International Journal of Basic and Applied Biology
Print ISSN: 2394-5820, Online ISSN: 2349-2539, Volume 2, Number 6; July-September, 2015, pp. 394-398

http://www.krishisanskriti.org/ijbab.html

Study of Differentially Expressed Genes in Cervical Cancer

Shalaka Joshi[1], Sneha Patil[2] and Riya Rane[3]

[1,2,3]School of Biotechnology & Bioinformatics, D. Y. Patil University Belapur, Navi Mumbai
E-mail: [1]shalakajoshi25@gmail.com, [2]snp2312@gmail.com, [3]riya.rane9@gmail.com

Abstract—*Cervical cancer is the fourth most common cancer in women, and the seventh overall, with an estimated 528,000 new cases and 266,000 deaths from cervical cancer worldwide in 2012. The causative agent of cervical cancer is Human Papillomavirus (HPV) and HPV 16 and 18 are the main causative types .There are reports dedicated towards the characterization of the genetic contribution to the disease variance, with many modifiers gene for cervical cancer. In this study we have undertaken a normalization approach to identify the differential regulation of genes in cervical cancer by analyzing cervical cancer microarray data sets from GEO database. EPHA3 gene belongs to the ephrin receptor subfamily of the protein-tyrosine kinase family and its expression has been reported in many types of cancers. We investigated the role of EPHA3 gene; found to be 3-fold down regulated, in cervical cancer. Our results suggest that the lower levels of EPHA3 may lead to RAS activation, causing uncontrolled differentiation and proliferation, thus playing a key role in metastasis.*

1. INTRODUCTION

Cervical cancer is the fourth most common cancer in women, and the seventh overall, with an estimated 528,000 new cases and 266,000 deaths from cervical cancer worldwide in 2012 [1]. The causative agent of cervical cancer is Human Papillomavirus (HPV) and HPV 16 and 18 are the main causative types [2, 3, 4-6].

The mechanisms by which HPV integrates its DNA into the human genome are not fully understood although there are a number of hypotheses including DNA instability, and transcriptional regulation of integrants [7]. Cervical carcinomas are associated with certain specific types of HPV, particularly type 16, 18, 33 and 42.

The association of HPV with cervical cancer has provided the background and the scientific justification for improving screening programs and for developing HPV vaccines [8]. Two broad classification of HPV vaccines are known; a. Prophylatic HPV vaccine and b. Therapeutic HPV vaccines. DNA-free virus-like particles (VLP) synthesized by the self-assembled viral particles of the main structural HPV proteins, L1 protein (or L1 and L2 protein), induce strong humoral responses from neutralizing antibodies and are thus the best candidate immunogens currently available for HPV vaccine trials.

Comparative genomic hybridization supports the evidence that *PIK3CA* (located in 3q26.3) is an oncogene in cervical cancer and that its amplification may be linked to cervical tumorigenesis [9]. *AGRN, CCL18, FERMT1, ILB, MELK, NUP210,* etc. are some of the genes reported to be up regulated in cervical cancers, while, *CRNN, HOPX, DALP1, HEBP2, KRTDAP,* etc. are some of the genes reported to be down regulated in cervical cancers [10]. For this, we have studied the regulation of cervical cancer its physiological functions, and the differentially expressed genes in cervical cancer.

The studies observed the down regulation of EPHA3 gene is may be responsible for variability in frequency of infections.

With intent to treat cervical cancer patients in a better manner, the present study aims to understand the upregulating and downregulating genes during the disease progression. The study also addresses the possibility of genes other than EPHA3 which may be responsible for variability in frequency of infections.

2. MATERIALS AND METHODOLOGY

2.1 Gene datasets

The Gene Expression Omnibus (GEO) is a public repository that archives and freely distributes microarray, next-generation sequencing, and other forms of high- throughput functional genomic data submitted by the scientific community. Two gene datasets were obtained from GEO viz. GSE29216 and GSE29570. This database stores curated gene expression Datasets, as well as original series and platform records in the Gene Expression Omnibus (GEO) repository. Dataset records contain additional resources including cluster tools and differential expression queries.

2.2 DNA microarray

DNA microarray known as gene chips, bio- chips or silicon-chips, are solid support usually made up of glass or silicon,

holding DNA that represent thousands of genes that act as probes for mRNA.

2.3 Analyzing the datasets

R is a free software environment for statistical computing and graphics. It comprises and runs on a wide variety of UNIX platforms, Windows and MacOS. R provides a wide variety of statistical and graphical techniques, including linear and nonlinear modelling, classical statistical tests, time-series analysis, classification, clustering, and others. It is an environment within which statistical techniques were implemented. Bioconductor is a free, open source and open development software project for the analysis and comprehension of genomic data generated by wet lab experiments in molecular biology. Bioconductor is based primarily on the statistical R programming.

MATLAB is high level language and interactive environment for numerical computation visualization, and programming. Using MATLAB, we can analyze data, develop algorithms, and create models and applications. Normalization was done for resolving the systematic errors and bias introduced by the microarray experimental platform. Normalization is useful for a number of situations including within-slide comparison, multiple-slide comparison, and paired-slide comparison for dye exchange experiments. Normalization removes unwanted systematic variability in hybridization and to adjust the spatial effects from microarray data by adjusting the intensities of hybridization to balance them appropriately to make decisive analysis. It identifies the differences introduced in labelling or detection efficiencies when different fluorescent dyes are used. It also adjusts to bias introduced because of scanner settings. The microarray data generated by the feature extraction software is typically in the form of one or more text files.

Multi Experiment Viewer (MeV) is an application that allows the user to view processed microarray slide representations and identifies genes and expression patterns of interest. This Crude Analysis was used to perform the back validation of genes which were obtained by using SAM test and T-test. In this, the selection of upregulated and downregulated genes was done using MS Excel.

3. RESULTS

3.1 Box plot analysis

Box plot analysis is method for visualizing the distribution of the data values throughout the dataset providing information about the spread and skewness in the dataset. Fig. 3.1 and Fig. 3.2 represent the box plot analysis of the gene datasets GSE29216 and GSE29570 respectively as before and after normalization.

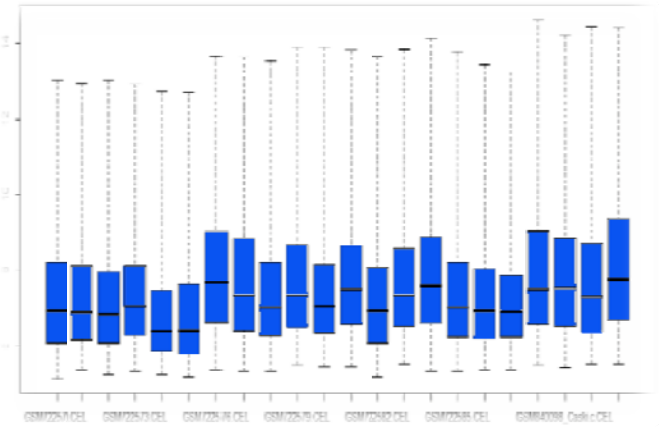

Fig. 3.1: Box plot of Dataset GSE29216 before normalization

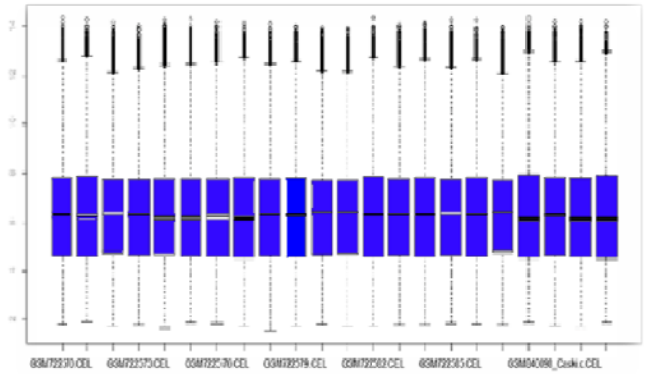

Fig. 3.2: Box plot of Dataset GSE29216 after normalization

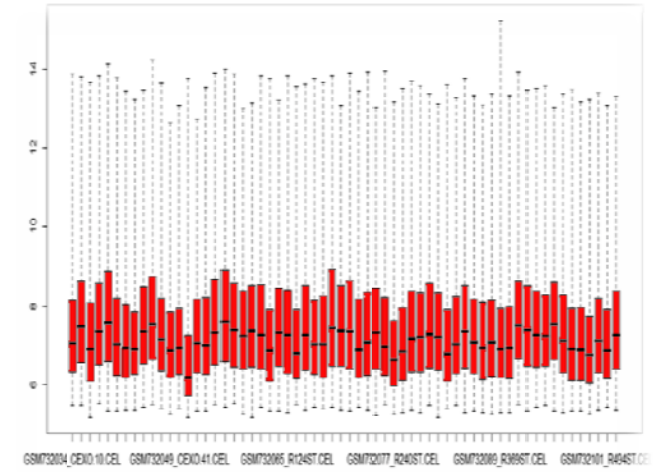

Fig. 3.3: Box plot of Dataset GSE29750 before normalization

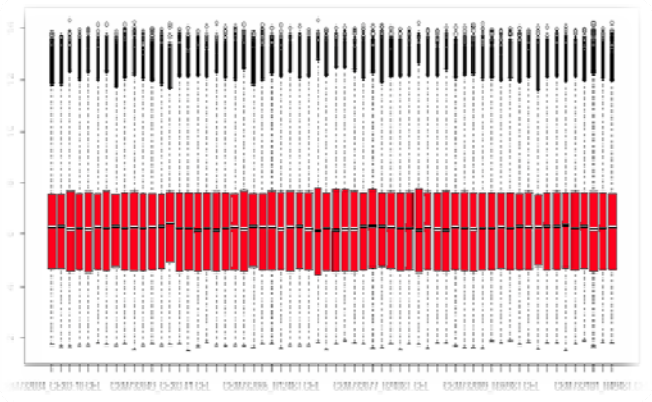

Fig. 3.4: Box plot of Dataset GSE29750 after normalization

3.2 Histogram analysis

A histogram is a representation of a statistical graph of a frequency distribution in which vertical rectangles of different heights are proportionate to corresponding frequencies. Fig. 3.3 and Fig. 3.4 represent the histogram curve of the datasets each before and after normalization. The microarray data shows histogram as frequency on y-axis and log transform of pixel intensities along the x-axis.

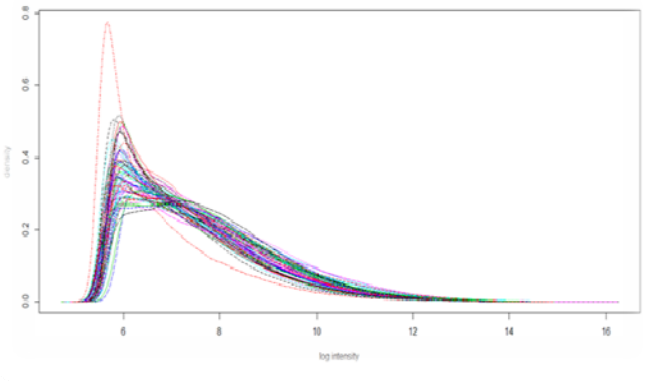

Fig. 3.5: Histogram of GSE29216) before normalization

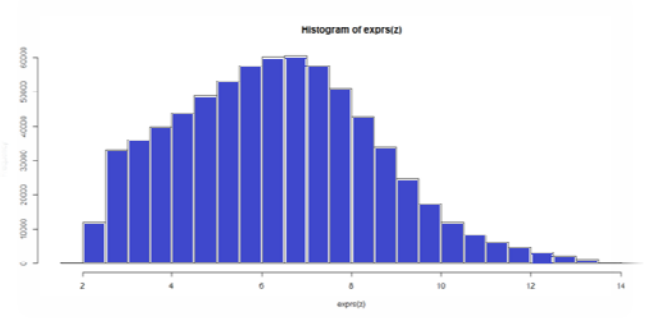

Fig. 3.6: Histogram of GSE29216 after normalization

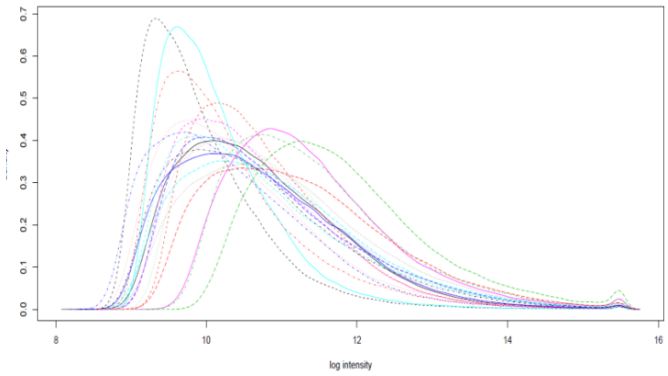

Fig. 3.7: Histogram of GSE29570 before normalization

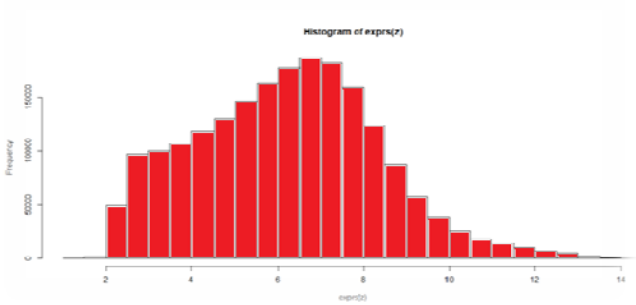

Fig. 3.8: Histogram of GSE29570 after normalization

3.3 Sam Results & T-test analysis

The SAM test analysis [**Fig.3.9**] and *t-test* Analysis **[Fig.3.10]** of the properly normalized each data set indicated that there are some up-regulated as well as down-regulated genes in the data sets. A list of up-regulated and down-regulated genes which were common in all the data sets was prepared. [Table-1, 2]

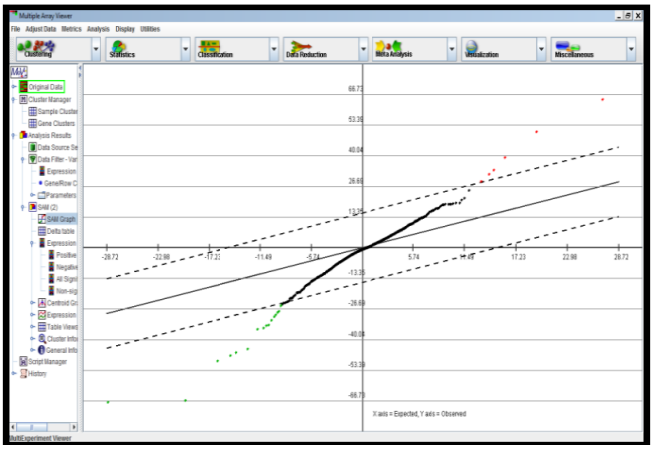

Fig. 3.9: SAM graph showing upregulated genes in red of HG-U133-A chip: GSE29216 and GSE29570

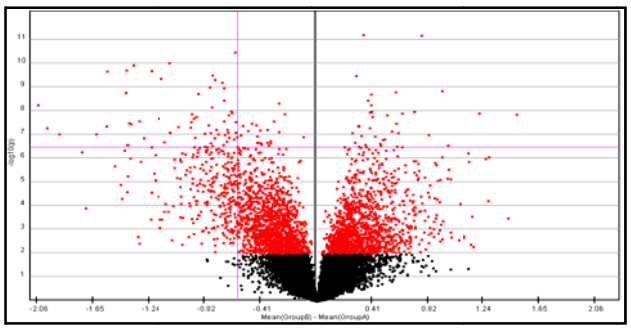

Fig. 3.10: Volcano plot for A Chip data: GSE29216 and GSE29570

Table 1: Up regulated genes in Cervical Cancer

Sr. No	Up-regulated Probe sets	Gene Description	Gene Name
1	7899192	Ribosomal protein S6 kinase, 90kDa, polypeptide 1	RPS6KA1
2	7901192	KAD54-like (S. cerevisiae)	KAD54L
3	7912956	Regulator of chromosome condensation 2	RCC2
4	7913852	Chromosome 1 open reading frame 155	C1orf155
5	7924733	Poly (ADP-ribose) polymerase 1	PARP1
6	7953351	Non-SMC condensin I complex, subunit D2	NCAPD2
7	7959052	Replication factor C (activator 1) 3, 36 5kDa	RFC3
8	7982294	Similar to nest domain and KLD 2	LOC39505 61
9	7992775	Progestin and adipoQ receptor family member IV	PAQR4
10	8002987	Chromosome 10 open reading frame 61	C10orf61
11	8003844	Germ cell associated 2 (haspin)	GSG2
12	8007509	Chromosome 17 open reading frame 55	C17orf55
13	8008510	Essential meiotic endonuclease 1 homolog 1 (S. pombe)	ENM1
14	8010770	Solute carrier family 16, member 3 (monocarboxylic acid transporter 4)	SLC16A3
15	8014454	Myosin XIX	MYO19
16	8015042	PSMG3 interacting protein	PSMG3IP
17	8024845	Chromatin assembly factor 1, subunit A (p150)	CHAF1A
18	8030429	BCL2-like 12 (proline rich)	BCL2L12
19	8033912	DNA (cytosine-5-)-methyltransferase 1	DNMT1
20	8035256	HAUS augmin-like complex, subunit 8	HAUS8
21	8037855	Solute carrier family 1 (neutral amino acid transporter), member 5	SLC1A5
22	8040578	Centromere protein O	CENPO
23	8040712	Centromere protein A	CENPA
24	8041888	MutS homolog 6 (E. coli)	MSH6
25	8047187	Coiled-coil domain containing 150	CCDC150
26	8062506	KAE1 NA export 1 homolog (S. pombe)	KAE1
27	8069953	Chromosome 21 open reading frame 45	C21orf45
28	8070129	Downstream neighbor of SON	DONSON
29	8071552	KAN binding protein 1	KANBP1
30	8073858	G-2 and S-phase expressed 1	GTSE1
31	8082504	Chromosome 3 open reading frame 57	C3orf57
32	8083228	Guanine monophosphate synthetase	GMPS
33	8084423	Proteasome (prosome, macropain) 26S subunit, non-ATPase, 2	PSMD2
34	8084654	DnaJ (Hsp40) homolog, subfamily B, member 11	DNAJB11

Table 2: Down regulated genes in Cervical Cancer

8081081	EPH receptor A3	EPHA3

4. DISCUSSION

The differentially expressed genes, their regulation in cervical cancer were studied.

According to the KEGG pathway, shown below ephrin binds to EPHA3 receptor. Ephexin acts as a Guanine Nucleotide Exchange factor (GEFs), which differentially activates GTPases RHOA, RAC1 and CDC42. Upon activation by Ephrin, GEF activity switches towards RHOA resulting in its activation [11]. GTPase activating proteins (GAP) negatively regulate Ras by stimulating Ras' intrinsic GTP hydrolyzing activity. Ras has an intrinsic GTPase activity, which means the protein on its own will hydrolyze a bound GTP molecule into GDP. However, this process is too slow for efficient function and hence GAP for Ras, RasGAP (mol. Wt. 120KD) may bind to and stabilize catalytic machinery of Ras. GAPs accelerate Ras inactivation [12]. It has been shown that there is strong binding between EPHA3 and RasGAP[13]. Thus, when EPHA3 is activated, it also activates RasGAP, which in turn inhibits, Ras, thus inactivating the MAPK signaling pathway [14].

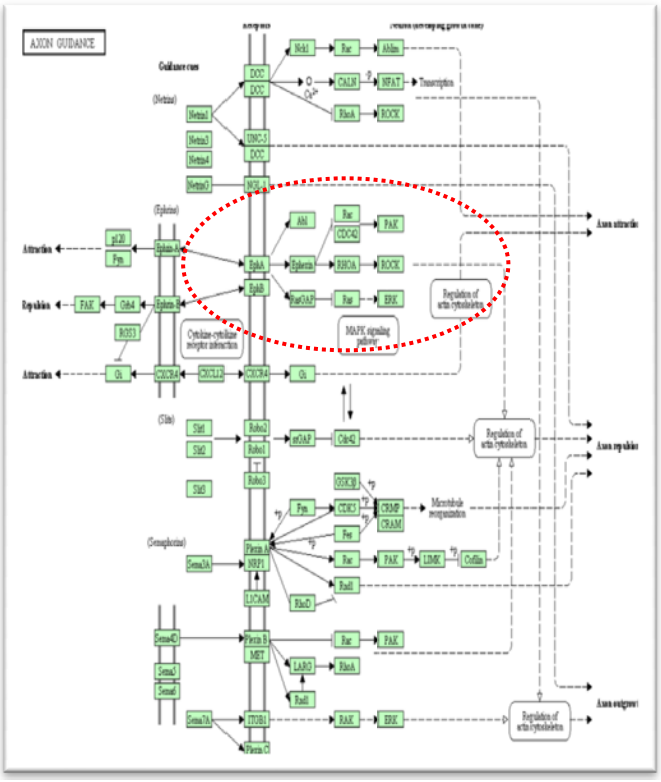

Fig. 4: KEGG pathway for EPHA3 gene

EPHA3 expression has been reported in many types of cancers. Eph receptors and their membrane bound ligands have limited expression and function in normal adult tissues[15]. On the other hand, during adulthood, many Ephs and ephrins are preferentially expressed in malignant tissue, where they are thought to participate in progressing invasive and metastatic cancers [16]. Eph and Ephrin over-expression in human cancers often correlates with aggressive, invasive and metastatic phenotypes [17], while deregulated expression or function of EPHs and ephrins is thought to play a role in the initial stages of epithelial neoplasia [18]. There has been contradictory reports stating that EPHA3 is a proto-oncogene and a tumor suppressor [18]. Eph/Ephrin- facilitated adhesive cell-cell contacts are critical during normal and oncogenic development suggesting a molecular switch controlling the direction of cellular responses to eph ligation [15]. There has been increasing evidence suggesting that stimulated Eph receptors and ephrins modulate integrin function .Thus deregulation of adhesion or loss of adhesion dependence is a hallmark of tumor progression and metastasis. It has been reported that silencing of EPHA3 through DNA methylation may be a late event allowing free movement of the tumor cells and hence contributing to the development of the metastatic phenotype [17].

Our studies observed the down regulation of EPHA3 gene may be responsible for variability in frequency of infections.

We can link the down regulation of EPHA3 may lead to RAS activation, causing uncontrolled differentiation and proliferation, thus playing a key role in metastasis. Therefore it may be concluded that down-regulation of EPHA3 in Cervical Cancer may also lead to loss of cell-to-cell adhesion and hence could play a key role in metastasis.

5. ACKNOWLEDGEMENT

The authors are grateful to the Head of the School of Biotechnology & Bioinformatics, D. Y. Patil University for providing the facilities to carry out the project.

REFERENCES

[1] http://globocan.iarc.fr/old/FactSheets/cancers/cervix-new.asp

[2] Jo, H. & Kim, J.W. Implications of HPV infection in uterine cervical cancer Review Article. *Cancer Therapy* **3**, 419-434 (2005).

[3] Del Toro-Arreola, S. *et al.* MHC class I-related chain A and B ligands are differentially expressed in human cervical cancer cell lines. *Cancer cell international* **11**, 15 (2011).

[4] Hamashima, C. *et al.* The Japanese guideline for cervical cancer screening. *Japanese journal ofclinical oncology* **40**, 485-502

[5] Vazquez-Mena, O. *et al.* Amplified genes may be overexpressed, unchanged, or downregulated in cervical cancer cell lines. *PloS one* **7**, e32667 (2012).

[6] Deshpande, A. *et al.* TNF-a Promoter Polymorphisms and Susceptibility to Human Papillomavirus 16 – Associated Cervical Cancer. *Cancer* **92037**, 969-976 (2005).

[7] Raybould, R., Fiander, A. & Hibbitts, S. Human Papillomavirus Integration and its Role in Cervical Malignant Progression. Oncogene 1-7 (2011).

[8] Nair, S. & Pillai, R.M. Human Papilloma Virus , Cellular Genetics and Susceptibility to Cervical Cancer *International Journal of Human Genetics* **5**, 11-17 (2005)

[9] Ma, Y.-Y. *et al.* PIK3CA as an oncogene in cervical cancer. *Oncogene* **19**, 2739-2744 (2000).

[10] Jo, H. & Kim, J.W. Implications of HPV infection in uterine cervical cancer Review Article. *Cancer Therapy* **3**, 419-434 (2005)

[11] Ephexin-1 - Homo sapiens (Human). *UniProt KB* (2012)

[12] Sundaram, M.V. RTK/Ras/MAPK signaling. *WormBook : the online review of C. elegans biology* 1-19 (2006).doi:10.1895/wormbook.1.80.

[13] Hu, T. *et al.* Regulation of process retraction and cell migration by EphA3 is mediated by the adaptor protein Nck1. *Biochemistry* **48**, 6369-6378 (2009).

[14] Ding, L. *et al.* Somatic mutations affect key pathways in lung adenocarcinoma. *Nature* **455**, 1069-1075 (2009).

[15] Wimmer-Kleikamp, S.H. *et al.* Elevated protein tyrosine phosphatase activity provokes Eph/ephrin-facilitated adhesion of pre-B leukemia cells. *Blood* **112**, 721-32 (2008).

[16] Vearing, C. et al. Concurrent binding of anti-EphA3 antibody and ephrin-A5 amplifies EphA3 signaling and downstream responses: potential as EphA3-specific tumor-targeting reagents. Cancer researh 65, 6745-54 (2005)

[17] Dottori, M. *et al.* Cloning and characterization of EphA3 (hek) Gene Promoter: DNA Methylation Regulates Expression in Hematopoietic Tumor Cells. *Cell* **3**, 2477-2486 (2012).

[18] Bonifaci, N. *et al.* Exploring the link between germline and somatic genetic alterations in breast carcinogenesis. *PloS one* **5**, e14078 (2010).