PNAS

## Supporting Information for

# Syndecan-4 is a maestro of gastric cancer cell invasion and communication that underscores poor survival

Juliana Poças[1,2,3], Catarina Marques[1,2,3], Catarina Gomes[1,2], Andreia Hanada Otake[4,♦], Filipe Pinto[1,2], Mariana Ferreira[1,2], Tiago Silva[1,2], Isabel Faria-Ramos[1,2], Rita Matos[1,2], Ana Raquel Ribeiro[1], Emanuel Senra[1], Bruno Cavadas[1], Sílvia Batista[4], Joana Maia[4], Joana A. Macedo[1,2], Luís Lima[5], Luís Pedro Afonso[5], José Alexandre Ferreira[5], Lúcio Lara Santos[5], António Polónia[1,2], Hugo Osório[1,2,6], Mattias Belting[7,8,9], Celso A. Reis[1,2,3,6], Bruno Costa-Silva[4], Ana Magalhães[1,2,3*]

[1] i3S - Instituto de Investigação e Inovação em Saúde, Universidade do Porto, 4200-135 Porto, Portugal.
[2] IPATIMUP - Instituto de Patologia e Imunologia Molecular da Universidade do Porto, 4200-465 Porto, Portugal.
[3] ICBAS - Instituto de Ciências Biomédicas Abel Salazar, Universidade do Porto, 4050-313 Porto, Portugal.
[4] Champalimaud Physiology and Cancer Programme, Champalimaud Foundation, 1400-038 Lisbon, Portugal.
[5] Experimental Pathology and Therapeutics Group, Research Center of IPO Porto (CI-IPOP)/RISE@CI-IPOP (Health Research Network), Portuguese Oncology Institute of Porto (IPO Porto)/Porto Comprehensive Cancer Centre (Porto.CCC), Porto, Portugal.
[6] FMUP - Faculdade de Medicina da Universidade do Porto, 4200-319 Porto, Portugal.
[7] Department of Clinical Sciences, Section of Oncology, Lund University, Lund, Sweden.
[8] Department of Immunology, Genetics and Pathology, Science for Life Laboratory, Uppsala University, Uppsala, Sweden.
[9] Skåne University Hospital, Lund, Sweden.
♦current address: Instituto do Cancer do Estado de Sao Paulo, Hospital das Clinicas HCFMUSP, Faculdade de Medicina, Universidade de Sao Paulo, 01246-000, Sao Paulo, Brazil.

***To whom correspondence may be addressed. Email:** amagalhaes@ipatimup.pt

**This PDF file includes:**

Figures S1 to S12
Tables S1 to S4
SI Material and Methods

**SI also includes Excel file: SI_EV signature risk.xlsx**

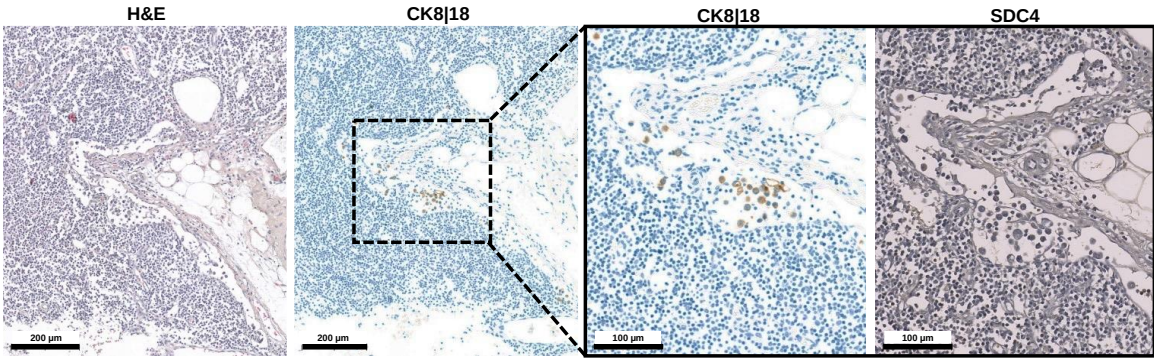
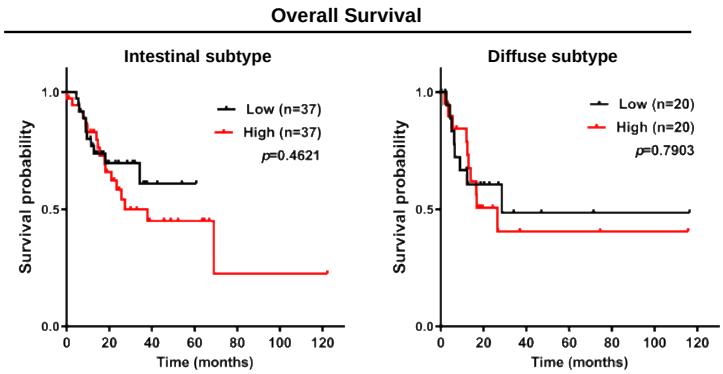
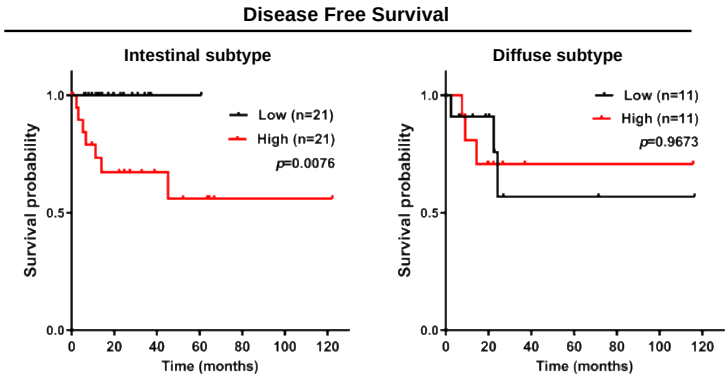



**Fig. S1. Higher magnification of a lymph node metastasis from diffuse gastric cancer.** Magnification images of the R3 area shown in Figure 1b. Hematoxylin-eosin (H&E) staining and cytokeratin 8 and 18 (CK8/18) allow the identification of the tumour epithelial cells and higher magnification of SDC4 staining confirms that these tumour cells are negative for SDC4 expression.

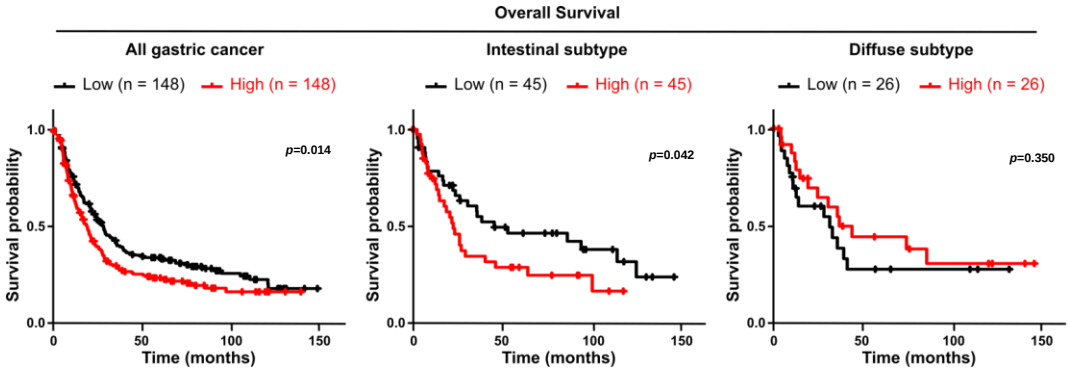
**a.**


Overall Survival
Intestinal subtype
Low (n=37)
High (n=37)
p=0.4621
Survival probability
Time (months)
Diffuse subtype
Low (n=20)
High (n=20)
p=0.7903
Survival probability
Time (months)

**b.**


Disease Free Survival
Intestinal subtype
Low (n=21)
High (n=21)
p=0.0076
Survival probability
Time (months)
Diffuse subtype
Low (n=11)
High (n=11)
p=0.9673
Survival probability
Time (months)

**c.**


Overall Survival
All gastric cancer
Low (n = 148)
High (n = 148)
p=0.014
Survival probability
Time (months)
Intestinal subtype
Low (n = 45)
High (n = 45)
p=0.042
Survival probability
Time (months)
Diffuse subtype
Low (n = 26)
High (n = 26)
p=0.350
Survival probability
Time (months)

**d.**

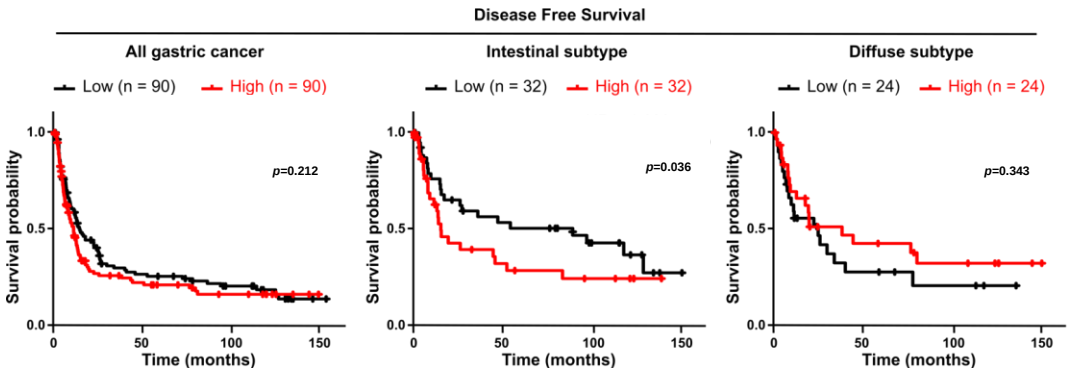

Disease Free Survival
All gastric cancer
Low (n = 90)
High (n = 90)
p=0.212
Survival probability
Time (months)
Intestinal subtype
Low (n = 32)
High (n = 32)
p=0.036
Survival probability
Time (months)
Diffuse subtype
Low (n = 24)
High (n = 24)
p=0.343
Survival probability
Time (months)

**Fig. S2. *SDC4* expression associates with worse survival in gastric cancer patients.** Kaplan-Meier (KM) analysis between *SDC4* mRNA levels and gastric cancer **(a, c)** overall survival (OS) or **(b, d)** disease-free survival (DFS). **(a, b)** Refer to TCGA database and **(c, d)** to GEO cohort extracted from KM Plotter database. The categorization of patients' samples was divided into low (lowest 25%, black lines) and high (highest 25%, red lines) subgroups according to *SDC4* mRNA expression levels. Comparison between curves in KM plots was performed using log-rank test.

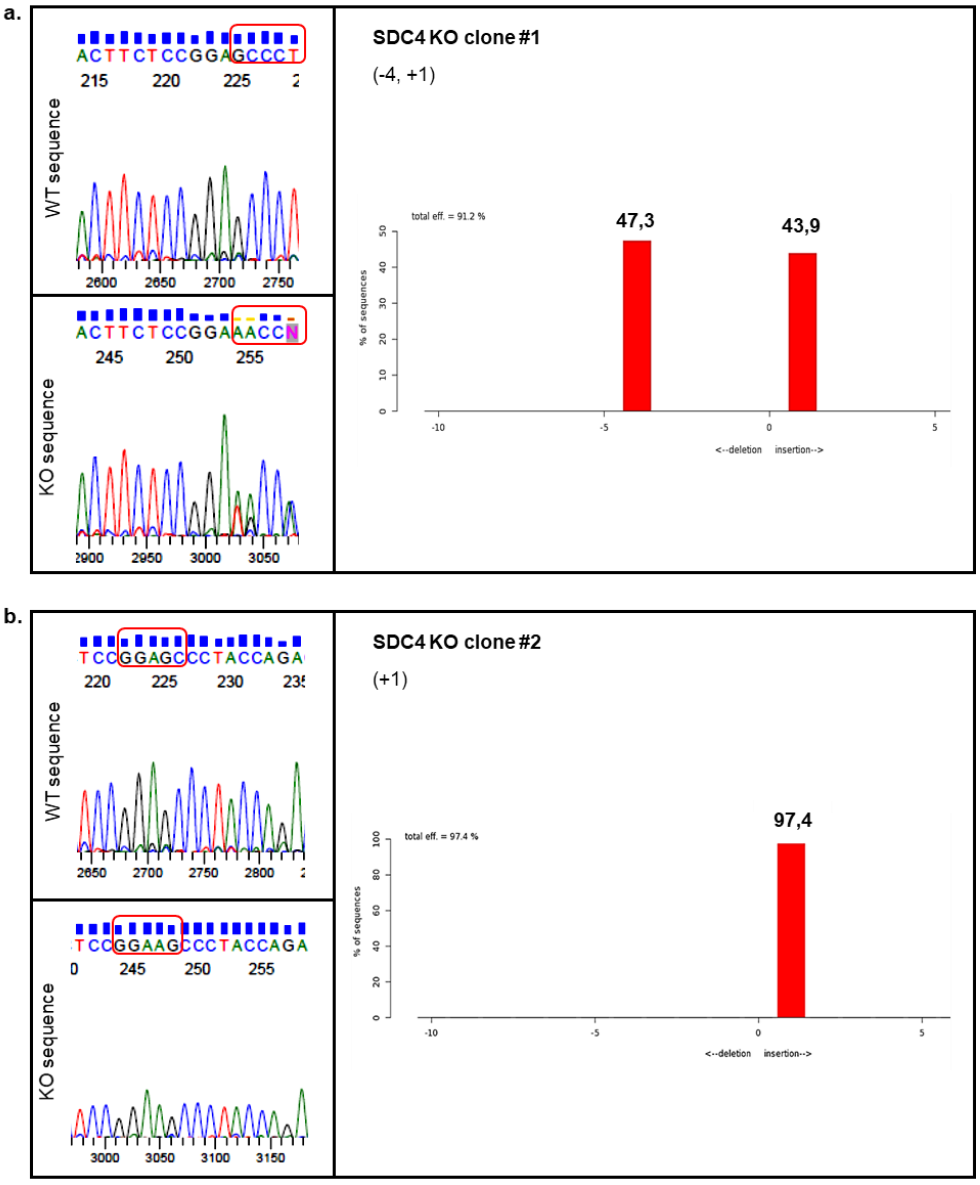



**Fig. S3. *SDC4* KO validation by indel sequencing.** *SDC4* KO validation through Sanger sequencing for **(a)** *SDC4* KO #1 and **(b)** *SDC4* KO #2 clones, from MKN74 WT cell line (Control). Indel sequencing was validated through Tracking of Indels by Decomposition (TIDE) analysis.

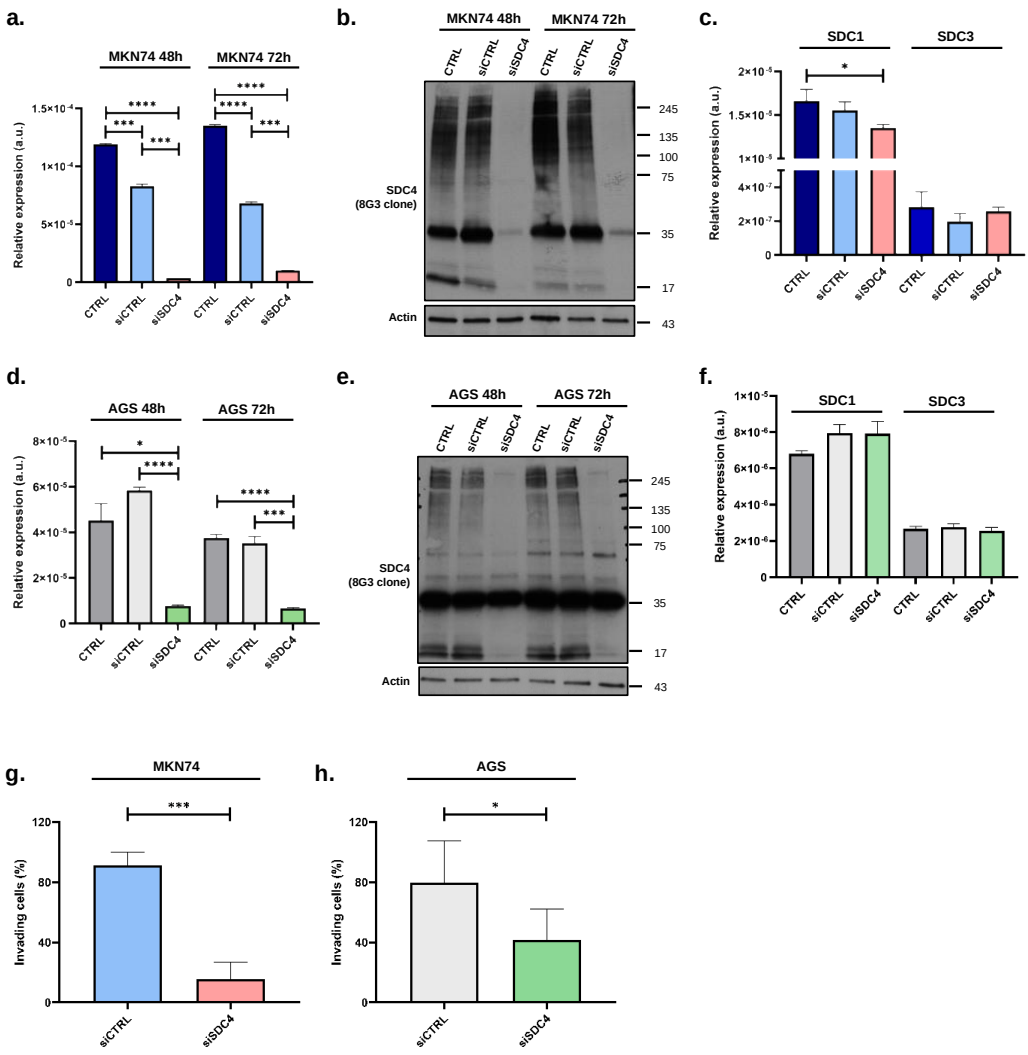



**Fig. S4. *SDC4* silencing impairs gastric cancer cell invasion. (a, d)** Bar graphs display *SDC4* mRNA levels in **(a)** MKN74 and **(d)** AGS cell lines: CTRL (Lipofectamine), siCTRL(scramble) and si*SDC4*, 48h and 72h after transfection. **(b, e)** WB analysis of SDC4 on **(b)** MKN74 and **(e)** AGS cell lines. Actin was used as loading control. **(c, f)** Bar graphs display *SDC1* and *SDC3* mRNA levels in **(c)** MKN74 and **(f)** AGS cell lines: CTRL (Lipofectamine), siCTRL(scramble) and si*SDC4*, 72h after transfection. **(g, h)** Invasion capacity of siCTRL (scramble) and si*SDC4* was evaluated on **(g)** MKN74 and **(h)** AGS cells. Bar graphs show the average of the percentage of invading cells + SD. A total of n= 6 from 3 independent experiments is shown. Statistical significance was determined using student's t-test with Welch's correction. *p≤0.05, ***p<0.001, ****$p$<0.0001.

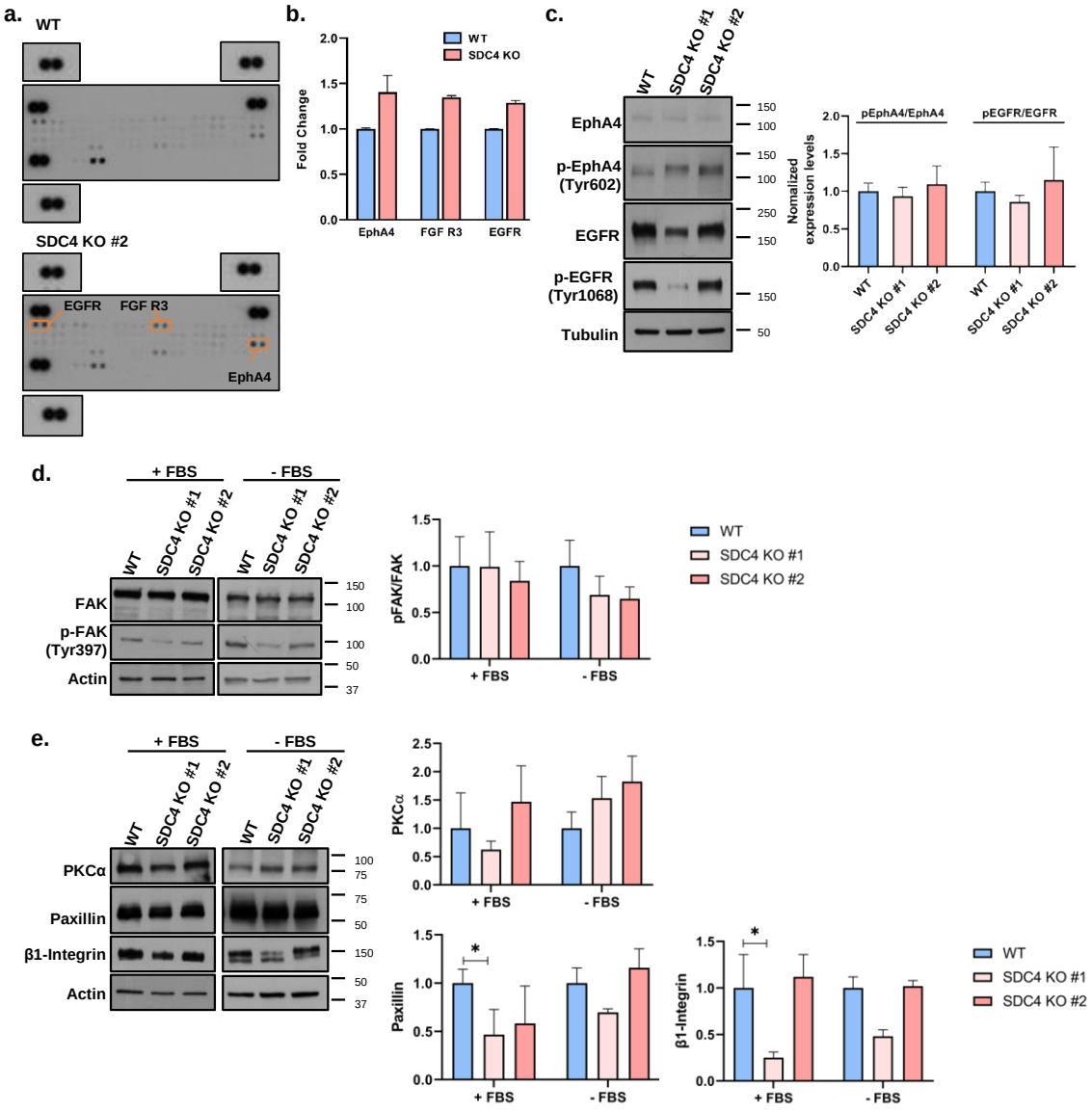



**Fig. S5. Assessment of *SDC4* KO impact on key cellular signalling pathways. (a)** Evaluation of the phosphorylation state of 50 Receptor Tyrosine Kinases (RTKs) in WT and *SDC4* KO cells using the Proteome Profiler Human p-RTK array. Each couple of dots represents one p-RTK. Reference spots are shown. The RTKs that presented a fold change ≥ 0.15 when comparing *SDC4* KO and WT cells are highlighted **(b)** Bar graphs show relative phosphorylation levels of *SDC4* KO cells normalized to WT values for the RTKs highlighted in (a). **(c)** The activation of EphA4 and EGFR in WT and *SDC4* KO cell clones was measured by WB analysis. α-tubulin was

used as loading control. Band density values of pEphA4 and pEGFR were normalized to the levels of the total receptor. Bar graph shows the mean values of the normalized phosphorylated receptor + SD. Data from 3 independent biological replicates. **(d)** The activation of FAK in WT and *SDC4* KO cell clones was measured by WB in protein extracts collected from cells grown in complete medium (+FBS) and simple medium (-FBS). Bar graph shows the mean values of the normalized phosphorylated receptor (pFAK/FAK) and normalized total form to β-actin. Data from 4 (+FBS) and 3 (-FBS) independent biological replicates **(e)** Evaluation of the expression levels of PKCα, paxillin and β1-Integrin in protein extracts collected from cells grown in complete medium (+FBS) and simple medium (-FBS). β-actin was used as loading control. Bar graphs show the mean values normalized to β-actin + SD. Data from 4 (+FBS) and 3 (-FBS) independent biological replicates, and 4 (+FBS) and 2 (-FBS) for β1-Integrin. One representative image from each experiment is shown. Statistical significance was determined using one-way ANOVA. *p≤0.05.

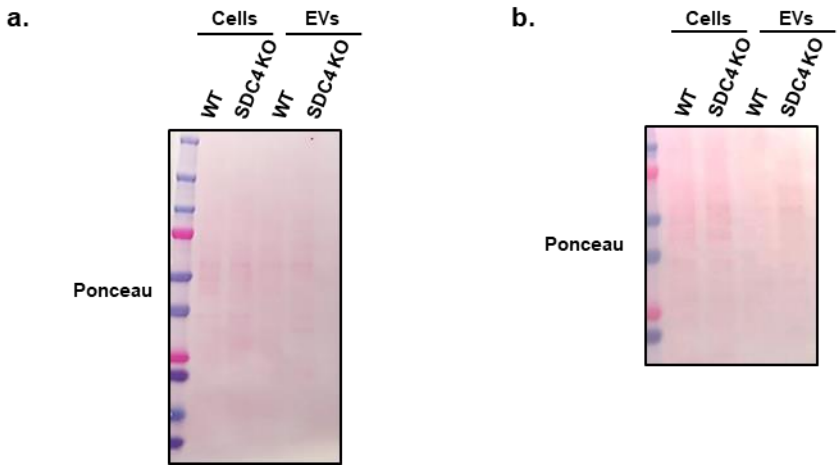



**Fig. S6. Ponceau staining of protein lysates from cells and EVs. (a)** Ponceau staining of the membrane used to evaluate SDC4 expression (relatively to Fig. 3a). **(b)** Ponceau staining of the membrane used to evaluate classical EV-markers (relatively to Fig. 3j).

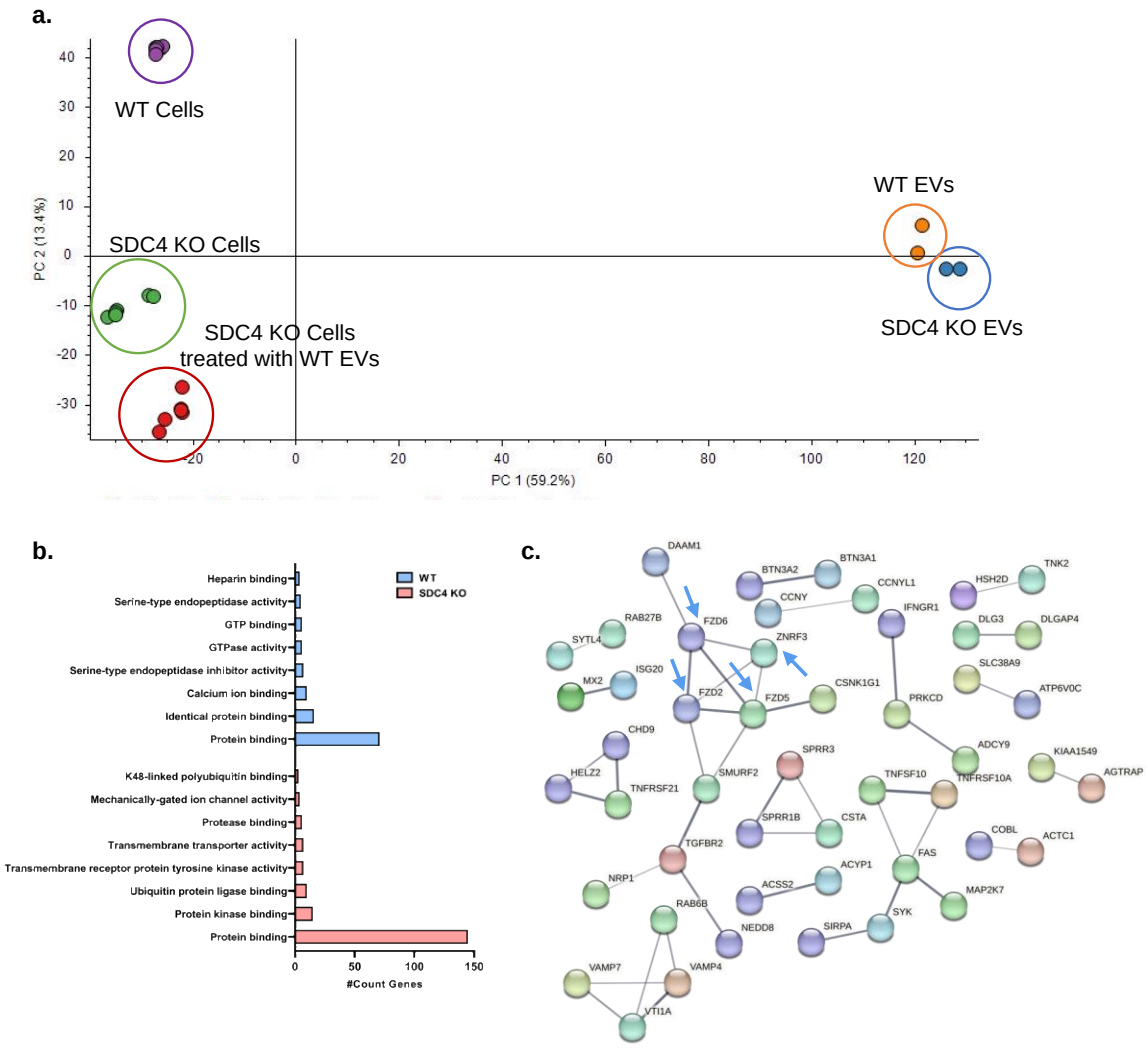



**Fig. S7. Data from MS-based proteomic analysis. (a)** Principal Component Analysis (PCA) based normalized abundance. **(b)** Functionally grouped network of gene ontology (GO) molecular function terms for the SDC4-specific proteome in WT EVs and exclusive proteins in KO EVs (Benjamini p<0.01). **(c)** Protein-protein interaction for the exclusive KO EVs proteome. Blue arrows highlight Frizzled receptors (FDZ2, FDZ5, and FDZ6) and E3 Ubiquitin-Protein Ligase (ZNRF3).

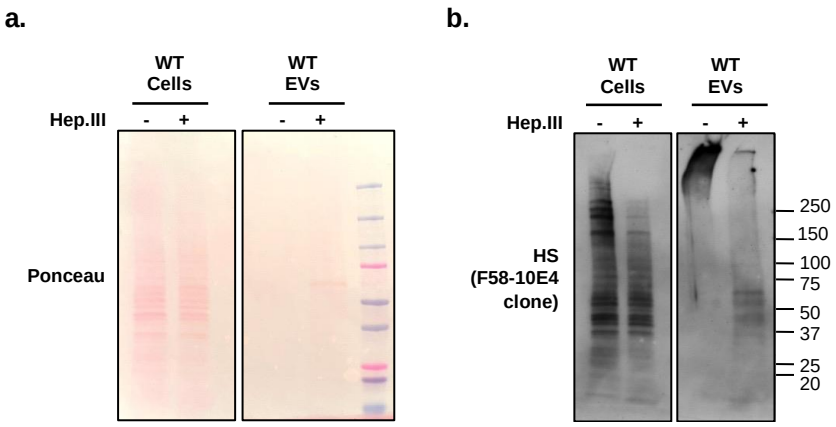



**Fig. S8. WB for heparan sulfate (HS) in WT cells and EVs treated with Heparinase III (Hep. III). (a)** Ponceau staining of the membrane used to evaluate HS in WT cells and EVs, with (+) and without (-) Hep.III digestion. **(b)** WB analysis of HS in WT cells and EVs, with (+) and without (-) Hep.III digestion.

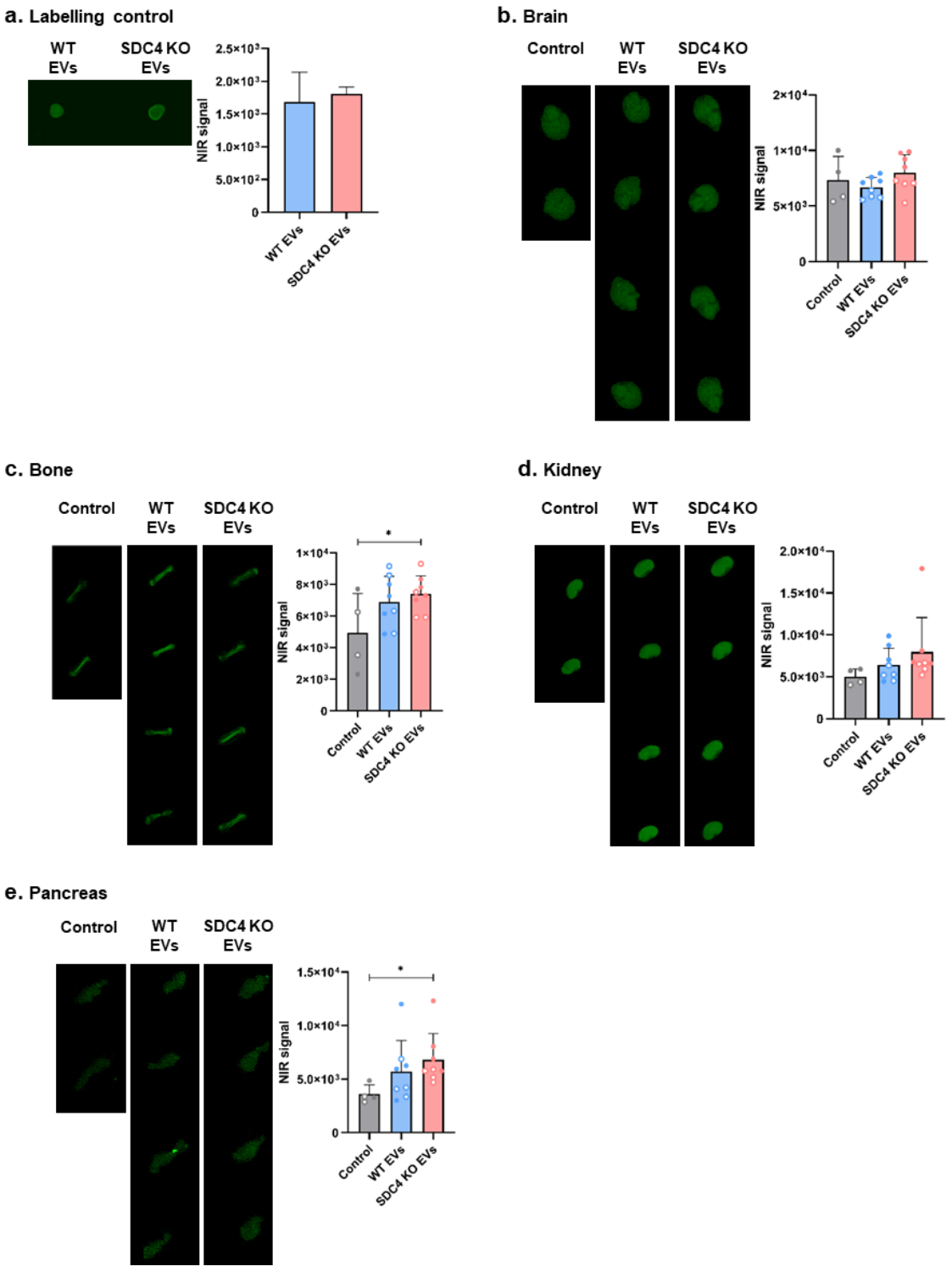
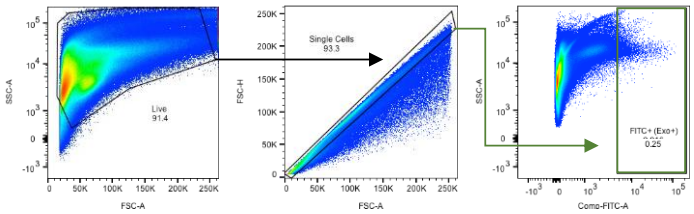



**Fig. S9. Biodistribution of EVs derived from WT and *SDC4* KO cells. (a)** Labelling control of WT and *SDC4* KO EVs was performed with Odyssey® CLx Imaging System (Licor) by fluorescence quantification. Bar graphs represents the average of NIR fluorescence signal

quantification + SD. It is shown n=2 from 2 independent experiments. Statistical significance was assessed using student's t-test. One representative image is shown. EVs biodistribution was assed using immunodeficient NOD SCID mice by measurement of fluorescence intensity of the collected organs **(b)** brain, **(c)** bone, **(d)** kidney and **(e)** pancreas. The NIR fluorescence quantification was assessed using Odyssey® CLx Imaging System. Bar graphs represent the average of fluorescence quantification of the organs + SD. Data from 2 independent biological replicates for WT and *SDC4* KO (each n=4, per condition group), and for control (each n=2). Statistical significance was assessed using student's t-test.*$p \leq 0.05$; ** $p < 0.01$; ***$p < 0.001$.

**a. Sample gating strategy for Exo FITC positive cells (same for all organs tested)**




**b. Sample gating strategy for Exo FITC positive cells (all organs tested → Tim-4+ cells were considered just for liver)**

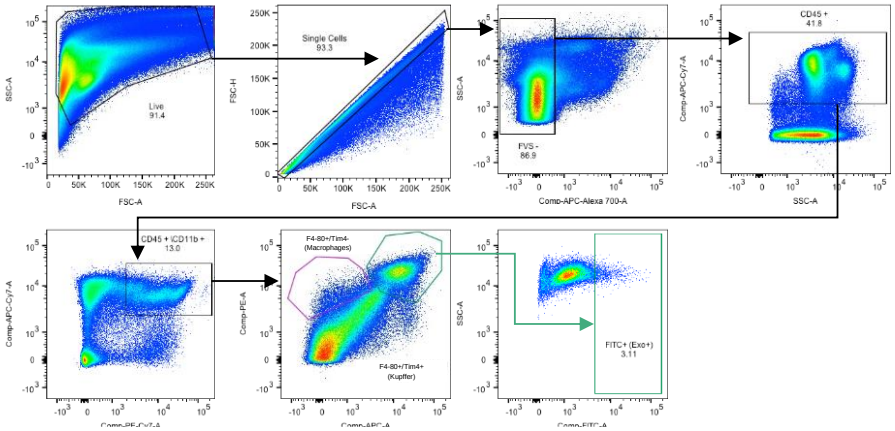



**Fig. S10. Sample gating strategy for FITC positive cells (Exo+), in flow cytometry assays. (a)** Direct gating strategy to evaluate all the live cells that are FITC positive (incorporated PKH67-labelled EVs) Single Cells, Live cells, FITC+ (Exo+). **(b)** Gating strategy used to evaluate Tim4-macrophages cell population that is FITC positive (incorporated PKH67-labelled EVs) - Single cells, FVS-, CD45+, CD45+ /CD11b+, F4/80+ /Tim4-, FITC+ (Exo+) and Kupffer cell population that is FITC positive (incorporated PKH67-labelled EVs) - Single cells, FVS-, CD45+, CD45+ /CD11b+, F4/80+ /Tim4+, FITC+ (Exo+).

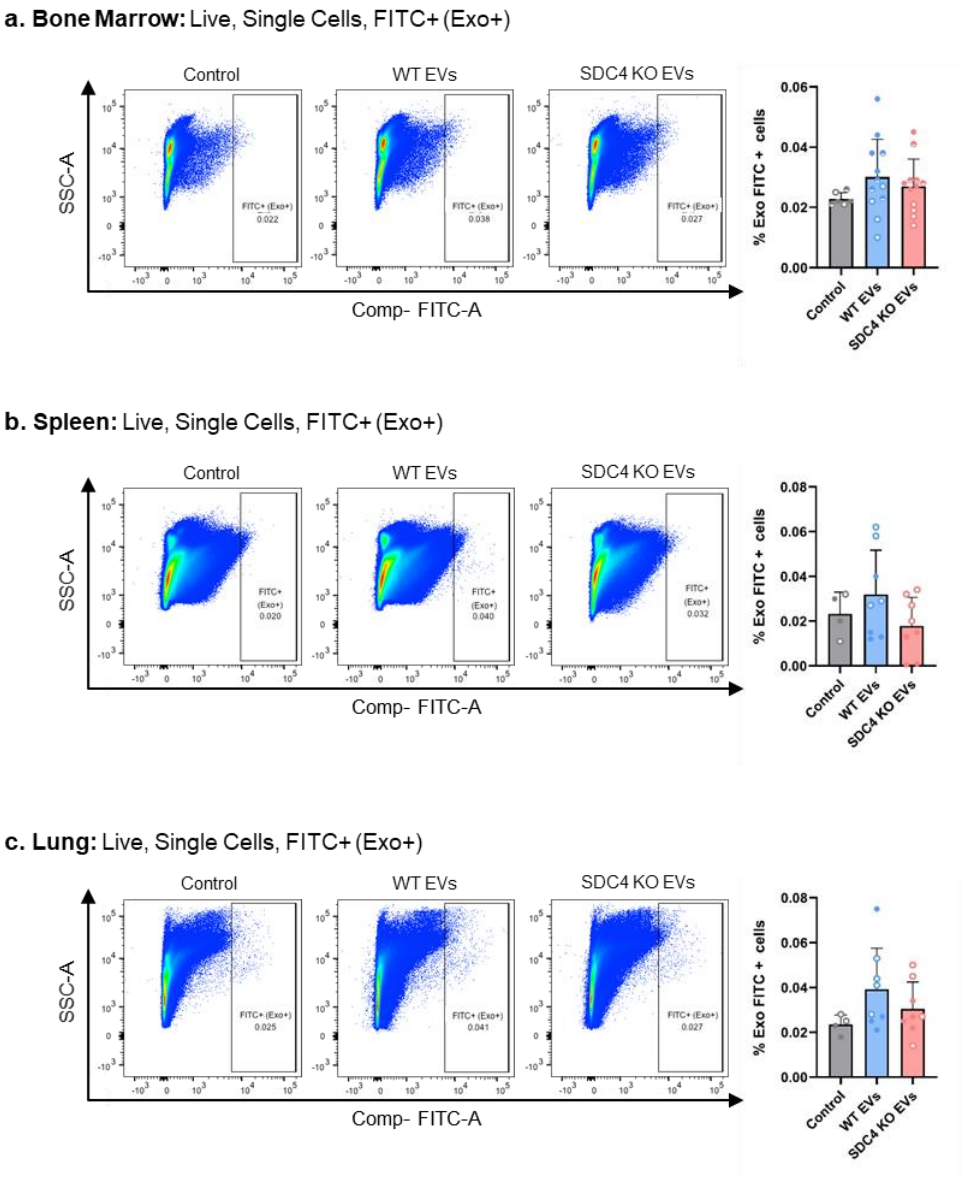



**Fig. S11. Determination of EV uptake by NOD SCID mice organs.** FACS analysis of **(a)** bone marrow, **(b)** spleen and **(b)** lung cells from mice injected with PKH67-labelled EVs isolated from either WT or *SDC4* KO cells. Cells were analyzed using the following gating strategy: Single Cells, Live cells, FITC+ (Exo+) cells. Bar graphs represent the average of the percentage of cells that incorporated EVs (FITC+ cells). Data for bone marrow from 3 independent biological replicates for WT and KO EVs (each n=4, per condition group) and for control (each n=2). Data for spleen and lung from 2 independent biological replicates for WT and KO EVs (each n=4, per condition group) and for control (each n=2). Statistical significance was assessed using student's t-test.

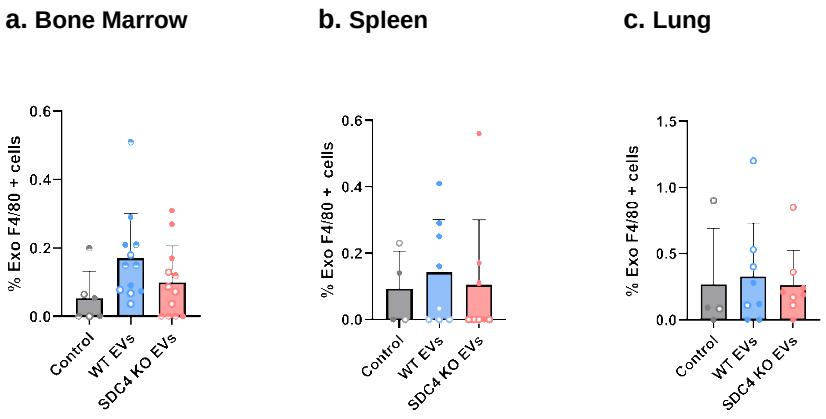



**Fig. S12. EV uptake by macrophages.** FACS analysis of **(a)** bone marrow, **(b)** spleen and **(c)** lung cells from mice injected with PKH67-labelled WT and KO EVs. Cells were analyzed using the following gating strategy: Single cells, FVS-, CD45+, CD45+ /CD11b+, F4/80+ /Tim4-, FITC+ (Exo+). Bar graphs represent the average of the percentage of cells that incorporated EVs (FITC+ cells). Data for bone marrow from 3 independent biological replicates for WT and KO EVs (each n=4, per condition group) and for control (each n=2). Data for spleen and lung from 2 independent biological replicates for WT and KO EVs (each n=4, per condition group) and for control (each n=2). Statistical significance was assessed using student's t-test with Welch's correction.

**Table S1. Association of SDC4 expression with clinicopathological variables in gastric cancer patients.**

| Clinicopathological variable | Total | SDC4 (8G3) | | | | |
|---|---|---|---|---|---|---|
| | | Negative | | Positive | | |
| | | n | % | n | % | *p*-value* |
| *Age (median, 35-89)* | | | | | | |
| <66 | 68 | 27 | 39.7 | 41 | 60.3 | 0.085 |
| ≥66 | 83 | 22 | 26.5 | 61 | 73.5 | |
| *Gender* | | | | | | |
| Male | 82 | 28 | 34.1 | 54 | 65.9 | 0.586 |
| Female | 70 | 21 | 30.0 | 49 | 70.0 | |
| *Stage* | | | | | | |
| IA-IB | 17 | 9 | 52.9 | 8 | 47.1 | 0.225 |
| II | 55 | 17 | 30.9 | 38 | 69.1 | |
| IIIA-IIIB | 57 | 15 | 26.3 | 42 | 73.7 | |
| IV | 23 | 8 | 34.8 | 15 | 65.2 | |
| *Primary tumour* | | | | | | |
| T1 | 11 | 6 | 54.5 | 5 | 45.5 | 0.178 |
| T2 | 17 | 7 | 41.2 | 10 | 58.8 | |
| T3 | 100 | 27 | 27.0 | 73 | 73.0 | |
| T4 | 20 | 8 | 40.0 | 12 | 60.0 | |
| *Regional lymph nodes* | | | | | | |
| pN0 | 44 | 18 | 40.9 | 26 | 59.1 | 0.083 |
| pN1 | 37 | 14 | 37.8 | 23 | 62.2 | |
| pN2 | 25 | 3 | 12.0 | 22 | 88.0 | |
| pN3 | 41 | 13 | 31.7 | 28 | 68.3 | |
| *Metastasis* | | | | | | |
| pM0 | 114 | 41 | 36.0 | 73 | 64.0 | 0.914 |
| pM1 | 23 | 8 | 34.8 | 15 | 65.2 | |
| *Lauren classification* | | | | | | |
| Intestinal (primary tumour) | 119 | 34 | 28.6 | 85 | 71.4 | **0.031** |
| Diffuse (primary tumour) | 21 | 11 | 52.4 | 10 | 47.6 | |
| Intestinal (LN metastasis) | 10 | 4 | 40 | 6 | 60 | 1.000 |
| Diffuse (LN metastasis) | 2 | 1 | 50 | 1 | 50 | |
| *Treatment* | | | | | | |
| Surgery alone | 105 | 33 | 31.4 | 72 | 68.6 | 0.750 |
| Surgery plus other tx | 47 | 16 | 34.0 | 31 | 66.0 | |

*Pearson Chi-Square

**Table S2. Kaplan-Meier analysis for overall survival stratified for subgroups of gastric cancer patients (represented values correspond to the mean).**

| | | Overall survival (95% CI) | | |
|---|---|---|---|---|
| **Subgroup** | **Total** | **SDC4 (8G3)** | | |
| | | **SDC4-negative** | **SDC4-positive** | **p-value*** |
| All patients | 67.0 (57.4-76.6) | 80.4 (64.2-96.7) | 59.7 (48.7-70.8) | **0.025** |
| *Age* | | | | |
| <66 | 64.0 (50.6-77.5) | 89.0 (67.6-110.4) | 42.1 (31.2-53.0) | **0.005** |
| ≥66 | 68.4 (55.5-81.2) | 66.3 (44.3-88.2) | 66.4 (55.5-81.2) | 0.672 |
| *Gender* | | | | |
| Male | 59.6 (48.8-70.4) | 78.7 (59.5-97.8) | 49.7 (38.0-61.3) | **0.028** |
| Female | 70.6 (56.2-85.1) | 76.9 (51.7-102.1) | 66.5 (49.6-83.4) | 0.378 |
| *Stage* | | | | |
| IA-IB | - | - | - | 0.480 |
| II | 81.9 (68.8-94.9) | 100.6 (81.2-120.1) | 69.3 (54.7-83.8) | 0.073 |
| IIIA-IIIB | 42.8 (31.1-54.5) | 52.1 (26.2-78.0) | 34.9 (25.7-44.0) | 0.464 |
| IV | 17.09 (10.9-23.3) | 18.0 (7.08-28.9) | 16.6 (8.76-24.4) | 0.618 |
| *Primary tumour* | | | | |
| T1 | - | - | - | 0.273 |
| T2 | 87.1 (64.4-109.6) | 111.4 (83.2-139.7) | 60.7 (43.6-77.7) | 0.129 |
| T3 | 57.9 (47.3-68.5) | 74.8 (53.5-96.1) | 47.8 (37.6-58.0) | 0.071 |
| T4 | 56.2 (32.5-79.8) | 41.1 (7.8-74.4) | 52.6 (30.9-74.3) | 0.302 |
| *Regional lymph nodes* | | | | |
| pN0 | 94.1 (79.8-108.3) | 115.8 (101.1-130.4) | 74.6 (57.1-92.1) | **0.016** |
| pN1 | 87.3 (67.6-106.9) | 57.2 (39.2-75.1) | 87.1 (62.2-112.1) | 0.967 |
| pN2 | 46.4 (30.3-62.5) | 46.7 (0-106.9) | 42.6 (28.9-56.3) | 0.867 |
| pN3 | 31.8 (20.4-43.2) | 43.3 (15.0-71.6) | 24.9 (17.2-32.6) | 0.469 |
| *Metastasis* | | | | |
| pM0 | 76.4 (66.0-86.8) | 90.9 (74.0-107.8) | 62.9 (52.0-73.7) | **0.044** |
| pM1 | 17.1 (10.9-23.3) | 18.0 (7.08-28.9) | 16.6 (8.76-24.4) | 0.618 |
| *Lauren classification* | | | | |
| Intestinal | 69.3 (58.4-80.1) | 88.8 (69.8-107.7) | 60.6 (48.5-72.8) | **0.018** |
| Diffuse | 56.9 (35.2-78.6) | 52.7 (23.1-82.3) | 47.6 (29.2-65.9) | 0.624 |
| *Treatment* | | | | |
| Surgery alone | 19.8 (18.3-21.2) | 20.5 (17.9-23.1) | 19.4 (17.7-21.2) | **0.026** |
| Surgery plus other tx | 18.3 (16.4-20.3) | 18.8 (15.6-22.0) | 18.1 (15.6-20.5) | 0.371 |

*log-rank (Mantel-Cox)

**Table S3. SDC4 abundance through proteomic analysis in WT and *SDC4* KO #2 cells and respective EVs (WT and KO); and in *SDC4* KO #2 cells treated with WT EVs.**

| Identification | Condition | Abundance (Normalized) | Abundance Average | Position and Peptide Sequence |
|---|---|---|---|---|
| EVs_WT1 | WT EVs_1 | 3,43E+08 | 2,45E+08 | [19-36]: A.ESIREREVIDPQDLLEGR.Y<br>[23-36]: R.ETEVIDPQDLLEGR.Y<br>[93-104]: R.AGSGSQVPTEPK.K<br>[115-128]: K.RISPVEESEDVSNK.V<br>[116-128]: K.ISPVEESEDVSNK.V<br>[129-143]: K.VSMSSTVQGSNIFER.T<br>[132-143]: M.SSTVQGSNIFER.T<br>[175-184]: K.KDEGSYDLGK.K<br>[176-184]: K.DEGSYDLGK.K<br>[190-198]: K.KAPTNEFYA.- |
| EVs_WT2 | WT EVs_2 | 1,47E+08 | | |
| EVs_KO1 | SDC4 KO EVs_1 | Not Found | n.a. | |
| EVs_KO2 | SDC4 KO EVs_2 | Not Found | | |
| WT1 | WT Cells_1_TR1 | 1,67E+07 | 1,67E+07 | [19-36]: A.ESIREREVIDPQDLLEGR.Y<br>[23-36]: R.ETEVIDPQDLLEGR.Y<br>[115-128]: K.RISPVEESEDVSNK.V<br>[116-128]: K.ISPVEESEDVSNK.V<br>[176-184]: K.DEGSYDLGK.K<br>[190-198]: K.KAPTNEFYA.- |
| WT1 | WT Cells_1_TR2 | 1,70E+07 | | |
| WT2 | WT Cells_2_TR1 | 1,76E+07 | | |
| WT2 | WT Cells_2_TR2 | 1,78E+07 | | |
| WT3 | WT Cells_3_TR1 | 1,52E+07 | | |
| WT3 | WT Cells_3_TR2 | 1,52E+07 | | |
| KO1 | SDC4 KO Cells_1_TR1 | Not Found | n.a. | |
| KO1 | SDC4 KO Cells_1_TR2 | Not Found | | |
| KO2 | SDC4 KO Cells_2_TR1 | Not Found | | |
| KO2 | SDC4 KO Cells_2_TR2 | Not Found | | |
| KO3 | SDC4 KO Cells_3_TR1 | Not Found | | |
| KO3 | SDC4 KO Cells_3_TR2 | Not Found | | |
| KO_WT1 | SDC4 KO Cells + WT EVs_1_TR1 | 1,40E+06 | 1,17E+06 | [93-104]: R.AGSGSQVPTEPK.K<br>[176-184]: K.DEGSYDLGK.K |
| KO_WT1 | SDC4 KO Cells + WT EVs_1_TR2 | 8,84E+05 | | |
| KO_WT2 | SDC4 KO Cells + WT EVs_2_TR1 | 1,01E+06 | | |
| KO_WT2 | SDC4 KO Cells + WT EVs_2_TR2 | 1,30E+06 | | |
| KO_WT3 | SDC4 KO Cells + WT EVs_3_TR1 | 1,08E+06 | | |
| KO_WT3 | SDC4 KO Cells + WT EVs_3_TR2 | 1,34E+06 | | |

**Table S4. Syntenin-1, Alix and SDC1 abundance through proteomic analysis in WT and *SDC4* KO #2 cells and respective EVs (WT and KO); and in *SDC4* KO #2 cells treated with WT EVs.**

| | | Syntenin-1 | | ALIX | | SDC1 | |
|---|---|---|---|---|---|---|---|
| **Identification** | **Condition** | **Abundance (Normalized)** | **Abundance Average** | **Abundance (Normalized)** | **Abundance Average** | **Abundance (Normalized)** | **Abundance Average** |
| **EVs_WT1** | **WT EVs_1** | 1,54E+10 | 1,54E+10 | 6,23E+09 | 6,22E+09 | 2,31E+07 | 3,58E+07 |
| **EVs_WT2** | **WT EVs_2** | 1,54E+10 | | 6,20E+09 | | 4,85E+07 | |
| **EVs_KO1** | **SDC4 KO EVs_1** | 1,75E+10 | 1,75E+10 | 7,93E+09 | 8,46E+09 | 1,15E+08 | 1,19E+08 |
| **EVs_KO2** | **SDC4 KO EVs_2** | 1,75E+10 | | 8,99E+09 | | 1,24E+08 | |
| **WT1** | **WT Cells_1_TR1** | 1,10E+10 | 1,93E+09 | 5,35E+08 | 5,59E+08 | 1,18E+07 | 1,35E+07 |
| **WT1** | **WT Cells_1_TR2** | 1,09E+08 | | 5,60E+08 | | 1,27E+07 | |
| **WT2** | **WT Cells_2_TR1** | 1,11E+08 | | 5,52E+08 | | 1,21E+07 | |
| **WT2** | **WT Cells_2_TR2** | 1,15E+08 | | 5,54E+08 | | 1,27E+07 | |
| **WT3** | **WT Cells_3_TR1** | 1,12E+08 | | 5,68E+08 | | 1,54E+07 | |
| **WT3** | **WT Cells_3_TR2** | 1,15E+08 | | 5,84E+08 | | 1,59E+07 | |
| **KO1** | **SDC4 KO Cells_1_TR1** | 1,12E+08 | 1,14E+08 | 6,20E+08 | 6,40E+08 | 1,13E+07 | 1,15E+07 |
| **KO1** | **SDC4 KO Cells_1_TR2** | 1,14E+08 | | 6,30E+08 | | 1,12E+07 | |
| **KO2** | **SDC4 KO Cells_2_TR1** | 1,18E+08 | | 6,52E+08 | | 1,05E+07 | |
| **KO2** | **SDC4 KO Cells_2_TR2** | 1,19E+08 | | 6,53E+08 | | 1,12E+07 | |
| **KO3** | **SDC4 KO Cells_3_TR1** | 1,09E+08 | | 6,36E+08 | | 1,22E+07 | |
| **KO3** | **SDC4 KO Cells_3_TR2** | 1,12E+08 | | 6,48E+08 | | 1,24E+07 | |
| **KO_WT1** | **SDC4 KO Cells + WT EVs_1_TR1** | 1,30E+08 | 1,26E+08 | 5,46E+08 | 5,48E+08 | 1,07E+07 | 1,06E+07 |
| **KO_WT1** | **SDC4 KO Cells + WT EVs_1_TR2** | 1,35E+08 | | 5,28E+08 | | 8,31E+06 | |
| **KO_WT2** | **SDC4 KO Cells + WT EVs_2_TR1** | 1,29E+08 | | 5,25E+08 | | 8,89E+06 | |
| **KO_WT2** | **SDC4 KO Cells + WT EVs_2_TR2** | 1,23E+08 | | 5,78E+08 | | 1,25E+07 | |
| **KO_WT3** | **SDC4 KO Cells + WT EVs_3_TR1** | 1,17E+08 | | 5,45E+08 | | 1,00E+07 | |
| **KO_WT3** | **SDC4 KO Cells + WT EVs_3_TR2** | 1,20E+08 | | 5,66E+08 | | 1,31E+07 | |

**SI Materials and Methods**

**Gastric cancer tissue samples.** This study was performed retrospectively in a series of 152 formalin-fixed paraffin-embedded (FFPE) gastric tumour tissues obtained from archived blocks at the Portuguese Institute of Oncology-Porto (IPO-Porto), Portugal. Clinicopathological information was obtained from patients' clinical records, upon IPO ethics committee approval (reference 87/17 approved on 23 March 2017). All procedures were carried out following the rules of the Declaration of Helsinki of 1975, revised in 2013.The gastric tumours obtained from IPO-Porto were surgically removed from 82 men and 70 women, ranging from 39 to 89 years of age with a mean age of 64.2 ± 12.1 (median 66 years), admitted and treated at the IPO-Porto between 2004 and 2009 (Supplementary Table 1). Approximately 70% of the patients were treated only with surgery, 17% performed chemotherapy in palliative setting, 10% as adjuvant treatment and only 3% were treated with neoadjuvant chemotherapy and/or radiotherapy prior to the surgery. For analysis purposes we dichotomized treatment variable into "surgery alone" and "surgery plus other tx". Overall survival (OS) was defined as the period between the date of surgery and the date of patient death by cancer or last follow-up.

**Mice.** Male and female NOD SCID mice were purchased from the Charles River Laboratories (Barcelona, Spain) and used at 6-8 weeks of age. NOD SCID mice were bred and maintained at the Champalimaud Foundation animal facility.

**Immunohistochemistry.** Expression of SDC4 was assessed in 152 cases of human gastric carcinoma and 12 lymph node metastases. Tissue slides were dewaxed and rehydrated. Treatment with 3% $H_2O_2$ in methanol for 20min was performed to block endogenous peroxidase. After blocking, tissues were incubated with the primary antibody anti-SDC4 8G3 clone (1)in 5% bovine serum albumin (BSA) in phosphate-buffered saline (PBS). Sections were then incubated with 1:200 biotin-labelled secondary rabbit anti-mouse antibody (Dako) in PBS with 5% BSA for 30min and with ABC kit (Vector Labs) for 30min. Sections were stained for 0.05% 3,3'-diaminobenzidinetetrahydrochloride (DAB) containing 0.01% $H_2O_2$ and counterstained with Mayers' haematoxylin solution, dehydrated and mounted. SDC4 expression was viewed and

assessed by two sample-blind pathologists and classified as either negative or positive (whenever at least 1% of the cells were positive). Slides scanning was performed using D-Sight F 2.0 (Menarini Diagnostics).

***In silico* analysis of *SDC4* expression in gastric cancer tissues.** mRNA *SDC4* transcript data was downloaded from the Cancer Genome Atlas (TCGA; https://gdc.cancer.gov/; data release 36), comprised a total of 412 stomach adenocarcinoma samples and 36 normal gastric tissues.

RNA sequencing data was quantified as FPKM (fragments per kilobase million) by STAR. *SDC4* transcript levels were correlated with the available clinicopathological data, including Lauren classification and TCGA molecular subtype, presence of metastasis, recurrence, MSI and aneuploid status.

***In silico* analysis of the prognostic impact of *SDC4* in gastric cancer.** To determine the clinical relevance of *SDC4* expression in gastric cancer we used the TCGA-STAD dataset. Patients were stratified in two groups with high or low expression. The top 25% of patient samples with highest expression were considered as high *SDC4* expression group. The bottom 25% of patients with the lowest expression values were considered as low *SDC4* expression group. Survival information was extracted from clinical files and information on recurrence was downloaded from the cBioPortal. Kaplan–Meier analysis was used to calculate the survival rate of stratified patients and plot the survival curve. For overall survival (OS) analysis, only individuals with survival information and follow-up information above to zero were keep (n= 387). In disease-free survival (DFS) plots, deceased individuals without information on recurrence were removed from the analysis (n= 240). Multivariate Cox regression with backward elimination was performed to assess the independent prognostic value (OS and DFS) of *SDC4* expression in all GC patients and in the intestinal subtype tumour patients. The clinicopathological parameters such as Lauren classification, gender, age and disease stage were used as covariables. The prognostic value of SDC4 expression was further validated in the KM plotter database (2), with expression data for 740 patients and survival data for 592 patients (GSE14210, GSE15459, GSE22377, GSE29272

and GSE51105). Patients were stratified in two groups with high or low expression, as described above. OS and disease-free survival (DFS) were analysed for all gastric cancer patients and then divided according to the Lauren subtype classification (intestinal n=179, diffuse n=106).

***In silico* analysis of the prognostic impact of EV signatures.** Survival analysis of signature genes was assessed for the WT (93 genes) and *SDC4* KO (194 genes) signature sets using the STAD-TCGA database. Each signature was used to construct the predictive component by Cox proportion hazard regression (3). The risk scores were calculated by a linear combination of the mean normalized gene expression values for the selected genes, weighted by their estimated Cox regression coefficients. Kaplan-Meier survival analyses for OS and DFS was carried for the samples classified into “high” and “low” risk groups according to the upper and lower risk score quartiles of all samples. All analyses were implemented in R (v4.1.2) using the “survival” and “survminer” packages.

**Cell culture.** The cell lines AGS and MKN45 were obtained from the Japanese Collection of Research Bioresources (JCRB); NCI-N87, MKN74 and RAW 264.7 were obtained from the American Type Culture Collection (ATCC); Kupffer cells (HUKCCS) were obtained from Life Technologies. All cells lines were grown in RPMI 1640 Medium (Thermofisher Scientific), supplemented with 10% FBS (Invitrogen) and maintained at 37ºC in an atmosphere of 5% $CO_2$. Cultures were routinely screened for mycoplasma. Cell line identity validation was performed by short tandem repeat (STR) profiling using the PowerPlex® 16 HS System kit (Promega).

***SDC4* KO by CRISPR/Cas9.** Briefly, 3 different gRNA targeting exon1 of *SDC4* were design, using DESKGEN platform, and gRNA (5’-CGGAGCCCTACCAGACGATG-3’) was validated. Then, cells were transfected with a plasmid containing Cas9 endonuclease + GFP and a plasmid containing the validated gRNA. Fluorescence-activated cell sorting was used to obtain GFP-positive single cell clones. Indel detection was performed by Restriction Fragment Length Polymorphism (RFLP) analysis by tri primer PCR (SDC4fw_ext: AGCTGACCGGCAGCAAAAAAGGACACGCGGTGAGTTAG; SDC4rv: AGTCTCAAGGCATGGTCACC; Famfw: 6-FAM-AGCTGACCGGCAGCAAAATTG) (4), and indels

of the selected clones were validated by Sanger sequencing and Tracking of Indels by Decomposition (TIDE) analysis.

**siRNAs transfection.** Cells were transfected with siRNA triplets targeting *SDC4*: siRNAA (5'-AGAAACUAGAGGAG- AAUGAG GUUAT-3'), siRNAB (5'-CUAUUCUAGAGAACUAAACUGGCTT-3') and siRNAC (5'-GGAUUGGAUCACUUCUUAAACUUCC-3') (SR304301, OriGene Technologies) or siScramble (5'-CGUUAAUCGCGUAUAAUACGCGUAT-3') (SR30004, OriGene Technologies), using Lipofectamine 2000 reagent (Invitrogen) according to the manufacturer's instructions. Cells were analyzed 48h and 72h after transfection.

**RNA extraction and quantification.** RNA extraction was performed using TRI Reagent® (T9424, Sigma-Aldrich®) according to manufacturer's instructions. RNA quantification was done using NanoDrop ND-1000® UV-Vis spectrophotometer (NanoDrop Technologies).

**Reverse Transcriptase PCR (RT-PCR) and Real Time PCR (qPCR)**. The syndecan members were amplified using the primer sequences SDC1 (fw: 5'-ATGGCTCTGGGGATGACTCT-3'; rv: 5'-GTGGGAATAGCCGTCAGGAG-3'), SDC2 (fw: 5'-TCAGACAAAGTCACCTGAAGA-3'; rv: 5'-AACTCCACCAGCAATGAC-3'), SDC3 (fw: 5'-TGACATCCCTGAGAGGAGCA-3'; rv: 5'-GCTACCACCTCATTGGCTGT-3') and SDC4 (fw: 5'-CCGGAGCCCTACCAGACGAT-3'; rv: 5'-AGGCACCAAGGGATGGACAA-3'). The 18S gene was used as control (fw: 5′-CGCCGCTAGAGGTGAAATTC-3′; rv: 5′-CATTCTTGGCAAATGCTTTCG-3′). 1 µg of RNA from each cell condition was converted into cDNA via reverse transcription by incubation with random hexamers/primers, using the SuperScript™ IV Reverse Transcriptase Kit (Invitrogen). After RNA to cDNA conversion, real-time PCR was performed using for each condition: 4 µL of cDNA diluted 1:10 in ultrapure water, 0.6 µL of each primer at 10 mM (Table 1), 10 µL of PowerUp™ SYBR™ Green Master Mix (Thermo Fischer Scientific) and ultrapure water up to a final volume of 20 µL. The real-time PCR run was performed using a 7500 Fast Real-Time PCR System (Applied

Biosystems). RQ values were determined for each gene and 18S ribosomal RNA was used as a housekeeping gene to normalize relative gene expression.

**Western Blotting (WB).** Protein extracts were prepared from EVs and cells grown in complete medium (+ FBS) or grown in simple medium for 48h (-FBS). Briefly, cells were washed and scraped on ice in lysis buffer 17 (R&D Systems) supplemented with PhosSTOP and Complete Phosphatase/Protease inhibitor cocktails (Roche Diagnostics GmbH). The whole lysate was centrifuged at 16000g for 15min and stored at -80°C. Protein concentration was determined using the BCA protein kit (Pierce, Thermo Fisher Scientific). Protein lysates were separated by gradient (4-15%) (BioRad) gel electrophoresis. After electrophoresis mediated separation, proteins were blotted to nitrocellulose membranes, which were blocked with 5% Milk or BSA (diluted in PBS-Tween 0.1%) for 1h and incubated overnight at 4°C with the following primary antibodies: anti-SDC4 mouse (8G3 clone) (1); anti-EphA4 (1:250) (Invitrogen); anti-phosphorylated EphA4 – Tyr602 (1:1000) (ECM Biosciences); anti-EGFR rabbit (1:1000); anti-phosphorylated EGFR – Tyr1068 (1:1000); anti-FAK rabbit (1:1000); anti-phosphorylated FAK rabbit – Tyr397 (1:1000); anti-Alix mouse (1:1000) (Cell Signalling); anti-β1-Integrin mouse (1:10000) (Biosciences 610468); anti-Paxillin (1:600) (Merck Millipore); anti-syntenin-1 (1:200); anti-Cytochrome-c (1:200); anti-PKCα (1:1000) (Santa Cruz Biotechnology); anti-HSP70 (1:1000); anti-CD9 (1:1000); anti-CD81 (1:1000) (System Biosciences),anti-SDC1 (1:125) (Abcam) and anti-HS (1:600) (Amsbio) primary antibodies. As loading control, membranes were incubated with anti-β-Actin (1:2000) (Santa Cruz Biotechnology) or anti-α-Tubulin (1:10000) (Sigma-Aldrich). After incubation overnight, membranes were washed three times and incubated 1h with either anti-mouse IgG (1:5000), anti-mouse IgG1 (1:10000) or anti-rabbit IgG (1:25000) secondary antibodies (all from Jackson ImmunoResearch Laboratories). Immunodetection of proteins was performed using ECL Western Blotting Detection Reagents (GE Healthcare Life Sciences). WB analysis was performed using Quantity One 1-D Analysis Software (Bio-Rad). The phospho-protein band quantification was normalized for the respective total protein amount.

**Flow cytometry.** Cell pellets were suspended in 5% BSA/PBS and incubated with SDC4 antibody 8G3 clone (1:200) for 45min at room temperature (RT). Incubation with rabbit anti-mouse FITC (Dako) (1:50) for 45min was carried out on RT. Negative control was performed using only the secondary antibody. Washing steps were performed with 2% BSA/PBS. The stained cells were analysed on a FACS (Fluorescence-Activated Cell Sorting) Calibur flow cytometer and data processed by using CellQuest Pro v8.0 (Becton Dickinson) and FlowJo v10.0 software.

**Immunofluorescence.** Cells were fixed with 4% paraformaldehyde (PFA) (Alfa Aesar) for 20min. After the blocking procedure, with goat anti-mouse serum (1:250) (Dako), permeabilization of cell membranes was performed using 0.2% triton X-100 in PBS for 10 min. PBS-Tween 0.01% was used for washing steps. The cells were stained with SDC4 8G3 clone antibody (1:200) overnight at 4ºC. Primary antibodies were diluted in 5% BSA/PBS solution. Goat anti-mouse Alexa 488 (Thermo Fisher Scientific) was used as secondary antibody, for 1h at RT. Phalloidin 568 (Invitrogen) was used to stain actin filaments and nuclei were stained with DAPI (Sigma-Aldrich). Cells were visualized using a Zeiss Imager Z.1 microscope (Zeiss).

**Enzymatic digestions.** HS enzymatic digestion was performed in WT and *SDC4* KO protein extracts and EVs. Protein extracts were treated with calcium acetate and Hep.III (0.005 mU/µL). After incubation overnight at 37°C with agitation, samples were stored at -80°C.

**Annexin V viability assay.** Cell viability was evaluated by the Annexin V-FITC staining assay. $1.5x10^5$ cells were seeded per well on 6-well plates (Corning Incorporated Costar). 48h following cell seeding, cells were enzymatically detached with trypsin (Biowest), resuspended in their culture medium and centrifuged at 100g for 5min. The supernatant was rejected, and the cell pellet was resuspended in fresh medium. Cells were washed twice in PBS, followed by another two sets of washes with Annexin V Binding Buffer (BioLegend). Each wash was followed by centrifugation at 100g for 5min. The resulting cell pellet was resuspended and incubated with Annexin V-FITC (BioLegend), diluted in a ratio of 1:40 in Annexin V Binding Buffer, for 15min at

RT. Cells were filtered and data was acquired via FACS Canto II cytometer (Becton, Dickinson and Company) and analysed with the FlowJo Software v10.0.

**Proliferation assay.** Cell proliferation was determined using Click-iT™ Plus EdU Alexa Fluor™ 647 Flow Cytometry Assay Kit (Molecular Probes), following manufacturer's instructions. Briefly, at day one $2.5x10^5$ cells/mL were seeded in a T25 flask in RPMI supplemented with 10% FBS and left to grow at 37ºC with 5% $CO_2$. In parallel, cells were also grown in simple media (without FBS). After 48h, 10µM of EdU was added to cells for a period of 1h30min prior to harvesting. EdU non-treated cells served as control while cells in simple media served as a cell arrest control. Detection of EdU incorporation into the DNA was performed according to the manufacturer's instructions. Briefly, harvested cells were washed in 1% BSA/PBS and fixed in 100 µL Click-iT1 fixative. After a 15min incubation step at RT in the dark, cells were washed again in 1% BSA/PBS and then resuspended in 100 µL saponin-based permeabilization and wash reagent. Click-iT1 EdU reaction cocktail was prepared according to the manufacturer's instructions and added. Samples were incubated for 30min at RT in the dark and washed with saponin-based permeabilization and wash reagent. Cell pellet was resuspended in 300µL saponin-based permeabilization and wash reagent and acquired using FACS Canto II cytometer, and further analysed using FlowJo software v10.0.

***In vitro* migration assay.** Migration assays were performed using the Ibidi culture-inserts family. ibidi µ-Slide 8 well with different coatings, ibiTreat surface (polymer) (Cat. No 80826), collagen IV (Cat. No 80822), fibronectin (Cat. No 80823) and poly-D-lysine (Cat. No 80824) were used with culture-inserts for self-insertion (Cat. No 80209). WT and SDC4 KO cells were seeded at a density of $2x10^4$ cells/insert with RPMI supplemented with 10% FBS. Cells were monitored during 18h for wound healing. The percentage of free area was calculated by measuring the free area at each time-point and normalizing for the initial free area measured using Image J software. The statistical analysis was performed using GraphPad Prism (version 8.0). Results are shown as average ± SD. Two-way ANOVA was used as statistical test.

***In vitro* invasion assay.** Invasion assays were performed in BD Biocoat Matrigel (Corning) invasion chambers with 8µm of diameter pore size, in 24-well plate. WT and SDC4 KOcells were seeded at a density of $7x10^4$ cells/insert with RPMI medium, using RPMI medium supplemented with 10% FBS as chemo-attractive in the lower part of the insert, for 18h. After incubation, non-invading cells were carefully removed from the upper part of the insert. The inserts were washed and fixed with cold methanol for 10min on ice. The membranes were mounted on slides with Vectashield with DAPI (Vector Laboratories). Assay with EV pre-treatment was performed as described above with a previous 24h co-culture of the cells with with $8x10^8$ particles/mL WT or *SDC4* KO #2 EVs. The total number of invasive nuclei was counted using a Leica DM2000 microscope under 200x magnification (Leica). Results are shown as average ± SD. T-student test was used for statistical analysis performed on GraphPad version 8.0.

**RTK phosphorylation array.** The activation state of important RTKs was determined by using the Proteome Profiler Human Phospho-RTK Array Kits (R&D Systems #ARY001B) following the manufacturer's instructions. Cells proteome was isolated from confluent cells (1x107 cells/mL) as described above. The RTK array membranes were incubated with the Array Buffer 1 for 1h at RT with continuous shaking. 300µg of whole cell lysates were diluted in Array buffer 1 to a final volume of 1.5mL and incubated on the RTK array membranes overnight at 4°C with continuous shaking. Membranes were then washed thrice with 1x Wash Buffer, 10min each wash, followed by incubation with 1:5000 Anti-Phospho Tyrosine-HRP detection antibody diluted in Array buffer 2, for 2h at RT with continuous shaking. Membranes were washed again thrice with 1x Wash Buffer and finally were developed and visualized with ECL chemiluminescent detection reagent. For signal quantification, densitometry was evaluated and the phosphorylation fold change of each RTK was calculated for KO cell model by comparison with the RTK phosphorylation values determined for the WT sample.

**Isolation of EVs by ultracentrifugation and size exclusion chromatography.** EV isolation and purification was performed by Ultracentrifugation (UC), followed by Size Exclusion Chromatography (SEC). Specifically, conditioned medium was collected from WT and *SDC4* KO

#2 cells (cells seeded at a concentration of $2x10^6$ cells/ 150 $cm^2$ plate) grown without FBS for 48h. Conditioned medium was submitted to two initial centrifugations, at 500g for 10min and 3000 g for 20min (at 10ºC) to remove cell debris. The supernatant was then filtered using 0.22µm filtration system (Corning). Media was ultracentrifuged for 2h20min at 100000 g (at 10ºC) and supernatant was discarded. For EV-tracking experiments, EVs were fluorescently labelled using either PKH67 membrane dye (PKH67GL-1KT, Sigma-Aldrich), or NIR815 membrane dye (CellVue® NIR815, Licor). A washing with saline solution (0.9% of NaCl) was performed for 2h at 100000 g (at 10ºC), after 2h at 4ºC to "auto-resuspend". Pellets were pooled together and resuspended in saline solution. All ultracentrifugations were performed with 45Ti fixed angle rotor (Beckman-Coulter) and washing steps with 70Ti fixed angle rotor. After a "auto-resuspension" step, during 2h at 4ºC, SEC was performed according to manufacturer's instructions of the qEV column (iZon). Briefly, 500µL of resuspended pellet in saline solution was overlaid on top of the qEV column. Fractions from 6 to 12 were collected and pooled together. After, samples were concentrated by ultrafiltration with 10K Amicon Ultra-15 centrifugal filter units (EMD Milipore).

**Nanoparticle Tracking Analysis (NTA)**. All EV samples were analysed for particle concentration (particles per millilitre) and size distribution using a NanoSight NS300 system (Malvern Technologies). The NanoSight NS300 system was configured with a Red laser (638 nm) and a high sensitivity scientific sCMOS camera. For NTA analysis of the EVs, saline solution (0.9% of NaCl) as diluent and a syringe pump with constant flow injection were used. Five to nine videos of 30s using a camera level of 16 and a threshold between 5 and 7 were captured per sample. The videos were recorded and analysed by NTA software version 3.4.

**EV protein concentration quantification by Bicinchoninic Acid (BCA).** Protein quantification of the EV preparations was assessed by BCA assay (PierceTM BCA Protein Assay kit, Thermo Fisher Scientific) following manufacturer's instructions. Absorbance was measured at the wavelength of 562nm.

**Negative-staining Transmission Electron Microscopy (TEM).** EVs were visualized by TEM using negative staining. For this, 10µL of undiluted sample was mounted on Formvar/carbon film-

coated mesh nickel grids (Electron Microscopy Sciences). After 20min settling, the liquid excess was removed with filter paper and 10µL of 1% uranyl acetate was added onto the grids. Visualization was carried out on a JEOL JEM 1400 TEM at 120 kV (Tokyo, Japan). Images were digitally recorded using a CCD digital camera (Orious 1100W Tokyo, Japan).

**Gold-immunolabeling microscopy.** WT and *SDC4* KO #2 EVs were deposited on Formvar/carbon film-coated mesh nickel grids (Electron Microscopy Sciences, Hatfield, PA, SA) and let to settle for 20min. EVs were fixed with 2% PFA/PBS for 20min at RT. After, EVs were incubated with 50mM glycine/PBS (4x 4 min), and then with 10% BSA/PBS for 10 min. After 2h of anti-SDC4 (8G3 clone) primary antibody incubation, EVs were washed with 0.5% BSA/PBS (6x 3 min). EVs were incubated with anti-mouse Gold-conjugated secondary antibody (BBI Solutions, EM.GMHL15) in 5% BSA/PBS for 1h and then washed with 0.5% BSA/PBS (6x 3 min). Next, 10µL of 1% uranyl acetate was added onto the grids for 10sec. Visualization was carried out on a JEOL JEM 1400 TEM at 120 kV (Tokyo, Japan). Images were digitally recorded using a CCD digital camera.

**Protein identification using LC-MS/MS.** Proteins were solubilized with 100 mM Tris pH 8.5, 1% sodium deoxycholate, 10 mM tris(2-carboxyethyl)phosphine (TCEP) and 40 mM chloroacetamide for 10 minutes at 95ºC at 1000 rpm (Thermomixer, Eppendorf). Each sample was processed for proteomic analysis following the solid-phase-enhanced sample-preparation (SP3) protocol as described in (5). Enzymatic digestion was performed with Trypsin/LysC overnight at 37ºC at 1000 rpm. Protein identification and quantitation was performed by nanoLC-MS/MS. This equipment is composed by an Ultimate 3000 liquid chromatography system coupled to a Q-Exactive Hybrid Quadrupole-Orbitrap mass spectrometer (Thermo Scientific, Bremen, Germany). Samples were loaded onto a trapping cartridge (Acclaim PepMap C18 100Å, 5 mm x 300 µm i.d., 160454, Thermo Scientific) in a mobile phase of 2% ACN, 0.1% FA at 10 µL/min. After 3 min loading, the trap column was switched in-line to a 50 cm by 75µm inner diameter EASY-Spray column (ES803, PepMap RSLC, C18, 2 µm, Thermo Scientific) at 250 nL/min. Separation was generated by mixing A: 0.1% FA, and B: 80% ACN, with the following gradient: 5 min (2.5% B to 10% B),

120 min (10% B to 30% B), 20 min (30% B to 50% B), 5 min (50% B to 99% B) and 10 min (hold 99% B). Subsequently, the column was equilibrated with 2.5% B for 17 min. Data acquisition was controlled by Xcalibur 4.0 and Tune 2.11 software (Thermo Scientific). The mass spectrometer was operated in data-dependent (dd) positive acquisition mode alternating between a full scan (m/z 380-1580) and subsequent HCD MS/MS of the 10 most intense peaks from full scan (normalized collision energy of 27%). ESI spray voltage was 1.9 kV. Global settings: use lock masses best (m/z 445.12003), lock mass injection Full MS, chrom. peak width (FWHM) 15s. Full scan settings: 70k resolution (m/z 200), AGC target 3e6, maximum injection time 120 ms. dd settings: minimum AGC target 8e3, intensity threshold 7.3e4, charge exclusion: unassigned, 1, 8, >8, peptide match preferred, exclude isotopes on, dynamic exclusion 45s. MS2 settings: microscans 1, resolution 35k (m/z 200), AGC target $2x10^5$, maximum injection time 110 ms, isolation window 2.0 m/z, isolation offset 0.0 m/z, spectrum data type profile. The raw data was processed using Proteome Discoverer 2.5.0.400 software (Thermo Scientific) and searched against the UniProt database for the Homo sapiens reviewed Proteome 2021_03 with 20371 entries and the NIST human spectral library. A common protein contaminant list from MaxQuant was also considered. The MSPepSearch and Sequest HT search engines were used to identify tryptic peptides. The ion mass tolerance was 10 ppm for precursor ions and 0.02 Da for fragment ions for both software. Maximum allowed missing cleavage sites was set 2. Cysteine carbamidomethylation was defined as constant modification. Methionine oxidation, deamidation of glutamine and asparagine, peptide terminus glutamine to pyroglutamate, and protein N-terminus acetylation, Met-loss and Met-loss+acetyl were defined as variable modifications. Peptide confidence was set to high. The processing node Percolator was enabled with the following settings: maximum delta Cn 0.05; decoy database search target FDR 1%, validation based on q-value. Protein label free quantitation was performed with the Minora feature detector node at the processing step. Precursor ions quantification was performing at the processing step with the following parameters: Peptides to use all peptides, precursor abundance based on intensity, and normalization based on total peptide amount. For data analysis the following filters

were applied: (i) common contaminants were removed; (ii) protein identification p-value ≤ 0.01; (iii) unique/razor peptides were equal or greater than 1.

Venny (version 2.1.0) plug-in was used to build the Venn diagram (6) and Database for Annotation, Visualization and Integrated Discovery (DAVID) GO (version2022q4) was used to visualize the functionally organized Gene Ontology (GO) terms for the exclusive proteomes of WT and KO EVs. Analyses were conducted for GO_Biological Process and Molecular Function with Benjamini step down correction and $p<0.01$, minimum number of proteins 2, in order of genes count. Protein-protein interaction for the specific proteome of WT and KO EVs was performed using String (version 11.5).Only know interactions from databases and experimentally determined were considered (line thickness indicates the strength of data support). The minimum required interaction score was 0.400.

***In vitro* EV uptake assay.** EVs derived from WT and *SDC4* KO #2 cells were labelled with PKH67 green fluorescent dye. WT, RAW and Kupffer cells were seeded on 24-well plates ($10\times10^4$ cells/mL). After incubation overnight, cells were treated with EVs ($2.14x10^8$ particles/mL). EVs were untreated, or treated with Hep.III (as described above), or blocked with HS antibody (F58-10E4, Amsbio) for 48h, or incubated with Heparin (H3393, Sigma). After 2h of incubation, cells were fixed with 4% PFA, washed, and labelled with Hoechst for nuclei stain, and with HCS Cell mask plasma membrane stain deep red for 15min (both diluted in PBS). Images of 25 fields/well and per condition were taken by IN Cell Analyzer 2000 (GE Healthcare) or Operetta CLS High Content Analysis System (PerkinElmer).

**Biodistribution of EVs in mice.** NOD SCID mice were retro-orbitally injected with NIR815-stained or PKH67-labelled EVs, WT and *SDC4* KO #2, and PBS (control), using $4.6x10^8$ particles/mice. Mice were sacrificed 18h after the NIR815-stained EV injection. The liver, lung, spleen, pancreas, bone (femur), brain and kidney were isolated to measure the fluorescence signal from EVs trapped in the organs collected, using Odyssey® CLx Imaging System (Licor) with solid-state diode laser at 785nm. For labelling control, EVs previously labelled with PKH67 were seeded on a nitrocellulose membrane at a concentration of $4.6x10^8$ particles/mL (same

concentration of mice injection). The measurements of fluorescent signal were evaluated by Image Studio™ Software accompanied with Odyssey. The software removes automatically the background signal.

**Flow cytometry of cells from mice.** NOD SCID mice were retro-orbitally injected with PKH67-stained EVs, WT and *SDC4* KO #2, and PBS (control), at the amount of $4.6x10^8$ particles/mice. Mice were sacrificed 18h after injection, the lung and liver were isolated and perfused with Collagenase D and H, respectively. Both femurs and spleen of the mice were isolated and washed in once in PBS. Cells were isolated from bone marrow, spleen, lung and liver of NOD SCID, filtered through a 70-mm cell strainer, treated with FcBlock (BD Biosciences) for 15min, on ice in the dark, and then labelled with Fixable Viability Stain 700 (BD Horizon); CD45-APC-Cy7 (BD Pharmigen); F4/80-PE (BioLegend); CD11b-PE-Cy7 (BioLegend) and Tim4-APC Alexa 647 (BioLegend) for 1h, on ice in the dark. After washing, cells were resuspended in FACS buffer and analysed with LSRFortessa TM X-20 flow cytometer (BD Biosciences). List mode data files were analysed using FlowJo v10.0 software.

**Statistical analysis.** Prism software (GraphPad, v8.0) and IBM® SPSS® Statistics v.22 were used for statistical analysis. All values are displayed as mean ± SD, unless otherwise specified. Significance was calculated using two-way ANOVA or student t-test with Welch's correction for continuous variables and Chi-square test for categorical variables, as indicated. Kaplan-Meier (KM) survival curves were generated to compare the OS and DFS between patients with positive and negative SDC4 expression. The statistical significance between the curves was determined using the log-rank test. Multivariate Cox proportional hazard regression models were performed to determine independent factors associated with prognosis. In all statistical tests $p \leq 0.05$ was considered to be statistically significant (ns – non significant; *$p \leq 0.05$; **$p < 0.01$ ***$p < 0.001$; ****$p < 0.0001$), with 95% confidence interval.

**References to SI Materials and Methods**


1. G. David, B. van der Schueren, P. Marynen, J. J. Cassiman, H. van den Berghe, Molecular cloning of amphiglycan, a novel integral membrane heparan sulfate proteoglycan expressed by epithelial and fibroblastic cells. *J Cell Biol* **118**, 961-969 (1992).
2. A. M. Szász *et al.*, Cross-validation of survival associated biomarkers in gastric cancer using transcriptomic data of 1,065 patients. *Oncotarget* **7**, 49322-49333 (2016).
3. H. Li, J. Gui, Partial Cox regression analysis for high-dimensional microarray gene expression data. *Bioinformatics* **20 Suppl 1**, i208-215 (2004).
4. Y.-H. Chen *et al.*, The GAGOme: a cell-based library of displayed glycosaminoglycans. *Nature Methods* **15**, 881-888 (2018).
5. C. S. Hughes *et al.*, Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. *Nat Protoc* **14**, 68-85 (2019).
6. P. Bardou, J. Mariette, F. Escudié, C. Djemiel, C. Klopp, jvenn: an interactive Venn diagram viewer. *BMC Bioinformatics* **15**, 293 (2014).