이학석사 학위논문

Differential Expression Profiling of Cell Surface Antigens of CSC like Cells Derived from a Breast Cancer Cell Line Using Mass Spectrometry

질량 분석기를 이용한 유방암세포에서 유래된
유사암줄기세포의 표면항원 발현차이 규명에 관한 연구

2012 년 8 월

서울대학교 대학원

분자의학 및 바이오제약학 전공

이 경 섭

Differential Expression Profiling of Cell Surface Antigens of CSC like Cells Derived from a Breast Cancer Cell Line Using Mass Spectrometry

질량 분석기를 이용한 유방암세포에서 유래된
유사암줄기세포의 표면항원 발현차이 규명에 관한 연구

지도교수 이 유 진
이 논문을 이학석사 학위논문으로 제출함
2012년 8월

서울대학교 대학원
분자의학 및 바이오제약학과
이 경 섭

이경섭의 이학석사 학위논문을 인준함
2012년 6월

위 원 장 _____ (인)

부위원장 _____ (인)

위 원 _____ (인)

Differential Expression Profiling of Cell Surface Antigens of CSC like Cells Derived from a Breast Cancer Cell Line Using Mass Spectrometry

By

Kyeong-seob Lee

(Directed by Eugene C. Yi, Ph.D)

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science in Molecular Medicine and Biopharmaceutical Sciences at the Seoul National University, Seoul, Korea

August, 2012

Approved by thesis committee:

Professor _____ Chairperson

Professor _____ Vice Chairperson

Professor _____

Abstract

Differential Expression Profiling of Cell Surface Antigens of CSC like Cells Derived from a Breast Cancer Cell Line Using Mass Spectrometry

Kyeong-seob Lee

Department of Molecular Medicine and Biopharmaceutical Sciences

WCU Graduate School of Convergence Science and Technology

Seoul National University

Cancer stem cells (CSCs) are a specific sub-population of tumor cells that are exhibiting self-renewal ability, tumorigenesis, chemotherapy resistance and metastasis. As their unique stemness properties are enabling them to escape conventional anti-cancer therapies, CSCs have become a prime interest for the development of novel therapeutic strategies for cancer treatment. However, comprehensive CSCs surface marker discovery has been limited due to their rapid progression and rare population. Recently we have obtained CSC-like cells derived from a human breast cancer cell line. The CSC-like cells exhibit not only CSC-like

properties, but they maintain undifferentiated state. To discover potential CSCs specific surface markers, we carried out in-depth analyses of cell surface protein profiling of CSC-like cells compared with the Non CSC-like cells. To improve the coverage of cell surface proteins representing a "snapshot" of a live cell surface, we adapted the advantages of using a cell-surface biotinylation technique followed by affinity purification, SDS-PAGE protein fractionation and mass spectrometry analysis. This strategy allowed the identification of proteins unique and up-regulated in the CSC-like cells. Among the total 1525 proteins identified, we found 94 membrane proteins including known CSC surface markers (such as CD44, Integrin β-1, CD133) that are up-regulated in the CSC-like cells compared to the Non CSC-like cells. We also identified several interesting proteins that are known to be involved in tumorigenesis, chemoresistance, and metastasis which are related with CSC properties by adapting PLGEM based signal to noise ratio into Ingenuity Pathway Analysis. We have validated those up-regulated surface antigens by immunoblot assays and found that differentially expressed membrane proteins of the CSC-like cells are related in the known functions of CSCs.

Key words: Cancer stem cell, Cell surface antigen, Biotinylation, Mass spectrometry, Protein

Student Number: 2010-24238

List of Abbreviations

Ab Antibody

ACN Acetonitrile, CH_3CN

ANXA1 Annexin A1

ATP1B1 Sodium/potassium-transporting ATPase subunit beta-1

CAV1 Caveolin-1

CD Cluster of differentiation

CD147 Basigin

CD44 Hyaluronate receptor

CD9 Cell growth-inhibiting gene 2 protein

CDH1 Cadherin-1

CLTCL1 Clathrin heavy chain 2

CSC Cancer stem cell

DB Database

DTT Dithiothreitol

EGFR Epidermal growth factor receptor

EPCAM Epithelial cell adhesion molecule

EPHA2 Ephrin type-A receptor 2

ERBB2 Receptor tyrosine-protein kinase erbB-2

EZR Ezrin

FA Formic acid, HCO_2H

FDR False discovery rate

GO Gene ontology

HRP Horseradish peroxidase

IAA Iodoacetamide

IPA Ingenuity pathway analysis

ITGA3 Integrin alpha3

ITGAV Integrin alpha V

ITGB4 Integrin beta4

LC-MS/MS Liquid chromatography and tandem mass spectrometry

LTQ Linear ion trap quadrupole

MeOH Methanol

MS Mass spectrometry

MS/MS Tandem mass spectrometry

NH_4HCO_3 Ammonium bicarbonate

NP-40 Nonyl phenoxypolyethoxylethanol

OG Oxidized glutathione

PAGE Poly acrylamide gel electrophoresis

PLGEM Power law global error model

PTM Post translational modification

RTK Receptor tyrosine kinase

SDS Sodium dodecyl sulfate

SLC16A3 Monocarboxylate transporter 4

SLC1A5 Neutral amino acid transporter B

SLC2A1 Solute carrier family 2, facilitated glucose transporter member 1

TACSTD2 Tumor-associated calcium signal transducer 2

WB Western blot

Table of Contents

List of Tables

List of Figures

1. Introduction

Cancer is a pathological status initiated by genetic instability and characterized by uncontrolled proliferation. According to the CSC hypothesis, cancer is initiated and driven by aberrant stem cells. These tumor initiating cells have self renewal properties, the capacity to generate progenitor stem cells and resistance to chemotherapeutic agents (Natarajan and FitzGerald 2007). Current cancer therapies frequently fail to eradicate tumors because of two reasons. One is ineffective targeting of anti-cancer drugs and the other is CSC's resistance to chemotherapeutic agents and radiation (Boman and Wicha 2008). Therefore, presenting CSC specific markers shed new light on the tumor treatment issues by initiating effective therapeutic approach for inhibition of CSC proliferation, induction of differentiation, and stimulation of apoptosis. However, analyzing CSC is a challenging subject in proteomic study due to their extremely low availability which was caused by rare population in tumor tissue and rapid progression. In the present study, we have employed CSC-like cells derived from human a breast cancer cells. The CSC-like cells remained not only CSC like properties which are characterized by chemoresistance, survival in hypoxic and acidotic environments, and high tumor initiating abilities, but also they maintained undifferentiated state (Sajithlal, Rothermund et al. 2010). These cell lines are also enable to

accommodate the enough sample size for various proteomic approaches.

To address effective cancer therapies with a paradigm shift in the mechanism of the therapeutic resistance and recurrence of tumor, we focused on an improved understanding of the cell surface proteins in CSC-like cells. The proteins anchored to the cell membrane and transmembrane proteins occupy more than 30 % of all proteins in the human genome (Wallin and von Heijne 1998), more than two third of the known protein targets for drugs (Stevens and Arkin 2000). In spite of their relevance, a general analysis and quantification of membrane proteins have been hampered due to their hydrophobicity and relatively low abundance (Scheurer, Rybak et al. 2004). The biotin directed affinity purification with chemical derivatives of biotin which have water-soluble reactive moiety and cleavable linker has been examined as effective method to overcome the restriction of systemic analysis of membrane proteins. The noteworthy advantage in this method is enabling analysis of *in situ* distribution of cell surface proteins through membrane protein isolation.

After reduction with a cleavable reactive derivative of biotin, cells were lysed in the presence of stringent detergents and membrane proteins were purified on streptavidin. Then, proteins were eluted and their amount was determined by reducing agent compatible BCA protein assay. The resulting proteins were then in-gel tryptic digested after fractionation by SDS-PAGE. The peptides were introduced into a tandem mass spectrometry analysis to identify the proteins present. Then, the protein identification was performed by SEQUEST algorithm-

based sorcerer platform. A label free protein quantification using spectral counts was employed to measure the signal to noise ratio between the CSC-like and Non CSC-like cells by Power Law Global Error Model (PLGEM). Furthermore, to increase reliability of quantitative information and the coverage of differentially expressed proteins, we applied technical triplet replications. Ultimately, the significant proteins detected by this method were uploaded to Ingenuity Pathway Analysis (IPA) for functional analysis to identify remarkable biological functions in CSC-like cells compared to Non CSC-like cells as well as meaningful proteins which are related in CSC properties. Then, the immunoblot assay was used to confirm the quantitative mass spectrometry results about the significantly expressed proteins.

2. Materials and Methods

2.1 Cell Culture

Cancer Stem Cell-like cells have been established from MDA-MB453. These sub-clones of MDA-MB453 were grown at 37 ℃ in a humidified 5 % CO_2 incubator in Dulbecco's Modified Eagle Medium (DMEM)-high glucose (Invitrogen, Gibco) supplemented with 10 % fetal bovine serum (FBS) (Invitrogen, Gibco) until approaching confluency (~90 %).

2.2 Biotinylation of Cell Surface Proteins

CSC-like cells (~10^8cells from MDA-MB453) were used for cell surface biotinylation as described previously (Scheurer, Rybak et al. 2005; Kischel, Guillonneau et al. 2008). CSC-like cells were rinsed three times with PBS, and incubated with 0.5 mg/ml EZ-link Sulfo-NHS-SS-biotin (Thermo Fisher Scientific, Bonn, Germany) for 30 min at RT on a rocking platform. The biotinylation reaction was terminated by addition of 50 mM Tris-HCl (pH 7.5) in PBS for 10 min. The cells were washed two times with PBS then harvested in 0.1 mM Oxidized glutathione (Sigma-Aldrich) solution in PBS. After centrifugation for 5 min at 500 G, the cells were lysed in lysis buffer containing 2 % NP-40, 0.2 % SDS, Protease inhibitor cocktail (Roche, Vilvoorde, Belgium) and 0.1 mM Oxidized glutathione

solution with sonication. The cell lysate was centrifuged for 10 min at 16,100 G, and the cleared supernatant contained 10 mg of total protein on each cell lines (determined by BCA protein assay, Pierce). Then, 1.5 ml of streptavidin Sepharose high performance (GE Healthcare, Piscataway, NJ, USA) slurry was washed four times with 1 ml of washing buffer A (1 % w/v NP-40 and 0.1 % w/v SDS in 0.1 mM Oxidized glutathione solution in PBS) before the cleared lysate was added. The samples were tumbled for 2 hours at 4 ℃ while rotating. Unbounded proteins were removed by washing two times with wash buffer A, two times with wash buffer B (0.1 % w/v NP-40 and 1 M NaCl in 0.1 mM Oxidized glutathione solution in PBS), two times with wash buffer C (100 mM Na_2CO_3 in PBS), and two times with wash buffer D (50 mM Tris-HCl in PBS, pH 7.5). Captured proteins were eluted from streptavidin Sepharose by incubation with elution buffer (100 mM DTT, 0.5 % w/v SDS and 0.1 % w/v NP-40 in PBS) for 30 min at 60 ℃. This step was repeated. To wipe samples, the samples were vortexed in 0.1X PBS then sonicated. This step was also repeated. Protein concentration in elutes was calculated by reducing agent compatible BCA protein assay. The collected elute volume reduced in a Speed Vac concentrator (Rybak, Scheurer et al. 2004).

2.3 Electrophoresis

The biotinylated protein samples were heated with 4X sample buffer and 50 mM DTT (final concentration) at 70 ℃ for 10 min prior to separation in a 4-12 % Bis-Tris NuPAGE minigel (Invitrogen, Carlsbad, CA). Total proteins (1 mg) of each

cell lines were separated. The gel was run at 150 V for 70 min. The gel was stained using Coomassie Blue stain for 15 min.

2.4 In-Gel Tryptic Digestion

Total ten fractions were destained in destaining buffer (15 % v/v Methanol and 10 % v/v Acetic acid in water). After excision of the gel slice, the slices were washed with 100 mM NH_4HCO_3 (pH 8.0) for 15 min on shaker, and soaked in 100 % ACN to dehydrate the gel for 15 min on shaker. This wash cycle was repeated three times. ACN was removed and the gel slices were dried in a Speed Vac concentrator (Labconco, Kansas City, MO) for 10-20 min. The dried gel pieces was reduced with 20 mM DTT in 100 mM NH_4HCO_3 (pH 8.0) at 60 ℃ for 1 hour and alkylated with 55 mM Iodoacetamide in 100 mM NH_4HCO_3 at RT for 45 min on rocker. Iodoacetamide was removed and the gel slices were dried in a Speed Vac concentrator for 20-30 min. Gel pieces were rehydrated in 50 mM NH_4HCO_3 containing 13.3 ng/μL trypsin (Sequencing grade, Promega , Madison, WI) on ice for 30 min, then incubated at 37 ℃ overnight. For extraction of peptides from gel pieces, extraction buffer A (5 % ACN, 0.1 % Formic acid) was added and sample tubes were incubated while shaking at room temperature for 15 min. Supernatants transferred to new tubes. Then, extraction buffer B (50 %ACN, 0.1 % Formic acid) was added and tubes were incubated while shaking for 15 min. Supernatants were collected and second extraction step was repeated. The extracts were taken to dryness in the Speed Vac concentrator and the residues resuspended in mass

spectrometry analysis buffer (2 % ACN, 0.1 % Formic acid).

2.5 LC-MS/MS Analysis

Tryptic peptides from each of the gel fraction were injected into LC-MS/MS system. LC-MS/MS was performed using an integrated system, which consisted of the EASY-nLC system (Proxeon Biosystems, Odense, Denmark) and LTQ-Velos mass spectrometer (Thermo Fisher Scientific, San Jose, California). The samples were loaded onto a C_{18} trap column (inner diameter of 75 μm) which was packed in-house. Eluted peptides from the trap column were subsequently loaded onto a resolving analytical column (inner diameter of 75 μm) which was packed in-house. The mobile phases consisted of 2 % ACN with 0.1 % Formic acid (Solvent A) and 98 % ACN with 0.1 % Formic acid (Solvent B). The samples were separated using a linear gradient 2~38 % for 90 min at a flow rate of 300 nL/min. Data-dependent MS acquisition conditions were as follows: the spray voltage was set at 1.9 kV and the heated capillary was set at 325 ℃. Each cycle of survey consisted of full MS scan at the mass range 400-1400 m/z and MS/MS scan for five most intense ions. Peptides were fragmented in the LTQ section using collision-induced dissociation and the normalized collision energy value was set at 35 %. In addition, previously fragmented peptides were excluded for 30 sec.

2.6 Database Search

LTQ raw data were converted to mzxml file which is search-available format using Trans Proteomic Pipeline provided by Institute for Systems Biology (TPP, version 4.4 http://www.proteomecenter.org). The transformed data were searched against the decoyed human IPI database (IPI.human.v3.87) by using the SEQUEST algorithm-based sorcerer platform with the following parameters: mass tolerance 1.5 amu, variable modification on methionine of 16 Da (oxidation) and cystein of 57 Da (carbamidomethylation). The specrta were confidently filtered with PeptideProphet probability ≧ 0.95, FDR < 1 % (FDR formula: $N_{\text{reverse database}}$ X2 / [$N_{\text{reverse database}}$ + $N_{\text{real database}}$]), and spectral counts ≧ 2 peptide sequence.

2.7 Data Analysis

The spectral count is the total number of spectra for one protein identification and used for label free quantification in proteomic analysis (Liu, Sadygov et al. 2004). Because Power Law Global Error Model (PLGEM, http://www.bioconductor.org) provides signal to noise (STN) ratio and *p*-value which demand the calculation of variance, PLGEM requires at least two replicates for the relative comparing analysis. Spectral count data from triplicates of CSC-like and Non CSC-like cells was examined using a PLGEM to identify statistically significant protein changes. To investigate detailed molecular information from output data, these files were uploaded to IPA. The uploaded Microsoft Excel spreadsheet file includes the relevant proteins with their spectral count, PLGEM

based *p*-value, PLGEM based signal to noise, and IPI accession number. The significance values for molecular function were calculated using the Fisher's exact test by comparing the number of proteins that participate in given function compared to the total number of occurrences of these proteins in all functional annotations stored in the ingenuity pathway knowledge base.

2.8 Immunoblot Assay

Each thirty μg of purified membrane faction product on two cell lines was loaded onto a 4-12 % Bis-Tris NuPAGE minigel and electrophoresed. Proteins were transferred to PVDF using the Invitrogen Xcell II blot module (Invitrogen, Carlsbad, CA). After blocking, Membranes were incubated for overnight at 4 ℃ with the following primary antibodies: anti-EGFR Ab (SC-71032, Santa Cruz), anti-CD147 Ab (MEM-M6/1, Genetax), anti-CD44 Ab (SC-53298, Santa Cruz), anti-VAMP3 Ab (SC-136162, Santa Cruz). Then, membranes were washed and incubated with corresponding HRP-conjugated secondary antibody. The washed membranes were developed with enhanced chemiluminescence reagent and the images were captured with LAS-4000 instrument (Fuji, Japan).

3. Results

3.1 Experimental Strategies

The purpose of this experiment was to profile the differential expression of CSC-like cell surface antigens. Our experimental strategies were divided as three stages (Fig.1). In first stage, membrane protein isolation on the live cells was examined using biotinylation based on previous study (Scheurer, Rybak et al. 2005). The biotinylation method had crucial benefit which enabled to profile steady state distribution of specific antigens for mining novel cell surface biomarkers compared to conventional sucrose gradient centrifugation method. After biotin directed affinity purification, electrophoresis was performed to avoid the sample loss in mass spectrometry analysis as well as to overcome limitation of identifying low abundance proteins due to sample complexity. In second stage, the fractionated proteins were in-gel trypsinized for analyzing by LC-MS/MS. Database search was fulfilled using Trans Proteomic Pipeline provided by Institute for Systems Biology (TPP, version 4.4), the SEQUEST algorithm-based sorcerer platform and IPI protein database. To obtain improved dataset, the output data was applied to signal to noise ratios which had been calculated by PLGEM based standard deviations with R-package program (ver.2.15). Then the acquired data were investigated by Ingenuity Pathway Analysis in order to further investigate functional clustering of

significantly expressed proteins. In final stage, the significant up-regulated membrane proteins which were related in cancer stemness and satisfied with PLGEM based p-value < 0.05 were chosen for verification. The verification of meaningful marker candidates was carried out by immunoblot assay to confirm the mass spectrometry quantitative results.

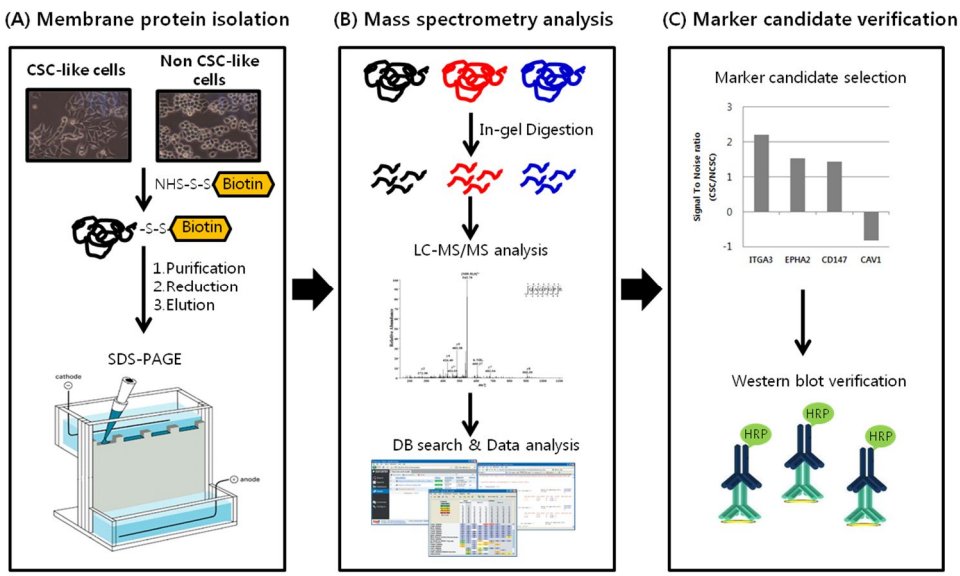


Figure 1. Outline of experimental strategies

Our experimental strategies were divided as three stages which consist of membrane protein isolation, mass analysis and verification of marker candidate proteins. (A)first stage, membrane protein isolation was examined using biotinylation and electrophoresis. (B)second stage, LTQ-mass spectrometry and software tools were used to profile significant proteins. (C)final stage, marker candidates were verified by Western blot assay.

3.2 Proteome Analysis of CSC-like Cell Specific Gel Bands

To screen proteins which were differentially expressed on CSC-like cell surface in comparison with Non CSC-like cells, equal amounts of 60 μg protein elutes were separated in an SDS-PAGE gel after biotin directed affinity purification. As a result, we visualized unique and up-regulated bands on CSC-like cell compared to Non CSC-like cells (Fig.2). Among the appeared bands, the nine distinguishable bands were selected to perform in-gel digestion. Peptides assignment was performed with the Trans Proteomic Pipeline, the Sorcerer platform and IPI database. These peptides with probabilities ≧ 0.95, FDR < 1 % and spectral counts ≧ 2 were included in the following PeptideProphet search. After analyzing with mass spectrometry data, remarkable proteins were identified and the currently used cancer stem cell markers were included (Natarajan and FitzGerald 2007). In addition, cancer physiology related proteins were identified. Figure 3 shows the representative LTQ MS/MS spectra of in-gel digested peptides of CSC-like cell specific bands. The peptides were selected by criteria which contain accordance of molecular weight on identified protein with distinguishable band, uniqueness on the identified protein, meaningful signal intensity, and the total number of fragment ions.

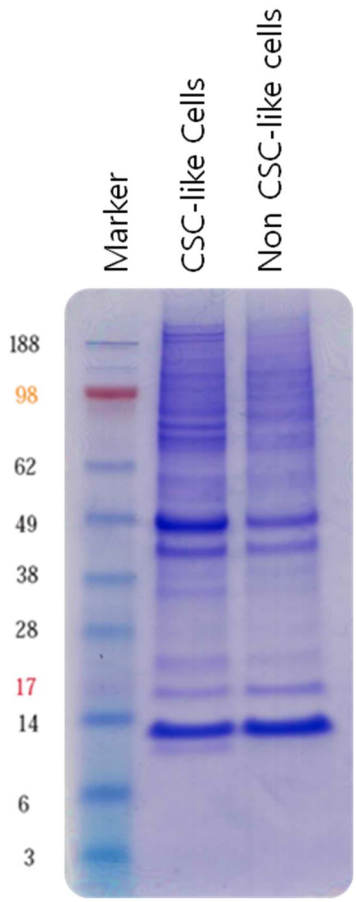
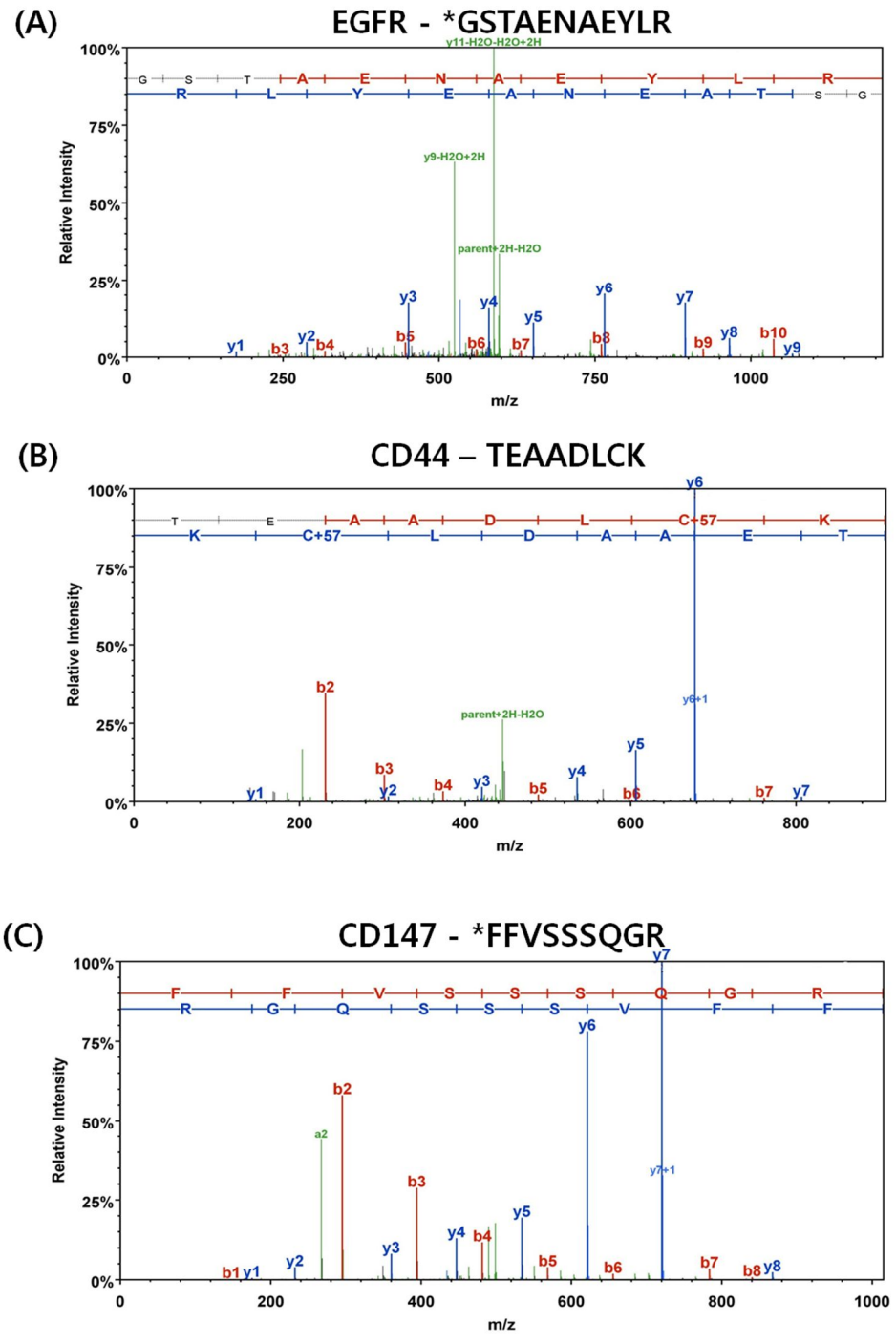


Figure 2. Visualization of CSC-like cell specific bands

After BCA protein quantification, 60 µg of total protein was separated by preparative 4-12 % Bis-Tris SDS-PAGE. Stained with Coomassie Blue, nine gel band regions were excised and subjected to an in-gel digestion with trypsin. Lane 1 molecular weight marker, lane 2 Oct3/4 transfected MDA-MB-453 GFP (+), lane 3 Oct3/4 transfected MDA-MB-453 GFP (-), CSC-like cells showed different protein profiles to those of Non CSC-like cells.


(A)
EGFR - *GSTAENAEYLR
y11-H2O-H2O+2H
G S T A E N A E Y L R
R L Y E A N E A T S G
y9-H2O+2H
parent+2H-H2O
Relative Intensity
m/z
(B)
CD44 – TEAADLCK
T E A A D L C+57 K
K C+57 L D A A E T
y6
y6+1
b2
parent+2H-H2O
Relative Intensity
m/z
(C)
CD147 - *FFVSSSQGR
F F V S S S Q G R
R G Q S S S V F F
y7
y6
b2
a2
y7+1
Relative Intensity
m/z

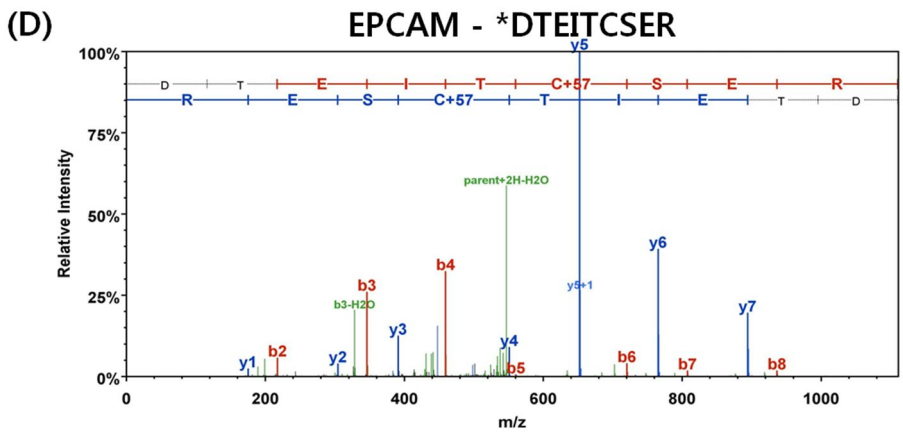


Figure 3. Tandem mass spectra of the peptides in the CSC-like cell specific bands detected by LC-MS/MS

(A-D) MS/MS spectra of the peptides, GSTAENAEYLR, TEAADLCK, FFVSSSQGR, and DTEITCSER that represent EGFR, CD44, CD147 and EPCAM. Criteria of these peptides selection are the identified protein's molecular weight accordance about distinguishable band, uniqueness on corresponding protein, significant signal intensity, and the total number of fragment ions. Red and blue colors present the b fragment ions and y fragment ions, which were assigned by SEQUEST protein database search. The asterisk indicates unique peptide corresponding identified protein.

3.3 Indepth Analysis of CSC-like and Non CSC-like Cells

To study further distribution on cell surface antigens in CSC-like cells and Non CSC-like cells based on screening data, we examined in-depth analysis about each 1 mg eluent by separating into ten fractions with SDS-PAGE. We performed in-gel digestion to ten fractions then analyzed with mass spectrometry. In order to check the difference in variation from digested peptides population during technical processes, we conducted protein profiling with triplet replications. Peptides assignment was carried out with the Trans Proteomic Pipeline, the Sorcerer platform and IPI database. These peptides were filtered by probabilities ≧ 0.95, FDR < 1 % and spectral counts ≧ 2. The average number of total identified proteins in each triplet replication was 1100. Also, 51 % proteins of total identified protein were overlapped between CSC-like and Non CSC-like cells on average (Fig.4). Overall, 54.4 % proteins were overlapped in total 1525 proteins identified from the three experiments (Fig.5, C). In triplet replications, total 1208 proteins were identified in the CSC-like cells, and total 1146 proteins were identified in Non CSC-like cells (Fig.5, A and B). Spectral count and signal to noise ratio from PLGEM were used to measure relatively differential expression of targeted proteins. The Power Law Global Error Model (PLGEM) is useful method for spectral count quantification by providing an efficient and informative screening strategy as well as reliable and robust biosignature mapping of samples (Lee, McKinney et al. 2011). Signal to noise ratio means that the standard deviation of a protein's spectral abundance factor value which dependant on the average spectral

abundance factor value of the protein itself, following a power law. The important common feature identified by this method is that PLGEM derived signal to noise ratio allowed the detection of a higher number of truly differentially expressed proteins without increasing the false positive rate (Pavelka, Pelizzola et al. 2004). Whereas the fold change statistic value which was employed in classical method missed detection of the meaningful proteins and embarrassed comparison of some protein counted by zero, PLGEM overcame not only these problems, but limitation of the coefficient of variation which becomes progressively smaller as the average of replicated abundance values increase more. PLGEM also showed reasonable stability to decreasing numbers of replicates (Pavelka, Fournier et al. 2008). We applied PLGEM in the investigation of total 1525 proteins that identified by triplet replications. Among the investigated proteins, the significantly quantifiable 93 proteins that satisfied the PLGEM p-value < 0.05 and membrane fraction were identified (Table1). The meaningfully up-regulated 34 proteins among the 93 proteins and breast cancer stem cell markers (such as CD44, Integrin β-1, CD133) were revealed in CSC-like cells compared Non CSC-like cells (Table2). The down regulated 59 proteins were also identified (Table3).

To obtain overview into the biological changes of expressed proteome in the CSC-like cells versus Non CSC-like cells, the differentially expressed proteins were classified by Ingenuity Pathway Analysis (IPA) (Ingenuity® Systems, www.ingenuity.com.) (Dai, Li et al. 2010). Also, to gain the biological interpretation of these proteins, we applied Human Protein Reference Database

(HPRD) (http://www.hprd.org, release 8, July 2009). The overrepresented processes in the biological process categorized according to the integrated database based on Gene Ontology (GO) were involved in cellular movement, cellular growth and proliferation, molecular transport, cell to cell signaling and interaction, cell death, immune response, and cell morphology (Fig.6). These categories were selected by cutoff criteria which had applied a threshold p-value < 0.001 based on Fisher's exact test. The enriched categories showed three major functions related to the characteristics of cancer stem cells: tumorigenesis, chemoresistance, and metastasis. The top molecular function in membrane fraction using IPA is cellular movement related to metastasis. Moreover, cellular growth and proliferation is associated with tumorigenesis, and molecular transport is involved to chemoresistance.

Figure 7 shows the top four significantly expressed proteins in CSC-like cells, related to colony formation (ITGB4, EGFR, ANXA1, EPCAM), survival of tumor cells (CD44, CDH1, SLC2A1, EGFR), transportation of molecules (SLC2A1, SLC16A3, ATP1B1, CLTCL1) and invasion of tumor cells (ITGA3, EZR, EPHA2, CD147) in subcategories of the three major molecular functions related to CSC properties. The proteins related to colony formation play a role in tumor initialization. Epithelial cell adhesion molecule which is a well known tumor marker has a direct impact on cell proliferation and the ability to rapidly up-regulate the proto-oncogene *c-myc*. Human epithelial 293 cells as well as murine NIH3T3 fibroblasts expressing EPCAM had a decreased requirement for growth

factors, enhanced metabolic activity and colony formation capacity (M, C et al. 2004). Annexin family have been implicated in cellular processes, including modulation of phospholipase A2 activity and inflammation, immune response, proliferation, differentiation, membrane skeletal linkage, and intracellular signal transduction. Annexin A1 is a good substrate of receptor- or non-receptor tyrosine kinases and interacts calcium-dependently with protein kinase C. Annexin A1 is strongly involved in the modulation of cell functions by controlling intracellular calcium release. MCF-7 subclones which overexpress Annexin A1 displayed an enhanced proliferation rate, anchorage independent growth and dramatic morphological alterations including loss of epithelioid appearance (BM, BF et al. 1999).

The molecular functions of survival of tumor cells and transportation of molecules reflect chemoresistance property in cancer stem cells. CD44 is a multifunctional proteins involved in cell adhesion, migration and drug resistance as well as one of the important cell surface markers for cancer stem cell (Ezeh, Turek et al. 2005). Hyaluronan binding to CD44 up-regulates p300 expression and its acetyltransferase activity that promotes acetylation of β-catenin and NFκB-p65 leading to activation of β-catenin associated T-cell factor/lymphocyte enhancer factor transcriptional co-activation and NFκB specific transcriptional up-regulation, respectively. These change cause the expression of the MDR1 (multidrug resistance1) gene and the anti-apoptotic gene *Bcl-x*$_L$ resulting in chemoresistance in breast cancer cells (LY, W et al. 2009). SLC16A3 is monocarboxylate transporter 4

which is one of the solute carrier proteins. According to the study of the ovarian cancer chemotherapy responses, many common genes may be the key players in conferring chemotherapy resistance (Cheng, Lu et al. 2010). For example, the solute carrier proteins such as SLC16A3 belong to the common gene list, suggesting that solute carrier proteins may modulate the influx/efflux of drugs from cells and thus modulate chemotherapy response. Moreover, SLC16A3 and SLC16A1 co-localized with CD147, CD44, and MDR1 in metastatic prostate cancer cell lines, and demonstrated that overexpression of CD147, CD44, MDR1, SLC16A1, and SLC16A3 is related to the drug resistance in prostate cancer progression (Hao, Chen et al. 2010).

Invasion of tumor cells is a representative characteristic associated with cancer metastasis. Degradation of basement membrane by Matrix MetalloProteinases (MMPs) has been implicated in several aspects of tumor invasion, metastasis, progression, growth, and angiogenesis. We found the proteins correlated with MMPs among the significantly expressed proteins in this study as well as the proteins in invasion of tumor cells category. CD147 is a heavily glycosylated protein enriched on the surface of tumor cells and stimulates the production of MMPs. CD147 induced the production of secreted MMP not only by dermal fibroblasts but also by breast cancer cell line MDA-435, suggesting homophilic CD147 binding may play a key role in MMP-2 (gelatinase A) production and tumor cell invasion (J and ME 2001). In connection with MMP-2, the production of MMP-2 by hyaluronan stimulation in a human cancer cell line

QG90 that expresses a large amount of CD44s was investigated. Treatment of QG90 with hyaluronan strongly activated MMP-2 secretion in a time- and dose-dependent manner. Moreover, pre-treatment of cells with the neutralizing anti-CD44 antibody significantly inhibited both the hyaluronan dependent MMP-2 secretion (Y, AA et al. 2002). SLC2A1 is also related to MMP-2. SLC2A1 is glucose transporter 1 expressed in cancer with high levels. MMP-2 and SLC2A1 are co-expressed in human cancer cell lines. Overexpression of SLC2A1 in the human rhabdomyosarcoma cell line RD increased MMP-2 expression 4.3 fold and invasiveness 3.2 fold relative to control cells. Conversely, suppression of SLC2A1 expression decreased MMP-2 expression by 71.5 % and invasiveness by 53.0 % (S, T et al. 2002). Thus, MMPs are crucial proteins correlated with not only metastasis but also cancer stemness. Ezrin is a member of the ezrin/radixin/moesin (ERM) family of proteins that regulate cytoskeletal related functions such as cell adhesion, cell survival, and cell motility, all of which are important in tumor progression. High expression of Ezrin is associated with a more aggressive phenotype in human osteosarcomas, rhabdomyosarcomas, and melanomas as well as in endometrial cancers (Yu, Khan et al. 2004). Recently, abnormal Ezrin localization has been associated with adverse tumor characteristics in breast cancer (Sarrio, Rodriguez-Pinilla et al. 2006). Ezrin has been shown to bind directly to phosphatidylinositol 3-kinase (PI3K) and influence a number of signaling pathways that affect cellular functions related to tumorigenesis and metastasis, including the mitogen activated protein kinase / extracellular signal regulated kinase 1/2 (MAPK-ERK1/2), PI3K-

AKT, and Rho pathways. Ezrin mediated effects on AKT and ERK1/2 activity have been linked to the ability of Ezrin to promote tumor progression and metastasis (Elliott, Meens et al. 2005). EPHA2 is ephrin type-A receptor normally expressed at low levels in adult epithelial tissues (Lindberg and Hunter 1990) and is also a powerful onco-protein sufficient to confer malignant potential. Suppression of EPHA2 expression impairs *in vitro* cellular invasiveness and anoikis resistance (MS, H et al. 2004). In developing embryonic tissues, EPH and ephrin expression occurs in reciprocal embryonic compartments, with ligand-receptor interaction taking place at tissue boundaries (Gale, Holland et al. 1996). These interactions are postulated to prevent aberrant cellular migration across embryonic boundaries, thus maintaining normal tissue organization. In malignant tissues characterized by deranged tissue organization, differential expression and functional alterations of EPH have been reported to be prevalent (Kinch and Carles-Kinch 2003).

Furthermore, we also found meaningful down-regulated proteins related to cancer stemness. Caveolin-1 expression correlates with the level of oncogenic transformation, suggesting that caveolin may regulate cancer cell proliferation. The disruption of caveolae integrity appears to be a general feature of oncogenic transformation of fibroblast (Koleske, Baltimore et al. 1995). Also, p53-deficient cells showed a loss in caveolin expression (SW, CL et al. 1998). Caveolin-1 is associated with MMPs as well. The inhibition of anchorage-independent growth induced by caveolin-1 is shown with significant reduction in MMP-2 release and impaired invasive potential. In addition, plating MCF-7 cells on a laminin matrix

resulted in activation of ERK1/2, which was dramatically inhibited in MCF-7 cells transfected with caveolin-1 (G, D et al. 2002). Down-regulation of Caveolin-1 is involved in tumor growth and movement. ErbB2 is a member of receptor tyrosine kinase ErbB family. Especially, EGFR (ErbB1) is associated with relatively aggressive tumors of the stomach, bladder, lung, and breast. Decreased heterodimerization of ErbB2 and ErbB3 led to a reorganization in ErbB function as heterodimerization between EGFR and ErbB3 increase (YP, S et al. 2005). In other words, elimination of ErbB2 signaling results in and increase in EGFR expression and activation. Increased EGFR activation contributes to the sustained cell survival.

Collectively, the significantly expressed membrane proteins in triplet replications are categorized into three major molecular functions: tumorigenesis, chemoresistance, and metastasis, which are strongly correlated with cancer stem cell properties (Fig.8).

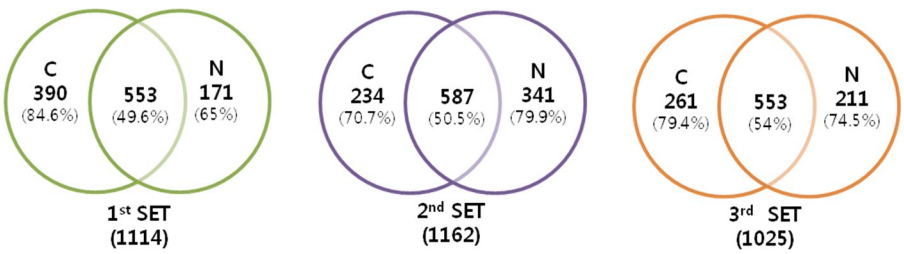


Figure 4. Venn diagrams displaying of identified proteins in triplet replications

Venn diagrams depict the overlap of proteins identified from the each triplet replication sets. The average 1100 proteins on each set were identified, and the overlapped proteins occupied 51 percentages of total identified proteins in each set.

C: CSC like cells, N: Non CSC-like cells

Criteria of identification: probability ≧ 0.95, FDR < 1 %, Spectral counts ≧ 2

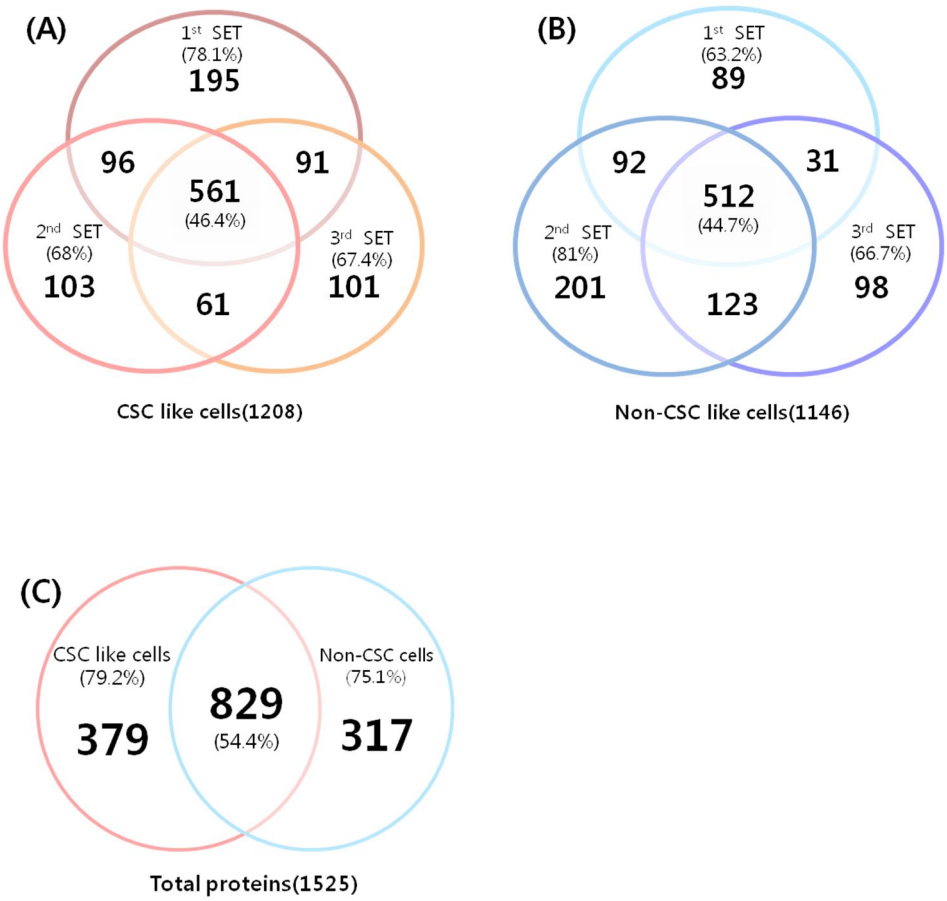


Figure 5. Overlapped proteins of triplicated experiments in the CSC-like and Non CSC-like cells

(A)Total 1208 proteins were identified in the CSC-like cells. (B)Total 1146 proteins were identified in Non CSC-like cells. (C)Total 1525 proteins were identified in three replications, only 829 proteins were identified between the CSC-like and Non CSC-like cells.

Criteria of identification: probability ≧ 0.95, FDR < 1 %, Spectral counts ≧ 2

Table 1. Membrane proteins identified from the CSC-like and Non CSC-like cells

No.	ID	Gene symbol	Protein name	Spectral counts of Non-CSC like cells			Spectral counts of CSC like cells				
				[a]N1	N2	N3	[b]C1	C2	C3	[c]STN	[d]ρ-value
1	IPI00027493	SLC3A2	4F2 cell-surface antigen heavy chain	820	656	376	275	280	180	-7.1537	0.0003
2	IPI00297160	CD44	CD44	0	0	0	120	107	138	7.2084	0.0003
3	IPI00028911	DAG1	Dystroglycan	95	136	63	6	5	3	-5.7656	0.0004
4	IPI00027422	ITGB4	Integrin beta-4	0	2	0	62	57	70	4.8163	0.0005
5	IPI00027505	ITGAV	Integrin alpha-V	101	108	84	23	20	8	-4.2940	0.0005
6	IPI00015688	GPC1	Glypican-1	80	82	71	12	5	8	-4.3930	0.0005
7	IPI00015102	ALCAM	CD166	149	85	101	37	21	32	-3.8113	0.0009
8	IPI00015476	SLC1A4	Neutral amino acid transporter A	35	36	46	0	0	0	-3.5287	0.0011
9	IPI00002406	BCAM	Basal cell adhesion molecule	35	31	53	5	2	5	-3.1407	0.0012
10	IPI00019146	CXADR	Coxsackievirus and adenovirus receptor	56	50	82	11	14	17	-3.0854	0.0012
11	IPI00383597	SLC4A2	Anion exchange protein 2	25	31	32	0	0	0	-2.9051	0.0013
12	IPI00003373	OCLN	Occludin	36	45	67	6	6	18	-2.8385	0.0013
13	IPI00152540	CD109	CD109	31	39	51	6	9	12	-2.4534	0.0019
14	IPI00029741	ITGB5	Integrin beta-5	27	20	17	0	0	0	-2.3108	0.0024
15	IPI00014371	CDH18	Cadherin-18	18	19	26	0	0	0	-2.2839	0.0024
16	IPI00001754	F11R	Junctional adhesion molecule A	60	40	32	17	10	10	-2.2917	0.0024
17	IPI00215995	ITGA3	Integrin alpha-3	4	0	3	25	27	22	2.3402	0.0026
18	IPI00006482	ATP1A1	Sodium/potassium-transporting ATPase subunit alpha-1	12	17	34	43	50	77	2.1771	0.0027
19	IPI00645614	CDH3	Isoform 2 of Cadherin-3	0	0	0	34	9	15	2.1465	0.0028
20	IPI00026952	PKP3	Plakophilin 3b (Fragment)	27	30	30	8	4	8	-2.0383	0.0028
21	IPI00394851	KLRG2	Killer cell lectin-like receptor subfamily G member 2	13	17	24	0	0	0	-2.0325	0.0028
22	IPI00000513	CDH1	Cadherin-1	7	6	9	33	28	34	2.1267	0.0029
23	IPI00220194	SLC2A1	Solute carrier family 2, facilitated glucose transporter member 1	4	3	7	18	25	34	2.1060	0.0029
24	IPI00843975	EZR	Ezrin	9	12	5	40	44	12	1.9781	0.0031
25	IPI00396658	ITFG3	Protein ITFG3	18	24	5	0	0	0	-1.8229	0.0035
26	IPI00006666	SLC16A3	Monocarboxylate transporter 4	0	0	0	22	13	13	1.8537	0.0037
27	IPI00168813	PTK7	Inactive tyrosine-protein kinase 7	0	0	0	15	19	12	1.7918	0.0038
28	IPI00008986	SLC7A5	Large neutral amino acids transporter small subunit 1	39	34	88	14	16	44	-1.7442	0.0039
29	IPI00386170	SSFA2	Sperm-specific antigen 2	0	0	0	11	13	18	1.6641	0.0045
30	IPI00003021	ATP1A2	Sodium/potassium-transporting ATPase subunit alpha-2	3	5	9	15	21	29	1.6443	0.0046
31	IPI00021267	EPHA2	Ephrin type-A receptor 2	3	0	3	13	17	15	1.6047	0.0047
32	IPI00018274	EGFR	Epidermal growth factor receptor	0	0	0	16	12	12	1.5982	0.0048
33	IPI00022048	PTGFRN	Prostaglandin F2 receptor negative regulator	36	17	44	13	14	13	-1.4872	0.0053
34	IPI00298994	TLN1	Talin-1	7	12	5	27	28	18	1.5269	0.0055
35	IPI00002478	ECE1	Endothelin-converting enzyme 1	16	18	11	4	4	0	-1.4677	0.0056
36	IPI00017562	LPHN2	Latrophilin-2	6	13	17	0	0	0	-1.4614	0.0057
37	IPI00215997	CD9	CD9	85	31	34	30	29	20	-1.4331	0.0059
38	IPI00024307	EFNB1	Ephrin-B1	12	8	15	0	0	0	-1.4261	0.0059
39	IPI00232571	GPC4	Glypican-4	7	12	15	0	2	0	-1.3903	0.0061
40	IPI00437751	ACE	Angiotensin-converting enzyme	246	16	17	331	42	29	1.4613	0.0062
41	IPI00026602	HLA-B	HLA class I histocompatibility antigen, B-41 alpha chain	10	9	13	21	43	17	1.4049	0.0065
42	IPI00019906	BSG	CD147 (Basigin)	57	66	80	91	91	120	1.3582	0.0069
43	IPI00747849	ATP1B1	Sodium/potassium-transporting ATPase subunit beta-1	18	11	9	25	27	36	1.3548	0.0070
44	IPI00021975	TNFRSF10A	Tumor necrosis factor receptor superfamily member 10A	22	26	13	7	10	6	-1.2588	0.0081
45	IPI00747243	CDH3	Isoform 1 of Cadherin-3	0	0	0	0	17	12	1.2799	0.0082
46	IPI00300384	ERBB2	Receptor tyrosine-protein kinase erbB-2	9	8	13	0	0	0	-1.2420	0.0082
47	IPI00299412	CD97	CD97	0	0	0	15	12	3	1.2420	0.0086
48	IPI00334190	STOML2	Stomatin-like protein 2	22	25	11	8	10	5	-1.1767	0.0092
49	IPI00217766	SCARB2	Lysosome membrane protein 2	6	11	17	0	0	5	-1.1655	0.0094
50	IPI00219682	STOM	Erythrocyte band 7 integral membrane protein	13	12	3	0	0	0	-1.1643	0.0097

51	IPI00843765	SPTAN1	Spectrin alpha chain, brain	12	15	3	3	2	0	-1.1621	0.0097
52	IPI00022558	MPZL1	Myelin protein zero-like protein 1	12	12	7	3	0	3	-1.1255	0.0105
53	IPI00002483	SLC30A1	Zinc transporter 1	6	6	15	2	0	0	-1.1245	0.0107
54	IPI00025846	DSC2	Desmocollin-2	0	0	0	7	5	15	1.1245	0.0109
55	IPI00005614	SPTBN1	Spectrin beta chain, brain 1	12	9	15	4	5	0	-1.1068	0.0111
56	IPI00002070	LRRC8A	Leucine-rich repeat-containing protein 8A	7	8	13	0	0	3	-1.0845	0.0116
57	IPI00015756	PTPRK	Receptor-type tyrosine-protein phosphatase kappa	7	8	11	0	0	0	-1.0840	0.0121
58	IPI00009236	CAV1	Caveolin-1	30	24	30	11	16	18	-1.0354	0.0132
59	IPI00291175	VCL	Vinculin	3	6	15	0	0	0	-1.0008	0.0146
60	IPI00760554	HLA-A	HLA class I histocompatibility antigen, A-69 alpha chain	9	7	11	0	4	0	-0.9705	0.0157
61	IPI00026241	BST2	Bone marrow stromal antigen 2	9	5	9	0	0	0	-0.9579	0.0162
62	IPI00013449	TSPAN6	Tetraspanin-6	16	4	15	5	0	6	-0.9534	0.0163
63	IPI00297910	TACSTD2	Tumor-associated calcium signal transducer 2	0	0	0	6	5	12	0.9579	0.0164
64	IPI00024067	CLTC	Clathrin heavy chain 1	4	8	0	7	14	13	0.8621	0.0215
65	IPI00296379	TNFRSF10B	Tumor necrosis factor receptor superfamily member 10B	0	0	5	5	9	12	0.8604	0.0215
66	IPI00418169	ANXA2	Annexin A2	16	18	21	24	35	30	0.8524	0.0219
67	IPI00000070	LDLR	Low-density lipoprotein receptor	3	13	13	4	4	3	-0.8503	0.0220
68	IPI00009507	SYPL1	Synaptophysin-like protein 1	15	10	13	6	5	6	-0.8448	0.0222
69	IPI00218487	GJA1	Gap junction alpha-1 protein	0	13	5	0	0	0	-0.8238	0.0242
70	IPI00183082	TMEM51	Transmembrane protein 51	0	5	13	2	0	0	-0.8238	0.0242
71	IPI00218918	ANXA1	Annexin A1	31	24	30	41	47	36	0.8158	0.0250
72	IPI00022881	CLTCL1	Clathrin heavy chain 2	0	3	0	4	12	5	0.7901	0.0262
73	IPI00002459	ANXA6	Annexin A6	9	5	5	0	0	0	-0.7771	0.0273
74	IPI00170692	VAPA	Vesicle-associated membrane protein-associated protein A	15	10	17	8	6	8	-0.7433	0.0297
75	IPI00018860	ULBP2	NKG2D ligand 2	9	7	7	2	0	5	-0.7347	0.0307
76	IPI00290039	CDCP1	CUB domain-containing protein 1	33	18	34	24	10	22	-0.7331	0.0308
77	IPI00456359	ATXN2L	Ataxin-2-like protein	0	0	0	3	0	13	0.7291	0.0312
78	IPI00023942	SDC3	Syndecan	0	5	11	0	0	0	-0.7291	0.0316
79	IPI00022284	PRNP	Major prion protein	9	10	13	2	6	8	-0.6837	0.0360
80	IPI00943891	LOC100132336	GDNF family receptor alpha-2-like	28	0	0	12	0	0	-0.6837	0.0360
81	IPI00013744	ITGA2	Integrin alpha-2	0	0	0	5	2	10	0.6799	0.0361
82	IPI00793199	ANXA4	Annexin A4	0	0	0	8	6	3	0.6799	0.0361
83	IPI00019472	SLC1A5	Neutral amino acid transporter B	89	56	78	102	72	96	0.6498	0.0397
84	IPI00000059	CD58	CD58	6	5	7	3	2	0	-0.6499	0.0399
85	IPI00215948	CTNNA1	Catenin alpha-1	3	4	0	6	10	5	0.6465	0.0401
86	IPI00026650	HLA-C	HLA class I histocompatibility antigen, Cw-1 alpha chain	0	10	9	0	4	3	-0.6465	0.0410
87	IPI00018213	LRRC8D	Leucine-rich repeat-containing protein 8D	3	10	3	0	0	0	-0.6293	0.0454
88	IPI00168728	IGHM	FLJ00385 protein (Fragment)	0	12	0	0	2	0	-0.6293	0.0454
89	IPI00472013	HLA-A	HLA-A HLA class I histocompatibility antigen, A-33 alpha chain	15	7	7	6	4	5	-0.6228	0.0463
90	IPI00009111	TPBG	Trophoblast glycoprotein	6	3	0	11	6	6	0.6059	0.0486
91	IPI00006211	VAPB	Isoform 1 of Vesicle-associated membrane protein-associated protei	16	15	28	9	10	20	-0.6017	0.0488
92	IPI00472138	HLA-B	HLA class I histocompatibility antigen, B-58 alpha chain	3	5	9	0	2	3	-0.6007	0.0494
93	IPI00011578	NPTN	Neuroplastin	6	12	9	9	0	3	-0.5982	0.0498

[a]N1,N2,N3 : Spectral counts which were investigated in each triplet replication of Non-CSC like cells [b]C1,C2,C3 : Spectral counts which were investigated in each triplet replication of CSC like cells
[c]STN : Signal to noise raio generated by PLGEM [d]p-value calculated by PLGEM analysis

Table 2. Up-regulated membrane proteins in the CSC-like cells

No.	ID	Gene symbol	Protein name	Spectral counts of Non-CSC like cells			Spectral counts of CSC like cells				
				[a]N1	N2	N3	[b]C1	C2	C3	[c]STN	[d]p-value
1	IPI00297160	CD44	CD44	0	0	0	120	107	138	7.2084	0.0003
2	IPI00027422	ITGB4	Integrin beta-4	0	2	0	62	57	70	4.8163	0.0005
3	IPI00215995	ITGA3	Integrin alpha-3	4	0	3	25	27	22	2.3402	0.0026
4	IPI00006482	ATP1A1	Sodium/potassium-transporting ATPase subunit alpha-1	12	17	34	43	50	77	2.1771	0.0027
5	IPI00645614	CDH3	Isoform 2 of Cadherin-3	0	0	0	34	9	15	2.1465	0.0028
6	IPI00000513	CDH1	Cadherin-1	7	6	9	33	28	34	2.1267	0.0029
7	IPI00220194	SLC2A1	Solute carrier family 2, facilitated glucose transporter member 1	4	3	7	18	25	34	2.1060	0.0029
8	IPI00843975	EZR	Ezrin	9	12	5	40	44	12	1.9781	0.0031
9	IPI00006666	SLC16A3	Monocarboxylate transporter 4	0	0	0	22	13	13	1.8537	0.0037
10	IPI00168813	PTK7	Inactive tyrosine-protein kinase 7	0	0	0	15	19	12	1.7918	0.0038
11	IPI00386170	SSFA2	Sperm-specific antigen 2	0	0	0	11	13	18	1.6641	0.0045
12	IPI00003021	ATP1A2	Sodium/potassium-transporting ATPase subunit alpha-2	3	5	9	15	21	29	1.6443	0.0046
13	IPI00021267	EPHA2	Ephrin type-A receptor 2	3	0	3	13	17	15	1.6047	0.0047
14	IPI00018274	EGFR	Epidermal growth factor receptor	0	0	0	16	12	12	1.5982	0.0048
15	IPI00298994	TLN1	Talin-1	7	12	5	27	28	18	1.5269	0.0055
16	IPI00437751	ACE	Angiotensin-converting enzyme	246	16	17	331	42	29	1.4613	0.0062
17	IPI00026602	HLA-B	HLA class I histocompatibility antigen, B-41 alpha chain	10	9	13	21	43	17	1.4049	0.0065
18	IPI00019906	BSG	CD147 (Basigin)	57	66	80	91	91	120	1.3582	0.0069
19	IPI00747849	ATP1B1	Sodium/potassium-transporting ATPase subunit beta-1	18	11	9	25	27	36	1.3548	0.0070
20	IPI00747243	CDH3	Isoform 1 of Cadherin-3	0	0	0	0	17	12	1.2799	0.0082
21	IPI00299412	CD97	CD97	0	0	0	15	12	3	1.2420	0.0086
22	IPI00025846	DSC2	Desmocollin-2	0	0	0	7	5	15	1.1245	0.0109
23	IPI00297910	TACSTD2	Tumor-associated calcium signal transducer 2	0	0	0	6	5	12	0.9579	0.0164
24	IPI00024067	CLTC	Clathrin heavy chain 1	4	8	0	7	14	13	0.8621	0.0215
25	IPI00296379	TNFRSF10B	Tumor necrosis factor receptor superfamily member 10B	0	0	5	5	9	12	0.8604	0.0215
26	IPI00418169	ANXA2	Annexin A2	16	18	21	24	35	30	0.8524	0.0219
27	IPI00218918	ANXA1	Annexin A1	31	24	30	41	47	36	0.8158	0.0250
28	IPI00022881	CLTCL1	Clathrin heavy chain 2	0	3	0	4	12	5	0.7901	0.0262
29	IPI00456359	ATXN2L	Ataxin-2-like protein	0	0	0	3	0	13	0.7291	0.0312
30	IPI00013744	ITGA2	Integrin alpha-2	0	0	0	5	2	10	0.6799	0.0361
31	IPI00793199	ANXA4	Annexin A4	0	0	0	8	6	3	0.6799	0.0361
32	IPI00019472	SLC1A5	Neutral amino acid transporter B	89	56	78	102	72	96	0.6498	0.0397
33	IPI00215948	CTNNA1	Catenin alpha-1	3	4	0	6	10	5	0.6465	0.0401
34	IPI00009111	TPBG	Trophoblast glycoprotein	6	3	0	11	6	6	0.6059	0.0486
35	IPI00217561	ITGB1	Integrin beta-1	101	90	90	127	94	105	**Breast Cancer Stem Cell Marker**	
36	IPI00012540	PROM1	CD133 (Prominin-1)	0	0	0	4	4	3		

[a]N1,N2,N3 : Spectral counts which were investigated in each triplet replication of Non-CSC like cells [b]C1,C2,C3 : Spectral counts which were investigated in each triplet replication of CSC like cells
[c]STN : Signal to noise raio generated by PLGEM [d]p-value calculated by PLGEM analysis

Table 3. Down-regulated membrane proteins in the CSC-like cells

No.	ID	Gene symbol	Protein name	Spectral counts of Non-CSC like cells			Spectral counts of CSC like cells				
				[a]N1	N2	N3	[b]C1	C2	C3	[c]STN	[d]p-value
1	IPI00027493	SLC3A2	4F2 cell-surface antigen heavy chain	820	656	376	275	280	180	-7.1537	0.0003
2	IPI00028911	DAG1	Dystroglycan	95	136	63	6	5	3	-5.7656	0.0004
3	IPI00015688	GPC1	Glypican-1	80	82	71	12	5	8	-4.3930	0.0005
4	IPI00027505	ITGAV	Integrin alpha-V	101	108	84	23	20	8	-4.2940	0.0005
5	IPI00015102	ALCAM	CD166	149	85	101	37	21	32	-3.8113	0.0009
6	IPI00015476	SLC1A4	Neutral amino acid transporter A	35	36	46	0	0	0	-3.5287	0.0011
7	IPI00002406	BCAM	Basal cell adhesion molecule	35	31	53	5	2	5	-3.1407	0.0012
8	IPI00019146	CXADR	Coxsackievirus and adenovirus receptor	56	50	82	11	14	17	-3.0854	0.0012
9	IPI00383597	SLC4A2	Anion exchange protein 2	25	31	32	0	0	0	-2.9051	0.0013
10	IPI00003373	OCLN	Occludin	36	45	67	6	6	18	-2.8385	0.0013
11	IPI00152540	CD109	CD109	31	39	51	6	9	12	-2.4534	0.0019
12	IPI00029741	ITGB5	Integrin beta-5	27	20	17	0	0	0	-2.3108	0.0024
13	IPI00001754	F11R	Junctional adhesion molecule A	60	40	32	17	10	10	-2.2917	0.0024
14	IPI00014371	CDH18	Cadherin-18	18	19	26	0	0	0	-2.2839	0.0024
15	IPI00026952	PKP3	Plakophilin 3b (Fragment)	27	30	30	8	4	8	-2.0383	0.0028
16	IPI00394851	KLRG2	Killer cell lectin-like receptor subfamily G member 2	13	17	24	0	0	0	-2.0325	0.0028
17	IPI00396658	ITFG3	Protein ITFG3	18	24	5	0	0	0	-1.8229	0.0035
18	IPI00008986	SLC7A5	Large neutral amino acids transporter small subunit 1	39	34	88	14	16	44	-1.7442	0.0039
19	IPI00022048	PTGFRN	Prostaglandin F2 receptor negative regulator	36	17	44	13	14	13	-1.4872	0.0053
20	IPI00002478	ECE1	Endothelin-converting enzyme 1	16	18	11	4	4	0	-1.4677	0.0056
21	IPI00017562	LPHN2	Latrophilin-2	6	13	17	0	0	0	-1.4614	0.0057
22	IPI00215997	CD9	CD9	85	31	34	30	29	20	-1.4331	0.0059
23	IPI00024307	EFNB1	Ephrin-B1	12	8	15	0	0	0	-1.4261	0.0059
24	IPI00232571	GPC4	Glypican-4	7	12	15	0	2	0	-1.3903	0.0061
25	IPI00021975	TNFRSF10A	Tumor necrosis factor receptor superfamily member 10A	22	26	13	7	10	6	-1.2588	0.0081
26	IPI00300384	ERBB2	Receptor tyrosine-protein kinase erbB-2	9	8	13	0	0	0	-1.2420	0.0082
27	IPI00334190	STOML2	Stomatin-like protein 2	22	25	11	8	10	5	-1.1767	0.0092
28	IPI00217766	SCARB2	Lysosome membrane protein 2	6	11	17	0	0	5	-1.1655	0.0094
29	IPI00219682	STOM	Erythrocyte band 7 integral membrane protein	13	12	3	0	0	0	-1.1643	0.0097
30	IPI00843765	SPTAN1	Spectrin alpha chain, brain	12	15	3	3	2	0	-1.1621	0.0097
31	IPI00022558	MPZL1	Myelin protein zero-like protein 1	12	12	7	3	0	3	-1.1255	0.0105
32	IPI00002483	SLC30A1	Zinc transporter 1	6	6	15	2	0	0	-1.1245	0.0107
33	IPI00005614	SPTBN1	Spectrin beta chain, brain 1	12	9	15	4	5	0	-1.1068	0.0111
34	IPI00002070	LRRC8A	Leucine-rich repeat-containing protein 8A	7	8	13	0	0	3	-1.0845	0.0116
35	IPI00015756	PTPRK	Receptor-type tyrosine-protein phosphatase kappa	7	8	11	0	0	0	-1.0840	0.0121
36	IPI00009236	CAV1	Caveolin-1	30	24	30	11	16	18	-1.0354	0.0132
37	IPI00291175	VCL	Vinculin	3	6	15	0	0	0	-1.0008	0.0146
38	IPI00760554	HLA-A	HLA class I histocompatibility antigen, A-69 alpha chain	9	7	11	0	4	0	-0.9705	0.0157
39	IPI00026241	BST2	Bone marrow stromal antigen 2	9	5	9	0	0	0	-0.9579	0.0162
40	IPI00013449	TSPAN6	Tetraspanin-6	16	4	15	5	0	6	-0.9534	0.0163
41	IPI00000070	LDLR	Low-density lipoprotein receptor	3	13	13	4	4	3	-0.8503	0.0220
42	IPI00009507	SYPL1	Synaptophysin-like protein 1	15	10	13	6	5	6	-0.8448	0.0222
43	IPI00218487	GJA1	Gap junction alpha-1 protein	0	13	5	0	0	0	-0.8238	0.0242
44	IPI00183082	TMEM51	Transmembrane protein 51	0	5	13	2	0	0	-0.8238	0.0242
45	IPI00002459	ANXA6	Annexin A6	9	5	5	0	0	0	-0.7771	0.0273
46	IPI00170692	VAPA	Vesicle-associated membrane protein-associated protein A	15	10	17	8	6	8	-0.7433	0.0297
47	IPI00018860	ULBP2	NKG2D ligand 2	9	7	7	2	0	5	-0.7347	0.0307
48	IPI00290039	CDCP1	CUB domain-containing protein 1	33	18	34	24	10	22	-0.7331	0.0308
49	IPI00023942	SDC3	Syndecan	0	5	11	0	0	0	-0.7291	0.0316
50	IPI00022284	PRNP	Major prion protein	9	10	13	2	6	8	-0.6837	0.0360
51	IPI00943891	LOC100132336	GDNF family receptor alpha-2-like	28	0	0	12	0	0	-0.6837	0.0360
52	IPI00000059	CD58	CD58	6	5	7	3	2	0	-0.6499	0.0399
53	IPI00026650	HLA-C	HLA class I histocompatibility antigen, Cw-1 alpha chain	0	10	9	0	4	3	-0.6465	0.0410
54	IPI00018213	LRRC8D	Leucine-rich repeat-containing protein 8D	3	10	3	0	0	0	-0.6293	0.0454
55	IPI00168728	IGHM	FLJ00385 protein (Fragment)	0	12	0	0	2	0	-0.6293	0.0454
56	IPI00472013	HLA-A	HLA-A HLA class I histocompatibility antigen, A-33 alpha chain	15	7	7	6	4	5	-0.6228	0.0463
57	IPI00006211	VAPB	/form 1 of Vesicle-associated membrane protein-associated pro	16	15	28	9	10	20	-0.6017	0.0488
58	IPI00472138	HLA-B	HLA class I histocompatibility antigen, B-58 alpha chain	3	5	9	0	2	3	-0.6007	0.0494
59	IPI00011578	NPTN	Neuroplastin	6	12	9	9	0	3	-0.5982	0.0498

[a]N1,N2,N3 : Spectral counts which were investigated in each triplet replication of Non-CSC like cells [b]C1,C2,C3 : Spectral counts which were investigated in each triplet replication of CSC like cells

[c]STN : Signal to noise raio generated by PLGEM [d]p-value calculated by PLGEM analysis

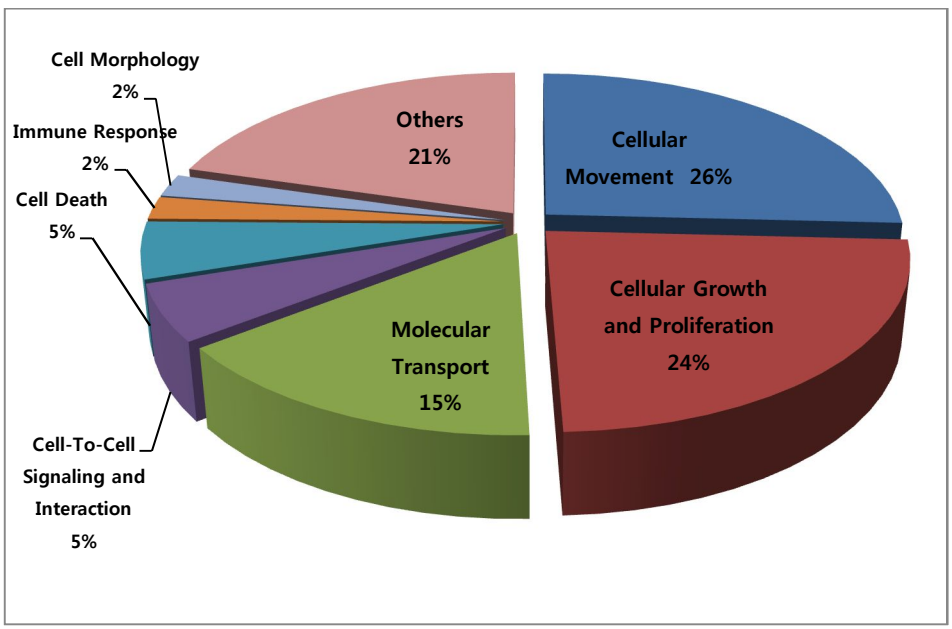


Figure 6. Molecular and cellular functions of significantly expressed proteins

The seven main molecular functions of the differentially expressed proteins were presented. The 93 proteins were classified according to molecular functions using Ingenuity Pathway Analysis. A threshold p-value < 0.001 is applied. The p-value calculated based on Fisher's exact test.

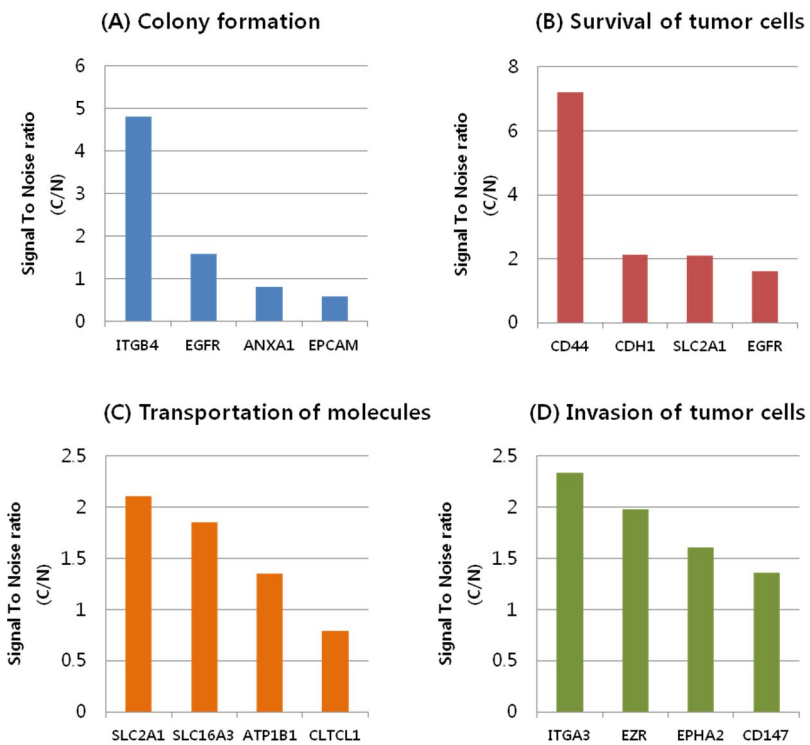


Figure 7. Relative ratios of differentially expressed proteins clustered by molecular function

Relative changes in the proteins differentially expressed in the CSC-like cells versus Non CSC- like cells are associated with (A)colony formation, (B)survival of tumor cells, (C)transportation of molecules, and (D)invasion of tumor cells. A threshold p-value < 0.05 is applied. The p-value calculated based on PLGEM. C : CSC-like cells N : Non CSC-like cells ITGB4 : Integrin beta4 ITGA3 : Integrin alpha3 EPCAM : Epithelial cell adhesion molecule EFGR : Epidermal growth factor receptor CDH1 : Cadherin-1 SLC2A1 : Solute carrier family 2, facilitated glucose transporter member 1 SLC16A3 : Monocarboxylate transporter 4 ATP1B1 : Sodium/potassium-transporting ATPase subunit beta-1 EPHA2 : Ephrin type-A receptor CD44 : Hyaluronate receptor CD147 : Basigin ANXA1 : Annexin A1 EZR : Ezrin CLTCL1 : Clathrin heavy chain 2

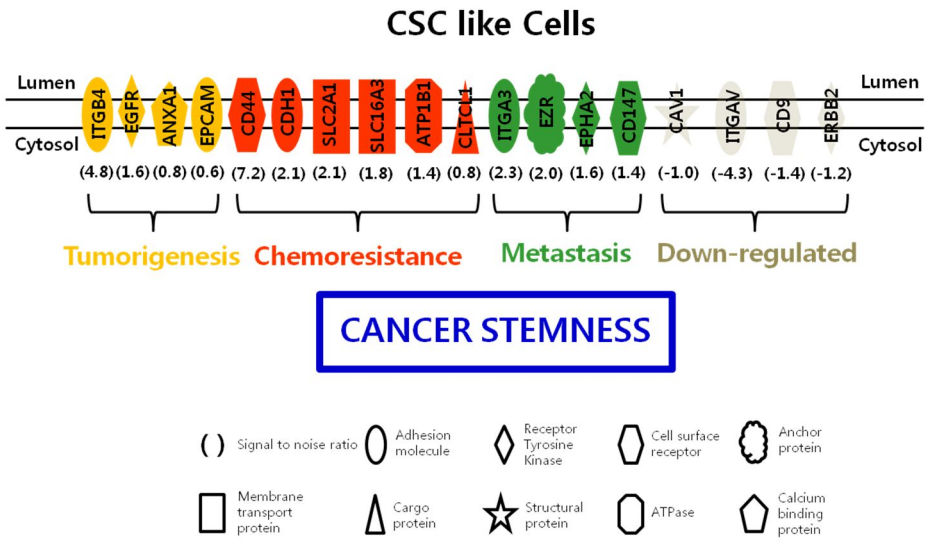


Figure 8. The schematic diagram of cell surface proteins which are differentially expressed in the CSC-like cells

The cell surface proteins and their functions in CSC-like cells represent cancer stem cell properties. ITGB4 : Integrin beta4 ITGA3 : Integrin alpha3 EPCAM : Epithelial cell adhesion molecule CAV1 : Caveolin-1 EFGR : Epidermal growth factor receptor CDH1 : Cadherin-1 SLC2A1 : Solute carrier family 2, facilitated glucose transporter member 1 SLC16A3 : Monocarboxylate transporter 4 ATP1B1 : Na/K-transporting ATPase subunit beta-1 ERBB2 : Receptor tyrosine-protein kinase erbB-2 EPHA2 : Ephrin type-A receptor CD44 : Hyaluronate receptor CD147 : Basigin ANXA1 : Annexin A1 EZR : Ezrin CLTCL1 : Clathrin heavy chain 2 CD9 : Cell growth-inhibiting gene 2 protein ITGAV : Integrin alpha V

3.4 Immunoblot Verification for Differentially Expressed Proteins in CSC-like and Non CSC -like Cells

To validate the differentially expressed proteins obtained from the biotin directed affinity purification based on proteomic study, Western blotting was performed on three proteins. According to the MS data, the significantly up-regulated proteins (EGFR, CD44, CD147) in the CSC-like cells were verified by Western blotting. The selection was mainly based on biological functions about cancer stem cell relatedness of proteins as well as availability of antibodies. These correlated biological functions involved major properties related with cancer stem cells, which are colony formation in tumorigenesis, survival of tumor cells in chemoresistance, and invasion of tumor cells in metastasis. EGFR, CD44, and CD147 were more highly expressed in the CSC-like cells (Fig.9). The signals of the VAMP3 were comparable for both cell lines, ascertaining loading of equal amounts of proteins. The actual expression of these proteins was verified, confirming that the MS analyses were able to depict correctly global protein expression pattern of CSC-like cells versus Non CSC-like cells.

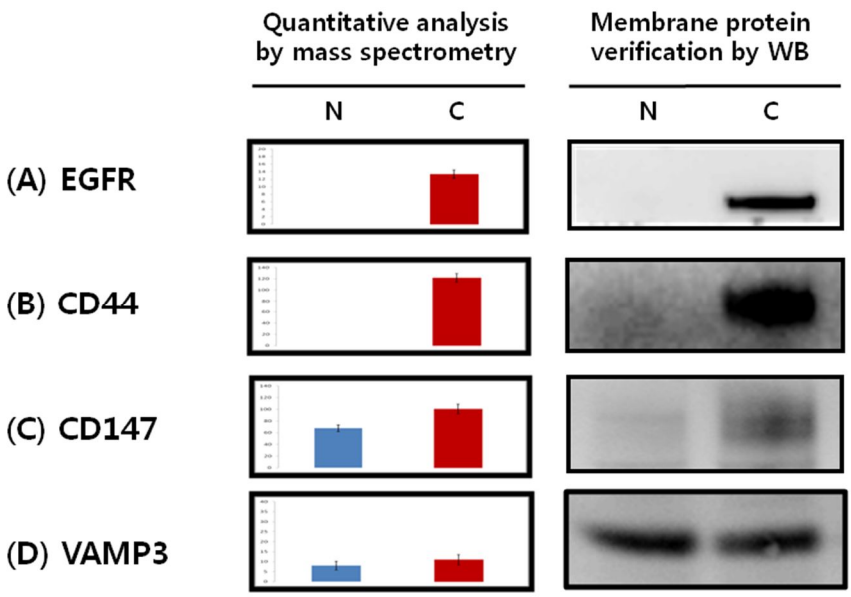


Figure 9. Representative Western blot images for differentially expressed proteins in the CSC-like and Non CSC-like cells

The representative proteins were extracted from the CSC-like and Non CSC-like cells, and the expression pattern of EGFR (Epidermal growth factor receptor, A), CD44 (Hyaluronate receptor, B), and CD147 (Basigin, C) were determined by Western blot. The amount of proteins loaded (30 μg) was normalized using VAMP3 (Vesicle-associated membrane protein3, D).

C: CSC-like cells N: Non CSC-like cells

4. Discussion

In this study, we profiled cell surface antigens of the CSC-like cells through the biotin directed affinity purification, mass spectrometry. To overcome limitations for studying cancer stem cells due to their rare population and phenotype alteration under in vitro conditions, the CSC-like cells were used for proteome analysis (Hart and El-Deiry 2008). The CSC-like cells are not only characterized by tumor initiating abilities, survival in hypoxic and acidotic environments, and chemoresistance but also enough to apply proteomic experiments (Sajithlal, Rothermund et al. 2010).

The biotinylation of the CSC-like cells was performed by incubating in sulfo-NHS-SS-biotin reagent which is a cleavable, charged biotin derivative. The reagent improved the yield of surface accessible proteins allowing the affinity purification and enrichment of labeled proteins by streptavidin Sepharose. Furthermore, we obtained meaningful information about the steady state distribution of cell surface antigens representing a “snapshot” of the live cell surface. To gain more purified membrane proteins, we optimized the four check points during this process. The first point is incubation time of biotinylation. The adequate incubation time leads to not only low interaction of biotin reagent with

cytosolic and nuclear proteins but also sufficient interaction with membrane proteins. We tested three different incubation time consisted of 5, 15, 30 minute. As a result, 30 minute incubation showed the best performance among the three different incubation times (data not shown). The second is cell lysis buffer and elution buffer which contain highly concentrated detergents for affinity purification. Since most membrane proteins are hydrophobic, anionic detergents and non-ionic detergents were added into the lysis buffer and elution buffer to increase the protein recovery. We confirmed that elution buffer with detergents has better achievement for membrane protein isolation than elution buffer without detergents (data not shown). Moreover, we optimized elution process by adding second elution step which contains only 0.1X PBS. This buffer was used to wash out membrane protein residues and to minimize the effect of salt during downstream SDS-PAGE assay. The third is adding the oxidized glutathione in buffers which are used in cell lysis and affinity purification processes. Because the oxidized glutathione also has disulfide bond, the biotinylated proteins are protected from reduction due to free reactive H+ ions which are originated from disrupted cells. We applied proper concentration of oxidized glutathione on reference (Scheurer, Rybak et al. 2005). The last point is stringent wash condition due to high affinity of biotin for streptavidin. The washing condition allows purification of the target proteins, thus reduces background bindings. We repeated total eight times of washing processes in four different conditions which include strict detergents and high ionic strength solutions on previous study (Kischel, Guillonneau et al. 2008).

To expand insight for indepth analysis of differential expressed proteins in the CSC-like compared to Non CSC-like cells, 1 mg of isolated proteins of each cell line were separated into ten fractions through SDS-PAGE. In consequence, additional CSC-like cell specific proteins were investigated such as CD133 (Prominin-1) which is one of the representative cancer stem cell surface markers (Hermann, Huber et al. 2007). Three technological replicates in the each cell line were evaluated to maximize the influence of reliable quantitative information and also to improve the coverage of differentially expressed proteins. In addition, the technical replicates which had consisted of the CSC-like cells and Non CSC-like cells as a one set were repeated three times not only to obtain preliminary data on first set experiment, but also to confirm the reproducibility of follow-up data. The reproducibility between triplet replications was shown as the total number of identified proteins in each set appeared similar to 1100, overlapped proteins in the CSC-like between Non CSC-like cells occupied 51 % of total identified proteins on average, and 54.4 % proteins were overlapped in the sum of triplet replications. Markers of protein identification increased on the CSC-like cells in the first set and Non CSC-like cells in the second set. Although we tried to do experiments on the CSC-like and Non CSC-like cells in same condition while mass spectrometry analysis, such results were presented because of deterioration of analytical column condition. As samples run, the condition of the column on latter sample worsens. In order to check reproducibility between each same cell line triplicates, we depicted Venn diagrams. The 561 proteins which equate to 46.4 % of total identified

proteins in the CSC-like cells overlapped, and the 512 proteins which correspond to 44.7 % of total investigated proteins in the Non CSC-like cells superimposed. Thus, there are two reasons why each overlapped proteins from triplicates cannot exceed 50 % of total proteins. One is that CSC-like and Non CSC-like cells were analyzed as a one technical replication set without serial triplet analysis in the same cells. As a result, due to the interval between replications in the same cells, sample concentration changed. Especially hydrophobic peptides were aggregated and attached on the inner wall of thread vial in the process of storage for mass spectrometry analysis. The difference between replications in the same sample occurred despite we applied enough sonication and shaking to the sample vial to overcome the problems. The other is that LTQ-Velos mass spectrometer has limitation. LTQ-Velos has crucial advantage for protein profiling based on fast scan time. However, it has weakness on reproducibility for selecting the target ions due to under sampling. For these reasons, the limitation of reproducibility was revealed in the triplet replications on the same sample.

Three proteins (EGFR, CD44, CD147) which had shown up-regulated levels in the CSC-like cells were examined by immunoblot for further verification. The equal loading amounts of proteins for Western blotting were assessed by VAMP3. Since the experimental strategy for analyzing steady state distribution of surface proteins on live cells is the membrane protein isolation, the generally used loading control proteins such as β-actin and α-actin which are cytosolic proteins are inadequate. In addition, the amounts of these proteins are unstable as the efficiency

of biotinylation. VAMP3 is a membrane protein as well as showed undifferentiated expression levels on MS data. Therefore, VAMP3 was employed as loading control proteins. In this study, EGFR, CD44 and CD147 revealed similar quantification results both in mass spectrometry and immunoblot. It is important to validate proteome profiling results using alternative methods such as Western blot. However, there are no other effective techniques with less than 1 μg of protein currently. To compensate for the lack of validation, we increased the confidence level of protein identification by applying statistical criteria. We instituted the cutoff to a peptide probability score ≧ 0.95, FDR < 1% and spectral counts ≧ 2 as well as PLGEM based *p*-value < 0.05. Moreover, multiple replicate corrections are employed to control false positives induced by performing significance tests on a large number of proteins. As a result, differentially expressed proteins identified in this way are more likely to be true positives.

In conclusion, because direct study of cell surface proteins of cancer stem cells is difficult for effective anti-cancer therapy, we carried out differentially expression profiling of the CSC-like cells using mass spectrometry. The 93 significantly expressed proteins were identified, including cancer stem cell markers. Out of 93 proteins, several proteins were known to be associated with cancer stem cell properties. Thus, the global analysis of differentially expressed membrane proteins in the CSC-like cells may improve our understanding and developing further management of cancer. To mine more effective therapeutic and diagnostic targets, we also need to increase the sensitivity and specificity of marker candidates

about cancer stem cells by comparing the candidates extracted from the CSC-like cells with the cell surface proteins existed on primary cancer cells and normal stem cells in same tissue.

5. References

BM, F., R. BF, et al. (1999). "Annexin I modulates cell functions by controlling intracellular calcium release." FASEB J 13(15): 2235-2245.

Boman, B. M. and M. S. Wicha (2008). "Cancer Stem Cells: A Step Toward the Cure." Journal of Clinical Oncology 26(17): 2795-2799.

Cheng, L., W. Lu, et al. (2010). "Analysis of chemotherapy response programs in ovarian cancers by the next-generation sequencing technologies." Gynecologic Oncology 117(2): 159-169.

Dai, L., C. Li, et al. (2010). "Quantitative proteomic profiling studies of pancreatic cancer stem cells." J Proteome Res 9(7): 3394-3402.

Elliott, B. E., J. A. Meens, et al. (2005). "The membrane cytoskeletal crosslinker ezrin is required for metastasis of breast carcinoma cells." Breast Cancer Res 7(3): R365-373.

Ezeh, U. I., P. J. Turek, et al. (2005). "Human embryonic stem cell genes OCT4, NANOG, STELLAR, and GDF3 are expressed in both seminoma and breast carcinoma." Cancer 104(10): 2255-2265.

G, F., R. D, et al. (2002). "Caveolin-1 inhibits anchorage-independent growth, anoikis and invasiveness in MCF-7 human breast cancer cells." Oncogene 21(15): 2365-2375.

Gale, N. W., S. J. Holland, et al. (1996). "Eph receptors and ligands comprise two major specificity subclasses and are reciprocally compartmentalized during embryogenesis." Neuron 17(1): 9-19.

Hao, J., H. Chen, et al. (2010). "Co-expression of CD147 (EMMPRIN), CD44v3-10, MDR1 and monocarboxylate transporters is associated with prostate cancer drug resistance and progression." British Journal of Cancer 103(7): 1008-1018.

Hart, L. S. and W. S. El-Deiry (2008). "Invincible, but not invisible: imaging

approaches toward in vivo detection of cancer stem cells." J Clin Oncol 26(17): 2901-2910.

Hermann, P. C., S. L. Huber, et al. (2007). "Distinct populations of cancer stem cells determine tumor growth and metastatic activity in human pancreatic cancer." Cell Stem Cell 1(3): 313-323.

J, S. and H. ME (2001). "Regulation of MMP-1 and MMP-2 production through CD147/extracellular matrix metalloproteinase inducer interactions." Cancer Res 61(5): 2276-2281.

Kinch, M. S. and K. Carles-Kinch (2003). "Overexpression and functional alterations of the EphA2 tyrosine kinase in cancer." Clin Exp Metastasis 20(1): 59-68.

Kischel, P., F. Guillonneau, et al. (2008). "Cell membrane proteomic analysis identifies proteins differentially expressed in osteotropic human breast cancer cells." Neoplasia 10(9): 1014-1020.

Koleske, A. J., D. Baltimore, et al. (1995). "Reduction of caveolin and caveolae in oncogenically transformed cells." Proc Natl Acad Sci U S A 92(5): 1381-1385.

Lee, Y. Y., K. Q. McKinney, et al. (2011). "Subcellular tissue proteomics of hepatocellular carcinoma for molecular signature discovery." J Proteome Res 10(11): 5070-5083.

Lindberg, R. A. and T. Hunter (1990). "cDNA cloning and characterization of eck, an epithelial cell receptor protein-tyrosine kinase in the eph/elk family of protein kinases." Mol Cell Biol 10(12): 6316-6324.

Liu, H., R. G. Sadygov, et al. (2004). "A model for random sampling and estimation of relative protein abundance in shotgun proteomics." Anal Chem 76(14): 4193-4201.

LY, B., X. W, et al. (2009). "Hyaluronan-mediated CD44 Interaction with p300 and SIRT1 Regulates {beta}-Catenin Signaling and NF{kappa}B-specific Transcription Activity Leading to MDR1 and Bcl-xL Gene Expression and Chemoresistance in Breast Tumor Cells." J Biol Chem 284(5): 2657-2671.

M, M., K. C, et al. (2004). "The carcinoma-associated antigen EpCAM upregulates c-myc and induces cell proliferation." Oncogene 23(34): 5748-5758.

MS, D., I. H, et al. (2004). "EphA2: a determinant of malignant cellular behavior and a potential therapeutic target in pancreatic adenocarcinoma." Oncogene 23(7): 1448-1456.

Natarajan, T. G. and K. T. FitzGerald (2007). "Markers in normal and cancer stem cells." Cancer Biomark 3(4-5): 211-231.

Pavelka, N., M. L. Fournier, et al. (2008). "Statistical similarities between transcriptomics and quantitative shotgun proteomics data." Mol Cell Proteomics 7(4): 631-644.

Pavelka, N., M. Pelizzola, et al. (2004). "A power law global error model for the identification of differentially expressed genes in microarray data." BMC Bioinformatics 5: 203.

Rybak, J.-N., S. B. Scheurer, et al. (2004). "Purification of biotinylated proteins on streptavidin resin: A protocol for quantitative elution." Proteomics 4(8): 2296-2299.

S, I., F. T, et al. (2002). "Coexpression of glucose transporter 1 and matrix metalloproteinase-2 in human cancers." J Natl Cancer Inst 94(14): 1080-1091.

Sajithlal, G. B., K. Rothermund, et al. (2010). "Permanently Blocked Stem Cells Derived From Breast Cancer Cell Lines." Stem Cells 28(6): 1008-1018.

Sarrio, D., S. M. Rodriguez-Pinilla, et al. (2006). "Abnormal ezrin localization is associated with clinicopathological features in invasive breast carcinomas." Breast Cancer Res Treat 98(1): 71-79.

Scheurer, S. B., J.-N. Rybak, et al. (2005). "Identification and relative quantification of membrane proteins by surface biotinylation and two-dimensional peptide mapping." Proteomics 5(11): 2718-2728.

Scheurer, S. B., J. N. Rybak, et al. (2004). "Modulation of gene expression by hypoxia in human umbilical cord vein endothelial cells: A transcriptomic and proteomic study." Proteomics 4(6): 1737-1760.

Stevens, T. J. and I. T. Arkin (2000). "Do more complex organisms have a greater proportion of membrane proteins in their genomes?" Proteins 39(4): 417-420.

SW, L., R. CL, et al. (1998). "Tumor cell growth inhibition by caveolin re-

expression in human breast cancer cells." Oncogene 16(11): 1391-1397.

Wallin, E. and G. von Heijne (1998). "Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms." Protein Sci 7(4): 1029-1038.

Y, Z., T. AA, et al. (2002). "Hyaluronan-CD44s signaling regulates matrix metalloproteinase-2 secretion in a human lung carcinoma cell line QG90." Cancer Res 62(14): 3962-3965.

YP, H., V. S, et al. (2005). "Reorganization of ErbB Family and Cell Survival Signaling after Knock-down of ErbB2 in Colon Cancer Cells." J Biol Chem 280(29): 27383-27392.

Yu, Y., J. Khan, et al. (2004). "Expression profiling identifies the cytoskeletal organizer ezrin and the developmental homeoprotein Six-1 as key metastatic regulators." Nat Med 10(2): 175-181.

6. Abstract in Korean

암줄기세포란 암세포들 중 자가재생, 종양생성, 화학치료저항과 같은 특성을 지닌 특이한 암세포이다. 전통적인 항암치료를 무력화시키는 독특한 줄기세포적 특성 때문에, 암줄기세포는 항암치료에 있어서 새로운 치료방법개발을 위한 중요한 표적이 되었다. 그러나 암줄기세포의 빠른 분화능력과 희박한 존재량에 기인하여 전반적인 암줄기세포 표면 표지자 발굴이 제한 되었다. 최근 본 연구팀은 인간 유방암세포에서 유래한 유사 암줄기세포를 획득할 수 있었고, 이 유사 암줄기세포는 암줄기세포의 특성뿐 만 아니라 미분화상태를 유지하는 특성도 가지고 있었다. 암줄기세포의 잠재적인 특이 표면 표지자 발굴을 위해서 비유사 암줄기세포와 비교한 유사 암줄기세포의 세포 표면 단백질 규명의 심층분석을 실행하였다. 살아있는 세포 표면에서 마치 스냅 촬영한 것과 같은 정보를 얻으면서 세포 표면 단백질 들의 회수율을 증가시키기 위해서 본 연구에서는 세포 표면 바이오티닐레이션 기술을 사용하였다. 이어서 strepavidin 정제 방법과 SDS전기영동을 이용한 단백질 분획, 질량 분석법을 적용하였다. 이러한 방법을 통해 유사 암줄기세포에서 특이적이고 과발현된 단백질을 찾아낼 수 있었다. 확인된 총 1525개 단백질 중에

CD44, Integrinβ-1, CD133 와 같이 이미 알려진 암줄기세포 표면 표지자를 포함하여 94개 단백질이 비유사 암줄기세포와 비교하여 유사 암줄기세포에서 과발현된 것을 살펴볼 수 있었다. 또한 PLGEM에 기반한 신호대잡음 비율을 Ingenuity Pathway Analysis에 적용시킴으로써 종양생성, 화학치료저항, 암전이와 같은 암줄기세포의 특성과 관계 깊은 몇 가지 흥미로운 단백질들도 확인할 수 있었다. 본 연구팀은 면역단백질검출검사법을 통하여 이러한 과발현 표면 항원들을 입증했고, 이로써 유사 암줄기세포 표면 항원 연구가 암줄기세포 표면 항원 연구와 매우 밀접한 관련이 있음을 알 수 있었다.

주요어: 암줄기세포,세포표면항원,바이오티닐레이션,질량분석기,단백질

학번: 2010-24238