HHS Public Access

Author manuscript

Chem Biol Interact. Author manuscript; available in PMC 2024 October 01.

Published in final edited form as:

Chem Biol Interact. 2023 October 01; 384: 110714. doi:10.1016/j.cbi.2023.110714.

Multi-omics Profiling Reveals Cellular Pathways and Functions Regulated by ALDH1B1 in Colon Cancer Cells

Yewei Wang[1], **Zeljka Popovic**[1], **Georgia Charkoftaki**[1], **Rolando Garcia-Milian**[1,2], **TuKiet T Lam**[3,4], **David C. Thompson**[5], **Ying Chen**[1,*], **Vasilis Vasiliou**[1,*]

[1]Department of Environmental Health Sciences, Yale School of Public Health, New Haven, CT

[2]Bioinformatics Support Program, Cushing/Whitney Medical Library, Yale University, New Haven, CT

[3]Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT

[4]Keck MS & Proteomics Resource, Yale School of Medicine, New Haven, CT

[5]Department of Clinical Pharmacy, University of Colorado Skaggs School of Pharmacy & Pharmaceutical Sciences, Aurora, CO

Abstract

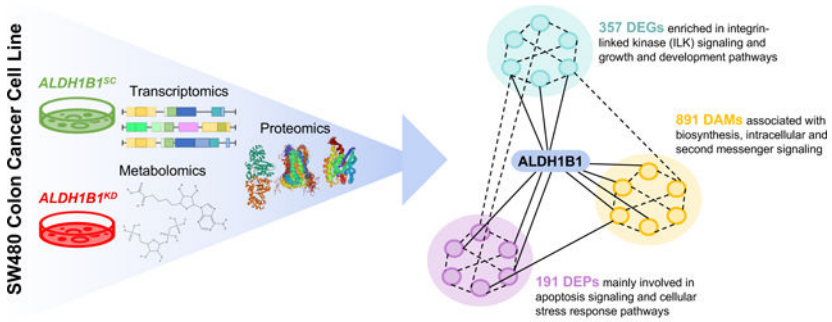
Colon cancer is the third leading cause of cancer death globally. Although early screenings and advances in treatments have reduced mortality since 1970, identification of novel targets for therapeutic intervention is needed to address tumor heterogeneity and recurrence. Previous work identified aldehyde dehydrogenase 1B1 (ALDH1B1) as a critical factor in colon tumorigenesis. To investigate further, we utilized a human colon adenocarcinoma cell line (SW480) in which the *ALDH1B1* protein expression has been knocked down by 80% via shRNA. Through multi-omics (transcriptomics, proteomics, and untargeted metabolomics) analysis, we identified the impact of *ALDH1B1* knocking down (KD) on molecular signatures in colon cancer cells. Suppression of ALDH1B1 expression resulted in 357 differentially expressed genes (DEGs), 191 differentially expressed proteins (DEPs) and 891 differentially altered metabolites (DAMs). Functional annotation and enrichment analyses revealed that: (1) DEGs were enriched in integrin-linked kinase (ILK) signaling and growth and development pathways; (2) DEPs were mainly involved in apoptosis signaling and cellular stress response pathways; and (3) DAMs were associated with biosynthesis, intercellular and second messenger signaling. Collectively, the present study provides new molecular information associated with the cellular functions of ALDH1B1, which helps to direct future investigation of colon cancer.

[*]**Corresponding authors:** Vasilis Vasiliou, Department of Environmental Health Sciences, Yale School of Public Health, 60 College Street, Rm. 511, New Haven, CT 06520-8034, Tel 203.737.8094, vasilis.vasiliou@yale.edu. Ying Chen, Department of Environmental Health Sciences, Yale School of Public Health, 60 College Street, Rm. 521, New Haven, CT 06520-8034, Tel 203.785.4694, ying.chen@yale.edu.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Graphical Abstract



Keywords

Colon Cancer; Mass Spectrometry; RNASeq; Proteomics; Metabolomics

1. Introduction

By 2030, it is expected that the global burden of colon cancer will increase by 60%, with 2.2 million new cases and 1.1 million deaths per year [1]. Due to improvements in early screening strategies and treatment, developed countries have experienced a decline in mortality rate and an increase in longevity of colon cancer patients [2,3]. Unfortunately, colon cancer incidence and death rate have steadily increased among those younger than 50 years of age for reasons unknown [1]. In addition, colon cancer patients who initially benefit from established treatments can develop resistance and relapse [4] In general, patients with colon cancer have a low 5-year survival rate (64%) compared to other cancers, such as prostate (98%) or breast (91%) [5,6]. Colon cancer is a highly heterogeneous and complex disease that involves changes in multiple biochemical pathways driven by various mutations [7–10]. The diversity of factors that promote colon cancer development is a major challenge for effectively treating colon cancer [11,12]. Colon cancer represents an important public health concern and while treatments have improved, they remain sub-optimal relative to other cancer treatments.

ALDHs are a family of enzymes that detoxify exogenous and endogenous aldehydes [13,14]. These enzymes have been implicated in the regulation of several important physiological processes, including cellular maintenance, differentiation, and proliferation [13]. Increased expression of ALDH enzymes is observed in various cancers and utilized as a marker of cancer cells and cancer stem cells (CSCs) [15–18]. Evidence suggests that increased expression of ALDH isozymes *ALDH1A1, ALDH1A3*, and *ALDH1B1* promote tumorigenesis and therapeutic resistance in colon cancers [19–24]. As such, ALDH represents an interesting new therapeutic target for the treatment of colon cancer. Studies from our group have indicated that ALDH1B1 is the sole ALDH isozyme consistently overexpressed in colorectal cancer tissues at both mRNA and protein levels, making it a potential biomarker of colon cancer [25]. Overexpression of *ALDH1B1* in human colorectal adenocarcinoma (HT29) cells is associated with alterations in cell morphology and cell cycle, increased chemoresistance and migratory potential [26]. Suppression of

ALDH1B1 expression in other human colorectal adenocarcinoma (SW480) cells inhibited spheroid formation *in vitro* and xenograft tumor growth *in vivo*, and downregulated cellular Wnt/β-catenin, Notch and PI3K/Akt-signaling pathways [23]. *ALDH1B1* promoter activity is observed in SW480 colon cancer cells that exhibit the high expression of *ALDH1B1* mRNA. Observations from human colon cancer patient-explants and numerous colon cancer cell lines both show increased promoter activity [23]. Collectively, these results support a functional role of ALDH1B1 in colon tumorigenesis. The molecular details of the role ALDH1B1 plays, however, remain to be elucidated.

The present study used a comprehensive multi-omics (transcriptomics, proteomics, and untargeted metabolomics) approach coupled with canonical pathway analysis to identify cellular pathways influenced by ALDH1B1 by suppressing *ALDH1B1* in SW480 cells, a colon cancer cell line that exhibits high constitutive expression of ALDH1B1 [21]. Our multi-omics cellular profiling identified cellular signatures of ALDH1B1 functioning in colon cancer cells. With a multi-omics approach, we were able to identify several pathways, genes, and proteins with known associations with colon cancer as well as some that have yet to be implicated.

2. Materials and Methods

2.1 Cell Culturing and Sample Preparation

Scramble ($ALDH1B1^{SC}$) and ALDH1B1 shRNA-transfected ($ALDH1B1^{KD}$) SW480 cell lines were generated and reported previously [23]. Cells at passage 4 (P4) were grown in a 150 mm Corning dish (Corning, NY, USA) in RPMI-1640 medium (Life Technologies, Carlsbad, CA, USA) containing 10% fetal bovine serum, 1% Antibiotic-Antimycotic (Gibco, Thermo Fisher Scientific Inc., Waltham, MA, USA), and 20 μg/mL puromycin for constant selection. P4 cells were grown to ~80% confluency and split at a 1:3 ratio to P5 cells and sequentially to P6 cells. The atmosphere was humidified with 5% CO2 and maintained at 37 °C. Cell medium containing puromycin was replaced every two days until cell harvesting. At ~ 90% confluency, one aliquot of P6 cells from each P5 dish at splitting was harvested, flash frozen in liquid nitrogen and stored in −80 °C for western immunoblotting analysis of ALDH1B1, as shown in Fig 1A. Briefly, one set of 3 × 150 mm dishes of P6 cells derived from the same dish of P5 cells were harvested using cell scrapers and the resultant cell pellets were flash frozen in liquid nitrogen and stored in −80 °C for metabolomics. The remaining two sets of P6 cells in 150 mm dishes were harvested, respectively, using 0.25% trypsin-EDTA (Gibco, Thermo Fisher Scientific Inc., Waltham, MA, USA) and brought to a final concentration of 10^6 cells/ml in cell medium. Cell pellets containing 10^6 cells from each dish were collected, flash frozen in liquid nitrogen and stored in −80 °C for proteomics and RNASeq analyses. All cell suspensions were centrifuged at 2,000 RPM for 5 mins at 4 °C to collect cell pellets.

2.2 Western Immunoblotting (WB)

WB was conducted exactly as described previously to measure ALDH1B1 expression levels in SW480 cells [23]. Briefly, frozen cell pellets were resuspended and lysed in RIPA buffer (150 mM NaCl, 1% TritonX-100, 0.25% sodium deoxycholate, 0.1% SDS, 50 mM Tris, 1

mM EDTA, 1 mM PMSF, protease inhibitor cocktail, pH 7.4), followed by centrifugation at 11,000 RPM for 10 mins at 4 °C. The supernatants were collected and measured for total protein quantification using the Pierce BCA Protein Assay Kit (Thermo Fisher Scientific Inc., Waltham, MA, USA) according to the manufacturer's protocol. Cell lysates containing 50 μg total proteins per sample and 2 ug of recombinant human ALDH1B1 protein were used for SDS-PAGE. Primary antibody was freshly made using an in-house rabbit polyclonal anti-ALDH1B1 antibody at a dilution of 1:5000. Expression of GAPDH protein was measured for protein loading control using anti-GAPDH antibody (Abcam, USA, # ab9485) at a dilution of 1:2500 [27]. Horseradish peroxide-conjugated goat anti-rabbit secondary antibody (Sigma, Burlington, MA) were used at a dilution of 1:5000. Quantitation of protein band density was conducted using NIH ImageJ software (version 1.52; NIH). The protein level of ALDH1B1 was normalized by the protein level of GAPDH for each sample. Relative ALDH1B1 expression in $ALDH1B1^{KD}$ cells was reported as the % of that in $ALDH1B1^{SC}$ cells.

2.3 RNASeq

Triplicate samples per cell line were processed for RNASeq analysis by the Yale Center for Genome Analyses previously described [28]. mRNA (about 500 ng) was isolated using oligo-dT beads and sheared by heating at 94 °C. The second strand synthesis was conducted using dUTP to generate strand-specific sequencing libraries. The cDNA library was then end-repaired, A-tailed adapters were ligated, and second-strand digestion was performed using uracil-DNA-glycosylase. The indexed libraries that met appropriate cut-offs were quantified and insert size distribution was determined. The samples were normalized and loaded onto Illumina flow cells at a concentration that yielded 150–250 million passing filter clusters per lane. Samples were sequenced using 75 bp single or paired-end sequencing on an Illumina HiSeq 2000 (Illumina Inc., San Diego, CA, USA), and the 6-bp index was read during an additional sequencing read. Genome peak alignment, quality assurance, quality control, expression quantification and annotation were carried out for RNAseq analysis using the Partek® Flow platform (Partek, St. Louis, MO, USA). Raw fastq files were trimmed (quality score >20 and read length >25) and aligned to the reference genome list (human genome hg38) following the Bowtie algorithm [29]. The reads per million (RPM) values of genes were normalized and used for downstream analyses. RNASeq data was deposited in NCBI's Gene Expression Omnibus and is accessible through GEO Series accession number GSE231706.

2.4 Label-free Quantitative (LFQ) Proteomics

LFQ proteomic analysis was performed by the Keck MS and Proteomics Resource at the Yale School of Medicine as previously described [30]. Briefly, cell pellets were reconstituted and then centrifuged at 14,000 RPM for 10 min at 4 °C. 100 μL of the supernatant was mixed with chloroform: methanol: water (1:4:3, v/v/v), followed by another centrifugation at 14,000 RPM for 1 minute at 4 °C. The supernatant was removed, and the pellets were reconstituted in 400 mL of methanol for an additional wash. The samples were centrifuged under the same conditions once again and then dried down under speed vac. Protein pellets were then resuspended in 25 μL RapiGest™ SF surfactant solution (Waters Corporation, Milford, MA, USA), 50 mM NH_4HCO_3, 5 μL dithiothreitol (45 mM) and 5

μL iodoacetamide (100 mM) solution. The mixture was then digested by dual enzymatic LysC and trypsin at room temperature overnight, followed by quenching with 0.1% formic acid (FA). A 2 μL aliquot of the digest solution was mixed with 3.2 μL Pierce Retention Time Calibration (Thermo Fisher Scientific Inc., Waltham, MA, USA, Cat#88321). Liquid chromatography (LC) separation was performed on a nanoACQUITY (Waters Corporation, Milford, MA, USA) system. Samples were injected in a randomized order of 5 μL injections followed by three blank solvent injections. The LC system was connected on-line to an Orbitrap Fusion mass spectrometer (Thermo Fisher Scientific Inc., Waltham, MA, USA). Intact mass spectra (MS) were collected in positive profile mode at resolution 120,000 @ m/z 400. Fragment spectra (MS/MS) were collected in positive centroid mode at a resolution of 60,000 @ m/z 400. Progenesis QI Proteomics software (Version 4.2, Nonlinear Dynamics, Waters Corporation, Milford, MA, USA) was used to process raw data dependent LC-MS/MS acquired proteomics data. MS spectral data were aligned based on peptide retention time prior to peak picking. Picked MS scan features were then filtered using MS/MS fragmentation scans to exclude MS/MS scans with spectral ranking value of 10 or higher. This was to ensure that the highest quality (top 10) MS/MS spectral data are utilized for subsequent protein database search and peptide assignments. The top 10 MS/MS fragments for each unique features from each comparison set were then exported as an ".mgf" (Mascot generic file) file for Mascot database searching [31,32]. An .xml file of the Mascot search result was created, and then imported into the Progenesis QI software, where search hits (peptides ID) were assigned to corresponding features, and protein quantitation were then calculated by the using a non-conflicting peptide assignment feature in the Progenesis QI software. Proteins with two or more unique quantifiable peptides were filtered and used in downstream analyses. The MS proteomics data have been deposited in the ProteomeXchange Consortium *via* the PRIDE partner repository with the identified number PXD011197.

2.5 Untargeted Metabolomics

Triplicate samples per cell line were processed in house for untargeted metabolomics analyses following an established protocol [28]. Briefly, frozen cell pellets were resuspended in 1 ml ice-cold acetonitrile:methanol:water (2:2:1 %v/v/v). Cells were then lysed through three snap-freeze and thaw cycles, followed by centrifugation at 15,000 RPM for 10 min at 4 °C. The supernatants containing soluble metabolites were transferred to clean Eppendorf tubes and evaporated in a vacuum concentrator (SPD111V Savant, Thermo Fisher Scientific, Waltham, MA, USA). The remaining pellets were processed for total protein quantification using the Pierce BCA Protein Assay Kit (Thermo Fisher Scientific Inc., Waltham, MA, USA). Each dried metabolite sample was then resuspended in appropriate volume of isopropanol/acetonitrile/water (40%:40%:20%) mixture to make it equivalent to 5 mg protein/mL, and 100 uL of this final solution was spiked with 2 μL TruQuant Yeast Extract Semi-targeted QC Workflow Kit (IROA Technologies, Ann Arbor, Michigan, USA) as a reference standard and used for metabolomics analysis. A pooled sample was prepared using 10 μL aliquots from each sample. All samples were stored at −80 °C. Chromatographic separation was performed as in our previous work, using an Acquity UPLC BEH C18 column (particle size, 1.7 μm; 50 mm [length] × 2.1 mm [i.d.]) (Waters Corporation) for RPLC-MS [28]. An Acquity BEH Amide column (particle size, 1.7 μm; 100 mm [length] ×

2.1 mm [i.d.]) (Waters Corporation, Milford, Massachusetts) was used for chromatographic separation for HILIC-MS. The mobile phase for RPLC-MS analysis consisted of A (0.1% formic acid in water) and B (0.1% formic acid in acetonitrile) delivered at a flow rate of 0.5 ml/min with column temperature at 35¼. The linear gradient elution started at 1% B (0–1 min), 1%–100% B (1–8 min), 100% B (8–10 min), 100%–1% B (10.0–10.1 min) and continued at 1% B (10.1–12.0 min). The mobile phase for HILIC-MS analysis consisted of A (25 mM ammonium hydroxide and 25 mM ammonium acetate in water) and B (acetonitrile) delivered at a flow rate of 0.5 ml/min. The linear gradient elution started at 95% B (0–0.5 min), 95%–65% B (0.5–7 min), 65%–40% B (7–8 min), 40% B (8–9 min), 40%–95% B (9–9.1 min) and continuing at 95% B (9.1–12.0 min) with column temperature at 25¼. The injection volume for all samples and standard solutions was 3 μL. Positive and negative mode electrospray ionization was performed on-line with a Waters quadrupole time-of-flight mass spectrometry (QTOF-MS, Milford, MA, USA). The pooled sample was injected after every five sample injections for quality control. ProteoWizard MSConvert (Version 3.0.6150, Palo Alto, CA, USA) was used to convert raw MS data into .mzML files. The data was filtered by removing peaks with a missing rate greater than 30%. Truncated peaks were then normalized, and missing values were imputed using the k-nearest neighbor method with k = 5. A generalized log transformation was applied to the data. Principal component analysis (PCA) was used to identify variables and the Wilcoxon rank sum test, combined with FDR correction (α= 5%), was used for univariate analysis across groups. The Mummichog algorithm was applied for metabolite annotation based on accurate mass m/z, to predict pathways based on the human metabolic network. All parameters and thresholds, and analysis steps were repeated as described previously [28]. Metabolomics data are available at the NIH Metabolomics workbench (DataTrack ID 3910, http://www.metabolomicsworkbench.org/).

2.6 Statistical and bioinformatics analyses

Statistical analyses and heatmap visualization were performed using the Qlucore Omics Explorer platform (Qlucore AB, Lund, Sweden). The differential expression of genes (DEGs) or proteins (DEPs) and differentially altered metabolites (DAMs) in $ALDH1B1^{KD}$ relative to $ALDH1B1^{SC}$ cells was expressed as fold-change (FC), which was calculated based on the $ALDH1B1^{KD}/ALDH1B1^{SC}$ ratio of normalized TPMs for genes or of normalized peptides abundance for proteins or normalized metabolite peaks. Specifically, if the ratio value is > 1, fold change equals to the ratio; if the ratio is between 0 and 1, fold change equals to −1*(1/ratio) [30]. Therefore, the FCs of up-regulated genes, proteins or metabolites are positive numbers, whereas the FCs of down-regulated genes, proteins or metabolites are negative numbers. The cut-off criteria were FC ≥ ± 2 for DEGs, FC ≥ ± 1.5 for DEPs and DAMs, and a Benjamini-Hochberg False Discovery Rate (FDR) adjusted p-value < 0.05 by a two-tailed Student’s unpaired t-test.

Bioinformatic analyses of DEGs, DEPs and DAMs were performed by the Ingenuity Pathway Analysis software (IPA, version 01–20-04, Qiagen, CA). Individual omics data lists were imported into IPA and z-scores were automatically calculated based on our omics dataset and the IPA knowledge database. A z-score reflects the likelihood that a particular biological pathway is affected in our experiments, such that the magnitude of the z-score

represents the significance and strength of the pathway activation with a positive z-score or pathway inhibition with a negative z-score [33]. Fisher's exact test and Benjamin-Hochberg (B-H) FDR were used to determine the levels of significance of generated pathways.

3. Results

3.1 Quantification of ALDH Genes and Proteins Differentially Expressed

Western blot analyses revealed an approximately 80% reduction in ALDH1B1 expression in *ALDH1B1*^{KD} cells relative to *ALDH1B1*^{SC} control cells (Fig. 1B and 1C). In SC and KD cells, seventeen *ALDH* genes were detected by RNASeq and ranked based on mRNA abundance in control SC cells (Fig. 2A). A total of ten ALDH proteins were detected by LFQ proteomics in both SC and KD cells, corresponding to the top ten ranked in mRNA abundance (Fig. 2B). Of these, seven of the seventeen *ALDH* genes and three of the ten ALDH proteins were expressed differently in the two cell populations. As expected, the largest differences occurred for ALDH1B1 where the gene and protein were suppressed by 50% and 70% in KD cells (Figs. 2A, B). Decrease in gene expression in KD cells was observed for *ALDH18A1, ALDH9A1, ALDH6A1, and ALDH4A1* (Fig. 2A). By contrast, there were no differences in the protein expression for any of these isozymes except for ALDH18A1 which was increased in KD cells (Fig. 2B). Although increases in gene expression occurred for *ALDH3B1* and *ALDH3A1* in KD cells (Fig. 2A), the corresponding proteins were not detected in either SC or KC cells (Fig. 2B). While *ALDH2* gene expression was no different in KD cells than in SC cells, the levels of ALDH2 protein were higher in KD cells (Figs. 2A, B).

3.2 Multi-omics Profiling Overview of ALDH1B1^{KD} cells versus ALDH1B1^{SC} cells

The -omics data sets were analyzed individually, and each data set was curated to identify differences in genes, proteins, metabolites, and pathways that distinguished *ALDH1B1*^{KD} versus *ALDH1B1*^{SC} cells. A total of 22,233 transcripts, 3,772 proteins and 3,753 metabolomic features were detected. Among these, 357 DEGs (FC ≥ ± 2, FDR $q < 0.05$), 191 DEPs (FC ≥ ± 1.5, FDR $q < 0.05$) and 891 DAMs (FC ≥ ± 1.5, FDR $q < 0.05$) were identified as being significantly different between the two cell lines. The heatmaps in Figs. 3A and 4A show the relative differences in mRNA abundance of DEGs and protein abundance of DEPs, respectively.

3.3 Canonical Pathways and Cellular Functions Enriched by DEGs, DEPs and DAMs

To understand the potential functional impact of the identified 357 DEGs, 191 DEPs and 891 DAMs, IPA canonical pathway analyses were performed. IPA provides a platform that allows us to do pathway analysis on all three -omics data sets. IPA performed several analyses of the data, but we investigated the direct pathways identified (Tables S1, S2, and S3), upstream regulators, and casual network analysis [33]. Figs. 3A, 4A, 5A bubble charts graph functional class categories (y-axis) of identified canonical pathways (bubbles) based on –log(p-values). In Figs. 3B, 4B and 5B, 10 canonical pathways identified in the analysis were extracted due to their strong connection to colon cancer in previous studies, but all pathways were investigated.

IPA casual network analysis goes beyond the data's initial pathway analysis by accounting for regulator molecules not directly connected to the targets initially identified. Upstream regulators analysis is performed by IPA and predicts behavior of upstream (pathway) molecules that may cause observed expression changes. One potential upstream regulator, beta-catenin (CTNNB1), was identified as being downregulated (activation Z-score = −2.795) based on observed gene activity. A network of 33 genes, identified in this study, that are potentially regulated by CTNNB1 are shown in Fig. 6. Lastly, IPA matched eighteen proteins to the corresponding genes from the DEGs and DEPs identifications. There was a consistent pattern of alterations in gene and protein expression of the overlapping entries, shown in Fig. 7 with calculated fold-changes.

4. Discussion

Through our multi-omics method, we were able to study and comprehensively characterize molecular changes in colon cancer SW480 cells caused by the knock-down of the *ALDH1B1* gene for the first time. Multi-omics (transcriptomic, proteomic, and metabolomic) profiles were compared between SW480 control cells and ALDH1B1 knockdown cells. In the knockdown cells, the ALDH1B1 protein expression was suppressed by 80% which is consistent with previous work by Singh et al. [23]. We identified 357 DEGs, 191 DEPs, and 891 DAMs and linked them to their respective canonical pathways. Functional annotation and enrichment analyses revealed up- and down-regulated pathways at the gene, protein, and metabolite levels, which is a key advantage of our comprehensive multi-omics approach.

RNAseq profiling and proteomics analysis allows for unbiased determination of ALDH isozymes expression in $ALDH1B1^{KD}$ and $ALDH1B1^{SC}$ SW480 cells. ALDH2 is the closest family member to ALDH1B1, sharing 72 % homology in protein sequence and having overlapped substrate profiles [34]. As such, we speculate that ALDH2 upregulation may compensate for ALDH1B1 suppression in SW480 cells. Changes in regulation of ALDH2 expression and activity has been recorded in cancers [35]. Due to the known non-selective nature of *ALDH2* expression in cancers, it is not of particular interest for more narrowed studies on colon cancer [36]. The downregulation of ALDH18A1 mRNA and upregulation in protein expression observed in $ALDH1B1^{KD}$ cells is an intriguing finding. The *ALDH18A1* has been shown to be upregulated in hepatocellular carcinoma tumors and circulating tumor cells of prostate cancer, but further studies are needed to better understand its role [37]. Changes in the expressions of other *ALDH* genes due to ALDH1B1 suppression caused either no alterations in protein levels (*ADLH9A1, ALDH6A1,* and *ADLH4A1*) or non-detectable proteins (*ADLH3B1* and *ALDH3A1*); therefore, it is unlikely they have functional significance.

Previous work done in our group by Singh et. al. took a more targeted approach and focused on specific genes and proteins associated with three pathways, Wnt/β-catenin, Notch and PI3K/Akt [23]. Four genes, *AXIN2, CTNNB1, cMyc,* and *LGR5* were measured by RT-PCR with *AXIN2* being the only gene upregulated in $ALDH1B1^{KD}$ cells. Additionally, WB analysis was performed on several proteins (active B-catenin, LEF1, C-Myc, JAG1, c-Notch-1, FABP5, PPAR-delta, PI3Kp85, Akt, and MMP2) involved in these pathways,

all of which showed to be downregulated in *ALDH1B1*^{KD} cells. Our study did not target any of these genes or proteins specifically, but through IPA we gained insight into their predicted behavior. IPA upstream regulators prediction data from the RNASeq analysis gave many useful insights. The analysis predicted CTNNB1 (Z-score: −2.795), LEF-1 (−1.964), JAG-1 (−0.991), PPAR-delta (−0.552), PI3Kp85 (PI3k (family); −1.348), and Akt (−0.813) are all downregulated as expected. MMP2 was identified in the RNASeq IPA upstream casual networks and predicted to be downregulated as well (−0.745). Additionally, AXIN2 was predicted to be downregulated (−1.000) from the IPA upstream casual networks data in the proteomics dataset. C-Myc (MYC) activity was observed and/or predicted in all three datasets but varied, making any conclusions difficult to determine. We were unable to find any supporting data for LGR5, c-Notch-1 and FABP5 activity to make any predictions.

The IPA pathway predictions based on raw data identifications gave results opposite to that of what was expected. Specifically, the Wnt/β-catenin pathway was predicted to be upregulated (Z-score = 2.000) based on the identification of 5 genes within the pathway, shown in Fig 3B. Of these, four are downregulated and predicted to downregulate the pathway. *CDH1* encodes E-cadherin proteins that are calcium-dependent cell-cell adhesion proteins in epithelial cells that take part in the Wnt/β-catenin pathway [38]. The *CDH1* gene has been studied extensively over decades due to its potential role in carcinogenesis [39]. Mutations and loss of function of the *CDH1* gene have been linked to an increased risk in the development of certain cancers, such as gastric, breast, and colon by increasing proliferation, invasion and/or metastasis during cancer progression [40]. Increased activity of the CDH1 would negatively correlate with the activity of *ALDH1B1* and the Wnt/β-catenin pathway, which aligns with previous SW480 *ALDH1B1*^{KD} studies and is contrary to the IPA prediction from this study [23]. The *CDH1* gene was also implicated in another pathway, the pulmonary fibrosis idiopathic signaling pathway, to be downregulated, which indicates potential misassignments or false positives. Similarly, the PI3K/Akt pathway was predicted to be upregulated (Z-score = 1) even though both the PI3K and Akt genes were predicted to be downregulated. We were unable to make any strong observations regarding the Notch pathway and its activity.

It is well-established that CTNNB1 plays a critical oncogenic role in colon cancer. The *CTNNB1* gene encodes β-catenin (CTNNB1) protein that regulates and coordinates cell-cell adhesion and gene transcription [41]. Additionally, CTNNB1 plays a critical role as a major mediator of the Wnt signaling pathway that ignites transcription of Wnt-specific genes [42]. In the current study, various regulatory activities of CTNNB1 were observed. Thirty-three DEGs were predicted to be regulated by CTNNB1 and showed an expression profile in consistency with a decreased activity of CTNNB1 in *ALDH1B1*^{KD} cells (Fig. 6). However, the protein expression of CTNNB1 by proteomics analysis showed to be upregulated by 1.9-fold in *ALDH1B1*^{KD} cells. Note that the abundance of CTNNB1 total protein does not necessarily correspond to the level of transcriptionally active CTNNB1, which is dephosphorylated form at specific serine/threonine residues [43]. Future studies are warranted to validate the level of active CTNNB1 using phosphorylation-specific immunoblotting or proteomics. Nevertheless, the RNASeq results in this study indicate that ALDH1B1 may promote CTNNB1 transcriptional activity in colon cancer cells and thereby colon tumorigenesis. This important finding may serve to direct treatment of colon cancers

featured with high ALDH1B1 expression given that some novel inhibitors of CTNNB1 have already been identified [44,45].

Downregulation of ILK signaling was predicted for the *ALDH1B1*^{KD} cells (DEGs; Z-score = −0.82) and this is supported by previous studies [46]. The ILK pathway is responsible for several cellular processes such as proliferation, survival, and invasion. ILK has been shown to be overexpressed in colon cancer and a driver of tumorigenesis of the colon [47]. The exact mechanistic role of ILK signaling in colon cancer has yet to be determined, but the gene's linked by IPA provide specific targets for future studies. Several other pathways and molecules stood out due to their significance (p-values) and known associations. Pathways involving the biosynthesis and degradation of melatonin were identified in all three datasets. This high and varying activity of melatonin associated molecules may allude to a potential influential role of melatonin and ALDH1B1 in colon cancer. Previous studies have shown that melatonin plays an anti-cancer role but has not been linked to ALDH1B1 and needs further investigation.[48] Dopamine, Rho Family GTPases, and amino acid-related pathways (I.e., Histidine Degradation III pathway) also showed varying activity across the analyses and should be further investigated for their potential connection to ALDH1B1 and role in colon cancer.

Eighteen molecules, as shown in Fig 7, were identified in both the DEGs and DEPs data. Several of the genes and proteins have been previously found in colon cancer studies as playing a potential role in disease progression. *PODXL, ANXA1,* and *MARCKS* all show to be upregulated in colon cancer tumors with increased *ALDH1B1* activity [49–51]. *NT5E*, *VIL1, CFL2, CD74, S100A3, CSRP2, CADM1, CKB,* and *DPYSL3* (also known as *CRMP4*) have all been connected to colon cancer through numerous studies but more extensive studies are needed for better comparison [52–59]. Limited experimental research has been done to investigate the role these molecules play in cancer and more specifically colon cancer. The behavior of these molecules, as observed by multi-omics, provides evidence that their role in colon cancer is potentially significant and therefore needs to be further studied. Lastly, *TXNDC12, NPC2, ARG2, EPS8L3, RAB3B,* and *FSCN1* have all been associated with cancers, with some showing to be non-specific cancer markers [60–64].

Among the identified DAMs (Fig. 5B and Tables S1), some are aldehyde metabolites, a majority of which show elevations in ALDH1B1^{KD} relative to control ALDH1B1^{SC} SW480 cells. These aldehydes represent intermediate metabolites of numerous metabolic pathways and have been implicated as endogenous substrates of different ALDH isozymes. For example, 3,4-dihydroxyphenylacetaldehyde, 3-methoxy-4-hydroxyphenylglycolaldehyde and 3,4-dihydroxymandelaldehyde are metabolites of dopamine, noradrenaline, and adrenaline degradations; metabolism of these aldehyde metabolites has been associated with the dehydrogenase activities of ALDH1A1, ALDH2 and ALDH3B1 [65]. (S)-methylmalonate semialdehyde, a metabolite of valine degradation, is the substrate of ALDH6A1 [66]. Glutamic acid gamma-semialdehyde is an intermediate metabolite of proline metabolism, which involves ALDH18A1 in proline synthesis and ALDH4A1 in proline degradation [67,68]. In the current study, the observation of increased expression of ALDH18A1 protein, along with elevations in glutamic acid gamma-semialdehyde, (S)-1-pyrroline-5-carboxylic acid and L-proline, supports enhanced proline biosynthesis

in ALDH1B1^{KD} SW480 cells. Given the proposed role of proline in tumor growth and metastasis [69], this metabolic alteration may be linked to the phenotypic change in the tumorigenicity of SW480 cells resulting from ALDH1B1 knockdown. Nevertheless, the findings from this multi-omics high-throughput investigation should be perceived as informative but not conclusive, since: (i) the catalytic activities of identified ALDH isoenzymes, including ALDH1B1, have not been examined to be related to their protein expression levels, and (ii) metabolite quantitation has not been validated via targeted metabolomics. Future studies are warranted to address these potential caveats.

5. Conclusions

Our findings suggest that ALDH1B1 plays a regulatory role in pathways involved in cellular stress, injury, growth, and proliferation. The Wnt/β-catenin pathway results provided insight into additional genes of the pathway that may play a critical role in colon cancer. CTNNB1 activity was particularly interesting and when considering other published studies, it shows itself to be a promising cancer therapeutic target. Additionally, we speculate that the molecular features mechanistically linked to ALDH1B1 could further serve as evidence that ALDH1B1 is a valuable therapeutic target in colon cancer treatments.

The multi-omics approach used in this study provides key insights into molecular features associated with ALDH1B1 and colon cancer cells. It also serves as a screening method that can identify many potential genes, proteins, metabolites, and pathway candidates for future more focused studies. Additionally, the number of genes and proteins that have already been connected to pathways in IPA is significantly greater than the number of metabolites. The hope is that with the move to multi-omics approaches, we can expand current metabolomics databases and better integrate that data with gene and protein data.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We would acknowledge the Yale Keck Foundation Biotechnology Resource Laboratory and the Yale Center for Genome Analyses for their support throughout this work. We thank Wendy Wang and Jean Kanyo for preparing the samples and for LC-MS/MS data collection, respectively.

Funding Sources

This work was supported in part by the National Institutes of Health grants R01AA021724, R24AA022057 and R01AA028859. The mass spectrometers and associated biotechnology tools at the Keck MS & Proteomics Resource at Yale School of Medicine were funded in part by the Yale School of Medicine and by the Office of The Director, National Institutes of Health (S10OD02365101A1, S10OD019967, and S10OD018034). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

[1]. Siegel RL, Miller KD, Fuchs HE, Jemal A, Cancer statistics, 2022, CA Cancer J Clin. 72 (2022) 7–33. 10.3322/caac.21708. [PubMed: 35020204]

[2]. Rawla P, Sunkara T, Barsouk A, Epidemiology of colorectal cancer: incidence, mortality, survival, and risk factors, Gastroenterology Review. 14 (2019) 89–103. 10.5114/pg.2018.81072. [PubMed: 31616522]

[3]. Chakrabarti S, Peterson CY, Sriram D, Mahipal A, Early stage colon cancer: Current treatment standards, evolving paradigms, and future directions, World J Gastrointest Oncol. 12 (2020) 808–832. 10.4251/wjgo.v12.i8.808. [PubMed: 32879661]

[4]. Mo S, Dai W, Xiang W, Li Y, Feng Y, Zhang L, Li Q, Cai G, Prognostic and predictive value of an autophagy-related signature for early relapse in stages I–III colon cancer, Carcinogenesis. 40 (2019) 861–870. 10.1093/carcin/bgz031. [PubMed: 30933267]

[5]. Hangaard Hansen C, Gögenur M, Tvilling Madsen M, Gögenur I, The effect of time from diagnosis to surgery on oncological outcomes in patients undergoing surgery for colon cancer: A systematic review, European Journal of Surgical Oncology. 44 (2018) 1479–1485. 10.1016/j.ejso.2018.06.015. [PubMed: 30251641]

[6]. Ahmed M, Colon Cancer: A Clinician's Perspective in 2019, Gastroenterology Res. 13 (2020) 1–10. 10.14740/gr1239. [PubMed: 32095167]

[7]. Balmain A, The critical roles of somatic mutations and environmental tumor-promoting agents in cancer risk, Nat Genet. 52 (2020) 1139–1143. 10.1038/s41588-020-00727-5. [PubMed: 33106632]

[8]. Pavlenko E, Cabron A-S, Arnold P, Dobert JP, Rose-John S, Zunke F, Functional Characterization of Colon Cancer-Associated Mutations in ADAM17: Modifications in the Pro-Domain Interfere with Trafficking and Maturation., Int J Mol Sci. 20 (2019). 10.3390/ijms20092198.

[9]. Zhang M, Yang D, Gold B, The Adenomatous Polyposis Coli (APC) mutation spectra in different anatomical regions of the large intestine in colorectal cancer, Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis. 810 (2018) 1–5. 10.1016/j.mrfmmm.2018.04.003. [PubMed: 29751128]

[10]. Wiesweg M, Kasper S, Worm K, Herold T, Reis H, Sara L, Metzenmacher M, Abendroth A, Darwiche K, Aigner C, Wedemeyer HH, Helfritz FA, Stuschke M, Schumacher B, Markus P, Paul A, Rahmann S, Schmid KW, Schuler M, Impact of RAS mutation subtype on clinical outcome—a cross-entity comparison of patients with advanced non-small cell lung cancer and colorectal cancer, Oncogene. 38 (2019) 2953–2966. 10.1038/s41388-018-0634-0. [PubMed: 30568222]

[11]. Jahanafrooz Z, Mosafer J, Akbari M, Hashemzaei M, Mokhtarzadeh A, Baradaran B, Colon cancer therapy by focusing on colon cancer stem cells and their tumor microenvironment, J Cell Physiol. 235 (2020) 4153–4166. 10.1002/jcp.29337. [PubMed: 31647128]

[12]. Orangio GR, The Economics of Colon Cancer, Surg Oncol Clin N Am. 27 (2018) 327–347. 10.1016/j.soc.2017.11.007. [PubMed: 29496093]

[13]. Jackson B, Brocker C, Thompson DC, Black W, Vasiliou K, Nebert DW, Vasiliou V, Update on the aldehyde dehydrogenase gene (ALDH) superfamily, Hum Genomics. 5 (2011) 283. 10.1186/1479-7364-5-4-283. [PubMed: 21712190]

[14]. Vasiliou V, Nebert DW, Analysis and update of the human aldehyde dehydrogenase (ALDH) gene family., Hum Genomics. 2 (2005) 138–43. 10.1186/1479-7364-2-2-138. [PubMed: 16004729]

[15]. Toledo-Guzmán ME, Hernández MI, Gómez-Gallegos ÁA, Ortiz-Sánchez E, ALDH as a Stem Cell Marker in Solid Tumors, Curr Stem Cell Res Ther. 14 (2019) 375–388. 10.2174/1574888X13666180810120012. [PubMed: 30095061]

[16]. Nakahata K, Uehara S, Nishikawa S, Kawatsu M, Zenitani M, Oue T, Okuyama H, Aldehyde Dehydrogenase 1 (ALDH1) Is a Potential Marker for Cancer Stem Cells in Embryonal Rhabdomyosarcoma, PLoS One. 10 (2015) e0125454. 10.1371/journal.pone.0125454.

[17]. Chen Y, Orlicky DJ, Matsumoto A, Singh S, Thompson DC, Vasiliou V, Aldehyde dehydrogenase 1B1 (ALDH1B1) is a potential biomarker for human colon cancer, Biochem Biophys Res Commun. 405 (2011) 173–179. 10.1016/j.bbrc.2011.01.002. [PubMed: 21216231]

[18]. 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, ALDH1 Is a

Marker of Normal and Malignant Human Mammary Stem Cells and a Predictor of Poor Clinical Outcome, Cell Stem Cell. 1 (2007) 555–567. 10.1016/j.stem.2007.08.014. [PubMed: 18371393]

[19]. Dehghan Harati M, Rodemann HP, Toulany M, Nanog Signaling Mediates Radioresistance in ALDH-Positive Breast Cancer Cells, Int J Mol Sci. 20 (2019) 1151. 10.3390/ijms20051151. [PubMed: 30845764]

[20]. Mori Y, Yamawaki K, Ishiguro T, Yoshihara K, Ueda H, Sato A, Ohata H, Yoshida Y, Minamino T, Okamoto K, Enomoto T, ALDH-Dependent Glycolytic Activation Mediates Stemness and Paclitaxel Resistance in Patient-Derived Spheroid Models of Uterine Endometrial Cancer, Stem Cell Reports. 13 (2019) 730–746. 10.1016/j.stemcr.2019.08.015. [PubMed: 31564647]

[21]. Vishnubalaji R, Manikandan M, Fahad M, Hamam R, Alfayez M, Kassem M, Aldahmash A, Alajez NM, Molecular profiling of ALDH1+ colorectal cancer stem cells reveals preferential activation of MAPK, FAK, and oxidative stress pro-survival signalling pathways, Oncotarget. 9 (2018) 13551–13564. 10.18632/oncotarget.24420. [PubMed: 29568377]

[22]. 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, Büchler M, Weitz J, Expression Analysis of Aldehyde Dehydrogenase 1A1 (ALDH1A1) in Colon and Rectal Cancer in Association with Prognosis and Response to Chemotherapy, Ann Surg Oncol. 19 (2012) 4193–4201. 10.1245/s10434-012-2518-9. [PubMed: 22878609]

[23]. Singh S, Arcaroli J, Chen Y, Thompson DC, Messersmith W, Jimeno A, Vasiliou V, ALDH1B1 Is Crucial for Colon Tumorigenesis by Modulating Wnt/β-Catenin, Notch and PI3K/Akt Signaling Pathways, PLoS One. 10 (2015) e0121648. 10.1371/journal.pone.0121648.

[24]. Russo J, Barnes A, Berger K, Desgrosellier J, Henderson J, Kanters A, Merkov L, 4-(N,N-dipropylamino)benzaldehyde inhibits the oxidation of all-trans retinal to all-trans retinoic acid by ALDH1A1, but not the differentiation of HL-60 promyelocytic leukemia cells exposed to all-trans retinal., BMC Pharmacol. 2 (2002) 4. 10.1186/1471-2210-2-4. [PubMed: 11872149]

[25]. Matsumoto A, Arcaroli J, Chen Y, Gasparetto M, Neumeister V, Thompson DC, Singh S, Smith C, Messersmith W, Vasiliou V, Aldehyde dehydrogenase 1B1: a novel immunohistological marker for colorectal cancer, Br J Cancer. 117 (2017) 1537–1543. 10.1038/bjc.2017.304. [PubMed: 28881356]

[26]. Tsochantaridis I, Roupas A, Voulgaridou G-P, Giatromanolaki A, Koukourakis MI, Panayiotidis MI, Pappa A, Aldehyde Dehydrogenase 1B1 Is Associated with Altered Cell Morphology, Proliferation, Migration and Chemosensitivity in Human Colorectal Adenocarcinoma Cells, Biomedicines. 9 (2021) 44. 10.3390/biomedicines9010044. [PubMed: 33419031]

[27]. Stagos D, Chen Y, Brocker C, Donald E, Jackson BC, Orlicky DJ, Thompson DC, Vasiliou V, Aldehyde Dehydrogenase 1B1: Molecular Cloning and Characterization of a Novel Mitochondrial Acetaldehyde-Metabolizing Enzyme, Drug Metabolism and Disposition. 38 (2010) 1679–1687. 10.1124/dmd.110.034678. [PubMed: 20616185]

[28]. Charkoftaki G, Thompson DC, Golla JP, Garcia-Milian R, Lam TT, Engel J, Vasiliou V, Integrated multi-omics approach reveals a role of ALDH1A1 in lipid metabolism in human colon cancer cells, Chem Biol Interact. 304 (2019) 88–96. 10.1016/j.cbi.2019.02.030. [PubMed: 30851239]

[29]. Langmead B, Trapnell C, Pop M, Salzberg SL, Ultrafast and memory-efficient alignment of short DNA sequences to the human genome, Genome Biol. 10 (2009) R25. 10.1186/gb-2009-10-3-r25. [PubMed: 19261174]

[30]. Wang Y, Chen Y, Garcia-Milian R, Golla JP, Charkoftaki G, Lam TT, Thompson DC, Vasiliou V, Proteomic profiling reveals an association between ALDH and oxidative phosphorylation and DNA damage repair pathways in human colon adenocarcinoma stem cells, Chem Biol Interact. 368 (2022) 110175. 10.1016/j.cbi.2022.110175.

[31]. Cappadona S, Levander F, Jansson M, James P, Cerutti S, Pattini L, Wavelet-Based Method for Noise Characterization and Rejection in High-Performance Liquid Chromatography Coupled to Mass Spectrometry, Anal Chem. 80 (2008) 4960–4968. 10.1021/ac800166w. [PubMed: 18510348]

[32]. Silva JC, Gorenstein MV, Li G-Z, Vissers JPC, Geromanos SJ, Absolute Quantification of Proteins by LCMSE, Molecular & Cellular Proteomics. 5 (2006) 144–156. 10.1074/mcp.M500230-MCP200. [PubMed: 16219938]

[33]. Krämer A, Green J, Pollard J, Tugendreich S, Causal analysis approaches in Ingenuity Pathway Analysis, Bioinformatics. 30 (2014) 523–530. 10.1093/bioinformatics/btt703. [PubMed: 24336805]

[34]. Jackson BC, Holmes RS, Backos DS, Reigan P, Thompson DC, Vasiliou V, Comparative genomics, molecular evolution and computational modeling of ALDH1B1 and ALDH2, Chem Biol Interact. 202 (2013) 11–21. 10.1016/j.cbi.2012.11.022. [PubMed: 23247008]

[35]. Zhang H, Fu L, The role of ALDH2 in tumorigenesis and tumor progression: Targeting ALDH2 as a potential cancer treatment, Acta Pharm Sin B. 11 (2021) 1400–1411. 10.1016/j.apsb.2021.02.008. [PubMed: 34221859]

[36]. Ma B, Liu Z, Xu H, Liu L, Huang T, Meng L, Wang L, Zhang Y, Li L, Han X, Molecular Characterization and Clinical Relevance of ALDH2 in Human Cancers, Front Med (Lausanne). 8 (2022). 10.3389/fmed.2021.832605.

[37]. Yao S, Chen W, Zuo H, Bi Z, Zhang X, Pang L, Jing Y, Yin X, Cheng H, Comprehensive Analysis of Aldehyde Dehydrogenases (ALDHs) and Its Significant Role in Hepatocellular Carcinoma, Biochem Genet. 60 (2022) 1274–1297. 10.1007/s10528-021-10178-0. [PubMed: 34928471]

[38]. Prasad CP, Mirza S, Sharma G, Prashad R, DattaGupta S, Rath G, Ralhan R, Epigenetic alterations of CDH1 and APC genes: Relationship with activation of Wnt/β-catenin Pathway in invasive ductal carcinoma of breast, Life Sci. 83 (2008) 318–325. 10.1016/j.lfs.2008.06.019. [PubMed: 18662704]

[39]. Luo W, Fedda F, Lynch P, Tan D, CDH1 Gene and Hereditary Diffuse Gastric Cancer Syndrome: Molecular and Histological Alterations and Implications for Diagnosis And Treatment., Front Pharmacol. 9 (2018) 1421. 10.3389/fphar.2018.01421. [PubMed: 30568591]

[40]. Semb H, Christofori G, The Tumor-Suppressor Function of E-Cadherin, The American Journal of Human Genetics. 63 (1998) 1588–1593. 10.1086/302173. [PubMed: 9837810]

[41]. van Schie EH, van Amerongen R, Aberrant WNT/CTNNB1 Signaling as a Therapeutic Target in Human Breast Cancer: Weighing the Evidence, Front Cell Dev Biol. 8 (2020). 10.3389/fcell.2020.00025.

[42]. Valenta T, Hausmann G, Basler K, The many faces and functions of β-catenin, EMBO J. 31 (2012) 2714–2736. 10.1038/emboj.2012.150. [PubMed: 22617422]

[43]. van der Wal T, van Amerongen R, Walking the tight wire between cell adhesion and WNT signalling: a balancing act for β-catenin., Open Biol. 10 (2020) 200267. 10.1098/rsob.200267.

[44]. Babic I, Yenugonda VM, Kesari S, Nurmemmedov E, Wnt pathway: a hallmark of drug discovery challenge, Future Med Chem. 10 (2018) 1399–1403. 10.4155/fmc-2018-0084. [PubMed: 29779411]

[45]. Cheltsov A, Nomura N, Yenugonda VM, Roper J, Mukthavaram R, Jiang P, Her N-G, Babic I, Kesari S, Nurmemmedov E, Allosteric inhibitor of β-catenin selectively targets oncogenic Wnt signaling in colon cancer, Sci Rep. 10 (2020) 8096. 10.1038/s41598-020-60784-y. [PubMed: 32415084]

[46]. Almasabi S, Ahmed AU, Boyd R, Williams BRG, A Potential Role for Integrin-Linked Kinase in Colorectal Cancer Growth and Progression via Regulating Senescence and Immunity, Front Genet. 12 (2021). 10.3389/fgene.2021.638558.

[47]. Almasabi S, Boyd R, Ahmed AU, Williams BRG, Integrin-Linked Kinase Expression Characterizes the Immunosuppressive Tumor Microenvironment in Colorectal Cancer and Regulates PD-L1 Expression and Immune Cell Cytotoxicity, Front Oncol. 12 (2022). 10.3389/fonc.2022.836005.

[48]. Iravani S, Eslami P, Dooghaie Moghadam A, Moazzami B, Mehrvar A, Hashemi MR, Mansour-Ghanaei F, Mansour-Ghanaei A, Majidzadeh-A K, The Role of Melatonin in Colorectal Cancer, J Gastrointest Cancer. 51 (2020) 748–753. 10.1007/s12029-019-00336-4. [PubMed: 31792737]

[49]. Liang Z, Li X, Identification of ANXA1 as a potential prognostic biomarker and correlating with immune infiltrates in colorectal cancer, Autoimmunity. 54 (2021) 76–87. 10.1080/08916934.2021.1887148. [PubMed: 33596760]

[50]. Rombouts K, Carloni V, Mello T, Omenetti S, Galastri S, Madiai S, Galli A, Pinzani M, Myristoylated Alanine-Rich protein Kinase C Substrate (MARCKS) expression modulates the metastatic phenotype in human and murine colon carcinoma in vitro and in vivo, Cancer Lett. 333 (2013) 244–252. 10.1016/j.canlet.2013.01.040. [PubMed: 23376641]

[51]. Le Tran N, Wang Y, Nie G, Podocalyxin in Normal Tissue and Epithelial Cancer, Cancers (Basel). 13 (2021) 2863. 10.3390/cancers13122863. [PubMed: 34201212]

[52]. Luk IY, Jenkins LJ, Schoffer KL, Ng I, Tse JWT, Mouradov D, Kaczmarczyk S, Nightingale R, Burrows AD, Anderson RL, Arango D, Dopeso H, Croft L, Richardson MF, Sieber OM, Liao Y, Mooi JK, Vukelic N, Reehorst CM, Afshar-Sterle S, Whitehall VLJ, Fennell L, Abud HE, Tebbutt NC, Phillips WA, Williams DS, Shi W, Mielke LA, Ernst M, Dhillon AS, Clemons NJ, Mariadason JM, Epithelial de-differentiation triggered by co-ordinate epigenetic inactivation of the EHF and CDX1 transcription factors drives colorectal cancer progression, Cell Death Differ. 29 (2022) 2288–2302. 10.1038/s41418-022-01016-w. [PubMed: 35606410]

[53]. Maharshak N, CD74 is a survival receptor on colon epithelial cells, World J Gastroenterol. 16 (2010) 3258. 10.3748/wjg.v16.i26.3258. [PubMed: 20614481]

[54]. LIU B, SUN W-Y, ZHI C-Y, LU T-C, GAO H-M, ZHOU J-H, YAN W-Q, GAO H-C, Role of S100A3 in human colorectal cancer and the anticancer effect of cantharidinate, Exp Ther Med. 6 (2013) 1499–1503. 10.3892/etm.2013.1344. [PubMed: 24255681]

[55]. Chen L, Long X, Duan S, Liu X, Chen J, Lan J, Liu X, Huang W, Geng J, Zhou J, CSRP2 suppresses colorectal cancer progression via p130Cas/Rac1 axis-meditated ERK, PAK, and HIPPO signaling pathways, Theranostics. 10 (2020) 11063–11079. 10.7150/thno.45674. [PubMed: 33042270]

[56]. Wang G, Fu S, Li D, Chen Y, Expression and clinical significance of serum NT5E protein in patients with colorectal cancer, Cancer Biomarkers. 24 (2019) 461–468. 10.3233/CBM-182207. [PubMed: 30932882]

[57]. Li H, Gao J, Zhang S, Functional and Clinical Characteristics of Cell Adhesion Molecule CADM1 in Cancer, Front Cell Dev Biol. 9 (2021). 10.3389/fcell.2021.714298.

[58]. Mooney SM, Rajagopalan K, Williams BH, Zeng Y, Christudass CS, Li Y, Yin B, Kulkarni P, Getzenberg RH, Creatine kinase brain overexpression protects colorectal cells from various metabolic and non-metabolic stresses, J Cell Biochem. 112 (2011) 1066–1075. 10.1002/jcb.23020. [PubMed: 21308735]

[59]. Wu C-C, Chen H-C, Chen S-J, Liu H-P, Hsieh Y-Y, Yu C-J, Tang R, Hsieh L-L, Yu J-S, Chang Y-S, Identification of collapsin response mediator protein-2 as a potential marker of colorectal carcinoma by comparative analysis of cancer cell secretomes, Proteomics. 8 (2008) 316–332. 10.1002/pmic.200700819. [PubMed: 18203259]

[60]. Liu H, Zhang Y, Li L, Cao J, Guo Y, Wu Y, Gao W, Fascin actin-bundling protein 1 in human cancer: Promising biomarker or therapeutic target?, Mol Ther Oncolytics. 20 (2021) 240–264. 10.1016/j.omto.2020.12.014. [PubMed: 33614909]

[61]. Tsunedomi R, Yoshimura K, Kimura Y, Nishiyama M, Fujiwara N, Matsukuma S, Kanekiyo S, Matsui H, Shindo Y, Watanabe Y, Tokumitsu Y, Yoshida S, Iida M, Suzuki N, Takeda S, Ioka T, Hazama S, Nagano H, Elevated expression of RAB3B plays important roles in chemoresistance and metastatic potential of hepatoma cells, BMC Cancer. 22 (2022) 260. 10.1186/s12885-022-09370-1. [PubMed: 35277124]

[62]. Li P, Hu T, Wang H, Tang Y, Ma Y, Wang X, Xu Y, Chen G, Upregulation of EPS8L3 is associated with tumorigenesis and poor prognosis in patients with liver cancer, Mol Med Rep. (2019). 10.3892/mmr.2019.10471.

[63]. Liao Y-J, Lin M-W, Yen C-H, Lin Y-T, Wang C-K, Huang S-F, Chen K-H, Yang C-P, Chen T-L, Hou M-F, Arthur Chen Y-M, Characterization of Niemann-Pick Type C2 Protein Expression in Multiple Cancers Using a Novel NPC2 Monoclonal Antibody, PLoS One. 8 (2013) e77586. 10.1371/journal.pone.0077586.

[64]. Lee E, Lee DH, Emerging roles of protein disulfide isomerase in cancer, BMB Rep. 50 (2017) 401–410. 10.5483/BMBRep.2017.50.8.107. [PubMed: 28648146]

[65]. Marchitti SA, Deitrich RA, Vasiliou V, Neurotoxicity and metabolism of the catecholamine-derived 3,4-dihydroxyphenylacetaldehyde and 3,4-dihydroxyphenylglycolaldehyde: the role of aldehyde dehydrogenase., Pharmacol Rev. 59 (2007) 125–50. 10.1124/pr.59.2.1. [PubMed: 17379813]

[66]. Shin H, Cha H-J, Lee MJ, Na K, Park D, Kim C-Y, Han DH, Kim H, Paik Y-K, Identification of ALDH6A1 as a Potential Molecular Signature in Hepatocellular Carcinoma via Quantitative Profiling of the Mitochondrial Proteome., J Proteome Res. 19 (2020) 1684–1695. 10.1021/acs.jproteome.9b00846. [PubMed: 31985234]

[67]. Vettore LA, Westbrook RL, Tennant DA, Proline metabolism and redox; maintaining a balance in health and disease., Amino Acids. 53 (2021) 1779–1788. 10.1007/s00726-021-03051-2. [PubMed: 34291343]

[68]. Püschel J, Dubrovska A, Gorodetska I, The Multifaceted Role of Aldehyde Dehydrogenases in Prostate Cancer Stem Cells., Cancers (Basel). 13 (2021). 10.3390/cancers13184703.

[69]. D'Aniello C, Patriarca EJ, Phang JM, Minchiotti G, Proline Metabolism in Tumor Growth and Metastatic Progression, Front Oncol. 10 (2020). 10.3389/fonc.2020.00776.

Highlights:

- Suppression of *ALDH1B1* in SW480 cells resulted in 357 DEGs, 191 DEPs and 891 DAMs.
- DEGs were enriched in pathways of ILK signaling and growth and development.
- DEPs were enriched in pathways of apoptosis signaling and cellular stress response.
- DAMs were associated with biosynthesis, intercellular and second messenger signaling.

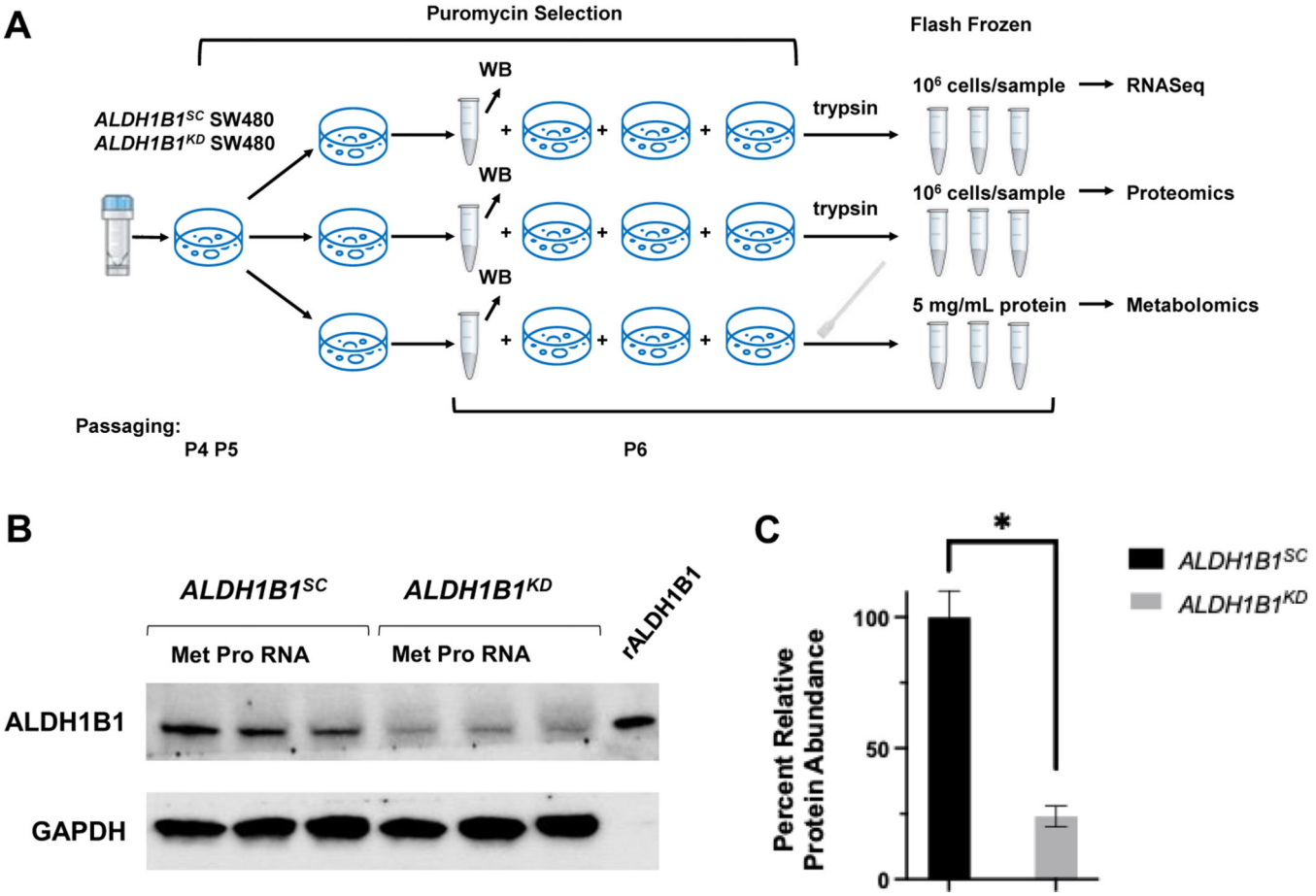


Figure 1. Multi-omics Workflow and *ADLDH1B1* Expression. (A)
Flow chart of the *ALDH1B1*^{SC} and *ALDH1B1*^{KD} SW480 sample preparation used for the multi-omics approach (metabolomics, proteomics, and transcriptomics). **(B)** Image of Western blot analysis of ALDH1B1 protein extracted from *ALDH1B1*^{SC} and *ALDH1B1*^{KD} SW480 cells. GAPDH serves as a loading control and recombinant human ALDH1B1 (rALDH1B1) was used as a positive control. The intensity of the ALDH1B1 bands were normalized to the intensity of the GAPDH bands and the control group was set as 100%. **(C)** Bar graph of the ALDH1B1 protein expression in *ALDH1B1*^{SC} and *ALDH1B1*^{KD} SW480 cells (*p<0.05).

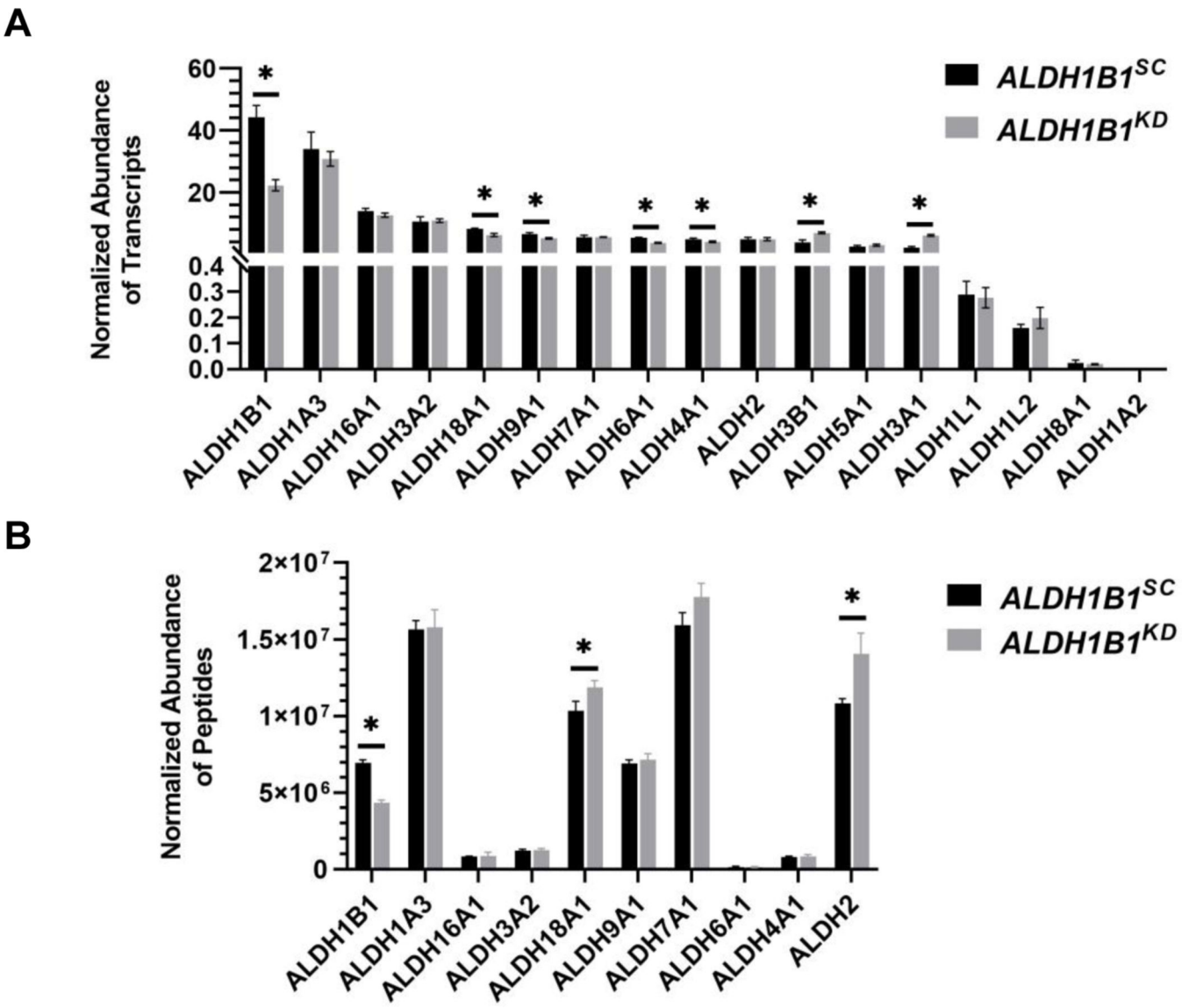
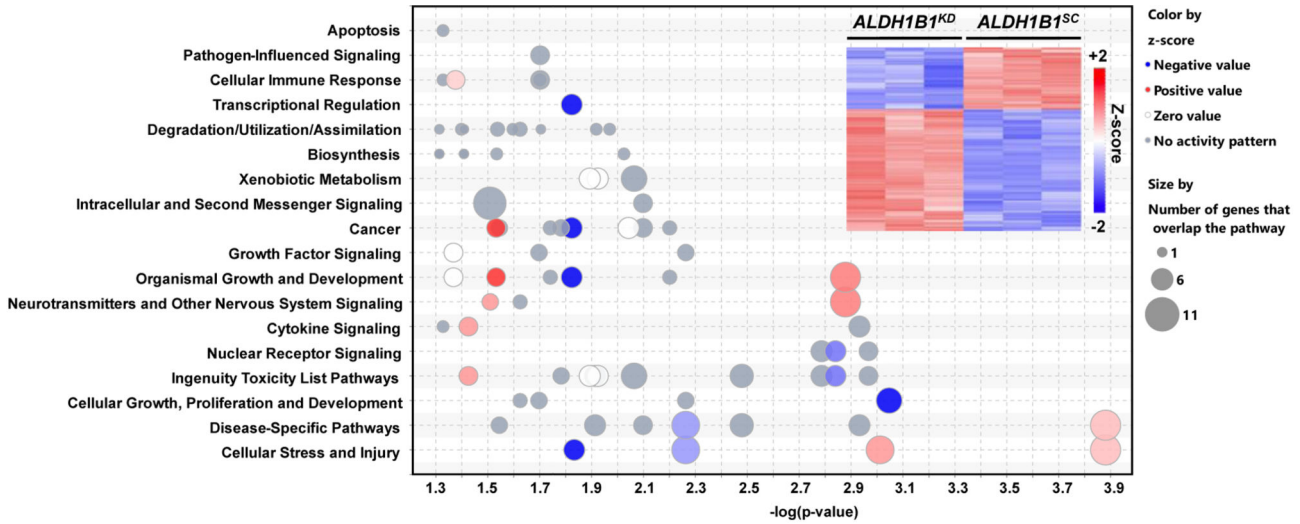


Figure 2. Expression of ALDH isozymes in *ALDH1B1*^{SC} and *ALDH1B1*^{KD} SW480 cells.

(A) Normalized abundance of transcripts of the seventeen *ALDH* genes detected by RNASeq profiling in *ALDH1B1*^{SC} (black) and *ALDH1B1*^{KD} (grey) SW480 cells. **(B)** Peptides normalized abundance of the ten ALDH isozymes detected by proteomics profiling in *ALDH1B1*^{SC} (black bar) and *ALDH1B1*^{KD} (grey bar) SW480 cells. Data are presented as the mean and associated standard deviation from triplicate measurements (*P < 0.05, unpaired student t-test) for both graphs.

A



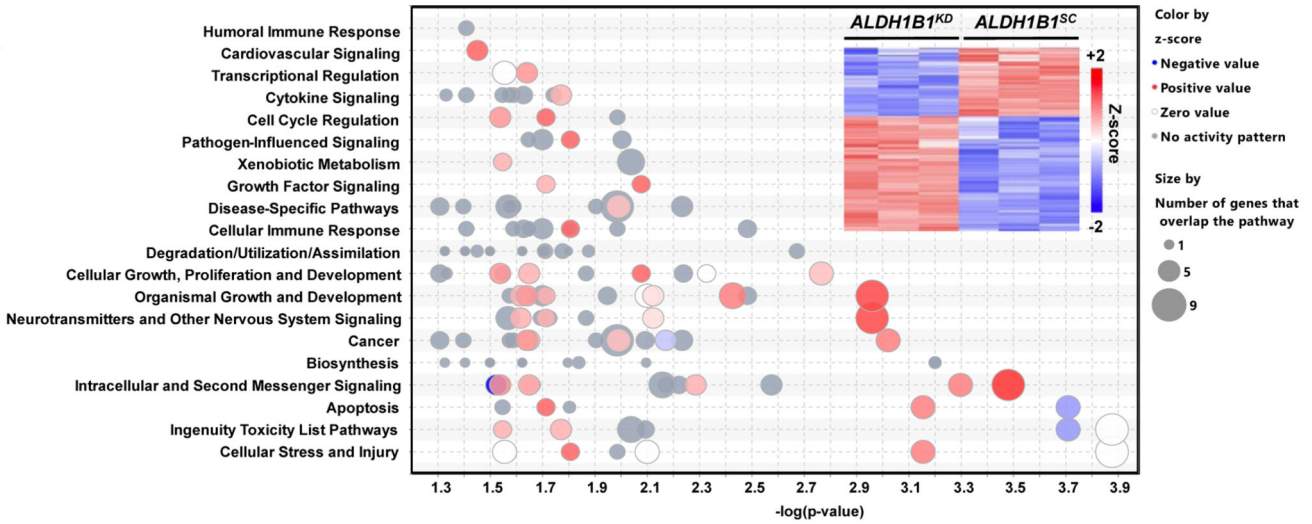
B

Ingenuity Canonical Pathways	-log(p-value)	z-score	Molecules
WNT/β-catenin Signaling	1.53	2	CDH1, DKK1, SOX14, SOX2, SOX4
Superpathway of Melatonin Degradation	1.54	N/A	CYP2U1, CYP3A5, SULT1A2
Melatonin Degradation I	1.63	N/A	CYP2U1, CYP3A5, SULT1A2
MSP-RON Signaling Pathway	1.7	N/A	KLK10, NOS2, TNF
IGF-1 Signaling	1.7	N/A	CCN1, CCN2, IGFBP3, IGFBP7
p53 Signaling	1.78	N/A	GADD45G, SERPINE2, SNAI2, THBS1
Xenobiotic Metabolism Signaling	2.06	N/A	ALDH3A1, CAMK1, CITED2, CYP3A5, NOS2, PRKCG, SULT1A2, TNF
Notch Signaling	2.2	N/A	HES1, HES7, JAG1
FXR/RXR Activation	2.79	N/A	C3, FGF19, MLXIPL, SAA1, TNF, VLDLR
ILK Signaling	3.05	-0.816	TNF, CDH1, BMP2, CFL2, FN1, NOS2, RND3, SNAI2

Figure 3. RNASeq profiling in *ALDH1B1*^{SC} and *ALDH1B1*^{KD} SW480 cells.

(A) A bubble chart of all canonical pathways identified (represented by bubbles) grouped into functional categories (y-axis) by IPA and graphed based on −log(p-value). The legend on the right, provided by IPA, explains bubble (pathway) coloring and sizing. On the top-right corner, a heat map of the relative mRNA abundance of 357 differentially expressed genes (DEGs) in *ALDH1B1*^{SC} and *ALDH1B1*^{KD} SW480 cells based on RNASeq analysis is shown. The hierarchical heat map scale of Z-scores ranges from −2 (blue) to 2 (red) with a midpoint of 0 (white) is included. The intensity of color reflects the Z-score of protein expression. **(B)** Ten pathways enriched by IPA analysis based on DEGs with calculated z-scores and −log(p-values). The font colors of the DEGs indicate the fold change, with red indicating up-regulation and blue indicating down-regulation.

A



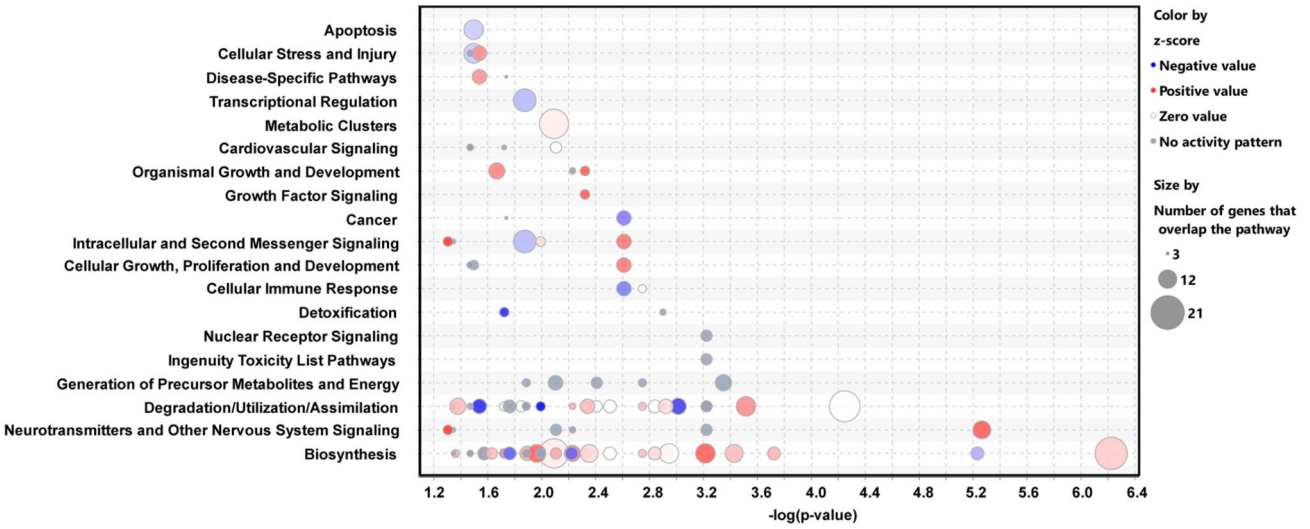
B

Ingenuity Canonical Pathways	-log(p-value)	z-score	Molecules
Signaling by Rho Family GTPases	3.48	2.646	ARHGEF16, ARHGEF7, ITGB4, ROCK2, SEPTIN10, SEPTIN2, SEPTIN8, SEPTIN9, CFL2
MSP-RON Signaling In Cancer Cells Pathway	3.02	1.633	CTNNB1, ITGB4, KLK6, RALB, RRAS, SFN
PI3K/AKT Signaling	1.65	1	CTNNB1, ITGB4, RALB, RRAS, SFN
Endocannabinoid Cancer Inhibition Pathway	2.17	-0.447	CASP2, CTNNB1, ROCK2, AKT1S1, CASP7
RHOGDI Signaling	1.52	-2	ARHGEF16, ARHGEF7, ITGB4, ROCK2, CFL2
Epithelial Adherens Junction Signaling	2.76	0.816	CTNNB1, RALB, ROCK2, RRAS, SFN, CFL2
IL-3 Signaling	1.59	N/A	RALB, RRAS, STAT6
JAK/STAT Signaling	1.55	N/A	RALB, RRAS, STAT6
Dopamine Degradation	2.67	N/A	ALDH1B1, MAOB, SULT1A1
Superpathway of Melatonin Degradation	1.77	N/A	CYP2S1, MAOB, SULT1A1

Figure 4. Proteomics profiling in ALDH1B1^{SC} and ALDH1B1^{KD} SW480 cells.

(A) A bubble chart of all canonical pathways identified (represented by bubbles) grouped into functional categories (y-axis) by IPA and graphed based on –log(p-value). The legend on the right, provided by IPA, explains bubble (pathway) coloring and sizing. On the top-right corner, a heat map of the relative protein abundance of 191 differentially expressed proteins (DEPs) in ALDH1B1^{SC} SW480 and ALDH1B1^{KD} SW480 cells based on proteomics analysis. Hierarchical clustering was constructed using Qlucore software with unsupervised hierarchical classification. The hierarchical heat map scale of Z-scores ranges from 2 (blue) to 2 (red) with a midpoint of 0 (white). The intensity of color reflects the Z-score of protein expression. **(B)** Ten pathways enriched by IPA analysis based on DEPs with calculated z-scores and −log(p-values). The font colors of the DEPs indicate the fold change, with red indicating up-regulation and blue indicating down-regulation.

A



B

Ingenuity Canonical Pathways	-log(p-value)	z-score	Molecules
Dopamine Receptor Signaling	5.27	2	(6R)-6-(L-erythro-1,2-dihydroxypropyl)-5,6,7,8-tetrahydro-4a-hydroxypterin, 3,4-dihydroxyphenylacetic acid, 3-methoxytyramine, 4-hydroxy-3-methoxyphenylacetic acid, 7,8-dihydrobiopterin, 7,8-dihydroneopterin 3'-triphosphate, cyclic AMP, dopamine, dysproptерin, GTP, L-tyrosine, levodopa
UDP-N-acetyl-D-glucosamine Biosynthesis II	3.72	0.816	acetyl-coenzyme A, coenzyme A, fructose-6-phosphate, glucosamine-6-phosphate, L-glutamic acid, L-glutamine, N-acetyl-D-glucosamine 1-phosphate, N-acetyl-D-glucosamine 6-phosphate, UDP-N-acetylglucosamine
Serotonin and Melatonin Biosynthesis	3.72	1.134	4a-hydroxytetrahydrobiopterin, 5-hydroxytryptamine, acetyl-coenzyme A, coenzyme A, L-tryptophan, melatonin, N-acetylserotonin, S-adenosylhomocysteine, S-adenosylmethionine
FXR/RXR Activation	3.22	N/A	androsterone, chenodeoxycholic acid, cholic acid, cyclic AMP, D-glucose, deoxycholate, lithocholic acid, xylitol
Histidine Degradation III	3.22	N/A	10-formyltetrahydrofolic acid, 4-imidazolone-5-propanoic acid, 5,10-methenyltetrahydrofolate, 5-formiminotetrahydrofolic acid, formiminoglutamic acid, L-glutamic acid, L-histidine, urocanic acid
Serotonin Receptor Signaling	3.22	N/A	4a-hydroxytetrahydrobiopterin, 5-hydroxyindol-3-acetic acid, 5-hydroxytryptamine, 7,8-dihydrobiopterin, 7,8-dihydroneopterin 3'-triphosphate, dysproptерin, GTP, L-tryptophan
UDP-N-acetyl-D-galactosamine Biosynthesis II	2.92	1	acetyl-coenzyme A, beta-D-glucose, beta-D-glucose 6-phosphate, coenzyme A, fructose-6-phosphate, glucosamine-6-phosphate, N-acetyl-D-glucosamine 1-phosphate, N-acetyl-D-glucosamine 6-phosphate, UDP-N-acetyl-D-galactosamine, UDP-N-acetylglucosamine
Pentose Phosphate Pathway (Non-oxidative Branch)	2.75	N/A	D-ribose-5-phosphate, D-xylulose 5-phosphate, erythrose 4-phosphate, fructose-6-phosphate, glyceraldehyde-3-phosphate, sedoheptulose 7-phosphate
AMPK Signaling	2.61	1.667	acetyl-coenzyme A, AICAR, AMP, carnitine, creatine, fructose-2,6-diphosphate, GDP, malonyl-coenzyme A, palmitoyl-coenzyme A, phosphocreatine
Melatonin Degradation I	1.99	-1.342	3',5'-ADP, 6-hydroxymelatonin, formaldehyde, melatonin, N-acetylserotonin, UDP, UDP-glucuronic acid

Figure 5. Untargeted metabolomics profiling of ALDH1B1^{SC} and ALDH1B1^{KD} SW480 cells.
(A) A bubble chart of all canonical pathways identified (represented by bubbles) grouped into functional categories (y-axis) by IPA and graphed based on −log(p-value). The legend on the right, provided by IPA, explains bubbles (pathway) coloring and sizing. **(B)** A table of ten pathways enriched by IPA analysis based on DAMs results with calculated z-scores and – log(p-values). The font colors of the DAMs indicate the fold change, with red indicating upregulation and blue indicating downregulation.

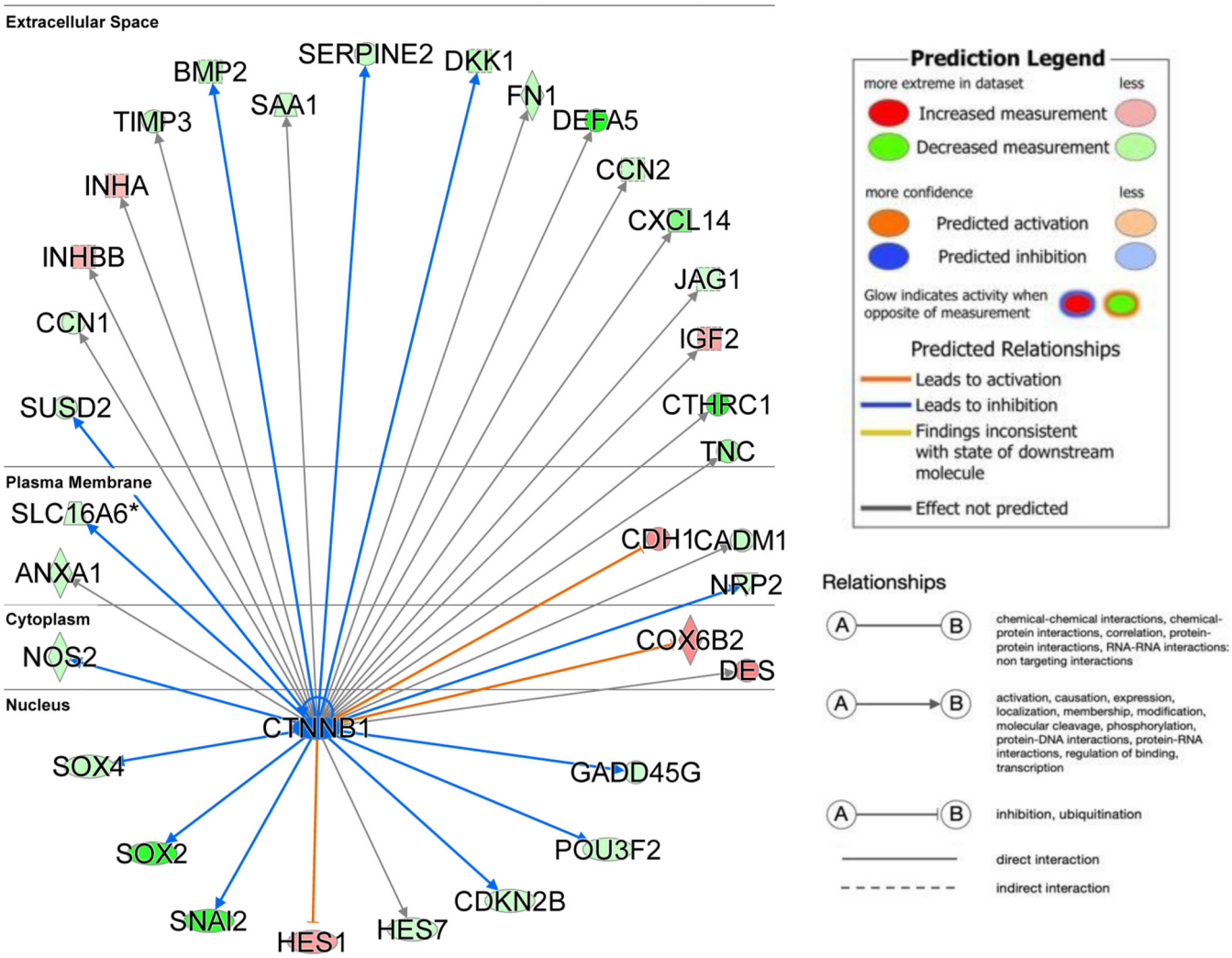


Figure 6. RNASeq Upstream Regulator: CTNNB1.

IPA results identified CTNNB1 as a key upstream transcriptional regulator of the DEGs based on the genes identified through RNASeq. IPA predicts how upstream regulators affect the activity of genes identified (predicted relationship lines/arrows) and compares it to measured data (red/green). The cellular location of gene expression is indicated on the left.

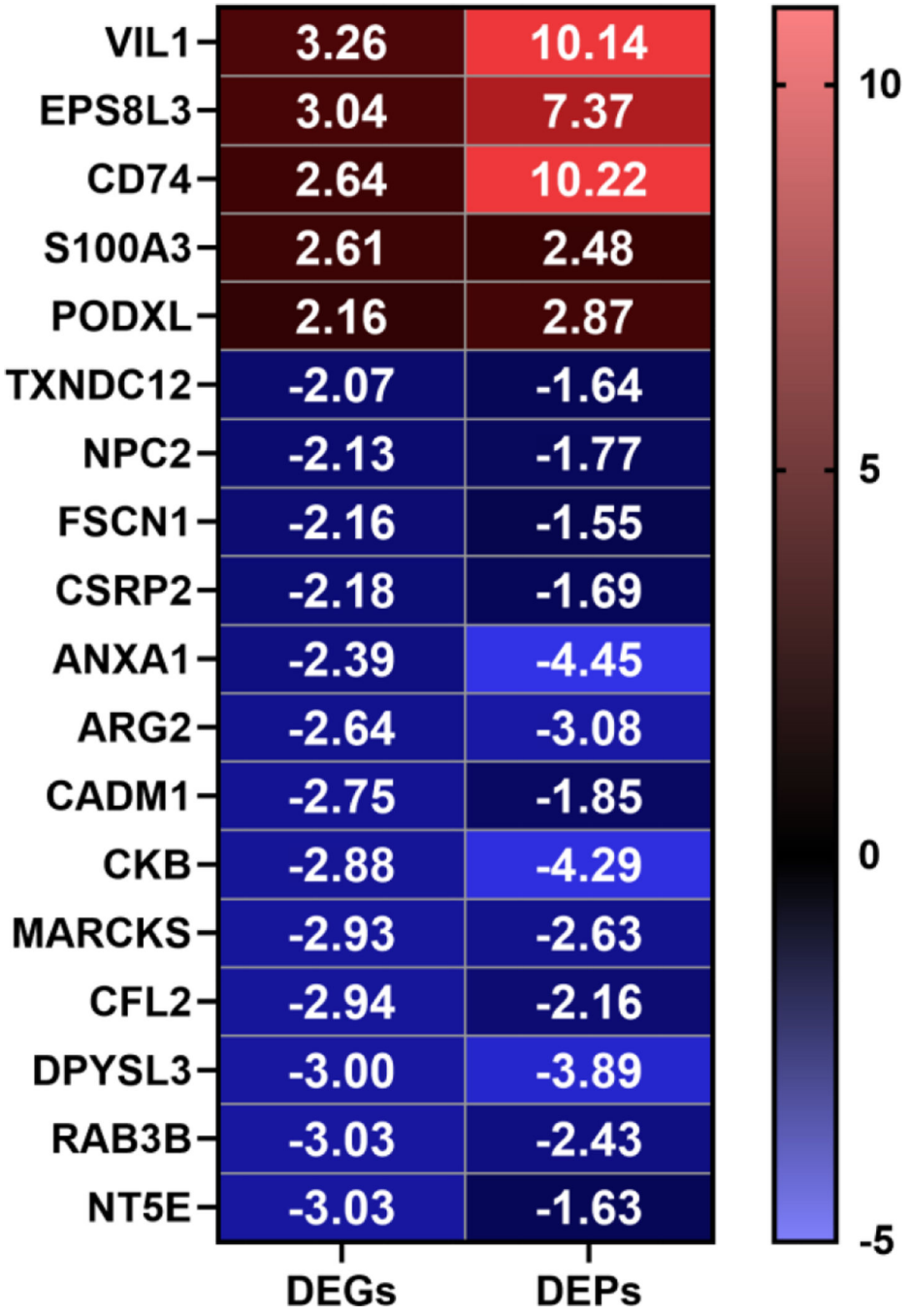


Figure 7. Overlap of DEGs and DEPs.

Of 358 DEGs and 191 DEPs identified, eighteen molecules were detected in both data sets. The heatmap shows the expression levels of shared molecules identified (left y-axis) in the two data sets. The color scale indicates the fold change of expression levels (−5 to 10). Red indicates upregulation and blue indicates downregulation.