GENE EXPRESSION PROFILING OF HEDGEHOG PATHWAY GENES IN EPITHELIAL CANCERS

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

BY
ÖZLEM AKILLI ÖZTÜRK
SEPTEMBER 2006

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. Işık G. YULUĞ

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.

Asst. Prof. Özlen KONU

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. Hilal ÖZDAĞ

Approved for the Institute of Engineering and Science

Prof. Dr. Mehmet Baray
Director of Institute of Engineering and Science

ABSTRACT

GENE EXPRESSION PROFILING OF HEDGEHOG PATHWAY GENES IN EPITHELIAL CANCERS

ÖZLEM AKILLI ÖZTÜRK

M. Sc. in Molecular Biology and Genetics

Supervisor: Assoc. Prof. Işık G. Yuluğ

September 2006, 127 Pages

The Hedgehog (Hh) signaling pathway has been investigated in many cancer types and shown to have important effects, but not effectively studied in breast, colon and liver cancers. In this study, gene expression profile of *BCL2*, *SHH, SMO, IHH*, *PTCH1, GLI1, GLI2* and *GLI3* were analyzed in 15 breast, 14 colon, and 15 hepatocellular carcinoma (HCC) cell lines and 29 primary breast tumor samples and three matched normal tissue sample pools by quantitative real-time RT-PCR. Breast cell lines have different levels of target gene expression in the Hh pathway, such that this pathway is activated only in some of the breast carcinoma cell lines possibly through a ligand-independent pathway. In all the breast cancer cell lines studied, *PTCH1, SMO, GLI3* and *BCl2* had high expression. Expression profiles of the target genes predicted the estrogen receptor status correctly in 93% of the breast cell lines studied. In HCC cell lines, Hh pathway gene expression profile differentiates the poorly differentiated HCC cell lines from well differentiated ones in Discriminant Function Analysis (DFA) perfectly. High *SMO* and *IHH* expressions have been found to be markers for aberrant Hh pathway activity in the well differentiated HCC cell lines. In colon cancer cell lines deregulated expression profile among the genes were observed. In primary breast tumor samples, there was a very strong prediction for ER status of the samples with the expression of the genes included in this study.

This is the first comprehensive study that shows the transcriptional gene expression profiles of the main target genes of the Hh pathway in cancer cell lines and breast cancer tissue samples

ÖZET

EPİTEL KANSERLERDE HEDGEHOG YOLAĞI GENLERİNİN GEN İFADE PROFİLİ

ÖZLEM AKILLI ÖZTÜRK

Master Tezi, Moleküler Biyoloji ve Genetik Bölümü

Doç. Dr. Işık G.Yuluğ

Eylül 2006, 127 Sayfa

Hedgehog (Hh) sinyal yolu birçok kanser tipinde incelenmiş ve önemli etkileri olduğu gösterilmiştir ama meme, kolon ve karaciğer kanserlerinde kapsamlı olarak çalışılmamıştır. Bu çalışmada *BCL2*, *SHH, SMO, IHH*, *PTCH1, GLI1, GLI2* ve *GLI3* gen ifadesi profili 15 meme, 14 kolon ve 15 hepatoselüler karsinom (HSK) hücre hattında, 29 primer meme tümörü örneğinde ve 3 eşleştirilmiş normal doku örneği havuzunda nicel gerçek zamanlı RT-PCR ile analiz edilmiştir. Meme hücresi hatlarının Hh yolundaki hedef gen ifadesi düzeyi farklıdır ve bu yol sadece bazı meme karsinomu hücre hatlarında muhtemelen liganddan bağımsız bir yolla aktive edilir. Çalışılan tüm meme kanseri hatlarında *PTCH1, SMO, GLI3* ve *BCL2* ifadesi yüksek bulunmuştur. Hedef genlerin ifade profilleri çalışılan meme hücresi hatlarının %93'ünde estrojen reseptörü durumunu doğru tahmin etmiştir. HSK hücre hatlarında Hh yolu gen ifadesi profili iyi farklılaşmamış HSK hücre hatlarını iyi farklılaşmış olanlardan Ayırdedici Fonksiyon Analizi (DFA) kullanıldığında kusursuz olarak ayırmıştır. Yüksek SMO ve IHH ifadelerinin iyi farklılaşmış HSK hücre hatlarında normal dışı Hh yolu aktivitesi için belirteçler olduğu bulunmuştur. Kolon kanseri hücre hatlarında bazı genler için düzenlemesi bozulmuş ifade profili gözlenmiştir. Primer meme tümörü örneklerinde örneklerin ER durumunun öngörülmesi ile bu çalışmaya dahil edilen genlerin ifadesi arasında çok kuvvetli bir ilişki bulunmuştur.

Bu çalışma kanser hücre hatları ve meme kanseri doku örneklerinde Hh yolunun ana hedef genlerinin transkripsiyonel gen ifadesi profillerini gösteren ilk kapsamlı çalışmadır.

to efulim, to dost, and to my family.....

ACKNOWLEDGEMENTS

First of all, I would like to thank Assoc. Prof. Işık G. YULUĞ for her valuable supervision during my study. She supported me both theoretically and morally during my studies in MBG Department of Bilkent University.

I would like to thank Asst. Prof. Özlen Konu for her instructive comments in statistical analysis, and supporting me during graduate life at Bilkent.

I would like to thank deeply to Bâlâ for sharing her experience and knowledge regarding Real-Time RT-PCR and her valuable friendship.

I would like to thank Murat and İsmail for sharing their valuable senior project study with me, and for helping me in primer design.

I would like to express my special thanks to the Scientific and Technical Research Council of Turkey and the Turkish Association for Cancer Research and Control, Terry Fox Programme for their financial support.

I would like to thank deeply to my friends, Fatma, Bahar, and Aslı for sharing their valuable time, and encouraging me and I wish them best luck and happiness in their lives. I also would like to thank all members of the MBG Department of Bilkent University.

I would like to thank deeply to efulim, and to dost for their support in every step of my life.

Of course, my deepest gratitudes go to my family, who have been far away from me for 7 years, but I always felt that they were near me. I thank them for their love and support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

ACTB	Actin, beta
APC	Adenomatosis Polyposis Coli
A.T.	Annealing Temperature
bp	Base Pairs
BCC	Basal Cell Carcinoma
BCL2	B-cell CLL/lymphoma 2
BFB	Bromophenol Blue
BRCA	Breast Cancer
BMP	Bone Morphogenetic Protein
BWS	Beckwith-Wiedemann Syndrome
cDNA	Complementary DNA
CHEK	CHK checkpoint homolog (S. pompe)
Ci	Cubitus Interruptus
CNS	Central Nervous System
COS2	Kinesin-related protein Costal 2
CP	Crossing Point
Ct	Cycle Threshold
ddH_2O	Double distilled water
DCIS	Ductal Carcinoma *In Situ*
DEPC	Diethylpyrocarbonate
DFA	Discriminant Function Analysis
DHH	Desert Hedgehog (Drosophila)
DISP	Dispatched (Drosophila)
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic Acid
dNTP	Deoxyribonucleotide triphosphate
dsDNA	double-stranded DNA
E	Embryonic day

EDTA	Diaminoethane Tetra-acetic Acid
EGF	Epidermal Growth Factor
ERBB2	v-erb-b2 Erythroblastic Leukemia Viral Oncogene homolog 2, neuro/glioblastoma derived oncogene homolog (avian)
ER	Estrogen Receptor
EtBr	Ethidium Bromide
EXT	Exostoses Protein Family
FAM	6-carboxyfluorescein
FAP	Familial Adenomatous Polyposis
FBS	Fetal Bovine Serum
FGF	Fibroblast Growth Factor
FRET	Fluorescent Resonance Energy Transfer
Fu	Serine/Threonine Kinase Fused
GAPDH	Glyceraldehyde-3-Phosphate Dehydrogenase
GLI1	Glioma-Associated Oncogene homolog 1
GLI2	Gli-Kruppel Family Member GLI2
GLI3	Gli-Kruppel Family Member GLI3
HCC	Hepatocellular Carcinoma
HBS	Hepes Buffered Saline
HBV	Hepatis B Virus
HCV	Hepatis C Virus
HER2	ERBB2
HGF	Hepatocyte Growth Factor
Hh	Hedgehog
HIP1	Hedgehog Interacting Protein 1
HMBS	Hydroxymethylbilane Synthase
HNF	Hepatic Nuclear Factor
HNPCC	Hereditary Nonpolyposis Colorectal Cancer
HPE	Holoprosencephaly
HPRT	Hypoxanthine Guanine Phosphoribosyl Transferase

HSPG	Heparin Sulfate Proteoglycan
IDC	Infiltrating Ductal Carcinoma
IHH	Indian Hedgehog homolog (Drosophila)
IGF	Insulin-like Growth Factor
ILC	Infiltrating Lobular Carcinoma
LA	Lobulo-Alveolar
LCIS	Lobular Carcinoma *In Situ*
LOF	Loss of Function
LOH	Loss of Heterozygosity
MDS	Multidimensional Scaling
MIN	Microsatellite Instability
MMP	Matrix Metallopeptidase
MMR	Mismatch Repair
µg	Microgram
mg	Miligram
min	Minute
µl	Microliter
ml	Mililiter
µm	Micrometer
mm	Milimeter
µM	Micromolar
mM	Milimolar
MPF	Mitosis Promoting Factor
mRNA	Messenger RNA
MYC	Myelocytomatosis Viral Oncogene homolog (avian)
NBCCS	Naevoid Basal Cell Carcinoma Syndrome
NOS	No Special Type
Oligo(dT)	Oligodeoxythymidylic acid
P21/CIP1	Cyclin-dependent Kinase Inhibitor 1A
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction

Pg	Progesterone
pmol	Picomole
PR	Progesterone Receptor
Prl	Prolactin
PTCH	Patched homolog (Drosophila)
PTEN	Phosphatase and Tensin homolog
PTHrP	Parathyroid Hormone-Related Protein
q-rt-RT-PCR	Quantitative real time RT-PCR
RB1	Retinoblastoma 1
RFU	Relative Fluorescence Unit
Rpm	Revolutions Per Minute
RT PCR	Reverse Transcription PCR
RTK	Receptor Tyrosine Kinase
SCLC	Small Cell Lung Cancer
SDHA	Succinate Dehydrogenase Complex, Subunit A, flavoprotein (Fp)
Sec	Second
SHH	Sonic Hedgehog homolog (Drosophila)
SMO	Smoothened homolog (Drosophila)
SUFU	Suppressor of Fused homolog (Drosophila)
TAE	Tris-Acetate-EDTA buffer
TBP	TATA Box Binding Protein
TDLU	Terminal Duct Lobular Unit
TGF-β	Transforming Growth Factor-β
Tm	Melting Temperature
TP53	Tumor protein p53
Tris	Tris (Hydroxymethyl)- Methylamine
TTV	Tout-Velu
UV	Ultraviolet
VEGF	Vascular Endothelial Growth Factor
v/v	volume/volume

w/v	weight/volume
Wnt	Wingless
XC	Xylene Cyanol
YO	Oxazole Yellow
ZPA	Zone of Polarizing Activity

CHAPTER 1

INTRODUCTION

Cancer is abnormal proliferation of the cells that loss control over the cell growth. Cancers that arise from epithelial surfaces are called carcinoma. If the surface is glandular type of surface then it is called adenocarcinoma. Carcinomas and adenocarcinomas are the most common types of the cancer. They include breast, lung, prostate, intestinal, skin, ovary, kidney, and liver. Cancer develops as a result of loss of control over the mechanisms regulating the normal organ functioning. Signaling pathways are one side of these mechanisms. Hedgehog signaling pathway is one of the developmentally most important pathways with a role in many cancer types. It is important to identify the role of Hedgehog signaling in epithelial cancer progression. Studies up to now have shown that Hh pathway induce cell proliferation in carcinomas. In the future, studying the global gene expression of different types of tumors will allow the identification of expression profiles unique for the epithelial cancers, their stage, and grade, leading to the finding of new targets for the treatment of these cancers.

1.1. EPITHELIAL CANCERS

Carcinoma is the general name given to epithelial cancers. Carcinomas invade surrounding tissues and organs, and may spread to lymph nodes and distal sites of the body. When the carcinoma is not invasive it is called Carcinoma *in situ*. Carcinomas may have the name of the organ of the origin (pancreas, breast) or the putative cell of the origin (hepatocellular carcinoma).

1.1.1. Breast Cancer

A woman's breast is made up of glands that make breast milk (lobules), ducts (small tubes that connect lobules to the nipple), fatty and connective tissue, blood vessels, and lymph vessels (Figure1.1).

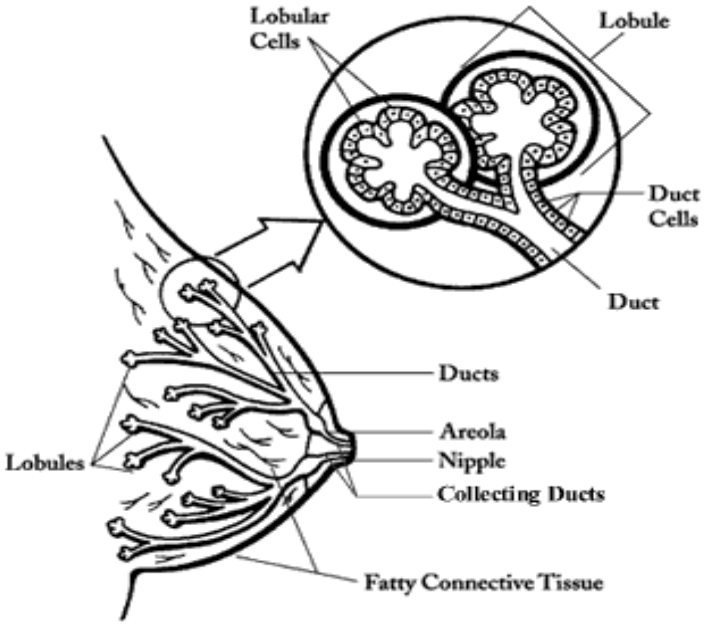


Figure 1.1: Anatomy of the breast. Woman breast is formed by ducts, lobules, fatty and connective tissues [http://www.cancer.org].

Breast cancer is one of the major cancer types women suffer from in the Unites States and Western Europe. After lung cancer, it is the second leading cause of cancer death in women. Nearly 212,920 women in the United States will be found to have invasive breast cancer in 2006 and about 40,970 of them will die from the disease. Right now there are slightly over 2 million women living in the United States who have been treated for breast cancer.

The chance of a woman having invasive breast cancer some time during her life is about 1 in 8. The chance of dying from breast cancer is about 1 in 33. Breast cancer death rates are going down. This decline is probably due to earlier diagnosis and improved treatment of the disease.

Risk factors for breast cancer are age, genetic background, reproductive history, radiation, socio-economic status, place of residence, and ethnicity [Medina, 2005; http://www.cancer.org]. Reproductive history has an important impact on the onset of breast cancer. Early menarche, late menopause, and late age of first pregnancy increase the risk of getting breast cancer [Medina, 2005].

It is thought that breast cancer is a heterogeneous group of diseases with each subtype having its own stable phenotype maintained during tumor progression rather than a single disease with a single tumorigenesis pathway [Mallon *et al.,* 2000; Polyak, 2006]. The most important determinants of these subtypes found are estrogen receptor (ER) and progesterone receptor (PR) status of tumor cells, the amplification and overexpression of the HER2 oncogene, and histologic grade. Considering these features, breast tumors are divided into luminal A, B, and C, HER2 (+), and basal subtypes. Tumor subtypes may reflect the cell of origin of the tumor with ER (-) and ER (+) tumors initiating from the earliest progenitors and more committed transit amplifying cells (cells with more limited proliferative capacity), respectively. In accordance with the proposed existence of subtype specific tumor progression pathways, it is increasingly evident that risk factors are also different for each tumor type [http://www.cancer.org].

1.1.1.1. Pathology of Breast Cancer

Most breast lumps, areas of thickening, are benign; that is, they are not cancer. Benign breast tumors are abnormal growths, but they do not spread outside of the breast and they are not life-threatening. Some benign breast lumps can increase a woman's risk of getting breast cancer. Lumps are formed by fibrocystic changes in most cases. These changes include stromal fibrosis, cyst formation, and adenosis.

Adenomas are also common benign lesions characterized by well-circumscribed benign epithelial elements with a variable amount of stroma.

Epithelial hyperplasia may be one of the initiating steps of breast carcinoma. Atypical hyperplasia is an epithelial proliferation in which some features of ductal carcinoma are seen in epithelial tissues [Beckmann *et al.,* 1997; Mallon *et al.,* 2000].

1.1.1.1.1. Ductal Carcinoma in Situ (DCIS)

It is the most common type of noninvasive breast cancer. Nearly all women with cancer at this stage can be cured. DCIS is a morphologically identifiable, preinvasive malignant proliferation of the breast epithelial cells [Mallon *et al.,* 2000; http://www.cancer.org]. The abnormal cells are contained within the mammary epithelial structures. No invasion of the basement membrane and no infiltration of the breast stroma are apparent. With a true *in situ* carcinoma, malignant epithelial cells do not have access to the lymphatic or vascular channels present within the breast stroma. Classifications are performed according to the degree of nuclear pleomorphism (often graded on a scale of 1–3), the presence or absence of necrosis, and the mitotic activity. The most characteristic feature of DCIS is that the cells composing the intraluminal proliferation are morphologically similar to each other, but have nuclear abnormalities associated with malignancy.

1.1.1.1.2. Lobular Carcinoma in Situ (LCIS)

LCIS is a neoplastic proliferation of epithelial cells in the terminal duct lobular unit with specific morphological features and therapeutic implications [Beckmann *et al.,* 1997; Mallon *et al.,* 2000]. LCIS is a proliferation of neoplastic, epithelial cells which expand the individual acini of the lobular units involving more than 50% of the acini in a lobular unit. Both LCIS and DCIS are observed more in premenopausal women, suggesting that these lesions regress after menopause and that they are hormone dependent. This idea is supported by the ER positivity of these lesions.

1.1.1.1.3. Infiltrating Ductal Carcinoma (IDC)

IDC is the most common type of breast cancer. It accounts for about 80% of invasive breast cancers [http://www.cancer.org]. If a tumor does not show the morphological features of a special type of invasive carcinoma or the characteristics of invasive lobular carcinoma, it is an IDC. Therefore, it is also called carcinoma of no special type (NOS) [Mallon *et al.,* 2000]. This group of tumors is morphologically heterogeneous. IDC tumors have very variable growth patterns and stromal responses. They are often hard and fibrous. The pattern of infiltration of the edge of the tumor is variable. In some instances these tumors have a well-circumscribed pushing margin which abuts on and distorts the surrounding normal stroma, resulting in a smooth, rounded interface between the tumor and the surrounding tissue. In these instances there is often an associated dense lymphatic infiltrate at the margin. In other cases, the tumor infiltrates the surrounding tissue diffusely enveloping preexisting normal structures in its path, resulting in an irregular infiltrative margin.

The stage is determined by spread of the tumor to the body. However, grade is determined by how the tumor cells appear under the microscope, growth rate of the tumor cells, and the tendency of tumor to spread other parts of the body. There are four stages of breast cancer [http://www.cancer.gov]. If the tumor size is less than 2 centimeters and there is no metastasis, it is a stage I tumor. As it progresses to stage IV, tumor size and metastasis levels increase. In stage IIIB and IV, the metastasis spreads to other parts of the body rather than lymph nodes. As the stage increases, the severity of the disease increases, as well. It is possible to separate IDC into three grades based on the degree of tubule formation, nuclear pleomorphism, and mitotic activity. Each of the three parameters is given a score of 1–3 and the individual scores are then added together. A score of 3–5 indicates Grade 1, 6–7 indicates Grade 2, and 8–9 indicates Grade 3. The first parameter, tubule formation, is assessed on the basis of percentage of the tumor showing distinct tubules: a score of 1 is assigned if 75% or more, a score of 2 if 10–75%, and a score of 3 if less than 10%. Nuclear pleomorphism is the second component. If the nuclei are small, with regular outlines, uniform chromatin, and little variation in size, they are assigned a

score of 1. The cytoplasm of the tumor cells may also show considerable variation, with some cells having little cytoplasm and others having abundant cytoplasm that can be eosinophilic and granular, or foamy and basophilic, or midway between the two. The third parameter is an assessment of the proliferation rate determined by counting the number of mitoses in 10 high-power fields at the periphery of the tumor. The method is standardized for each microscope objective and the tumor scored on a scale of 1–3.

1.1.1.1.3. Infiltrating Lobular Carcinoma (ILC)

ILC is the second most common type of invasive breast carcinoma and accounts for approximately 10% of all invasive breast malignancies [Mallon *et al.,* 2000; http://www.cancer.org]. The cellular morphology of the tumor and the pattern of infiltration are very important in diagnosis. The tumor cells of lobular carcinoma are found in association with foci of typical LCIS and infiltrate in a very characteristic way with one cell behind the other in a defined pattern called the Indian filing pattern. They often form concentric rings around blood vessels and lobules producing a targetoid pattern. In classical lobular carcinoma, the tumor cells are relatively small. They have regular rounded nuclei with dense, evenly staining chromatin. Nucleoli are not prominent. ILC has a higher incidence of multicentricity and bilaterality than IDC and a slightly better overall survival rate than tumors in the NOS category. There are also several other less common types of breast cancer including tubular, mucinous, medullary, papillary, invasive cribriform, and secretory carcinoma.

1.1.1.1.4. Invasion and Metastasis

Breast carcinoma metastasizes more widely and more frequently than other malignant tumors [Mallon *et al.,* 2000]. When this occurs, 95% of the patients so affected die of metastatic disease. Breast carcinoma metastasizes to distant sites by way of the angiolymphatic system. The first lymph nodes to be seeded with tumor cells are the axillary nodes, followed by the supraclavicular nodes and the mammary

nodes. Distant sites of metastases include the opposite breast, skeletal system, lungs and pleurae, liver, ovaries, adrenal glands, and central nervous system including the leptomeninges and eyes. Lymphovascular channel invasion is used as a pathological prognostic indicator.

1.1.1.2. Genetics of Breast Cancer

Cancer is a genetic disease and breast cancer has all the hallmarks of multistep genetic disease. Progression of breast cancer occurs through the accumulation of various genetic alterations [Beckmann *et al.,* 1997] (Table 1.1). A normal epithelial cell develops into a premalignant atypical cell and after clonal expansion becomes a premalignant lesion called carcinoma *in situ*. As cells detach from the basement membrane and invade the stroma, the carcinoma becomes invasive [Kenemans *et al.,* 2004]. Gain of function genetic events develop in oncogenes by DNA mutations, rearrangements or amplifications, and loss-of-function mutations in tumor suppressor genes are important molecular events. Mutations in tumors suppressor genes mostly occur in inherited cancers while mutations of oncogenes are mostly observed in sporadic cancer development.

Hereditary breast cancer forms nearly half of familial breast cancer cases [Kenemans *et al.,* 2004]. Most inherited cases of breast cancer have been associated with two genes: BRCA1, which stands for BReast CAncer gene one, and BRCA2, or BReast CAncer gene two [Futreal *et al.,* 1994; Miki *et al.,* 1994; Shattuck *et al.,* 1995; Wooster *et al.,* 1994; Wooster *et al.,* 1995]. The function of these genes is to keep breast cells growing normally and to prevent any cancer cell growth. But when these genes contain abnormalities, or mutations, they are associated with an increased breast cancer risk. Abnormal *BRCA1* and *BRCA2* genes may account for up to 2-3% of all breast cancers. Women diagnosed with breast cancer that have an abnormal BRCA1 or BRCA2 gene often have a family history of breast cancer, ovarian cancer, or both. Mutations in the cell cycle check point kinase gene (*CHEK2*) are observed in approximately 5% of all familial breast cancer cases [Kenemans *et al.,* 2004].

However, it is also important to remember that most women with breast cancer have no family history of the disease.

Only 4-10% of breast cancers have a germ line mutation that predispose to breast cancer [Lerebours *et al.,* 2002]. The remaining part is somatic breast cancers resulting from the genetic alterations. 70% of breast cancers are aneuploid, with DNA amplifications in oncogenes or chromosomal deletions in tumor suppressor genes. The consistently mutated tumor suppressor gene in sporadic breast cancer is *TP53* with point mutations in approximately 22-34% of the cases. *TP53* mutations are observed mostly in ductal carcinoma [Beckmann *et al.,* 1997]. Point mutations are the common form of mutation for the *TP53* leading to proteins defective for sequence-specific DNA binding and activation of target genes. *TP53* mutation is a rare event in the progress of breast cancer, and it is rarely observed in hereditary breast cancer [Kenemans *et al.,* 2004]. *MYC* is also altered in the breast cancer. It is amplified and thus overexpressed in approximately 15-20% of breast tumors. *MYC* may be involved in the early stages of cancer progression since it is a key transcription factor having role in cell growth, differentiation, and apoptosis [Lerebours *et al.,* 2002]. In addition, down regulation of the expression of the *RB1* tumor suppressor gene is frequent in breast cancer. Deletions of the chromosomal region containing the *RB1* gene, 13q14.1, are found in 25-35% of breast cancers [Lerebours *et al.,* 2002]. Although *BRCA1* and *BRCA2* mutations are rare in sporadic breast cancers, the down regulation of BRCA1 expression by hyper-methylation is observed [Lerebours *et al.,* 2002]. The respective locus at 17q22 is frequently deleted in sporadic breast cancers. P21 has a role in the regulation of DNA methylation. In the normal cells it prevents DNA methylation. In the cancer cells it loses its control and the level of DNA methylation decreases overall, although at some major genes there is dense methylation [Salisbury, 2001]. P21 itself is under epigenetic control in the breast cancer by histone deacetylation. Beside these, there are many genes that are deregulated in breast cancer. Microarray technology will allow more progress in the understanding of the genetics of breast cancer.

Table 1.1: Genes involved in breast carcinogenesis [Kenemans et al., 2004]

Gene	Locus	Role in hereditary breast cancer	Role in sporadic breast cancer
BRCA1	17q12-21	Germline mutation (hereditary breast ovarian cancer syndrome)	Inactivation by hypermethylation of the BRCA1 promotor region
BRCA2	13q12-13	Germline mutation (hereditary breast ovarian cancer syndrome)	Silenced by overexpressed EMSY
TP53 (p53)	17p13.1	Germline mutation (Li-Fraumeni syndrome) TP53 mutations frequent in BRCA1 and BRCA2 mutant breast cancers	Late event
Rb1	13q14.1	No specific role	Late event
PTEN (MMAC1)	10q.23-24	Germline mutation (Cowden disease syndrome)	Rare
MYC	8q24	No specific role	Overexpressed in 25–30%
ERBB2/Her2/neu	17q21	Frequently underexpressed in BRCA1 mutant breast cancers	Overexpressed in 25–30%
CDH1 (E-Cadherin)	16q22.1	No specific role	Early event in lobular breast cancer
CCND1 (Cyclin D1)	11q13	Frequently underexpressed in BRCA1 mutant breast cancers	Overexpressed in 30–40%
ERα	6q25.1	Frequently underexpressed in BRCA1 mutant breast cancers	Underexpressed in 25%
ERβ	14q22-24	Not known	Not known

1.1.1.3. Signaling Pathways in Breast Cancer

There is a hypothesis about the cellular origin of cancer in which stem cells are thought to be responsible for cancer progression, since they are quiescent over long periods of time, and could accumulate mutations, and thus when stimulated to proliferate enrich abnormal cell growth [Li *et al.,* 2005; Liu *et al.,* 2005; Woodward *et al.,* 2005]. Normal stem cells and cancer stem cells share some properties that prove this hypothesis, in which self-renewal, the ability to differentiate, active telomerase and anti-apoptotic pathways, anchorage independence and, ability to migrate and form metastasis are important ones. Mammary gland stem cells function in tissue remodeling during cycles of pregnancy, which make them good candidate for the transfer of the mutations to the remodeled tissues. Understanding the signaling pathways involved in the self-renewal of both normal and cancer stem cells is important for anti-cancer therapies [Woodward *et al.,* 2005]. There are several key signaling pathways governing this process in stem cells. These are Wnt/β-catenin, Hedgehog (Hh), transforming growth factor (TGF)-β, PTEN and Bmi signaling

pathways. Deregulation of these pathways is shown to cause mammary tumors in transgenic mice [Li *et al.,* 2005; Liu *et al.,* 2005].

1.1.2. Hepatocellular Carcinoma

The liver plays a central role in the metabolism of an organism. In this organ, several highly-specialized cell types interact by taking over complex and life-essential metabolic functions such as glucose and protein metabolism. Since the liver is made up of different types of cells, different types of tumors can form in the liver [http://www.cancer.org]. Some of these are cancerous and some are benign, and they treated differently. While there are other types of liver cancers, the most common form in adults is called hepatocellular carcinoma (HCC). It begins in the hepatocytes, the main type of liver cell. About 3 out of 4 primary liver cancers are HCC. Hepatocellular carcinoma (HCC) is one of the most common malignant tumors in some areas of the world with an extremely poor prognosis. The major etiologic risk factors for HCC development include toxins (alcohol, aflatoxin B1), androgens and estrogens, hepatitis B virus (HBV) and hepatitis C virus (HCV) infection as well as various inherited metabolic disorders, such as alpha-1-antitrypsin deficiency and hemochromatosis [Blum, 2005].

18,510 new cases of primary liver cancer and bile duct cancer are estimated to be diagnosed in the United States during 2006. It is about twice as common in men as in women. About 16,200 people will die of liver cancer in the United States during 2006. This cancer is many times more common in developing countries in Africa, and East Asia than in the United States. In many of these countries it is the most common type of cancer [http://www.cancer.org].

1.1.2.1. Pathology of Hepatocellular Carcinoma

Hepatocellular carcinoma is heterogeneous both phenotypically and genetically. Assessment of the differentiation grade of hepatocellular carcinomas (HCCs) is important for evaluation of the pathological diagnosis, prognosis and therapeutic

treatment [Ishiyama *et al.,* 2003]. Early well-differentiated tumors are highly proliferative and become less differentiated when they reach 1-1.5 cm. At this stage, HCC cells start to acquire the abilities of angiogenesis, tissue invasion and metastasis. Later on, they become undifferentiated and are able to invade vessels and form extra-hepatic metastases [Bruix *et. al.,* 2004]. This dedifferentiation process is associated with a progressive accumulation of genomic changes including chromosomal gains and losses, as well as *p53* mutations [Thorgeirsson *et al.,* 2002]. The comprehensive expression of HNF-3β, HNF-4α, HNF-1α, and C/EBPα corresponds well with differentiated functions and morphology of HCC. Expressions of HNF-4α and HNF-1α among these four transcription factors showed synchronism and had a close relationship with HCC differentiation. HNF-4α and HNF-1α are useful markers to evaluate differentiation of HCC [Ishiyama *et al.,* 2003].

Malignant transformation occurs through a pathway of increased liver cell turnover, induced by chronic liver injury and regeneration in a context of inflammation, immune response, and oxidative DNA damage which cause mutations of various genes and a genetic instability [Blum, 2005].

1.1.2.2. Genetics of Hepatocellular Carcinoma

Numerous genetic alterations are accumulated during the progression of HCC. There are mainly two types of genetic alterations of which the first one is genetic alterations specific of hepatocellular tumor risk factors as shown in Figure 1.2 [Laurent-Puig *et al.,* 2006]. It includes integration of hepatitis B virus (HBV) DNA, R249S *TP53* mutation in aflatoxin B1-exposed patients, *KRAS* mutations related to vinyl chloride exposure, hepatocyte nuclear factor 1a, *HNF1a*, mutations associated to hepatocellular adenomas and adenomatosis polyposis coli, APC, germline mutations predisposing to hepatoblastomas. The second type of genetic alterations are etiological nonspecific, it includes recurrent gains and losses of chromosomes, alteration of *TP53* gene, activation of WNT/β-catenin pathway through CTNNB1/β-catenin and axis inhibition protein, AXIN, mutations, inactivation of retinoblastoma and insulin-like growth factor 2 receptor, IGF2R, pathways through inactivation of

RB1, P16 and IGF2R. Based on the presence or the absence of chromosomal instability two pathways of hepatocarcinogenesis have been found as shown in Figure 1.2. Hepatitis B virus and poorly differentiated tumors are related to chromosome instable tumors associated with frequent TP53 mutations, whereas non-HBV and well-differentiated tumors are related to chromosomal stable samples that are frequently β-catenin activated.

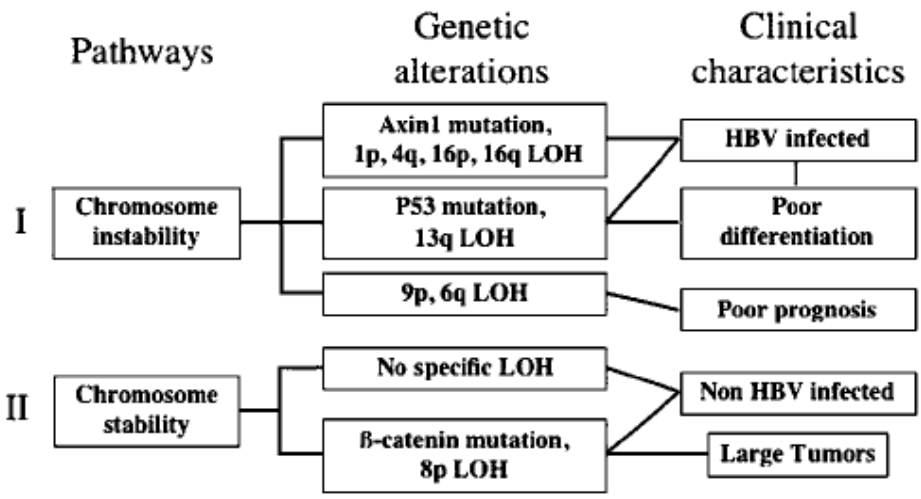


Figure 1.2: Major hepatocarcinogenesis pathways defined by genetic alterations with clinical parameters. Lines joining boxes indicate significant correlations [Laurent-Puig *et al.,* 2006].

Beckwith–Wiedemann syndrome (BWS) or in the familial adenomatous polyposis (FAP) are two cancer predispositions leading to hereditary HCC [Laurent-Puig *et al.,* 2006]. FAP is owing to *APC* germline mutation. Nuclear/cytoplasmic accumulation of the β-catenin is observed in most hepatoblastomas and sequence analysis of the β-catenin N-terminal domain has revealed interstitial deletions or missense mutations in the GSK3β phosphorylation motif in 48–67% of sporadic hepatoblastoma tumors. Activation of the WNT/β-catenin pathway may be related to axis inhibition protein 2, AXIN2, mutations in other familial cases.

1.1.2.3. Signaling Pathways in Hepatocellular Carcinoma

Dysregulation of growth factors, receptors and their downstream signaling pathway components represent a central pro-tumorigenic principle in human hepatocarcinogenesis. Especially the insulin-like growth factor/IGF-1 receptor (IGF/IGF-1R), hepatocyte growth factor (HGF/MET), wingless (Wnt/β-catenin/FZD), transforming growth factor alpha/epidermal growth factor receptor (TGFα/EGFR) and Transforming Growth Factor β (TGF-β/TGF-β Receptor) pathways contribute to proliferation, anti-apoptosis and invasive behavior of tumor cells.

1.1.3. Colorectal Cancer

Colorectal cancer is the third most common cancer found in men and women in this United States other than skin cancer [http://www.cancer.org]. It is estimated that there will be about 106,680 new cases of colon cancer and 41,930 new cases of rectal cancer in 2006 in the United States. Combined, they will cause about 55,170 deaths.

The death rate from colorectal cancer has been decreasing for the past 15 years. One reason is that there are fewer cases. Thanks to colorectal cancer screening, polyps can be found and removed before they turn into cancer [http://www.cancer.org].

1.1.3.1. Genetics of Colorectal Cancer

Hereditary nonpolyposis colorectal cancer (HNPCC) and familial adenomatous polyposis (FAP) are inherited cancer syndromes that account for approximately five percent of all colon cancers [http://www.cancer.org]. In HNPCC, families typically have at least three members with a history of colorectal cancer, and patients commonly develop cancers at an early age. Patients with FAP also have a strong family history of colon cancer and develop multiple polyps at a young age, some of which ultimately turn malignant if preventive measures are not taken. FAP is caused

by mutations in the APC gene, whereas HNPCC is caused by muttaions in several genes, including MSH2, MLH1, PMS1, PMS2, MSH6, TGFBR2, and MLH3 [http://www.cancer.org].

Due to the accumulation of mutations in a number of oncogenes and tumor suppressor genes transformation of colorectal epithelial cells into the cancer cells occur as shown in Figure 1.3 [Wang *et al.,* 2006]. Approximately 15% of sporadic colorectal cancers are caused by somatic inactivation of mismatch repair (MMR) genes, which leads to a microsatellite instability (MIN) phenotype. Interestingly, a much higher incidence (~85%) of MIN has been found in colorectal cancer patients with HNPCC tumors. The predominant mechanism responsible for the inactivation of MMR in these tumors is epigenetic silencing through promoter methylation with the additional involvement of somatic mutations [Wang *et al.,* 2006].

The tumor suppressor APC was first identified by genetic analysis in patients with FAP. Adenomatous polyposis coli, which is encoded by the FAP locus on chromosome 5, is mutated in most sporadic forms of colon cancer as well. This invariably leads to hyperproliferation of colon crypt cells and to the formation of polyps. K-ras is the well known oncogene functioning in 47-50% of colorectal cancers. When mutated it stimulates the Wnt signaling through suppression of GSK-3β [Wang *et al.,* 2006]. Matrix metalloproteinases have an important role in colon cancer progression. Increased levels of MMP-1, -2, -3, -7, -9, and -13 are found in the colon cancer samples as compared to normal mucosa. MMP may function in the tumor progression not only through ECM degradation but also through liberating growth factors and/or cytokines, suppressing the immune response, and modulating the angiogenesis. In addition, MMP-7 may have a role in EGF receptor activation in colon cancer cells.

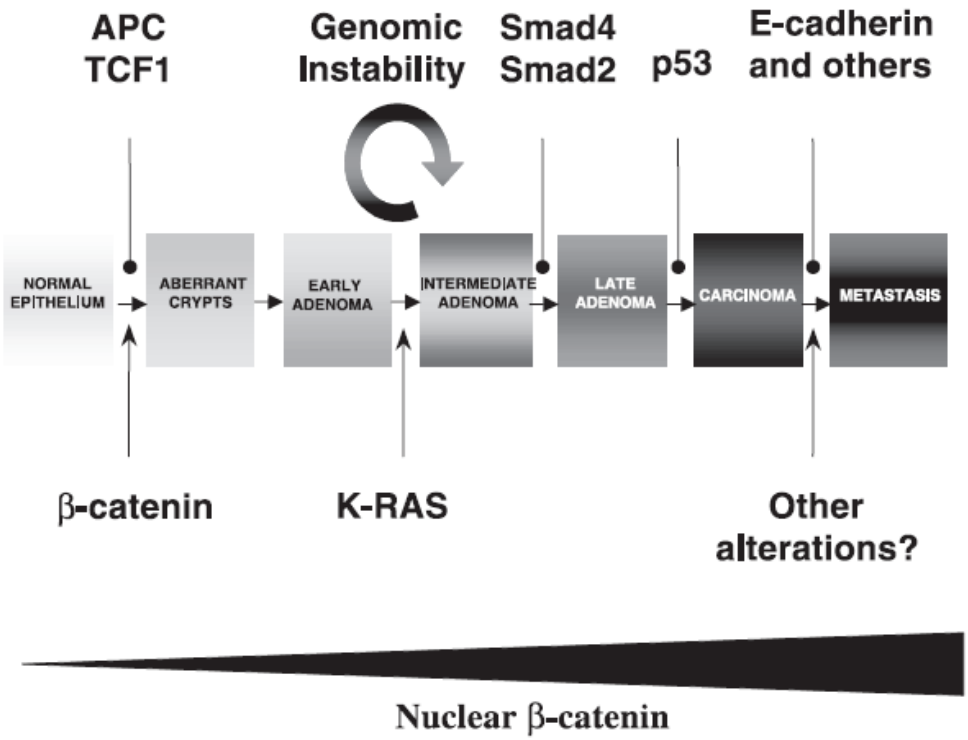


Figure 1.3: The adenoma-carcinoma sequence for colorectal cancer. A mutation in APC or β-catenin results in the activation of the Wnt signaling pathway, triggering tumor formation. Subsequent progression towards malignancy is accompanied by sequential mutations in k-RAS, deletion of chromosome 18q affecting genes encoding SMAD2 and SMAD4, p53, and genes involved in tumor invasiveness such as E-cadherin. Tumor suppressor proteins are represented above the adenoma-carcinoma sequence, whereas oncogenes are depicted below. Increasing levels of nuclear β-catenin accompany tumor progression [Giles *et al.,* 2003].

1.1.3.2. Signaling Pathways in Colorectal Cancer

Stem cells residing in the crypt base of the colon have been thought to be responsible for colon cancer progression since they are the cells providing the balance between asymmetrical cell division and cell proliferation which can be disrupted by several genetic and/or epigenetic alterations (Figure 1.4) [Wang *et al.,* 2006]. Activating

mutations of the Wnt pathway causes 90% of the colorectal cancer [Giles *et al.,* 2003]

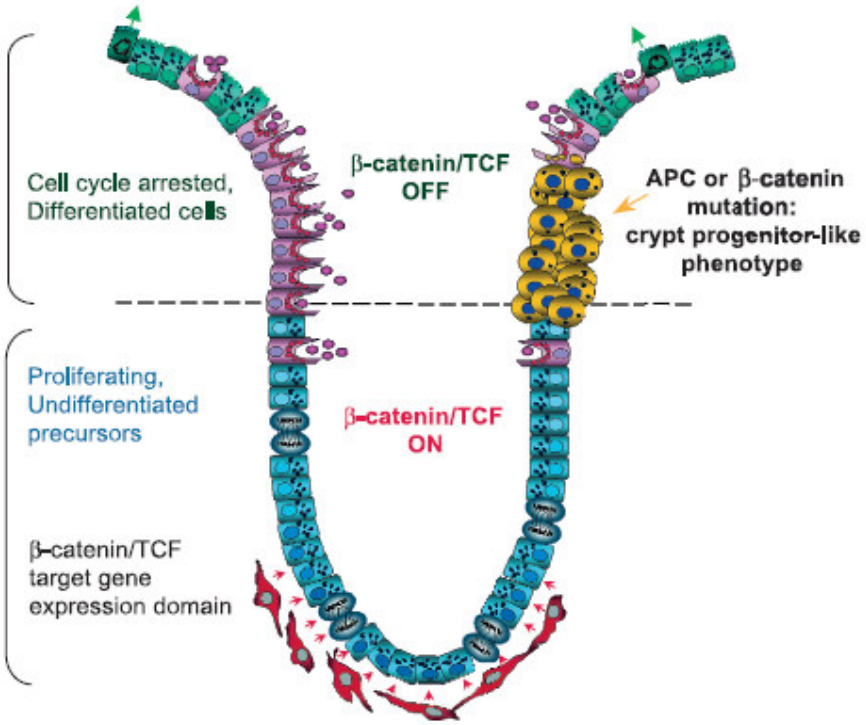


Figure 1.4: Schematic representation of a colon crypt and proposed model for polyp formation. At the bottom third of the crypt, the progenitor proliferating cells accumulate nuclear β-catenin. Consequently, they express β-catenin/TCF target genes. An uncharacterized source of WNT factors likely resides in the mesenchymal cells surrounding the bottom of the crypt, depicted in red. As the cells reach the mid-crypt region, β-catenin/TCF activity is downregulated and this results in cell cycle arrest and differentiation. Cells undergoing mutation in APC or β-catenin become independent of the physiological signals controlling β-catenin/TCF activity. As a consequence, they continue to behave as crypt progenitor cells in the surface epithelium giving rise to aberrant crypt foci [Giles *et al.,* 2003].

1.2. HEDGEHOG SIGNALING PATHWAY

Hedgehogs (Hh) comprise a family of secreted signaling molecules that direct numerous developmental and patterning events in mammalian tissues [Ingham *et al.,* 2001; Taipale *et al.,* 2001]. Both the distribution of Hh molecules and the cellular response to Hh are tightly regulated. Research over the past years has shed light on the complex Hh signaling pathway.

1.2.1. Hedgehog Signaling Pathway and the Genes

The Hh protein family includes Sonic (SHH), Indian (IHH), and Desert (DHH) hedgehogs [Ingham *et al.,* 2001]. The Hh signaling pathway starts with expression of one of these proteins (Figure 1.5). SHH is the predominant signaling molecule in lung, brain and limb development and is the most extensively studied Hh protein in vertebrates. The receptor that SHH binds is called patched (PTCH). PTCH is a 12-pass transmembrane protein. In quiescent cells PTCH catalytically acts on Smoothened (SMO) to inhibit it. In this form a cytoplasmic complex made up from atypical kinesin costal2, Ser/Thr kinase fused, and suppressor of fused (SUFU) favors the repressor form of the GLI transcription factors. Binding of SHH to PTCH inactivates it so that SMO is relieved and COS2/FU/SUFU cannot process the GLI transcription factor family of proteins. SUFU has a role in subcellular localization of GLI proteins. Mutations of SUFU results in a truncated form which is unable to transport GLI proteins, leading to constitutively active Hh pathway [Wetmore, 2003]. Processed GLI proteins result in N-terminal truncated activator and C-terminal truncated repressor fragments [Magliano *et al.,* 2003]. Active GLI transcription factor family proteins translocate to nucleus, bind to Hh target gene promoters and induce their expression [Watkins *et al.,* 2004]. In vertebrates there are three GLI transcription factor family proteins. Glioma-associated oncogene (GLI1) is an onco-protein and positive regulator of the pathway. It is target of itself in the Hh signaling and thought not to be regulated by processing. GLI-Kruppel family member Gli2 (GLI2) and GLI-Kruppel family member Gli3 (GLI3) contain potent repressor

domains and behave both as activators and repressors [Toftgard, 2000]. GLI1 is thought to be the main transcription factor activating the pathway in both development and disease. GLI2 has a role in ectopic expression of the pathway [Wetmore, 2003].

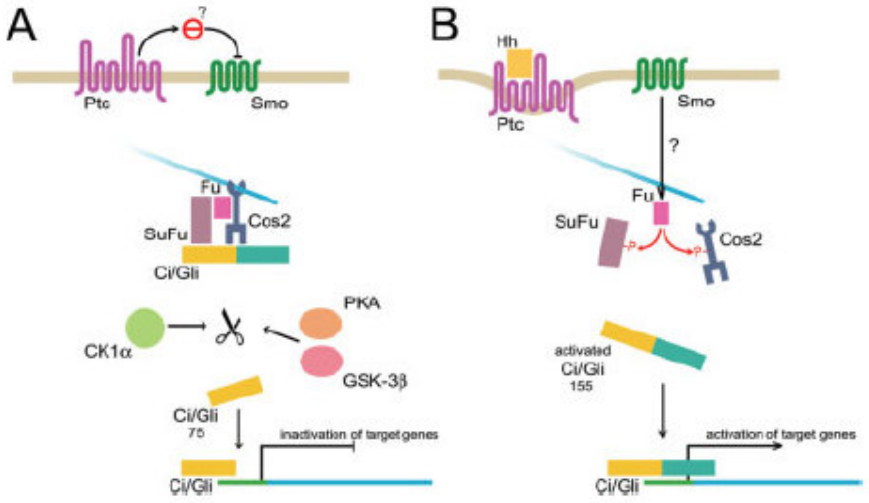


Figure 1.5: Hedgehog Signaling Pathway. [Biljsma *et al.,* 2004]

The active SHH signaling peptide is formed by an autoprocessing reaction that converts a 45 kDa precursor into a 20 kDa signaling peptide that is doubly lipid modified, with palmitate and cholesterol residues at the N and C termini, respectively (Figure 1.6). The 25kDa C-terminal part has the autoproteolytic activity, and the N-terminal domain has the signaling activity. Transport of the ligand requires transmembrane protein Dispatched (DISP) [Ma *et al.,* 2002], which is a 12-pass transmembrane protein with sequence similarity to PTCH. DISP makes Hh ligand available for long range signaling by releasing it from the plasma membrane [Cohen, 2003]. The movement of processed ligand to the responding cell occurs by help of tout-velu (TTV) protein. It belongs to the EXT protein family functioning in heparin sulfate proteoglycan (HSPG) biosynthesis. HSPG may function in the carriage of ligand to the PTCH receptor or moving of it from cell to cell [Cohen, 2003].

SHH is the most broadly expressed hedgehog gene and responsible for the major effects on the development of the brain, spinal cord, axial skeleton, and limbs. *IHH* is found to be functioning in regulation of cartilage differentiation in the long bones. *DHH* is expressed mainly in developing germ line and Schwann cells of the peripheral nervous system [Bale, 2003].

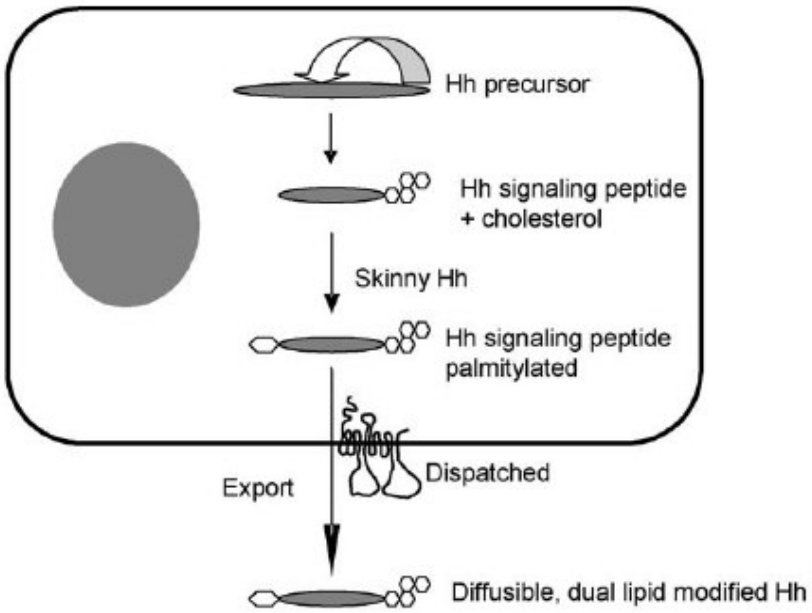


Figure 1.6: Synthesis of Hedgehog ligand in the signaling cells [Watkins *et al.,* 2004]. Hedgehog signaling protein is formed by autoprocessing reaction in which Hh precursor protein is converted to dual lipid modified active protein.

Gli transcription factor family is composed of three GLI proteins in vertebrates (GLI1, GLI2, and GLI3). These proteins have several regions with sequence homology, including a centrally located DNA-binding domain with five C2-H2 zinc fingers and a C-terminal transcription activation domain [Ruiz i Altaba, 1999]. These proteins have distinct activities and are not functionally equivalent. Nevertheless, their partial redundancy and often overlapping domains of expression has made it difficult to define precisely their individual features and functions.

GLI3 is the transcription factor that most resembles the Ci transcription factor of *Drosophila*. Like Ci, it is proteolytically cleaved to form a repressor form in the absence of the ligand. Hh stimulation prevents the cleavage; the repressor effect of GLI3 is relieved [Ruiz i Altaba, 1999], and contributes to target gene activation at least by a direct derepression and possibly by direct transcriptional activation. Hh ligand binding is thought to induce activation of GLI2 mediated transcriptional activity. GLI1 is primary target of the pathway and its expression is regulated by both GLI2 and GLI3. In the absence of GLI2 and GLI3, it can mediate the target gene activation, however in the absence of GLI2 and GLI3 there is no GLI1 expression in the cell [Lipinski *et al.,* 2006] (Figure 1.7).

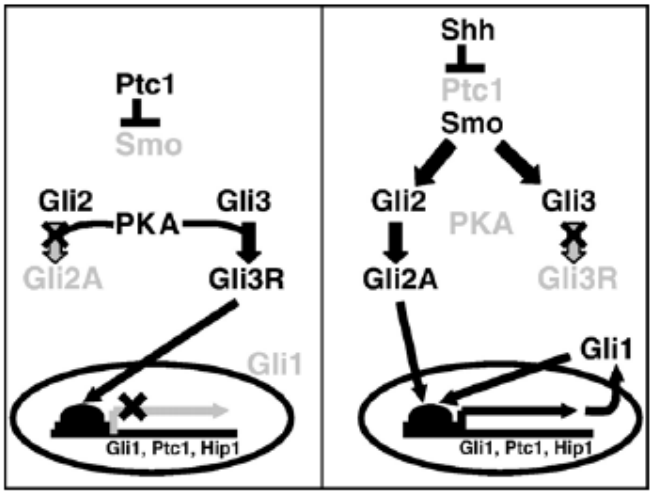


Figure 1.7: Upstream regulation of the Gli transcription factors and their individual and combined roles in regulating Hh target gene expression [Lipinski *et al.,* 2006]. In the absence of the signal, Gli3 functions as a repressor of the Hh signaling pathway. When there is signaling, repression of Gli3 is relieved, and activator form of the Gli2 becomes active and induces target gene activation.

Several antagonists of Hh signaling have been determined that regulate the Hh pathway during development of certain organs. The most important one of these is hedgehog interacting protein 1 (HIP1). It is also a target gene of the pathway, and

functions in feedback inhibition [Toftgard, 2000]. HIP1 binds to all Hh ligands with high affinity and sequesters them in the target cell membrane, thus preventing the signaling. HIP1 modulates the Hh signaling in vertabrates. Its binding to the ligand attenuates the signal, and allows differential responses to be generated [Chuang *et al.,* 1999]. Ectopic expression of HIP prevents the Hh signaling [Kayed *et al.,* 2005].

1.2.2. Hedgehog Signaling in Development

Cell-cell signaling is a crucial aspect of development and yet just five signal transduction pathways mediate the early development of most animals. These intercellular signaling pathways consist of the Wnt, TGF-β, Notch, RTK (receptor tyrosine kinase) and Hedgehog pathways. The most common target of signaling in development is transcription. Different pathways activate or repress different genes at distinct times and places in the embryo. Signaling pathways have important roles in determining embryonic patterning and cell fate decisions. Analysis of *Drosophila* development has been vital in elucidating the components and functions of these signaling pathways. Research in vertebrates revealed that vertebrates not only have the same signaling components as *Drosophila*, but also often the developmental roles of these signaling components are similar to those in *Drosophila*. Hedgehog signaling was first identified in *Drosophila* embryo development with a role in segmental patterning.

The three mammalian Hh genes are broadly expressed throughout embryonic development with Shh showing the widest range of expression [Bitgood *et al.,* 1995]. Hedgehog signaling pathway has many crucial functions in the growth, patterning, and morphogenesis of many regions in the vertebrates and insects as shown in Table 1.2. It can act both as a morphogen to induce different cell fates and as a mitogen to regulate cell proliferation. Signaling can be short- and long-range, direct and indirect, and importantly, concentration-dependent, evoking distinct molecular responses at discrete concentration thresholds. It functions in the segmental patterning in the fruit fly embryo, digit patterning in the chick limb bud, and left-right asymmetry of vertebrate embryos. In addition it is crucial in the maintenance of tissue patterns in

adults [Bijlsma *et al.,* 2004; Harfe *et al.,* 2004]. Hh signaling functions in limb and forebrain development by regulating organ size and cell proliferation [Watkins *et al.,* 2004]. Shh is involved in development of CNS [Dahmane *et.al,* 2001]. Shh acts as a morphogen throughout the dorso-ventral axis of the embryo along the ventral neural tube and induces expression of homeobox genes [Ruizi i Altaba et al., 2004]. It is required for differentiation of floor plate cells and ventral neurons in the early neural tube. In addition Hh signaling regulates stem cell fate in the cerebellum [Dahmane *et al.,* 2001; Ruizi i Altaba *et al.,* 2004; Watkins *et al.,* 2004]. In the limb, Shh is expressed from the cells found in the zone of polarizing activity (ZPA) and specifies the digit identity in the anterior-posterior axis. If the temporal gradient of Shh changes, digit identity also changes [Chuang *et al.,* 2000; Harfe *et al.,* 2004].

In studies of mouse embryo it has been shown that Hh signaling is required for foregut development. Experiments performed in both wild-type and *Shh*-/- mutant mice embryos show that *Shh*-/- mutant embryos have abnormal foregut and lung development. These results show that *Shh* is required for normal development of oesophagus and trachea and mutations of *SHH* and/or its signaling components may contribute to foregut defects in humans [Litingtung *et al.,* 1998].

Shh and *Ihh* mutant mice embryos show defects in gastrointestinal organogenesis [Ramalho-Santos *et al.,* 2000]. Hh signaling is important for anterior-posterior patterning, radial patterning, and epithelial stem cell differentiation and proliferation in gastrointestinal development. Radial axis patterning in the developing intestine is governed by the ratio of lamina propria and submucosa to smooth muscle and enteric neuronal cell neurons which is determined by Shh and Ihh expression [Watkins *et al.,* 2004]. In the developing mouse pancreas, increased Hh signaling prevents the organogenesis [Kawahira *et al.,* 2003]. Hh signaling regulates organ size, morphogenesis, and function of the pancreas.

Table 1.2: Vertebrate Hedgehog Functions [Ingham *et al.,* 2001].

Tissue/cell type/organ	Ligand	**Nature of role**
Angiogenesis/vasculogenesis	Ihh	Stimulates endothelial cell production in yolk sac
	Shh	Induces angiogenesis
Blood cells	Shh	Proliferation of stem cells, modulation hematopoietic and thymocyte differentiation
Bone and cartilage	Shh	Induction of early chondrogenic factors in somite proliferation/survival axial chondrogenic precursors
	Ihh	Coordination of proliferation/ differentiation in endochondral skeleton
Cerebellum	Shh	Proliferation of granule cell precursors, differentiation of Bergmann glia
Eye	Shh (Ihh, Dhh, twhh)	Stimulates and inhibits retinal neurogenesis astrocyte proliferation in optical nerve, retinal precursor proliferation
	Shh and/or twhh	Induction of proximal fates in the eye fields
Gut	Shh	Separation of trachea and esophagus, A-P patterning gut tube
	Ihh and/or Shh	Proliferation and inhibition of mesenchyme differentiation, radial patterning of gut tube
Gonads/external genitalia	Dhh	Peritubular cell development, maturation of testes, Sertoli Leydig cell interactions, male germ line development, masculinization
Hair/feather	Shh	Follice/feather morphogenesis, polarity of feather, Telogen to anogen switch in hair follicle
Heart	Shh or Ihh	Cardiac morphogenesis
Lateral Asymmetry	Shh, Shh or Ihh	Regulation of L/R asymmetry (left pathway)
Limbs	Shh	A-P patterning of skeleton, outgrowth of limb-bud
Lung	Shh	Branching epithelium, proliferation/ survival of mesenchyme
Muscle	Shh	Induction/proliferation/survival of epaxial muscle precursors, fiber-type identity, regulation of smooth muscle differentiation
Neural Crest	Shh	Survival of cranial neural crest, cranial facial morphogenesis, proliferation/ differentiation of sympathetic cells
Neurons	Shh	Induction of specific ventral neural cell types, proliferation/survival/death neural precursors
Olfactory	Shh	Olfactory pathway formation
Oligodendrocytes	Shh	Proliferation/differentiation/survival of precursors
Pancreas	Shh	Inhibition of pancreatic anlagen formation, insulin production Specification of pancreas anlagen (in fish)
Peripheral nerves	Dhh	Formation of peripheral nerve sheath
Pituitary	Shh	Inhibits transdifferentiation to lens, proliferation/cell type determination
Prostate	Shh	Growth and ductal morphogenesis
Tooth	Shh	Growth, polarity and morphogenesis

1.2.3. Hedgehog Signaling in the Mammary Gland Development

In mouse mammary gland development, both Shh and Ihh can be detected in the mammary epithelium in E12. After placode formation, Ptch1 expression can be observed in both ectodermal and mesenchymal tissues. Gli1 expression cannot be observed in early development. Only after E13.5 it is expression is detectable in the epithelium. Gli2 and Gli3 expressions can be seen after placode formation at E11.5-E12.5 in epithelial tissues [Lewis *et al.,* 2004; Michno *et al.,* 2003]. Gli3 is shown to be important for mammary placode formation although detectable expression is seen only after placode formation. In ductal development, *Ptch1* and *Gli2* genes have been shown to be important [Lewis *et al.,* 2004].

During the pregnancy and lactation periods Ptch is highly expressed. Haploinsufficiency of Ptch1 causes severe histological defects in ductal structure, and small morphological changes in terminal end buds in virgin mice [Lewis *et al.,* 2004; Michno *et al.,* 2003]. Gli2 is thought to be the main Hh signaling transcription factor in the mammary gland development. Gli2 null phenotype mice had misshapen ducts showing hyperplasia similar to human micropapillary ductal hyperplasia. While it is expressed in the stroma during mammary gland development, it is expressed in both stroma and epithelial tissues during pregnancy and lactation [Gallego *et al.,* 2002]. Single gene loss of function mutation studies in mice for *Shh, Ihh, Gli1*, and *Gli3* showed that these genes are not required for ductal development individually. Shh and Ihh may have redundancy in ductal development in virgin mice. Indeed, high expression of Ihh is observed in the Shh deficient mice. Loss of Shh expression does not have any effect on the expression profile of cyclin D1, a target of the pathway [Michno *et al.,* 2003].

The working model for the role of Hh signaling in ductal development is that in the ducts of the virgin mice Ptch1 inhibits the Smo in the presence of low levels of Ihh in the epithelium [Lewis *et al.,* 2004] (Figure 1.8). Thus, Gli2 cannot be expressed in that part of the mammary gland. However, in the stroma Ptch activates the pathway, and Gli2 acts as a transcriptional activator. Ptch1 can bind to Hh ligand and sequester

ligand away from the cells to prevent respond to stimulus. Loss of Ptch1 allows Hh ligand to signal to the stroma which is not exposed to high levels of ligand. In the absence of the Hh ligand Ptch1 can initiate apoptosis. If loss of Ptch1 occurs, it may lead to a failure in the apoptotic program, and cells accumulate in the duct.

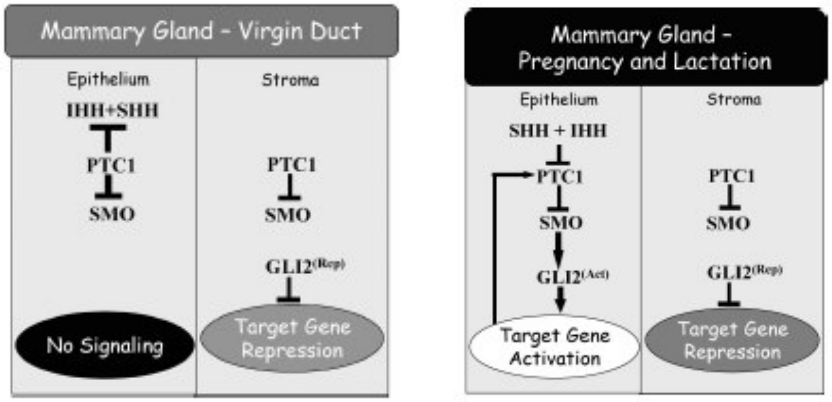


Figure 1.8: Hedgehog network model in virgin mice duct and alveolar development in pregnancy and lactation. Hypothesized functional relationships among network components are shown for both the stroma and the epithelium [Lewis *et al.,* 2004].

There is also a revised model for ductal development in which ductal development is ligand independent [Lewis *et al.,* 2004]. Pcth1 is expressed in both epithelium and stroma to prevent signaling in the presence of low level of ligand expression by Smo inhibition, sequestration of the ligand, and induction of apoptosis, or combination of these. In this case Gli2 activity should be regulated by other pathways like Wnt. If this is the case, then Gli2 should act as a transcriptional repressor.

Ihh expression seems to be hormonally regulated because expression increases dramatically in the pregnancy, and becomes maximal during the lactation [Walterhouse *et al.,* 2003]. Expressional regulation of Shh is not clear yet [Lewis *et al.,* 2004]. That is why Ihh is the primary active Hh ligand in mammary development. All three Hh ligand expression are observed in postnatal mice

mammary gland by RT-PCR, whereas in the *in situ* hybridization experiments only Ihh expression is detected, which proves the importance of the Ihh in the mammary development [Michno *et al.,* 2003]. In the alveolar epithelium Ptch1 mRNA expression increases with the pregnancy, but stays same in the stroma. Gli2 expression becomes both stromal and epithelial in the pregnancy and lactation. Increased expression of Ihh in the pregnancy may inhibit the Ptch1 and activate the downstream pathway. As a result Gli2 expression may be induced in the epithelium and acts as transcriptional activator for the target genes.

Effects of Hh signaling on the proliferation and differentiation of the mammary mesenchyme have not been described.

1.2.4. Hedgehog Signaling and Cancer

Distinct tissues require specific levels of Hh signaling for proper functioning. Any increase or decrease in the signaling activity results in severe defects. Although Hh signaling is one developmental pathway, it is expressed in some mature cell types and its deregulation leads to tumor progression in these tissues.

Table 1.3, Table 1.4 and Table 1.5 shows the Hh pathway with relation to cancer. The importance of hedgehog signaling in carcinogenesis has been demonstrated by the fact that many of the genes involving Hh signaling are known oncogenes, including *SMO, SHH, GLI1,* and *GLI2*, or that *PTCH1* can function as a tumor suppressor. Mutations in these genes causes to the development of many common cancers, which are shown to be dependent on activated Hh signaling. These mutations both include inactivating mutations in the *PTCH* and activating mutations in *SHH* and *SMO*.

GLI1 is first identified as an amplified gene in malignant glioma cells. Overexpression of the Gli1 in mouse skin results in rapid and spontaneous development of basal cell carcinoma (BCC), cylindromas, and trichoblastomas [Toftgard, 2000]. These tumors show increased expression of Ptch1 and Sufu.

Individuals with naevoid basal cell carcinoma syndrome (NBCCS) have an inherited mutation in one of the PTCH1 alleles. Cancer develops by inactivation of the remaining allele. In nearly 40% of sporadic BCC, a *PTCH1* mutation is detected. Most of the mutations lead to premature protein production [Wicking *et al.,* 2001]. In animal models, any activating mutation of the pathway is sufficient for the development of BCC. Ectopic expression of Shh, Gli1, or Gli2 results in lesions similar to human BCC [Wetmore, 2003].

Table 1.3: Animal models of Hedgehog-dependent tumors [Magliano *et al.,* 2003]

Animal model	**Phenotype**	**Species**
Shh overexpression in the skin K14 promoter	BCC	Mouse
Ptch inactivation	Medulloblastoma and other tumors	Mouse
Smo-M2 overexpression in the skin K5 promoter	BCC	Mouse
Gli1 overexpression in the skin	BCC, trichoepithelioma	Mouse
Gli2 overexpression in the skin	BCC	Mouse
Gli2 overexpression in the skin	Skin tumors with BCC-like characteristics	Xenopus
Gli1 overexpression in the brain	Hyperproliferation of progenitor cells	Xenopus

Human patients with Gorlin's syndrome have inherited mutations in *PTCH* gene (Table 1.4). Patients suffering from Gorlin's syndrome develop BCC and sporadic medulloblastomas [Ruizi i Altaba *et al.,* 2004]. *PTCH* mutations are also observed in desmoplastic sporadic medulloblastoma. Most of these mutations are loss of heterozygosity (LOH) [Toftgard, 2000]. There are additional mutated genes of the Hh pathway playing a role in development of medulloblastoma, in which mutations of *SUFU* is important. In 9% of sporadic human medulloblastomas truncating mutations of *SUFU* are observed. There is upregulation of *GLI1* and *GLI2* mRNA level in medulloblastomas showing the active pathway in these tumors.

Table 1.4: Mutations in the PTCH1 gene [Hahn *et al.,* 1996]

Sample Type	Inheritance	Exon	Type of Mutation	Designation
NBCCS	Familial	8	Premature stop	C1081T
NBCCS	Familial	6	37 bp deletion	del 804-840
NBCCS	Familial	8	Premature stop	G1148A
NBCCS	Familial	13	2 bp insertion	2047insC
NBCCS	Sporadic	13	1 bp insertion	2000insC
NBCCS	Sporadic	15	1 bp deletion	2583delC
BCC	Sporadic	8	Premature stop	CC1081TT
BCC	Sporadic	16	14 bp deletion	del2704-2717

Sporadic and inherited mutations of human *SHH* gene cause holoprosencephaly (HPE), a severe midline defect including cleft lip and palate, single maxillary incisor, impaired CNS septation, and phenotypes ranging from hypotelorism to a single cyclopic eye.

In normal conditions in the mature cell types, Hh signaling regulates cell proliferation through expressional regulation of cyclin D and cyclin E [Magliano *et al.,* 2003]. Myc transcription factors are transcriptional targets of the Hh signaling and they are important inducers of cell proliferation. Berman *et.al.* (2002) showed that expression of c-MYC, l-MYC, and n-MYC is decreased in medulloblastomas by cyclopamine treatment. Furthermore, PTCH regulates the activity of cyclin B, a part of mitosis-promoting-factor (MPF), by interacting in the cytoplasm, hence preventing the nuclear localization of the activated complex [Barnes *et al.,* 2001; Magliano *et al.,* 2003]. In response to SHH stimulation PTCH relieves the cyclin B, and the cell enters mitosis. Dysregulation of Hh signaling results in uncontrolled cell proliferation in these cells.

Several foregut-derived tumors require ligand-dependent activation of the pathway (Table 1.5). In cancers of lung, esophagus, stomach, and pancreas SHH and IHH are expressed from the tumor cells and taken by the same cells as well [Watkins *et al.,* 2004; Berman *et.al,* 2002].

Table 1.5: Hedgehog-dependent primary human tumors

Gene	Tissue	Expression	Method	% of sample	References
SHH	Pancreas	upregulated	In situ, IHC	70% in 20	Thayer *et al.,* 2003
PTCH1, GLI1	Stomach	upregulated	In situ	64% in 99	Ma *et al.,* 2005
PTCH1	Prostate	upregulated	IHC	70% in 22 high score	Sheng *et al.,* 2004
GLI1	Lung (SCLC)	upregulated	IHC	65% in 40	Vestergaard *et al.,* 2006
SHH	Colon	upregulated	q-rt-RT-PCR	82% in 57 (t-n)	Manzo *et al.,* 2006
GLI1	Breast	upregulated	IHC	100% in 52	Kubo *et al.,* 2004
SHH	Colon	upregulated	RT-PCR	86% in 44 (t-n)	Douard *et al.,* 2006
GLI1	Skin	upregulated	q-rt-RT-PCR	47% in 68	Hatta *et al.,* 2005
SHH	Pituitary adenoma	downregulated	IHC	100% in 55	Vila *et al.,* 2005
SHH	Liver	upregulated	In situ	60% in 115	Huang *et al.,* 2006
PTCH1	Liver	upregulated	In situ	56% in 115	Huang *et al.,* 2006
GLI1	Liver	upregulated	In situ	70% in 115	Huang *et al.,* 2006
PTCH1	Pancreas	upregulated	RT-PCR, IHC	71% in 28	Shao *et al.,* 2006
SMO	Pancreas	upregulated	RT-PCR, IHC	54% in 28	Shao *et al.,* 2006
IHH	Pancreas	upregulated	q-rt-RT-PCR	100% in 31	Kayed *et al.,* 2004
PTCH1	Skin	upregulated	q-rt-RT-PCR	100% in 27	Bonifas *et al.,* 2001
GLI1	Skin	upregulated	q-rt-RT-PCR	100% in 12	Bonifas *et al.,* 2001

Despite the liver's requirement for Hh signaling during embryogenesis, mature hepatocytes lack Hh pathway activity [Sicklick *et al.,* 2006]. Expression of SHH was detectable in about 60% of 115 cases of hepatocellular carcinoma (HCC) examined. Furhermore, Hh target genes PTCH1 and GLI1 were expressed in over 50% of the tumors, concluding that the Hh pathway is frequently activated in HCCs [Huang *et al.,* 2006]. Overexpression of the SMO, as well as an increase in the stoichiometric ratio of SMO to PTCH1 mRNA levels, correlated with tumor size. Overexpression and/or tumorigenic activation of the SMO mediate c-MYC overexpression which plays a critical role in HCC progress and makes SMO a good prognostic factor in HCC tumorigenesis [Sicklick *et al.,* 2006]. Quantitative real-time RT-PCR also revealed the decreased expression of HIP1 and increased expression of GLI1 and SMO in HCC samples compared with nontumor liver tissues [Patil *et al.,* 2006].

Hh signaling is also required for the maintenance of the tumor. Transformed cells require Hh activity for survival and growth. Cyclopamine treatment induces apoptosis and loss of proliferation in the cancer cells. Cyclopamine is a natural steroidal alkaloid that induces Hh blockage without interfering with the generation or processing of the ligand [Cooper *et al.,* 1998; Incardona *et al.,* 1998]. It directly binds to the heptahelical bundle of the Smo protein [Watkins *et al.,* 2004].

1.2.5. Hedgehog Signaling and Breast Cancer

Several signaling pathways function in the regulation of mammary stem cells such as Wnt, Notch, Hedgehog, and TGF-β. Dysregulation of these pathways in the mammary gland induce mammary tumors in transgenic mice [Li *et al.,* 2005].

Similarities between hedgehog mutation-induced ductal dysplasias and human breast pathologies suggest a role for altered hedgehog signaling in the development of mammary cancer [Lewis *et al.,* 2004]. There is also evidence that altered hedgehog signaling has a direct role in the neoplastic progression of the mammary gland. One study showed *PTCH1* mutation in two of seven human breast cancers [Xie *et al.,* 1997]. Recently, a polymorphism in the 3'end of the *PTCH1* coding region (C3944T;Pro1315→Leu) has been linked to increased breast cancer risk associated with oral contraceptive use [Chang-Claude *et al.,* 2003]. Evidence for a role in breast cancer also comes from published genetic studies in mice showing hyperplastic defects in the mammary gland of ΔPtch1 + and ΔGli1 mutants [Lewis *et al.,* 2004]. When compared to adjacent tissues breast cancer samples have higher intensity GLI1 staining in immunohistochemistry experiments. Cyclopamine is able to inhibit the growth of mammary carcinoma cells *in vitro* [Kubo *et al.,* 2004].

1.3. MEASUREMENTS OF GENE EXPRESSION WITH QUANTITATIVE REAL TIME RT-PCR

Many applications in medicine or research require detection of the number of specific targets in the specimen [Mocellin *et al.,* 2003]. Northern blotting, RNase protection assays, and polymerase chain reaction (PCR)-based methods are applied to quantify mRNA amount. The most sensitive one of these methods is PCR based ones with its combination with reverse transcription (RT) [Fronhoffs *et al.,* 2002]. Since it requires a minute amount of mRNA it is more advantageous to perform PCR rather than other methods.

PCR is the method to detect as little as a single copy of a particular sequence of DNA and RNA. In theory there is a quantitative relationship between the amount of starting sequence and amount of PCR product at a given cycle. This is not the case in practice since the PCR product increases in every cycle. Reverse transcription PCR (RT-PCR) is a kind of PCR used to detect expression level of a gene at mRNA level [Pfaffl, 2001]. Gene expressional differences between tissues, disease states, and treatments can be revealed by using quantitative RT-PCR [Pfaffl, 2001; Bustin *et al.,* 2004a].

Quantitative real-time RT-PCR (q-rt-RT-PCR) is one of the methods that is developed for quantitative measurement of the gene expression levels [Pfaffl, 2001; Mocellin *et al.,* 2003]. It has many advantages over the traditional RT-PCR [Fronhoffs *et al.,* 2002; Wong *et al.,* 2005]. Firstly it is simple, has a wide dynamic range of quantitation and allows high throughput screening at one time. Secondly, it is reproducible, accurate, specific, and sensitive. It can detect as few as 10 initial template copies [Fronhoffs *et al.,* 2002]. Besides, gene expression differences as small as 23% between samples can be detected by q-rt-RT-PCR [Wong *et al.,* 2005].

In a PCR reaction there are four major phases: the linear ground phase, early exponential phase, log-linear (exponential) phase, and plateau phase [Wong *et al.,* 2005]. In the linear ground phase PCR just begins and fluorescence does not exceed background. Calculation of baseline fluorescence is performed at that phase (Figure 1.9). In early exponential phase fluorescence exceeds the background and Ct value is measured. The Ct value is used as measurement unit in the real-time RT-PCR. It is the cycle number at which fluorescence reaches a threshold value of ten times the standard deviation of baseline fluorescence emission [Mocellin *et al.,* 2003; Wong *et al.,* 2005]. The Ct value is inversely proportional to the amount of starting material. The threshold value is the point at which a reaction reaches a fluorescent intensity above background. It should be in the linear part of the reaction. The program automatically determines the Ct value of the sample. In the third phase, the PCR reaction reaches its optimal amplification period in which PCR doubling in each cycle is ideal. In the plateau phase reaction components become rate limiting agents.

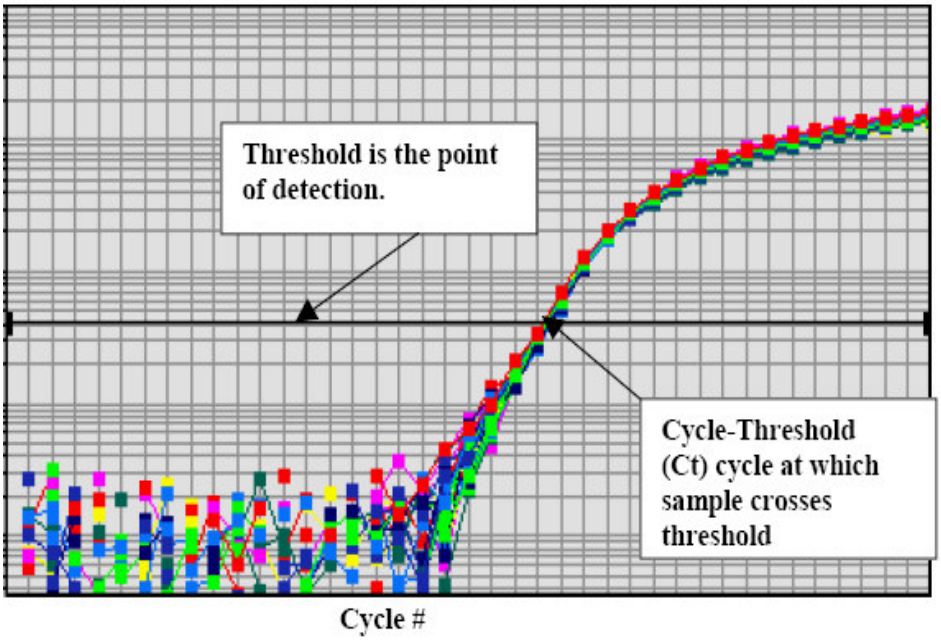


Figure 1.9: Amplification Curve: When fluorescent signal reaches to detectable level it is displayed as an amplification curve. The point at which the amplification curve reaches to the threshold level is called the Ct value [http://www.appliedbiosystems.com]

There are several considerations when designing primers for q-rt-RT-PCR. The optimal PCR product length is approximately 100 to 200bp for SYBR Green. The forward and the reverse primer should have similar melting temperatures within 0.5^{o}C of each other. The primers should have low or no self complementarities in order to avoid the formation of primer dimers. Primers that span introns or cross intron/exon boundaries are advantageous because they allow the distinction of cDNA from genomic DNA contaminations.

1.3.1. Detection Chemistries

There are specific and non-specific detection chemistries used in real-time RT-PCR. In non-specific method, DNA intercalating dyes are used. Probes and molecular beacons are the molecules used in the specific chemistry. In the specific chemistry, fluorescent resonance energy transfer (FRET) or similar interactions between the donor and quencher molecules form the basis of detection system [Bustin *et al.,* 2004a].

The most simple detection method in real-time RT-PCR requires a dye that emits fluorescent light when intercalated into double-stranded DNA (dsDNA) but not to single-stranded DNA. The intensity of the fluorescence signal is proportional to the amount of all double-stranded DNA present in the reaction [Wong *et al.,* 2005]. In the first experiments ethidium bromide and YO-PRO-1 were used as intercalating dyes. However, by ethidium bromide it is not possible to distinguish between the amounts of the ssDNA and the dsDNA. Currently, SYBR Green I is the most frequently used one. SYBR Green I is a binding dye for ds DNA. It binds to the minor groove of the ds DNA and increases the fluorescence over a hundred fold. SYBR Green I is preferred to other methods in most cancer studies due to the usage of high number of genes and sample input. It is more precise, and produces more linear decay plot than the TaqMan detection system [Bustin *et al.,* 2004b]. Since these intercalating dyes can bind to every ds-DNA this technique can be used in

different gene assays [Wong *et al.,* 2005]. In addition they are cheaper because of the lack of probe associated cost [Bustin *et al.,* 2004a].

1.3.2. Quantification Strategies

The quantification strategy is the most important part of the real-time RT-PCR. It depends on the target sequence, expected amount of the mRNA, and degree of accuracy required [Pfaffl, 2001].

Two strategies can be performed in real-time RT-PCR: absolute and relative quantitative real-time RT-PCR [Pfaffl, 2001]. Absolute quantification relates the PCR signal to input copy number using a calibration curve, while relative quantification measures the relative change in mRNA expression levels. The reliability of an absolute real-time RT-PCR assay depends on the condition of 'identical' amplification efficiencies for both the native target and the calibration curve in RT reaction and in the following kinetic PCR. Therefore, it should be used in the analysis where the determination of exact copy number is necessary. In most analyses it is enough to determine the relative change in the expression of the gene. In these cases, relative quantification is easier to perform than absolute quantification since there is no need for a calibration curve. It is based on the expression levels of a target gene versus a housekeeping gene (reference gene). The units used to express relative quantities are irrelevant and the relative quantities can be compared across multiple real-time RT-PCR experiments.

Relative quantification determines the changes in steady-state mRNA levels of a gene across multiple samples and expresses it relative to the levels of an internal control RNA [Pfaffl, 2001]. This reference gene is often a housekeeping gene and can be co-amplified in the same tube in a multiplex assay or can be amplified in a separate tube [Huggett *et al.,* 2005]. Therefore, relative quantification does not require standards with known concentrations and the reference can be any transcript, as long as its sequence is known. To calculate the expression of a target gene in relation to an adequate reference gene various mathematical models are established.

Calculations are based on the comparison of the distinct cycle determined by various methods, e.g. crossing points (CP) and threshold values (Ct) at a constant level of fluorescence; or CP acquisition according to established mathematic algorithm. To date, several mathematical models that calculate the relative expression ratio have been developed. Relative quantification models without efficiency correction are available (equations 1.1, 1.2).

$$\text{Equation 1.1:} \quad R=2^{-[\Delta Ctsample-\Delta Ctcontrol]}$$

$$\text{Equation 1.2:} \quad R= 2^{-\Delta\Delta Ct}$$

In equation 1.2, the first ΔCt is for the target gene and the second ΔCt is for the housekeeping gene.

With kinetic PCR efficiency correction, the relative expression ratio of a target gene is calculated based on its real time PCR efficiency (E) and the difference between Ct (CP) values of the unknown sample versus the control sample. By dividing expression ratio of the target gene to the reference gene relative quantification is performed [Pfaffl, 2001].

$$\text{Equation 1.3:} \quad R= [(E_{\text{target}})^{\Delta Ct_{\text{target}}(\text{control-sample})}] / [(E_{\text{reference}})^{\Delta Ct_{\text{reference}}(\text{control-sample})}]$$

For a standard curve with ½ serial dilutions:

$$\text{Equation 1.4: Exponential amplification} = 2^{(-1/\text{slope})}$$

$$\text{Equation 1.5: Efficiency} = [2^{(-1/\text{slope})}]-1$$

Since there is no need for an external calibration curve, hence artificial nucleic acids, relative expression studies are rapid and sensitive methods. They are adequate for most studies of physiological expression changes with minimal error rate [Pfaffl, 2001].

1.3.3. Normalization

Housekeeping genes are expressed in all nucleated cells and required for cell survival [Thellin *et al.,* 1999]. Thus, accurate normalization of quantitative RT-PCR data requires knowledge of which housekeeping gene or genes are expressed at equal or similar levels within a group of samples [Vandesompele *et al.,* 2002]. Many techniques developed for mRNA quantification use housekeeping genes as internal standards. Beta actin (ACTB) and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) have been historically considered to be adequate housekeepers for normalization of gene expression. However, studies show that their expression may be regulated as well between different tissues [Bereta *et al.,* 1995; Chang *et al.,* 1998]. This is partly explained by their participation to other functions in different cell types [Singh *et al.,* 1993; Ishitani *et al.,* 1996]. GAPDH has been found to be regulated by biphosphonates in breast cancer cell lines [Valenti *et al.,* 2006]. The researcher should find the most suitable reference gene for his/her experimental setup. For this purpose, several excel based programs have been developed [Huggett *et al.,* 2005]. The program, geNorm, finds the best reference gene by using the geometric mean of the reference gene expression of sample cDNAs. The software is free and can be downloaded from http://medgen.ugent.be/~jvdesomp/genorm/. BestKeeper and Norm-Finder are other software developed for the same purpose [Huggett *et al.,* 2005]. They also use the geometric mean of the expressions. In addition, Norm-Finder ranks the potential reference genes by how much they differ between the samples. It is available at http://www.mdl.dk.

1.3.4. Melt Curve Analysis

In real-time PCR, the melting point of the PCR product is determined by raising the temperature by a fraction of a degree and measuring the changes in fluorescence [Ririe *et al.,* 1997]. Plotting the fluorescence as a function of temperature during the thermal cycler heats up to dissociation temperature gives the melting curve. The software plots the rate of change of the relative fluorescence units (RFU) with time (T) (-d (RFU)/dT) on the Y-axis versus the temperature on the X-axis, and this will

peak at the melting temperature (Tm) [http://pathmicro.med.sc.edu/pcr/realtime-home.htm]. The melting point of a DNA sample depends on its base composition and length. Therefore all PCR products for a particular primer pair should have the same melting point. If there are different melting points that means there is contamination, mispriming, and primer-dimer artifacts. In real-time PCR reactions performed with SYBR green specificity is provided by drawing the melting curve.

1.4. AIM AND STRATEGY

Signaling pathways having role in development cross talks each other and deregulation in one pathway affects the others. For this reason it is important to study developmentally important signaling pathways in cancer since deregulation of these pathways is crucial for cancer development. One pathway that is important in regulating patterning, proliferation, survival and growth in the embryo and adults is the Sonic Hedgehog (Shh)-Gli signaling pathway. Uncontrolled activation of this pathway results in cancer development.

We aimed to identify the expression profile of the genes that play major roles in the Hedgehog signaling pathway in epithelial cancers; breast, colon, and hepatocellular carcinoma by using real time quantitative-reverse-transcriptase PCR technique.

The strategy was as follows:

- Seven genes that play major roles in the Hh pathway (*SHH, IHH, SMO, PTCH1, GLI1, GLI2* and *GLI3*), and a downstream target of Hh pathway which has an important role in cancer progression, *BCL2,* and four house keeping genes, *GAPDH*, *ACTB, TBP,* and *SDHA,* were selected for the gene expression analysis.
- Fifteen breast, 14 colon, and 15 hepatocellular carcinoma (HCC) cell lines and 29 primary breast tumor samples and three matched normal tissue sample pools were used in this study.

- The total RNAs were isolated from each sample and used for OligodT primed-cDNA synthesis.
- The gene specific primers were designed for q-rt-RT-PCR analysis.
- Primer amplification efficiencies were determined for each primer pair.
- Melt curve analysis of each gene were performed with q-rt-RT-PCR by using Bio-Rad icycler machine for each sample.
- Quantification of mRNA expression values of each gene was calculated by using efficiency corrected relative quantification method.
- Geometrical mean of *ACTB, SDHA,* and *TBP1* reference gene expression values were used for normalization of the breast tissue samples and GADPH expression values were used for normalization of the cell lines.
- Multivariate statistical methods such as Cluster, Discriminant Function Analysis, and Multidimensional Scaling were used to analyze the data.

CHAPTER 2

MATERIALS & METHODS

2.1. TISSUE CULTURE

Breast carcinoma, hepatocellular carcinoma (HCC), and colon carcinoma cell lines used in this thesis are shown in Table 2.1, Table 2.2, and Table 2.3 respectively. All cell lines were grown in 150 mm culture dishes as monolayer. Cells were grown in Dulbecco's Modified Eagle Medium (DMEM) supplied with 10% FBS and 50mg/ml penicillin/streptomycin.

HME1 (Clontech, California, USA), a normal epithelial cell line immortalized with hTERT expression, was used as normal control for breast cancer cell lines. The growth medium of HME1 was different from others and prepared with equal amounts of HAM'S F-12 Medium and DMEM supplemented with 10% FBS, 100U/ml penicillin, 100µg/ml streptomycin, 2mM L-glutamine, 0.1mM non-essential amino acids, 10mM HEPES Buffer, 3.5ng/ml insuline, 0.1mg/ml EGF, and 0.5µg/ml hydrocortisone.

Table 2.1: Breast carcinoma cell line information

Breast Cancer Cell Lines	ATCC Number	Cancer Type	ER Status[a]
MDA MB 157	HTB 24	Medulallary carcinoma	N
MDA MB 231	HTB 26	Adenocarcinoma	N
MDA MB 361	HTB 27	Adenocarcinoma	P
MDA MB 453	HTB 131	Metastatic carcinoma	N
MDA MB 468	HTB 132	Adenocarcinoma	N
BT 474	HTB 20	Ductal carcinoma	P
BT 20	HTB 19	Carcinoma	P
MCF 7	HTB 22	Adenocarcinoma	P
MCF 12A	CRL 10782	Non-tumorigenic epithelial cell line	P
HCC 1937	CRL 2336	Primary ductal carcinoma	N
CAMA 1	HTB 21	Adenocarcinoma	P
SK BR 3	HTB 30	Adenocarcinoma	N
ZR 75.1	CRL 1500	Ductal carcinoma	P
T47D	HTB 133	Ductal carcinoma	P
HBL100	HTB 124	Tumorigenic epithelial cell line	N
Htert- HME 1	CRL 4010	Epithelial; immortalized with hTERT	N

[a]ER: Estrogen receptor

Table 2.2: Hepatocellullar carcinoma cell line information

HCC Cell Lines	ATCC Number	Differentiation Status
SNU182	CRL-2235	Poor
SNU387	CRL-2237	Poor
SNU398	CRL-2233	Poor
SNU423	CRL-2238	Poor
SNU449	CRL-2234	Poor
SNU475	CRL-2236	Poor
HEP3B	HB 8064	Well
HEP3B TR	M. Ozturk	Well
HEP40	M. Ozturk	Well
HEPG2	HB 8065	Well
SKHEP1	HTB52	Poor
HUH7	JCRB database	Well
PLC/PRL/5	CRL-8024	Well
FOCUS	M. Ozturk	Poor
MAHLAVU	M. Ozturk	Poor

Table 2.3: Colon carcinoma cell line information

Colon Cancer Cell Lines	ATCC Number
SW620	CCL-227
SW480	CCL-228
SW48	CCL-231
HCT8	CCL-244
HCT15	CCL-225
HT29	HTB-38
TC71	R. Hamelin
T7	R. Hamelin
LOVO	CCL-229
LS411	CRL-2159
COLO205	CCL-222
CO115	R. Hamelin
W05	R. Hamelin
KM12	R. Hamelin

All cell lines were cultured at a 37^{o}C incubator with an atmosphere of 5% CO_2 in air (Heto-Holten, Surrey, UK). All applications were performed under sterile laminair hoods (Heto-Holten, Surrey, UK) in cell culture facility. All media and solutions used for tissue culture were kept at 4^{o}C and preheated to 37^{o}C before use.

2.1.1. Cryopreservation of cell lines

Exponentially growing cells were harvested by trypsinization and neutralized by adding fresh growth medium. The cells were counted and precipitated at 1500 rpm for 5 minutes. The cells were resuspended with freezing medium at a concentration of $4x10^6$ cells/ml/vial. Freezing medium was made up 90% FBS and 10% DMSO and stored in the cryotubes. Cryotubes were incubated at -20^{o}C for 1 hour, at -80^{o}C overnight and transferred to the liquid nitrogen tank for long term storage.

2.1.2. Cell Line Culturing

The frozen stocks of the cell lines were kept in a liquid nitrogen tank for long term storage. The cells were thawed quickly at 37^{o}C and mixed with 5 µl growth medium and centrifuged at 1500 rpm for 5 minutes at room temperature. Supernatant was

removed and the cells were resuspended with fresh complete growth medium. Cells were grown in the 10 mm cell culture plates in the incubator.

2.1.3. Subculturing the cells

The cells were grown at a confluence of approximately 80%. The old medium was aspirated and the cells were washed with 1X PBS two times. 2µL Trypsin/EDTA solution was added and the cells were left for incubation for 3 to 4 min. When cells were detached from the surface of the plate, trypsin was inactivated by adding fresh growth medium. Cells were then transferred into the new plates.

2.1.4. Preparation of cell pellets

The cell pellets were collected to freeze down and stored for mRNA or protein preparation. When cells were 70-80% confluent, medium was aspired and cells were washed with ice cold 1X PBS once. Then 3ml of cold 1X PBS was added and cells were collected by using a scrapper. Then cells were centrifuged at 1500 rpm for 5 minutes at $4^{o}C$ and supernatant was removed. Cell pellet were stored at $-80^{o}C$ for further experiments.

2.1.5. Tumor Samples

The primary breast tumor and matched normal tissue samples were obtained from Ankara Numune Education and Research Hospital. The tissue samples were collected freshly after surgical operation and then snap frozen in liquid nitrogen. Tissue specimens were stored at $-80^{o}C$. Then the frozen tissues were used for further experiments. The pathological analysis of primary breast tumor samples is given in the Table 2.4.

Table 2.4: Primary breast tumor sample information

Name	Cancer Type	[b]LN	Grade	Stage	Age	[e]ER status	[f]PR status	ErbB2 status	[g]Tumor %
MFT14	[a]IDC	[c]P		4	40	P	P		90
MFT16	IDC	[d]N	2	2A	39	P	P	N	100
MFT21	IDC	P	3	4	60	N	N	N	90
MFT25	IDC	P	3	3A	51	P	P	N	100
MFT40	IDC	P	3	4	48				100
MFT41	IDC	P	2	2B	28	N	P	P	100
MFT49	IDC	P	2	2A	43	P	P	P	100
MFT78	IDC	P	2		69	P	P	P	60
MFT79	IDC	P	3	2A	24	N	N	P	90
MFT90	IDC	P	2	2B	32	P	P	P	100
MFT94	IDC	N	3	2A	37	N	N	N	100
MFT95	IDC	P	3		54	P	P	N	80
MFT97	IDC	P	2	3B	43				80
MFT113	IDC	P	1	3B	63	P	N	P	100
MFT115	IDC	N	3	2A					80
MFT116	IDC	P	1	1	74	N	N	P	90
MFT117	IDC		2	4		N	N	P	90
MFT120	IDC	N	2	2B	50				100
MFT124	IDC	P	1	2A	57	P	N	P	90
MFT125	IDC	P	3	2A	30	N	N	P	100
MFT127	IDC	P	2	2A		N	N		50
MFT129	IDC	P	2	2A	58	N	N	N	80
MFT131	IDC	P	2	2B	53	N	P	P	90
MFT132	IDC	N	3	2B	68	N	N	P	100
MFT149	IDC	P	3	2B	56				95
MFT154	IDC	P	3	3B	32	N	N	P	95
MFT155	IDC	N	3	1	54				95
MFT173	IDC	P	3	3B	60	P	P	N	90
MFT174	IDC	P	1	2B	44	N	N	P	90

[a]IDC: Infiltrating ductal carcinoma; [b]LN: lymph node; [c]P: positive; [d]N: negative; [e]ER: estrogen receptor; [f]PR: progesterone receptor; [g]Tumor %: tumor percentage in pathological sections

The tissue sections were cut by using a cryostat and 5µm sections were transferred on to the slides. The slides were stained with hematoxylin for 2 minutes washed under the tap water and rinsed with acid alcohol (1%) respectively. Then the slides were gently washed with 1% ammonia water (v/v) and rinsed under the tap water and stained with Eosin for one minute followed by tap water rinses. The slides were immersed in 70%, 90% and 100% ethanol respectively and air-dried. As a last step

the slides were rinsed in xylene and mounted with coverslips with mounting medium. The samples were sent to a pathologist for the evaluation of the tumor percentage in the tissue sections. The tumor samples with high tumor percentage and no tumor cells in the normal counterparts were chosen for further analysis.

2.2. RNA ISOLATION

Before RNA isolation all the solutions and materials were treated with DEPC in order to avoid RNase contamination and hence degradation of RNA. Total RNA of tumor tissues was isolated with TRI reagent (AppliChem, Germany). Tissues were lysed in 1ml TRI reagent and passed through a 21-gauge needle several times. After 5 min incubation at room temperature, 0.2ml chloroform was added per ml of TRI reagent. Tubes were shaken vigorously by hand for 15 seconds and incubated at room temperature for 2-3 min. After incubation the mixture was centrifuged at 12000xg for 15 min at 4^{o}C and then aqueous phase was collected into a new tube. 0.5ml isopropanol was added per 1ml of TRI reagent. The mixture was incubated at room temperature for 10 min and then centrifuged at 12000xg for 10 min at 4^{o}C to recover RNA. The supernatant was removed and pellet was washed with 75% ethanol twice. Mixture was centrifuged at 7500xg for 5 min at 4^{o}C and pellet was air-dried. Dried pellet was dissolved in ddH_2O and to remove contaminating DNA RNA isolation was repeated by use of NucleoSpin RNA II kit (Macherey & Nagel, Germany). The total RNA isolation of cell line pellets was performed directly by use of NucleoSpin RNA II kit. Isolated RNAs were stored at -80^{o}C. The concentrations of RNA samples were measured by ND-1000 Spectrophotometer (Nanodrop Technologies, Inc., Wilmington, USA).

2.3. cDNA PREPARATION

The cDNA samples were synthesized from the total RNA samples with the RevertAid First Strand cDNA Synthesis Kit (MBI Fermentas, Ontario, Canada) according to manufacturer's protocol. All cDNAs were made from 1 μg of RNA and stored at -20^{o}C.

2.4. OLIGONUCLEOTIDES

The gene specific primers used in RT-PCR and real-time RT-PCR were designed by Primer 3 program [http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi] and purchased from Iontek Inc. (Bursa, Turkey). The forward and reverse primers were designed to intron/exon boundaries wherever it was possible, otherwise designed to take an intronic sequence between them. The primer sequences are given in Table 2.5.

Table 2.5: Oligonucleotide sequences

Gene	Accession Number	Sequence (5' to 3')	Size (bp)[a]
GLI1-F	NM_002181	CACCGAGGGCCCACTCTTTT-3'	154
GLI1-R		TCGCGAGTTGATGAAAGCTA-3'	
GLI2-F	NM_005270	CCCACTCCAACGAGAAACCC-3'	96
GLI2-R		GGACCGTTTTCACATGCTTCC-3'	
GLI3-F	NM_000168	CAGCAAGTGGCTCCTATGGT	116
GLI3-R		TTGTCGGCTTAGGATCTGCT	
SHH-F	NM_000193	GTGGCCGAGAAGACCCTA	177
SHH-R		CAAAGCGTTCAACTTGTCCTTA	
IHH-F	NM_002181	TCACCACATCAGACCGCGAC	127
IHH-R		GAGTGCTCGGACTTGACGGA	
PTCH1-F	NM_000264	CTCACGTCCATCAGCAATGT	147
PTCH1-R		CATGCTGAGAATTGCAGGAA	
SMO-F	NM_005631	CTGCACGAATGAGGTGCAGA	149
SMO-R		GCTCAGCCTCTGTGAAGAGC	
BCL2-F	NM_000657	ACTGAGTACCTGAACCGGCA	123
BCL2-R		CTTCAGAGACAGCCAGGAGA	
GAPDH-F	NM_002046	GGCTGAGAACGGGAAGCTTGTCAT	150
GAPDH-R		CAGCCTTCTCCATGGTGGTGAAGA	
TBP-F	NM_003194	TGCACAGGAGCCAAGAGTGAA	132
TBP-R		CACATCACAGCTCCCCACCA	
ATCB-F	NM_001101	CCAACCGCGAGAAGATGACC	124
ACTB-R		GGAGTCCATCACGATGCCAG	
SDHA1-F	NM_004168	TGGGAACAAGAGGGCATCTG	86
SDHA1-R		CCACCACTGCATCAAATTCATG	

[a]bp: base pair

2.5. REVERSE TRANSCRIPTION POLYMERASE CHAIN REACTION (RT-PCR)

Primers designed for the real-time RT-PCR were first checked with PCR to determine their optimal working conditions. Polymerase Chain Reaction (PCR) was performed to amplify the desired DNA fragments from cDNAs using the thermal cycler TechGene (Techne Inc., New Jersey, USA). A reaction mixture of 2.5ml 10X reaction buffer, 1.5ml $MgCl_2$ (25mM), 0.5ml dNTP (10µM), 1ml of each primer (10 pmol), and 0.5ml Taq DNA polymerase (5u/µL) was prepared per 1ml cDNA.

Initial denaturation	95.0°C	5 min	
Denaturation	95.0°C	30 sec	30 cycles
Annealing	A.T.	30 sec	
Extension	72.0°C	30 sec	
Final extension	72.0°C	5 min	

A.T. is the annealing temperature at which the primer pair binds to its specific gene.

2.6. QUANTITATIVE REAL TIME RT-PCR

Amplification efficiency for each primer was tested by using standard two-fold dilution series of the breast cancer cell line cDNAs. The general approach is to choose a cell line that shows the highest expression level for a particular gene to determine the amplification efficiency. The expression of the genes varies between the cell lines. For each gene a different cell line was used for amplification efficiency analysis. The real-time RT-PCR reactions were carried out in iCycler (Bio-Rad, California, USA). After the amplification steps, a melting curve analysis was by performed raising the temperature by 0.5°C and measuring the change in fluorescence. Starting temperature was 55°C and ending temperature was 94°C. Each PCR reaction was made in duplicates from the same reaction mixture.

Six candidate reference genes which belong to different functional classes were chosen for stability analysis, which are *ACTB, SDHA, TBP, HPRT, GAPDH*, and *HMBS*. 30 tumor and matched normal tissues were analyzed by geNorm and Norm-Finder software tools to find out most stably expressed housekeeping genes for breast tumor samples studied. *ACTB, SDHA*, and *TBP* were the housekeeping genes with most stable expression pattern in primary breast tumor samples.

Analysis of relative expressions were performed by use of the comparative C_t method. For the cell lines, *GAPDH* was used as reference gene. For the breast cancer tumor samples, geometric mean of the expression values of *ACTB*, TATA-binding protein (*TBP*), and succinate dehydrogenase complex subunit A (*SDHA*) were used as reference gene. For the colon and hepatocellular carcinoma cell lines, geometric average of expression level of these cell lines were used as the control sample. hTERT cell line was used in the breast cancer cell lines, and the geometric mean of three normal sample pools was used in the breast tumor samples as the control sample.

Table 2.6: Breast normal sample pool information

Pool Name	Normal braest samples
MFN P1	MFN21, MFN41, MFN 49
MFN P2	MFN95, MFN113, MFN116
MFN P3	MFN120, MFN125, MFN129

The real-time RT-PCR reaction mixture was made up of 6.25µl SYBR Green mix (Bio-Rad, California, USA), 0.5µl of each primer (10 pmol), and 4.25µl ddH_2O. 12.5µl mineral oil was added to cover top of the mixture to prevent evaporation.

Initial denaturation	94.0°C	3 min	
Denaturation	94.0°C	30 sec	50 cycles
Annealing	A.T.	30 sec	
Extension	72.0°C	30 sec	

2.7. AGAROSE GEL ELECTROPHORESIS

PCR products were fractionated by horizontal gel electrophoresis in 2% (w/v) agarose gel by using 1xTAE buffer. Agarose was completely dissolved in 1xTAE electrophoresis buffer in the desired percentages. Ethidium bromide solution was added to final concentration of 30ng/ml for visualization. 2 µl 6X DNA loading dye was added to 10 µl of quantitative real time RT-PCR (q-rt-RT-PCR) products and 12 µl of normal PCR products and total volume was loaded to each well and run at 100 V for 30 minutes. Puc8 marker (MBI Fermentas, Ontario, Canada) was used as DNA size marker. Transilluminator (Bio-Rad, California, USA) was used to visualize the DNA bands under ultraviolet light (long wave, 340 nm). MultiAnalyst (Bio-Rad, California, USA) software was used to take photographs of the gels.

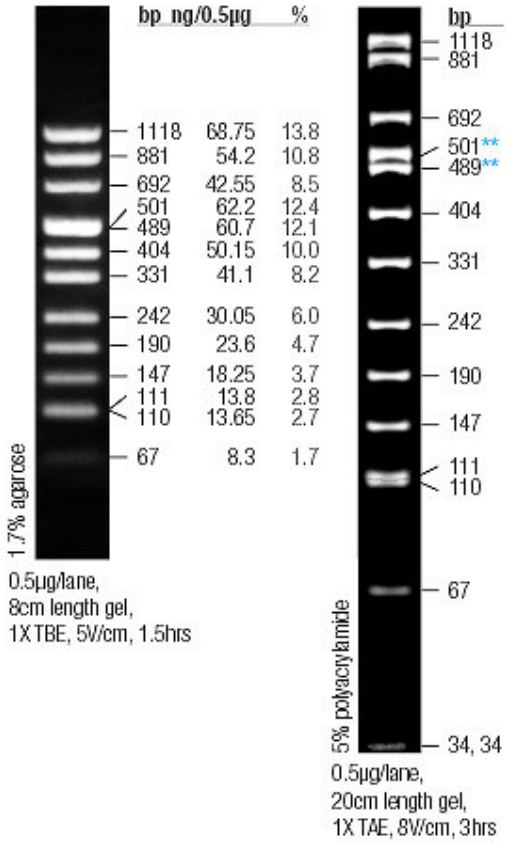


Figure 2.1: pUC Mix Marker, 8

2.8. SOLUTIONS AND BUFFERS

PBS

Stock solution (10XPBS) was prepared by dissolving 80g NaCl, 2g KCl, 11.5g $Na_2HPO_4.7H_2O$, and 2g KH_2PO_4 in 1lt ddH_2O. Working solution (1XPBS) was prepared by dilution of 10XPBS to 1X with ddH_2O. pH of the working solution was adjusted to 7.4.

TAE

Stock solution (50XTAE) was prepared by addition of 121g Tris-base, 18,6g EDTA, and 28.55ml glacial acetic acid to 500ml ddH_2O. pH of the stock solution was adjusted to 8.5. Working solution (1XTAE) was prepared by dilution of 50XTAE to 1X with ddH_2O.

DEPC-Treated Water

1ml DEPC was added to 1lt ddH_2O and stirred under hood overnight. DEPC was inactivated by autoclaving.

6X Agarose Gel Loading Dye

A mixture of 0.009g bromophenol blue (BFB), 0.009g xylene cyanol (XC), 2.8ml ddH_2O, 1.2 ml 0.5M EDTA was prepared. The total volume was brought to 15ml by addition of glycerol.

2.9. STATISTICAL ANALYSIS

Multivariate statistical methods such as Cluster, Discriminant Function Analysis, and Multidimensional Scaling provide a reduction of variables and an objective classification of samples.

2.9.1. Cluster Analysis

Cluster Analysis makes grouping of the objects according to degree of association between the objects. If the degree of association is high between the two objects, they are put in the same group, and if the association is low, they do not belong the same group. Hierarchical cluster analysis was chosen for clustering. In the hierarchical cluster, each case has a separate cluster at the beginning, and then clusters are combined according to their degree of association. This hierarchical clustering process can be represented as a tree, or dendrogram. Average linkage method of hierarchical cluster was chosen for clustering. It is based on the average distance between the objects. The clustering method represents the relationships of genes as a tree structure by connecting genes using their similarity scores based on the Pearson correlation coefficient. Both gene and sample based cluster analysis was performed. Cluster analysis and imaging were carried out by Eisen Lab Cluster Version 2.11 and Treeview Version 1.60 softwares [http://rana.lbl.gov/EisenSoftware.htm].

2.9.2. Pearson Correlation

Pearson correlation reflects the degree of linear relationship between the two variables. It ranges from +1 to -1. +1 reflects the perfect positive linear relationship and -1 represents the negative linear relationship between the two variables. Pearson correlation was carried out in Minitab™. The cut off p value used in the Pearson correlation test is 0.05.

2.9.3. Multidimensional Scaling (MDS)

Multidimensional Scaling refers to a family of models where the structure in a set of data is represented in a map based on the relationships between a set of points in a space [http://www.stat.psu.edu/~chiaro/BioinfoII/mds_sph.pdf]. MDS can be used on a variety of data, using different models and allowing different assumptions about the level of measurement. It makes it easier for the researcher to analyze data by

providing more user friendly summary of the analysis. There are two important things to consider in MDS. First the axes are meaningless, and the second the orientation of picture is arbitrary. There are two parameters of an MDS analysis: clusters and dimensions. Clusters are formed by the objects that are grouped together and the dimensions order the objects in the map along a continues line. This line may be from left to right and from down to up.

MDS requires an input of matrix showing the proximity between the interested variables. A matrix of Pearson correlation between the interested variables was used as input in our analysis. It is a useful tool to analyze similarities and dissimilarities between the interested variables. MDS is carried out in Matlab© for both the proximities between the samples and between the target genes. MDS function library for Matlab© was downloaded from Assoc. Prof. Rich Strauss's web site http://www.biol.ttu.edu/Strauss/Matlab/matlab.htm.

2.9.4. Discriminant Function Analysis

Discriminant function analysis shows how independent variables contribute to the categorical dependent variable. All variables were assumed to have the same covariance matrix, thus, linear discriminant analysis was performed. Discriminant analysis was performed in Minitab ™ for several clinical properties of the samples used in the analysis.

2.9.5. Mann Whitney Test

The Mann Whitney test is used to identify if the two populations are different from each other. In the two-tailed test, null hypothesis is that median of the two populations are equal. If the p value is less than the chosen α level there is sufficient evidence to reject null hypothesis. The Mann Whitney test was performed in Minitab™ for ER status of primary breast tumor samples used in this study with a 95% confidence interval for each gene.

CHAPTER 3

RESULTS

We aimed to identify the expression profile of the genes that play major roles in the Hedgehog signaling pathway in epithelial cancers, primarily in breast cancer. It is important to detect the expressional levels of the Hh signaling elements to explain this pathway's role in tumor development.

The expression levels of the genes can be determined by determining the protein or mRNA levels of the genes. In this study the mRNA expression levels of the genes were determined with q-rt-RT-PCR, which is the most sensitive method to determine the expression level of the genes.

3.1. RNA INTEGRITY

The integrity of the purified total RNA from the tissue samples and the cell lines were checked by using ND-1000 Spectrophotometer. The $\lambda_{260}/\lambda_{280}$ ratio of the RNAs was between 1.8-2.0. This range shows the high quality of isolated total RNAs. The total RNA was used for cDNA synthesis.

3.2. FIRST STRAND cDNA SYNTHESIS

The purified total RNAs were used in first strand cDNA synthesis. The integrity and amplification efficiency of the cDNA samples were determined by amplifying the samples with housekeeping gene specific primers for *GADPH*. The total RNAs isolated both from the primary tumor and the cell lines were synthesized with the

same efficiency and gave the expected 150bp amplification product in RT-PCR with *GADPH* primers. This result showed that the cDNAs were intact and can be used for further analysis (Figure 3.1).

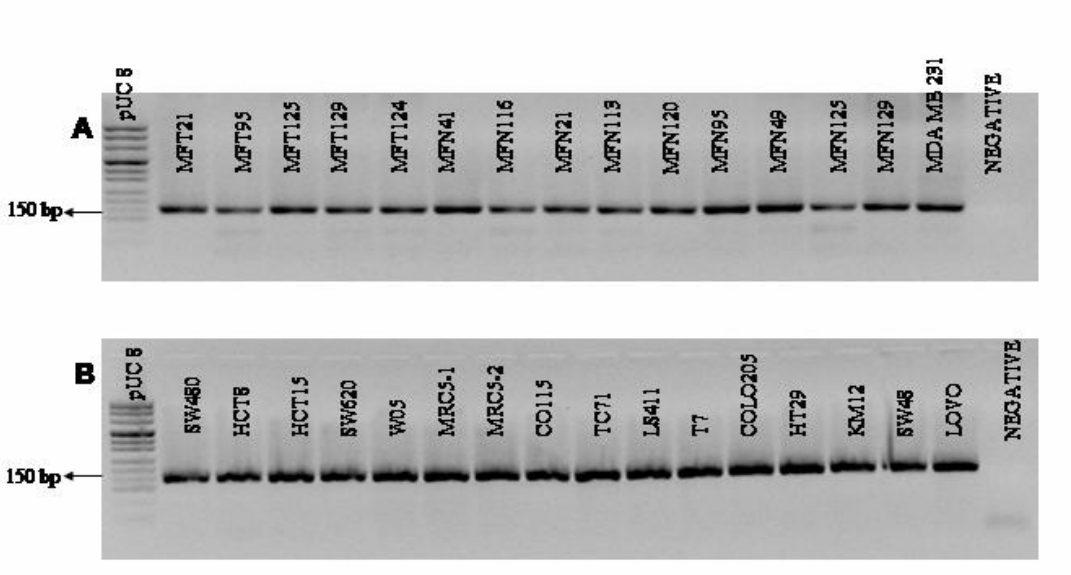


Figure 3.1: *GADPH* amplified cDNA samples. (A) Primary breast tumor tissue RNAs and (B) cell line RNAs were used for cDNA synthesis and amplified with *GADPH* gene specific primers. The DNA size marker is pUC8 and the product size is 150 bp.

3.3. MELT CURVE ANALYSIS

The real-time RT-PCR reaction uses SYBR Green I as a detection system. SYBR Green I is a DNA intercalating agent which binds to double stranded DNA during the product amplification at q-rt-RT-PCR.

The melt curve analysis is an important step to show the specificity of the q-rt-RT-PCR reaction. In a typical melt curve result, we expect to see a single peak for each sample at the same dissociation temperature. Any extra peak means that nonspecific amplicons, or primer dimers occurred in the reaction. Therefore, in each real-time RT-PCR reaction, melt curve analysis was carried out to ensure the specificity of the

reaction. Figure 3.2 A shows a typical melt curve result, and Figure 3.2 B shows a melt curve graph that contains non-specific amplicon and primer dimer in the reaction.

In this study, the PCR conditions for each gene were optimized to give the desired melt curve with real-time RT-PCR. If the desired melt curve was not achieved, new primer pairs were designed for that gene. *IHH* gene specific primers were redesigned because the first set of primers gave non-specific amplicon (Figure 3.2.B).

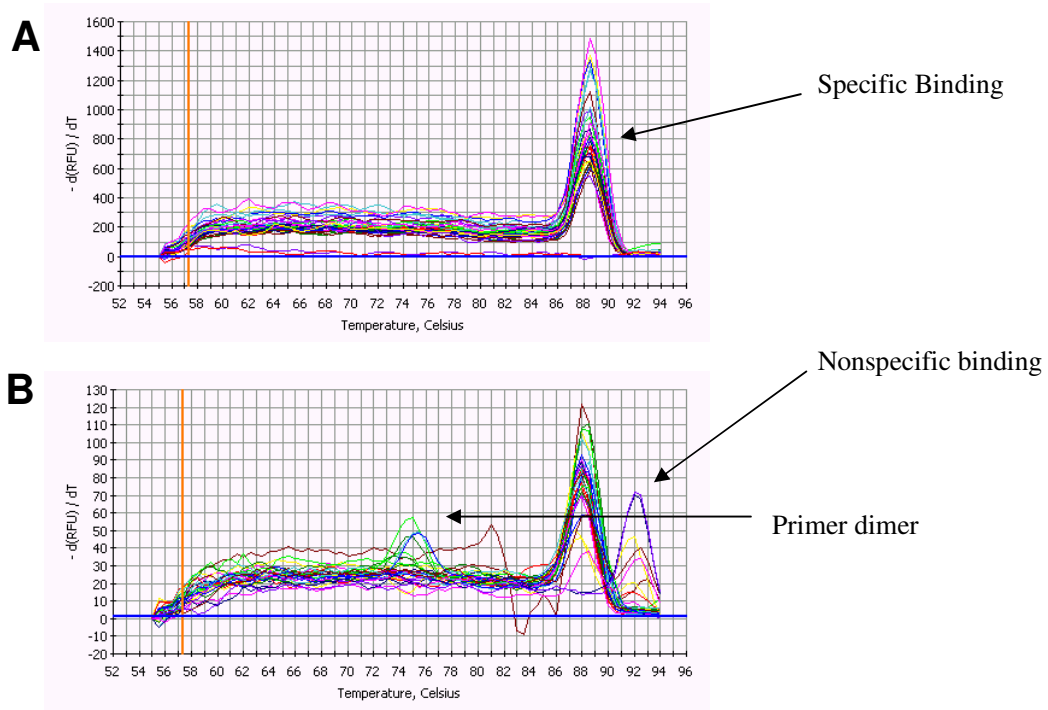


Figure 3.2: Melt curve graph: (A) Melt curve for q-rt-RT-PCR reaction of *GLI1* gene for breast carcinoma cell lines gave single peak for each sample at the same temperature. (B) Melt curve for the first set *IHH* primer pair gave nonspecific PCR product in q-rt-RT-PCR reaction for breast carcinoma cell lines.

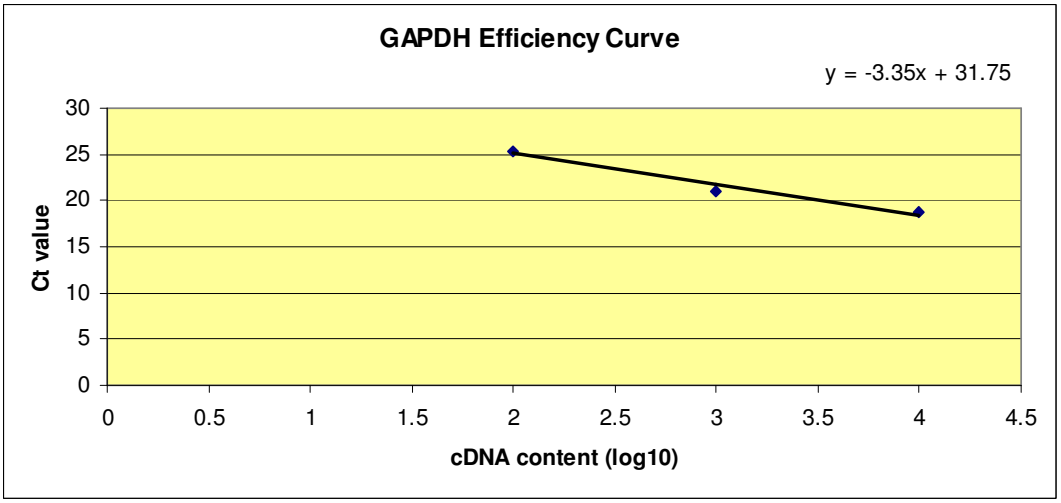
3.4. STANDARD CURVES AND AMPLIFICATION EFFICIENCIES

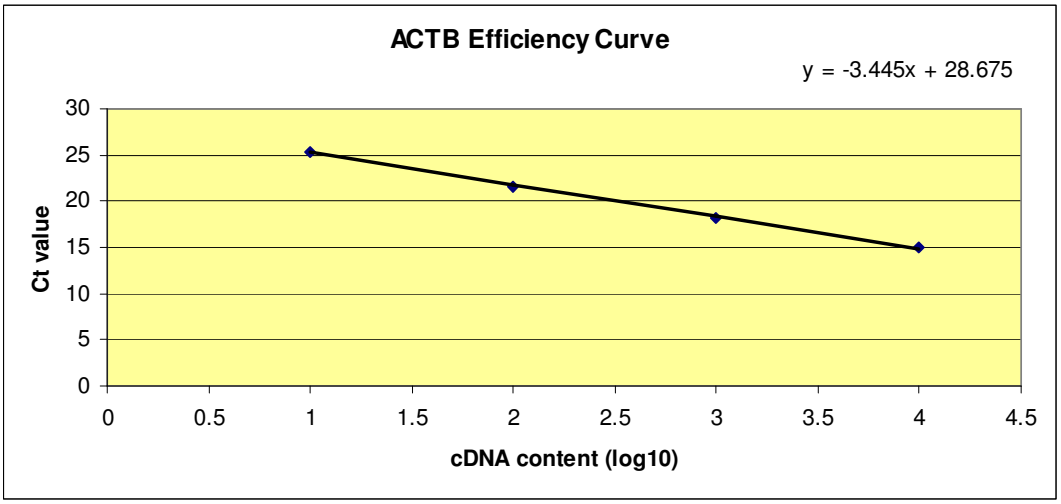
Amplification efficiency calculation is an important step in the q-rt-RT-PCR analysis. In relative quantification measurements, it is used in the calculation of the comparative expressional level of the samples. If the amplification efficiency for both reference gene and the target gene was 100%, equation 1.2 (section 1.3.2.1) was used for the calculations. In other cases, equation 1.3 (section 1.3.2.1) was used.

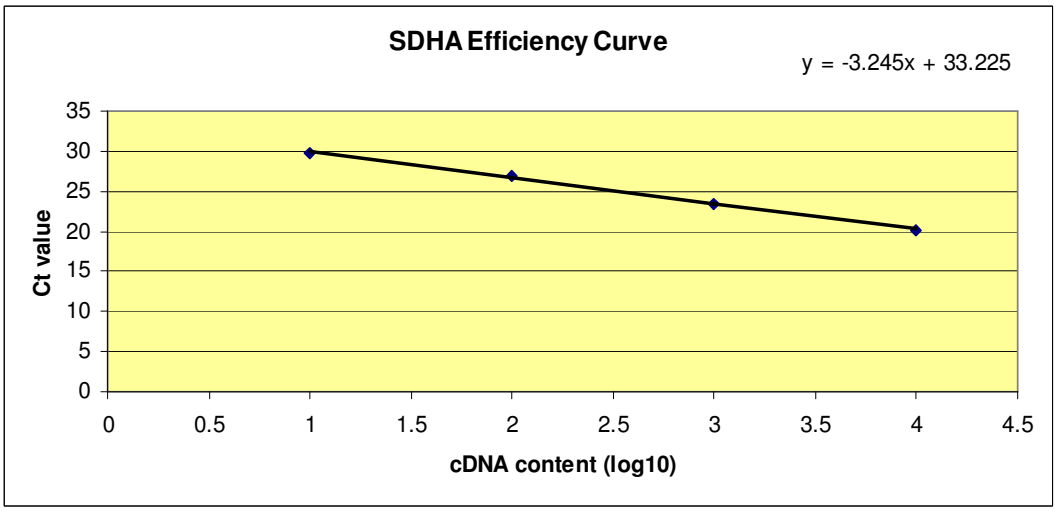
A cell line that shows the highest expression level for a particular gene was chosen to determine the amplification efficiency of the primers for that particular gene. Standard curves for each gene specific primer pairs were generated based on the data obtained by performing the q-rt-RT-PCR with a two-fold serial dilution of the cDNAs. Theoretically, a 100% efficient amplification of a PCR product equals one cycle of increase in each two-fold dilution series.

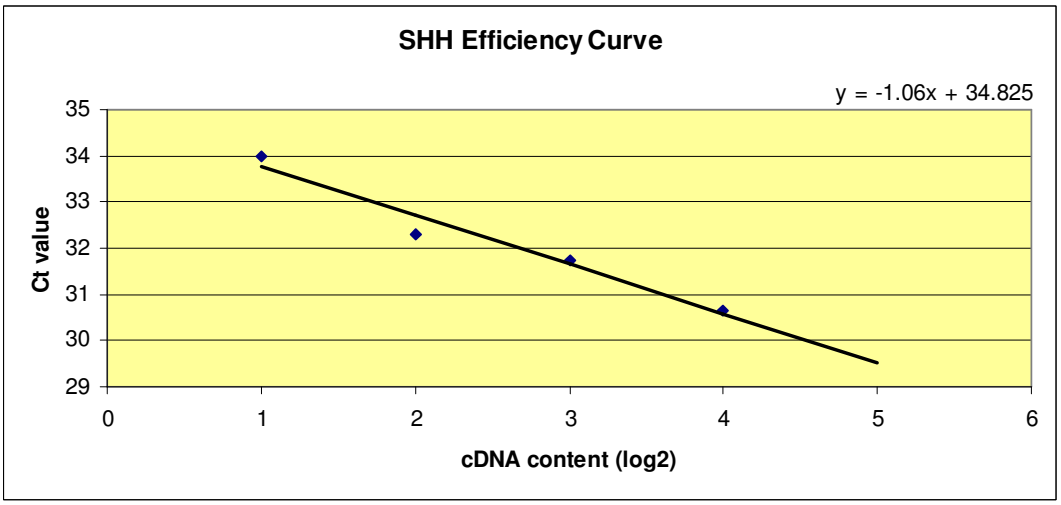
GAPDH was used as a housekeeping gene for the carcinoma cell lines and *ACTB, TBP*, and *SDHA* were used for the primary breast tumor and normal samples. The amplification efficiencies for housekeeping genes were calculated previously by Bala Gur on a series of ten-fold dilution of cDNA (unpublished data). 100% efficient amplification equals to 3.33 cycle of increase in each ten-fold dilution series. Figure 3.3 shows the results of the amplification efficiency standard curves of the genes used in the study.

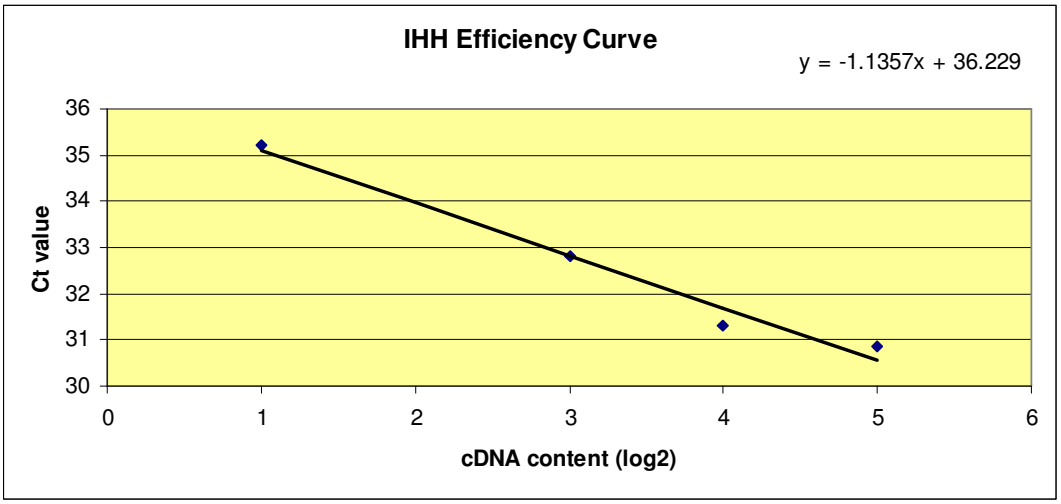
The slopes obtained from the dilution curves were used to calculate the amplification efficiency of the primers with the equation 1.4 (the equation was given in (section 1.3.2.1). The amplification efficiency of the all the housekeeping gene primers were 100 percent. Table 3.1 shows the summary of the amplification efficiencies of the primers used in this study.

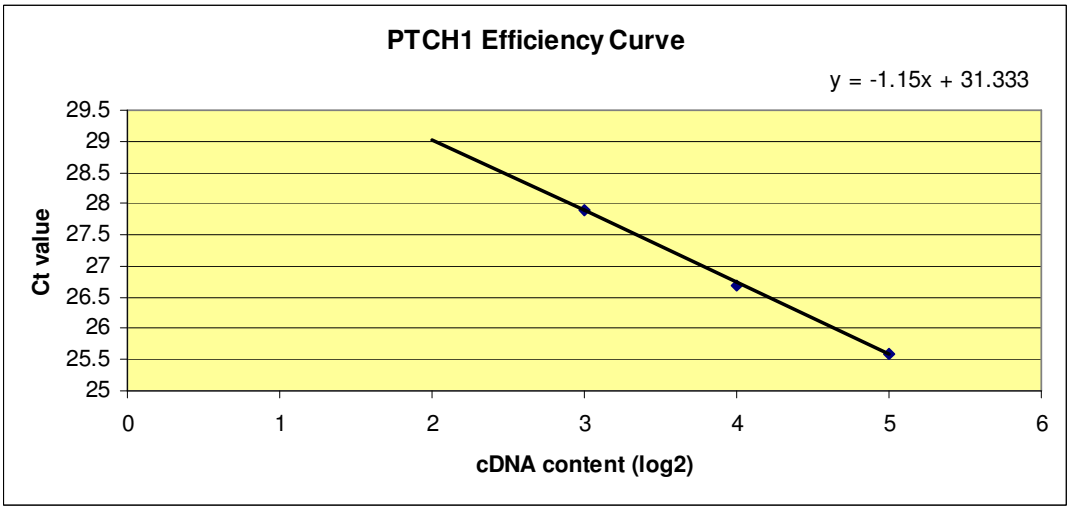

GAPDH Efficiency Curve
y = -3.35x + 31.75
Ct value
cDNA content (log10)

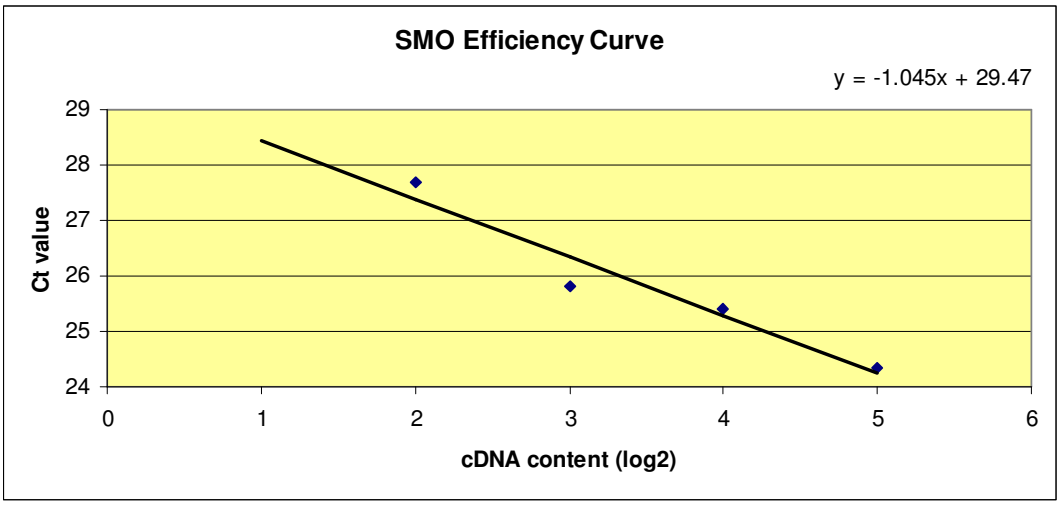

ACTB Efficiency Curve
y = -3.445x + 28.675
Ct value
cDNA content (log10)

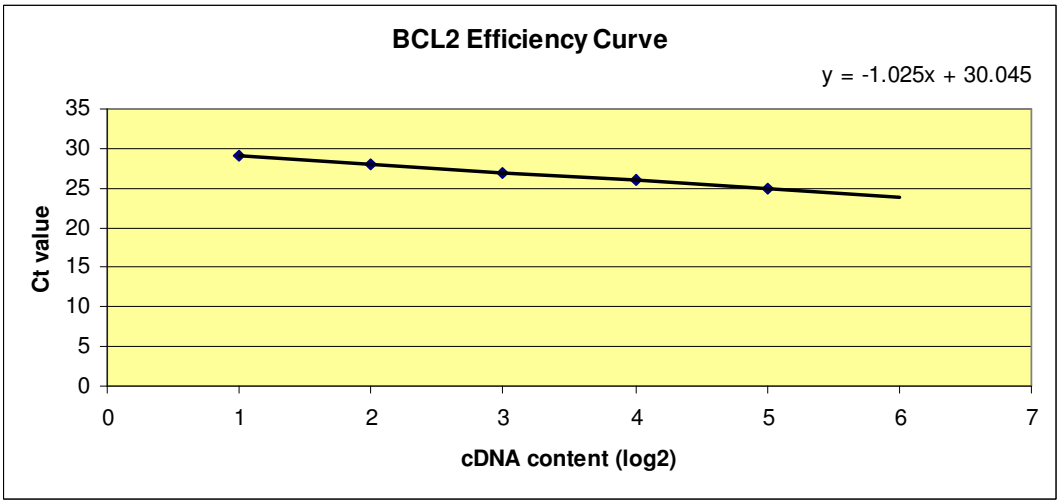

SDHA Efficiency Curve
y = -3.245x + 33.225
Ct value
cDNA content (log10)

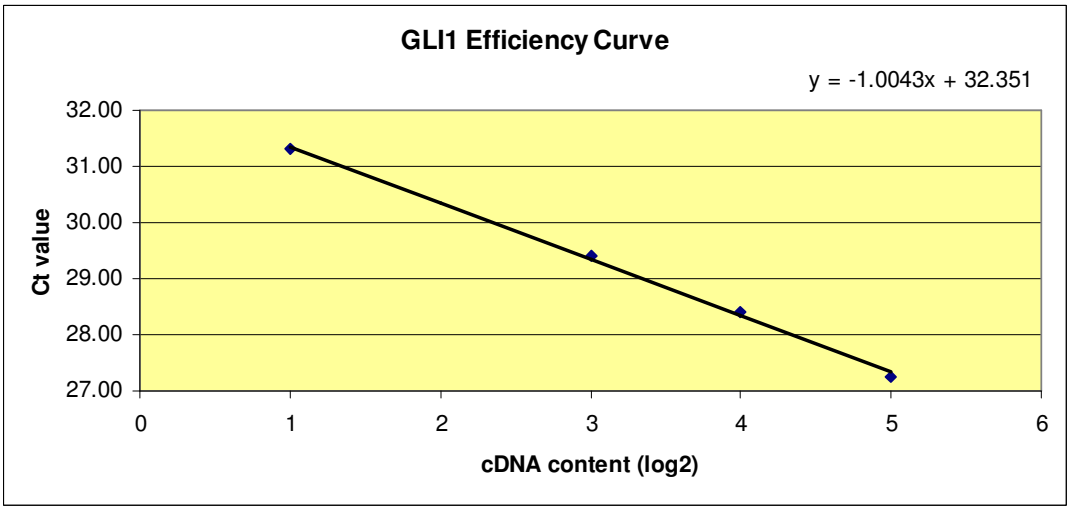

SHH Efficiency Curve
y = -1.06x + 34.825
Ct value
35
34
33
32
31
30
29
0
1
2
3
4
5
6
cDNA content (log2)

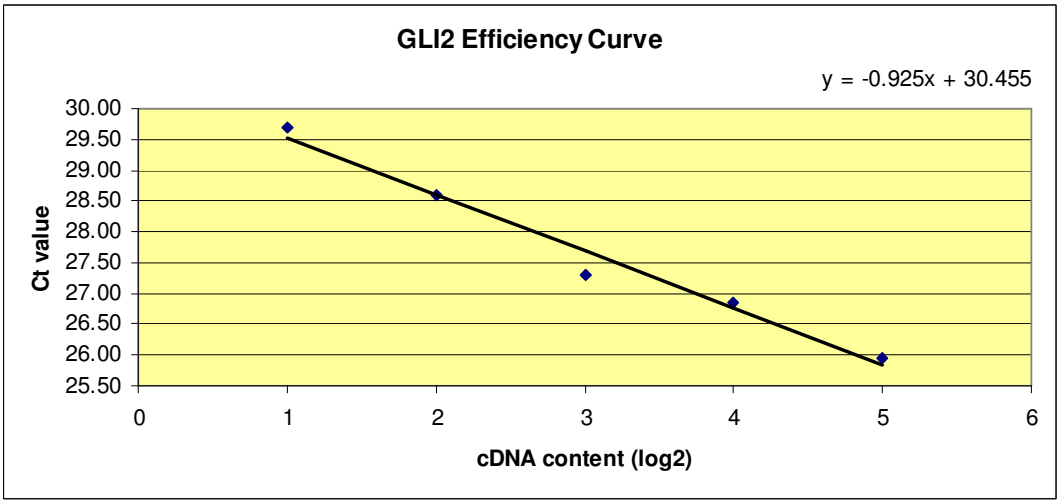

IHH Efficiency Curve
y = -1.1357x + 36.229
Ct value
36
35
34
33
32
31
30
0
1
2
3
4
5
6
cDNA content (log2)


PTCH1 Efficiency Curve
y = -1.15x + 31.333
Ct value
29.5
29
28.5
28
27.5
27
26.5
26
25.5
25
0
1
2
3
4
5
6
cDNA content (log2)


SMO Efficiency Curve
y = -1.045x + 29.47
Ct value
29
28
27
26
25
24
0
1
2
3
4
5
6
cDNA content (log2)


BCL2 Efficiency Curve
y = -1.025x + 30.045
Ct value
35
30
25
20
15
10
5
0
0
1
2
3
4
5
6
7
cDNA content (log2)


GLI1 Efficiency Curve
y = -1.0043x + 32.351
Ct value
32.00
31.00
30.00
29.00
28.00
27.00
0
1
2
3
4
5
6
cDNA content (log2)



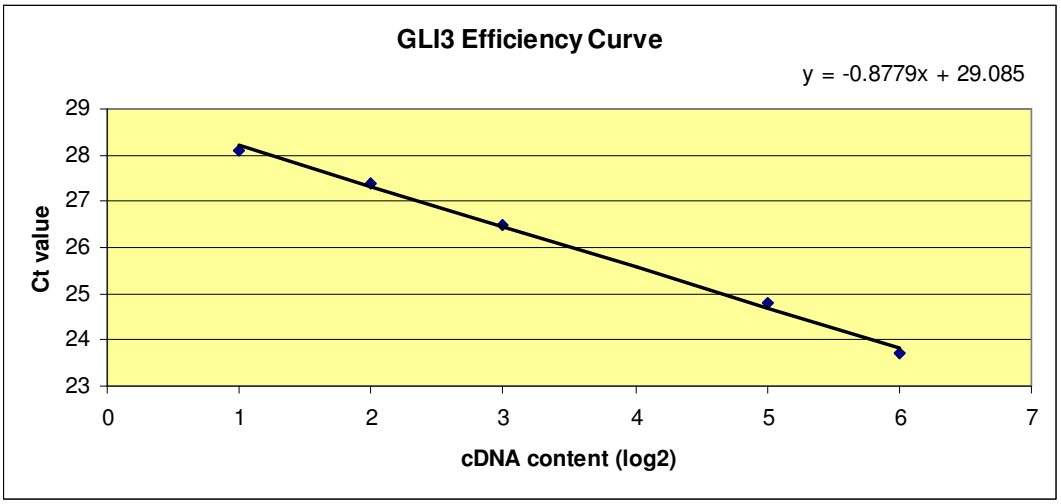


Figure 3.3: The amplification efficiency graphs of the gene specific primers. The amplification efficiency was calculated from the 10-fold dilution series of cDNAs for housekeeping genes and 2-fold serial dilution series of cDNAs for the Hh pathway specific genes.

Table 3.1: Amplification efficiencies of the gene-specific primers

Gene	Slope	Efficiency
GAPDH	-3.35	99
ACTB	-3.45	95
SDHA	-3.25	103
TBP	-3.44	95
SHH	1.06	92
IHH	1.14	84
PTCH1	1.15	83
SMO	1.05	94
GLI1	1	100
GLI2	0.93	112
GLI3	0.88	120
BCL2	1.025	97

3.5. QUANTITATIVE REAL-TIME RT-PCR

The Hh signaling pathway expression profile of the carcinoma cell lines of epithelial origin was determined by q-rt-RT-PCR analysis. In this study, human breast, hepatocellular, and colon carcinoma cell lines, and human primary breast tumor samples were used for the expression profiling of the target genes (Table 2.1, 2.2, 2.3, and 2.4 respectively, in section 2.1).

All the q-rt-RT-PCR reactions were run on a BioRad iCycler. The software displays an amplification graph, Ct values and the melt curve at the end of each amplification reaction. Figure 3.5 shows a typical q-rt-RT-PCR amplification and melt curve graphs for a target gene. All the reactions were target gene specific and displayed the expected single melt curve peak in the graphs.

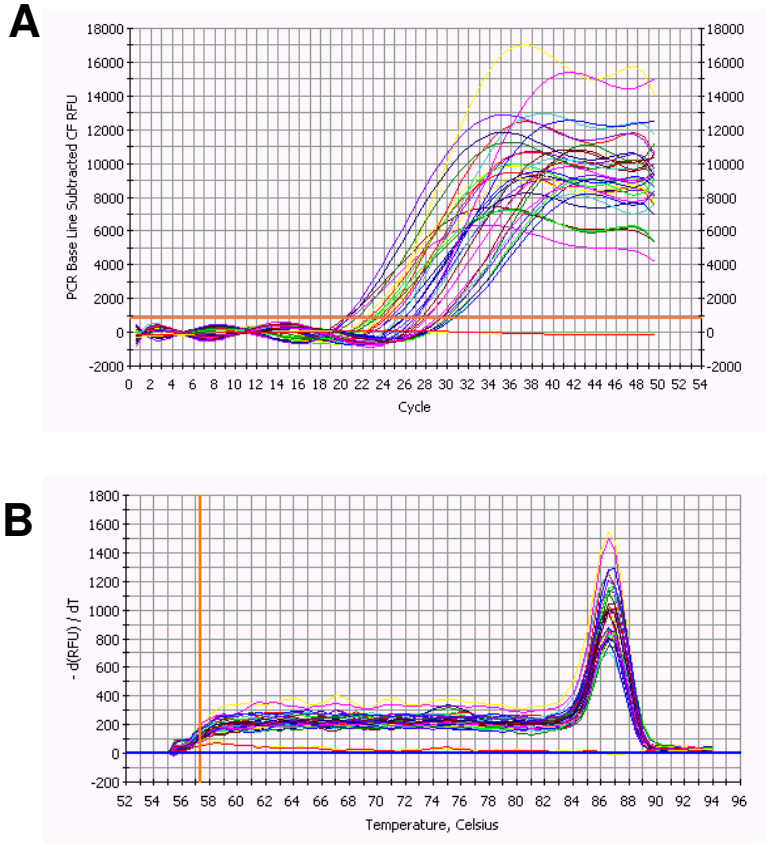


Figure 3.4: q-rt-RT-PCR amplification graph and the melt curve. (A) The amplification graph of *PTCH1* gene in breast carcinoma cell lines, and (B) melt curve graph of the same reaction.

Q-rt-RT-PCR reaction for each gene was performed in duplicates, including positive and negative control samples. The duplicate PCR products were run on 2% agarose gel to observe the specific product size and confirm the melt curve analysis obtained with q-rt-RT-PCR reactions.

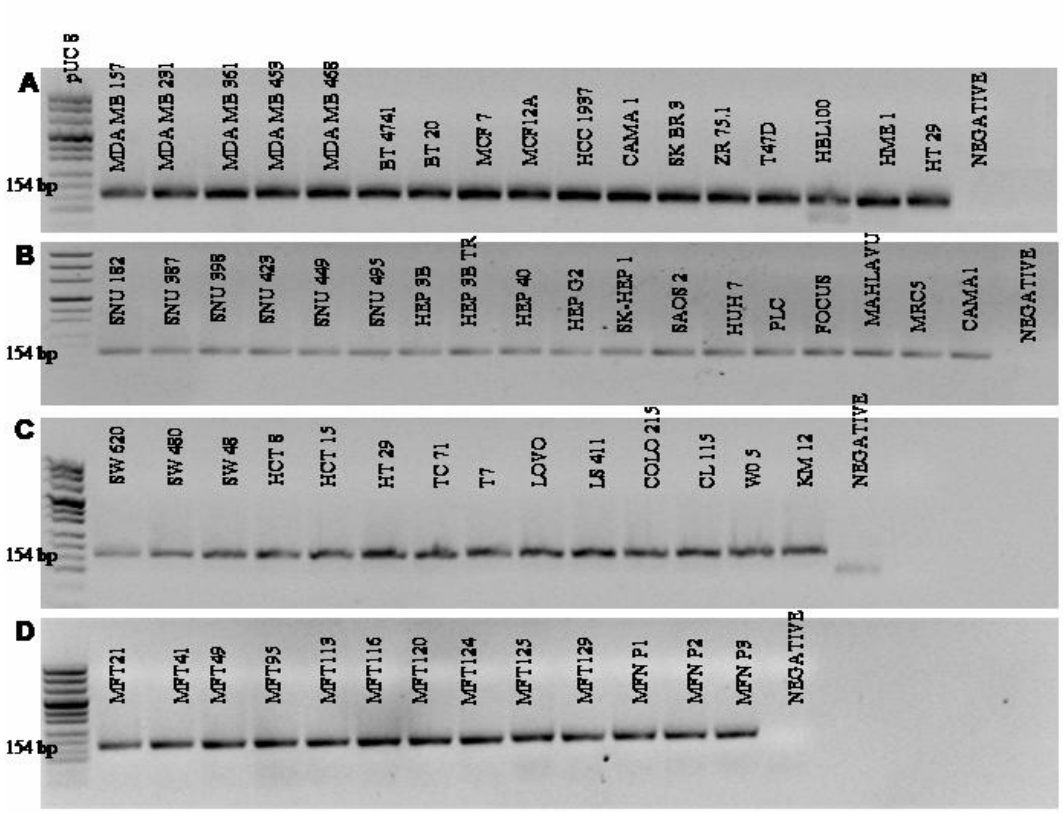


Figure 3.5: Selected q-rt-RT-PCR agarose gel electrophoresis photographs for *GLI1*. (A) PCR products of *GLI1* gene in breast carcinoma cell lines, (B) hepatocellular carcinoma cell lines, (C) colon carcinoma cell lines, and (D) primary breast tumor samples. For each q-rt-RT-PCR reaction, negative control was used. The DNA size marker is pUC8 and the *GLI1* product size is 154 bp.

3.6. BREAST CARCINOMA CELL LINES

Quantification of mRNA expression values of each gene is calculated by using the equations 1.2 and 1.3 that are given in section 1.3.2.1. The expression values that were obtained with these calculations were then transformed to log2 scale. Log2 transformation makes upregulation and downregulation values more symmetrical.

Table 3.2: Fold expression values of the Hh pathway genes in breast carcinoma cell lines. The value in bold may represent an outlier.

GENE	*GLI1*	*GLI2*	*GLI3*	*SHH*	*IHH*	*PTCH1*	*SMO*	*BCL2*
MDA MB157	-2.20	6.52	4.80	0.20	2.48	**7.70**	2.77	3.25
MDA MB231	-1.75	4.51	0.05	0.92	-1.07	2.50	0.59	3.50
MDA MB361	-2.55	-2.92	0.39	1.48	-2.46	3.97	0.23	2.30
MDA MB453	-0.10	**-20.59**	4.54	-0.56	-1.56	3.98	0.66	1.70
MDA MB468	-0.80	0.12	1.80	0.03	-0.72	5.17	2.49	3.60
BT474	1.45	-5.43	0.08	1.41	4.80	1.43	4.31	5.70
BT20	-3.15	-2.25	0.00	-0.07	0.31	4.32	1.66	0.80
MCF7	-0.50	-5.57	2.89	-0.21	0.93	4.88	1.59	8.00
MCF12A	-2.15	3.19	2.41	-0.57	-3.96	2.99	2.12	0.30
HCC1937	-3.70	6.98	1.26	-1.51	1.38	4.44	0.59	1.40
CAMA1	1.70	-0.19	4.31	0.69	4.41	1.45	4.32	5.85
SKBR3	-2.90	-3.66	-2.69	-0.18	-3.27	1.58	-0.55	-2.55
ZR75.1	-3.35	-4.58	-0.85	-0.52	-1.67	1.96	-0.17	4.60
T47D	-0.80	-17.39	3.54	0.18	-0.29	1.92	4.02	-0.50
HBL100	-0.30	3.19	-0.35	2.71	-2.17	2.38	-0.01	1.90

3.6.1. Expression Status of Hh Pathway Target Genes

The expression status of the Hh pathway target genes in breast carcinoma cell lines were determined by q-rt-RT-PCR. Fifteen human breast carcinoma cell lines and an hTERT immortalized normal human mammary epithelial cell line, hTERT-HME1, were used to obtain the Ct values of each target gene. Each gene's Ct value was normalized to that of the housekeeping gene *GADPH* for a given sample. Then, the normalized values compared to the normalized Ct value of the hTERT-HME1 cell line by using the equations 1.2 and 1.3 (section 1.3.2.1.). Figure 3.6 shows the expression levels of the breast carcinoma cell lines for Hh pathway target genes. Observed log fold differences ranged between 7.7 and -20.6; accordingly, a higher variability in gene expression has been observed for *GLI2* whereas *SHH* and *SMO* exhibited the least amount of variation in terms of mRNA expression (Table 3.2; Figure 3. 6).

The raw Ct values are given in Appendix XXX.

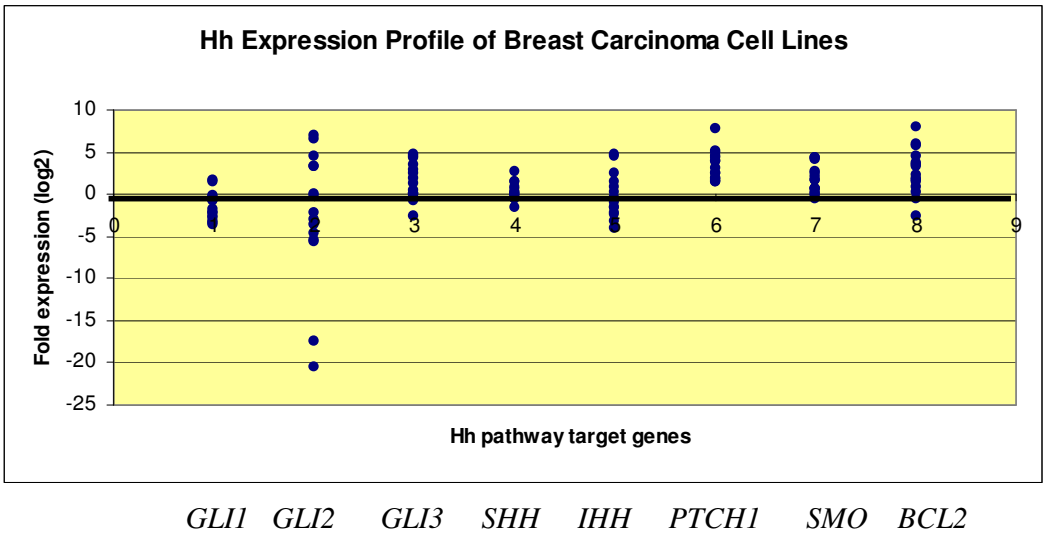


Figure 3.6: Expression profile of the Hedgehog pathway target genes in human breast carcinoma cell lines. Fold expressions were given in log2 scale.

3.6.2. Hierarchical Cluster Analysis

Cluster analysis makes grouping of the objects according to association between the objects. Hierarchical Cluster Analysis was carried out to observe possible grouping of the cell lines based on their gene expression profiles.

Expression values of all the target genes in 15 human breast carcinoma cell lines were clustered by using the Cluster 2.11 program created by Eisen Lab [http://rana.lbl.gov/]. Average cluster analysis was performed both in a gene and array based manner and represented as a tree and dendrogram in Figure 3.7.

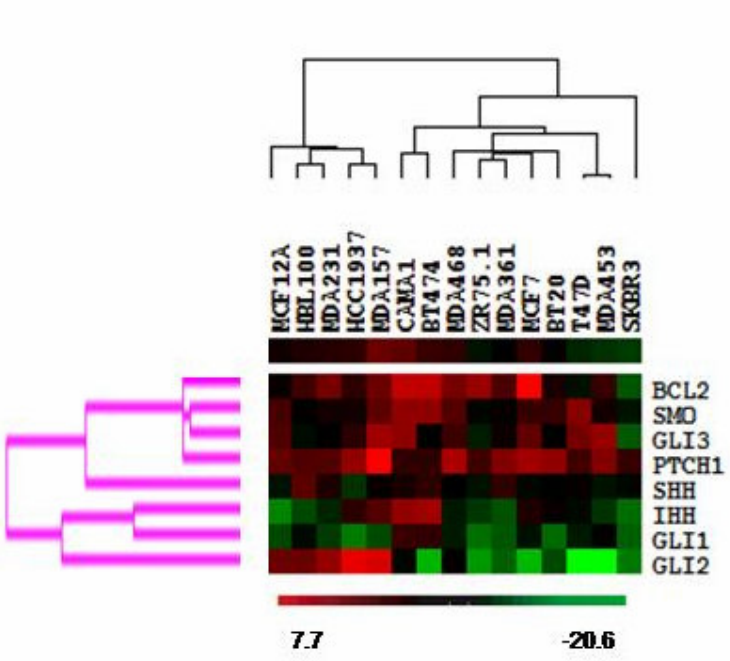


Figure 3.7: Two-dimensional Cluster Analysis of genes in human breast carcinoma cell lines. A dendogram obtained from hierarchical clustering, using expression data from 8 genes in 15 breast carcinoma cell lines. The red and the green colors indicate the high and the low expressions values respectively.

3.6.3. Pearson Correlation Analysis

Pearson correlation analysis was used to determine if there is a linear relationship between the two genes' expression pattern in the cell lines studied. This analysis was applied to the data obtained from 15 human breast carcinoma cell lines by using statistical tools in Minitab™ Program. A correlation of +1 represents a perfect positive relationship with a p value less than 0.05, which means that the performed analysis is significant. This data shows that *IHH, SMO* and *GLI1* gene expression profiles were found to be highly correlated with each other in all the breast carcinoma cell lines studied. Significant correlations included those between *GLI1* and *IHH* (r=0.521, p=0.047); *SMO* and *GLI1* (r=0.637, p=0.011); and *SMO* and *IHH* (r=0.708, p=0.003). No multiple test correction has been made.

Table 3.3: Pearson Correlation for human breast carcinoma cell lines. The first value represents the Pearson correlation and the second value represents the p value for each pair of gene.

	GLI1	GLI2	GLI3	SHH	IHH	PTCH1	SMO
GLI2	−0.316						
	0.251						
GLI3	0.417	−0.230					
	0.122	0.410					
SHH	0.472	0.109	−0.216				
	0.076	0.699	0.440				
IHH	**0.521**	0.088	0.382	0.073			
	0.047	0.756	0.160	0.795			
PTCH1	−0.301	0.268	0.444	−0.275	0.088		
	0.276	0.334	0.098	0.322	0.756		
SMO	**0.637**	−0.131	**0.579**	0.082	**0.708**	−0.028	
	0.011	0.643	**0.024**	0.771	**0.003**	0.921	
BCL2	0.496	0.108	0.281	0.177	**0.599**	0.126	0.323
	0.060	0.703	0.310	0.528	**0.018**	0.655	0.240

3.6.4. Multidimensional Scaling (MDS)

MDS, a data analysis tool to make the data much easier to comprehend, provides a visual representation of the pattern of proximity of a set of objects. The proximity is a number which indicates how similar or how different these objects from each other. MDS plots the objects in a map in which objects with high similarity are placed near each other. Interpoint distances correspond to experimental dissimilarities between the points. In this study the proximity calculation was performed by Pearson correlation as given in section 3.6.3.

MDS was performed to 15 human breast carcinoma cell lines with an input matrix of Pearson correlation of genes, or cell lines in the Matlab© program. This analysis was performed to examine the clustering among genes upon reduction of the data dimensions. All the Hh target gene expression values and the cell lines used in this

study were taken into consideration for MDS analysis. The results of MDS analysis of the Hh target genes showed that there was not a strong grouping among the genes (Figure 3.8); yet when the same analysis was performed for the cell lines for their target genes' expression profiles, the pattern of dissimilarity has led to the identification of mainly two major groupings. The most striking expression similarities between cell lines included those between MDA MB453 and T47D; MDA MB361, MDA MB468, MCF7, ZR75.1 and BT20 (Figure 3.9). In addition in the right part of the Figure 3.9, MDA MB157, MDA MB231, MCF12A, HCC1937, and HBL100 were clustered together which were also clustered in the Hierarchical Cluster analysis (Figure 3.7).

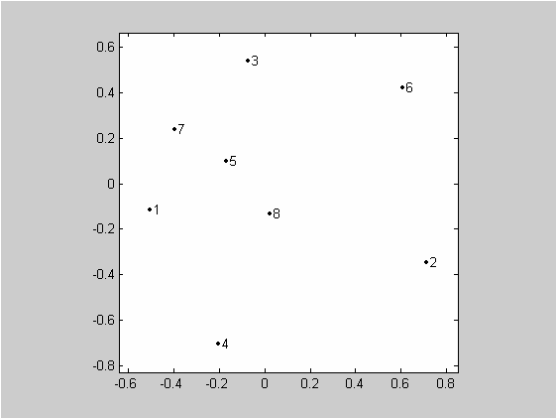


Figure 3.8: Gene-based MDS Analysis for breast carcinoma cell lines. Correlation values between the expression levels of genes were used as the input data. Each point shown in the map by a number represents one gene. The numbering of genes used in the analysis from number 1 to 8 is *GLI1*, *GLI2*, *GLI3*, *SHH*, *IHH*, *PTCH1*, *SMO*, *BCL2* respectively.

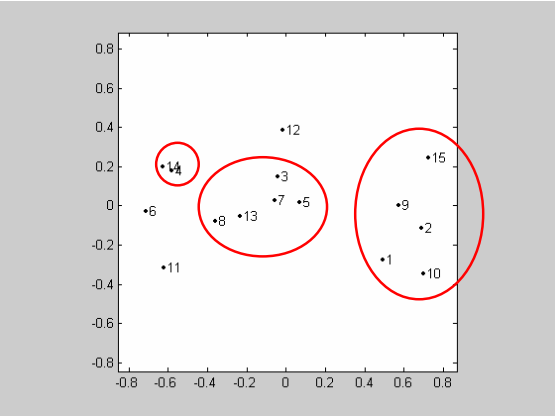


Figure 3.9 Cell line-based MDS Analysis for breast carcinoma cell lines. Correlation values between the expression levels of cell lines were used as the input data. Each point shown in the map by a number represents one cell line. The numbering of cell lines used in the analysis from number 1 to 15 is in the same order as given in Table 3.2.

3.6.5. Discriminant Function Analysis (DFA)

DFA differentiates the separate populations by combining a number of variables and maximizing the ratio of the differences. A prior knowledge of grouping of the samples is required. The grouping was carried out for the estrogen receptor (ER) status of the cell lines. The ER status of the cell lines were given in Table 2.1.

Linear discriminant function analysis (DFA) was performed for 15 human breast carcinoma cell lines. There are 7 ER (-) and 8 ER (+) cell lines available for the study. When all the target gene expression values were taken into consideration for the DFA, the analysis predicted the ER status correctly in 93% of the cell lines (Table 3.4). Cross-validation has been performed to assess the variability in misclassification rate. Cross-validation of the DFA results indicated that the percentage of correct classification might vary depending on the selection of cell lines since the misclassification rate increased to 40% with cross-validation.

Table 3.4: DFA of human breast carcinoma cell lines based on their ER status.

A Classification

Put into Group	..True Group.. ER−	ER+
Total N	8	7
N Correct	7	7
Proportion	0.875	1.000

Proportion Correct= 0.933

B Classification with Cross-validation

Put into Group	..True Group.. ER−	ER+
Total N	8	7
N Correct	5	4
Proportion	0.625	0.571

Proportion Correct = 0.600

C Linear Discriminant Function for Group

	ER−	ER+
Constant	−9.269	−20.673
GLI1	−4.567	−8.801
GLI2	−0.384	−0.628
GLI3	−0.081	1.034
SHH	2.971	4.693
IHH	−1.573	−3.020
PTCH1	0.099	−2.968
SMO	4.444	8.143
BCL2	2.027	4.431

3.7. HEPATOCELLULAR CARCINOMA CELL LINES

Quantification of mRNA expression values of each gene is calculated by using the equations 1.2 and 1.3 that are given in section 1.3.2.1. The expression values that were obtained with these calculations were then transformed to log2 scale.

Table 3.5: Fold expression values of the Hh pathway genes in HCC cell lines. The value in bold may represent an outlier.

GENE	*GLI1*	*GLI2*	*GLI3*	*SHH*	*IHH*	*PTCH1*	*SMO*	*BCL2*
SNU182	0.86	7.89	8.52	6.72	-8.84	3.16	5.34	4.61
SNU387	2.46	7.55	6.52	4.00	-6.29	1.79	-10.76	-9.59
SNU398	1.86	0.74	5.75	-0.72	-11.44	1.56	-15.91	-14.74
SNU423	0.16	2.15	5.97	7.11	-0.87	-0.33	4.84	4.11
SNU449	0.86	0.88	3.49	0.47	-1.04	-1.64	3.49	3.46
SNU475	5.06	3.01	-14.84	8.93	6.72	1.26	2.04	**10.76**
HEP3B	-2.04	-6.22	-10.58	5.21	0.39	-1.05	3.44	-2.44
HEP3B TR	-0.94	-4.84	-12.88	5.41	2.93	-1.63	4.94	-1.44
HEP40	-2.99	-6.98	1.12	-10.78	5.82	-0.73	3.39	1.36
HEPG2	-2.79	-6.35	-5.16	2.07	2.33	-1.56	2.74	0.61
SKHEP1	-3.79	0.42	2.38	-11.63	-2.64	-2.28	2.44	4.36
HUH7	-1.24	-5.77	-8.49	8.01	6.85	0.44	5.74	-2.59
PLC/PRL/5	1.36	-6.62	4.22	-10.58	1.77	-2.01	1.99	-6.29
FOCUS	0.76	1.54	3.30	-3.26	-3.35	-0.39	-13.86	7.41
MAHLAVU	-1.49	0.91	4.12	-10.05	-4.84	-1.16	**-18.36**	4.31

3.7.1. Expression status of Hh pathway target genes in Hepatocellular Carcinoma Cell Lines (HCC)

The expression status of the Hh pathway target genes in hepatocellular carcinoma cell lines were determined by q-rt-RT-PCR. Fifteen human HCC cell lines were used to obtain the Ct values of each target gene. Each genes Ct value was normalized to the Ct value of the housekeeping gene *GADPH* for a given sample. The Ct values from all the samples were used to calculate the geometrical mean of each target and *GADPH* genes. The geometric mean of the each target gene was normalized to geometric mean of *GADPH* gene. The first normalized values compared to the normalized geometric mean Ct value of the each gene for a given cell line by using the equations 1.2 and 1.3 (section 1.3.2.1.). Observed log fold differences ranged between 10.76 and -18.36; accordingly, almost all of Hh genes exhibited substantial variability in terms of mRNA expression while *GLI1* and *PTCH1* showed a relatively more stable expression level among HCC cell lines (Table 3.5; Figure 3. 10).

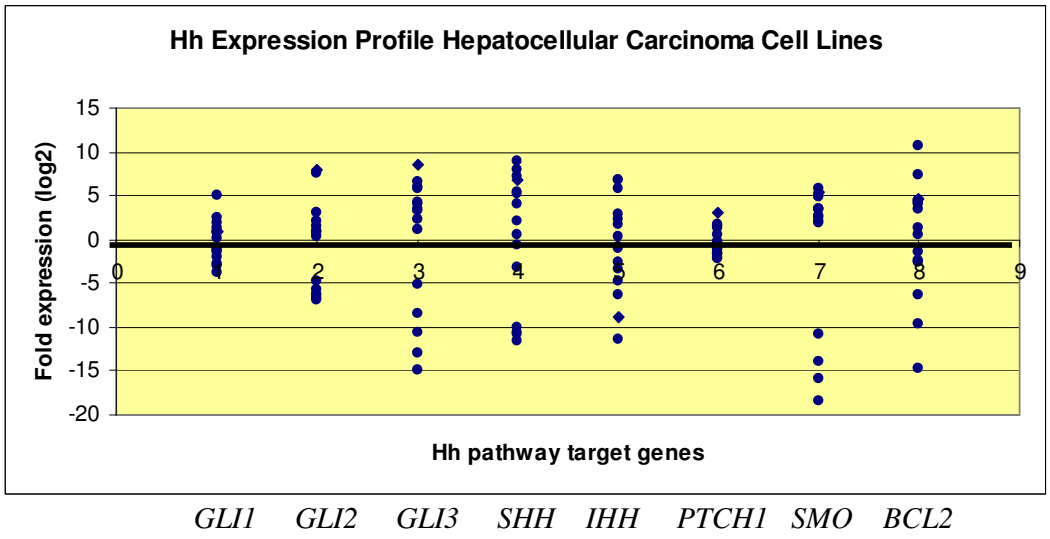


Figure 3.10: Expression profile of the Hedgehog pathway target genes in human HCC cell lines. The fold expressions were given in log2 scale.

3.7.2. Hierarchical Cluster Analysis

Expression values of all the target genes in 15 human HCC cell lines were clustered by using the Cluster 2.11 program created by Eisen Lab. Average cluster analysis was performed both in a gene and array based manner and represented as a tree and dendrogram in Figure 3.11.

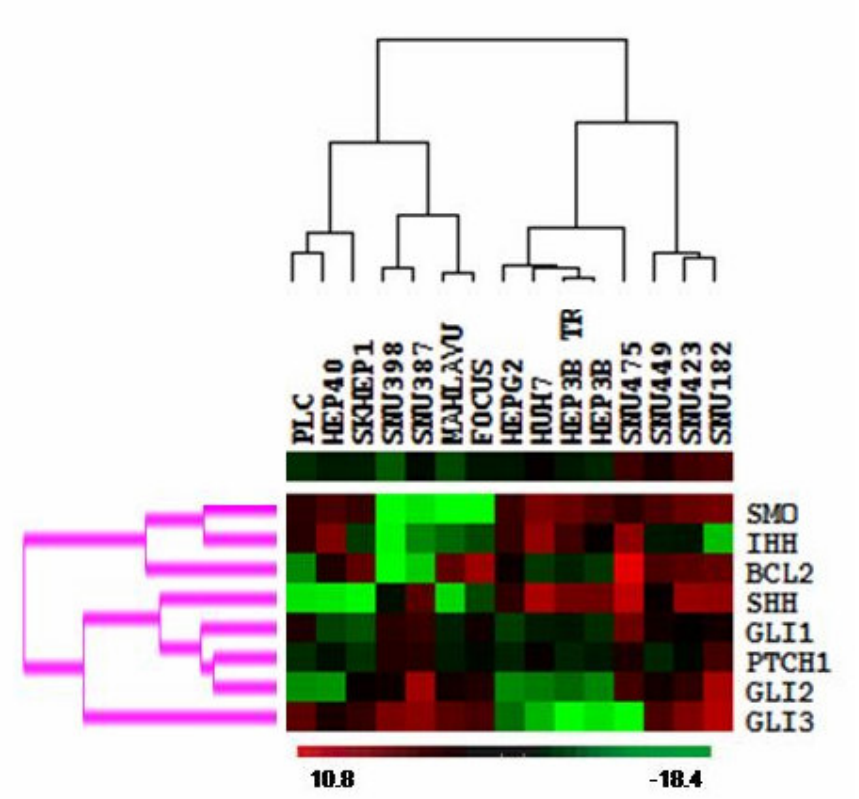


Figure 3.11: Two-dimensional Cluster Analysis of genes in human HCC cell lines. A dendogram obtained from hierarchical clustering, using expression data from 8 genes in 15 HCC cell lines. The red and the green colors indicate the high and low expression values respectively.

3.7.3. Pearson Correlation Analysis

Pearson correlation analysis was applied to the data that was obtained from 15 human HCC cell lines by using statistical tools in Minitab™ Program. Significant negative correlations were observed between *IHH* and *GLI2* (r=-0.647, p=0.009) and *GLI3* (r=-0.721, p=0.002) as shown in Table 3.6. High positive correlations were obtained for *PTCH1* and *GLI1* (r=0.586, p=0.022), and *GLI2* (r=0.661, p=0.007). No multiple test correction has been made.

Table 3.6: Pearson Correlation for human HCC cell lines. The first value represents the Pearson correlation and the second value represents the p value.

	GLI1	GLI2	GLI3	SHH	IHH	PTCH1	SMO
GLI2	**0.560**						
	0.030						
GLI3	0.041	0.494					
	0.886	0.061					
SHH	0.417	0.253	-0.427				
	0.122	0.364	0.112				
IHH	-0.174	**-0.647**	**-0.721**	0.111			
	0.535	**0.009**	**0.002**	0.694			
PTCH1	**0.586**	**0.661**	0.199	**0.533**	-0.412		
	0.022	**0.007**	0.477	**0.041**	0.127		
SMO	-0.230	-0.353	-0.405	0.330	**0.611**	-0.187	
	0.409	0.196	0.134	0.230	**0.015**	0.505	
BCL2	-0.048	0.176	-0.191	0.034	0.340	-0.141	0.265
	0.866	0.531	0.495	0.905	0.216	0.615	0.340

3.7.4. Multidimensional Scaling

MDS was performed to 15 human HCC cell lines with an input matrix of Pearson correlation of genes (Table 3.6), and cell lines in the Matlab© program. This analysis was performed to examine the clustering among genes upon reduction of the data dimensions.

All of the Hh pathway gene expression values and the 15 cell lines used in this study were taken into consideration for MDS analysis. When MDS was performed with respect to the Hh genes, the resulting graph showed that there was not a strong grouping among genes (Figure 3.12). MDS analysis of the 15 HCC cell lines, on the other hand, revealed that there were several subgroups of cell lines (Figure 3.13). The most striking expression similarities between cell lines included those between

FOCUS and MAHLAVU; SNU387 and SNU398; and HEP3B, HEP3B TR, SKHEP1 and PLC/PRL/5.

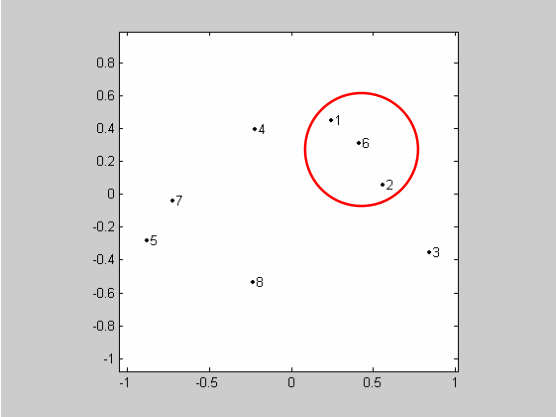


Figure 3.12: Gene-based MDS Analysis for HCC cell lines. Correlation values between the expression levels of genes were used as the input data. Each point shown in the map by a number represents one gene. The numbering of genes used in the analysis from number 1 to 8 is *GLI1*, *GLI2*, *GLI3*, *SHH*, *IHH*, *PTCH1*, *SMO*, *BCL2* respectively.

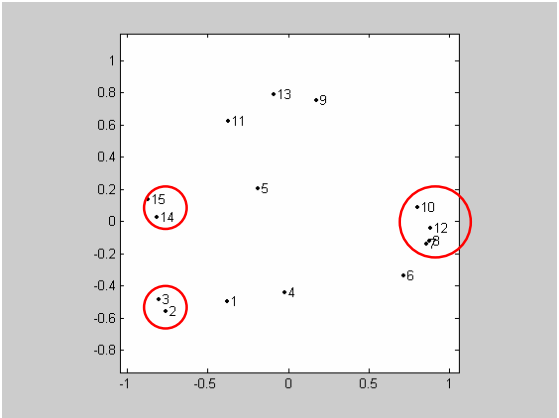


Figure 3.13: Cell line-based MDS Analysis for HCC cell lines. Correlation values between the expression levels of cell lines were used as the input data. Each point shown in the map by a number represents one cell line. The numbering of cell lines used in the analysis from number 1 to 15 is in the same order as given in Table 3.5.

3.7.5. Discriminant Function Analysis

Linear discriminant function analysis was performed for 15 human HCC cell lines. The grouping was performed for the differentiation status of the cell lines. There were 6 well differentiated and 9 poorly differentiated HCC cell lines (Table 2.2). When all the target gene expression values were taken into consideration for the DFA, it predicted the differentiation status of the cell lines perfectly (Table 3.7). Cross-validation of the DFA results indicated the percentage of correct classification might vary depending on the selection of cell lines since the misclassification rate increased to 27% with cross-validation.

Table 3.7: DFA of human HCC cell lines based on their differentiation status

A **Summary of Classification**

Put into Group	..True Group.. Poor	Well
Total N	9	6
N Correct	9	6
Proportion	1.000	1.000

Proportion Correct = 1.000

B **Summary of Classification with Cross-validation**

Put into Group....	....True Poor	Well
Total N	9	6
N Correct	6	5
Proportion	0.667	0.833

Proportion Correct = 0.733

C **Linear Discriminant Function for Group**

	Poor	Well
Constant	-2.7461	-6.8274
GLI1	0.5564	-0.4646
GLI2	0.9089	-2.0837
GLI3	0.0414	-0.0605
SHH	0.0326	0.0955
IHH	-0.4068	0.1212
PTCH1	-1.2270	1.5965
SMO	-0.1538	0.2074
BCL2	0.3182	-0.2604

3.8. COLON CARCINOMA CELL LINES

The list of colon carcinoma cell lines used in this study was given in Table 2.3. The fold expression values of the target genes were calculated and transformed to log scale (Table 3.8).

Table 3.8: Fold expression values of the Hh pathway genes in colon carcinoma cell lines. The value in bold may represent an outlier.

GENE	*GLI1*	*GLI2*	*GLI3*	*SHH*	*IHH*	*PTCH1*	*SMO*	*BCL2*
SW620	-2.70	-6.36	2.44	-0.60	1.99	0.30	3.94	-0.03
SW480	-2.05	4.29	9.91	-0.14	-2.45	-0.75	8.44	-4.08
SW48	-0.70	4.47	**14.12**	-0.52	-2.95	1.77	-8.16	-0.68
HCT8	-0.90	-0.42	-7.93	0.56	-2.84	0.09	-5.56	2.72
HCT15	-1.65	3.58	-1.01	2.94	0.20	2.20	0.29	3.27
HT29	3.80	-1.66	-6.78	4.06	0.91	0.20	-7.01	-2.03
TC71	3.30	0.23	8.87	1.19	-3.00	-2.40	1.19	1.52
T7	1.35	-2.80	-6.19	1.11	4.98	-0.13	0.84	1.27
LOVO	1.65	1.77	5.48	2.23	3.68	-0.08	9.74	4.07
LS411	1.15	0.12	-1.68	-0.27	1.77	0.08	-0.16	0.52
COLO205	-3.05	**-10.80**	-6.61	-4.79	-4.21	1.09	-1.56	-2.38
CL115	-1.45	2.66	-9.37	-2.74	-2.54	-2.15	0.54	-2.63
W05	2.15	-0.16	-4.31	-3.75	0.48	0.85	-0.46	-2.13
KM12	-1.95	2.13	-8.83	-0.11	2.29	-0.81	-7.61	-0.23

3.8.1. Expression status of Hh pathway target genes

The expression status of the Hh pathway target genes in colon carcinoma cell lines were determined by q-rt-RT-PCR. Fourteen human colon carcinoma cell lines were used to obtain the Ct values of each target gene. Each genes Ct value was normalized to the Ct value of the housekeeping gene *GADPH* for a given sample. The Ct values from all the samples were used to calculate the geometrical mean of each target and *GADPH* genes. The geometric mean of the each target gene was normalized to geometric mean of *GADPH* gene. The first normalized values compared to the normalized geometric mean Ct value of the each gene for a given cell line by using the equations 1.2 and 1.3 (section 1.3.2.1.). Observed log fold differences ranged

between 14.12 and -10.80; accordingly, *GLI1, SHH, PTCH1*, and *BCL2* showed relatively more stable expression levels among the cell lines whereas *GLI3* gene expression showed high variability (Table 3.8; Figure 3. 14).

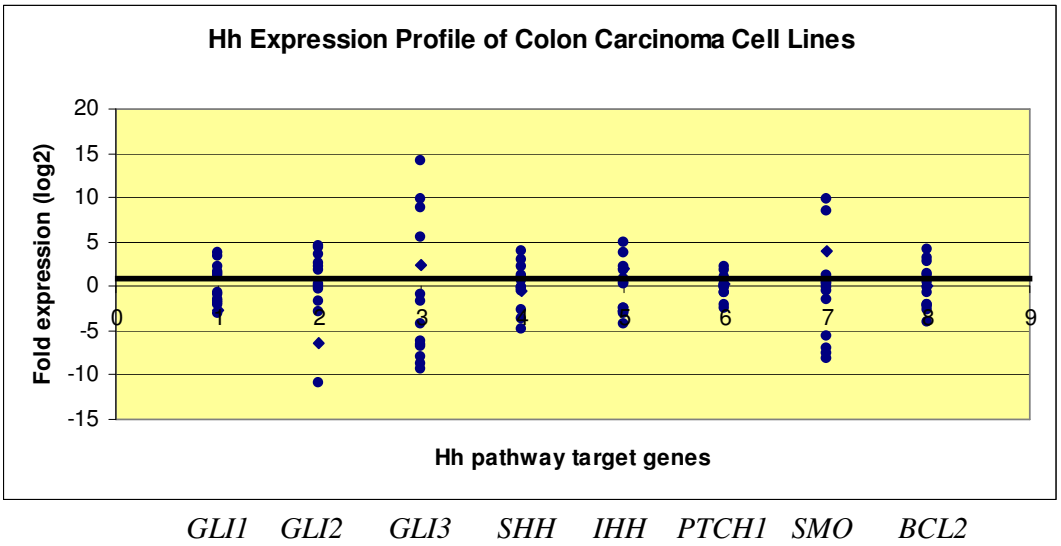


Figure 3.14: Expression profile of the Hedgehog pathway target genes in human colon carcinoma cell lines. Fold expressions were given in log2 scale.

3.8.2. Hierarchical Cluster Analysis

Expression values of all the target genes in 14 human colon carcinoma cell lines were clustered by using the Cluster 2.11 program created by Eisen Lab. Average cluster analysis was performed both in a gene and array based manner and represented as a tree and dendrogram in Figure 3.15.

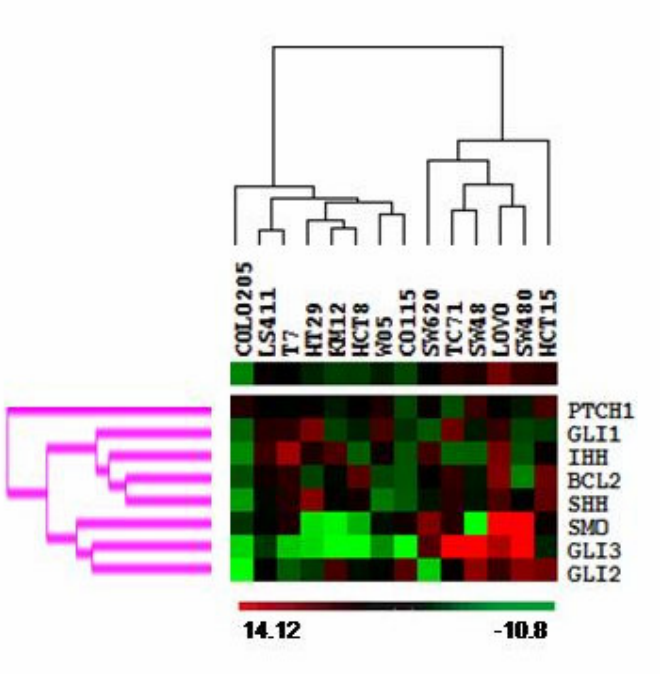


Figure 3.15: Two-dimensional Cluster Analysis of genes in human colon carcinoma cell lines. A dendogram obtained from hierarchical clustering, using expression data from 8 genes in 14 colon carcinoma cell lines. The red and the green colors indicate the high and the low expression values respectively.

3.8.3. Pearson Correlation Analysis

Pearson correlation analysis was applied to the data that was obtained from 14 human colon carcinoma cell lines by using statistical tools in Minitab™ Program. Significant positive correlation was observed between *SHH* and *BCL2* (r=0.547, p=0.043) (Table 3.9). No multiple test correction has been made.

Table 3.9: Pearson Correlation for human colon carcinoma cell lines. The first value represents the Pearson correlation and the second value represents the p value.

	GLI1	GLI2	GLI3	SHH	IHH	PTCH1	SMO
GLI2	0.143						
	0.626						
GLI3	0.066	0.365					
	0.822	0.199					
SHH	0.441	0.371	0.181				
	0.114	0.191	0.535				
IHH	0.299	-0.005	-0.198	0.414			
	0.299	0.988	0.497	0.141			
PTCH1	-0.221	-0.131	0.074	0.018	0.070		
	0.447	0.654	0.802	0.952	0.812		
SMO	-0.037	0.019	0.359	0.014	0.214	-0.236	
	0.900	0.948	0.207	0.963	0.463	0.416	
BCL2	0.199	0.119	0.068	**0.547**	0.392	0.153	0.105
	0.494	0.684	0.817	**0.043**	0.166	0.601	0.721

3.8.4. Multidimensional Scaling

MDS was performed to 14 human colon carcinoma cell lines with an input matrix of Pearson correlation of genes (Table 3.9), and cell lines in the Matlab© program. This analysis was performed to examine the clustering among genes upon reduction of the data dimensions. All of the Hh pathway gene expression values and the 14 cell lines used in this study were taken into consideration for MDS analysis. When MDS was performed with respect to the Hh genes, the resulting graph showed that there was not a strong grouping among genes (Figure 3.16). MDS analysis of the cell lines, on the other hand, revealed mainly two subgroups of cell lines (Figure 3.17). The most striking expression similarities between cell lines included those between HCT8, HT29 and KM12; T7, LS411, and W05.

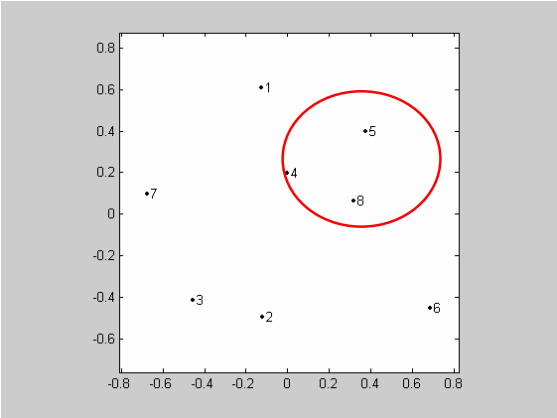


Figure 3.16: Gene-based MDS Analysis for colon cancer cell lines. Correlation values between the expression levels of genes were used as the input data. Each point shown in the map by a number represents one gene. The numbering of genes used in the analysis from number 1 to 8 is *GLI1*, *GLI2*, *GLI3*, *SHH*, *IHH*, *PTCH1*, *SMO*, *BCL2* respectively.

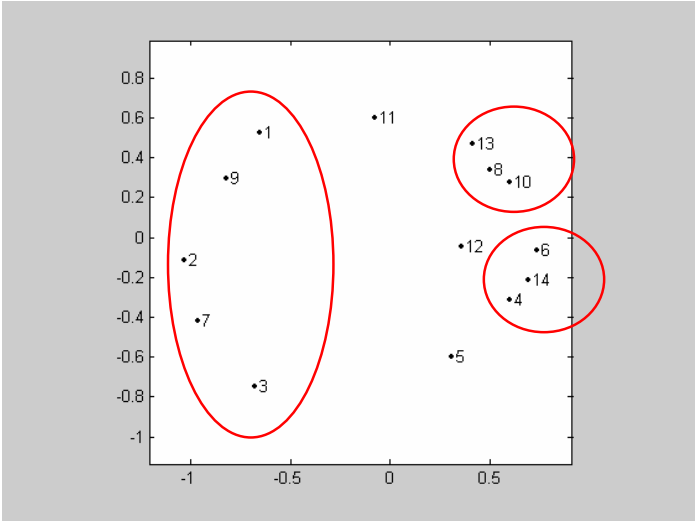


Figure 3.17: Cell line-based MDS Analysis for colon cancer cell lines. Correlation values between the expression levels of cell lines were used as the input data. Each point shown in the map by a number represents one cell line. The numbering of cell lines used in the analysis from number 1 to 14 is in the same order as given in Table 3.8.

3.9. PRIMARY BREAST TUMOR SAMPLES

The list of primary breast tumor samples used in this study was given in Table 2.4. The fold expression values of the target genes were calculated and transformed to log scale (Table 3.10).

Table 3.10: Fold expression values of the Hh pathway genes in primary breast tumor samples. The value in bold may represent an outlier. N.D: not detected. +: fold expression could not be calculated due to lack of expression in normal sample pools.

GENE	*GLI1*	*GLI2*	*GLI3*	*SHH*	*IHH*	*PTCH*	*SMO*	*BCL2*
MFT14	1.98	-0.69	-1.71	N.D.	1.19	-1.13	-3.21	1.96
MFT16	0.83	3.54	2.05	N.D.	-1.41	0.31	3.84	3.31
MFT21	-1.42	2.51	2.05	N.D.	-2.27	-1.98	-1.18	-3.10
MFT25	-0.12	-2.88	-1.99	N.D.	-0.53	-2.31	-6.36	1.16
MFT40	3.73	-2.13	1.31	N.D.	-3.61	-0.04	4.74	-1.04
MFT41	-0.11	-3.85	-2.28	N.D.	-3.62	-3.47	-9.87	-0.44
MFT49	3.01	-3.50	-1.79	N.D.	1.57	-1.55	-5.60	2.98
MFT78	-0.77	1.13	4.21	N.D.	-5.10	-2.27	1.64	-0.29
MFT79	0.38	-2.77	-3.41	N.D.	-5.10	-3.49	-9.36	1.21
MFT90	-1.07	-0.90	-1.14	N.D.	-5.28	-3.14	-6.91	0.51
MFT94	0.43	4.51	0.46	N.D.	-1.23	2.35	3.84	-2.04
MFT 95	4.97	-10.96	3.66	N.D.	1.26	-0.31	-7.79	5.14
MFT97	-1.52	-3.74	-1.42	N.D.	-2.46	-4.05	-7.01	-3.04
MFT113	3.22	1.24	0.73	N.D.	3.10	0.07	-4.64	0.04
MFT115	0.63	1.56	1.59	N.D.	1.41	2.14	4.59	1.21
MFT116	1.13	-0.94	0.89	N.D.	0.78	0.66	-6.03	0.90
MFT117	-2.57	-3.41	-3.81	N.D.	-3.52	-5.01	-7.41	-5.79
MFT127	-1.37	-1.49	-1.99	+	-3.91	-3.62	-5.56	-3.39
MFT120	2.22	0.51	2.42	N.D.	2.77	1.53	-2.94	-0.61
MFT124	0.15	2.11	2.72	N.D.	-1.64	2.95	-1.56	0.57
MFT125	0.11	1.10	1.78	N.D.	0.30	-0.64	-0.25	0.33
MFT129	0.82	2.35	1.53	N.D.	-4.71	-0.15	0.41	-0.06
MFT131	-1.67	-2.02	**29.92**	+	-7.57	-5.97	-17.36	-3.04
MFT132	-2.22	1.94	-1.19	N.D.	-3.96	0.26	-1.41	-1.89
MFT149	27.38	4.29	3.24	N.D.	-3.21	1.44	2.49	2.61
MFT154	-7.62	2.42	3.07	N.D.	-4.71	0.26	-1.86	-10.09
MFT155	-0.57	4.24	4.95	N.D.	-2.29	2.88	2.74	2.56
MFT173	-0.32	3.06	0.11	N.D.	-2.95	-0.13	3.39	0.56
MFT174	-1.17	-13.48	-1.65	N.D.	**-18.47**	-5.45	-14.16	-4.84

3.9.1. Expression Status of Hh Pathway Target Genes

The expression status of the Hh pathway target genes in primary breast tumor samples were determined by q-rt-RT-PCR. Twenty nine primary breast tumor samples were used to obtain the Ct values of each target gene. Each genes Ct value was normalized to the geometric mean of Ct values of the housekeeping genes *ACTB, SDHA*, and *TBP* for a given sample. Then, the normalized values compared to the normalized Ct value of the normal tissue pools by using the equations 1.2 and 1.3 (section 1.3.2.1.). Figure 3.18 shows the expression levels of the primary breast tumor tissues for Hh pathway target genes. Observed log fold differences ranged between 29.92 and -18.47; accordingly, *SMO* exhibited the least amount of variation in terms of mRNA expressions (Table 3.10; Figure 3. 18).

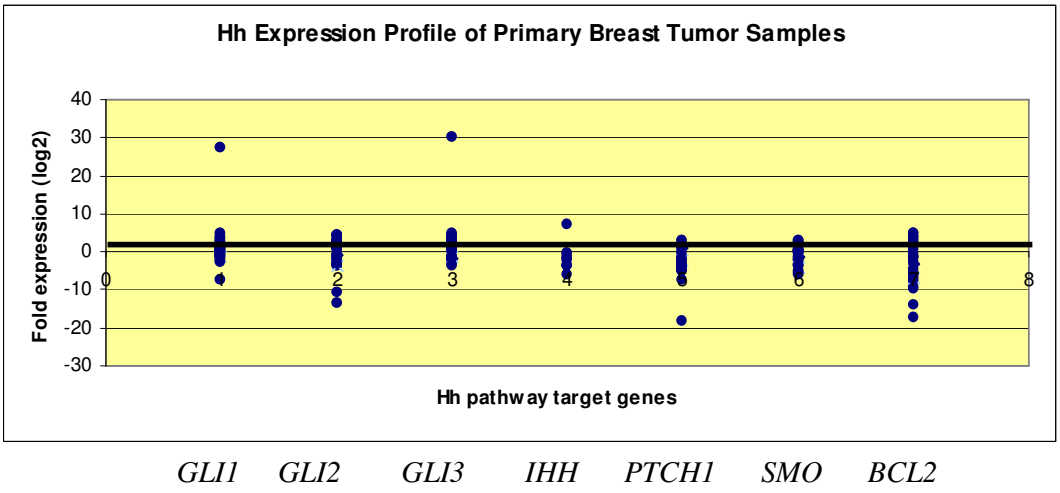


Figure 3.18: Expression profile of the Hedgehog pathway target genes in primary breast tumor samples. Fold expressions were given in log2 scale.

3.9.2. Hierarchical Cluster Analysis

Expression values of all the target genes in 29 human primary breast tissues were clustered by using the Cluster 2.11 program created by Eisen Lab. Average cluster analysis was performed both in a gene and array based manner and represented as a tree and dendrogram in Figure 3.19.

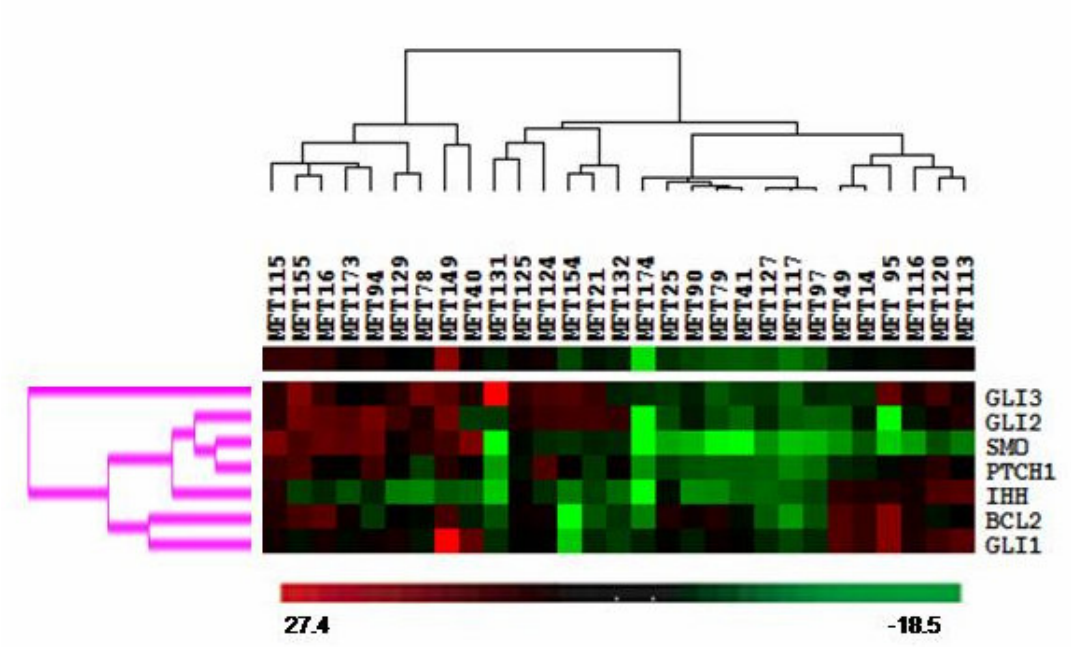


Figure 3.19: Two-dimensional Cluster Analysis of genes in primary breast tumor tissues. A dendogram obtained from hierarchical clustering, using expression data from 8 genes in 29 primary tumor samples. The red and the green colors indicate the high and the low expression values respectively.

3.9.3. Pearson Correlation Analysis

Pearson correlation analysis was applied to the data that was obtained from 29 primary breast tumor samples by using statistical tools in Minitab™ Program. High positive correlations were obtained for *PTCH1* and *SMO* (r=0.772, p=0.000); *PTCH1* and *GLI2* (r=0.725, p=0.000); *SMO* and *GLI2* (r=0.725, p=0.000) (Table 3.11). No multiple test correction has been made.

Table 3.11: Pearson Correlation for primary breast tumor tissues. The first value represents the Pearson correlation and the second value represents the p value.

	GLI1	GLI2	GLI3	IHH	PTCH1	SMO
GLI2	0.119					
	0.540					
GLI3	0.011	0.104				
	0.954	0.591				
IHH	0.170	**0.400**	−0.132			
	0.377	**0.031**	0.496			
PTCH	0.284	**0.621**	−0.093	**0.554**		
	0.135	**0.000**	0.631	**0.002**		
SMO	0.224	**0.725**	−0.231	**0.396**	**0.772**	
	0.242	**0.000**	0.227	**0.033**	**0.000**	
BCL2	**0.497**	0.056	−0.059	**0.482**	**0.380**	0.239
	0.006	0.772	0.759	**0.008**	**0.042**	0.211

3.9.4. Multidimensional Scaling

MDS was performed to 29 primary breast tumor tissues with an input matrix of Pearson correlation of genes (Table 3.11), and cell lines in the Matlab© program. This analysis was performed to examine the clustering among genes upon reduction of the data dimensions. All of the Hh pathway target gene expression values and the 29 tumor samples used in this study were taken into consideration for MDS analysis. When MDS was performed with respect to the Hh genes, the resulting graph showed that there was not a strong grouping among genes except for the *GLI2, PTCH1*, and *SMO* (Figure 3.20). MDS analysis of the 29 primary breast tumor tissues, on the other hand, revealed that there were several subgroups of samples (Figure 3.21). The most striking expression similarities between samples included those between MFT14, MFT25, and MFT49; MFT16 and MFT173; MFT97, MFT113, MFT116, MFT117, and MFT120.

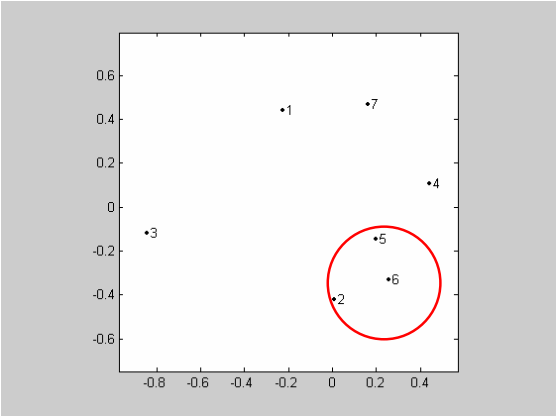


Figure 3.20 Gene-based MDS Analysis for primary breast tumor tissues. Correlation values between the expression levels of genes were used as the input data. Each point shown in the map by a number represents one gene. The numbering of genes used in the analysis from number 1 to 7 is *GLI1*, *GLI2*, *GLI3*, *IHH*, *PTCH1*, *SMO*, *and BCL2* respectively.

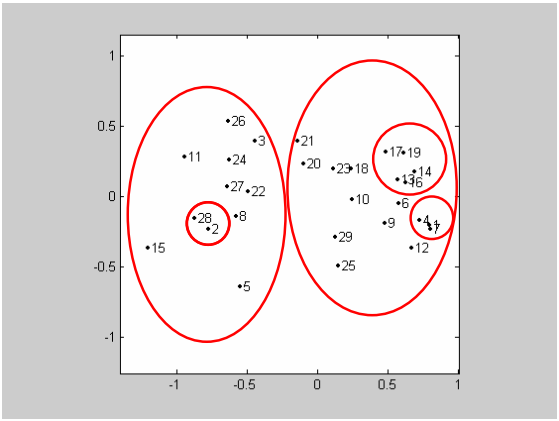


Figure 3.21: Sample-based MDS Analysis for primary breast tumor tissues. Correlation values between the expression levels of tumor samples were used as the input data. Each point shown in the map by a number represents one tumor sample. The numbering of tumor samples used in the analysis from number 1 to 29 is in the same order as given in Table 3.10. Two large clusters are apparent while highly similar sets of cell lines also could be detected within each cluster (shown by red circles).

3.9.5. Discriminant Function Analysis

Linear DFA was performed for 29 primary breast tumor samples. The grouping was performed for the estrogen receptor (ER) status and grade status of the human breast cancer tumor samples (Table 2.4). There were 13 ER (-) and 10 ER (+) tumor samples (Table 2.4). When all the target gene expression values were taken into consideration for ER status, DFA predicted the ER status correctly in 87% of the samples (Table 3.12A). Cross-validation of the DFA results indicated the percentage of correct classification might vary depending on the selection of cell lines since the misclassification rate increased to 35% with cross-validation (Table 3.12B). DFA predicted the grade status correctly in 71% of the samples as shown in Table 3.13A. Correct prediction rate decreased to 57% after cross-validation test (Table 3.13B).

Table 3.12: DFA of primary breast tumor tissues based on their ER status

A **Summary of Classification**

Put into	..True Group..	
Group	**ER-**	**ER+**
Total N	13	10
N Correct	11	9
Proportion	0.846	0.900
Proportion Correct= 0.870		

B **Summary of Classification with Cross-validation**

Put into	..True Group..	
Group	**ER-**	**ER+**
Total N	13	10
N Correct	8	7
Proportion	0.615	0.700
Proportion Correct= 0.652		

C **Linear Discriminant Function for Group**

	ER-	**ER+**
Constant	-2.0837	-0.6783
GLI1	0.8247	0.3559
GLI2	0.6961	0.3972
GLI3	-0.1445	-0.0775
IHH	-0.4129	-0.2446
PTCH	0.1301	0.0424
SMO	-0.5633	-0.2981
BCL2	-0.6052	0.1753

Table 3.13: DFA of primary breast tumor tissues based on their grade status

A **Summary of Classification**

Put into	..True Group..		
Group	**1**	**2**	**3**
Total N	4	11	13
N Correct	4	10	6
Proportion	1.000	0.909	0.462

Proportion Correct= 0.714

B **Summary of Classification with Cross-validation**

Put into	..True Group..		
Group	**1**	**2**	**3**
Total N	4	11	13
N Correct	1	10	5
Proportion	0.250	0.909	0.385

Proportion Correct= 0.571

C **Linear Discriminant Function for Group**

	1	**2**	**3**
Constant	-2.3198	-1.3272	-0.5479
GLI1	0.1414	-0.0132	0.0766
GLI2	0.0113	0.3809	0.1174
GLI3	-0.1326	0.0039	-0.0349
IHH	-0.2477	-0.0873	-0.2644
PTCH1	1.4716	-0.9646	0.6211
SMO	-0.6777	-0.0726	-0.2235
BCL2	-0.2657	0.2707	-0.0329

3.9.6. Mann-Whitney U Test

Mann-Whitney U test was performed to primary breast tumor samples with known ER status in Minitab™. The grouping was performed for the estrogen receptor (ER) status (Table 2.4). There were 13 ER- and 10 ER+ tumor samples (Table 2.4). For each gene Mann-Whitney U test was performed in 95% confidence interval.

Table 3.14: Mann-Whitney test for primary breast tumor tissues based on their ER status. Bold p value represents the significant results.

GENE	ER (-) Median	ER (+) Median	W Value	P Value
GLI1	-1.17	0.49	120	**0.0277**
GLI2	-0.94	0.22	152	0.8282
GLI3	0.46	0.42	145	0.5149
IHH	-3.91	-0.97	124.5	**0.0545**
PTCH1	-1.98	0.72	139	0.3062
SMO	-5.56	-3.925	142.5	0.4201
BCL2	-2.04	0.865	105	0.0017

CHAPTER 4

DISCUSSION

Signaling pathways with a role in development are known to interact with each other and disturbance in one pathway can influence the regulation of others. It is therefore important to study these signaling pathways in cancer.

Dysregulation of Hh pathway has been implicated in the formation of different types of cancer. The research describing the role of Hh pathway in cancer mainly focus on particular genes of the pathway, namely *SHH, PTCH1*, and *GLI1*. The observed phenotypes of this pathway in different types of cancers are the mutations in the *PTCH1* gene and/or overexpression of the other genes in this pathway. However misregulation of the remaining genes having role in this pathway may disturb the homeostasis of the pathway.

In this study, the expression profiles of seven genes, *SHH, IHH, SMO, PTCH1, GLI1, GLI2* and *GLI3*, that play major roles in the Hh pathway, and a downstream target of Hh pathway which has an important role in cancer progression, *BCL2,* were analyzed in three different types of human epithelial cancer cell lines and primary breast tumor samples by quantitative real-time RT-PCR technique. Fifteen breast, 14 colon, and 15 hepatocellular carcinoma (HCC) cell lines were used in this study. To the best of our knowledge, this is the first comprehensive study that shows the expression profiles of the main target genes of the Hh pathway in cancer cell lines that have an epithelial cell origin.

An association between the Hh pathway activation and the initiation of human tumors have been shown [Magliano *et al.,* 2003], but not effectively studied in breast cancer. In addition to the cell lines, the expression profiles of the genes named in this study were analyzed in 29 primary breast tumor samples and three matched normal tissue sample pools by quantitative real-time RT-PCR. Evidences indicate that aberrant Hh pathway signaling may be acting in a ligand-dependent or ligand-independent manner depending on the type of cancer [Regl *et al.,* 2004]. Our results showed that each target gene analyzed in this study was differentially co-regulated in different types of carcinoma cell lines. Hierarchical cluster and Pearson correlation analyses results indicated that different components of the pathway may be responsible for activating this pathway in each cancer type.

4.1. EXPRESSION ANALYSIS IN BREAST CARCINOMA CELL LINES

One type of epithelial cancer cell line used in this study was the human breast carcinoma cell lines (Table 2.1, Section 2.1). Fifteen cell lines were used for quantitative real-time RT-PCR and the results were analyzed by using different statistical analysis. Expression of *GLI3, PTCH1, SMO* and *BCL2* were increased in all of the cell lines analyzed except for SKBR3 as shown in Figure 3.6 and 3.7. In SKBR3 cell line, *BCL2, SMO* and *GLI3* expressions were found to be low. Kubo *et.al.* (2004) showed increased expression of *SHH, PTCH1*, and *GLI1* in BT474, SKBR3, MDA-MB231, and MCF7 cell lines by immunostaining. The authors also observed clear nuclear staining of GLI1 in three cell lines except MDA-MB231. However, increase in the mRNA level of *SHH* and *GLI1* were not observed in our study for SKBR3 and MCF7 cell lines. In MDA MB231 cell line, one-fold increase was detected in *SHH* expression whereas *GLI1* expression was low. *PTCH1* was upregulated in all the cell lines, and all these genes were upregulated in BT474. This finding may be explained by the basal level of mRNA expression may be enough for the detected protein expressions or proteins accumulate in these cells.

GLI2 mRNA expression showed high variability in breast carcinoma cell lines. In MDA MB453 and T47D cell lines, expression was decreased over 15-fold. Explanation for these findings may come from the studies on the inhibitors of the Gli transcription factors such as protein kinase A and analysis of epigenetic changes in the *GLI2* promoter region in these cell lines [Magliano *et al.,* 2003]. *SHH* and *IHH* gene expressions, which are the ligands of the Hh pathway, were not observed in most of the cell lines; this finding suggests that if there is any pathway activity it is likely to be through a ligand-independent manner in these cell lines or basal level expression of the ligands are sufficient for an aberrant pathway activity. Ligand-independent signaling has been shown to be the result of either the loss of function mutations of *PTCH1* or the gain of function mutations of *SMO* [Regl *et al.,* 2004].

In most cases of the cancer types where Hh pathway is a cause for the disease, expression of the *GLI1, PTCH1*, and *SMO* have been found to be upregulated [Toftgard, 2000]. *SMO* and *GLI1* are thought to be oncogenes responsible for the aberrant pathway activity. The receptor *PTCH1* is also a downstream target of the pathway and its increased transcription indicates an active pathway. In the breast carcinoma cell lines, high *GLI1* expression was not detected in most of the cell lines, only in CAMA1 and BT474 cell lines had increased expression of *GLI1*. If the Hh pathway is active, it should be through other transcription factors of the pathway, such as GLI2. GLI2 was found to be highly expressed in five cell lines, MDA MB157, MDA MB231, MCF12A, HCC1937, and HBL100, and other cell lines had low *GLI2* expression. Gli-independent Hh signaling has also been implicated in some developmental processes [Kato *et al.,* 2001], which may be responsible for the active pathway in the breast carcinoma cell lines.

Expression values of all the target genes for 15 human breast carcinoma cell lines were clustered by using two dimensional Hierarchical cluster analyses. Clustering resulted in two major groups as shown in Figure 3.7. In the small cluster formed by MCF12A, HBL100, MDA MB231, HCC1937, and MDA MB157 target genes had upregulated expression profile except for *IHH* and *GLI1*. *GLI2* expression seems to

be the most definitive factor in the groping of the cell lines. When *GLI2* expression was not included clustering changes completely. However when *PTCH1* or *GLI1* expressions were not included there was no change in the two major clusters, but only minor changes in small subgroups. This finding was also confirmed by stepwise regression analysis (data not shown). This data shows that at least in some breast cell lines aberrant Hh pathway activity is present and induction of target gene expression might occur through the *GLI2* transcription factor. It was indicated that *GLI1* and *GLI2* regulate a similar set of target genes during mouse embryogenesis. But later in development, new gain-of-function defects can emerge, suggesting that context-dependent regulation of biological activities of GLI factors [Bai *et al.,* 2001]. Whether *GLI1* or *GLI2* or a combination of both factors is required for Hh-induced tumorigenesis is not clear yet.

Pearson correlation analysis was used to determine if there is a linear relationship between the two genes' expression patterns in the breast carcinoma cell lines studied. The analysis showed that *IHH, SMO* and *GLI1* gene expression profiles were found to be highly correlated with each other in all the breast carcinoma cell lines studied. There was a strong positive correlation between *SMO* and *GLI1* (r=0.637, p=0.011) and *SMO* and *IHH* (r=0.708, p=0.003) and weak correlation between *GLI1* and *IHH* (r=0.521, p=0.047) (Table 3.3). Although there was a correlation among these genes, not all of them were clustered together in the Hierarchical cluster analysis (Figure 3.7). In the cell lines where *IHH* and *GLI1* were increased, *SMO* expression was also high. But since *SMO* was increased in all the cell lines they did not cluster together in the Hierarchical cluster analysis. Hierarchical cluster analysis not only includes the Pearson correlation values, but also expression levels of the cell lines. That is why it is meaningful that these genes did not cluster together in the hierarchical clustering. *GLI1* is the downstream target of the Hh pathway. So when there is a high *IHH* expression it is expected to have an upregulated *GLI1* expression. *SMO* upregulation thus should correlate with *GLI1* expression as well since *SMO* is the receptor that mediates the downstream signaling when the ligand is bound.

To identify the distance between the genes and the distance between the cell lines, Multidimensional Scaling (MDS) was performed. The correlated genes *IHH, SMO*, and *GLI1* did not cluster together in the gene based MDS analysis although they were closer to each other than they were to other genes based on their correlation coefficients (Figure 3.8). The small group formed in Hierarchical cluster analysis was also formed in the cell line based MDS analysis as shown in Figure 3.9. There were two small groupings as well which were formed in Hierarchical cluster analysis. MDS analysis confirmed the Hierarchical cluster and Pearson correlation analysis.

ER status is an important parameter in the prognosis of the breast cancer which helps in determination of treatment strategy for breast cancer patients. Patients with high ER level respond to hormone therapy better than the others. Recent studies showed that spatiotemporal expression of *SHH* can be regulated by steroid hormones in the development of several organs [Pu *et al.,* 2004]. Estrogen exposure can suppress the *Shh* expression in the developing rat prostate [Pu *et al.,* 2004]. Genes can be differentially expressed during the progression from estrogen dependent state to estrogen independent state. To investigate the presence of a potential role of Hh pathway target gene expressions with respect to ER status of the cell lines, linear Discriminant Function Analysis was performed. Expression profiles of the target genes predicted the ER status correctly in 93% of the cell lines studied (Figure 3.4A) although cross-validation seemed to have decreased prediction power to 60%. *GLI1, SMO,* and *BCL2* expression patterns affected the ER status the most (Figure 3.4B). *BCL2* expression was upregulated two fold in the ER (-) cell lines while induced four-fold in the ER (+) cell lines. Similarly, *SMO* expression showed a four-fold and eight-fold increase in ER (-) and ER (+) cell lines, respectively. Generally these results show that, Hh pathway target genes might exhibit higher expression in ER (+) cell lines. ER positivity is frequently associated with well-differentiated tumors, while more poorly differentiated cancers, and cancers with significant lymphocytic infiltration, are more likely to be ER (-) [Stanford *et al.,* 1986]. In Sorlie's microarray study higher *SMO* expression has been observed in non-metastatic tumor samples than metastatic ones ($t=2.937$; $p=0.012$) [Sorlie *et al.,* 2003], which

correlates with our study since ER (+) tumor samples have been thought to be less metastatic than ER (-) tumor samples.

In conclusion, human breast carcinoma cell lines have different levels of target gene expression in the Hh pathway, such that this pathway is activated only in some of the breast carcinoma cell lines possibly through a ligand-independent pathway. It is difficult, however, to draw a definite conclusion from the results obtained in this study based on the breast cancer cell lines that a specific gene or a set of genes of the Hh pathway is the main cause of the breast cancer formation. One way to obtain more insight to this question can be to use primary tumor samples, which were also included in this study.

4.2. EXPRESSION ANALYSIS IN HEPATOCELLULAR CARCINOMA CELL LINES

In this study 15 human hepatocellular carcinoma cell lines were used to analyze the expression level of the defined target genes with q-rt-RT PCR. In SNU182, SNU423, and SNU449 cell lines, all the genes were upregulated as shown in Figure 3.11. This may indicate the active Hh pathway in these cell lines. *GLI1* is thought to be main transcription factor which shows the aberrant pathway activity [Toftgard, 2000; Wetmore 2003]. In our study, GLI expression was high in 40% of the HCC cell lines and low in 47% of the cell lines, and did not change in the remaining 13% (Figure 3.10). In the cell lines where *GLI1* expression was low, the expression level of other transcription factors were also low in general. The cell lines that express high *GLI1* and *GLI2* were SNU182, SNU387, SNU423, and SNU475. *SHH* and *PTCH1* expressions were also found to be increased in these cell lines. This indicates that the ligand dependent pathway might be a more likely route in these HCC cell lines. *GLI3* mRNA expression varied between the cell lines extensively, and it is almost downregulated in the four well differentiated HCC cell lines and SNU475. Microarray study performed by Chen *et al*. (2002) has showed that *GLI2, PTCH1*, and *BCL2* expressions were decreased in HCC tumor tissues as compared to non-tumor tissues while the expression of *SMO* increased. We found correlated results with Chen's microarray study: *PTCH1* expression decreased in 60% of the cell lines,

however expression of *GLI2* and *BCL2* were decreased only in 40% of the cell lines. *SMO* expression was increased in 73% of the HCC cell lines in our study. Increased *SMO* expression may lead to constitutive activation of the Gli transcription factors although their expressional levels were low.

Mature hepatocytes lack Hh pathway activity despite the liver's requirement for Hh signaling during embryogenesis [Sicklick *et al.,* 2006]. Previous studies have shown that *GLI1* and *SMO* expressions were higher in HCC tumor samples compared to paired-normal tissues whereas *SHH, IHH*, and *PTCH1* expressions did not change between the tumor-normal pairs [Patil *et al.,* 2006]. Wang *et al.,* (2006) showed *GLI1* and *PTCH1* expression in 50% of the hepatocellular carcinoma samples by *in situ* hybridization. In another study, *SHH, IHH, PTCH, SMO,* and *GLI1* expression were upregulated in HEP3B and HEPG2 cell lines [Sicklick *et al.,* 2006]. Our findings were correlated with Sicklick *et al.,* for the *SHH, IHH,* and *SMO* genes expression profiles. However upregulation of *PTCH1* and *GLI1* in HEP3B and HEPG2 were not observed in our study.

Hierarchical cluster analysis differentiated the cell lines into four subgroups based on their expression profile. In the small group at the right part of the Figure 3.11 all the target genes analyzed except for *IHH* had increased expression. These cell lines had elevated Hh pathway activity through increased *SHH* expression. The group formed by HEPG2, HUH7, HEP3B TR, HEP3B, and SNU 475 had decreased transcription factor expression. In the aggressive tumors, *GLI3* expression may be lost to induce Hh target gene expression since in the cell lines where *GLI3* expression is downregulated target gene induction occurs through the derepression of the promoter region [Ruiz I Altaba, 1999]. The third subgroup was formed by SNU398, SNU387, MAHLAVU, and FOCUS. These cell lines showed elevated expression of transcription factors *GLI2* and *GLI3* whereas reduced expression of *SMO* and *IHH*. The fourth group was formed by PLC, HEP40, and SKHEP1 of which *PTCH1* and *SHH* expressions were reduced.

Pearson correlation analysis was performed to find out the linear relationship between the expression patterns of the genes. There was a strong positive correlation between *PTCH1* and *GLI1* (r=0.586, p=0.022) and *PTCH1* and *GLI2* (r=0.661, p=0.007) (Table 3.6). These genes clustered together in the Hierarchical cluster analysis (Figure 3.11). *GLI1* and *PTCH1* are the downstream targets of the Hh pathway, and hence increased *GLI2* can bind to their promoter region and induce their expressions. Significant negative correlation between *IHH* and *GLI2* (r=-0.647, p=0.009) and *IHH* and *GLI3* (r=-0.721, p=0.002) genes were identified in the HCC cell lines studied and as expected these genes did not cluster together in the Hierarchical cluster analysis. Having a negative correlation between the expressional regulation of ligands and transcription factors is not an expected finding. Instead, an increased ligand expression should have led to increased transcription factor expression not to the decreased one unless there is an alternative pathway in which the ligand is used to activate other downstream targets while default transcription factors are repressed.

Multidimensional Scaling (MDS) was performed to determine the distance between the genes and the distance between the cell lines based on the correlation between their expression profiles. The correlated genes *PTCH1, GLI1*, and *GLI2* clustered together in the gene based MDS analysis (Figure 3.12). Three small clusters were formed in the cell line based MDS analysis as shown in Figure 3.13. Two of the clusters were also formed in the Hierarchical cluster analysis. However, the group formed by HEP3B, HEP3B TR, SKHEP1, and PLC/PRL/5 did not clustered in the Hierarchical clustering (Figure 3.11), but belong to two different clusters. SKHEP1 and PLC/PRL/5 clustered with HEP40 in the Figure 3.11 although HEP40 was far from these cell lines in the MDS analysis. These findings confirmed the Hierarchical cluster analysis in most respect.

The hepatocellular carcinoma cell lines can be defined as two groups based on their differentiation status. HEP3B, HEP3B TR, HEP40, HEPG2, HUH7, and PLC/PRL/5 are well differentiated cell lines while others, SNU 182, SNU 387, SNU 398, SNU 423, SNU 449, SNU 475, SKHEP1, FOCUS, and MAHLAVU, are poorly

differentiated (Table 2.2). While *GLI2* gene had decreased expression in all well differentiated cell lines, *SMO* and *IHH* genes were highly expressed. These expression patterns differentiate the well differentiated cell lines from poorly differentiated ones in DFA analysis (Table 3.7). High *SMO* and *IHH* expressions have been found to be markers for aberrant Hh pathway activity in the well differentiated cell lines. However, due to reduced transcription factor expression Gli-independent signaling pathway may be responsible for the active pathway in the well differentiated HCC cell lines. In view of the fact that these expression patterns differentiate the poorly differentiated cell lines from well differentiated ones in DFA analysis perfectly (Table 3.7), this finding may help to identify differentiation status of the HCC samples, and treatment of the disease with further experiments. In the microarray study performed by Okada *et al.,* (2003) HCC grade 1 tumors have been found to have decreased *IHH* expression than grade 3 ($r=0.35$; $p=0.006$). Since grade 3 tumors are more undifferentiated than grade 1 tumors this finding does not in accordance with our results. Because carcinoma cell lines may not reflect the real tumor environment, having a different phenotype is not surprising.

To conclude, human HCC cell lines are thought to have Hh pathway target gene expressions, but that does not mean that Hh pathway is a cause for the hepatocellular carcinoma formation. To further investigate the role of Hh pathway in HCC, same study should be carried out in the tumor samples, and also functional protein detection should be performed for the pathway components.

4.3. EXPRESSION ANALYSIS IN COLON CARCINOMA CELL LINES

Human colon carcinoma cell lines were the third type of epithelial cell lines used in this study. In the adult human colon, *SHH* mRNA expression was determined in the epithelial cells at the base of normal intestinal crypts, which suggest a role for Hh signaling in the stem cell maintenance of colon [Qualthrough *et al.,* 2004] similar to Wnt pathway. Dysregulated Wnt pathway is a cause for the colon carcinoma. The Hh

pathway may have a role in colon carcinoma. There are contrary findings on the role of the Hh signaling in colon carcinoma. One study suggests an inhibitory role for *GLI1* in the progression of the colon cancer through inhibition of Wnt signaling [Akiyoshi *et al.,* 2006]. The opponents propose an autocrine Hh pathway signaling that can increase cell survival in the colorectal tumor cells [Qualthrough *et al.,* 2004]. Qualthrough *et.al.* showed expression of *SHH, IHH, PTCH1, SMO* and *GLI1* in the carcinoma-derived cell lines HT29 and SW480 in both mRNA and protein level by RT-PCR and western blotting respectively. Qualthrough *et.al.* also proposed that there is a change in the role of the signaling during progression to invasive carcinoma because adenoma-derived cell lines had more *IHH* expression than carcinoma-derived ones.

Our findings do not correlate with the study of Qualthrough *et al.,* (2004). In SW480 colon carcinoma cell line, only for *SMO* gene increased expression was observed: SHH*, IHH, PTCH1,* and *GLI1* expressions were decreased. In HT29 cell line, *SHH, IHH,* and *GLI1* showed increased expression while *PTCH1* and *SMO* expressions were much lower (Table 3.8). This contradiction between the two studies may partly be explained by the experimental strategy followed. By performing RT-PCR, only the presence or absence of the mRNA can be identified. To determine the upregulation or downregulation as compared to normal tissue, quantitative real time RT-PCR should be performed. Qualthrough *et al.,* (2004) carried out RT-PCR reactions for 35 cycles, which is above the Ct values we obtained for these genes in our q-rt-RT-PCR reactions in these cell lines except for the *SMO* expression in HT29 (Appendix XX). In our study no *SMO* expression was detected in HT29 cell line at all, which was also determined by Zhu *et al.,* (2003) by RT-PCR. They showed that *SMO* is fully methylated in the HT29 and HCT15 cell lines, and partially methylated in the LOVO cell line. In our study we did not detect *SMO* expression in HCT15 as well, but 10-fold increased expression was determined in LOVO as shown in Table 3.8. SW480 was almost completely unmethylated, and showed 8-fold increased *SMO* expression in our study.

To further evaluate our findings, microarray studies for Hh pathway expression were analyzed. Only three studies with significant results were obtained. *GLI1* and *BCL2* expressions have been shown to be decreased in the adenocarcinoma samples as compared to normal colon tissues in Alon's (1999) and Notterman's (2001) microarray studies respectively. In addition, Scherf's (2000) cell line microarray study has showed that *SMO* expression was lower in colon carcinoma cell lines as compared to other carcinoma cell lines (t=-4.754; p=3.1E-4). Only in four cell lines we observed increased *SMO* expression, in others it was decreased or not changed in our q-rt-RT-PCR assays. *GLI1* and *BCL2* expressions were decreased in 50% and 43% of the cell lines respectively in our study while in 43% of the cell lines *GLI1* expression was high, and in 36% of the cell lines *BCL2* had increased expression.

Human colon carcinoma cell lines showed deregulated expression profile among the genes. Two dimensional Hierarchical cluster analyses were performed to colon carcinoma cell lines to observe the data obtained from the q-rt-RT PCR reactions easier. Clustering resulted in two major groups as shown in Figure 3.15. In the cluster formed in the right part of the figure by SW620, TC71, SW48, LOVO, SW480 and HCT15, *SMO* and *GLI3* genes had upregulated expression profile. *GLI3* expression seems to be the most definitive factor in the grouping of the cell lines. In general, Hh pathway components studied in this research did not show significant upregulation except for LS411, LOVO, and HCT15 cell lines. Aberrant pathway activity was observed only in these cell lines.

Pearson correlation analysis was performed to human colon carcinoma cell lines and a significant positive correlation only between *BCL2* and *SHH* (r=0.547, p=0.043) (Table 3.9) was observed. These genes clustered together in the Hierarchical cluster analysis (Figure 3.15) as well. There was no significant negative correlation between any genes in the colon carcinoma cell lines studied, which may mean that in the colon carcinoma cell lines, Hh pathway genes may act independently from each other.

Multidimensional Scaling (MDS) was performed to determine the distance between the genes and the distance between the colon carcinoma cell lines. In this study, Pearson correlation was used as the input matrix for the distance. The correlated genes *SHH* and *BCL2* clustered together in the gene based MDS analysis although *IHH* was also the same distance to them (Figure 3.16) which were clustered together in Hierarchical cluster analysis, too. The cluster formed in the right part of the Hierarchical cluster analysis was also formed in the cell line based MDS analysis as shown in Figure 3.17. There were two small groupings as well which were formed in Hierarchical cluster analysis. The cell lines clustered together in the MDS analysis were also clustered together in the Hierarchical cluster analysis. However for the group formed by T7, LS411, CL115, and W05, hierarchical clustering separated these cell lines into two different groups. Since it was based on expression profile of the cell lines along with the correlation coefficients it is reasonable to have different grouping there.

To sum up, human colon carcinoma cell lines do not seem to have aberrant Hh pathway target gene expression, however we can not drawn a conclusion from the cell lines' study whether Hh pathway is active or inactive in the colon carcinoma Cell line studies generally do not reflect what happens *in vivo* in most cases. Hence, the same study should be performed in the tumor samples, and also functional protein detection should be performed for the pathway components along with mutation analysis.

4.5. EXPRESSION ANALYSIS IN PRIMARY BREAST TUMOR TISSUES

Pathways with a role in cancer formation have been shown mostly by studies done in tumor samples, which is the most reliable way to identify causes of cancer. Observing the increased Hh pathway component expression in human breast carcinoma cell lines prompted us to further evaluate the pathway in the primary breast tumor tissues. For this purpose q-rt-RT-PCR analysis was performed on 29

primary breast tumor samples with known clinical properties. All the tumor samples were chosen from invasive ductal carcinoma patients because the Hh pathway has been implicated in the development of mouse mammary ductal histoarchitecture [Lewis *et al.,* 2004].

Kubo *et.al.* (2004) showed the constitutive activation of the Hh pathway in human breast carcinoma for the first time by performing immunohistochemistry to 52 tumor samples. Carcinoma tissues overexpressed SHH, PTCH1, and GLI1 proteins in all the samples studied whereas no protein was detected in normal tissues. In our study, we did not detect any *SHH* mRNA expression in either in tumor samples or in normal sample pools as shown in Table 3.10. Besides, *PTCH1* and *GLI1* mRNA expressions were highly increased in only 6 and 8 of the samples respectively. Accumulated protein levels may be responsible for the finding of Kubo *et al.,* (2004) or since tumor samples used are different this contradictory findings are reasonable. Since we did not detect any *SHH* expression in tumor samples and normal sample pools we studied, we searched for possible alternative transcripts of the *SHH* in the literature and ENSEMBL database, which may not be recognized by our primers. The primers we used for SHH corresponds to exon 1 and exon 2 of the *SHH* gene. We did not found any alternative transcripts of the gene. However, *SHH* gene has a high CpG island content which may be responsible for hypermethylation of the gene in the primary breast tumor samples we studied which should be checked with methylation specific PCR.

GLI1 and GLI2 are the transcriptional activators in the Hh pathway. When we look at individual tumor samples, in 69% of the samples at least one of the genes had increased expression. Hence, in 69% of the tumor samples Hh target gene induction may occur. Only in 28% expressions of both transcription factors were low. There may be functional redundancy between these genes in breast cancer, and lack of increased expression of one gene may result from the epigenetic changes in the promoter region of the genes.

Expression values of all the target genes for primary breast tumor samples were clustered by using two dimensional Hierarchical cluster analyses. Clustering resulted in three major groups as shown in Figure 3.19. The cluster formed in the left side of the figure had elevated expressions of the target genes except for the *IHH*. There is a small subgroup in the cluster at the right part of the figure, in which all the genes' expressions were downregulated. There was no significant pathological property that differentiates these groups from each other. Hierarchical cluster analysis shows that in most of the tumor samples Hh pathway target gene expression was decreased. Gene based clustering showed that *GLI2* expression was correlated with receptor expressions in primary breast tumor tissues and *BCL2* was correlated with *GLI1*. *GLI1* may be the transcription factor inducing *BCL2* expression in these tumor samples.

Pearson correlation analysis was performed for primary breast tumor samples to determine the linear relationship between the expression patterns of the genes, and to confirm the data obtained from the Hierarchical cluster analysis. Strong positive relationships were found between *PTCH1, SMO*, and *GLI2* expressions (Table 3.11), which were clustered together in the Hierarchical cluster as well. This finding may be explained by sequential induction of the genes. If the *SMO* expression is increased, *GLI2* transcription factor will be active and induce the expression of *PTCH1*, which is a downstream target of the Hh pathway. *GLI2* may be the major transcription factor inducing Hh target gene induction in the samples where *SMO* and *PTCH1* expressions were increased. *IHH* had a correlation with these genes as well, but this correlation was not highly significant as shown in Table 3.11. Correlation between the *BCL2* and *GLI1* was confirmed by Pearson correlation analysis, too. *BCL2* had a correlation with *IHH* and *PTCH1* as well, but these values were not highly significant although had a significant p value. There was no significant negative correlation between any genes in the tumor samples used in this study which shows that genes do not act opposite to each other.

MDS was performed to Hh pathway expression profile in a sample- and gene-based manner to visualize the distance between the samples, and between the genes. In the

gene-based MDS analysis, the correlation between the *PTCH1, SMO* and *GLI2* was confirmed as shown in Figure 3.20. Other genes were at very distant from this cluster based on their expression profile. *GLI3* was the outlier of the genes, which was also placed at a distance from the other genes in the Hierarchical cluster. Since *GLI3* is thought to have a repressor activity when the pathway is active, it is expected to have adverse expression profile with the other genes studied.

Sample-based MDS analysis gave several small clusters of the samples based on the correlation between their expression profiles. The most striking clusters were those between MFT16 and MFT173; MFT14, MFT25, and MFT49; and MFT97, MFT113, MFT116, MFT117, and MFT120. These clusters did not correspond to the Hierarchical cluster analysis. Besides, the most obvious clusters formed in the Hierarchical clustering were spread in the MDS distance map, and mixed with other samples. MDS was not sufficient to differentiate the groups clearly. The large sample size may be responsible for this outcome.

Knowledge of the steroid receptor levels of human breast cancer is important for diagnosis and for deciding the appropriate treatment for breast cancer In a study, a positive correlation between percentage of nuclear staining of *GLI1* and ER status was found by performing Mann-Whitney U test [Kubo *et al.,* 2004]. This corresponds to our finding with the *GLI1* gene and ER (+) status in our tumor samples which showed a positive correlation with Mann-Whitney test (W=120, p=0.0277). In addition, a positive correlation was found for *BCL2* and ER (+) with the same statistical test (W=105, p=0.0017) (Table 3.14). To determine if the expression profiling of Hh pathway genes discriminate the ER status of the tumor samples Discriminant Function analysis was performed. Gene expression profiles of the samples predicted the ER status correctly in 87% of the tumor samples studied (Figure 3.12A). Although the cross-validation test decreased prediction power to 65%, there was still a very strong prediction for ER status of the samples with the expression of the genes included in this study.

PTCH1 and *BCL2* expression patterns seem to have an effect for the ER status (Figure 3.12B) since their slopes are opposite to each other. Their expression values still did not discriminate too much between ER (+) and ER (-) tumor samples. This finding does not fully correlate with the cell lines' DFA analysis. In the cell lines, *PTCH1* expression was not affected in ER (-) cells whereas in the tumor samples it was increased in ER (-) tumors relative to the ER (+) cells which was also shown in Sorlie's microarray study [Sorlie *et al.,* 2003]. *BCL2* expression was decreased in ER (-) tumor samples, while it did not change in ER (+) ones. However, in the breast carcinoma cell lines, BCL2 expression was increased 4-fold in ER (+) cell lines, and 2-fold in ER (-) cell lines. Indeed, almost all the genes studied seemed to have basal level expression in ER (+) samples. Accordingly, it is possible that tumor samples exhibit larger variability with respect to the link between Hh signaling and ER status as suggested by the cell line study; this might be because multiple signaling pathways might carry out the downstream effect of ER status. Moreover, it is well known that cell line studies cannot fully explain real cancer cases due to factors such as divergence in the microenvironments tumors encounter.

There are lots of microarray studies that aim to identify expression profiles specific to grade of the tumor samples. With knowledge of the genes that differentiates the grade; it will be easier to treat the disease by targeting these genes specifically. Microarray studies performed by Zhao *et al.,* (2004) and Van't Veer *et al.,* (2002) have showed that expression of *GLI1, GLI2*, and *BCL2* are higher in grade 1 tumors than in grade 3 tumors. Hence, DFA was performed for the grade status of the samples to find out whether Hh pathway had any predictive effect. DFA predicted the grade status correctly in 71% of the samples as shown in Table 3.13A, which decreased to 57% after cross-validation test (Table 3.13B). 71% is a high percentage to predict the grade status yet its power may be increased by including other genes' expression profiles. *PTCH1* gene expression profile was the most predictive value in the discrimination of the grade status (Table 3.13C). In the grade 1 tumors, its expression was 1.5 fold increased and 1 fold decreased in the grade 2 tumors. Since the invasiveness of the disease increase with higher grade, loss of *PTCH1* expression in grade 2 tumors may be explained by early role of hedgehog signaling in the cancer

progression. In the grade 3 tumors there was still 0.6 fold increased *PTCH1* expression but this was not significant. Since we did not detect increased expression profile for Hh pathway target genes, these predictions may not reflect the grade status perfectly.

In conclusion, primary breast tumor samples did not have aberrant Hh pathway gene expressions although in a small group there were elevated expressions, but that does not mean that Hh pathway is not active in these samples. To make such conclusion functional protein detection should be performed for the pathway components and possible mutation analysis should be carried out. In the tumor samples where expression of pathway components was high, there was no *SHH* expression and *IHH* expression was decreased. Thus, ligand-independent pathway activity may exist in these samples due to mutations in the pathway components, which was also found in the breast carcinoma cell lines. One other possibility is elevated expression of *DHH*, which has a very restricted expression profile in embryonic development [Bale, 2002].

CHAPTER 5

FUTURE PERSPECTIVES

Signaling pathways generally reactivated in cancer due to increased expression of the ligands and/or mutations in the pathway genes. Gain of function genetic events in oncogenes by DNA mutations, rearrangements or amplifications, and loss-of-function mutations in tumor suppressor genes are the molecular events leading to aberrant pathway activity. In the Hh signaling, *SMO, SHH, IHH, GLI1,* and *GLI2* are known oncogenes and *PTCH1* and *SUFU* can function as tumor suppressors [Magliano *et al.,* 2003; Wetmore, 2003]. There have been studies showing *PTCH1* mutation in two of seven human breast cancers [Xie *et al.,* 1997], and linking the polymorphism in the 3'end of the *PTCH1* to increased breast cancer risk associated with oral contraceptive use [Chang-Claude *et al.,* 2003]. In addition, mice showing hyperplastic defects in the mammary gland of ΔPtch1 + and ΔGli1 mutants suffer from breast cancer [Lewis *et al.,* 2004]. In the cell lines and tumor samples studied, possible mutations of Hh signaling should be checked to investigate the contribution of mutations to the expression profiles of the genes and to cancer formation. Such a study will show the mechanism by which Hh pathway is activated in the epithelial cell lines and tumor samples studied.

Furthermore, epigenetic changes in the Hh pathway may be responsible for the aberrant pathway activity. There have been studies showing that *SMO* is fully methylated in the HT29 and HCT15 colon carcinoma cell lines, and partially methylated in the LOVO cell line [Zhu *et al.,* 2003]. Besides, *SUFU* expression may be downregulated through promoter methylation [Sheng *et al.,* 2004]. Analysis of these epigenetic changes in the epithelial cell lines studied may help in understanding

of the observed gene expression profiles. Especially SKBR3 is a good candidate cell line for this study since all the genes shown to have low expression in this cell line.

BCL2 expression was not correlated with *GLI2* expression in all the cell lines and tumor samples studied, which was unexpected since activation of the *BCL2* expression in response to Hh signaling occurs through the binding of *GLI2* to the promoter of *BCL2* [Regl et al., 2004]. There should a positive correlation between expression values of these two genes if the *BCL2* is activated in response to Hh signaling. There is also a study showing that *BCL2* expression can be induced by *GLI1* in mouse keratinocytes [Bigelow *et al.,* 2004]. However, there was no positive correlation between expressional levels of these two genes in our analysis. Hence we think that *BCL2* may not be regulated through Hh pathway but regulated by other pathways in the cancer formation in these cell lines and tumor tissues. Although there was no significant correlation, regulation of the *BCL2* promoter by *GLI* transcription factors should be checked with knock down and cotransfection studies to the cell lines where *BCL2* and transcription factor expressions were high. Bioinformatics analysis of the coexpressed genes based on their promoter sequences also may provide insight into the shared regulatory regions/modules by which Hh signaling is synchronized or sequentially regulated.

Hh signaling has a role in cell proliferation by directly activating the expression of genes that have a role in cell cycle control [Magliano *et al.,* 2003]. In the cell lines where pathway thought to be active expression of several of these target genes should be checked such as *MYC*, cyclin D, cyclin E, and cyclin B expressions to show the Hh pathway is responsible for the cancer progression through induction of cell proliferation. Cyclin D1 is downstream target of the *SHH* signaling in developing CNS [Lobjois *et al.,* 2004] although *SHH* expression does not have any effect on the expression profile of cyclin D1 in ductal development in virgin mice [Michno *et al.,* 2003]. Expression of *c-MYC, l-MYC*, and *n-MYC* have found to be decreased in medulloblastomas by cyclopamine treatment [Berman *et al.,* 2002]. The same study should be performed to the cell lines we studied with elevated expressions of the target genes.

Blocking the pathway with cyclopamine treatment is a way of determining the role of the Hh signaling in the cell lines studied. There are several inhibitors of the Hh pathway acting at different levels for the components of the pathway [Magliano *et al.,* 2003]. For example anti-SHH antibody blocks SHH and inhibits the cell proliferation *in vivo*. These inhibitors can be used to study the effect of Hh signaling in the epithelial cell lines in which pathway thought to be active such as MDA MB157, BT474, CAMA1 breast carcinoma cell lines. For HCC, SNU 182 and SNU 475 and for colon carcinoma TC71 and LOVO are best candidate cell lines suitable for this analysis. Inhibition of the Hh pathway by inhibitors would suppress the cell proliferation and cell invasiveness. Thus, cell invasion and apoptosis assays should be performed. LOVO is an aggressive cell line and it can be a good candidate cell line to identify the role of Hh pathway in the metastasis by performing cell migration assays.

For the primary breast tumor samples, there were three clusters formed in the hierarchical cluster analysis, which could not differentiated based on their known pathological properties. Determination of subtype of each sample used in this analysis may help in understanding of these clusters better. Since breast cancer is a heterogeneous disease each subtype has its own stable phenotype. Furthermore, the study should be performed in primary hepatocellular carcinoma and colon carcinoma samples to find out gene expression profile of Hh pathway in the real tumor environment for these carcinomas, which has been shown previously by in situ hybridization in liver cancer and by q-rt-RT PCR in colon cancer for some members of the pathway as shown in Table 1.5 in section 1.2.4 [Huang *et al.,* 2006; Manzo *et al.,* 2006].

Having an increased gene expression profile does not mean functional protein expression is found in most of the cases. Gene expression profiling should be supported by protein expression profiling. The best way of showing the protein expression profile is to perform western blot and/or immunohistochemistry studies. By showing protein expression, the role of Hh pathway can be identified more

clearly. Increased expression of *GLI1, PTCH1, SHH*, and *HIP* indicate the activation of the Hh pathway.

In conclusion, the Hh signaling pathway is activated in some epithelial carcinomas and that the components of the Hh pathway can be potential therapeutic targets, making it possible to develop new treatment strategies against epithelial cancers.

REFERENCES

Alon, U., Barkai, N., Notterman, D.A., Gish, K., Ybarra, S., Mack, D., and Levine, A.J.. Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by ologonucleotide arrays. Proceedings of the National Academy of Sciences 1999; 96: 6745-6750.

Bai, CB., Joyner, AL.. Gli1 can rescue in vivo function of Gli2. Development 2001; 128: 5161-5172.

Bale, A.E.. Hedgehog signaling and human disease. Annual Review of Genomics and Human Genetics 2003; 3: 47-65.

Barnes, E.A., Kong, M., Ollendorff, V., and Donoghue, D.J.. Patched1 interacts with cyclin B1 to regulate cell cycle progression. EMBO Journal 2001; 20: 2214-2223.

Beckmann, M.W., Niederacher, D., Schnurch, H.G., Gusterson, B.A., Bender, H.G.. Multistep carcinogenesis of breast cancer and tumor heterogeneity. Journal of Molecular Medicine 1997; 75: 429-439.

Bereta, J., Bereta, M.. Stimulation of glyceraldehyde-3- phosphate dehydrogenase mRNA levels by endogenous nitric oxide in cytokine-activated endothelium. Biochemistry and Biophysical Research Communinication 1995; 217: 363–369.

Berman, D.M., Karhadkar, S.S., Hallahan, A.R., Pritchard, J.I., Eberhart, C.G., Watkins, D.N., Chen, J.K., Cooper, M.K., Taipale, J., Olson, J.M., Beachy, P.A.. Medulloblastoma growth inhibition by hedgehog pathway blockade. Science 2002; 297:1559–61.

Bigelow, R.L., Chari, N.S., Unden, A.B., Spurgers, K.B., Lee, S., Roop, D.R., Toftgard, R., McDonnell, T.J.. Transcriptional regulation of Bcl-2 mediated by the sonic hedgehog signaling pathway through Gli-1. Journal of Biological Chemistry 2004; 279: 1197-205.

Bijlsma, M.F., Spek, C.A., and Peppelenbosch, M.P.. Hedgehog: an unusual signal transducer. BioEssays 2004; 26: 387-394.

Bitgood, M.J. and McMahon, A.P.. Hedgehog and BMP genes are coexpressed at many diverse sites of cell-cell interaction in the mouse embryo. Developmental Biology 1995; 172:126–138.

Blum, H.E.. Hepatocellular carcinoma: Therapy and preventation. World Journal of Gastroenterology 2005; 11: 7391-7400.

Bonifas, J.M., Pennypacker, S., Chuang, P.T., McMahon, A.P., Williams, M., Rosenthal, A., Sauvage, F.J., and Epstein Jr., E.H.. Activation of expression of hedgehog target genes in basal cell carcinomas. Journal of Investigate Dermatology 2001; 116: 739-742.

Bruix, J., Sala, M., and Llovet, J.M.. Focus on hepatocellular carcinoma. Cancer Cell 2004; 5: 215-219.

aBustin, S.A., and Nolan, T.. Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction. Journal of Biomolecular Techniques 2004; 15: 155-166.

bBustin, S.A.. A-Z of Quantitative PCR. La Jolla, CA: International University Line, 2004.

Chang, T.J., Juan, C.C., Yin, P.H., Chi, C.W., Tsay, H.J.. Up-regulation of beta-actin, cyclophilin and GAPDH in N1S1 rat hepatoma. Oncology and Reproduction 1998; 5: 469–471.

Chang-Claude, J., Dunning, A., Schnitzbauer, U., Galmbacher, P., Tee, L., Wjst, M., Chalmers, J., Zemzoum, I., Harbeck, N., Pharoah, P.D., Hahn, H.. The patched polymorphism Pro1315Leu (C3944T) may modulate the association between use of oral contraceptives and breast cancer risk. International Journal of Cancer 2003; 103:779-783.

Chen, X., Cheung, S.T., So, S., Fan, S.T., Barry, C., Higgins, J., Lai, K.M., Ji, J., Dudoit, S., Ng, I.O.L., van de Rijn, M., Botstein, D., and Brown, P.O.. Gene expression patterns in human liver cancers. Molecular Biology of the Cell 2002; 13: 1929-1939.

Chuang, P.T., and McMahon, A.P.. Vertebrate hedgehog signaling modulated by induction of a hedgehog-binding protein. Nature 1999; 397: 617-621.

Chuang, P.T., and Kornberg, T.B.. On the range of hedgehog signaling. Current Opinion in Genetics and Development 2000; 10: 515-522.

Cohen, M.M.. The Hedgehog signaling network. American Journal of Medical Genetics 2003; 123A: 5-28.

Cooper, M.K., Porter, J.A., Young, K.E., Beachy, P.A.. Teratogen-mediated inhibition of target tissue response to Shh signaling. Science 1998; 280:1603–1607.

Dahmane, N., Sanchez, P., Gitton, Y., Palma, V., Sun, T., Beyna, M., Weiner, H., and Ruizi i Altaba, A.. The sonic hedgehog-gli pathway regulates dorsal brain growth and tumorigenesis. Development 2001; 128: 5201-5212.

Douard, R., Moutereau, S., Pernet, P., Chimingqi, M., Allory, Y., Manivet, P., Conti, M., Vaubourdolle, M., Cugnenc, P.H., and Loric, S.. Sonic hedgehog-dependent proliferation in a series of patients with colorectal cancer. Surgery 2006; 139: 665-670.

Fronhoffs, S., Totzke, G., Stier, S., Wernert, N., Rothe, M., Bruning, T., Koch, B., Sachinidis, A., Vetter, H., and Ko, Y.. A method for the rapid contruction of cRNA standard curves in quantitative real-time reverse transcription polymerase chain reaction. Molecular and Cellular Probes 2002; 16: 99-110.

Futreal, P.A., Liu, Q., Shattuck-Eidens, D., Cochran, C., Harshman, K., Tavtigian, S., Bennett, L.M., Haugen-Strano, A., Swensen, J., Miki, Y., Eddington, K., McClure, M., Frye, C., Weaver-Feldhaus, J.D.W., Gholami, Z., Söderquist, P., Terry, L., Jhanwar, S., Berchuck, A., Iglehart, J.D., Marks, J., Ballinger, D.G., Barrett, J.C., Skolnick, M.H., Kamb, A., Wiseman, R.. BRCA1 mutations in primary breast cancer and ovarian carcinomas. Science 1994; 266: 120–122.

Gallego, M.I., Beachy, P.A., Hennighausen, L., Robinson, G.W.. Differential requirements for Shh in mammary tissue and hair follicle morphogenesis. Developmental Biology 2002; 249: 131-139.

Giles. R.H., van Es, J.H., and Clevers, H.. Caught up in a Wnt storm: Wnt signaling in cancer. Biochimica et Biophysica Acta 2003; 1653: 1-24.

Harfe, B.D., Scherz, P.J., Nissim, S., Tian, H., McMahon, A.P., and Tabin, C.J.. Evidence for an expansion-based temporal Shh gradient in specifying vertebrate digit identities. Cell 2004; 118: 517-528

Hatta, N., Hirano, T., Kimura, T., Hashimoto, K., Mehregan, D.R., Ansai, S., Takehara, K., and Takata, M.. Molecular diagnosis of basal cell carcinoma and other basaloid cell neoplasms of the skin by the quantification of Gli1 transcript level. Journal of Cutaneous Pathology 2005; 32: 131-136.

Huang, S., He, J., Zhang, X., Bian, Y., Yang, L., Xie, G., Zhang, K., Tang, W., Stelter A.A., Wang, Q., Zhang, H., and Xie, J.. Activation of the hedgehog pathway in human hepatocellular carcinoma. Carcinogenesis 2006; 27: 1334-1340.

Huggett, J., Dheda, K., Bustin, S., and Zumla, A.. Real-time RT-PCR normalisation; strategies and considerations. Genes and Immunity 2005; 6: 279-284.

Ingham, P.W., McMahon, A.P.. Hedgehog signaling in animal development: paradigms and principles. Genes and Development 2001; 15: 3059–3087.

Ishitani, R., Sunaga, K., Hirano, A., Saunders, P., Katsube, N., Chang, D.M.. Evidence that glyceraldehyde-3- phosphate dehydrogenase is involved in age-induced apoptosis in mature cerebellar neurons in culture. Journal of Neurochemistry 1999; 66: 928–935.

Ishiyama, T., Kano, J., Minami, Y., Iijima, T., Morishita, Y., and Noguchi, M.. Expression of HNFs and C/EBPα is correlated with immunocytochemical differentiation of cell lines derived from human hepatocellular carcinomas, hepatoblastomas and immortalized hepatocytes. Cancer Science 2003; 94: 757-763.

Kato, M., Seki, N., Sugano, S., Hashimoto, K., Masuha, Y., Muramatsu, M., Kaibuchi, K., and Nakafuku, M.. Identification of sonic hedgehog-responsive genes using cDNA microarray. Biochemical and Biophysical Research Communications 2001; 289: 472-478.

Kayed, H., Kleeff, J.,Keleg, J., Guo, J., Ketterer, K., Berberat, P.O., Giese,N., Esposito,I ., Giese, T., Büchler, M.W., and Friess, H.. Indian Hedgehog Signaling Pathway: Expression and Regulation in Pancreatic Cancer. International Journal of Cancer: 2004; 110: 668–676.

Kayed, H., Kleeff, J., Esposito, I., Giessa, T., Keleg, S., Giese, N., Buchler, M.W., and Friess, H.. Localization of the human Hedgehog-Interacting Protein (Hip) in the normal and Diseased Pancreas. Molecular Carcinogenesis 2005; 42: 183-192.

Kawahira, H., Ma, N.H., Tzanakakis, E.S., McMahon, A.P., Chuang, P.T., and Hebrok, M.. Combined activities of hedgehog signaling inhibitors regulate pancreas development. Development 2003; 130: 4871-4879.

Kenemans P., Verstraeten R.A., Verheijen R.H.M.. Oncogenic pathways in hereditary and sporadic breast cancer. The European Menopause Journal 2004; 49: 34-43.

Kubo, M., Nakamura, M., Tasaki, A., Yamanaka, N., Nakashima, H., Nomura, M., Kuroki, S., Katano, M.. Hedgehog signaling pathway is a new therapeutic target for patients with breast cancer. Cancer Research 2004; 64: 6071-6074.

Laurent-Puig, P., and Zucman-Rossi, J.. Genetics of hepatocellular tumors. Oncogene 2006; 25: 3778-3786.

Lerebours, F., and Lidereau, R.. Molecular alterations in sporadic breast cancer. Oncology Hematology 2002; 44: 121-141.

Lewis, M.T. and Veltmaat, J.M.. Next stop, the twilight zone: hedgehog network regulation of mammary gland development. Journal of Mammary Gland Biology and Neoplasia 2004; 9: 165-181.

Li, Y., and Rosen, J.M.. Stem/progenitor cells in mouse mammary gland development and breast cancer. Journal of Mammary Gland Biology and Neoplasia 2005; 10: 17-24.

Lipinski, R.J., Gipp, J.J., Zhang, J., Doles, J.D., and Bushman, W.. Unique and complimentary activities of Gli transcription factors in hedgehog signaling. Experimental Cell Research 2006; 312: 1925-1938.

Litingtung, Y., Lei, L., Westphal, H., and Chiang C.. Sonic hedgehog is essential to foregut development. Nature Genetics 1998; 20: 58-61.

Liu, S., Dontu, G., and Wicha, M.S.. Mammary stem cells, self-renewal pathways, and carcinogenesis. Breast Cancer Research 2005; 7: 86-95.

Lobjois, V., Benazeraf, B., Bertrand, N., Medevielle, F., and Pituello, F.. Specific regulation of cyclins D1 and D2 by FGF and Shh signaling coordinates cell cycle progression, patterning, and differentiation during early steps of spinal cord development. Developmental Biology 2004; 273: 195– 209.

Ma, Y., Erkner, A., Gong, R., Yao, S., Taipale, J., Basler, K., and Beachy, P.A.. Hedgehog-mediated patterning of the mammalian embryo requires transporter-like function of Dispatched. Cell 2002; 111: 63–75.

Ma, X., Chen, K., Huang, S., Zhang, X., Adegboyega, P.A., Evers, B.M., Zhang, H., and Xie, J.. Frequent activation of the hedgehog pathway in advanced gastric adenocarcinomas. Carcinogenesis 2005; 26: 1698-1705.

Magliano, M.P., and Hebrok, M.. Hedgehog signaling in cancer formation and maintenance. Nature Reviews Cancer 2003; 3: 903-911.

Mallon, E., Osin, P., Nasiri, N., Blain, I., Howard, B., and Gusterson, B.. The basic pathology of human breast cancer. Journal of Mammary Gland Biology and Neoplasia 2000; 5: 139-163.

Manzo, M., Moreno, I., Artells, R., Ibeas, R., Navarro, A., Moreno, J., Hernandez, R., Granell. M., and Pie, J.. Sonic hedgehog mRNA expression by real-time quantitative PCR in normal and tumor tissues from colorectal cancer patients. Cancer Letters 2006; 233: 117-123.

Medina, D.. Mammary developmental fate and breast cancer risk. Endocrine-Related Cancer 2005; 12: 483-495.

Michno, K., Boras, Granic, K., Mill, P., Hui, C.C., and Hamel, P.A.. Shh expression is required for embryonic hair follicle but not mammary gland development. Developmental Biology 2003; 264:153-65.

Miki, Y., Swensen, J., Shattuck-Eidens, D., Futreal, P.A., Harshman, K., Tavtigian, S., Liu, Q., Cochran, C., Bennett, L.M., Ding, W., Bell, R., Rosenthal, J., Hussey, C., Tran, T., McClure, M., Frye, C., Hattier, T., Phelps, R., Haugen-Strano, A., Katcher, H., Yakumo, K., Gholami, Z., Shaffer, D., Stone, S., Bayer, S., Wray, C., Bodgen, R., Dayananth, P., Ward, J., Tonin, P., Narod, S., Bristow, P.K., Norris, F.H., Helvering, L., Morrisson, P., Rosteck, P., Lai, M., Barrett, J.C., Lewis, C., Neuhausen, S., Cannon-Albright, L., Goldgar, D., Wiseman, R., Kamb, A., and Skolnick, M.H.. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. Science 1994; 266: 66–71.

Mocellin, S., Rossi, C.R., Pilati, P., Nitti, D., and Marincola, M.. Quantitative real-time PCR: a powerful ally in cancer research. Trends in Molecular Medicine 2003; 9: 189-195.

Notterman, D.A., Alon, U., Sierk, A.J., and Levine, A.J.. Transcriptional gene expression profiles of colorectal adenoma, adenocarcinoma, and normal tissue examined by oligonucleotide arrays. Cancer Research 2001; 61: 3124-3130.

Okada, T., Iizuka, N., Yamada-Okabe, H., Mori, N., Tamesa, T., Takemoto, N., Tangoku, A., Hamada, K., Nakayama, H., Miyamoto, T., Uchimura, S., Hamamoto, Y., and Oka, M.. Gene expression profile linked to p53 status in hepatitis C virus-related hepatocellular carcinoma. FEBS 2003; 555: 583-590.

Patil, M.A., Zhang, J., Ho, C., Cheung, S.T., Fan, S.T., and Chen, X.. Hedgehog signaling in human hepatocellular carcinoma. Cancer Biology and Therapy 2006; 5: 111-117.

Pfaffl, M.W.. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acid Research 2001; 29: 2002-2007.

Polyak, K.. Pregnancy and breast cancer: the other side of coin. Cancer Cell 2006; March: 151-153.

Pu, Y., Huang, L., and Prins, G.S.. Sonic hedgehog-patched Gli signaling in the developing rat prostate gland: lobe-specific suppression by neonatal estrogens reduces ductal growth and branching. Developmental Biology 2004; 273: 257-275.

Qualthruogh, D., Buda, A., Gaffield, W., Williams, A.C., and Pareskeva, C. Hedgehog signaling in colorectal tumor cells: Induction of apoptosis with cyclopamine treatment. International Journal of Cancer 2004; 110: 831-837.

Ramalho-Santos, M., Melton, D.A., and McMahon, A.P.. Hedgehog signals regulate multiple aspects of gastrointestinal development. Development 2000; 127: 2763-2772.

Regl, G., Kasper, M., Schnidar, H., Eichberger, T., Neill, G.W., Phillpott, M.P., Esterbauer, H., Hauser-Kronberger, C., Frischauf, A.M., and Aberger, F.. Activation of the BCL2 promoter in response to hedgehog/GLI signal transduction is predominantly mediated by GLI2. Cancer Research 2004; 64: 7724-7731.

Ririe, K.M., Rasmussen, R.P., and Wittwer, C.T.. Product differentiation by analysis of DNA melting curves during the polymerase chain reaction. Analytical Biochemistry 1997; 245: 154-160.

Ruiz I Altaba, A.. Gli proteins and hedgehog signaling: development and cancer. Trends in Genetics 1999; 15: 418-425.

Ruiz i Altaba, A., Stecca, B., and Sanchez, P.. Hedgehog-Gli signaling in brain tumors: stem cells and paradevelopmental programs in cancer. Cancer Letters 2004; 204: 145-157.

Salisbury, J.L.. The contribution of epigenetic changes to abnormal centrosomes and genomic instability in breast cancer. Journal of Mammary Gland Biology and Neoplasia 2001; 6: 203-212.

Scherf, U., Ross, D.T., Waltham, M., Smith, L.H., Lee, J.K., Tanabe, L., Kohn, K.W., Reinhold, W.C., Myers, T.G., Andrews, D.T., Scudiero, D.A., Eisen, M.B., Sausville, E.A., Pommier, Y., Botstein, D., Brown, P.O., and Weinstein, J.N.. A gene expression database for the molecular pharmacology of cancer. Nature Genetics 2000; 24: 236-244.

Shao, J., Zhang, L., Gao, J., Li, Z., and Chen, Z.. Aberrant expression of PTCH and Smo in human pancreatic cancerous tissues and its association with hyperglycemia. Pancreas 2006; 33: 38-44.

Shattuck-Eidens, D., McClure, M., Simard, J., Labrie, F., Narod, S., Couch, F., Hoskins, K., Weber, B., Castilla, L., Erdos, M., Brody, L., Friedman, L., Ostermeyer, E., Szabo, C., King, M.C., Jhanwar, S., Offit, K., Norton, L., Gilewski, T., Lubin, M., Osborne, M., Black, D., Boyd, M., Steel, M., Ingles, S., Haile, R., Lindblom, A., Olsson, H., Borg, A., Bishop, T., Solomon, E., Radice, P., Spatti, G., Gayther, S., Ponder, B., Warren, W., Stratton, M., Liu, Q., Fujimura, F., Lewis, C., Skolnick, M.H., and Goldgar, D.E.. A collaborative survey of 80 mutations in the BRCA1 breast and ovarian cancer susceptibility gene. JAMA 1995; 273: 535–541.

Sheng, T., Li, C., Zhang, X., Chi, S., He, N., Chen, K., McCormick, F., Gatalica, Z., and Xie, J., Activation of the hedgehog pathway in advanced prostate cancer. Molecular Cancer 2004; 3: 29.

Sicklick, J.K., Li, Y.X., Jayaroma, A., Kannangai, R., Qi, Y., Vivekanandan, P., Ludlow, J.W., Owzar, K., Chen, W., Torbenson, M.S., and Diehl, A.M.. Dysregulation of the hedgehog pathway in human hepatocellular carcinoma. Carcinogenesis 2006; 27: 748-757.

Singh, R., and Green, M.R.. Sequence-specific binding of transfer RNA by glyceraldehyde-3-phosphate dehydrogenase. Science 1993; 259: 365–368.

Sorlie, T., Tibshirani, R., Parker, J., Hastie, T., Marron, J.S., Nobel, A., Deng, S., Johnsen, H., Pesich, R., Geisler, S., Demeter, J., Perou, C.M., Lonning, P.E., Brown, P.O., Borresen-Dale, A.L., and Botstein, D.. Repeated observation of breast tumor subtypes in independent gene expression data sets. Proceedings of the National Acedemy of Sciences 2003; 100: 8418-8423.

Stanford, J.L., Szklo, M., and Brinton, L.A.. Estrogen receptors and breast cancer. Epidemiologic Reviews 1986; 8: 42-59.

Taipale, J., and Beachy, P.A.. The Hedgehog and Wnt signaling pathways in cancer. Nature 2001; 411:349 –54.

Thayer, S.P., Magliano, M.P., Heiser, P.W., Nielsen, C.M., Roberts, D.J., Lauwers, G.Y., Qi Y.P., Gysin, S., Castillo, C.F., Yajnik, V., Antoniu, B., McMahon, M., Warshaw, A.L., and Hebrok, M.. Hedgehog is an early and late mediator of pancreatic cancer tumorigenesis. Nature 2003; 425: 851-856.

Thellin, O., Zorzi, W., Lakaye, B., De Borman, B., Coumans, B., Hennen, G., Grisar, T., Igout, A., and Heinen, E.. Housekeeping genes as internal standards: use and limits. Journal of Biotechnology 1999; 75: 291-295.

Thorgeirsson, S.S., and Grisham, J.W.. Molecular pathogenesis of human hepatocellular carcinoma. Nature Genetics 2002; 31: 339-346.

Toftgard, R.. Hedgehog signaling in cancer. Cellular and Molecular Life Sciences 2000; 57: 1720-1731.

Valenti, M.T., Bertoldo, F., Dalle Carbonare, L., Azzarello, G., Zenari, S., Zanatta, M., Balducci, E., Vinante, O., and Lo Cascio, V.. The effect of biphosphates on gene

expression: GAPDH as a housekeeping or a new target gene? BMC Cancer 2006; 3:6:49.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A. and Speleman, F.. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biology 2002; 3: 0034.

Van't Veer, L.J., Dai, H., van de Vijver, M.J., He, Y.D., Hart, A.A.M., Mao, M., Peterse, H.L., van der Kooy, K., Marton, M.J., Witteven, A.T., Schreiber, G.J., Kerkhoven, R.M., Roberts, C., Linsley, P.S., Bernards, R., and Friend, S.H.. Gene expression profiling predicts clinical outcome of breast cancer. Nature 2002; 415: 530-536.

Vestergaard, J., Pedersen, M.W., Pedersen, N., Ensinger, C., Tumer, Z.,Tommerup, N., Poulsen, H.S., and Larsen, L.A.. Hedgehog signaling in small-cell lung cancer: frequent in vivo but a rare event in vitro. Lung Cancer 2006; 52: 281-290.

Vila, G., Theodoropoulou, M., Stalla, J., Tonn, J.C., Losa, M., Renner, U., Stalla, G.K., and Paez-Pereda, M.. Expression and function of sonic hedgehog pathway components in putiatary adenomas: evidence for a direct role in hormone secretion and cell proliferation. Journal of Clinical Endocrinology and Metabolism 2005; 90: 6687-6694.

Wang, W.S., Chen, P.M., and Su, Y.. Colorectal carcinoma: from tumorigenesis to treatment. Cellular and Molecular Life Sciences 2006; 63: 663-671.

Walterhouse, D.O., Lamm, M.L.G., Villavicencio, E., and Iannaccone, P.M.. Emerging roles for Hedgehog-Patched-Gli signal transduction in reproduction. Biology of Reproduction 2003; 69: 8-14.

Watkins, D.N., and Peacock, C.D.. Hedgehog signalling in foregut malignancy. Biochemical Pharmacology 2004; 68: 1055–1060.

Wetmore, C.. Sonic hedgehog in normal and neoplastic proliferation: insight gained from human tumors and animal models. Current Opinion in Genetics and Development 2003; 13: 34-42.

Wicking, C., and McGlinn, E.. The role of hedgehog signaling in tumorigenesis. Cancer Letters 2001; 173: 1-7.

Woodward, W.A., Chen, M.S., Behbod, F., and Rosen J.M.. On mammary stem cells. Journal of Cell Science 2005; 118: 3585-3594.

Wong, M.L., and Medrano, J.F.. Real-time PCR for mRNA quantitation. Biotechniques 2005; 39:1.

Wooster, R., Neuhausen, S.L., Mangion, J., Quirk, Y., Ford, D., Collins, N., Nguyen, K., Seal, S., Tran, T., Averill, D., Fields, P., Marshall, G., Narod, S., Lenoir, G.M., Lynch, H., Feunteun, J., Devilee, P., Cornelisse, C.J., Menko, F.H., Daly, P.A., Ormiston, W., McManus, R., Pye, C., Lewis, C.M., Cannon-Albright, L.A., Peto, J., Ponder, B.A.J., Skolnick, M.H., Easton, D.F., Goldgar, D.E., and Stratton, M.R.. Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13. Science 1994; 265: 2088–2090.

Wooster, R., Bignell, G., Lancaster, J., Swift, S., Seal, S., Mangion, P., Collins, N., Gregory, S., Gumbs, C., Micklem, G., Barfoot, R., Hamoudi, R., Pantel, S., Rice, C., Biggs, P., Hashim, Y., Smith, A., Connor, F., Arason, A., Gudmundsson, J., Ficenec, D., Kelsell, D., Ford, D., Tonin, P., Bishop, D.T., Spurr, N.K., Ponder, B.A.J., Eeles, R., Peto, J., Devilee, P., Cornelisse, C., Lynch, H., Narod, S., Lenoir, G., Egilsson, V., Barkadottir, R.B., Easton, D.F., Bentley, D.R., Futreal, P.A., Ashworth, A., and Stratton, M.R.. Identification of the breast cancer gene BRCA2. Nature 1995; 378: 789–792.

Xie, J, Johnson, R.L., Zhang, X., Bare, J.W., Waldman, F.M., Cogen, P.H., Menon, A.G., Warren R.S., Chen L.C., Scott, M.P., and Epstein E.H.Jr.. Mutations of the PATCHED gene in several types of sporadic extracutaneous tumors. Cancer Research 1997; 57: 2369-2372.

Zhao, H., Langerod, A., Ji, Y., Nowels, K.W., Nesland, J.M., Tibshirani, R., Bukholm, I.K., Karesen, R., Botstein, D., Borresen-Dale, A.L., and Jeffrey, S.S.. Different gene expression patterns in invasive lobular and ductal carcinomas of the breast. Molecular Biology of the Cell 2004; 15: 2523-2536.

Zhu, Y., James, R.M., Peter, A., Lomas, C., Cheung, F., Harrison, D.J., and Bader, S.A.. Functional *Smoothened* is required for expression of *GLI3* in colorectal carcinoma cells. Cancer Letters 2004; 207: 205-214.

http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi

http://rana.lbl.gov

http://www.appliedbiosystems.com

http://www.biol.ttu.edu/Strauss/Matlab/matlab.htm

http://www.cancer.org

http://www.cancer.gov

http://www.genomebiology.com/2002/3/7/research/0034

http://www.mdl.dk

http://www.stat.psu.edu/~chiaro/BioinfoII/mds_sph.pdf

APPENDICES

APPENDIX A

Quantitative Real Time RT-PCR Ct Values for Breast Carcinoma Cell Lines

Cell Line	GLI1	GLI2	GLI3	SHH	IHH	PTCH1	SMO	BCL2	GAPDH-1	GAPDH-2
MDA157	27.4-27.4	25.9-26.7	20.3-22	29.4-30.8	28.3-27.7	20.3-20.8	27.1-26.9	25-25.3	14.9-14.2	15.4-15.6
MDA231	27.1-27.2	26.5-26.6	25.2-25.8	29.3-29.8	30.2-29.9	26.4-25.7	43-42.4	25.1-25.1	14.3-15.2	13.2-14.3
MDA361	25.6-25.4	32.8-32.5	23.1-23	26.3-26.4	30.4-30.8	22.4-22.1	24.9-24.4	24.2-23.5	12.4-12	13-12.7
MDA453	25.2-25.1	NA	21.1-21.4	30.8-30.7	30.4-30.8	24.5-24.8	39-39.7	26.7-26.4	14-14.8	13.5-14
MDA468	26.6-26	32-32	24.1-24	30.3-30.9	31.3-31.3	23.8-23.8	26-26.5	25.1-25.2	14.7-15	15.1-15.3
BT474	25.7-25.9	38.9-39.3	27.7-26.5	30.6-31.4	27.2-27.5	29.8-30.4	32.3-33.5	24.6-24.9	16.5-16.7	17.2-17.3
BT20	28.5-28.5	34.3-33.1	25.4-25.6	30.3-30.8	30-29	24.2-23.4	34.5-34.8	28.1-27.4	14.6-14.8	14.6-14.7
MCF7	25.5-24.7	36.2-37.5	22.4-22.2	30.1-29.7	28.7-29	23.2-23	44.4-39.7	19.9-19.7	14.5-13.4	15-14.4
MCF12A	27.1-26.8	29-28.8	22.8-23	30.9-30.1	35.3-34.1	25.8-25.2	29.7-29.3	27.6-27.8	14.5-13.8	14.7-15.2
HCC1937	27-26.8	23.9-24.4	22.9-22.1	29.6-31	27.1-27.2	21.3-22.7	37.7-37.4	25.1-24.9	12.6-12.5	13.7-13.6
CAMA1	25.7-25.1	34.2-34.3	23.4-23.1	31.3-31.9	28-27.7	30.6-29.2	34.7-33.9	24.3-24.6	16.6-16.3	17.2-17.4
SKBR3	27.6-28.3	33.5-32.9	27.2-28	30.4-30.3	31-31.7	27.8-27	48-48.4	31.3-30.3	14.7-14.1	12.9-12.5
ZR75.1	28.8-27.7	34.1-34	25.7-26	31.1-30	29.3-30	26.8-26.8	33-34.7	23.5-23.5	14.2-14.3	12.6-13
T47D	26.7-26.2	NA	23-22.3	31.2-30	32.7-32.9	27.4-28	32.7-32.4	28.9-29.8	16.1-15.9	17-16.9
HBL100	28.2-28.2	26.8-26.8	29.2-28.9	30.2-30.4	30-30.2	29.4-30.1	26.3-26.6	28.8-29.6	17.7-16.8	13.1-12.3
HME1	25.1-24.8	30.9-30.9	25.2-25.1	29.9-30.2	31.3-31.9	29.3-28.9	31.7-31.3	27.9-28.4	14.4-14.2	14.2-13.6

NA: no amplification

APPENDIX B

Quantitative Real Time RT-PCR Ct Values for Hepatocellular Carcinoma Cell Lines

Cell Line	GLI1	GLI2	GLI3	SHH	IHH	PTCH1	SMO	BCL2	GAPDH
SNU182	25.3-25.2	22.6-22.1	21.2-21.9	28.7-29.3	NA	21.1-22	25.8-26.9	28.1-28.4	16.7-16.9
SNU387	26.7-26.2	24.7-25.4	25.9-25.2	35.2-34	NA	25.9-26.2	NA	NA	19.3-19.4
SNU398	20-22.8	26.5-26.7	21.7-21.7	33.1-35.2	NA	20.6-20.2	NA	NA	14.2-14.2
SNU423	24.8-24.9	26.8-26.1	22.7-22.5	27.1-27.2	29-29.8	24.3-23.7	25.7-25.3	27.5-27.3	15.6-15.3
SNU449	22.6-22.2	26.3-25.7	23.4-23.1	32.3-32.4	27.6-27.6	23.8-23.2	24.8-25.4	26.3-26.3	13.5-13.9
SNU475	24.7-24.3	30.1-29.7	47.7-44.9	29.9-30.2	26-25.9	27.5-27.3	32.8-32.9	24.8-25.8	19.8-20.2
HEP3B	26-25.7	32.7-33.6	37.1-36.1	28.2-27.6	26.6-26.6	23.4-23.5	25.4-26	32.2-33.3	14.2-14.3
HEP3B TR	25-24.9	32.8-31.3	36.4-38.5	28-27.8	24.2-23.7	24.6-24.1	24.7-24.4	32.3-31.6	14.414.5
HEP40	26.9-26.9	33.8-34.1	25.9-25.9	NA	20.8-20.3	23.7-22.7	26.3-25.4	28.9-29.2	14.4-14.3
HEPG2	27.7-27.7	34.7-33.9	31.7-32.9	31.7-33.1	25.4-25.9	25.3-25.3	28-27	31-30.6	15.3-15.4
SKHEP1	26.6-27.1	26.4-26.1	24.5-23.6	NA	29.5-28.9	24-24	26.4-25.5	25.6-24.8	13.9-13.1
SAOS2	22.6-22.6	21.1-21.2	21.6-21.4	46.5-45.4	22.6-23	18.3-18.4	25.9-25.7	25.7-25.7	14.3-14.1
HUH7	24.6-24.9	32.3-32.6	35-35	24.6-24.6	19.6-19.6	21.2-21.6	23.3-24.1	32.7-33.7	14.4-14.7
PLC	22.5-23	33.5-34.1	34-34	NA	24.7-24.7	24.7-25.1	27-26.7	36.1-36.5	13.9-14
FOCUS	23.1-23	26-25.8	23.1-23.6	36.9-36.9	31.1-31.3	23.1-22.3	43.3-46.7	22.8-23.6	14.8-14.3
MAHLAVU	25.4-24.8	26.4-26.2	23.9-23.9	NA-43.9	32.5-32.6	23.2-23.5	47.5-48	25.9-26.1	14.4-14.1

NA: no amplification

APPENDIX C

Quantitative Real Time RT-PCR Ct Values for Colon Carcinoma Cell Lines

Cell Line	GLI1	GLI2	GLI3	SHH	IHH	PTCH1	SMO	BCL2	GAPDH-1	GAPDH-2
SW 620	28.9-29.3	36-37.7	31-30.6	29.4-29.8	25.9-26.3	23.6-23.4	33.3-33	28.1-29.7	14.8-15.1	12.9-12.7
SW 480	27.3-26.8	26.8-26.7	22.8-23.2	30-29.8	29.1-30	23.7-22.5	29.2-29.6	31.6-31.5	13.5-13.6	13.8-13.3
SW 48	27.2-27.5	25.7-25.6	21-20.5	29.1-29.4	32.4-31.6	22.2-22	NA	29.7-29.9	15.6-14.8	12.5-12.6
HCT8	27.6-27.2	33.4-33.2	NA	32.5-30.7	31.4-32	24.4-23.3	NA-45.7	26.2-26.3	15.3-14.8	15.7-16
HCT15	30.5-29.4	27.6-27.7	NA	27.5-26.3	30.8-29.8	23.3-23.7	37.9-37.7	27.3-27.7	16.8-16.9	13.6-14
HT29	23.9-23.8	31.8-33.1	NA	25.4-25.8	29.1-28.4	25.2-24.9	NA	32.3-32	16.6-15.8	14.2-13.2
TC 71	21.9-21.7	27.9-28.8	24.2-23.8	25.8-26.2	30.7-29.9	25.2-25	34.3-34.3	25.8-26.3	14.3-13	11.4-11
T7	25.8-25.7	33.3-32.8	39.5-38.5	31-29.3	23.6-23.4	25.2-24.4	37.3-36	28.5-28.1	15.8-15.5	13.3-13.1
LOVO	24.1-24.2	28.3-27.3	27.9-27.3	25.8-26	23.5-23.5	23.5-23	26.8-26.6	25.5-24.9	14.4-14.3	12.1-12.2
LS411	24.5-24.6	29.6-30.3	47.2-33.8	29.2-29.3	25.3-25.8	22.7-23	36.7-37.8	27.9-27.4	14.4-14.1	12.9-12.7
COLO205	30.2-29.5	40.2-43.1	39-39.1	33.6-34.7	33.3-33.9	23.2-22.9	39.4-38	32.1-31.2	15.6-15.1	13-12.7
CL115	26.3-26.5	26.5-26.6	39.8-40	30.3-31.1	29.4-29.8	24.4-24.9	35.4-35.5	29.6-30.5	13.8-13.2	11.6-11.8
W 05	24.1-24.1	30.1-30.6	36.6-36.6	24.7-24.4	27.6-27.7	22.8-22.6	37.5-37.9	30.8-30.9	14.9-14.7	12.8-13.1
KM 12	27.7-27.4	28.9-27.9	NA	29.8-29	24.8-24.9	23.9-23.8	NA	28.5-28.1	14.2-14.1	13.2-13

NA: no amplification

APPENDIX D

Quantitative Real Time RT-PCR Ct Values for Primary Breast Tumor Samples

Tumor Sample	GLI1	GLI2	GLI3	SHH	IHH	PTCH1	SMO	BCL2	ACTB	TBP	SDHA
MFT14	25.1-25.7	32.3-33.6	28.2-27.4	NA	27.6-27.7	29.9-29.8	33.6-34.1	22.7-23.2	21.4-22.7	20.7-20.1	23.5-24
MFT16	26.6-26.5	29.2-28.8	24.5-24.5	NA	30.8-30.4	27.9-28.5	27.2-26.4	21.7-21.5	16.1-16	21.1-21.1	19.2-19.1
MFT21	27.1-28	30-29.3	23.1-23.3	35.1-45	28.9-29.2	27.4-27.7	27.3-27.6	25-26.4	14-13.7	19.7-20.2	19.4-20.1
MFT25	27.5-27.5	45.5-35	28.7-27.4	NA	29.6-33.8	30.9-31.5	36.1-37.9	24-23.5	26.4-26	23.9-23.8	27.4-27.7
MFT40	23.8-23.5	34.3-47.6	25.5-24.8	35.7	33.1-43.3	28.6-28.4	26-25.8	26.2-25.7	15.6-14.2	20.2-19.9	19.3-18.9
MFT41	28.3-28.7	34.2-35.2	28.6-29.4	NA	33.4-32.9	31.5-32.2	38-38.8	26-25.5	15.7-15.4	22.2-22.4	22.7-22.8
MFT49	25.3-25.2	34.7-33.8	28.6-28.3	NA	27.1-27.1	29.4-29.6	33.9-34.1	22.5-21.9	16.1-16	20.7-20.8	22.9-23.6
MFT78	28.2-28.1	31.4-31.1	22.4-22.8	NA	36.9-34.8	31.1-31.2	29-29	25.7-24.7	17.1-15.7	22.7-22.4	20.8-20.3
MFT79	27.2-26.8	NA	29.6-29	NA	32.2-34.8	32.5-32.6	42.5-40	23.6-23.8	25.3-25.5	23-22.9	26.8-27.4
MFT90	28.2-28.7	34-32.3	27.2-27.4	NA	35-33.2	31.4-32.9	38.4-36.7	24.7-24.1	21.6-20.5	21.7-21.8	23.2-22.8
MFT94	26.9-27	27.8-28.4	26-25.8	36	30.2-30.6	25.7-26	26.8-26.8	27-26.9	14.8-15.6	21.2-21.5	20.3-19.8
MFT95	26.8-26.8	44.50	26.5-26.9	NA	31.5-31.4	31.9-32.3	39.7-43.6	23.5-23.6	20.2-19.9	23.6-23.8	26.7-26.5
MFT97	28.8-29	35.6-36	27.5-27.6	NA	32.6-31	33.2	37.1-38.2	27.9-28	24.4-23.7	25.1-25	27.7-28.1
MFT113	27.2-26.9	32.3-31.1	28-28	NA	27.6-27.7	30.1-29.8	35.7-34.4	27.3-27	16.8-16.6	23.2-23.5	26.8-26.3
MFT115	26.9-26.6	31-30.7	25-24.8	NA	27.7-27.1	26.1-26.1	26-26.1	24.1-23.3	17-15.2	20.8-20.7	20.2-20.1
MFT116	30.9-30.3	35.3-34.9	29.4-28.9	NA	31.8-32.1	31.2-30.7	37.9-40.4	27.4-28.1	18.9-19	25-25.2	27.1-27
MFT117	29.9-30	35.5	30.1-29.2	NA	33.2-32.8	34.3	38.2-37.9	31.6-29.9	24.7-23.5	26.6-26.8	27.4-27.3
MFT120	29.1-28.1	33.1-32.7	26.7-27.3	NA	28.5-28.8	29.2-28.6	34.1-33.7	28.4-28.3	18.2-17.9	23.9-24.1	26-25.5
MFT124	30.1-29.1	31.6-29.2	25.7-25.9	NA	32-32.9	26.3-25.8	31.7-31.2	26.4-25.8	17.5-17.3	22.4-22.5	24.8-24.5
MFT125	27.5-27.4	29.2-29.4	24.8-24.6	32.8-32.7	27.8-27.7	27.7-27.6	28-27.9	24.1-24.2	14.6-14.6	21.3-21.2	22.9-21.9
MFT127	28.5-29	NA	28.2-27.9	33.6-35.1	33.2-33.7	32.9-32.5	43.9-36.2	28.1-28.5	24.7-24.9	24.8-24.6	27.1-27.6
MFT129	27.3-27.5	28.7-28.8	25.2-25.8	NA	33.6-34.8	27.7-28	28-27.9	25.3-25.1	15-15.2	22.1-21.9	23-23.3
MFT131	29.3-28.8	34.2-39.1	29.4-26.8	33.7-34.6	37.3-37.9	35.3-35.7	NA	28.3-27.6	24-24.2	24.1-23.9	26.1-25.3
MFT132	29.4-29.8	30.5-30.5	27.4-27.3	34.8	34.8-33.5	28.5-28	32.5-31.6	26.8-26.8	17.6-16.8	23.2-23.7	22.3-22.2

MFT149	31.2-28.7	28.9-27.7	22.8-24.1	NA	32.5-32.8	27-26.8	28.1-28.1	22.6-22	16.5-15.1	20.8-20.6	21.3-20.7
MFT154	NA	31.2-30.9	23.3-23.9	48.6	34.2-34.5	28.3-28.2	32.8-32.2	NA	NA	24.1-24	28-27.3
MFT155	27.9-28	28.4-28.3	21.8-22.1	NA	33-31.6	25.1-25.4	28-27.8	22.4-22.3	16.7-16.3	21.8-21.7	21-20.9
MFT173	27.8-27.6	29.7-29.2	26.8-25.8	NA	32.6-32.1	28.7-28.7	27.6-26.8	24.8-23.9	16.7-16.9	22.7-22.1	21.1-20.7
MFT174	28.5-28.6	48.1-44.9	27.5-28	NA	50-50	34.8	44.8	29.5-30	27.5-27.3	25-25.2	26.8-27.5
MFN P1-1	28.2-27.4	31.9-30.2	26.7-26.3	NA	28.1-27.4	27.1-26.1	29-28.2	24.9-24	16.3-15.9	21.9-21.4	22.1-21.1
MFNP1	27.2-27.3	32.9-32.5	26.1-25.2	NA	28.4-28	27.6-27.2	29.7-29.7	24-23.7	18.2-17	22.7-22.8	22-21.8
MFN P2-1	27.8-27.3	31.3-32.1	27.9-28.5	NA	28.6-28.4	29.4-28.4	29.1-29.4	26.4-26.5	15.1-15.2	21.9-21.3	23-20.6
MFNP2	27.3-27.4	34-34.4	27.7-27.8	NA	31.8-29.1	31.7-30.3	32.7-32.6	27.1-26.6	17.8-16.3	23.4-22.7	22.9-22
MFN P3-1	28.1-28.2	29.2-28.9	24.9-25	32.6-30.9	28.9-29	26.3-26.3	25.8-26.5	23.2-23.7	15.8-15.8	21.3-21.4	22-22.5
MFNP3	27.6-27.5	30.7-29.6	25.6-25.5	NA	30.2-29.3	27.3-27.5	29.5-29.8	24-24.3	16.7-18.2	22.5-22.2	21.6-21.6

NA: no amplification