Biochimie xxx (xxxx) xxx

Contents lists available at ScienceDirect

Biochimie

journal homepage: www.elsevier.com/locate/biochi

Low expression of CD24 is associated with poor survival in colorectal cancer

Stepan Nersisyan [a, b, 1], Ann-Kristin Ahlers [c, 1], Tobias Lange [c], Daniel Wicklein [c], Alexei Galatenko [d, e], Hanibal Bohnenberger [f], Omar Elakad [f], Lena-Christin Conradi [g], Sandra Genduso [c], Hanna Maar [c], Alina Schiecke [c], Diana Maltseva [a, b], Maria Raygorodskaya [a, b], Julia Makarova [a], Udo Schumacher [c], Alexander Tonevitsky [a, b, *]

[a] Faculty of Biology and Biotechnology, HSE University, Moscow, Russia
[b] SRC Bioclinicum, Moscow, Russia
[c] Institute of Anatomy and Experimental Morphology, University Medical Center, Hamburg-Eppendorf, Hamburg, Germany
[d] Faculty of Mechanics and Mathematics, Lomonosov Moscow State University, Moscow, Russia
[e] Moscow Center for Fundamental and Applied Mathematics, Moscow, Russia
[f] Institute of Pathology, University Medical Center, Göttingen, Germany
[g] Clinic for General, Visceral and Pediatric Surgery, University Medical Center, Göttingen, Germany

ARTICLE INFO

Article history:
Received 29 July 2021
Received in revised form
24 September 2021
Accepted 6 October 2021
Available online xxx

Keywords:
CD24
Colorectal cancer
Gene expression
HT-29
TCGA

ABSTRACT

In this study we analyzed expression of CD24 in a cohort of colorectal cancer patients using immunohistochemistry staining of CD24. We found a significant association between absence or low expression of CD24 (10% of membranous and 55% of cytoplasmic staining) and shortened patient survival. Protein localization played a crucial role in the prognosis: membranous form was the major and prognostic one in primary tumors, while cytoplasmic expression was elevated in liver metastases compared to the primary tumors and contained prognostic information. Then, using The Cancer Genome Atlas Colon Adenocarcinoma (TCGA-COAD) RNA-seq data, we showed that *CD24* mRNA level was two-fold decreased in primary colorectal cancers compared to adjacent normal mucosa. Like the protein staining data, ten percent of patients with the lowest mRNA expression levels of *CD24* in primary tumors had reduced survival compared to the ones with higher expression. To explain these findings mechanistically, shRNA-mediated CD24 knockdown was performed in HT-29 colorectal cancer cells. It resulted in the increase of cell migration *in vitro*, no changes in proliferation and apoptosis, and a slight decrease in cell invasion. As increased cell migration is a hallmark of metastasis formation, this finding corroborates the association of a decreased CD24 expression with poor prognosis. Differential gene expression analysis revealed upregulation of genes involved in cell migration in the group of patients with low *CD24* expression, including integrin subunit α3 and α3, β3 subunits of laminin 332. Further co-expression analysis identified SPI1, STAT1 and IRF1 transcription factors as putative master-regulators in this group.

1. Introduction

CD24 is a small GPI-anchored cell surface glycoprotein with diverse roles in several malignancies. Initially, this protein was found to be a ligand for P-selectin expressed on tumor cells [1], while more recent studies suggest that CD24 is a stem cell surface marker in many cancer entities [2] and plays a crucial role in tumor cells immune evasion via interactions with Siglec-10 [3,4]. Many authors showed that high protein expression levels of CD24 were associated with shortened patient overall survival and disease-free survival in breast [5,6], ovarian [7,8], colorectal [9,10] and other cancers [11,12]. In contrast, clinical and experimental studies showed that low or even absent CD24 expression is related to increased metastasis formation in breast [13,14] and pancreatic [15] cancers.

* Corresponding author. Faculty of Biology and Biotechnology, HSE University, Moscow, Russia.
E-mail address: atonevitsky@hse.ru (A. Tonevitsky).
[1] These authors have contributed equally to this work.

https://doi.org/10.1016/j.biochi.2021.10.004
0300-9084/

Please cite this article as: S. Nersisyan, A.-K. Ahlers, T. Lange *et al.*, Low expression of CD24 is associated with poor survival in colorectal cancer, Biochimie, https://doi.org/10.1016/j.biochi.2021.10.004

The existing studies of CD24 expression and its association to clinicopathological features of colorectal cancer are contradictive: there are several reports showing an association of poor survival with either high [9,10] or low [16,17] CD24 expression. To address this scientific question, we evaluated CD24 protein expression in formalin-fixed tissue microarrays (TMAs) of the primary tumors and liver metastases of colorectal carcinoma patients. The expression was scored both for apical membrane staining and intracytoplasmic staining to assess the role of CD24 protein localization on patient overall survival. In addition, we carried out bioinformatics analysis to study the distribution of *CD24* mRNA expression in primary colorectal tumors and adjacent normal tissues, as well as to detect possible associations of *CD24* expression with patients survival using RNA sequencing data from The Cancer Genome Atlas Colon Adenocarcinoma (TCGA-COAD) [18]. Finally, we performed shRNA-mediated knockdown of CD24 in human HT-29 colorectal cancer cells to study phenotypic effects related to downregulated CD24 expression.

2. Materials and methods

2.1. RNA sequencing data processing

Bulk RNA sequencing profiles of colorectal cancers and adjacent normal tissues from TCGA-COAD cohort were downloaded from GDC Data Portal (https://portal.gdc.cancer.gov) in format of count matrices. Raw FASTQ files corresponding to five matched primary tumors and liver metastases were downloaded from Sequence Read Archive (SRA) under SRP060016 accession number [19]. Reads were mapped to the reference genome (GENCODE GRCh38.p13) with STAR 2.7.5b [20]. GENCODE genome annotation (release 34) [21] was used to generate the count matrix.

RNA-seq library sizes were normalized with the Trimmed Mean of M-values (TMM) algorithm implemented in edgeR v3.30.3 package [22], default filtering of genes with near-zero read counts was applied. The same edgeR package was used to generate TMM-normalized Fragments Per Kilobase of transcript per Million mapped reads (TMM-FPKM) matrices, obtained values were $\log_2$-transformed. Low expressed genes were detected using the first quartile of TMM-FPKM medians distribution and were removed from the downstream analysis.

2.2. Differential gene expression and enrichment analyses

Differential expression analysis was conducted with the use of DESeq2 v1.28.1 [23], *p*-values were adjusted by the Benjamini-Hochberg procedure. Shrinkage of logarithmic fold changes was performed using apeglm method available in DESeq2 [24], default 0.005 threshold was set on *s*-values and 0.5 threshold was set on absolute values of $\log_2$ fold changes (examples of the similar standard fold change threshold could be found, e.g., in Refs. [25,26]). Sets of differentially expressed genes were uploaded to DAVID v6.8 [27] for Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways enrichment analysis.

2.3. Survival analysis

TCGA-COAD cohort was randomly split into equal-sized training and validation sets with a stratification by outcome and survival time (either date of death or date of the last follow-up). Threshold for determination of $CD24^{low}$ and $CD24^{high}$ groups was selected by the search over a grid of quantile levels (0.1, 0.2, …, 0.9) with minimization of logrank test *p*-value on the training patients cohort (all patients with *CD24* expression level below the threshold were marked as $CD24^{low}$, while the rest of patients were classified as $CD24^{high}$). Lifelines implementations of logrank test and Kaplan-Meier estimation were used (https://doi.org/10.5281/zenodo.805993).

2.4. Transcription factor analysis

JASPAR 2020 database [28] was used to download a complete list of human transcription factors (TFs). Identification of TFs regulating *CD24* expression was conducted with hTFtarget web application [29] (http://bioinfo.life.hust.edu.cn/hTFtarget#!/).

2.5. Cell cultures

The human colon adenocarcinoma cell line HT-29 was purchased from the European Cell Culture Collection (Porton Down, Wiltshire, UK). Cells were grown in RPMI 1640 + L-Glutamine (Gibco, Thermo Fisher Scientific, MA, USA) supplemented with 10% fetal calf serum (Gibco), 1% penicillin and streptomycin (Gibco) and cultured under standard cell culture conditions (37 °C, 100% relative humidity, 5% CO_2). The HT-29 cells were additionally tested for mycoplasma contamination.

2.6. ShRNA-mediated CD24 knockdown

CD24 knockdown in HT-29 cells was achieved by a shRNA-mediated approach: a 65 bp DNA oligomer containing a 19 bp anti-*CD24* sequence (AGGCCAAGAAACGTCTTCT) was inserted into the pLVX-Puro vector (Clontech, Takara Bio Europe, Saint-Germain-en-Laye, France). The anti-CD24 sequence was checked for potential off-target effects using NCBI BLAST (plus/plus and plus/minus strands down to an E value of 170, see Ref. [30]) and no off-target sequences were found. The same vector containing a sequence against firefly luciferase was used to generate a transduced control cell line (HT-29 shLuc). Viral particles were produced as cell-free supernatants by transient transfection of HEK-293T packaging cells as described [31]. In brief, lentiviral vectors based on pLVX-shRNA1 were packaged using the second-generation packaging plasmid psPAX2 (Addgene #12260, Addgene, Teddington, UK) and phCMV-VSV-G [32] expressing the envelope protein of vesicular stomatitis virus. The supernatant was harvested 24 h after transfection, 0.45 μm filtered and stored at −80 °C. Target cells were plated at 5×10^4 cells in 0.5 mL medium in each well of a 24 well plate. After addition of viral particles containing supernatant to the cells, the medium was replaced next day and puromycin was added the second day after transduction at a concentration of 1 μg/mL. The puromycin selection was carried out for at least one week. For MACS sorting, cells were trypsinized and resuspended in the MACS buffer. 8 μl of mouse anti-human PE-labeled CD24 (clone SN3 A5-2H10, eBioscience™, Fisher Scientific, Schwerte, Germany) were added and incubated on ice for 10 min. After washing, cells were incubated with anti-PE MicroBeads (Miltenyi Biotec, Bergisch Gladbach, Germany). Cells were then separated on a MACS separator, using LD columns (Miltenyi Biotec), and collected in 6-well plates (Sarstedt, Nümbrecht, Germany). The efficiency of CD24 knockdown was confirmed by fluorescence-activated cell sorting (FACS).

2.7. Flow cytometry

For verification of the CD24 knockdown, 5×10^4 - 1×10^6 cells were incubated with PE-labeled anti-CD24 (mAB mouse anti-human CD24, clone SN3 A5-2H10, eBioscience™, Fisher Scientific, Schwerte, Germany) or corresponding isotype control (mouse IgG1-PE, eBioscience™) diluted in PBS with 1% BSA in a final concentration of 1 μg/mL at 4 °C. For the integrin flow cytometry

analysis, 5×10^4 HT-29 shLuc and shCD24 cells were incubated with FITC-labeled anti-CD49C (1:50; Integrin α3, clone REA360, Miltenyi Biotec, Bergisch-Gladbach, Germany). As isotype control, the recombinant human IgG1 antibody (1:50; clone REA293, Miltenyi Biotec) was used. All measurements were compensated. Labeled cells were subjected to fluorescence assisted flow cytometry on the MACSQuant Analyzer 10 (Miltenyi Biotec, Bergisch-Gladbach, Germany). Files were analyzed with the FlowLogic Software (Innovai, Sydney, Australia).

2.8. Migration, invasion and proliferation in vitro assays

For the migration assay, the Oris Platypus Universal Cell Migration Kit (Platypus Technologies LLC, Fitchburg, WI, USA) was used in accordance with the manufacturer's instructions. A cell-seeding stopper ($d = 2$ mm^2) was positioned in each well on the surface of the 96-well plate. 5×10^4 HT-29 shLuc and shCD24 cells (Sarstedt, Nümbrecht, Germany) were seeded in a volume of 200 µl per well (in supplemented 1640 RPMI medium) on the outer annular region of the stopper. Cells were allowed to adhere at 37 °C under 5% CO_2 for 24 h. After 24 h the stoppers were removed to permit cell migration into the detection zone. Cell migration was photographed daily on the microscope (Axio, Zeiss; AxioCam ERc5s, Zeiss), over a period of four days. Migration was further analyzed as percent closure by determining the area of the detection zone after migration (96 h) compared to pre-migration (0 h), using the ImageJ software (http://rsb.info.nih.gov/ij/). For the analysis, 12 replicates were carried out for both shLuc and shCD24.

As an additional migration assay, the Corning® FluoroBlok™ Transwell Migration Assay (Corning USA, NY) was used. HT-29 shLuc and shCD24 cells were harvested and diluted in the appropriate serum-free media (without supplements) to a concentration of 5×10^4 cells/mL. 500 µl of cell suspension were seeded in each well of the 24-Multiwell Insert Plate. 750 µl of supplemented media were placed in the lower wells. The FluoroBlok plates containing media and cells were incubated for 22 h at 37 °C. The next day, the lower plate was replaced with a new one. 500 µl of HBSS (Gibco) and 2 µl Calcein AM fluorescent dye (BD Pharmingen™ #564061, NJ, USA) were added to each well, followed by incubation of the cells for 1 h at 37 °C. Labeled migratory cells were measured in a fluorescence microplate reader (TECAN GENios™ plate reader, Tecan Trading AG, Switzerland) at wavelengths 485 nm and 535 nm. Five replicates were done per cell line.

Invasion test was performed using a Corning® BioCoat™ Tumor Invasion 24-well Plate (Corning, #354166, MA, USA). Before use, the plate was rehydrated with PBS for 2 h at 37 °C in a non-CO_2 environment. The BioCoat FluoroBlok invasion system was incubated for 22 h at 37 °C and 5% CO_2. Fluorescence of invaded cells was measured in a fluorescence microplate reader (TECAN GENios™ plate reader) at wavelengths 485 nm and 535 nm. The rest of the protocol was completely similar to the one described in the previous paragraph, except the use of 4 µg/mL Calcein AM fluorescent dye. As in case of Transwell Migration Assay, five replicates were done per cell line.

Proliferation of HT-29 shLuc and shCD24 cells was assayed using the CellTrace CFSE kit (Thermo Fisher Scientific, #C34554, MA, USA). The cell pellets were resuspended in 2 mL of CellTrace CFSE staining solution supplemented with 0.1% FCS and incubated for 20 min at 37 °C using the MACSmix™ tube rotator (Miltenyi Biotec, Bergisch-Gladbach, Germany). After centrifugation (1500 rpm, 5 min), cell pellets were resuspended in 2 mL RPMI 1640 without supplements. 6×10^4 cells were seeded in a volume of 1000 µl (in RPMI 1640 medium without supplements) in a 12-well plate (Sarstedt, Nümbrecht, Germany). Cells were allowed to adhere at 37 °C under 5% CO_2 for 24 h and proliferation was measured every day over a period of three days using the MACSQuant Analyzer 10 (Miltenyi Biotec). All measurements were compensated. Files were analyzed with the FlowLogic Software.

2.9. Apoptosis assay

For the detection and discrimination of apoptotic, necrotic and dead cells, the Annexin V-FITC kit (Miltenyi Biotec, #130-092-052) was used according to manufacturer's instructions. 10^6 cells were washed in 1× Binding Buffer and centrifuged at 300×*g* for 10 min. For gating of dead cells, 10^6 cells were heated at 95 °C for 10 min. The pellets of all samples were resuspended in 100 µl of 1× Binding Buffer and stained with 10 µl of Annexin V-FITC, followed by incubation for 15 min at RT. After washing with 1× Binding Buffer and centrifugation at 300×*g* for 10 min, the cells were resuspended in 100 µl of 1× Binding Buffer and analyzed using the MACSQuant Analyzer 10 (Miltenyi Biotec). Dead cells were positively stained with propidium iodide (100 µg/mL).

2.10. QRT-PCR and primers used

For cDNA synthesis, purified RNA of HT-29 shLuc and shCD24 cells was reverse transcribed with the RT^2 First Strand Kit (Qiagen, Hilden, Germany). Real-time PCR amplification of genes was performed on 96-well plates in a LightCycler® 480 Real-Time PCR System (Roche, Basel, Switzerland). QRT-PCR was performed according to manufacturer's instructions using the SYBR Green I Master kit. 60 ng of cDNA were applied per reaction. The samples were run in triplicate following the MIQE guidelines [33]. A no-template control with RNase-free water instead of cDNA was included. Expression levels were calculated as relative values using the mean of both reference genes *GAPDH* and *ACTB*. The difference of the C_T-values (ΔC_T) from the target gene and the mean of the C_T-values from both reference genes were determined to quantify the target gene expression. Expression levels were further related ($\Delta\Delta C_T$) to control samples using the difference of the ΔC_T -value from the sample (ΔC_T sample) as well as the ΔC_T -value from the control (ΔC_T control) and the relative values were calculated as the $2^{-\Delta\Delta C_T}$ according to Livak and Schmittgen [34]. Oligonucleotide sequences of primers are listed in Supplementary Table 1.

2.11. Patient samples and immunohistochemistry

Formalin-fixed tissue microarray (TMA) sections from colorectal carcinoma patients were used for immunohistochemical analysis. These were retrospectively collected samples of cases where primary tumor tissue as well as tissue of resected liver metastases was available from the University Medical Center, Göttingen, Germany. TMAs were generated with a core needle diameter of 1 mm and according to the approval by the local ethics committee (application number 21/3/11). The mAB SWA11 (kindly provided by Prof. Dr. Peter Altevogt) specifically reacts with the LAP amino acid motif in the protein core to human CD24 [35]. Deparaffinized freshly cut sections of TMAs were rehydrated, followed by antigen retrieval with Fast Enzyme (BIOZOL, Eching, Germany) for 5 min at RT. The primary antibody mAB SWA11 was diluted with the Antibody Diluent (Dako #S3022, Agilent Technologies, CA, USA) to a final concentration of 28 µg/mL and incubated for 1 h at RT. In parallel, control sections were incubated with isotype control mouse IgG2a (1:3; Dako #X0943, Agilent Technologies, CA, USA). The secondary antibody (biotinylated goat anti-mouse, Agilent #E0433) was used in a final concentration of 3,25 µg/mL and incubated for 30 min at RT. The immunohistochemical reaction was visualized using the detection system ABC-AP (Vectastain® #AK-5000, Vector Laboratories, CA, USA) for 30 min at RT, followed by Permanent Red (up to

15 min at RT; Dako #K0640, Agilent Technologies, CA, USA) and counterstaining of nuclei with hemalum staining (5–10 s; Mayer's hemalum solution #109249, Merck, Darmstadt, Germany). Expression intensity was examined in a semiquantitative manner (score 0: no staining, score 1: weak, score 2: moderate and score 3: strong staining).

2.12. Statistical analysis

Comparisons of migration, invasion, proliferation, apoptosis, FACS and qRT-PCR data between HT-29 shLuc and shCD24 cells were performed with Student's *t*-test. Intensities of CD24 protein expression levels quantified by immunohistochemical analyses were compared between primary tumors and matched liver metastases with Wilcoxon's signed-rank test. All code for the statistical analysis was written in Python programming language with the use of SciPy stats module [36].

3. Results

Decreased CD24 protein expression in apical membrane of primary tumor and cytoplasm of liver metastases as well as total mRNA expression is associated with poor survival in colorectal cancer.

Formalin-fixed tissue microarrays (TMAs) of primary tumors and liver metastases of colorectal carcinoma patients were evaluated for CD24 protein expression (clinicopathological data of the study cohort are presented in Table 1 and Supplementary Table 2). As CD24 showed a cytoplasmic and apical membrane staining, we scored both staining qualities and classified them into negative (0), weakly (1), moderately (2) or strongly (3) immunoreactive. Interestingly, we observed a clear and statistically significant decrease of CD24 immunoreactivity in apical tumor cell membranes of liver metastases compared to matched primary tumors ($p = 0.0256$, Fig. 1A–C) and an increase of cytoplasmic CD24 immunoreactivity in liver metastases ($p = 3.10 \times 10^{-3}$, Fig. 1D–F).

For both cytoplasmic and apical membrane staining, the absence or weak expression of CD24 were associated with a statistically significant decrease in survival of patients: logrank test $p = 3.93 \times 10^{-3}$ for apical membrane staining (low expression group was 10.3% of total cases, Fig. 2A) and $p = 0.0198$ for cytoplasm (55.2% of total cases were from the low expression group, Fig. 2B). Then, the same analysis was separately performed for primary tumor and metastasis samples. In the case of primary tumors apical membrane CD24 expression was prognostic (7.9% of tumors, $p = 2.23 \times 10^{-6}$, Fig. 2C), while cytoplasmic expression had not resulted in a statistically significant divergence of Kaplan-Meier curves ($p = 0.534$, Fig. 2D). To the contrary, cytoplasmic expression was the prognostic one for liver metastases ($p = 0.0258$, Fig. 2F) unlike the membrane staining ($p = 0.172$, Fig. 2E). These data fully matched the one presented in the previous paragraph: prognostic information was concentrated in the majorly expressed CD24 form (membranous for primary tumors and cytoplasmic for liver metastases).

Aside from our immunohistochemistry analyses, we assessed mRNA expression levels of *CD24* in 429 primary tumors and 38 adjacent normal tissues of patients from The Cancer Genome Atlas Colon Adenocarcinoma (TCGA-COAD). As a result, we identified decreased *CD24* expression in cancer (1.9 folds compared to the adjacent normal mucosa, $p = 1.78 \times 10^{-6}$) without any significant differences between stages of cancer (Fig. 3). To study the role of *CD24* mRNA expression in colorectal cancer prognosis, we identified a threshold for defining classes of low and high *CD24* expression (from here onwards we refer to these classes as $CD24^{low}$ and $CD24^{high}$), which minimized logrank test *p*-value associated with separation of $CD24^{low}$ and $CD24^{high}$ classes on the training set. The best divergence of Kaplan-Meier curves on the training set was observed when $CD24^{low}$ class was defined by taking 10% of primary tumors with the lowest expression estimates ($p = 0.0118$). At the same time, separation was also statistically significant for the validation set ($p = 4.03 \times 10^{-4}$), confirming the association between low *CD24* expression and poor prognosis on the transcriptional level (Fig. 4). Notably, the percentage of tumors with low CD24 expression (approximately 10%) matched between our immunohistochemical staining and TCGA transcriptomics data.

Finally, we compared expression patterns of *CD24* mRNA

Table 1
Clinicopathological data of the study cohort of colorectal cancer patients.

	Total	Apical membrane			Cytoplasm		
		CD24 low (n = 6)	CD24 high (n = 52)	*p*-value	CD24 low (n = 32)	CD24 high (n = 26)	*p*-value
Age of diagnosis, completed years, median (Q25–Q75)	61.5 (55–65.75)	59 (53.25–67)	61.5 (56.5–65.25)	0.918	61.5 (56.25–66.5)	61.5 (55–64)	0.567
Age							
≤60	27	3 (50%)	24 (46.1%)	1	15 (46.9%)	12 (46.2%)	1
>60	31	3 (50%)	28 (53.9%)	1	17 (53.1%)	14 (53.8%)	1
Gender							
Male	43	4 (66.7%)	39 (75%)	0.643	21 (66%)	22 (84.6%)	0.136
Female	15	2 (33.3%)	13 (25%)	0.643	11 (34%)	4 (15.4%)	0.136
Sample type							
Primary tumor	38	3 (50%)	35 (67.3%)	0.405	27 (84.4%)	11 (42.3%)	**1.84×10^{-3}**
Liver metastases	20	3 (50%)	17 (32.7%)	0.405	5 (15.6%)	15 (57.7%)	**1.84×10^{-3}**
Lymph node metastases							
No	21	0 (0%)	21 (40.4%)	0.0774	12 (37.5%)	9 (34.6%)	1
Yes	37	6 (100%)	31 (59.6%)	0.0774	20 (62.5%)	17 (65.4%)	1
Clinical stage							
I, II	11	0 (0%)	11 (21.1%)	0.583	6 (18.7%)	5 (19.2%)	1
III, IV	47	6 (100%)	41 (78.9%)	0.583	26 (81.3%)	21 (80.8%)	1
Resection status							
R0	56	5 (83.3%)	51 (98.1%)	0.198	30 (93.7%)	26 (100%)	0.497
R1	2	1 (16.7%)	1 (1.9%)	0.198	2 (6.3%)	0 (0%)	0.497

Feature percentages were compared between low and high CD24 expression classes with Fisher's exact test. Bold values indicate *p*-value < 0.05.

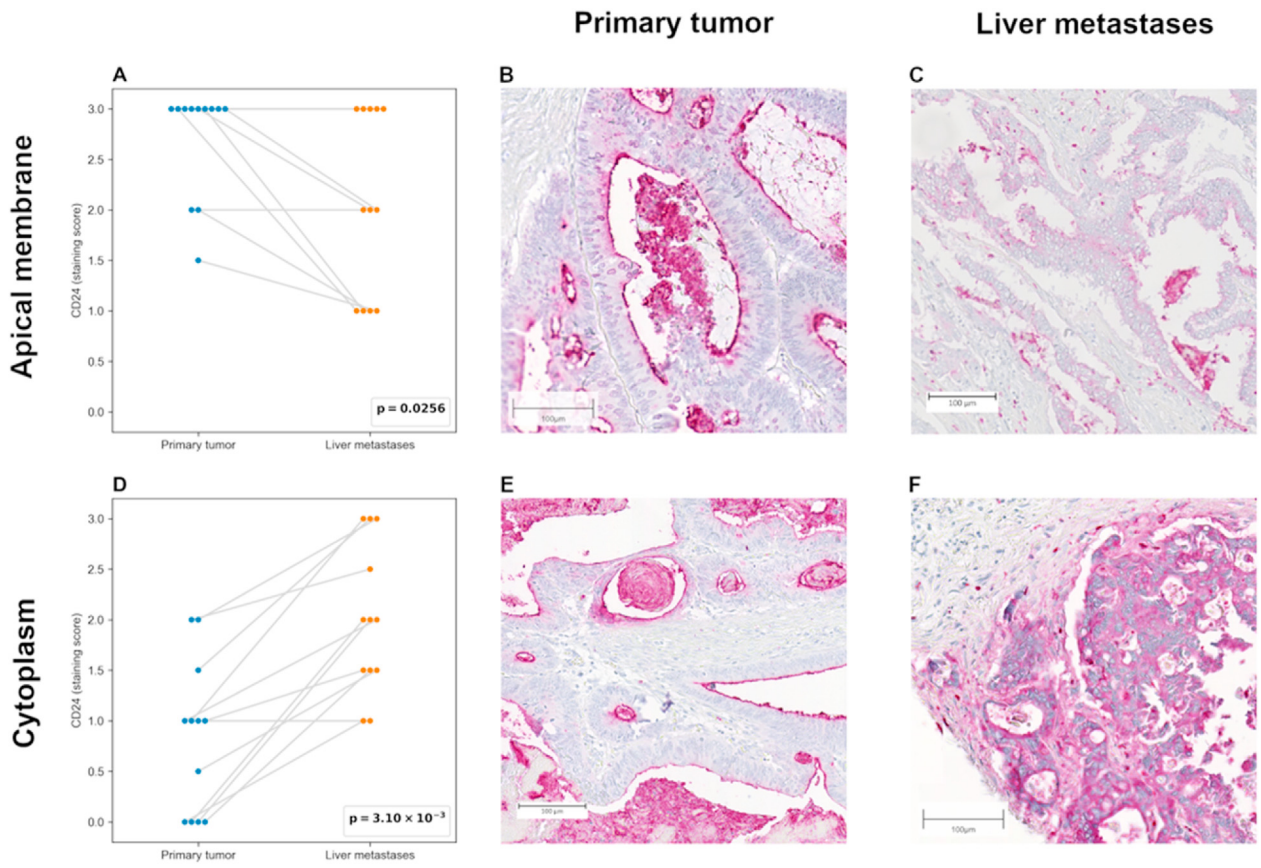

Fig. 1. CD24 protein levels are decreased in apical membranes and increased in cytoplasm of liver metastases compared to primary tumors of analyzed TMAs from colon carcinoma patients. (A–C) CD24 expression was decreased at the apical membrane of liver metastases (**C**) compared to primary tumor (**B**), as shown for the immunohistochemical staining for CD24 ($p = 0.0256$). **(D–F)** In the cytoplasm of liver metastases (**F**), a significant increase of CD24 expression was noted compared to the primary tumor (**E**; $p = 3.10 \times 10^{-3}$).

between primary tumors and matched liver metastases. Bulk RNA-seq data of five primary tumors and matched metastases samples generated by Lee et al. [19] were processed. *CD24* expression was found to be 2.3-fold downregulated in liver metastases ($p = 0.00137$). Apical membrane/cytoplasm staining patterns, together with the decrease of total mRNA expression in metastases, uncover complex pathways involving not only transcriptional regulation, but also the independent regulation of CD24 protein sorting within the cancer cells.

3.1. Knockdown of CD24 from HT-29 cells results in increased cell migration

To uncover the possible mechanisms underlying decreased survival associated with CD24 down-regulation, we performed shRNA-mediated knockdown of CD24 in HT-29 human colorectal adenocarcinoma cells. Knockdown was confirmed by FACS analysis showing a 98% decrease in CD24 expression (Fig. 5A). The functional role of CD24 for human colorectal cancer cell migration was further determined by Platypus and Transwell migration assays. The Platypus assay was used to analyze the extent of migration into the detection zone over time. In the shCD24 group, the relative migration (in %) of the central cell-free detection zone was significantly increased after 96 h of documentation as compared to the control (shLuc) group ($p = 8.12 \times 10^{-10}$; Fig. 6AB). The increased motility of HT-29 shCD24 cells was further confirmed by the Transwell migration assay ($p = 7.87 \times 10^{-5}$; Fig. 6C). Thus, the observed differences in survival can be due to the increased migration of cancer cells with the lowest expression levels of *CD24*.

In addition to migration, we also performed invasion, proliferation and apoptosis tests. With the use of Transwell invasion assay we observed a very slight (though statistically significant) decrease of cell invasion upon CD24 knockdown ($p = 0.0282$, Supplementary Fig. 1A). While cell migration and invasion are usually co-regulated, there exist well described examples of uncoupled migratory/invasive phenotypes in cancer [37]. The proliferation index measured by CFSE assay showed no statistically significant differences between HT-29 shLuc and shCD24 cells (Supplementary Fig. 1B). Apoptopic HT-29 cells, grown as 3D culture, were measured by an Annexin V-FITC staining. No statistically significant differences in apoptosis were detected between shLuc and shCD24 cells (Supplementary Fig. 2).

3.2. Genes involved in cell migration and adhesion are upregulated in $CD24^{low}$ group

Using differential expression analysis, we assessed differences in transcriptomic landscapes of $CD24^{low}$ and $CD24^{high}$ primary tumors from TCGA-COAD in order to discover possible mechanisms and pathways responsible for decreased survival and increased migration associated with low *CD24* expression. As a result, we identified 604 differentially expressed genes (DEGs) and applied enrichment analysis to detect pathways associated with these genes.

Focal adhesion and extracellular matrix-receptor (ECM-receptor) interaction pathways were highly enriched in the list of DEGs

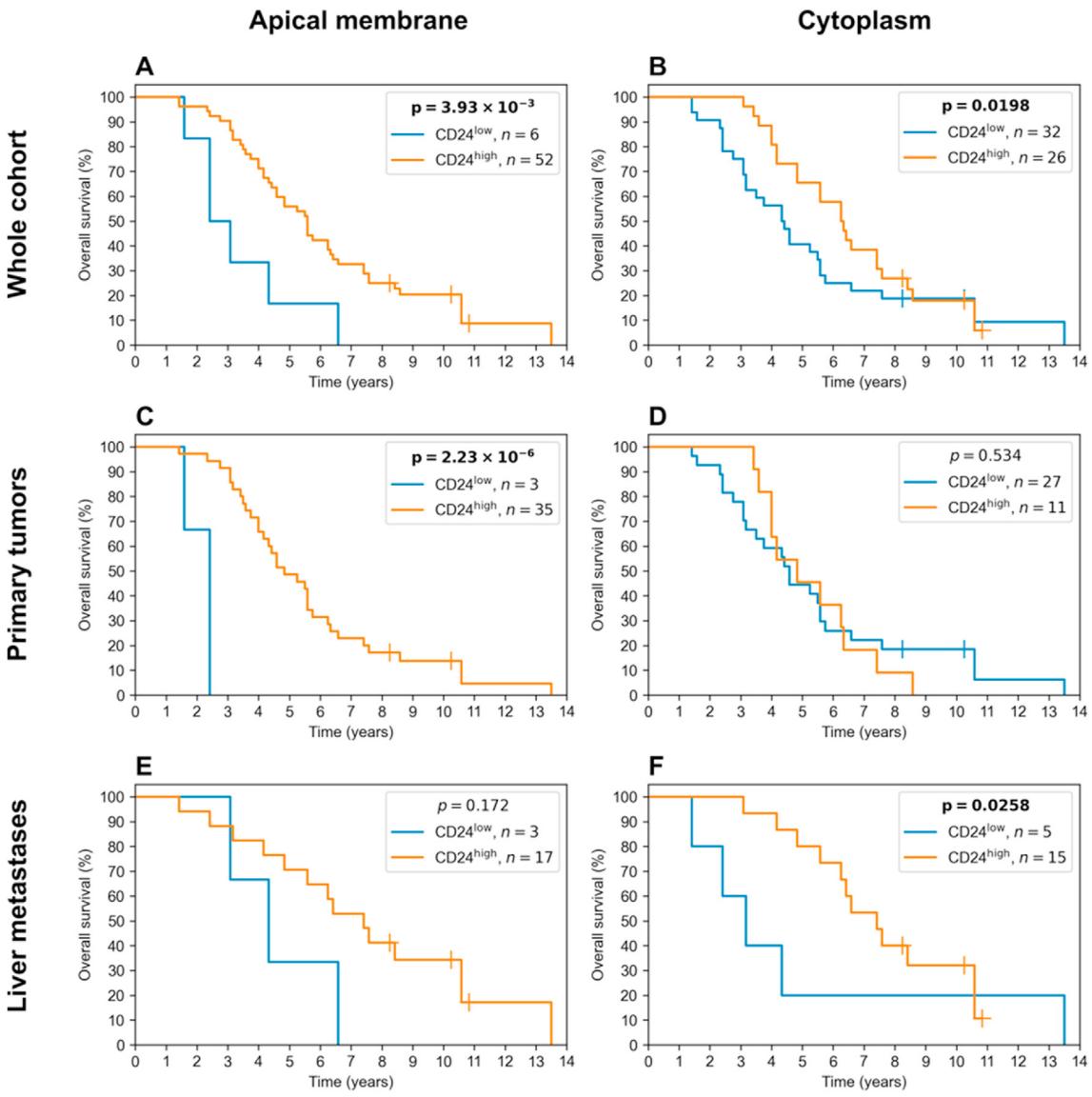

Fig. 2. Low levels of CD24 protein abundance in apical membrane of primary tumors and cytoplasm of liver metastases are unfavorable markers for colorectal cancer patients.

(Table 2). Interestingly, most of such genes (19 out of 21) were upregulated in the *CD24*^{low} group. We paid special attention to upregulation of *ITGA3*, *LAMA3* and *LAMB3* genes encoding integrin subunit α3 and α3, β3 subunits of laminin 332, since the direct interaction between these two common adhesion molecules could serve as an explanation for increased migration related to low CD24 expression. Notably, integrin subunit α3 protein expression level measured by FACS was 1.5-fold increased upon CD24 knockdown in HT-29 cells ($p = 0.0287$, Fig. 5B). Thus, the negative correlation between *CD24* and *ITGA3* expression levels in TCGA-COAD was also observed *in vitro*.

3.3. SPI1 transcription factor is a potential master-regulator in *CD24*^{low} colorectal cancer

To identify potential transcription regulators in the *CD24*^{low} group, we quantified the number of genes co-expressed with known transcription factors (TFs). For this analysis, we selected TF-gene pairs with strong correlation in the TCGA-COAD dataset (different thresholds for Spearman correlation coefficient were used). Next, for each differentially expressed TF we calculated the number of co-expressed DEGs as well as the total number of co-expressed genes. Finally, we ranked the list of TFs according to the enrichment of their co-expressed DEGs in the *CD24*^{low} group using hypergeometric test.

As a result, we identified several TFs with a highly over-represented number of co-expressed *CD24*^{low} DEGs (Table 3). Interestingly, while for some TFs (such as SPI1 or STAT1) the absolute number of co-expressed DEGs was high for medium correlation cutoff (0.5), more stringent cutoff (0.7) resulted in especially high percentage of co-expressed DEGs. Most genes co-expressed with *SPI1* (71%), *STAT1* (90%) and *IRF1* (88%) were deregulated in the *CD24*^{low} colon cancer. A statistically significant increase of *STAT1* (2.1 folds, $p = 7.43 \times 10^{-3}$) and *IRF1* (1.5 folds, $p = 8.03 \times 10^{-3}$) mRNA expression levels was detected upon CD24 knockdown in HT-29 cells by qRT-PCR, while *SPI1* was not expressed in the considered cell lines. Since the analysis of TFs and co-expressed genes was conducted in the set of DEGs, distribution of calculated correlation coefficients could be highly biased from the random background level: if two genes appear to be deregulated in the

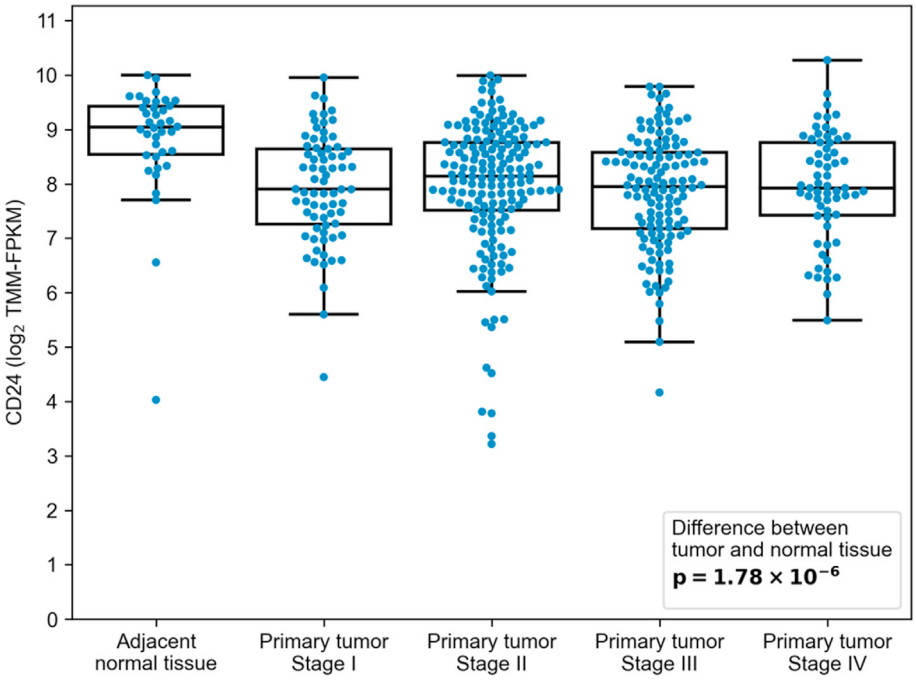

Fig. 3. Distribution of *CD24* mRNA expression levels in primary tumors and adjacent normal tissues from TCGA-COAD.

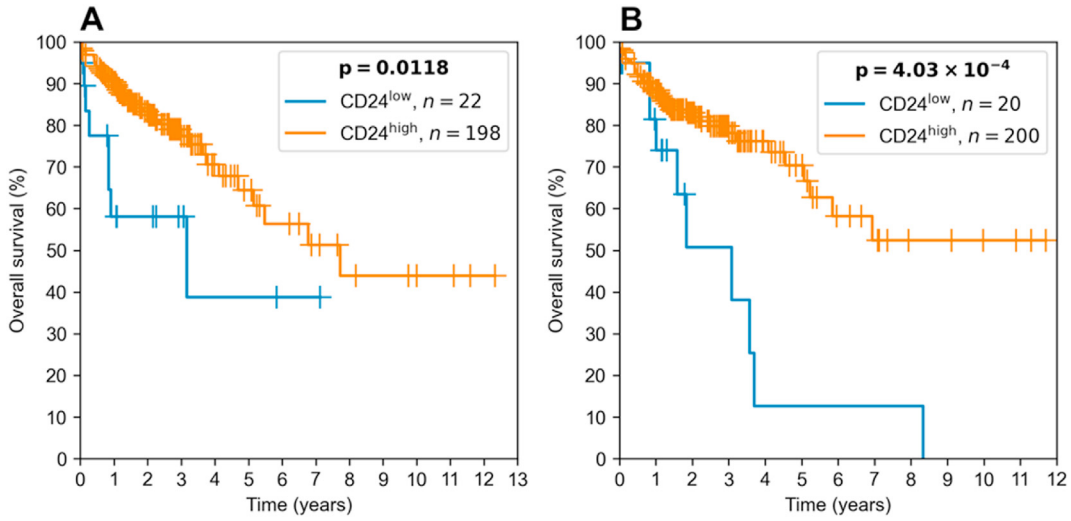

Fig. 4. Low levels of *CD24* mRNA expression are associated with poor prognosis in TCGA-COAD patients. (A) Training set ($p = 0.0118$). **(B)** Validation set ($p = 4.03 \times 10^{-4}$).

CD24^{low} group, it is highly probable to observe significant correlation between their expression levels. To take a possibility of such bias into account, we performed the same procedure for non-TF DEGs. It turned out that *SPI1* was co-expressed with the maximum number of genes compared both to TFs and non-TFs ($p = 0.00171$). Thus, SPI1 is a hub TF potentially responsible for transcriptomic changes in the *CD24*^{low} colon cancer.

Finally, we searched for possible transcription regulators of *CD24*. Notably, *CD24* was predicted to be a direct target of SPI1 both by analysis of TF binding motif and ChIP-seq data (according to hTFtarget database). Thus, SPI1 can be considered not only as a hub in the *CD24*^{low} group, but also as a possible direct regulator of *CD24* mRNA expression.

4. Discussion

In this contribution we demonstrate the association between low protein and mRNA expression levels of CD24 and poor prognosis of colorectal cancer using immunohistochemical staining in patient material from our own cohort and RNA-seq data from TCGA-COAD. These data were collected as multiple existing reports contain contradictory information on the role of CD24 expression in colorectal cancer. Specifically, some authors demonstrate an association between high CD24 expression and shortened patients survival and an increase in the metastatic load [9,10], while others show no significant prognostic value [16] or inverse association [17].

CD24 protein expression was separately evaluated for apical membranous and intracytoplasmic staining, which allowed us to

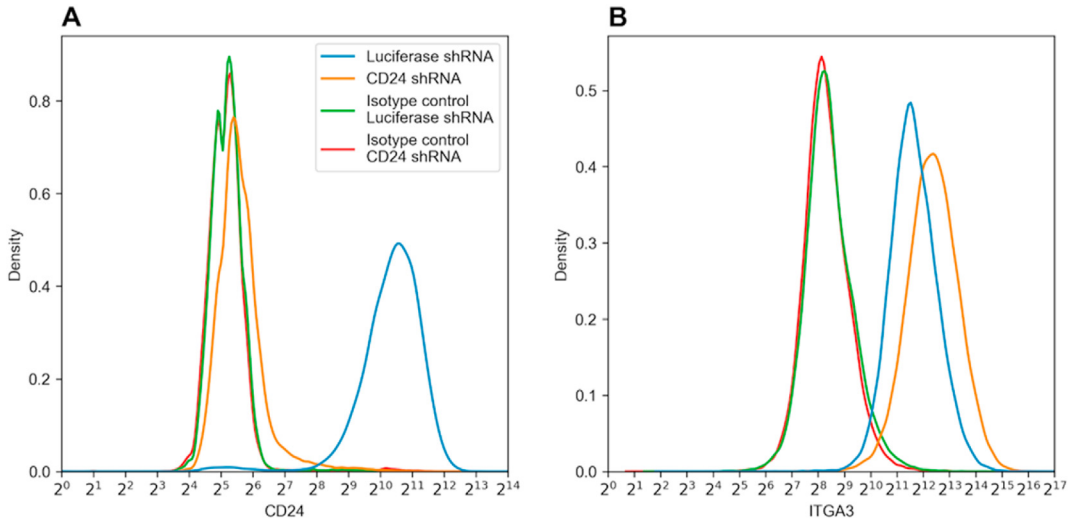

Fig. 5. FACS analysis of CD24 and ITGA3 proteins after CD24 knockdown. (A) The CD24 knockdown efficiency was evaluated by FACS and determined as 98% decrease in expression. **(B)** Protein expression level of integrin subunit α3 was increased upon CD24 knockdown (representative replicate).

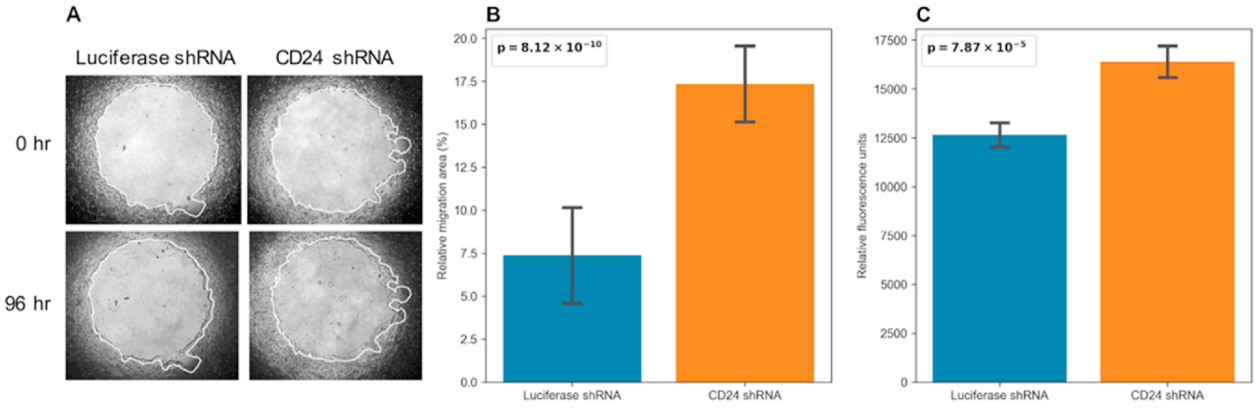

Fig. 6. CD24 knockdown leads to increased migration of HT-29 cells. (A) Representatives of the pre-migration (0 h) and post-migration (96 h) images of the Platypus assay are shown for HT-29 shLuc and shCD24. White polygons indicate the central cell-free zone that was covered by stoppers before cell migration. Migrated areas were determined as the percentage of migration area of HT-29 shCD24 compared to control (shLuc) cells. **(B)** A significant increase of migration (%) was determined upon CD24 knockdown in the Platypus assay at 96 h ($p = 8.12 \times 10^{-10}$). **(C)** The increase in migration was validated by the Transwell migration assay ($p = 7.87 \times 10^{-5}$).

get insights on the role of CD24 localization within different cellular compartments. It turned out that CD24 expression was decreased in apical membranes and significantly increased in cytoplasm site of liver metastases compared to the primary tumors. Concordantly, membranous CD24 expression was prognostic in primary tumors, while cytoplasmic form had the prognostic ability in liver metastases. The absence of CD24 on the apical membrane in the liver metastasis is of interest as only those CD24 molecules which are present on the cell surface and carry P-selectin binding sites can serve as P-selectin ligands on the endothelium. Hence this finding indicates either that P-selectin sites are irrelevant for liver metastasis formation or that CD24 can be up or down regulated during epithelial mesenchymal transition. As CD24 is highly expressed in cells of the immune system and hence mesenchymal, the down-regulation of CD24 within liver metastases may indicate a reversal towards a more epithelial phenotype, the so called mesenchymal to epithelial transition.

The information about the cytoplasmic CD24 as a prognostic marker in colorectal cancer is also conflictive. It was shown that the strong cytoplasmic CD24 expression significantly correlated with shortened patient survival in the cohort of patients without distant metastases [9]. However, Choi with co-authors did not detect any positive correlation between CD24 expression and patient survival, nodal status [38]. In the recent report, upregulation of cytoplasmic CD24 expression was connected to malignant transformation but favorable prognosis of colorectal adenocarcinoma [39]. To the best of our knowledge, there were no studies with the comparison of the expression levels of cytoplasmic CD24 in primary tumors and matched liver metastases. In addition, there is no consensus on molecular mechanisms of membranous/cytoplasmic CD24 regulation; understanding these mechanisms can uncover new roles of CD24 in colorectal cancer.

Knockdown of CD24 in HT-29 colorectal cancer cell lines resulted in a significant increase of cell migration, a process which is integral to metastasis formation and as such for a poor prognosis of a patient. Increased migration can thus serve as a possible explanation of poor prognosis associated with low *CD24* expression. Moreover, analysis of TCGA-COAD patient transcriptomes revealed the increase in integrin subunit α3 and α3, β3 subunits of laminin 332, as well as other genes directly involved in cell adhesion and migration (Table 2). FACS analysis of integrin subunit α3 revealed a significant upregulation of protein expression level upon

Table 2
Differentially expressed genes (*CD24*^{low} vs *CD24*^{high}) involved in focal adhesion (KEGG hsa04510) and ECM-receptor interaction (KEGG hsa04512) pathways.

Gene	CD24^{high} mean expression (TMM-FPKM)	Fold change	Adjusted *p*-value
BIRC3	4.2	1.8	5.98×10^{-6}
CAV1	7.7	1.6	5.54×10^{-4}
CAV2	2.1	1.9	3.43×10^{-7}
COL1A1	169.4	1.7	0.00170
COL1A2	98.5	1.5	0.00565
COL3A1	168.8	1.6	0.00113
COL5A3	2.6	1.7	3.33×10^{-5}
COL6A1	43.7	1.9	8.40×10^{-7}
COL6A2	56.1	1.5	0.00157
FN1	32.3	1.6	0.00679
ITGA3	13.2	1.5	7.72×10^{-5}
ITGA5	7.6	1.8	6.10×10^{-5}
JUN	74.2	−1.4	8.08×10^{-5}
LAMA3	5.2	1.8	1.75×10^{-5}
LAMB2	12.2	1.5	7.75×10^{-5}
LAMB3	37.7	1.6	9.34×10^{-6}
RAC2	14.3	1.6	6.74×10^{-6}
RAC3	3.1	1.6	2.30×10^{-4}
TNC	7.8	2.4	0.0282
VAV3	8.6	−2.8	8.10×10^{-12}
VWF	8.1	1.4	0.00292

Positive and negative fold changes indicate increase and decrease in mRNA expression in the *CD24*^{low} group, respectively. Benjamini-Hochberg procedure was used to adjust *p*-values.

Table 3
Co-expression analysis of transcription factors and differentially expressed genes (*CD24*^{low} vs *CD24*^{high}).

TF	Medium correlation cutoff (|r| > 0.5)			Stringent correlation cutoff (|r| > 0.7)		
	Number of co-expressed *CD24*^{low} DEGs	Total number of co-expressed genes	*p*-value	Number of co-expressed *CD24*^{low} DEGs	Total number of co-expressed genes	*p*-value
SPI1	184	483	**1.83×10^{-110}**	95	134	**2.86×10^{-89}**
STAT1	89	124	**2.47×10^{-84}**	9	10	**5.35×10^{-11}**
IRF1	73	103	**2.26×10^{-68}**	7	8	**1.37×10^{-8}**
RUNX3	58	81	**9.21×10^{-55}**	0	0	1
BCL6	43	89	**2.40×10^{-30}**	0	0	1
NFKB2	39	113	**4.49×10^{-21}**	0	0	1
ETV5	21	30	**4.08×10^{-20}**	0	0	1
FOXD2	24	63	**1.46×10^{-14}**	0	0	1
PLAGL2	43	245	**2.09×10^{-11}**	4	21	**0.0282**
CDX2	25	95	**5.33×10^{-11}**	0	0	1
ISX	14	34	**1.45×10^{-9}**	0	0	1
ETS1	97	938	**2.06×10^{-9}**	12	42	**2.24×10^{-6}**
PPARG	13	30	**2.64×10^{-9}**	0	0	1
CEBPA	22	95	**1.04×10^{-8}**	0	0	1
SOX13	6	9	**2.33×10^{-6}**	0	0	1
BATF	3	3	**1.80×10^{-4}**	0	0	1
FOSL1	1	14	0.558	0	0	1
HNF1B	0	13	1	0	0	1
JUN	0	3	1	0	0	1
MYB	0	7	1	0	0	1

Bold values indicate *p*-value < 0.05.

CD24 knockdown, confirming the finding from TCGA-COAD RNA sequencing data. Several studies already showed the crucial role of integrin subunit α3 [40,41] as well as laminin 332 and integrin α3β1 interaction in cancer cell migration and metastasis [42,43]. Thus, reduced expression of CD24 might turn on compensatory mechanisms promoting metastases of colon cancer. The obtained results uncover new possible roles of CD24 in colorectal cancer pathogenesis in addition to already known functions such as P-selectin [1] and Siglec-10 [4] interactions.

Gene co-expression analysis revealed the transcription factor SPI1 as a potential master-regulator in *CD24*^{low} colorectal cancer. It turned out that *SPI1* is upregulated in *CD24*^{low} colorectal cancer, and most genes co-expressed with *SPI1* fell into the group of genes exhibiting differential expression in *CD24*^{low} patients compared to *CD24*^{high} ones. Furthermore, bioinformatics studies (motif search and ChIP-seq data analysis) suggests that SPI1 can be a direct inhibitor of *CD24* expression. Though the role of SPI1 in colorectal cancer is not well-studied, few reports contain information on the poor prognosis of patients with high SPI1 expression levels [44]. It is, however, interesting to note that SPI1 is a “critical regulator of cellular communication in the immune system” [45] and that interaction of circulating tumor cells with the endothelium of the target organ depends on molecules of the leukocyte adhesion cascade, which are also of importance for leukocyte trafficking [46]. While expression of SPI1 was not detectable in HT-29 cells, two other TFs (STAT1 and IRF1) were also major regulators in *CD24*^{low} colorectal cancer and were significantly upregulated on mRNA level upon knockdown of CD24 *in vitro*. Consistently with these findings,

in a recent study Tanaka with co-authors showed that high cytoplasmic expression of STAT1 was unfavorable marker for colorectal cancer [47].

We note several limitations of our work. First, *in vivo* experiments with CD24-deficient colorectal cancer cells should be carried out to study the mechanisms underlying our findings more deeply. Special attention should be paid to the method of CD24 silencing: while shRNAs are widely used, there is no guarantee that they will result in sufficiently strong downregulation of CD24 which is required for the decreased survival (based on our results, this is applicable only for the 10% of tumors with the lowest CD24 protein or mRNA expression levels). Thus, we believe that CRISPR/Cas9-mediated CD24 knockout is a preferable strategy for these aims. Finally, one should consider the use of multiple colorectal cancer cell lines, since the expression of some CD24-related molecules could be specific to HT-29 cells.

Author contributions

Conceptualization: S.N., A.A., T.L., D.M., J.M., U.S., A.T. Bioinformatic analysis: S.N., A.G., M.R. *In vitro* experiments: A.A., T.L., D.W., S.G., H.M., A.S. Patient materials and immunohistochemistry: H.B., O.E., L.C. Writing – Original Draft: S.N., A.A., L.C. Writing – Review & Editing: S.N., A.A., T.L., D.W., A.G., H.B., O.E., L.C., S.G., H.M., A.S., D.M., M.R., J.M., U.S., A.T.

Funding

The study was supported by the Russian Science Foundation (project 17-14-01338).

Declaration of competing interest

None.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.biochi.2021.10.004.

References

[1] S. Aigner, Z.M. Sthoeger, M. Fogel, E. Weber, J. Zarn, M. Ruppert, Y. Zeller, D. Vestweber, R. Stahel, M. Sammar, P. Altevogt, CD24, a mucin-type glycoprotein, is a ligand for P-selectin on human tumor cells, Blood 89 (1997) 3385–3395. http://www.ncbi.nlm.nih.gov/pubmed/9129046.

[2] A. Jaggupilli, E. Elkord, Significance of CD44 and CD24 as cancer stem cell markers: an enduring ambiguity, Clin. Dev. Immunol. (2012) 1–11, https://doi.org/10.1155/2012/708036, 2012.

[3] A.A. Barkal, R.E. Brewer, M. Markovic, M. Kowarsky, S.A. Barkal, B.W. Zaro, V. Krishnan, J. Hatakeyama, O. Dorigo, L.J. Barkal, I.L. Weissman, CD24 signalling through macrophage Siglec-10 is a target for cancer immunotherapy, Nature 572 (2019) 392–396, https://doi.org/10.1038/s41586-019-1456-0.

[4] S.-S. Yin, F.-H. Gao, Molecular mechanism of tumor cell immune escape mediated by CD24/Siglec-10, Front. Immunol. 11 (2020), https://doi.org/10.3389/fimmu.2020.01324.

[5] G. Kristiansen, K.-J. Winzer, E. Mayordomo, J. Bellach, K. Schlüns, C. Denkert, E. Dahl, C. Pilarsky, P. Altevogt, H. Guski, M. Dietel, CD24 expression is a new prognostic marker in breast cancer, Clin. Canc. Res. 9 (2003) 4906–4913. PMID: 14581365.

[6] P. Athanassiadou, D. Grapsa, M. Gonidi, A.-M. Athanassiadou, A. Tsipis, E. Patsouris, CD24 expression has a prognostic impact in breast carcinoma, Pathol. Res. Pract. 205 (2009) 524–533, https://doi.org/10.1016/j.prp.2009.01.008.

[7] G. Kristiansen, C. Denkert, K. Schlüns, E. Dahl, C. Pilarsky, S. Hauptmann, CD24 is expressed in ovarian cancer and is a new independent prognostic marker of patient survival, Am. J. Pathol. 161 (2002) 1215–1221, https://doi.org/10.1016/S0002-9440(10)64398-2.

[8] P. Surowiak, V. Materna, I. Kaplenko, M. Spaczynski, M. Dietel, G. Kristiansen, H. Lage, M. Zabel, Unfavorable prognostic value of CD24 expression in sections from primary and relapsed ovarian cancer tissue, Int. J. Gynecol. Canc. 16 (2006) 515–521, https://doi.org/10.1111/j.1525-1438.2006.00555.x.

[9] W. Weichert, C. Denkert, M. Burkhardt, T. Gansukh, J. Bellach, P. Altevogt, M. Dietel, G. Kristiansen, Cytoplasmic CD24 expression in colorectal cancer independently correlates with shortened patient survival, Clin. Cancer Res. 11 (2005) 6574–6581, https://doi.org/10.1158/1078-0432.CCR-05-0606.

[10] E. Sagiv, L. Memeo, A. Karin, D. Kazanov, J. Jacob–Hirsch, M. Mansukhani, G. Rechavi, H. Hibshoosh, N. Arber, CD24 is a new oncogene, early at the multistep process of colorectal cancer carcinogenesis, Gastroenterology 131 (2006) 630–639, https://doi.org/10.1053/j.gastro.2006.04.028.

[11] P. Altevogt, M. Sammar, L. Hüser, G. Kristiansen, Novel insights into the function of <scp>CD24</scp> : a driving force in cancer, Int. J. Cancer 148 (2021) 546–559, https://doi.org/10.1002/ijc.33249.

[12] J.-H. Lee, S.-H. Kim, E.-S. Lee, Y.-S. Kim, CD24 overexpression in cancer development and progression: a meta-analysis, Oncol. Rep. 22 (2009) 1149–1156, https://doi.org/10.3892/or_00000548.

[13] C. Sheridan, H. Kishimoto, R.K. Fuchs, S. Mehrotra, P. Bhat-Nakshatri, C.H. Turner, R. Goulet, S. Badve, H. Nakshatri, CD44+/CD24-breast cancer cells exhibit enhanced invasive properties: an early step necessary for metastasis, Breast Cancer Res. 8 (2006) R59, https://doi.org/10.1186/bcr1610.

[14] B.K. Abraham, P. Fritz, M. McClellan, P. Hauptvogel, M. Athelogou, H. Brauch, Prevalence of CD44+/CD24-/low cells in breast cancer may not be associated with clinical outcome but may favor distant metastasis, Clin. Cancer Res. 11 (2005) 1154–1159. http://www.ncbi.nlm.nih.gov/pubmed/15709183.

[15] K. Taniuchi, I. Nishimori, M.A. Hollingsworth, Intracellular CD24 inhibits cell invasion by posttranscriptional regulation of BART through interaction with G3BP, Cancer Res. 71 (2011) 895–905, https://doi.org/10.1158/0008-5472.CAN-10-2743.

[16] M.A.H. Ahmed, A. Al-Attar, J. Kim, N.F.S. Watson, J.H. Scholefield, L.G. Durrant, M. Ilyas, CD24 shows early upregulation and nuclear expression but is not a prognostic marker in colorectal cancer, J. Clin. Pathol. 62 (2009) 1117–1122, https://doi.org/10.1136/jcp.2009.069310.

[17] A. Nosrati, F. Naghshvar, I. Maleki, F. Salehi, Cancer stem cells CD133 and CD24 in colorectal cancers in Northern Iran, Gastroenterol. Hepatol. from Bed to Bench 9 (2016) 132–139. http://www.ncbi.nlm.nih.gov/pubmed/27099673.

[18] D.M. Muzny, M.N. Bainbridge, K. Chang, H.H. Dinh, J.A. Drummond, G. Fowler, C.L. Kovar, L.R. Lewis, M.B. Morgan, I.F. Newsham, J.G. Reid, J. Santibanez, E. Shinbrot, L.R. Trevino, Y.Q. Wu, M. Wang, P. Gunaratne, L.A. Donehower, C.J. Creighton, D.A. Wheeler, R.A. Gibbs, M.S. Lawrence, D. Voet, R. Jing, K. Cibulskis, A. Sivachenko, P. Stojanov, A. McKenna, E.S. Lander, S. Gabriel, L. Ding, R.S. Fulton, D.C. Koboldt, T. Wylie, J. Walker, D.J. Dooling, L. Fulton, K.D. Delehaunty, C.C. Fronick, R. Demeter, E.R. Mardis, R.K. Wilson, A. Chu, H.J.E. Chun, A.J. Mungall, E. Pleasance, A. Gordon Robertson, D. Stoll, M. Balasundaram, I. Birol, Y.S.N. Butterfield, E. Chuah, R.J.N. Coope, N. Dhalla, R. Guin, C. Hirst, M. Hirst, R.A. Holt, D. Lee, H.I. Li, M. Mayo, R.A. Moore, J.E. Schein, J.R. Slobodan, A. Tam, N. Thiessen, R. Varhol, T. Zeng, Y. Zhao, S.J.M. Jones, M.A. Marra, A.J. Bass, A.H. Ramos, G. Saksena, A.D. Cherniack, S.E. Schumacher, B. Tabak, S.L. Carter, N.H. Pho, H. Nguyen, R.C. Onofrio, A. Crenshaw, K. Ardlie, R. Beroukhim, W. Winckler, M. Meyerson, A. Protopopov, A. Hadjipanayis, E. Lee, R. Xi, L. Yang, X. Ren, N. Sathiamoorthy, P.C. Chen, P. Haseley, Y. Xiao, S. Lee, J. Seidman, L. Chin, P.J. Park, R. Kucherlapati, J. Todd Auman, K.A. Hoadley, Y. Du, M.D. Wilkerson, Y. Shi, C. Liquori, S. Meng, L. Li, Y.J. Turman, M.D. Topal, D. Tan, S. Waring, E. Buda, J. Walsh, C.D. Jones, P.A. Mieczkowski, D. Singh, J. Wu, A. Gulabani, P. Dolina, T. Bodenheimer, A.P. Hoyle, J.V. Simons, M. Soloway, L.E. Mose, S.R. Jefferys, S. Balu, B.D. O'Connor, J.F. Prins, D.Y. Chiang, D. Neil Hayes, C.M. Perou, T. Hinoue, D.J. Weisenberger, D.T. Maglinte, F. Pan, B.P. Berman, D.J. Van Den Berg, H. Shen, T. Triche, S.B. Baylin, P.W. Laird, G. Getz, M. Noble, D. Voat, N. Gehlenborg, D. Dicara, J. Zhang, H. Zhang, C.J. Wu, S.Y. Liu, S. Shukla, L. Zhou, P. Lin, R.W. Park, M.D. Nazaire, J. Robinson, H. Thorvaldsdottir, J. Mesirov, V. Thorsson, S.M. Reynolds, B. Bernard, R. Kreisberg, J. Lin, L. Iype, R. Bressler, T. Erkkilä, M. Gundapuneni, Y. Liu, A. Norberg, T. Robinson, D. Yang, W. Zhang, I. Shmulevich, J.J. De Ronde, N. Schultz, E. Cerami, G. Ciriello, A.P. Goldberg, B. Gross, A. Jacobsen, J. Gao, B. Kaczkowski, R. Sinha, B. Arman Aksoy, Y. Antipin, B. Reva, R. Shen, B.S. Taylor, M. Ladanyi, C. Sander, R. Akbani, N. Zhang, B.M. Broom, T. Casasent, A. Unruh, C. Wakefield, S.R. Hamilton, R. Craig Cason, K.A. Baggerly, J.N. Weinstein, D. Haussler, C.C. Benz, J.M. Stuart, S.C. Benz, J. Zachary Sanborn, C.J. Vaske, J. Zhu, C. Szeto, G.K. Scott, C. Yau, S. Ng, T. Goldstein, K. Ellrott, E. Collisson, A.E. Cozen, D. Zerbino, C. Wilks, B. Craft, P. Spellman, R. Penny, T. Shelton, M. Hatfield, S. Morris, P. Yena, C. Shelton, M. Sherman, J. Paulauskis, J.M. Gastier-Foster, J. Bowen, N.C. Ramirez, A. Black, R. Pyatt, L. Wise, P. White, M. Bertagnolli, J. Brown, T.A. Chan, G.C. Chu, C. Czerwinski, F. Denstman, R. Dhir, A. Dörner, C.S. Fuchs, J.G. Guillem, M. Iacocca, H. Juhl, A. Kaufman, B.K. Iii, X. Van Le, M.C. Mariano, E.N. Medina, M. Meyers, G.M. Nash, P.B. Paty, N. Petrelli, B. Rabeno, W.G. Richards, D. Solit, P. Swanson, L. Temple, J.E. Tepper, R. Thorp, E. Vakiani, M.R. Weiser, J.E. Willis, G. Witkin, Z. Zeng, M.J. Zinner, C. Zornig, M.A. Jensen, R. Sfeir, A.B. Kahn, A.L. Chu, P. Kothiyal, Z. Wang, E.E. Snyder, J. Pontius, T.D. Pihl, B. Ayala, M. Backus, J. Walton, J. Whitmore, J. Baboud, D.L. Berton, M.C. Nicholls, D. Srinivasan, R. Raman, S. Girshik, P.A. Kigonya, S. Alonso, R.N. Sanbhadti, S.P. Barletta, J.M. Greene, D.A. Pot, K.R.M. Shaw, L.A.L. Dillon, K. Buetow, T. Davidsen, J.A. Demchok, G. Eley, M. Ferguson, P. Fielding, C. Schaefer, M. Sheth, L. Yang, M.S. Guyer, B.A. Ozenberger, J.D. Palchik, J. Peterson, H.J. Sofia, E. Thomson, Comprehensive molecular characterization of human colon and rectal cancer, Nature 487 (2012) 330–337, https://doi.org/10.1038/nature11252.

[19] J.-R. Lee, C.H. Kwon, Y. Choi, H.J. Park, H.S. Kim, H.-J. Jo, N. Oh, D.Y. Park,

Transcriptome analysis of paired primary colorectal carcinoma and liver metastases reveals fusion transcripts and similar gene expression profiles in primary carcinoma and liver metastases, BMC Canc. 16 (2016) 539, https://doi.org/10.1186/s12885-016-2596-3.

[20] A. Dobin, C.A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, T.R. Gingeras, STAR: ultrafast universal RNA-seq aligner, Bioinformatics 29 (2013) 15–21, https://doi.org/10.1093/bioinformatics/bts635.

[21] A. Frankish, M. Diekhans, A.M. Ferreira, R. Johnson, I. Jungreis, J. Loveland, J.M. Mudge, C. Sisu, J. Wright, J. Armstrong, I. Barnes, A. Berry, A. Bignell, S. Carbonell Sala, J. Chrast, F. Cunningham, T. Di Domenico, S. Donaldson, I.T. Fiddes, C. García Girón, J.M. Gonzalez, T. Grego, M. Hardy, T. Hourlier, T. Hunt, O.G. Izuogu, J. Lagarde, F.J. Martin, L. Martínez, S. Mohanan, P. Muir, F.C.P. Navarro, A. Parker, B. Pei, F. Pozo, M. Ruffier, B.M. Schmitt, E. Stapleton, M.M. Suner, I. Sycheva, B. Uszczynska-Ratajczak, J. Xu, A. Yates, D. Zerbino, Y. Zhang, B. Aken, J.S. Choudhary, M. Gerstein, R. Guigó, T.J.P. Hubbard, M. Kellis, B. Paten, A. Reymond, M.L. Tress, P. Flicek, GENCODE reference annotation for the human and mouse genomes, Nucleic Acids Res. 47 (2019) D766–D773, https://doi.org/10.1093/nar/gky955.

[22] M.D. Robinson, D.J. McCarthy, G.K. Smyth, edgeR, A Bioconductor package for differential expression analysis of digital gene expression data, Bioinformatics 26 (2009) 139–140, https://doi.org/10.1093/bioinformatics/btp616.

[23] M.I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2, Genome Biol. 15 (2014), https://doi.org/10.1186/s13059-014-0550-8.

[24] A. Zhu, J.G. Ibrahim, M.I. Love, Heavy-Tailed prior distributions for sequence count data: removing the noise and preserving large differences, Bioinformatics 35 (2019) 2084–2092, https://doi.org/10.1093/bioinformatics/bty895.

[25] W. Yang, P.C. Rosenstiel, H. Schulenburg, ABSSeq: a new RNA-Seq analysis method based on modelling absolute expression differences, BMC Genom. 17 (2016) 541, https://doi.org/10.1186/s12864-016-2848-2.

[26] N.E. Ilott, T. Schneider, J. Mill, L. Schalkwyk, G. Brolese, L. Bizarro, I.P. Stolerman, E. Dempster, P. Asherson, Long-term effects of gestational nicotine exposure and food-restriction on gene expression in the striatum of adolescent rats, PLoS One 9 (2014), e88896, https://doi.org/10.1371/journal.pone.0088896.

[27] D.W. Huang, B.T. Sherman, R.A. Lempicki, Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources, Nat. Protoc. 4 (2009) 44–57, https://doi.org/10.1038/nprot.2008.211.

[28] O. Fornes, J.A. Castro-Mondragon, A. Khan, R. van der Lee, X. Zhang, P.A. Richmond, B.P. Modi, S. Correard, M. Gheorghe, D. Baranašić, W. Santana-Garcia, G. Tan, J. Chèneby, B. Ballester, F. Parcy, A. Sandelin, B. Lenhard, W.W. Wasserman, A. Mathelier, JASPAR 2020: update of the open-access database of transcription factor binding profiles, Nucleic Acids Res. 48 (D1) (2019) D87–D92, https://doi.org/10.1093/nar/gkz1001.

[29] Q. Zhang, W. Liu, H.-M. Zhang, G.-Y. Xie, Y.-R. Miao, M. Xia, A.-Y. Guo, hTFtarget: a comprehensive database for regulations of human transcription factors and their targets, genomics, Proteomics Bioinform. 18 (2020) 120–128, https://doi.org/10.1016/j.gpb.2019.09.006.

[30] A.-K. Ernst, A. Putscher, T.R. Samatov, A. Suling, V.V. Galatenko, M.Y. Shkurnikov, E.N. Knyazev, A.G. Tonevitsky, T. Haalck, T. Lange, H. Maar, J. Schröder- Schwarz, K. Riecken, U. Schumacher, D. Wicklein, Knockdown of L1CAM significantly reduces metastasis in a xenograft model of human melanoma: L1CAM is a potential target for anti-melanoma therapy, PLoS One 13 (2018), e0192525, https://doi.org/10.1371/journal.pone.0192525.

[31] K. Weber, U. Mock, B. Petrowitz, U. Bartsch, B. Fehse, Lentiviral gene ontology (LeGO) vectors equipped with novel drug-selectable fluorescent proteins: new building blocks for cell marking and multi-gene analysis, Gene Ther. 17 (2010) 511–520, https://doi.org/10.1038/gt.2009.149.

[32] W.R. Beyer, M. Westphal, W. Ostertag, D. von Laer, Oncoretrovirus and lentivirus vectors pseudotyped with lymphocytic choriomeningitis virus glycoprotein: generation, concentration, and broad host range, J. Virol. 76 (2002) 1488–1495, https://doi.org/10.1128/JVI.76.3.1488-1495.2002.

[33] S.A. Bustin, V. Benes, J.A. Garson, J. Hellemans, J. Huggett, M. Kubista, R. Mueller, T. Nolan, M.W. Pfaffl, G.L. Shipley, J. Vandesompele, C.T. Wittwer, The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments, Clin. Chem. 55 (2009) 611–622, https://doi.org/10.1373/clinchem.2008.112797.

[34] K.J. Livak, T.D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method, Methods 25 (2001) 402–408, https://doi.org/10.1006/meth.2001.1262.

[35] D. Jackson, R. Waibel, E. Weber, J. Bell, R.A. Stahel, CD24, a signal-transducing molecule expressed on human B cells, is a major surface antigen on small cell lung carcinomas, Canc. Res. 52 (1992) 5264–5270. http://www.ncbi.nlm.nih.gov/pubmed/1327504.

[36] P. Virtanen, R. Gommers, T.E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S.J. van der Walt, M. Brett, J. Wilson, K.J. Millman, N. Mayorov, A.R.J. Nelson, E. Jones, R. Kern, E. Larson, C.J. Carey, İ. Polat, Y. Feng, E.W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E.A. Quintero, C.R. Harris, A.M. Archibald, A.H. Ribeiro, F. Pedregosa, P. van Mulbregt, A. Vijaykumar, A. Pietro Bardelli, A. Rothberg, A. Hilboll, A. Kloeckner, A. Scopatz, A. Lee, A. Rokem, C.N. Woods, C. Fulton, C. Masson, C. Häggström, C. Fitzgerald, D.A. Nicholson, D.R. Hagen, D.V. Pasechnik, E. Olivetti, E. Martin, E. Wieser, F. Silva, F. Lenders, F. Wilhelm, G. Young, G.A. Price, G.L. Ingold, G.E. Allen, G.R. Lee, H. Audren, I. Probst, J.P. Dietrich, J. Silterra, J.T. Webber, J. Slavič, J. Nothman, J. Buchner, J. Kulick, J.L. Schönberger, J.V. de Miranda Cardoso, J. Reimer, J. Harrington, J.L.C. Rodríguez, J. Nunez-Iglesias, J. Kuczynski, K. Tritz, M. Thoma, M. Newville, M. Kümmerer, M. Bolingbroke, M. Tartre, M. Pak, N.J. Smith, N. Nowaczyk, N. Shebanov, O. Pavlyk, P.A. Brodtkorb, P. Lee, R.T. McGibbon, R. Feldbauer, S. Lewis, S. Tygier, S. Sievert, S. Vigna, S. Peterson, S. More, T. Pudlik, T. Oshima, T.J. Pingel, T.P. Robitaille, T. Spura, T.R. Jones, T. Cera, T. Leslie, T. Zito, T. Krauss, U. Upadhyay, Y.O. Halchenko, Y. Vázquez-Baeza, SciPy 1.0: fundamental algorithms for scientific computing in Python, Nat. Methods 17 (2020) 261–272, https://doi.org/10.1038/s41592-019-0686-2.

[37] D. Schaeffer, J.A. Somarelli, G. Hanna, G.M. Palmer, M.A. Garcia-Blanco, Cellular migration and invasion uncoupled: increased migration is not an inexorable consequence of epithelial-to-mesenchymal transition, Mol. Cell Biol. 34 (2014) 3486–3499, https://doi.org/10.1128/MCB.00694-14.

[38] D. Choi, H.W. Lee, K.Y. Hur, J.J. Kim, G.-S. Park, S.-H. Jang, Y.S. Song, K.-S. Jang, S.S. Paik, Cancer stem cell markers CD133 and CD24 correlate with invasiveness and differentiation in colorectal adenocarcinoma, World J. Gastroenterol. 15 (2009) 2258, https://doi.org/10.3748/wjg.15.2258.

[39] M.-K. Yeo, Y.-M. Lee, I.-O. Seong, S.-Y. Choi, K.-S. Suh, K.S. Song, C.-S. Lee, J.M. Kim, K.-H. Kim, Up-regulation of cytoplasmic CD24 expression is associated with malignant transformation but favorable prognosis of colorectal adenocarcinoma, Anticanc. Res. 36 (2016) 6593–6598, https://doi.org/10.21873/anticanres.11264.

[40] J.-R. Wang, B. Liu, L. Zhou, Y.-X. Huang, MicroRNA-124-3p suppresses cell migration and invasion by targeting ITGA3 signaling in bladder cancer, Cancer Biomarkers 24 (2019) 159–172, https://doi.org/10.3233/CBM-182000.

[41] K. Koshizuka, T. Hanazawa, N. Kikkawa, T. Arai, A. Okato, A. Kurozumi, M. Kato, K. Katada, Y. Okamoto, N. Seki, Regulation of ITGA3 by the anti-tumor miR-199 family inhibits cancer cell migration and invasion in head and neck cancer, Cancer Sci. 108 (2017) 1681–1692, https://doi.org/10.1111/cas.13298.

[42] P. Rousselle, J.Y. Scoazec, Laminin 332 in cancer: when the extracellular matrix turns signals from cell anchorage to cell movement, Semin. Cancer Biol. 62 (2020) 149–165, https://doi.org/10.1016/j.semcancer.2019.09.026.

[43] D.V. Maltseva, S.A. Rodin, Laminins in metastatic cancer, Mol. Biol. 52 (2018) 350–371, https://doi.org/10.1134/S0026893318030093.

[44] D. Luo, W. Ge, MeCP2 promotes colorectal cancer metastasis by modulating ZEB1 transcription, Cancers (Basel) 12 (2020) 758, https://doi.org/10.3390/cancers12030758.

[45] S.A. Turkistany, R.P. DeKoter, The transcription factor PU.1 is a critical regulator of cellular communication in the immune system, Arch. Immunol. Ther. Exp. (Warsz). 59 (2011) 431–440, https://doi.org/10.1007/s00005-011-0147-9.

[46] C. Strell, F. Entschladen, Extravasation of leukocytes in comparison to tumor cells, Cell Commun. Signal. 6 (2008) 10, https://doi.org/10.1186/1478-811X-6-10.

[47] A. Tanaka, Y. Zhou, M. Ogawa, J. Shia, D.S. Klimstra, J.Y. Wang, M.H. Roehrl, STAT1 as a potential prognosis marker for poor outcomes of early stage colorectal cancer with microsatellite instability, PLoS One 15 (2020), e0229252, https://doi.org/10.1371/journal.pone.0229252.