Pathol. Oncol. Res. (2017) 23:537–544
DOI 10.1007/s12253-016-0133-y

ORIGINAL ARTICLE

Expression Data Analysis for the Identification of Potential Biomarker of Pregnancy Associated Breast Cancer

Raja Rajeswary Thanmalagan[1] · **Leimarembi Devi Naorem**[1] · **Amouda Venkatesan**[1]

Received: 11 April 2016 / Accepted: 12 October 2016 / Published online: 10 November 2016

Abstract Breast cancer affects every 1 of 3000 pregnant women or in the first post-partum year is referred as Pregnancy Associated Breast Cancer (PABC) in mid 30s. Even-though rare disease, classified under hormone receptor negative status which metastasis quickly to other parts by extra cellular matrix degradation. Hence it is important to find an optimal treatment option for a PABC patient. Also additional care should be taken to choose the drug; in order to avoid fetal malformation and post-partum stage side-effects. The adaptation of target based therapy in the clinical practice may help to substitute the mastectomy treatment. Recent studies suggested that certain altered Post Translational Modifications (PTMs) may be an indicative of breast cancer progression; an attempt is made to consider the over represented PTM as a parameter for gene selection. The public dataset of PABC from GEO were examined to select Differentially Expressed Genes (DEG). The corresponding PTMs for DEG were collected and association between them was found using data mining technique. Usually clustering algorithm has been applied for the study of gene expression with drawback of clustering of gene products based on specified features. But association rule mining method overcome this shortcoming and determines the useful and in depth relationships. From the association, genes were selected to study the interactions and pathways. These studies emphasis that the genes KLF12, FEN1 MUC1 and SP110, can be chosen as target, which control cancer development, without any harm to pregnancy as well as fetal developmental process.

✉ Amouda Venkatesan
amouda@bicpu.edu.in

[1] Centre for Bioinformatics, Pondicherry University, Puducherry 605014, India

Keywords Network analysis · PTM · Pregnancy-associated breast cancer · Association rule mining

Introduction

Breast cancer is one of the most commonly diagnosed cancer during pregnancy is referred as pregnancy–associated breast cancer. Every 1 of 3000 pregnant women are associated with poor prognosis [1]. It is much important to find appropriate treatment for the breast cancer during pregnancy. Protein post-translational modification (PTMs) is recognized as key regulators of protein functions, led to diverse functions of proteins.

PTM can occur at any stage of protein to enhance or reduce their property and functionality or sometimes totally degrade the protein. So, as a whole, PTM are important controllers of the whole cellular functions. For example phosphorylation, glycosylation, ubiquitination etc., are the manipulators of many cellular events such as signal transduction, protein-protein interactions etc. Consequently the alteration in the PTM leads to affect the cellular growth mechanisms, which may in turn leads to abnormal cellular proliferation [2]. Its alteration also have progressive associations with many disease and disorders [3]. Thus, understanding the post translational modification is important to characterize the cancer biology.

Data mining method association rule mining is applied to select the specific set of gene from the DEG. Most of the previous studies applied clustering to analyze microarray gene expression data, in order to find out the group of gene expressions, in different biological situations. Usually clustering algorithm, groups the genes based on the similarities in two or more biological constraints. Accordingly a single gene cannot be present in two or more group even it has some similarity. Thus the main drawback of clustering is gene single grouping, which lacks the information that the single gene can interact

with different sets of gene; consequently unsupervised data mining technique describes the relationships among genes. Association rule mining is used to search for the frequent patterns and to be applied to gene expression data, in alternative to cluster technique [4–6]. Apriori algorithm is the competent algorithm [7] for predictive association mining of unknown knowledge from categorical data, which is applied to find the frequently associated PTMs for the differentially expressed genes of pregnancy-associated breast cancer (PABC). However the power of rule mining technique for the genes expression study is explored in this study. An attempt has been made to provide in-depth insight of the PTMs and their contribution in pregnancy associated breast cancer by an integrative analysis of important genes and pathways. Furthermore protein – protein interaction network has been constructed to study and identify the target genes.

Methods

Data Collection and Preprocessing

To study about the breast cancer impacts during pregnancy the dataset GSE31192 [8] was downloaded from the publicly available genomic data repository NCBI GEO (Gene Expression Omnibus) database, which is GPL570 (HG-U133 Plus 2) Affymetrix Human Genome U133 Plus 2.0 Array platform. In that 13 were normal and 20 were tumor samples includes both PABC and non- PABC patients. The raw dataset was preprocessed using limma package in R (V.3.10.1). The DEGs were selected from the normalized data of tumor and normal PABC samples using p-value <0.05 and |log FC| > = 1.5 as the threshold.

Association Rule Mining of PTM

For each of the DEG the PTMs were collected from UniProt [9] and organized as the transaction data set, which is compatible for Apriori algorithm. Using the Apriori algorithm implemented in WEKA, the frequently associated PTMs for DEGs have been identified. This algorithm search for the subsets of transaction from the item sets [10], can be applied to the expression data [5]. Here, datasets are organized in the form of transaction datasets. Each transaction contains a list of items; (differentially expressed genes and their corresponding PTMs). For an example, the rule (A → B, C) interpreted as B and C are frequently associated with A. hence the rule Phosphorylation → Nitrosylation, Acetylation., can be interpreted as nitrosylation and acetylation are frequently associated with phosphorylation and this association is overrepresented in DEGs. The quality of the associations was measured using two indexes: support and confidence. Hence the rules with high support and confidence were considered and their corresponding genes were collected for the further analysis like pathway enrichment analysis, network interaction, etc.,

Pathway Enrichment Analysis

Identification of pathways and gene networks for DEGs in cancer progression would yield biologically significant information of the underlying cellular mechanism. The KEGG PATHWAY database is a standard and comprehensive which provides a valuable resource for various biological networks [11]. The clustered genes were analyzed for pathway enrichment analysis using DAVID functional enrichment tool with the threshold of p-value <0.05 and gene count >2.

Interaction Network Analysis

For both up and down regulated DEGs the protein-protein interaction network were constructed using online GeneMANIA [12] program with attributes namely pathway, co-expression, genetic interaction, physical interaction and shared protein domain. The functional enrichment analysis for the network was done by DAVID tool [13]. The network was visualized by Cytoscape [14], the centrality measures were calculated by a plugin Network Analyzer. The network topological parameter, betweenness [15] centrality measures the node's control over the information flow in the network and the node with high betweenness centrality can influence the information flow by altering or hindering the communication in the network [16]. Hence here it is used as measure to select the hub gene in the network [17].

Result

Differentially Expressed Genes

Based on p value <0.05 and |log FC| > = 1.5 as threshold, 352 down-regulated genes and 321 up-regulated genes were identified. Out of it, top 10 DEGs were shown in tables 1 and 2. The corresponding PTMs for both up and down-regulated genes were retrieved from UniProt. Most of the up regulated gene products are ribonucleotide binding proteins, involved in cell-cycle process and many of them resides in cytoplasm. The down regulated gene products are membrane proteins, having calcium ion binding and kinase activity and involved in receptor linked signal transduction process.

Association Rule Mining of PTM

The up and down regulated genes with PTMs are organized as transaction dataset to find the frequently associated pattern using Apriori algorithm are shown in tables 3 and 4. With confidence threshold 1.0 the top ten rules were selected.

Table 1 Top ten up regulated DEG

ID	Gene symbol	Gene title	P Value	log FC
204712_at	WIF1	WNT inhibitory factor 1	3.93E-10	6.347919
208399_s_at	EDN3	Endothelin 3	2.69E-11	5.744123
223623_at	C2orf40	chromosome 2 open reading frame 40	3.54E-11	5.508598
202037_s_at	SFRP1	secreted frizzled-related protein 1	0.000204	5.128947
205051_s_at	KIT	v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog	2.53E-06	5.065702
1552509_a_at	CD300LG	CD300 molecule-like family member g	2E-07	4.998486
1553602_at	MUCL1	mucin-like 1	1.1E-06	4.994673
231535_x_at	ROPN1	rhophilin associated tail protein 1	0.000173	4.747366
227194_at	FAM3B	family with sequence similarity 3, member B	8.83E-06	4.701641
226147_s_at	PIGR	polymeric immunoglobulin receptor	3.64E-06	4.639175

Tables 5 and 6 shows some of the association rules for up and down regulated genes respectively. It shows that, up regulated genes are enriched in phosphorylation, associated with Sumoylation and Caspase. Also down regulated genes enriched in phosphorylation and associated with ubiquitilation, N-Linked glycosylation and Acetylation. The genes involving in the top ten associations were selected for the further analysis.

Pathway Enrichment Analysis

Pathway enrichment analysis reveals the whole set of interconnected events and their biological interactions of cluster identified. It clearly shows that most of the differentially expressed genes are enriched in Cytokine-cytokine receptor interaction, Focal adhesion, Chemokine signaling pathway. Deficiency in cytokine leads susceptibility to viral infections as well as tumor growth [18]. Focal adhesion pathway is important in cell proliferation, cell survival and cell migration. Altered activities of focal adhesion kinases are associated with cancer cells [19]. Table 7 shows that most of the growth related pathways are altered in this data set. Most of the genes are enriched in development process, since these are pregnant patient's samples. Further analysis is required to eliminate genes involved in such developmental process during pregnancy in order to find the cancer related gene by constructing an interaction network.

Interaction Network Analysis

The protein-protein interaction network provides the topological and dynamic features of gene products involved in the disease mechanisms. The interaction network for the selected up and down regulated genes was constructed by

Table 2 Top ten down regulated DEG

ID	Gene symbol	Gene title	P Value	logFC
205242_at	CXCL13	chemokine (C-X-C motif) ligand 13	1.92E-05	-5.06924
213909_at	LRRC15	leucine rich repeat containing 15	9.03E-06	-5.00452
213831_at	HLA-DQA1	major histocompatibility complex, class II, DQ alpha 1	0.00156	-4.89899
203915_at	CXCL9	chemokine (C-X-C motif) ligand 9	0.000394	-4.81126
203936_s_at	MMP9	matrix metallopeptidase 9 (gelatinase B, 92 kDa gelatinase, 92 kDa type IV collagenase)	4.72E-08	-4.75841
242579_at	BMPR1B	bone morphogenetic protein receptor, type IB	0.000879	-4.68206
211122_s_at	CXCL11	chemokine (C-X-C motif) ligand 11	6.77E-05	-4.66402
204533_at	CXCL10	chemokine (C-X-C motif) ligand 10	7.5E-06	-4.53807
209480_at	HLA-DQB1	major histocompatibility complex, class II, DQ beta 1	0.00266	-4.48946
210523_at	BMPR1B	bone morphogenetic protein receptor, type IB	0.0021	-4.48795
209773_s_at	RRM2	ribonucleotide reductase M2	5.43E-07	-4.41353
229975_at	BMPR1B	bone morphogenetic protein receptor, type IB	0.000378	-4.35931
206134_at	ADAMDEC1	ADAM-like, decysin 1	2.45E-05	-4.22363

Table 3 Up – regulated DEG enriched in acetylation, ubiquitylation, methylation and citrulination

Gene symbol	Post translational modification
AADAT	Acetylation, Pyridoxal phosphate
ABCA6	Phosphorylation
ABCB1	Phosphorylation, Ubiquitylation
ABCG2	Phosphorylation
ACADSB	Phosphorylation, Ubiquitylation, Acetylation
ACTA2	Phosphorylation, Ubiquitylation, Acetylation, S-nitrosylation, Nitration
ACTG2	Phosphorylation, Ubiquitylation, Acetylation

GeneMANIA, which consist of 86 node and 652 edges. Figure 1 shows the gene product interaction network for up and down regulated genes. The gene products are differentiated based on their betweenness centrality score. The network has been visualized and betweenness centrality of each node of the network was calculated by Cytoscape visualization tool, shown in Table 7. The genes KLF12, COL17A1, MKI67, BLM, FEN1, SP110, MUC1, TFAP2C, EGFR, TFRC, IRF1, TTK, STAT1, KIRREL, PDZRN3, RRM2, FYB with high betweenness centrality were selected. The contribution of selected genes towards cancer progression and pregnancy are tabulated (Table 8), which helps to find the genes involving in development of cancer which is not involved in any form of fetal development or any other pregnancy related process. And the roles of those genes are discussed briefly.

Discussion

It is a fact that pregnancy lowers lifetime risk of developing breast cancer. And it is evident that incidence of breast cancer observed in nulliparous women and women giving birth

Table 4 Down - regulated genes enriched in ubiquitilation, phosphorylation and caspase

Gene symbol	Post translational modification
ABHD2	Ubiquitylation
ACER3	Phosphorylation
ACP2	N-linked Glycosylation
ACP5	N-linked Glycosylation
ADAMDEC1	N-linked Glycosylation
ADCY7	Ubiquitylation, Acetylation
AMPH	Phosphorylation, Acetylation
ANLN	Phosphorylation, Acetylation
AP1S2	Ubiquitylation

Table 5 Best association rule - up regulated genes

1. N-linked Glycosylation = Ubiquitylation 11 == > Phosphorylation 11 conf:(1)
2. N-linked Glycosylation = Acetylation 7 == > Phosphorylation 7 conf:(1)
3. Ubiquitylation = Sumoylation 7 == > Phosphorylation 7 conf:(1)
4. Dephosphorylation 6 == > Phosphorylation 6 conf:(1)
5. Methylation 6 == > Phosphorylation 6 conf:(1)

at late 30s [20–22]. Approximately 7 % of women with breast cancer are diagnosed before the age of 40 years, this disease accounts for more than 40 % of all cancer in women in this age group [23]. Rare breast cancer in young women is worthy of special attention due to the unique and complex issues that are raised [24]. Thus, a better understanding of driver pathways and genes of PABCs is imperative for improved diagnosis and therapeutic strategies for pregnant and lactating women [20]. Our study is to find specific genes and pathways in PABC tissue expression. From the association rules, the PTM pattern of the PABC was explored, which contributes in every stage of protein's lifetime to regulate their function. The phosphorylation of proteins regulates almost all aspects of all living cell, modification in the ratio of phosphorylation results in modifications in their function which reflect in the cellular such as cancer evaluation [25]. Phosphorylation and Sumoylation of progesterone receptor involve in the regulation of mammary gland development. Poor Sumoylation of progesterone receptors significantly associated with cancer metastasis and shorten the survival [26]. Phosphorylated progesterone receptor might be under sumoylated during the development of breast cancer or mammary gland development [27]. Caspase cleavage is regulating the apoptotic cell death, change in caspase activity leads to disease such as cancer [28, 29]. N-linked glycosylation is important for the stability of the ATP Binding Cassette (ABC) transporter. The increased expression of ABCG2 results in resistance to chemotherapy [30, 31]. The de-glycosylated ABC transporters, which are known as multidrug resistance proteins in cancer cells, are degraded by Ubiquitylation. Ubiquitylation regulates the stability of glycol proteins, so that they affect the functions of the membrane proteins that mediate multi drug resistance

Table 6 Best association rule - down regulated genes

1. Sumoylation 17 == > Phosphorylation 17 conf:(1)
2. Caspase 16 == > Phosphorylation 16 conf:(1)
3. Methylation 16 == > Phosphorylation 16 conf:(1)
4. Ubiquitylation 11 = Methylation 13 == > Phosphorylation 13 conf:(1)
5. Palmitoylation 12 == > Phosphorylation 12 conf:(1)

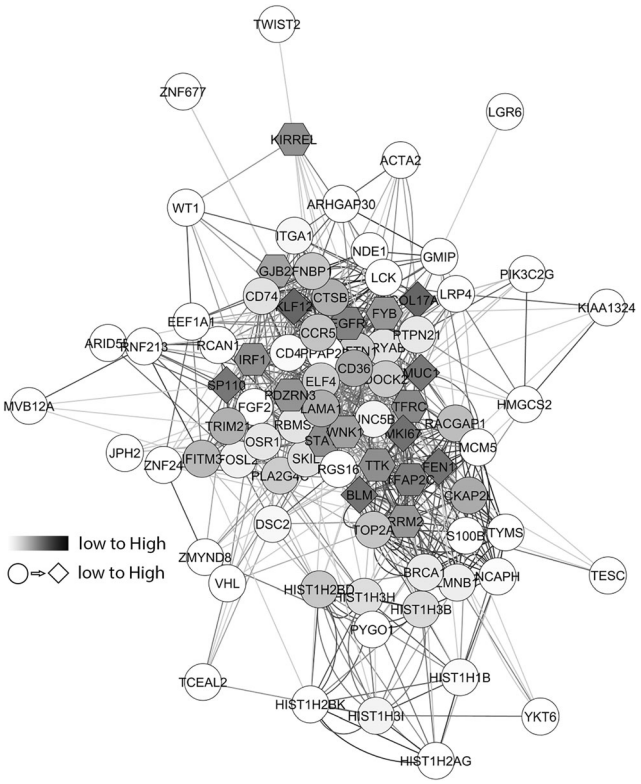

Fig. 1 Protein – protein interaction network constructed based on the Betweenness Centrality. Cytoscape tool used to visualize the interactions and VizMapper graphics plugin is utilized to highlight the network with different shapes and shades. The proteins with higher betweenness centrality are highlighted by dark shades as well as different shapes

[32]. It is found that glycoproteins constitutively ubiqutinated in cancer cells. These relationships between the PTM are mined by association rules. The over represented relationships and their corresponding genes of PABC were selected from the top 10 rules and their interaction networks were obtained from GeneMANIA tool. From the network, the DEGs, KLF12, COL17A1, MKI67, BLM, FEN1, SP110, MUC1, TFAP2C, EGFR, TFRC, IRF1, TTK, STAT1, KIRREL, PDZRN3, RRM2, FYB were selected as hub genes by using network topological parameter betweenness centrality which is the node's centrality in a network. Majority of the hub genes are related to pregnancy, fetal development and also related to cancer initiation, progression and metastasis. Among interacts, SP110 and KIRREL are expressed in cancer tissues but still the role of these genes in cancer is unclear. MKI67 is involved in cellular proliferation and reported as potential target for HR positive breast cancer [33] and also expressed in normal pregnant patients [34], hence targeting it would harm the fetal development. BLM, TFAP2C, EGFR, LAMA1, CTSB, TTK, KIRREL, PDZRN3 and RRM2 were known to be expressed in cancer tissues and also previously reported as biomarkers but they are involving in critical roles such as pronephros, brain, eyes, embryonic angiogenic remodeling of fetal development process [35–43]. Accordingly inhibition or regeneration of these genes will affect the fetal development. So the genes such as KLF12, FEN1 SP110 and MUC1, which are involving in cancer and not harm to pregnancy, lactation as well as fetal developmental related process were chosen for the further studies. KLF12 is the transcription factor reported as potential target for gastric cancer and also a negative regulator of decidualization and implantation of maternal endometrium development. Hence down regulation of KLF12 may improve the growth and development of the conceptus and also it prevents the cancer growth. FEN1 is the tumor suppressor gene and overexpression of FEN1 leads resistance to chemotherapy. Its overexpression during pregnancy leads to embryonic lethality and normal fetal development was observed in FEN−/− in mouse model [44]. SP110 is involved in chromatin remodeling and formation, but up-regulation of SP110 results in hepatic veno-occlusive disease with immunodeficiency for fetus; hence inhibition of its expression would help in the progressive fetal development [45]. MUC1 is the important

Table 7 Pathway enrichment score

KEGG ID	Term	Count	%	P-Value	FDR
hsa05200	Pathways in cancer	51	3.63766	7.78E-04	0.944223
hsa04060	Cytokine-cytokine receptor interaction	48	3.42368	2.05E-05	0.02504
hsa04510	Focal adhesion	43	3.067047	1.07E-06	0.001303
hsa04514	Cell adhesion molecules (CAMs)	39	2.78174	3.44E-10	4.19E-07
hsa04062	Chemokine signaling pathway	37	2.639087	4.32E-05	0.052655
hsa04810	Regulation of actin cytoskeleton	31	2.211127	0.0274	28.7232
hsa05322	Systemic lupus erythematosus	27	1.92582	1.64E-06	0.001993
hsa04612	Antigen processing and presentation	26	1.854494	1.50E-07	1.83E-04
hsa05416	Viral myocarditis	25	1.783167	2.21E-08	2.70E-05
hsa04670	Leukocyte trans-endothelial migration	25	1.783167	3.53E-04	0.42899

Table 8 Network topological parameters

ID	Betweenness Centrality	Closeness Centrality	Clustering Coefficient	Degree	Radiality
KLF12	0.039352	0.548387	0.213439	23	0.794118
COL17A1	0.038248	0.53125	0.176471	17	0.779412
MKI67	0.034449	0.574324	0.333333	25	0.814706
BLM	0.034054	0.551948	0.303333	25	0.797059
FEN1	0.033575	0.559211	0.332016	23	0.802941
SP110	0.032975	0.521472	0.236842	20	0.770588
MUC1	0.03264	0.566667	0.221014	24	0.808824
TFAP2C	0.030914	0.559211	0.272727	23	0.802941
EGFR	0.029995	0.551948	0.251082	22	0.797059
TFRC	0.028109	0.551948	0.209524	21	0.797059
IRF1	0.026723	0.548387	0.264069	22	0.794118
TTK	0.026655	0.574324	0.356667	25	0.814706
STAT1	0.026586	0.537975	0.236842	20	0.785294
KIRREL	0.026485	0.469613	0.138889	9	0.717647
PDZRN3	0.025385	0.548387	0.226316	20	0.794118
RRM2	0.024582	0.559211	0.343874	23	0.802941
FYB	0.024419	0.53125	0.304094	19	0.779412
WNK1	0.022901	0.537975	0.191176	17	0.785294
GJB2	0.022342	0.508982	0.171429	15	0.758824

gene in preventing embryo implantation and developing Ectopic Pregnancy. It also interacts with EGRF and other receptor tyrosine kinases in the cell membrane and activates the PI3K/ AKT which is the most altered pathway in cancer development. It is localizes in the nucleus and activates the Wnt/B-catenin, Signal transducer and activation of transcription (STAT) and involves in the self-renewal of breast cancer cells NF-κB➔IL-8/CXCR1 pathway [46, 47]. Thus MUC1 may act as the potential 5target for the pregnancy associated breast cancer. Hence this study uncovers that four genes (KLF12, FEN1, SP110 and MUC1) might be the potential target for the PABC, which are not affect the fetal development and improves fetal implantation as well. For further validation the expression of the above said four genes were compared between the cancer samples (antibody staining) and breast and female reproductive system tissues from Human Protein Atlas (HPA) [48–50] Database. The gene SP110 expressed low level in breast and female reproductive tissues and medium in breast cancer tissue samples. And the gene FEN1 in expressed high in cancer tissues as well as endometrium, ovary and placenta and low in normal breast tissue. The Gene MUC1 shows high expression in antibody staining of breast cancer tissues but very low or not detected in ovary placenta and breast tissues. Hence this observation supports that targeting these may not disturb the normal tissues where they are expressed very less and can control the cancer progression.

Acknowledgment We thank Pranitha Jenardhanan, Manivel Panneerselvam and Kannan Muthu for their valuable suggestions.

Support and Funding This work was carried out in Centre for Bioinformatics, Pondicherry University under the UGC funded project.

Compliance with Ethical Standards

Competing Interests The authors declare that they have no competing interests.

Authors' Contributions RT designed the study, performed the analysis and wrote the manuscript. AV, NL contributed to critical review of the manuscript. All authors read and approved the final manuscript.

References

1. Keyser EA, Staat BC, Fausett MB, Shields AD (2012) Pregnancy-associated breast cancer. Rev Obstet Gynecol 5(2):94–99
2. Jin H, Zangar RC (2009) Protein modifications as potential biomarkers in breast cancer. Biomark Insights 4:191–200
3. Wang Y-C, Peterson SE, Loring JF (2014) Protein post-translational modifications and regulation of pluripotency in human stem cells. Cell Res 24(2):143–160. doi:10.1038/cr.2013.151
4. Becquet C, Blachon S, Jeudy B, Boulicaut J-F, Gandrillon O (2002) Strong-association-rule mining for large-scale gene-expression data analysis: a case study on human SAGE data. Genome Biol 3(12) research0067.0061-research0067.0016

5. Creighton C, Hanash S (2003) Mining gene expression databases for association rules. Bioinformatics (Oxford, England) 19(1):79–86
6. Seeja KR, Alam MA, Jain SK (2009) An association rule mining Approach for co-regulated Signature genes identification in cancer. J Circ Sys Comp 18(08):1409–1423. doi:10.1142/S0218126609005757
7. Pugazhendi D (2013) Apriori algorithm on marine fisheries biological data. Int J Comp Sci & Eng Technol 4(12):1409–1411
8. Harvell DM, Kim J, O'Brien J, Tan AC, Borges VF, Schedin P, Jacobsen BM, Horwitz KB (2013) Genomic signatures of pregnancy-associated breast cancer epithelia and stroma and their regulation by estrogens and progesterone. Horm & cancer 4(3):140–153. doi:10.1007/s12672-013-0136-z
9. The UniProt C (2012) Reorganizing the protein space at the universal protein resource (UniProt). Nucleic Acids Res 40(D1):D71–D75
10. Agrawal R, Srikant R (1994) Fast algorithms for mining association rules in large databases. Paper presented at the proceedings of the 20th international conference on very large data bases
11. Kanehisa M, Goto S, Kawashima S, Nakaya A (2002) The KEGG databases at GenomeNet. Nucleic Acids Res 30(1):42–46
12. Zuberi K, Franz M, Rodriguez H, Montojo J, Lopes CT, Bader GD, Morris Q (2013) GeneMANIA prediction server 2013 update. Nucleic Acids Res 41(W1):W115–W122
13. Huang DW, Sherman BT, Tan Q, Kir J, Liu D, Bryant D, Guo Y, Stephens R, Baseler MW, Lane HC, Lempicki RA (2007) DAVID bioinformatics resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. Nucleic Acids Res 35(Web Server issue):W169–W175. doi:10.1093/nar/gkm415
14. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13(11):2498–2504. doi:10.1101/gr.1239303
15. Freeman LC (1978) Centrality in social networks conceptual clarification. Soc Networks 1(3):215–239. doi:10.1016/0378-8733(78)90021-7
16. Valente TW, Coronges K, Lakon C, Costenbader E (2008) How correlated are network centrality measures? Connections (Toronto, Ont) 28(1):16–26
17. Azuaje FJ (2014) Selecting biologically informative genes in co-expression networks with a centrality score. Biol Direct 9:12. doi:10.1186/1745-6150-9-12
18. Chen HM, Tanaka N, Mitani Y, Oda E, Nozawa H, Chen JZ, Yanai H, Negishi H, Choi MK, Iwasaki T, Yamamoto H, Taniguchi T, Takaoka A (2009) Critical role for constitutive type I interferon signaling in the prevention of cellular transformation. Cancer Sci 100(3):449–456. doi:10.1111/j.1349-7006.2008.01051.x
19. Golubovskaya VM, Kweh FA, Cance WG (2009) Focal adhesion kinase and cancer. Histol Histopathol 24(4):503–510
20. McCready J, Arendt LM, Glover E, Iyer V, Briendel JL, Lyle SR, Naber SP, Jay DG, Kuperwasser C (2014) Pregnancy-associated breast cancers are driven by differences in adipose stromal cells present during lactation. Breast Cancer Res : BCR 16(1):R2. doi:10.1186/bcr3594
21. Russo J, Moral R, Balogh GA, Mailo D, Russo IH (2005) The protective role of pregnancy in breast cancer. Breast Cancer Res : BCR 7(3):131–142. doi:10.1186/bcr1029
22. Britt K, Ashworth A, Smalley M (2007) Pregnancy and the risk of breast cancer. Endocr Relat Cancer 14(4):907–933. doi:10.1677/erc-07-0137
23. Anders CK, Johnson R, Litton J, Phillips M, Bleyer A (2009) Breast cancer before age 40 Years. Semin Oncol 36(3):237–249. doi:10.1053/j.seminoncol.2009.03.001
24. Gabriel CA, Domchek SM (2010) Breast cancer in young women. Breast Cancer Res : BCR 12(5):212–212. doi:10.1186/bcr2647
25. Pawson T, Scott JD (2005) Protein phosphorylation in signaling – 50 years and counting. Trends Biochem Sci 30(6):286–290. doi:10.1016/j.tibs.2005.04.013
26. Knutson TP, Daniel AR, Fan D, Silverstein KA, Covington KR, Fuqua SA, Lange CA (2012) Phosphorylated and sumoylation-deficient progesterone receptors drive proliferative gene signatures during breast cancer progression. Breast Cancer Res : BCR 14(3):R95. doi:10.1186/bcr3211
27. Daniel AR, Faivre EJ, Lange CA (2007) Phosphorylation-dependent antagonism of sumoylation derepresses progesterone receptor action in breast cancer cells. Mol endocrinol (Baltimore, Md) 21(12):2890–2906. doi:10.1210/me.2007-0248
28. McIlwain DR, Berger T, Mak TW (2013) Caspase functions in cell death and disease. Cold Spring Harb Perspect Biol 5(4):a008656. doi:10.1101/cshperspect.a008656
29. Elmore S (2007) Apoptosis: a review of programmed cell death. Toxicol Pathol 35(4):495–516. doi:10.1080/01926230701320337
30. Nakagawa H, Wakabayashi-Nakao K, Tamura A, Toyoda Y, Koshiba S, Ishikawa T (2009) Disruption of N-linked glycosylation enhances ubiquitin-mediated proteasomal degradation of the human ATP-binding cassette transporter ABCG2. FEBS J 276(24):7237–7252. doi:10.1111/j.1742-4658.2009.07423.x
31. Nakanishi T, Ross DD (2012) Breast cancer resistance protein (BCRP/ABCG2): its role in multidrug resistance and regulation of its gene expression. Chinese J Cancer 31(2):73–99. doi:10.5732/cjc.011.10320
32. Nakagawa H, Tamura A, Wakabayashi K, Hoshijima K, Komada M, Yoshida T, Kometani S, Matsubara T, Mikuriya K, Ishikawa T (2008) Ubiquitin-mediated proteasomal degradation of non-synonymous SNP variants of human ABC transporter ABCG2. Biochem J 411(3):623–631. doi:10.1042/bj20071229
33. Zhang MH, Man HT, Zhao XD, Dong NI, Ma SL (2014) Estrogen receptor-positive breast cancer molecular signatures and therapeutic potentials (review). Biomed Rep 2(1):41–52. doi:10.3892/br.2013.187
34. Schraenen A, de Faudeur G, Thorrez L, Lemaire K, Van Wichelen G, Granvik M, Van Lommel L, in't Veld P, Schuit F (2010) mRNA expression analysis of cell cycle genes in islets of pregnant mice. Diabetologia 53(12):2579–2588. doi:10.1007/s00125-010-1912-8
35. Zhang H, Zhu X, Chen J, Jiang Y, Zhang Q, Kong C, Xing J, Ding L, Diao Z, Zhen X, Sun H, Yan G (2015) Kruppel-like factor 12 is a novel negative regulator of forkhead box O1 expression: a potential role in impaired decidualization. Reprod Biol Endocrinol : RB&E 13:80. doi:10.1186/s12958-015-0079-z
36. Chester N, Kuo F, Kozak C, O'Hara CD, Leder P (1998) Stage-specific apoptosis, developmental delay, and embryonic lethality in mice homozygous for a targeted disruption in the murine Bloom's syndrome gene. Genes Dev 12(21):3382–3393
37. Li W, Cornell R (2007) Redundant activities of Tfap2a and Tfap2c are required for neural crest induction and development of other non-neural ectoderm derivatives in zebrafish embryos. Dev Biol 304(1):338–354. doi:10.1016/j.ydbio.2006.12.042
38. Piccinni SA, Bolcato-Bellemin A-L, Klein A, Yang VW, Kedinger M, Simon-Assmann P, Lefebvre O (2004) Krüppel-like factors regulate the Lama1 Gene encoding the laminin α1 chain. J Biol Chem 279(10):9103–9114. doi:10.1074/jbc.M305804200
39. Spencer TE, Bazer FW (2004) Uterine and placental factors regulating conceptus growth in domestic animals. J Anim Sci 82(E-Suppl):E4–13
40. Iwase T, Tanaka M, Suzuki M, Naito Y, Sugimura H, Kino I (1993) Identification of protein-tyrosine kinase genes preferentially expressed in embryo stomach and gastric cancer. Biochem Biophys Res Commun 194(2):698–705. doi:10.1006/bbrc.1993.1878

41. Hayata T, Blitz IL, Iwata N, Cho KWY (2009) Identification of embryonic pancreatic genes using Xenopus DNA microarrays. Dev Dyn 238(6):1455–1466. doi:10.1002/dvdy.21868
42. Sewduth RN, Jaspard-Vinassa B, Peghaire C, Guillabert A, Franzl N, Larrieu-Lahargue F, Moreau C, Fruttiger M, Dufourcq P, Couffinhal T, Duplaa C (2014) The ubiquitin ligase PDZRN3 is required for vascular morphogenesis through Wnt/planar cell polarity signalling. Nat Commun 5:4832. doi:10.1038/ncomms5832
43. De Hertogh R, Ekka E, Vanderheyden I, Glorieux B (1986) Estrogen and progestogen receptors in the implantation sites and interembryonic segments of rat uterus endometrium and myometrium. Endocrinology 119(2):680–684. doi:10.1210/endo-119-2-680
44. Larsen E, Gran C, Sæther BE, Seeberg E, Klungland A (2003) Proliferation failure and gamma radiation sensitivity of Fen1 null mutant mice at the blastocyst stage. Mol Cell Biol 23(15):5346–5353. doi:10.1128/MCB.23.15.5346-5353.2003
45. Wang T, Ong P, Roscioli T, Cliffe ST, Church JA (2012) Hepatic veno-occlusive disease with immunodeficiency (VODI): first reported case in the U.S. and identification of a unique mutation in Sp110. Clin Immunol (Orlando, Fla) 145(2):102–107. doi:10.1016/j.clim.2012.07.016
46. Horm TM, Schroeder JA (2013) MUC1 and metastatic cancer: expression, function and therapeutic targeting. Cell Adhes Migr 7(2):187–198. doi:10.4161/cam.23131
47. Alam M, Rajabi H, Ahmad R, Jin C, Kufe D (2014) Targeting the MUC1-C oncoprotein inhibits self-renewal capacity of breast cancer cells. Oncotarget 5(9):2622–2634. doi:10.18632/oncotarget.1848
48. Uhlen M, Bjorling E, Agaton C, Szigyarto CA, Amini B, Andersen E, Andersson AC, Angelidou P, Asplund A, Asplund C, Berglund L, Bergstrom K, Brumer H, Cerjan D, Ekstrom M, Elobeid A, Eriksson C, Fagerberg L, Falk R, Fall J, Forsberg M, Bjorklund MG, Gumbel K, Halimi A, Hallin I, Hamsten C, Hansson M, Hedhammar M, Hercules G, Kampf C, Larsson K, Lindskog M, Lodewyckx W, Lund J, Lundeberg J, Magnusson K, Malm E, Nilsson P, Odling J, Oksvold P, Olsson I, Oster E, Ottosson J, Paavilainen L, Persson A, Rimini R, Rockberg J, Runeson M, Sivertsson A, Skollermo A, Steen J, Stenvall M, Sterky F, Stromberg S, Sundberg M, Tegel H, Tourle S, Wahlund E, Walden A, Wan J, Wernerus H, Westberg J, Wester K, Wrethagen U, Xu LL, Hober S, Ponten F (2005) A human protein atlas for normal and cancer tissues based on antibody proteomics. Mol Cell Proteomics : MCP 4(12):1920–1932. doi:10.1074/mcp.M500279-MCP200
49. Uhlen M, Oksvold P, Fagerberg L, Lundberg E, Jonasson K, Forsberg M, Zwahlen M, Kampf C, Wester K, Hober S, Wernerus H, Bjorling L, Ponten F (2010) Towards a knowledge-based human protein atlas. Nat Biotechnol 28(12):1248–1250. doi:10.1038/nbt1210-1248
50. Ponten F, Jirstrom K, Uhlen M (2008) The human protein atlas–a tool for pathology. J Pathol 216(4):387–393. doi:10.1002/path.2440