

Keywords: ALDH1B1; colorectal cancer; ALDH1A1; ALDH2; CD44

Aldehyde dehydrogenase 1B1: a novel immunohistological marker for colorectal cancer

Akiko Matsumoto^{*1,6}, John Arcaroli², Ying Chen^{1,7}, Maura Gasparetto^{3,8}, Veronique Neumeister⁴, David C Thompson⁵, Surendra Singh^{1,7}, Clayton Smith^{3,8}, Wells Messersmith² and Vasilis Vasiliou^{*1,7}

¹Department of Pharmaceutical Sciences, University of Colorado Denver, Aurora, CO 80045, USA; ²Division of Medical Oncology, University of Colorado School of Medicine, Aurora, CO 80045, USA; ³British Columbia Cancer Agency/Terry Fox Laboratory, Vancouver, BC V5Z 1L3, Canada; ⁴Department of Pathology, Yale University, School of Medicine, New Haven, CT 06510, USA and ⁵Department of Clinical Pharmacy, University of Colorado Denver, Aurora, CO 80045, USA

Background: Aldehyde dehydrogenase (ALDH) 1A1 is an immunohistological biomarker of various solid tumours, but has not been successfully proved as a colorectal cancer (CRC) marker. We recently reported that ALDH1B1, which has functional roles in tumourigenesis, may be a better CRC marker than ALDH1A1.

Methods: Human CRC explants and cell lines were analysed to identify candidate CRC markers from eight ALDH isozymes including ALDH1A1 and ALDH1B1. A tissue microarray, including paired specimens of normal and tumour tissues, was subsequently analysed to determine if candidate ALDHs could distinguish CRC from normal tissue.

Results: Based on mRNA analysis, *ALDH1B1* and *ALDH2* were selected as suitable candidates. These were strongly and regularly expressed in tumour tissue and cell lines, including highly tumourigenic cell populations (ALDH⁺CD44⁺ cells), while other *ALDHs*, including *ALDH1A1*, showed differential or low expression. No genetic alteration of *ALDH1B1* in CRC was suggested by the relationships between mRNA and protein levels/enzymatic activities, and cDNA sequences of CRC cell lines. Tissue microarray findings showed that *ALDH1B1*, but not *ALDH2*, could distinguish CRC from normal tissue. Furthermore, ratios of *ALDH1B1* to *ALDH1A1* or *ALDH2* were found to be powerful CRC indicators.

Conclusions: These results suggest that *ALDH1B1* is a novel human CRC biomarker.

Aldehyde dehydrogenases (ALDHs) are enzymes that oxidise endogenous and exogenous aldehydes. To date, 19 ALDH isozymes have been identified in the human ALDH superfamily (Marchitti *et al*, 2008). Of these, *ALDH1A1* has been shown to be associated with highly malignant types of solid tumours, such as cancers of the breast, lung, prostate, and ovary (Ginestier *et al*, 2007; Patel *et al*, 2008; Li *et al*, 2010; Meng *et al*, 2014), making this isozyme a potential predictive biomarker of these cancers. Expression of ALDHs in

colorectal cancer (CRC) has also been investigated, but the evidence does not support *ALDH1A1* as an effective biomarker for CRC (Kahlert *et al*, 2012; Zhou *et al*, 2015). We previously reported that another ALDH1 family member, *ALDH1B1*, might be a better CRC biomarker than *ALDH1A1* (Chen *et al*, 2011); a role for *ALDH1B1* in CRC tumourigenesis was subsequently identified (Singh *et al*, 2015).

The present study was designed to extend these previous studies by investigating which ALDH isozyme is most consistently

*Correspondence: Professor V Vasiliou; E-mail: vasilis.vasiliou@yale.edu or Professor A Matsumoto; matsumoa@cc.saga-u.ac.jp

⁶Current address: Department of Social Medicine, Saga Medical School, Saga 849-8501, Japan.

⁷Current address: Department of Environmental Health Sciences, Yale School of Public Health, New Haven, CT 06520, USA.

⁸Current address: Division of Hematology, University of Colorado Denver, Aurora, CO 80045, USA.

Revised 2 July 2017; accepted 8 August 2017; published online 7 September 2017

© 2017 Cancer Research UK. All rights reserved 0007–0920/17

expressed by or associated with CRC. First, we analysed mRNA levels of eight ALDH isozymes previously reported to be potential tumour markers (Marchitti *et al*, 2010; Marcato *et al*, 2011; Seidensaal *et al*, 2015) or to be carcinogenic (Nishiyama *et al*, 2015), including ALDH1A1 and ALDH1B1. Second, we investigated the possibility of post-translational or genetic modification of candidate ALDHs by analysing mRNA and protein expression, enzymatic activity levels, and cDNA sequences. Finally, the potential of candidate ALDHs for detecting CRC was tested using tissue microarray, which included normal/CRC paired samples.

MATERIALS AND METHODS

Explants and cell lines. Patient-derived CRC tissue samples were collected from consenting patients between 2007 and 2011 at the University of Colorado Hospital in accordance with protocols approved by the Colorado Multiple Institutional Review Board. They were obtained from 26 patients aged 32–73 (mean, 54.5 years) who had primary or metastatic CRC of tumour-node-metastasis (TNM) stages II–IV.

Immortalised human CRC cell lines, that is, BE, Caco-2, COLO320 DM, HCT116, HT29, and SW480 cells at passage 4 (p4), p26, p2, p4, p4, and p4 (respectively), were kindly provided by Dr David Ross (University of Colorado School of Pharmacy). All cell lines were cultured in plastic flasks or Petri dishes in RPMI 1640 medium supplemented with 10% foetal bovine serum, 100 U ml⁻¹ penicillin, 100 µg ml⁻¹ streptomycin, and 0.25 µg ml⁻¹ amphotericin B (Sigma-Aldrich, St Louis, MO, USA). Cells were maintained at 37 °C in an environment of 5% CO₂ in air. Media and supplement exchanges were performed twice weekly.

Messenger RNA sequencing. Total RNA from each explant was extracted using the RNeasy kit (Qiagen, Germantown, MD, USA). Total RNA libraries were constructed according to the Illumina TruSeq RNA Sample Preparation v2 Guide (Illumina Inc., San Diego, CA, USA). Complementary DNA library integrity was determined on an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). The Illumina cBot system was used for cluster generation, and sequencing was performed on the Illumina HiSeq 2000 (Illumina). A de-multiplexing step was carried out before the alignment step for multiplexed lanes/samples. Mapping against the human genome (Trapnell *et al*, 2009) was performed with TopHat (version 1.3.2). Gene expression was estimated as the sum of the normalised fragment counts (fragment number/kilobase of exon/million mapped fragments (FPKM)) of multiple transcripts that represent the same gene (Garber *et al*, 2011). The fragment count was normalised to the exon length and total fragment amount to achieve a meaningful comparison between samples and genes.

Quantitative PCR. Total RNA from CRC cell lines was extracted as described above. The cDNA was then synthesised using a SuperScript I First-Strand Synthesis System (Invitrogen, Frederick, MD, USA). Ribonucleic acid from sorted cell populations was extracted using TRIzol (Invitrogen) in a different laboratory, transported at –20 °C, and stored at –80 °C until cDNA synthesis was conducted. Messenger RNA levels were analysed with respective primers (Supplementary Table S1) and the SYBR Green PCR Master Mix using ABI PRISM 7000 (Applied Biosystems, Foster City, CA, USA). Individual ALDH mRNA copy numbers were quantified using a standard curve generated from respective cDNA constructs of the gene.

Protein assays. Cells were scraped from culture dishes, resuspended in RIPA buffer (25 mM Tris-HCl pH 7.6, 150 mM NaCl, 1%

NP-40, 1% sodium deoxycholate, 0.1% sodium dodecyl sulphate) containing proteinase inhibitors at the concentration recommended by the manufacturer (Promega, Madison, WI, USA), and sonicated (10 s, three times) to obtain homogenates. Protein concentrations in the homogenate were assessed using a Bio-Rad Protein Assay Kit (Bio-Rad, Hercules, CA, USA). Protein separation and western blotting were performed using protocols and antibodies previously described (Chen *et al*, 2011). β -actin was used as a loading control. Quantitative analysis of the signal intensity of western blot protein bands was performed using ImageJ (National Institutes of Health, Bethesda, MD, USA).

Enzyme activity assay. Freshly collected lysates of CRC cell lines obtained by scraping were washed with phosphate-buffered saline (4 °C, pH 7) and homogenised in 0.1 M potassium phosphate buffer (pH 7.5) containing 1 mM EDTA and 1 µM 2-mercaptoethanol. Aldehyde dehydrogenase enzymatic activity of whole homogenates was determined by monitoring the formation of NADH (when acetaldehyde and propionaldehyde were used as substrates) and NADPH (when benzaldehyde was used as substrate) at 340 nm during the oxidation of aldehyde substrates for 5 min at room temperature using a spectrophotometer (Beckman DU-640) (molar extinction coefficient: 6.22 mM⁻¹ cm⁻¹). The reaction was initiated by adding 0.1 ml of substrate (1 mM acetaldehyde, 100 µM propionaldehyde, or 500 µM benzaldehyde) to 0.9 ml of reaction mixture containing 0.1 M sodium pyrophosphate (pH 8.0), 20 mM NAD or 50 mM NADP, 20 mM pyrazole, 1 mM 2-mercaptoethanol, and 50 µl of cell homogenate (5–8 mg protein ml⁻¹).

Flow cytometry and cell sorting. After gentle cellular detachment from the culture plate using trypsin (0.04%)/EDTA (0.03%) solution, cells were resuspended and washed in phosphate-buffered saline, and pelleted by centrifugation. CD44 staining was performed using a monoclonal antibody and reagent from BD Biosciences (San Jose, CA, USA). Aldehyde dehydrogenase detection was performed using the Aldefluor assay (StemCell Technologies, Vancouver, BC, Canada) as follows: a total of 1 × 10⁶ cells ml⁻¹ was resuspended in Aldefluor buffer and incubated at 37 °C for 30 min with 2.5 µl ml⁻¹ activated Aldefluor reagent, which includes a fluorescent-labelled substrate of ALDHs. Baseline fluorescence was established using control cells in the presence of 5 µl ml⁻¹ of 1.5 mM dimethylaminobenzaldehyde (DEAB), an ALDH inhibitor. Cell staining and analytical procedures were performed either manually or using high-content semi-automated flow cytometry as described previously (Gasparetto *et al*, 2004). Flow cytometric analysis and cell sorting were performed on a FACSAria device equipped with 405 nm violet, 488 nm argon, and 633 nm HeNe lasers (Becton Dickinson, San Jose, CA, USA). Data were analysed using FlowJo software (TreeStar, Palo Alto, CA, USA).

Tissue microarray. A tissue microarray (TMA) was constructed at the University of Colorado Pathology Department from previously formalin-fixed paraffin-embedded specimens (22 normal and 32 tumour tissues, including 22 matched pairs of normal and tumour tissue from the same patient) in accordance with protocols approved by the Colorado Multiple Institutional Review Board. A hollow needle was used to remove representative 0.6-mm diameter tissue cores (in duplicate for each specimen) from different locations while avoiding necrotic tissue. The TMA was designed with control normal tissue samples that included: appendix, bladder, breast, colon, liver, ovary, pancreas, and spleen. The 0.6 mm cores were cut into 5 µm sections and adhered to glass slides.

Slides were deparaffinised with xylene and rehydrated with ethanol. Antigen retrieval was performed using citrate buffer

(pH 6) at a temperature of 97 °C for 20 min. After blocking of endogenous peroxidase with methanol and hydrogen peroxide, slides were preincubated with 0.3% bovine serum albumin in 0.1 M Tris-buffered saline for 30 min at room temperature. Slides were then incubated with cytokeratin and the primary antibody for ALDH1A1, ALDH1B1, and ALDH2 at previously optimised staining conditions. Mouse/rabbit EnVision reagent (Agilent Technologies, Santa Clara, CA, USA) and Alexa 546-conjugated goat anti-rabbit/mouse secondary antibody (Molecular Probes, Eugene, OR, USA; 1:100 dilution) were used as secondary antibodies that were visualised using Cy5-tyramide (PerkinElmer, Waltham, MA, USA). 4'-diamidino-2-phenylindole (DAPI) staining was used to identify tissue nuclei.

Automated quantitative analysis. Automated quantitative analysis (AQUA) was performed to allow the objective and accurate measurement of protein expression within defined tumour areas and subcellular compartments as described previously (Camp *et al*, 2002; McCabe *et al*, 2005). Briefly, after immunofluorescent staining of the TMA, a series of monochromatic, high-resolution images were captured using an Olympus AX-51 epi-fluorescent microscope (Olympus, Tokyo, Japan). For each histospot represented on a TMA, images for three different channels were obtained to visualise nuclei, cytokeratin, and ALDHs using DAPI, Alexa 546, and Cy5-tyramide, respectively. To set the area for quantification, the cytokeratin signal was binarised. The pixel intensity of antibody-labelled ALDHs was measured within these compartments and divided by the compartment area, resulting in a continuous scoring system, which was directly proportional to the concentration of ALDHs.

Statistical analyses. Statistical analyses were performed using StatView software for Windows Version 5.0 (SAS Institute Inc., Cary, NC, USA). The Mann–Whitney *U*-test was used to compare two groups of mRNA levels of human CRC explants. Regression analyses were conducted to detect significant associations among mRNA and protein expression and enzymatic activity levels across all cell lines studied. Paired or unpaired *t*-tests (as appropriate) were performed to detect differences in immunohistochemical signal intensity between normal tissue and CRC tissue. A value of $P < 0.05$ was considered to be significant.

RESULTS

ALDH2 and ALDH1B1 mRNA is highly expressed in human CRC tissue and cell lines. Of the eight ALDHs investigated, the highest mRNA expression was observed for *ALDH1A1* in a human tissue (329 FPKM); however, it showed great expression variability among the tumour samples, with the lowest level < 1 FPKM. Expression of *ALDH1B1*, *ALDH2*, *ALDH3A2*, and *ALDH3B1* was > 1 FPKM in all tumour tissue samples (each minimum value was 3, 36, 5, 2, respectively), while *ALDH1A1*, *ALDH1A2*, *ALDH1A3*, and *ALDH3A1* showed < 1 FPKM in some or all samples. Cancer stage (TNM II/III vs TNM IV) had no impact ($P > 0.05$) on mRNA expression levels of the ALDH isozymes examined (Figure 1A). Similarly, no differences ($P > 0.05$) were found in mRNA expression levels between primary CRC and metastatic CRC samples (Figure 1B).

As seen for the explants, *ALDH1B1* and *ALDH2* mRNA was regularly expressed in the six cell lines (Figure 2A). *ALDH1A1* mRNA expression levels varied widely among the cell lines. Other ALDHs showed different patterns of expression between explants and cell lines and were not consistently detected in all samples. For example, *ALDH3A2* and *ALDH3B1* mRNA expression was detectable in tumour tissues but undetectable in almost all cell lines. *ALDH1A2* and *ALDH1A3* expression was low and undetectable in many tumour tissue samples (median, < 1 FPKM) but showed high expression in some cell lines.

Relationship between mRNA, protein expression, and enzyme activity levels in human CRC cell lines. Aldehyde dehydrogenase 1A1 protein expression was not detected in BE, HCT116, or SW480 cells. ALDH1B1 and ALDH2 were detected in all of the cell lines (Figure 2B). Regression analysis of protein and mRNA levels revealed a significant relationship for ALDH1A1 ($r = 0.88$, $P = 0.02$) and ALDH2 ($r = 0.82$, $P = 0.05$) in all cell lines. Aldehyde dehydrogenase 1B1 showed a weaker correlation between its mRNA and protein expression levels ($r = 0.65$, $P = 0.17$).

The ALDH enzymatic activities were assessed in the cell lines using three major substrates: acetaldehyde and propionaldehyde (being reflective of ALDH1A1, ALDH1B1, and ALDH2 activity) (Lassen *et al*, 2005; Manzer *et al*, 2003; Stagos *et al*, 2010), and benzaldehyde (indicative of ALDH3 activity) (Pappa *et al*, 2003).

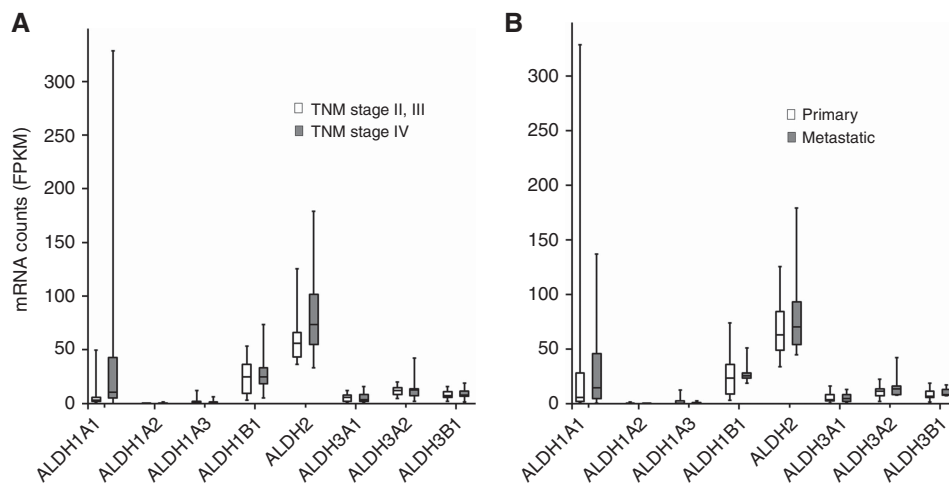


Figure 1. ALDH mRNA levels in human CRC tissue samples by RNA sequencing. (A) Messenger RNA expression levels according to TNM classification. TNM stage II or III (white boxes, $n = 9$) and IV (grey boxes, $n = 17$). (B) Messenger RNA expression levels according to primary tumours (white boxes, $n = 18$) or metastases (grey boxes, $n = 8$). Bottom and top of each box represents the first and third quartiles, respectively. Horizontal line in the box indicates the median. The ends of the whiskers are maximum and minimum values.

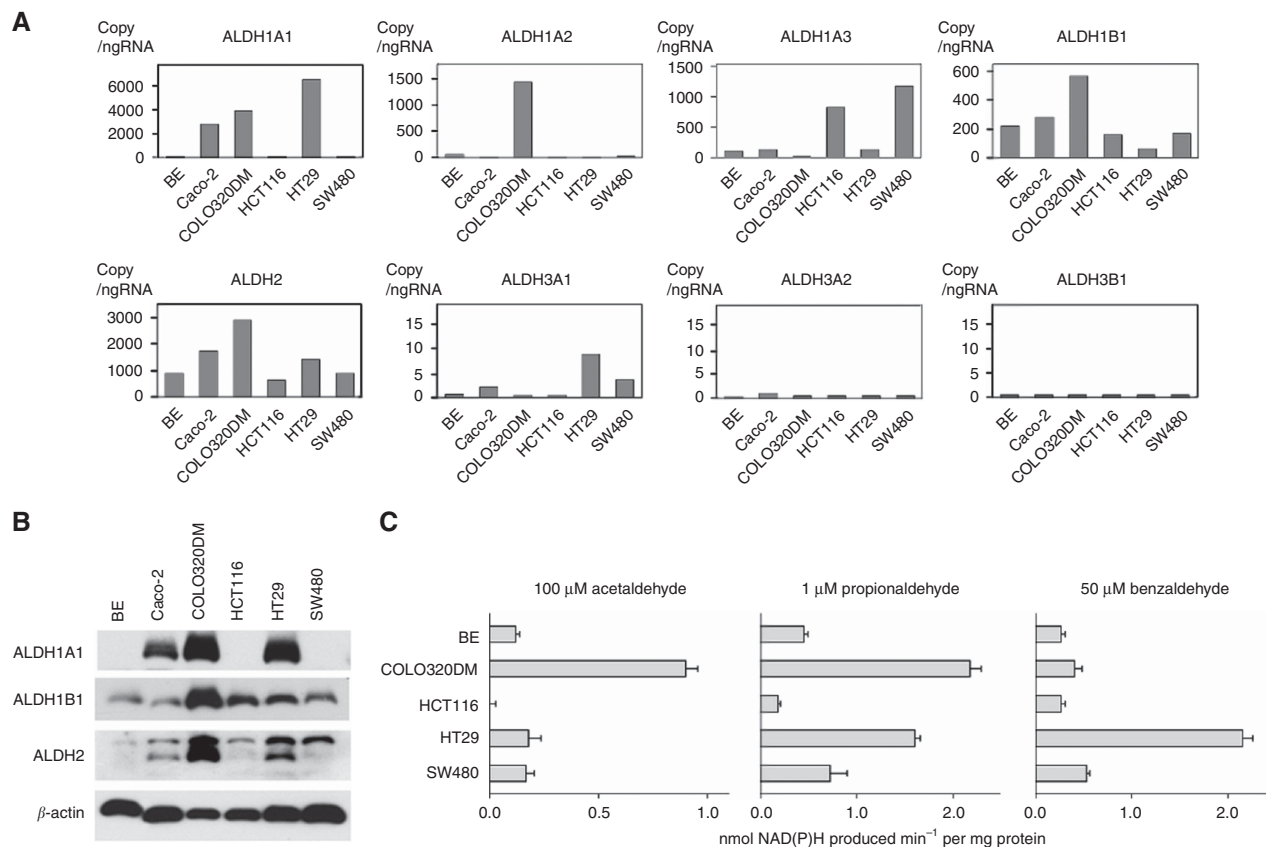


Figure 2. Messenger RNA and protein expression and enzymatic activity of ALDH isozymes in human CRC cell lines. **(A)** Messenger RNA expression levels were measured by real-time PCR. Data are presented as the mean of three replicates. **(B)** Western blot analysis of ALDH1A1, ALDH1B1, ALDH2, and β -actin. Ten μ g of protein was applied to each lane. **(C)** ALDH enzymatic activity. Substrates tested include acetaldehyde, propionaldehyde, and benzaldehyde. Data are presented as the mean and associated s.e.m. ($n = 3$).

Acetaldehyde metabolising activities correlated with mRNA levels of *ALDH1B1* ($r = 0.92$, $P = 0.04$) and *ALDH2* ($r = 0.98$, $P < 0.01$). Similarly, metabolism of propionaldehyde correlated with *ALDH2* mRNA levels ($r = 0.93$, $P = 0.02$), but not with *ALDH1B1* mRNA levels ($r = 0.60$, $P = 0.29$). *ALDH1A1* mRNA levels showed no significant correlation with any examined enzymatic activities, while *ALDH3A1* mRNA levels correlated with enzymatic activity for benzaldehyde ($r = 0.95$, $P = 0.01$). The relationship between protein levels and enzymatic activity for acetaldehyde was almost significant for ALDH2 ($r = 0.88$, $P = 0.05$) and ALDH1B1 ($r = 0.86$, $P = 0.06$), no correlation was observed for ALDH1A1 ($r = 0.80$, $P = 0.10$). In the case of propionaldehyde enzymatic activity, ALDH1A1 and ALDH2 were significantly correlated ($r = 0.97$, $P < 0.01$ or $r = 0.97$, $P < 0.01$, respectively), while ALDH1B1 showed a weaker correlation between protein and enzymatic activity levels ($r = 0.78$, $P = 0.12$).

Taken together, mRNA and protein expression and enzymatic activities were correlated for ALDH2. Although ALDH1A1 and ALDH1B1 showed the same tendency, the level of correlation was much weaker than that for ALDH2. To examine the possibility that ALDH1B1 was functionally modified by genetic alterations in CRC, we examined the entire *ALDH1B1* cDNA sequence of the six CRC cell lines but found no differences among any of the cells (Supplementary Table S2).

ALDH isozyme mRNA expression in a highly tumorigenic cell population. In a xenograft model, cells exhibiting high ALDH enzymatic activity (ALDH⁺) and expressing CD44 (CD44⁺) or CD133 (CD133⁺) were shown to have a stronger potential to initiate and expand tumours than cells that had high ALDH enzymatic activity alone (Huang *et al.*, 2009). In the present study,

mRNA expression for all ALDH isozymes was higher in ALDH⁺ CD44⁺ cells than in the total ALDH⁺ cell population. However, only *ALDH1B1* and *ALDH2* mRNA was expressed in all cell lines (Figure 3). The expression of *ALDH3A1*, *ALDH3A2*, and *ALDH3B1* mRNA was not analysed because levels in the cell populations were low, that is, less than 10 copies ng⁻¹ RNA.

AQUA analysis of ALDH isozymes in CRC. To determine the utility of using ALDHs (e.g., ALDH1B1 and ALDH2), for CRC detection, the TMA was tested using the AQUA system (Figure 4). Among the 22 matched specimens (i.e., normal and tumour tissue from the same donor), nine pairs for ALDH1B1 and 10 pairs for ALDH1A1 and ALDH2 were successfully quantified with AQUA. Unexpectedly, the AQUA score for ALDH1A1 was significantly lower in CRC than in normal tissue ($P = 0.012$), while that for ALDH1B1 was higher in CRC than in normal tissue ($P = 0.048$). No difference was detected for ALDH2 ($P = 0.122$). The ratio of ALDH1B1 to ALDH1A1 or ALDH1B1 to ALDH2 was found to be powerful in detecting differences between normal and CRC tissue ($P = 0.018$ and 0.034 , respectively) (Figure 4C), which was reproduced in an unmatched comparison with whole samples that were successfully quantified in normal tissue ($n = 13$) and CRC tissue ($n = 24$) ($P < 0.0001$ for both ratios) (Figure 4D).

DISCUSSION

ALDH1B1 and *ALDH2* mRNA expression in human CRC tissue samples and cells was found to be consistently high. This was not the case for other ALDH isozymes, which were undetectable or low in some or all samples. Western blot and ALDH enzymatic activity

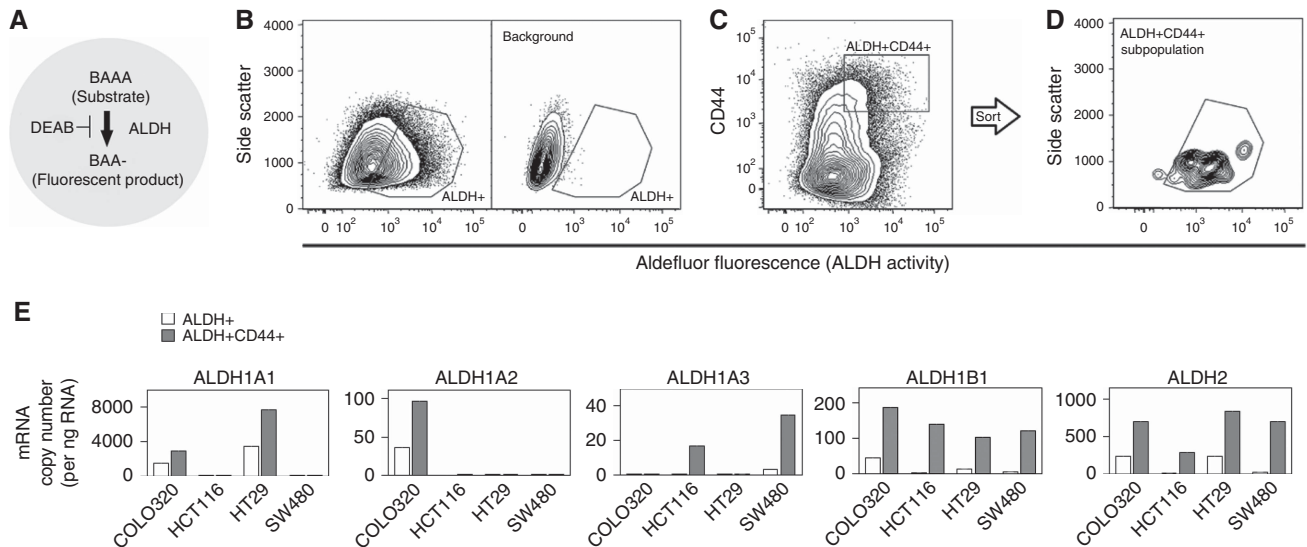


Figure 3. ALDH⁺CD44⁺ cells showed high expression of *ALDH1B1* and *ALDH2*. **(A)** Schematic of Aldefluor system. The ALDH⁺ population was selected by its ability to metabolise aminoacetaldehyde. Fluorescent dye-modified aminoacetaldehyde (BAAA) was applied to cells in the presence or absence of ALDH inhibitor (DEAB). Cells that fluoresced in the presence of DEAB were excluded (background). **(B–D)** Selection of subpopulations. Representative two-parameter scatter plots of flow cytometry are shown for COLO320 DM cells. Blue box shows the selected population. **(B)** Selection of ALDH⁺ subpopulation. Right panel shows background fluorescence, that is, cells fluorescing in the presence of DEAB. **(C and D)** Selection of ALDH⁺CD44⁺ subpopulation. The ALDH⁺CD44⁺ subpopulation was selected from the whole population, that is, no background subtraction step was included **(C)**. ALDH⁺CD44⁺ cells were post-tested with ALDH-SSC two-parameter histograms (confirmation of background exclusion) **(D)**. **(E)** Aldehyde dehydrogenase isozyme mRNA levels in the subpopulations. Messenger RNA expression levels were measured in ALDH⁺ and ALDH⁺CD44⁺ subpopulations separated from the whole population on separate days. Data are the mean of three replicates. A full colour version of this figure is available at the *British Journal of Cancer* journal online.

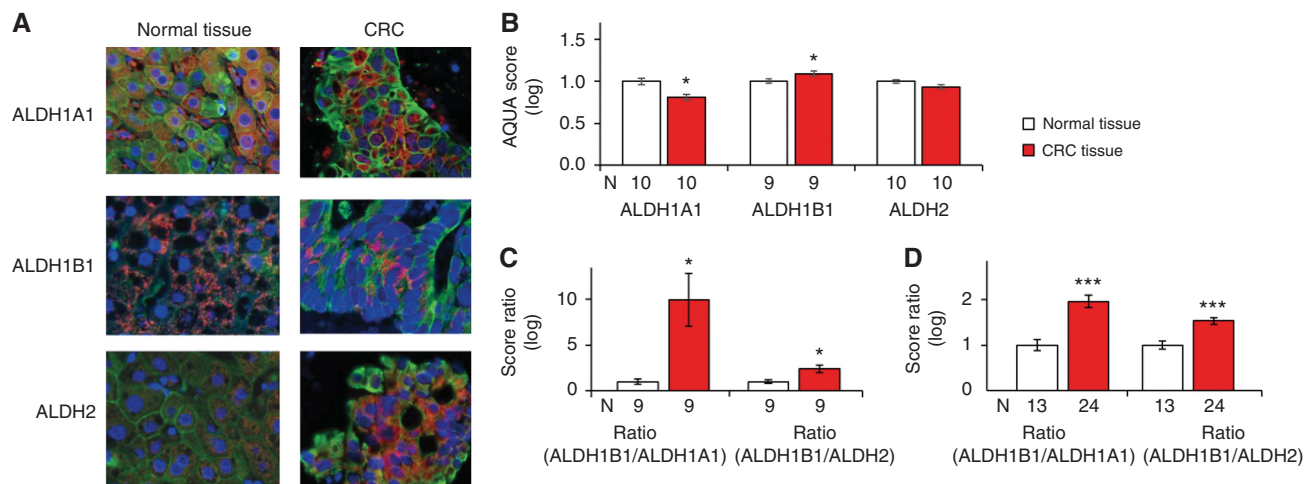


Figure 4. Fluorescent multiplexed analysis of ALDHs. **(A)** Representative histological images (magnification = × 400) of normal and CRC tissue. Blue: nuclei, green: cytokeratin, red: ALDH. **(B)** Automated quantitative analysis scores for ALDH1A1, ALDH1B1, and ALDH2 in matched pairs of normal and CRC tissue obtained from the same donors (**P* < 0.05, paired *t*-test). **(C)** Automated quantitative analysis score ratio of ALDH1B1 to ALDH1A1 or ALDH1B1 to ALDH2 in normal and CRC matched pairs of tissue (**P* < 0.05, paired *t*-test). **(D)** Ratio of ALDH1B1 to ALDH1A1 or ALDH1B1 to ALDH2 in normal and CRC tissue obtained from 24 donors (****P* < 0.0001, unpaired *t*-test). Data represent the mean ± s.e.m. from *N* samples.

analysis of CRC cell lines revealed significant correlations among mRNA and protein expression and enzymatic activity for ALDH2, and weaker correlations for ALDH1B1. Based on this observation, cDNA sequences were analysed in all of the cell lines to determine whether ALDH1B1 was functionally modified by genetic alterations in CRC; however, no mutation was found. *ALDH1B1* and *ALDH2* mRNA was highly expressed in ALDH⁺CD44⁺ cells (a proposed highly tumorigenic stem cell population (Huang *et al*, 2009)) in all human cell lines examined. Finally, paired normal and

CRC tissues from the same donors were compared using the AQUA system, which showed that ALDH1B1, but not ALDH2, is expressed at higher levels in CRC than in normal tissue. Notably, the ratio of ALDH1B1 to ALDH1A1 or ALDH2 showed a strong ability to distinguish CRC tissue from normal tissue.

ALDH1B1 is a promising CRC marker. Among the eight isozymes tested, only the expression of ALDH1B1 was consistently found to correlate with the presence of CRC in cells and tissue

samples. Furthermore, the ratios of ALDH1B1/ALDH1A1 and ALDH1B1/ALDH2 were found to be powerful CRC detectors. In support of this proposal, a previous study investigating potential CRC markers (i.e., CD29, CD44, ALDH1A1, ALDH1B1, EpCam, and CD166) showed that only ALDH1B1 could distinguish CRC from normal tissue (Langan *et al*, 2012). ALDH2 showed high expression in the human CRC cell lines but was less able to distinguish CRC tissue from normal tissue in human CRC explants. This may relate to ALDH2 being expressed in various normal tissues (Stewart *et al*, 1996; Stagos *et al*, 2010; Chiang *et al*, 2012). Given the inconsistent or low mRNA profiles of other ALDHs, including ALDH1A1, it may be speculated that their contribution to CRC is relatively small. The present results do not support a role for ALDH1A1 in CRC detection because its expression in human cell lines was inconsistent and its AQUA scores decreased in CRC tissues. As such, we suggest measurement of ALDH1B1 expression using the AQUA system could be a useful marker for the detection of CRC tissue.

ALDH1B1 is a functionally important cancer-related factor. We recently reported that *ALDH1B1*-silenced SW480 cells showed inhibited proliferation in a spheroid assay and xenograft model (Singh *et al*, 2015). The effect of ALDH1B1 in tumorigenesis appears to involve modulation of the Wnt-, Notch-, and PI3K/Akt signalling pathways (Singh *et al*, 2015). Our current finding that a supposedly highly tumorigenic CRC subpopulation shows high expression of *ALDH1B1* mRNA lends further support to this hypothesis.

Based on its possible ability to enzymatically convert retinal to retinoic acid, ALDH1B1 is presumed to promote differentiation of stem cells (Chute *et al*, 2006; Jackson *et al*, 2015). Indeed, a previous study reported that ALDH1B1 protein expression in human CRC was increased in line with the extent of differentiation (Langan *et al*, 2012). Given that differentiation reduces the self-renewing capacity of cells (Todaro *et al*, 2010), ALDH1B1 would be predicted to decrease the degree of malignancy. However, such an anti-cancer action would be contrary to the tumorigenic role of ALDH1B1 proposed herein. Recently, we found that the PI3K/Akt pathway could be accelerated by retinoic acid (Singh *et al*, 2015). This pathway regulates cell growth under normal conditions; however, in CRC, it becomes hyper-activated by the aberration of associated genes, resulting in the promotion of malignant growth (Danielsen *et al*, 2015). Retinoic acid may therefore have two contradictory roles in CRC: one in the differentiation of stem cells that reduces the potential of malignancy, and another in promoting tumour growth through a hyper-activated PI3K/Akt pathway.

Future challenges in the application of ALDH1B1 as a CRC marker. For the practical application of a proposed marker, a cutoff value to distinguish CRC from normal colorectal tissue needs to be established; this will require evaluation in a larger number of specimens.

The possibility of using ALDH1B1 as a predictive biomarker for CRC should be examined in future studies. A pilot study by Langan and colleagues (2012) suggested an association between ALDH1B1 expression and CRC differentiation, but it failed to detect an association between ALDH1B1 disease stage; ALDH1A1 was shown to be linked with increased disease stage. The present study also failed to show an association between ALDH1B1 and cancer stage (TNM II/III vs TNM IV).

In conclusion, our present work suggests that cellular expression of ALDH1B1 (as determined by immunohistology) or the ratio of ALDH1B1 to ALDH1A1 (or ALDH2) in tissue samples could be used as a novel biomarker for identifying CRC.

ACKNOWLEDGEMENTS

This work was supported by the National Institutes of Health (NIH) Grant AA017754, AA022057. The authors thank Dr Joe Gomez, Jamie Betker, Jamie Bunker, and Dr Nicole M Payton for their technical assistance.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- Camp RL, Chung GG, Rimm DL (2002) Automated subcellular localization and quantification of protein expression in tissue microarrays. *Nat Med* **8**: 1323–1327.
- Chen Y, Orlicky DJ, Matsumoto A, Singh S, Thompson DC, Vasilou V (2011) Aldehyde dehydrogenase 1B1 (ALDH1B1) is a potential biomarker for human colon cancer. *Biochem Biophys Res Commun* **405**: 173–179.
- Chiang CP, Jao SW, Lee SP, Chen PC, Chung CC, Lee SL, Nieh S, Yin SJ (2012) Expression pattern, ethanol-metabolizing activities, and cellular localization of alcohol and aldehyde dehydrogenases in human large bowel: association of the functional polymorphisms of ADH and ALDH genes with hemorrhoids and colorectal cancer. *Alcohol* **46**: 37–49.
- Chute JP, Muramoto GG, Whitesides J, Colvin M, Safi R, Chao NJ, McDonnell DP (2006) Inhibition of aldehyde dehydrogenase and retinoid signaling induces the expansion of human hematopoietic stem cells. *Proc Natl Acad Sci USA* **103**: 11707–11712.
- Danielsen SA, Eide PW, Nesbakken A, Guren T, Leithe E, Lothe RA (2015) Portrait of the PI3K/AKT pathway in colorectal cancer. *Biochim Biophys Acta* **1855**: 104–121.
- Garber M, Grabherr MG, Guttman M, Trapnell C (2011) Computational methods for transcriptome annotation and quantification using RNA-seq. *Nat Methods* **8**: 469–477.
- Gasparetto M, Gentry T, Sebt S, O'Bryan E, Nimmanapalli R, Blaskovich MA, Bhalla K, Rizzieri D, Haaland P, Dunne J, Smith C (2004) Identification of compounds that enhance the anti-lymphoma activity of rituximab using flow cytometric high-content screening. *J Immunol Methods* **292**: 59–71.
- Ginestier C, Hur MH, Charafe-Jauffret E, Monville F, Dutcher J, Brown M, Jacquemier J, Viens P, Kleer CG, Liu S, Schott A, Hayes D, Birnbaum D, Wicha MS, Dontu G (2007) ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. *Cell Stem Cell* **1**: 555–567.
- Huang EH, Hynes MJ, Zhang T, Ginestier C, Dontu G, Appelman H, Fields JZ, Wicha MS, Boman BM (2009) Aldehyde dehydrogenase 1 is a marker for normal and malignant human colonic stem cells (SC) and tracks SC overpopulation during colon tumorigenesis. *Cancer Res* **69**: 3382–3389.
- Jackson BC, Reigan P, Miller B, Thompson DC, Vasilou V (2015) Human ALDH1B1 polymorphisms may affect the metabolism of acetaldehyde and all-trans retinaldehyde—in vitro studies and computational modeling. *Pharm Res* **32**: 1648–1662.
- Kahlert C, Gaitzsch E, Steinert G, Mogler C, Herpel E, Hoffmeister M, Jansen L, Benner A, Brenner H, Chang-Claude J, Rahbari N, Schmidt T, Klupp F, Grabe N, Lahrmann B, Koch M, Halama N, Buchler M, Weitz J (2012) Expression analysis of aldehyde dehydrogenase 1A1 (ALDH1A1) in colon and rectal cancer in association with prognosis and response to chemotherapy. *Ann Surg Oncol* **19**: 4193–4201.
- Langan RC, Mullinax JE, Ray S, Raiji MT, Schaub N, Xin HW, Koizumi T, Steinberg SM, Anderson A, Wiegand G, Butcher D, Anver M, Bilchik AJ, Stojadinovic A, Rudloff U, Avital I (2012) A pilot study assessing the potential role of non-CD133 colorectal cancer stem cells as biomarkers. *J Cancer* **3**: 231–240.
- Lassen N, Estey T, Tanguay RL, Pappa A, Reimers MJ, Vasilou V (2005) Molecular cloning, baculovirus expression, and tissue distribution of the zebrafish aldehyde dehydrogenase 2. *Drug Metab Dispos* **33**: 649–656.
- Li T, Su Y, Mei Y, Leng Q, Leng B, Liu Z, Stass SA, Jiang F (2010) ALDH1A1 is a marker for malignant prostate stem cells and predictor of prostate cancer patients' outcome. *Lab Invest* **90**: 234–244.

- Manzer R, Qamar L, Estey T, Pappa A, Petersen DR, Vasiliou V (2003) Molecular cloning and baculovirus expression of the rabbit corneal aldehyde dehydrogenase (ALDH1A1) cDNA. *DNA Cell Biol* **22**: 329–338.
- Marcato P, Dean CA, Pan D, Araslanova R, Gillis M, Joshi M, Helyer L, Pan L, Leidal A, Gujar S, Giacomantonio CA, Lee PW (2011) Aldehyde dehydrogenase activity of breast cancer stem cells is primarily due to isoform ALDH1A3 and its expression is predictive of metastasis. *Stem Cells* **29**: 32–45.
- Marchitti SA, Brocker C, Stagos D, Vasiliou V (2008) Non-P450 aldehyde oxidizing enzymes: the aldehyde dehydrogenase superfamily. *Expert Opin Drug Metab Toxicol* **4**: 697–720.
- Marchitti SA, Orlicky DJ, Brocker C, Vasiliou V (2010) Aldehyde dehydrogenase 3B1 (ALDH3B1): immunohistochemical tissue distribution and cellular-specific localization in normal and cancerous human tissues. *J Histochem Cytochem* **58**: 765–783.
- McCabe A, Dolled-Filhart M, Camp RL, Rimm DL (2005) Automated quantitative analysis (AQUA) of in situ protein expression, antibody concentration, and prognosis. *J Natl Cancer Inst* **97**: 1808–1815.
- Meng E, Mitra A, Tripathi K, Finan MA, Scalici J, McClellan S, Madeira da Silva L, Reed E, Shevde LA, Palle K, Rocconi RP (2014) ALDH1A1 maintains ovarian cancer stem cell-like properties by altered regulation of cell cycle checkpoint and DNA repair network signaling. *PLoS One* **9**: e107142.
- Nishiyama M, Nita A, Yumimoto K, Nakayama KI (2015) FBXL12-mediated degradation of ALDH3 is essential for trophoblast differentiation during placental development. *Stem Cells* **33**: 3327–3340.
- Pappa A, Estey T, Manzer R, Brown D, Vasiliou V (2003) Human aldehyde dehydrogenase 3A1 (ALDH3A1): biochemical characterization and immunohistochemical localization in the cornea. *Biochem J* **376**: 615–623.
- Patel M, Lu L, Zander DS, Sreerama L, Coco D, Moreb JS (2008) ALDH1A1 and ALDH3A1 expression in lung cancers: correlation with histologic type and potential precursors. *Lung Cancer* **59**: 340–349.
- Seidensaal K, Nollert A, Feige AH, Muller M, Fleming T, Gunkel N, Zaoui K, Grabe N, Weichert W, Weber KJ, Plinkert P, Simon C, Hess J (2015) Impaired aldehyde dehydrogenase 1 subfamily member 2A-dependent retinoic acid signaling is related with a mesenchymal-like phenotype and an unfavorable prognosis of head and neck squamous cell carcinoma. *Mol Cancer* **14**: 204.
- Singh S, Arcaroli J, Chen Y, Thompson DC, Messersmith W, Jimeno A, Vasiliou V (2015) ALDH1B1 is crucial for colon tumorigenesis by modulating Wnt/beta-catenin, notch and PI3K/Akt signaling pathways. *PLoS One* **10**: e0121648.
- Stagos D, Chen Y, Brocker C, Donald E, Jackson BC, Orlicky DJ, Thompson DC, Vasiliou V (2010) Aldehyde dehydrogenase 1B1: molecular cloning and characterization of a novel mitochondrial acetaldehyde-metabolizing enzyme. *Drug Metab Dispos* **38**: 1679–1687.
- Stewart MJ, Malek K, Crabb DW (1996) Distribution of messenger RNAs for aldehyde dehydrogenase 1, aldehyde dehydrogenase 2, and aldehyde dehydrogenase 5 in human tissues. *J Investig Med* **44**: 42–46.
- Todaro M, Francipane MG, Medema JP, Stassi G (2010) Colon cancer stem cells: promise of targeted therapy. *Gastroenterology* **138**: 2151–2162.
- Trapnell C, Pachter L, Salzberg SL (2009) TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**: 1105–1111.
- Zhou Y, Wang Y, Ju X, Lan J, Zou H, Li S, Qi Y, Jia W, Hu J, Liang W, Zhang W, Pang L, Li F (2015) Clinicopathological significance of ALDH1A1 in lung, colorectal, and breast cancers: a meta-analysis. *Biomark Med* **9**: 777–790.

This work is published under the standard license to publish agreement. After 12 months the work will become freely available and the license terms will switch to a Creative Commons Attribution-NonCommercial-Share Alike 4.0 Unported License.

Supplementary Information accompanies this paper on British Journal of Cancer website (<http://www.nature.com/bjc>)