The regulatory role of the CD2AP/TKS4 interaction in EMT and its potential as a biomarker for colon cancer

Anita Kurilla [1*], Loretta László [1,2], Tamás Takács [1,2], Álmos Tilajka [1,2], Laura Lukács[1], Julianna Novák[1], Rita Pancsa[1], László Buday [1], Virág Vas [1*]

Affiliations

[1] Institute of Enzymology, Research Centre for Natural Sciences, 1117 Budapest, Hungary

[2] Doctoral School of Biology, Institute of Biology, ELTE Eötvös Loránd University,1117 Budapest, Hungary

*** Correspondence:**

Anita Kurilla

kurilla.anita@ttk.hu

Virág Vas

vas.virag@ttk.hu

Abstract

Colon cancer is a leading cause of death worldwide. Although several molecular biomarkers have been developed for colon cancer, more sensitive and specific methods are needed for its early detection. The TKS4 and CD2AP scaffold proteins have been linked to dynamic actin assembly-related processes, such as cell migration and metastasis, although their co-instructive role during colon cancer development remains unknown. Therefore, this study was designed to identify the TKS4-interacting proteins and to investigate the role of the newly discovered TKS4-interacting CD2AP adaptor protein in colon cancer. Furthermore, we studied the interdependence of TKS4 and CD2AP during colon cancer progression as well as their involvement in EMT (epithelial-mesenchymal transition). This is the first study demonstrating a TKS4-CD2AP protein-protein interaction in vitro and in vivo. We describe that CD2AP and TKS4 interdependently inhibit a mesenchymal transition-like process and impair the migration ability of HCT116 colon cancer cells. Here we show that CD2AP is a novel TKS4 interacting partner, and together they form a scaffold complex that potentially regulates EMT-related signalling. Based on our analysis of the relative TKS4 and CD2AP expression levels, we propose that their levels might cooperatively serve as a reliable biomarker for colon cancer detection.

Keywords: scaffold proteins, CD2AP, TKS4, epithelial mesenchymal transition, colon cancer biomarker

1. Introduction

Colon (colorectal) cancer is one of the most common cancer types and is the second leading cause of cancer death worldwide. Mortality and survival rates mostly depend on appropriate screening and treatment techniques. As these methods have

improved in the last decade, death rates have dropped significantly (seer.cancer.gov). Currently, several blood-based and molecular biomarkers have been developed and used clinically, including the presence of *KRAS* mutations as well as the presence of circulating tumour DNA (ctDNA) or circulating tumour cells (CTCs). However, more sensitive and specific methods are needed to detect colon cancer at its early stages (1).

Several molecular events are known to drive colon cancer formation, including the accumulation of oncogenic mutations, development of drug resistance, and initiation of the cellular invasion cascade, leading to the appearance of CTCs and metastasis (2). However, most of the molecular mechanisms underlying oncogenesis are not yet fully understood; therefore, identifying novel molecular players that regulate colon cancer invasiveness will help us understand its formation and progression, ultimately contributing to the discovery of novel biomarkers and therapeutic targets. The EMT process is the central regulator of metastasis during which epithelial cells transform to mesenchymal state (3). Although EMT has been generally considered as a binary process, lately it is proven that rather a gradual transformation event takes place during transition. In this spectrum, EMT includes an intermediate partial EMT (pEMT) state in which cells maintain both epithelial and mesenchymal phenotypes, implying that co-expression of epithelial and mesenchymal markers can be observed (4). It is known that tumor cells can be heterogenous with epithelial, mesenchymal, and hybrid E/M (epithelial/mesenchymal) phenotypes simultaneously. Likewise, tumors with pEMT phenotype can also show intra-tumoral heterogeneity and epithelial-mesenchymal plasticity. The pEMT process includes different steps, such as early, intermediate, and late states that cover the spectrum from epithelial to completely mesenchymal phenotype (5). Furthermore,

pEMT has a high clinical significance as it was associated with worse prognosis and higher metastatic risk (6) (7). Presently, there is no clear consensus in defining pEMT phenotype, however, co-expression of different epithelial and mesenchymal markers is considered as an intermediate cell state (8).

CD2-associated protein (CD2AP) and tyrosine kinase substrate with four SH3 domains (TKS4) have recently emerged as potential cancer formation-regulating factors. These proteins, which lack enzymatic activity, are scaffolds/adaptors that serve as molecular hubs within cells by binding and releasing multiple signalling molecules via phosphorylation-dependent or -independent interactions mainly mediated by SH3 domains. CD2AP was originally implicated in dynamic actin remodelling via direct binding of filamentous actin. Later, it was found that CD2AP also plays a regulatory role in epithelial cell junction formation (9), and it was also linked to gastric cancer development as it can inhibit metastasis by promoting cellular adhesion and cytoskeleton assembly (10). In a previous study, Bazalii et al. (11) have already demonstrated that CIN85 (also called SH3KBP1), a close paralog of CD2AP, interacts with TKS4 through its three SH3 domains; however, they did not identify the exact interacting residues within TKS4. CD2AP and CIN85 have three SH3 domains each, which are special in the sense that they recognize non-canonical PxP/A/V/IxPR motifs instead of the canonical PxxP SH3 binding sites (12) (13).

TKS4 also serves as a docking site for molecules involved in cell movement, e.g., cortactin or N-WASP, and it regulates podosome formation in normal cells and invadopodia formation in cancer cells (14)(15). The involvement of TKS4 in EGF-EGFR-SRC signalling has been previously described, and regulatory roles in mesenchymal stem cell differentiation, adipose tissue, and bone homeostasis have recently been discovered (16)(17)(18). Recent progress in TKS4-related research

has revealed that TKS4 might also be involved in several cancer types, including breast cancer (19), melanoma (15), acute myeloid leukaemia (20), and prostate cancer (21). Moreover, it was recently studied the effect of TKS4 in a colon cancer cell line and they concluded that reduced level of TKS4 leads to disturbed EMT-like processes (22).

Here we show that CD2AP and TKS4 can form a complex in several different cancer cell lines. The fact that both proteins have been implicated in colon cancer cell biology raises the possibility that their interaction is fundamental to colon cancer development; therefore, we analysed both the individual and relative expression levels of CD2AP and TKS4 in colon cancer samples to shed light on the potential effects of their interplay during cancer development. Moreover, we used siRNA-based silencing to downregulate their transcript levels in HCT116 colon cancer cells and measured the resulting changes in EMT process-related gene expression levels and cell migration potential.

We found that CD2AP and TKS4 individually and interdependently inhibit pEMT and impair colon cancer cell migration. We found that the ratio of CD2AP and TKS4 expression levels has an instructive role in tumour phenotypes and that CD2AP/TKS4 ratios are outstanding biomarkers for the identification of tumorous colon tissue.

2. Materials and Methods

2.1. Cell culture

All cell lines were obtained from ATCC and were tested for mycoplasma contamination using MycoAlert Mycoplasma Detection Kit (Lonza, Basel, Switzerland). HCT116 (RRID:CVCL_0291) cells were grown in McCoy's 5A Medium

(Thermo Fisher Scientific, 16600082). DLD-1 (RRID:CVCL_0248) cells were grown in RPMI medium. MCF-7 (RRID:CVCL_0031), HEK293 (RRID:CVCL_0045), HPAC (RRID:CVCL_3517), and NIH3T3 (RRID:CVCL_0594) cells were grown in DMEM. All media were supplemented with 10% FBS (Thermo Fisher Scientific, 26400044) and penicillin/streptomycin, and cells were incubated at 37 °C in 5% CO_2.

2.2. siRNA and DNA plasmid transfection

Small interfering RNA (siRNA) transfection was performed with Lipofectamine RNAiMax Transfection Reagent (Thermo Fisher Scientific, 13778030) according to the manufacturer's instructions. SiRNAs targeting CD2AP and TKS4 were purchased from Thermo Fisher Scientific and used at 10 nM. After transfection, cells were incubated for 48 h prior to RNA or protein isolation.

For DNA plasmid transfection, V5-TKS4 or/and MYC-CD2AP (Origene Technologies Inc., Rockville, MD)) plasmids were transfected using X-tremeGENE Hp DNA Transfection Reagent (Roche) following the manufacturer's instructions. Cells were incubated with transfection reagent for 48 h before cell lysis.

2.3. RNA isolation and quantitative real-time polymerase chain reaction (qRT-PCR)

RNA was extracted using Direct-zol RNA Miniprep Kit (Zymo Research). First Strand cDNA Synthesis Kit (Roche) and Fast SYBR Green Master Mix (Thermo Fisher Scientific) were used in qRT-PCR. Data were normalized to the *GAPDH* and *PUM1* transcript levels, and the $2^{-\Delta\Delta Ct}$ method was used to calculate relative gene expression levels. Primers used in this study are listed in Supplementary Table 1.

2.4. Western blotting and immunoprecipitation

For protein extraction, cells were first washed with ice-cold phosphate-buffered saline (PBS). Next, ice-cold harvest buffer (30 mM Tris pH 7.5, with 100 mM NaCl, 1% Triton X-100, 10 mM NaF, 1 mM Na_3VO_4, 1 mM EGTA, 2 mM 4-nitrophenyl phosphate, 10 mM benzamidine, 1 mM phenylmethylsulphonyl fluoride, 25 µg/mL Pepstatin A, 25 µg/mL trypsin inhibitor, and 25 µg/mL aprotinin) was added to the cells. Lysates were centrifuged at 14,000 rpm, 10 min at 4 °C.

Immunoprecipitation was performed using Protein-A Sepharose from *Staphylococcus aureus* (Merck Millipore) and anti-TKS4 antibody (23). Cell lysates, agarose beads, and anti-TKS4 antibody were incubated for 1 hour at 4 °C and then washed three times with PBS+0.4% Triton-X at 14000 rpm for 1 min at 4 °C.

Sample loading buffer was added to the supernatants, followed by incubation at 95 °C for 5 min. Each protein sample was subjected to 10% SDS-PAGE gel (BioRad). Nitrocellulose membranes were blocked and incubated for 60 min with anti-TKS4 (1:5000) and anti-CD2AP (1:2000) antibody at room temperature or overnight at 4 °C with anti-E-cadherin, or anti-vimentin antibody. After three washes, the membranes were incubated for 60 min with horseradish peroxidase-conjugated secondary antibody (GE Healthcare) and then washed four times for 15 min each. The proteins of interest were visualized using enhanced chemiluminescence (ECL) detection reagents (Amersham, Little Chalfont, UK). Chemiluminescent imaging was performed with a ChemiDoc MP system (Bio-Rad). The analysed protein levels were normalized to total protein staining.

2.5. Mass spectrometry analyses

After immunoprecipitation, samples were loaded onto 10% SDS-PAGE gels. Next, proteins were visualized via Coomassie Brilliant Blue dye, and gels were fractionized

horizontally based on molecular weight (300-100 kDa, 100-50 kDa, 50-25 kDa, 25-10 kDa). Mass spectrometry analyses were performed by UD-GenoMed Medical Genomic Technologies Kft (Debrecen, Hungary).

2.6. Wound healing assays

HCT116 cells were forward-transfected with siRNAs or plasmid DNA. Next, $5x10^5$ cells were cultured in Ibidi Culture-Insert well. After 24 hours, cells reached confluency and the silicone inserts were removed. The dishes were filled with McCoy's Medium, and images were acquired every 24 hours.

2.7. DuoLink Proximity Ligation Assays

In situ protein-protein interactions between CD2AP and TKS4 were detected via proximity ligation assays (Duolink Mouse/Rabbit, DUO92101−1KT, Sigma-Aldrich). The DPLA probe anti-rabbit plus binds to the TKS4 primary antibody (HPA036471, Sima Aldrich, 1:100 dilution), whereas the DPLA probe anti-mouse minus binds to the CD2AP antibody (clone 2A2.1, Millipore, 1:250 dilution). The DPLA secondary antibodies generate a signal only when the two DPLA probes interact, indicating a CD2AP-TKS4 interaction. HCT116 cells were plated in 12-well Ibidi removable silicone cultivation chambers and cultured for 2 days. Next, the cells were fixed in 4% paraformaldehyde-PBS for 15 min on ice and further processed according to the manufacturer's instructions: permeabilization with 0.1% Triton X-100 dissolved in PBS for 5 min and blocking with Duolink Blocking Solution for 30 min at 37 °C. The secondary antibodies were conjugated to oligonucleotides that form a closed DNA circle via base pairing when the antibodies are in close proximity. Signal detection was conducted by ligation using Duolink ligase and rolling circle amplification using DNA polymerase and fluorescently labelled nucleotides hybridized into the

amplification products. The amplified fluorescent DNA resulted in bright red dots in the samples, with each dot representing an individual CD2AP/TKS4 interaction event. Control samples were processed similarly without addition of the primary antibodies. Staining and image acquisition were performed as previously described (22). Briefly, nuclei were visualized via DAPI staining (with DuoLink mounting medium), and images were acquired on a Zeiss LSM710 inverted confocal microscope with a 40× objective (Carl Zeiss).

2.8. Peptide array

PepStar Peptide Microarray experiments were performed based on the method of Harnos and colleagues (2018) with some modifications (24). Microarrays were purchased from JPT (JPT Peptide Technologies GmbH). The TKS4 protein sequence (A1X283) was printed on a glass slide (25x75 mm) in the form of 15-mer peptides with 11 overlapping residues. The peptide microarrays were printed in three identical subarrays. Several controls were also printed on the glass slides to exclude non-specific binding events or to confirm positive signals (Myc-tag, anti-Myc antibody, human serum albumin, human IgG, rabbit IgG). When screening the binding ability of CD2AP to TKS4 peptides, Myc-tagged CD2AP protein was used. Myc-DDK tagged human CD2-associated protein (RC210191) was purchased from Origene Technologies. The screening experiments were performed using two glass slides. As a control, a slide was incubated with Alexa Fluor® 488-conjugated Myc-Tag (71D10) rabbit monoclonal antibody purchased from CST (Cell Signaling Technology). For the experimental slides, peptides were incubated with Myc-CD2AP recombinant protein. The screening experiments and signal visualization were performed by Diagnosticum Zrt using a sandwich-like format in a 300 ul reaction volume in an incubation chamber. First, slides were blocked with SmartBlock solution (CANDOR Bioscience,

113125) for 1 h at 30 °C. Next, the experimental slides were incubated with recombinant Myc-CD2AP protein for 1 h at 30 °C in SmartBlock solution in a final volume of 300 ul. Slides were washed four times in 300 ul TBS buffer for 10 min at room temperature then incubated with 1 ug/ml Myc-tag Alexa Fluor 488 antibody in SmartBlock solution in a final volume of 300 ul volume for 1 h at 30 °C. To remove unbound fluorescent antibodies, slides were washed four times with TBS buffer for 10 min at room temperature, and the slides were dried via centrifugation in 50 ml falcon tubes at 1200 rcf for 2 min. Recombinant protein binding was detected by measuring the fluorescence intensity of the peptide spots using a Molecular Devices Axon Genepix 4300A laser scanner. The results were analysed with PepSlide Analyzer Software (Sicasys). Peptide sequence of TKS4 and the results of the peptide array are in Supplementary Table 2.

2.9. Tissue cDNA array and database analyses

The Colon Cancer Tissue cDNA Array IV panel (HCRT104) was purchased from Origene Technologies (Origene Technologies Inc., Rockville, MD). Each plate contains 48 samples from patients with cancer whose clinical pathological features are freely available at the following link from Origene: https://www.origene.com/catalog/tissues/tissuescan/hcrt104/tissuescan-colon-cancer-cdna-array-iv. The three cecum samples (one from a normal group and two from a tumorous group) were excluded from the analyses. qRT-PCR was performed with FastStart Green PCR Master Mix, and gene expression levels were normalized to the *GAPDH* and *Pum1* levels (25)based on the method of Hellemans and colleagues (26). Database analyses were performed using GENT2 database (27) and The Cancer Genome Atlas Colon Adenocarcinoma (TCGA-COAD) using the Xena Platform (28). *CD2AP* and *TKS4 (SH3PXD2B)* mRNA expression level data

were collected from tumour or matched normal tissues. From GENT2 database, datasets including the same nomenclatures (normal, stage I, stage II, stage III and stage IV samples) were exported. Thus, GSE21510 and GSE32323 datasets were omitted from our analyses. Statistically significant differences were calculated using Student's t-test ($p<0.05$). Analysis of variance (ANOVA) followed by the Tukey HSD (Honest Significant Difference) test were used for statistical analysis. ROC curve analyses were performed using the Wilson/Brown method. All statistical analyses were performed in GraphPad Prism 8.0.1 (GraphPad Prism, Inc. San Diego, CA). Protein interaction network of TKS4-interacting partners were created using the STRING database (29).

3. Results

3.1. CD2AP is a novel TKS4 interacting partner

To identify specific TKS4-interacting partners, a high-throughput approach was performed using immunoprecipitation (IP) and mass spectrometry (MS). We identified various novel interacting partners, including transmembrane receptors, RNA-binding proteins, phosphatases, and cytoskeletal motor proteins. Six proteins consistently appeared as TKS4-interacting partners in all analysed cancer cells, i.e., breast (MCF-7), colon (HCT 116, DLD-1), and pancreas (HPAC) cancer types, namely CD2AP, CAPZA, CAPZB, KIF5B, CCAR1, and WARS (Figure 1A). Furthermore, these proteins were also present in the precipitate in two studied non-cancerous cell lines (NIH 3T3, HEK 293). Among them, CD2AP was previously shown to be a cancer-related scaffold protein (10); thus, we focused on the CD2AP-TKS4 interaction in this study. The TKS4-CD2AP interaction was also validated by IP-western blot (WB) (Figure 1B).

To confirm the presence of the CD2AP-TKS4 interaction in colon cancer cells and to visualize the resulting complexes, proximity ligation assays (PLA) were used (DuoLink). Micrographs of the assays revealed that CD2AP and TKS4 colocalized and are in close proximity (at a distance of less than 40 nm) as indicated by the appearance of red fluorescent dots in HCT116 cells (Figure 1C).

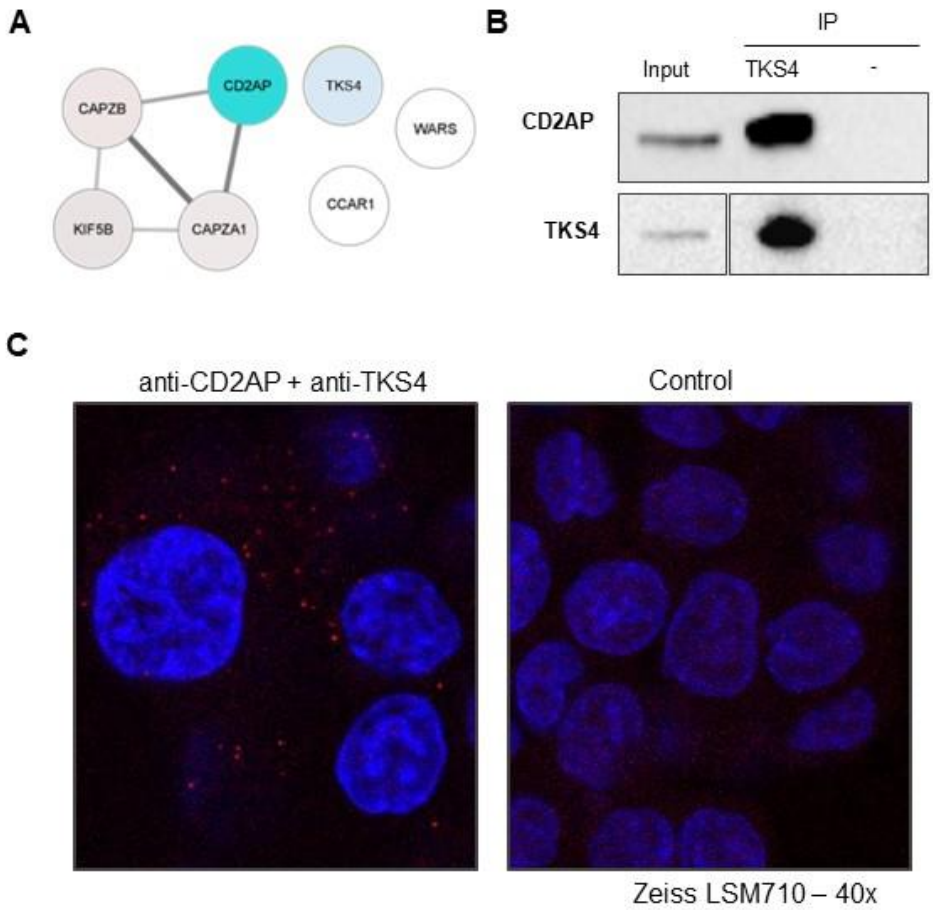


Figure 1. A combination of co-IP and MS was used to identify new TKS4-interacting partners, and the validation of CD2AP-TKS4 interaction via IP-WB and DuoLink Proximity Ligation Assay. **(A)** Six consistently detected TKS4-interacting partners and their interaction network created via the STRING database. Lines indicate previously identified interactions. **(B)** Validation of CD2AP as a TKS4-interacting partner via IP-WB. TKS4: Immunprecipitated protein lysates with TKS4-specific antibody; -:Immunpreciptated protein lysates without TKS4-specific antibody. **(C)** Detecting the CD2AP-TKS4 interaction in HCT116 cells: fluorescence signal is generated only

when the two DPLA probes are in proximity. Colocalization appears as red dots in the cytoplasm and close to the cell membrane. No colocalization was seen in control staining that did not include the primary antibodies in the DPLA reaction. Colours: red dots = CD2AP-TKS4 colocalization, blue = DAPI-stained nuclei;

Next we decided to map the interaction sites and determine the location of the CD2AP-binding sites in TKS4 by performing a Pepstar Peptide Microarray analysis. Binding of full-length CD2AP protein was screened against fragments containing 15mer peptides overlapping TKS4. Since 15mer peptides cannot preserve conformational features, we calculated the relative accessible surface area (rASA) of each amino acid of the TKS4 fragments that showed binding to see if those are accessible in the folded protein. The results of microarray analysis showed that the 188 and 189 TKS4 peptides showed the strongest interactions with the full length CD2AP. Both TKS4 peptides contain a known short linear motif (SLIM) binding site ^{755}PVVPPR^{760}; in addition, these peptides had the highest surface accessibility values (10 of 15 peptides with >50% rASA) (Figure 2). The other TKS4 peptides that bound CD2AP mainly reside within the four SH3 domains of TKS4 with low or medium accessibility values (Figure 2), thus these binding events are probably artefacts stemming from domain fragmentation and exposure of their hydrophobic cores. Based on our results, we conclude that CD2AP is a novel direct TKS4-interacting partner and that it binds to the ^{755}PVVPPR^{760} motif in TKS4 through the SH3 domains of CD2AP.

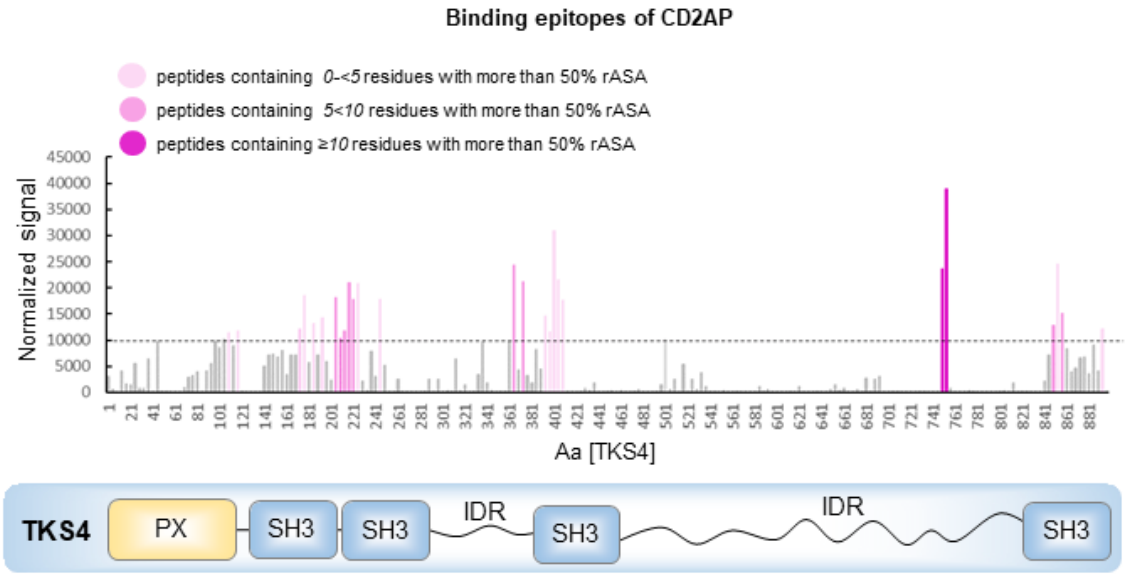


Figure 2. Identification of potential CD2AP binding sites within TKS4. Columns indicate the detected normalized fluorescent signal from the peptide microarray, and the dashed line represents the threshold signal intensity (10000 U). The X axis shows the TKS4 aminoacids numbered by the position of their first residue in the TKS4 sequence. Pink columns represent normalized signal above threshold (dashed line). The pink colour shading indicates the degree of the relative surface accessibility value (number of residues with >50% rASA) of each 15mer peptide. (*Light pink* columns represent peptides containing *0-<5* residues with more than 50% rASA, *medium pink* columns represent peptides containing *5<10* residues with more than 50% rASA, *dark pink* columns represent peptides containing ≥*10* residues with more than 50% rASA.)

3.2. Absence of CD2AP and TKS4 promote colon cancer cell migration separately and simultaneously through pEMT process

Although it was already reported that TKS4 regulates cell migration and plays a role in EMT-like process in colon cancer cells (22), and CD2AP has also been connected to cancer development and the inhibition of metastasis in gastric cancer, the mechanism how these proteins exerts their effect on EMT processes remained

elusive. Therefore, we hypothesized that the TKS4/CD2AP complex cooperatively regulates cytoskeletal rearrangements during cancer progression and participates in EMT process initiation. To investigate the cooperative role of the CD2AP/TKS4 adhesion-related adapter proteins in colon cancer progression, *CD2AP* and *TKS4* were silenced via siRNA treatment either separately or together in HCT116 colon cancer cells, and cell motility assays were performed in the temporarily modified cells. We confirmed that CD2AP and TKS4 protein levels were efficiently knocked down after the siRNA treatments (Figure 3B). To explore whether CD2AP levels modulate cell migration, as was previously reported in the absence of TKS4 protein, wound healing assays were performed. *CD2AP* silencing alone or in combination with *TKS4* silencing enhanced HCT116 cell migration compared with that of the control cells (Figure 3A).

To further examine whether increased cell migration stems from an EMT process, we measured the expression levels of EMT-related genes. In TKS4-silenced cells, epithelial marker E-cadherin expression slightly increased, while mesenchymal marker vimentin was unchanged (Figure 3B, C). In line with the EMT-related Transcription factors (TF), *Twist* and *Snai1* mRNA levels were elevated. Thus, moderate expressional changes of EMT-related genes were observed. In siCD2AP-treated cells, the RNA and protein levels of the epithelial marker E-cadherin increased (Figure 3B, C), while the mesenchymal marker vimentin decreased. Besides, we measured reduced level of *Zeb1 m*RNA than in the control cells, however, *Snai1* level was unchanged and *Twist* transcripts were upregulated (Figure 3C). In short, both epithelial-, and mesenchymal-related gene expressions were observed in CD2AP-silenced cells. The simultaneously reduced level of TKS4/CD2AP resulted in decreased E-cadherin and vimentin levels. However, all

analysed TFs (*Zeb1, Twist, Snai1*) mRNA levels were upregulated (Figure 3B, C). To conclude, different co-expressional profiles of epithelial and mesenchymal markers were observed depending upon which genes were silenced.

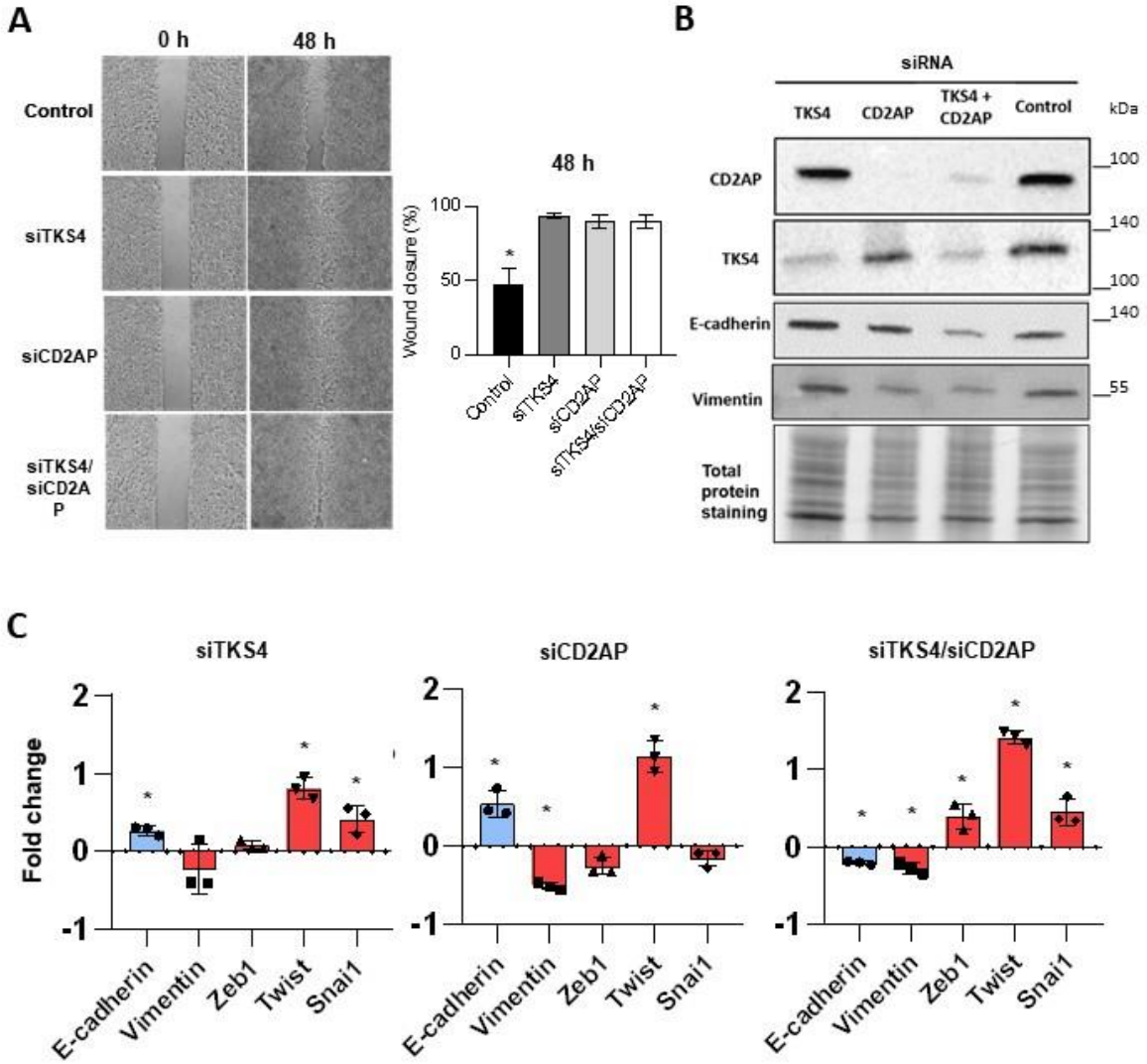


Figure 3. Depletion of *TKS4* and/or *CD2AP* expression in HCT116 colon cells. **(A)** Wound healing assay in *CD2AP*-, *TKS4*-, and *CD2AP/TKS4*-silenced HCT116 cells. Between-group comparisons were performed using ANOVA and Tukey's test, and statistical significance was set at *P<0.05. **(B)** EMT marker protein levels as measured via western blotting (E-cadherin and vimentin). **(C)** qPCR analyses of EMT markers in si*TKS4*-, si*CD2AP*- and si*CD2AP/TKS4*-treated HCT116 cells. Fold changes show the expressions relative to the mRNA level in negative control siRNA-

treated cells. P values were calculated with Student's t-test, and statistical significance was set at *P<0.05.

To test the hypothesis that CD2AP and TKS4 levels modify cancer cell motility, HCT116 cells were transduced with *CD2AP*- and/or *TKS4*-overexpressing plasmids. We confirmed that CD2AP and TKS4 expressions were efficiently enhanced after transfection (Figure 4B). The results showed that *CD2AP, TKS4* and *CD2AP/TKS4* overexpression resulted in slower cell migration (Figure 4A). In the case of EMT-related gene expressions, *CD2AP*-, *TKS4*-, and *CD2AP/TKS4*-overexpressing HCT116 cells both have normal E-cadherin and vimentin protein levels (Figure 4B). Based on these results, it appears that either TKS4 or CD2AP inhibits cell migration and can prevent the formation of partial EMT in colon cancer cells.

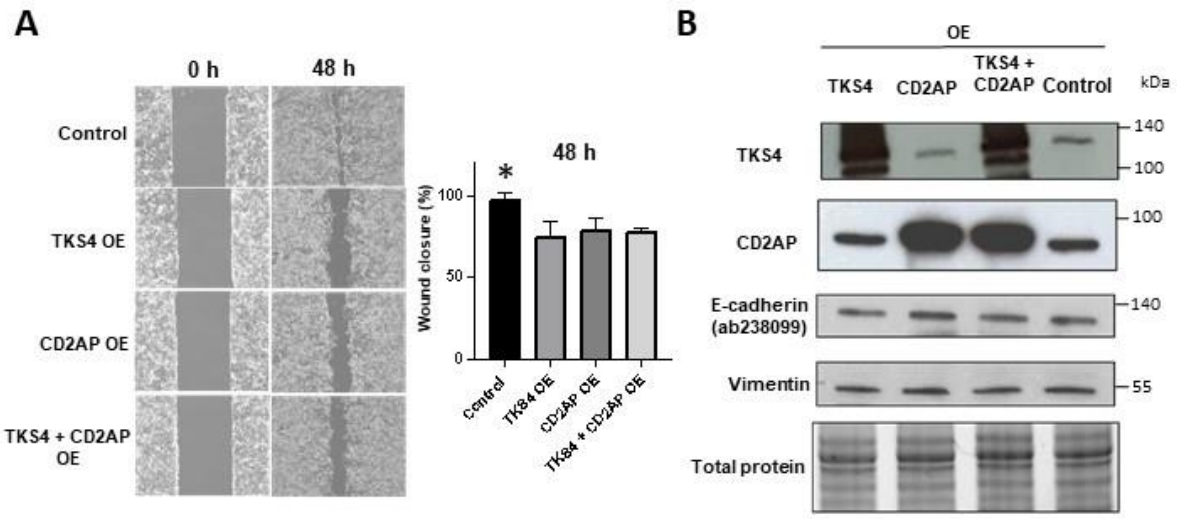
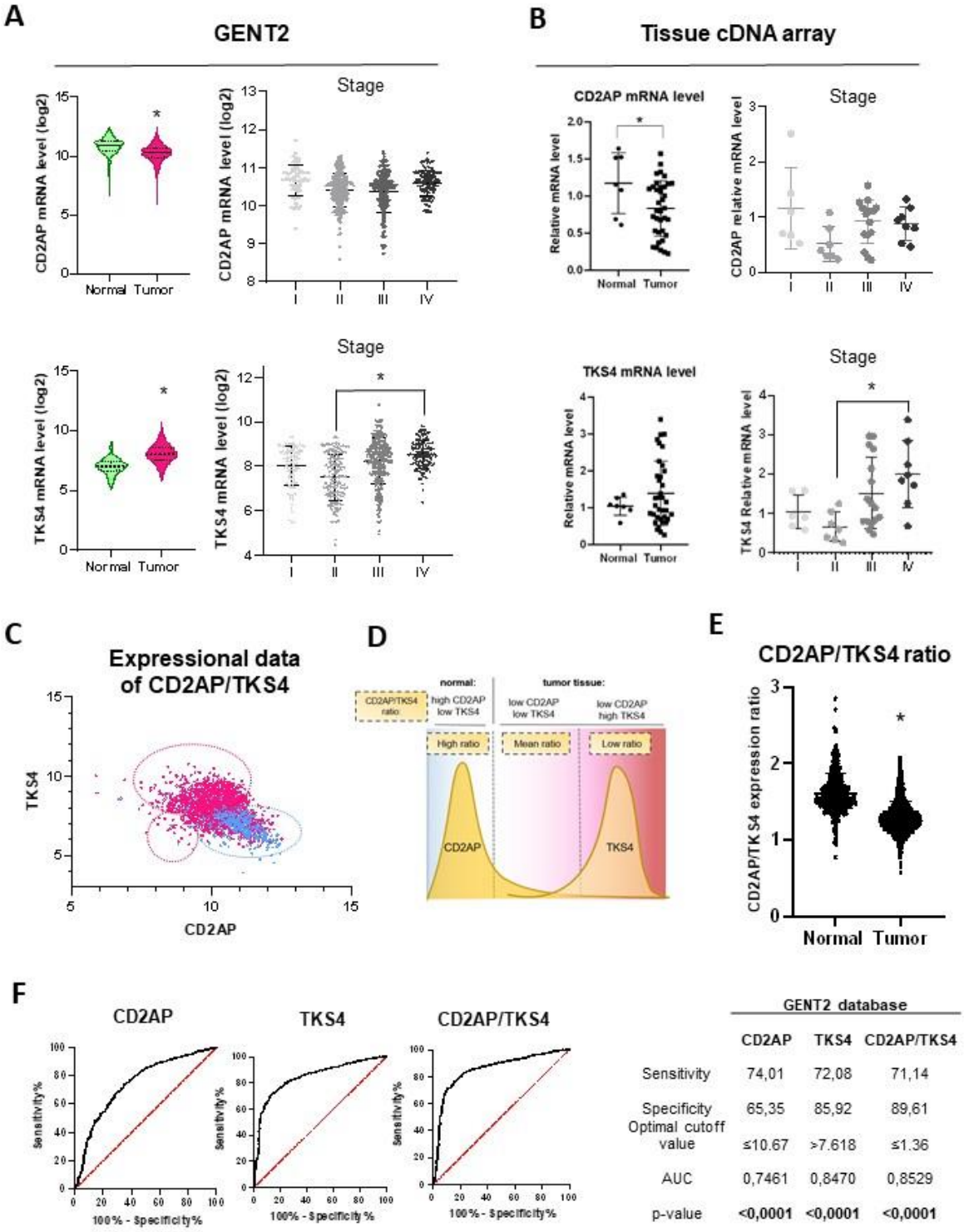


Figure 4. Overexpression of *TKS4* and/or *CD2AP* in HCT116 colon cells **(A)** Wound healing assays with *CD2AP-, TKS4-,* and *CD2AP/TKS4*-overexpressing HCT116 cells**.** (left panel); Results of wound closure after 48 hours (% relative to control) (right panel). Between-group comparisons were performed using ANOVA and Tukey's test,

and statistical significance was set at *P<0.05. **(B)** EMT marker protein levels as measured via western blotting. Protein levels of EMT markers (E-cadherin and vimentin) normalized to total protein staining.

3.4. CD2AP and TKS4 expression levels are a useful colon cancer biomarker

To investigate the expression levels of CD2AP and TKS4 in colon cancer patients, we collected and analysed mRNA data from GENT2 database. We found lower CD2AP levels and higher TKS4 levels in tumour tissue compared to normal tissue from cancer patients (Figure 5A). To test if further expressional changes present during progression, GENT2 datasets were grouped to stage I, stage II, stage III and stage IV. While CD2AP levels were equal in all stages, TKS4 expressions were significantly increased in Stage IV compared to Stage II cancer (Figure 5A). We also wanted to validate the results of the database analysis and to further explore the effects of the expression levels on disease progression in human colon cancer patients. Therefore, mRNA expression levels were measured using an qRT-PCR method with a Human Colon Cancer Tissue cDNA Array. Similar to the database results, the tissue cDNA array analysis revealed that the *CD2AP* mRNA level was decreased in cancerous colon tissue compared to normal tissue (Figure 5B). Furthermore, significantly increased *TKS4* expression levels were detected in Stage II and Stage IV cancer, and the *TKS4* level increased during disease progression, similar to findings gleaned from the database analyses (Figure 5B). In summary, the *CD2AP* level decreases while the *TKS4* level increases in colon cancer tissue.



	GENT2 database		
	CD2AP	TKS4	CD2AP/TKS4
Sensitivity	74,01	72,08	71,14
Specificity	65,35	85,92	89,61
Optimal cutoff value	≤10.67	>7.618	≤1.36
AUC	0,7461	0,8470	0,8529
p-value	**<0,0001**	**<0,0001**	**<0,0001**

Figure 5. (A) *CD2AP* and *TKS4* expression levels in colon cancer patients taken from the GENT2 database (left panel). *CD2AP* and *TKS4* expression levels in Stage I-IV from GENT2 (right panel) **(B)** Colon Cancer Tissue cDNA array results of *CD2AP* and *TKS4* mRNA expression levels in normal and tumour samples (left panel).

CD2AP and *TKS4* expression levels in tumour tissue from stage I, II, III and IV cancers (right panel). P values were calculated with Student's t-test or Tukey's test (with an ANOVA post-hoc test in the comparison of stages). Statistical significance was set at *P<0.05. **(C)** Merged expressional data of *CD2AP* and *TKS4* expression levels in non-tumorous (blue dots) and tumorous colon (pink dots) tissue from the GENT2 database. Transcript levels are shown as gene expression RNA-seq unitlog2(norm_count+1) values. **(D)** CD2AP/TKS4 ratio clusters in normal and tumorous colon tissues. **(E)** The CD2AP/TKS4 gene expression ratio in colon cancer patients from GENT2 database. P values were calculated with Student's t-test, and statistical significance was set at *P<0.05. **(F)** ROC curve analyses of TKS4, CD2AP and CD2AP/TKS4 ratios from GENT2 data (left panel); Results of ROC curve analyses from GENT2 datasets (right panel).

It was previously reported that the gene expression levels of interacting partners might be correlated (30) (31), implying that members of the TKS4 interactome might be co-regulated, particularly CD2AP. Therefore, we investigated the expressional relationships between TKS4 and CD2AP in normal or tumorous tissues. Expressional data from GENT2 were plotted to explore whether there is any clustering based on the *TKS4* and *CD2AP* levels This analysis revealed three clusters defined by different *CD2AP* and *TKS4* levels, namely CD2AP/TKS4 ratios (Figure 5C). Normal samples formed a single "low*TKS4* - high*CD2AP*" cluster (high ratio), while tumour samples formed two distinct clusters: a "low*TKS4* - low*CD2AP*" (mean ratio) and a "high*TKS4* - low*CD2AP*" cluster (low ratio) (Figure 5D). Similar results have been observed investigating data from TCGA-COAD and cDNA Tissue Array (Suppl. Fig. 1.). Thus, the discrimination of tumorous and normal tissue based on analysis of *CD2AP* or *TKS4* expression alone might result in a false negative detection, while co-analysis of *CD2AP* and *TKS4* expression might represents a useful diagnostic marker.

Our results suggest that *CD2AP* and *TKS4* expression levels are associated with colon cancer. As loss of CD2AP or loss of TKS4 alone resulted in different expression patterns of EMT markers *in vitro* (Figure 3), we hypothesized that the *CD2AP/TKS4* expression ratio might have an instructive role in tumour phenotypes. To test this possibility, *CD2AP/TKS4* gene expression ratios were calculated using GENT2 data (Figure 5E) The results showed that the CD2AP/TKS4 gene expression ratio is consistently decreased in tumour tissue (Figure 5E).

To further test the power of the *CD2AP/TKS4* ratio as a biomarker, ROC curve analyses were performed based on our database analyses and our colon cancer tissue cDNA array data to compare the clinical sensitivity and specificity of *CD2AP* or *TKS4* expression alone as well as the *CD2AP/TKS4* expression ratio in distinguishing normal and tumorous tissue (Figure 5F; Suppl. Fig. 1.). The ROC curve analysis of GENT2 data showed that the *CD2AP* is acceptable, while the *TKS4* levels as well as the *CD2AP/TKS4* ratio are excellent markers for distinguishing tumour and normal phenotypes. For the *CD2AP* and *TKS4* expression levels and the *CD2AP/TKS4* ratio, the area under curve (AUC) values were 0.7461, 0.8470, and 0.8529, respectively (Figure 5F). Of these three features, the *CD2AP/TKS4* ratio was the most specific, with an AUC of nearly 0.9 (Figure 5F, right panel). Similar results have been observed using data from TCGA-COAD and cDNA Tissue Array (Suppl. Fig. 1.).

In conclusion, the *CD2AP/TKS4* expression ratio is an outstanding biomarker for the identification of tumorous colon tissue.

4. Discussion

Several biomarkers have been developed and used for colon cancer detection; however, due to the high prevalence of this cancer and its poor prognosis, more sensitive and specific biomarkers are needed. Here we show that the previously identified gastric cancer-related CD2AP adaptor protein also has a role in colon cancer development. In addition, we found that CD2AP and the EMT regulator TKS4 form a complex in several different cancer cell types. Besides, we confirmed that CD2AP and TKS4 are in close proximity via Duolink Proximity Ligation Assays.

Via peptide array analysis, we confirmed that CD2AP binds to the ^{755}PVVPPR^{760} motif located between the last two SH3 domains in TKS4. The ^{755}PVVPPR^{760} sequence in TKS4 represents an optimal recognition motif for the first two SH3 domains of CD2AP. Our results regarding this interaction are reinforced by the results of a previous study that confirmed an interaction between TKS4 and CIN85 (11). CIN85 is a paralog of CD2AP and both have three SH3 domains (32), which are special in the sense that they recognize atypical polyproline/arginine motifs (PxP/A/V/IxPR) instead of the canonical PxxP SH3 binding sites (12)(13). Based on these similarities, we propose that CD2AP and CIN85 likely interact with TKS4 via the same mechanism.

We also showed that TKS4 and CD2AP individually and simultaneously modify cell motility and can regulate EMT-related processes in colon cancer. As co-expression of both epithelial and mesenchymal genes were observed in cells lacking TKS4 and/or CD2AP, we suggest that these proteins parallelly inhibit the formation of pEMT process (Figure 6). Furthermore, different stages of pEMT were observed depending upon which genes were silenced. For instance, individual gene silencing resulted in

elevated E-cadherin level, however, some of the mesenchymal-like TFs were upregulated as well. Therefore, we propose that an early pEMT occurs due to the decrease of TKS4 or CD2AP proteins (Figure 6). On the contrary, the simultaneously depleted TKS4 and CD2AP seems to result in an alternate EMT phenotype. The decreased E-cadherin level is accompanied by the increase of all the analysed EMT-related TFs such as *Zeb1, Twist, Snai1*. Thus, we propose that TKS4/CD2AP can simultaneously inhibit an advanced stage of pEMT.

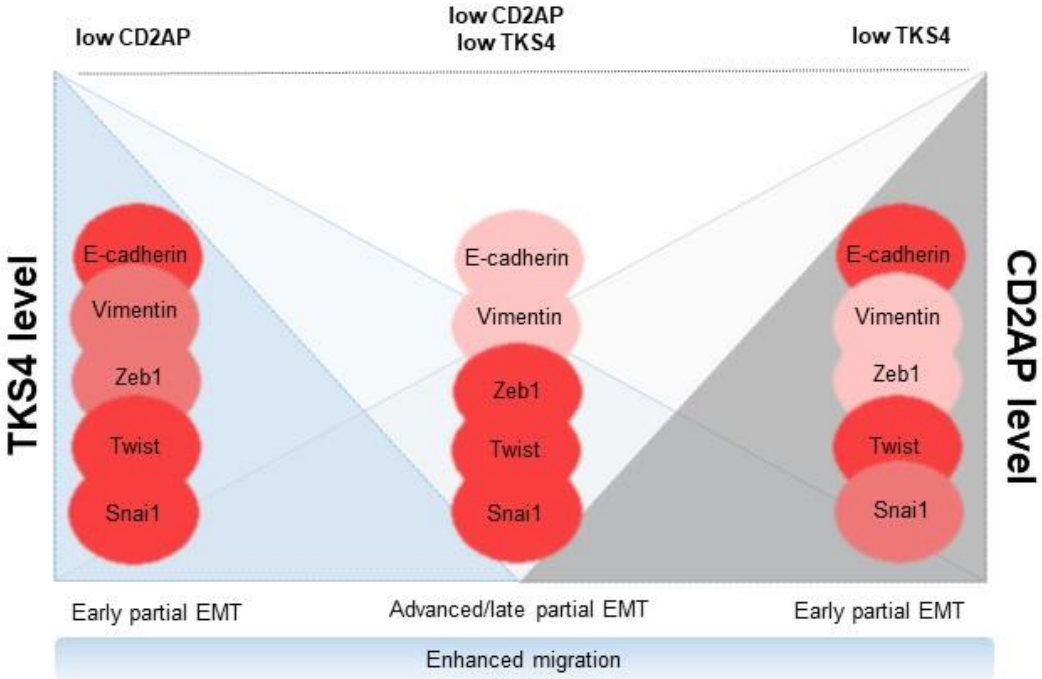


Figure 6. EMT phenotypes, cell migrations and expression levels of EMT-related genes grouped according to the levels of the CD2AP and TKS4 proteins based on gene silencing experiments. Shadings of red indicate the expressional level changes of the EMT-related gene.

Our results suggest that as adaptor proteins, TKS4 and CD2AP form a complex that organises spatially distinct signal transduction pathways by bringing signalling molecules into close proximity. This type of signal control is not unique, as previous studies have also demonstrated that adaptor proteins can functionally interact with each other to co-organize the members of different signalling pathways (33)(34). As both TKS4 and CD2AP bind to the actin cytoskeleton rearranging proteins, we established a model in which TKS4 and CD2AP cooperatively regulate actin cytoskeleton and actin-based cell motility processes (Figure 7). It was previously confirmed that TKS4 could be membrane-bound (via the PX-domain) and that it co-localizes with cortactin, an actin-binding scaffold protein (14)(23)(35). In addition, CD2AP can also interact with cortactin (36)(37). Thus, we hypothesize that CD2AP and TKS4 form a stable complex at the cell membrane and that they both conditionally interact with cortactin. This web of interactions might bring actin filaments, actin-regulating proteins and the membrane together to create membrane protrusions, which are necessary for the formation of membrane ruffles, lamellipodia, or invadopodia to support efficient metastasis. As TKS4 and CD2AP function synergistically to inhibit pEMT, we suggest that the TKS4/CD2AP complex might regulate actin rearrangement and actin-based cell motility processes, including EMT (Figure 7).

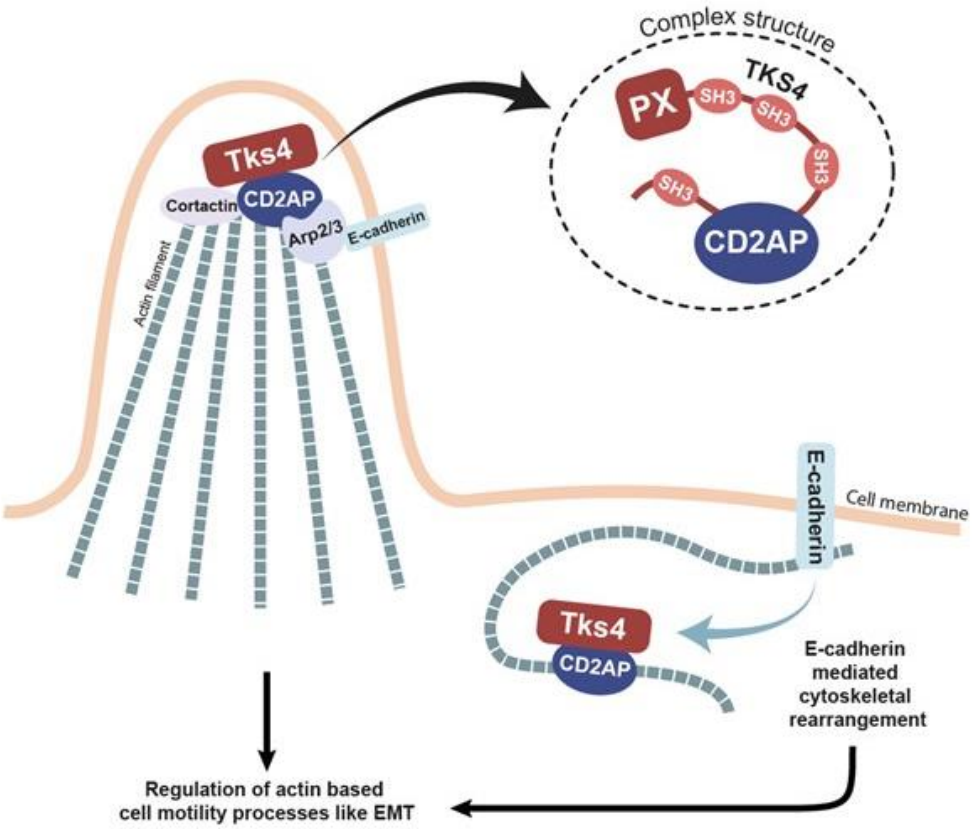


Figure 7. Model of the role of the TKS4/CD2AP complex in cytoskeleton rearrangement and actin-based cell motility process.

Previous studies showed that molecular interactions between EMT-related proteins can regulate metastasis. It was proposed that impairing the molecular interactions between adaptor proteins involved in metastasis might be a novel approach for treating aggressive cancer (38). Similarly, our results revealed that the interaction between CD2AP and TKS4 might have a cooperative role in tumour development, particularly in regulating a pEMT state in colon cancer cells. Thus, targeting this protein-protein interaction might be a novel approach for inhibiting colon cancer metastasis.

As we detected the CD2AP/TKS4 interaction in various cancer types, we hypothesize that the cooperative function of CD2AP and TKS4 in metastasis might be conserved in other cancerous and normal cells. Using cDNA samples from colon tumour samples and RNA-seq data from databases, we found that the *CD2AP* level

decreases as the *TKS4* level increases in colon cancer and that further expression levels changes can be detected during disease progression. We further found that the CD2AP/TKS4 expression level ratio has an instructive role in tumour phenotypes. We hypothesize that disturbed EMT signalling regulation induced by the decreased CD2AP/TKS4 expressional ratio in colon cancer cells might enhance their migratory capacity. Our ROC curve analyses suggest that the CD2AP/TKS4 expression ratio is an outstanding biomarker for the identification of tumorous colon tissue. Further exploration of the molecular interaction between CD2AP and TKS4, their cooperative roles in EMT, as well as the implications of the decreased CD2AP/TKS4 expression ratio in tumours might reveal additional details that could be used to develop novel approaches for detecting colon cancer. From a broader perspective, our study highlights the possibility that gene expression ratios might be more reliable biomarkers than the expression levels of single genes for the diagnosis of early-stage colon cancer and perhaps additional cancer types, such as breast or pancreatic adenocarcinoma.

Conclusions

Our study revealed that CD2AP is a novel TKS4-interacting protein and together they form a scaffold complex. They function synergistically to inhibit partial EMT as they potentially regulate EMT-related signalling in HCT116 cells. Furthermore, our work provided a new insight into that multiple genes' expressional ratio – as CD2AP/TKS4 ratio - might be more reliable biomarker than single gene expression for the diagnosis of early-stage colon cancer. At last, as the interaction between CD2AP and TKS4 might have a cooperative role in tumour development, targeting this protein-protein interaction might be a novel approach for inhibiting colon cancer metastasis.

Additional files

Additional file 1: Additional Tables 1-2.

Additional file 2: Figure S1.

Ethical approval

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author contributions

AK initiated the project, designed the study, carried out the experiments, analysed the data, supervised students, and wrote the manuscript. VV participated in the coordination, conceptualization, and implementation of the study, helped in data analysis as well as writing and editing of the manuscript. LL contributed to the experimental work, especially the DuoLink assay, data interpretation, and manuscript preparation. TT and JN helped in cell-based experiments and in data visualization. ÁT contributed to the analysis steps and supported manuscript preparation. LL contributed to the western blot experiments and data interpretation. RP contributed to data interpretation and visualization, as well as editing of the manuscript. LB supervised the project and obtained the funding. All authors read and approved the final manuscript.

Acknowledgments

We thank Zita Haidar for technical assistance and Éva Csősz for valuable help with the mass spectrometry experiments. The authors thank Rita Lózsa, Dávid Szüts, Gyöngyi Bézsenyi and László Homolya for the technical help and the tasks related to cell lines. We also thank János Fekete for his help with bioinformatic analyses. We are especially grateful to Balázs Győrffy for his useful comments about the manuscript.

Funding

This research was funded by grants from the National Research, Development, and Innovation Fund of Hungary K124045 (L.B.), FIEK_16-1-2016–0005, HunProtExc 2018-1.2.1-NKP-2018–00005 (L.B.) and. The authors are thankful for the financial support received as being a Centre of Excellence of the Hungarian Academy of Sciences. The work of TT was supported by the KDP-2021 program of the Ministry of Innovation and Technology from the source of the National Research, Development and Innovation Fund. LL is especially thankful for the support from an award from the Dr. Bodzsár Éva Foundation.

Data Availability Statement

The data and materials supporting the conclusions of this study are included within the article and its additional files.

References

1. Jelski W, Mroczko B. Biochemical Markers of Colorectal Cancer – Present and Future. *Cancer Manag Res* (2020) Volume 12:4789–4797. doi: 10.2147/CMAR.S253369

2. László L, Kurilla A, Takács T, Kudlik G, Koprivanacz K, Buday L, Vas V. Recent Updates on the Significance of KRAS Mutations in Colorectal Cancer Biology. *Cells* (2021) 10:667. doi: 10.3390/cells10030667

3. Ye X, Weinberg RA. HHS Public Access. *Trends Cell Biol* (2015) 25:675–686. doi: 10.1016/j.tcb.2015.07.012.Epithelial-Mesenchymal

4. Yang J, Antin P, Berx G, Blanpain C, Brabletz T, Bronner M, Campbell K, Cano A, Casanova J, Christofori G, et al. Guidelines and definitions for research on epithelial–mesenchymal transition. *Nat Rev Mol Cell Biol* (2020) 21:341–352. doi: 10.1038/s41580-020-0237-9

5. Pastushenko I, Brisebarre A, Sifrim A, Fioramonti M, Revenco T, Boumahdi S, Van Keymeulen A, Brown D, Moers V, Lemaire S, et al. Identification of the tumour transition states occurring during EMT. *Nature* (2018) 556: doi: 10.1038/s41586-018-0040-3

6. Saitoh M. JB special review-cellular plasticity in epithelial homeostasis and diseases: Involvement of partial EMT in cancer progression. *J Biochem* (2018) 164:257–264. doi: 10.1093/jb/mvy047

7. Yagasaki R, Noguchi M, Minami M, Earashi M. Clinical significance of E-cadherin and vimentin co-expression in breast cancer. *Int J Oncol* (1996) 9:755–761. doi: 10.3892/ijo.9.4.755

8. Aggarwal V, Montoya CA, Donnenberg VS, Sant S. Interplay between tumor microenvironment and partial EMT as the driver of tumor progression. *iScience* (2021) 24:102113. doi: 10.1016/j.isci.2021.102113

9. Tang VW, Brieher WM. FSGS3/CD2AP is a barbed-end capping protein that stabilizes actin and strengthens adherens junctions. *J Cell Biol* (2013) 203:815–833. doi: 10.1083/jcb.201304143

10. Xie W, Chen C, Han Z, Huang J, Liu X, Chen H, Zhang T, Chen S, Chen C, Lu M, et al. CD2AP inhibits metastasis in gastric cancer by promoting cellular adhesion and cytoskeleton assembly. *Mol Carcinog* (2020) 59:339–352. doi: 10.1002/mc.23158

11. Bazalii A V., Samoylenko AA, Petukhov DM, Rynditch A V., Redowicz MJ, Drobot LB. Interaction between adaptor proteins Ruk/CIN85 and Tks4 in normal and tumor cells of different tissue origins. *Biopolym Cell* (2015) 30:37–41. doi: 10.7124/bc.00087A

12. Kowanetz K, Szymkiewicz I, Haglund K, Kowanetz M, Husnjak K, Taylor JD, Soubeyran P, Engstrom U, Ladbury JE, Dikic I. Identification of a Novel Proline-Arginine Motif Involved in CIN85-dependent Clustering of Cbl and Down-regulation of Epidermal Growth Factor Receptors. *J Biol Chem* (2003) 278:39735–39746. doi: 10.1074/jbc.M304541200

13. Rouka E, Simister PC, Janning M, Kumbrink J, Konstantinou T, Muniz JRC, Joshi D, O'Reilly N, Volkmer R, Ritter B, et al. Differential Recognition Preferences of the Three Src Homology 3 (SH3) Domains from the Adaptor CD2-associated Protein (CD2AP) and Direct Association with Ras and Rab Interactor 3 (RIN3). *J Biol Chem* (2015) 290:25275–25292. doi:

10.1074/jbc.M115.637207

14. Buschman MD, Bromann PA, Cejudo-Martin P, Wen F, Pass I, Courtneidge SA. The novel adaptor protein Tks4 (SH3PXD2B) is required for functional podosome formation. *Mol Biol Cell* (2009) 20:1302–11. doi: 10.1091/mbc.e08-09-0949

15. Iizuka S, Abdullah C, Buschman MD, Diaz B, Courtneidge SA. The role of Tks adaptor proteins in invadopodia formation, growth and metastasis of melanoma. *Oncotarget* (2016) 7:78473–78486. doi: 10.18632/oncotarget.12954

16. Dülk M, Kudlik G, Fekete A, Ernszt D, Kvell K, Pongrácz JE, Mero BL, Szeder B, Radnai L, Geiszt M, et al. The scaffold protein Tks4 is required for the differentiation of mesenchymal stromal cells (MSCs) into adipogenic and osteogenic lineages. *Sci Rep* (2016) 6:1–9. doi: 10.1038/srep34280

17. Vas V, Háhner T, Kudlik G, Ernszt D, Kvell K, Kuti D, Kovács KJ, Tóvári J, Trexler M, Merő BL, et al. Analysis of Tks4 knockout mice suggests a role for Tks4 in adipose tissue homeostasis in the context of beigeing. *Cells* (2019) 8:1–16. doi: 10.3390/cells8080831

18. Vas V, Kovács T, Körmendi S, Bródy A, Kudlik G, Szeder B, Mező D, Kállai D, Koprivanacz K, Merő BL, et al. Significance of the Tks4 scaffold protein in bone tissue homeostasis. *Sci Rep* (2019) 9:1–10. doi: 10.1038/s41598-019-42250-6

19. Maziveyi M, Dong S, Baranwal S, Alahari SK. Nischarin regulates focal adhesion and Invadopodia formation in breast cancer cells. *Mol Cancer* (2018) 17:21. doi: 10.1186/s12943-018-0764-6

20. Li Z-H, Liu Y, Gao S-Y. Correlation between IL-7 genomic protein methylation level and acute myeloid leukemia. *Eur Rev Med Pharmacol Sci* (2019) 23:1196–1202. doi: 10.26355/eurrev_201902_17012

21. Caires-dos-Santos L, da Silva S V., Smuczek B, de Siqueira AS, Cruz KSP, Barbuto JAM, Augusto TM, Freitas VM, Carvalho HF, Jaeger RG. Laminin-derived peptide C16 regulates Tks expression and reactive oxygen species generation in human prostate cancer cells. *J Cell Physiol* (2020) 235: doi: 10.1002/jcp.28997

22. Szeder B, Tárnoki-Zách J, Lakatos D, Vas V, Kudlik G, Merő B, Koprivanacz K, Bányai L, Hámori L, Róna G, et al. Absence of the Tks4 Scaffold Protein Induces Epithelial-Mesenchymal Transition-Like Changes in Human Colon Cancer Cells. *Cells* (2019) 8: doi: 10.3390/cells8111343

23. Lányi Á, Baráth M, Péterfi Z, Bőgel G, Orient A, Simon T, Petrovszki E, Kis-Tóth K, Sirokmány G, Rajnavölgyi É, et al. The Homolog of the Five SH3-Domain Protein (HOFI/SH3PXD2B) Regulates Lamellipodia Formation and Cell Spreading. *PLoS One* (2011) 6:e23653. doi: 10.1371/journal.pone.0023653

24. Harnoš J, Ryneš J, Víšková P, Foldynová-Trantírková S, Bajard-Ešner L, Trantírek L, Bryja V. Analysis of binding interfaces of the human scaffold protein AXIN1 by peptide microarrays. *J Biol Chem* (2018) 293:16337–16347. doi: 10.1074/jbc.RA118.005127

25. Vermani L, Kumar R, Senthil Kumar N. GAPDH and PUM1: Optimal Housekeeping Genes for Quantitative Polymerase Chain Reaction-Based Analysis of Cancer Stem Cells and Epithelial-Mesenchymal Transition Gene Expression in Rectal Tumors. *Cureus* (2020) doi: 10.7759/cureus.12020

26. Hellemans J, Mortier G, De Paepe A, Speleman F, Vandesompele J. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biol* (2007) 8:R19. doi: 10.1186/gb-2007-8-2-r19

27. Park S-J, Yoon B-H, Kim S-K, Kim S-Y. GENT2: an updated gene expression database for normal and tumor tissues. *BMC Med Genomics* (2019) 12:101. doi: 10.1186/s12920-019-0514-7

28. Goldman MJ, Craft B, Hastie M, Repečka K, McDade F, Kamath A, Banerjee A, Luo Y, Rogers D, Brooks AN, et al. Visualizing and interpreting cancer genomics data via the Xena platform. *Nat Biotechnol* (2020) 38:675–678. doi: 10.1038/s41587-020-0546-8

29. Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, Doncheva NT, Legeay M, Fang T, Bork P, et al. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res* (2021) 49:D605–D612. doi: 10.1093/nar/gkaa1074

30. Grigoriev A. A relationship between gene expression and protein interactions on the proteome scale: analysis of the bacteriophage T7 and the yeast Saccharomyces cerevisiae. *Nucleic Acids Res* (2001) 29:3513–9. doi: 10.1093/nar/29.17.3513

31. Ge H, Liu Z, Church GM, Vidal M. Correlation between transcriptome and interactome mapping data from Saccharomyces cerevisiae. *Nat Genet* (2001) 29:482–6. doi: 10.1038/ng776

32. Gaidos G, Soni S, Oswald DJ, Toselli PA, Kirsch KH. Structure and function analysis of the CMS/CIN85 protein family identifies actin-bundling properties and heterotypic-complex formation. *J Cell Sci* (2007) 120: doi: 10.1242/jcs.004333

33. Yablonski D. Bridging the Gap: Modulatory Roles of the Grb2-Family Adaptor, Gads, in Cellular and Allergic Immune Responses. *Front Immunol* (2019) 10: doi: 10.3389/fimmu.2019.01704

34. Buday L, Khwaja A, Sipeki S, Faragó A, Downward J. Interactions of Cbl with two adaptor proteins, Grb2 and Crk, upon T cell activation. *J Biol Chem* (1996) 271: doi: 10.1074/jbc.271.11.6159

35. Bögel G, Gujdár A, Geiszt M, Lányi Á, Fekete A, Sipeki S, Downward J, Buday L. Frank-ter Haar syndrome protein Tks4 regulates epidermal growth factor-dependent cell migration. *J Biol Chem* (2012) 287:31321–31329. doi: 10.1074/jbc.M111.324897

36. Lynch DK, Winata SC, Lyons RJ, Hughes WE, Lehrbach GM, Wasinger V, Corthals G, Cordwell S, Daly RJ. A Cortactin-CD2-associated Protein (CD2AP) Complex Provides a Novel Link between Epidermal Growth Factor Receptor Endocytosis and the Actin Cytoskeleton. *J Biol Chem* (2003) 278:21805–21813. doi: 10.1074/jbc.M211407200

37. Zhao J, Bruck S, Cemerski S, Zhang L, Butler B, Dani A, Cooper JA, Shaw AS. CD2AP Links Cortactin and Capping Protein at the Cell Periphery To Facilitate Formation of Lamellipodia. *Mol Cell Biol* (2013) 33:38–47. doi: 10.1128/MCB.00734-12

38. Feldker N, Ferrazzi F, Schuhwerk H, Widholz SA, Guenther K, Frisch I, Jakob K, Kleemann J, Riegel D, Bönisch U, et al. Genome-wide cooperation of EMT transcription factor ZEB 1 with YAP and AP -1 in breast cancer . *EMBO J* (2020) 39: doi: 10.15252/embj.2019103209

Figure titles and legends

Figure 1. A combination of co-IP and MS was used to identify new TKS4-interacting partners, and the validation of CD2AP-TKS4 interaction via IP-WB and DuoLink Proximity Ligation Assay. **(A)** Six consistently detected TKS4-interacting partners and their interaction network created via the STRING database. Lines indicate previously identified interactions. **(B)** Validation of CD2AP as a TKS4-interacting partner via IP-WB. TKS4: Immunprecipitated protein lysates with TKS4-specific antibody; -: Immunpreciptated protein lysates without TKS4-specific antibody. **(C)** Detecting the CD2AP-TKS4 interaction in HCT116 cells: fluorescence signal is generated only when the two DPLA probes are in proximity. Colocalization appears as red dots in the cytoplasm and close to the cell membrane. No colocalization was seen in control staining that did not include the primary antibodies in the DPLA reaction. Colours: red dots = CD2AP-TKS4 colocalization, blue = DAPI-stained nuclei;

Figure 2. Identification of potential CD2AP binding sites within TKS4. Columns indicate the detected normalized fluorescent signal from the peptide microarray, and the dashed line represents the threshold signal intensity (10000 U). The X axis shows the TKS4 aminoacids numbered by the position of their first residue in the TKS4 sequence. Pink columns represent normalized signal above threshold (dashed line). The pink colour shading indicates the degree of the relative surface accessibility value (number of residues with >50% rASA) of each 15mer peptide. (*Light pink* columns represent peptides containing *0-<5* residues with more than 50% rASA, *medium pink* columns represent peptides containing *5<10* residues with more than 50% rASA, *dark pink* columns represent peptides containing *≥10* residues with more than 50% rASA.)

Figure 3. Depletion of *TKS4* and/or *CD2AP* expression in HCT116 colon cells. **(A)** Wound healing assay in *CD2AP-*, *TKS4-*, and *CD2AP/TKS4*-silenced HCT116 cells. Between-group comparisons were performed using ANOVA and Tukey's test, and statistical significance was set at *P<0.05. **(B)** EMT marker protein levels as measured via western blotting (E-cadherin and vimentin). **(C)** qPCR analyses of EMT markers in si*TKS4-*, si*CD2AP-* and si*CD2AP/TKS4*-treated HCT116 cells. Fold changes show the expressions relative to the mRNA level in negative control siRNA-treated cells. P values were calculated with Student's t-test, and statistical significance was set at *P<0.05.

Figure 4. Overexpression of *TKS4* and/or *CD2AP* in HCT116 colon cells **(A)** Wound healing assays with CD2AP-, TKS4-, and CD2AP/TKS4-overexpressing HCT116 cells**.** (left panel) ; Results of wound closure after 48 hours (% relative to control) (right panel). Between-group comparisons were performed using ANOVA and Tukey's test, and statistical significance was set at *P<0.05. **(B)** EMT marker protein levels as measured via western blotting. Protein levels of EMT markers (E-cadherin and vimentin) normalized to total protein staining.

Figure 5. (A) *CD2AP* and *TKS4* expression levels in colon cancer patients taken from the GENT2 database (left panel). *CD2AP* and *TKS4* expression levels in Stage I-IV from GENT2 (right panel) **(B)** Colon Cancer Tissue cDNA array results of *CD2AP* and *TKS4* mRNA expression levels in normal and tumour samples (left panel). *CD2AP* and *TKS4* expression levels in tumour tissue from stage I, II, III and IV cancers (right panel). P values were calculated with Student's t-test or Tukey's test (with an

ANOVA post-hoc test in the comparison of stages). Statistical significance was set at *P<0.05**. (C)** Merged expressional data of *CD2AP* and *TKS4* expression levels in non-tumorous (blue dots) and tumorous colon (pink dots) tissue from the GENT2 database. Transcript levels are shown as gene expression RNA-seq unitlog2(norm_count+1) values**. (D)** CD2AP/TKS4 ratio clusters in normal and tumorous colon tissues. **(E)** The CD2AP/TKS4 gene expression ratio in colon cancer patients from GENT2 database. P values were calculated with Student's t-test, and statistical significance was set at *P<0.05. **(F)** ROC curve analyses of TKS4, CD2AP and CD2AP/TKS4 ratios from GENT2 data (left panel); Results of ROC curve analyses from GENT2 datasets (right panel).

Figure 6. EMT phenotypes, cell migrations and expression levels of EMT-related genes grouped according to the levels of the CD2AP and TKS4 proteins based on gene silencing experiments. Shadings of red indicate the expressional level changes of the EMT-related genes.

Figure 7. Model of the role of the TKS4/CD2AP complex in cytoskeleton rearrangement and actin-based cell motility process.

Supplementary Figure S1. (A) The CD2AP/TKS4 gene expression ratio in colon cancer patients from TCGA-COAD database and **(B)** based on the Colon Cancer Tissue cDNA array results. P values were calculated with Student's t-test, and statistical significance was set at *P<0.05. **(C)** ROC curve analyses of TKS4, CD2AP and CD2AP/TKS4 ratios from TCGA-COAD database. **(D)** ROC curve analyses of TKS4, CD2AP, and TKS4/CD2AP ratios to discriminate normal and tumour tissue from the Colon Cancer Tissue cDNA array. **(E)** Results of ROC curve analyses from TCGA or Tissue cDNA array datasets.

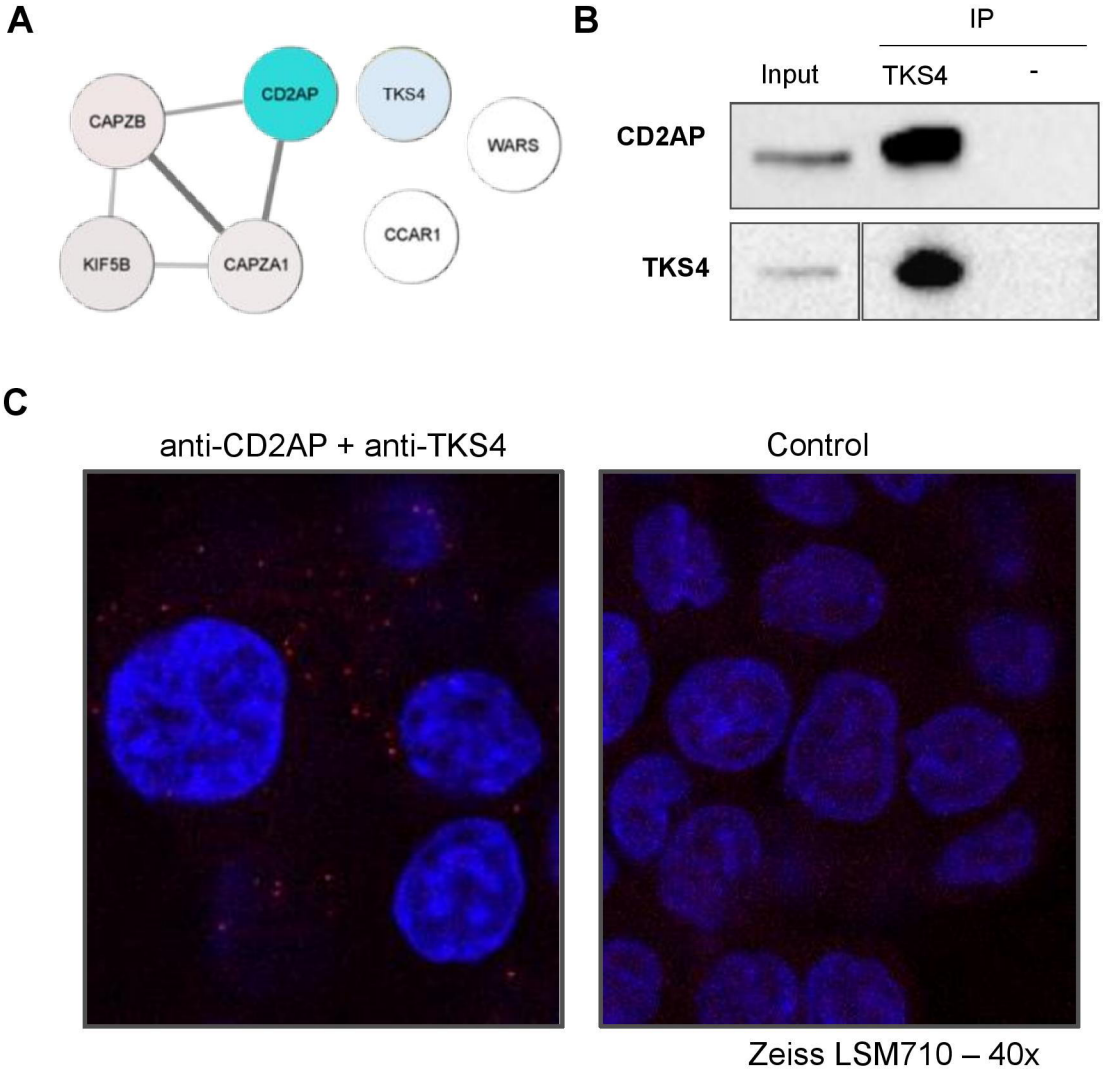
Supplementary Table S1. Oligos used in the study

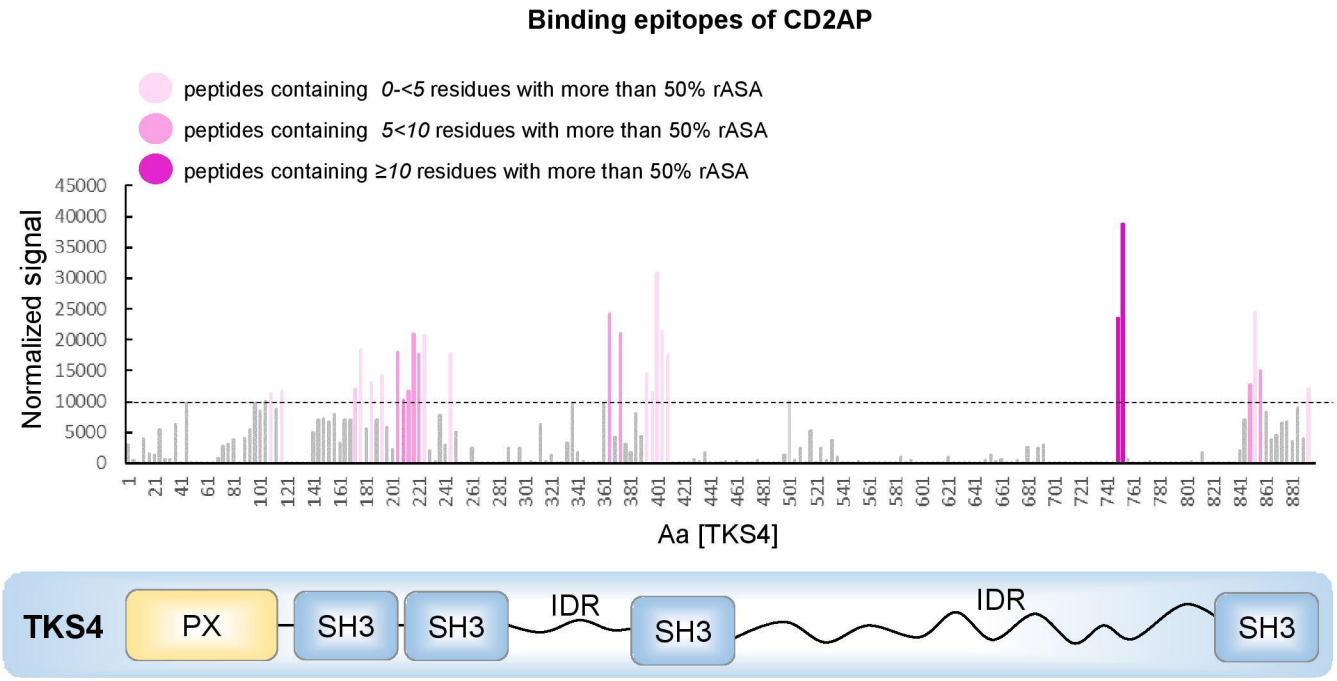
Supplementary Table S2. TKS4 and CD2AP peptide array. PX and SH3 domains are shown with bold letters, signals up to 10000 UA are grey.

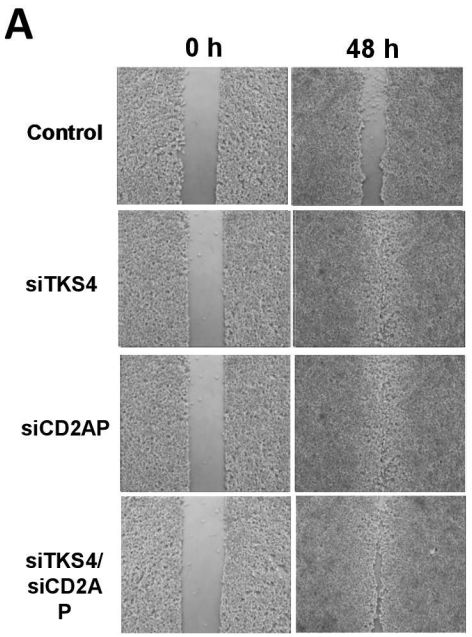
Additional files

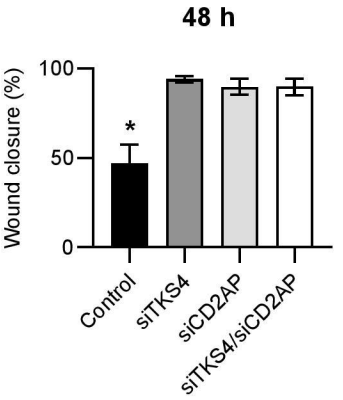
Supplementary Tables 1-2

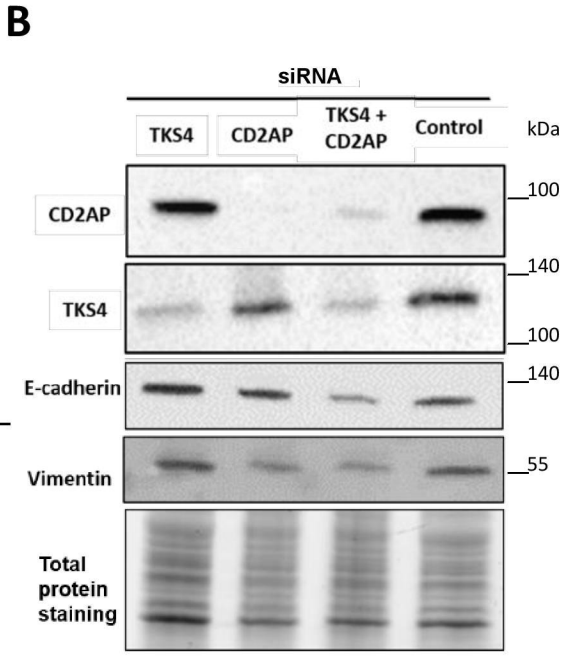
Supplementary Figures S1

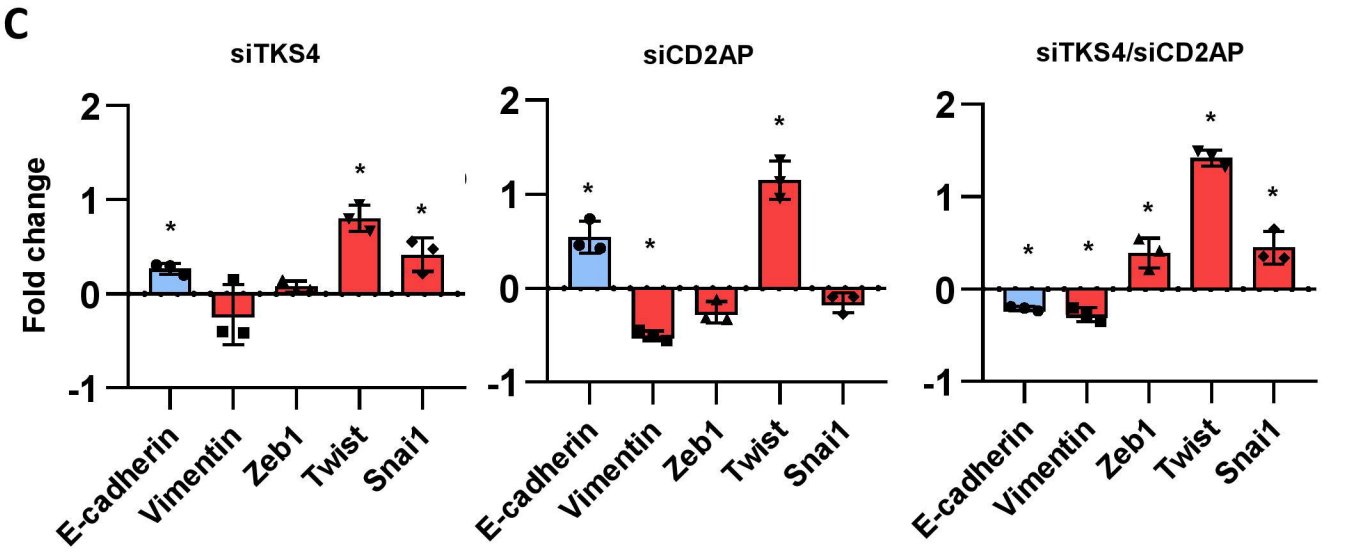

A
CAPZB
CD2AP
TKS4
WARS
KIF5B
CAPZA1
CCAR1
B
IP
Input
TKS4
-
CD2AP
TKS4
C
anti-CD2AP + anti-TKS4
Control
Zeiss LSM710 – 40x

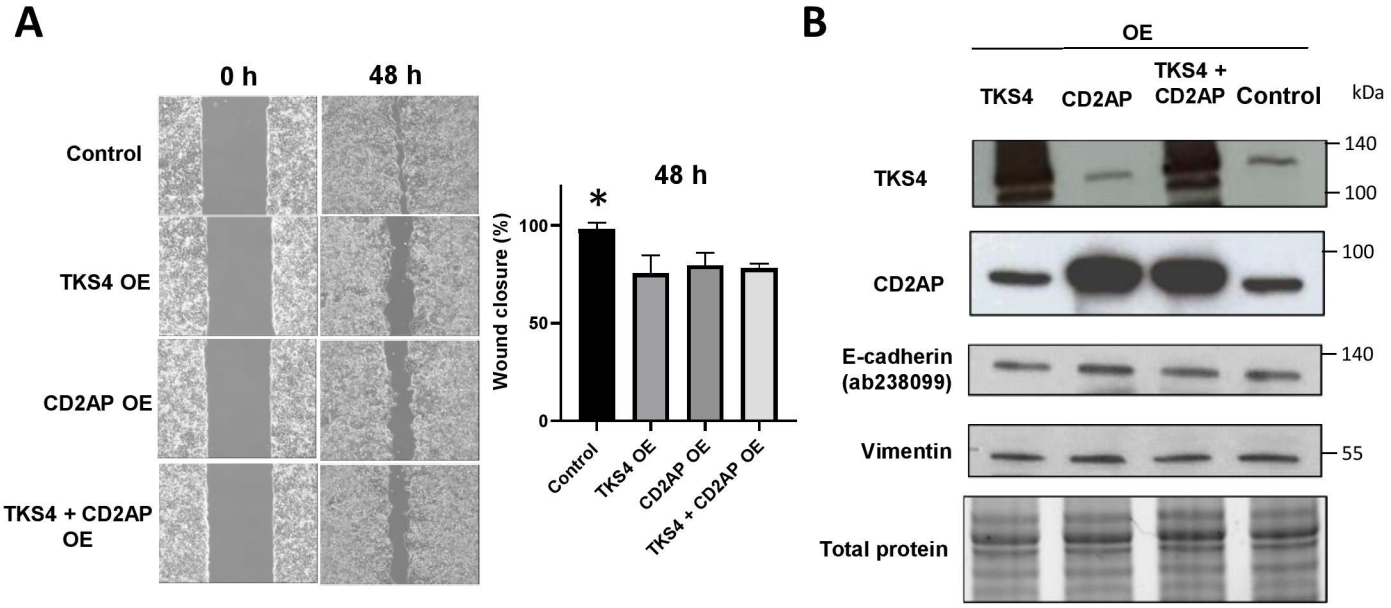

Binding epitopes of CD2AP
peptides containing 0-<5 residues with more than 50% rASA
peptides containing 5<10 residues with more than 50% rASA
peptides containing ≥10 residues with more than 50% rASA
Normalized signal
Aa [TKS4]
TKS4
PX
SH3
SH3
IDR
SH3
IDR
SH3

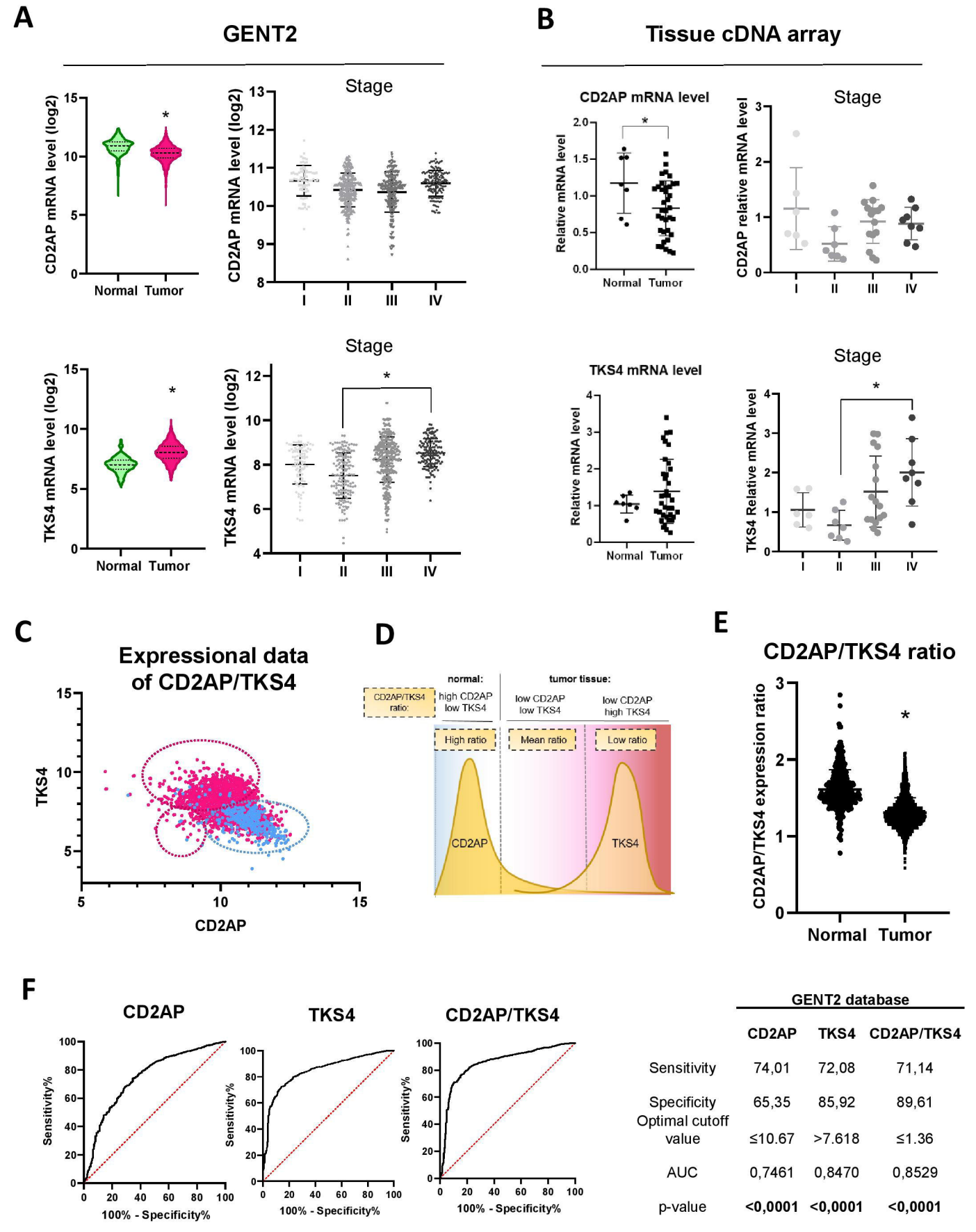

A
0 h
48 h
Control
siTKS4
siCD2AP
siTKS4/
siCD2A
P


48 h
Wound closure (%)
100
50
0
*
Control
siTKS4
siCD2AP
siTKS4/siCD2AP

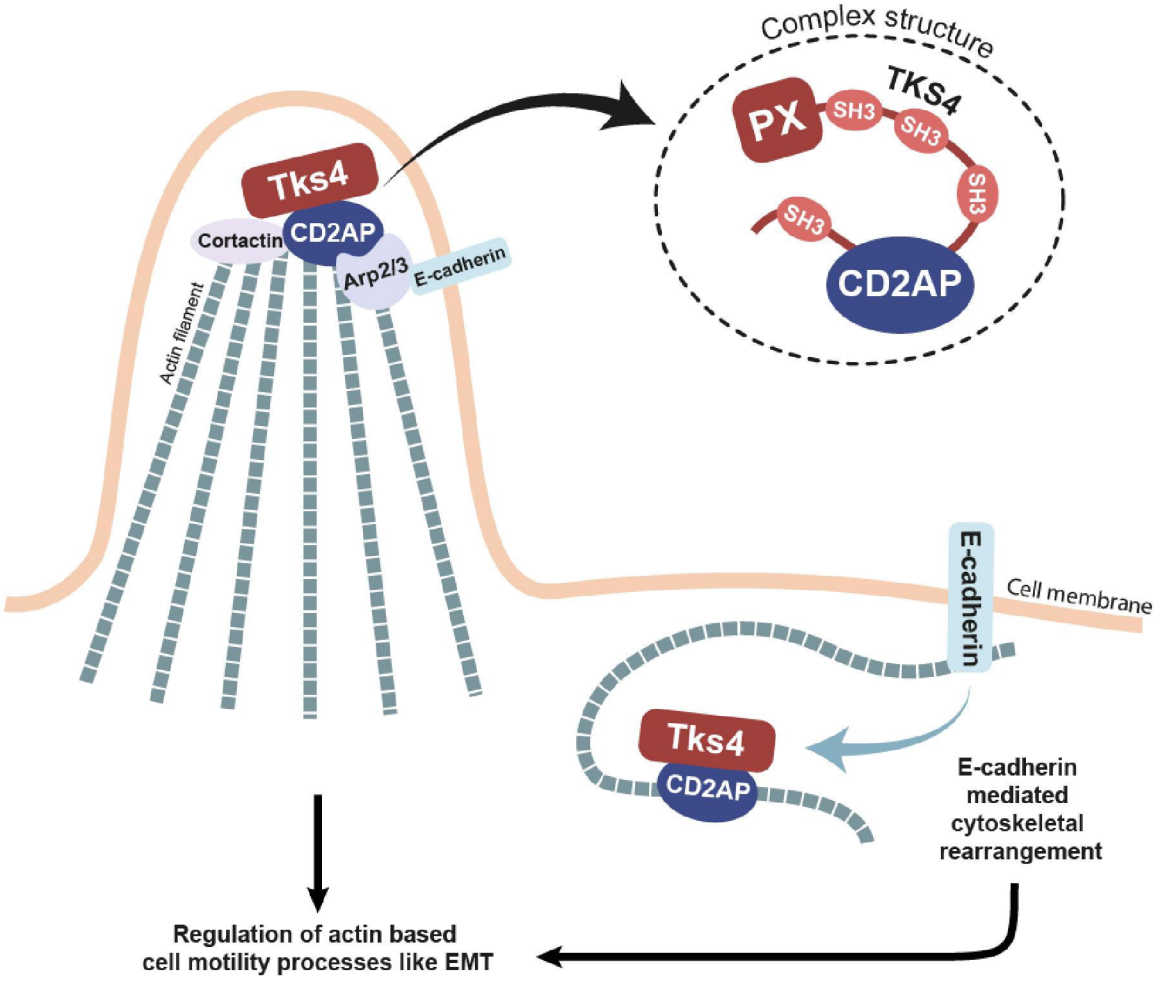

B
siRNA
TKS4
CD2AP
TKS4 +
CD2AP
Control
kDa
CD2AP
TKS4
E-cadherin
Vimentin
Total
protein
staining
100
140
100
140
55

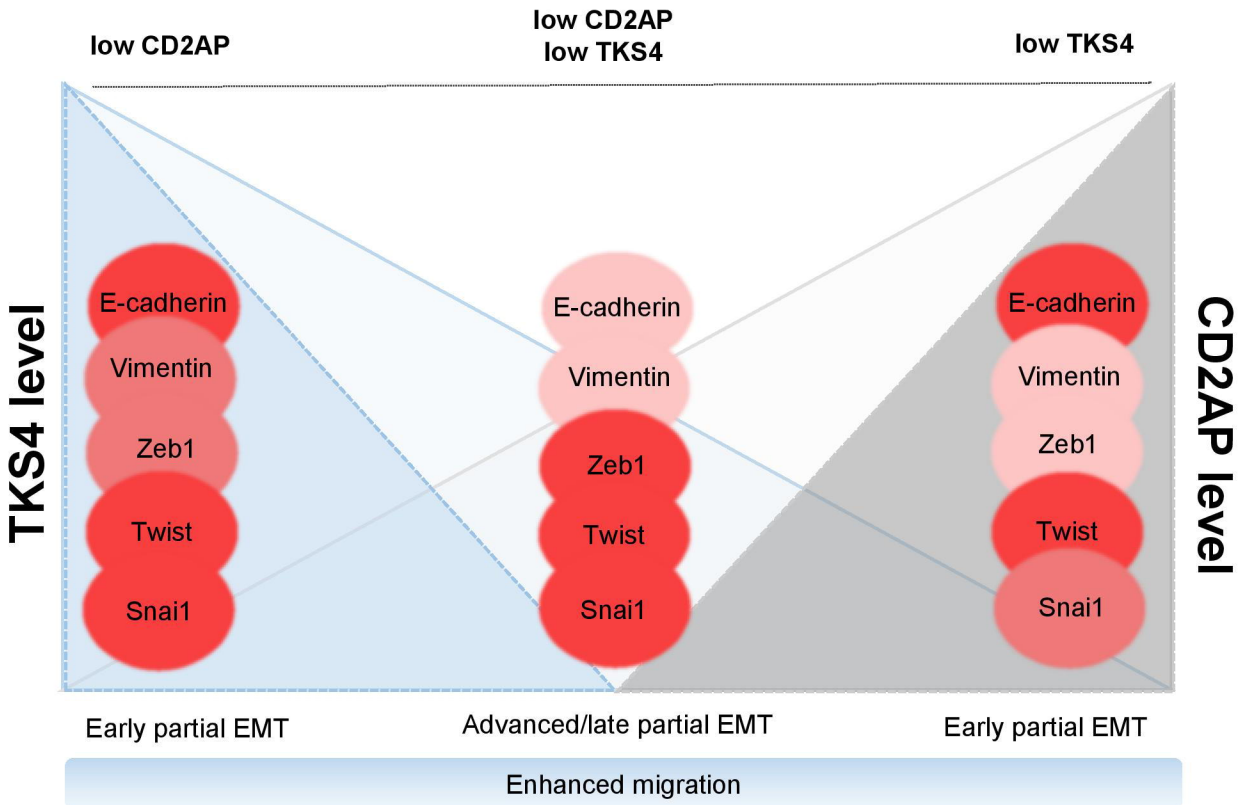

C
siTKS4
siCD2AP
siTKS4/siCD2AP
Fold change
2
1
0
-1
*
E-cadherin
Vimentin
Zeb1
Twist
Snai1


A
0 h
48 h
Control
TKS4 OE
CD2AP OE
TKS4 + CD2AP OE
48 h
Wound closure (%)
100
50
0
*
Control
TKS4 OE
CD2AP OE
TKS4 + CD2AP OE
B
OE
TKS4
CD2AP
TKS4 + CD2AP
Control
kDa
TKS4
140
100
CD2AP
100
E-cadherin (ab238099)
140
Vimentin
55
Total protein



GENT2 database	CD2AP	TKS4	CD2AP/TKS4
Sensitivity	74,01	72,08	71,14
Specificity	65,35	85,92	89,61
Optimal cutoff value	≤10.67	>7.618	≤1.36
AUC	0,7461	0,8470	0,8529
p-value	**<0,0001**	**<0,0001**	**<0,0001**


Complex structure
TKS4
PX
SH3
SH3
SH3
SH3
CD2AP
Tks4
Cortactin
CD2AP
Arp2/3
E-cadherin
Actin filament
E-cadherin
Cell membrane
Tks4
CD2AP
E-cadherin mediated cytoskeletal rearrangement
Regulation of actin based cell motility processes like EMT


low CD2AP
low CD2AP
low TKS4
low TKS4
TKS4 level
CD2AP level
E-cadherin
Vimentin
Zeb1
Twist
Snai1
E-cadherin
Vimentin
Zeb1
Twist
Snai1
E-cadherin
Vimentin
Zeb1
Twist
Snai1
Early partial EMT
Advanced/late partial EMT
Early partial EMT
Enhanced migration