RESEARCH ARTICLE

Proteomic profiling of cancer stem cells derived from primary tumors of HER2/Neu transgenic mice

Deepak Kanojia[1], Weidong Zhou[2], Jiajia Zhang[3], Chunfa Jie[4], Pang-Kuo Lo[1], Qian Wang[5] and Hexin Chen[1]

[1] Department of Biological Science, Centre for Colon Cancer, University of South Carolina, Columbia, SC, USA
[2] Center for Applied Proteomics and Molecular Medicine, George Mason University, Manassas, VA, USA
[3] Arnold School of Public Health, Health Sciences Building, University of South Carolina, Columbia, SC, USA
[4] Department of Surgery, Transplant Surgery Division, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA
[5] Department of Chemistry and Biochemistry, University of South Carolina, Columbia, SC, USA

Received: March 7, 2012
Revised: July 9, 2012
Accepted: August 23, 2012

Human epidermal growth factor receptor 2 (HER2) overexpression leads to mammary tumorigenesis and its elevated levels lead to increase in cancer stem cells (CSCs), invasion, and metastasis. CSCs are resistant to radiation/chemotherapeutic drugs and are believed to be responsible for recurrence/relapse of cancer. CSCs are isolated using flow cytometry based sorting, although reliable, this technology hinders the convenient identification of molecular targets of CSCs. Therefore to understand the molecular players of increased CSC through HER2 overexpression and to develop meaningful targets for combination therapy, we isolated and characterized breast CSCs through convenient tumorsphere culture. We identified the altered protein expression in CSC as compared to non-CSC using LC-MS/MS and confirmed those results using qRT-PCR and Western blotting. Ferritin heavy chain 1 (FTH1) was identified as a candidate gene, which is involved in iron metabolism and iron depletion significantly decreased the self-renewal of CSCs. We further performed in silico analysis of altered genes in tumorsphere and identified a set of genes (PTMA, S100A4, S100A6, TNXRD1, COX-1, COX-2, KRT14, and FTH1), representing possible molecular targets, which in combination showed a promise to be used as prognostic markers for breast cancer.

Keywords:
Biomedicine / Breast cancer stem cells / FTH1 / HER2 / Tumorsphere

1 Introduction

Cancer comprises of a heterogeneous population of cells within which exist a subset of cells called the cancer stem cells (CSCs). These cells have higher tumorogenic ability and generate the original tumor heterogeneity when injected in non-obese diabetic/ severe combined immunodeficiency (NOD/SCID) mice. After their initial identification in breast cancer [1], a number of reports have highlighted its impact in various other solid tumors [2]. CSCs are resistant to radiation and chemotherapeutic drugs and are believed to be responsible for recurrence/relapse of cancer. Elimination of these cells through CSC-specific drugs will open avenues for eradication of cancer. Hence, a fundamental understanding of differences between CSC and non-CSC warrants investigation.

Human epidermal growth factor receptor 2 (HER2) is a tyrosine kinase receptor, which is overexpressed in ~30% of breast cancer patients. Its overexpression leads to mammary tumorigenesis and its elevated levels lead to increased invasion, metastasis, and CSC population [3]. The molecular mechanism that leads to increase in CSC through HER2 overexpression is not entirely understood. Breast cancer patients with overexpression of HER2 are at a high risk of recurrence [4]. Anti-HER2 drugs such as trastuzumab and lapatinib are in clinics for breast cancer treatment, however

Correspondence: Dr. Hexin Chen, Department of Biological Sciences, University of South Carolina, 715 Sumter Street, PSC621, Columbia, SC 29205, USA
E-mail: hchen@biol.sc.edu
Fax: +1 803 777 4002

Abbreviations: CSCs, cancer stem cells; **DFO**, deferoxamine mesylate; **FTH1**, ferritin heavy chain; **HER2**, human epidermal growth factor receptor 2; **TICs**, tumor initiating cells

resistant tumors progress, which further increases the risk of recurrence [4]. Hence, it is necessary to target the breast cancer with HER2 overexpression with a combination of drugs, which can target both the sensitive and resistant cell types.

Increase in CSCs/TICs (where TICs is tumor initiating cells) through overexpression of HER2 has been documented through recent reports [3, 5, 7]. Therefore to understand the increased CSC through HER2 overexpression and to identify novel targets, we isolated and characterized breast CSCs from primary cultures of spontaneous tumors from of HER2/Neu transgenic mice. We identified the altered protein expression in CSC as compared to non-CSC using LC-MS/MS and confirmed those results using quantitative RT-PCR and Western blotting. We further performed in silico analysis and identified a set of genes (PTMA, S100A4, S100A6, TNXRD1, COX-1, COX-2, KRT14, and FTH1 [where FTH1 is ferritin heavy chain]), which are possible target genes and in combination they showed a promise to be used as prognostic markers for breast cancer.

2 Materials and methods

2.1 Isolation of primary tumor cells

Primary tumor cells were isolated as described previously [6–8]. Briefly, spontaneous mammary tumors from FVB/N-Tg(MMTVneu)202Mul/J (MMTV-Neu) transgenic mice were minced and digested in DMEM/F12 medium (Invitrogen, CA, USA) with collagenase and hyaluronidase. Single cells obtained were transferred into DMEM with 10% FBS and 10 μg/mL insulin (Sigma, MO, USA) to grow adherent monolayer of parental cells.

2.2 Isolation of tumorspheres

Adherent cells were dissociated by 0.25% trypsin-EDTA (Invitrogen, NY, USA) and suspended in serum-free medium [DMEM/F12, 20 ng/mL EGF (Sigma), 10 ng/mL bFGF (Sigma), 0.4% BSA (Sigma), 5 μg/mL insulin (Sigma) and B27 supplement (Invitrogen)] in Petri dish for 6 days to generate tumorospheres [6, 7].

2.3 In vivo tumorigenecity assay for tumorspheres and parental cells

MMTV-Neu transgenic mice were housed in Animal Resource Facilities (University of South Carolina). Adherent tumor cells and tumorspheres were dissociated using 0.25% trypsin and washed with PBS at 4°C. Cells were injected into the fourth pair of mammary glands of 6∼8 week old HER2/Neu transgenic. Upon tumor onset, length and width of palpable tumors were measured and tumor volume was calculated [7]. All the animal protocols were approved by University of South Carolina.

2.4 Trypsin digestion and desalting

Trypsin digestion and desalting was performed as described previously [9]. Briefly, the cells were resuspended in 8 M urea and protein concentration was measured by Bradford Assay (BioRad, CA, USA). After reducing and alkylating the proteins, samples were digested by trypsin at 37°C for 6 h. The digestion mixture was then acidified by adding glacial acetic acid and desalted by ZipTip (Millipore, MA, USA).

2.5 MS for peptide identification

The peptides were analyzed by highly sensitive RP LC coupled nanospray MS/MS (LC-MS/MS) using an LTQ-Orbitrap mass spectrometer (Thermo Fisher, FL, USA) as described previously [9]. The reversed-phase LC column was slurry packed in house with 5 m, 200 Å pore size C18 resin (Michrom BioResources, CA, USA) in a 100-m id 10-cm-long piece of fused silica capillary (Polymicro Technologies, AZ, USA) with a laser-pulled tip. After sample injection, the column was washed for 5 min with mobile phase A (0.1% formic acid), and peptides were eluted using a linear gradient of 0% mobile phase B (0.1% formic acid, 80% ACN) to 50% B in 120 min at 200 nL/min, then to 100% B in an additional 10 min for the proteomics analysis. The LTQ-Orbitrap mass spectrometer was operated in a data-dependent mode in which each full MS scan (60 000 resolving power) was followed by eight MS/MS scans where the eight most abundant molecular ions were dynamically selected and fragmented by CID using a normalized collision energy of 35%. The dynamic exclusion time was 30 s, and the dynamic exclusion size was 200. The "FT master scan preview mode," "Charge state screening," "Monoisotopic precursor selection," and "Charge state rejection" were enabled so that only the 1^+, 2^+, and 3^+ ions were selected and fragmented by CID.

2.6 MS data analysis

Tandem mass spectra collected by Xcalibur (version 2.0.2) were searched against the NCBI human protein database using SEQUEST (Bioworks software from ThermoFisher, version 3.3.1) with full tryptic cleavage constraints, static cysteine alkylation by iodoacetamide, and variable methionine oxidation. Mass tolerance for precursor ions was 5 ppm and mass tolerance for fragment ions was 0.25 Da. The SEQUEST search results of proteomics data were filtered by the criteria "Xcorr versus charge 1.9, 2.2, 3.0 for 1^+, 2^+, 3^+ ions; $\Delta Cn > 0.1$; probability of randomized identification of peptide <0.01." Confident peptide identifications were determined using these stringent filter criteria for database match scoring followed by manual evaluation of the results. The "false discovery rate (FDR)" was estimated by searching a combined forward-reversed database as described by Elias

Table 1. Eight-gene signature overexpressed in tumorspheres

No.	Gene name	Gene symbol	Accession no.
1	Ferritin heavy chain 1	FTH-1	6753912
2	Prostaglandin endoperoxide synthase-1	PTGS1	6679537
3	Prostaglandin endoperoxide synthase-2	PTGS2	31981525
4	Prothymosin alpha	PTMA	7110705
5	S100 Calcium binding protein A4	S100A4	33859624
6	S100 Calcium binding protein A6	S100A6	6755392
7	Thioredoxin reductase 1	TXNRD1	13569841
8	Keratin 14	KRT14	13097093

and Gygi [10]. The SEQUEST search results were exported to Excel files and compared.

2.7 Quantitative real-time RT-PCR analysis and Western blotting

Total RNA was isolated was reverse-transcribed using M-MLV RT (Invitrogen) and quantitative real-time RT-PCR was performed using RT^2 Fast SYBR® Green/ROXTM qPCR Master Mix (SABiosciences, MD, USA) and data were analyzed as described previously [7]. GAPDH expression value was used as internal reference control.

Western blotting was performed as described previously [7]. Antibodies used were PTGS2 (Cell signaling, MA, USA), FTH1 (Cell signaling), GAPDH (Santacruz, CA, USA), and G6PD (Cell signaling).

2.8 Deferoxamine mesylate (DFO) treatment

Adherent cells were seeded at a density of 8×10^5 cells in 100-mm tissue culture dish for 24 h. The cells were treated with 50 and 100 μM DFO (Calbiochem, CA, USA) for 2 days. Following partial trypsin digestion (0.25% Invitrogen) for 1–2 min at 37°C and vigorously tapping the culture dish against the palm of your hand to dissociate the cells, single cells were isolated and tumorsphere formation was performed using 5000 cells in 35-mm suspension culture dishes. Primary tumorspheres were digested with Accutase (Innovative Cell Technologies, San Diego, CA, USA) and seeded for secondary tumorsphere cultures, with and without DFO.

2.9 In silico statistical analysis

To comprehend the clinical relevance of altered proteins, we used a list of eight genes from Table 1 to perform a statistical analysis of published microarray data. After translating the signature genes into UnigeneIDs, we extracted the gene expression information from published datasets of breast cancer patients [11–13]. A hierarchical cluster analysis, which determines clusters by the minimum variance method, is applied to divide the patients into two clusters [8, 14, 15]. The association between the different clusters and survival time are estimated through the Kaplan–Meier method. We define the cluster with the high survival probability as the low-risk group and the low survival probability as the high-risk group. The *p*-values were calculated using the logrank test, and differences were considered statistically significant at $p < 0.05$.

3 Results

3.1 Tumorspheres derived from adherent tumor cells of HER2/Neu Mice are enriched in breast CSCs

A number of studies have identified and characterized the breast TICs from tumors and cancer cell lines using tumorsphere/mamosphere culture system [16, 17]. To enrich breast CSCs, we isolated cells from primary tumors of MMTV-Neu transgenic mice and propagated them in adherent cell culture, which was employed to induce differentiation [16, 18]. We then generated primary tumorspheres by culturing them for 6 days in suspension culture medium (Fig. 1A). Both the adherent culture cells and tumorspheres were injected into the mammary gland of MMTV-Neu mice where 50 cells from tumorspheres were sufficient to generate tumors whereas 50/500 adherent cells were unable to generate tumors in syngeneic MMTV-Neu mice (Fig. 1C). Moreover the tumor latency for tumorspheres was much shorter as compared to adherent tumor cells. Lung metastasis was also observed in tumorsphere injected mice, which were apparently absent in the mice injected with adherent cells (Fig. 1C). Moreover to establish the CSC nature of these tumorspheres, we injected the single cells obtained from primary tumors, which generated tumors when injected as low as 50 cells. These results clearly reveal that the tumorspheres derived from primary cells of HER2/Neu-induced tumors are enriched in breast CSCs. We employed syngeneic mice (MMTV-Neu) for our transplantation experiments as it represents the correct genetic background and the right environment required for the growth of CSCs and in turn their proper identification.

3.2 Identification of differentially expressed proteins in tumorspheres using LC-MS/MS, and functional analysis of genes altered in tumorspheres

To identify differentially expressed proteins in breast CSCs, proteins from adherent cells and tumorspheres were digested with trypsin and extracted peptides were subjected to LC-MS/MS using LTQ-Orbitrap. The SEQUEST search results were filtered by stringent criteria which yielded 2276 proteins

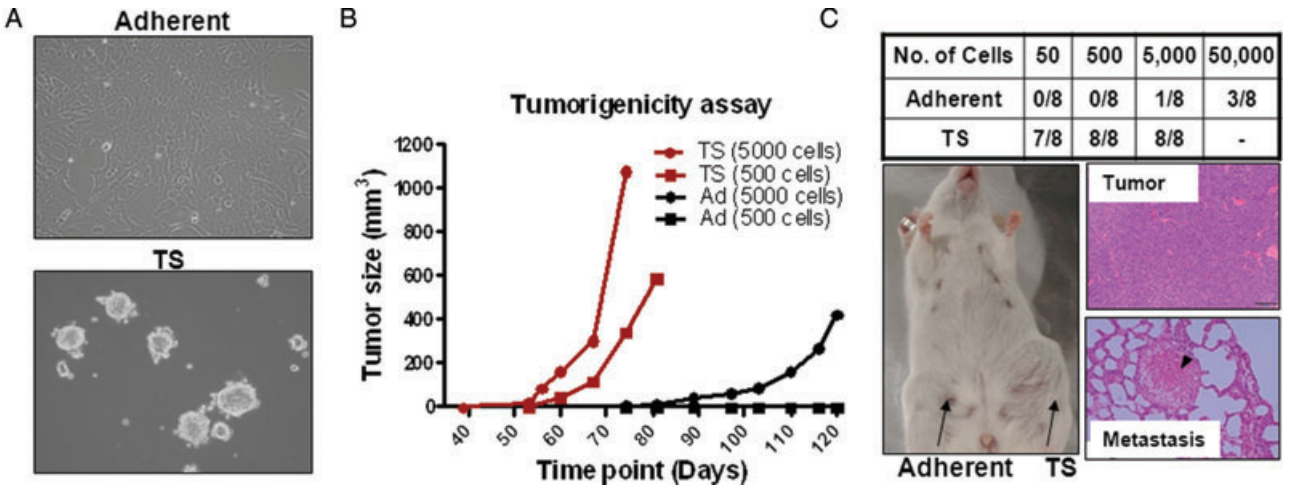


Figure 1. Tumorspheres derived from adherent HER2/Neu tumor cells are enriched in breast cancer stem cells. (A) Morphology of adherent cells and tumorspheres. (B) Tumorspheres exhibit early tumor onset. A total of 5000 and 500 cells were injected to mammary glands of syngeneic mice and the tumor onset and growth was monitored for 120 days. (C) The upper panel depicts tumor incidence of tumorspheres and adherent cells. Arrows in the left lower panel indicate the tumors that arose from adherent and tumorspheres, right panel indicates histology of tumor, and arrow heads indicate lung metastasis. (E). Ad, adherent cells; TS, tumorspheres.

(including homologous isoforms) from adherent cells and 2182 proteins (including homologous isoforms) from tumorspheres with 1% FDR (Supporting Information Table S2). On the basis of MS spectra count (Supporting Information Table S3), we performed relative quantitative analysis of these identified proteins and identified dozens of differentially expressed proteins (Supporting Information Table S4). One of the significantly overexpressed protein in the tumorsphere was FTH1 (Supporting Information Fig. 1) whose CID spectrum is depicted in Fig. 2.

We analyzed the altered proteins in tumorspheres and found alterations in protein synthesis and ribosomal components, metabolism-related proteins, cytoskeleton remodeling proteins, redox regulators, calcium-binding proteins, and proteins of ubiquitin proteasome system (Supporting Information Table S5). We further discovered alterations in eukaryotic initiation factor 2 (eIF2) signaling in tumorspheres, where eIF3B, eIF3C/eIF3C1, RPS25, RPS27, RPS27L were found to be decreased and RPL28, RPL34, RPL23A, RPL23A were found to be upregulated. We also noted alterations in proteins involved in oxidative stress (FTH1, PTGS1, PTGS2, S100A6, TXNRD1, XDH, TRAP1, G6PD, and CAT) and free radical scavengers (PRDX5, SOD3, and XDH) (Supporting Information Table S5).

3.3 RT- PCR, Western blotting, and quantitative RT-PCR based confirmation of genes altered in tumorspheres

Through proteomic screening we identified differentially expressed proteins in tumorspheres. Further to understand their transcript levels we performed RT-PCR of 20 different genes (FTH1, PTGS-1, PTGS2, S100A6, TNXRD1, PTMA, RPS27, RPS27L, G6PD, PCK2, GLG1, PCD6IP, RPL34, RPL28, G6PDX, CYB5R3, eIF3B, IDH1, KRT14, and eIF3D) with GAPDH as internal reference control. The upregulated genes were selected based on their association with cancer and the downregulated genes were selected based on the spectral count to validate the proteomics screening. The results indicate that using semi-quantitative RT-PCR, FTH1, IDH1, TNXRD1, PTMA, NADH, PTGS2, S100A4, GLG1, PTGS-2, S100A6, PTGS1 were overexpressed and PCK2, RPS27L, RPS27, eIF3B, PCD6IP, G6PD, and eIF3D were downregulated in tumorspheres, validating the proteomics results (Fig. 3A and B). Further RPL34, CYB5R3, and RPL28 were not found to be altered using semi-quantitative RT-PCR. These can be explained by the fact that many proteins are overexpressed without the increase in their transcript levels [19]. To further determine the precise levels of gene expression, we performed qRT-PCR studies and found that GLG1, RPL34, FTH1, PTGS-1, PTGS2, S100A6, S100A4, KRT14, PTMA, and TXNRD1 were overexpressed, whereas RPS27, RPS27L, G6PD, and PCK2 were downregulated in tumorspheres (Fig. 3 D and E).

We also performed Western blotting experiments to elucidate the protein levels of FTH1, COX2, G6PD, and GAPDH and found that FTH1 and COX2 were overexpressed whereas G6PD was downregulated in tumorspheres (Fig. 3C), substantiating the proteomics screening.

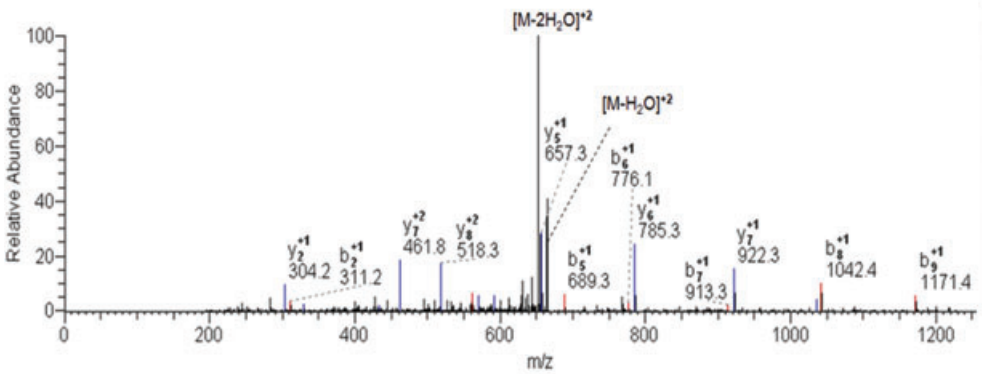
3.4 Deferoxamine mesylate decreases tumorsphere formation efficiency

In the proteomics screening, we identified high levels of FTH1 expression both at protein and mRNA levels. FTH1 is

A

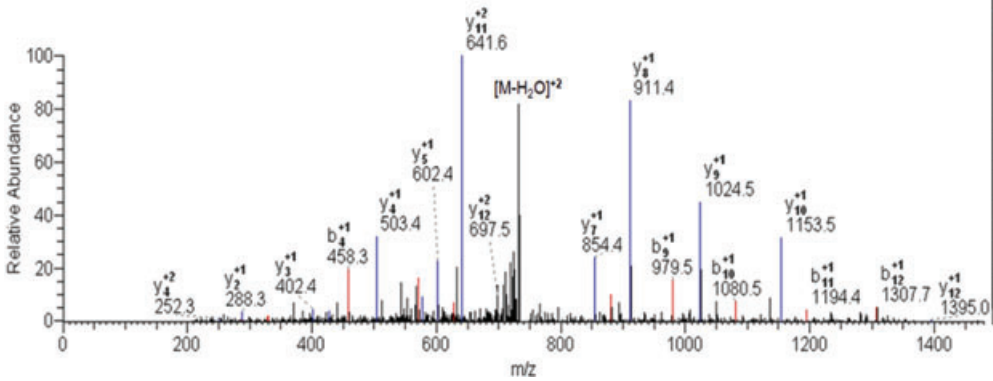
	1-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80	81-90	91-100
1	MTTASPSQVR	QNYHQDAEAA	INRQINLELY	ASYVYLSMSC	YFDRDDVALK	NFAKYFLHQS	HEEREHAEKL	MKLQNQRGGR	IFLQDIKKPD	RDDWESGLNA
101	MECALHLEKS	VNQSLLELHK	LATDKNDPHL	CDFIETYYLS	EQVKSIKELG	DHVTNLRKMG	APEAGMAEYL	FDKHTLGHGD	ES	

B



	AA	B	Y	
1	Y	164.07060	-	10
2	F	311.13902	1182.56505	9
3	L	424.22308	1035.49663	8
4	H	561.28199	922.41257	7
5	Q	689.34057	785.35366	6
6	S	776.37260	657.29508	5
7	H	913.43151	570.26305	4
8	E	1042.47411	433.20414	3
9	E	1171.51670	304.16155	2
10	R	-	175.11895	1

C



	AA	B	Y	
1	S	88.03930	-	13
2	I	201.12337	1394.77504	12
3	K	329.21833	1281.69097	11
4	E	458.26092	1153.59601	10
5	L	571.34499	1024.55342	9
6	G	628.36645	911.46935	8
7	D	743.39340	854.44789	7
8	H	880.45231	739.42095	6
9	V	979.52072	602.36204	5
10	T	1080.56840	503.29362	4
11	N	1194.61133	402.24594	3
12	L	1307.69539	288.20302	2
13	R	-	175.11895	1

Figure 2. Identification of ferritin heavy chain 1 (FTH1) by LC-MS/MS. (A) The primary sequence of FTH1 (accession number 6753912) is presented with identified peptides highlighted in red. A total of 11 MS2 spectra (corresponding to ten unique full tryptic peptides, 60.4% coverage by amino acid residues) were matched to FTH1 in tumorspheres, but zero in adherent cells; (B) example of CID spectrum of the identified peptide YFLHQSHEER (2^+ ion *m/z* 673.3167) from tumorspheres; (C) CID spectrum of the identified peptide SIKELGDHVTNLR (2^+ ion *m/z* 741.4063) from tumorspheres. The spectrum (left panel) is labeled to show b ions, y ions, and neutral loss of water from parent ions. The right panel is the table of the fragment assignments of the peptide, in which the matched b ions are red and y ions blue.

an iron-binding protein and is involved in regulation of iron metabolism, cellular proliferation, and apoptosis [20]. Excess iron leads to carcinogenesis in animal studies and epidemiological studies demonstrate that decrease in iron through phlebotomy decrease the risk of cancer in supposedly normal human population [21]. Moreover iron deprivation has been shown to suppress hepatocellular carcinoma in mouse models [22].

Given the fact that FTH1 is an iron-binding protein, we altered the balance of iron in the cells through iron chelator DFO. The cells were treated for 2 days with different concentrations of DFO and seeded for tumorsphere formation (Fig. 4A). We found iron depletion of adherent HER2/Neu tumor cells significantly decreases the viability in cell culture and dramatically reduces their ability to form tumorspheres (Fig. 4B). Further to test if DFO has the ability to directly inhibit the formation of tumorsphere, cells cultured in suspension medium were treated with DFO and it was found that DFO decreased/inhibited the growth of tumorspheres (Fig. 4A). Furthermore, we generated secondary tumorspheres from the DFO-treated primary tumorspheres. We found that cells treated with DFO had significantly reduced ability to form secondary tumorspheres (Fig. 4A and B).These results indicate that DFO treatment depletes iron and decreases the number of breast CSCs.

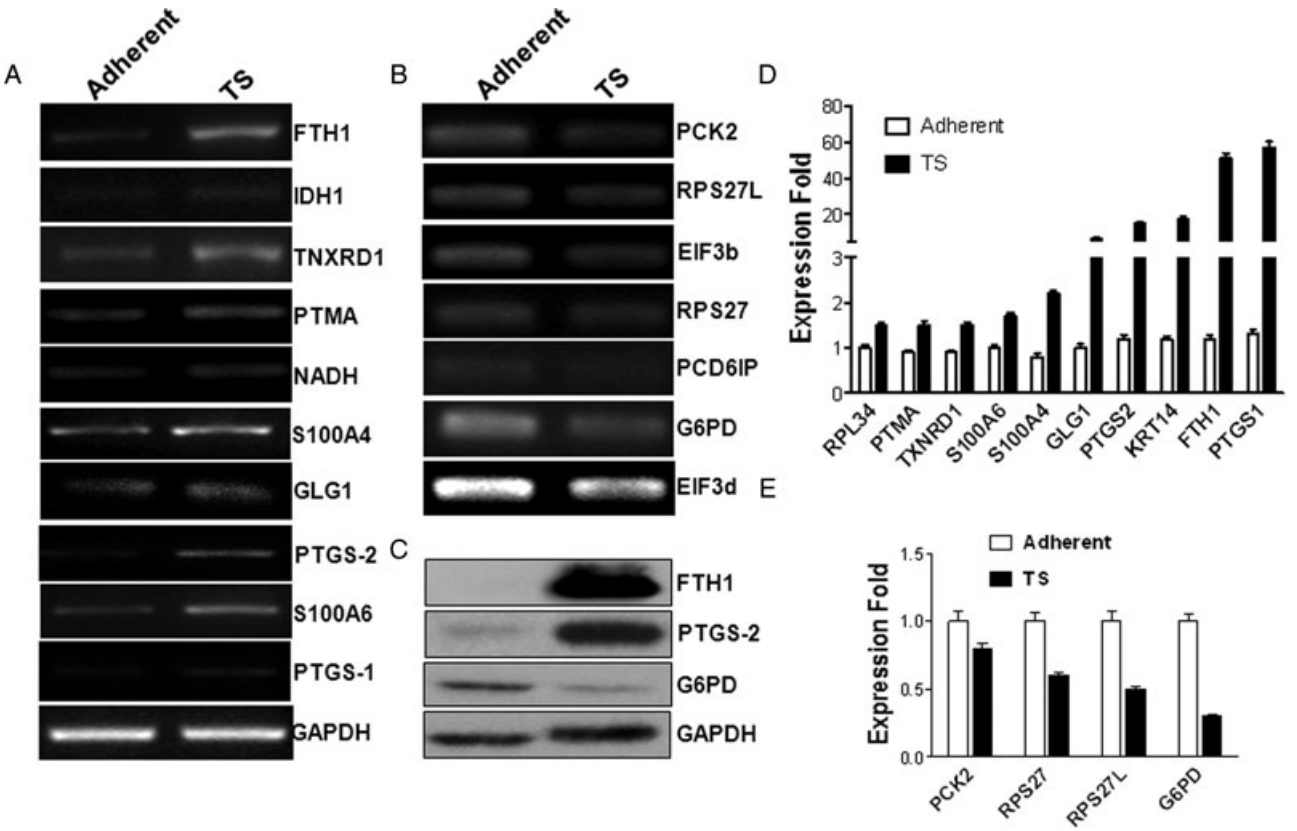


Figure 3. Confirmation of identified alterations using RT-PCR, qRT-PCR, and Western blotting. (A) RT-PCR demonstrating upregulation of genes, (B) RT-PCR demonstrating downregulation of genes, (C) Western blot analysis showing downregulation of G6PD and upregulation of FTH1 and PTGS2. (D) qRT-PCR analysis showing upregulation of genes. (E) qRT-PCR analysis showing downregulation of genes. The mean and SE of each qRT-PCR dataset were derived from triplicate experiments. TS, tumorspheres.

3.5 In silico analysis of potential markers for breast cancer prognosis

To understand the clinical importance of the identified alterations specifically in enriched breast CSCs, we performed statistical analysis using supervised principle component method to assess the performance of eight genes using published microarray data (Table 1). Regression model was derived to predict the survival probability of breast cancer patient. Statistical significance was not obtained using dataset in Wang et al. [11] (Fig. 4C). However that eight-gene signature could predict the survival probability of breast cancer patients with statistical confidence in the dataset published by van de Vijver et al. [12] (Fig. 4D). When the patients were further divided into HER2-positive and HER2-negative groups, we found that the eight-gene signature was a strong prognostic biomarker for relapse-free survival rate in the HER2-positive patients ($p = 0.0137$) (Fig. 4E). In the HER2-negative patients, the clinical outcome between the high- and low-risk subgroups was not significantly different. These data suggested that this gene signature can serve as a prognostic marker to identify a subgroup of HER2-positive patients with worse clinical outcome.

4 Discussion

Proteomics of CSC and non-CSC holds the potential to differentiate the two cell types and such technology can provide novel targets for cancer therapy. Conventionally scientists have isolated CSCs/non-CSCs from cell lines/primary tumor cells, based on the cell surface expression such as CD24/CD44, CD133, ALDH+ cells activity and side population studies [23]. These studies have shortcomings such as (1) it demands culture of a large number of cells, (2) it involves flow cytometry based separation that inflicts damage to the cells, and (3) low yield of CSC are obtained. Tumorspheres have been cultured from cancer cells [16, 17] and appears to be a practical approach to enrich the TICs. We identified CSCs from HER2/Neu transgenic mice using tumorsphere cultures where 50 cells were able to generate tumors in syngeneic HER2/Neu with lung metastasis (Fig. 1B and C) demonstrating the powerful methodology that can be used for studying the CSCs.

We identified 2709 differentially expressed peptides through proteomic profiling (Supporting Information Table S2) and within these we filtered 188 proteins based on identification of fair difference in the spectra counts (Supporting

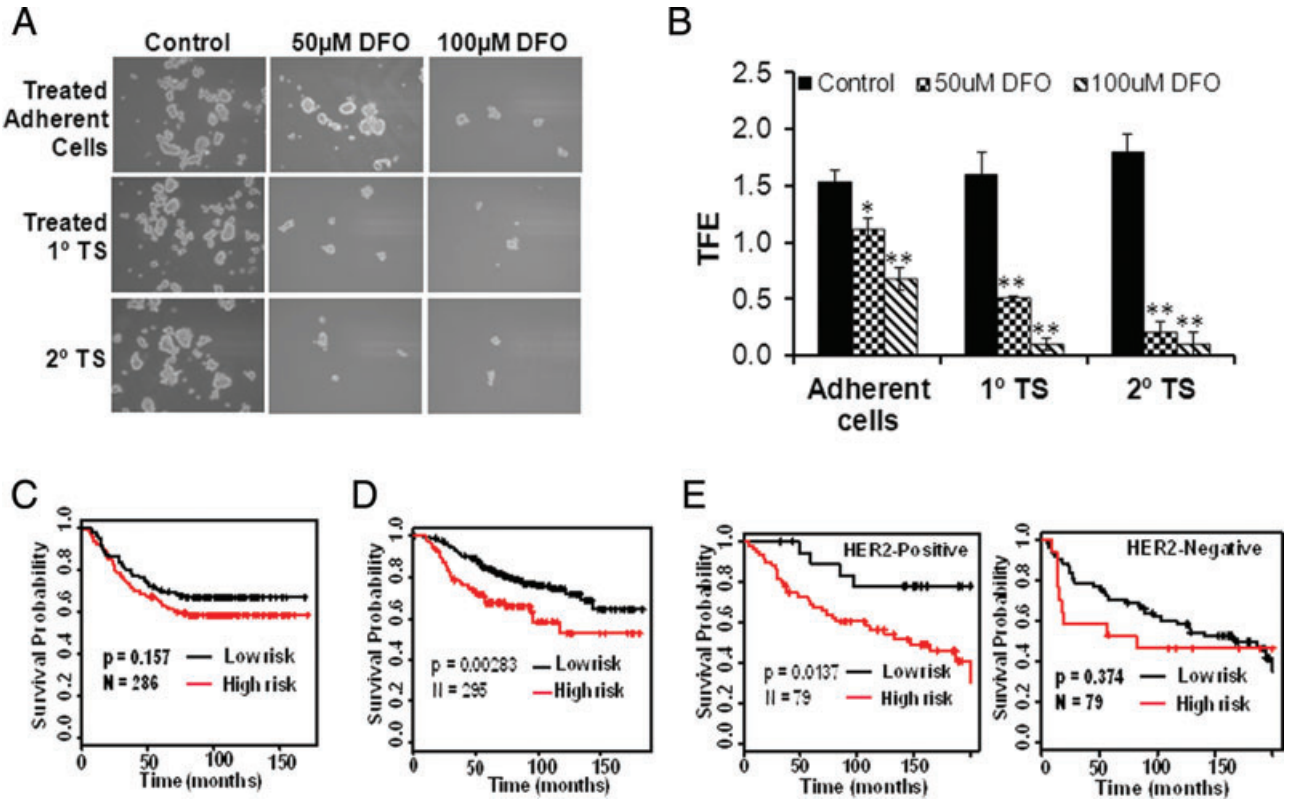


Figure 4. Deferoxamine mesylate (DFO) treatment decreases tumorsphere formation efficiency (TFE). (A) The upper panel shows tumorspheres derived from adherent control tumor cells and cells treated with 50 and 100 μM of DFO. The middle panel shows primary tumorspheres cultured in the presence of DFO for 6 days. The lower panel shows the secondary tumorspheres derived from the primary tumorspheres with pretreatment of DFO. 1° TS, primary tumorsphere; 2° TS, secondary tumorsphere. (B) TFE of DFO-treated adherent tumor cells and primary tumorspheres. Note: $*p < 0.05$; $**p < 0.01$; TFE, tumorsphere formation efficiency. The mean and SE of control and DFO-treated cells were derived from triplicate experiments. (C) Prediction of clinical outcome based on eight-gene signature. Kaplan–Meier analysis was performed to predict the survival probability of breast cancer patients using published microarray datasets. Microarray dataset from van de Vijver group [12]. A total of 295 patients had stage I or II breast cancer, 151 had lymph-node-negative disease and 144 had lymph-node-positive disease. *y*-axis represents survival probability and *x*-axis represents the time in months. (D) Microarray dataset from Wang group [11]. This dataset includes 286 lymph-node-negative breast cancer patients who received no adjuvant treatment when clinical samples were collected. TS, tumorspheres. (E) Prediction of clinical outcome in the HER2-positive and -negative patients [13]. The patients were statistically stratified into HER2-positive and -negative subgroups according to the expression status of HER2. The *y*-axis represents the relapse-free survival probability.

Information Tables S3 and S4). The differentially expressed proteins were analyzed manually and alterations in genes involved in translation, initiation, elongation, and termination process in tumorspheres were observed (Supporting Information Table S5). We also observed alteration in genes involved in reactive oxygen species (ROS) regulation (FTH1, PTGS1, PTGS2, S100A6, TXNRD1, XDH, TRAP1, G6PD, CAT, PRDX5, SOD3, and XDH).

In cells, the ROS are kept in check by free radical scavenging system and free radical scavengers, such as PRDX5, SOD3, XDH, were found to be upregulated in tumorspheres. This corroborates the earlier report that the breast CSCs have lower levels of ROS partly because of overexpression of these free radical scavengers [24].

We observed downregulation of a number of genes associated with eIF2 signaling (eIF3B, eIF3C/eIF3C1, RPS25, RPS27, and RPS27L). In the past, cancer cells have been treated based on high proliferative ability, which has been exploited for cancer treatment, however CSCs have been shown to be slow cycling [25], which renders them to be resistant to chemotherapeutic drugs. eIF signaling is also implicated in treatment of cancer [26], however in this proteomics screen, eIF2 signaling molecules are downregulated indicating that putative slow cycling cells in the tumorsphere culture may be resistant to inhibitors of eIF2 signaling.

PTGS-2 [27], TXNRD1 [28], and KRT14 [29] genes have already been shown to be associated with prognosis of breast cancer, except FTH1. Also there are no reports that document the role of FTH1, PTMA, S100A6, S100A4, and TNXRD1 in CSCs, which warrants investigation. Also higher expression of this set of genes is associated with poor survival of breast cancer patients implicates higher percentage of CSC

population in these patients and that it identifies high-risk population.

FTH1 was found to be overexpressed in tumorspheres (Fig. 3A and C), which is involved in iron metabolism. It has been shown to play an anti-apoptotic role and its up-regulation by NFκB inhibits TNF-α-induced apoptosis [20]. FTH1 overexpression in erythroid cells induces multidrug resistance [30] and depletion of iron decreases the expression of multidrug resistant gene-1 hence iron depletion might sensitize cells to common chemotherapeutic drugs [31]. Moreover, when FTH1 was silenced through liposomal-mediated siRNA delivery, it increased the sensitivity of cancer cells to chemotherapeutic drugs [32]. We found over-expression of FTH1 both at protein and mRNA levels in tumorspheres.

DFO is an iron chelator whose treatment to breast cancer cells induces apoptosis both in vitro and in vivo [33]. DFO treatment to gastrointestinal cancer cells decreases the expression of CD133, which implicates its role in CSCs [34]. We observed significant decrease in tumorsphere formation in DFO pretreated cells (Fig. 4A and B), which indicates that iron chelator DFO targets the CSC self-renewal and survival.

We found that eight-gene signature (FTH1, KRT14, PTGS-1, PTGS2, PTMA, S100A6, S100A4, and TXNRD1) could predict the survival probability of breast cancer patients with statistical confidence in the dataset published by van de Vijver et al. [12] (Fig. 4D). However, statistical significance was not obtained using dataset in Wang et al. [11] (Fig. 4C) possibly because of difference in the patient population, patient selection (>40% of patients were of age > 56 years), and array platform. Further, we have identified only a limited set of genes that do not have a large amount of overlap with other studies reported earlier on CSC [35] possibly because of different methodologies and that this study identifies HER2-specific genes involved in tumorsphere formation and CSC enrichment.

In summary, through proteomic profiling we identified molecular targets of CSCs and a set of proteins, which can predict the survival of breast cancer patients. Iron metabolism is important for survival of CSCs and iron chelator reduces the self-renewal ability of CSCs. The eight-gene signature in CSCs was a good predictor of survival of breast cancer patients, some of them have already been shown to have a role in tumor progression and metastasis. Hence, we believe that establishment of small antibody based array can be developed to assess the prognostic value of breast cancer patients, which can assist in modifying treatment regimen of high-risk breast cancer patients with CSC-specific inhibitors.

This work was supported by the ASPIRE-I grant from the Office of the Vice President for Research at USC, the Elsa U Pardee Cancer Foundation grant (B94AFFAA), the American Cancer Society Research Award (RSG-10–067-01-TBE) to H.C.

The authors have declared no conflict of interest.

5 References

[1] Al-Hajj, M., Wicha, M. S., Benito-Hernandez, A., Morrison, S. J. et al., Prospective identification of tumorigenic breast cancer cells. *Proc. Natl. Acad. Sci. USA* 2003, *100*, 3983–3988.

[2] Lobo, N. A., Shimono, Y., Qian, D., Clarke, M. F., The biology of cancer stem cells. *Ann. Rev. Cell Dev. Biol.* 2007, *23*, 675–699.

[3] Korkaya, H., Paulson, A., Iovino, F., Wicha, M. S., HER2 regulates the mammary stem/progenitor cell population driving tumorigenesis and invasion. *Oncogene* 2008, *27*, 6120–6130.

[4] Jones, K. L., Buzdar, A. U., Evolving novel anti-HER2 strategies. *Lancet Oncol.* 2009, *10*, 1179–1187.

[5] Magnifico, A., Albano, L., Campaner, S., Delia, D. et al., Tumor-initiating cells of HER2-positive carcinoma cell lines express the highest oncoprotein levels and are sensitive to trastuzumab. *Clin. Cancer Res* 2009, *15*, 2010–2021.

[6] Gu, Y., Fu, J., Lo, P., Wang, S. et al., The effect of B27 supplement on promoting in vitro propagation of Her2/neu-transformed mammary tumorspheres. *J. Biotech. Res.* 2011, *3*, 7–18.

[7] Lo, P. K., Kanojia D., Liu X., Singh U. P. et al., CD49f and CD61 identify Her2/neu-induced mammary tumor initiating cells that are potentially derived from luminal progenitors and maintained by the integrin-TGF beta signaling. *Oncogene* 2012, *24*, 2614–2626.

[8] Chen, H., Pimienta, G., Gu, Y., Sun, X. et al., Proteomic characterization of Her2/neu-overexpressing breast cancer cells. *Proteomics* 2010, *10*, 3800–3810.

[9] Zhou, W., Capello, M., Fredolini, C., Piemonti, L. et al., Proteomic analysis of pancreatic ductal adenocarcinoma cells reveals metabolic alterations. *J. Proteome Res.* 2011, *10*, 1944–1952.

[10] Elias, J. E., Gygi, S. P., Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat. Methods* 2007, *4*, 207–214.

[11] Wang, Y., Klijn, J. G., Zhang, Y., Sieuwerts, A. M. et al., Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet* 2005, *365*, 671–679.

[12] van de Vijver, M. J., He, Y. D., van't Veer, L. J., Dai, H. et al., A gene-expression signature as a predictor of survival in breast cancer. *N. Engl. J. Med.* 2002, *347*, 1999–2009.

[13] Desmedt, C., Piette, F., Loi, S., Wang, Y. et al., Strong time dependence of the 76-gene prognostic signature for node-negative breast cancer patients in the TRANSBIG multicenter independent validation series. *Clin. Cancer Res* 2007, *13*, 3207–3214.

[14] Chen, H., Lee, J. S., Liang, X., Zhang, H. et al., Hoxb7 inhibits transgenic HER-2/neu-induced mouse mammary tumor onset but promotes progression and lung metastasis. *Cancer Res.* 2008, *68*, 3637–3644.

[15] Gong, Y., Yan, K., Lin, F., Anderson, K. et al., Determination of oestrogen-receptor status and ERBB2 status of breast carcinoma: a gene-expression profiling study. *Lancet Oncol.* 2007, *8*, 203–211.

[16] Ponti, D., Costa, A., Zaffaroni, N., Pratesi, G. et al., Isolation and in vitro propagation of tumorigenic breast cancer cells with stem/progenitor cell properties. *Cancer Res.* 2005, *65*, 5506–5511.

[17] Cariati, M., Naderi, A., Brown, J. P., Smalley, M. J. et al., Alpha-6 integrin is necessary for the tumourigenicity of a stem cell-like subpopulation within the MCF7 breast cancer cell line. *Int. J. Cancer* 2008, *122*, 298–304.

[18] Vermeulen, L., Todaro, M., de Sousa Mello, F., Sprick, M. R. et al., Single-cell cloning of colon cancer stem cells reveals a multi-lineage differentiation capacity. *Proc. Natl. Acad. Sci. USA* 2008, *105*, 13427–13432.

[19] Foss, E. J., Radulovic, D., Shaffer, S. A., Ruderfer, D. M. et al., Genetic basis of proteome variation in yeast. *Nat. Genet.* 2007, *39*, 1369–1375.

[20] Pham, C. G., Bubici, C., Zazzeroni, F., Papa, S. et al., Ferritin heavy chain upregulation by NF-kappaB inhibits TNFalpha-induced apoptosis by suppressing reactive oxygen species. *Cell* 2004, *119*, 529–542.

[21] Zacharski, L. R., Chow, B. K., Howes, P. S., Shamayeva, G. et al., Decreased cancer risk after iron reduction in patients with peripheral arterial disease: results from a randomized trial. *J. Natl. Cancer Inst.* 2008, *100*, 996–1002.

[22] Ba, Q., Hao, M., Huang, H., Hou, J. et al., Iron deprivation suppresses hepatocellular carcinoma growth in experimental studies. *Clin. Cancer Res.* 2011, *17*, 7625–7633.

[23] Alison, M. R., Islam, S., Wright, N. A., Stem cells in cancer: instigators and propagators? *J. Cell Sci.* 2010, *123*, 2357–2368.

[24] Diehn, M., Cho, R. W., Lobo, N. A., Kalisky, T. et al., Association of reactive oxygen species levels and radioresistance in cancer stem cells. *Nature* 2009, *458*, 780–783.

[25] Roesch, A., Fukunaga-Kalabis, M., Schmidt, E. C., Zabierowski, S. E. et al., A temporarily distinct subpopulation of slow-cycling melanoma cells is required for continuous tumor growth. *Cell* 2010, *141*, 583–594.

[26] Bi, M., Naczki, C., Koritzinsky, M., Fels, D. et al., ER stress-regulated translation increases tolerance to extreme hypoxia and promotes tumor growth. *EMBO J.* 2005, *24*, 3470–3481.

[27] Holmes, M. D., Chen, W. Y., Schnitt, S. J., Collins, L. et al., COX-2 expression predicts worse breast cancer prognosis and does not modify the association with aspirin. *Breast Cancer Res. Treat. 130*, 657–662.

[28] Cadenas, C., Franckenstein, D., Schmidt, M., Gehrmann, M. et al., Role of thioredoxin reductase 1 and thioredoxin interacting protein in prognosis of breast cancer. *Breast Cancer Res. 12*, R44.

[29] Abd El-Rehim, D. M., Pinder, S. E., Paish, C. E., Bell, J. et al., Expression of luminal and basal cytokeratins in human breast carcinoma. *J. Pathol.* 2004, *203*, 661–671.

[30] Epsztejn, S., Glickstein, H., Picard, V., Slotki, I. N. et al., H-ferritin subunit overexpression in erythroid cells reduces the oxidative stress response and induces multidrug resistance properties. *Blood* 1999, *94*, 3593–3603.

[31] Fang, D., Bao, Y., Li, X., Liu, F. et al., Effects of iron deprivation on multidrug resistance of leukemic K562 cells. *Chemotherapy 56*, 9–16.

[32] Liu, X., Madhankumar, A. B., Slagle-Webb, B., Sheehan, J. M. et al., Heavy chain ferritin siRNA delivered by cationic liposomes increases sensitivity of cancer cells to chemotherapeutic agents. *Cancer Res. 71*, 2240–2249.

[33] Jiang, X. P., Wang, F., Yang, D. C., Elliott, R. L. et al., Induction of apoptosis by iron depletion in the human breast cancer MCF-7 cell line and the 13762NF rat mammary adenocarcinoma in vivo. *Anticancer Res.* 2002, *22*, 2685–2692.

[34] Matsumoto, K., Arao, T., Tanaka, K., Kaneda, H. et al., mTOR signal and hypoxia-inducible factor-1 alpha regulate CD133 expression in cancer cells. *Cancer Res.* 2009, *69*, 7160–7164.

[35] Liu, R., Wang, X., Chen, G. Y., Dalerba, P. et al., The prognostic role of a gene signature from tumorigenic breast-cancer cells. *New Engl. J. Med.* 2007, *356*, 217–226.

