Title: CD44v8-10 is a cancer-specific marker for gastric cancer stem cells

Authors: Wen Min Lau¹,⁴, Eileen Teng¹,⁴, Hui Shan Chong¹, Kirsten Anne Pagaduan Lopez¹, Amy Yuh Ling Tay², Manuel Salto-Tellez³,¹, Asim Shabbir², Jimmy Bok Yan So²,⁵ and Shing Leng Chan¹,⁵,⁶

1. Cancer Science Institute of Singapore, National University of Singapore. 14 Medical Drive, MD6, #12-01, Singapore 117599
2. Department of Surgery, National University of Singapore, Singapore. NUHS Tower Block Level 8,1E Kent Ridge Road, Singapore 119228.
3. Centre for Cancer Research & Cell Biology, Queen's University Belfast, 97 Lisburn Road, Belfast BT9 7BL, Northern Ireland, UK.
4. These authors contributed equally to this work.
5. The project is conceptualized by these authors.
6. Corresponding author.

Running title: CD44v8-10 marks gastric cancer stem cells

Keywords: stomach cancer, cancer stem cell, carcinogenesis, tumor development, patient-derived xenograft

Financial support: This work is supported by grants to S.L. Chan from the Ministry of Education, Singapore and Singapore Stem Cell Consortium, A*Star Singapore.

Corresponding author: Shing Leng Chan, PhD
Cancer Science Institute of Singapore
14 Medical Drive, MD6, #12-01, Singapore 117599
Tel: 65-65167211, FAX: 65-68739664.
Email: csicsl@nus.edu.sg

Disclosure: The authors do not have any conflicts of interest.

Word Count: 4477

Figures: 5 Figures, 2 tables
ABSTRACT

The surface marker CD44 has been identified as one of several markers associated with cancer stem cells (CSCs) in solid tumors but its ubiquitous expression in many cell types including hematopoietic cells has hindered its use in targeting CSCs. In this study, 28 paired primary tumor and adjacent non-tumor gastric tissue samples were analyzed for cell surface protein expression. Cells that expressed pan-CD44 were found to occur at significantly higher frequency in gastric tumor tissues. We identified CD44v8-10 as the predominant CD44 variant expressed in gastric cancer cells and verified its role as a gastric CSC marker by limiting dilution and serial transplantation assays. Parallel experiments using CD133 failed to enrich for gastric CSCs. Analyses of another 26 primary samples showed significant CD44v8-10 upregulation in gastric tumor sites. Exogenous expression of CD44v8-10 but not CD44 standard (CD44s) increased the frequency of tumor initiation in immune-compromised mice. Reciprocal silencing of total CD44 resulted in reduced tumor initiating potential of gastric cancer cells that could be rescued by CD44v8-10 but not CD44s expression. Our findings provide important functional evidence that CD44v8-10 marks human gastric CSCs and contributes to tumor initiation, possibly through enhancing oxidative stress defense. In addition, we showed that CD44v8-10 expression is low in normal tissues. Since CD44 also marks CSCs of numerous human cancers, many of which may also overexpress CD44v8-10, CD44v8-10 may provide an avenue to target CSCs in other human cancers.
INTRODUCTION

Gastric cancer is a major cause of cancer-related death worldwide, with low survival and high recurrence rates for patients with advanced disease (1-3). Therefore, new therapies for treatment of gastric cancer are urgently needed. The cancer stem cell hypothesis proposes that a specific subset of cancer stem cells (CSCs) are primarily responsible for initiating and maintaining tumor growth (4, 5), a proposition which may explain the high frequency of relapse and resistance to current therapeutic modalities which focus on reduction of tumor bulk without considering tumor heterogeneity. Definitive CSC markers for gastric cancer have yet to be characterized, although several studies have identified gastric CSCs from cell lines (6-9).

CD44 is a cell surface transmembrane glycoprotein encoded by the CD44 gene, a 20-exon DNA segment (10) of which exons 1-5 and 16-20 are spliced together and translated into CD44s, the standard isoform. In addition, the variant exons 6-15 can be alternatively spliced and assembled in different combinations with the standard exons to generate other variant (CD44v) protein isoforms. More recently CD44 has been recognized as a CSC marker in several types of cancer (11). As CD44s is ubiquitously expressed in many cell types (12-16), its usefulness as a CSC marker may be limited. Furthermore, conflicting data in the field implicates CD44 in both tumor suppression and progression (17-19), largely attributed to the expression of alternatively spliced variants. In light of this, we investigated the role of CD44 splice variants in gastric cancer. In this study, we isolated gastric CSCs using two well known markers for CSCs, EpCAM and CD44. Further investigation of the expression of CD44 variants in gastric cancer revealed that CD44v8-10 is specifically upregulated in tumor cells, and its expression confers an advantage in tumor initiation.
MATERIALS AND METHODS

Study Approval

All patient tissue samples were collected with informed patient consent from National University Hospital Singapore according to Institutional Review Board guidelines. All animal experiments were conducted with the approval of Institutional Animal Care and Use Committee in the National University of Singapore.

Isolation and purification of single cells from dissociated tumor tissue

Samples were washed twice in sterile HBSS before mincing finely with scalpel blades. Tissue pieces were incubated in 0.1% DTT/PBS to remove mucus and washed twice in HBSS before incubation at 37°C for 2 hours in 0.2μm sterile filtered DMEM/F12 media (Invitrogen) containing 100U/mL hyaluronidase (H1136, Sigma), 300U/mL collagenase type II (C6885, Sigma), 5ng/mL human recombinant insulin (12585, Invitrogen) and 1ng/mL hydrocortisone (H2270, Sigma). Digestion mixture was washed twice in HBSS by centrifugation at 1,200g for 8 minutes. Pelleted material was resuspended in HBSS and filtered using a 40μm cell strainer (Falcon) to collect single cells. Single cell suspensions were passed through LD and dead cell removal columns (Miltenyi Biotec) to remove cellular debris and dead cells. All procedures were carried out under sterile conditions.

Flow cytometric analysis and cell sorting

Single cells were counted with a hemocytometer using trypan blue exclusion of dead cells. 1x10^5 and 3x10^5 cells were used for direct and indirect staining respectively. For cells isolated from clinical gastric samples, CD45+ blood cells and GlyA+ erythroid precursors were gated out and excluded from analysis (Supplementary Figure S1). For xenograft-derived cells, murine H-2Kd MHC class I alloantigen-positive (H-2Kd+) cells were excluded from
analysis. Cells were analyzed on a BD LSRII flow cytometer (BD Biosciences). Median fluorescence intensity (MFI) values were calculated using FlowJo software (Tree Star). Cell sorting was performed on a BD FACS Aria (BD Biosciences). Sytox Blue (S34857) from Invitrogen was used to exclude dead cells in all analyses. Post-sort analysis was performed to ensure purity of cell fractions was ≥90%. Cells were washed twice in sterile PBS and counted before resuspension in matrigel (BD Biosciences) for xenograft transplant assays. Rat anti-human CD44v8-10 specific antibody was kindly provided by H. Saya and was generated as previously described (20). Briefly, CD44v8-10 was overexpressed in McA RH-7777 rat epithelial cells which do not express any endogenous isoforms of CD44. These cells were then used to immunize rat host in order to raise CD44v8-10 antibodies. Other antibodies used are described in supplementary methods.

Patient-derived xenograft models

Single cell suspensions or minced tissue pieces were mixed 2:1 with sterile HBSS/Matrigel (BD Biosciences) in a final volume of 200μL per injection and administered subcutaneously. For CD44 overexpression studies, vector control and CD44-transfected cells were injected on opposite flanks of the same animals to minimize experimental variability from differences in recipient mouse hosts. For cell dosing experiments, 20 to 25,000 sorted or unsorted cells were injected per site. Six to eight week old NOD/SCID or NOD/SCID/IL2Rγc (NSG) mice were used for xenograft transplantations and anesthetized intraperitoneally with a cocktail of xylazine (11 mg/kg) and ketamine (72 mg/kg). Mice were checked weekly for tumor formation and sacrificed when tumors were 1.5cm in diameter (usually 6-10 weeks) or after 20 weeks (if no tumor was observed).

Primary gastric samples and cell lines
Primary gastric cancer cells used in our study include GC59, GC71, GC101, GC121 and GC123. Established patient-derived xenograft lines include GC16, GC21, GC38, GC45, GC84 and GC119. Primary cell line GC38-adh was established in vitro from GC38 xenograft line as described below. Other cell lines used were MKN28 (JCRB cell bank, Japan), SNU5 (ATCC) and TMK1 (provided by Dr. Patrick Tan, Duke-NUS, Singapore). These cell lines were authenticated by 16-loci STR profiling (LGC standards; Middlesex, UK) in November 2013.

**Primary cell culture**

All cultures were maintained in humidified 37°C incubators supplemented with 5% CO₂. Primary cell cultures were established from GC38 xenograft tumors. Tumor spheroids were first derived from freshly minced tumor pieces in a 10cm tissue culture dish in serum-free low glucose DMEM (Invitrogen; details in Supplementary Material). Spheroids were propagated by dissociation with TrypLE (Invitrogen) for 10 minutes at 37°C, centrifugation at 1200g for 8 minutes and resuspension in PBS before passing them through a cell strainer (30μm) to remove dead cell aggregates. Live single cells were then re-plated in serum-free DMEM at a density of 10,000 single cells per 35mm dish. GC38 tumor spheroids were EpCAM⁺ as characterized by flow cytometric analysis, with 55-85% CD44⁺ (EpCAM⁺CD44⁺) cell populations. Established cultures of tumor spheroids were used to derive adherent monolayer primary cultures of GC38 cells. Spheroids were transferred to DMEM containing 10% FBS, which facilitated the growth of adherent cells after a period of 3-4 weeks. These adherent cells were passaged and expanded as the GC38-adh primary cell line.

*In vivo* limiting dilution assays
Non-mouse cells from each xenograft were sorted based on the indicated surface phenotypes and injected subcutaneously in NOD/SCID/IL2Rγ− (NSG) mice. Each mouse received the same cell dose of indicated fractions from the same xenograft and were either harvested when tumor generated from any fraction reached 1.5 cm diameter (usually 6-10 weeks), or 20 weeks later. Mice were considered negative for tumor formation when there was no palpable tumor. For functional assays of CD44s and CD44v8-10 overexpression, cells were sorted for the respective population and injected into mice at indicated doses. The frequency of cancer-initiating cells was calculated using the web-based tool Extreme Limiting Dilution Assay (ELDA; Walter and Eliza Hall Institute).

**Cell proliferation assays**

GC38-adh cells stably transfected with pcDNA3 vector, CD44s or CD44v8-10 were fractionated for the relevant population and 3000 cells were seeded in each of five replicate wells in a 96-well plate. Cell proliferation over a 5-day period was measured using WST-1 cell proliferation reagent (Roche Diagnostics) according to manufacturer’s instructions. Briefly, 10μL WST-1 reagent was added per 100μL culture medium and incubated for 2 hours before measurement of absorbance at 460nm in a plate reader (Tecan). Absorbance readings were normalized against blank wells and growth curves were plotted with data points showing mean ± SD.

**Depletion of CD44 in TMK1 cells**

shRNA that targets CD44 and scrambled control shRNA were obtained from Origene (TG314080). CD44 shRNA consisted of 2 different shRNA sequences targeted against human CD44 (NM_000610). Cells were transfected with shRNA plasmids using Lipofectamine 2000 (Invitrogen) as described above. For CD44 ‘rescue’ experiments, TMK1
cells were co-transfected with 6μg each of CD44 shRNA plasmids (Origene) and either CD44s, CD44v8-10 or pcDNA3 vector as shRNA-resistant rescue constructs. These shRNA-resistant rescue constructs were engineered by mutating relevant residues in the CD44 plasmid sequences targeted by the shRNA constructs. The mutagenesis primers and shRNA sequences are provided in Supplementary Table M1. Transfections were performed as described in the section above. TMK1 cells were sorted by flow cytometry for the relevant populations prior to injecting them into NSG mice for tumor formation assay.

**ELISA**

ELISA was performed using human CD44 ELISA kit (Ab45912, Abcam). 100-300μL conditioned media from cells cultured in 10cm dishes was assayed according to manufacturer’s protocol.

**Histopathology**

Tissue samples were fixed in 10% formalin and embedded in paraffin before sectioning at 4μm thickness. Hematoxylin and eosin (Sigma) staining was performed on all sections and evaluated by a pathologist.

**Statistical analysis**

All error bars represent the mean ± SD. Paired two-tailed Student’s t-tests and Fisher's exact test were performed using GraphPad Prism (GraphPad Software). Chi-squared tests were performed using the web-based software for Extreme Limiting Dilution Analysis (21) (http://bioinf.wehi.edu.au/software/elda/, Walter and Eliza Hall Institute of Medical Research). For all statistical tests, $P<0.05$ was considered to be significant. For flow
cytometric data, median fluorescence intensity (MFI) was calculated using FlowJo software (Tree Star) where relevant.
RESULTS

Identification of EpCAM and CD44 as surface markers for gastric cancer stem cells

A candidate approach was undertaken to identify putative gastric cancer stem cells from primary gastric cancer specimens because various surface markers have previously been used to describe CSCs from different cancers. Cells from 28 paired primary tumor and adjacent non-tumor tissues from gastric cancer patients (Supplementary Table S1) were isolated and analyzed by flow cytometry for potential cancer stem cell markers including CD44, CD133, CD34, CD117, CD90, CD166 and EpCAM, as well as CD45 (hematopoietic cells), CD31 (endothelial cells) and CD140b (fibroblasts). Of these, gastric cancer cells were positive for EpCAM, CD44, CD133 and CD166 expression. In particular, EpCAM levels were significantly higher in cells from 28 gastric tumors compared to paired adjacent non-tumor tissue ($P=0.0006$, Figure 1A). EpCAM is a well-established epithelial marker that is highly expressed in most carcinomas including gastric cancer (22, 23). Also, EpCAM+ but not EpCAM− cells initiated tumors in mice (Supplementary Fig. S1C) suggesting that gastric CSCs should be a subpopulation within the EpCAM expressing tumor cells, hence we used EpCAM as an identifier for gastric tumor epithelial cells and to exclude non-tumor cells.

Following this, flow cytometric analyses of EpCAM-expressing cells indicated the absence of CD90+, CD34+ and CD117+ cells, and a high proportion of CD166+ cells (60-70%) in both tumor and non-tumor sites. CD44 and CD133-expressing cells were found in the range of 0.8 to 54% and 0.6 to 45% respectively. Analyses of CD44+ or CD133+ cells in the EpCAM+ fraction isolated from 21 paired tumor and non-tumor patient samples revealed a consistently higher proportion of CD44+ cells ($P=0.0124$), but not CD133+ cells at the tumor site ($P=0.861$) (Fig. 1B, C&D). Nevertheless CD44 and CD133 were further analyzed as potential markers for gastric CSCs.
We established a robust gastric cancer xenograft model by subcutaneous implantation of primary tumors taken from gastric cancer patients. Histological analysis verified that all xenograft tumors recapitulated the original primary tumors, as shown for GC16 (well-differentiated), GC21 (moderately differentiated) and GC38 (poorly differentiated) (Fig. 2A). The proportion of EpCAM⁺CD44⁺ cells in GC16, 21 and 38 xenografts was 3.6%, 58% and 44% respectively (Fig. 2B). Limiting dilution assays were performed to determine the frequency of CSCs in EpCAM⁺CD44⁺ fractions from GC16, GC21 and GC38 xenografts (Table 1, Fig. 2C). The estimated number of CSCs in EpCAM⁺CD44⁺ fractions was significantly enriched by 204-fold, 17-fold, and 8-fold in GC16, GC21 and GC38 xenografts respectively, compared to EpCAM⁺CD44⁻ fractions which were depleted of CD44⁺ cells. Parallel experiments using EpCAM⁺CD133⁺ cells showed that gastric CSCs are not enriched in this fraction as both CD133⁺ and CD133⁻ cells have similar tumor initiating potential (Supplementary Table S2). Hence, we concluded that EpCAM and CD44 may be better surface markers for identifying gastric CSCs than CD133.

Serial transplantations were also performed by re-isolating EpCAM⁺CD44⁺ cells from first transplants in NSG mice and injecting these cells into recipient mice as second transplants. The proportion of CD44⁺ cells in xenograft tumors of both first and second transplants remained unchanged from the original tumors (Fig. 2D). Similarly, xenograft tumors of both first and second transplants showed the same histology (Fig. 2E). Since the injected cells are ≥95% CD44⁺, this suggests that EpCAM⁺CD44⁺ cells are able to differentiate to re-establish tumor heterogeneity of the original patient tumor as well as to self-renew and sustain tumor growth in serially passaged xenografts. This also demonstrates that CD44⁺ cells retained
tumor forming potential during serial transplantation. Taken together, these data indicate that gastric cancer stem cells are enriched in the EpCAM^+CD44^+ fraction of gastric tumors.

**CD44v8-10 is the predominant CD44 variant in gastric cancer xenografts and marks gastric cancer stem cells**

To determine the expression of CD44 variants in xenograft tumors, polymerase chain reaction (PCR) primers were designed in constant exons 5 and 16/17, flanking the variable region of the CD44 gene (Fig. 3A). PCR amplification of cDNA from xenograft tumors GC16, GC21 and GC38 generated 2 main products - 324bp and 720bp fragments whose sequences were verified to be CD44s, the standard transcript in many cell types, and CD44v8-10, a variant first identified in colon carcinomas respectively (24) (Fig. 3A). By quantitative PCR analysis we found that the level of CD44v8-10 transcript is highest in GC21, followed by GC38 and GC16 (Fig. 3B). Most importantly, levels of CD44v8-10 transcripts correlated with the estimated number of gastric CSCs in the unsorted fractions (Fig. 2C, Table 1).

We next used antibodies specific to CD44v8-10 to determine if the EpCAM^+CD44^+ cells identified in clinical samples using pan-CD44 antibody consist of CD44v8-10 expressing cells. The specificity of CD44v8-10 antibody was verified as follows; firstly, the antibody binds to CD44v8-10 but not CD44s when either of these was transiently expressed in MKN28, a gastric cancer cell line devoid of CD44 expression (Supplementary Fig. S2A). Secondly, the extracellular domain of CD44v8-10 (sol-CD44v8-10) but not sol-CD44s blocked the binding of the CD44v8-10 antibody to gastric cancer cells (Supplementary Fig. S2B) indicating that binding of the antibody to these cells is specific to the CD44v8-10 protein.
Using this CD44v8-10 specific antibody, we found that the majority of EpCAM^+CD44^+ cells identified using pan-CD44 antibodies are CD44v8-10 cells. This is shown in a representative sample GC101, where in the EpCAM^+ fraction, 79% of pan-CD44 expressing cells are CD44v8-10 positive (Fig. 3C; isotype controls are shown in Supplementary Fig. S2C). This observation was further substantiated by analysis of two other clinical samples GC121 and GC123 showing that majority of the CD44 population in the tumor consists of CD44v8-10 (Fig. 3D&E). Strikingly, CD44v8-10 is absent in hematopoietic CD45^+ cells (Fig. 3C), in contrast to CD44s which is expressed in hematopoietic cells.

In order to determine that CD44v8-10 marks gastric cancer stem cells, we prospectively isolated the CD44v8-10 cells and subjected them to cancer initiation assay by limiting dilution analyses. CD44v8-10 specific antibody was used to fractionate CD44v8-10^+ cells from gastric tumor xenografts GC45 and GC84 to high purity (Fig. 4A) before performing limiting dilution assays. We injected the unsorted, CD44v8-10^+ and CD44v8-10^- cell fractions into NSG mice at cell doses of 2000, 200 and 20. In the GC45 xenograft, the CD44v8-10^- fraction showed a 51-fold \( (P=2.5 \times 10^{-19}) \) and 3.8-fold \( (P=0.00217) \) enrichment of CSCs compared to the CD44v8-10^- and unsorted fractions respectively (Fig. 4B). Similarly for GC84, CD44v8-10^- fractions were enriched for CSCs by 84-fold \( (P=1.4 \times 10^{-18}) \) and 6-fold \( (P=0.00242) \) compared to CD44v8-10^- and unsorted cell fractions respectively (Table 2, Fig. 4B). We also serially transplanted CD44v8-10^+ cells re-isolated from CD44v8-10^- first transplant tumors of GC45 and GC84 into second transplant recipients, and observed the same percentage of CD44v8-10^+ and CD44v8-10^- cells as the first transplant xenograft tumors (Fig. 4C). Histology of serially transplanted xenograft tumors recapitulated that of the initial xenografts (Fig. 4D), confirming that CD44v8-10 indeed marks a population of cancer stem cells.
stem cells, which are able both to form tumors as well as regenerate the original tumor heterogeneity. However, since CD44v8-10\(^+\) cells may consist of both CD44\(^+\) and CD44\(^-\) populations, we next compared the tumor initiating potential between CD44\(^+\) and CD44v8-10\(^+\) fractions and observed that when injected at a dose of 200 cells, CD44v8-10\(^+\) cells formed tumors with higher frequency than CD44\(^+\) cells (Fig. 4E). This clearly demonstrates that CD44v8-10\(^+\) cells within the CD44\(^+\) fraction have enhanced tumor initiating capability compared to CD44\(^+\) cells.

The sphere formation assay is widely used as \textit{in vitro} method to evaluate stem cell activity in normal tissue as well as putative CSCs (25), hence we established tumor spheres from explants of GC38 xenograft tumor and confirmed that the sphere-forming cells are indeed human epithelial cells that express EpCAM (Supplementary Fig. S3C, D). However, when we isolated CD44v8-10\(^+\) and CD44v8-10\(^-\) cells from the spheroid culture, we found that both isolated subpopulations formed spheres in serum-free medium (Supplementary Fig. S3E, F). Moreover, other gastric xenograft lines did not survive as long-term tumor sphere cultures regardless of CD44v8-10 expression. As it has been noted that not all spheres originate from stem cells (26), our results suggest that sphere-forming \textit{in vitro} limiting dilution assay may not be a reliable surrogate for the more definitive \textit{in vivo} limiting dilution assay in the identification of surface markers for gastric CSCs.

**CD44v8-10 is upregulated in gastric tumor samples and plays a functional role in tumor initiation**

\textit{CD44} variant-specific primers were designed for quantitative real-time PCR (qPCR) analysis on another cohort of 26 gastric tumors and matched adjacent non-tumor tissues. The CD44v6 variant has also been reported to promote tumor growth and metastasis (27, 28), hence we
also investigated its expression level in gastric tumors. Our results show that CD44v8-10 expression was significantly elevated ($P=0.0029$) in tumor compared to adjacent non-tumor tissue, whereas CD44s and CD44v6 expression levels were similar in both tumor and adjacent non-tumor tissue (Fig. 4F).

To determine if CD44v8-10 plays a functional role in gastric cancer initiation, we next established and characterized a clonal primary gastric cancer cell line GC38-adh from a gastric cancer patient. We used GC38-adh cells for overexpression studies as the percentage of CD44 expressing cells was relatively low (5%) in this cell line. CD44s and CD44v8-10 was stably overexpressed in GC38-adh cells (Fig. 5A) and their proliferation rate was monitored over a period of 5 days, during which we observed that the \textit{in vitro} growth rate of these cells was not significantly different (Fig. 5A). We next sorted CD44s overexpressing and CD44v8-10 overexpressing cells and injected these fractions into NSG mice at limiting cell doses to determine their ability to initiate tumors. Our data show that CD44v8-10 overexpressing cells were able to initiate tumors with higher efficiency than vector control cells ($P=0.0168$) (Supplementary Table S3A, Fig. 5B).

In reciprocal experiments, we performed shRNA-mediated knockdown of total CD44 in gastric cancer cells. We used TMK1 cells because $>99.9\%$ of the cells express the CD44v8-10 isoform (Fig. 5C). We observed that injection of 200 CD44-depleted cells (Supplementary Fig. S4B) did not form tumors, compared to scrambled control shRNA (Supplementary Fig. S4C). To “rescue” this knockdown phenotype, we expressed shRNA-refractory CD44v8-10 as well as CD44s and observed that CD44v8-10 but not CD44s expression rescued the tumor initiation potential of TMK1 cells in which total CD44 was depleted (Fig. 5D). As shown, the percentage of tumor forming mice was higher when injected with CD44v8-10 rescued
cells compared to CD44s rescued cells, and the tumors formed were also larger (Supplementary Fig. S4D). In summary, these findings support the notion that CD44v8-10 plays an important role in tumor initiation.
DISCUSSION

Recent developments in the stem cell field have further validated the cancer stem cell concept. By using genetic techniques to trace the cells during tumor progression in mice, three research groups have provided clear experimental evidence that cancer stem cells exist and drive tumor growth (29-31), further emphasizing the need to eliminate CSCs for effective eradication of cancer.

We identified and enriched for gastric cancer stem cells (CSCs) in primary gastric cancer using EpCAM and CD44 cell surface markers, and further characterized these cells. Several other studies have reported the use of CD44 as a gastric CSC marker. Takaishi et al (6) have previously identified CD44+ cell populations from gastric cancer cell lines which possess features of CSCs, namely tumor sphere formation and tumor formation in vivo. Chen et al. (32) identified CSCs in a subpopulation of CD44+CD54+ cells from gastric tumors and peripheral blood of gastric cancer patients based on their capability to self-renew and differentiate in vitro and in vivo. However, thrice the number of CD44+CD54− and CD44−CD54+ cells could also form tumors in mice, suggesting that the combination of CD44 and CD54 may not be sufficient to enrich for CSCs in gastric tumors. CD44+CD54+ cells were more efficient in cancer initiation only when compared to total cells from biopsy samples, which consisted of both tumor and non-tumor cells. More importantly, although CD44 marks gastric CSCs, it is also highly expressed in hematopoietic cells. Ideally, a targeted therapy should inhibit functions of CSCs but not normal cells. Hence there is a need to identify cancer-specific or CSC markers that are upregulated in tumors.

CD44 and its variants have been reported in gastric cancer (33), and CD44 expression as determined by immunohistochemistry was found to correlate with poor prognosis in patients.
with intestinal type gastric adenocarcinoma (34). The relevance and significance of specific CD44 variants in cancer have not been fully explored, and the nomenclature used to describe these variants can be confusing. We isolated and identified the transcripts of a predominant CD44 variant containing variable exons 8, 9 and 10 that was significantly elevated in gastric tumors compared to normal gastric tissue, referred to in this study as CD44v8-10 (Fig. 3A). The CD44v8-10 surface protein can be recognized and identified using a CD44v9-specific antibody (20, 35) (Supplementary Fig. S2A, 2B), which allowed us to isolate and study CD44v8-10 expressing cells.

We used several patient-derived xenografts models in our identification of gastric CSCs and showed that pan-CD44 or CD44v8-10 expression alone was sufficient to predict enrichment of gastric CSCs. As CD44 is a known Wnt-signaling target (36), we characterized the xenograft models by microarray analyses and found that despite differences in expression of Wnt target genes such as LGR5, ZNF43, ASCL2, CLDN2, SP5 and AXIN2 (36, 37), there was no correlation between gene expression profiles of the Wnt pathway and CD44v8-10 expression (data not shown), indicating that molecular differences in these xenografts had no bearing on the function of CD44v8-10 in tumor initiation. We also evaluated p53 mutation status in a selected number of xenograft lines because p53 has been reported to inhibit expression of the CD44 cell-surface molecule (38). Although we observed p53 mutations in some xenograft lines, we found no correlation between p53 mutation status and proportion of CD44 expressing cells in the xenograft tumor (Supplementary Table S4).

CD44v8-10 (CD44R1) was first described in colon cancer and was found to be upregulated in primary and metastatic tumors but rarely expressed in normal mucosa of adults (24). In the Gan mouse, a genetic model for gastric tumorigenesis, the mouse homolog of human
CD44v8-10 is expressed in precancerous regions of the stomach from which tumor-initiating cells are thought to arise, suggesting that CD44v8-10 may be a marker for gastric tumor-initiating cells in the mouse (39). CD44v8-10 also appears to promote metastasis in colon, pancreatic and breast cancer cells (24, 40, 41), with significant positive correlation to tumor recurrence and mortality (24). Using antibodies directed against the CD44 variant exon 9, CD44v9 expression in primary tumors was significantly and positively associated with tumor recurrence and mortality (35, 42). These studies indicate the importance of CD44v9 as a prognostic marker but functional studies that define CD44v9 expressing cells as gastric CSCs are lacking. Since the CD44v8-10 isoform includes exon 9, our current findings may provide further functional and mechanistic explanation for these correlative observations.

Overexpression of CD44v8-10 but not CD44s resulted in cells with higher tumor initiating potential in vivo although we did not observe any significant changes in proliferation upon ectopic expression of either protein in vitro. This suggests that CD44v8-10 may enhance tumor initiation capability of gastric cancer cells by increasing their resilience to adverse conditions such as hypoxia or oxidative stress. Indeed, there is evidence that CD44v8-10 stabilizes the cystine transporter xCT, thereby promoting the ability of cancer cells to defend themselves against reactive oxygen species (20). In light of these observations, CD44v8-10+ cells may also be responsible for conferring resistance to therapeutic treatment; therefore it is of paramount importance to eradicate this subpopulation of cells for effective treatment of gastric cancer.

In conclusion, we have shown that CD44v8-10 is the major variant found in CD44 expressing gastric cancer cells. The CD44v8-10+ cell population is enriched with gastric cancer stem cells, and CD44v8-10 plays a functional role in tumor initiation. More importantly, CD44v8-
expression is low in normal tissues (both epithelial and hematopoietic cells, Fig. 3C), thus making it an ideal target for directed therapy against gastric cancer stem cells. In addition, our study has broader implications as CD44 also marks CSCs of numerous cancers (11), many of which may also overexpress CD44v8-10. However, as pan-CD44 antibodies were used for prospective isolation of CSCs in most studies, this remains to be investigated. CD44 plays diverse physiological roles in many different tissues (43), thereby undermining its suitability as a CSC target. For instance, injection of antagonizing anti-CD44 antibody induces systemic shock in mice (44), and other side effects of anti-CD44 treatment have also been documented (45). Unlike CD44s that is expressed in many normal tissues, CD44v8-10 is cancer specific; therefore our discovery of CD44v8-10 as a gastric CSC marker opens a critical window for therapeutic targeting of CSCs in other human cancers.
ACKNOWLEDGEMENTS

We gratefully acknowledge H. Saya, O. Nagano and T. Masuko for providing us with the CD44v8-10 antibody. We thank M. Teh for assisting us with histological analysis of patient specimens and T.L.M. Chia for technical assistance with preparation of histology sections and breeding of NSG mice. We also thank M.G.K. Tan and V. Krishnan for critical review of this manuscript.

GRANT SUPPORT

This research was supported by the National Research Foundation Singapore and the Singapore Ministry of Education under the Research Centres of Excellence initiative and Singapore Stem Cell Consortium, A*Star Singapore.

AUTHOR CONTRIBUTIONS

REFERENCES


Table 1. Limiting dilution assay for EpCAM+CD44+ cells

<table>
<thead>
<tr>
<th>Gastric cancer cell source</th>
<th>Cell dose</th>
<th>Number of mice with tumors/Number of mice injected (%)</th>
<th>Estimated Gastric CSC in 10,000 cells</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Unsorted Ep+CD44+ Ep+CD44-</td>
<td>1.2</td>
</tr>
<tr>
<td>GC16</td>
<td>25000</td>
<td>5/5 (100%) ND 3/5 (60%)</td>
<td>1 in 8140 (4782-13855)</td>
</tr>
<tr>
<td></td>
<td>10000</td>
<td>4/5 (80%) ND 4/5 (80%)</td>
<td>1 in 72 (35-148)</td>
</tr>
<tr>
<td></td>
<td>5000</td>
<td>5/12 (42%) ND 3/5 (60%)</td>
<td>1 in 14770 (7586-28758)</td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>2/9 (22%) ND 0/9 (0%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>0/5 (0%) ND 0/5 (0%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>0/3 (0%) 3/3 (100%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0/8 (0%) 6/8 (75%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>0/8 (0%) 2/8 (25%)</td>
<td></td>
</tr>
<tr>
<td>Estimated stem cell</td>
<td></td>
<td>frequency</td>
<td></td>
</tr>
<tr>
<td>(Confidence intervals)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estimated Gastric CSC</td>
<td></td>
<td>in 10,000 cells</td>
<td>1.2</td>
</tr>
<tr>
<td>in 10,000 cells</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GC21</td>
<td>5000</td>
<td>2/2 (100%) 2/2 (100%)</td>
<td>1 in 267 (156-458)</td>
</tr>
<tr>
<td></td>
<td>4000</td>
<td>3/3 (100%) 3/3 (100%)</td>
<td>1 in 83 (48-144)</td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>12/12 (100%) 17/17 (100%)</td>
<td>1 in 1455 (882-2400)</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>3/5 (60%) 5/5 (100%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>400</td>
<td>8/10 (80%) 14/15 (93%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>5/10 (50%) 13/15 (87%)</td>
<td></td>
</tr>
<tr>
<td>Estimated stem cell</td>
<td></td>
<td>frequency</td>
<td></td>
</tr>
<tr>
<td>(Confidence intervals)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estimated Gastric CSC</td>
<td></td>
<td>in 10,000 cells</td>
<td>37.5</td>
</tr>
<tr>
<td>in 10,000 cells</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>37.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GC38</td>
<td>2000</td>
<td>12/13 (92%) 13/13 (100%)</td>
<td>1 in 498 (282-880)</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>4/7 (57%) 6/7 (86%)</td>
<td>1 in 140 (76-259)</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>5/6 (83%) 6/6 (100%)</td>
<td>1 in 1142 (660-1977)</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>2/13 (15%) 5/13 (38%)</td>
<td></td>
</tr>
<tr>
<td>Estimated stem cell</td>
<td></td>
<td>frequency</td>
<td></td>
</tr>
<tr>
<td>(Confidence intervals)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estimated Gastric CSC</td>
<td></td>
<td>in 10,000 cells</td>
<td>20.1</td>
</tr>
<tr>
<td>in 10,000 cells</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 in 1142 (660-1977)</td>
</tr>
<tr>
<td>ND: Not determined.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2. Limiting dilution assay for EpCAM+CD44v8-10+ (Ep+CD44v+) cells

<table>
<thead>
<tr>
<th>Gastric cancer cell source</th>
<th>Cell dose</th>
<th>Number of mice with tumors/Number of mice injected (%)</th>
<th>Estimated stem cell frequency (Confidence intervals)</th>
<th>Estimated Gastric CSC in 10,000 cells</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Unssorted                 Ep+CD44v+          Ep+CD44v-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GC45</td>
<td>2000</td>
<td>13/13 (100%)               13/13 (100%)              8/13 (62%)</td>
<td>1 in 139 (82-237)</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>13/19 (68%)                 19/19 (100%)              3/19 (16%)</td>
<td>1 in 36 (17-78)</td>
<td>277.8</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>4/10 (40%)                  4/10 (40%)               0/10 (0%)</td>
<td>1 in 1844 (988-3440)</td>
<td>5.4</td>
</tr>
<tr>
<td>GC84</td>
<td>20000</td>
<td>4/4 (100%)                  4/4 (100%)               3/4 (75%)</td>
<td>1 in 454 (187-1108)</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>7/7 (100%)                  7/7 (100%)               2/7 (29%)</td>
<td>1 in 75 (34-166)</td>
<td>133.3</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>2/8 (25%)                   7/8 (88%)                1/8 (13%)</td>
<td>1 in 8859 (3320-23641)</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>1/8 (13%)                   3/8 (38%)                0/8 (0%)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
FIGURE LEGENDS

Figure 1. EpCAM expressing cells in gastric cancer tissues are tumor epithelial cells containing subpopulations of CD44+ and CD133+ cells. Flow cytometric analyses of surface marker expression in cells isolated from 28 paired gastric tumors and adjacent non-tumor tissue. (A) EpCAM levels in 28 gastric tumors compared to adjacent non-tumor tissue. Median fluorescence intensity (MFI) was calculated using FlowJo software (Supplementary Fig. S1B). Horizontal bars indicate the mean value. Patient clinical information is included in Supplementary Table S1. (B) Proportion of CD44+ and CD133+ cells in gastric cancer compared to adjacent non-tumor tissue. Representative flow cytometry data plots from 3 gastric cancer patients (GC16, 59 and 71) are shown. Proportion (%) of CD44+ or CD133+ cells within the EpCAM+ population is shown in the top right quadrant. (C) & (D) Proportion of CD44 and CD133 cells in EpCAM expressing cells isolated from 21 paired tumor and non-tumor sites. Cells were stained with antibodies against EpCAM and CD44 or CD133 and analyzed by flow cytometry. Data were analyzed for statistical significance using Student’s paired t-test.

Figure 2. Gastric cancer stem cells are enriched in EpCAM+CD44+ cell fractions. (A) Gastric cancer xenograft lines from 3 clinical biopsies of well differentiated (GC16), moderately differentiated (GC21) and poorly differentiated (GC38) gastric cancer. Upper panel: H&E staining of primary gastric cancer biopsies. Lower panel: H&E staining of respective xenogeneic patient-derived tumors. Scale bars represent 200µm. (B) Proportion of CD44+ cells in respective xenograft lines. (C) Summary of limiting dilution assays estimating the frequency of CSCs in unsorted, EpCAM+CD44+ and EpCAM+CD44- cell fractions from GC16, GC21 and GC38 xenografts. Differences between groups were tested for statistical significance using chi-squared test (degrees of freedom=1, 95% confidence intervals).
Asterisks denote statistically significant P values. Detailed data for limiting dilution analysis and estimation of gastric CSC frequency are shown in Table 1. **(D)** Flow cytometric analysis of EpCAM^+CD44^+ population in primary and secondary xenograft tumors from serial transplantation experiments generated from sorted EpCAM^+CD44^+ cells. **(E)** H&E staining showing histology and cellular composition of tumors initiated by EpCAM^+CD44^+ cells from first and second xenograft transplants. Scale bars represent 100µm.

**Figure 3. CD44v8-10 is the major CD44 variant found on cells that express surface CD44.** **(A)** PCR of cDNA from xenograft lines GC16 (Lane 1), GC21 (Lane 2) and GC38 (Lane 3). MW denotes molecular weight marker. Diagram of human CD44 gene structure showing alternatively spliced variant exons v1-10 in colored boxes and constant exons in grey boxes. CD44v8-10 comprises all constant exons and variant exons 8-10 while CD44s contains only constant exons. PCR primers are located in constant exons 5 and 16/17. **(B)** Real-time qPCR of CD44v8-10 expression in xenograft tumors of GC16, 21 and 38 normalized to 18S rRNA internal control. Error bars represent Mean ± SD. **(C)** Flow cytometric analyses of surface expression of CD44v8-10 in EpCAM^+CD45^- and CD45^+ cells isolated from tumor and adjacent normal tissues of patient GC101. Proportion (%) of CD44v8-10 positive (CD44v+) and CD44v8-10 negative cells within the EpCAM^+CD45^- or CD45^+ populations are shown in the top and lower right quadrants respectively. **(D)** Percentage of CD44v8-10^- (CD44v+) cells in EpCAM^+CD45^- population from tumor and normal tissues of 3 gastric cancer patients GC101, GC121 and GC123. **(E)** Percentage of CD44v8-10^- cells within the CD44^+ population of gastric cancer patients GC101, GC121 and GC123.
Figure 4. CD44v8-10 marks gastric cancer stem cells. 
(A) Flow cytometric analysis of CD44v8-10+ and CD44v8-10- (CD44v+ or CD44v- respectively) cells in unsorted GC45 and GC84 xenograft tumors and the purity of sorted fractions. 
(B) Summary of limiting dilution assays estimating CSC frequency in unsorted, CD44v+ and CD44v- cell fractions from GC45 and GC84 xenografts. Differences between groups were tested for statistical significance using chi-squared test (degrees of freedom=1, 95% confidence intervals). Asterisks denote statistically significant P values. Detailed data for limiting dilution analysis and estimation of gastric CSC frequency are shown in Table 2. 
(C) Flow cytometric analysis of CD44v+ population in first and second serially transplanted xenograft tumors generated from primary CD44v+ cells. 
(D) H&E staining showing histology of tumors from first and second xenograft transplants initiated by CD44v+ cells. Scale bars represent 100µm. 
(E) Comparison of tumor initiating potential between CD44+, CD44+ and CD44v8-10+ fractions from xenograft GC119. 200 cells of the indicated cell fraction were injected and proportion of tumor-forming NSG mice is plotted. 
(F) qPCR of CD44s and CD44v8-10 expression in 26 primary tumor samples versus adjacent non-tumor tissue, normalized to 18S rRNA internal control. Horizontal bars represent the mean value. Paired Student’s t-test was performed for statistical significance.

Figure 5. CD44v8-10 is upregulated in clinical gastric cancer and plays a functional role in tumor initiation. 
(A) Immunoblot of primary gastric cancer cells GC38-adh stably expressing CD44s and CD44v8-10. Cell proliferation assay was performed over 5 days. Data points represent mean ± SD of five replicates. Difference in slopes indicates cell proliferation rate. Results are representative of two independent experiments. 
(B) Xenograft tumor formation potential of fractionated GC38-adh cells overexpressing CD44s and CD44v8-10. Chi-squared tests were performed for statistical significance (degrees of freedom = 1). Asterisks denote statistically significant P values. Data from two independent experiments are
represented, with cell doses and number of mice used reflected in Supplementary Table S3A. (C) Upper panels: CD44v8-10 expression in TMK1 cells analyzed by flow cytometry. Lower panels: post-sort analyses of CD44-silenced TMK1 cells rescued with indicated shRNA-refractory genes. Percentage purity of sorted cells is indicated. These cells were used in xenograft tumor formation assay. (D) Xenograft tumor formation in NSG mice by CD44-silenced TMK1 cells harboring the indicated shRNA-refractory CD44 variants. Xenograft tumor formation potential for the indicated cell population was determined by limiting dilution assay. Differences between groups were tested for statistical significance using chi-squared test (degrees of freedom=1, 95% confidence intervals). Asterisks denote statistically significant P values. Detailed data for limiting dilution analysis is presented in Supplementary Table S3B.
Figure 2

A

Primary Patient Biopsy

Xenograft

B

Pairwise tests for difference in stem cell frequencies

<table>
<thead>
<tr>
<th>Group 1</th>
<th>Group 2</th>
<th>$\chi^2$</th>
<th>$P$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>GC16</td>
<td>CD44+</td>
<td>1469</td>
<td>2.7e-321</td>
</tr>
<tr>
<td>CD44+</td>
<td>CD44-</td>
<td>66.2</td>
<td>1.5e-16</td>
</tr>
<tr>
<td>CD44+</td>
<td>Unsorbed</td>
<td>1.95</td>
<td>0.163</td>
</tr>
<tr>
<td>GC21</td>
<td>CD44-</td>
<td>69</td>
<td>9.7e-17</td>
</tr>
<tr>
<td>CD44+</td>
<td>Unsorbed</td>
<td>10.4</td>
<td>2.0e-07</td>
</tr>
<tr>
<td>CD44-</td>
<td>Unsorbed</td>
<td>21.6</td>
<td>3.3e-06</td>
</tr>
<tr>
<td>GC38</td>
<td>CD44-</td>
<td>27</td>
<td>2.0e-07</td>
</tr>
<tr>
<td>CD44+</td>
<td>Unsorbed</td>
<td>9.96</td>
<td>0.0016</td>
</tr>
<tr>
<td>CD44-</td>
<td>Unsorbed</td>
<td>4.9</td>
<td>0.028</td>
</tr>
</tbody>
</table>

C

D

Sorted EpCAM+CD44+

First transplant

Second transplant

E

Sorted EpCAM+CD44+

Unsorted

First transplant

Second transplant
Figure 3

A

B

C Primary cells from GC101

Further gating of EpCAM+CD45- cells

Further gating of CD45+ cells

Tumor

Normal

D E

% of CD44+ cells in Ep+CD44+ population

% of CD44+ cells in CD44+ population

0 10 20 30 40 50

0 20 40 60 80 100

GC101 GC121 GC123

GC101 GC121 GC123
CD44v8-10 is a cancer-specific marker for gastric cancer stem cells

Wen Min Lau, Eileen Teng, Hui Shan Chong, et al.

Cancer Res  Published OnlineFirst March 11, 2014.

Updated version  Access the most recent version of this article at: doi:10.1158/0008-5472.CAN-13-2309

Supplementary Material  Access the most recent supplemental material at: http://cancerres.aacrjournals.org/content/suppl/2014/03/12/0008-5472.CAN-13-2309.DC1

Author Manuscript  Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

E-mail alerts  Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions  To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions  To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.