BBA Clinical 8 (2017) 56–65

Contents lists available at ScienceDirect

BBA Clinical

journal homepage: www.elsevier.com/locate/bbaclin

Signatures derived from increase in SHARPIN gene copy number are associated with poor prognosis in patients with breast cancer

Diane Ojo[a,b,c], Maryam Seliman[a,b,c,d], Damu Tang[a,b,c,*]

[a] Division of Nephrology, Department of Medicine, McMaster University, St. Joseph's Hospital, Hamilton, Ontario, Canada
[b] Father Sean O'Sullivan Research Institute, St. Joseph's Hospital, Hamilton, Ontario, Canada
[c] The Hamilton Center for Kidney Research, St. Joseph's Hospital, Hamilton, Ontario, Canada
[d] School of Medicine, National University of Ireland, Galway, Ireland

ARTICLE INFO

Keywords:
SIPL1/SHARPIN
Cancer
Breast cancer
Overall survival

ABSTRACT

We report three signatures produced from *SHARPIN* gene copy number increase (GCN-Increase) and their effects on patients with breast cancer (BC). In the Metabric dataset (n = 2059, cBioPortal), *SHARPIN* GCN-Increase occurs preferentially or mutual exclusively with mutations in *TP53*, *PIK3CA*, and *CDH1*. These genomic alterations constitute a signature (SigMut) that significantly correlates with reductions in overall survival (OS) in BC patients (n = 1980; p = 1.081e − 6). Additionally, *SHARPIN* GCN-Increase is associated with 4220 differentially expressed genes (DEGs). These DEGs are enriched in activation of the pathways regulating cell cycle progression, RNA transport, ribosome biosynthesis, DNA replication, and in downregulation of the pathways related to extracellular matrix. These DEGs are thus likely to facilitate the proliferation and metastasis of BC cells. Additionally, through forward (FWD) and backward (BWD) stepwise variate selections among the top 160 downregulated and top 200 upregulated DEGs using the Cox regression model, a 6-gene (SigFWD) and a 50-gene (SigBWD) signature were derived. Both signatures robustly associate with decreases in OS in BC patients within the Curtis (n = 1980; p = 6.16e − 11 for SigFWD; p = 1.06e − 10, for SigBWD) and TCGA cohort (n = 817; p = 4.53e − 4 for SigFWD and p = 0.00525 for SigBWD). After adjusting for known clinical factors, SigMut (HR 1.21, p = 0.0297), SigBWD (HR 1.25, p = 0.0263), and likely SigFWD (HR 1.17, p = 0.062) remain independent risk factors of BC deaths. Furthermore, the proportion of patients positive for these signatures is significantly increased in ER −, Her2-enriched, basal-like, and claudin-low BCs compared to ER + and luminal BCs. Collectively, these *SHARPIN* GCN-Increase-derived signatures may have clinical applications in management of patients with BC.

1. Introduction

Breast cancer (BC) is a leading cause of cancer death in women with approximately 1.7 million new cases and 500,000 fatalities annually [1]. Clinically, BC is classified based on the expression of estrogen receptor (ER) and Her2 and is categorized into ER +, Her2 +, or triple negative (TN; negative for ER, progesterone receptor (PR), and Her2) subclasses. Additionally, it is also classified according to gene expression profiles. Six intrinsic subtypes have been identified: luminal A and B (ER +), normal-like, Her2-enriched, basal-like, and claudin-low (the latter two being TN) BCs [2–5]. Furthermore, combination of gene copy number (GCN) variation and gene expression profile has divided BC into 10 integrative clustering sub-groups [6]. In addition to GCN variation, BC contains several common mutations including TP53, PIK3CA, GATA3, and CDH1 [7–9]. Despite this rich knowledge of BC tumorigenesis, our understanding on the etiology leading to BC tumorigenesis and progression remains limited.

SHARPIN (Shank-associated RH domain interacting protein) or SIPL1 (Shank-Interacting Protein-Like 1) is a Shank-binding protein in the postsynaptic density [10]. SHARPIN is also a major component of the linear ubiquitin chain assembly complex (LUBAC), an E3 ubiquitin-protein ligase complex that activates NF-κB signalling [11–14]. Additionally, SHARPIN/SIPL1 is physically associated with PTEN, resulting in PTEN inactivation [15,16]. In line with the important contributions of NF-κB signalling and PTEN inactivation in tumorigenesis [15,17], SHARPIN possesses multiple oncogenic activities in vitro, including suppression of apoptosis [18,19], enhancement of cell detachment and migration [20,21], and AKT activation [15]. In vivo, SHARPIN promotes the tumorigenesis of cervical cancer [15] and hepatocellular carcinoma [22]. In patients, upregulation of SHARPIN

* Corresponding author at: St. Joseph's Hospital, T3310, 50 Charlton Ave East, Hamilton, Ontario L8N 4A6, Canada.
E-mail address: damut@mcmaster.ca (D. Tang).

http://dx.doi.org/10.1016/j.bbacli.2017.07.004
Received 27 June 2017; Received in revised form 28 July 2017; Accepted 31 July 2017
Available online 24 August 2017

occurs in ovarian cancer, cervical cancer, liver cancer, and prostate cancer [15,20,22,23].

NF-κB signalling, PI3K activation, and PTEN inactivation are well demonstrated oncogenic events that occur during BC tumorigenesis [24]. Of note, upregulation of SHARPIN expression was detected in BC [16,25,26]. SHARPIN plays a role in BC metastasis [25] and inhibition of p53-mediated tumor suppression [26]. Elevations in SHARPIN expression modestly associate with reductions in overall survival (OS) in patients with breast cancer [16,25,26]. In supporting these observations, an increase in SHARPIN gene copy number (GCN) was also observed and this genomic alteration modestly correlates with shortening of OS in patients with BC [16].

To further examine the association of *SHARPIN* GCN increase with BC prognosis, we have taken a thorough in silico investigation of *SHARPIN* GCN increase, its-associated enrichment in mutations and gene expression, and the impact of these events on OS in patients with breast cancer. The two most comprehensive datasets, the Metabric (n = 2509) and TCGA-Cell (n = 817) cohorts within the cBioPortal database were used. We report here the identification of *SHARPIN* GCN increase-associated mutations of *TP53*, *PI3KCA*, *GATA3*, *CDH1*, *AKT*, and *ASXL1* and > 4000 differentially expressed genes (DEGs). These DEGs function in multiple aspects of cell proliferation and extracellular matrix processes. Furthermore, these enrichments in mutations and DEGs form three signatures that robustly associate with shortening of OS in patients with breast cancer.

2. Materials and methods

2.1. cBioPortal

The Metabric and TCGA-Cell 2015 (TCGA) datasets within cBioPortal [27,28] (http://www.cbioportal.org/index.do) contain 2509 and 815 patients with breast cancer, respectively. The Metabric dataset constitutes two sub-datasets with one being the Curtis [6] containing 1980 patients with a follow-up period up to 350 months (http://www.cbioportal.org/index.do). The TCGA dataset has 816 tumors (from 815 patients) with RNA sequencing and copy number variation data and a follow-up period of 300 months [29] (http://www.cbioportal.org/index.do).

2.2. Establishment of three SHARPING GCN increase-derived signatures: SigMut, SigFWD, and SigBWD

Enrichment data of mutations and gene expression (differentially expressed genes/DEGs) with respect to *SHARPIN* GCN increases were extracted from the Metabric dataset. The six enriched mutations (Table 1) were analyzed for contributions to hazard ratio (HR) using the Cox regression model (SPSS Statistics version 23). The top 160 down-regulated DEGs and top 200 upregulated DEGs were selected for their effects on HR by either forward (FWD) stepwise or backward (BWD) stepwise method, using the multivariate Cox regression model (SPSS Statistics version 23). These analyses resulted in three signatures: SigMut, SigWFD, and SigBWD.

2.3. Gene set and pathway enrichment analyses

The GAGE package in R was used to analyze DEGs for gene set enrichment within the KEGG (Kyoto Encyclopedia of Genes and Genomes) and GO (gene ontology) databases [30]. Pathway enrichment was performed using the Reactome package in R [31].

2.4. Statistical analysis

Fisher's exact test was performed using the GraphPad Prism 5 software. Kaplan-Meier surviving curves and log-rank test were carried out using the R survival package, SPSS Statistics version 23, and tools provided by cBioPortal. Multivariate Cox regression analysis was performed using the R survival package and SPSS Statistics version 23. A value of $p < 0.05$ is considered statistically significant.

3. Results

3.1. Increase in SHARPIN GCN associates with mutations in TP53, PIK3CA, CDH1, and others

The SHARPIN gene resides at 8q24.3 which is frequently amplified in breast cancers (BC) [32–37]. *SHARPIN* GCN increase was reported to correlate with a reduction in OS in patients with BC [16]. In accordance with these observations, we observed *SHARPIN* GCN amplification in 403 breast tumors among 1980 BCs in the Curtis sub-dataset [6] of the Metabric dataset within the cBioPortal database; the amplification modestly associates with a decrease in OS in BC patients [38] (see Fig. 1 in Ref [38]). This modest association is consistent with the consensus that individual factors have limited biomarker values and that combination of multiple factors is needed to yield a robust signature. To improve *SHARPIN* GCN amplification-derived association with the OS reduction, we have extracted the genomic events (mutations and gene copy number variation) from the Metabric dataset, which are enriched with *SHARPIN* GCN amplification. Since the 8q24 region is commonly amplified in BC, a large number of loci are co-amplified with *SHARPIN*, which is attributable to “the neighboring effect”. As such, their co-amplification may not contribute to the biomarker value of *SHARPIN* GCN increase. We thus focused on enrichment in gene mutations. With enrichment being defined at q value < 0.05, we have extracted a set of enriched mutations with either co-occurrence or mutual exclusiveness with *SHARPIN* GCN increase (Table 1). Interestingly, these enriched mutations include *TP53*, *PIK3CA*, *GATA3*, and *CDH1* (Table 1). These genes are frequently mutated in BC and play important roles in BC tumorigenesis [6,7,29,39]. Nonetheless, the potential role of ASXL1 (additional sex comb-like 1) in BC remains unclear.

Table 1
Co-alteration of mutations with *SHARPIN* GCN amplification.

Gene	locus	AMP + [a]	AMP − [a]	Log R [c]	p-Value	q-Value
TP53 [b]	17q13.1	232 (53.33%) [d]	509 (31.50%) [e]	0.76	9.27e − 17	1.60e − 14
PIK3CA [b]	3q26.3	136 (31.26%)	719 (44.49%)	− 0.51	3.36e − 7	2.91e − 15
GATA3 [b]	10p15	24 (5.52%)	217 (13.43%)	− 1.28	8.45e − 7	4.87e − 5
CDH1 [b]	16q22.1	23 (5.29%)	175 (10.83%)	− 1.03	1.75e − 4	7.56e − 3
AKT1 [b]	14q32.32	6 (1.38%)	80 (4.95%)	− 1.84	2.40e − 4	8.31e − 3
ASXL1 [b]	20q11	2 (0.46%)	44 (2.72%)	− 2.57	1.40e − 3	0.0403

[a] SHARPIN genomic amplification positive and negative.
[b] These mutations were co-altered with SHARPIN genomic amplification.
[c] Log2-based ratio of percentage in altered group/percentage in unchanged group; positive and negative ratios are for co-occurrence and mutual exclusiveness, respectively.
[d] Number of mutation cases/number of cases with SHARPIN amplification × 100.
[e] Number of mutation cases/number of cases without SHARPIN amplification × 100.