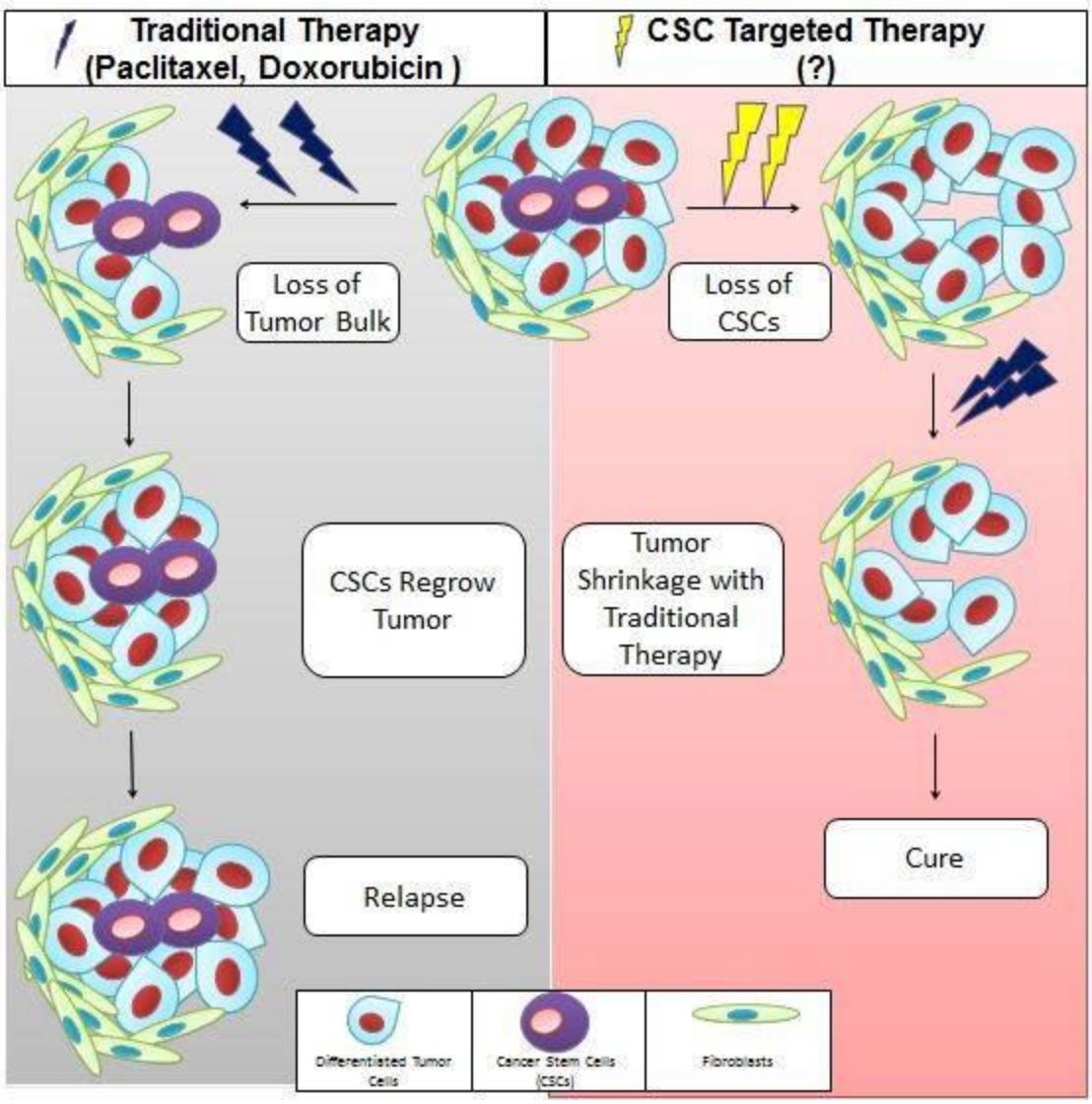



**Figure 1.2: Targeting cancer stem cells and non-cancer stem cells**.

### 1.9 Cancer Stem Cells in Cell Lines

It has been shown in many studies that Breast cancer cell lines also contain CSCs. *Fillmore et al* showed presence of stem cells in eight different breast cancer cell lines using CD24, CD44 and epithelial-specific antigen (ESA) expression. Stem cells present in cell lines were referred as CSC like cells and these cells when cultured, can grow to parental cell line[36]. *Jauffret et al* later showed presence of stem cells in 33 different breast cancer cell lines using ALDH1 marker. And when ALDH1+ cells were injected in mice, very small number of these cells could generate tumors in mice[37].

### 1.10 Cancer Stem Cell Culturing

Sorted CSCs are mostly cultured in absence of attachment as spheres *in vitro* (also referred as 3D culturing). In breast cancer cells these spheres are called as mammosphere. This culturing method enriches CSCs but when cells from these spheres are re-cultured in form of monolayers (2D culturing) they lose their stem cell phenotype[16]. Among several methods to culture spheres, some are: gel based culture, low binding plates, rotary culture and hanging drop method[38]. Culturing spheres is a faster and cheaper method but it has some limitations as important factors required for in vivo growth for these cells are not provided and time required for these cells to grow and differentiate is more as compared to in vivo systems. CSCs are also transplanted in immunodeficient mice where even in smaller numbers they can generate tumors. And these tumors can metastasize and resemble parental cancer cells. But this method also has some limitations such as mice short life, altered niche for CSCs and lack of cytokines etc.[39].

### 1.11 Cancer Cell Line Encyclopedia and Cancer Genome Project

Role of human cancer cell lines in understanding tumor biology and various discoveries of drug is irrefutable. In 2012 two databases were published containing genomic and cytotoxicity data for cell lines. One of these data is Cancer Cell Line Encyclopedia (CCLE). CCLE contains genomic data for 947 human cancer cell lines. These cell lines cover 36 different types of tumor. Further, around 500 of these cell lines are profiled for 24 different pharmacological compounds. Mutational data for cell lines is also provided. Only for breast cancer, this database contains 56 cell lines[40]. Second database is Cancer Genome Project (CGP). This database contains genomic information for 639 cell lines and these cell lines were screened with 130 different compounds. Mutational data for these cell lines was also provided. This database contains 39 breast cancer cell lines[41]. In our project both of these databases were used for their gene expression datasets and drug profiling.

In CCLE, Dose responses curves were fitted using two different models (called as 3 and 4 parameter). Additionally, four different parameters were used to describe cytotoxic effects of drugs. These parameters were half maximal inhibitory concentration (IC50), half maximal effective concentration (EC50), maximum activity (Amax) and Activity area. Amax is the maximal activity reached within a dose response curve model and Activity area is referred to as area between activities at lowest concentration up to maximum tested concentration.

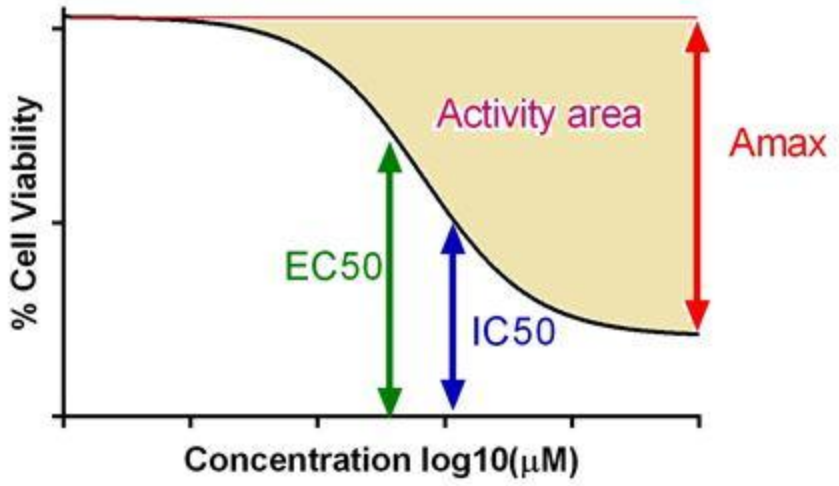



**Figure 1.3: Parameters used for determination of drug's cell cytotoxicity.**

### 1.12 Patentable Gene List and Non Patentable Gene Lists

Previously our group has published a paper based only on *in silico* data only, which shows that based upon their gene expression, some breast cancer cell lines also behave as stem cell like. Later breast cancer cell lines were divided into Basal A, Basal B and Luminal subtypes. And to our surprise, all cell lines in Basal B subtype showed CSC like characteristics[42]. Only CCLE data for 27 cell lines, was included in this paper.

In this project, we used more number of breast cancer cell lines as compared to previous, 56 cell lines from CCLE and also included CGP data, 39 cell lines. To characterize breast cancer cell lines into stem cell like and non-stem cell like groups, two gene lists were proposed. These gene lists were named as Non Patentable and Patentable gene lists. Non Patentable gene list was the one which contained best genes which can differentiate these groups into two but genes in this list were found to be described previously for their role in separation of CSC like cells from non CSC like. But we wanted to patent some genes which can separate these groups, so for this reason other genes were determined which could do the same and named this gene list as Patentable gene list.

# 2 Aims of the study

CSCs have been identified in many cancers including breast, heart, pancreas, lung, liver, skin and testes[43]. To get rid of tumors, targeting these cells is prerequisite. Like other cancers, in breast cancer, these cells lead to self- renewability, giving rise to different phenotypes causing tumor heterogeneity and resistance to therapies. When breast tumors are treated with traditional therapy, these cells get away unharmed and then result into relapse of tumor. So targeting these cells with traditional therapy rotationally seems to be perfect solution for CSC problem.

Generally BCSCs are sorted through FACS and then sorted cells are used for further studies. But in this project, one of our objectives is to characterize breast cancer cell lines in terms of their CSC behavior. To characterize these cells based upon said characters, we will define set of biomarker genes by which cell lines can be clustered in CSC or non CSC like phenotypes. Additionally other objective is to predict drugs which can specifically target those CSC like and non CSC like cells separately. And in the end these findings (both genes and drugs) will be validated in breast cancer cell lines.

## 2.1 Scheme of the Study

Breast cancer cell lines will be first stratified into CSC like and non CSC like by using a previously published gene list which is capable of dividing CSC from non CSC cells. Then differentially expressed genes will be determined between those stratified cell lines. Minimum number of genes will be determined which can cluster those cell lines into same manner. Additionally drugs will be identified which can target these subclasses specifically. Both drugs and genes will be then validated *in vitro*.

# 3 MATERIALS

## 3.1 General Laboratory Reagents and Equipment

Various reagents were bought from different sources. Chloroform, Isopropanol and Ethanol were bought from SIGMA-ALDRICH (St. Louis, MO, USA). DMSO (Dimethyl sulfoxide) was procured from AppliChem (Darmstadt, Germany). RNAse Zap and Nuclease Free Water were purchased from Ambion (Carlsbad, CA, U.S.). Qubit™ RNA BR Assay Kit and DNA-free™ Kit were purchased from Invitrogen (Carlsbad, CA, USA). Agilent RNA 6000 Nano Kit was purchased from Agilent Technologies (Santa Clara, CA, USA). RevertAid First Strand cDNA Synthesis Kit was purchased from Thermo Scientific (Waltham, MA, USA). OneTaq Quick-Load 2X Master Mix with Standard Buffer was purchased from New England BioLabs (Ipswich, MA, USA). TaqMan Gene Expression Assays and TaqMan Universal PCR Master Mix were used from Applied Biosystems (Carlsbad, CA, U.S.).

To evaluate the purity and quantity of RNA samples NanoDrop (Thermo Scientific, Wilmington, USA) was utilized. Additionally to measure the more explicit value of RNA samples Qubit Fluorometer (Invitrogen, Carlsbad, CA, USA) was used. Gel electrophoresis system was supplied by Hoefer Inc. (Holliston, MA).

## 3.2 General Laboratory Solutions

For preparation of 100 ml of 1% Agarose gel, 5 µl of Ethidium bromide was used from 10mg/ml stocks.

50X Tris Acetate EDTA was prepared by dissolving 37.2 g of Tritiplex III (EDTA), 242 g of Tris Base and 57.1 ml of glacial acetic acid in 1 liter of $ddH_2O$.

## 3.3 Cell Culture Materials and Reagents

**Table 3.1: Materials purchased for use in general cell culture procedures and the companies.**

| Material | Company |
|---|---|
| Serological pipettes | Costar Corning Incorporated (NY, USA) |
| Cell culture scrapers | Sarstedt (Nümbrecht, Germany) |
| Cell culture flasks (25, 75 $cm^2$) | Greiner Bio One (Monroe, NC, USA) |
| SFCA (Surfactant-Free Cellulose Acetate) membrane Serum filter | Thermo Fisher – Nalgene (Waltham, MA, USA) |
| Millex-FG Syringe Filter | Merck MilliPore (Billerica, MA, USA) |
| 96 well plates | Costar Corning Incorporated (NY, USA) |
| Ultra-Low attachment 96 well plates | Costar Corning Incorporated (NY, USA) |

| | |
|---|---|
| Ultra-Low attachment 75 cm cell culture flasks | Costar Corning Incorporated (NY, USA) |
| Cell Strainer, 40 µm | Becton Dickinson (NJ, USA) |
| 10, 20 , 200 and 1000 µl filtered tips for micropipettes | Greiner Bio One (Monroe, NC, USA) |

Cell culture medium and reagents were purchased from mentioned companies for routine cell culture as mentioned in table 3.2.

**Table 3.2: Cell Culture Medium and Reagents**

| Reagents | Company |
|---|---|
| Dulbecco's Modified Eagle Medium (DMEM) | Lonza (Basel, Switzerland) |
| Rosswell Park Memorial Institute Medium (RPMI) | GIBCO (Carlsbad, CA, USA) |
| B27 spheroid media supplement | Invitrogen, Carlsbad, CA, USA |
| Epidermal Growth Factor human (EGF) | SIGMA-ALDRICH (St. Louis, MO, USA) |
| Fibroblast Growth Factor-Basic human (FGF) | SIGMA-ALDRICH (St. Louis, MO, USA) |
| Heparin sodium salt | SIGMA-ALDRICH (St. Louis, MO, USA) |
| Trypsin-EDTA | SIGMA-ALDRICH (St. Louis, MO, USA) |
| Fetal Bovine Serum (FBS) | GIBCO (Carlsbad, CA, USA) |
| L-Glutamine | HyClone (Rockford, USA) |
| Fetal Bovine Serum (FBS) | (Carlsbad, CA, USA) |
| Penicillin –Streptomycin | HyClone (Rockford, USA) |
| CellTiter-Glo® Luminescent Cell Viability Assay | Promega (Fitchburg, Wisconsin, USA) |
| TrypLE™ Select Enzyme (1X), no phenol red | Life Technologies (Waltham, MA,USA) |
| TRIzol® Reagent | Life Technologies (Waltham, MA,USA) |

### 3.4 Cell Culture Media and Solutions

Breast cancer cell lines, MDA-MB-157, MDA-MB-231, MDA-MB-453, MDA-MB-468, CAMA-1 and MCF7 were cultured in DMEM media. T47D and HCC1937 cell lines were cultured in RPMI media.

Growth media for 2D cell culture was prepared by filtering 10% FBS, 1% L-Glutamine and 1% Penicillin – Streptomycin of the total media, through 0.2 µm syringe filters. This filtrate was then added to respective media.

EGF stock solution was prepared in $ddH_2O$ as 1 µg/µl. And from this intermediate and working stock solutions were prepared as 20ng/µl and 4ng/µl respectively.

FGF stock solution was prepared in Phosphate Buffer as 1 µg/µl. And from this intermediate and working stock solutions were prepared as 20ng/µl and 2ng/µl respectively.

Heparin stock solution was prepared in $ddH_2O$ as 100µg/µl. And from this intermediate and working stock solutions were prepared as 10µg/µl and 4µg/µl respectively.

To prepare Phosphate buffer, 5.52 g of monobasic sodium phosphate ($NaH_2PO_4$) was dissolved in 100 ml $ddH_2O$ and 5.68 g of dibasic sodium phosphate ($Na_2H_2PO_4$) was dissolved in 100 ml $ddH_2O$ separately. Then both solutions were added to each other to make 0.2 M Phosphate buffer which was then diluted 1:1 to make 0.1 M Phosphate buffer.

To prepare Growth media for 3D culture 1% L-Glutamine and 1% Penicillin –Streptomycin of the total media was filtered through 0.2 µm syringe filters. Later to prepare working media, EGF, FGF, Heparin and B27 supplement were added in the amount described in the following table 3.3.

**Table 3.3: Composition of media for 3D culture**

| Reagent | Quantity to be added<br><br>(from working stocks) | Final Concentration to be achieved |
|---|---|---|
| B27 (50X) | 1 ml from working socks | 1 X |
| FGF | 250 µl | 20 ng/ml |
| EGF | 250 µl | 10 ng/ml |
| Heparin | 200 µl | 2 µg/ml |
| Media | 48.3 ml | |
| Total | 50 ml | |

Media prepared for 3D culturing of cells was used only for 7 days, after which new media was prepared.

10X Phosphate buffered saline (PBS) was prepared by mixing 80g of Sodium Chloride (NaCl), 2g of Potassium Chloride (KCl), 2.4g of Mon potassium Phosphate ($KH_2PO_4$) and 14.4g of sodium phosphate ($Na_2HPO_4$) in 1 liter of dd$H_2O$. 10X PBS was then diluted to 1X and then autoclaved. Before use in cell culture, it was also filtered through SFCA membrane Serum filter.

Freezing medium for cells was prepared by adding 10% DMSO to 90% FBS of total volume of media.

For cell cytotoxicity analysis of 4 drugs, Panobinostat, Lapatinib, Elesclomol and Midostaurin were purchased from Selleck Chem, CAYMAN, Medchemexpress and Sigma-Aldrich respectively. All drugs were dissolved in DMSO. Stock concentration of Panobinostat, Lapatinib and Elesclomol was 50mM and Midostaurin stocks were made at 5mM.

# 4 METHODS

## 4.1 Cell Culture Techniques

Breast cancer cell lines used in this study are MDA-MB-157, MDA-MB-231, MDA-MB-453, MDA-MB-468, CAMA-1, T47D and HCC1937. These cell lines were grown in both 2D and 3D culturing conditions. All cell lines were incubated in 5% $CO_2$ incubator at 37 $^0$C.

## 4.2 Culturing Preserved Cells

Cryovials containing specified cell lines were previously stored in liquid nitrogen Tank. These were removed from there and frozen cells were thawed immediately by placing them in water at 37 $^0$C. These cells were then added to 15ml falcon tubes and 5 ml pre warmed growth medium was added on top of them very slowly. Each falcon tube was centrifuged for 3 minutes at 1500 rev/min. The Cells settled down and formed a pellet. Supernatant freezing medium containing DMSO was removed by aspirator and cell's pellet was re-suspended in 1 ml of fresh medium and this was added to 25 $cm^2$ flasks containing 5 ml media. Then that flask was placed in $CO_2$ Incubator at 37 $^0$C. After achieving 80-90% confluency, cells were transferred to 75$cm^2$ flasks.

## 4.3 Culturing Cells in 2D Culture

Cell lines were cultured in either DMEM or RPMI-1640 as illustrated in materials section. On achieving more than 70% confluency, cells were passaged. Passaging ratios was based upon the population doubling time of those cell lines. Fast growing cell lines were passaged 1:5 while cell lines whose doubling time was high were passaged 1:2. For passaging cells, growth media was removed by aspirator and 1X PBS was used to wash the cells. To detach cells from flask surface, pre-warmed 1ml Trypsin-EDTA was poured on the top of cells and spread to each corner of flask. Then flask was placed in $CO_2$ incubator at 37 $^0$C for 1-2 minutes. After observing detached cells through microscope, 5 ml fresh media was added to that flask to inactivate trypsin. This media now containing cells was pipetted up and down a few times to disperse cells. Then this media was transferred to 15ml falcon and centrifuged at 1500 rev/min for 3 minutes. Later supernatant was removed and cell pellet was re-suspended in 2 ml media and transferred in 75 $cm^2$ flasks containing 8ml media.

## 4.4 Cryopreservation of Cells

After passaging cells were incubated for 24 hrs. Growth media was removed confluent cell lines flasks and then cells were washed with 1X PBS. This was followed by trypsinization of those cells by using 2ml trypsin-EDTA. Cells were incubated for 2 min to detach. Later 5 ml fresh media was added to inactivate trypsin and this was transferred to 15ml falcon tubes. This tube was the centrifuged at 1500 rev/min for 3

minutes. Supernatant was removed and pellet of cells was re-suspended in freezing medium. The amount of freezing media was adjusted relative to confluency of cells. 1 ml form this was then transferred to each cryovial. These vials were labeled with name, passage number and date. After that, these vials were stored at -20 $^0$C for 1-2 hrs. and then at -80$^0$C or liquid nitrogen for long term storage.

### 4.5 Culturing Cells in 3D Culture

Sphere media or 3D media was prepared as described in materials part. For initiating 3D culturing of cells, cells were plated into ultra-low attachment flasks. To initiate this culture cells were first cultured in 2D culture. On achieving 80-90% confluency, these cells were trypsinized and centrifuged at 800rev/min for 4 minutes. Supernatant was removed and cells were re-suspended in 2 ml of sphere media. Then this was transferred to 75cm$^2$ ultra low attachment flask containing 8 ml media. This flask was then placed in $CO_2$ incubator. After 3 days, cells had formed spheres but there were single cells as well. To separate these spheroids from single cells and subculture them, a 40 µm cell strainer was put on top of a 50 ml falcon tube. Then spheroid suspension was taken out from flask and poured onto strainer. This way single cells having size less than 40 µm, went through strainer with media but spheroids left on the strainer. These spheroids were then transferred to a petri dish by inverting that strainer over petri dish and 1X PBS was poured on top of strainer. PBS containing spheroids was collected form petri dish and transferred to 15 ml falcon tube. Strainer was washed one more time, the same way, with 5 ml of PBS and that was also collected into same falcon tube. Then that tube was centrifuged at 800rev/min for 4 minutes. Later supernatant was removed. 3 ml of TrypLE was added and pipetted up and down to dissociate spheres both chemically and physically. Then this suspension was placed in incubator at 37 $^0$C. Then 6 ml of 1X PBS was added and mixed by pipetting. That tube was again centrifuged at 800rev/min for 4 minutes to collect pellet of spheroid cells. Then supernatant was removed and cells were re-suspended in 2 ml of sphere media. Then this was transferred to 75cm$^2$ ultra low attachment flask containing 8 ml media. Same procedure was repeated for sub culturing of spheroids one more time but for later passages, only filtration step was removed from protocol. Spheres were passaged 12 times post 1$^{st}$ filtration. Cells were not diluted for passaging, as number of cells decreased for some cell lines along the time. 1/3$^{rd}$ of pellet was separated for RNA isolation at 3$^{rd}$, 6$^{th}$, 9$^{th}$ and 12$^{th}$ passages and after these passages, cell cytotoxicity values were measured with these cells for four drugs[38].

### 4.6 Collection of Cells for RNA Extraction

For collection of cells from monolayer adherent culture, to 75cm$^2$ flask containing 80% confluent cell lines growth media was removed and washed with 1X PBS to remove dead cells. Then 6 ml ice cold 1X PBS was added and cells were scratched using cell scratcher. Scratched cells were later collected into PBS and transferred into 15 ml falcon tubes which were later centrifuged at 1500 rev/min for 4 minutes at 4$^o$C. PBS was then removed by aspirator and pellet was stored at -80 $^o$C for later RNA extraction.

For RNA extraction from spheroid, cells were separated during passaging protocol. For this purpose, 1/3$^{rd}$ spheroid suspension was separated and centrifuged at 1500 rev/min for 4 minutes at 4$^{o}$C. Later by removing supernatant, 6 ml ice cold 1X PBS was used to re-suspend pellet and then again centrifuged at 1500 rev/min for 4 minutes at 4$^{o}$C. Then supernatant PBS was removed and cell pellet was stored at -80 $^{o}$C for later RNA extraction.

### 4.7 RNA Extraction

The pellet of each cell line was homogenized with 1 ml TRIzol® Reagent (Life Technologies, Waltham, MA, USA), transferred to separate eppendorf for each sample and incubated for 5 minutes at room temperature. Later this was centrifuged at 13000rev/min for 15 minutes at 4$^{o}$C. On top this then chloroform was added to each sample and 10 minutes incubation was given followed by centrifugation at 13000rev/min for 15 minutes at 4$^{o}$C. Upper phases containing RNAs from each sample was transferred to a separate eppendorf. Then Isopropanol was added to each tube which was followed by 10 minute incubation at room temperature then centrifugation at 13000rev/min for 15 minutes at 4$^{o}$C. Now supernatant was removed and Ethyl Alcohol was added to wash extracted RNA. Pellet was incubated for 10 minutes at room temperature and centrifuged at 13000rev/min for 15 minutes at 4$^{o}$C. Then ethyl alcohol was removed and pellets were air dried. Then depending upon the sizes of pellets, these were dissolved in relative Nuclease Free Water (Ambion Carlsbad, CA, USA). Then each sample was incubated for 15 min at Thermolyne heater. To determine quality and quantity of each sample, RNA concentration was measured by Nanodrop spectrophotometer ND-1000 (NanoDrop Technologies; Wilmington, DE). Then samples were stored at -80 $^{o}$C.

### 4.8 DNAse Treatment

RNA isolation can also result in possible contamination of DNA in RNA samples. So to avoid such contamination, DNAse treatment was performed with each sample using DNA-Free kit (Ambion, Carlsbad, CA, USA). Samples were diluted to 200ng/µl. To each sample, 1µl rDNAse I and 5µl DNAse Buffer was added to 44µl of RNA sample. Total reaction of 50µl was prepared. Samples were then incubated at 37 $^{o}$C for 30 minutes on Thermolyne heater. To terminate the reaction, 5µl DNAse inactivation was added to each sample and pipetted for 2 minutes. Then each sample was centrifuged at 10000rev/min for 2 minutes. Supernatants contained DNA free RNAs. These were separated in new labeled eppendorfs.

### 4.9 RNA Quantification by Qubit

After DNAse treatment, quantity of RNA samples were quantified again by Qubit Fluorometer (Invitrogen) as it is more explicit to NanoDrop. The advantage of Qubit over NanoDrop comes as fluorescents bind directly to nucleic acid. Qubit® RNA BR Assay Kit (Invitrogen, Carlsbad, CA, USA) was supplied with

Qubit™ RNA BR Reagent, Standards and Buffer. Working solutions was prepared by adding 1 µl Qubit BR reagent to 190 µl buffers for every sample. Both standards were prepared by adding 10µl to 190 µl of working solution. 2 minutes incubation was given to both standards and instrument was calibrated. Then 2 µl of each sample was added to 198 µl working solution. 2 minutes incubation time was given to all samples and measured.

### 4.10 cDNA Synthesis

For each sample, 500ng of total RNA was used to synthesize cDNAs. RevertAid First Strand cDNA Synthesis Kit (Thermo Fisher Scientific, Waltham, MA, USA) was used for this purpose. For each sample triplicates of 20µl was prepared. Random hexamer primers were used. RNA sample, primer and water were incubated for 5 minutes at 65 $^{o}$C. For RT+ve samples, Master mix was prepared including 5X Reaction Buffer, Ribolock RNase Inhibitor, 10 mMdNTP Mix and Reverse Transcriptase was prepared. For RT-ve controls, separate Master mix was prepared including 5X Reaction Buffer, Ribolock RNase Inhibitor, 10 mMdNTP Mix and water was prepared. To these controls, instead of Reverse Transcriptase, water was added. cDNA synthesis reaction conditions were at 25°C for 10 min, 42°C for 60 min and 70°C for 5 min.

### 4.11 End Point Polymerase Chain Reaction

To make sure that product of cDNAs were free of genomic DNA. Primers used for this reaction were 18S as they consist of one exon only so these can identify genomic DNA contamination. Sequences of forward and reverse primer were 5 'CGTGCATTTATCAGATCAAAACCAACC-3' and 5'-ATGGTAGGCACGGCGACTAC-3' respectively. OneTaq™ Quick Load® Master Mix (New England BioLabs) Nuclease free water and primers were mixed and allocated to separate reaction tubes. To these tubes, RT+ve and RT-ve samples were added. Conditions for this reaction were initial denaturation at 94°C for 5 minutes, followed by 30 cycles of denaturation at 94°C for 30 seconds, annealing at 60°C for 30 seconds and extension at 72°C for 30 seconds. Final extension was done at 72°C for 5 minutes.

### 4.12 Agarose Gel Electrophoresis

To detect genomic DNA contamination and visualize End point Polymerase Chain Reaction products, these were run on Agarose Gel. 1% agarose gel solution was prepared in 1X TAE and heated in microwave so that agarose gets dissolve in TAE completely. To this solution then Ethidium Bromide was added as described in materials section. This solution was then poured into gel tray so that it can solidify and gel can be formed. Then this tray containing gel was placed in 1X TAE filled tank. Samples, 8µl, were loaded to wells after mixing with 2µl of loading dye. Gene ruler (#SM0373 Thermo Scientific Waltham, MA, USA), 50bp DNA ladder, was also added, 3µl to one well. The gel was run at 110volt for 30 minutes

and then visualized. Samples having band in RT +ve samples and no band in RT-ve controls were further used.

### 4.13 Quantitative Real time Polymerase Chain Reaction (qRT-PCR)

Exicycler™ 96 Real-Time Quantitative Thermal Block (Bioneer, Korea) was used for performing qRT-PCR experiments. Both SybrGreen and TaqMan assays were performed for different genes. Every gene was run in triplicate for each sample. Every sample was diluted 1:4. Each TaqMan reaction consisted of 2µl of diluted sample, 10µl TaqMan Universal Master Mix, 1µl TaqMan gene expression primer probes (Applied Biosystems) and 7µl of Nuclease free water. Predesigned TaqMan probes were ordered for 9 genes whose assay IDs, catalog no and probe dyes are given in table 4.1.

**Table 4.1: Predesigned TaqMan probes**

| Gene | Assay ID | Catalog No. | Probe |
|---|---|---|---|
| GAPDH | Hs02758991_g1 | 4453320 | FAM |
| RAB25 | Hs01040784_m1 | 4448892 | FAM |
| GRHL2 | Hs00227745_m1 | 4448892 | FAM |
| SPINT1 | Hs00173678_m1 | 4448892 | FAM |
| PVRL3 | Hs00210043_m1 | 4448892 | FAM |
| FBN1 | Hs00171191_m1 | 4453320 | FAM |
| CLDN7 | Hs00600772_m1 | 4453320 | FAM |
| EMP3 | Hs00171319_m1 | 4453320 | FAM |
| GNG11 | Hs00914578_m1 | 4448892 | FAM |

For TaqMan assay, reaction conditions were holding stage at 50°C for 2 min to block reamplification of carryover PCR products by activating AmpErase UNG enzyme, incubation for inactivation of former mentioned enzyme, followed by 45 cycles of incubation at 95°C for 15 seconds and at 60°C for 1 minute. Reaction was lastly incubated for 1 minute at 60°C. In every PCR plate, for every gene, one template control was run as well.

For cyber green assay, primers were custom designed for specific probesets of genes. Sequence of each probe set was downloaded from Batch Query, NetAffyx Analysis Center, Affymetrix (https://www.affymetrix.com/analysis/netaffx/batch_query.affx?netaffx=netaffx4_annot). Then using these sequences, primers were designed using Primer3 online tool (http://primer3.ut.ee/) and then validated again with NCBI Primer-Blast Tool (http://www.ncbi.nlm.nih.gov/tools/primer-blast/). Each qPCR reaction

contained 10µl of SYBR® Green Real-Time PCR Master Mix (Life Technologies), 0.6µl of both forward and reverse primer and 6.8µl of nuclease free water. For each reaction, first incubation was at 50°C for 2 minutes, then at 95°C for 10 minutes followed by 40 cycles of incubation at 95°C for 15 seconds and 60°C for 1 minute. Melting stage was 60°C to 95°C, Every 0.5°C per 1 second.

**Table 4.2: Sequence and melting temperature of each gene used with SYBR Green assay**

| Gene | Primer | Sequence (5'->3') | Length | Tm (°C) | probe |
|---|---|---|---|---|---|
| ST14 | F | AGAAACCGGCAGAGTACAGC | 20 | 60.04 | ROX |
| | R | TTGATGACGCGGATCTCACC | 20 | 60.18 | ROX |
| BSPRY | F | CAAGGGTTCTGGCAGTGACT | 20 | 59.89 | ROX |
| | R | GGAAGGACACATGATGGGCA | 20 | 60.03 | ROX |
| IRF6 | F | CGTGCACTATGATGCTTGGC | 20 | 59.97 | ROX |
| | R | CCCGACACAGACAGATAGGC | 20 | 59.9 | ROX |
| PVRL3 | F | GTGGAGCAGGTTGGATGGAC | 20 | 60.68 | ROX |
| | R | TGCTAGATCCTCGATGTCAGC | 21 | 59.39 | ROX |
| DDR2 | F | CTTACCTCCCTCAACCAGCC | 20 | 59.75 | ROX |
| | R | GCATGGGTGAGTGGTAGGTC | 20 | 60.11 | ROX |
| BNC2 | F | TCCTTGACTTGAGCACCACC | 20 | 59.89 | ROX |
| | R | ATGATCCCACCATTGCTCCC | 20 | 59.81 | ROX |
| ZNF165 | F | GGGCTGTCCTACTGATCCTG | 20 | 59.24 | ROX |
| | R | GGTTGTCCCCAAGTGTCCTAC | 21 | 60.27 | ROX |
| AP1M2 | F | GAAACAGTCAGTGGCCAACG | 20 | 59.69 | ROX |
| | R | AGTGGGCTCGCATCAAGTAC | 20 | 60.11 | ROX |
| SLIT2 | F | TGACCAACGGACCAATGACC | 20 | 60.25 | ROX |
| | R | CCCATGCTTGCACTTGATCG | 20 | 59.9 | ROX |
| DKK3 | F | AGTTTCCCCTCTGGCTTGAC | 20 | 59.6 | ROX |
| | R | ACTGGTAGAGGCAAAGCAGC | 20 | 60.32 | ROX |

| | | | | | |
|---|---|---|---|---|---|
| TMEM158 | F | ACGTGCCCTAGATTCATGGC | 20 | 60.18 | ROX |
| | R | AAATCCTTCCCATGCCCTCC | 20 | 59.74 | ROX |
| GAPDH | F | TTCTTTTGCGTCGCCAGCCG | 20 | 61.4 | ROX |
| | R | CGACCAAATCCGTTGACTCCGACC | 24 | 66.1 | ROX |
| FN1 | F | TGTGATCCCGTCGACCAATGCC | 22 | 59.23 | ROX |
| | R | TGCCACTCCCCAATGCCACG | 20 | 59.62 | ROX |
| VIM | F | CCAAGACACTATTGGCCGCCTGC | 23 | 60.36 | ROX |
| | R | GCAGAGAAATCCTGCTCTCCTCGC | 24 | 59.42 | ROX |
| CLDN4 | F | ACCTGTCCCCGAGAGAGAGTGC | 22 | 59.4 | ROX |
| | R | GATTCCAAGCGCTGGGGACGG | 21 | 60.11 | ROX |
| E-CAD | F | TGGGCCAGGAAATCACATCCTACA | 24 | 57.57 | ROX |
| | R | TTGGCAGTGTCTCTCCAAATCCGA | 24 | 57.8 | ROX |

GAPDH was used as endogenous reference control for normalization sake. All the data was calculated manually using ΔΔCT Method.

### 4.14 Cell Counting

To determine the approximate number of cells for culturing, haemocytometer was used. Cells were trypsinized and re-suspended as described previously. Haemocytometer and coverslip, both were cleaned gently with 70% ethyl alcohol. Coverslip was then placed gently over chamber area. The cell suspension was gently mixed so that cells should get distributed equally in suspension. Then 10µl of cell suspension was then loaded into each chamber. The cells in each corner square of chamber were counted under light microscope. Average of these values was calculated and multiplied with $10^4$. This is the approximate number of cells per ml of cell suspension.

### 4.15 Drug Cytotoxicity

For Drug cytotoxicity in 2D culture, cell lines were cultured into 96 well plates after counting with haemocytometer in separate plates (1200 cells/well in 90µl). 24 hours after seeding, cells were treated with four different drugs (Panobinostat, Lapatinib, Elesclomol and Midostaurin). Different concentrations

of drugs were prepared using 1% DMSO. 10µl of drugs were added to each well. 72 hours after drug treatment, plates were taken out of incubator and placed at room temperature. Then using CellTiter-Glo® Luminescent Cell Viability Assay (Promega, Madison, WI, United States of America), OD values were taken using The Reporter Microplate Luminometer (Turner Designs). OD results were then used to calculate percentage cell viability by using formula given below.

% cell Viability = (OD value of drug treated well / OD of 1% DMSO treated control well)* 100

These values were then used to draw dose-inhibition response curves using GraphPad software. Concentrations were transformed into log scale and six models were used for this purpose. For each curve model with least error was selected and graphs were drawn. R based program developed by my lab member "Murat İşbilen", was used to calculate IC50, EC50, $A_{MAX}$, Activity Area, IC90 and IC95[44]. 6 models used were 3 parameter, 3 parameter Top 100, 3 parameter Bottom 0, 4 parameter, 4 parameter Top 100 and 4 parameter Bottom 0.

Same procedure was used to calculate cytotoxicity results for 3D culture except cells were cultured in 96 well ultra-low attachment plates.

### 4.16 Calculation of Cytotoxicity Parameters

Optical density (OD) values were obtained from The Reporter Microplate Luminometer (Turner Designs).

**Table 4.3: OD values for HCC1937 treated with Lapatinib. R1, R2, R3 and R4 show replicates.**

| Drug Concentration (µM) | R1 | R2 | R3 | R4 |
|---|---|---|---|---|
| 50 | 646.639 | 658.559 | 1095.47 | 625.725 |
| 10 | 78573 | 107601 | 112638 | 99009 |
| 5 | 328355 | 345084 | 312809 | 376827 |
| 1 | 539296 | 495448 | 523102 | 491441 |
| 0.5 | 607376 | 518335 | 581933 | 508540 |
| 0.1 | 593425 | 544754 | 565369 | 593047 |
| 0.05 | 608690 | 519001 | 601370 | 620710 |
| 0.01 | 536828 | 584786 | 574911 | 612565 |
| 0.005 | 567194 | 559140 | 573009 | 574903 |
| 0.001 | 522998 | 569093 | 622908 | 564012 |
| control | 639313 | 619749 | 662154 | 618879 |

Using % cell viability formula OD values were converted into % cell viability values.

**Table 4.4: % viability values for HCC1937 treated with Lapatinib.** R1, R2, R3 and R4 show replicates.

| Drug Concentration log(µM) | R1 | R2 | R3 | R4 |
|---|---|---|---|---|
| 1.69897 | 0.104013 | 0.105931 | 0.176209 | 0.100649 |
| 1 | 12.63866 | 17.30789 | 18.1181 | 15.92584 |
| 0.69897 | 52.81671 | 55.50762 | 50.3161 | 60.61356 |
| 0 | 86.7471 | 79.69404 | 84.14225 | 79.0495 |
| -0.30103 | 97.69793 | 83.37547 | 93.60536 | 81.79992 |
| -1 | 95.45388 | 87.62503 | 90.941 | 95.39308 |
| -1.30103 | 97.90929 | 83.4826 | 96.73185 | 99.84274 |
| -2 | 86.35011 | 94.06428 | 92.47586 | 98.5326 |
| -2.30103 | 91.23456 | 89.93905 | 92.16992 | 92.47457 |
| -3 | 84.12552 | 91.54002 | 100.1963 | 90.72273 |

6 Graphs using different model were drawn using % cell viability values and model with least error was selected.

**Table 4.5: Standard error of model, log (IC50), log(EC50), Activity area and Amax values for HCC1937 treated with Lapatinib.** Parameter with least standard error was selected (4 parameter bottom 0).

| Model | Sample | Std. Err of Model | log(IC50) | log(EC50) | Activity Area | Amax |
|---|---|---|---|---|---|---|
| 3-Parameter | Lapatinib HCC1937 | 5.938 | 0.646 | 0.823 | 102.393 | 110.452 |
| 3-Parameter Bottom 0 | Lapatinib HCC1937 | 7.067 | 0.608 | 0.656 | 101.929 | 94.685 |
| 3-Parameter Top 100 | Lapatinib HCC1937 | 7.434 | 0.628 | 0.733 | 114.402 | 113.623 |
| 4-Parameter | Lapatinib HCC1937 | 3.696 | 0.726 | 0.763 | 85.514 | 91.731 |
| 4-Parameter Bottom 0 | Lapatinib HCC1937 | 3.435 | 0.727 | 0.760 | 84.961 | 90.833 |
| 4-Parameter Top 100 | Lapatinib HCC1937 | 7.942 | 0.629 | 0.723 | 113.796 | 112.438 |

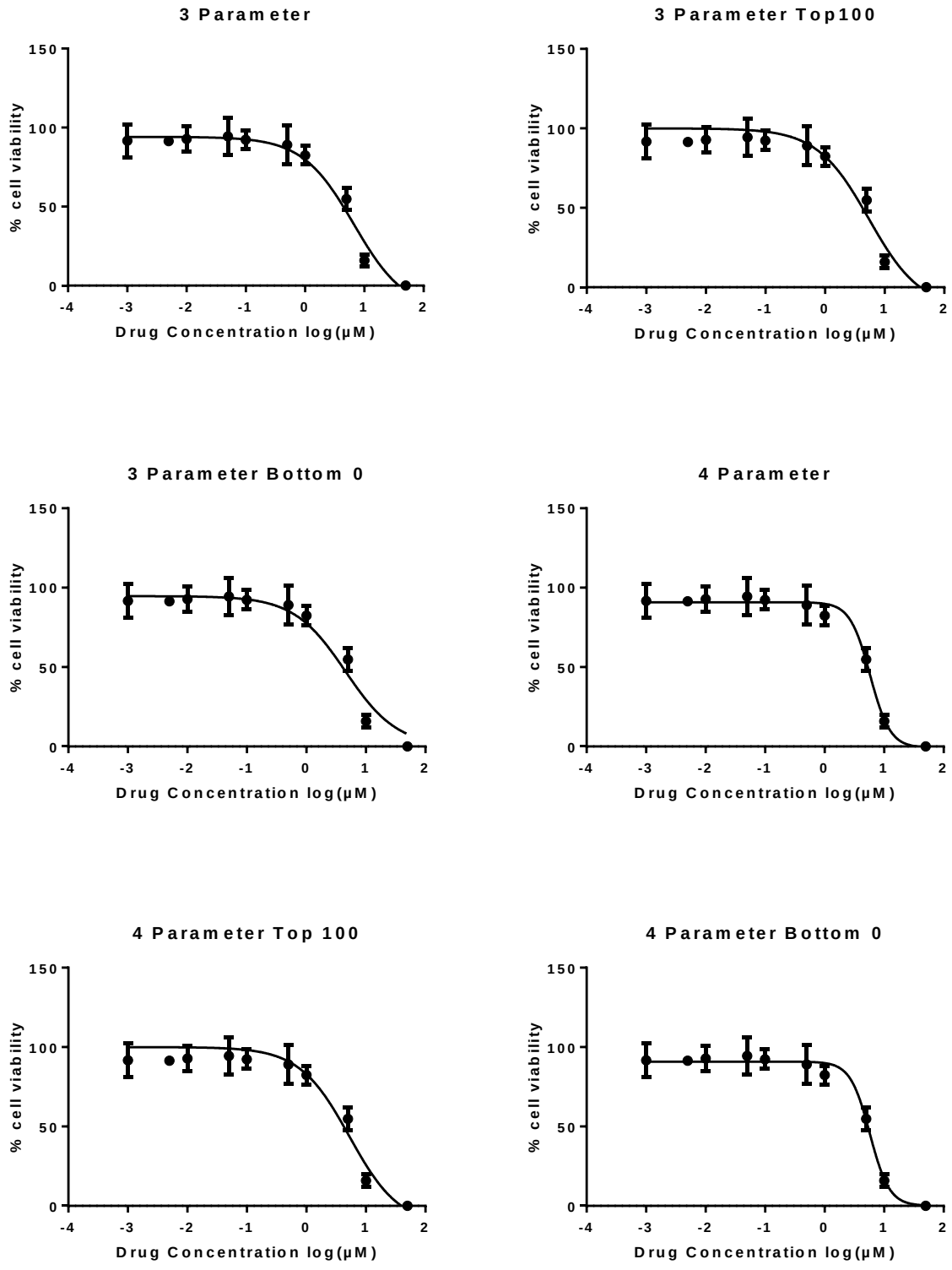



**Figure 4.1: Percent Viability curves for HCC1937 treated with Lapatinib.** Error bars show mean and error with 95% confidence interval. Parameter with least standard error was selected which was 4 parameter bottom 0 for this case.

### 4.17 Software Programs Used in this Study

Microarray data downloaded from genomic data hosting websites, ArrayExpress (http://www.ebi.ac.uk/arrayexpress/) and Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE7513) was RMA normalized using BRB tool[45]. Cluster 3.0 program was used to hierarchically cluster data then heatmaps were generated using Java Treeview. For both genes and samples, Euclidean distances were calculated using complete linkage.

### 4.18 GraphPad Prism 6.0

GraphPad 6s.0 (GraphPad Software Inc., La Jolla, CA, USA) was used to draw graphs for drug cytotoxicity assays and gene expression analysis. Pearson r correlation was used to relate gene expression and cytotoxicity data. And results were plotted using GraphPad 6.0.

### 4.19 Principle Component Analysis

Principle component analysis (PCA) was performed to find combination of genes which can be associated with drug cytotoxicity parameters. R based program[44] developed by my lab member “Murat İşbilen” was used for this analysis. This program performs PCA with expression of genes. For each possible combination of genes, first principle component (PC1) was determined and this was then correlated with drug sensitivity values.

# 5 RESULTS

## 5.1 *In Silico* Results

### 5.1.1 Determination of Gene lists differentiating CSC and Non CSC like breast cancer cell lines

Aim of this study was to find such genes by which breast cancer cell lines can be divided into two groups as CSC and non CSC like. Previously in literature some gene lists have been reported which can differentiate between isolated CSC and Non CSC cells. Among the most recent gene lists, the one reported by Gupta *et al* was chosen. In this study, authors enriched HMLER cells for CSC and non CSC by treating them with Paclitaxel (which targets Non CSC cells and enriches CSC cells) and Salinomycin (which targets CSC cells and enriches non CSC cells) respectively. By microarray technology, they determined the differentially expressed probesets for different genes and reported the most significant ones as gene list. So by using this gene list we decided to divide breast cancer cell lines into two groups as CSC like and non CSC like.

Next the expression dataset for CCLE, Geo accession number GSE36139, was downloaded from GEO. This dataset contained gene expression data for fifty six breast cancer cell lines. The dataset was then normalized. By using Gupta *et al* gene list, CCLE dataset was hierarchically clustered into CSC like and non CSC like groups and distinct clusters were formed. Then to find most differentially expressed genes between CSC like and non CSC like breast cancer cell lines, GSEA analysis was performed with the same dataset. And from the results top 100 upregulated and 100 downregulated genes were selected. Same analysis was repeated with CGP dataset; accession number E-MTAB-783, containing expression data for thirty nine breast cancer cell lines. From CGP's GSEA results again top 100 upregulated and 100 downregulated genes were selected. At that time, we had 200 genes for each dataset. Afterwards for each gene we found all the corresponding probesets. These probesets were ranked using feature selection algorithm which uses Maximum Relevance Minimum Redundancy (MRMR) approach. Aim of this approach was to sort genes according to their separation ability in both datasets.From these ranked lists, common probesets were selected and rank values for those common probesets were summed up. Then these probesets were sorted according to their rank sum values. The resulting gene list was named as Rank Sum Gene list. From rank sum gene list top 4 upreglated and 4 downregulated genes were selected. This gene list was named as Non-Patentable Gene List as we found later that some of genes were previously reported as related with CSC feature and we wanted to patent this gene list. To find the genes which can be patented, we removed all the genes previously associated with CSC feature. Remaining genes were again sorted according to rank sum values and top 6 upregulated and 6 downregulated genes were selected. This gene list was named as Patentable Gene List.

At that point we had two gene lists, one comprising of 8 genes and other of 12 genes. To find that these gene lists were capable of clustering CSC and non CSC like cell lines separately, we performed

hierarchical clustering with CCLE and CGP datasets. Distinct clusters were seen for both gene lists in said datasets.

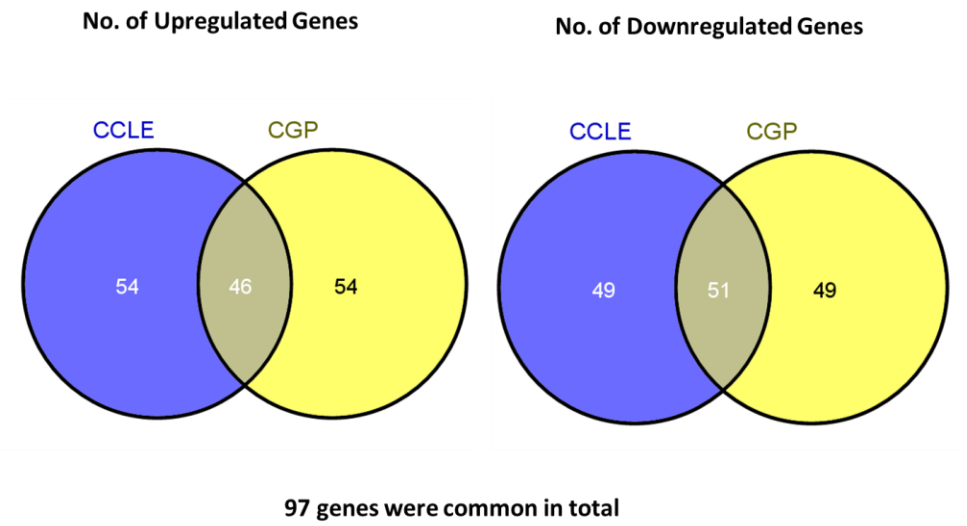



**Figure 5.1: Common genes between genes selected from CCLE and CGP datasets.**

**Table 5.1 : Non-Patentable Gene List**. Downregulated and Upregulated pattern of genes describes CSC like cell lines behavior.

| Gene Name | Gene Symbol | Probeset ID | Rank Sum Score | Behavior in CSC like cells | T Test p-value | Fold change |
|---|---|---|---|---|---|---|
| Ras-related protein Rab-25 | RAB25 | 218186_at | 4 | Downregulated | 3.67E-21 | -5.6 |
| Grainy head-Like 2 (Drosophila) | GRHL2 | 219388_at | 5 | Downregulated | 3.74E-11 | -3.9 |
| Serine Peptidase Inhibitor, Kunitz Type 1 | SPINT1 | 202826_at | 7 | Downregulated | 1.66E-15 | -4.4 |
| Claudin 7 | CLDN7 | 202790_at | 15 | Downregulated | 6.48E-14 | -4 |
| Poliovirus Receptor-Related 3 | PVRL3 | 213325_at | 9 | Upregulated | 1.61E-15 | 2.3 |
| Fibrillin 1 | FBN1 | 202766_s_at | 10 | Upregulated | 1.20E-15 | 4.6 |
| Epithelial membrane protein 3 | EMP3 | 203729_at | 16 | Upregulated | 9.33E-16 | 4.9 |
| Guanine nucleotide binding protein (G protein), gamma 11) | GNG11 | 204115_at | 38 | Upregulated | 8.84E-08 | 3.6 |

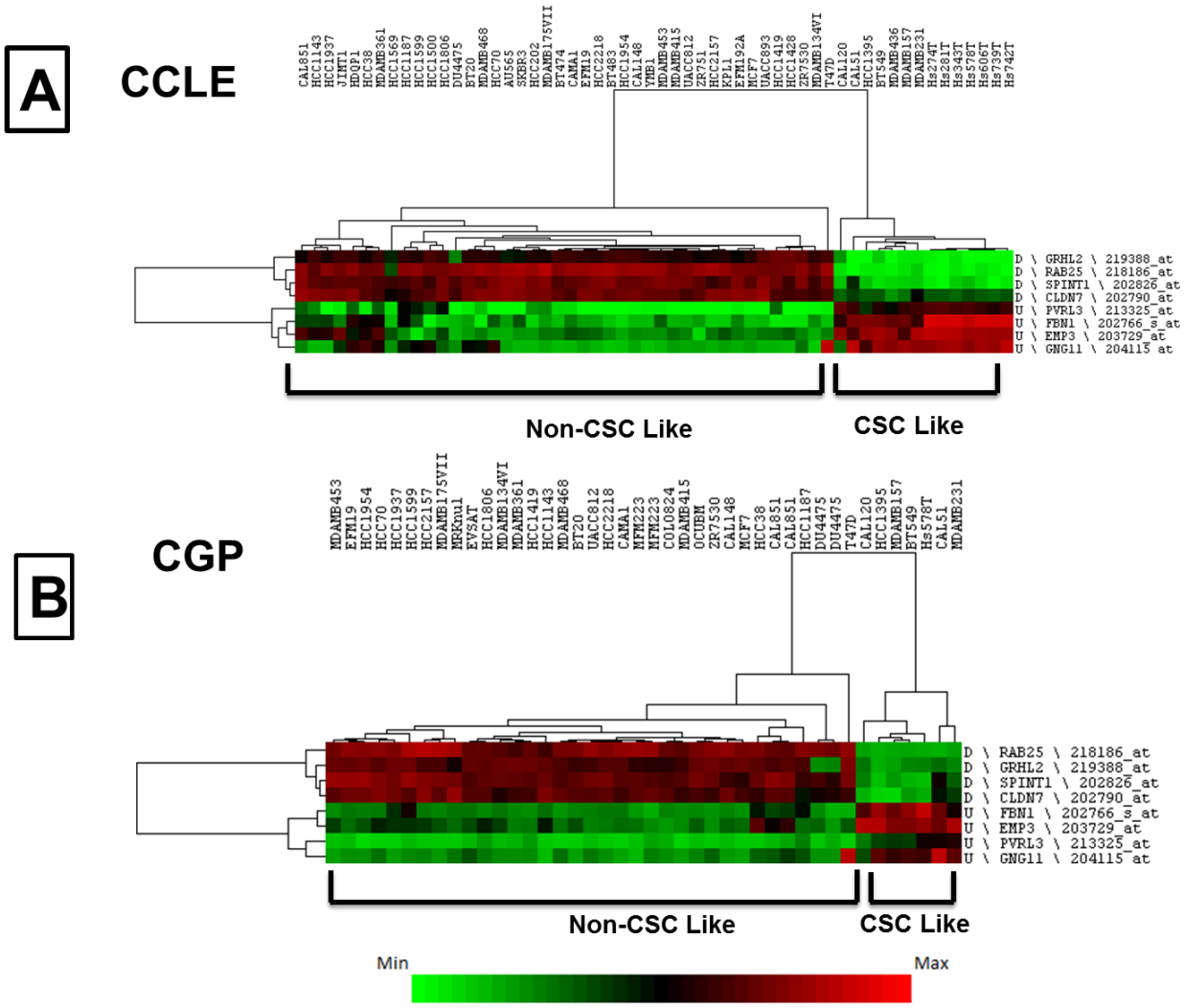



**Figure 5.2: (A) CCLE and (B) CGP Datasets were clustered by using Non-Patentable Gene List.** CSC like breast cancer cell lines are on the right and non CSC like cell lines are on the left. Gene expression datasets were normalized and hierarchical clustering was performed using 8 genes. Probeset IDs are on the left of heat maps and cell line names are on the top. Scale is given below. Red shows maximum value and green shows minimum value.

**Table 5.2: Patentable Gene List**. Downregulated and Upregulated pattern of genes describes CSC like cell lines behavior.

| Gene Name | Gene Symbol | Probeset ID | Rank Sum Score | Behavior in CSC like cells | T Test p-value | Fold change |
|---|---|---|---|---|---|---|
| Interferon Regulatory Factor 6 | IRF6 | 202597_at | 23 | Downregulated | 1.43E-19 | -3.5 |
| Suppression Of Tumorigenicity 14 | ST14 | 202005_at | 26 | Downregulated | 6.70E-19 | -4.22 |
| B-Box And SPRY Domain Containing protein | BSPRY | 218792_s_at | 47 | Downregulated | 4.02E-20 | -4.2 |
| Lethal Giant Larvae Homolog 2 | LLGL2 | 203713_s_at | 74 | Downregulated | 3.43E-14 | -3.37 |
| Adaptor-related Protein Complex 1, mu 2 Subunit | AP1M2 | 65517_at | 78 | Downregulated | 1.11E-13 | -3.37 |
| Zinc finger protein 165 (CT gene) | ZNF165 | 206683_at | 84 | Downregulated | 3.39E-12 | -3.05 |
| Poliovirus Receptor-Related 3 | PVRL3 | 213325_at | 9 | Upregulated | 2.11E-18 | 4.38 |
| Slit Homolog 2 | SLIT2 | 209897_s_at | 48 | Upregulated | 7.43E-11 | 3.63 |
| Basonuclin 2 | BNC2 | 220272_at | 50 | Upregulated | 2.56E-15 | 2.92 |
| Discoidin Domain Receptor Tyrosine | DDR2 | 205168_at | 77 | Upregulated | 8.13E-17 | 4.12 |
| Transmembrane Protein 158 | TMEM158 | 213338_at | 103 | Upregulated | 1.21E-11 | 4.16 |
| Dickkopf WNT Signaling Pathway Inhibitor 3 | DKK3 | 202196_s_at | 141 | Upregulated | 6.59E-13 | 4.42 |

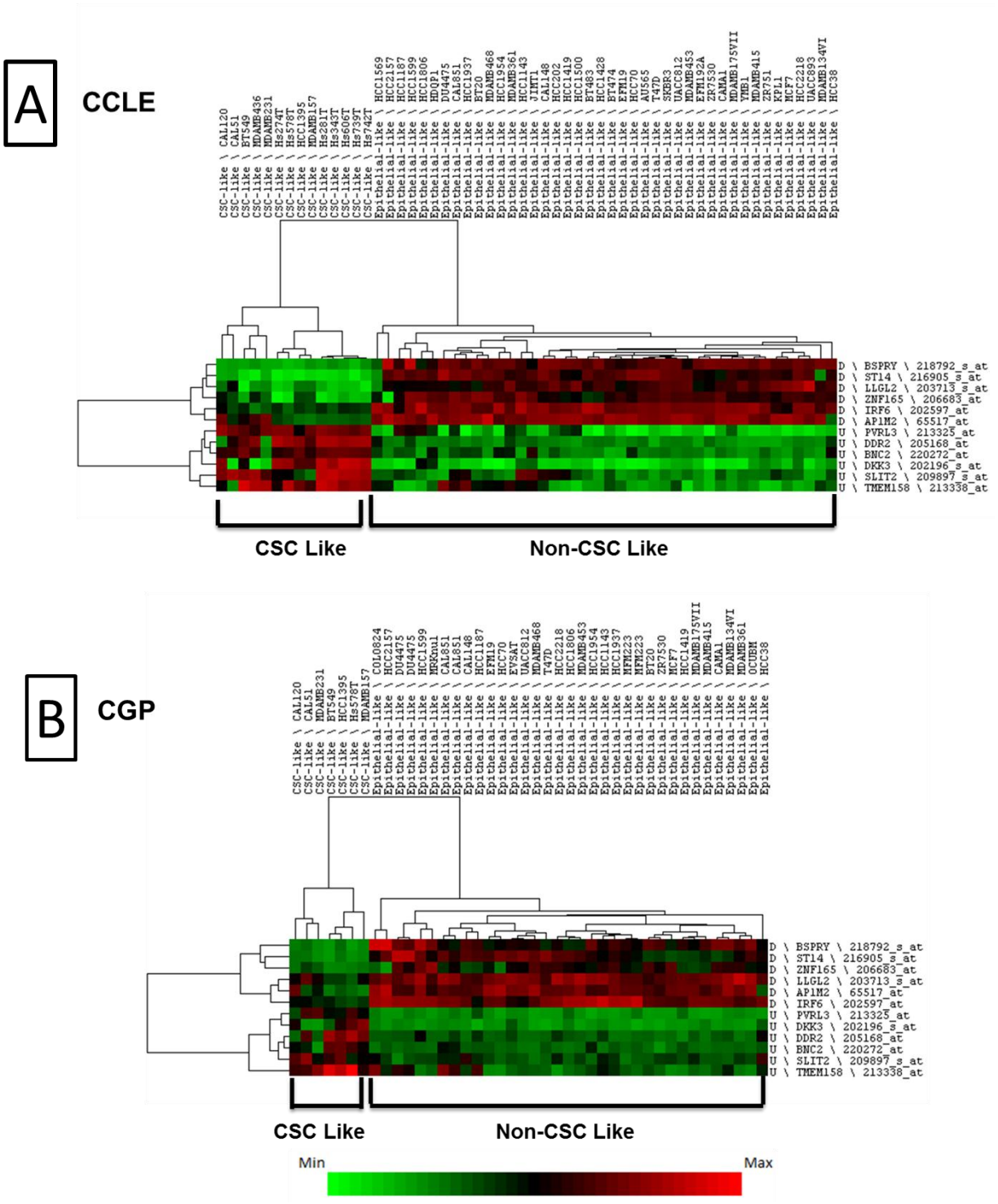



**Figure 5.3: (A) CCLE and (B) CGP Datasets were clustered by using Patentable Gene List.** CSC like breast cancer cell lines are on the left and non CSC like cell lines are on the right. Gene expression datasets were normalized and hierarchical clustering was performed using 12 genes. Probeset IDs are on the left of heatmaps and cell line names are on the top. Scale is given below. Red shows maximum value and green shows minimum value.

In both gene lists four genes were common. BNC2, PVRL3 and SLIT2 were common among upregulated genes and only IRF6 was common in downregulated genes.

### 5.1.2 Validation of Gene Lists

Both gene lists were validated by using a recently published dataset published by Bhat-Nakshatri, P. et al (Geo accession number GSE15192) . In this study authors overexpressed SLUG in MCF10A cell line by transfection. These cells expressed the behavior of stem cells as compared to control cells. This dataset was normalized and then clustered using both gene lists. In both scenarios, CSC cells were clustered separately from non CSC cells.

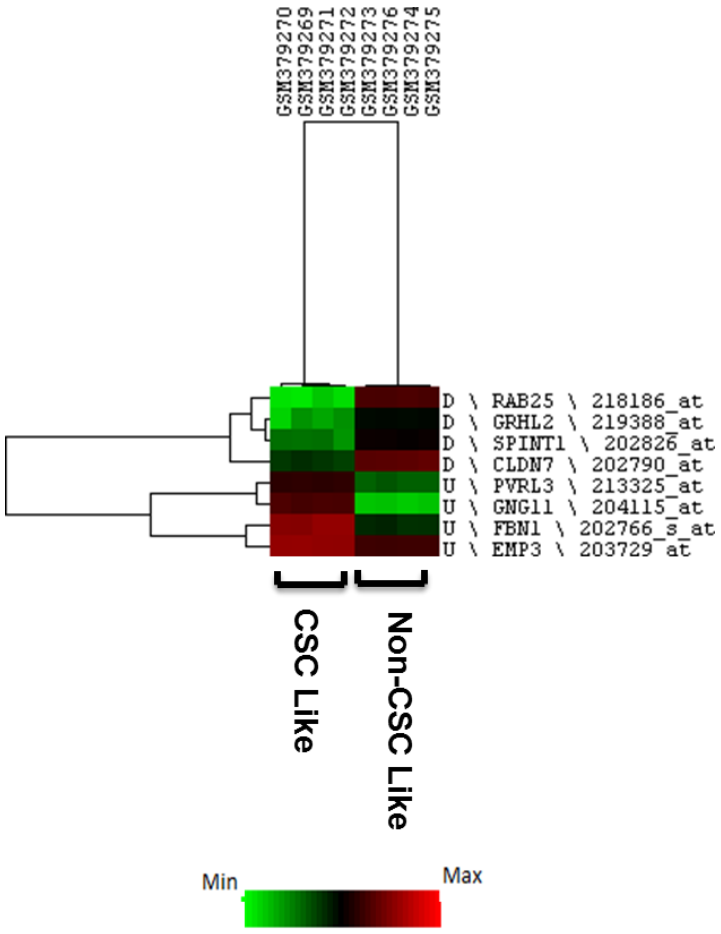



**Figure 5.4: Validation with Bhat-Nakshatri et al dataset with Non-Patentable Gene List.** CSC like breast cancer cell lines (transfected with SLUG) are on the left and non CSC like cell lines (control) are on the left. Gene expression datasets were normalized and hierarchical clustering was performed using 12 genes. Probeset IDs are on the left of heatmaps and Sample names are on the top. Scale is given below. Red shows maximum value and green shows minimum value.

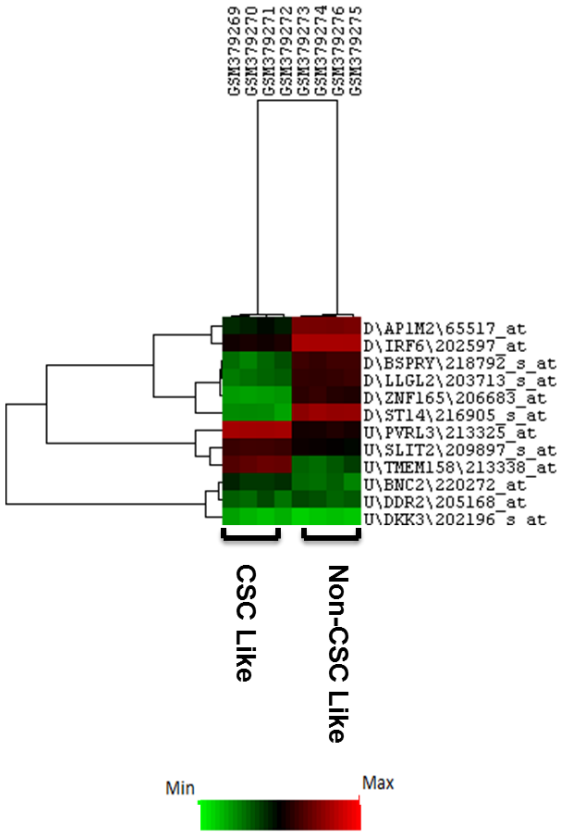



**Figure 5.5: Validation with Bhat-Nakshatri et al dataset with Patentable Gene List.** CSC like breast cancer cell lines (transfected with SLUG) are on the left and non CSC like cell lines (control) are on the right. Gene expression datasets were normalized and hierarchical clustering was performed using 12 genes. Probeset IDs are on the left of heatmaps and Sample names are on the top. Scale is given below. Red shows maximum value and green shows minimum value.

Genes from both lists were further validated in eight different datasets. Details of these datasets are given in table 5.3. Datasets were downloaded and normalized. GSEA was used to analyze each dataset. From results top 1000 downregulated and top 1000 upregulated genes were selected. Then behavior of our selected genes was determined in those datasets. If gene was present, it was given 1 score and if it was absent then it was given 0 score. All scores were summed up.

**Table 5.3: Details of the datasets selected for further validation.**

| Dataset number | Author | Type | Description of comparison groups | Accession number |
|---|---|---|---|---|
| D1 | Suling Liu et al[46] | Tumor | ALDH+ Tumor sorted cells versus ALDH- cells | GSE52327 |
| D2 | Chad J. Creighton et al[13] | Tumor | Sorted CSC cells versus Non CSC cells | GSE7513 |
| D3 | Chad J. Creighton et al[13] | tumor | Mammosphere formed from tumor cells (CSC) versus Primary tumor (non CSC) | GSE7515 |
| D4 | Chad J. Creighton et al[13] | Tumor | Letrozole and Docetaxel treated versus non treated patients cells | GSE10281 |
| D5 | Gupta et al[47] | Cell line | HMLER breast cancer cells treated with Salinomycin (non CSC) or Paclitaxel (CSC) | GSE17215 |
| D6 | Salvatore Pece et al[19] | Cell line | hNMSC sorted into PKH+ (CSC) vs. PKH- (Non-CSC) | GSE18931 |
| D7 | Anna Maria Calcagno et al[48] | Cell line | MCF7 Doxorubicin treated (CSC) vs. Non treated (Non CSC) | GSE24460 |
| D8 | Venkata Lokesh Battula et al[17] | Cell line | Breast cancer cell lines were clustered based upon gene expression | GSE24717 |

**Table 5.4: Comparison of selected genes behavior among different datasets (Downregulated in CSC like cells)**

| (Downregulated in CSC like cells) | Rank Sum Score | D1 | D2 | D3 | D4 | D5 | D6 | D7 | D8 | Score Sum |
|---|---|---|---|---|---|---|---|---|---|---|
| RAB25 | 4 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 6 |
| GRHL2 | 5 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 5 |
| SPINT1 | 7 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 4 |
| CLDN7 | 15 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 5 |
| IRF6 | 23 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 4 |
| ST14 | 26 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 3 |
| BSPRY | 47 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 6 |
| LLGL2 | 74 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 4 |
| AP1M2 | 78 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 4 |
| ZNF165 | 84 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 6 |

**Table 5.5: Comparison of selected genes behavior among different datasets (Upregulated in CSC like cells).**

| (Upregulated in CSC like cells) | Rank Sum Score | D1 | D2 | D3 | D4 | D5 | D6 | D7 | D8 | Score Sum |
|---|---|---|---|---|---|---|---|---|---|---|
| PVRL3 | 9 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 4 |
| FBN1 | 10 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 5 |
| EMP3 | 16 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 6 |
| GNG11 | 38 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 4 |
| SLIT2 | 48 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 4 |
| BNC2 | 50 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 2 |
| DDR2 | 77 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 4 |
| TMEM158 | 103 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 4 |
| DKK3 | 141 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 3 |

### 5.1.3 Molecular Subtyping of Breast Cancer Cell Lines

Our group has previously published a paper in which molecular subtyping of breast cancer cell lines, present in CCLE dataset, was performed using previously published gene list by Kao et al. Then breast cancer cell lines molecular subtypes (Luminal, Basal A and B) were compared with CSC like characterization for both CCLE and CGP datasets. It was found that all CSC like cell lines were Basal B except one which was Basal A.

**Table 5.6: Comparison of Molecular subtypes of breast cancer cell lines with CSC like characterization for both CCLE and CGP datasets.**

| CCLE | Luminal | Basal A | Basal B |
|---|---|---|---|
| **CSC-like** | 0 | 1 | 13 |
| **Non CSC-like** | 26 | 16 | 0 |

| CGP | Luminal | Basal A | Basal B |
|---|---|---|---|
| **CSC-like** | 0 | 0 | 7 |
| **Non CSC-like** | 18 | 14 | 0 |

### 5.1.4 Prediction of Drug Response

Using CCLE and CGP drug sensitivity data, drugs were predicted to target all molecular subtypes of breast cancer cell lines. For this purpose, drug sensitivity values (Activity area or IC50) of all molecular subtypes against different drugs, were compared with each other in set of two (Basal A vs. Basal B, Basal A vs. Luminal and Basal B vs. Luminal). From these comparisons, four most effective drugs based upon p-values were selected for different molecular subtypes. Out of these four drugs selected, Panobinostat and Lapatinib target Basal A and Luminal subtype (non CSC like) while Elesclomol and Midostaurin target Basal B subtype (CSC like).

**Table 5.7: Selected drugs, their mode of action and molecular subtype of breast cancer cell lines targeted.**

| Drug Name | Mode of Action | Effective for subtype of breast cancer |
|---|---|---|
| Elesclomol | Reactive Oxygen Species | Basal B |
| Midostaurin | Tyrosine Kinase Inhibitor | Basal B |
| Panobinostat | Histone Deacetylase Inhibitor | Luminal |
| Lapatinib | EGFR/ERBB2 | Basal A / Luminal |

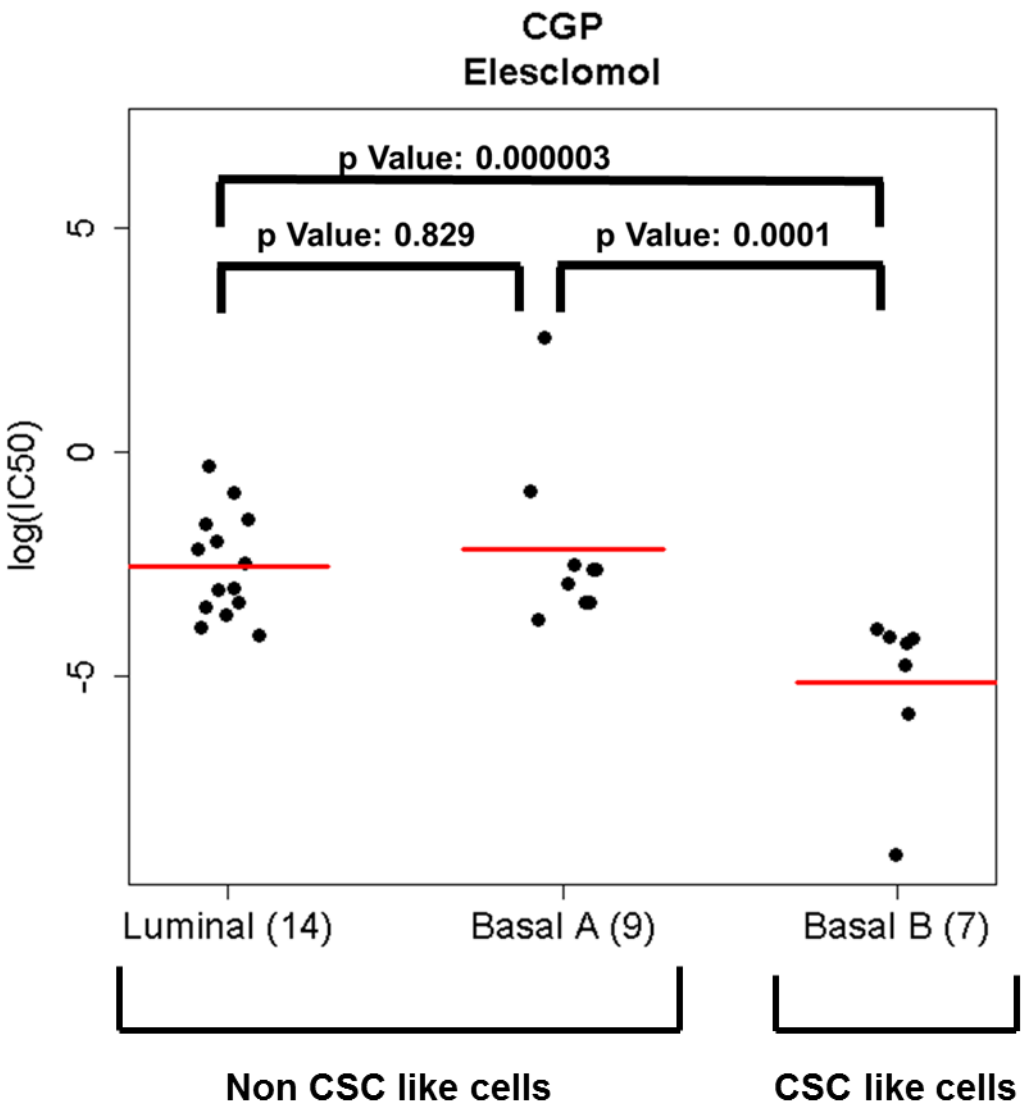



**Figure 5.6: Comparison of Elesclomol cytotoxicity against Breast cancer cell lines.** Breast cancer subtypes are shown on x axis and number next to each subtype shows number of breast cancer cell lines for that type. Y-axis shows log(IC50). Drug sensitivity increases with decrease in log(IC50). For each comparison p-values are mentioned.

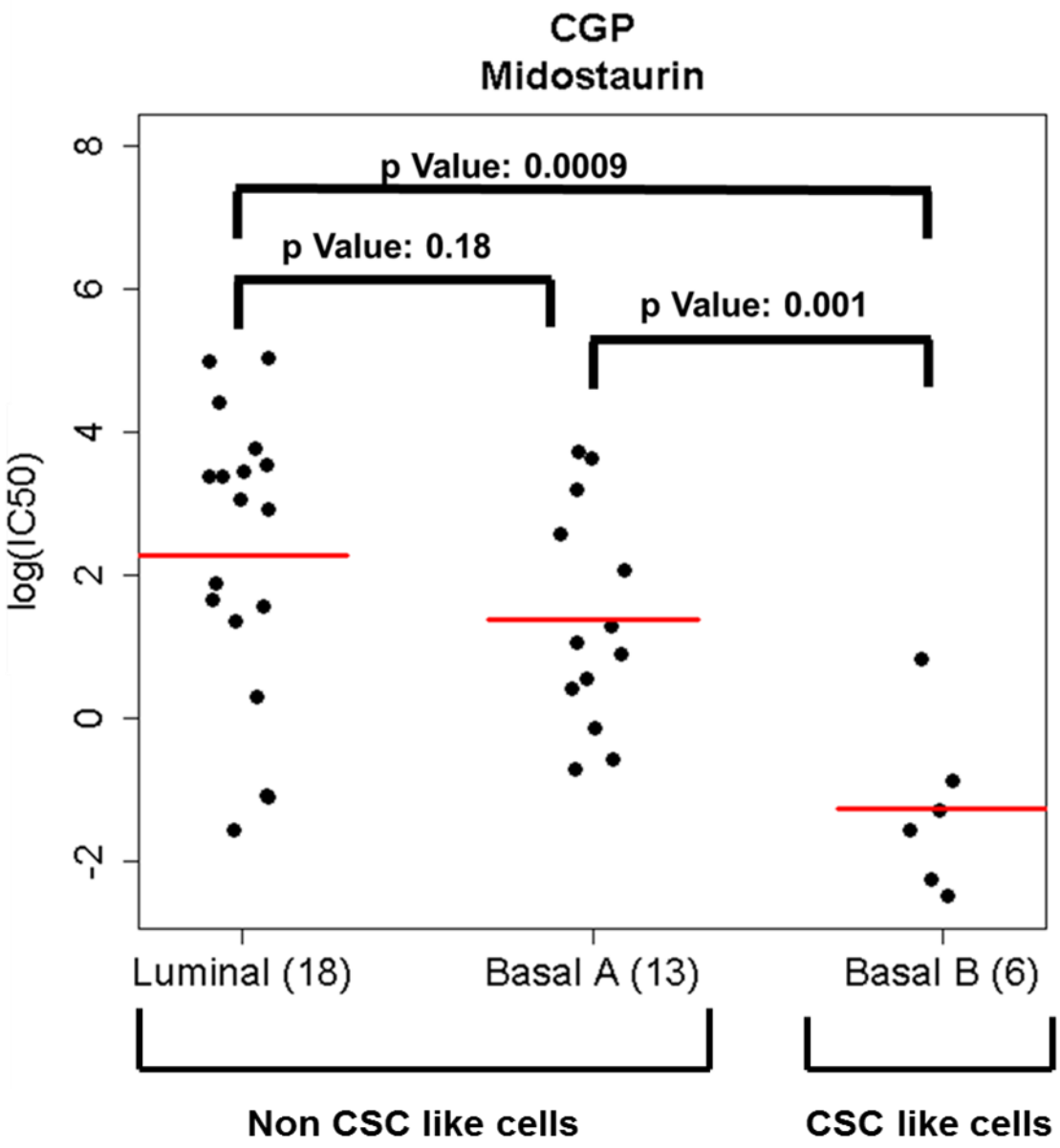



**Figure 5.7: Comparison of Midostaurin cytotoxicity against Breast cancer cell lines.** Breast cancer subtypes are shown on x axis and number next to each subtype shows number of breast cancer cell lines for that type. Y-axis shows log(IC50). Drug sensitivity increases with decrease in log(IC50). For each comparison p-values are mentioned.

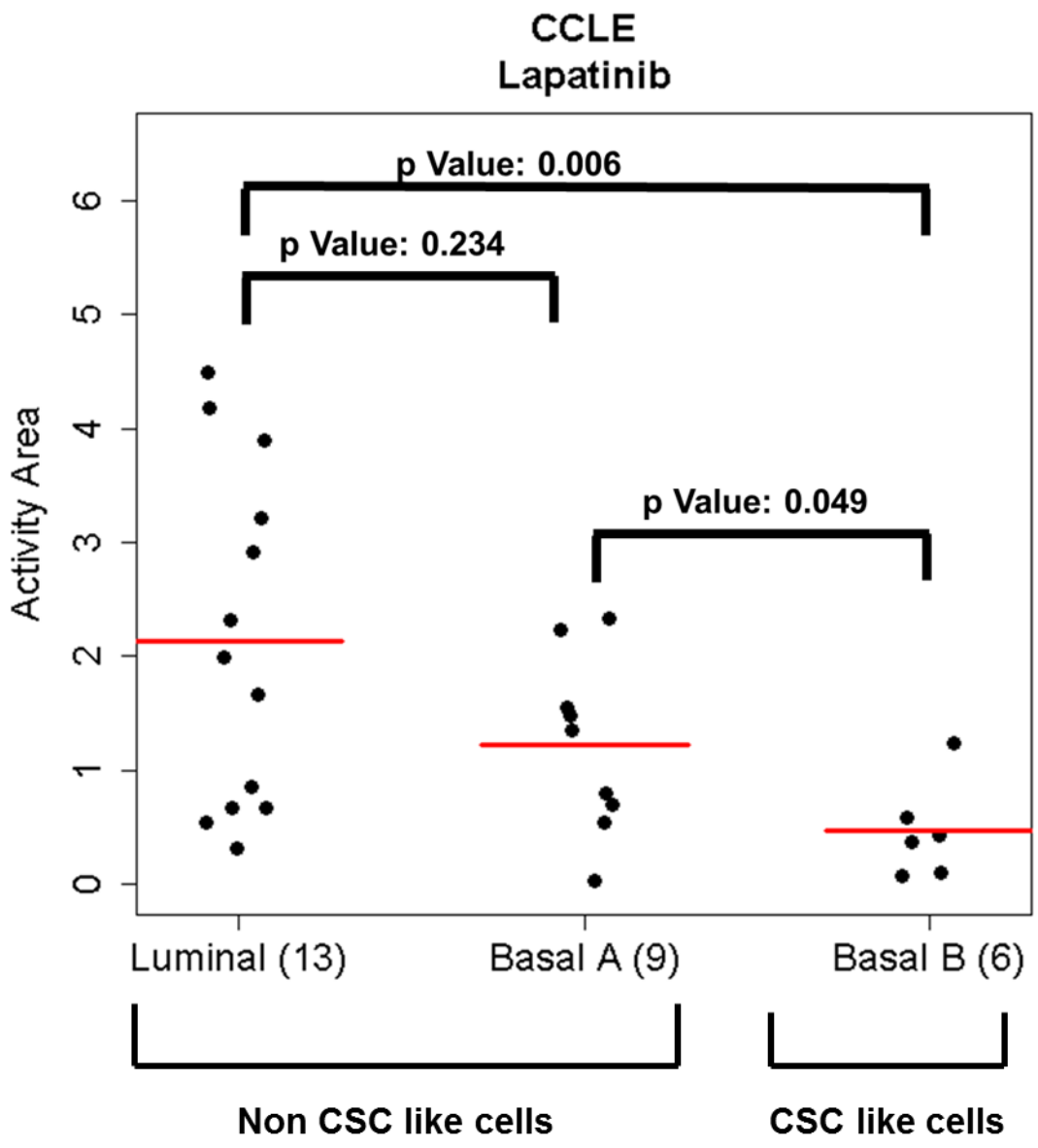



**FIGURE 5.8: Comparison of Lapatinib cytotoxicity against breast cancer cell lines.**

Breast cancer subtypes are shown on x axis and number next to each subtype shows number of breast cancer cell lines for that type. Y-axis shows Activity area. Drug sensitivity increases with increase in Activity area. For each comparison p-values are mentioned.

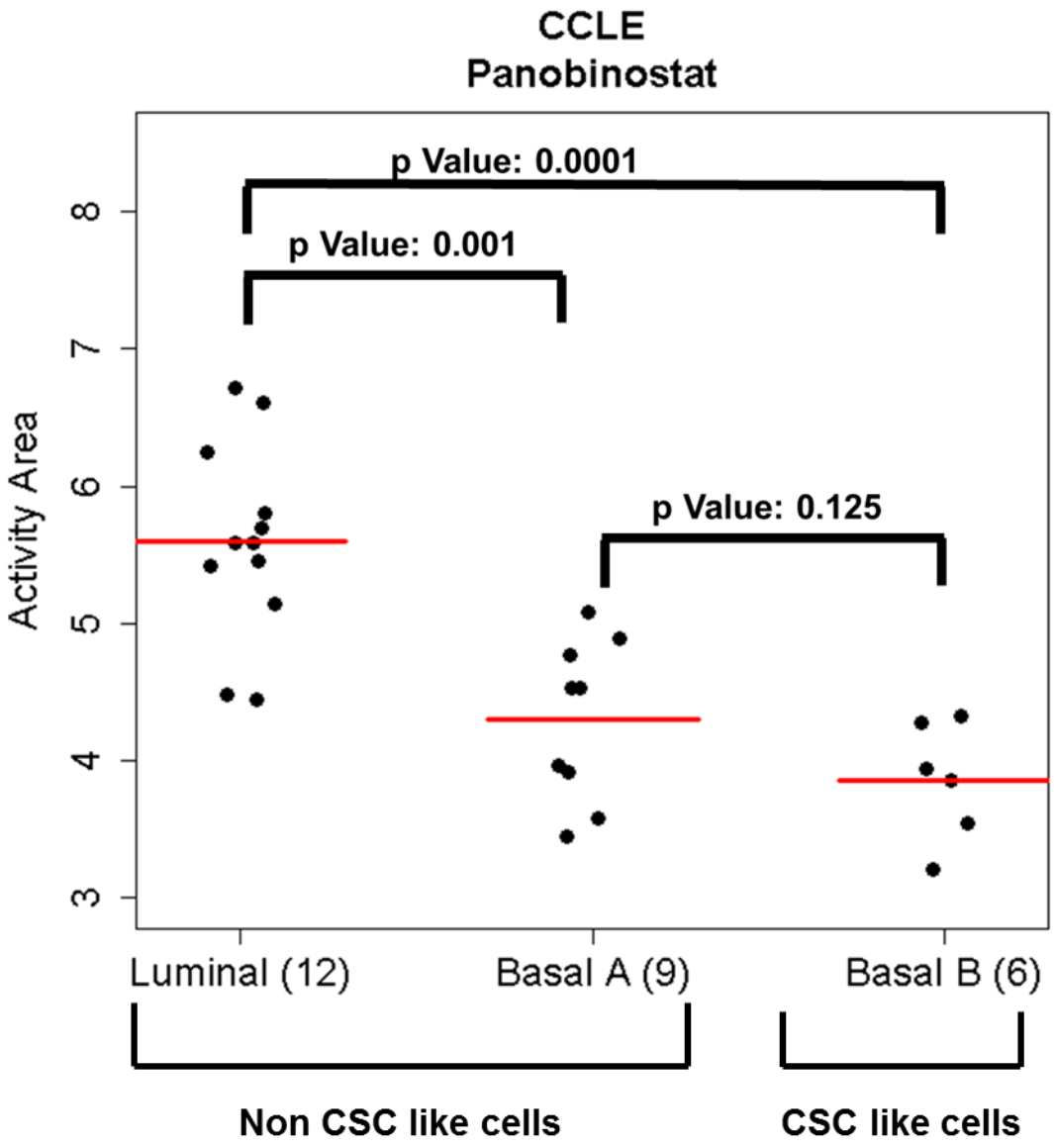



**Figure 5.9: Comparison of Panobinostat cytotoxicity against Breast cancer cell lines.** Breast cancer subtypes are shown on x axis and number next to each subtype shows number of breast cancer cell lines for that type. Y-axis shows Activity area. Drug sensitivity increases with increase in Activity area. For each comparison p-values are mentioned.

### 5.2 *In vitro* Validation

I validation was done with both 2D and 3D cell culture conditions.

### 5.2.1 Validation in 2D Cell Culture Conditions

Cells were grown in 2D culturing conditions to validate both gene lists (Non Patentable and Patentable) and Drugs predicted by *in silico* analysis.

#### 5.2.1.1 Validation of Non-Patentable Gene List by qPCR

Non-Patentable gene list comprised of 4 upregulated and 4 downregulated genes. TaqMan technology was used for qPCR as described in Materials and Methods section.

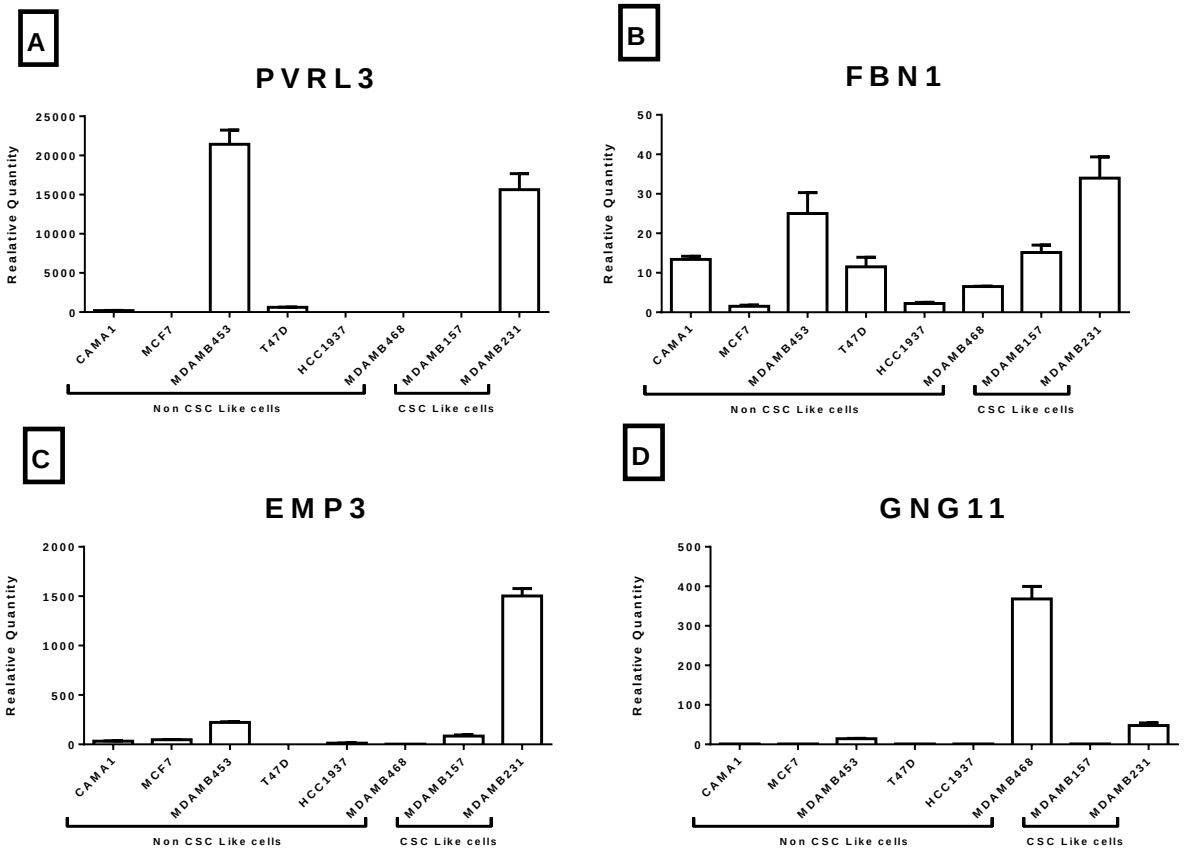



**Figure 5.10: Comparison of expression of (A) PVRL3, (B) FBN1, (C) EMP3 and (D) GNG11 across breast cancer cell lines.** These genes were selected as upregulated ones in CSC like cell lines. Error bars show standard error of mean.

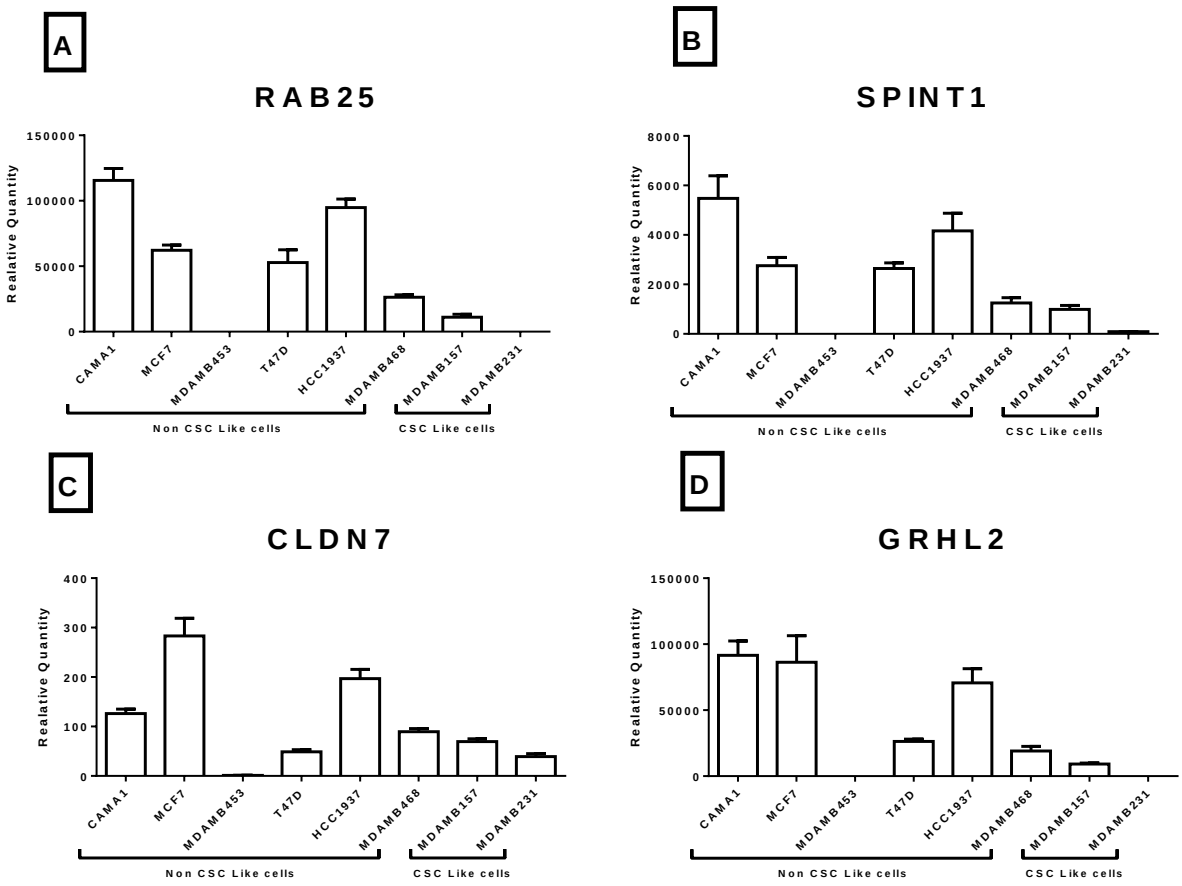



**Figure 5.11: Comparison of expression of (A) RAB25, (B) SPINT1, (C) CLDN7 and (D) GRHL2 across breast cancer cell lines.** These genes were selected as downregulated ones in CSC like cell lines. Error bars show standard error of mean.

CAMA-1, MCF7, MDA-MB-453, T47D are luminal, HCC1937 and MDA-MB-468 are Basal A while MDA-MB-157 and MDA-MB-231 are Basal B. Luminal and Basal A show non CSC like behavior and Basal B show CSC like behavior. From the comparison of the qPCR data, expression of PVRL3, FBN1, EMP3 and GNG11 was upregulated and expression of RAB25, SPINT1, CLDN7 and GRHL2 was downregulated in CSC like cells as compared to non CSC like cells as predicted with exception of MDA-MB-453.

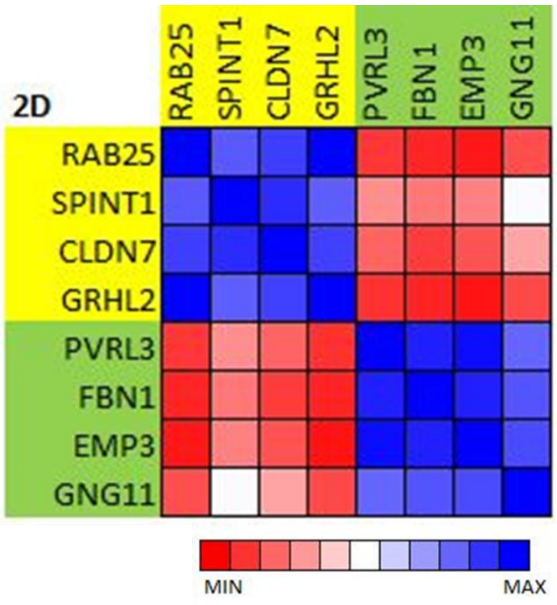



**Figure 5.12: Correlation of genes of Non Patentable genes with each other.** Gene names highlighted with yellow color are downregulated and with green are upregulated in CSC like cells.

When genes of Non Patentable gene list were correlated with each other using Pearson r, correlated pattern among them was seen. But when two sided T.Test was used to determine the capability of genes to separate CSC like from non CSC like, none of the genes was found significant.

**Table 5.8: Capability of Non Patentable genes to separate CSC like from non CSC like.** Genes are sorted according to p-values.

| Gene name | T Test p-value | Fold Change |
|---|---|---|
| EMP3 | 0.1 | 4.54 |
| GRHL2 | 0.15 | -7.26 |
| RAB25 | 0.16 | -7.38 |
| FBN1 | 0.19 | 1.78 |
| GNG11 | 0.24 | 2.2 |
| PVRL3 | 0.67 | 2.08 |
| SPINT1 | 0.69 | -1.44 |
| CLDN7 | 0.94 | -0.17 |

Then concordance of gene expression was calculated with microarray data from CCLE and it was low.

**Table 5.9: Concordance between microarray and qPCR gene expression data for Non Patentable gene list.**

| Gene Name | Pearson r | P-value |
|---|---|---|
| RAB25 | 0.03 | 0.927 |
| SPINT1 | 0.03 | 0.940 |
| CLDN7 | 0.11 | 0.786 |
| GRHL2 | 0.08 | 0.832 |
| PVRL3 | 0.67 | 0.063 |
| FBN1 | 0.45 | 0.253 |
| EMP3 | 0.61 | 0.107 |
| GNG11 | NA | NA |

So for Non Patentable gene, none of the genes was found significant in CSC like cells as compared to non CSC like cells.

#### 5.2.1.2 Validation of Patentable Gene List by qPCR

Patentable gene list comprised of 6 upregulated and 6 downregulated genes. All genes from this list were selected for qPCR validation except LLGL2. All the primers for this analysis were custom designed. The reason for not choosing LLGL2 was that primers could not be designed which hit the same transcript as microarray probe. Four other genes were also selected as established EMT markers which were Fibronectin 1 (FN1), Vimentin (VIM), E cadherin (E-CAD) and Claudin 4 (CLDN4). These genes were also present in Rank Sum Gene List but their rank values were lower as compared to other selected genes. SybrGreen technology was used for this validation analysis.

Gene expression of 11 selected genes and 4 EMT markers was compared between CSC and non CSC like cells.

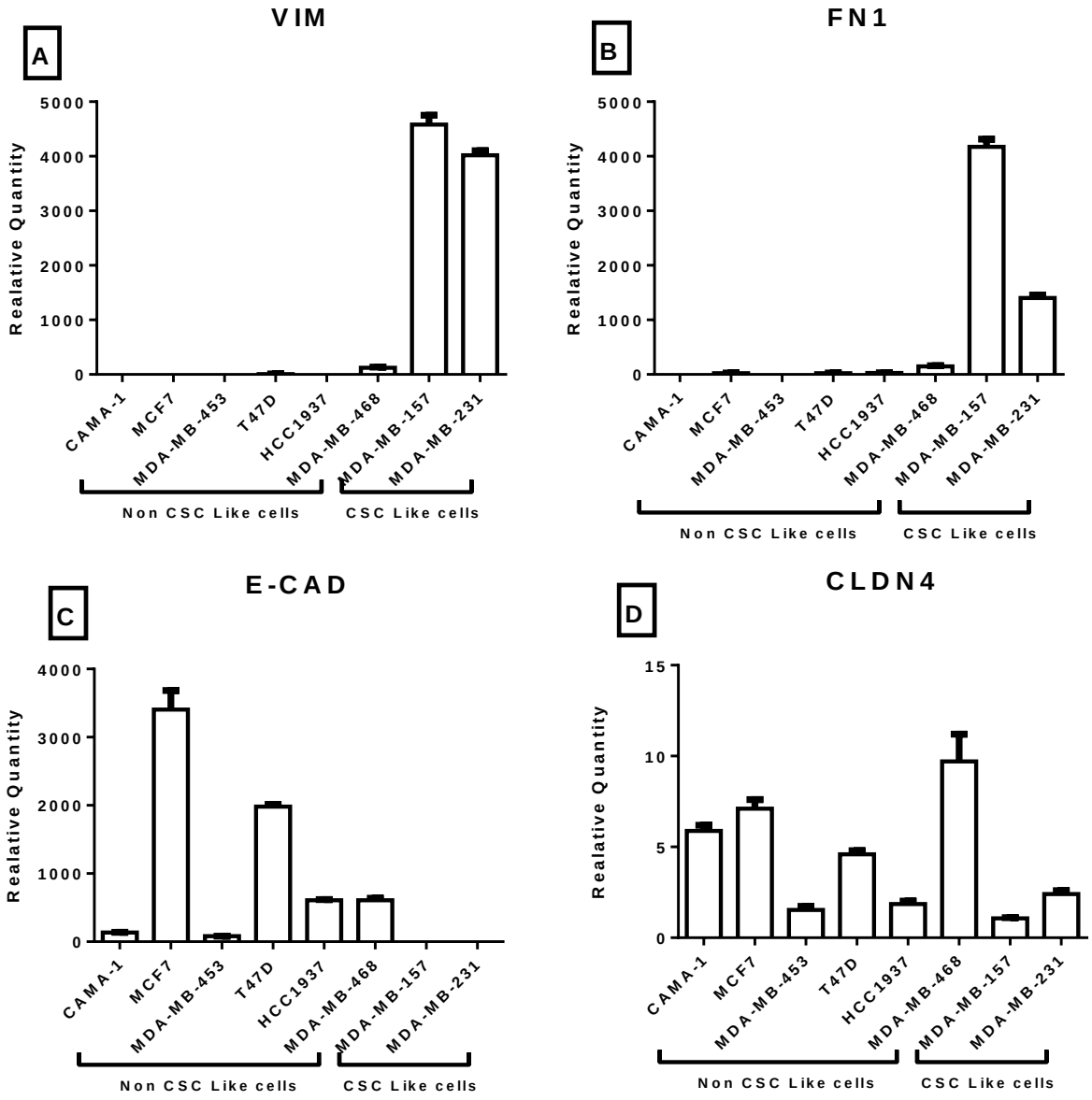



**Figure 5.13: Comparison of (A) VIM, (B) FN1, (C) E-CAD and (D) CLDN4 between breast cancer cell lines.** Error bars show standard error of mean.

MDA-MB-157 and MDA-MB-231 showed mesenchymal behavior while other cell lines showed epithelial behavior. Cell lines showing mesenchymal behavior are considered as CCS like cell lines and the ones showing epithelial behavior are considered as non CSC like cell lines.

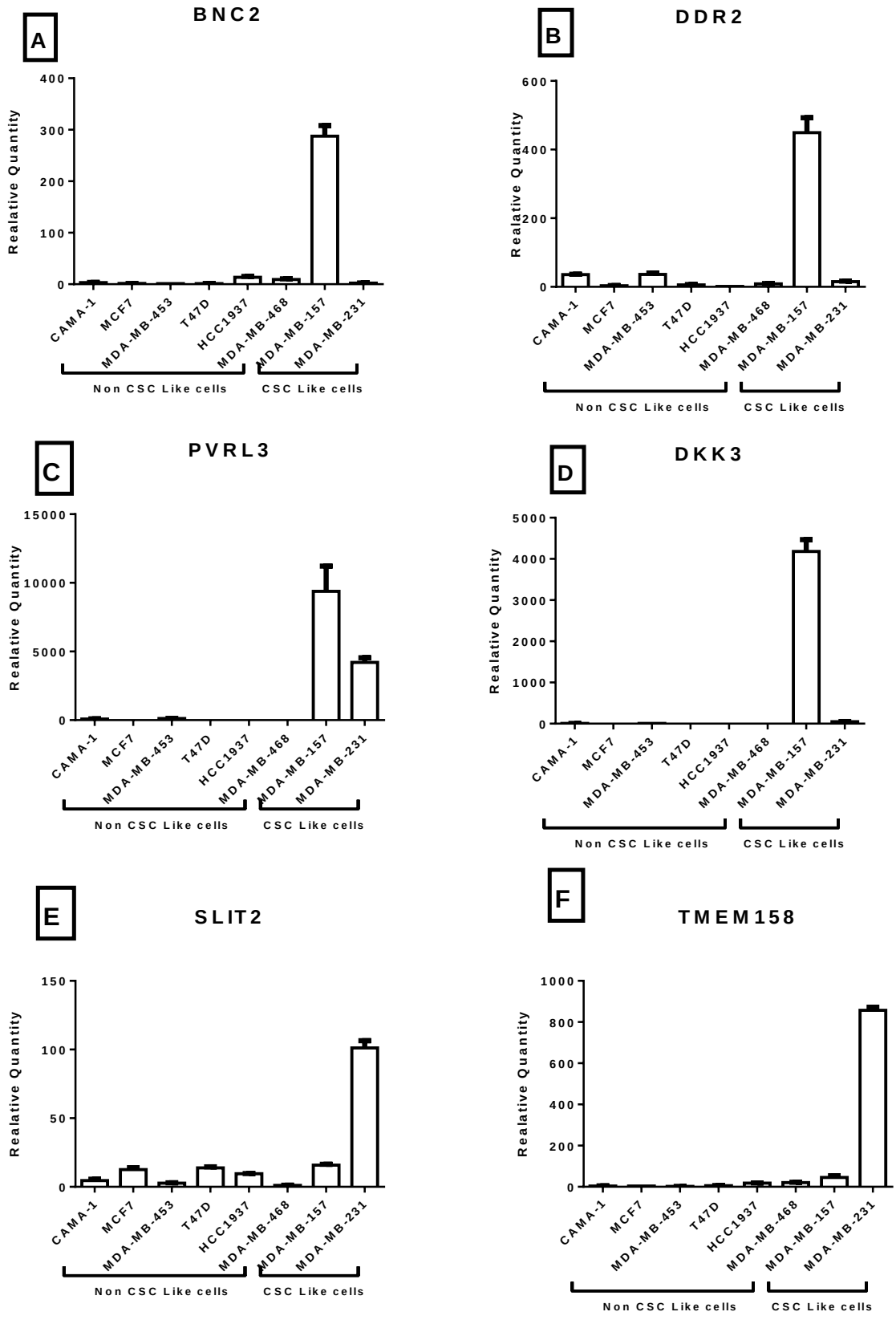

A
BNC2
Realative Quantity
400
300
200
100
0
CAMA-1
MCF7
MDA-MB-453
T47D
HCC1937
MDA-MB-468
MDA-MB-157
MDA-MB-231
Non CSC Like cells
CSC Like cells
B
DDR2
Realative Quantity
600
400
200
0
CAMA-1
MCF7
MDA-MB-453
T47D
HCC1937
MDA-MB-468
MDA-MB-157
MDA-MB-231
Non CSC Like cells
CSC Like cells
C
PVRL3
Realative Quantity
15000
10000
5000
0
CAMA-1
MCF7
MDA-MB-453
T47D
HCC1937
MDA-MB-468
MDA-MB-157
MDA-MB-231
Non CSC Like cells
CSC Like cells
D
DKK3
Realative Quantity
5000
4000
3000
2000
1000
0
CAMA-1
MCF7
MDA-MB-453
T47D
HCC1937
MDA-MB-468
MDA-MB-157
MDA-MB-231
Non CSC Like cells
CSC Like cells
E
SLIT2
Realative Quantity
150
100
50
0
CAMA-1
MCF7
MDA-MB-453
T47D
HCC1937
MDA-MB-468
MDA-MB-157
MDA-MB-231
Non CSC Like cells
CSC Like cells
F
TMEM158
Realative Quantity
1000
800
600
400
200
0
CAMA-1
MCF7
MDA-MB-453
T47D
HCC1937
MDA-MB-468
MDA-MB-157
MDA-MB-231
Non CSC Like cells
CSC Like cells

**Figure 5.14: Comparison of (A) BNC2, (B) DDR2, (C) PVRL3, (D) DKK3, (E) SLIT2 and (F) TMEM158 in breast cancer cell lines.** These genes were selected as upregulated ones in CSC like cell lines. Error bars show standard error of mean.

BNC2, DDR2, PVRL3, DKK3, SLIT2 and TMEM158 are genes which were found upregulated in CSC like cells and it was validated as well. MDA-MB-157 and MDA-MB-231 (CSC like cell lines) showed upregulation of these genes while other cell lines, CAMA-1, MCF7, MDA-MB-453, T47D, HCC1937 and MDA-MB-468 (non CSC like cell lines) showed downregulation of these genes.

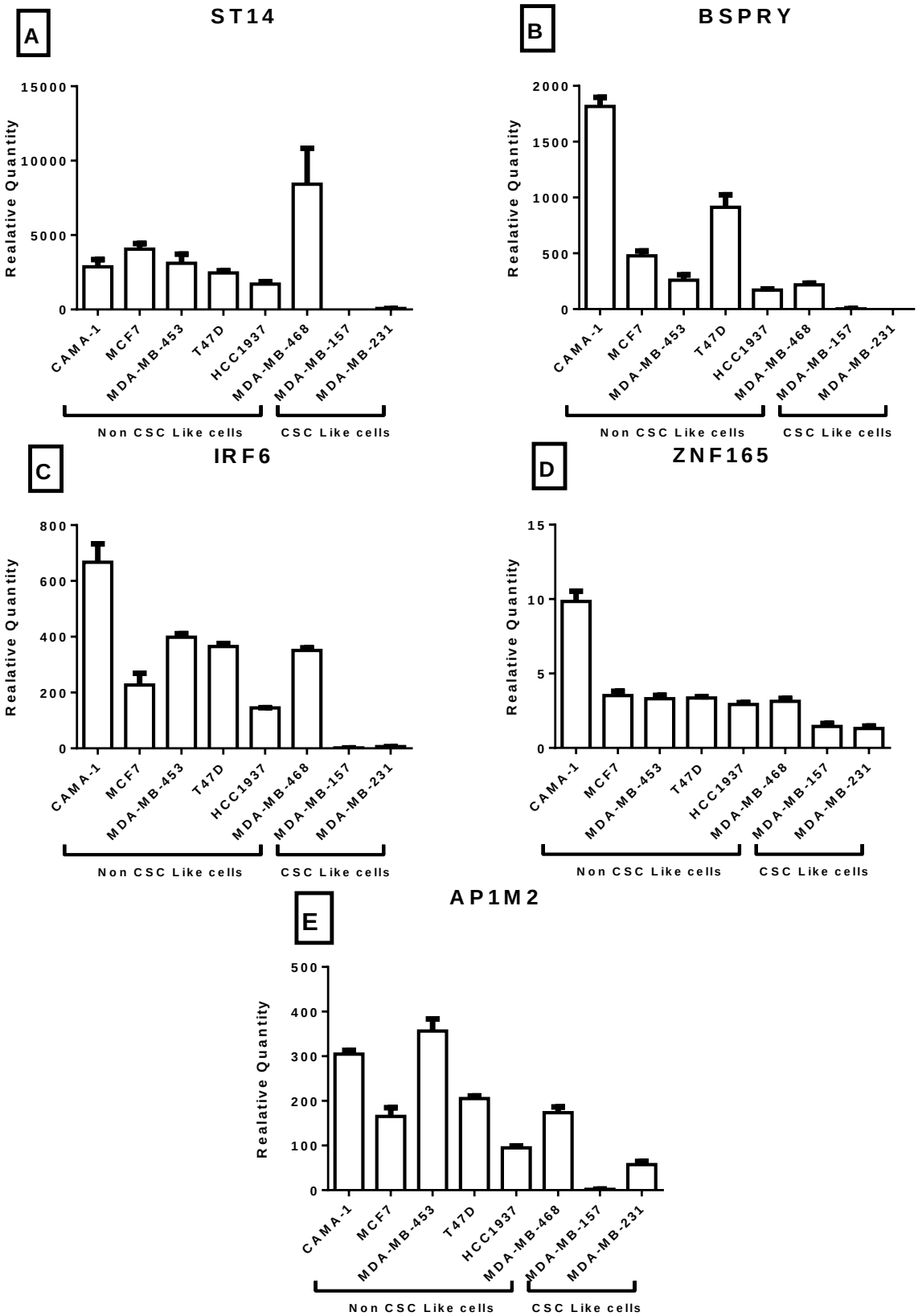

A
ST14
Realative Quantity
15000
10000
5000
0
CAMA-1
MCF7
MDA-MB-453
T47D
HCC1937
MDA-MB-468
MDA-MB-157
MDA-MB-231
Non CSC Like cells
CSC Like cells
B
BSPRY
Realative Quantity
2000
1500
1000
500
0
CAMA-1
MCF7
MDA-MB-453
T47D
HCC1937
MDA-MB-468
MDA-MB-157
MDA-MB-231
Non CSC Like cells
CSC Like cells
C
IRF6
Realative Quantity
800
600
400
200
0
CAMA-1
MCF7
MDA-MB-453
T47D
HCC1937
MDA-MB-468
MDA-MB-157
MDA-MB-231
Non CSC Like cells
CSC Like cells
D
ZNF165
Realative Quantity
15
10
5
0
CAMA-1
MCF7
MDA-MB-453
T47D
HCC1937
MDA-MB-468
MDA-MB-157
MDA-MB-231
Non CSC Like cells
CSC Like cells
E
AP1M2
Realative Quantity
500
400
300
200
100
0
CAMA-1
MCF7
MDA-MB-453
T47D
HCC1937
MDA-MB-468
MDA-MB-157
MDA-MB-231
Non CSC Like cells
CSC Like cells

**Figure 5.15: Comparison of (A) ST14, (B) BSPRY, (C) IRF6, (D) ZNF165 and (E) AP1M2 in breast cancer cell lines.** These genes were selected as upregulated ones in CSC like cell lines. Error bars show standard error of mean.

ST14, BSPRY, IRF6, ZNF165 and AP1M2 are genes which were found downregulated in CSC like cells and it was validated as well. MDA-MB-157 and MDA-MB-231 (CSC like cell lines) showed downregulation of these genes while other cell lines, CAMA-1, MCF7, MDA-MB-453, T47D, HCC1937 and MDA-MB-468 (non CSC like cell lines) showed upregulation of these genes.

Gene expression of downregulated and upregulated genes was further correlated with each other using Pearson r and results are shown in heat map below:

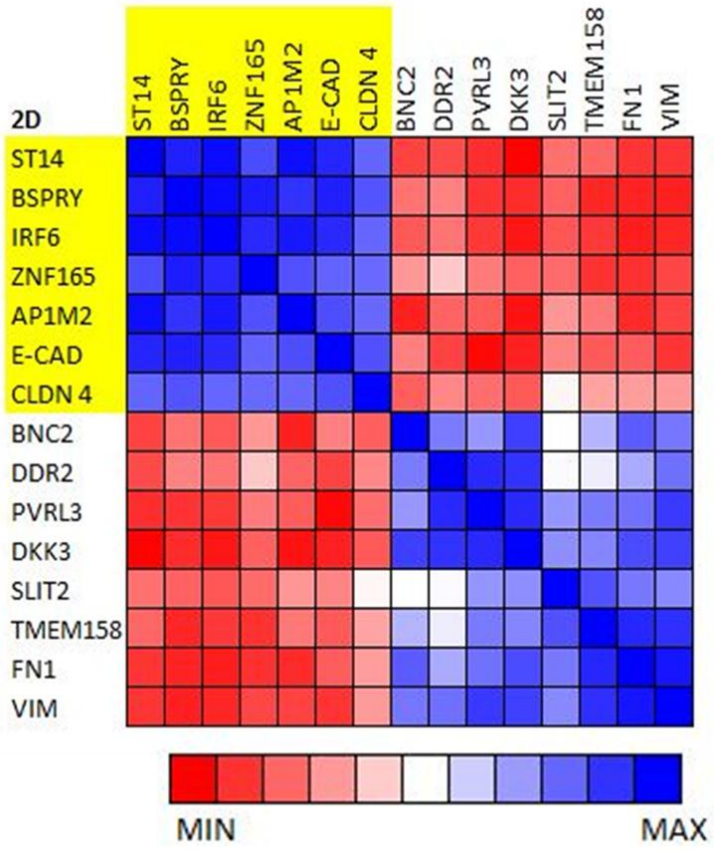



**Figure 5.16: Correlation between 2D and 3D gene expression for Patentable gene list.** Gene names highlighted with yellow color are downregulated and others are upregulated in CSC like cells.

When genes of Patentable gene list were correlated with each other using Pearson r, correlated pattern among them was seen. And with two sided T.Test which was used to determine the capability of genes to separate CSC like from non CSC like, 11 out of 15 genes were significant.

**Table 5.10: Capability of Non Patentable genes to separate CSC like from non CSC like.** Genes are sorted according to p-values.

| Sample Name | T Test p-Value | Fold Change |
|---|---|---|
| BSPRY | 0.0001 | -8.42 |
| IRF6 | 0.0001 | -6.69 |
| ST14 | 0.001 | -8.48 |
| E-CAD | 0.001 | -8.81 |
| VIM | 0.002 | 9.85 |
| PVRL3 | 0.004 | 9.45 |
| DKK3 | 0.004 | 7.29 |
| FN1 | 0.008 | 7.44 |
| AP1M2 | 0.015 | -4.34 |
| TMEM158 | 0.015 | 5.09 |
| ZNF165 | 0.021 | -1.53 |
| SLIT2 | 0.057 | 2.94 |
| DDR2 | 0.119 | 3.39 |
| BNC2 | 0.195 | 2.98 |
| CLDN 4 | 0.293 | 1.01 |

Then concordance of gene expression was calculated with microarray data from CCLE and it was high.

**Table 5.11: Concordance between microarray and qPCR gene expression data for Patentable gene list.**

| Gene | Pearson r | P-value |
|---|---|---|
| BSPRY | 0.98 | 2.7E-05 |
| TMEM158 | 0.93 | 9.8E-04 |
| DKK3 | 0.92 | 1.2E-03 |
| AP1M2 | 0.91 | 1.6E-03 |
| PVRL3 | 0.90 | 2.1E-03 |
| BNC2 | 0.90 | 2.2E-03 |
| IRF6 | 0.88 | 4.1E-03 |
| ST14 | 0.82 | 1.3E-02 |
| DDR2 | 0.78 | 2.1E-02 |
| ZNF165 | 0.76 | 3.0E-02 |
| SLIT2 | 0.53 | 1.7E-01 |

Out of both gene lists, genes of Non Patentable gene list were found insignificant and data was not concordant to microarray data as well. And from Patentable gene list all the genes were concordant with microarray data and 11 out of 15 genes were found significant.

But out of these 11 statistically significant genes, 7 showed clear difference in graph bars between CSC like and non CSC like.

**Table 5.12: Genes showed clear difference in graph bars between CSC like and non CSC like.**

| Sample Name | T Test p-Value | Fold Change |
|---|---|---|
| BSPRY | 0.0001 | -8.42 |
| IRF6 | 0.0001 | -6.69 |
| ST14 | 0.001 | -8.48 |
| E-CAD | 0.001 | -8.81 |
| VIM | 0.002 | 9.85 |
| PVRL3 | 0.004 | 9.45 |
| FN1 | 0.008 | 7.44 |

#### 5.2.1.3 Validation of Drug Response

From *in silico* analysis, four drugs were chosen whose cytotoxicity were checked using 2D culture conditions. Cells were grown in 96 well plates and their cytotoxicity for Elesclomol, Midostaurin, Lapatinib and Panobinostat was measured. For each cell line IC50, EC50, Activity area and Amax was calculated. All the drug response curves are mentioned in Appendix chapter.

**Table 5.13: IC50, EC50, Activity area and Amax values for Panobinostat, Lapatinib, Midostaurin and Elesclomol in 2D cell culture conditions.**

| Drug Name | Cell Line | IC50 (μM) | EC50 (μM) | Activity Area | Amax |
|---|---|---|---|---|---|
| Panobinostat | CAMA1 | 0.0110 | 0.0103 | 307.34 | 83.31 |
| Panobinostat | HCC1937 | 0.0296 | 0.0280 | 314.77 | 97.51 |
| Panobinostat | MCF7 | 0.0124 | 0.0146 | 313.70 | 89.00 |
| Panobinostat | MDAMB157 | 0.0120 | 0.0160 | 295.47 | 84.88 |
| Panobinostat | MDAMB231 | 0.0235 | 0.0255 | 277.11 | 84.29 |
| Panobinostat | MDAMB453 | 0.0188 | 0.0170 | 325.07 | 94.08 |
| Panobinostat | MDAMB468 | 0.0191 | 0.0206 | 324.08 | 95.63 |
| Panobinostat | T47D | 0.0172 | 0.0166 | 338.69 | 97.71 |
| Lapatinib | CAMA1 | 10.5741 | 11.0675 | 63.55 | 95.81 |
| Lapatinib | HCC1937 | 5.3322 | 5.7571 | 84.96 | 90.83 |
| Lapatinib | MCF7 | 6.3249 | 6.4390 | 86.64 | 97.69 |
| Lapatinib | MDAMB157 | 5.9305 | 5.3202 | 107.23 | 109.95 |
| Lapatinib | MDAMB231 | 20.0385 | 110.1227 | 50.43 | 307.25 |
| Lapatinib | MDAMB453 | 2.8093 | 14.9435 | 150.53 | 157.40 |
| Lapatinib | MDAMB468 | 2.2571 | 2.2884 | 131.48 | 98.44 |
| Lapatinib | T47D | 5.2815 | 5.8206 | 84.87 | 91.00 |
| Midostaurin | CAMA1 | 32.8046 | 27.1126 | 43.38 | 104.03 |
| Midostaurin | HCC1937 | 2.3328 | 0.5407 | 68.09 | 67.67 |
| Midostaurin | MCF7 | 1.3268 | 0.7236 | 58.04 | 68.55 |
| Midostaurin | MDAMB157 | 0.5435 | 0.3933 | 87.19 | 78.99 |
| Midostaurin | MDAMB231 | 4.3369 | 0.5330 | 48.92 | 50.46 |
| Midostaurin | MDAMB453 | NA | 1.1511 | 20.28 | 31.32 |
| Midostaurin | MDAMB468 | 0.3354 | 0.2259 | 100.36 | 74.78 |
| Midostaurin | T47D | 2.2270 | 1.0890 | 46.88 | 70.48 |
| Elesclomol | CAMA1 | 0.0029 | 0.0011 | 475.21 | 177.50 |
| Elesclomol | HCC1937 | 0.0043 | 0.0035 | 277.04 | 117.57 |
| Elesclomol | MCF7 | 0.0104 | 0.0103 | 201.73 | 101.11 |
| Elesclomol | MDAMB157 | 0.0123 | 0.0083 | 231.19 | 113.18 |
| Elesclomol | MDAMB231 | 0.0072 | 0.0063 | 241.03 | 109.43 |
| Elesclomol | MDAMB453 | 0.0021 | 0.0026 | 231.94 | 89.57 |
| Elesclomol | MDAMB468 | 0.0031 | 0.0031 | 250.68 | 99.65 |
| Elesclomol | T47D | 0.0070 | 0.0077 | 217.35 | 105.07 |
| Midostaurin | CAMA1 | 32.8046 | 27.1126 | 43.38 | 104.03 |

Then drug cytotoxicity values for cell lines were compared between CSC like and non CSC like groups.

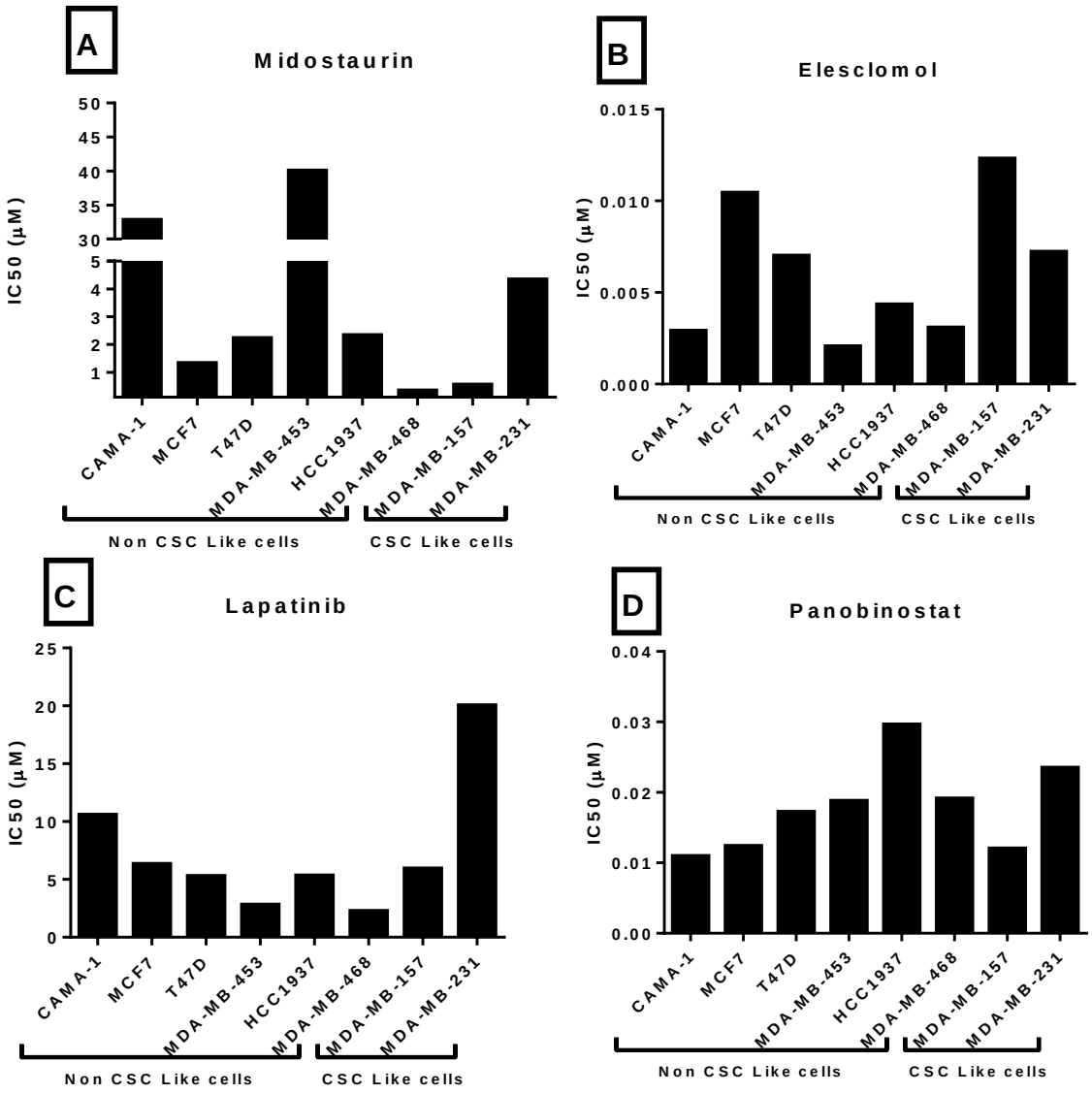



**Figure 5.17: Bar plots showing Comparison of Drug IC50 (A) Midostaurin, (B) Elesclomol, (C) Lapatinib and (D) Panobinostat against CSC like and non CSC like cell lines.** Y-axis shows IC50.

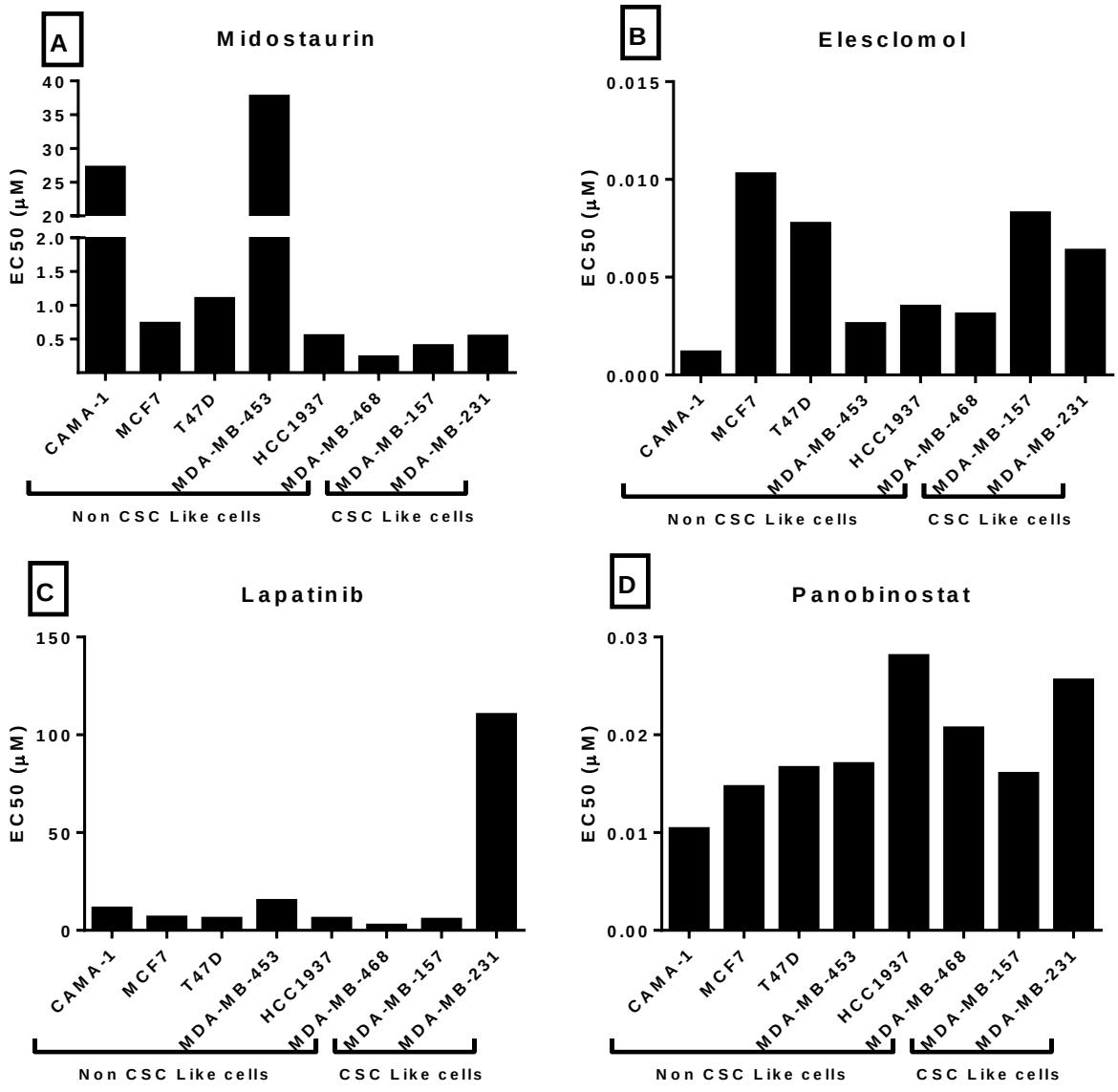



**Figure 5.18: Bar plots showing Comparison of Drug EC50 (A) Midostaurin, (B) Elesclomol, (C) Lapatinib and (D) Panobinostat against CSC like and non CSC like cell lines.** Y-axis shows EC50.

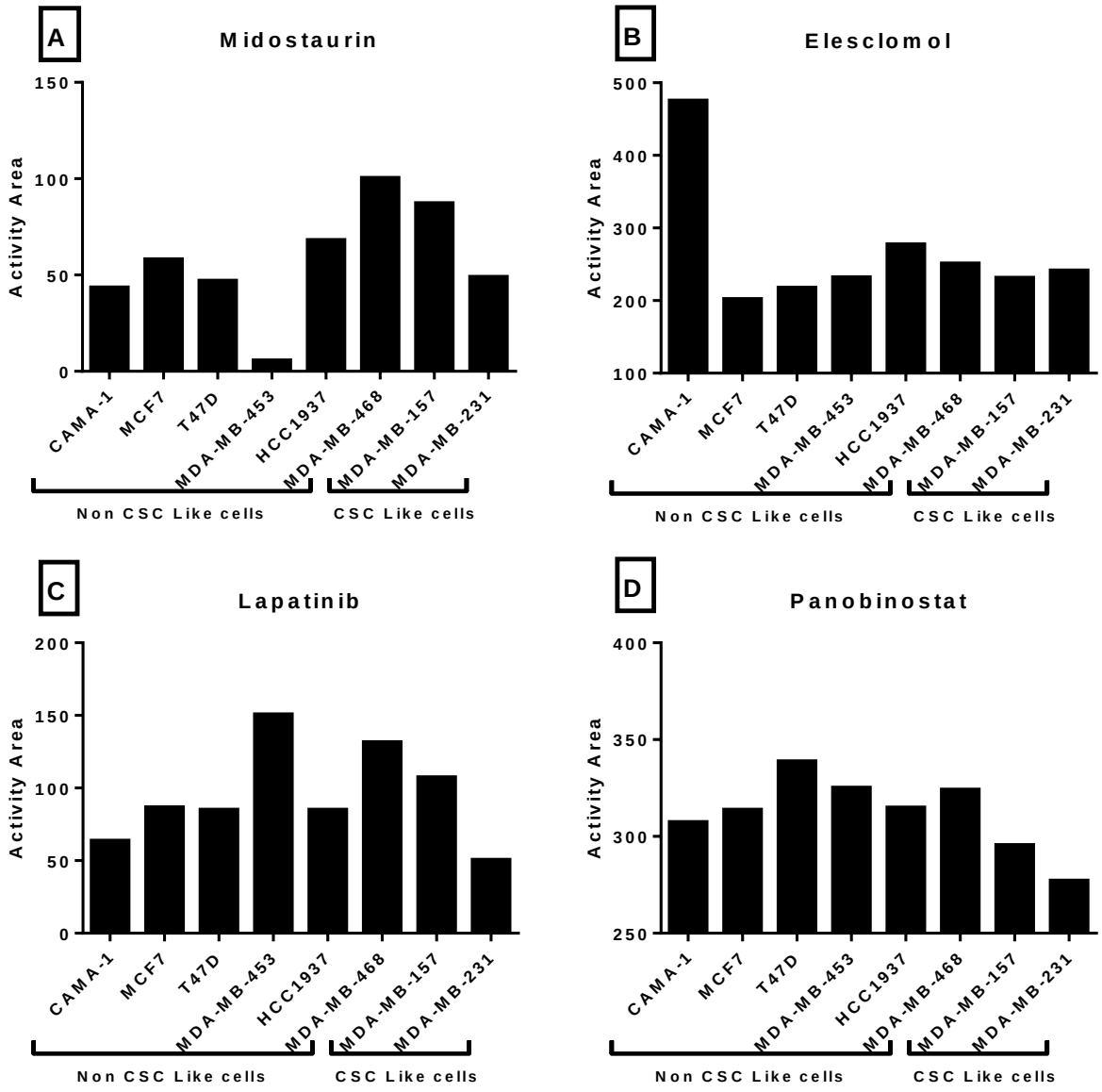



**Figure 5.19: Bar plots showing Comparison of Drug Activity area (A) Midostaurin, (B) Elesclomol, (C) Lapatinib and (D) Panobinostat against CSC like and non CSC like cell lines.** Y-axis shows Activity area.

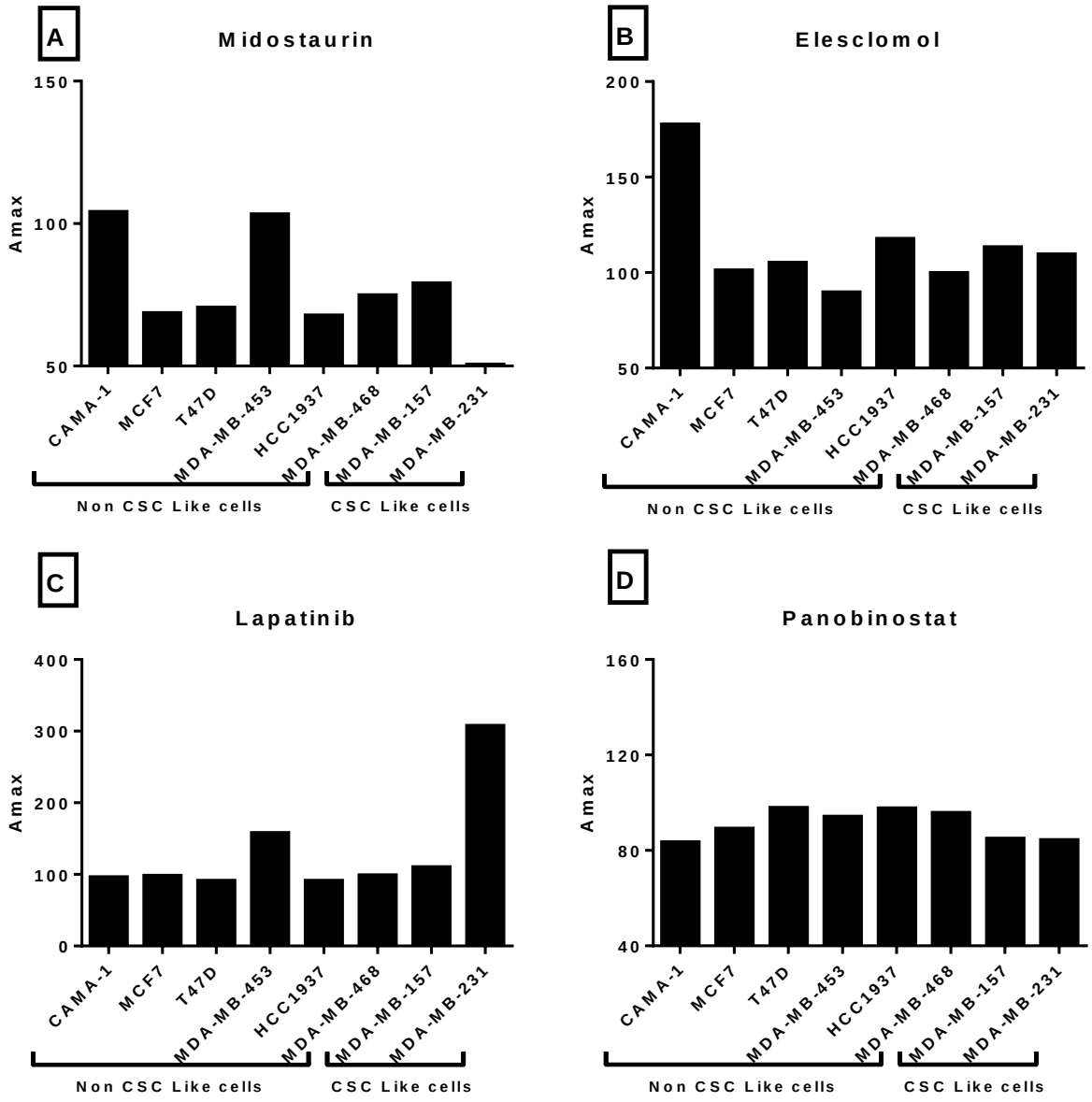



**Figure 5.20: Bar plots showing Comparison of Drug Amax (A) Midostaurin, (B) Elesclomol, (C) Lapatinib and (D) Panobinostat against CSC like and non CSC like cell lines.** Y-axis shows Amax.

Then cumulative drug response was seen by comparing mean of drug parameters of CSC like cell lines versus mean of non CSC like cell lines. Such cutoff was used that if one drug is showing effect against a specific group by three parameters then that drug was considered as toxic against that group.

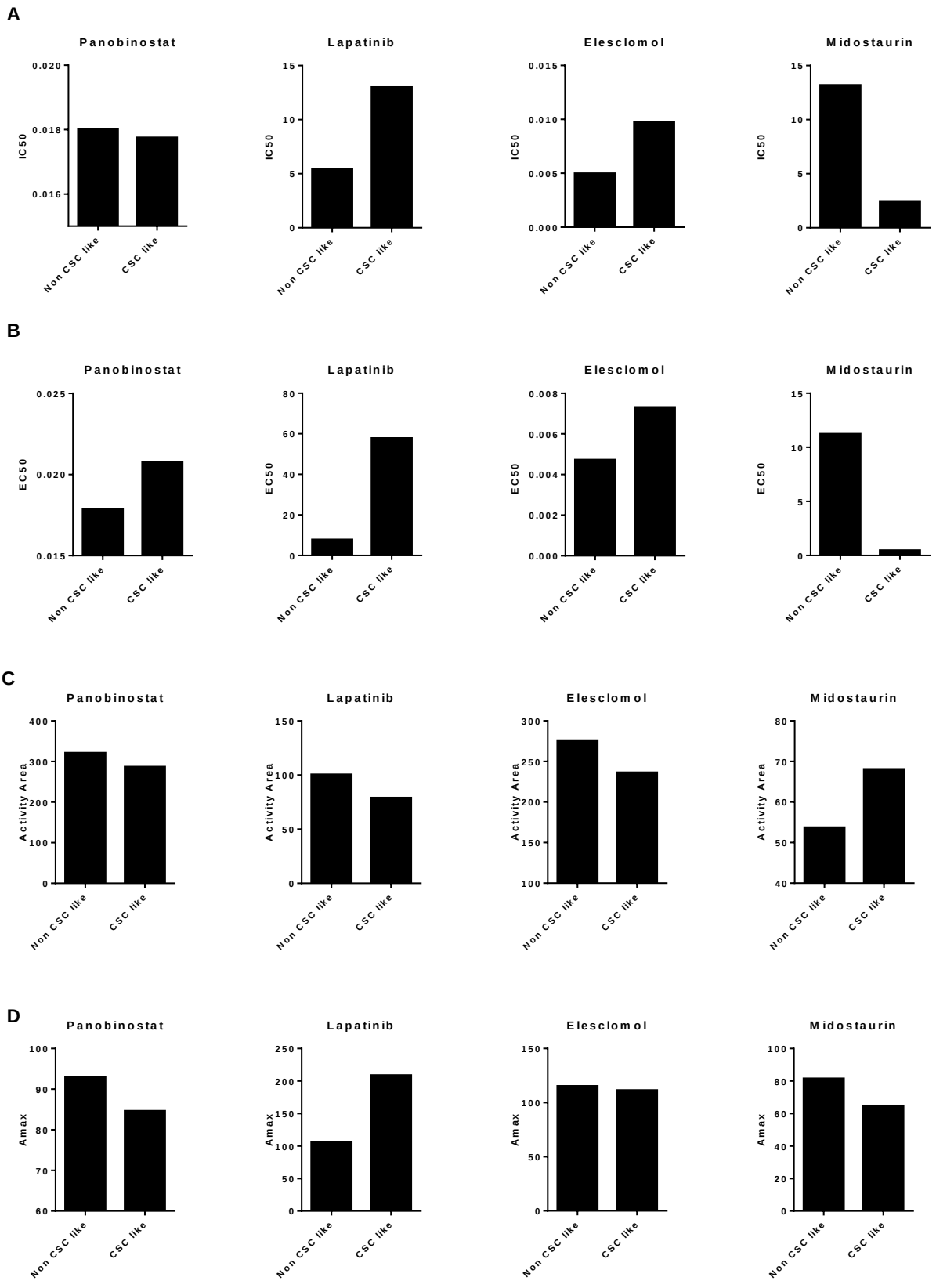

A
Panobinostat
IC50
0.020
0.018
0.016
Non CSC like
CSC like
Lapatinib
IC50
15
10
5
0
Non CSC like
CSC like
Elesclomol
IC50
0.015
0.010
0.005
0.000
Non CSC like
CSC like
Midostaurin
IC50
15
10
5
0
Non CSC like
CSC like
B
Panobinostat
EC50
0.025
0.020
0.015
Non CSC like
CSC like
Lapatinib
EC50
80
60
40
20
0
Non CSC like
CSC like
Elesclomol
EC50
0.008
0.006
0.004
0.002
0.000
Non CSC like
CSC like
Midostaurin
EC50
15
10
5
0
Non CSC like
CSC like
C
Panobinostat
Activity Area
400
300
200
100
0
Non CSC like
CSC like
Lapatinib
Activity Area
150
100
50
0
Non CSC like
CSC like
Elesclomol
Activity Area
300
250
200
150
100
Non CSC like
CSC like
Midostaurin
Activity Area
80
70
60
50
40
Non CSC like
CSC like
D
Panobinostat
Amax
100
90
80
70
60
Non CSC like
CSC like
Lapatinib
Amax
250
200
150
100
50
0
Non CSC like
CSC like
Elesclomol
Amax
150
100
50
0
Non CSC like
CSC like
Midostaurin
Amax
100
80
60
40
20
0
Non CSC like
CSC like

**Figure 5.21: Cumulative drug response of drugs parameters against non CSC like vs. CSC like (A) IC50, (B) EC50, (C) Activity area and (D) Amax.**

Panobinostat showed increased activity against non CSC like cells as compared to CSC like cells by three parameters (EC50, Activity area and Amax) as predicted by *in silico* analysis but only IC50 showed opposite results. Lapatinib showed increased activity against non CSC like cells as compared to CSC like cells by three parameters (IC50, EC50 and Activity area) as predicted by *in silico* analysis but only Amax showed opposite results. Elesclomol showed increased activity against non CSC like cells as compared to CSC like cells by all four parameters (IC50, EC50, Activity area and Amax) as opposed to prediction by *in silico* analysis. Midostaurin showed increased activity against CSC like cells as compared to non CSC like cells by three parameters (IC50, EC50 and Amax) as predicted by *in silico* analysis but only Activity area showed opposite results. But comparisons between CSC like and non CSC like resulted insignificant by all 4 parameters using 2 sided T.Test as shown in table below due to the less number of samples used..

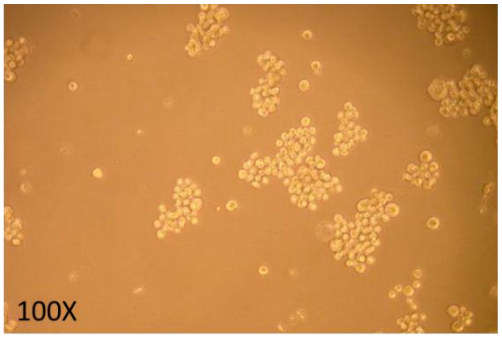
**Table 5.14: Comparison of drug parameters for CSC like cells vs. non CSC like cells using 2 sided T.Test.**

| Drug name | IC50 p-value | EC50 p-value | Activity area p-value | Amax p-value |
|---|---|---|---|---|
| Elesclomol | 0.122092 | 0.36569 | 0.619088 | 0.87917 |
| Lapatinib | 0.107603 | 0.091704 | 0.473401 | 0.085631 |
| Midostaurin | 0.45886 | 0.42123 | 0.586546 | 0.294687 |
| Panobinostat | 0.96431 | 0.584722 | 0.010398 | 0.097471 |

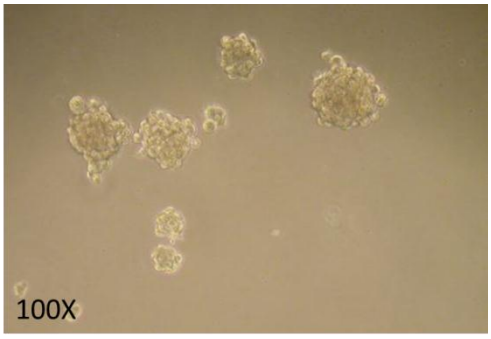
So we concluded that Panobinostat and Lapatinib act against non CSC like cells and only Midostaurin work against CSC like cells.

### 5.2.2 Validation in 3D Cell Culture Conditions

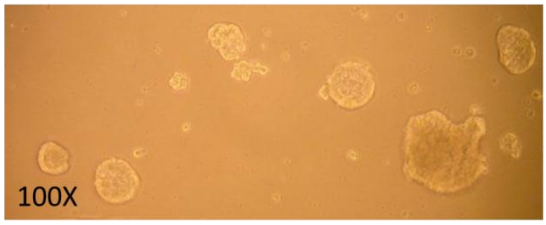
For enriching of CSCs, 3D cell culture conditions are the gold standard. So next we cultured CSC like cells and non CSC like cells in 3D cell culture conditions to see if these cell lines show more mesenchymal behavior or switch from epithelial to mesenchymal behavior. All the cell lines formed mammosphere. MDA-MB-453, MDA-MB-468, MCF7, HCC1937 and T47D cell lines were cultured for 12 passages in these conditions while MDA-MB-157 and MDA-MB-231 were cultured for 6 passages and CAMA-1 was cultured for 9 passages. Less no of passages was used for MDA-MB-157, MDA-MB-231 and CAMA-1, because their cell number decreased a lot. qPCR analysis was performed for both gene lists. Previously no gene from Non Patentable gene list was significant, so its results are mentioned in appendix.



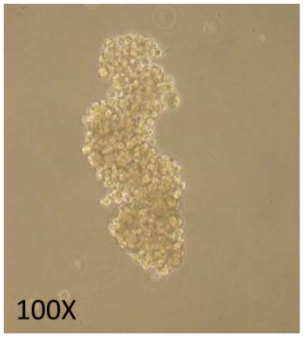

CAMA-1



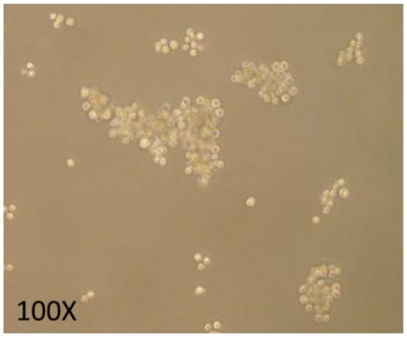

HCC1937



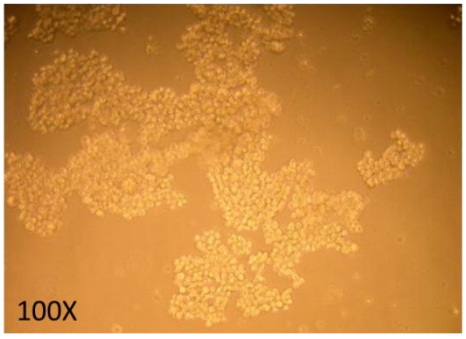

MCF7



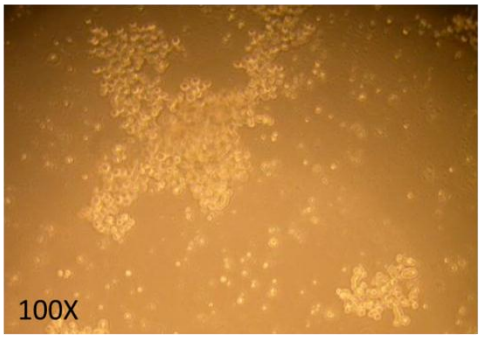

MDA-MB-157



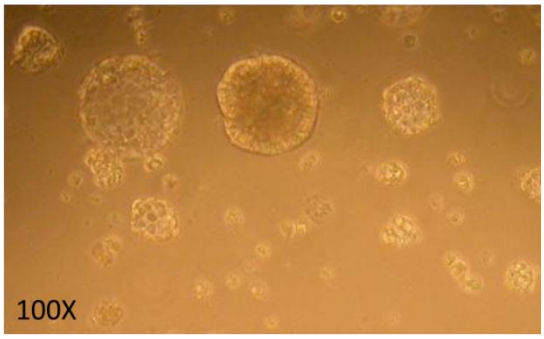

MDA-MB-231




MDA-MB-453




MDA-MB-468

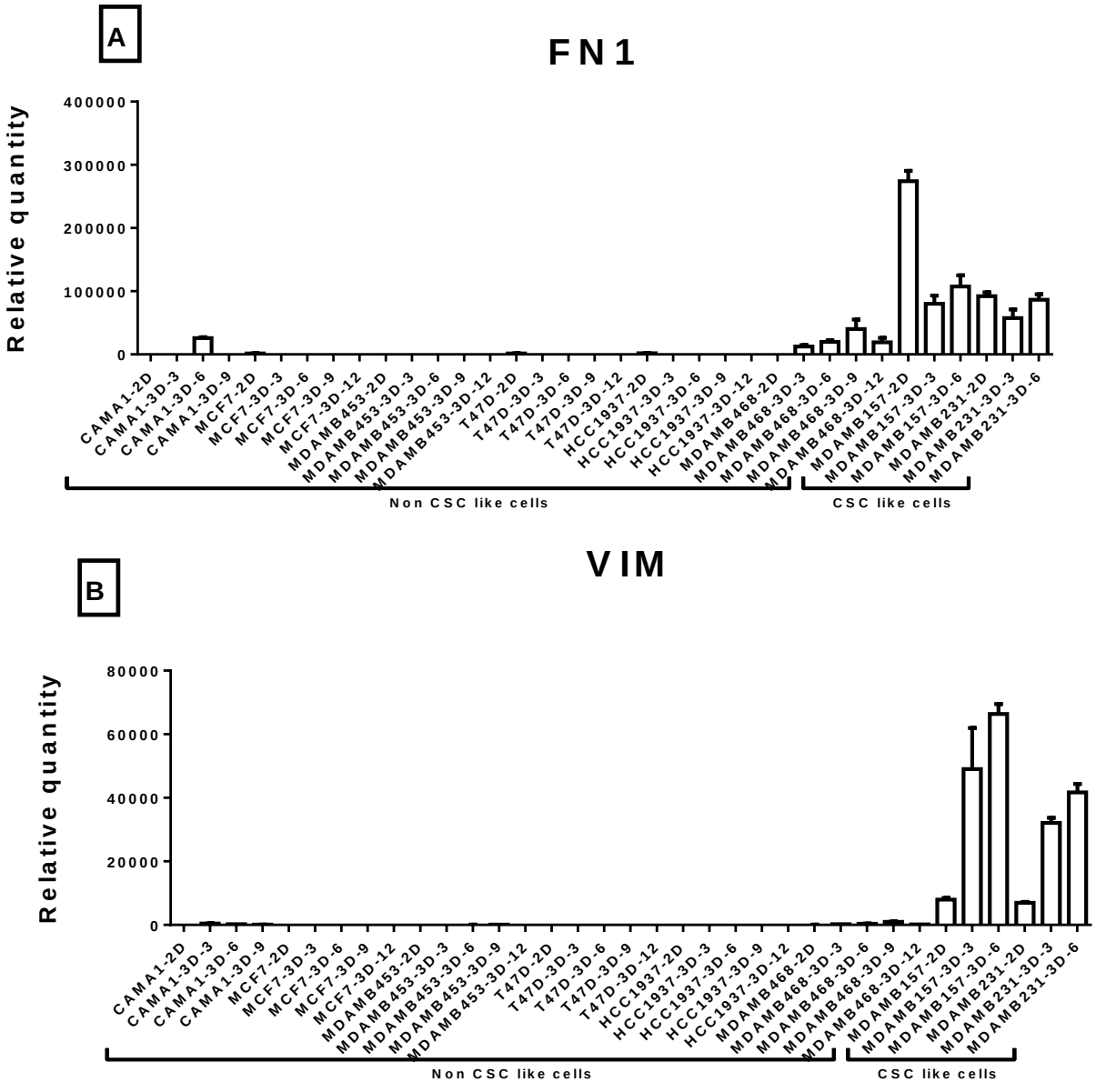



T47D

**Figure 5.22: Mammosphere formed by Breast cancer cell lines. (A) CAMA-1, (B) HCC1937, (C) MCF7, (D) MDA-MB-157, (E) MDA-MB- 231, (F) MDA-MB-453 (G) MDA-MB-468 and (H)T47D.** Scale used was 100X.

#### 5.2.2.1 Expression of Patentable Gene List in 3D

When cells were grown in 3D cell culture conditions, none of the cell line showed increase in stemness qualities except MDA-MB-231.



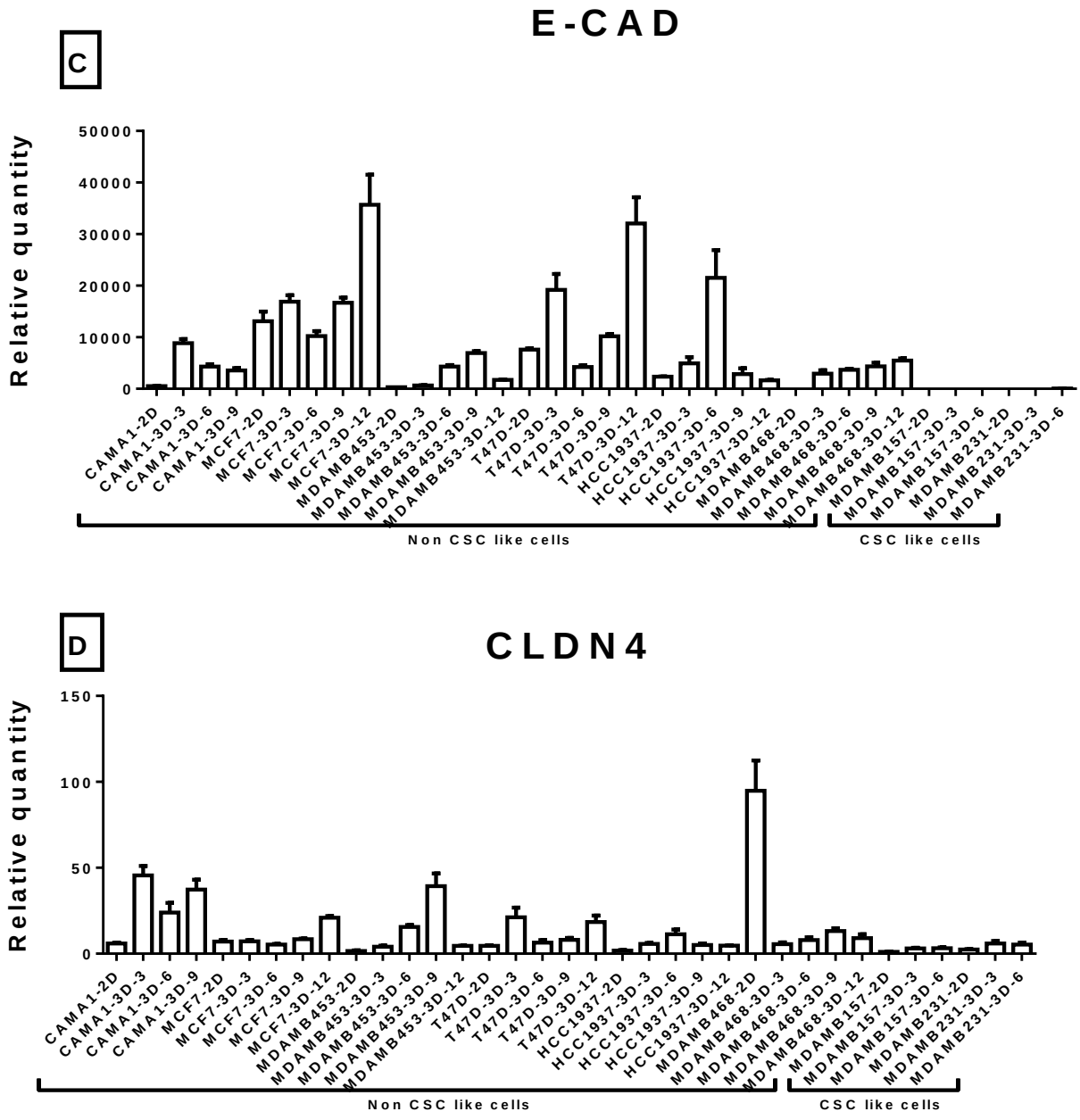



**Figure 5.23: Comparison of (A) VIM, (B) FN1, (C) E-CAD and (D) CLDN4 between breast cancer cell lines**. 2D and 3D mentioned after name shows cell culture conditions. Passage number for each sample, are given for 3D after sample name. Error bars show standard error of mean.

As there was no consistent and expected change in gene expression behavior in cell lines except MDA-MB153, so rest of the results are given in Appendix. Additionally we narrowed down Patentable gen list to 7 genes so these results are mentioned below.

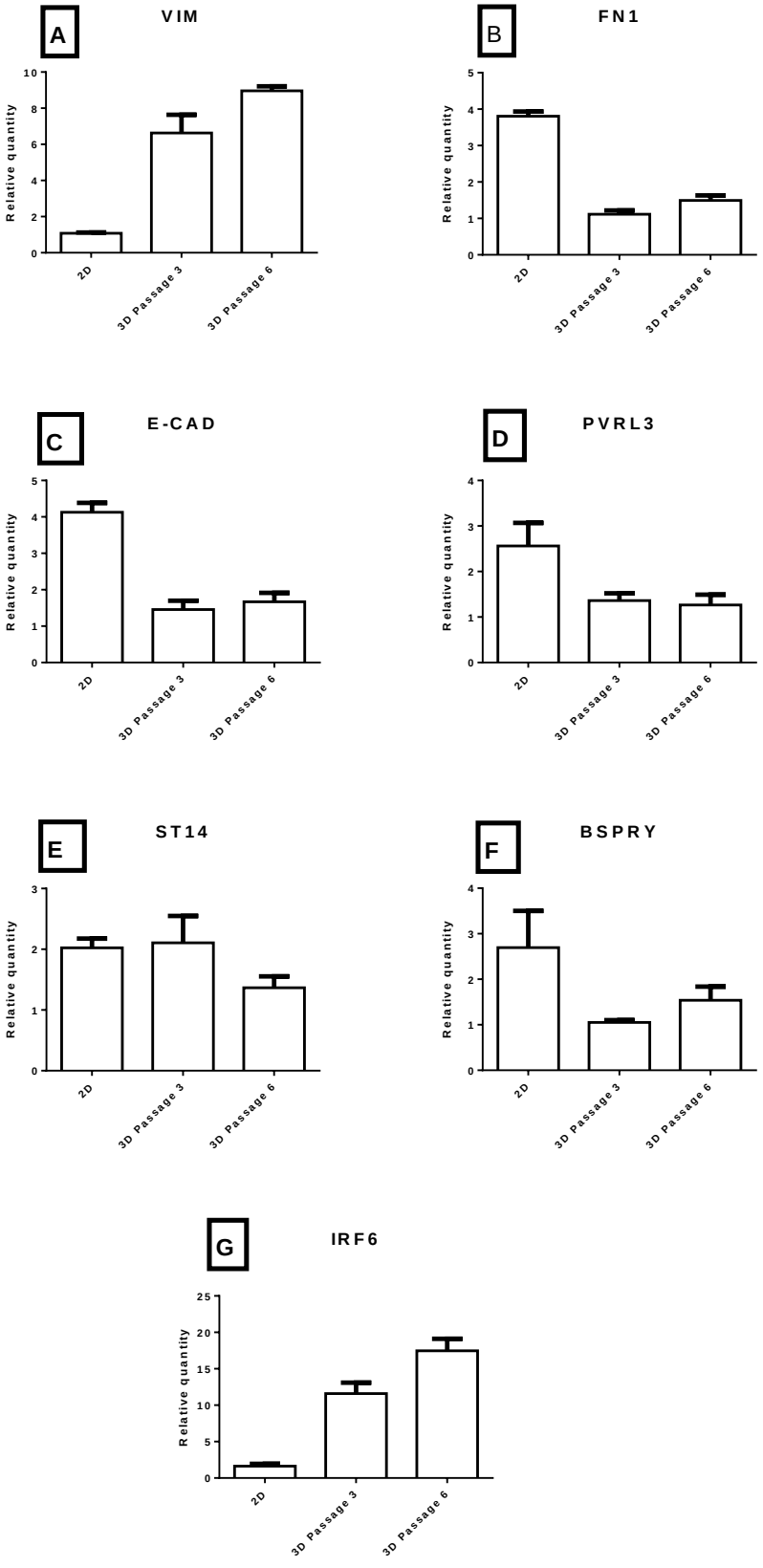

A
VIM
Relative quantity
10
8
6
4
2
0
2D
3D Passage 3
3D Passage 6
B
FN1
Relative quantity
5
4
3
2
1
0
2D
3D Passage 3
3D Passage 6
C
E-CAD
Relative quantity
5
4
3
2
1
0
2D
3D Passage 3
3D Passage 6
D
PVRL3
Relative quantity
4
3
2
1
0
2D
3D Passage 3
3D Passage 6
E
ST14
Relative quantity
3
2
1
0
2D
3D Passage 3
3D Passage 6
F
BSPRY
Relative quantity
4
3
2
1
0
2D
3D Passage 3
3D Passage 6
G
IRF6
Relative quantity
25
20
15
10
5
0
2D
3D Passage 3
3D Passage 6

**Figure 5.24: Comparison of selected genes (A) VIM, (B) FN1, (C) E-CAD and (D) PVRL3, (E) ST14, (F) BSPRY and (G) IRF6 in MDA-MB-157 breast cancer cell line**. 2D and 3D mentioned after name shows cell culture conditions. Passage number for each sample, are given for 3D after sample name. Error bars show standard error of mean.

Out of these 7 genes only 4 genes showed expected behavior. VIM, FN1 and PVRL3 expression should increase according to our predictions but only VIM showed such behavior. And E-CAD, ST14, IRF6 and BSPRY expression should decrease according to prediction all of them showed decrease in expression with only exception of IRF6. Behavior of other genes was also determined, but as we dropped other genes in 2D analysis so their results are shown in Appendix. Among these 8 genes, DDR2, SLIT2 and TMEM158 showed increased expression and, ZNF165 and AP1M2 showed decrease in expression from 2D to 3D transition as expected. And BNC2 and DKK3 showed no change. While only CLDN4 showed opposite behavior as from expected.

In conclusion VIM, E-CAD, ST14, and BSPRY showed increase in stemness in MDA-MB-157 cell line.

#### 5.2.2.2 Drug response of Breast cancer cell lines cultured in 3D

All breast cancer cell lines were cultured in 96 well ultra-low attachment plates and treated with Midostaurin, Elesclomol, Lapatinib and Panobinostat. But as only MDA-MB-157 was able to increase its stemness properties, so its results are mentioned below only for Midostaurin, Lapatinib and Panobinostat. Elesclomol was excluded from this analysis as it did not behaved for 2D cultured cells opposite of our *in* silico analysis. All drug response curves are mentioned in Appendix section.

**Table 5.15: IC50, EC50, Activity area and Amax values for Panobinostat, Lapatinib, Midostaurin and Elesclomol in 3D cell culture conditions.**

| Drug Name | Cell Line | IC50 (μM) | EC50 (μM) | Activity Area | Amax |
|---|---|---|---|---|---|
| Panobinostat | HCC1937 | 0.0674 | 0.0739 | 247.86 | 87.71 |
| Panobinostat | MCF7 | 0.0865 | 0.0615 | 313.47 | 108.48 |
| Panobinostat | MDAMB157 | 0.0344 | 0.0340 | 323.38 | 102.25 |
| Panobinostat | MDAMB231 | 0.5647 | 0.2904 | 199.99 | 89.56 |
| Panobinostat | MDAMB453 | 0.0121 | 0.0057 | 546.05 | 141.23 |
| Panobinostat | MDAMB468 | 0.0755 | 0.0742 | 279.00 | 98.81 |
| Panobinostat | T47D | 0.0574 | 0.0484 | 328.20 | 109.34 |
| Lapatinib | CAMA1 | 48.716 | 197.668 | 27.71 | 285.31 |
| Lapatinib | HCC1937 | 7.123 | 6.939 | 83.04 | 97.04 |
| Lapatinib | MCF7 | 9.486 | 17.307 | 72.69 | 123.87 |
| Lapatinib | MDAMB157 | 17.079 | 56.807 | 56.88 | 208.85 |
| Lapatinib | MDAMB231 | 29.676 | 29.451 | 43.01 | 100.38 |
| Lapatinib | MDAMB453 | 2.600 | 4.763 | 461.96 | 596.40 |
| Lapatinib | MDAMB468 | 5.621 | 5.492 | 99.19 | 103.83 |
| Lapatinib | T47D | 15.703 | 54.018 | 53.02 | 187.56 |
| Midostaurin | CAMA1 | 16.207 | 14.473 | 13.55 | 105.99 |
| Midostaurin | HCC1937 | NA | 0.443 | 38.31 | 35.28 |
| Midostaurin | MCF7 | 5.447 | 4.807 | 40.34 | 105.25 |
| Midostaurin | MDAMB157 | 0.399 | 0.229 | 81.24 | 60.87 |
| Midostaurin | MDAMB231 | NA | 0.002 | NA | NA |
| Midostaurin | MDAMB453 | NA | 0.869 | 47.30 | 56.89 |
| Midostaurin | MDAMB468 | 0.564 | 0.356 | 103.50 | 90.36 |
| Midostaurin | T47D | NA | 0.618 | NA | NA |
| Elesclomol | CAMA1 | NA | 0.091 | 87.38 | 80.73 |
| Elesclomol | HCC1937 | 15.886 | 14.572 | 3.04 | 104.51 |
| Elesclomol | MCF7 | NA | 0.001 | 130.02 | 45.81 |
| Elesclomol | MDAMB157 | NA | NA | NA | NA |
| Elesclomol | MDAMB231 | NA | 0.008 | 14.49 | 6.96 |
| Elesclomol | MDAMB453 | 0.105 | 0.009 | 92.67 | 45.56 |
| Elesclomol | MDAMB468 | NA | 0.047 | 49.85 | 36.85 |
| Elesclomol | T47D | NA | NA | NA | NA |

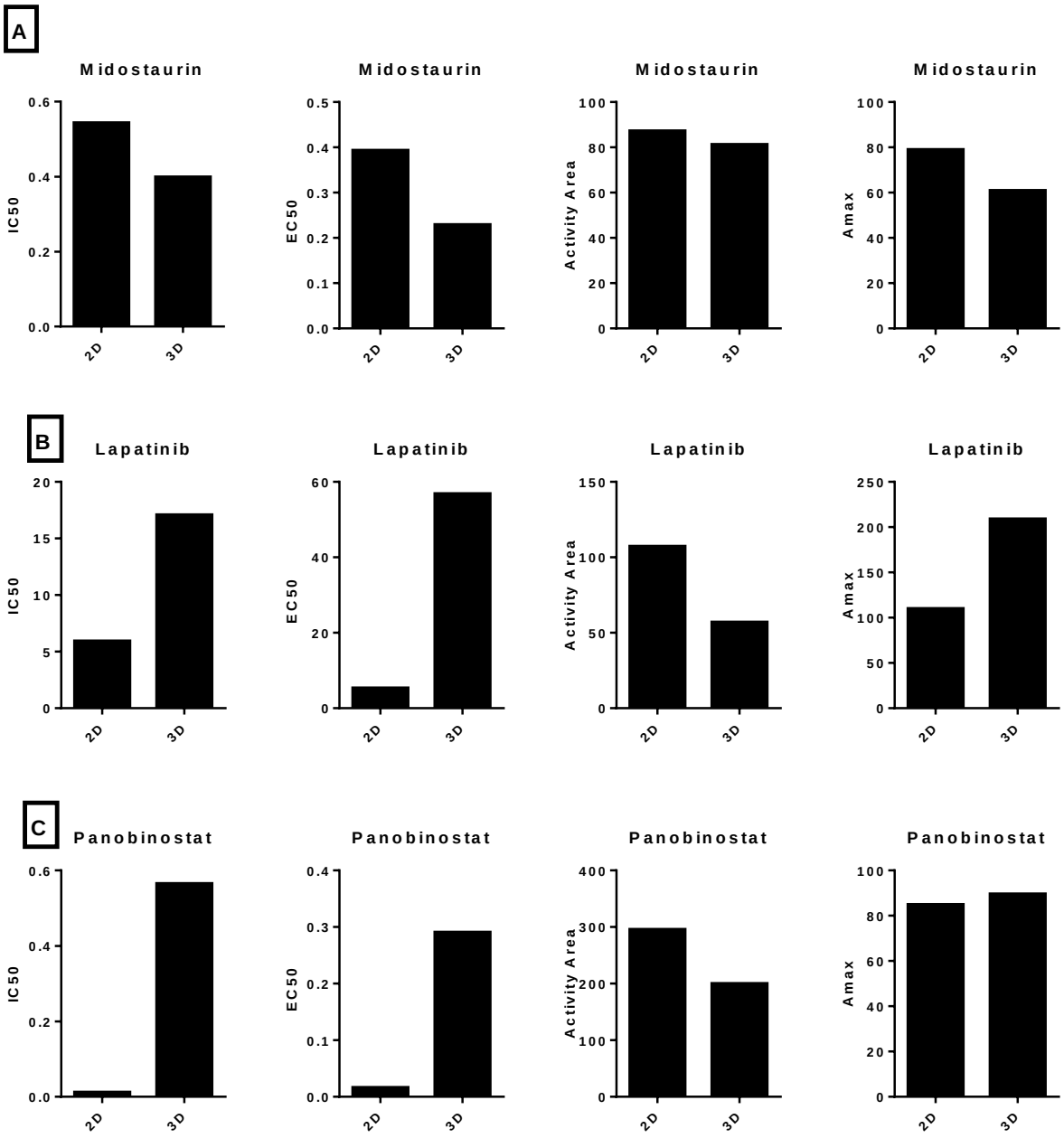



**Figure 5.25: Bar plots showing Comparison of Drug parameters for (A) Midostaurin, (B) Lapatinib and (C) Panobinostat against MDA-MB157 cells cultured in 2D and 3D culture conditions.** Y-axis show drug parameters used.

Midostaurin by all four parameters targeted 3D cultured cells as compared to 2D cultured cells as we expected. And Panobinostat and Lapatinib targeted 2D cultured cells as compared to 3D cultured cells by three parameters (IC50, EC50 and Activity area) but not by Amax.

So we concluded MDA-MB-157 showed increased stemness behavior in 3D culture conditions and became more sensitive to Elesclomol and resistant to Lapatinib and Panobinostat as expected.

### 5.3 Correlation between Gene Expression and Cytotoxicity Data

Pearson correlation between gene expression and drug cytotoxicity for 2D cultured cells was determined. And expression values for all genes data from Patentable gene list was used as this was concordant with microarray data.

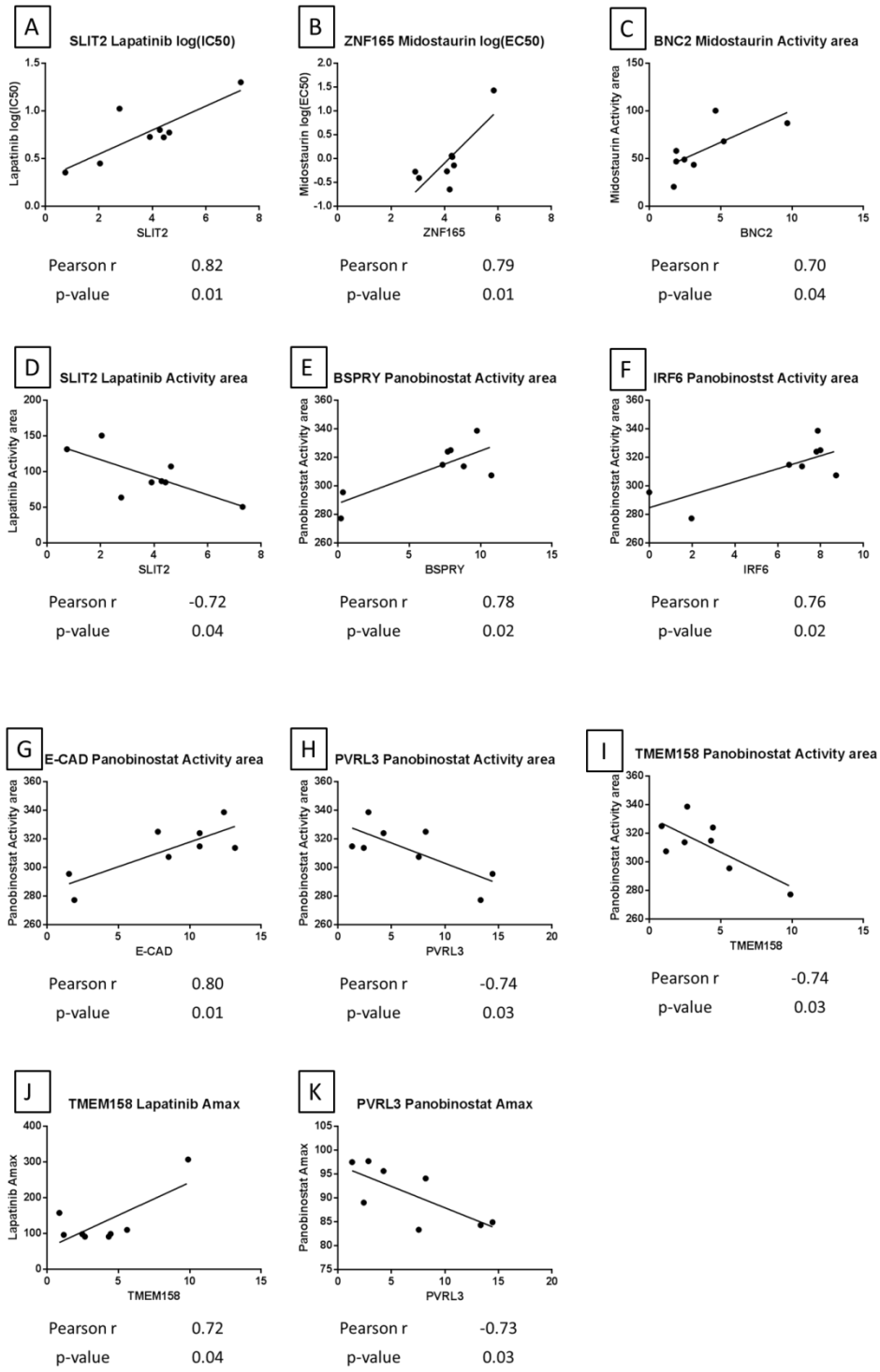


**Figure 5.26: Pearson correlation analysis results between gene expression and drug cytotoxicity parameters.** (A) SLIT2 expression correlation with Lapatinib log(IC50) value, (B) ZNF165 expression correlation with Midostaurin log(EC50) value, (C) BNC2 expression correlation with Midostaurin Activity area value, (D) SLIT2 expression correlation with Lapatinib Activity area value (E) BSPRY expression correlation with Panobinostat Activity area value, (F) IRF6 correlation with Panobinostat Activity area value, (G) E-CAD expression correlation with Panobinostat Activity area value, (H) PVRL3 expression correlation with Panobinostat Activity area value, (I) TMEM158 expression correlation with Panobinostat Activity area value, (J) PVRL3 expression correlation with Panobinostat Amax value.

SLIT2 gene expression has a positive correlation with Lapatinib IC50. So if SLIT2 expression is upregulated in CSC like cells then Lapatinib will not be a good drug for this group but if its expression is low as in non CSC like cells then Lapatinib will give better results. ZNF165 has a positive correlation with MidostaurinEC50. So if ZNF165 expression goes down, as in CSC like cells, then Midostaurin will affect these cells more as compared to non CSC. BNC2 gene has a positive correlation with Midostaurin Activity area. So if BNC2 expression is more as in CSC like cells then Midostaurin will be a good choice to target these cells. SLIT2 gene has negative correlation with Lapatinib Activity area. As SLIT2 is upregulated in CSC like cells so Lapatinib will not affect these cells but non CSC like cells can be targeted very well with this drug. BSPRY gene has positive correlation with Panobinostat Activity area. As BSPRY expression will increase as in non CSC like cells then Panobinostat is a good option against this kind of cells. IRF6 gene has positive correlation with Panobinostat Activity area. As happens in case of non CSC like cells as compared to CSC like cells, expression of BSPRY is high so Panobinostat can act better against non CSC like cells. E-Cad expression is positively correlated with Panobinostat Activity area. So if E-CAD expression is high as in non CSC like cells considering its level in CSC like cells, then Panobinostat will be our drug of choice. PVRL3 gene expression is negatively correlated with Panobinostat Activity area. So if PVRL3 expression is high, as in CSC like cells, then Panobinostat will not be effective against these but against non CSC like cells this drug will be more effective. TMEM158 gene is upregulated in CSC like cells and it is negatively correlated with Panobinostat Activity area. So Panobinostat will be more effective against non CSC like cells. PVRL3 expression is upregulated in CSC like cells as compared to non CSC like cells. So Panobinostat will be more effective against non CSC like cells due to its negative correlation with PVRL3.

## 5.4 Principle Component Analysis

To determine further that if we can use more than one gene, as combination, to describe the correlation with drug cytotoxicity, Principle Component analysis (PCA) was performed. Only SybrGreen qPCR data was used for this analysis. There were 32767 combinations of genes. For each combination of genes First component value (PC1) was determined and this value was then correlated with drug cytotoxicity data of cell lines cultured in 2D conditions. All 4 parameters determined previously were used for this analysis.

When PCA analysis was performed correlating with log(IC50) of drugs, 29 different combinations of genes were significant.

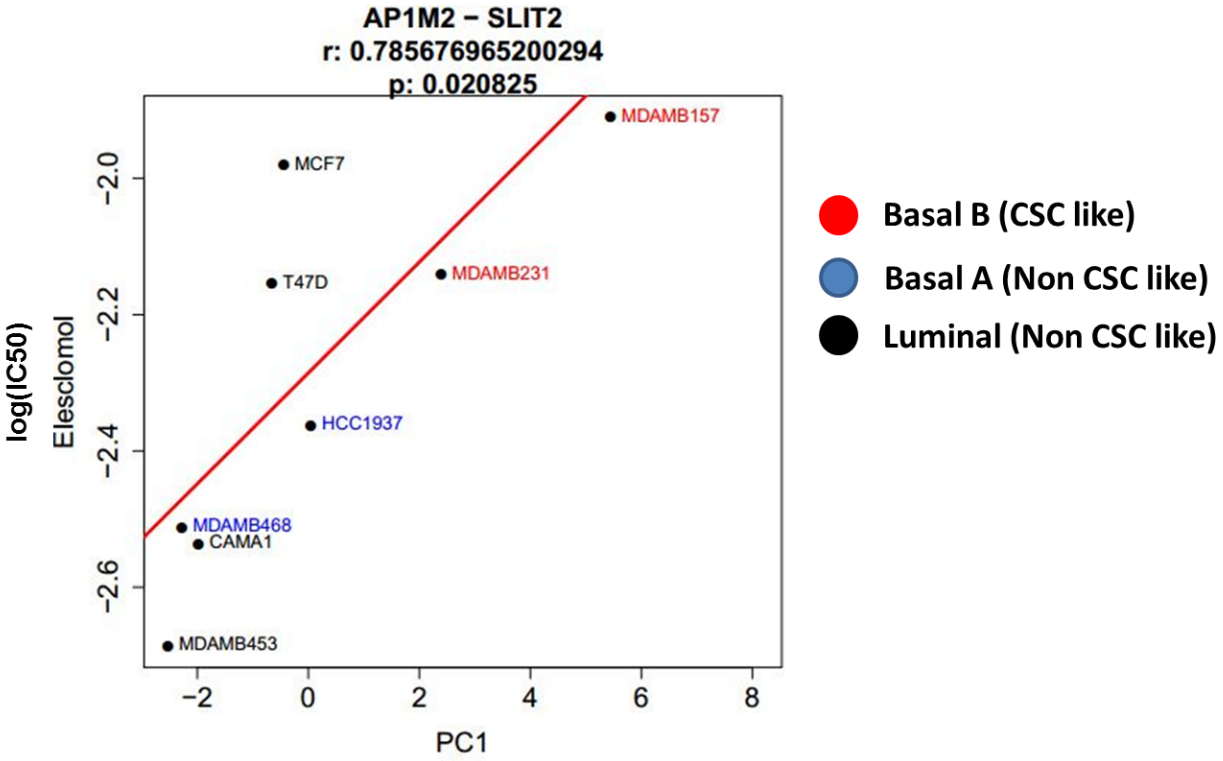



**Figure 5.27: Correlation of PC1 of AP1M2 and SLIT2 with Elesclomol IC50.** X axis shows PC1 for AP1M2 and SLIT2 and y axis shows log(IC50) values of Elesclomol obtained against breast cancer cell lines. Cell lines names shown in red color are Basal B (CSC like), in blue color are Basal A (Non CSC like) and in black color are Luminal (Non CSC like). r value and p-value are also mentioned. Percentage of variation explained by PC1 was 0.73.

AP1M2 and SLIT2 combination was most significant when analysis with log (IC50) was performed. So by using these two genes expression, sensitivity of Elesclomol can be predicted. PC1 value for this combination will increase, if expression of AP1M2 is decreased and SLIT2 expression is increased. And increase in PC1 will result into increased sensitivity of cells for Elesclomol.

When PCA analysis was performed correlating with log (EC50) of drugs, only one significant result was achieved.

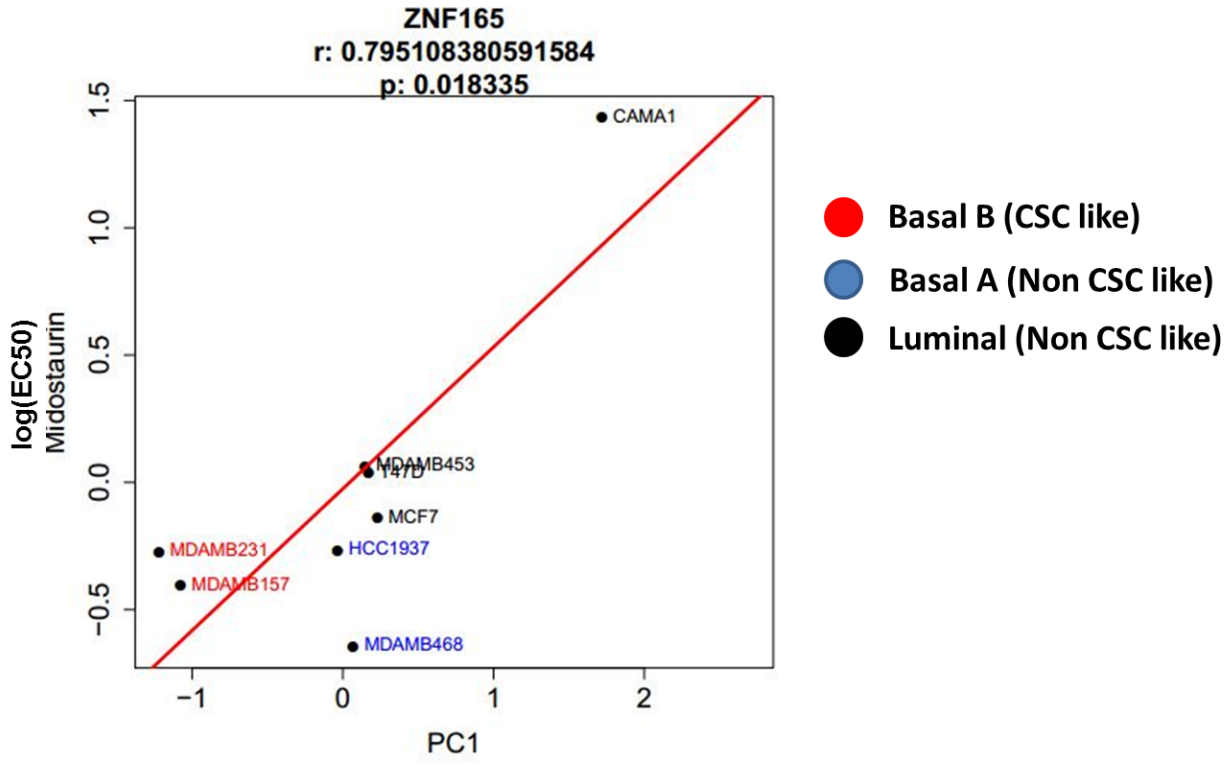



**Figure 5.28: Correlation of PC1 of ZNF165 with Midostaurin EC50.** X axis shows PC1 for ZNF165 and y axis shows log(EC50) values of Midostaurin obtained against breast cancer cell lines. Cell lines names shown in red color are Basal B (CSC like), in blue color are Basal A (Non CSC like) and in black color are Luminal (Non CSC like). r value and p-value are also mentioned. Percentage of variation explained by PC1 was 0.73.

Correlation of ZNF165 with log (EC50) was the only significant result when PCA analysis was performed with EC50 values of all drugs. So by using this gene expression, sensitivity of Midostaurin can be predicted. PC1 value for this ZNF165 will increase, if its expression is decreased. And increase in PC1 will result into increased sensitivity of cells for Elesclomol.

When PCA analysis was performed correlating with Activity area of drugs, 30070 significant results were achieved.

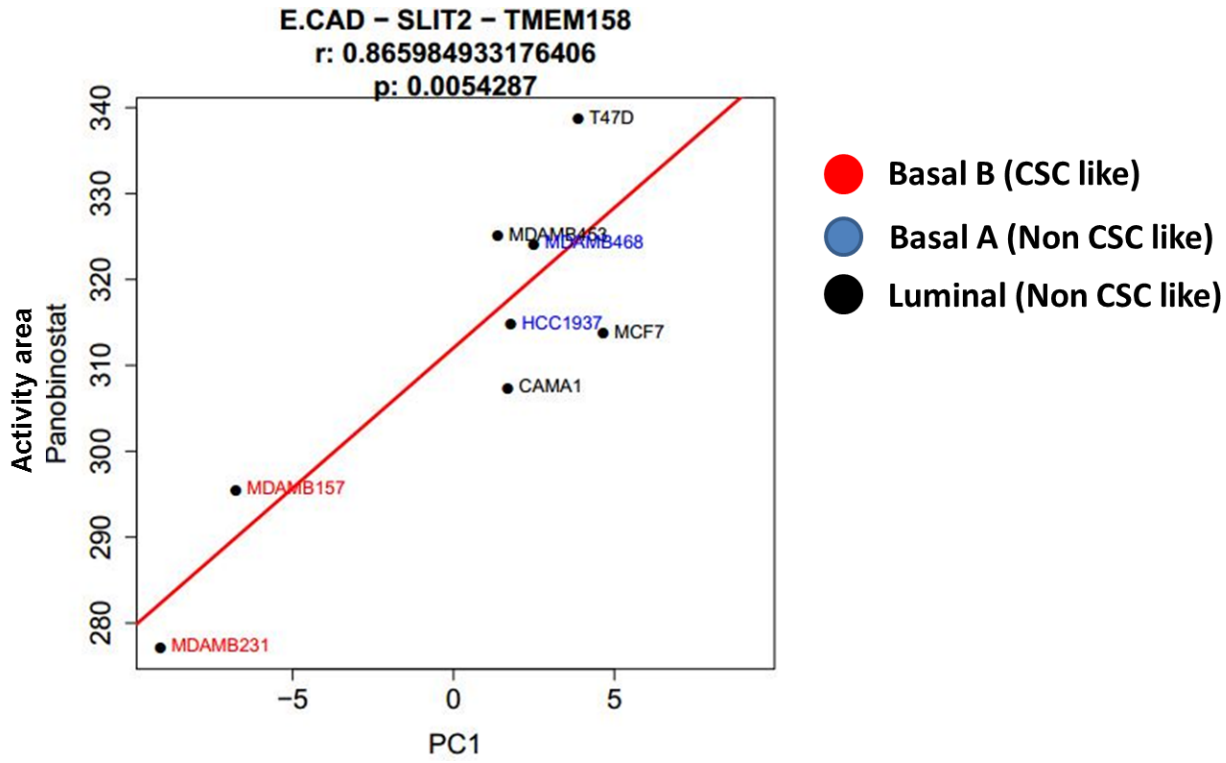



**Figure 5.29: Correlation of PC1 of E-CAD, SLIT2 and TMEM158 with Panobinostat Activity area.** X axis shows PC1 for E-CAD, SLIT2 and TMEM158 and y axis shows Activity area values of Panobinostat obtained against breast cancer cell lines. Cell lines names shown in red color are Basal B (CSC like), in blue color are Basal A (Non CSC like) and in black color are Luminal (Non CSC like). r value and p-value are also mentioned. Percentage of variation explained by PC1 was 0.79.

E-CAD, SLIT2 and TMEM158 combination was significant when analysis with Panobinostat Activity area was performed. So by using these three genes expression, sensitivity of Panobinostat can be predicted. PC1 value for this combination will increase, if expression of E-CAD is decreased and SLIT2 and TMEM158 expression is increased. And increase in PC1 will result into increased sensitivity of cells for Elesclomol.

When PCA analysis was performed correlating with Amax of drugs, 98 significant results were achieved.

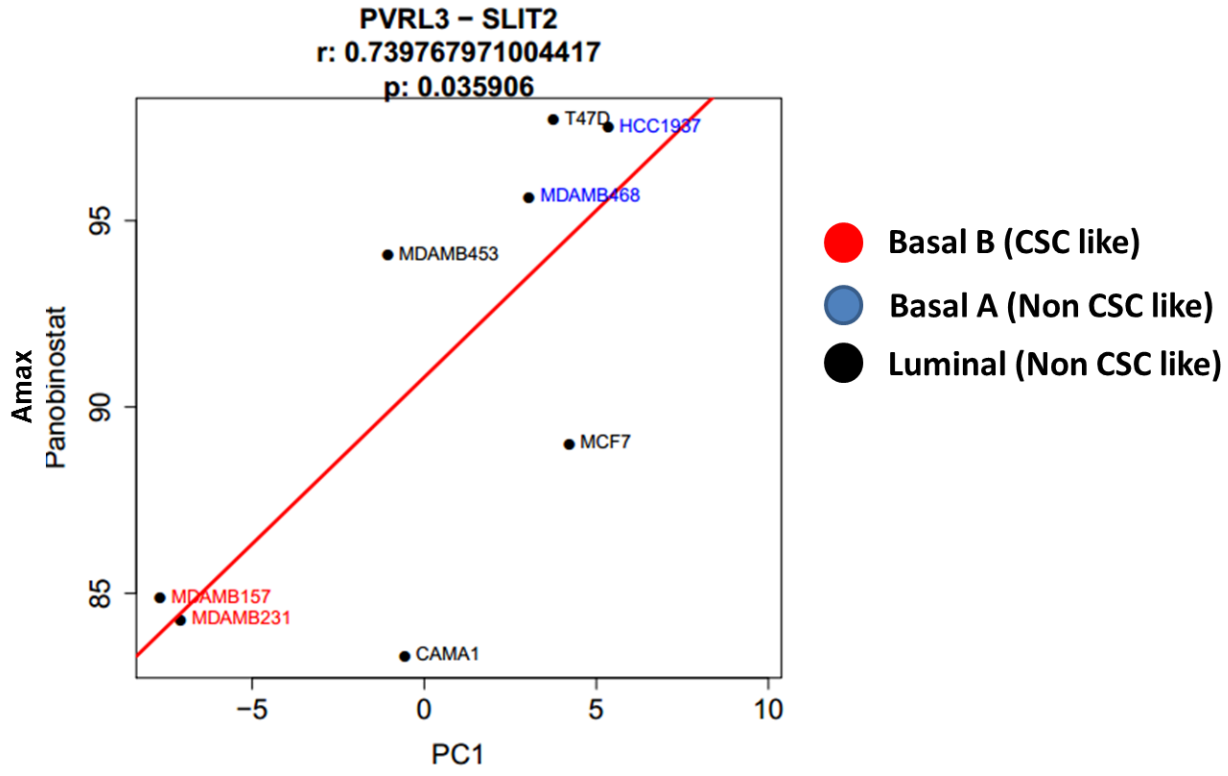



**Figure 5.30: Correlation of PC1 of PVRL3 and SLIT2 with Panobinostat Amax.** X axis shows PC1 for PVRL3 and SLIT2 and y axis shows Amax values of Panobinostat obtained against breast cancer cell lines. Cell lines names shown in red color are Basal B (CSC like), in blue color are Basal A (Non CSC like) and in black color are Luminal (Non CSC like). r value and p-value are also mentioned. Percentage of variation explained by PC1 was 0.89.

PVRL3 and SLIT2 combination was significant when analysis with Panobinostat Amax was performed. So by using these three genes expression, sensitivity of Panobinostat can be predicted. PC1 value for this combination will increase, if expression of SLIT2 and PVRL3 expression is increased. And increase in PC1 will result into increased sensitivity of cells for Elesclomol.

## 5.5 Concordance of Microarray Data with SybrGreen and TaqMan qPCR Data

SybrGreen qPCR data was concordant with microarray data and there was low concordance between TaqMan qPCR data and microarray data. To determine the reason behind these findings, another analysis was done. Motive of this analysis was to check whether primers of SybrGreen and probes for TaqMan hit the same transcript or not. From Ensembl Genome Browser, all sequence of cDNAs for genes were downloaded. And using R based program, cDNA sequences were analyzed for SybrGreen primers, TaqMan probes and microarray Affymetrix probes.

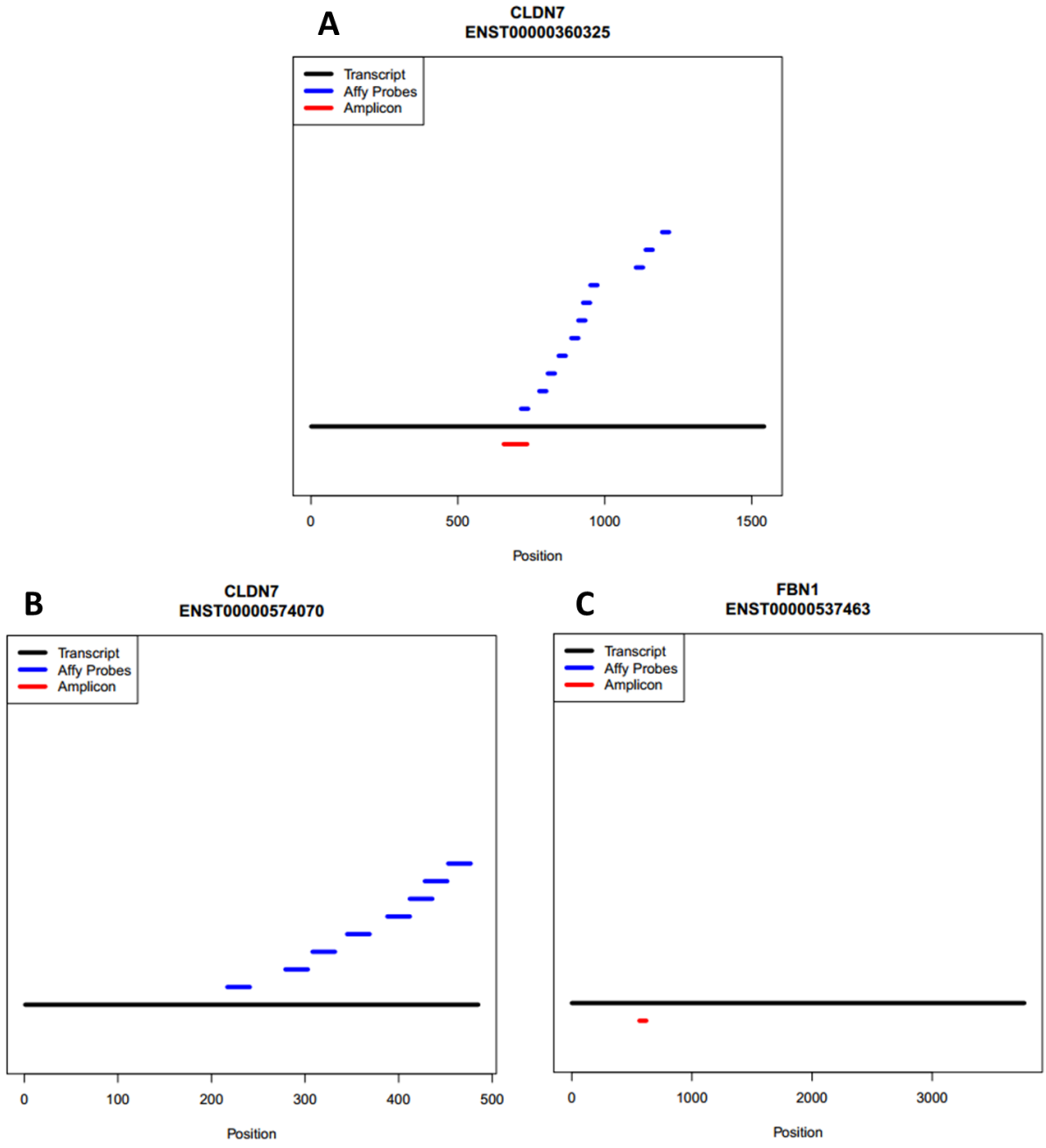



**Figure 5.31: Analysis for explaining concordance between microarray data and TaqMan qPCR data.** In each figure gene name is given at the top. Ensembl ID is given below gene name. Black line shows transcript, Blue lines show Affymetrix probes of microarray and red shows TaqMan probes. (A) TaqMan probes and Affymetrix probes hitting the same transcript of CLDN7 gene, (B) Only Affymetrix probes can detect this transcript of CLDN7 and (C) Only TaqMan probes are detecting this transcript of FBN1. Additionally there were 18 conditions where either Affymetrix probe or TaqMan probe detects cDNA.

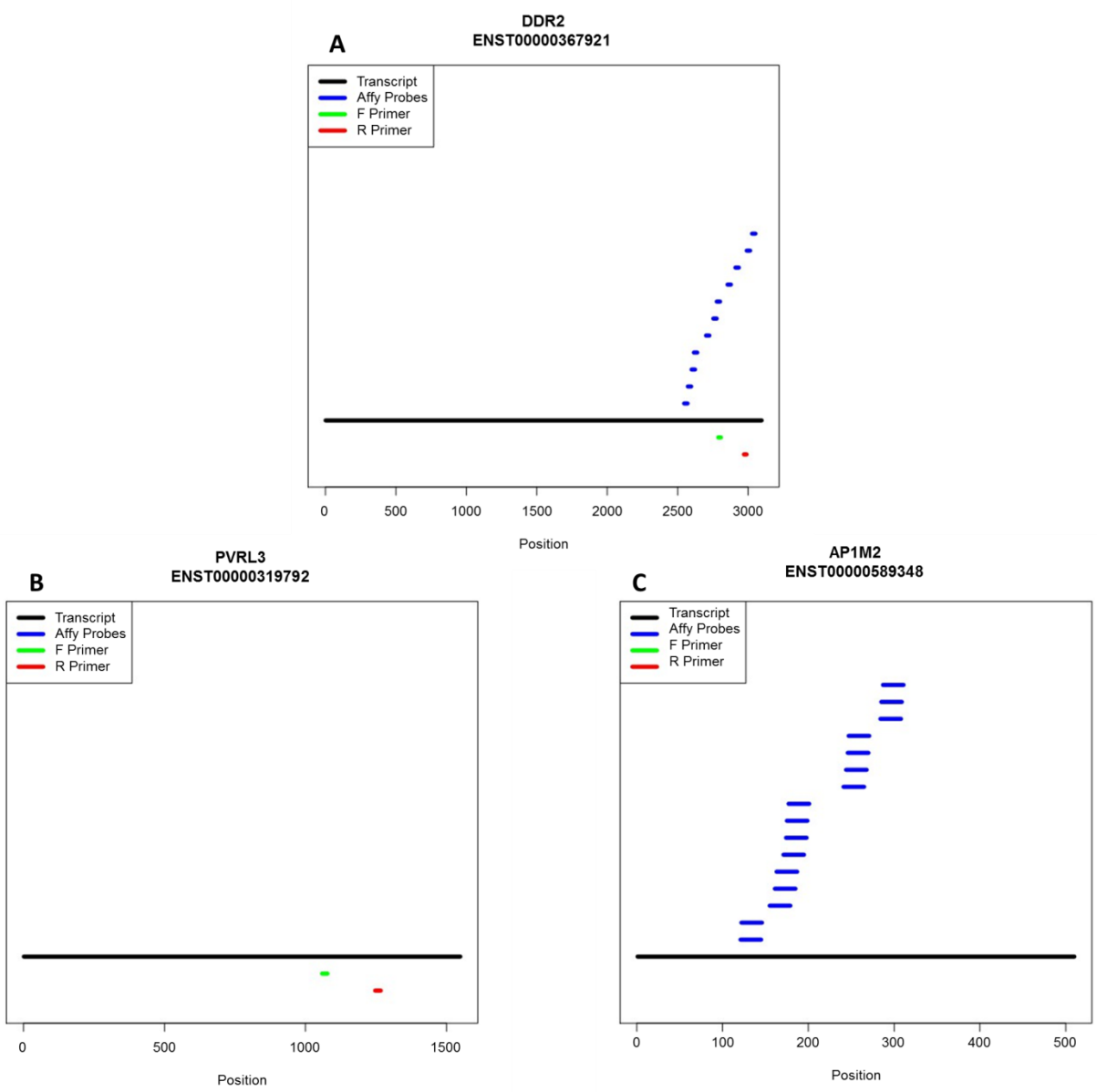



**Figure 5.32: Analysis for explaining concordance between microarray data and SybrGreen qPCR data.** In each figure gene name is given at the top. Ensembl ID is given below gene name. Black line shows transcript, Blue lines show Affymetrix probes of microarray, green line shows forward primer and red shows reverse primer. (A) SybrGreen primers and Affymetrix probes hitting the same transcript of DDR2 gene, (B) Only SybrGreen primers are detecting this transcript of PVRL3 and (C) Only Affymetrix probes can detect this transcript of AP1M2. Additionally B and C are the only conditions where either Affymetrix probe or SybrGreen primers detect cDNA.

**Table 5.16: Concordance analysis of Microarray with TaqMan Probes result**

| Gene | Total transcripts | Transcript detected | | | Not Detected |
|---|---|---|---|---|---|
| | | Only by Affymetrix Probes | Only by TaqMan Probes | By Both | |
| CLDN7 | 8 | 2 | 0 | 5 | 1 |
| EMP3 | 9 | 1 | 0 | 8 | 0 |
| FBN1 | 8 | 1 | 1 | 1 | 5 |
| GNG11 | 1 | 0 | 0 | 1 | 0 |
| GRHL2 | 6 | 0 | 3 | 1 | 2 |
| PVRL3 | 11 | 0 | 4 | 1 | 6 |
| RAB25 | 5 | 0 | 0 | 4 | 1 |
| SPINT1 | 11 | 1 | 6 | 2 | 2 |

**Table 5.17: Concordance analysis of Microarray with SybrGreen result**

| Gene | Total Transcripts | Transcript detected | | | Not Detected |
|---|---|---|---|---|---|
| | | Only by Affymetrix Probes | Only by SybrGreen Primers | By Both | |
| AP1M2 | 10 | 1 | 0 | 2 | 7 |
| BNC2 | 11 | 0 | 0 | 6 | 5 |
| BSPRY | 2 | 0 | 0 | 2 | 0 |
| DDR2 | 6 | 0 | 0 | 2 | 4 |
| DKK3 | 17 | 0 | 0 | 3 | 14 |
| IRF6 | 4 | 0 | 0 | 1 | 3 |
| PVRL3 | 11 | 0 | 1 | 1 | 9 |
| SLIT2 | 11 | 0 | 0 | 4 | 7 |
| ST14 | 4 | 0 | 0 | 1 | 3 |
| TMEM158 | 1 | 0 | 0 | 1 | 0 |
| ZNF165 | 1 | 0 | 0 | 1 | 0 |

As we expected, SybrGreen primers hit the same transcript which was hit by Affymetrix probes with exception of only 2 genes. But for TaqMan probes, 18 different conditions were found where some transcripts were hit by either Affymetrix probe or TaqMan probe separately.

To conclude this project, we predicted two gene lists (Patentable and Non Patentable) which are capable of separating cells with CSC like characteristics from the ones with non CSC like characteristics. From these two gene lists, only Patentable Gene List when validated showed concordance with microarray data. 11 genes were statistically significant and 7 genes showed clear difference between CSC like cells and non CSC like cells. Additionally Midostaurin showed increased activity against CSC like cells and Lapatinib and Panobinostat showed increased activity against non CSC like cells. When CSC like and non CSC like cells were grown in 3D cell culture conditions only MDA-MB-157 showed consistent and predicted change in gene expression. Further this cell line also became more sensitive to Midostaurin and resistant to Panobinostat and Lapatinib. By statistical analysis, various genes and combination of genes were identified as biomarkers for Lapatinib, Panobinostat and Midostaurin. Additionally this was observed that if such primers are designed which target the same transcript, being targeted by microarray probes, then chances for concordant data between qPCR and microarray are high.

So based upon these facts, it can be proposed that by using BSPRY, E-CAD, VIM and FN1, CSC like cells can be separated from non CSC like cells and can be potential biomarkers for Midostaurin, Panobinostat and Lapatinib. Additionally Midostaurin can target CSC like cells while Lapatinib and Panobinostat can target non CSC like cells. And MDA-MB-157 can be a good model to study EMT.

# 6 DISCUSSION & FUTURE PERSPECTIVES

The aim of this study was to cluster cell lines based on characteristics common with CSC. Cell lines having common expression of genes were named as CSC like cell lines and Non CSC like cell lines for vice versa. We came up with two gene lists which can cluster these cell lines into two defined groups. One was named as non-Patentable gene list and other was named as Patentable gene list. And we also predicted four drugs which can target these groups separately.

## 6.1 Non-Patentable Gene List

Non-patentable gene list has 8 genes out of which 4 genes are downregulated (RAB25, GRHL2, SPINT1 and CLDN7) and 4 are upregulated (PVRL3, FBN1, EMP3 and GNG11) in CSC like cells.

RAB25 (Ras-related protein) is among the downregulated ones in CSC like cell lines and is at the top of Rank sum list. In our validation analysis in 8 different datasets, this gene showed same downregulated behavior in CSCs in 6 datasets. It is expressed in breast tissue and ovarian epithelium. In previous studies its expression is related with aggressiveness and progression in epithelial cancers in ovary and breast[49].

GRHL2 (Grainy head-Like 2, Drosophila) is second downregulated gene in this gene list. Its decreased expression is previously related with mesenchymal phenotype. On the other hand, GRHL2 expression inhibits EMT. In breast cancer its expression decreases $CD44^{+}/CD24^{-}$ expressing cells (CSC). Its modest expression is also observed in Luminal A, B and HER2+ve molecular subtypes[50]. In our validation analysis with different datasets, its behavior was same for 5 out of 8 datasets.

SPINT1 (Serine Peptidase Inhibitor, Kunitz Type 1) is also among downregulated genes. Its expression is related with inhibition of hepatocyte growth factor (HGF). HGF activation results in invasion and metastasis of cancer cells. HGF is also responsible for EMT in cancer cells[51, 52]. In our validation analysis with different datasets, its behavior was same for 4 out of 8 datasets..

Last downregulated gene in this gene list is CLDN7 (Claudin 7). Claudin family of genes is responsible for making tight junctions between epithelial cells. Claudin low tumors are reported as highly enriched in CSCs[53-55]. In our validation analysis with different datasets, its behavior was same for 5 out of 8 datasets.

Among 4 genes, reported as upregulated in CSC like cells, PVRL3 (Poliovirus Receptor-Related 3) is at the highest position in Rank sum gene list. Its expression is uniquely identified in human embryonic stem cells[56]. In our validation analysis with different datasets, its behavior was same for 4 out of 8 datasets.

Second upregulated gene in this gene list is FBN1 (Fibrillin 1). Its expression is considered as EMT marker[57]. In our validation analysis with different datasets, its behavior was same for 5 out of 8 datasets.

EMP3 (Epithelial membrane protein 3) is also among the upregulated ones. It plays role in regulation of apoptosis, invasion and differentiation of cancer cells. Additionally it is also associated with EMT induction[58]. In our validation analysis with different datasets, its behavior was same for 6 out of 8 datasets.

GNG11 (Guanine nucleotide binding protein G protein, gamma 11) is the last upregulated gene in this gene list. In one previous study it was used as EMT marker for prostate cancer[59]. In our validation analysis with different datasets, its behavior was same for 4 out of 8 datasets.

Although significant validation could not be achieved by qPCR experiments but based upon the facts described above, this gene list has the potential for discriminating CSC like cells from non CSC like cells. Additionally in qPCR validation analysis, genes showed difference between CSC like cells and non CSC like cells in expected way if MDA-MB-453 is excluded.

### 6.2 Patentable Gene List

Patentable gene list consists of 12 genes, 6 of which are upregulated and 6 are downregulated in CSC like cells.

IRF6 (Interferon Regulatory Factor 6) is one of the downregulated genes. Previously it has been reported that this gene is expressed in normal mammary epithelial cells and its expression is reduced when cells progress to a neoplastic or metastatic state[60]. In our validation analysis with different datasets, its behavior was same for 4 out of 8 datasets.

ST14 (Suppression of Tumorigenicity 14) is also shown as downregulated one in CSC like cells. In one previous study, ST14 expression was upregulated in MCF10A cells when CSCs were compared with non CSCs but when this data was validated then ST14 expression was found as downregulated[33]. In our validation analysis with different datasets, its behavior was same for 3 out of 8 datasets.

BSPRY (B-Box And SPRY Domain Containing protein) is downregulated in CSC like cells. In our validation analysis with different datasets, its behavior was same for 6 out of 8 datasets.

LLGL2 (Lethal Giant Larvae Homolog 2) is downregulated in CSC like cell lines. Previously SNAI1 expression is related with repression of LLGL2 expression[61]. In colorectal cancer, it also has been shown that loss of epithelial markers also results in loss of LLGL2 expression as well[62]. In our validation analysis with different datasets, its behavior was same for 4 out of 8 datasets.

AP1M2 (Adaptor-related Protein Complex 1, mu 2 Subunit) is downregulated in CSC like cells. In our validation analysis with different datasets, its behavior was same for 4 out of 8 datasets.

ZNF165 (Zinc finger protein 165) is downregulated gene in CSC like cell lines. It is previously described as downregulated one migratory breast tumor cells[63]. In our validation analysis with different datasets, its behavior was same for 6 out of 8 datasets.

SLIT2 (Slit Homolog 2) is upregulated one CSC like cells. In contrary to this analysis, SLIT2 has been previously reported as tumor suppressor and its overexpression in MCF7 caused loss in colony formation properties of cells over soft agar[64, 65]. But in our validation analysis with different datasets, its behavior was same for 4 out of 8 datasets.

BNC2 (Basonuclin 2) also came up as upregulated gene in CSC like cell lines. BNC2 is considered as germ cell marker[66]. Additionally its expression is reported to be as significantly higher in basal cell carcinoma as compared to normal skin[67].

DDR2 (Discoidin Domain Receptor Tyrosine) was also found as upregulated gene in CSC like cells. DDR2 expression is reported as responsible for maintenance of SNAIL1 level and activity in tumor cells which have experienced EMT in breast cancer[68]. In our validation analysis with different datasets, its behavior was same for 4 out of 8 datasets.

TMEM (Trans membrane protein 158 (gene/pseudogene) was also found to be upregulated in CSC like cell lines and in our validation analysis, same pattern was seen for 4 out of 8 datasets.

DKK3 (Dickkopf WNT Signaling Pathway Inhibitor 3) is upregulated in CSC like cells. Previously in mouse model of glioma, it has been shown that DKK3 expression is upregulated in primary and secondary CSCs[69]. In our validation analysis with different datasets, its behavior was same for 3 out of 8 datasets.

All the genes for Patentable gene list were validated positively for their expression in 2D culturing conditions. 11 out of 15 genes showed statistical significant distinction. And out of these 11 significant genes, 7 genes showed enormous difference between CSC like cells and non CSC like cells.

### 6.3 Drugs selected for CSC like and non CSC like cell lines

Elesclomol and Midostaurin were selected for CSC like cell lines and Lapatinib and Panobinostat were selected for non CSC like cell lines.

Elesclomol was found to be effective against CSC like cell lines. Elesclomol induces HSP70 expression and targets cancer cells through generation of reactive oxygen species (ROS) causing oxidative stress[70].

In a study, Phase I/II clinical trials against metastatic melanoma, it was reported that Elesclomol alone or in combination with paclitaxel showed promising results[71]. In another study against breast cancer cell lines, terminating cancer cells intensity of Elesclomol was increased when used simultaneously with Paclitaxel or Doxorubicin[72]. Previously it also has been shown that CSCs have lower levels of ROS and reason behind this is that these cells might use redox regulatory machinery to increase their survival[73]. In a study, CSCs were isolated from MCF7 and MDA-MB-231. Cells were grown in 2D and 3D culture as well. On treating these cells with radiation therapy, the cells grown in 2D were more affected as compared to 3D cells[74]. Additionally, same behavior is reported in human breast tumors. CSC lineage enriched cells also contain less level of ROS as compared to normal counterparts and are less prone to radiotherapy[75]. These evidences explain the reason as cells grown in 2D were sensitive to Elesclomol as compared to cells grown in 3D. So based upon these evidences it can be proposed that although Elesclomol affects CSC like cells in 2D cell culture conditions but it is not suitable drug for cells cultured in 3D conditions.

Midostaurin is a protein kinase inhibitor which inhibits protein kinase C (PKC), vascular endothelial growth factor receptor 2 (VEGFR2), platelet derived growth factor (PDGFR), Fms-related tyrosine Kinase (Fit-3) and kit. It was found to be effective against CSC like cell lines in our study. Although Midostaurin is still facing trials to be used for acute myeloid leukemia[76]. But it has been reported previously that PKC level (one of the targets of Midostaurin) is significantly high in CSC enriched cells as compared to non CSCs[77]. As this drug targets PKC which is upregulated in CSCs, so it is proposed that it can work against these cells. Upregulation of PKC also has been previously reported in triple negative breast cancer with poor outcome and less survival rate. And in this study, Safingol which is also a PKC inhibitor was reported as the potential compound to target triple negative breast cancer cells with stem cell like features[78]. Same thing was reported in a recent study in which it was shown that when Midostaurin was used to target MDA-MB-231 cells grown in 3D culture, it decreased their VIM expression[79]. In our project, we proposed Midostaurin as potential drug to target CSC like cells and its cytotoxicity even increases when cells are cultured in 3D conditions which is a better representative of tumor environment.

Lapatinib is dual kinase inhibitor which interferes with HER2 and epidermal growth factor receptor (EGFR) pathways and it came up as drug which affects mostly Luminal and Basal A phenotypes in our analysis. Many combinational therapies have been developed with Lapatinib to target breast cancer which show HER2 overexpression[80]. These evidences also support our results as we in this project showed that Lapatinib can target only cells with non CSC like features.

Panobinostat is histone deacetylase inhibitor (HDAC) which was found to be effective against non CSC like cell lines in our study. HDAC1 is overexpressed in breast, gastric, prostrate and colon cancers[81]. Currently due to its effectiveness this drug is being tested for effectiveness against various cancers. Previously Panobinostat has proven to be effective against triple negative breast cancer cells when transplanted as xenografts in mice[82]. HDAC inhibitors can also induce differentiation in CSCs converting

them into non CSCs[83]. As we concluded from this study that Panobinostat target non CSC like cells more as compared to CSC like cells and its cytotoxicity for CSC like cells decreases when cells are cultured in 3D conditions.

In our validation analysis we found, MDA-MB-157 which is CSC like cell line and when grown in 3D culture conditions, it became more CSC like. Additionally activity of Midostaurin increased against this cell line in 3D cultured cells as compared to 2D cultured cells. Lapatinib and Panobinostat activity against this cell line decreased in 3D cultured cells as compared to 2D cells.

### 6.4 Correlation Analysis

Pearson correlation analysis and Principle component analysis showed many biomarkers for drug cytotoxicity. PCA analysis showed clearly three clusters in graphs showing Luminal, Basal A (both non CSC like) and Basal B (CSC like) separately. Additionally if we use BSPRY, E-CAD, FN1 and VIM for this correlation analysis then these genes can be better markers for Lapatinib, Elesclomol and Panobinostat activity.

### 6.5 Concordance of Microarray Data with qPCR Validation Data

In this study we saw increased concordance between microarray data and SybrGreen data as compared with TaqMan data. The reason for these results was found to be that SybrGreen primers were hitting same transcript which was hit by microarray probes but this was not the case with TaqMan probes as they were targeting other transcripts not targeted by microarray probes. So if we design our primers considering transcripts and microarray probes, then there are high chances for validating microarray data.

More number of cell lines is needed for this project as there were only two cell lines which showed CSC like behavior which can be a reason for some insignificant results. Additionally, cytotoxicity of Midostaurin, Panobinostat and Lapatinib should be again validated in *in vivo* models like mice by injecting CSC and non CSC like cells. And expression of selected genes should also be determined in tumors developed by mice. Moreover these experiments were carried out only on cell lines which were not a true model for breast cancer. So using primary breast cancer cell lines for this study may lead us to a better understanding and true understanding of CSC and non CSC like features.

# 7 REFERENCES


[1] DeSantis C, Ma J, Bryan L, Jemal A: Breast cancer statistics, 2013. CA: a cancer journal for clinicians 2014, 64:52-62.

[2] Siegel R, Ma J, Zou Z, Jemal A: Cancer statistics, 2014. CA: a cancer journal for clinicians 2014, 64:9-29.

[3] Olsson E, Honeth G, Bendahl PO, Saal LH, Gruvberger-Saal S, Ringner M, Vallon-Christersson J, Jonsson G, Holm K, Lovgren K, Ferno M, Grabau D, Borg A, Hegardt C: CD44 isoforms are heterogeneously expressed in breast cancer and correlate with tumor subtypes and cancer stem cell markers. BMC cancer 2011, 11:418.

[4] Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, Thorsen T, Quist H, Matese JC, Brown PO, Botstein D, Lonning PE, Borresen-Dale AL: Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proceedings of the National Academy of Sciences of the United States of America 2001, 98:10869-74.

[5] Prat A, Parker JS, Karginova O, Fan C, Livasy C, Herschkowitz JI, He X, Perou CM: Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. Breast cancer research : BCR 2010, 12:R68.

[6] Clarke MF, Dick JE, Dirks PB, Eaves CJ, Jamieson CH, Jones DL, Visvader J, Weissman IL, Wahl GM: Cancer stem cells—perspectives on current status and future directions: AACR Workshop on cancer stem cells. Cancer research 2006, 66:9339-44.

[7] Velasco-Velázquez MA, Homsi N, De La Fuente M, Pestell RG: Breast cancer stem cells. The international journal of biochemistry & cell biology 2012, 44:573-7.

[8] Croker AK, Goodale D, Chu J, Postenka C, Hedley BD, Hess DA, Allan AL: High aldehyde dehydrogenase and expression of cancer stem cell markers selects for breast cancer cells with enhanced malignant and metastatic ability. Journal of cellular and molecular medicine 2009, 13:2236-52.

[9] Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF: Prospective identification of tumorigenic breast cancer cells. Proceedings of the National Academy of Sciences of the United States of America 2003, 100:3983-8.

[10] Nowell PC: The clonal evolution of tumor cell populations. Science 1976, 194:23-8.

[11] Hill RP: Identifying cancer stem cells in solid tumors: case not proven. Cancer research 2006, 66:1891-6.

[12] Ablett MP, Singh JK, Clarke RB: Stem cells in breast tumours: are they ready for the clinic? European journal of cancer 2012, 48:2104-16.

[13] Creighton CJ, Li X, Landis M, Dixon JM, Neumeister VM, Sjolund A, Rimm DL, Wong H, Rodriguez A, Herschkowitz JI, Fan C, Zhang X, He X, Pavlick A, Gutierrez MC, Renshaw L, Larionov AA, Faratian D, Hilsenbeck SG, Perou CM, Lewis MT, Rosen JM, Chang JC: Residual breast cancers after conventional therapy display mesenchymal as well as tumor-initiating features. Proceedings of the National Academy of Sciences of the United States of America 2009, 106:13820-5.

[14] Ginestier C, Hur MH, Charafe-Jauffret E, Monville F, Dutcher J, Brown M, Jacquemier J, Viens P, Kleer CG, Liu S, Schott A, Hayes D, Birnbaum D, Wicha MS, Dontu G: ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. Cell stem cell 2007, 1:555-67.

[15] Stingl J, Eirew P, Ricketson I, Shackleton M, Vaillant F, Choi D, Li HI, Eaves CJ: Purification and unique properties of mammary epithelial stem cells. Nature 2006, 439:993-7.

[16] Wright MH, Calcagno AM, Salcido CD, Carlson MD, Ambudkar SV, Varticovski L: Brca1 breast tumors contain distinct CD44+/CD24-and CD133+ cells with cancer stem cell characteristics. Breast cancer research : BCR 2008, 10:R10.
[17] Battula VL, Shi Y, Evans KW, Wang R-Y, Spaeth EL, Jacamo RO, Guerra R, Sahin AA, Marini FC, Hortobagyi G: Ganglioside GD2 identifies breast cancer stem cells and promotes tumorigenesis. The Journal of clinical investigation 2012, 122:2066-78.
[18] Kim J, Villadsen R, Sørlie T, Fogh L, Grønlund SZ, Fridriksdottir AJ, Kuhn I, Rank F, Wielenga VT, Solvang H: Tumor initiating but differentiated luminal-like breast cancer cells are highly invasive in the absence of basal-like activity. Proceedings of the National Academy of Sciences 2012, 109:6124-9.
[19] Pece S, Tosoni D, Confalonieri S, Mazzarol G, Vecchi M, Ronzoni S, Bernard L, Viale G, Pelicci PG, Di Fiore PP: Biological and molecular heterogeneity of breast cancers correlates with their cancer stem cell content. Cell 2010, 140:62-73.
[20] Chen D, Bhat-Nakshatri P, Goswami C, Badve S, Nakshatri H: ANTXR1, a Stem Cell-Enriched Functional Biomarker, Connects Collagen Signaling to Cancer Stem-like Cells and Metastasis in Breast Cancer. Cancer research 2013, 73:5821-33.
[21] Ricardo S, Vieira AF, Gerhard R, Leitao D, Pinto R, Cameselle-Teijeiro JF, Milanezi F, Schmitt F, Paredes J: Breast cancer stem cell markers CD44, CD24 and ALDH1: expression distribution within intrinsic molecular subtype. Journal of clinical pathology 2011, 64:937-46.
[22] Takebe N, Warren RQ, Ivy SP: Breast cancer growth and metastasis: interplay between cancer stem cells, embryonic signaling pathways and epithelial-to-mesenchymal transition. Breast cancer research : BCR 2011, 13:211.
[23] Floor S, van Staveren WC, Larsimont D, Dumont JE, Maenhaut C: Cancer cells in epithelial-to-mesenchymal transition and tumor-propagating-cancer stem cells: distinct, overlapping or same populations. Oncogene 2011, 30:4609-21.
[24] Frisch SM, Schaller M, Cieply B: Mechanisms that link the oncogenic epithelial–mesenchymal transition to suppression of anoikis. Journal of cell science 2013, 126:21-9.
[25] Mani SA, Guo W, Liao MJ, Eaton EN, Ayyanan A, Zhou AY, Brooks M, Reinhard F, Zhang CC, Shipitsin M, Campbell LL, Polyak K, Brisken C, Yang J, Weinberg RA: The epithelial-mesenchymal transition generates cells with properties of stem cells. Cell 2008, 133:704-15.
[26] Puisieux A, Brabletz T, Caramel J: Oncogenic roles of EMT-inducing transcription factors. Nature cell biology 2014, 16:488-94.
[27] Asiedu MK, Ingle JN, Behrens MD, Radisky DC, Knutson KL: TGFβ/TNFα-mediated epithelial–mesenchymal transition generates breast cancer stem cells with a claudin-low phenotype. Cancer research 2011, 71:4707-19.
[28] Liu S, Ginestier C, Charafe-Jauffret E, Foco H, Kleer CG, Merajver SD, Dontu G, Wicha MS: BRCA1 regulates human mammary stem/progenitor cell fate. Proceedings of the National Academy of Sciences 2008, 105:1680-5.
[29] Liu S, Ginestier C, Ou SJ, Clouthier SG, Patel SH, Monville F, Korkaya H, Heath A, Dutcher J, Kleer CG: Breast cancer stem cells are regulated by mesenchymal stem cells through cytokine networks. Cancer research 2011, 71:614-24.
[30] Chen L, Xiao Z, Meng Y, Zhao Y, Han J, Su G, Chen B, Dai J: The enhancement of cancer stem cell properties of MCF-7 cells in 3D collagen scaffolds for modeling of cancer and anti-cancer drugs. Biomaterials 2012, 33:1437-44.
[31] Fernando RI, Castillo MD, Litzinger M, Hamilton DH, Palena C: IL-8 signaling plays a critical role in the epithelial–mesenchymal transition of human carcinoma cells. Cancer research 2011, 71:5296-306.
[32] Korkaya H, Kim G-i, Davis A, Malik F, Henry NL, Ithimakin S, Quraishi AA, Tawakkol N, D'Angelo R, Paulson AK: Activation of an IL6 inflammatory loop mediates trastuzumab resistance in HER2+ breast cancer by expanding the cancer stem cell population. Molecular cell 2012, 47:570-84.

[33] Bhat-Nakshatri P, Appaiah H, Ballas C, Pick-Franke P, Goulet R, Jr., Badve S, Srour EF, Nakshatri H: SLUG/SNAI2 and tumor necrosis factor generate breast cells with CD44+/CD24- phenotype. BMC cancer 2010, 10:411.
[34] Liu S, Wicha MS: Targeting breast cancer stem cells. Journal of Clinical Oncology 2010, 28:4006-12.
[35] Chen K, Huang Y-h, Chen J-l: Understanding and targeting cancer stem cells: therapeutic implications and challenges. Acta Pharmacologica Sinica 2013, 34:732-40.
[36] Fillmore CM, Kuperwasser C: Human breast cancer cell lines contain stem-like cells that self-renew, give rise to phenotypically diverse progeny and survive chemotherapy. Breast cancer research : BCR 2008, 10:R25.
[37] Charafe-Jauffret E, Ginestier C, Iovino F, Wicinski J, Cervera N, Finetti P, Hur M-H, Diebel ME, Monville F, Dutcher J: Breast cancer cell lines contain functional cancer stem cells with metastatic capacity and a distinct molecular signature. Cancer research 2009, 69:1302-13.
[38] Vinci M, Gowan S, Boxall F, Patterson L, Zimmermann M, Lomas C, Mendiola M, Hardisson D, Eccles SA: Advances in establishment and analysis of three-dimensional tumor spheroid-based functional assays for target validation and drug evaluation. BMC biology 2012, 10:29.
[39] Valent P, Bonnet D, De Maria R, Lapidot T, Copland M, Melo JV, Chomienne C, Ishikawa F, Schuringa JJ, Stassi G: Cancer stem cell definitions and terminology: the devil is in the details. Nature Reviews Cancer 2012, 12:767-75.
[40] Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, Wilson CJ, Lehár J, Kryukov GV, Sonkin D: The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature 2012, 483:603-7.
[41] Garnett MJ, Edelman EJ, Heidorn SJ, Greenman CD, Dastur A, Lau KW, Greninger P, Thompson IR, Luo X, Soares J: Systematic identification of genomic markers of drug sensitivity in cancer cells. Nature 2012, 483:570-5.
[42] Isbilen M, Mert Senses K, Osmay Gure A: Predicting Chemotherapy Sensitivity Profiles for Breast Cancer Cell Lines with and Without Stem Cell-Like Features. Current Signal Transduction Therapy 2013, 8:268-73.
[43] Dean M, Fojo T, Bates S: Tumour stem cells and drug resistance. Nature Reviews Cancer 2005, 5:275-84.
[44] Team RC: R: A language and environment for statistical computing. 2012.
[45] Simon R, Lam A, Li M-C, Ngan M, Menenzes S, Zhao Y: Analysis of gene expression data using BRB-array tools. Cancer informatics 2007, 3:11.
[46] Liu S, Cong Y, Wang D, Sun Y, Deng L, Liu Y, Martin-Trevino R, Shang L, McDermott SP, Landis MD: Breast Cancer Stem Cells Transition between Epithelial and Mesenchymal States Reflective of their Normal Counterparts. Stem cell reports 2014, 2:78-91.
[47] Gupta PB, Onder TT, Jiang G, Tao K, Kuperwasser C, Weinberg RA, Lander ES: Identification of selective inhibitors of cancer stem cells by high-throughput screening. Cell 2009, 138:645-59.
[48] Calcagno AM, Salcido CD, Gillet J-P, Wu C-P, Fostel JM, Mumau MD, Gottesman MM, Varticovski L, Ambudkar SV: Prolonged drug selection of breast cancer cells and enrichment of cancer stem cell characteristics. Journal of the National Cancer Institute 2010, 102:1637-52.
[49] Cheng KW, Lahad JP, Kuo W-l, Lapuk A, Yamada K, Auersperg N, Liu J, Smith-McCune K, Lu KH, Fishman D: The RAB25 small GTPase determines aggressiveness of ovarian and breast cancers. Nature medicine 2004, 10:1251-6.
[50] Cieply B, Riley P, Pifer PM, Widmeyer J, Addison JB, Ivanov AV, Denvir J, Frisch SM: Suppression of the epithelial–mesenchymal transition by Grainyhead-like-2. Cancer research 2012, 72:2440-53.
[51] Gherardi E, Birchmeier W, Birchmeier C, Woude GV: Targeting MET in cancer: rationale and progress. Nature Reviews Cancer 2012, 12:89-103.

[52] Morel A-P, Lièvre M, Thomas C, Hinkal G, Ansieau S, Puisieux A: Generation of breast cancer stem cells through epithelial-mesenchymal transition. PloS one 2008, 3:e2888.
[53] Creighton CJ, Li X, Landis M, Dixon JM, Neumeister VM, Sjolund A, Rimm DL, Wong H, Rodriguez A, Herschkowitz JI: Residual breast cancers after conventional therapy display mesenchymal as well as tumor-initiating features. Proceedings of the National Academy of Sciences 2009, 106:13820-5.
[54] Park SY, Lee HE, Li H, Shipitsin M, Gelman R, Polyak K: Heterogeneity for stem cell–related markers according to tumor subtype and histologic stage in breast cancer. Clinical Cancer Research 2010, 16:876-87.
[55] Katz E, Dubois-Marshall S, Sims AH, Gautier P, Caldwell H, Meehan RR, Harrison DJ: An in vitro model that recapitulates the epithelial to mesenchymal transition (EMT) in human breast cancer. PloS one 2011, 6:e17083.
[56] Binato R, de Souza Fernandez T, Lazzarotto-Silva C, Du Rocher B, Mencalha A, Pizzatti L, Bouzas L, Abdelhay E: Stability of human mesenchymal stem cells during in vitro culture: considerations for cell therapy. Cell proliferation 2013, 46:10-22.
[57] Lehmann BD, Bauer JA, Chen X, Sanders ME, Chakravarthy AB, Shyr Y, Pietenpol JA: Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. The Journal of clinical investigation 2011, 121:2750.
[58] Blick T, Hugo H, Widodo E, Waltham M, Pinto C, Mani SA, Weinberg RA, Neve RM, Lenburg ME, Thompson EW: Epithelial mesenchymal transition traits in human breast cancer cell lines parallel the CD44HI/CD24lO/-stem cell phenotype in human breast cancer. Journal of mammary gland biology and neoplasia 2010, 15:235-52.
[59] Albino D, Longoni N, Curti L, Mello-Grand M, Pinton S, Civenni G, Thalmann G, D'Ambrosio G, Sarti M, Sessa F: ESE3/EHF controls epithelial cell differentiation and its loss leads to prostate tumors with mesenchymal and stem-like features. Cancer research 2012, 72:2889-900.
[60] Bailey CM, Khalkhali-Ellis Z, Kondo S, Margaryan NV, Seftor RE, Wheaton WW, Amir S, Pins MR, Schutte BC, Hendrix MJ: Mammary Serine Protease Inhibitor (Maspin) Binds Directly to Interferon Regulatory Factor 6 IDENTIFICATION OF A NOVEL SERPIN PARTNERSHIP. Journal of Biological Chemistry 2005, 280:34210-7.
[61] Moreno-Bueno G, Portillo F, Cano A: Transcriptional regulation of cell polarity in EMT and cancer. Oncogene 2008, 27:6958-69.
[62] Spaderna S, Schmalhofer O, Wahlbuhl M, Dimmler A, Bauer K, Sultan A, Hlubek F, Jung A, Strand D, Eger A: The transcriptional repressor ZEB1 promotes metastasis and loss of cell polarity in cancer. Cancer research 2008, 68:537-44.
[63] Patsialou A, Wang Y, Lin J, Whitney K, Goswami S, Kenny PA, Condeelis JS: Selective gene-expression profiling of migratory tumor cells in vivo predicts clinical outcome in breast cancer patients. Breast cancer research : BCR 2012, 14:R139.
[64] Prasad A, Paruchuri V, Preet A, Latif F, Ganju RK: Slit-2 induces a tumor-suppressive effect by regulating β-catenin in breast cancer cells. Journal of Biological Chemistry 2008, 283:26624-33.
[65] Dallol A, Da Silva NF, Viacava P, Minna JD, Bieche I, Maher ER, Latif F: SLIT2, a human homologue of the Drosophila Slit2 gene, has tumor suppressor activity and is frequently inactivated in lung and breast cancers. Cancer research 2002, 62:5874-80.
[66] Anisimov SV, Christophersen NS, Correia AS, Li J-Y, Brundin P: " NeuroStem Chip": a novel highly specialized tool to study neural differentiation pathways in human stem cells. BMC genomics 2007, 8:1-15.
[67] O'Driscoll L, McMorrow J, Doolan P, McKiernan E, Mehta JP, Ryan E, Gammell P, Joyce H, O'Donovan N, Walsh N: Investigation of the molecular profile of basal cell carcinoma using whole genome microarrays. Molecular cancer 2006, 5:74.

[68] Zhang K, Corsa CA, Ponik SM, Prior JL, Piwnica-Worms D, Eliceiri KW, Keely PJ, Longmore GD: The collagen receptor discoidin domain receptor 2 stabilizes SNAIL1 to facilitate breast cancer metastasis. Nature cell biology 2013, 15:677-87.
[69] Harris MA, Yang H, Low BE, Mukherje J, Guha A, Bronson RT, Shultz LD, Israel MA, Yun K: Cancer stem cells are enriched in the side population cells in a mouse model of glioma. Cancer research 2008, 68:10051-9.
[70] Kirshner JR, He S, Balasubramanyam V, Kepros J, Yang C-Y, Zhang M, Du Z, Barsoum J, Bertin J: Elesclomol induces cancer cell apoptosis through oxidative stress. Molecular cancer therapeutics 2008, 7:2319-27.
[71] O'Day S, Gonzalez R, Lawson D, Weber R, Hutchins L, Anderson C, Haddad J, Kong S, Williams A, Jacobson E: Phase II, randomized, controlled, double-blinded trial of weekly elesclomol plus paclitaxel versus paclitaxel alone for stage IV metastatic melanoma. Journal of Clinical Oncology 2009, 27:5452-8.
[72] Qu Y, Wang J, Sim M-S, Liu B, Giuliano A, Barsoum J, Cui X: Elesclomol, counteracted by Akt survival signaling, enhances the apoptotic effect of chemotherapy drugs in breast cancer cells. Breast cancer research and treatment 2010, 121:311-21.
[73] Trachootham D, Alexandre J, Huang P: Targeting cancer cells by ROS-mediated mechanisms: a radical therapeutic approach? Nature reviews Drug discovery 2009, 8:579-91.
[74] Kundu N, Ma XR, Kochel T, Goloubeva O, Staats P, Thompson K, Martin S, Reader J, Take Y, Collin P, Fulton A: Prostaglandin E receptor EP4 is a therapeutic target in breast cancer cells with stem-like properties. Breast cancer research and treatment 2014, 143:19-31.
[75] Diehn M, Cho RW, Lobo NA, Kalisky T, Dorie MJ, Kulp AN, Qian D, Lam JS, Ailles LE, Wong M: Association of reactive oxygen species levels and radioresistance in cancer stem cells. Nature 2009, 458:780-3.
[76] Stone R, Fischer T, Paquette R, Schiller G, Schiffer C, Ehninger G, Cortes J, Kantarjian H, Deangelo D, Huntsman-Labed A: Phase IB study of the FLT3 kinase inhibitor midostaurin with chemotherapy in younger newly diagnosed adult patients with acute myeloid leukemia. Leukemia 2012, 26:2061-8.
[77] Scheel C, Eaton EN, Li SH-J, Chaffer CL, Reinhardt F, Kah K-J, Bell G, Guo W, Rubin J, Richardson AL: Paracrine and autocrine signals induce and maintain mesenchymal and stem cell states in the breast. Cell 2011, 145:926-40.
[78] Hsu Y-H, Yao J, Chan L-C, Wu T-J, Hsu JL, Fang Y-F, Wei Y, Wu Y, Huang W-C, Liu C-L: Definition of PKC-α, CDK6, and MET as Therapeutic Targets in Triple-Negative Breast Cancer. Cancer research 2014.
[79] Li Q, Chen C, Kapadia A, Zhou Q, Harper MK, Schaack J, Labarbera DV: 3D Models of Epithelial-Mesenchymal Transition in Breast Cancer Metastasis High-Throughput Screening Assay Development, Validation, and Pilot Screen. Journal of biomolecular screening 2011, 16:141-54.
[80] Sahin O, Wang Q, Brady SW, Ellis K, Wang H, Chang C-C, Zhang Q, Priya P, Zhu R, Wong ST: Biomarker-guided sequential targeted therapies to overcome therapy resistance in rapidly evolving highly aggressive mammary tumors. Cell research 2014, 24:542-59.
[81] Atadja P: Development of the pan-DAC inhibitor panobinostat (LBH589): successes and challenges. Cancer letters 2009, 280:233-41.
[82] Rao R, Nalluri S, Kolhe R, Yang Y, Fiskus W, Chen J, Ha K, Buckley KM, Balusu R, Coothankandaswamy V: Treatment with panobinostat induces glucose-regulated protein 78 acetylation and endoplasmic reticulum stress in breast cancer cells. Molecular cancer therapeutics 2010, 9:942-52.
[83] Singh A, Settleman J: EMT, cancer stem cells and drug resistance: an emerging axis of evil in the war on cancer. Oncogene 2010, 29:4741-51.

# 8 APPENDIX

## 8.1 Validation of Drug Response in 2D Cultured Cells

Dose response curves of selected drugs for breast cancer cell lines.

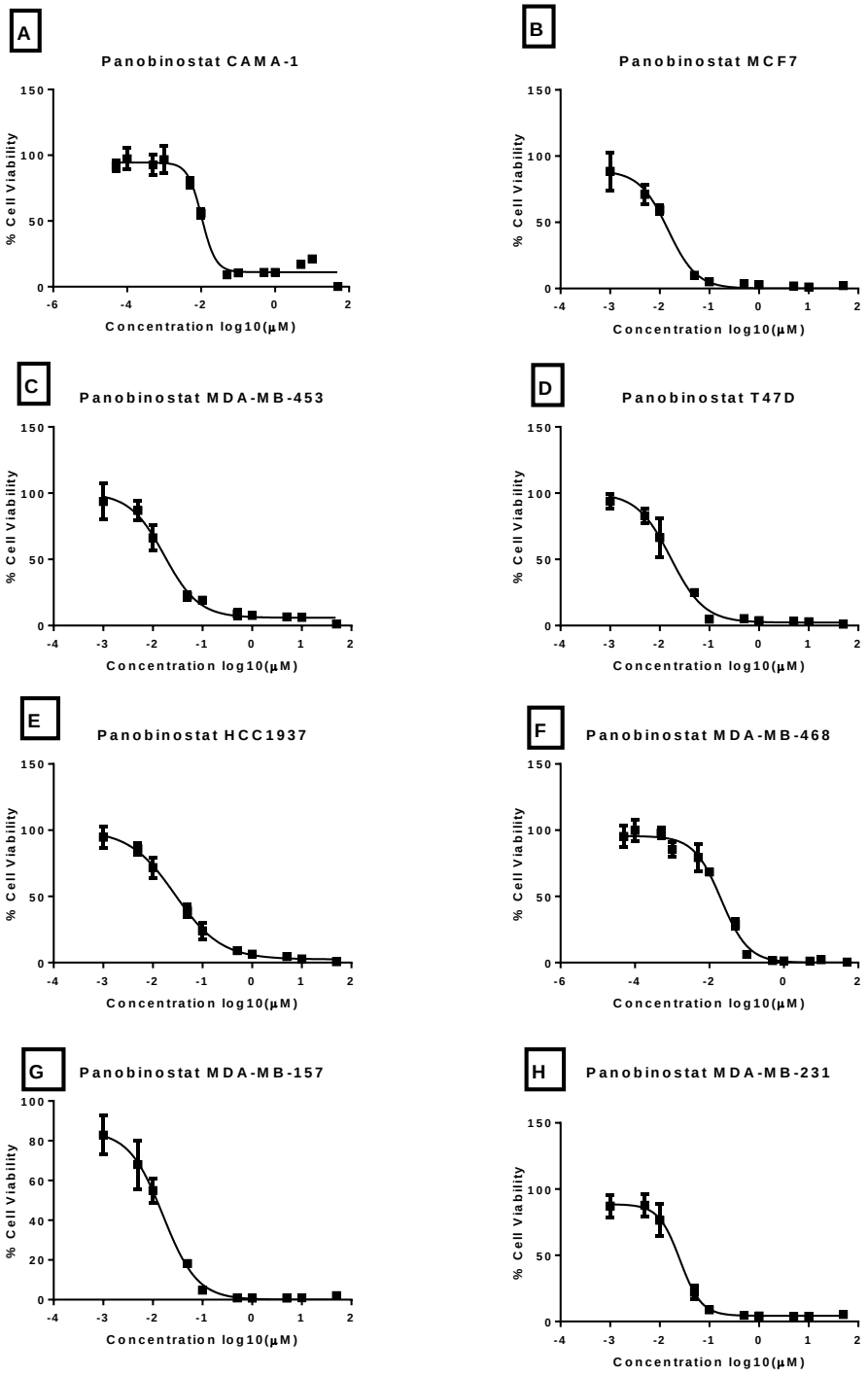


**Figure 8.1: Percent Cell Viability curves for Panobinostat against breast cancer cell lines.** X-axis shows log concentration of drugs in µM. And y-axis shows percent Cell Viability of cells by drug. Error bars show mean with 95% confidence interval. Drug was used in different concentration with highest concentration as 50 µM.

While treating cells with Panobinostat, MCF7, MDA-MB-453, T47D, HCC1937, MDA-MB-157 and MDA-MB-231 were treated with 10 concentrations as 50, 10, 5, 1, 0.5, 0.1, 0.05, 0.01, 0.005 and 0.00100µM. CAMA-1 and MDA-MB-468 was treated with 13 concentrations as 50, 10, 5, 1, 0.5, 0.1, 0.05, 0.01, 0.005 and 0.00100, 0.00050, 0.00010 and 0.00005µM.Whenever 100% cell viability was not seen so more concentrations were added to the experiments.

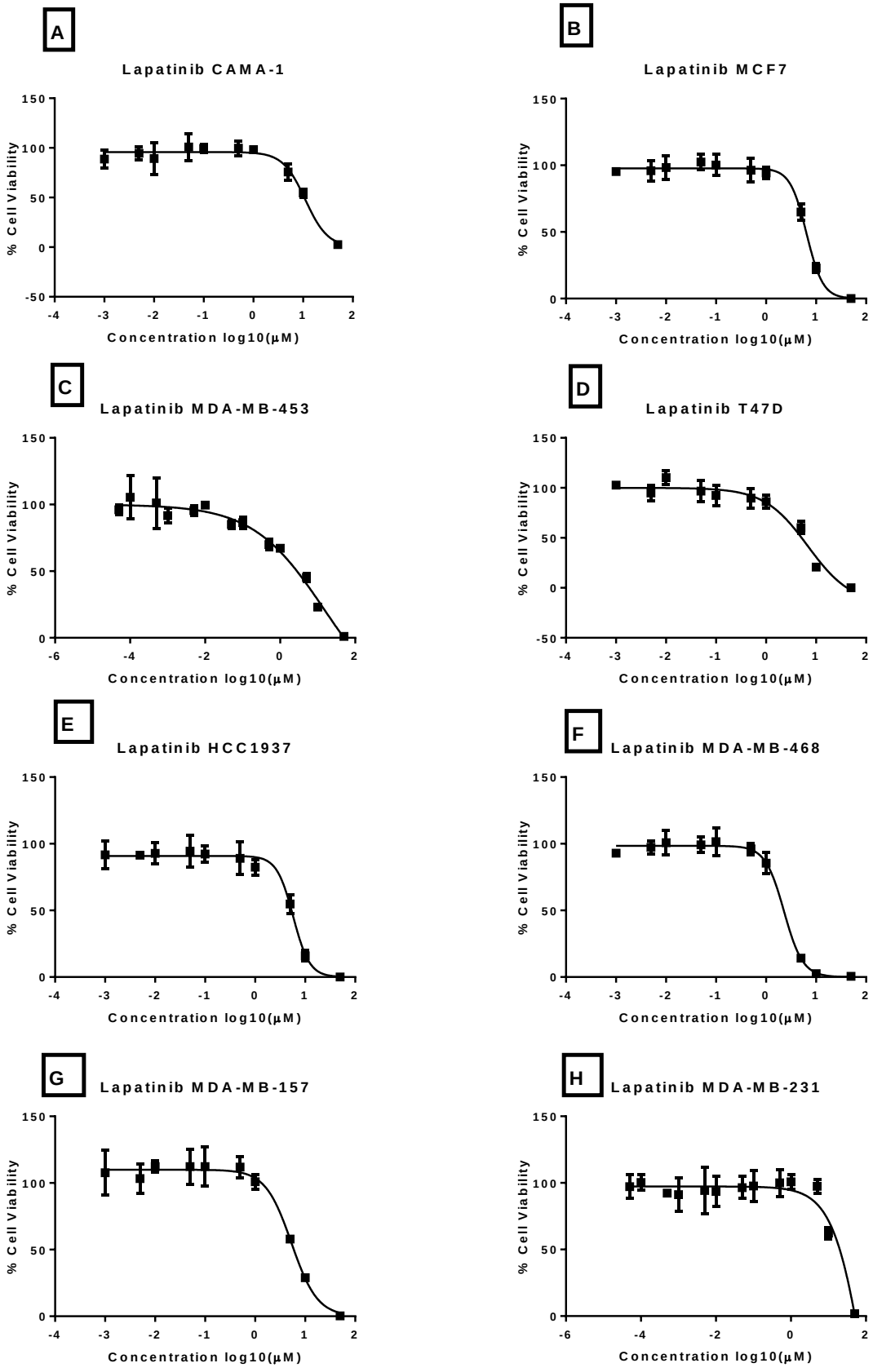

A
Lapatinib CAMA-1
% Cell Viability
Concentration log10(μM)
B
Lapatinib MCF7
% Cell Viability
Concentration log10(μM)
C
Lapatinib MDA-MB-453
% Cell Viability
Concentration log10(μM)
D
Lapatinib T47D
% Cell Viability
Concentration log10(μM)
E
Lapatinib HCC1937
% Cell Viability
Concentration log10(μM)
F
Lapatinib MDA-MB-468
% Cell Viability
Concentration log10(μM)
G
Lapatinib MDA-MB-157
% Cell Viability
Concentration log10(μM)
H
Lapatinib MDA-MB-231
% Cell Viability
Concentration log10(μM)

**Figure 8.2: Percent Cell Viability curves for Lapatinib against breast cancer cell lines.** X-axis shows log concentration of drugs in µM. And y-axis shows percent Cell Viability of cells by drug. Error bars show mean with 95% confidence interval. Drug was used in different concentration with highest concentration as 50 µM.

While treating cells with Lapatinib, CAMA-1, MCF7, MDA-MB-468, T47D, HCC1937 and MDA-MB-157 were treated with 10 concentrations as 50, 10, 5, 1, 0.5, 0.1, 0.05, 0.01, 0.005 and 0.00100µM. MDA-MB-231 and MDA-MB-453 was treated with 13 concentrations as 50, 10, 5, 1, 0.5, 0.1, 0.05, 0.01, 0.005 and 0.00100, 0.00050, 0.00010 and 0.00005µM. whenever 100% cell viability was not seen so more concentrations were added to experiments.

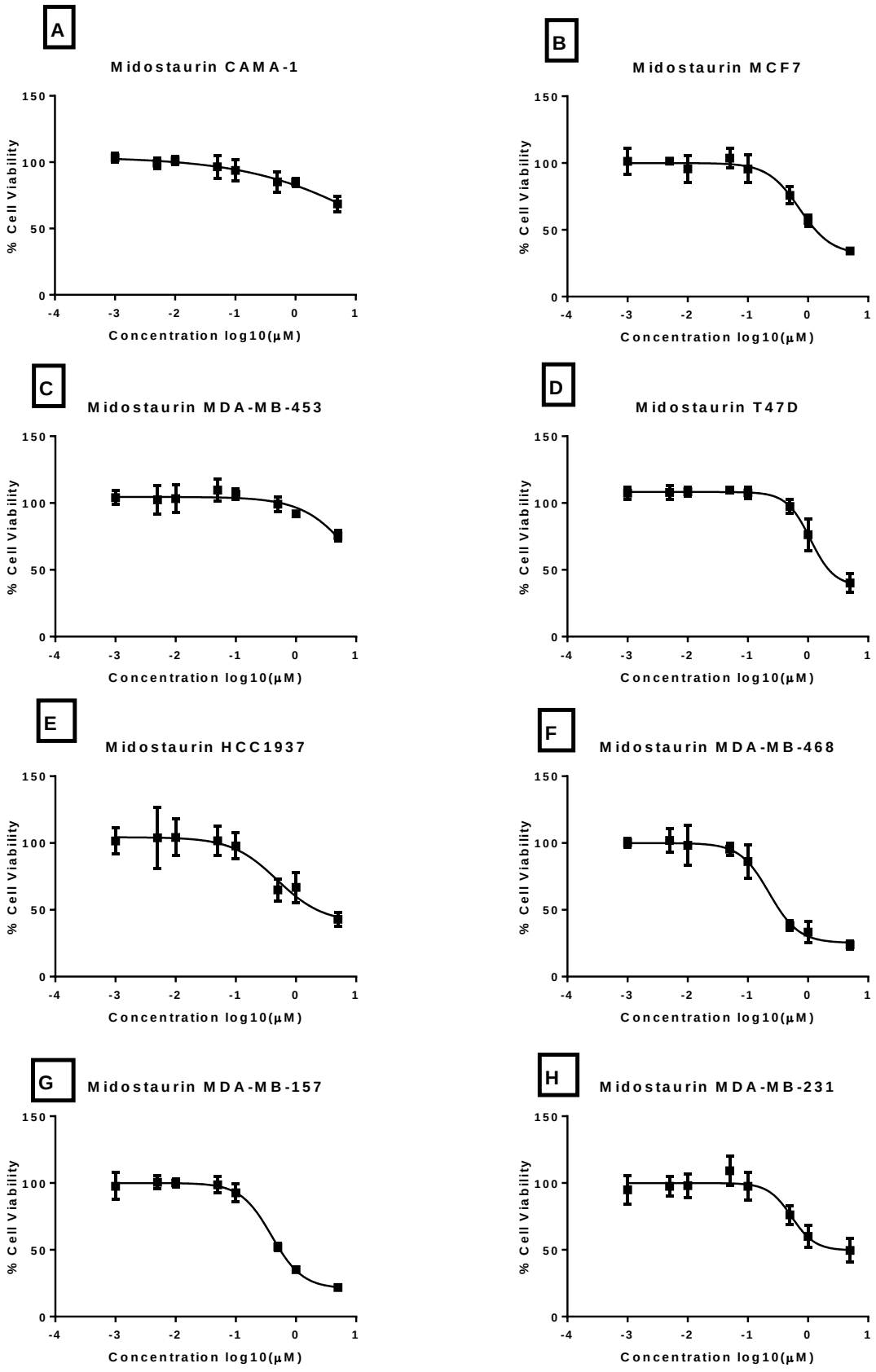

A
Midostaurin CAMA-1
B
Midostaurin MCF7
C
Midostaurin MDA-MB-453
D
Midostaurin T47D
E
Midostaurin HCC1937
F
Midostaurin MDA-MB-468
G
Midostaurin MDA-MB-157
H
Midostaurin MDA-MB-231
% Cell Viability
Concentration log10(μM)
150
100
50
0
-4
-3
-2
-1
0
1

**Figure 8.3: Percent Cell Viability curves for Midostaurin against breast cancer cell lines.** X-axis shows log concentration of drugs in µM. And y-axis shows percent Cell Viability of cells by drug. Error bars show mean with 95% confidence interval. Drug was used in different concentration with highest concentration as 5 µM. Drug concentrations were 5, 1, 0.5, 0.1, 0.05, 0.01, 0.005 and 0.00100µM.

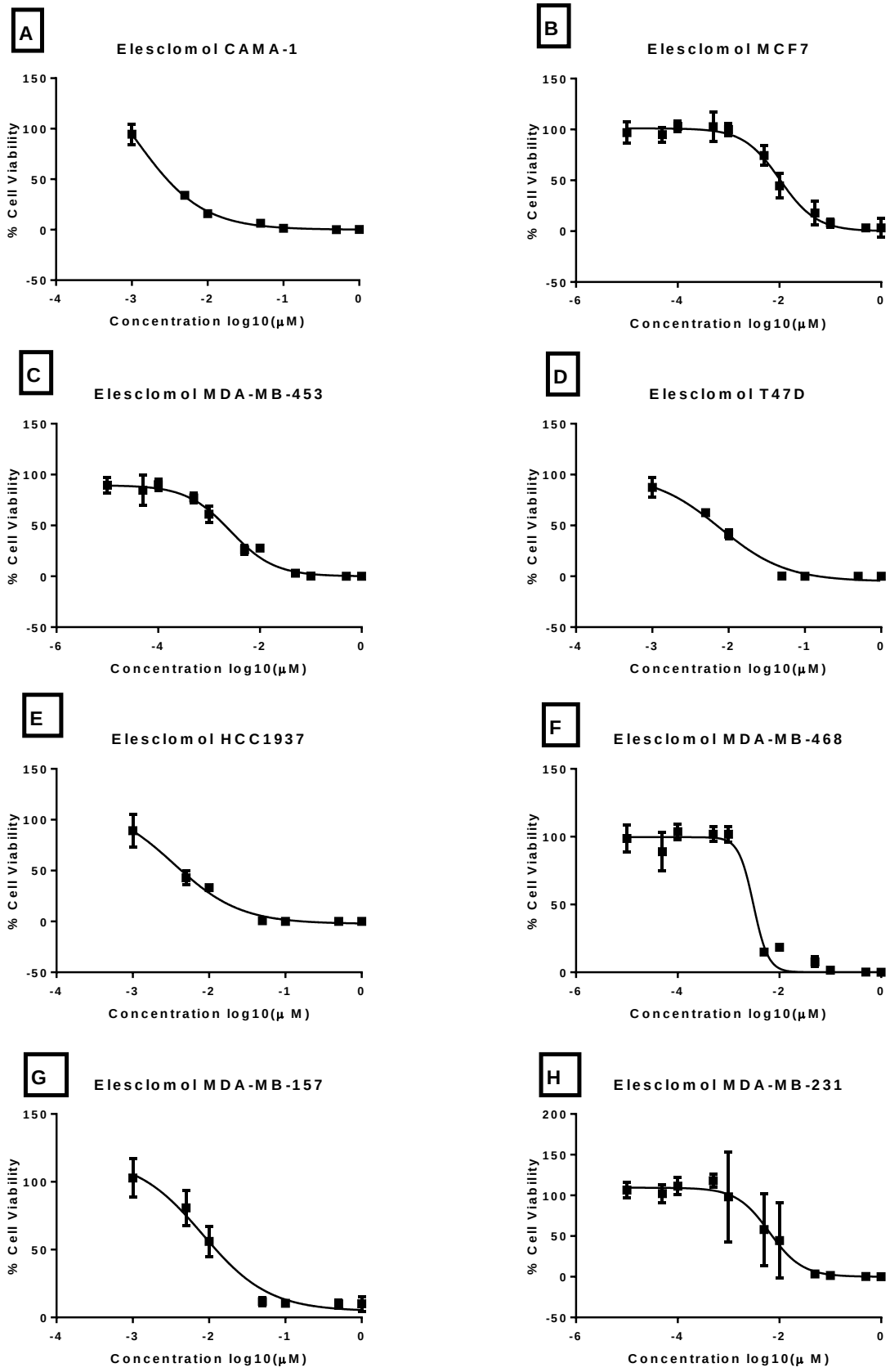

A
Elesclomol CAMA-1
% Cell Viability
Concentration log10(μM)
B
Elesclomol MCF7
% Cell Viability
Concentration log10(μM)
C
Elesclomol MDA-MB-453
% Cell Viability
Concentration log10(μM)
D
Elesclomol T47D
% Cell Viability
Concentration log10(μM)
E
Elesclomol HCC1937
% Cell Viability
Concentration log10(μ M)
F
Elesclomol MDA-MB-468
% Cell Viability
Concentration log10(μM)
G
Elesclomol MDA-MB-157
% Cell Viability
Concentration log10(μM)
H
Elesclomol MDA-MB-231
% Cell Viability
Concentration log10(μ M)

**Figure 8.4: Percent Cell Viability curves for Elesclomol against breast cancer cell lines.** X-axis shows log concentration of drugs in µM. And y-axis shows percent Cell Viability of cells by drug. Error bars show mean with 95% confidence interval. Drug was used in different concentration with highest concentration as 1 µM.

While treating cells with Elesclomol, CAMA-1, T47D, HCC1937 and MDA-MB-157 were treated with 7 concentrations as 1, 0.5, 0.1, 0.05, 0.01, 0.005 and 0.00100µM. For MCF7, MDA-MB-231, MDA-MB-468 and MDA-MB-453 was treated with 11 concentrations as 1, 0.5, 0.1, 0.05, 0.01, 0.005, 0.00100, 0.00050, 0.00010, 0.00005 and 0.000010µM. Whenever 100% cell viability was not seen so more concentrations were added to experiments.

For each cell line IC50, EC50, Activity area and Amax was calculated.

## 8.2 Comparison of Gene Expression between 2D and 3D Culture Conditions of Non Patentable Gene List

Then gene expression was compared between 2D and 3D samples of Non Patentable genes. For determining gene expression for 3D cultured cells, samples were used from $6^{th}$ passage for MDA-MB-157 & MDA-MB-231, from $9^{th}$ passage for CAMA-1 and from $12^{th}$ passage for MDA-MB-453, MDA-MB-468, MCF7 & T47D.

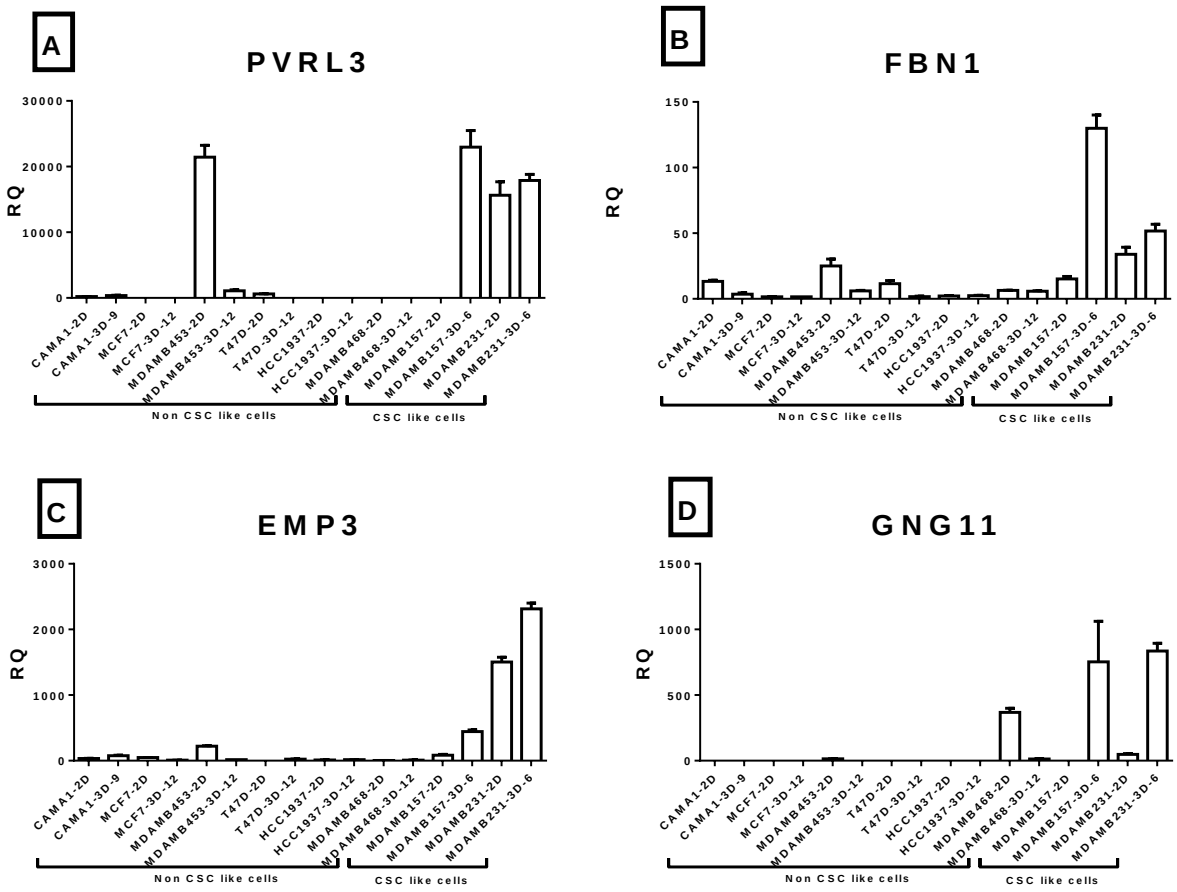



**Figure 8.5: Comparison of (A)PVRL3,(B) FBN1, (C)EMP3 and (D) GNG11 in both 2D and 3D.** 2D and 3D mentioned after name shows cell culture conditions. Passage number for each sample, are given for 3D after sample name. Error bars show standard error of mean.

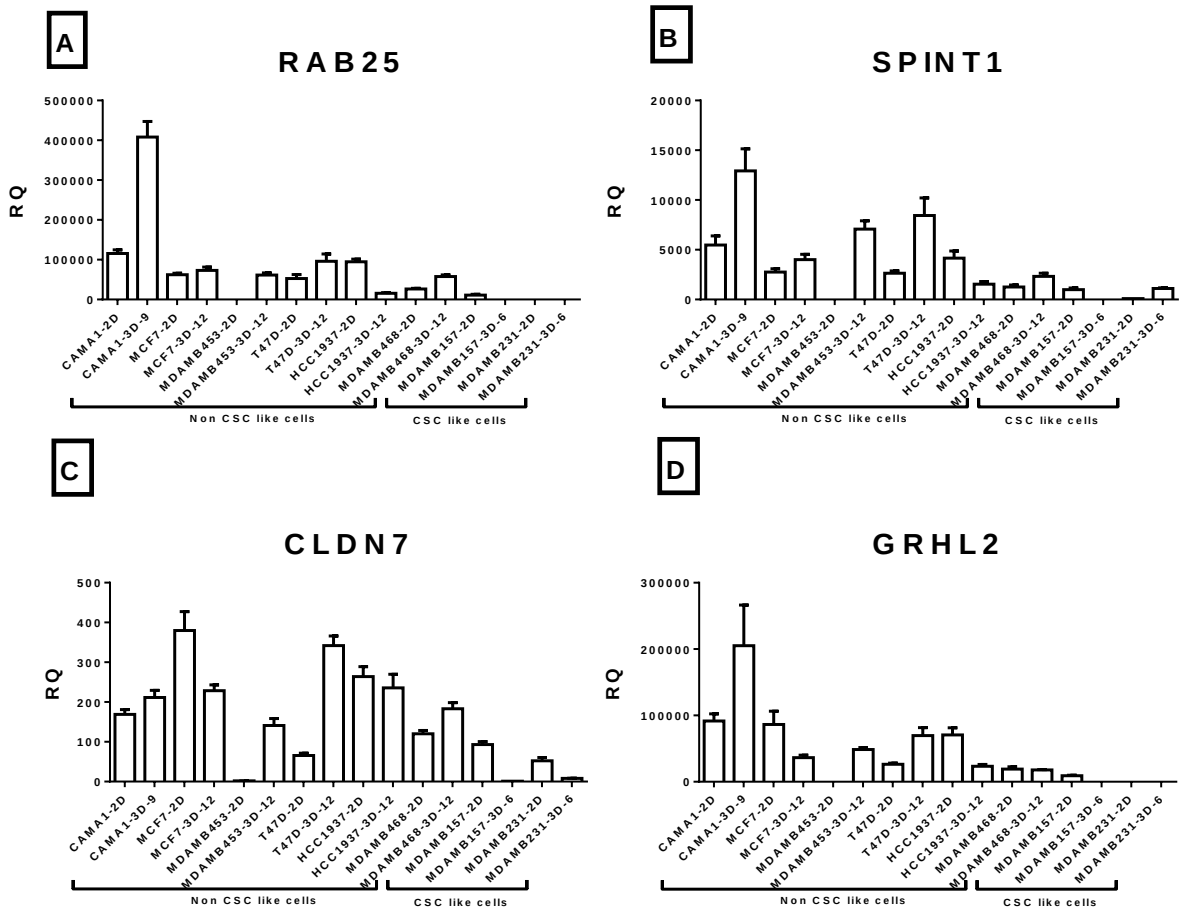
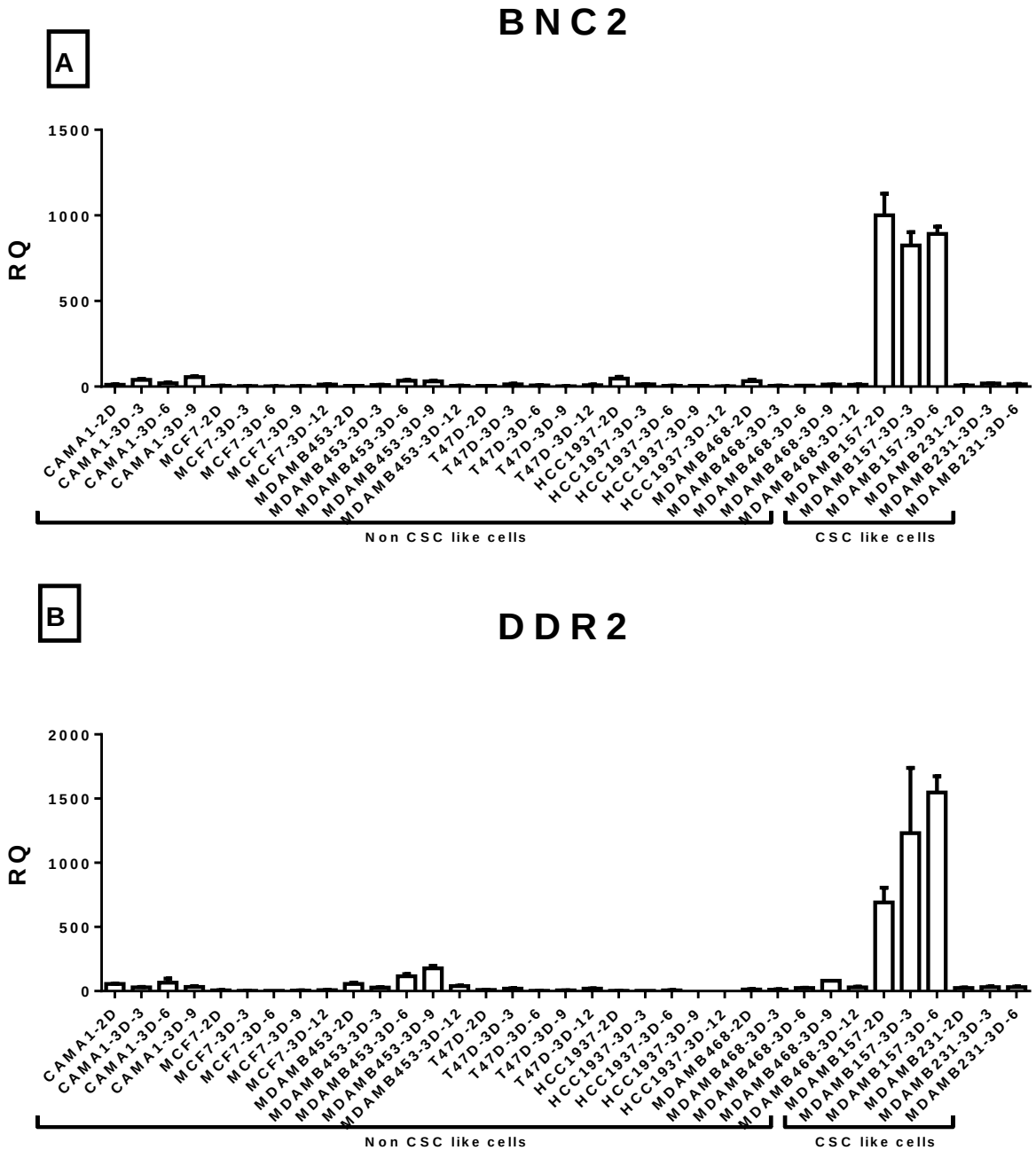


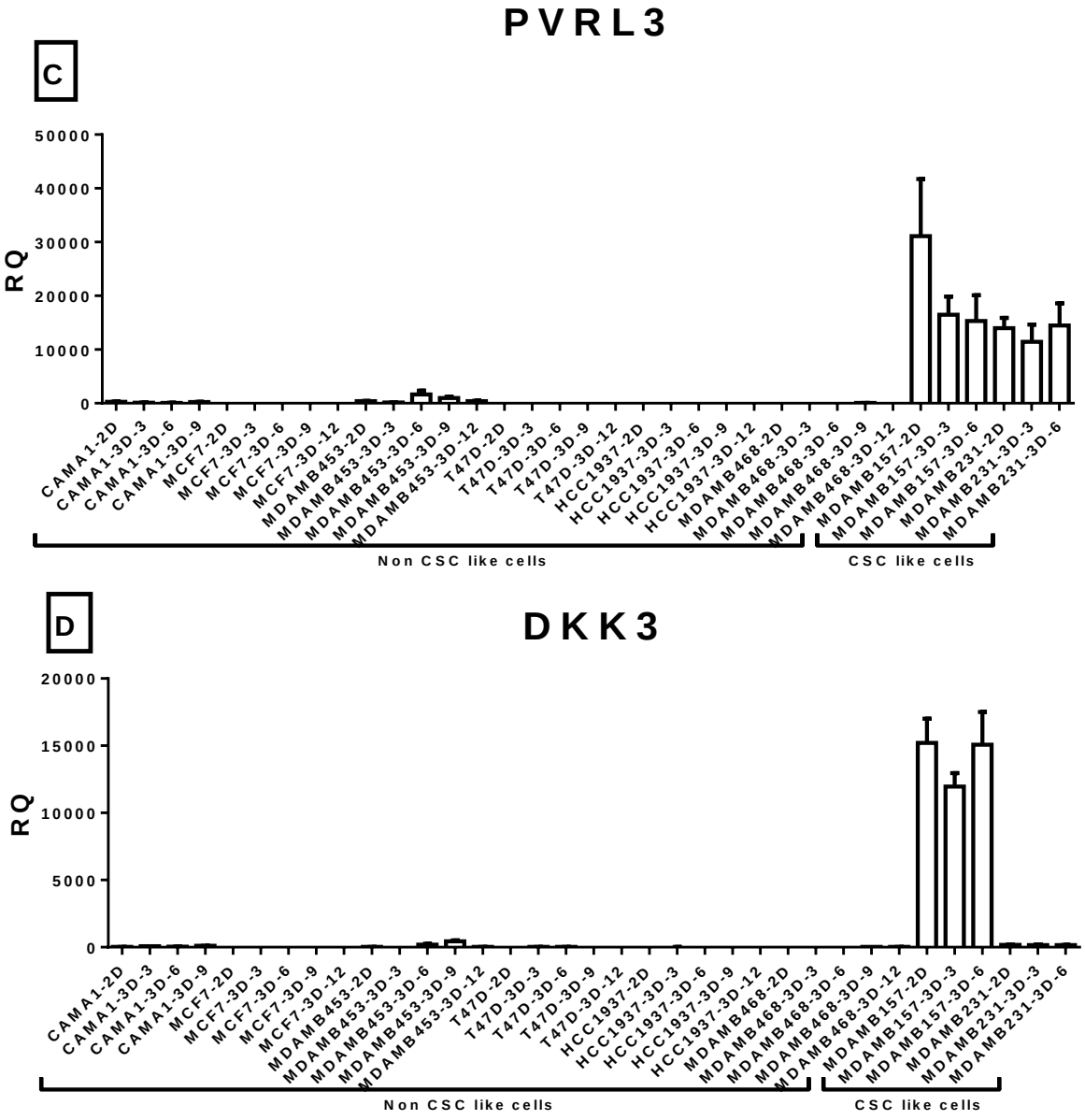

**Figure 8.6: Comparison of (A)RAB25, (B) SPINT1, (C) CLDN7 and (D) GRHL2 in both 2D and 3D.** 2D and 3D mentioned after name shows cell culture conditions. Passage number for each sample, are given for 3D after sample name. Error bars show standard error of mean.

Out of all cell lines, only HCC1937 and MDA-MB-157 showed changed gene expression in sense of CSC gene expression.

## 8.3 Comparison of Gene Expression between 2D and 3D Culture Conditions of Non Patentable Gene List

Gene expression was compared between 2D and 3D samples of Patentable genes. For determining gene expression for 3D cultured cells, samples were used from different passages. Cells from 3$^{rd}$ & 6$^{th}$ passage were used for MDA-MB-157 & MDA-MB-231, from 3$^{rd}$, 6$^{th}$ & 9$^{th}$ passage for CAMA-1 and from 3$^{rd}$, 6$^{th}$, 9$^{th}$ & 12$^{th}$ passage for MDA-MB-453, MDA-MB-468, MCF7 & T47D were used.


A
BNC2
RQ
1500
1000
500
0
CAMA1-2D
CAMA1-3D-3
CAMA1-3D-6
CAMA1-3D-9
MCF7-2D
MCF7-3D-3
MCF7-3D-6
MCF7-3D-9
MCF7-3D-12
MDAMB453-2D
MDAMB453-3D-3
MDAMB453-3D-6
MDAMB453-3D-9
MDAMB453-3D-12
T47D-2D
T47D-3D-3
T47D-3D-6
T47D-3D-9
T47D-3D-12
HCC1937-2D
HCC1937-3D-3
HCC1937-3D-6
HCC1937-3D-9
HCC1937-3D-12
MDAMB468-2D
MDAMB468-3D-3
MDAMB468-3D-6
MDAMB468-3D-9
MDAMB468-3D-12
MDAMB157-2D
MDAMB157-3D-3
MDAMB157-3D-6
MDAMB231-2D
MDAMB231-3D-3
MDAMB231-3D-6
Non CSC like cells
CSC like cells
B
DDR2
RQ
2000
1500
1000
500
0
CAMA1-2D
CAMA1-3D-3
CAMA1-3D-6
CAMA1-3D-9
MCF7-2D
MCF7-3D-3
MCF7-3D-6
MCF7-3D-9
MCF7-3D-12
MDAMB453-2D
MDAMB453-3D-3
MDAMB453-3D-6
MDAMB453-3D-9
MDAMB453-3D-12
T47D-2D
T47D-3D-3
T47D-3D-6
T47D-3D-9
T47D-3D-12
HCC1937-2D
HCC1937-3D-3
HCC1937-3D-6
HCC1937-3D-9
HCC1937-3D-12
MDAMB468-2D
MDAMB468-3D-3
MDAMB468-3D-6
MDAMB468-3D-9
MDAMB468-3D-12
MDAMB157-2D
MDAMB157-3D-3
MDAMB157-3D-6
MDAMB231-2D
MDAMB231-3D-3
MDAMB231-3D-6
Non CSC like cells
CSC like cells


PVRL3
C
RQ
50000
40000
30000
20000
10000
0
CAMA1-2D
CAMA1-3D-3
CAMA1-3D-6
CAMA1-3D-9
MCF7-2D
MCF7-3D-3
MCF7-3D-6
MCF7-3D-9
MCF7-3D-12
MDAMB453-2D
MDAMB453-3D-3
MDAMB453-3D-6
MDAMB453-3D-9
MDAMB453-3D-12
T47D-2D
T47D-3D-3
T47D-3D-6
T47D-3D-9
T47D-3D-12
HCC1937-2D
HCC1937-3D-3
HCC1937-3D-6
HCC1937-3D-9
HCC1937-3D-12
MDAMB468-2D
MDAMB468-3D-3
MDAMB468-3D-6
MDAMB468-3D-9
MDAMB468-3D-12
MDAMB157-2D
MDAMB157-3D-3
MDAMB157-3D-6
MDAMB231-2D
MDAMB231-3D-3
MDAMB231-3D-6
Non CSC like cells
CSC like cells
D
DKK3
RQ
20000
15000
10000
5000
0
CAMA1-2D
CAMA1-3D-3
CAMA1-3D-6
CAMA1-3D-9
MCF7-2D
MCF7-3D-3
MCF7-3D-6
MCF7-3D-9
MCF7-3D-12
MDAMB453-2D
MDAMB453-3D-3
MDAMB453-3D-6
MDAMB453-3D-9
MDAMB453-3D-12
T47D-2D
T47D-3D-3
T47D-3D-6
T47D-3D-9
T47D-3D-12
HCC1937-2D
HCC1937-3D-3
HCC1937-3D-6
HCC1937-3D-9
HCC1937-3D-12
MDAMB468-2D
MDAMB468-3D-3
MDAMB468-3D-6
MDAMB468-3D-9
MDAMB468-3D-12
MDAMB157-2D
MDAMB157-3D-3
MDAMB157-3D-6
MDAMB231-2D
MDAMB231-3D-3
MDAMB231-3D-6
Non CSC like cells
CSC like cells

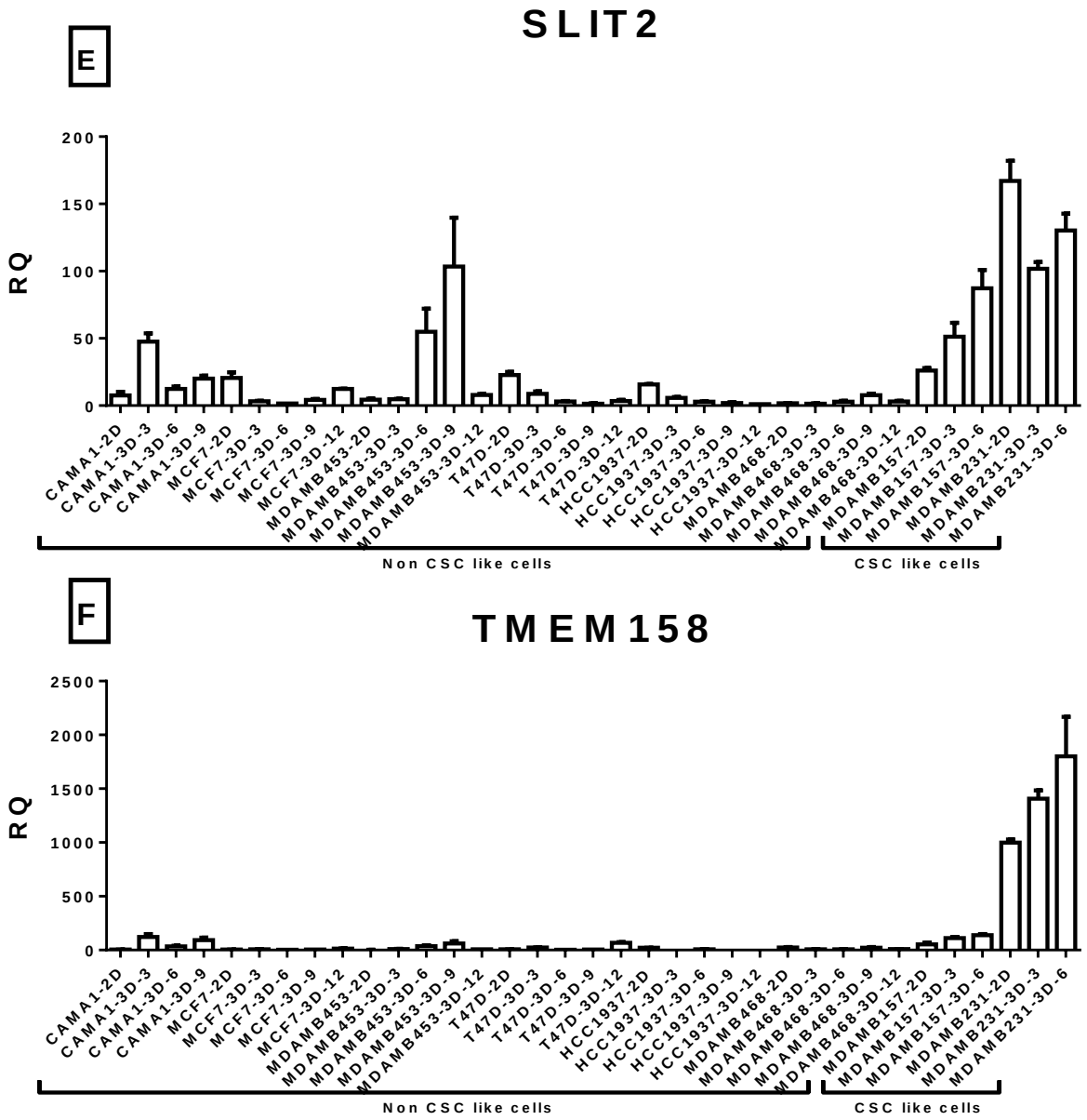
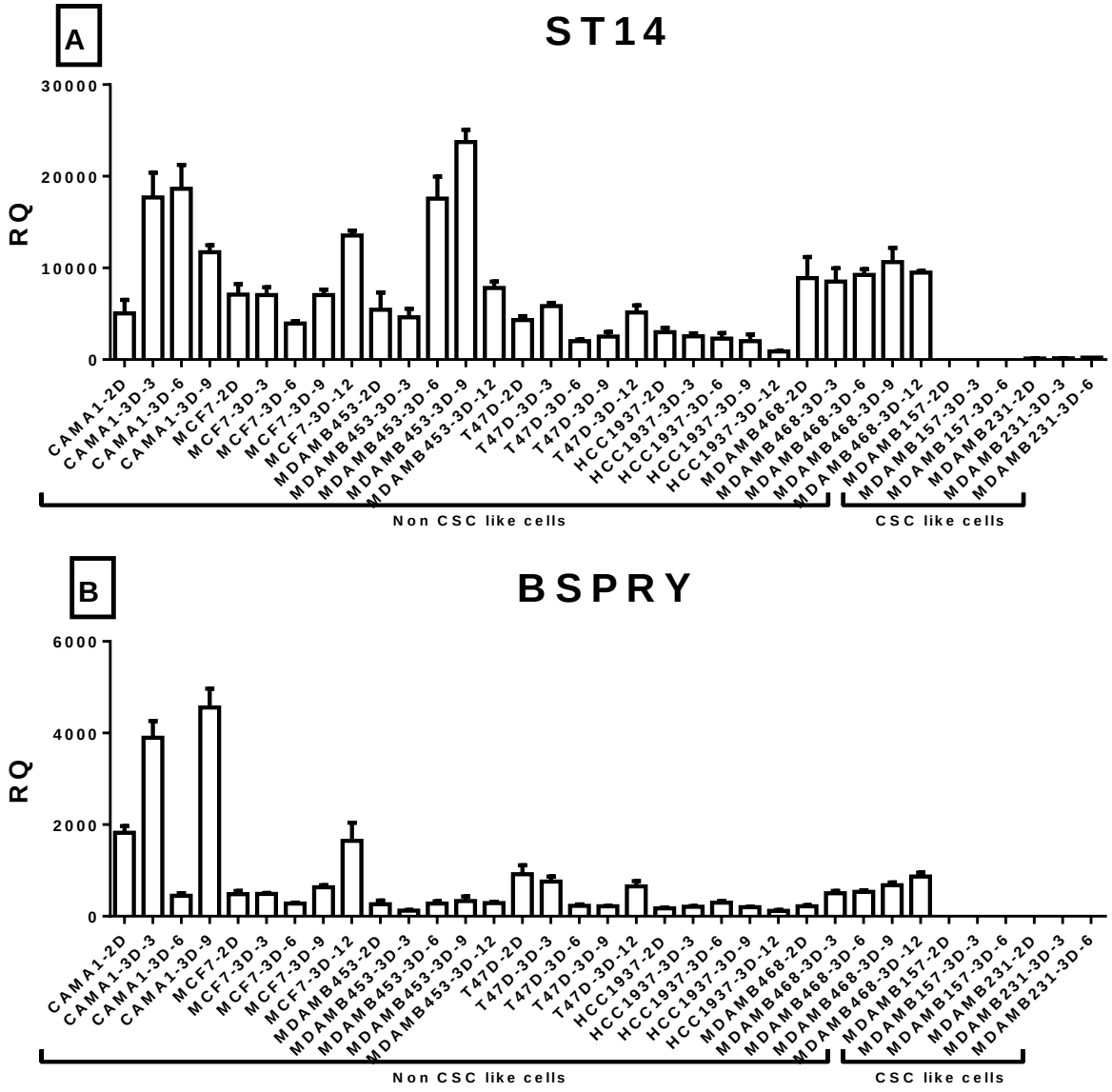
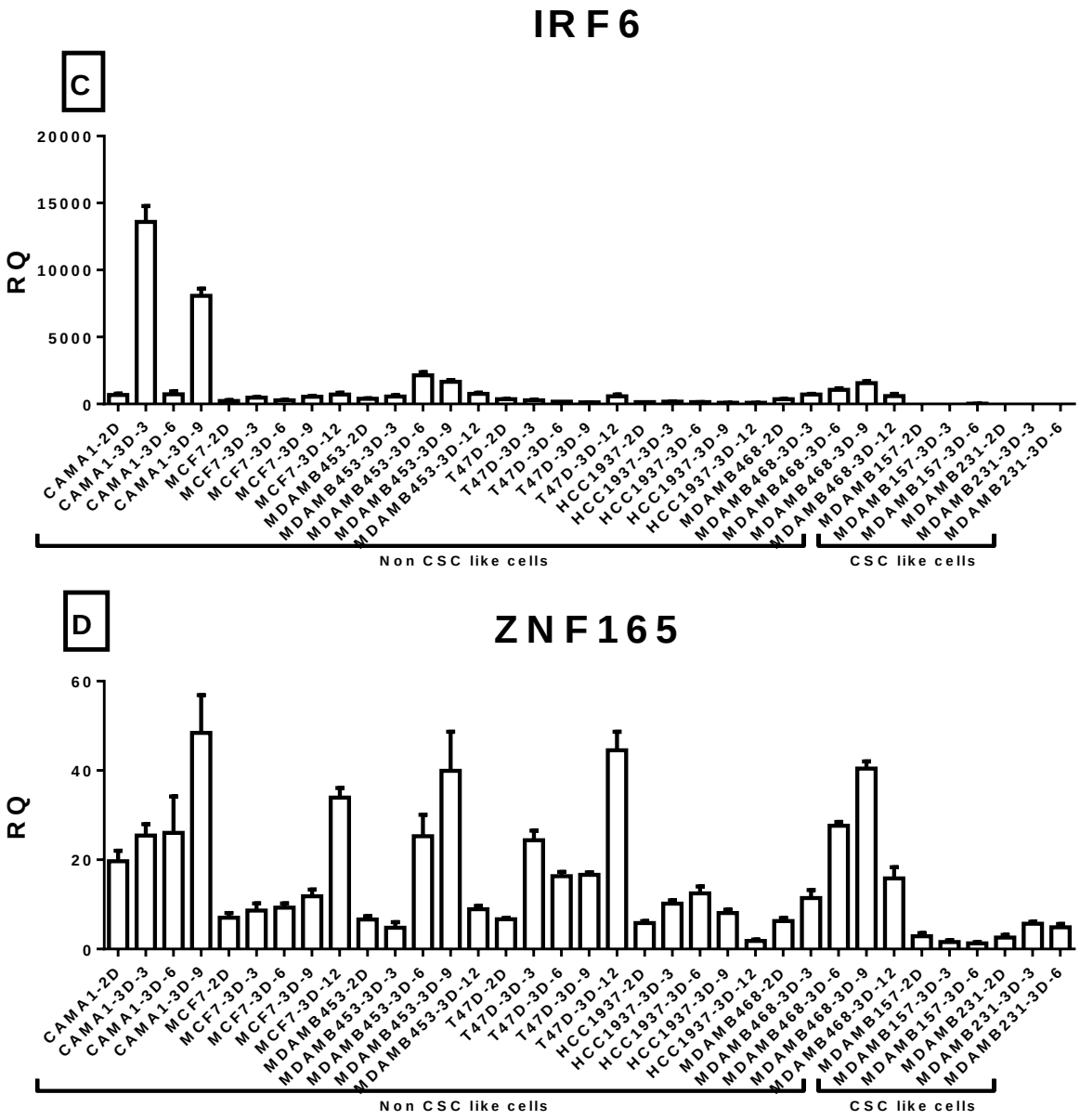



**Figure 8.7: Comparison of (A) BNC2, (B) DDR2, (C) PVRL3, (D)DKK3, (E) SLIT2 and (F) TMEM158 in breast cancer cell lines in both 2D and 3D.** 2D and 3D mentioned after name shows cell culture conditions. Passage number for each sample, are given for 3D after sample name. Error bars show standard error of mean.


A
ST14
RQ
30000
20000
10000
0
CAMA1-2D
CAMA1-3D-3
CAMA1-3D-6
CAMA1-3D-9
MCF7-2D
MCF7-3D-3
MCF7-3D-6
MCF7-3D-9
MCF7-3D-12
MDAMB453-2D
MDAMB453-3D-3
MDAMB453-3D-6
MDAMB453-3D-9
MDAMB453-3D-12
T47D-2D
T47D-3D-3
T47D-3D-6
T47D-3D-9
T47D-3D-12
HCC1937-2D
HCC1937-3D-3
HCC1937-3D-6
HCC1937-3D-9
HCC1937-3D-12
MDAMB468-2D
MDAMB468-3D-3
MDAMB468-3D-6
MDAMB468-3D-9
MDAMB468-3D-12
MDAMB157-2D
MDAMB157-3D-3
MDAMB157-3D-6
MDAMB231-2D
MDAMB231-3D-3
MDAMB231-3D-6
Non CSC like cells
CSC like cells
B
BSPRY
RQ
6000
4000
2000
0
CAMA1-2D
CAMA1-3D-3
CAMA1-3D-6
CAMA1-3D-9
MCF7-2D
MCF7-3D-3
MCF7-3D-6
MCF7-3D-9
MCF7-3D-12
MDAMB453-2D
MDAMB453-3D-3
MDAMB453-3D-6
MDAMB453-3D-9
MDAMB453-3D-12
T47D-2D
T47D-3D-3
T47D-3D-6
T47D-3D-9
T47D-3D-12
HCC1937-2D
HCC1937-3D-3
HCC1937-3D-6
HCC1937-3D-9
HCC1937-3D-12
MDAMB468-2D
MDAMB468-3D-3
MDAMB468-3D-6
MDAMB468-3D-9
MDAMB468-3D-12
MDAMB157-2D
MDAMB157-3D-3
MDAMB157-3D-6
MDAMB231-2D
MDAMB231-3D-3
MDAMB231-3D-6
Non CSC like cells
CSC like cells


IRF6
C
RQ
20000
15000
10000
5000
0
CAMA1-2D
CAMA1-3D-3
CAMA1-3D-6
CAMA1-3D-9
MCF7-2D
MCF7-3D-3
MCF7-3D-6
MCF7-3D-9
MCF7-3D-12
MDAMB453-2D
MDAMB453-3D-3
MDAMB453-3D-6
MDAMB453-3D-9
MDAMB453-3D-12
T47D-2D
T47D-3D-3
T47D-3D-6
T47D-3D-9
T47D-3D-12
HCC1937-2D
HCC1937-3D-3
HCC1937-3D-6
HCC1937-3D-9
HCC1937-3D-12
MDAMB468-2D
MDAMB468-3D-3
MDAMB468-3D-6
MDAMB468-3D-9
MDAMB468-3D-12
MDAMB157-2D
MDAMB157-3D-3
MDAMB157-3D-6
MDAMB231-2D
MDAMB231-3D-3
MDAMB231-3D-6
Non CSC like cells
CSC like cells
D
ZNF165
RQ
60
40
20
0
CAMA1-2D
CAMA1-3D-3
CAMA1-3D-6
CAMA1-3D-9
MCF7-2D
MCF7-3D-3
MCF7-3D-6
MCF7-3D-9
MCF7-3D-12
MDAMB453-2D
MDAMB453-3D-3
MDAMB453-3D-6
MDAMB453-3D-9
MDAMB453-3D-12
T47D-2D
T47D-3D-3
T47D-3D-6
T47D-3D-9
T47D-3D-12
HCC1937-2D
HCC1937-3D-3
HCC1937-3D-6
HCC1937-3D-9
HCC1937-3D-12
MDAMB468-2D
MDAMB468-3D-3
MDAMB468-3D-6
MDAMB468-3D-9
MDAMB468-3D-12
MDAMB157-2D
MDAMB157-3D-3
MDAMB157-3D-6
MDAMB231-2D
MDAMB231-3D-3
MDAMB231-3D-6
Non CSC like cells
CSC like cells

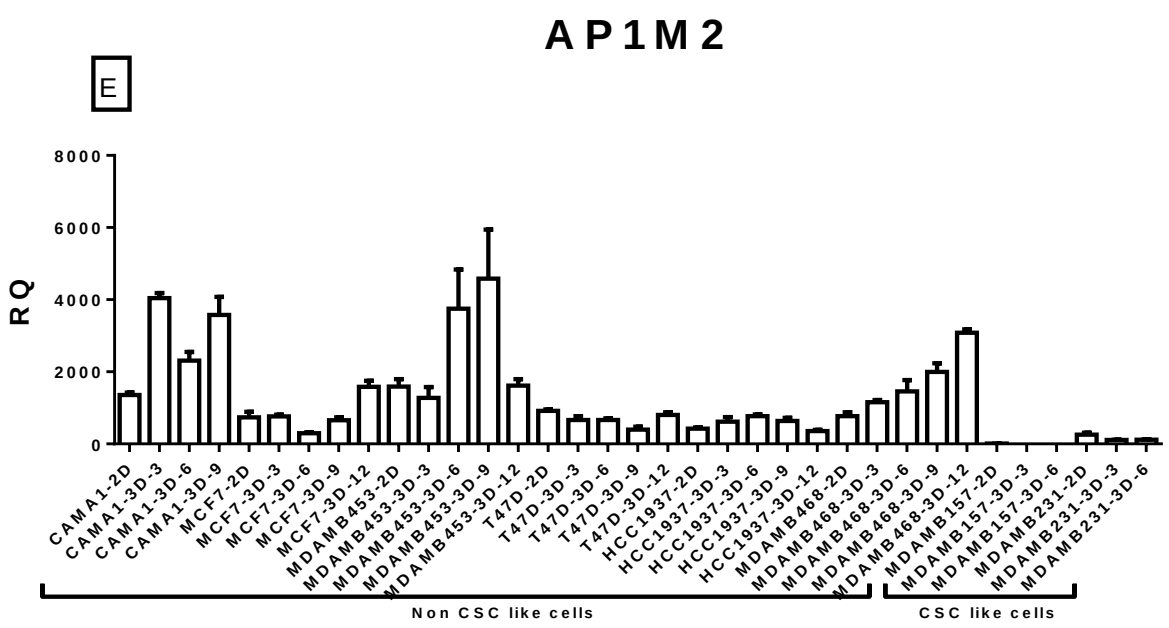



**Figure 8.8: Comparison of (A) ST14, (B) BSPRY, (C) IRF6, (D) ZNF165 and (E) AP1M2 in breast cancer cell lines in both 2D and 3D.** 2D and 3D mentioned after name shows cell culture conditions. Passage number for each sample, are given for 3D after sample name. Error bars show standard error of mean.

### 8.4 Drug Cytotoxicity for 3D Cell Culture Conditions

Cells were grown in low attachment 96 well plates and their cytotoxicity for Elesclomol, Midostaurin, Lapatinib and Panobinostat was calculated.

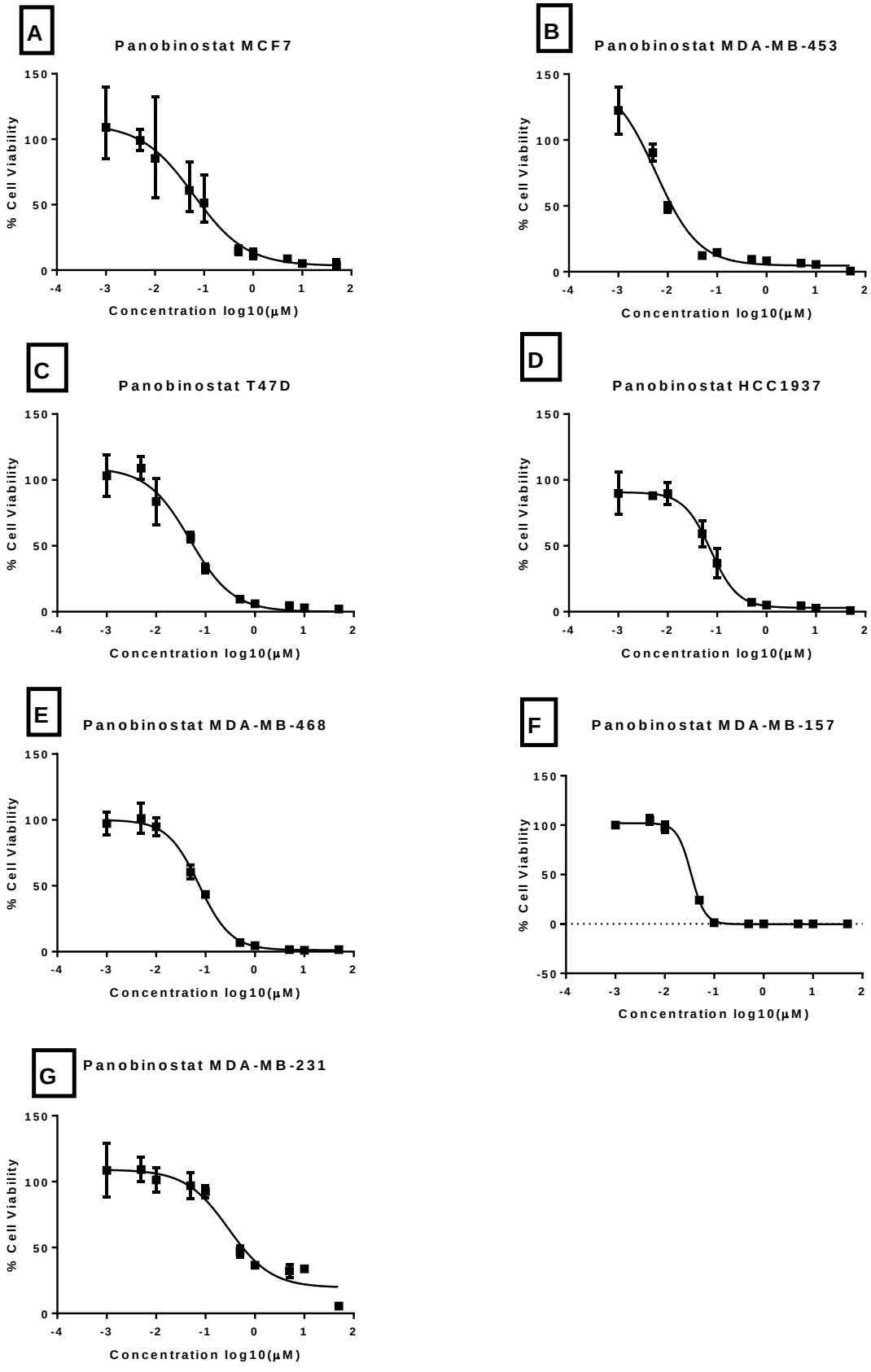

A
Panobinostat MCF7
B
Panobinostat MDA-MB-453
C
Panobinostat T47D
D
Panobinostat HCC1937
E
Panobinostat MDA-MB-468
F
Panobinostat MDA-MB-157
G
Panobinostat MDA-MB-231
% Cell Viability
Concentration log10(μM)

**Figure 8.9: Percent Cell Viability curves for Panobinostat against breast cancer cell lines.** X-axis shows log concentration of drugs in μM. And y-axis shows percent Cell Viability of cells by drug. Error bars show mean with 95% confidence interval. Drug was used in different concentration with highest concentration as 50 μM. Cells were treated with 10 concentrations as 50, 10, 5, 1, 0.5, 0.1, 0.05, 0.01, 0.005 and 0.00100μM.

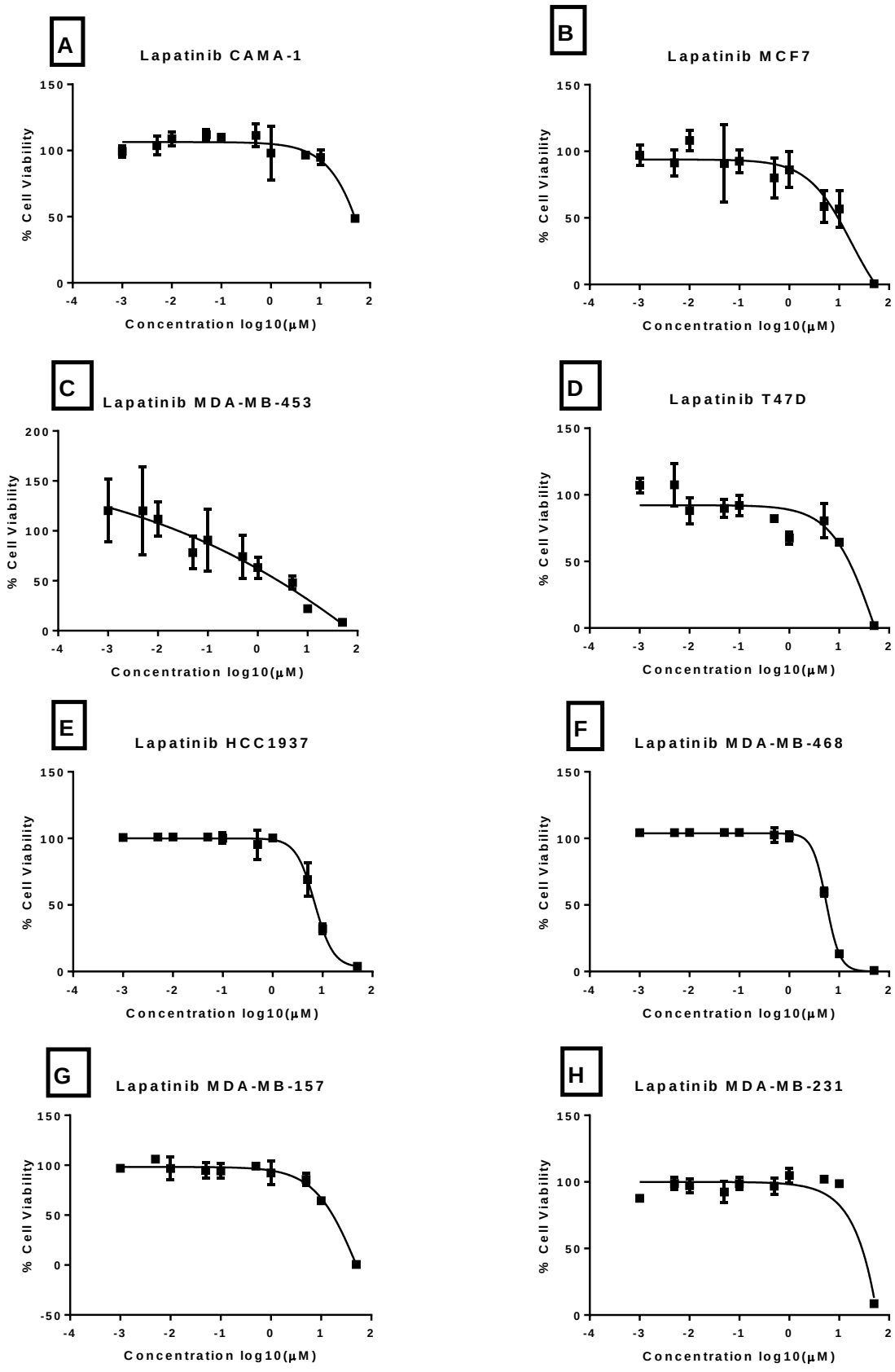

A
Lapatinib CAMA-1
% Cell Viability
Concentration log10(μM)
B
Lapatinib MCF7
% Cell Viability
Concentration log10(μM)
C
Lapatinib MDA-MB-453
% Cell Viability
Concentration log10(μM)
D
Lapatinib T47D
% Cell Viability
Concentration log10(μM)
E
Lapatinib HCC1937
% Cell Viability
Concentration log10(μM)
F
Lapatinib MDA-MB-468
% Cell Viability
Concentration log10(μM)
G
Lapatinib MDA-MB-157
% Cell Viability
Concentration log10(μM)
H
Lapatinib MDA-MB-231
% Cell Viability
Concentration log10(μM)

**Figure 8.10: Percent Cell Viability curves for Lapatinib against breast cancer cell lines.** X-axis shows log concentration of drugs in µM. And y-axis shows percent Cell Viability of cells by drug. Error bars show mean with 95% confidence interval. Drug was used in different concentration with highest concentration as 50 µM. Cells were treated with 10 concentrations as 50, 10, 5, 1, 0.5, 0.1, 0.05, 0.01, 0.005 and 0.00100µM.

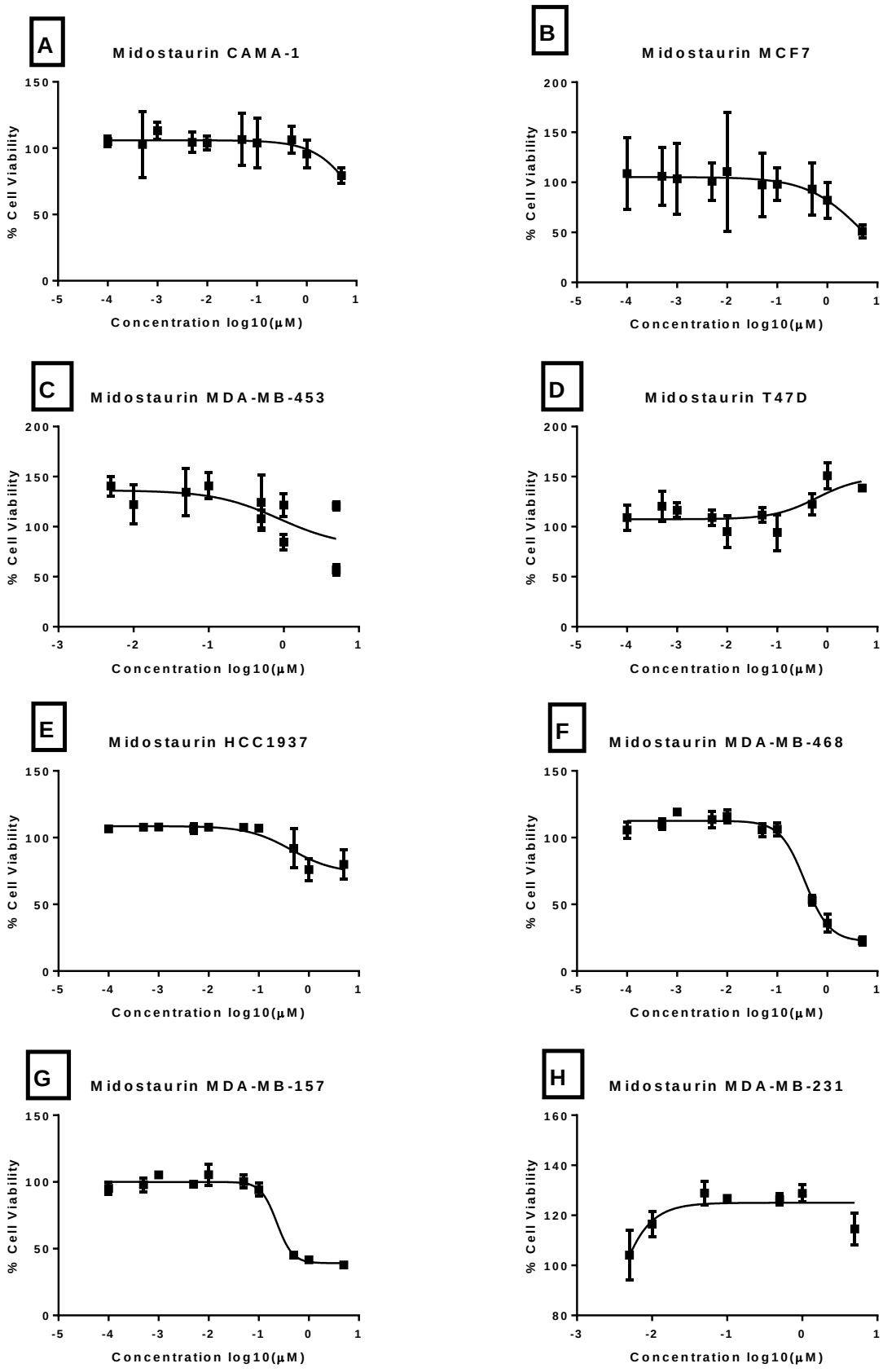

A
Midostaurin CAMA-1
% Cell Viability
Concentration log10(µM)
B
Midostaurin MCF7
% Cell Viability
Concentration log10(µM)
C
Midostaurin MDA-MB-453
% Cell Viability
Concentration log10(µM)
D
Midostaurin T47D
% Cell Viability
Concentration log10(µM)
E
Midostaurin HCC1937
% Cell Viability
Concentration log10(µM)
F
Midostaurin MDA-MB-468
% Cell Viability
Concentration log10(µM)
G
Midostaurin MDA-MB-157
% Cell Viability
Concentration log10(µM)
H
Midostaurin MDA-MB-231
% Cell Viability
Concentration log10(µM)

**Figure 8.11: Percent Cell Viability curves for Midostaurin against breast cancer cell lines.** X-axis shows log concentration of drugs in µM. And y-axis shows percent Cell Viability of cells by drug. Error bars show mean with 95% confidence interval. Drug was used in different concentration with highest concentration as 5 µM. Cells were treated with 10 concentrations as 5, 1, 0.5, 0.1, 0.05, 0.01, 0.005, 0.00100, 0.00050 and 0.00010µM.

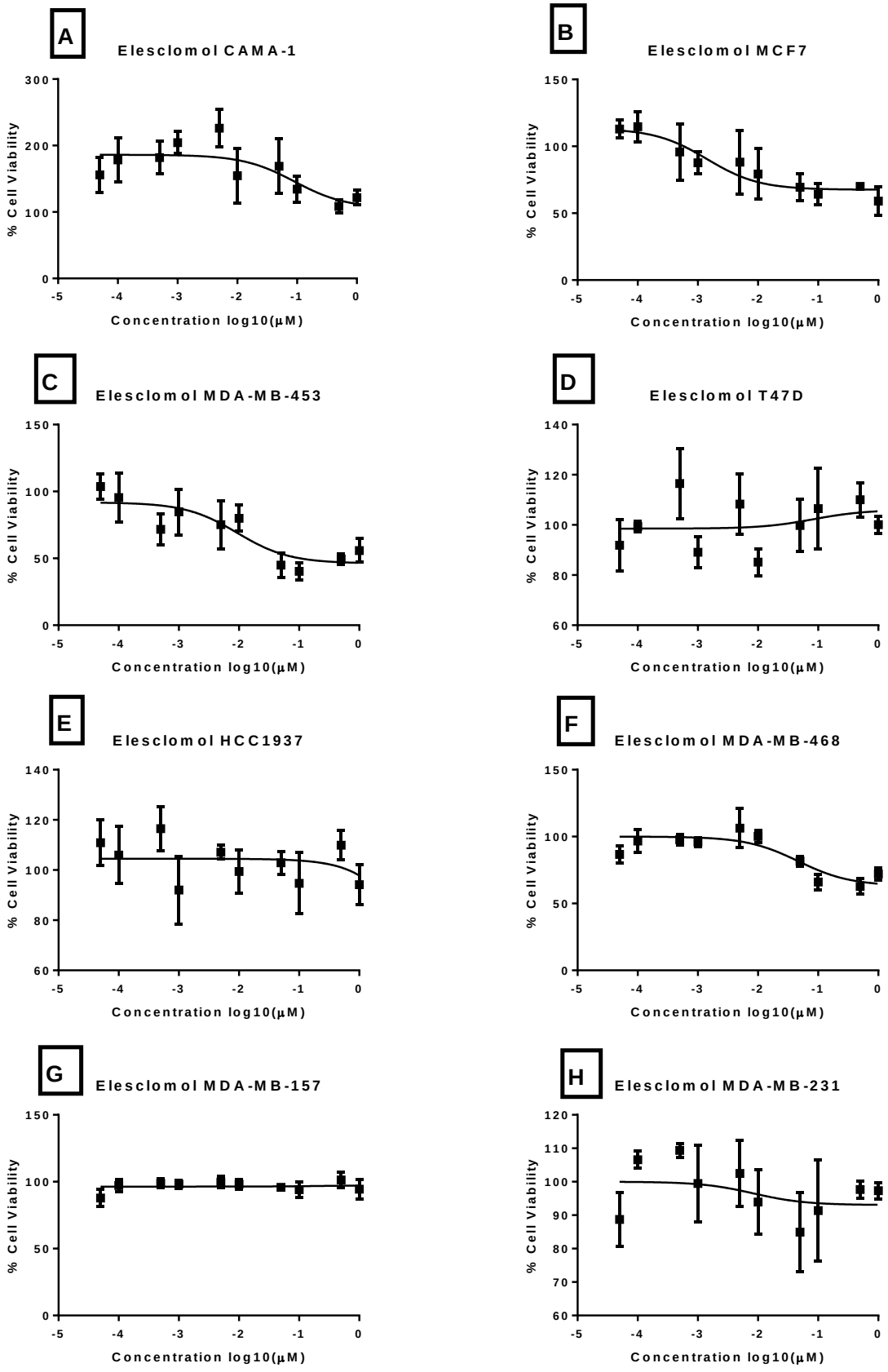

A
Elesclomol CAMA-1
% Cell Viability
Concentration log10(μM)
B
Elesclomol MCF7
% Cell Viability
Concentration log10(μM)
C
Elesclomol MDA-MB-453
% Cell Viability
Concentration log10(μM)
D
Elesclomol T47D
% Cell Viability
Concentration log10(μM)
E
Elesclomol HCC1937
% Cell Viability
Concentration log10(μM)
F
Elesclomol MDA-MB-468
% Cell Viability
Concentration log10(μM)
G
Elesclomol MDA-MB-157
% Cell Viability
Concentration log10(μM)
H
Elesclomol MDA-MB-231
% Cell Viability
Concentration log10(μM)

**Figure 8.12: Percent Cell Viability curves for Elesclomol against breast cancer cell lines.** X-axis shows log concentration of drugs in µM. And y-axis shows percent Cell Viability of cells by drug. Error bars show mean with 95% confidence interval. Drug was used in different concentration with highest concentration as 1 µM. Cells were treated with 10 concentrations 1, 0.5, 0.1, 0.05, 0.01, 0.005, 0.00100, 0.00050, 0.00010 and 0.00005µM.

## 8.5 Comparison of Gene Expression of MDA-MB-153 between 2D and 3D Culture Conditions

The genes from Patentable gene list not mentioned in Results section for MDA-MB-157 are given below.

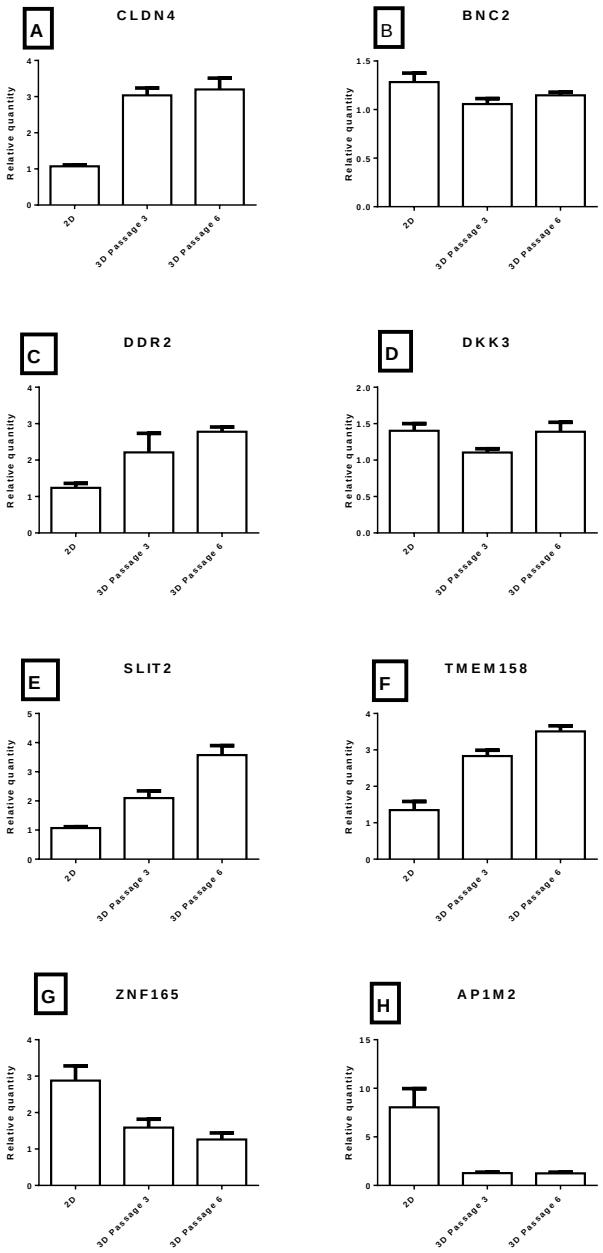


**Figure 8.13: Comparison of not selected genes (A) CLDN4, (B) BNC2, (C) DDR2 (D) DKK3, (E) SLIT2, (F) TMEM158, (G) IRF6 and (H) AP1M2 in MDA-MB-157 breast cancer cell line**. 2D and 3D mentioned after name shows cell culture conditions. Passage number for each sample, are given for 3D after sample name. Error bars show standard error of mean.