The 4th Iranian Conference on Systems Biology

A Framework for Improving the Predictive Model of Breast Cancer Recurrence

Parvin Zarei[1], Mohammad reza Sehhati[2]*

[1] *Student Research Committee, department of Bioinformatics, School of Advanced Technologies in Medicine, Isfahan University of Medical Sciences, Isfahan, Iran*

[2] *Department of Bioinformatics, School of Advanced Technologies in Medicine, Isfahan University of Medical Sciences, Isfahan, Iran*

**mr.sehhati@amt.mui.ac.ir*

Abstract

Objective and background

Breast cancer recurrence (BCR) is therapeutically a giant difficulty. To precisely predict prognosis and recognize patients who can take survival advantages from adjuvant therapies we need to identify biomarkers for the risk of BCR. Previous studies on prediction of BCR applied various feature selection algorithms to discover oncogenes from the gene expression microarray data. They are not reliable for biological interpretation [1]. Recently, we proposed a feature-scoring criterion to select a stable gene set from microarray data with a good prediction power in independent datasets [2]. In this paper, we attempt to draw a more organized picture of the BCR hallmarks to provide a systematic viewpoint of gene signature sets, which obtained by the mentioned in-house developed method.

Keywords:

Breast cancer recurrence, prediction, cancer hallmarks, biological pathways

Dataset	#Samples [a]	#High-risk	#Low-risk	Source
GSE2034	286	95	169	[2]
GSE390	198	54	100	[33]
GSE6532	244	65	123	[34]
GSE4922	249	70	139	[35]
GSE3494	236	37	158	[3]
GSE2990	125	30	67	[36]
GSE11121	200	28	136	[37]
metadata	1,538	579	959	

Table 1. Summary of the Utilized Microarray Datasets

Introduction

TPX2 is a microtubule-associated protein that promotes tumor malignancy by activating of PI3K/AKT/Bcl_2 pathway and inhibition of P_{21}, P_{53} and caspase 3. Inhibition of P_{21}, P_{53} and caspase 3 causes BCR by increasing the expression of kinesin family member 2C (KIF2C) and ubiquitin-conjugating enzyme E2C (UBE2C) genes [2]. The lack of retinoblastoma (Rb) protein activity and increased expression of forkhead box M1 transcription factor (FOXM1) directly activates transcription centromere protein A (CENPA) which leads to improper chromosome segregation, aneuploidy and multiple spindle poles that induces BCR [3]. CENPA leads to aberrant formation of multicentric chromosomes. These aberrant structures causing aneuploidy and prone cells to tumor and progression of malignancy. High expression of ASPM and cyclinB_2 enhances tumor metastasis that induce the BCR [4]. Maternal fetal leucine zinc kinase (MELK) is an oncogenic kinase that is essential for the progression of mitosis in breast cancer cells and is a key regulator of malignancy and cell proliferation [5].

Frequency	Gene Symbol	Entrez ID	Gene Name
838	TTK	7,272	TTK protein kinase
837	KIF2C	11,004	Kinesin family member 2C
836	CENPA	1,058	Centromere protein A
836	CCNB2	9,133	Cyclin B2
836	FOXM1	2,305	Forkhead box M1
835	TPX2	22,974	TPX2, microtubule-associated, homolog
834	ASPM	259,266	ASP (abnormal spindle) homolog, microcephaly associated (Drosophila)
834	UBE2C	11,065	Ubiquitin-conjugating enzyme E2C
833	MELK	9,833	Material embryonic leucine zipper kinase
833	C10orf3	55,165	Centrosomal protein 55kDa

Table 2. Top Frequent Genes Selected by the Proposed Method

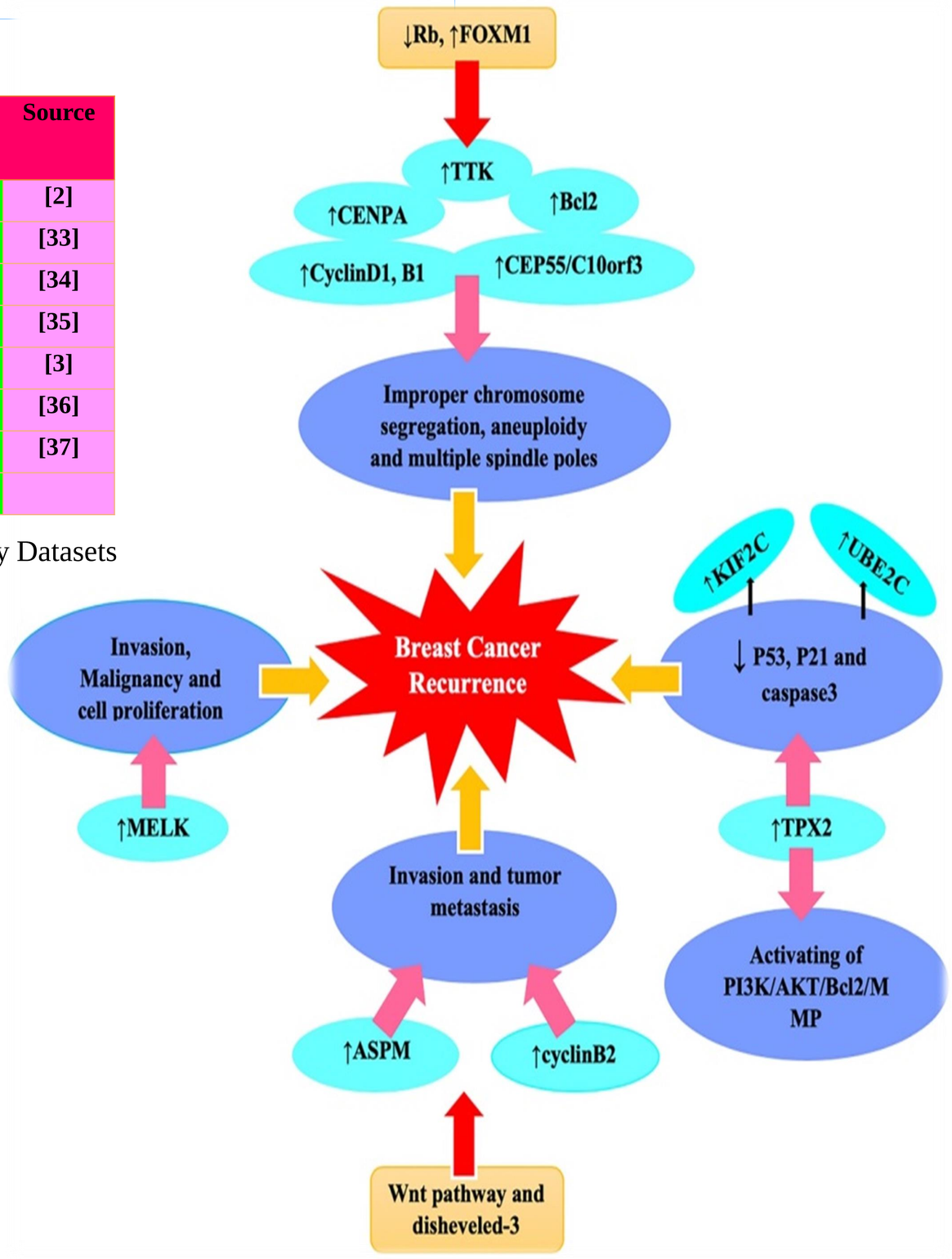

Figure 1. Schematic for demonstration of interaction among BCR hallmarks

Methodology

Using our previously developed technique for gene selection from microarray data [2], we reached to a 50-genes signature set for prediction of BCR in seven independent gene expression datasets. Also, a list of the 10 most frequently selected genes from 1000 datasets subsampled from combination of all datasets was reported as potential biomarkers of BCR (Table 1). For updating our scoring criterion to construct and rank the gene neighborhood regions, analysis of function of genes in different biological processes and pathways are described next.

Results

The schematic drawn in Figure 1 demonstrate how the cancer hallmarks, which selected by an artificial intelligence approach, act and cooperate in order to develop BCR.

Discussion, Conclusion and Suggestions

As PPI network previously used for enhancing the prediction accuracy of BCR, the proposed schematic introduces a framework for including other new gene sets in the model that may lead to more improvement in its predictive power.

References

1. Stable Gene Signature Selection for Prediction of Breast Cancer Recurrence Using Joint Mutual Information. IEEE/ACM Transactions on Computational BiSehhati M, Mehridehnavi A, Rabbani H, Javanmard SH. Using protein interactions database and support vector machine to improve gene signatures for prediction of breast cancer recurrence. Journal of Medical Signals and Sensors. 2013; 3(2):87–93.
2. Sehhati M, Mehridehnavi A, Rabbani H, Pourhossein M.ology and Bioinformatics. 2015; 12(6): 1440-48.
3. Chen, M, Zhang H, Zhang G, Zhong A, Ma Q, Kai J, Tong Y, Xie S, Wang Y, Zheng H, Guo L and Lu R. "Targeting TPX2 suppresses proliferation and promotes apoptosis via repression of the PI3k/AKT/P21 signaling pathway and activation of p53 pathway in breast cancer." Biochemical and biophysical research communications (2018):74.
4. Shen, R, Wu T, Huang P, Shao Q, Chen M. "The clinicopathological significance of ubiquitin-conjugating enzyme E2C, leucine-rich repeated-containing G protein-coupled receptor, WW domain-containing oxidoreductase, and vasculogenic mimicry in invasive breast carcinoma." Medicine (2019).
5. Tang, J, Lu M, Cui Q, Zhang D, Kong D, Liao X, et al. "Overexpression of ASPM, CDC20, and TTK Confer a Poorer Prognosis in Breast Cancer Identified by Gene Co-expression Network Analysis." Frontiers in oncology (2019).
6. Li, G, Yang M, Zuo L, Wang MX. "MELK as a potential target to control cell proliferation in triple-negative breast cancer MDA-MB-231 cells." Oncology letters (2018):9934.