Published in final edited form as:

Stem Cells. 2008 December ; 26(12): 3037–3046. doi:10.1634/stemcells.2008-0397.

Quantitative Mass Spectrometry Identifies Drug Targets in Cancer Stem Cell Containing Side Population

Sebastian C.J. Steiniger¹, Judith A. Coppinger³, Jörg A. Krüger³, John R. Yates III³, and Kim D. Janda^{1,2,*}

¹The Skaggs Institute for Chemical Biology and Departments of Chemistry and Immunology, La Jolla, California, 92037

²Worm Institute of Research and Medicine (WIRM), and Department of Immunology, La Jolla, California, 92037

³The Scripps Research Institute, La Jolla, California, 92037

Abstract

A multifaceted approach is presented as a general strategy to identify new drug targets in a breast cancer stem cell containing side population. The approach we have utilized combines side population cell sorting and stable isotope labeling by amino acids in cell culture with mass spectrometry to compare and identify proteins with differential expression profiles between side population cells, known to be enriched in cancer stem cells and non-side population cells, which are depleted in cancer stem cells, for two breast cancer cell lines, MCF7 and MDA-MB231. Almost 900 proteins were quantified and several important proteins in cell cycle control and differentiation, were found to be upregulated in the cancer stem cell containing side population. Most interestingly, a splice isoform of pyruvate kinase M2 was found to be downregulated. The differential levels of three of these proteins, Thymosin beta four, PA2G4 and SIAH interacting protein, were validated using Western blot. Furthermore, functional validation provided clear evidence that an elevated TB4 expression contributes to drug resistance in the stem cell population. SiRNA silencing of TB4 led to a loss of chemoresistance in two separate breast cancer populations. These proteins likely contribute to the resistance in the cancer stem cell containing side population and their altered expression in a tumor causes clinical resistance to chemotherapy. The ability to combine SILAC with mass spectrometry has enabled the identification of a series of proteins that could serve as future therapeutic targets.

Keywords

proteomics; side population; SILAC; breast cancer stem cells; Thymosin beta four; Proliferation associated protein 2G4; SIAH interacting protein; SIN3

Introduction

Interest in cancer stem cells (CSC) continues to expand as a fundamental understanding in CSC mechanism of development and disease progression moves forward. Stem cells are defined as

*Address correspondence to this author: Kim D. Janda, Departments of Chemistry and Immunology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, Tel. 858 784-2516; Fax. 858 784-2590 kdjanda@scripps.edu.

Authorships contributions. S.C.J.S. designed the study and wrote the paper. S.C.J.S. and J.A.C. performed the research and were analyzing the data. All authors participated in interpretation of data and approval of the final manuscript.

The authors declare no competing financial interests.

undifferentiated cells that are capable of self-renewing and differentiating into a large number of heterogeneous mature progeny. The capability of extensive proliferation of a distinct subpopulation was first described and extensively documented for hematopoietic tumors [1-3]. Later, it was demonstrated that only a subpopulation of tumor cells is enriched for clonogenic capacity and that the vast majority of cells lack this ability. Furthermore, this was not only shown for hematopoietic tumors, but also for solid cancers [4]. Overall, this led to the conclusion that tumors (malignant) contain stem-like cancer cells. This has been the case for solid tumors, for cancer cell lines in xenograft models as well as for mouse tumor cells in syngeneic models [5]. However, recent experiments show that the milieu in which the cancer cells are growing may also influence their tumorigenicity [6].

The concept of CSCs provides a new era interims of clinical implications, treatment options and could grant new insights into tumor development and metastasis [7]. Indeed, the lack of specificity seen with anticancer drugs towards tumors maybe a direct consequence of CSC in tumor tissue [8]. To address this issue, methods of delivering cytotoxic drugs specifically targeting CSC will have to be investigated [9]. In breast cancer, unlike colon cancer and pancreatic cancer, no single dominant pathway or histological presentation has emerged [10]. Human breast carcinomas have heterogeneous pathologies and molecular profiles, and in this respect it is plausible that breast cancer might be more similar to hematological malignancies than to other common epithelial cancers. Characterization of chromosomal aberrations, gene mutations and gene expression profiles of breast cancers and tumors have shown that breast tumorigenesis does not necessarily progress in a stepwise linear fashion from well-differentiated to poorly differentiated tumors [11]. Changes in gene expression that occur between different cell populations have been extensively studied using DNA microarrays [12]. This method offers the ability to compare the expression levels of thousands of genes in a single experiment, however, the observed expression levels do not always correspond with the protein levels because other control mechanisms exist, like post-transcriptional regulatory mechanism [13]. Protein expression can only be determined by direct measurement of protein levels.

To further delineate the importance of breast cancer stem cells we used a multifaceted approach to characterize the SP (side populations-known to be enriched in primitive cells) of two breast cancer cell lines *in vitro* and in a mouse model. In conduit we undertook a quantitative proteomics approach that combined cell sorting and SILAC labeling with mass spectrometry to compare and identify proteins with differential expression profiles between the side population and non-side population cells to determine proteins that may be responsible for the self propagating and tumorigenic properties of breast CSCs. With this combined research strategy we have uncovered proteins that play a role in drug resistance and possibly, tumorigenicity.

Experimental Methods

Animals and Cell Lines

Female CB17/SCID mice were obtained from Taconic Farms (Germantown, NY). All animal experiments were performed according to the National Institutes of Health Guides for the Care and Use of Laboratory Animals. Human mammary adenocarcinoma cells MCF7 and MDA-MB231 were obtained from ATCC (Manassas, VA) MCF7 cells were cultured in RPMI 1640 (Invitrogen, Carlsbad, CA) supplemented with 10% fetal bovine serum (Invitrogen) and 2 mM L-Glutamine (Invitrogen) and MDA-MB231 were cultured in supplemented DMEM. All cell cultures were maintained at 37 °C in a humidified atmosphere of 5% CO₂ in air.

Flow cytometry and cell sorting

Tumor cells were incubated with Bisbenzimidazole H33342 dye (Fluka, Buchs, Switzerland), or H33342 dye plus Verapamil (Sigma) or Reserpine (Sigma) and incubated at 37°C for 90 min. with periodic agitation as previously described [14]. After incubation cells were centrifuged, resuspended in cold buffer and kept at 4°C until analysis. PI was added at a concentration of 5 µg/ml to exclude dead cells. Cells were further stained with antibodies against sca-1, c-Kit, CD44 and CD24 (BD Pharmingen, San Diego, CA) and the appropriate isotype controls. FACS data were analyzed with FlowJo software (Tree Star, Ashland, OR).

Reverse transcription PCR analysis

Total RNA from three independently sorted MCF7 and MDA-MB231 cell samples were isolated using an RNeasy Kit (Qiagen) and the RNA was amplified with the Invitrogen Superscript First Strand Kit (Invitrogen) according to the manufacturer's instructions. RT-PCR was performed using Superscript II RT. The primers used are as follows: *ABCG2* (sense, 5'-GGCTTCTACCTGCACGAAAACCAGTTGAG-3', antisense, 5'-ATGGCGTTGAGACCAG-3'), *Stratifin* (sense, 5'-GGCCATGGACATCAGCAAGAA-3'; antisense, 5'-CGAAAGTGGTCTTGGCCAGAG-3'), *TB4* (sense, 5'-CTGAGATCGAGAAATTCGATAAGTC-3'; antisense, 5'-CCTGCTTGCTTCTCCTGTTC-3'), *PA2G4* (sense, 5'-AGCGACCAGGATTATATTCTCAAG-3'; antisense, 5'-ACAAAGGTGAGCTGCCTTAATAAC-3'), *SIAH-interacting protein* (sense, 5'-CTCCCATTAACAACGGGCTATAC-3', antisense, 5'-CTCAGTGGGAACCTTGATGAAGTC-3') etc. Ten nanograms RNA/sample were amplified using 35 cycles (Stratifin) and 40 cycles (SIAH, PA2G4, TB4) in a PTC-200 Thermal Cycler.

MTT assay

Stem cells have been found to be more resistant toward radiotherapy and chemotherapy and therefore to test the relative resistance of MCF7 SP cells to chemotherapy we performed a cytotoxic assay on the SP cell populations found within the MCF7 and MDA-MB231 cells. MCF7 cells stained with Hoechst 33342 dye were sorted into SP and Non-SP populations and plated in 96 well plates at a 10,000 cells/well concentration in the presence of Doxorubicin (2.5 mg/ml), Methotrexate (4mg/ml), 5-Fluorouracil and additionally Paclitaxel for siRNA studies. The MTT assay was performed 24 later according to the manufacturer's protocol (Promega).

Tumor induction

Cells stained with Hoechst 33342 dye were sorted into SP and Non-SP populations and injected into the flank of female CB17/SCID mice. Prior to xenograft implantation of MCF7 cells animals were implanted with estrogen pellets (17β-Estradiol, 0.72 mg/pellet, 60-day release, Innovative Research of America, Sarasota, FL) using a subcutaneous trocar (Innovative Research of America). For tumorigenicity assays cells were resuspended in a mixture of RPMI or DMEM and Matrigel™ (reconstituted basement membrane, BD Biosciences) 1:1, at variable cell densities. Mice were evaluated daily and the size was measured with a caliper, tumor size was calculated with the following formula: Tumor size (mm) = (a + b)/2, where a is the larger and b is the smaller diameter.

Cell Labeling and Fractionation

Cells were maintained in RPMI 1640 containing 10% FBS and 2 mM L-Glutamine. To initiate the incorporation of light or heavy labels, cancer cells (10⁴ cells) were harvested and resuspended in 10 mL of modified RPMI 1640 supplemented with 10% dialyzed FBS, and 0.1 mg/mL light L-lysine and light L-arginine, and an equal quantity of cells were resuspended in

10 mL of modified RPMI 1640 supplemented with 10% dialyzed FBS, and 0.1 mg/mL heavy [U-¹³C₆]L-lysine and heavy [U-¹⁵N₄, ¹³C₆]L-arginine (Invitrogen). Initially, labeled and non-labeled cells were cultured in two separate flasks. Every 3-4 days, the cells were split or the media was replaced with either the light or heavy labeling medium. The cells were expanded at least six doubling times to achieve complete incorporation of the labeled amino acid.

Preparation of protein samples

Cells from both populations were scraped in a lysis buffer containing 1% SDS, 1% Nonidet P-40, 50 mM Tris, pH 7.5, 150 mM NaCl and protease inhibitors (Complete™ tablets; Roche Diagnostics, Mannheim, Germany). The lysate was sonicated for two cycles of 30 s each and centrifuged to pellet cellular debris. Protein quantitation was performed using the Bradford protein assay, and mixtures of lysates were combined in protein concentration ratios of 1:1. The resulting lysates were subjected to methanol/ chloroform precipitation and digested with trypsin (1.50) (Promega). Before digestion, protein mixtures were denatured with solid urea (final concentration of 8M in 100mM Tris pH 8.5), reduced with TCEP (Sigma, 3 mM final concentration for 20 min. at room temp.), and alkylated with iodoacetamide (Sigma, 10mM final concentration for 30 min at room temp.).

Mass Spectrometry and Data Analysis

(b) Multidimensional Liquid Chromatography Tandem Mass Spectrometry of Peptide Mixtures (MudPIT)—This approach has been described in detail by the Yates laboratory [15]. A three-phase microcapillary column was constructed and equilibrated with 5% acetonitrile/ 0.1% formic acid for 30 min before the peptide mixture was loaded. For analysis, the microcolumn was positioned in-line with an Eksigent nanoflow HPLC pump directly in front of the heated capillary opening of an LTQ-Orbitrap hybrid mass spectrometer (ThermoElectron). Peptide mixtures were analyzed using a sixstep separation procedure. As peptides were eluted and ionized into the mass spectrometer, acquisition of tandem mass spectra was repeated continuously during the course of the analysis. All tandem mass spectra were collected using a normalized collision energy of 35%, an isolation window of 3 *m/z*, and 1 μ scan. Data-dependent acquisition was employed and consisted of one full-scan mass spectra followed by four tandem mass spectra. Other instrumental parameters included a spray voltage of 2.4 kV, a maximum injection time of 100 ms, LTQ AGC target for MS/MS of 10 000 ions, and an FTMS AGC target for MS of 500 000 ions.

Software. (a) Interpretation of Tandem Mass Spectra—Both MS and tandem mass spectra were extracted from the Xcalibur data system format (.RAW) into MS1 and MS2 formats 19 using in-house software (RAW_Xtractor). Tandem mass spectra were interpreted by SEQUEST (ver 27) [16] which was parallelized on a Beowulf cluster of 100 computers and results were filtered, sorted, and displayed using the DTASelect (ver 2.0) program utilizing a decoy database strategy [17] that sets scoring threshold values based on a specific false positive rate (user supplied). Searches were performed against an EBI-IPI –human_3.05_04-04-2007 database concatenated to a decoy database in which the sequence for each entry in the original database was reversed.

Quantitative Analysis Using Census—After filtering the results from SEQUEST using DTASelect2, ion chromatograms were generated using an updated version of a program previously written in the Yates' laboratory. This software, called Census is composed of two parts: a quantitative analysis module that generates chromatograms from spectra and calculates peptide ion abundances, and a qualitative module that aids in tandem MS spectra visualization. Census is available from the authors for individual use and evaluation (<http://fields.scripps.edu/Census> for details) [18].

Western Blot Analysis—Following SDS-PAGE separation of equal amounts of protein from each sample on a 4-12% Bis-Tris gel, protein was transferred overnight at 4°C onto a PVDF nitrocellulose membrane. After transfer the membrane was blocked for one hour in 2% milk, membranes were washed in PBS-Tween 20 (PBS-T) and incubated with the following primary antibodies: mouse monoclonal GAPDH (1:1000) from Ambion, rabbit polyclonals EBP1 (1:1000), SIN3 (1/400) and mouse monoclonal SIAH Interacting protein (1:200) from Abcam and rabbit polyclonal TB4 (1/200) from Biodesign followed by incubation with goat anti rabbit IgG horse radish peroxidase (1:1000) and goat anti mouse (1:5000) from Pierce. Blots were developed with SuperSignal West Dura from Pierce.

Inhibition of TB4 expression by siRNA—The short interference RNA (siRNA) duplexes were constructed with Silencer™ siRNA Construction kit (Ambion, Austin, TX) following the manufacturer's instructions. Three siRNAs for human TB4 and SIN3 were synthesized and control siRNA oligonucleotide, which does not match any known human coding cDNA, was designed and used. The siRNA duplexes were reconstituted to 10 µmol/L in sterile RNase-free water. Transfection of siRNA for targeting endogenous genes was done using Lipofectamine RNAimax (Invitrogen) as per the manufacturer's instructions.

Results

Breast cancer cell lines contain a cancer stem cell containing side population specified by Hoechst 33342 dye exclusion technique

The Hoechst 33342 dye exclusion technique was utilized to identify side population cells in two different breast cancer carcinoma cell lines. This technique has been described previously in the literature [14,19]. Thus, putative breast cancer SP cells were expected to form a distinctly shaped population of H33342^{low} cells and to decrease in response to the ABC transport inhibitor verapamil. To this end, we have stained bone marrow as a positive control (Figure 1A,D) and MCF7 (Figure 1B,E), and MDA-MB231 (Figure 1C,F) cells as well. Our FACS data clearly show the side population in the bone marrow sample and the MCF7 cells and diffusely scattered H33342^{low} cells within MDA-MB231 cell line. To further verify our results, we used the verapamil as a Ca²⁺ blocker, and this reduced the number of side population cells. The results suggest, that the two cell lines, although to a different degree, contain side population cells similar to the bone marrow cells. The amount of side population cells was found to be greater in the MCF7 than in the MDA-MB231 cell line, therefore this cell line was examined for proteomic analysis.

MCF7 and MDA-MB231 cells express markers associated with stem cells

To further characterize the SP cell populations, the expression of markers that have been associated with stem cells or progenitor cells was investigated. Consequently, we examined the surface expression marker profiles of the MCF7 and MDA-MB231 cells, with antibodies against the surface markers as well as their isotypic controls (Figure 2A). FACS analysis revealed that H33342^{low} cells had a weak positive c-Kit profile. It also provided evidence that MCF7 SP cells are positive for CD44 and weakly positive for c-Kit compared to non-SP cells. MDA-MB231 SP cells also stained positive weakly positive for c-Kit. Recent reports indicate a CD44⁺/CD24^{-/low} subpopulation with stem cell properties in breast cancer cell lines. Consequently, we analyzed the co-expression of in both cell lines. The MCF7 cell line has little to none CD44⁺/CD24^{-/low} cells, with most of the cells positive for CD44 and CD24. The MDA-MB231 cell line has a higher amount of CD44⁺/CD24^{-/low} cells, but no difference in expression could be observed between the SP and non-SP population. To further examine the expression of markers associated with stem cells, total RNA was isolated and used for RT-PCR, which showed an upregulation of ABCG2 and Stratifin (Figure 2B).

MCF7 and MDA-MB231 SP cells are more resistant to chemotherapy

Stem cells have been found to be more resistant toward radiotherapy and chemotherapy. Therefore, to test the relative resistance of MCF7 SP cells to chemotherapy, the two cell lines stained and sorted for the SP cell population in MCF7 and MDA-MB231 cells which were then plated and a cytotox assay was conducted. After 24 h, the number of live cells was determined by the MTT assay. MCF7 SP and MDA-MB231 SP cells showed a significantly greater survival when treated with Doxorubicin, $91.2\% \pm 3\%$ and $99.5\% \pm 0.5\%$, respectively, than non-SP cells $50\% \pm 2.21\%$ and $52\% \pm 3.21\%$, respectively. Similar results were obtained with Methotrexate and 5-Flourouracil. SP cells were more resistant and showed greater survival (Figure 3). The number of cells treated with cytotoxic drugs showed a significantly greater survival, $91\% \pm 5\%$, respectively.

MCF7 and MDA-MB231 SP cells are more tumorigenic

The most significant characteristic of a stem cell is its ability to repopulate an entire cell population *in vivo*. To this end, MCF7 cells were stained with the H33342 dye; these were sorted into PI-/live SP, and non-SP cells; and injected 8×10^3 cells subcutaneously into the flank of SCID mice. SP cells were significantly more potent in inducing tumor formation (Figure 4). We also observed a similar result with MDA-MB231 cells.

Proteomic analysis of MCF7 adenocarcinoma cells

To identify differentially expressed proteins, SILAC analysis was engaged to examine the relative abundances of specific proteins in H33342^{high} (non-SP) and H33342^{low} (SP) populations (shown in Figure 5A) of the breast cancer cell line MCF7. Figure 5B provides a schematic diagram of the quantification strategy. Thus, two batches of cells were grown, one in a medium that contain 'light' (normal) amino acids and the other in a medium that contain 'heavy' amino acids, with a heavy amino acid containing ^{13}C instead of ^{12}C and ^{15}N instead of ^{14}N . Incorporation of the heavy amino acid into a peptide leads to a known mass shift compared to with the peptide that contains the light version of the amino acid, but no other chemical changes. After at least 6 cell doublings, both batches were subject to cell sorting, where the side population and the non-side population were collected (Figure 5A). The heavy labeled side population was combined 1:1 with the non labeled (light) parent cells and subjected to proteolysis and mass spectrometry analysis on an LTQ-Orbitrap hybrid mass spectrometer (ThermoElectron) (Figure 5B). Peptides were analysed using MudPIT and data-dependent acquisition was employed and consisted of one full-scan mass spectra (in the Orbitrap) followed by four tandem mass spectra in the LTQ. Higher quantification efficiency has been reported when Orbitrap MS scans were employed than when LTQ MS scans were used [20]. The relative abundances of each peptide identified in the mixture were then determined using the Census software platform. In the mass spectra, each peptide appears as a pair. The peptide with lower mass contains the light amino acid and originates from population A, and the peptide with higher mass contains the heavy amino acid and originates from population B. Census extracts the ion chromatogram and calculates peptide ion intensity ratios for each pair of extracted ion chromatograms. A SILAC peptide pair appears in a one-to-one ratio when there is no difference in the abundance of this protein between the proteomes. Experiments resulted in the quantification of 870 different proteins (Supplementary Table 1 and 2), and over 80% of these proteins were quantified by two or more peptides. Almost 70% of these proteins were quantified in a duplicate run, while 41 proteins were found to be upregulated in duplicate samples by greater than 1.5 fold in the side population (Figure 5C). A total of 17 proteins were found to be downregulated by 1.5 fold (Figure 5D). These proteins were classified by GOMiner and explored extensively in the literature, several proteins with roles in differentiation and cell cycle control were found to be upregulated e.g. neuroblast differentiation factor, TB4 and proliferation associated factor. Downregulated proteins, included Pyruvate kinase M2, and

peroxiredoxin 6, both have been reported to play a role in tumor progression. To confirm mass spectrometry data, western blot analysis was performed. Here, equal amounts of non-side and SP cell protein were loaded on a gel and upon transfer were probed with TB4, proliferation associated protein and SIAH interacting protein. An increase in the level of all three proteins was observed in the side population confirming the SILAC observations (Figure 6A).

MCF7 side population cells express high levels of TB4, Proliferation associated protein 2G4, and SIAH interacting protein

To examine whether protein alterations observed by proteomic analysis correlate with changes of mRNA at the transcription level, these three proteins with significant expression changes were chosen for further validation by semi-quantitative RT-PCR. As shown in Figure 6, our data demonstrate that the observed alterations in the expression of those genes are seen at both transcription and translational levels. In order to examine whether the upregulation of these proteins could also be found at the molecular level total RNA isolated from sorted H33342^{low} and H33342^{high} and RT-PCR, was performed for proliferation associated protein, TB4 and SIAH (Figure 6B). An upregulation in the level of expression of all three genes was observed in the side population.

siRNA mediated knockdown of endogenous TB4 decreases drug resistance of MCF7 cells

To investigate the role of TB4 in cancer cells, TB4's expression in MCF7 and MDA-MB231 cells was reduced by small interfering RNA (siRNA). Knock-down of TB4 was confirmed by western blots in these cell lines (Figure 7). Changes in cell morphology were observed where in the morphology of the cells became more fibroblast like, compared to mock-transfected cells. siRNA mediated knock-down of TB4 in MCF7 cells and MDA-MB231 cells resulted in augmented cell death and loss of chemoresistance when treated with paclitaxel and to a lesser extent doxorubicin (Figure 8). This finding suggests that cellular responsiveness to paclitaxel could be regulated by TB4 in MCF7 and MDA-MB231 cells. Additionally the expression of SIN3 was reduced by siRNA (Figure 7). Proliferation analysis under conditions where SIN3 levels are decreased by siRNA showed that SIN3 is required for proliferation of MCF7 cells and MDA-MB231 (data not shown).

Discussion

Breast cancer SP populations demonstrate stem-like properties

A growing body of evidence has been reported supporting the hypothesis that tumors are organized in a hierarchy of heterogeneous cell populations with different biological properties and that the capability to sustain tumor formation, and growth exclusively resides in a small proportion of tumor cells, termed cancer stem cells (CSCs) [21] [22,23]. In this study relevant models were successfully applied for the in-depth characterization of CSCs *in vitro* and in the tumor environment of immunodeficient mice. CSCs have been identified phenotypically through specific expression marker profiles. Our results demonstrate that using the Hoechst 33342 dye efflux method, we were able to detect and to analyze a Hoechst^{low} SP in two different breast cancer cell lines. FACS analysis and RT-PCR demonstrated that the SP cells found within the tested breast cancer cell lines also contain markers associated with stem cells, including sca-1, Stratifin and ABCG2. These observations are in agreement with previous studies that showed CD44⁺/CD24^{-/low} breast cancer cells displayed stem/progenitor cell markers [24]. The stem cell like phenotype of CSCs is thought to contribute to capability of certain tumors to be resistant to conventional drug therapies and several stem cell populations have shown resistance to chemotherapy [5]. We explored this concept further by analyzing the side populations in the presence of three different drugs, Doxorubicin, Methotrexate and 5-Fluorouracil. Our findings further supported these observations and in the presence of all three drugs, SPs were found to be more resistant and showed greater survival. We additionally

examined the ability of SP cells to drive tumor formation *in vivo* by injecting immunocompromised mice with both SP and non-SP cell populations. The SP population demonstrated greater ability to form tumors, an observation also echoed in a recent study, where in MCF7 SP cells upon injection into SCID mice showed efficient colonization of the mammary glands [25].

Proteomic analysis of SP and non- SP population reveals differential expression profiles

Over the past several years different cancer cell populations have been characterized using genomic tools. A variety of high-throughput techniques such as RT-PCR, DNA microarray and others have rendered it feasible to establish global quantitative mRNA expression profiles of cancer cells. In particular a comprehensive molecular and phenotypic analysis of CD44⁺ and CD45⁺ cells from breast carcinomas was performed. This study revealed that these cell populations have distinct gene expression, epigenetic, and genetic profiles [26]. In order to define the protein complement and differential profiles of CSC cells a proteomic approach was used. Proteomics can be defined as a systematic approach to analyze both qualitative and quantitative mapping of the whole proteome. [27,28]. Proteomic analysis maybe investigated on many different levels we selected SILAC analysis to examine the relative abundances of specific proteins in breast cancer cells exhibiting low (Hoechst33342^{high}) and high (Hoechst33342^{low}). Our choice of using SILAC to examine the CSC proteome derives from the high-throughput, fast-speed and easy manipulation of the proteome using SILAC, which is advantageous compared to 2D-PAGE or conventional western blot analysis. Using this technique the quantification of nearly 900 proteins was accomplished. Figure 5 shows the SILAC ratio distribution for all proteins upregulated or downregulated at least 1.5 fold in the stem cell population.

Using SILAC a total of the 41 upregulated proteins were identified to be upregulated and excitingly several of these proteins are known to play a role in differentiation and cell cycle control. The most interesting proteins identified include: (A) AHNAK – neuroblast associated differentiation factor, whose expression and cellular localization has been dynamically regulated during cell cycle progression and it has been localized to sites of major morphogenesis during mouse placentation [29]. (B) Proliferation associated protein, also known as Epb1, which is involved in the regulation of cell growth of differentiation in mammary epithelial cells through putative DNA binding properties after translocation into the nucleus [30]. (C) E3 ubiquitin-protein ligase TRIM33, known to play a role in the control of cell proliferation; while its association with SMAD2 and SMAD3 stimulates erythroid differentiation of hematopoietic stem/progenitor [31]. (D) Siah-1 interacting protein, was also discovered, Siah-1-interacting protein (CacyBP/SIP), a target protein of S100, has been identified as a component of a novel ubiquitylation complex leading to beta-catenin degradation, which was found to be related to the malignant phenotypes of gastric cancer [32]. As stem cells are a continually self-renewing proliferative population, control of growth may be a crucial step in moving away from the stem cell phenotype and some of these proteins may potentially serve as therapeutic targets.

In terms of downregulated proteins identified by SILAC pyruvate kinases were found to be prominent whose activity has been shown to be generally increased in malignant tumors such as carcinomas [33]. This result was of interest from the standpoint of tumorigenicity and malignancy of the side population cells. Another important molecule that was downregulated was Peroxiredoxin 6, which seems to be involved in tumor progression and metastasis [34]. To further compliment these findings the differential profiles of several targets by RT-PCR and immunoblotting were characterized and found to possess similar patterns.

We also tried to verify/confirm the results in the CD44⁺/CD24^{-low} population, since this another known population enriched for cancer stem cells. However, we could only confirm the

results for MDA-MB231 cells, since the population of CD44⁺/CD24^{-low} in MCF7 cells is very low.

SiRNA silencing of TB4 results in decreased drug resistance in MCF7 cells

Proteomic analysis of the SP cells also revealed an increase in expression of TB4. TB4 is a small, 43 amino acid, actin binding, intracellular protein involved in cell motility [35]. Increased expression of this protein has been found in various tumors, e.g. breast, ovary, uterus, colon and thyroid [36-39]. TB4 expression has been studied in breast cancer cell like MCF7 and MDA-MB231 cells and the levels of beta thymosins in tumor tissues have been proposed to constitute a potential diagnostic marker for the state of a tumor. Overexpression of TB4 has been associated with increased invasiveness and metastasis in colon cancer [40] therefore we postulated that decreasing the levels of thymosin expression in the breast cancer cells may alter growth rate and chemosensitivity of these cells. siRNA silencing of TB4 resulted in an increase in cell death upon paclitaxel and doxorubicin treatment in MCF7 and MDA-MB231 cells compared to the control siRNA indicating that chemoresistance of these cells had been decreased. This observation was echoed in a previous study that showed paclitaxel induced cell death was increased in Hela cells upon siRNA of thymosin [41]. As our findings demonstrate the SP population strongly contributes to the drug resistance and these results implicate TB4 in drug resistance of cancer stem cells.

Specifically, Pyruvate kinase M2 was found to be downregulated 1.5 fold. This result is interesting, considering the result of a recently published study [42]. This group could show that rapidly growing cancer cells have a high activity of PKM2. The activity of this glycolytic can be activated and controlled by tyrosine kinase signaling pathways. The regulation of PKM2 activity may provide a direct link between cell growth signals using tyrosine kinase and control of glycolytic metabolism. This regulation could be the key determinant on how cancer stem cells regulate their metabolism and their cell growth. Inhibiting this enzyme could be a way of dormant cell and prevent growth and regrowth of cancer stem cells.

Conclusion

In summary we have provided evidence/data/ that has enabled new insights into breast cancer stem cells populations. Our results are based on a research strategy of cell sorting combined with quantitative mass spectrometry that has allowed us to identify several new proteins that may play a role in directing stem cells and could serve as future therapeutic targets. Using this combination of cell sorting and quantitative mass spectrometry we identified 41 proteins that have altered expression in the cancer stem cell containing side population of two breast cancer cell lines. Furthermore, we showed that an elevated TB4 expression contributes significantly to the drug resistance phenotype of stem cells using siRNA technology. Because of the importance of CSC in tumor progression and metastasis our approach may represent a new pharmacological opportunity for therapeutic intervention. The current breast cancer stem cells models using SP, sphere culture and CD44⁺/CD24^{-low} surface markers, have limitations as each model probably only captures one subpopulation of the pure cancer stem cell population, and therefore, none these methods has been universally accepted. Although the SP does is not necessarily comprised of pure cancer stem cells and the Hoechst 33342 staining compromises the viability of the cells, the live stain represents a significant enrichment of rare cells responsible for initiating and maintaining cancer. Furthermore, other potential stem cell markers, including CD24, CD44 and Nestin do not fully associate with Hoechst^{low} cells and can therefore not replace the Hoechst 33342 efflux assay for isolating breast cancer stem cells. At this time, no breast cancer stem cell research can be performed at this time without the use of these models. However, the use of cell lines may not fully account for the differences

between primary patient material and longterm cell culture. Therefore, further validation of cancer stem cells from patient samples needs to be done in future studies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by The Skaggs Institute for Chemical Biology. We thank Dr. Alan Saluk, Cheryl Kim, Alex Ilic and Quyen Nguyen for the excellent assistance with FACS analysis and cell sorting at the flow cytometry core facility at TSRI. We would also like to thank Ralph A. Reisfeld for helpful discussion and support. John Yates acknowledges NIH P41 RR011823 and 1 R01 HL079442-01.

Abbreviations

MCF	Michigan Cancer Foundation
siRNAs	small interfering ribonucleic acids
TB4	Thymosin beta 4
SP	side population
CSC	cancer stem cells
H33342	Hoechst 33342
SILAC	stable isotope labeling with amino acids in cell culture
HSCs	human stem cells

References

1. Wodinsky I, Swiniarski J, Kensler CJ. Spleen colony studies of leukemia L1210. 3. Differential sensitivities of normal hematopoietic and resistant L1210 colony-forming cells to 6-mercaptopurine (NSC-755). *Cancer Chemother Rep* 1968;52(2):251–5. [PubMed: 5646852]
2. Bruce WR, Van Der Gaag H. A Quantitative Assay for the Number of Murine Lymphoma Cells Capable of Proliferation in Vivo. *Nature* 1963;199:79–80. [PubMed: 14047954]
3. Bergsagel DE, Valeriote FA. Growth characteristics of a mouse plasma cell tumor. *Cancer Res* 1968;28(11):2187–96. [PubMed: 5723963]
4. Bonnet D, Dick JE. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nat Med* 1997;3(7):730–7. [PubMed: 9212098]
5. Kruger JA, et al. Characterization of stem cell-like cancer cells in immune-competent mice. *Blood* 2006;108(12):3906–12. [PubMed: 16912222]
6. Kelly PN, et al. Tumor growth need not be driven by rare cancer stem cells. *Science* 2007;317(5836):337. [PubMed: 17641192]

7. Huff CA, et al. Strategies to eliminate cancer stem cells: clinical implications. *Eur J Cancer* 2006;42(9):1293–7. [PubMed: 16644203]
8. Langerak, AD. *Chemotherapy Regimens and Cancer Care*. Georgetown, TX: Landes Bioscience; 2001. p. 209
9. Zhou X, et al. Cell-specific delivery of a chemotherapeutic to lung cancer cells. *J Am Chem Soc* 2004;126(48):15656–7. [PubMed: 15571383]
10. Kinzler KW, Vogelstein B. Lessons from hereditary colorectal cancer. *Cell* 1996;87(2):159–70. [PubMed: 8861899]
11. Stingl J, Caldas C. Molecular heterogeneity of breast carcinomas and the cancer stem cell hypothesis. *Nat Rev Cancer* 2007;7(10):791–9. [PubMed: 17851544]
12. Piscaglia AC, et al. Microarray RNA/DNA in different stem cell lines. *Curr Pharm Biotechnol* 2007;8(3):167–75. [PubMed: 17584089]
13. Everley PA, et al. Assessing enzyme activities using stable isotope labeling and mass spectrometry. *Mol Cell Proteomics* 2007;6(10):1771–7. [PubMed: 17627935]
14. Goodell MA, et al. Isolation and functional properties of murine hematopoietic stem cells that are replicating in vivo. *J Exp Med* 1996;183(4):1797–806. [PubMed: 8666936]
15. Washburn MP, Wolters D, Yates JR 3rd. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol* 2001;19(3):242–7. [PubMed: 11231557]
16. MacCoss MJ, Wu CC, Yates JR 3rd. Probability-based validation of protein identifications using a modified SEQUEST algorithm. *Anal Chem* 2002;74(21):5593–9. [PubMed: 12433093]
17. Elias JE, Gygi SP. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat Methods* 2007;4(3):207–14. [PubMed: 17327847]
18. Park SK, et al. A quantitative analysis software tool for mass spectrometry-based proteomics. *Nature Methods*. 2008in press
19. Christgen M, et al. Identification of a distinct side population of cancer cells in the Cal-51 human breast carcinoma cell line. *Mol Cell Biochem* 2007;306(12):201–12. [PubMed: 17660947]
20. Venable JD, et al. Relative quantification of stable isotope labeled peptides using a linear ion trap-Orbitrap hybrid mass spectrometer. *Anal Chem* 2007;79(8):3056–64. [PubMed: 17367114]
21. Reya T, et al. Stem cells, cancer, and cancer stem cells. *Nature* 2001;414(6859):105–11. [PubMed: 11689955]
22. Kondo T, Setoguchi T, Taga T. Persistence of a small subpopulation of cancer stem-like cells in the C6 glioma cell line. *Proc Natl Acad Sci U S A* 2004;101(3):781–6. [PubMed: 14711994]
23. Zhou S, et al. The ABC transporter Bcrp1/ABCG2 is expressed in a wide variety of stem cells and is a molecular determinant of the side-population phenotype. *Nat Med* 2001;7(9):1028–34. [PubMed: 11533706]
24. Ponti D, et al. Isolation and in vitro propagation of tumorigenic breast cancer cells with stem/progenitor cell properties. *Cancer Res* 2005;65(13):5506–11. [PubMed: 15994920]
25. Yin L, Castagnino P, Assoian RK. ABCG2 expression and side population abundance regulated by a transforming growth factor beta-directed epithelial-mesenchymal transition. *Cancer Res* 2008;68(3):800–7. [PubMed: 18245481]
26. Shipitsin M, et al. Molecular definition of breast tumor heterogeneity. *Cancer Cell* 2007;11(3):259–73. [PubMed: 17349583]
27. Mann M. Functional and quantitative proteomics using SILAC. *Nat Rev Mol Cell Biol* 2006;7(12):952–8. [PubMed: 17139335]
28. Everley PA, et al. Quantitative cancer proteomics: stable isotope labeling with amino acids in cell culture (SILAC) as a tool for prostate cancer research. *Mol Cell Proteomics* 2004;3(7):729–35. [PubMed: 15102926]
29. Downs KM, et al. Multiple developmental roles of Ahnak are suggested by localization to sites of placentation and neural plate fusion in the mouse conceptus. *Gene Expr Patterns* 2002;2(12):27–34. [PubMed: 12617833]
30. Yoo JY, et al. Interaction of the PA2G4 (EBP1) protein with ErbB-3 and regulation of this binding by heregulin. *Br J Cancer* 2000;82(3):683–90. [PubMed: 10682683]

31. He W, et al. Hematopoiesis controlled by distinct TIF1gamma and Smad4 branches of the TGFbeta pathway. *Cell* 2006;125(5):929–41. [PubMed: 16751102]
32. Ning X, et al. Calcyclin-binding protein inhibits proliferation, tumorigenicity, and invasion of gastric cancer. *Mol Cancer Res* 2007;5(12):1254–62. [PubMed: 18171983]
33. Yilmaz S, Ozan S, Ozercan IH. Comparison of pyruvate kinase variants from breast tumor and normal breast. *Arch Med Res* 2003;34(4):315–24. [PubMed: 12957530]
34. Chang XZ, et al. Identification of the functional role of peroxiredoxin 6 in the progression of breast cancer. *Breast Cancer Res* 2007;9(6):R76. [PubMed: 17980029]
35. Safer D, Golla R, Nachmias VT. Isolation of a 5-kilodalton actin-sequestering peptide from human blood platelets. *Proc Natl Acad Sci U S A* 1990;87(7):2536–40. [PubMed: 2320573]
36. Hall AK. Differential expression of thymosin genes in human tumors and in the developing human kidney. *Int J Cancer* 1991;48(5):672–7. [PubMed: 2071228]
37. Santelli G, et al. Thymosin beta-10 gene overexpression is a general event in human carcinogenesis. *Am J Pathol* 1999;155(3):799–804. [PubMed: 10487837]
38. Clark EA, et al. Genomic analysis of metastasis reveals an essential role for RhoC. *Nature* 2000;406(6795):532–5. [PubMed: 10952316]
39. Xie D, et al. Discovery of over-expressed genes and genetic alterations in breast cancer cells using a combination of suppression subtractive hybridization, multiplex FISH and comparative genomic hybridization. *Int J Oncol* 2002;21(3):499–507. [PubMed: 12168092]
40. Wang WS, et al. Overexpression of the thymosin beta-4 gene is associated with malignant progression of SW480 colon cancer cells. *Oncogene* 2003;22(21):3297–306. [PubMed: 12761500]
41. Oh SY, et al. ERK activation by thymosin-beta-4 (TB4) overexpression induces paclitaxel-resistance. *Exp Cell Res* 2006;312(9):1651–7. [PubMed: 16515784]
42. Christofk HR, et al. The M2 splice isoform of pyruvate kinase is important for cancer metabolism and tumour growth. *Nature* 2008;452(7184):230–3. [PubMed: 18337823]

Figure 1. Bone marrow, MCF7 and MDA-MB231 cells contain a side population (SP)
 (A-C) The cell lines indicated were stained with Hoechst 33342 dye to determine the Hoechst effluxing SP. (D-F) These cells were stained with Hoechst 33342 in the presence of verapamil to verify the specificity of the SP population. Each experiment was carried out in triplicate.

Figure 2A.

Surface marker expression on MCF7 and MDA-MB231 SP and non-SP cells						
Marker	MCF7 SP	MCF7 non-SP	Isotype Control	MDA-MB231 SP	MDA-MB231 SP	Isotype Control
c-Kit	5.6 ± 1.1	1.8 ± 0.5	2.23 ± 0.4	8 ± 1.1	2 ± 0.1	1.8 ± 0.5
CD44	734 ± 13.83	945 ± 10.2	5.56 ± 0.32	321 ± 4.72	156 ± 1.2	2.8 ± 0.9
CD24	450 ± 0.8	635 ± 1.3	1.8 ± 0.5	11.9 ± 4.6	10.5 ± 3.2	1.8 ± 0.5

Figure 2B.

Figure 2. MCF7 and MDA-MB231 cells express markers associated with HSCs, progenitor cells and stem cells

(A) Values represent the mean ± SEM of the mean fluorescence intensity; cells were stained at least 3 times for all markers. (B). **2b RT-PCRS shows upregulation of stem cell markers.** RNA from SP and non-SP MCF7 and MDA-MB231 cells were amplified with primers targeting the indicated genes.

Figure 3. MCF7 and MDA-MB231 SP cells are more resistant to chemotherapy

Cells were stained with Hoechst 33342 dye, sorted into SP and non-SP samples and treated either with Doxorubicin, Methotrexate and 5-FU. The viability of the cells was assessed by a MTT assay. Results are expressed as percentages of cells still viable relative to non-treated control cells.

Figure 4A.

Figure 4B.

Figure 4. MCF 7 and MDA-MB231 SP cells are more tumorigenic than MCF7 (A) and MDA-MB231 non-SP cells (B)

Cells were stained with Hoechst 33342 and sorted into SP and non-SP populations. 10³ cells from each group were injected into the flank of SCID mice. Tumor volume was measured after tumor cell challenge.

Figure 5A.

Figure 5B.

Figure 5C.

PROTEINS WITH ELEVATED LEVELS IN MCF7 SIDE POPULATION CELLS			
Protein Accession number	Protein name	Upregulation factor	No. of peptides
AHNAK_HUMAN	Neuroblast differentiation associated protein AHNAK	3.76	17
TYBA_HUMAN	Thymosin beta 4	3.53	2
P62807	H2B histone family, member A	3.29	9
ENSP00000334063	Peroxiredoxin 2 isoform B	3.02	2
Q8WW97	DNA binding protein TAXREB107	3	7
Q8UPN9	Splice isoform beta of transcription intermediary factor 1 gamma	2.69	3
COX8B_HUMAN	Cytochrome c oxidase polypeptide Vb, mitochondrial precursor	2.68	3
Q6ZDN2	Hypothetical protein FLJ46846	2.58	5
HMG1_HUMAN	High mobility group protein 1	2.52	9
ENSP00000316411	10 kDa protein	2.44	3
ENSP00000346117	24 kDa protein	2.43	1
SAP18_HUMAN	Sin 3 associated polypeptide	2.43	2
PA2G4_HUMAN	Proliferation associated protein 2G4	2.34	2
KGPP_HUMAN	6-phosphofructokinase, type C	2.23	6
NCB5R_HUMAN	Cytochrome b5 reductase membrane-bound isoform	2.07	3
SF3A1_HUMAN	Splicing factor 3 subunit 1	2.07	4
ENSP00000331660	39 kDa protein	2	3
ENSP00000287144	PREDICTED: similar to 60S ribosomal protein L29	2.07	2
ENSP00000352557	PREDICTED: similar to ribosomal protein S12	1.98	3
Q06830	Peroxiredoxin 1	1.97	5
Q01082-1	Splice isoform Long of Spectrin beta chain, brain 1	1.89	2
ENSP00000346174	Splice isoform 2 of Host cell factor	1.88	5
HMG2_HUMAN	High-mobility group box 2	1.83	5
P02545	Splice isoform C of Lamin A/C	1.82	8
Q5R370	SIAH interacting protein	1.8	4
Q95292-1	Splice isoform 1 of Vesicle-associated membrane protein-associated protein B/C	1.8	2
P11940-2	Splice isoform 2 of Polyadenylate-binding protein 1	1.77	7
ENSP00000350882	Splice isoform 2 of Spectrin alpha chain, brain	1.76	6
FLNA_HUMAN	Filamin A	1.73	16
Q96543	JKTB1delta5	1.72	2
P45880-2	Splice isoform 2 of Voltage-dependent anion-selective channel protein 2	1.71	5
P48047	ATP synthase oligomycin sensitivity conferral protein, mitochondrial precursor	1.71	8
Q35R11	Interleukin enhancer binding factor 2, 45kDa	1.68	3
P62081	40S ribosomal protein S7	1.67	5
Q8NF52	FLJ00343 protein	1.64	14
P49755	Transmembrane protein Tmp21 precursor	1.64	4
P23284	Peptidyl-prolyl cis-trans isomerase	1.6	10
Q85383	Plectin 1	1.6	5
Q13232	Nucleoside diphosphate kinase 3	1.6	5
Q13151	Heterogeneous nuclear ribonucleoprotein G	1.62	5
Q9P2F1	SUP16H protein	1.51	4

Figure 5D.

PROTEINS WITH DECREASED LEVELS IN MCF7 SIDE POPULATION CELLS			
Protein Accession number	Protein name	Downregulation factor	No. of peptides
P04350	Tubulin beta-4 chain	1.53	11
Q9NY17	Pyruvate kinase M2	1.5	18
ENSP00000246019	18 kDa protein	1.58	3
P06744	Glucose phosphate isomerase	1.58	9
Q01105-2	Splice isoform 2 of SET protein	1.75	2
P30041	Peroxiredoxin 6	1.92	4
P21291	Hypothetical protein FLJ16726	1.63	2
P00966	Argininosuccinate synthase	1.85	4
Q9UBX3-2	Splice isoform 2 of Mitochondrial dicarboxylate carrier	1.58	2
P00505	Aspartate aminotransferase, mitochondrial precursor	1.78	2
Q961T0	Fatty acid synthase	1.63	12
Q8N9C4	ATP citrate lyase isoform 2	1.51	3
P49588	Alanyl-tRNA synthetase	1.92	3
nl	PREDICTED: similar to tubulin, beta 8	1.61	3
ENSP00000353707	29 kDa protein	1.61	2
Q96DH6-2	Splice isoform 2 of RNA-binding protein Musashi homolog	1.47	1
ENSP0000035461	Casein kinase II alpha 1 subunit isoform b	1.61	2

Figure 5. Proteomic Flow Chart

Two batches of cells were grown, one in a medium that contains a 'light' (normal) amino acid and the other in a medium that contains a 'heavy' amino acid, with the heavy amino acid containing ^{13}C instead of ^{12}C and ^{15}N instead of ^{14}N . After at least 6 cell doublings, both batches were subject to cell sorting, where the side population and the non-side population were collected (shown in Figure 5A). The heavy labeled side population was combined 1:1 with the non labeled (light) parent cells and subjected to proteolysis and mass spectrometry analysis on an LTQ-Orbitrap hybrid mass spectrometer (ThermoElectron). Peptides were analysed using MudPIT and data-dependent acquisition was employed and consisted of one full-scan mass spectra (in the Orbitrap) followed by four tandem mass spectra in the LTQ. The relative abundances of each peptide identified in the mixture were then determined using the Census software platform. Census extracts the ion chromatogram and calculate peptide ion intensity ratios for each pair of extracted ion chromatograms (as shown in Figure 5B). Proteins upregulated by 1.5 fold in the SP population in duplicates samples are listed in Figure 5C.

Proteins downregulated by 1.5 fold in the SP population in duplicates samples are listed in Figure 5D.

Figure 6A.

Figure 6B.

Figure 6. Detection of identified Proteins in Mammalian Cells by Western Blotting and RT-PCR
 Equal amounts of SP and non SP protein were loaded on a gel and transferred to a membrane which then was probed for Thymosin beta four, SIAH interacting protein and Proliferation associated protein. An upregulation of all proteins was observed in SP population (A). The same result was observed with RT-PCR (B).

Figure 7.

Figure 7. Detection of RNAi-induced Protein Knockdown in Mammalian Cells by Western Blotting MCF7 SP and MDA-MB231 SP cells were transfected with siRNA against TB4 and SIN3. After 48h cell extracts were subjected to SDS-PAGE, followed by immunoblotting with the polyclonal rabbit anti Thymosin and anti-SIN3 antibodies.

Figure 8.

Figure 8. Knock-down of TB4 increases chemosensitivity *in vitro*

Paclitaxel-induced cell death increase by siRNA mediated knock-down of TB4 expression. Cells were transfected with TB4 siRNA using Lipofectamine RNAiMax. 48 h later, cells were incubated with Paclitaxel or Doxorubicin and viability was measured using an MTT assay. Results are expressed as percentages of cells still viable relative to non-treated control cells.